

MULTIPLE-INPUT MULTIPLE-OUTPUT WIRELESS  
COMMUNICATION SYSTEMS WITH COCHANNEL  
INTERFERENCE

by

Yi Song

A thesis submitted to the  
Department of Electrical and Computer Engineering  
in conformity with the requirements  
for the degree of Doctor of Philosophy

Queen's University  
Kingston, Ontario, Canada  
September 2003

Copyright © Yi Song, 2003

# Abstract

To meet the requirement of very high data rates for wireless Internet and multimedia services, multiple transmitting and multiple receiving antennas have been proposed for fourth generation wireless systems. In cellular systems, performance is limited by fading and cochannel interference from other users. Most of the current studies on multiple-input multiple-output (MIMO) systems assume that the cochannel interference is both spatially and temporally white. In this thesis, we focus on MIMO systems under both spatially and temporally colored interference.

In MIMO systems, diversity gain is effectively achieved by the multiple receiving antennas. Outage performances of several receive diversity schemes are analytically compared for an interference-limited environment in a Rayleigh fading channel. We investigate three diversity schemes: a practical variation of maximal-ratio combining, equal-gain combining (EGC) and selection combining (SC). An exact outage probability expression is derived for EGC by accurately calculating the interference power at the output of the combiner. It is found that the relative performance between EGC and SC depends on the number of interferers and interferer power distribution.

Channel estimation and data detection for MIMO systems under both spatially and temporally colored interference are studied. By modelling interference statistics as being approximately temporally and spatially separable, we propose an algorithm to jointly estimate channel and spatial interference correlation matrices based on maximum likelihood

principle. Multi-vector-symbol data detection is developed to exploit the temporal interference correlation. In the case of one interferer operating at a lower data rate, the results show that significant improvement can be achieved by taking account of the temporal interference correlation in data detection.

Information capacities of MIMO channels under spatially and temporally correlated interference are derived. Capacity gains due to the knowledge of the channel matrix and interference statistics at the transmitter are assessed. To achieve these capacity gains, we propose an adaptive modulation scheme exploiting the channel matrix and interference statistics estimated by the receiver. In particular, the impact of channel estimation error and feedback quantization error on adaptive modulation is evaluated for MIMO systems.

# Acknowledgments

I am grateful to my thesis advisor Dr. Steven D. Blostein for his guidance, technical advice, encouragement, and financial support duration my Ph.D. study.

I would like to thank my thesis committee members, Dr. Jon W. Mark from University of Waterloo, Dr. Fady Alajaji from Department of Mathematics and Statistics, Dr. Saeed Gazor and Dr. Peter J. McLane, for their comments and suggestions with respect to this thesis.

I would like to thank my family for their support and understanding. In particular I can never thank my mother enough for her patience and her believing in me all these years.

I thank all my labmates, past and present, for their friendship and all the good time we had together. I would also like to thank Dr. Julian Cheng from University of Alberta for his constant encouragement and support when my confidence was down.

This work is in part sponsored by Ontario Graduate Scholarships, Canadian Institute for Telecommunications Research, Graduate Awards from Queen's University, Samsung Electronics, as well as Bell Canada. Their financial support are greatly appreciated.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>xvi</b>
<b>Acronyms</b>	<b>xvii</b>
<b>List of Important Symbols</b>	<b>xviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation and Thesis Overview . . . . .	3
1.2 Thesis Contributions . . . . .	5
<b>2 Background</b>	<b>7</b>
2.1 Circularly Symmetric Complex Gaussian Random Vectors . . . . .	7
2.2 MIMO Data Processing . . . . .	8
2.2.1 Zero-forcing detector . . . . .	10
2.2.2 MMSE detector . . . . .	10
2.2.3 Zero-forcing and MMSE detectors with ordering . . . . .	12

2.3	Channel Model . . . . .	13
2.3.1	Specular component . . . . .	14
2.3.2	Scattered component . . . . .	14
2.4	Properties of Kronecker Product . . . . .	16
<b>3</b>	<b>Outage Probability Comparisons for Diversity Systems with Cochannel Interference in Rayleigh Fading</b>	<b>18</b>
3.1	Introduction . . . . .	18
3.2	System Model . . . . .	22
3.3	Outage Probabilities of EGC, CMC and SC with CCI . . . . .	24
3.3.1	EGC Outage Probability . . . . .	25
3.3.2	CMC Outage Probability . . . . .	33
3.3.3	SC Outage Probability . . . . .	35
3.4	Analytical Outage Probability Comparisons . . . . .	36
3.4.1	Outage Probability Comparison for CMC and EGC . . . . .	36
3.4.2	Outage Probability Comparison for CMC and SC . . . . .	38
3.4.3	Outage Probability Comparison for EGC and SC . . . . .	41
3.5	Results and Discussions . . . . .	42
3.5.1	Coherent and Incoherent Interference Power Calculations for EGC . . . . .	42
3.5.2	Outage Probability Comparisons For CMC, EGC and SC . . . . .	47
3.5.3	Finite SNRs . . . . .	49
3.6	Conclusions . . . . .	51
<b>4</b>	<b>Channel Estimation and Data Detection for MIMO Systems under Spatially and Temporally Colored Interference</b>	<b>53</b>
4.1	Introduction . . . . .	53

4.2	System Model . . . . .	55
4.3	Joint Estimation of Channel Matrix and Spatial Interference Statistics . . . . .	57
4.3.1	Maximum-likelihood solution . . . . .	57
4.3.2	Special case: temporally white interference . . . . .	61
4.3.3	Whitening filter interpretation . . . . .	62
4.4	Data Detection . . . . .	64
4.5	Applications . . . . .	65
4.5.1	System model . . . . .	65
4.5.2	Interference statistics . . . . .	67
4.5.3	Data detection without estimating channel and interference . . . . .	72
4.5.4	Simulation results . . . . .	73
4.6	Conclusions . . . . .	89
<b>5</b>	<b>Information Capacity of MIMO Systems with Spatially and Temporally Colored Interference</b>	<b>91</b>
5.1	Introduction . . . . .	91
5.2	Channel Capacity: General Case . . . . .	92
5.3	Special Cases . . . . .	97
5.3.1	Transmitter knows channel and spatial interference correlation matrices . . . . .	97
5.3.2	Transmitter knows channel matrix . . . . .	98
5.3.3	Transmitter has no knowledge of channel and interference . . . . .	99
5.4	Numerical Results . . . . .	100
5.5	Conclusions . . . . .	105
<b>6</b>	<b>Adaptive Modulation for MIMO Systems</b>	<b>107</b>

6.1	Introduction . . . . .	107
6.2	Diagonalization of MIMO Channels . . . . .	109
6.2.1	Temporally white interference . . . . .	109
6.2.2	Temporally colored interference . . . . .	110
6.3	Adaptive Modulation . . . . .	112
6.3.1	Adaptive modulation algorithm . . . . .	113
6.3.2	Bit error rate of modulation schemes . . . . .	114
6.4	Effect of Imperfect Channel Estimates . . . . .	118
6.4.1	System model . . . . .	118
6.4.2	Simulation results . . . . .	120
6.5	Effect of Feedback Quantization . . . . .	126
6.5.1	Analysis of quantization error . . . . .	127
6.5.2	Simulation results . . . . .	128
6.6	Conclusion . . . . .	132
<b>7</b>	<b>Conclusions and Future Work</b>	<b>134</b>
7.1	Conclusions . . . . .	134
7.2	Future Work . . . . .	137
<b>A</b>	<b>Variance of <math>z_i[n]</math></b>	<b>138</b>
<b>B</b>	<b>Circular Symmetry of <math>\frac{c_s^\dagger c_i}{ c_s }</math></b>	<b>140</b>
<b>C</b>	<b>Alternative derivation of <math>\hat{H}</math> and <math>\hat{R}</math></b>	<b>142</b>
<b>D</b>	<b>Proof of <math>\det(I + AB) = \det(I + BA)</math></b>	<b>145</b>
	<b>Bibliography</b>	<b>146</b>



## List of Tables

4.1	Corresponding values of $Q$ for various confidence levels. . . . .	89
6.1	Required SNRs in dB for a predefined target BER. . . . .	118

# List of Figures

2.1	Geometric configuration of a $(2, 2)$ channel with local scatterers around the mobile user: $BS_p$ is the $p$ th antenna element at the base station, $U_l$ is the $l$ th antenna element at the mobile, $\delta_{pq}$ is the antenna spacing in wavelengths between the $p$ th and $q$ th antenna elements at the base station, $d_{lm}$ is the antenna spacing in wavelengths between the $l$ th and $m$ th antenna elements at the mobile, and $\Delta$ is the angle spread. . . . .	16
3.1	Regions of integration for the conditional outage probability of CMC and EGC. . . . .	37
3.2	Probability density functions of $\mu_1$ and $\zeta_1$ for $L = 1$ interferer and $N_r = 4$ antennas. . . . .	44
3.3	Outage probability comparison of coherent and incoherent interference power calculation for EGC with one interferer ( $L = 1$ ) and equal interferer powers ( $\Lambda = 10$ dB) for $N_r = 1, 2$ , and 4 antennas. . . . .	45
3.4	Outage probability comparison of coherent and incoherent interference power calculation for EGC with four antennas ( $N_r = 4$ ) and equal interferer powers ( $\Lambda = 10$ dB) for $L = 2, 6$ , and 10 interferers. . . . .	45

3.5	Analytical EGC outage probability using coherent interference power calculation and Monte Carlo simulated outage probability for equal ( $\Lambda = 10$ dB) and distinct ( $\Lambda_{avg} = 10$ dB) interference power distributions with $L = 2$ interferers and $N_r = 4$ antennas. The interference power vector for $L = 2$ is $[0.1, 0.9]$ . . . . .	46
3.6	Outage probability of CMC, EGC and SC with $N_r = 4$ antennas and equal interferer powers ( $\Lambda = 10$ dB) for $L = 1, 2,$ and $6$ interferers. . . . .	48
3.7	Outage probability of CMC, EGC and SC with $L = 3$ interferers and equal interferer powers ( $\Lambda = 10$ dB) for $N_r = 2$ and $4$ antennas. . . . .	48
3.8	Outage probability of CMC, EGC and SC with $N_r = 4$ antennas and distinct interferer powers ( $\Lambda_{avg} = 10$ dB) for $L = 2$ and $6$ interferers. The interference power vectors for $L = 2$ and $6$ are, respectively, $[0.1, 0.9]$ and $[0.05, 0.1, 0.15, 0.22, 0.23, 0.25]$ . . . . .	50
3.9	Outage probability of CMC, EGC and SC with $L = 3$ interferers and distinct interferer powers ( $\Lambda_{avg} = 10$ dB) for $N_r = 2$ and $4$ antennas. The interference power vector for $L = 3$ is $[0.1, 0.2, 0.7]$ . . . . .	50
3.10	Monte Carlo simulated outage probability of CMC, EGC and SC for a finite SNR: $N_r = 4$ antennas, $L = 1$ interferer, equal interferer powers $\Lambda = 10$ dB, and SNR=20 dB. . . . .	51
3.11	Monte Carlo simulated outage probability of CMC, EGC and SC for a finite SNR: $N_r = 4$ antennas, $L = 1$ interferer, equal interferer powers $\Lambda = 10$ dB, and SNR=10 dB. . . . .	52
4.1	Waveform of $g_I(t)$ for the case of lower-data-rate interferer with $T_I = 2T$ , $T = 1$ and $\beta = 1$ . . . . .	70

4.2	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $2N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate. . . . .	76
4.3	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $4N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate. . . . .	77
4.4	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $6N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate. . . . .	77
4.5	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $2N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user. . . . .	79
4.6	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $4N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user. . . . .	79
4.7	Average symbol error rate vs. SIR with $N_t = N_r = 5$ , $L = 6$ , and training length $6N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user. . . . .	80
4.8	Average symbol error rate vs. angle spread with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, and training length $4N_t$ . Correlated Rayleigh fading (Ricean factor $K = 0$ ) is assumed for the desired user. Both the desired and interfering users are at the same data rate. . . . .	82

4.9	Average symbol error rate vs. angle spread with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, and training length $4N_t$ . Correlated Rayleigh fading (Ricean factor $K = 0$ ) is assumed for the desired user. The data rate of the desired user is twice that of interfering user. . . . .	83
4.10	Average symbol error rate vs. $K$ -Ricean factor with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, training length $4N_t$ . The channel scattered components are independently faded for the desired user. Both the desired and interfering users are at the same data rate. . . . .	84
4.11	Average symbol error rate vs. $K$ -Ricean factor with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, and training length $4N_t$ . The channel scattered components are independently faded for the desired user. The data rate of the desired user is twice that of interfering user. . . . .	84
4.12	Average symbol error rate vs. INR with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, and training length $4N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate.	85
4.13	Average symbol error rate vs. INR with $N_t = N_r = 5$ , $L = 6$ , SIR=10dB, and training length $4N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user.	86
4.14	The improvement of estimating spatial correlation of interference-plus-noise in practical systems. With total interference power fixed, the solid lines are for one interferer, and the broken lines are for two interferers. Both the desired and interfering users employ a $(5, 5)$ MIMO link, the same data rate, total-interference-to-noise-ratio of 12dB, and training length of $4N_t$ . .	87

5.1	10% outage capacity versus SIR for temporally and spatially colored interference (interferer delay $0.4T$ ) with $N_t = N_r = L = 4$ . Independent Rayleigh fading is assumed for the desired user. . . . .	103
5.2	10% outage capacity versus SIR for spatially colored interference (synchronous interferer) with $N_t = N_r = L = 4$ . Independent Rayleigh fading is assumed for the desired user. . . . .	103
5.3	10% outage capacity versus $K$ -Ricean factor for temporally and spatially colored interference (interferer delay $0.4T$ ) with $N_t = N_r = L = 4$ and SIR=10dB. The scattered component of the desired user's channel is independently faded. . . . .	104
5.4	10% outage capacity versus angle spread for temporally and spatially colored interference (interferer delay $0.4T$ ) with $N_t = N_r = L = 4$ and SIR=10dB. Correlated Rayleigh fading (Ricean factor $K = 0$ ) is assumed for the desired user. . . . .	105
6.1	Diagonization of an MIMO channel under temporally white interference. . . . .	111
6.2	Diagonization of an MIMO channel under temporally colored interference. . . . .	112
6.3	QPSK signal constellation with Gray code mapping. . . . .	114
6.4	16-square QAM signal constellation with Gray code mapping. . . . .	115
6.5	64-square QAM signal constellation with Gray code mapping. . . . .	116
6.6	Bit error rates vs. SNR per symbol in AWGN channel for BPSK, QPSK, 16-square QAM and 64-square QAM. . . . .	117
6.7	Adaptive modulation with imperfect channel estimates. . . . .	119

6.8	The number of bits transmitted per Hz per sec for adaptive modulation and non-adaptive scheme assuming perfect channel knowledge at the transmitter. The target BER in adaptive modulation is $10^{-3}$ , $N_t = N_r = 4$ , and $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading. . . . .	121
6.9	Average bit error rate for adaptive modulation and non-adaptive scheme assuming perfect channel knowledge at the transmitter. The target BER in adaptive modulation is $10^{-3}$ , $N_t = N_r = 4$ , and $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading. . . . .	122
6.10	The achieved bit error rate versus training length of adaptive modulation with target BER $10^{-3}$ . It is assumed that $N_t = N_r = 4$ , and $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading. . . . .	123
6.11	The number of bits transmitted per Hz per sec versus SIR for adaptive modulation with $N_t = N_r = 4$ , $L = 5$ . The actual achieved BER is $10^{-3}$ . The desired user is assumed to experience i.i.d. Rayleigh fading. . . . .	124
6.12	The number of bits transmitted per Hz per sec versus Ricean factor $K$ for adaptive modulation with $N_t = N_r = 4$ , $L = 5$ , SIR=10 dB. The actual achieved BER is $10^{-3}$ . The scattered components of the desired user's channel are assumed to be independently faded. . . . .	125
6.13	The number of bits transmitted per Hz per sec versus angle spread for adaptive modulation with $N_t = N_r = 4$ , $L = 5$ , SIR=10 dB. The actual achieved BER is $10^{-3}$ . The desired user is assumed to experience correlated Rayleigh fading ( $K = 0$ ). . . . .	126
6.14	Adaptive modulation with feedback quantization. . . . .	129
6.15	Average bit error rate versus quantization level of feedback with $N_t = N_r = 4$ , SNR=15 dB and target BER $10^{-2}$ . . . . .	130

6.16	The number of bits transmitted per Hz per sec versus quantization level of feedback with $N_t = N_r = 4$ , SNR=15dB and target BER $10^{-2}$ . . . . .	131
6.17	Average bit error rate versus quantization level of feedback with $N_t = N_r = 4$ , SNR=15 dB and target BER $10^{-3}$ . . . . .	131
6.18	The number of bits transmitted per Hz per sec versus quantization level of feedback with $N_t = N_r = 4$ , SNR=15dB and target BER $10^{-3}$ . . . . .	132

# Acronyms

BER	Bit Error Rate
CCI	Cochannel Interference
CMC	Channel-Matched Combining
EGC	Equal Gain Combining
i.i.d.	independent identically distributed
LS	Least Squares
MIMO	Multiple-Input Multiple-Output
ML	Maximum-Likelihood
MMSE	Minimum Mean-Square Error
MRC	Maximal Ratio Combining
PDF	Probability Density Function
RV	Random Variable
SC	Selection Diversity
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
ZF	Zero-forcing

# List of Important Symbols

$(\cdot)^T$	Matrix or vector transpose
$(\cdot)^*$	Complex conjugate
$(\cdot)^\dagger$	Matrix of vector conjugate transpose
$\otimes$	Kronecker product
$*$	Signal convolution
$\Gamma(x)$	Gamma function
$\Lambda_N$	Temporal interference correlation matrix
$C$	Channel information capacity
$D_p(z)$	Parabolic cylinder function
$\det(\cdot)$	Determinant of a matrix
$E\{\cdot\}$	Expectation of random variables
${}_1F_1(\cdot; \cdot; \cdot)$	Degenerate hypergeometric function
$\mathbf{H}$	MIMO channel matrix of the desired user
$\hat{\mathbf{H}}$	Estimate of channel matrix $\mathbf{H}$
$\mathcal{H}(\cdot)$	Entropy
$\mathbf{I}_N$	$N \times N$ identity matrix
$\Im(\cdot)$	Imaginary part of a complex number
$\mathcal{I}(\cdot; \cdot)$	Mutual information
$L$	Number of transmitting antennas of the interferer

$N_t$	Number of transmitting antennas of the desired user
$N_r$	Number of receiving antennas of the desired user
$\mathbf{R}$	Spatial interference correlation matrix
$\hat{\mathbf{R}}$	Estimate of spatial interference correlation matrix $\mathbf{R}$
$\Re(\cdot)$	Real part of a complex number
$\text{tr}(\cdot)$	Trace of a matrix

# Chapter 1

## Introduction

Communicating over a wireless channel is highly challenging due to the complex propagation medium. The major impairments of the wireless channel are fading and cochannel interference. Due to ground irregularities and typical wave propagation phenomena such as diffraction, scattering, and reflection, when a signal is launched into the wireless environment, it arrives at the receiver along a number of distinct paths, referred to as multipath. Each of these paths has a distinct and time-varying amplitude, phase and angle of arrival. These multipaths add up constructively or destructively at the receiver. Hence, the received signal can vary over frequency, time, and space. These variations are collectively referred to as fading and deteriorate the link quality. Moreover, in cellular systems, to maximize the spectral efficiency and accommodate more users while maintaining the minimum quality of service, frequencies have to be reused in different cells that are separated sufficiently apart. Therefore, the desired user's signal may be corrupted by the interference generated by other users operating at the same frequency. This kind of interference is called cochannel interference (CCI). As a result, to increase capacity and spectral efficiency of wireless communication systems, it is crucial to mitigate fading and CCI.

One of the key technologies to mitigate fading and CCI is to implement antenna arrays

in the system [45, 46, 70, 83, 122]. Antenna arrays can be employed at the transmitter, or receiver, or both ends. With an antenna array at the receiver, fading can be reduced by diversity techniques, i.e., combining independently faded signals on different antennas that are separated sufficiently apart. If antennas receive independently faded signals, it is unlikely that all signals undergo deep fades, hence, at least one good signal can be received. Three common diversity schemes are selection combining (SC), equal gain combining (EGC) and maximal ratio combining (MRC). To reduce strong interference, appropriate combining weights can be chosen to maximize the signal-to-interference-plus-noise ratio (SINR), i.e., enhance the desired signal and suppress the interfering signals, as well as reduce fading. If the desired and interfering signals are highly directional, the array radiation pattern may form a beam, i.e., beamform, to the desired user and null the interfering signals.

Recently, antenna arrays located at transmitters have attracted much interest. Transmit diversity was first introduced by Wittneben in [123], and later popularized by Tarokh's space-time codes [113]. Similar to the receiver-based beamforming, if the channel information of the desired and cochannel users is available at the transmitter, transmit beamforming can be used to enhance the signal-to-noise ratio (SNR) for the intended user and minimize the interference energy sent towards cochannel users [32, 41, 92].

To meet the requirement of very high data rates for wireless Internet and multimedia services, multiple antennas at both the transmitter and receiver have been proposed for fourth generation broadband wireless systems [73, 80, 96]. In a rich scattering environment where channel links between different transmitters and receivers fade independently, it was shown that, by decomposing a multiple-input multi-output (MIMO) channel into several single-input single-output (SISO) subchannels, the Shannon's information capacity of a MIMO channel increases linearly with the smaller of the numbers of transmitting and receiving antennas [35, 114]. To realize this high capacity, various space-time transmission

schemes were investigated, including space-time trellis coding [112,113], space-time block coding [8, 110, 111], and space-time differential coding [57, 58]. Moreover, considerable work has been conducted to exploit the MIMO capacity using the already highly developed one-dimensional coding and decoding techniques. As a result, different layered space-time architectures were proposed, including Diagonal- [34], Vertical- [47], and Turbo-Bell Labs Layered Space-Time [99], also known as D-, V-, and T-BLAST, respectively. State-of-the-art research of MIMO systems was reviewed in [42]. Information capacity of MIMO channels under different environment has been summarized in [48].

## 1.1 Motivation and Thesis Overview

Most of the current studies on MIMO systems assume that the interference is both spatially and temporally white. For example, information channel capacity of a MIMO link under both spatially and temporally white interference was assessed in [35, 114], channel estimation was studied in [52, 71, 77, 100, 109] and data detection was investigated in [36, 47, 55]. Spatially white interference means that the interfering signals on different receiving antennas are uncorrelated with the same power. Temporally white interference implies that in the decision statistics for symbol detection, the interfering signals are uncorrelated from symbol to symbol with the same power. However, in cellular systems, the interference may be both spatially and temporally colored. The spatial correlation can be explained by the simple case of one interferer: the interfering signals at different antennas are different scaled versions of the same signal, hence they are correlated. The temporal correlation may be caused by the intersymbol interference as we will explain more in the thesis.

Recently, MIMO systems with spatially colored interference have attracted interest. Information channel capacity of MIMO systems under spatially colored interference was

studied in [15, 19, 30, 31]. Performance analysis of outage and error rate for MIMO systems with cochannel interference was given in [61]. In this thesis, we mainly focus on MIMO systems under both spatially and temporally colored interference in slow flat fading.

In Chapter 2, background on data detection algorithms for MIMO systems is first reviewed. The MIMO channel model used in this thesis is then described. Key properties of the Kronecker matrix product used extensively in the thesis are also reviewed.

We begin with a discussion on diversity systems, which are single-input multiple-output systems (SIMO), in Chapter 3. It is well known that receiver diversity can combat fading and, to some extent, reduce CCI. A comparative analysis of outage performance for MRC, EGC and SC in fading and CCI has not been attempted in the literature. In Chapter 3, we first derive an exact outage probability expression for EGC, then provide an analytical comparison on outage probability for MRC, EGC and SC with Rayleigh fading in an interference-limited environment.

Since the Shannon's channel information capacity can be greatly increased by employing multiple transmitting antennas, we move to describe MIMO systems. In Chapter 4, we present new algorithms for joint channel estimation and data detection under spatially and temporally colored interference. The impact of spatially and temporally colored interference on system performance are assessed by Monte Carlo simulations.

Chapter 4 focuses on the processing at the receiver-side and assumes that the transmitter has no knowledge of the channel matrix and interference statistics. In Chapter 5, we assess the potential impact of matrix channel knowledge as well as interference statistics at the transmitter on channel capacity under spatially and temporally correlated interference. Assuming the receiver has perfect knowledge of the channel matrix and interference statistics, we derive the channel capacities for different combinations of knowledge of the channel matrix and interference statistics at the transmitter.

In Chapter 6, we consider joint processing at the transmitter and receiver in practical systems. With the estimates of channel matrix and interference statistics in Chapter 4, we propose an adaptive modulation scheme to increase the system spectral efficiency. We also investigate the practical issue of channel estimation error and feedback quantization error for MIMO adaptive modulation.

Chapter 7 concludes this thesis and suggests future work.

## 1.2 Thesis Contributions

The primary contributions of this thesis are briefly summarized as follows.

- An exact outage probability expression is derived for EGC in Rayleigh fading with multiple interferers by accurately calculating the interference power at the output of the combiner. Using this exact expression, the accuracy of the approximate interference power calculation in the existing literature is assessed.
- Outage performances of several diversity schemes, including a practical variation of MRC that does not require signal-to-noise ratios at different antennas, EGC and SC, are compared analytically for an interference-limited environment in a Rayleigh fading channel. The analysis provides insight into performance of diversity schemes in the presence of CCI, as well as assesses the impact of cochannel interferer power distributions.
- Performance of a MIMO user in a multi-user environment is considered. An algorithm is proposed to jointly estimate the channel and spatial interference correlation matrices for the desired MIMO user under spatially and temporally colored interference. By exploiting temporal interference correlation, a multi-vector-symbol data

detection scheme is developed. The benefits of taking temporal and spatial interference correlation into account for channel estimation and data detection are evaluated through simulations.

- Assuming that the receiver has perfect knowledge of the channel matrix and interference statistics, information capacities of MIMO channels are derived for different combinations of knowledge of the channel matrix and interference statistics at the transmitter under both spatially and temporally colored interference. The benefits of taking the knowledge of the channel matrix and interference statistics at the transmitter are assessed.
- With the channel matrix and interference statistics estimated at the receiver, we propose an adaptive modulation scheme to increase the system spectral efficiency. The impact of channel estimation error and feedback quantization error on adaptive modulation is evaluated for MIMO systems. Rate-distortion theory is used to assess the achievable performance when feedback on channel information, from receiver to transmitter, is quantized.

# Chapter 2

## Background

In this chapter, we first review the definition of circularly symmetric complex Gaussian random vectors. Data detection algorithms for multiple-input multiple-output (MIMO) systems and the MIMO channel model used in this thesis are then described. Finally, we review key properties of the Kronecker matrix product that will be used in later chapters.

### 2.1 Circularly Symmetric Complex Gaussian Random Vectors

In this thesis, since we deal with the circularly symmetric complex Gaussian random vector quite often, we review its definition [86].

**Definition 1** *A complex random vector  $\mathbf{x}$  is circularly symmetric Gaussian with mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{R}$  if*

(1) *the elements in  $\tilde{\mathbf{x}} = \begin{bmatrix} \Re(\mathbf{x}) \\ \Im(\mathbf{x}) \end{bmatrix}$ , where  $\Re(\cdot)$  and  $\Im(\cdot)$  denote the real and imaginary part, respectively, are jointly Gaussian;*

(2)  $E\{\mathbf{x}\} = \boldsymbol{\mu}$ ;

$$(3) E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\dagger\} = \mathbf{R}, \text{ and } E\{(\tilde{\mathbf{x}} - E[\tilde{\mathbf{x}}])(\tilde{\mathbf{x}} - E[\tilde{\mathbf{x}}])^T\} = \frac{1}{2} \begin{bmatrix} \Re(\mathbf{R}) & -\Im(\mathbf{R}) \\ \Im(\mathbf{R}) & \Re(\mathbf{R}) \end{bmatrix} \text{ where}$$

$\dagger$  and  $T$  denote conjugate-transpose and transpose, respectively.

We denote the circularly symmetric complex Gaussian random vector as  $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ , i.e., a random vector with probability density function (PDF)

$$f(\mathbf{x}) = \frac{1}{\pi^N \det(\mathbf{R})} \exp\left\{-\mathbf{x}^\dagger \mathbf{R}^{-1} \mathbf{x}\right\}$$

where  $N$  is the dimension of the random vector. Note that only for the circularly symmetric complex Gaussian random vector can its PDF be completely specified by the mean vector  $\boldsymbol{\mu}$  and the covariance matrix  $\mathbf{R}$ . In general, the PDF of a complex Gaussian random vector is determined from the mean vector  $\boldsymbol{\mu}$  and the covariance matrix of  $\tilde{\mathbf{x}}$ , i.e.,  $E\{(\tilde{\mathbf{x}} - E[\tilde{\mathbf{x}}])(\tilde{\mathbf{x}} - E[\tilde{\mathbf{x}}])^T\}$ .

The random vector degenerates to the random variable if the dimension of the vector,  $N$ , is reduced to 1. If  $x = x_1 + jx_2$  is a circularly symmetric complex Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$ , then  $E(x) = \mu$ , and  $x_1$  and  $x_2$  are independent joint Gaussians each with variance  $\sigma^2/2$ .

## 2.2 MIMO Data Processing

Consider a MIMO link with  $N_t$  transmitting and  $N_r$  receiving antennas, denoted as  $(N_t, N_r)$ .

The baseband model of the received signal vector is expressed as [47]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.1}$$

where  $\mathbf{H}$  is the  $N_r \times N_t$  channel matrix, and  $\mathbf{x}$  is the  $N_t \times 1$  transmitted signal vector. The  $N_r \times 1$  noise vector  $\mathbf{n}$  is assumed to be circularly symmetric complex Gaussian with zero mean and covariance matrix  $\mathbf{R}$ .

Assume that the receiver has perfect knowledge of channel matrix  $\mathbf{H}$  and spatial noise covariance matrix  $\mathbf{R}$ . If the transmitted signal  $\mathbf{x}$  is chosen from the signal constellation with equal probability, the optimum receiver is a maximum-likelihood (ML) receiver that selects the most probable transmitted signal vector  $\mathbf{x}$  given the received signal vector  $\mathbf{y}$ . More specifically, the optimum ML receiver selects a transmitted signal vector that maximizes the conditional PDF

$$\Pr(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^{N_r} \det(\mathbf{R})} \exp \left\{ -(\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right\}. \quad (2.2)$$

Assuming the signal transmitted on each antenna is drawn from an  $M$ -ary signal constellation, there are  $M^{N_t}$  possible choices of the transmitted signal vector. The optimum receiver computes the conditional PDF for each possible transmitted signal vector, and selects the one that yields the largest conditional PDF. Hence, the complexity of the optimum ML receiver grows exponentially with the number of transmitting antennas,  $N_t$ .

Due to the high complexity of the optimum receiver, various suboptimal receivers which yield a reasonable tradeoff between performance and complexity have been investigated. Examples of nonlinear suboptimal detectors are the sphere detector [27] and detectors which combine linear processing with local ML search [69]. The linear suboptimal detectors usually used in practice are zero-forcing (ZF) and minimum mean-squared error (MMSE) detectors [36, 44, 47]. Data detection for MIMO systems is similar to multiuser detection for synchronous users [117], where in MIMO systems we consider one user having multiple transmitting antennas and in multi-user detection we consider multiple users each having one transmitting antenna. The ZF and MMSE MIMO detectors are akin to the decorrelating and MMSE multiuser detectors, respectively.

In the following, we briefly derive ZF and MMSE detectors which include the detection algorithms in [36, 44, 47] as special cases of spatially white noise. We assume that  $N_t \leq N_r$ . Note that these two detectors are valid even for non-Gaussian noise.

### 2.2.1 Zero-forcing detector

With Gaussian noise, the best linear estimate of  $\mathbf{x}$  is the value of  $\mathbf{x}$  that maximizes the conditional PDF in (2.2), which is equivalent to minimizing the term

$$L(\mathbf{x}) = (\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}). \quad (2.3)$$

Setting  $\partial L(\mathbf{x})/\partial \mathbf{x} = \mathbf{0}$  yields the soft estimate

$$\tilde{\mathbf{x}}_{\text{ZF}} = (\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{y}. \quad (2.4)$$

The detected signal vector is obtained by quantizing the soft estimate  $\tilde{\mathbf{x}}_{\text{ZF}}$  to the nearest value in the signal constellation.

Substituting (2.1) into (2.4), we obtain

$$\tilde{\mathbf{x}}_{\text{ZF}} = \mathbf{x} + (\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{n}. \quad (2.5)$$

From (2.5), we observe that, regardless of whether the noise is Gaussian or not,  $\tilde{\mathbf{x}}_{\text{ZF}}$  is a zero-forcing solution, completely nulling out signals from undesired transmitters. Hence,  $\tilde{\mathbf{x}}_{\text{ZF}}$  is an unbiased estimate of  $\mathbf{x}$ . It can be shown that the covariance matrix of the estimation error is

$$E\{(\mathbf{x} - \tilde{\mathbf{x}}_{\text{ZF}})(\mathbf{x} - \tilde{\mathbf{x}}_{\text{ZF}})^\dagger\} = (\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H})^{-1}. \quad (2.6)$$

For spatially white noise with  $\mathbf{R} = \sigma^2 \mathbf{I}_{N_r}$ , the soft estimate in (2.4) is reduced to  $\tilde{\mathbf{x}}_{\text{ZF}} = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{y}$  [47], where  $(\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger$  is the pseudo-inverse of  $\mathbf{H}$  if the rank of  $\mathbf{H}$  equals  $N_t$  [50].

### 2.2.2 MMSE detector

We seek linear estimate  $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{y}$  such that the mean square error (MSE)

$$J(\mathbf{A}) = \text{tr} \left\{ E [(\mathbf{x} - \mathbf{A}\mathbf{y})(\mathbf{x} - \mathbf{A}\mathbf{y})^\dagger] \right\} \quad (2.7)$$

is minimized. Without loss of generality, we assume that the transmitted signal vector is zero-mean and with covariance matrix  $E\{\mathbf{x}\mathbf{x}^\dagger\} = \mathbf{I}_{N_t}$ . It is also assumed that the transmitted signal vector is independent of the noise vector, i.e.,  $E\{\mathbf{x}\mathbf{n}^\dagger\} = \mathbf{0}$ . Substituting (2.1) into (2.7), the MSE becomes

$$J(\mathbf{A}) = \text{tr} \left\{ \mathbf{I}_{N_t} - \mathbf{A}\mathbf{H} - \mathbf{H}^\dagger\mathbf{A}^\dagger + \mathbf{A} \left( \mathbf{H}\mathbf{H}^\dagger + \mathbf{R} \right) \mathbf{A}^\dagger \right\}. \quad (2.8)$$

By setting  $\partial J(\mathbf{A})/\partial \mathbf{A} = \mathbf{0}$ , we obtain

$$\mathbf{A} = \mathbf{H}^\dagger \left( \mathbf{H}\mathbf{H}^\dagger + \mathbf{R} \right)^{-1} \quad (2.9)$$

$$= \left( \mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^\dagger \mathbf{R}^{-1} \quad (2.10)$$

where the second equality is due to the matrix identity in [78, p528, D.11]. Hence, the soft MMSE estimate is

$$\tilde{\mathbf{x}}_{\text{MMSE}} = \left( \mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{y}. \quad (2.11)$$

Again, the detected signal vector is obtained by quantizing the soft estimate  $\tilde{\mathbf{x}}_{\text{MMSE}}$  to the nearest point in the signal constellation.

Substituting (2.9) into the matrix of the trace operation in (2.8), we obtain the covariance matrix of the estimation error

$$\begin{aligned} E \left\{ (\mathbf{x} - \tilde{\mathbf{x}}_{\text{MMSE}})(\mathbf{x} - \tilde{\mathbf{x}}_{\text{MMSE}})^\dagger \right\} &= \mathbf{I}_{N_t} - \mathbf{H}^\dagger \left( \mathbf{H}\mathbf{H}^\dagger + \mathbf{R} \right)^{-1} \mathbf{H} \\ &= \mathbf{I}_{N_t} - \left( \mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \\ &= \left( \mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right)^{-1} \end{aligned} \quad (2.12)$$

where the second equality is due to the alternative expression of  $\mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger + \mathbf{R})^{-1}$  in (2.10), and the last equality comes from the fact that  $\mathbf{I}_{N_t} = (\mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H})$ . By substituting (2.1) into (2.11), it is easy to see that soft MMSE estimate  $\tilde{\mathbf{x}}_{\text{MMSE}}$  is a biased estimate of  $\mathbf{x}$ .

For spatially white noise with  $\mathbf{R} = \mathbf{I}_{N_r}$ , the estimate in (2.11) is reduced to  $\tilde{\mathbf{x}}_{\text{MMSE}} = (\mathbf{I}_{N_t} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{y}$  [55].

### 2.2.3 Zero-forcing and MMSE detectors with ordering

Analogous to the successive interference cancellation in multiuser detection [117], to enhance the receiver performance, successive symbol cancellation may be jointly used with ZF or MMSE MIMO detection. When successive cancellation is applied, the order in which the components of the transmitted signal vector are detected is important to the overall performance of the system. It is shown that post-detection signal-to-noise ratios (SNRs) should be used as the criterion for signal ordering [36]. Hence, assuming the components of the transmitted signal vector have the same power, we should first detect the signal component which has the smallest estimation error variance. With the estimation error covariance matrices in (2.6) and (2.12), the ZF and MMSE detection algorithms with ordering are described as follows [55].

Step 1 Initialization:  $k = 1$ ,  $\mathbf{H}_k = \mathbf{H}$ ,  $\tilde{\mathbf{x}}_k = \mathbf{x}$ ,  $\tilde{\mathbf{y}}_k = \mathbf{y}$ .

Step 2 Determine the ordering. Calculate the estimation error covariance matrix  $\mathbf{P}_k = (\mathbf{H}_k^\dagger \mathbf{R}^{-1} \mathbf{H}_k)^{-1}$  for ZF detector or  $\mathbf{P}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^\dagger \mathbf{R}^{-1} \mathbf{H}_k)^{-1}$  for MMSE detector. Find  $m = \arg \min_j \mathbf{P}_k(j, j)$  where  $\mathbf{P}_k(j, j)$  denotes the  $j$ th diagonal element of  $\mathbf{P}_k$ . Hence, the  $m$ th signal component of  $\tilde{\mathbf{x}}_k$  has the smallest estimation error variance.

Step 3 Calculate the weighting matrix  $\mathbf{A}_k = (\mathbf{H}_k^\dagger \mathbf{R}^{-1} \mathbf{H}_k)^{-1} \mathbf{H}_k^\dagger \mathbf{R}^{-1}$  for ZF detector or  $\mathbf{A}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^\dagger \mathbf{R}^{-1} \mathbf{H}_k)^{-1} \mathbf{H}_k^\dagger \mathbf{R}^{-1}$  for MMSE detector. The  $m$ th element of  $\tilde{\mathbf{x}}_k$  is estimated as  $\hat{x}_k^m = Q(\mathbf{A}_k(m, :) \tilde{\mathbf{y}}_k)$  where  $\mathbf{A}_k(m, :)$  denotes the  $m$ th row of matrix  $\mathbf{A}_k$  and  $Q(\cdot)$  denotes the quantization appropriate to the signal constellation.

Step 4 Assuming the detected signal is correct, remove the detected signal from the received signal,  $\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_k - \hat{x}_k^m \mathbf{H}_k(:, m)$  where  $\mathbf{H}_k(:, m)$  denotes the  $m$ th column of  $\mathbf{H}_k$ .

Step 5  $\mathbf{H}_{k+1}$  is obtained by eliminating the  $m$ th column of matrix  $\mathbf{H}_k$ .  $\tilde{\mathbf{x}}_{k+1}$  is obtained by eliminating the  $m$ th component of vector  $\tilde{\mathbf{x}}_k$ .

Step 6 If  $k < N_t$ , increment  $k$  and go to step 2.

Recall that the ZF and MMSE detection without ordering are described by (2.5) and (2.11), respectively.

Performances of ZF and MMSE detectors are compared in [9] for a (4, 4) MIMO system with QPSK modulation. It is shown that, at a bit error rate  $10^{-3}$ , compared to ZF detection, MMSE detection has a 3dB gain in SNR when no signal ordering is used and a 8dB gain for the case of signal ordering. The inferior performance of ZF detection is caused by the noise enhancement, a price paid for completely nulling out signals from undesired transmitters. Algorithms of ordered ZF or MMSE detectors with reduced computational complexity and improved numerical robustness have been investigated in [55] and [127].

## 2.3 Channel Model

To simulate the flat fading MIMO channel, we use a Ricean model [30, 93]. The channel matrix has two components: a deterministic specular (line-of-sight) component and a random scattered component. With the Ricean  $K$ -factor defined as the ratio of deterministic-to-scattered power, the channel matrix is given by

$$\mathbf{H} = \sqrt{\frac{K}{K+1}} \mathbf{H}^{sp} + \sqrt{\frac{1}{K+1}} \mathbf{H}^{sc}. \quad (2.13)$$

### 2.3.1 Specular component

The specular component is given by

$$\mathbf{H}^{sp} = \mathbf{a}_r(\theta_r)\mathbf{a}_t(\theta_t)^T \quad (2.14)$$

where  $\theta_t$  and  $\theta_r$  are the angles of departure and arrival of the specular signal at the transmitter and receiver, respectively. The array response vectors at the receiver and transmitter are  $\mathbf{a}_r(\cdot)$  and  $\mathbf{a}_t(\cdot)$ , respectively. For an  $N_r$ -element uniform linear array, the array response vector is given by

$$\mathbf{a}(\theta) = [1 \ \exp(-j2\pi d \sin\theta) \ \dots \ \exp(-j2\pi d(N_r - 1) \sin\theta)]^T \quad (2.15)$$

where  $\theta$  is the angle between the signal and the normal to the array, and  $d$  is the antenna spacing expressed in wavelengths.

### 2.3.2 Scattered component

The elements in the scattered component  $\mathbf{H}^{sc}$  are each zero-mean circularly symmetric complex Gaussian with cross-correlations determined by factors such as angle spread at the base station and mobile, antenna array geometry, and mean direction of signal arrivals. To derive the fading correlations among MIMO channels, the scattering model proposed by Jakes [59] is used to provide a reasonable description of the scattering environment around the transmitter and receiver [94]. It is usually assumed that the mobile is surrounded by local scatterers, and that the base station is elevated and unobstructed by local scatterers. In [103], the spatial fading correlation was derived for isotropic scattering (uniformly distributed scatterers) around the mobile. In [4], the space-time fading correlation of the general scenario of nonisotropic scattering around the mobile was discussed. For mathematical convenience and simplicity for simulation, the correlation matrix of MIMO links

may be approximated by a Kronecker product of fading correlation matrices for the antenna arrays at the base station and mobile [103, 104]. The validity of this approximation has been studied in [4, 22] from experimental measurements.

In this thesis, we adopt the spatial fading correlation presented in [4] due to its closed-form expression which is easy to calculate numerically (the expression in [103] involves integration). In Fig. 2.1, without loss of generality, we assume that the base station and mobile take on roles of receiver and transmitter, respectively. The base station receives the signal from a particular direction with an angle spread  $\Delta$ . Denote  $U_l - BS_p$  as the link between the  $l$ th antenna element at the mobile and the  $p$ th antenna element at the base station, and  $h_{lp}$  as the gain of the link  $U_l - BS_p$  due to scattering. For isotropic scattering around the mobile and a small angle spread  $\Delta$ , the correlation between link gains  $h_{lp}$  and  $h_{mq}$  is [4]

$$E\{h_{lp}h_{mq}^*\} \approx \exp[jc_{pq}\cos(\alpha_{pq})] \times I_0\left(\left\{-b_{lm}^2 - c_{pq}^2\Delta^2\sin^2(\alpha_{pq}) - 2c_{pq}b_{lm}\Delta\sin(\alpha_{pq})\sin(\beta_{lm})\right\}^{1/2}\right) \quad (2.16)$$

where  $I_0$  is the zeroth-order modified Bessel function;  $b_{lm} = 2\pi d_{lm}$  and  $c_{pq} = 2\pi\delta_{pq}$  where  $d_{lm}$  and  $\delta_{pq}$  are the antenna spacing in wavelengths; angles  $\alpha_{pq}$  and  $\beta_{lm}$  are shown in Fig. 2.1. Given spatial cross-correlations among MIMO links, we can simulate the elements of  $\mathbf{H}^{sc}$  by multiplying the square root of the cross-correlation matrix with a vector of independent identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussians.

The links from one mobile antenna element to two different base station elements are highly correlated at angle spread  $\Delta = 0$ , and become uncorrelated as  $\Delta$  increases. As the line-of-sight (LOS) component becomes prominent ( $K$  factor increases), the MIMO channel links become spatially correlated.

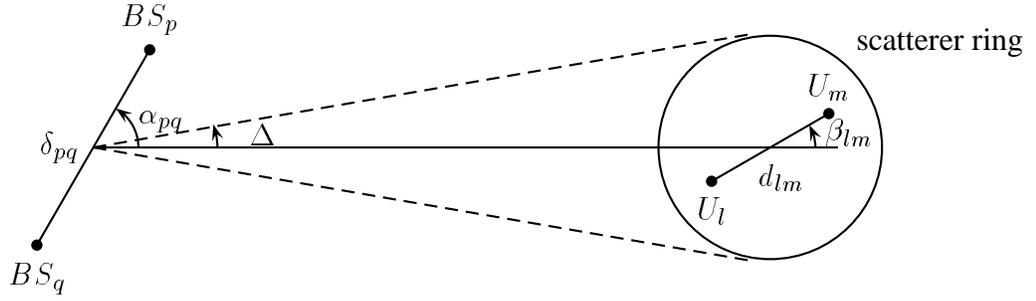


Figure 2.1. Geometric configuration of a  $(2,2)$  channel with local scatterers around the mobile user:  $BS_p$  is the  $p$ th antenna element at the base station,  $U_l$  is the  $l$ th antenna element at the mobile,  $\delta_{pq}$  is the antenna spacing in wavelengths between the  $p$ th and  $q$ th antenna elements at the base station,  $d_{lm}$  is the antenna spacing in wavelengths between the  $l$ th and  $m$ th antenna elements at the mobile, and  $\Delta$  is the angle spread.

## 2.4 Properties of Kronecker Product

Throughout this thesis, Kronecker product will be used extensively. Its definition and properties are summarized as follows [75].

- The Kronecker product of two matrices  $\mathbf{A}$  ( $m \times n$ ) and  $\mathbf{B}$  ( $p \times q$ ) is defined as

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

where  $a_{ij}$  and  $b_{ij}$  are the  $(i,j)$ th element of matrix  $\mathbf{A}$  and  $\mathbf{B}$ , respectively, and the dimension of  $\mathbf{A} \otimes \mathbf{B}$  is  $mp \times nq$ .

- For matrices  $\mathbf{A}(m \times n)$ ,  $\mathbf{B}(p \times q)$ ,  $\mathbf{C}(n \times r)$ , and  $\mathbf{D}(q \times s)$ ,

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}. \quad (2.17)$$

- For matrices  $\mathbf{A}(m \times n)$  and  $\mathbf{B}(p \times q)$ ,

$$(\mathbf{A} \otimes \mathbf{B})^\dagger = \mathbf{A}^\dagger \otimes \mathbf{B}^\dagger. \quad (2.18)$$

- For nonsingular square matrices  $\mathbf{A}(m \times m)$  and  $\mathbf{B}(n \times n)$ ,

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}. \quad (2.19)$$

- For square matrices  $\mathbf{A}(m \times m)$  and  $\mathbf{B}(n \times n)$ ,

$$\det(\mathbf{A} \otimes \mathbf{B}) = \det(\mathbf{A})^n \det(\mathbf{B})^m. \quad (2.20)$$

- If Hermitian matrices  $\mathbf{A}(m \times m)$  and  $\mathbf{B}(n \times n)$  can be eigenvalue-decomposed as  $\mathbf{A} = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^\dagger$  and  $\mathbf{B} = \mathbf{U}_2 \mathbf{\Lambda}_2 \mathbf{U}_2^\dagger$ , respectively, then  $\mathbf{A} \otimes \mathbf{B}$  can be eigenvalue-decomposed as

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{U}_1 \otimes \mathbf{U}_2)(\mathbf{\Lambda}_1 \otimes \mathbf{\Lambda}_2)(\mathbf{U}_1 \otimes \mathbf{U}_2)^\dagger. \quad (2.21)$$

## **Chapter 3**

# **Outage Probability Comparisons for Diversity Systems with Cochannel Interference in Rayleigh Fading**

### **3.1 Introduction**

Space diversity is an effective method to improve the performance of mobile radio systems. In space diversity, the received signals from all antenna branches are properly weighted and combined to combat fading, as well as cochannel interference (CCI) [107]. Three commonly used space diversity schemes are maximal ratio combining (MRC), equal gain combining (EGC), and selection combining (SC) [17]. For fading channels with only additive white Gaussian noise (AWGN) and no CCI, MRC is the optimal combining scheme; however, MRC carries high implementation complexity. The implementation costs for EGC and SC are much lower than that of MRC, but they are both suboptimal combining schemes in an AWGN environment. The optimal criterion is defined here in the sense of maximizing the signal-to-noise ratio (SNR) at the output of the diversity combiner, or equivalently, in the sense of minimizing bit error rate (BER). In the presence of cochannel

interfering signals at the receiving antennas, all aforementioned diversity schemes are, in fact, suboptimal. The optimal combining scheme in this scenario is called optimum combining (OC) [16, 120], which achieves the maximum signal-to-interference-plus-noise ratio (SINR) at the combiner output<sup>1</sup>. To fully implement OC, however, it is required to estimate the second-order statistics of interference and noise. In practical systems, for simplicity or due to the lack of knowledge of interference and noise statistics, suboptimal schemes, such as MRC, EGC and SC, may be used instead of OC. However, to the author's best knowledge, a comparative analysis of relative performance for these suboptimal schemes in fading and CCI has not been attempted. Such knowledge can be useful to better understand the design tradeoffs in practical cellular systems.

We assume that CCI is the primary source of system degradation [43]. Therefore, we ignore thermal noise in our analysis and consider an interference-limited environment [6, 21, 63, 102]. In an interference-limited environment, MRC, which maximizes output SNR and whose weights depend on noise powers on antenna branches [17], becomes invalid. Hence, we consider a variation of MRC, denoted as channel-matched combining (CMC), whose weights are given as the desired user's channel vector. In practical systems where diversity branches are usually assumed to have the same noise powers, MRC is reduced to CMC [7]. As a result, in this chapter, we provide a comparison study, both analytically and numerically, on the outage probability for CMC, EGC, and SC with CCI and flat Rayleigh fading in an interference-limited environment. Our analysis considers an arbitrary number of interferers, as well as arbitrary interferer power distributions.

Outage probability is an important performance criterion for mobile systems operating in the presence of interferers over fading channels. This criterion represents the probability

---

<sup>1</sup>For such a system, maximizing output SINR does not necessarily correspond to minimizing the BER unless the additive interference has a Gaussian distribution.

of unsatisfactory reception over the intended coverage area and can be used as a minimum quality of service requirement. In an interference-limited environment, the outage is defined as the event when the signal-to-interference ratio (SIR) at the combiner output drops below a threshold  $\beta$ , i.e.,  $P_{\text{OUT}} = \Pr\{\text{SIR} < \beta\}$ . Although bit or symbol error rates are more practical performance criteria, they are hard to calculate in some circumstances. There are quite a few papers on the analysis of average bit or symbol error rates of diversity systems under fading and CCI [2, 3, 67, 87, 101, 118]. However, most of these papers presented approximate analyses since CCI is explicitly assumed to be Gaussian. The exact calculation of bit or symbol error rates under CCI is, in general, very difficult. The difficulty arises from the fact that CCI may not be Gaussian. Therefore, we will study outage probability in our work.

In previous related work, Brennan [17], in his now classic paper, showed that, in the absence of CCI, the outage probability of MRC outperforms EGC and SC for an arbitrary fading distribution. The relative outage probability performances for EGC and SC, however, depend on the exact fading distribution. For Rayleigh fading, EGC has lower outage probability than that of SC. On the other hand, for a more disperse probability density function, the opposite is true. With the presence of CCI, an outage probability study for diversity systems is a completely different problem. This problem has received much interest in the past. In [26], Cui and Sheikh studied the outage probability of MRC with a small number of interfering signals in Rayleigh fading. This work was later extended by Aalo and Chayawan to an arbitrary number of interfering signals, for both equal and distinct interferer powers [1, 21]. More recently, Shah and Haimovich provided an alternative outage probability expression for CMC in Rayleigh fading with equal interferer powers [102]<sup>2</sup>.

---

<sup>2</sup>In [102], the combining scheme which the authors called MRC is really CMC.

In [6], Abu-Dayya and Beaulieu studied the EGC outage probability for an interference-limited environment with Nakagami- $m$  fading, a general model of fading amplitude which includes Rayleigh fading as the special case of  $m = 1$ . In that work, the interfering signal components are added incoherently across antenna array elements; hence, an approximation occurs in calculating the interfering power at the output of the combiner. Moreover, the analyses in [6] were restricted to equal interferer powers. In this work, we compute the outage probability for EGC using coherent interference power calculation (an exact analysis) over the diversity branches for both equal and distinct interferer powers. The outage probability for SC was studied by Sowerby and Williamson [106] in Rayleigh-distributed interference. Their work was later extended by Abu-Dayya and Beaulieu [6], Yao and Sheikh [125] to Nakagami- $m$  fading channels. The outage probability of OC can be found in [38, 39, 63, 87, 101, 118]

The main contributions of this chapter are: (1) we derive an exact outage probability expression for EGC in Rayleigh fading with multiple interferers by accurately calculating the interference power at the output of the combiner. Using this exact expression, we assess quantitatively the accuracy of the approximate interference power calculation used in [6]. (2) We provide a comparison study, both analytical and numerical, on the outage probability for CMC, EGC, and SC with CCI and flat Rayleigh fading in an interference-limited environment. In particular, we prove that CMC has a strictly lower outage probability than that of EGC, and that CMC has an outage probability no greater than that of SC. The relative performance of EGC and SC depends on factors such as the number of interferers and the interferer power distribution.

This chapter is organized as follows. In Section 3.2 we describe the system model which takes account of pulse shape, random delay of interfering signals, intersymbol interference, as well as both equal and distinct interferer powers. In Section 3.3, we derive new outage

probability expressions for EGC and CMC. In Section 3.4, we provide analytical outage probability comparisons for three diversity schemes. Numerical results are presented in Section 3.5.

## 3.2 System Model

The transmitted signals from the desired and the  $i$ th interfering users are, respectively,

$$s_s(t) = \sqrt{P_s T} \sum_{m=-\infty}^{+\infty} a_s[m] \tilde{g}(t - mT)$$

and

$$s_i(t) = \sqrt{P_i T} \sum_{m=-\infty}^{+\infty} a_i[m] \tilde{g}(t - mT)$$

where  $\tilde{g}(t)$  is the transmitter pulse response,  $1/T$  is the data transmission rate, and  $P_s$  and  $P_i$  are the transmitting powers of the desired and the  $i$ th interfering signals, respectively. The transmitter filter is assumed to have a square-root raised-cosine frequency response with a rolloff factor  $\rho$  ( $0 \leq \rho \leq 1$ ) [89]. The data symbols  $a_s[m]$  and  $a_i[m]$ 's are mutually independent with zero-mean and unit variance.

Here we show that the power of signal  $s_s(t)$  is  $P_s$ . The power spectrum density (PSD) of  $s_s(t)$  is [56, 89]

$$S(f) = P_s T \frac{1}{T} |\tilde{G}(f)|^2 S_a(f) = P_s |\tilde{G}(f)|^2$$

where  $\tilde{G}(f)$  is the Fourier transform of  $\tilde{g}(t)$ , and  $S_a(f)$  is the PSD of the data sequence  $a_s[m]$ ,  $m = -\infty, \dots, -1, 0, 1, \dots, \infty$ . Since the data symbols in the sequence are mutually independent and zero-mean with unit variance, it can be shown that  $S_a(f) = 1$ . Hence, the power of signal  $s_s(t)$  is

$$\int_{-\infty}^{\infty} S(f) df = P_s \int_{-\infty}^{\infty} |\tilde{G}(f)|^2 df = P_s$$

where the last equality comes from the fact that  $|\tilde{G}(f)|^2$  is a raised cosine with  $\int_{-\infty}^{\infty} |\tilde{G}(f)|^2 df = 1$ . Similarly, the transmit power of the  $i$ th interfering signal is  $P_i$ .

Ignoring thermal noise, the baseband received signal vector at an  $N_r$ -element receiver antenna array is

$$\mathbf{r}(t) = \sqrt{P_s T} \mathbf{c}_s \sum_{m=-\infty}^{+\infty} a_s[m] \tilde{g}(t - mT) + \sum_{i=1}^L \sqrt{P_i T} \mathbf{c}_i \sum_{m=-\infty}^{+\infty} a_i[m] \tilde{g}(t - mT - \tau_i) \quad (3.1)$$

where  $L$  is the number of interfering signals. The delay of the  $i$ th interfering signal relative to the desired user,  $\tau_i$ , is assumed to be uniformly distributed over the interval  $[0, T)$ . The channel vectors of the desired and the  $i$ th interfering users,  $\mathbf{c}_s$  and  $\mathbf{c}_i$ 's, are mutually independent. All channel vectors are assumed to be quasi-static (constant over a time frame [102]) and to have uncorrelated realizations in different frames. We further assume independent Rayleigh fading among diversity branches, i.e., the elements of  $\mathbf{c}_s$  and  $\mathbf{c}_i$  are i.i.d. circularly symmetric complex Gaussian random variables (RVs) with zero-mean and unit variance.

Passing  $\mathbf{r}(t)$  in (3.1) through a filter matched to  $\tilde{g}(t)$ , we have

$$\mathbf{r}_{\text{MF}}(t) = \sqrt{P_s T} \mathbf{c}_s \sum_{m=-\infty}^{+\infty} a_s[m] g(t - mT) + \sum_{i=1}^L \sqrt{P_i T} \mathbf{c}_i \sum_{m=-\infty}^{+\infty} a_i[m] g(t - mT - \tau_i) \quad (3.2)$$

where  $g(t) = \tilde{g}(t) * \tilde{g}(t)$  and  $*$  denotes convolution. Here,  $g(t)$  is a Nyquist pulse with a raised cosine spectrum and rolloff factor  $\rho$ .

Assuming perfect synchronization for the desired user, sampling the output of the receiver matched filter at  $t = nT$ , we obtain

$$\mathbf{r}[n] = \sqrt{P_s T} \mathbf{c}_s a_s[n] + \underbrace{\sum_{i=1}^L \sqrt{P_i T} \mathbf{c}_i \sum_{m=-\infty}^{+\infty} a_i[m] g(nT - mT - \tau_i)}_{z_i[n]} \quad (3.3)$$

where  $z_i[n]$  is the signal intersymbol interference (ISI) from the  $i$ th interferer. Note that no ISI exists for the desired user since  $g(t)$  is a Nyquist pulse. However, ISI exists for the

interferers due to delays. Since the zero-mean data symbols from different interferers are mutually independent, we have

$$E\{z_i[n]\} = 0$$

and

$$E\{z_i[n]z_j^*[n]\} = 0 \text{ for } i \neq j,$$

i.e.,  $z_i[n]$  and  $z_j[n]$  are uncorrelated for  $i \neq j$ . In Appendix A we show that the variance of  $z_i[n]$  is<sup>3</sup>

$$E\{|z_i[n]|^2\} = 1 - \rho/4.$$

We express, component-wise, the desired and the interfering channel vectors as

$$\mathbf{c}_s = \left[ \alpha_{s,1} e^{j\theta_{s,1}} \dots \alpha_{s,N_r} e^{j\theta_{s,N_r}} \right]^T$$

and

$$\mathbf{c}_i = \left[ \alpha_{i,1} e^{j\theta_{i,1}} \dots \alpha_{i,N_r} e^{j\theta_{i,N_r}} \right]^T.$$

The phase for the desired user channel,  $\theta_{s,j}$ , and the phase for the interfering user channel,  $\theta_{i,j}$ , are uniformly distributed over  $[0, 2\pi)$ . The fading amplitudes  $\alpha_{s,j}$  and  $\alpha_{i,j}$  are Rayleigh-distributed as

$$f_\alpha(\alpha) = 2\alpha e^{-\alpha^2}, \quad \alpha \geq 0.$$

### 3.3 Outage Probabilities of EGC, CMC and SC with CCI

In this section we derive an exact outage probability expression for EGC with cochannel interference in Rayleigh fading. In the case for CMC, a new alternative outage probability

---

<sup>3</sup>In [11], it was stated that  $E\{|z_i[n]|^2\} = 1 - \rho/4$  without showing it.

expression is derived. We will show that this new expression is more suitable for analytical comparison. The outage probability of SC is presented for completeness.

### 3.3.1 EGC Outage Probability

In equal gain combining, the outputs of all the branches are co-phased (with respect to the desired user signal) and weighted equally. The combining weight vector of EGC is  $\mathbf{w}_{\text{EGC}} = [e^{j\theta_{s,1}} \dots e^{j\theta_{s,N_r}}]^T$ , and the output of the combiner becomes

$$\begin{aligned} \mathbf{w}_{\text{EGC}}^\dagger \mathbf{r}[n] &= \sqrt{P_s T} (\mathbf{w}_{\text{EGC}}^\dagger \mathbf{c}_s) a_s[n] + \sum_{i=1}^L \sqrt{P_i T} (\mathbf{w}_{\text{EGC}}^\dagger \mathbf{c}_i) z_i[n] \\ &= \sqrt{P_s T} \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right) a_s[n] + \sum_{i=1}^L \sqrt{P_i T} \left( \sum_{j=1}^{N_r} \underbrace{\alpha_{i,j} e^{j(\theta_{i,j} - \theta_{s,j})}}_{g_{i,j}} \right) z_i[n]. \end{aligned} \quad (3.4)$$

It can be shown that  $(\theta_{i,j} - \theta_{s,j}) \bmod 2\pi$  is uniformly distributed over  $[0, 2\pi)$  and is independent of  $\alpha_{i,j}$ . Since  $\alpha_{i,j}$  is Rayleigh-distributed,  $g_{i,j}$  is circularly symmetric complex Gaussian with zero-mean and unit variance.

Since  $z_i[n]$  and  $z_j[n]$  are uncorrelated for  $i \neq j$ , the total interference power at the combiner output is obtained by adding interference powers from different interferers. For each interferer, interference from different antennas can combine either incoherently (see [6], Eqn. (8b)) or coherently. In the incoherent case, to compute the  $i$ th interferer's power, the channel amplitude of each diversity branch is first squared, and all branches are then summed, i.e.,  $\sum_{j=1}^{N_r} \alpha_{i,j}^2$ . If the interfering signals arriving at different antennas are mutually uncorrelated, the incoherent calculation is exact. However, these interfering signals are, in general, correlated. Thus, the incoherent calculation is only an approximation. In the *coherent* interference power calculation, *phasor* addition of each interfering signal is employed. Hence,  $\alpha_{i,j} e^{j(\theta_{i,j} - \theta_{s,j})}$ , are added first, and then squared, i.e.,  $\left| \sum_{j=1}^{N_r} g_{i,j} \right|^2$  in

(3.4). Coherent interference power calculation is an improved model of EGC over incoherent interference power calculation. Numerical results in Section 3.5 demonstrate cases where the two interference power calculation methods lead to significantly different outage probabilities.

The instantaneous SIR at the output of EGC, assuming coherent interference power calculation over the diversity branches, is

$$\text{SIR}_{\text{EGC}} = \frac{P_s \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2}{(1 - \rho/4) \sum_{i=1}^L P_i \left| \sum_{j=1}^{N_r} g_{i,j} \right|^2} = \frac{\left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2}{(1 - \rho/4) \sum_{i=1}^L \mu_i / \Lambda_i} \quad (3.5)$$

where  $\Lambda_i = P_s/P_i$ , for  $i = 1, \dots, L$ , is the power ratio of the desired signal to the  $i$ th interfering signal, and  $\mu_i = \left| \sum_{j=1}^{N_r} g_{i,j} \right|^2$ . Here,  $g_{i,1}, \dots, g_{i,N_r}$  are i.i.d. circularly symmetric complex Gaussian RVs with zero-mean and unit variance, thus,  $\sum_{j=1}^{N_r} g_{i,j}$  is also a circularly symmetric complex Gaussian RV with mean zero and variance  $N_r$ . It can be shown that  $\mu_i$  is exponentially distributed with mean  $N_r$  and its PDF is given by [89]

$$f_{\mu}(\mu) = \frac{1}{N_r} e^{-\mu/N_r}, \mu \geq 0. \quad (3.6)$$

We further note that the denominator and the numerator in (3.5) are independent. This is due to the independence assumption between the channel vectors for the desired and the interfering signals. This independence property can simplify the ensuing outage probability analyses.

Letting  $X \triangleq \sum_{j=1}^{N_r} \alpha_{s,j}$  and  $U \triangleq \sum_{i=1}^L \mu_i / \Lambda_i$  in (3.5), the output SIR of EGC can be rewritten as  $\text{SIR}_{\text{EGC}} = \frac{X^2}{(1-\rho/4)U}$ . The outage probability is expressed as

$$\begin{aligned} P_{\text{OUT,EGC}}(\beta) &= \Pr(\text{SIR}_{\text{EGC}} < \beta) \\ &= E_U \left\{ \Pr \left( X < \sqrt{\beta_0 U} \mid U \right) \right\} \\ &= E_U \left\{ \Pr \left( X < \sqrt{\beta_0 U} \right) \right\} \end{aligned} \quad (3.7)$$

where

$$\beta_0 = \left(1 - \frac{\rho}{4}\right) \beta \quad (3.8)$$

and (3.7) comes from the fact that  $X$  and  $U$  are independent.

Computation of the outage probability in (3.7) requires the knowledge of the cumulative distribution function (CDF) of  $X$ . We recall that  $X$  is a sum of  $N_r$  i.i.d. Rayleigh RVs and no known closed-form expression exists except for  $N_r = 2$ . In [10] Beaulieu derived an infinite series for the CDF of a sum of independent RVs. Essentially, this infinite series is a Fourier series. In [115] an alternative derivation was given which provided insights into the uses and limitations of the Beaulieu series. It can be shown that the Fourier series representation of CDF of  $X$  is [10, 115]

$$\Pr(X < x) = \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x} + \Delta \quad (3.9)$$

where  $\omega_0 = \frac{2\pi}{T_0}$ ,  $T_0$  is a parameter that controls the accuracy of result [10],  $\phi_X(\omega)$  is the characteristic function of  $X$ , and  $\Delta$  is an error term which tends to zero for large  $T_0$ .

Assuming  $T_0$  is large, we omit the error term  $\Delta$  in the following analysis.

The conditional outage probability in (3.7) hence can be expressed as

$$\begin{aligned} \Pr\left(X < \sqrt{\beta_0 U}\right) &= \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 \sqrt{\beta_0 U}} \\ &= \frac{1}{2} - \sum_{\substack{n=1 \\ n \text{ odd}}}^{+\infty} \frac{2 \Im \left\{ e^{-jn\omega_0 \sqrt{\beta_0 U}} \phi_X(n\omega_0) \right\}}{n\pi}, \end{aligned} \quad (3.10)$$

and the outage probability in (3.7) can be expressed as

$$P_{\text{OUT,EGC}}(\beta) = \frac{1}{2} - \sum_{\substack{n=1 \\ n \text{ odd}}}^{+\infty} \frac{2 \Im \left\{ E_U \left\{ e^{-jn\omega_0 \sqrt{\beta_0 U}} \right\} \phi_X(n\omega_0) \right\}}{n\pi}. \quad (3.11)$$

In obtaining (3.11), we have used

$$\begin{aligned} & \int_0^\infty \left[ \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0\sqrt{\beta_0}u} \right] f_U(u) du \\ &= \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \int_0^\infty \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0\sqrt{\beta_0}u} f_U(u) du \end{aligned} \quad (3.12)$$

where  $f_U(u)$  is the PDF of  $U$ , i.e., we have interchanged the integration and summation.

To justify this interchange, we introduce the following theorem [64, Theorem 15, p.423].

Let  $\{\phi_n(x)\}$  be a complete orthonormal system for the interval  $a \leq x \leq b$ . Let  $f(x)$  be piecewise continuous for  $a \leq x \leq b$  and let  $g(x)$  be piecewise continuous for  $x_1 \leq x \leq x_2$ , where  $a \leq x_1 < x_2 \leq b$ . Let  $\sum c_n \phi_n(x)$  be the Fourier series of  $f(x)$  with respect to  $\{\phi_n(x)\}$ . Then  $\int_{x_1}^{x_2} f(x)g(x)dx = \sum c_n \int_{x_1}^{x_2} g(x)\phi_n(x)dx$ .

In our case, using the substitution  $\sqrt{\beta_0}u = x$ , the left-hand side of (3.12) becomes

$$\begin{aligned} & \int_0^\infty \left[ \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x} \right] f_U\left(\frac{x^2}{\beta_0}\right) \frac{2x}{\beta_0} dx \\ &= \sum_{i=0}^\infty \int_{iT_0}^{(i+1)T_0} \left[ \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x} \right] f_U\left(\frac{x^2}{\beta_0}\right) \frac{2x}{\beta_0} dx \end{aligned} \quad (3.13)$$

Since  $\{e^{-jn\omega_0 x}\}$  ( $n = 0, \pm 1, \pm 2, \dots$ ) is a complete orthonormal system for  $iT_0 \leq x \leq (i+1)T_0$ , the functions  $\frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x}$  and  $f_U\left(\frac{x^2}{\beta_0}\right) \frac{2x}{\beta_0}$  are continuous for interval  $iT_0 \leq x \leq (i+1)T_0$ , and  $\frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x}$  is a Fourier series, according to the above introduced theorem, it is clear that we can interchange the summation in

Fourier series and the integration. Hence (3.13) becomes

$$\begin{aligned} & \sum_{i=0}^\infty \left[ \int_{iT_0}^{(i+1)T_0} \frac{1}{2} f_U\left(\frac{x^2}{\beta_0}\right) \frac{2x}{\beta_0} dx - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \int_{iT_0}^{(i+1)T_0} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0 x} f_U\left(\frac{x^2}{\beta_0}\right) \frac{2x}{\beta_0} dx \right] \\ &= \sum_{i=0}^\infty \left[ \int_{\frac{(iT_0)^2}{\beta_0}}^{\frac{[(i+1)T_0]^2}{\beta_0}} \frac{1}{2} f_U(u) du - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \int_{\frac{(iT_0)^2}{\beta_0}}^{\frac{[(i+1)T_0]^2}{\beta_0}} \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0\sqrt{\beta_0}u} f_U(u) du \right] \end{aligned} \quad (3.14)$$

where (3.14) is obtained through the substitution  $u = x^2/\beta_0$ . Since the series in the square brackets in (3.14) converges to  $\int_{(iT_0)^2/\beta_0}^{[(i+1)T_0]^2/\beta_0} \Pr(X < \sqrt{\beta_0 u}) f_U(u) du$ , (3.14) becomes

$$\begin{aligned} & \int_0^\infty \frac{1}{2} f_U(u) du - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \int_0^\infty \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0\sqrt{\beta_0 u}} f_U(u) du \\ &= \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{+\infty} \int_0^\infty \frac{\phi_X(n\omega_0)}{n\pi j} e^{-jn\omega_0\sqrt{\beta_0 u}} f_U(u) du \end{aligned}$$

which is the right-hand side of (3.12). Therefore, (3.12) is valid.

Now back to the expression of outage probability in (3.11). It can be shown that the characteristic function of  $X$  in (3.11) is [89, Eqn. (2-1-133)]<sup>4</sup>

$$\phi_X(\omega) = \left[ {}_1F_1\left(1; \frac{1}{2}; -\frac{\omega^2}{4}\right) + j\frac{\sqrt{\pi}}{2}\omega e^{-\omega^2/4} \right]^{N_r} \quad (3.15)$$

where  ${}_1F_1(a; b; z)$  is the degenerate hypergeometric function defined as [5]

$${}_1F_1(a; b; z) = 1 + \frac{az}{b} + \frac{(a)_2 z^2}{(b)_2 2!} + \cdots + \frac{(a)_n z^n}{(b)_n n!} + \cdots$$

and

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1), \quad (a)_0 = 1. \quad (3.16)$$

We recall that  $U$  is a weighted sum of  $L$  i.i.d. exponential RVs. The PDF of  $U$ , in the case of equal interferer powers,  $\Lambda_1 = \cdots = \Lambda_L = \Lambda$ , is given by [89, Eqn.(14-4-13)]

$$f_U(u) = \frac{1}{(L-1)! \left(\frac{N_r}{\Lambda}\right)^L} u^{L-1} e^{-\frac{\Lambda}{N_r} u}, \quad u \geq 0, \quad (3.17a)$$

and in the case of distinct interferer powers,  $\Lambda_i \neq \Lambda_j$  for  $i \neq j$ , is given by [89, Eqn.(14-5-26)]

$$f_U(u) = \sum_{k=1}^L \frac{\Lambda_k}{N_r} \pi_k e^{-\frac{\Lambda_k}{N_r} u}, \quad u \geq 0 \quad (3.17b)$$

---

<sup>4</sup>In [89, Eqn. 2-1-133], a minor typo needs to be corrected, i.e.,  $j\sqrt{\pi/2}v\sigma^2 e^{-v^2\sigma^2/2}$  should be  $j\sqrt{\pi/2}v\sigma e^{-v^2\sigma^2/2}$ .

where

$$\pi_k = \prod_{\substack{i=1 \\ i \neq k}}^L \frac{\Lambda_i}{\Lambda_i - \Lambda_k}. \quad (3.18)$$

In (3.11), we rewrite

$$E_U \left\{ e^{-jn\omega_0 \sqrt{\beta_0 U}} \right\} = \int_0^\infty \cos(n\omega_0 \sqrt{\beta_0 U}) f_U(u) du - j \int_0^\infty \sin(n\omega_0 \sqrt{\beta_0 U}) f_U(u) du. \quad (3.19)$$

In the case of equal interferer powers, by substituting (3.17a) into (3.19), we have

$$\begin{aligned} & E_U \left\{ e^{-jn\omega_0 \sqrt{\beta_0 U}} \right\} \\ &= \frac{1}{(L-1)! \left(\frac{\Lambda}{N_r}\right)^L} \left\{ \int_0^\infty \cos(n\omega_0 \sqrt{\beta_0 u}) u^{L-1} e^{-\frac{\Lambda}{N_r} u} du \right. \\ &\quad \left. - j \int_0^\infty \sin(n\omega_0 \sqrt{\beta_0 u}) u^{L-1} e^{-\frac{\Lambda}{N_r} u} du \right\} \\ &= \frac{2}{(L-1)! \left(\frac{\Lambda}{N_r}\right)^L} \left\{ \int_0^\infty \cos(n\omega_0 \sqrt{\beta_0 x}) x^{2L-1} e^{-\frac{\Lambda}{N_r} x^2} dx \right. \\ &\quad \left. - j \int_0^\infty \sin(n\omega_0 \sqrt{\beta_0 x}) x^{2L-1} e^{-\frac{\Lambda}{N_r} x^2} dx \right\} \\ &= {}_1F_1 \left( L; \frac{1}{2}; -\frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r \right) - j \frac{\Gamma(L + \frac{1}{2})}{\Gamma(L)} n\omega_0 \sqrt{\frac{\beta_0 N_r}{\Lambda}} e^{-\frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r} \\ &\quad \times {}_1F_1 \left( 1-L; \frac{3}{2}; \frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r \right) \Big\} \\ &= e^{-\frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r} \left\{ {}_1F_1 \left( \frac{1}{2} - L; \frac{1}{2}; \frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r \right) \right. \\ &\quad \left. - j \frac{\Gamma(L + \frac{1}{2})}{\Gamma(L)} n\omega_0 \sqrt{\frac{\beta_0 N_r}{\Lambda}} {}_1F_1 \left( 1-L; \frac{3}{2}; \frac{n^2 \omega_0^2}{4\Lambda} \beta_0 N_r \right) \right\} \\ &= \frac{2^L}{\sqrt{\pi}} \Gamma \left( L + \frac{1}{2} \right) e^{-\frac{n^2 \omega_0^2}{8\Lambda} \beta_0 N_r} D_{-2L} \left( jn\omega_0 \sqrt{\frac{\beta_0 N_r}{2\Lambda}} \right) \end{aligned} \quad (3.20a)$$

where the second equality follows from the substitution  $x = \sqrt{u}$ ; the third equality follows from [51, Eqn. 3.952(7), (8)]; the fourth equality follows from Kummer transformation [5]

${}_1F_1(a, b, z) = e^z {}_1F_1(b - a, b, -z)$ . To express the result in a compact form, the last equality follows from [51, Eqn. 9.240] where  $D_p(z)$  is the parabolic cylinder function. We use Kummer transformation since  $e^{-\frac{n^2\omega_0^2}{4\Lambda}\beta_0 N_r} {}_1F_1\left(\frac{1}{2} - L; \frac{1}{2}; \frac{n^2\omega_0^2}{4\Lambda}\beta_0 N_r\right)$  converges much more rapidly than  ${}_1F_1\left(L; \frac{1}{2}; -\frac{n^2\omega_0^2}{4\Lambda}\beta_0 N_r\right)$  in numerical calculation.

Similarly, in the case of distinct interferer powers, by substituting (3.17b) into (3.19), we have

$$\begin{aligned}
& E_U \left\{ e^{-jn\omega_0\sqrt{\beta_0 U}} \right\} \\
&= \sum_{k=1}^L \frac{\Lambda_k}{N_r} \pi_k \left\{ \int_0^\infty \cos(n\omega_0\sqrt{\beta_0 u}) e^{-\frac{\Lambda_k}{N_r}u} du - j \int_0^\infty \sin(n\omega_0\sqrt{\beta_0 u}) e^{-\frac{\Lambda_k}{N_r}u} du \right\} \\
&= \sum_{k=1}^L \frac{2\Lambda_k}{N_r} \pi_k \left\{ \int_0^\infty \cos(n\omega_0\sqrt{\beta_0 x}) e^{-\frac{\Lambda_k}{N_r}x^2} x dx - j \int_0^\infty \sin(n\omega_0\sqrt{\beta_0 x}) e^{-\frac{\Lambda_k}{N_r}x^2} x dx \right\} \\
&= \sum_{k=1}^L \pi_k e^{-\frac{n^2\omega_0^2}{4\Lambda_k}\beta_0 N_r} \left\{ {}_1F_1\left(-\frac{1}{2}; \frac{1}{2}; \frac{n^2\omega_0^2}{4\Lambda_k}\beta_0 N_r\right) - j \frac{\sqrt{\pi}}{2} n\omega_0 \sqrt{\frac{\beta_0 N_r}{\Lambda_k}} \right\} \\
&= \sum_{k=1}^L \pi_k e^{-\frac{n^2\omega_0^2}{8\Lambda_k}\beta_0 N_r} D_{-2}\left(jn\omega_0\sqrt{\frac{\beta_0 N_r}{2\Lambda_k}}\right). \tag{3.20b}
\end{aligned}$$

where the second equality follows from the substitution  $x = \sqrt{u}$ ; the third equality follows from [51, Eqn. 3.952(7), (8)]; the last equality follows from [51, Eqn. 9.240].

Substitution of (3.15) and (3.20) into (3.11) yields the outage probability of EGC for both equal and distinct interferer powers.

### 3.3.1.1 EGC outage probability: case of $N_r = 2$

The CDF of a sum of two i.i.d Rayleigh RVs is known [17] [53]. For  $N_r = 2$ , the conditional outage probability in (3.7) is

$$\Pr\left(X < \sqrt{\beta_0 U}\right) = 1 - e^{-\beta_0 U} - \sqrt{\frac{\pi}{2}} \beta_0 U e^{-\frac{1}{2}\beta_0 U} \operatorname{erf}\left(\sqrt{\frac{\beta_0 U}{2}}\right) \tag{3.21}$$

where the error function  $\operatorname{erf}(\cdot)$  is defined as  $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-x^2} dx$ .

In the case of equal interferer powers, averaging (3.21) with respect to (3.17a), we have

$$\begin{aligned}
P_{\text{OUT,EGC,N2=2}}(\beta) &= 1 - \frac{1}{(L-1)!} \left(\frac{\Lambda}{2}\right)^L \int_0^\infty e^{-(\beta_0 + \frac{\Lambda}{2})u} u^{L-1} du \\
&\quad - \frac{1}{(L-1)!} \left(\frac{\Lambda}{2}\right)^L \int_0^\infty \sqrt{\frac{\pi}{2}\beta_0 u} e^{-\frac{\beta_0 + \Lambda}{2}u} u^{L-1} \text{erf}\left(\sqrt{\frac{\beta_0 U}{2}}\right) du \\
&= 1 - \left(\frac{\Lambda}{2\beta_0 + \Lambda}\right)^L - \frac{1}{(L-1)!} \left(\frac{\Lambda}{2}\right)^L \sqrt{2\pi\beta_0} \\
&\quad \times \int_0^\infty \text{erf}\left(\sqrt{\frac{\beta_0}{2}}x\right) x^{2L} e^{-\frac{\beta_0 + \Lambda}{2}x^2} dx \\
&= 1 - \left(\frac{\Lambda}{2\beta_0 + \Lambda}\right)^L - \frac{\Gamma(L + \frac{1}{2})}{\Gamma(L)} \sqrt{\frac{\pi\beta_0}{\beta_0 + \Lambda}} \left(\frac{\Lambda}{\beta_0 + \Lambda}\right)^L \\
&\quad + \frac{2L}{2L+1} \left(\frac{\Lambda}{\beta_0}\right)^L {}_2F_1\left(L + \frac{1}{2}, L + 1; L + \frac{3}{2}; -\frac{\Lambda + \beta_0}{\beta_0}\right) \tag{3.22a}
\end{aligned}$$

where the second equality follows from [51, Eqn. 3.351(3)] and the substitution  $x = \sqrt{u}$ ; the third equality follows from [51, Eqn. 3.478(1), 6.286(1)] where  ${}_2F_1(a, b; c; z)$  is the hypergeometric function [5] defined as

$${}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$$

and  $(a)_n$  is defined as (3.16).

Similarly, in the case of distinct interferer powers, averaging (3.21) with respect to (3.17b), we have

$$\begin{aligned}
P_{\text{OUT,EGC,N2=2}}(\beta) &= 1 - \sum_{k=1}^L \frac{\Lambda_k}{2} \pi_k \int_0^\infty e^{-(\frac{\Lambda_k}{2} + \beta_0)u} du \\
&\quad - \sum_{k=1}^L \frac{\Lambda_k}{2} \pi_k \int_0^\infty \sqrt{\frac{\pi}{2}\beta_0 u} e^{-\frac{\beta_0 + \Lambda_k}{2}u} \text{erf}\left(\sqrt{\frac{\beta_0 U}{2}}\right) du
\end{aligned}$$

$$\begin{aligned}
&= 1 - \sum_{k=1}^L \pi_k \frac{\Lambda_k}{\Lambda_k + 2\beta_0} - \sum_{k=1}^L \Lambda_k \pi_k \sqrt{\pi} \frac{2}{\beta_0} \int_0^\infty \operatorname{erf}(x) x^2 e^{-\frac{\Lambda_k + \beta_0}{\beta_0} x^2} dx \\
&= 1 + \sum_{k=1}^L \Lambda_k \pi_k \left[ \sqrt{\beta_0} (\Lambda_k + \beta_0)^{-3/2} \arctan \left( \sqrt{1 + \frac{\Lambda_k}{\beta_0}} \right) \right. \\
&\quad \left. - \frac{1}{\Lambda_k + \beta_0} - \frac{\pi}{2} \sqrt{\beta_0} (\Lambda_k + \beta_0)^{-3/2} \right] \tag{3.22b}
\end{aligned}$$

where the second equality follows from the substitution  $x = \sqrt{\beta_0 u/2}$ , and the third equality follows from [51, Eqn. 3.478(1), 6.292].

### 3.3.2 CMC Outage Probability

In channel-matched combining, the desired user signal at all diversity branches are co-phased and weighted according to the desired user channel amplitudes. The combining weight vector is thus  $\mathbf{w}_{\text{CMC}} = \mathbf{c}_s$ , and the signal at the output of CMC becomes

$$\mathbf{w}_{\text{CMC}}^\dagger \mathbf{r}[n] = \sqrt{P_s T} (\mathbf{c}_s^\dagger \mathbf{c}_s) a_s[n] + \sum_{i=1}^L \sqrt{P_i T} (\mathbf{c}_s^\dagger \mathbf{c}_i) z_i[n].$$

The SIR at the output of CMC is

$$\text{SIR}_{\text{CMC}} = \frac{P_s |\mathbf{c}_s^\dagger \mathbf{c}_s|^2}{(1 - \rho/4) \sum_{i=1}^L P_i |\mathbf{c}_s^\dagger \mathbf{c}_i|^2} = \frac{|\mathbf{c}_s|^2}{(1 - \rho/4) \sum_{i=1}^L \frac{1}{\Lambda_i} \frac{|\mathbf{c}_s^\dagger \mathbf{c}_i|^2}{|\mathbf{c}_s|^2}} \tag{3.23}$$

where  $|\mathbf{c}_s| = \sqrt{\mathbf{c}_s^\dagger \mathbf{c}_s}$ . In [102], it was shown that  $\frac{\mathbf{c}_s^\dagger \mathbf{c}_i}{|\mathbf{c}_s|}$  is a zero-mean complex Gaussian RV with unit variance, and it is independent of  $\mathbf{c}_s$ . In Appendix B, it is further proved that  $\frac{\mathbf{c}_s^\dagger \mathbf{c}_i}{|\mathbf{c}_s|}$  is circularly symmetric. By denoting  $\eta_i$  as  $\frac{|\mathbf{c}_s^\dagger \mathbf{c}_i|^2}{|\mathbf{c}_s|^2}$ , we can rewrite the  $\text{SIR}_{\text{CMC}}$  as

$$\text{SIR}_{\text{CMC}} = \frac{\sum_{j=1}^{N_r} \alpha_{s,j}^2}{(1 - \rho/4) \sum_{i=1}^L \eta_i / \Lambda_i} \tag{3.24}$$

where  $\eta_i$  is exponentially distributed with unit mean. Since  $\frac{\mathbf{c}_s^\dagger \mathbf{c}_i}{|\mathbf{c}_s|}$  is independent of  $\mathbf{c}_s$ , the denominator and the numerator in (3.24) are, in fact, independent.

Let  $X_1 = \sum_{j=1}^{N_r} \alpha_{s,j}^2$  and  $U_1 = \sum_{i=1}^L \eta_i / \Lambda_i$ . Since  $X_1$  and  $U_1$  are independent, the outage probability of CMC with the outage threshold  $\beta$  can be calculated as

$$P_{\text{OUT,CMC}}(\beta) = P\left(\frac{X_1}{U_1} < \beta_0\right) = \int_0^\infty f_{X_1}(x_1) \int_{x_1/\beta_0}^\infty f_{U_1}(u_1) du_1 dx_1 \quad (3.25)$$

where  $f_{X_1}(x_1)$  and  $f_{U_1}(u_1)$  are the PDFs of  $X_1$  and  $U_1$ , respectively. Since  $X_1$  is chi-square distributed, we have [89]

$$f_{X_1}(x_1) = \frac{1}{\Gamma(N_r)} x_1^{N_r-1} e^{-x_1}, \quad x_1 > 0.$$

As  $U_1$  is a weighted sum of  $L$  i.i.d. exponential RVs, we have, in the case of equal interferer powers [89, (14-4-13)],

$$f_{U_1}(u_1) = \frac{\Lambda^L}{(L-1)!} u_1^{L-1} e^{-\Lambda u_1}, \quad u_1 > 0 \quad (3.26a)$$

and for the case of distinct interferer powers [89, (14-5-26)],

$$f_{U_1}(u_1) = \sum_{k=1}^L \Lambda_k \pi_k e^{-\Lambda_k u_1}, \quad u_1 > 0 \quad (3.26b)$$

where  $\pi_k$  is defined in (3.18).

In the case of equal interferer powers, using [51, 3.351(2)], we have

$$\int_{x_1/\beta_0}^\infty f_{U_1}(u_1) du_1 = \sum_{k=0}^{L-1} \frac{1}{k!} \left(\frac{\Lambda}{\beta_0}\right)^k x_1^k e^{-\frac{\Lambda}{\beta_0} x_1}; \quad (3.27a)$$

and in the case of distinct interferer powers, we have

$$\int_{x_1/\beta_0}^\infty f_{U_1}(u_1) du_1 = \sum_{k=1}^L \pi_k e^{-\frac{\Lambda_k}{\beta_0} x_1}. \quad (3.27b)$$

Substituting (3.27) into (3.25) and using [51, 3.351(3)], in the case of equal interferer powers, the outage probability of CMC is

$$P_{\text{OUT,CMC}}(\beta) = \left(\frac{\beta_0}{\beta_0 + \Lambda}\right)^{N_r} \sum_{k=0}^{L-1} \frac{(k + N_r - 1)!}{k!(N_r - 1)!} \left(\frac{\Lambda}{\beta_0 + \Lambda}\right)^k; \quad (3.28a)$$

and in the case of distinct interferer powers,

$$P_{\text{OUT,CMC}}(\beta) = \sum_{k=1}^L \pi_k \left( \frac{\beta_0}{\beta_0 + \Lambda_k} \right)^{N_r}. \quad (3.28b)$$

It can be verified that (3.28) is numerically equivalent to the outage expressions derived by Aalo and Chayawan [1, (13)-(14)] and another alternative expression derived by Shah and Haimovich [102, (43)]. However, as shown in Section 3.4, our present CMC outage probability expressions are more suitable for analytical outage probability comparison.

### 3.3.3 SC Outage Probability

Selection combining chooses the branch with the largest SIR. Hence, the outage probability of SC can be expressed as [17]

$$P_{\text{OUT,SC}}(\beta) = P(\text{SIR}_{\text{SC},1} < \beta, \dots, \text{SIR}_{\text{SC},N_r} < \beta) \quad (3.29)$$

where  $\text{SIR}_{\text{SC},i}$  is the SIR for the  $i$ th receiving antenna. Since  $\text{SIR}_{\text{SC},1}, \text{SIR}_{\text{SC},2}, \dots, \text{SIR}_{\text{SC},N_r}$  are i.i.d. RVs, we have

$$P_{\text{OUT,SC}}(\beta) = [P(\text{SIR}_{\text{SC},1} < \beta)]^{N_r}. \quad (3.30)$$

Without loss of generality, we consider the first antenna branch and write

$$\text{SIR}_{\text{SC},1} = \frac{\alpha_{s,1}^2}{(1 - \rho/4) \sum_{i=1}^L \alpha_{i,1}^2 / \Lambda_i} \quad (3.31)$$

where  $\alpha_{s,j}^2$  and  $\alpha_{i,j}^2$  are exponentially distributed with unit mean. The outage probability of SC can be obtained from [6], [106]<sup>5</sup>, [125], for the case of equal interferer powers, as

$$P_{\text{OUT,SC}}(\beta) = \left[ 1 - \left( \frac{\Lambda}{\beta_0 + \Lambda} \right)^L \right]^{N_r}, \quad (3.32a)$$

---

<sup>5</sup>Note that the SC outage probability expression (Eqn. (1)) presented in [106] is only valid for equal interference powers.

and, for the case of distinct interferer powers, as

$$P_{\text{OUT,SC}}(\beta) = \left[ \sum_{k=1}^L \pi_k \frac{\beta_0}{\beta_0 + \Lambda_k} \right]^{N_r}. \quad (3.32b)$$

## 3.4 Analytical Outage Probability Comparisons

### 3.4.1 Outage Probability Comparison for CMC and EGC

In this section, we use two methods to show that CMC has a strictly lower outage probability than that of EGC. We first rewrite the output SIR expression of CMC in (3.24) as

$$\text{SIR}_{\text{CMC}} = \frac{N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2}{(1 - \rho/4) \sum_{i=1}^L N_r \eta_i / \Lambda_i} = \frac{N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2}{(1 - \rho/4) \sum_{i=1}^L \nu_i / \Lambda_i} \quad (3.33)$$

where  $\nu_i = N_r \eta_i$ . Since  $\eta_i$  is exponentially distributed with unit mean,  $\nu_i$  is exponentially distributed with mean  $N_r$ . Comparing (3.33) with (3.5), we immediately recognize that the denominators  $\xi_1 \triangleq (1 - \rho/4) \sum_{i=1}^L \nu_i / \Lambda_i$  in (3.33) and  $\xi_2 \triangleq (1 - \rho/4) \sum_{i=1}^L \mu_i / \Lambda_i$  in (3.5) have the same distribution. We write the outage probabilities as

$$P_{\text{OUT,CMC}}(\beta) = \Pr \left\{ N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2 / \xi_1 < \beta \right\} = \int \Pr \left\{ \sum_{j=1}^{N_r} \alpha_{s,j}^2 < \beta \xi / N_r \right\} f_{\xi_1}(\xi) d\xi \quad (3.34)$$

and

$$P_{\text{OUT,EGC}}(\beta) = \Pr \left\{ \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2 / \xi_2 < \beta \right\} = \int \Pr \left\{ \sum_{j=1}^{N_r} \alpha_{s,j} < \sqrt{\beta \xi} \right\} f_{\xi_2}(\xi) d\xi \quad (3.35)$$

since the denominator and numerator are independent in (3.33) and (3.5).

We now provide a geometric interpretation to explain that CMC has a lower conditional outage probability, i.e.,  $\Pr \left\{ \sum_{j=1}^{N_r} \alpha_{s,j}^2 < \beta \xi / N_r \right\} < \Pr \left\{ \sum_{j=1}^{N_r} \alpha_{s,j} < \sqrt{\beta \xi} \right\}$ . This geometric argument is, in essence, same as the one used by Brennan [17]. However, we emphasize, the key difference is that the CCI is not considered in [17] but it is included in

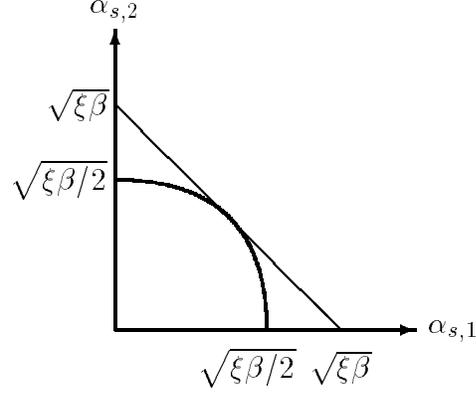


Figure 3.1. Regions of integration for the conditional outage probability of CMC and EGC.

our study. Consider  $N_r = 2$ , as shown in Fig. 3.1, for CMC,  $\Pr(\alpha_{s,1}^2 + \alpha_{s,2}^2 < \beta\xi/2)$  is obtained by integrating the joint density function of  $\alpha_{s,1}$  and  $\alpha_{s,2}$  over the interior of a quarter circle, while for EGC,  $\Pr(\alpha_{s,1} + \alpha_{s,2} < \sqrt{\beta\xi})$  is obtained by integrating the same density function over a triangular region. Since the integration region for CMC is smaller than that for EGC, it is obvious that for  $N_r = 2$ , CMC has a lower conditional outage probability. For  $N_r > 2$ , by integrating the joint density function of  $\alpha_{s,1}, \dots, \alpha_{s,N_r}$  over a space of  $N_r$ -dimensions, by the same arguments, it can also be shown that CMC has a lower conditional outage probability than that of EGC. Upon averaging the conditional outage probability with respect to the PDFs of  $\xi_1$  and  $\xi_2$  in (3.34) and (3.35), since PDFs  $f_{\xi_1}(\xi) = f_{\xi_2}(\xi)$ , it is clear that the outage probability for CMC is strictly lower than that for EGC.

We can also use the Cauchy-Schwarz inequality [54] to prove that CMC has a lower outage probability than that of EGC. We rewrite (3.34) and (3.35) as

$$P_{\text{OUT,CMC}}(\beta) = \int \Pr \left\{ N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2 < \beta\xi \right\} f_{\xi_1}(\xi) d\xi$$

and

$$P_{\text{OUT,EGC}}(\beta) = \int \Pr \left\{ \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2 < \beta\xi \right\} f_{\xi_2}(\xi) d\xi.$$

By applying the Cauchy-Schwarz inequality, we have

$$\left( \sum_{j=1}^{N_r} 1 \cdot \alpha_{s,j} \right)^2 \leq N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2.$$

Hence, the conditional outage probabilities

$$\Pr \left\{ N_r \sum_{j=1}^{N_r} \alpha_{s,j}^2 < \beta \xi \right\} \leq \Pr \left\{ \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2 < \beta \xi \right\}.$$

Since PDFs  $f_{\xi_1}(\xi) = f_{\xi_2}(\xi)$ , we have  $P_{\text{OUT,CMC}}(\beta) \leq P_{\text{OUT,EGC}}(\beta)$ , where equality is achieved when  $N_r = 1$  (single antenna). When  $N_r > 1$ , the outage probability for CMC is strictly lower than that for EGC. We remark that this conclusion is true for arbitrary interferer power distributions.

### 3.4.2 Outage Probability Comparison for CMC and SC

#### 3.4.2.1 $L = 1$ case

We first consider the case of one interfering signal. This approximates the case when the system has one strong dominant interfering user. By setting  $L = 1$  in (3.28a) and (3.32a), we obtain

$$P_{\text{OUT,CMC}}(\beta) \Big|_{L=1} = P_{\text{OUT,SC}}(\beta) \Big|_{L=1} = \left( \frac{\beta_0}{\beta_0 + \Lambda} \right)^{N_r}.$$

Therefore, the outage probabilities for CMC and SC are, in fact, the same for the case of one interfering signal.

An alternative way to show this equality is to observe that, for  $L = 1$ , in the denominators of (3.24) and (3.31),  $Y_1 \triangleq \eta_1/\Lambda_1$  and  $Y_2 \triangleq \alpha_{1,1}^2/\Lambda_1$  have the same distribution. Note that  $Y_1$  and  $Y_2$  are exponentially distributed with mean  $1/\Lambda_1$ . For CMC, the outage probability can be calculated as

$$\Pr(\text{SIR}_{\text{CMC}} < \beta) = \Pr \left( \frac{\sum_{j=1}^{N_r} \alpha_{s,j}^2}{Y_1} < \beta_0 \right)$$

$$\begin{aligned}
&= \mathbf{E}_{\alpha_{s,1}, \dots, \alpha_{s,N_r}} \left\{ \Pr \left( Y_1 > \frac{\sum_{j=1}^{N_r} \alpha_{s,j}^2}{\beta_0} \right) \right\} \\
&= \mathbf{E}_{\alpha_{s,1}, \dots, \alpha_{s,N_r}} \left\{ e^{-\frac{\Lambda_1}{\beta_0} \sum_{j=1}^{N_r} \alpha_{s,j}^2} \right\} \\
&= \left\{ \mathbf{E}_{\alpha_{s,1}} \left[ e^{-\frac{\Lambda_1}{\beta_0} \alpha_{s,1}^2} \right] \right\}^{N_r} \tag{3.36}
\end{aligned}$$

where the third equality follows the fact that  $Y_1$  is exponentially distributed with mean  $1/\Lambda_1$  and  $\Pr(Y_1 > y) = e^{-\Lambda_1 y}$ ; the last equality is due to the fact that  $\alpha_{s,1}, \dots, \alpha_{s,N_r}$  are i.i.d. RVs. For SC, the outage probability can be calculated as

$$\begin{aligned}
P(\text{SIR}_{\text{sc}} < \beta) &= \left[ \Pr \left( \frac{\alpha_{s,1}^2}{Y_2} < \beta_0 \right) \right]^{N_r} \\
&= \left\{ \mathbf{E}_{\alpha_{s,1}} \left[ \Pr \left( Y_2 > \frac{\alpha_{s,1}^2}{\beta_0} \right) \right] \right\}^{N_r} \\
&= \left\{ \mathbf{E}_{\alpha_{s,1}} \left[ e^{-\frac{\Lambda_1}{\beta_0} \alpha_{s,1}^2} \right] \right\}^{N_r}. \tag{3.37}
\end{aligned}$$

Comparing (3.36) to (3.37), we have that, when  $L = 1$ , the outage probabilities of CMC and SC are identical.

### 3.4.2.2 $L > 1$ case

When the number of interfering signals is greater than one, we prove that, for equal interferer powers, the outage probabilities for CMC is smaller than those of SC. For distinct interferer powers, our numerical results suggest that CMC still outperforms SC when  $L > 1$ .

To prove  $P_{\text{OUT,CMC}}(\beta) < P_{\text{OUT,SC}}(\beta)$  in the case of equal interferer powers, we need to show that, from (3.28a) and (3.32a),

$$\left( \frac{\beta_0}{\beta_0 + \Lambda} \right)^{N_r} \sum_{k=0}^{L-1} \frac{(k + N_r - 1)!}{k!(N_r - 1)!} \left( \frac{\Lambda}{\beta_0 + \Lambda} \right)^k < \left[ 1 - \left( \frac{\Lambda}{\beta_0 + \Lambda} \right)^L \right]^{N_r} \tag{3.38}$$

for  $L > 1$  and  $N_r > 1$ . Before proving this result, we introduce two lemmas.

**Lemma 1** For a positive integer  $N_r$  and a non-negative integer  $k$ ,

$$\sum_{j=0}^k \frac{(N_r + j - 1)!}{(N_r - 1)!j!} = \frac{(N_r + k)!}{N_r!k!}. \quad (3.39)$$

*Proof:* We prove Lemma 1 by induction. Clearly, (3.39) is true for  $k = 0$ . Assuming (3.39) holds for  $k$ , we shall show that the expression holds for  $k + 1$ , i.e.,

$$\sum_{j=0}^{k+1} \frac{(N_r + j - 1)!}{(N_r - 1)!j!} = \frac{(N_r + k + 1)!}{N_r!(k + 1)!}. \quad (3.40)$$

To show this, we expand the left side of (3.40) as

$$\begin{aligned} \sum_{j=0}^{k+1} \frac{(N_r + j - 1)!}{(N_r - 1)!j!} &= \sum_{j=0}^k \frac{(N_r + j - 1)!}{(N_r - 1)!j!} + \frac{(N_r + k)!}{(N_r - 1)!(k + 1)!} \\ &= \frac{(N_r + k)!}{N_r!k!} + \frac{(N_r + k)!}{(N_r - 1)!(k + 1)!} \\ &= \frac{(N_r + k + 1)!}{N_r!(k + 1)!}. \end{aligned}$$

Therefore, (3.40) holds. By induction, we have Lemma 1. ■

Lemma 1 can also be found in the exercise of [116, pp. 212] where no proof is given.

**Lemma 2** For positive integers  $L > 1$  and  $N_r > 1$ , we have

$$(1 + x + \cdots + x^{L-1})^{N_r} = \sum_{k=0}^{L-1} \frac{(N_r + k - 1)!}{(N_r - 1)!k!} x^k + \text{higher order terms}. \quad (3.41)$$

*Proof:* We prove Lemma 2 by induction. It can be shown easily that (3.41) is true for  $N_r = 2$ . Now assuming (3.41) holds, we need to show that the expression holds for  $N_r + 1$ , i.e.,

$$(1 + x + \cdots + x^{L-1})^{N_r+1} = \sum_{k=0}^{L-1} \frac{(N_r + k)!}{N_r!k!} x^k + \text{higher order terms}. \quad (3.42)$$

We write the left side of (3.42) as

$$(1 + x + \cdots + x^{L-1})^{N_r+1}$$

$$\begin{aligned}
&= (1 + x + \dots + x^{L-1})(1 + x + \dots + x^{L-1})^{N_r} \\
&= (1 + x + \dots + x^{L-1}) \left( 1 + N_r x + \dots + \frac{(N_r + k - 1)!}{(N_r - 1)!k!} x^k + \dots + \frac{(N_r + L - 2)!}{(N_r - 1)!(L - 1)!} x^{L-1} \right. \\
&\quad \left. + \text{higher order terms} \right) \\
&= 1 + (N_r + 1)x + \dots + \sum_{j=0}^k \frac{(N_r + j - 1)!}{(N_r - 1)!j!} x^j + \dots + \sum_{j=0}^{L-1} \frac{(N_r + j - 1)!}{(N_r - 1)!j!} x^{L-1} \\
&\quad + \text{higher order terms.}
\end{aligned}$$

Applying Lemma 1 in the last expression, we have (3.42). By induction, we have Lemma 2. ■

From Lemma 2, it follows that, for  $x > 0$ ,  $N_r > 1$ , and  $L > 1$ ,

$$(1 + x + \dots + x^{L-1})^{N_r} > \sum_{k=0}^{L-1} \frac{(N_r + k - 1)!}{(N_r - 1)!k!} x^k. \quad (3.43)$$

Denoting  $\frac{\Lambda}{\beta_0 + \Lambda}$  by  $x$ , for the case of equal interferer powers, we rewrite the outage probability for CMC as

$$P_{\text{OUT,CMC}}(x) = (1 - x)^{N_r} \sum_{k=0}^{L-1} \frac{(N_r + k - 1)!}{(N_r - 1)!k!} x^k, \quad (3.44)$$

and the outage probability for SC as

$$\begin{aligned}
P_{\text{OUT,SC}}(x) &= (1 - x^L)^{N_r} \\
&= (1 - x)^{N_r} (1 + x + \dots + x^{L-1})^{N_r}
\end{aligned} \quad (3.45)$$

where  $0 < x < 1$  (since both  $\beta_0$  and  $\Lambda$  are both positive). Now comparing (3.44) and (3.45) and using (3.43), we immediately obtain the inequality in (3.38), that is, the outage probability for CMC is strictly lower than that of SC when  $L > 1$ .

### 3.4.3 Outage Probability Comparison for EGC and SC

A general analytical outage probability comparison of EGC and SC in interference is difficult. As shown in Section 3.5, the relative performance of EGC and SC depends on factors such as the number of interferers and the interferer power distribution.

## 3.5 Results and Discussions

In this section, in addition to assessing the accuracy of incoherent interference power calculation for EGC in [6], we make quantitative comparisons of the outage probabilities for CMC, EGC, and SC with cochannel interference in a Rayleigh fading channel and an interference-limited environment. Outage probabilities for finite SNRs are also simulated.

The outage probabilities for EGC using incoherent interference power calculation are obtained from [6, (16)] by specializing the Nakagami- $m$  fading channels to a Rayleigh fading channel<sup>6</sup>. In obtaining our numerical results for the outage probabilities of EGC in (3.11) using coherent combining, we choose the period  $T_0$  in the range of 40 to 80. It was found that typically 64 or 128 terms in the series enable an accuracy of  $10^{-8}$  to be achieved. Unless otherwise specified, all EGC outage probabilities are obtained using coherent interference power calculation. We use (3.28) and (3.32) to compute the outage probabilities for CMC and SC, respectively. All outage probabilities are plotted on normal probability papers [65, Appendix 2B]. In obtaining the results, we set the rolloff factor  $\rho = 0$ .

### 3.5.1 Coherent and Incoherent Interference Power Calculations for EGC

The concept of coherent and incoherent interference power calculation over the diversity branches for EGC has been introduced in Section 3.3.1. Here, we compare the two calculation methods analytically for the case of one interferer. With the incoherent interference

---

<sup>6</sup>In [6], a minor typo in (16a) needs to be corrected, i.e.,  $A_n$  is to be substituted with  $A_n^L$ .

power calculation, the SIR at the EGC output becomes

$$\text{SIR}_{\text{EGC, incoherent}} = \frac{\frac{\Lambda_1}{1-\rho/4} \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2}{\sum_{j=1}^{N_r} \alpha_{s,j}^2} = \frac{\kappa}{\zeta_1} \quad (3.46)$$

where we have  $\kappa \triangleq \frac{\Lambda_1}{1-\rho/4} \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2$  and  $\zeta_1 \triangleq \sum_{j=1}^{N_r} \alpha_{s,j}^2$  in (3.46). It can be shown straightforwardly that  $\zeta_1$  has a chi-square distribution with  $2N_r$  degrees of freedom and its PDF is given by

$$p_{\zeta_1}(\zeta_1) = \frac{1}{\Gamma(N_r)} \zeta_1^{N_r-1} e^{-\zeta_1}. \quad (3.47)$$

With a coherent interference power calculation, the SIR at the output of EGC in (3.5) becomes

$$\text{SIR}_{\text{EGC, coherent}} = \frac{\frac{\Lambda_1}{1-\rho/4} \left( \sum_{j=1}^{N_r} \alpha_{s,j} \right)^2}{\mu_1} = \frac{\kappa}{\mu_1} \quad (3.48)$$

where  $\mu_1$  is exponentially distributed with mean  $N_r$  and its PDF is given by (3.6). We note that the numerators in (3.46) and (3.48) are identical. For a given outage threshold  $\beta$ , the outage probabilities for incoherent and coherent interference power calculation are, respectively, given by  $E_\kappa \{ \text{Pr}(\zeta_1 > \kappa/\beta) \}$  and  $E_\kappa \{ \text{Pr}(\mu_1 > \kappa/\beta) \}$ . For practical applications, the low outage probability region is of interest, i.e., small values of  $\beta$ , and it is sufficient to compare the tail probabilities of  $\zeta_1$  and  $\mu_1$ . Fig. 3.2 plots the PDFs of  $\zeta_1$  (incoherent calculation method) and  $\mu_1$  (coherent method) for four antennas and one interferer. By comparing tails, it is clear that in the low outage probability region, the coherent interference power calculation yields a higher outage probability.

Figs. 3.3 and 3.4 compare outage probabilities using the incoherent and coherent interference power calculations for EGC. Fig. 3.3 plots the outage probabilities for  $L = 1$  interferer and for  $N_r = 1, 2$ , and 4 antennas under equal interferer powers. As shown, for the trivial case of  $N_r = 1$ , as expected, both calculation methods give the same outage

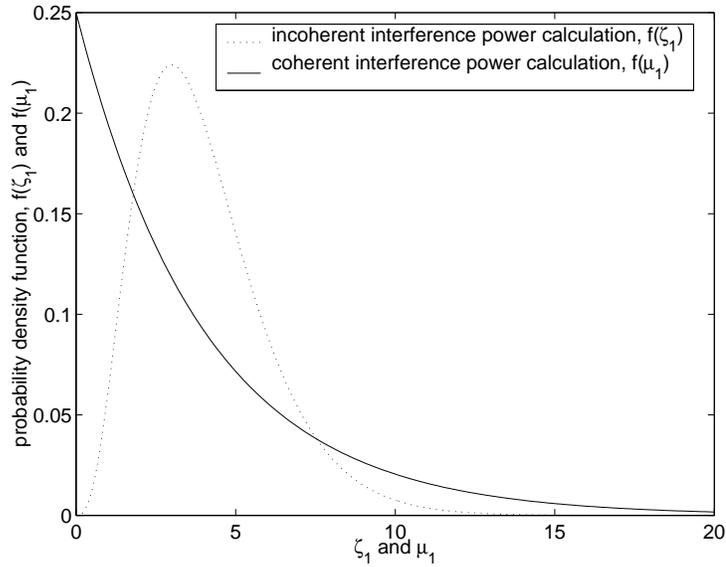


Figure 3.2. Probability density functions of  $\mu_1$  and  $\zeta_1$  for  $L = 1$  interferer and  $N_r = 4$  antennas.

probability. When  $N_r > 1$  antennas and for low outage probabilities, the coherent interference power calculation method predicts a higher outage probability. In other words, an outage analysis using incoherent interference power combining over the diversity branches can underestimate the outage probability. For example, for  $N_r = 4$ , at a 0.01 outage probability level, the incoherent interference power calculation overestimates the output SIR by about 1.5 dB. Similar observations can be made in Fig. 3.4 for a fixed number of antennas ( $N_r = 4$ ) and different numbers of interferers, where we note that the outage performance difference increases between these two interference power calculation methods as the number of interferers decreases.

To study the case of distinct interferer powers, we define the ratio of signal power to average interference power as

$$\Lambda_{avg}(\text{dB}) = 10 \log_{10} \frac{P_s}{(1/L) \sum_{i=1}^L P_i}.$$

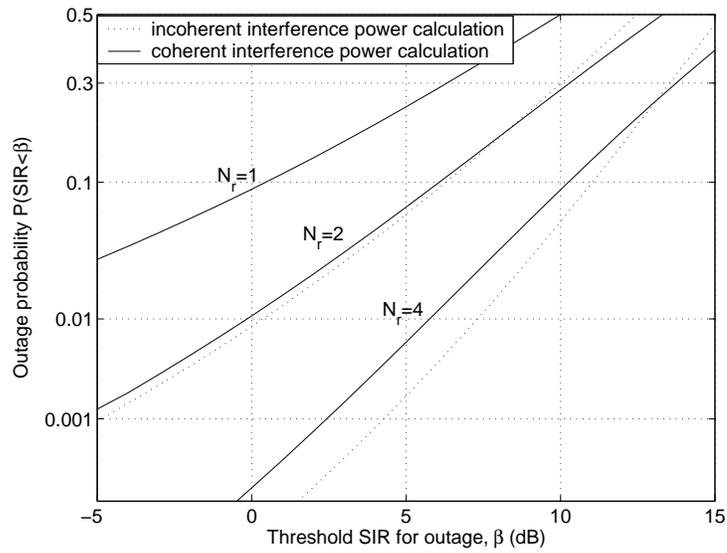


Figure 3.3. Outage probability comparison of coherent and incoherent interference power calculation for EGC with one interferer ( $L = 1$ ) and equal interferer powers ( $\Lambda = 10$  dB) for  $N_r = 1, 2$ , and 4 antennas.

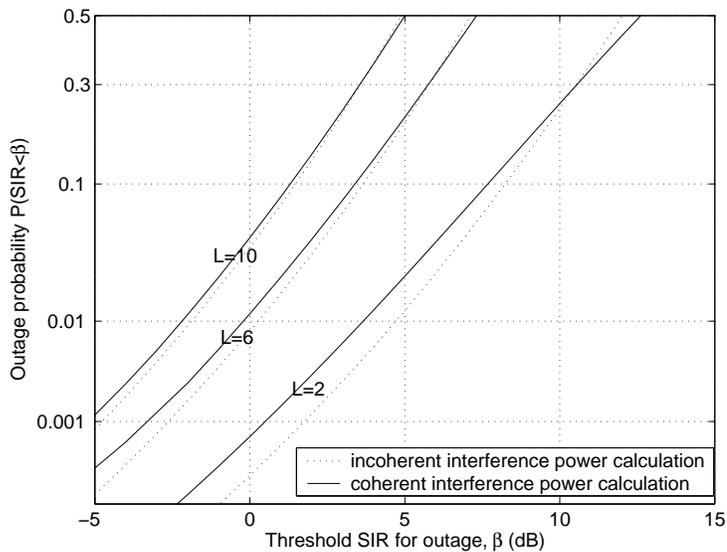


Figure 3.4. Outage probability comparison of coherent and incoherent interference power calculation for EGC with four antennas ( $N_r = 4$ ) and equal interferer powers ( $\Lambda = 10$  dB) for  $L = 2, 6$ , and 10 interferers.

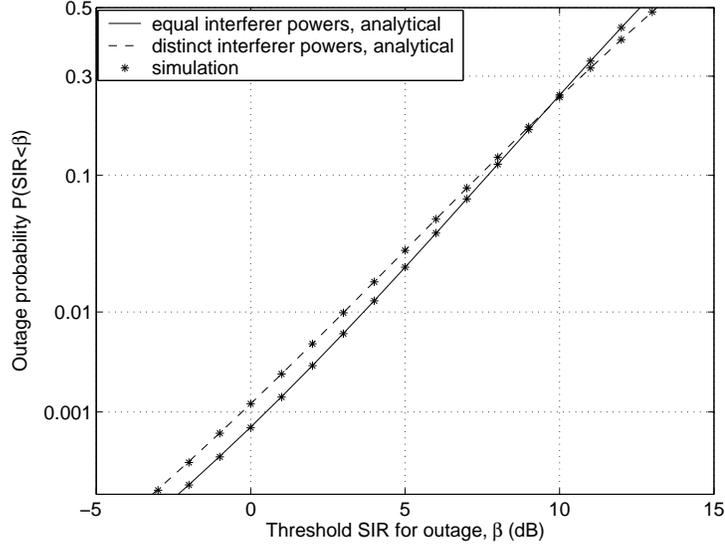


Figure 3.5. Analytical EGC outage probability using coherent interference power calculation and Monte Carlo simulated outage probability for equal ( $\Lambda = 10$  dB) and distinct ( $\Lambda_{avg} = 10$  dB) interference power distributions with  $L = 2$  interferers and  $N_r = 4$  antennas. The interference power vector for  $L = 2$  is  $[0.1, 0.9]$ .

Denote the normalized interference power vector by  $q = [q_1, q_2, \dots, q_L]$ , where  $\sum_{i=1}^L q_i = 1$ . Given  $q_i$  and  $\Lambda_{avg}$ , we can calculate the  $i$ th signal to interference power ratio  $\Lambda_i$  as

$$\Lambda_i(\text{dB}) = \frac{P_s}{P_i}(\text{dB}) = \Lambda_{avg}(\text{dB}) - 10 \log_{10}(Lq_i).$$

In Fig. 3.5, we compare the Monte Carlo simulation of EGC outage probability using (3.5) with the analytical outage probability using the new coherent interference power calculation method for the cases of both equal and distinct interferer powers. As shown, the analysis and simulation results agree.

### 3.5.2 Outage Probability Comparisons For CMC, EGC and SC

Fig. 3.6 plots the outage probabilities for CMC, EGC, and SC against the outage threshold  $\beta$  with four diversity branches and equal interferer power ( $\Lambda = 10$  dB) for  $L = 1, 2$ , and 6 interferers. Fig. 3.6 shows that the outage probabilities for all diversity schemes increase with an increasing number of interferers. We further confirm, as shown in Section 3.4.1, that the outage probabilities for CMC are smaller than those of EGC in all cases considered. However, the improvements of CMC over EGC are small. For example, for  $N_r = 4$ ,  $L = 2$  and at 1% outage probability level, the corresponding output SIR of EGC and CMC are, respectively, about about 3.5 dB and 4.5 dB. Therefore, with a four-element array, the advantage of CMC over EGC is only 1 dB. We further observe that this 1 dB improvement is approximately the same over a wide range of outage probability levels and is insensitive to the number of interferers. Fig. 3.6 also indicates that the outage probabilities of CMC are smaller than those obtained using SC, except that, in the special case of one interferer, these two diversity schemes yields the same outage probability. The improvements of CMC over SC, however, increase quickly with an increasing number of interferers.

Fig. 3.7 compares the outage probabilities of CMC, EGC, and SC with three interferers and equal interferer powers ( $\Lambda = 10$  dB) for  $N_r = 2$  and 4 diversity branches. As shown, CMC outperforms EGC and SC. However, when the number of antenna branches decreases, the performance differences among three diversity schemes are insignificant. For example, for a dual branch receiver, the improvements of CMC over EGC and SC are, respectively, 0.5 dB and 0.7 dB. We remark that in the case of equal interferer powers, changing the power ratio  $\Lambda$  simply causes the outage probability curves in Figs. 3.6 and 3.7 to shift horizontally.

Figs. 3.8 and 3.9 study the impacts of distinct interferer powers on the outage probabilities for CMC, EGC, and SC. With  $N_r = 4$  diversity branches, Fig. 3.8 compares the

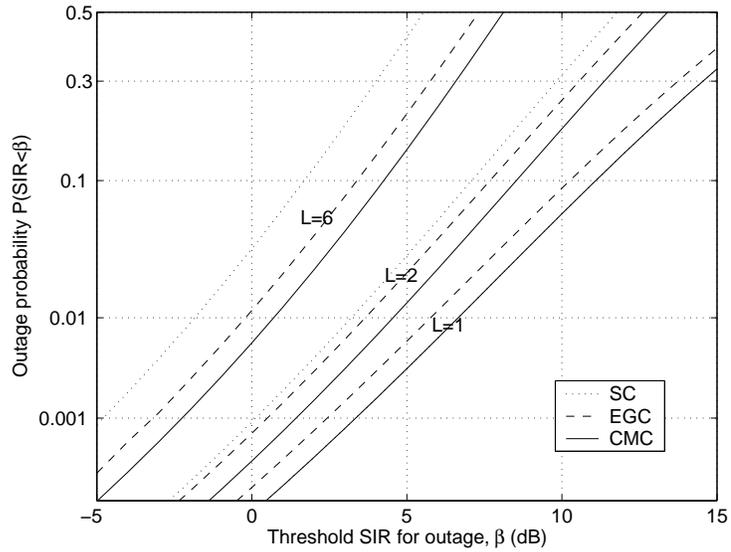


Figure 3.6. Outage probability of CMC, EGC and SC with  $N_r = 4$  antennas and equal interferer powers ( $\Lambda = 10$  dB) for  $L = 1, 2$ , and 6 interferers.

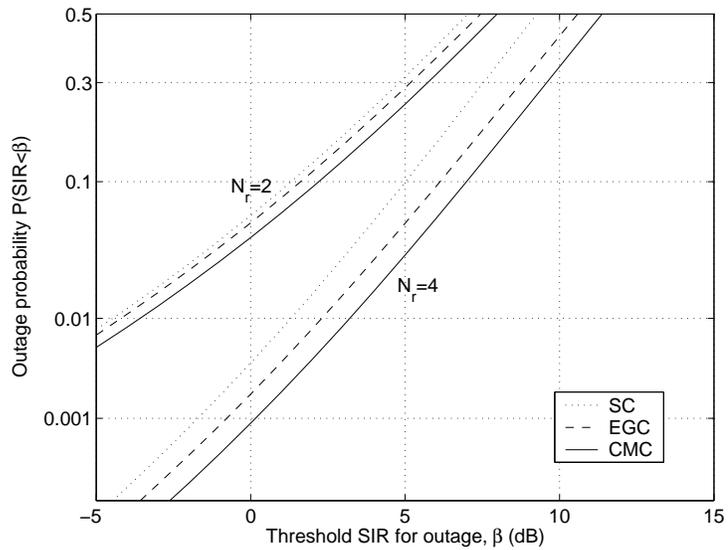


Figure 3.7. Outage probability of CMC, EGC and SC with  $L = 3$  interferers and equal interferer powers ( $\Lambda = 10$  dB) for  $N_r = 2$  and 4 antennas.

outage probabilities for  $L = 2$  interferers with a highly unbalanced interference power vector  $[0.1, 0.9]$  and for  $L = 6$  interferers with a more evenly-distributed interference power vector  $[0.05, 0.1, 0.15, 0.22, 0.23, 0.25]$ . In both cases, CMC outperforms EGC and SC. The relative performance for EGC and SC, in fact, depends on the interference power vector. For six interferers, Fig. 3.8 shows that EGC outperforms SC. This is similar to the case of equal interferer powers in Fig. 3.6. For two interferers, however, EGC is inferior to SC. This is in sharp contrast to the case of equal interferer powers. It is due to the fact that interference power vector  $[0.1, 0.9]$  represents the case of a strong dominant interferer, and SC performs almost as well as CMC for this scenario. With  $L = 3$  interferers and interference power vector  $[0.1, 0.2, 0.7]$ , Fig. 3.9 compares the outage probabilities for  $N_r = 2$  and 4 diversity branches. Comparing to the case of equal interferer powers in Fig. 3.7, the performance difference between EGC and SC is smaller. This is, again, due to the presence of a dominant interferer.

### 3.5.3 Finite SNRs

In the presence of noise, it is difficult to obtain an analytical outage expression for EGC which enables accurate numerical calculation. Therefore, we use Monte Carlo simulation to obtain the outage probability curves. Assuming that all antenna branches have the same noise powers, Figs. 3.10 and 3.11 show the outage probabilities of CMC<sup>7</sup>, EGC and SC at different SNRs for four antennas ( $N_r = 4$ ), one interferer ( $L = 1$ ) and equal interferer powers ( $\Lambda = 10$  dB). We observe that, for one interferer, as expected, SC outperforms EGC at high SNRs, but may not outperform EGC at lower SNRs. To investigate how closely the analysis of an interference-limited environment holds for finite SNRs, we compare the  $L = 1$  curves in Fig. 3.6 to those in Figs. 3.10 and 3.11. We note that the performance of

---

<sup>7</sup>In this case, CMC is equivalent to MRC.

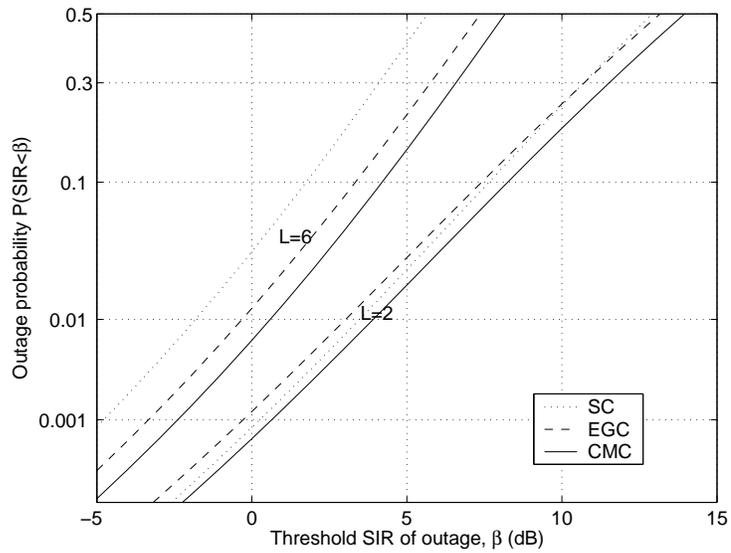


Figure 3.8. Outage probability of CMC, EGC and SC with  $N_r = 4$  antennas and distinct interferer powers ( $\Lambda_{avg} = 10$  dB) for  $L = 2$  and 6 interferers. The interference power vectors for  $L = 2$  and 6 are, respectively,  $[0.1, 0.9]$  and  $[0.05, 0.1, 0.15, 0.22, 0.23, 0.25]$ .

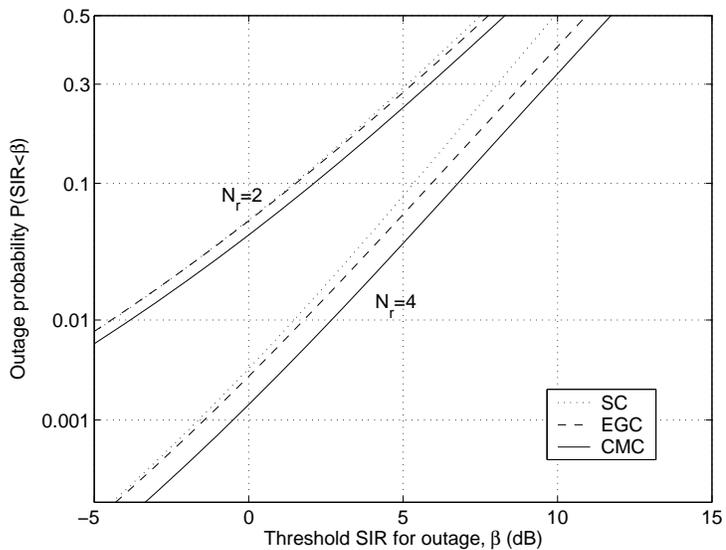


Figure 3.9. Outage probability of CMC, EGC and SC with  $L = 3$  interferers and distinct interferer powers ( $\Lambda_{avg} = 10$  dB) for  $N_r = 2$  and 4 antennas. The interference power vector for  $L = 3$  is  $[0.1, 0.2, 0.7]$ .

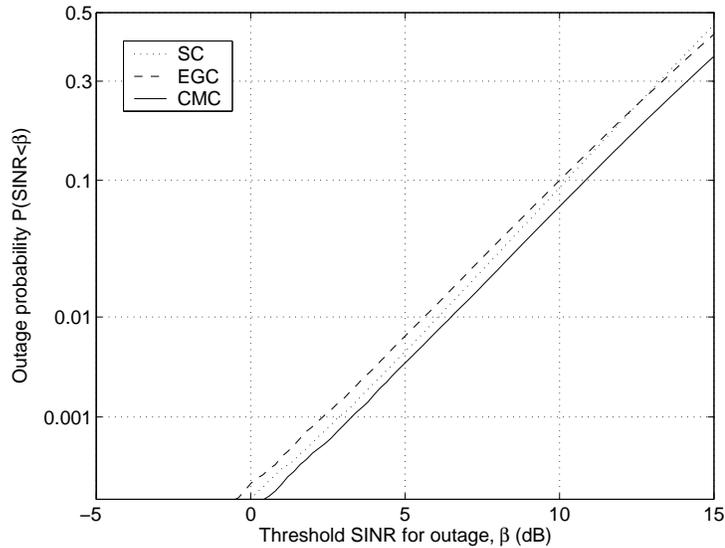


Figure 3.10. Monte Carlo simulated outage probability of CMC, EGC and SC for a finite SNR:  $N_r = 4$  antennas,  $L = 1$  interferer, equal interferer powers  $\Lambda = 10$  dB, and SNR=20 dB.

CMC and EGC is within 2 dB at low SNR (10 dB) and almost the same as that predicted by the interference-limited analysis at higher SNR (20 dB). The interference-limited analysis of SC is less accurate for finite SNRs.

### 3.6 Conclusions

In this chapter, a new outage probability expression for EGC, using the exact interference power calculations, has been derived for an interference-limited environment and flat Rayleigh fading. With this new analysis, we assess the accuracy of the existing method which calculates the interference power approximately. The numerical results show that with four receiving antennas, as much as 1.5 dB difference in output SIR of EGC combiner may exist at the same outage probability. Hence, the existing method can lead to overly

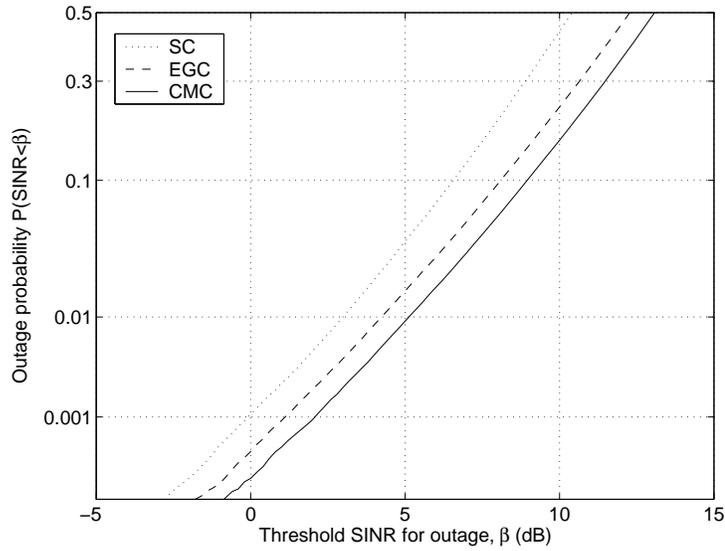


Figure 3.11. Monte Carlo simulated outage probability of CMC, EGC and SC for a finite SNR:  $N_r = 4$  antennas,  $L = 1$  interferer, equal interferer powers  $\Lambda = 10$  dB, and SNR=10 dB.

optimistic outage performance prediction.

We have analytically compared the outage probabilities for channel-matched combining (CMC, a practical variation of MRC that does not require SNRs at different antennas), EGC, and SC in an interference-limited environment over flat Rayleigh fading. We have shown that CMC has a lower outage probability than that of EGC, and that CMC has no greater outage probability than that of SC. The relative outage performance between EGC and SC, however, depends on the number of interferers and interferer power distribution. For finite SNRs, the simulation results show that the relative performance between EGC and SC is SNR-dependent.

## **Chapter 4**

# **Channel Estimation and Data Detection for MIMO Systems under Spatially and Temporally Colored Interference**

### **4.1 Introduction**

A large Shannon channel capacity exists for wireless systems with multiple transmitting and receiving antennas in a rich scattering environment [35, 114]. It was shown that the capacity increases linearly with the smaller of the numbers of transmitting and receiving antennas. While a substantial amount of research has been conducted for a single MIMO link without interference, the impact of interference on MIMO systems has recently attracted much interest. In a cellular environment, cochannel interference (CCI) from other cells exists due to channel reuse. In [30], channel capacities in the presence of spatially colored interference were derived with varying degrees of knowledge of the channel matrix and interference statistics at the transmitter. In [19] and [20], it was demonstrated that the capacity improvement of MIMO over SIMO (single-input multiple-output) is significantly reduced going from a noise-limited to an interference-limited environment due to the lack

of degrees of freedom at the receiver to combat CCI. In [105], with total interference power fixed, it was shown that the capacity degrades significantly as the number of interferers increases. The MIMO capacity with interference in the limit of a large number of antennas was studied in [74]. The spectral efficiency of a multicell system with MIMO links was assessed in [14, 15, 31]. The output signal-to-interference-power ratio (SIR) was analytically calculated in [60] when a single data stream is transmitted over independent Rayleigh MIMO channels. Space-time codes which suppress interference have been investigated in [37, 124]. While the majority of the studies deal with channel capacity, in this paper we will focus on the symbol error rate performance of a MIMO link with interference in practical systems.

The estimation of vector channels and spatial interference statistics for CDMA (code division multiple access) SIMO systems can be found in [108]. Most studies of channel estimation and data detection for MIMO systems assume spatially and temporally white interference. In [77], maximum-likelihood (ML) estimate of the channel matrix using training sequences was presented assuming temporally white interference. Assuming perfect knowledge of the channel matrix at the receiver, ordered zero-forcing (ZF) and minimum mean-squared error (MMSE) detection were studied for both spatially and temporally white interference in [47] and [55], respectively. However, in cellular systems, the interference is, in general, both spatially and temporally colored.

In this chapter, we propose an algorithm to jointly estimate the channel and spatial interference correlation matrices from training sequences. By exploiting the temporal interference correlation, we develop a multi-vector-symbol MMSE data detection scheme. In the case of a single interferer, we show that the spatial and temporal second-order interference statistics are separable, and that the temporal interference correlation can be determined a priori. The benefits of taking the temporal and spatial interference correlation into account

for channel estimation and data detection are evaluated through Monte Carlo simulation. The assumption of the decoupled structure of interference statistics in the presence of noise is also tested.

This chapter is organized as follows. In Section 4.2, we present a system model for interference statistics. In Section 4.3, we derive ML estimates of the channel and spatial interference correlation matrices assuming known temporal interference correlation. In Section 4.4, one-vector-symbol detection is extended to a multi-vector-symbol version by exploiting the temporal interference correlation. In Section 4.5, for the case of one interferer, we assess the benefits of taking temporal and spatial interference correlation into account for channel estimation and data detection, and test the assumption of the decoupled structure of interference statistics in the presence of noise.

## 4.2 System Model

We consider a single-user link consisting of  $N_t$  transmitting and  $N_r$  receiving antennas. The desired user transmits data frame by frame. Each frame has  $M$  data vectors. The first  $N$  data vectors are used for training so that the desired user's channel matrix and interference statistics can be estimated, and the remaining data vectors are for information transmission. Recall in Section 2.2, in a slow flat fading environment, the received signal vector at time  $j$  is expressed as

$$\mathbf{y}_j = \mathbf{H}\mathbf{x}_j + \mathbf{n}_j, \quad j = 0, \dots, M - 1 \quad (4.1)$$

where  $\mathbf{x}_j$  is the transmitted data vector,  $\mathbf{H}$  is the  $N_r \times N_t$  spatial channel gain matrix, and interference vector  $\mathbf{n}_j$  is zero-mean circularly symmetric complex Gaussian. We assume channel matrix  $\mathbf{H}$  is fixed during one frame. This is a reasonable assumption since high speed data services envisioned for MIMO systems are intended to low mobility users. By

the same argument, it is also assumed that the interference statistics are fixed during one frame.

In practice, the interference could be both spatially and temporally correlated. We assume that the cross-correlation between the interference vectors at times  $i$  and  $j$  is  $E \{ \mathbf{n}_i \mathbf{n}_j^\dagger \} = \Lambda_M(i, j) \mathbf{R}$  where  $\Lambda_M(i, j)$  is the  $(i, j)$ th element of the  $M \times M$  matrix  $\Lambda_M$ . As a result, the covariance matrix of the concatenated interference vector  $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{M-1}^T]^T$  is

$$E \{ \bar{\mathbf{n}} \bar{\mathbf{n}}^\dagger \} = \begin{bmatrix} \Lambda_M(0, 0) \mathbf{R} & \cdots & \Lambda_M(0, M-1) \mathbf{R} \\ \vdots & & \vdots \\ \Lambda_M(M-1, 0) \mathbf{R} & \cdots & \Lambda_M(M-1, M-1) \mathbf{R} \end{bmatrix} = \Lambda_M \otimes \mathbf{R} \quad (4.2)$$

where  $\otimes$  denotes Kronecker product, and matrices  $\Lambda_M$  and  $\mathbf{R}$  capture the temporal and spatial correlation of the interference, respectively. The  $(i, j)$ th element of matrix  $\mathbf{R}$  is the correlation between the  $i$ th and  $j$ th elements of interference vector  $\mathbf{n}_k$ ,  $k \in 0, \dots, M-1$ . The correlation matrices  $\Lambda_M$  and  $\mathbf{R}$  are determined by the application-specific signal model. Eqn. (4.2) implies that the spatial and temporal interference statistics are separable. In Section 4.5, we will give an example in which the interference covariance matrix has a Kronecker product form.

In addition to interference correlation, we remark that a decoupled temporal and spatial correlation structure arises in the statistics of fading vector channels consisting of a mobile with one antenna and a base station with an antenna array [81, 82]. Denoting the flat fading channel vector as  $\mathbf{a}(t)$ , it is shown that, for uniformly distributed scatterers around the mobile, the channel vector is zero-mean circularly symmetric complex Gaussian with correlation matrix

$$E \{ \mathbf{a}(t) \mathbf{a}^\dagger(t + \nu) \} = J_0(\omega_d \nu) \mathbf{R}_s$$

where  $\omega_d$  is the maximum Doppler frequency dependent on the mobile's speed, and  $J_0$  is the zero-order Bessel function of the first kind. The temporal and spatial correlation of the

fading channel vector are described by  $J_0(\omega_d \nu)$  and  $\mathbf{R}_s$ , respectively. Spatial correlation matrix  $\mathbf{R}_s$  depends on the antenna array's geometry, relative position between the mobile and the base station, and the radius of the scatterer ring around the mobile.

### 4.3 Joint Estimation of Channel Matrix and Spatial Interference Statistics

During a training period of  $N$  vector symbols, we concatenate the received signal vectors, the training signal vectors and the interference vectors as  $\bar{\mathbf{y}} = [\mathbf{y}_0^T \cdots \mathbf{y}_{N-1}^T]^T$ ,  $\bar{\mathbf{x}} = [\mathbf{x}_0^T \cdots \mathbf{x}_{N-1}^T]^T$  and  $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$ , respectively. The received signal in (4.1) is rewritten as

$$\bar{\mathbf{y}} = (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}} + \bar{\mathbf{n}}$$

where  $\bar{\mathbf{n}}$  is circularly symmetric complex Gaussian with zero-mean and covariance matrix  $\Lambda_N \otimes \mathbf{R}$ . Assuming prior knowledge of the temporal interference correlation matrix  $\Lambda_N$ , we need to estimate the channel matrix  $\mathbf{H}$  and the spatial interference correlation matrix  $\mathbf{R}$ . If  $\mathbf{R}$  and  $\Lambda_N$  are nonsingular, we have the conditional PDF

$$\Pr(\bar{\mathbf{y}}|\mathbf{H}, \mathbf{R}) = \frac{1}{\pi^{N \cdot N_r} \det(\Lambda_N \otimes \mathbf{R})} \exp \left\{ -[\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger (\Lambda_N \otimes \mathbf{R})^{-1} [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}] \right\}. \quad (4.3)$$

#### 4.3.1 Maximum-likelihood solution

The ML estimate of the pair of matrices  $(\mathbf{H}, \mathbf{R})$  is the value of  $(\mathbf{H}, \mathbf{R})$  that maximizes the conditional PDF in (4.3), which is equivalent to maximizing  $\ln \Pr(\bar{\mathbf{y}}|\mathbf{H}, \mathbf{R})$ . Using the inverse and determinant properties of Kronecker product in (2.19) and (2.20), it can be

shown that maximizing (4.3) is equivalent to minimizing

$$f(\mathbf{H}, \mathbf{R}) = \ln \det(\mathbf{R}) + \frac{1}{N} [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger (\Lambda_N^{-1} \otimes \mathbf{R}^{-1}) [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]. \quad (4.4)$$

Denoting the elements of  $\Lambda_N^{-1}$  as

$$\Lambda_N^{-1} = \begin{bmatrix} \alpha_{0,0} & \cdots & \alpha_{0,N-1} \\ \vdots & & \vdots \\ \alpha_{N-1,0} & \cdots & \alpha_{N-1,N-1} \end{bmatrix}, \quad (4.5)$$

we expand the second term in (4.4) as

$$\begin{aligned} & \frac{1}{N} [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger (\Lambda_N^{-1} \otimes \mathbf{R}^{-1}) [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}] \\ &= \frac{1}{N} [\bar{\mathbf{y}} - (\mathbf{I}_N \otimes \mathbf{H})\bar{\mathbf{x}}]^\dagger \begin{bmatrix} \alpha_{0,0}\mathbf{R}^{-1} & \cdots & \alpha_{0,N-1}\mathbf{R}^{-1} \\ \vdots & & \vdots \\ \alpha_{N-1,0}\mathbf{R}^{-1} & \cdots & \alpha_{N-1,N-1}\mathbf{R}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 - \mathbf{H}\mathbf{x}_0 \\ \vdots \\ \mathbf{y}_{N-1} - \mathbf{H}\mathbf{x}_{N-1} \end{bmatrix} \\ &= \frac{1}{N} [(\mathbf{y}_0 - \mathbf{H}\mathbf{x}_0)^\dagger \cdots (\mathbf{y}_{N-1} - \mathbf{H}\mathbf{x}_{N-1})^\dagger] \begin{bmatrix} \sum_{j=0}^{N-1} \alpha_{0,j}\mathbf{R}^{-1}(\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) \\ \vdots \\ \sum_{j=0}^{N-1} \alpha_{N-1,j}\mathbf{R}^{-1}(\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) \end{bmatrix} \\ &= \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j). \end{aligned} \quad (4.6)$$

Substituting (4.6) into (4.4), we obtain

$$\begin{aligned} f(\mathbf{H}, \mathbf{R}) &= \ln \det(\mathbf{R}) + \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) \\ &= \ln \det(\mathbf{R}) + \text{tr} \left\{ \mathbf{R}^{-1} \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i) (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)^\dagger \right\}. \end{aligned} \quad (4.7)$$

The last equality comes from the fact that  $(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)$  is a scalar, hence

$$(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) = \text{tr} \left\{ (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) \right\}$$

$$= \text{tr} \left\{ \mathbf{R}^{-1} (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j) (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i)^\dagger \right\}$$

due to  $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ .

To find the value of  $(\mathbf{H}, \mathbf{R})$  that minimizes  $f(\mathbf{H}, \mathbf{R})$  in (4.7), we set  $\partial f(\mathbf{H}, \mathbf{R})/\partial \mathbf{H} = \mathbf{0}$ .

The derivative of  $f(\mathbf{H}, \mathbf{R})$  with respect to matrix  $\mathbf{H}$  is defined as [76]

$$\frac{\partial f(\mathbf{H}, \mathbf{R})}{\partial \mathbf{H}} = \begin{bmatrix} \frac{\partial f}{\partial h_{11}} & \cdots & \frac{\partial f}{\partial h_{1N_t}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial h_{N_r 1}} & \cdots & \frac{\partial f}{\partial h_{N_r N_t}} \end{bmatrix} \quad (4.8)$$

where  $h_{ij}$  is the  $(i, j)$ th element of the complex matrix  $\mathbf{H}$ . The computation of (4.8) involves the derivative with respect to a complex number which is defined as, for example, the  $(1, 1)$ th element in (4.8),  $\frac{\partial f}{\partial h_{11}} = \frac{1}{2} \left[ \frac{\partial f}{\partial x} - j \frac{\partial f}{\partial y} \right]$  where  $h_{11} = x + jy$  [76].

Defining the weighted sample correlation matrices<sup>1</sup> as

$$\tilde{\mathbf{R}}_{yy} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} \mathbf{y}_i \mathbf{y}_j^\dagger, \quad (4.9)$$

$$\tilde{\mathbf{R}}_{xy} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} \mathbf{x}_i \mathbf{y}_j^\dagger, \quad (4.10)$$

and

$$\tilde{\mathbf{R}}_{xx} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} \mathbf{x}_i \mathbf{x}_j^\dagger, \quad (4.11)$$

from (4.7), we have

$$\frac{\partial f(\mathbf{H}, \mathbf{R})}{\partial \mathbf{H}} = \frac{\partial}{\partial \mathbf{H}} \left\{ -\text{tr}(\mathbf{H}\tilde{\mathbf{R}}_{xy}\mathbf{R}^{-1}) - \text{tr}(\mathbf{R}^{-1}\tilde{\mathbf{R}}_{xy}^\dagger\mathbf{H}^\dagger) + \text{tr}(\mathbf{H}^\dagger\mathbf{R}^{-1}\mathbf{H}\tilde{\mathbf{R}}_{xx}) \right\}.$$

Using the fact that

$$\frac{\partial \text{tr}(\mathbf{R}^{-1}\tilde{\mathbf{R}}_{xy}^\dagger\mathbf{H}^\dagger)}{\partial \mathbf{H}} = \mathbf{0}$$

---

<sup>1</sup>To distinguish weighted sample correlation matrices from conventional sample correlation matrices in Section 4.3.2, we denote the former by a tilde and the latter without a tilde.

and identities of matrix derivative [79]

$$\frac{\partial \text{tr}(\mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{A}^T \mathbf{B}^T$$

and

$$\frac{\partial \text{tr}(\mathbf{X}^\dagger \mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{A}^T \mathbf{X}^* \mathbf{B}^T,$$

we obtain

$$\hat{\mathbf{H}} = \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1}. \quad (4.12)$$

The spatial interference correlation matrix is estimated by setting  $\partial f(\hat{\mathbf{H}}, \mathbf{R}) / \partial \mathbf{R} = \mathbf{0}$ . Using identities of matrix derivative in [75]<sup>2</sup>

$$\frac{\partial \ln \det(\mathbf{X})}{\partial \mathbf{X}} = (\mathbf{X}^T)^{-1} \quad \text{for } \det(\mathbf{X}) > 0 \quad (4.13)$$

and

$$\frac{\partial \text{tr}(\mathbf{X}^{-1} \mathbf{B})}{\partial \mathbf{X}} = (\mathbf{X}^{-1} \mathbf{B} \mathbf{X}^{-1})^T \quad \text{for } \mathbf{X} \text{ nonsingular}, \quad (4.14)$$

we have

$$\frac{\partial f(\hat{\mathbf{H}}, \mathbf{R})}{\partial \mathbf{R}} = (\mathbf{R}^{-1})^T + \left( \mathbf{R}^{-1} \cdot \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \hat{\mathbf{H}} \mathbf{x}_i) (\mathbf{y}_j - \hat{\mathbf{H}} \mathbf{x}_j)^\dagger \cdot \mathbf{R}^{-1} \right)^T.$$

Therefore

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \hat{\mathbf{H}} \mathbf{x}_i) (\mathbf{y}_j - \hat{\mathbf{H}} \mathbf{x}_j)^\dagger \quad (4.15)$$

$$= \tilde{\mathbf{R}}_{yy} - \hat{\mathbf{H}} \tilde{\mathbf{R}}_{xy}. \quad (4.16)$$

---

<sup>2</sup>Although the identities of derivatives in [75] are for real matrices, it can be shown that they are also applicable to complex matrices.

In Appendix C, it is shown that the estimates  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{R}}$  can also be obtained by first setting  $\partial f(\mathbf{H}, \mathbf{R})/\partial \mathbf{R} = \mathbf{0}$  [68]. However, the derivation is much simpler by setting  $\partial f(\mathbf{H}, \mathbf{R})/\partial \mathbf{H} = \mathbf{0}$  first. It can be shown that  $\hat{\mathbf{H}}$  in (4.12) is an unbiased estimator, and that  $\hat{\mathbf{R}}$  in (4.16) is a biased estimator [88]. We remark that if  $\tilde{\mathbf{R}}_{xy}$  and  $\tilde{\mathbf{R}}_{xx}$  in (4.12) were instead known cross- and auto-correlation matrices, the estimate for  $\mathbf{H}$  would represent the Wiener solution.

### 4.3.2 Special case: temporally white interference

If interference is temporally white, with loss of generality, we may substitute  $\Lambda_N = \mathbf{I}_N$  into (4.9)-(4.16), and obtain estimates

$$\hat{\mathbf{H}}_w = \mathbf{R}_{xy}^\dagger \mathbf{R}_{xx}^{-1}, \quad (4.17)$$

and

$$\hat{\mathbf{R}}_w = \mathbf{R}_{yy} - \hat{\mathbf{H}}_w \mathbf{R}_{xy} \quad (4.18)$$

where the subscript  $w$  indicates temporally white interference, and the sample correlation matrices

$$\mathbf{R}_{yy} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{y}_i \mathbf{y}_i^\dagger, \quad (4.19)$$

$$\mathbf{R}_{xy} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{y}_i^\dagger, \quad (4.20)$$

and

$$\mathbf{R}_{xx} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{x}_i^\dagger. \quad (4.21)$$

Note that  $\hat{\mathbf{H}}_w$  in (4.17) is the same as the channel estimate used in [77]. We remark that both  $\hat{\mathbf{H}}$  in (4.12) and  $\hat{\mathbf{H}}_w$  in (4.17) are unbiased, however, the variances of  $\hat{\mathbf{H}}$  should be smaller than those of  $\hat{\mathbf{H}}_w$ .

### 4.3.3 Whitening filter interpretation

To obtain insight on the estimates in (4.12) and (4.16), we let the received signal vectors during the training period undergo a linear transformation where the transformed received signal vectors are

$$[\mathbf{y}'_0 \dots \mathbf{y}'_{N-1}] = [\mathbf{y}_0 \dots \mathbf{y}_{N-1}] \mathbf{\Lambda}_N^{-1/2}.$$

At the output of the transformation, we have

$$\mathbf{y}'_i = \mathbf{H}\mathbf{x}'_i + \mathbf{n}'_i, \quad i = 0, \dots, N-1, \quad (4.22)$$

where the transformed training signal vectors and interference vectors are

$$[\mathbf{x}'_0 \dots \mathbf{x}'_{N-1}] = [\mathbf{x}_0 \dots \mathbf{x}_{N-1}] \mathbf{\Lambda}_N^{-1/2}$$

and

$$[\mathbf{n}'_0 \dots \mathbf{n}'_{N-1}] = [\mathbf{n}_0 \dots \mathbf{n}_{N-1}] \mathbf{\Lambda}_N^{-1/2},$$

respectively. Concatenating the transformed interference vectors as  $\bar{\mathbf{n}}' = [\mathbf{n}'_0{}^T \dots \mathbf{n}'_{N-1}{}^T]^T$ , it can be shown that

$$\bar{\mathbf{n}}' = (\mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) \bar{\mathbf{n}}$$

where  $\bar{\mathbf{n}} = [\mathbf{n}_0{}^T \dots \mathbf{n}_{N-1}{}^T]^T$ . Since the covariance matrix of  $\bar{\mathbf{n}}$  is  $\mathbf{\Lambda}_N \otimes \mathbf{R}$ , the covariance matrix of  $\bar{\mathbf{n}}'$  is

$$\begin{aligned} \text{cov}(\bar{\mathbf{n}}') &= (\mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) \text{cov}(\bar{\mathbf{n}}) (\mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r})^\dagger \\ &= (\mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) (\mathbf{\Lambda}_N \otimes \mathbf{R}) (\mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{I}_{N_r}) \\ &= \mathbf{I}_N \otimes \mathbf{R} \end{aligned} \quad (4.23)$$

where we used the properties of Kronecker product in (2.17) and (2.18). We also used the fact that temporal correlation matrix  $\mathbf{\Lambda}_N$  is symmetric, so is  $\mathbf{\Lambda}_N^{-1/2}$ . From (4.23), it is

obvious that the transformed interference vectors  $\{\mathbf{n}'_0 \dots \mathbf{n}'_{N-1}\}$  are temporally white with spatial correlation matrix  $\mathbf{R}$ .

As a result, we can estimate  $\mathbf{H}$  and  $\mathbf{R}$  from the sample correlation matrices of transformed signal vectors as in Section 4.3.2. The sample correlation matrix

$$\begin{aligned}
\mathbf{R}_{y'y'} &= \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{y}'_i \mathbf{y}'_i{}^\dagger \\
&= \frac{1}{N} [\mathbf{y}'_0 \dots \mathbf{y}'_{N-1}] [\mathbf{y}'_0 \dots \mathbf{y}'_{N-1}]^\dagger \\
&= \frac{1}{N} [\mathbf{y}_0 \dots \mathbf{y}_{N-1}] \mathbf{\Lambda}_N^{-1/2} \mathbf{\Lambda}_N^{-\dagger/2} [\mathbf{y}_0 \dots \mathbf{y}_{N-1}]^\dagger \\
&= \frac{1}{N} [\mathbf{y}_0 \dots \mathbf{y}_{N-1}] \mathbf{\Lambda}_N^{-1} [\mathbf{y}_0 \dots \mathbf{y}_{N-1}]^\dagger = \tilde{\mathbf{R}}_{yy},
\end{aligned}$$

which shows that the weighted sample correlation matrix of  $\{\mathbf{y}_0 \dots \mathbf{y}_{N-1}\}$  is equivalent to the sample correlation matrix of  $\{\mathbf{y}'_0 \dots \mathbf{y}'_{N-1}\}$ . Similarly, the weighted sample correlation matrices  $\tilde{\mathbf{R}}_{xy}$  and  $\tilde{\mathbf{R}}_{xx}$  are equivalent to the sample correlation matrices  $\mathbf{R}_{x'y'}$  and  $\mathbf{R}_{x'x'}$ , respectively. Therefore, the estimates in (4.12) and (4.16) can also be realized by first temporally whitening the interference, and then forming the estimates from the sample correlation matrices of the transformed signal vectors.

If we drop the Gaussian assumption of interference statistics, the ML estimate  $\hat{\mathbf{H}}$  in (4.12) is still a least-squares (LS) estimate. During the training period, after whitening interference, the transformed received signal is shown in (4.22). The LS estimate of the channel matrix is the one minimizing

$$\begin{aligned}
f_3(\mathbf{H}) &= \text{tr} \left\{ \frac{1}{N} \sum_{i=0}^{N-1} (\mathbf{y}'_i - \mathbf{H}\mathbf{x}'_i) (\mathbf{y}'_i - \mathbf{H}\mathbf{x}'_i)^\dagger \right\} \\
&= \text{tr} \left\{ \mathbf{R}_{y'y'} - \mathbf{H}\mathbf{R}_{x'y'} - \mathbf{R}_{x'y'}^\dagger \mathbf{H}^\dagger + \mathbf{H}\mathbf{R}_{x'x'} \mathbf{H}^\dagger \right\} \\
&= \text{tr} \left\{ \tilde{\mathbf{R}}_{yy} - \mathbf{H}\tilde{\mathbf{R}}_{xy} - \tilde{\mathbf{R}}_{xy}^\dagger \mathbf{H}^\dagger + \mathbf{H}\tilde{\mathbf{R}}_{xx} \mathbf{H}^\dagger \right\}
\end{aligned}$$

By setting  $\partial f_3(\mathbf{H})/\partial \mathbf{H} = \mathbf{0}$ , we obtain the channel estimate in (4.12).

## 4.4 Data Detection

We focus on ordered MMSE detection due to the better performance of MMSE compared to ZF detection [9] as we mentioned in Section 2.2.3. For received signal vector  $\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \mathbf{n}_i$ , modifying the algorithm in [55], the steps of ordered MMSE detection of  $\mathbf{x}_i$  from  $\mathbf{y}_i$  with estimated channel and interference spatial correlation matrices are as follows:

Step 1 Initialization: set  $k = 1$ ,  $\mathbf{H}_k = \hat{\mathbf{H}}$ ,  $\tilde{\mathbf{x}}_k = \mathbf{x}_i$ ,  $\tilde{\mathbf{y}}_k = \mathbf{y}_i$ .

Step 2 Calculate the estimation error covariance matrix  $\mathbf{P}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^\dagger \hat{\mathbf{R}}^{-1} \mathbf{H}_k)^{-1}$ .

Find  $m = \arg \min_j \mathbf{P}_k(j, j)$  where  $\mathbf{P}_k(j, j)$  denotes the  $j$ th diagonal element of  $\mathbf{P}_k$ .

Hence, the  $m$ th signal component of  $\tilde{\mathbf{x}}_k$  has the smallest estimation error variance.

Step 3 Calculate the weighting matrix  $\mathbf{A}_k = (\mathbf{I}_{N_t+1-k} + \mathbf{H}_k^\dagger \hat{\mathbf{R}}^{-1} \mathbf{H}_k)^{-1} \mathbf{H}_k^\dagger \hat{\mathbf{R}}^{-1}$ . The  $m$ th

element of  $\tilde{\mathbf{x}}_k$  is estimated by  $\hat{x}_k^m = Q(\mathbf{A}_k(m, :) \tilde{\mathbf{y}}_k)$  where  $\mathbf{A}_k(m, :)$  denotes the  $m$ th row of matrix  $\mathbf{A}_k$  and  $Q(\cdot)$  denotes the quantization appropriate to the signal constellation.

Step 4 Assuming the detected signal is correct, remove the detected signal from the received signal,  $\tilde{\mathbf{y}}_{k+1} = \tilde{\mathbf{y}}_k - \hat{x}_k^m \mathbf{H}_k(:, m)$  where  $\mathbf{H}_k(:, m)$  denotes the  $m$ th column of  $\mathbf{H}_k$ .

Step 5  $\mathbf{H}_{k+1}$  is obtained by eliminating the  $m$ th column of matrix  $\mathbf{H}_k$ .  $\tilde{\mathbf{x}}_{k+1}$  is obtained by eliminating the  $m$ th component of vector  $\tilde{\mathbf{x}}_k$ .

Step 6 If  $k < N_t$ , increment  $k$  and go to Step 2.

We refer to this scheme as *one-vector-symbol detection* as we detect  $\mathbf{x}_i$  using  $\mathbf{y}_i$  only. Note that the above algorithm is the modified version of the algorithm in Section 2.2.3, where the true channel and interference spatial correlation matrices are replaced by the estimates.

When interference is temporally colored, there is performance to be gained by taking the temporal interference correlation into account. That is, we may use  $\mathbf{y}_{N+1}, \dots, \mathbf{y}_M$  to detect  $\mathbf{x}_{N+1}, \dots, \mathbf{x}_M$  jointly where  $N$  is the training length and  $M$  is the frame length. Due to the complexity of using all the received signal vectors and for simplicity of presentation, we consider *two-vector-symbol detection* in which  $(\mathbf{y}_i, \mathbf{y}_{i+1})$  is used to detect  $(\mathbf{x}_i, \mathbf{x}_{i+1})$  jointly. The one-vector-symbol algorithm can be easily extended to the two-vector-symbol version by writing

$$\underbrace{\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_{i+1} \end{bmatrix}}_{\check{\mathbf{y}}_i} = \underbrace{\begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}}_{\check{\mathbf{H}}} \underbrace{\begin{bmatrix} \mathbf{x}_i \\ \mathbf{x}_{i+1} \end{bmatrix}}_{\check{\mathbf{x}}_i} + \underbrace{\begin{bmatrix} \mathbf{n}_i \\ \mathbf{n}_{i+1} \end{bmatrix}}_{\check{\mathbf{n}}_i}.$$

With the estimated channel, an estimate of  $\check{\mathbf{H}}$ , denoted as  $\hat{\check{\mathbf{H}}}$ , can be obtained. Using the estimated spatial interference correlation and the known temporal interference correlation, we are able to estimate the covariance matrix of  $\check{\mathbf{n}}_i$ , denoted as  $\hat{\check{\mathbf{R}}}$ . Replacing  $\mathbf{x}_i, \mathbf{y}_i, \hat{\mathbf{H}}$  and  $\hat{\mathbf{R}}$  in the one-vector-symbol algorithm by  $\check{\mathbf{x}}_i, \check{\mathbf{y}}_i, \hat{\check{\mathbf{H}}}$  and  $\hat{\check{\mathbf{R}}}$ , respectively, we obtain the two-vector-symbol detection algorithm.

## 4.5 Applications

In this section, we apply the channel estimation in Section 4.3 and data detection in Section 4.4 to the case of a single-user link with one dominant cochannel interferer operating at different data rates.

### 4.5.1 System model

Consider a desired user with one dominant cochannel interferer. The assumption of one cochannel interferer is reasonable in cellular TDMA or FDMA systems especially when

sectoring is used. In 7-cell reuse systems, with 60 degree sectors, the number of co-channel interfering cells would be reduced to one [91]. Hence, there will be only one co-channel interferer. We assume that the desired and interfering users have  $N_t$  and  $L$  transmitting antennas, respectively, and that there are  $N_r$  receiving antennas. Assuming thermal noise is small relative to interference, we ignore thermal noise in the formulation of the problem. The accuracy of this assumption in a noisy channel will be verified in Section 4.5.4. In a slow flat fading environment, the vector signal at receiving antennas is

$$\mathbf{y}(t) = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \sum_{k=0}^{M-1} \mathbf{x}_k \tilde{g}(t - kT) + \sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k \tilde{g}_I(t - kT_I - \tau) \quad (4.24)$$

where  $M$  is the frame length, and  $\mathbf{H}$  ( $N_r \times N_t$ ) and  $\mathbf{H}_I$  ( $N_r \times L$ ) are the channel matrices of the desired and interfering users, respectively. The channel matrices are assumed fixed over a frame, and have independent realizations from frame to frame. The data transmission rates of the desired and interfering users are  $1/T$  and  $1/T_I$ , respectively. The spectra of transmit impulse responses  $\tilde{g}(t)$  and  $\tilde{g}_I(t)$  are square-root raised cosines with parameters  $T$  and  $T_I$ , respectively. The same rolloff factor,  $\beta$ , is assumed for both  $\tilde{g}(t)$  and  $\tilde{g}_I(t)$ . The data vectors of the desired and interfering users are  $\mathbf{x}_k$  ( $N_t \times 1$ ) and  $\mathbf{b}_k$  ( $L \times 1$ ), respectively. We assume that data symbols in  $\mathbf{x}_k$ 's and  $\mathbf{b}_k$ 's are mutually independent, zero-mean and with unit variance. We denote  $P_s$  and  $P_I$  as the transmit powers of the desired and interfering users, respectively. The delay of the interfering user relative to the desired user is  $\tau$ , assumed to lie in  $0 \leq \tau < \max(T, T_I)$ .

Passing  $\mathbf{y}(t)$  in (4.24) through a filter matched to the transmit impulse response of the desired user,  $\tilde{g}(t)$ , the vector signal at the output of the matched filter is

$$\mathbf{y}_{\text{MF}}(t) = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \sum_{k=0}^{M-1} \mathbf{x}_k g(t - kT) + \sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k g_I(t - kT_I - \tau) \quad (4.25)$$

where  $g(t) = \tilde{g}(t) * \tilde{g}(t)$  and  $g_I(t) = \tilde{g}_I(t) * \tilde{g}_I(t)$ . Recall that  $\tilde{g}(t)$  has a square-root raised-cosine spectrum, hence,  $g(t)$  has a raised-cosine spectrum and satisfies Nyquist condition

for zero intersymbol interference.

Assuming perfect synchronization for the desired user, as we sample the output of the matched filter (4.25) at time  $t = jT$ , we obtain

$$\mathbf{y}_j = \sqrt{\frac{P_s T}{N_t}} \mathbf{H} \mathbf{x}_j + \underbrace{\sqrt{\frac{P_I T_I}{L}} \mathbf{H}_I \sum_{k=-\infty}^{\infty} \mathbf{b}_k g_I(jT - kT_I - \tau)}_{\mathbf{n}_j}. \quad (4.26)$$

The interference vector  $\mathbf{n}_j$  is zero-mean as the data vector of interfering user  $\mathbf{b}_k$  is zero-mean. Note that there is no intersymbol interference for the desired user. However, for the interferer, due to the delay and/or mismatch between the transmit and receive impulse responses, intersymbol interference exists.

## 4.5.2 Interference statistics

The cross-correlation between the interference vectors in (4.26) at times  $jT$  and  $qT$  is

$$\begin{aligned} & E \left\{ \mathbf{n}_j \mathbf{n}_q^\dagger \right\} \\ &= \frac{P_I T_I}{L} \mathbf{H}_I \cdot E \left\{ \left( \sum_{k_1=-\infty}^{\infty} \mathbf{b}_{k_1} g_I(jT - k_1 T_I - \tau) \right) \left( \sum_{k_2=-\infty}^{\infty} \mathbf{b}_{k_2}^\dagger g_I(qT - k_2 T_I - \tau) \right) \right\} \mathbf{H}_I^\dagger \\ &= \frac{P_I T_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger \cdot \sum_{k=-\infty}^{\infty} \left\{ g_I(jT - kT_I - \tau) g_I(qT - kT_I - \tau) \right\}, \end{aligned}$$

where the last equality is due to the fact that  $E \left\{ \mathbf{b}_{k_1} \mathbf{b}_{k_2}^\dagger \right\} = \mathbf{0}$  for  $k_1 \neq k_2$  and that  $E \left\{ \mathbf{b}_k \mathbf{b}_k^\dagger \right\} = \mathbf{I}_L$ .

During a training period of  $N$  vector symbols, the covariance matrix of the concatenated interference vector  $\bar{\mathbf{n}} = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$  has the form of a Kronecker product in (4.2) where

$$\mathbf{\Lambda}_N(j, q) = \sum_{k=-\infty}^{\infty} \left\{ g_I(jT - kT_I - \tau) g_I(qT - kT_I - \tau) \right\}, \quad 0 \leq j, q \leq N-1 \quad (4.27)$$

and

$$\mathbf{R} = \frac{P_I T_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger. \quad (4.28)$$

The  $N_r \times N_r$  spatial correlation matrix  $\mathbf{R}$  is determined by the interferer's channel matrix. The  $N \times N$  temporal correlation matrix  $\mathbf{\Lambda}_N$  depends on parameters  $T$  and  $T_I$ , delay  $\tau$  and pulse  $g_I(t)$ , and can be calculated a priori if we have knowledge of these parameters. The temporal correlation is due to the intersymbol interference in the sampled interfering signal. We remark that for the case of multiple interferers with the same delay, the covariance matrix of interference also has the form of a Kronecker product.

We study the temporal interference correlation matrices of two cases: (1) interferer at the same data rate as the desired signal ( $T = T_I$ ), and (2) interferer at a lower data rate than the desired signal, i.e., the data rate of the desired user is an integer multiple of that of the interferer ( $T_I = mT$ ,  $m > 1$ ).

#### 4.5.2.1 Interferer at the same data rate as desired signal

The desired and interfering users have the same transmit impulse response whose spectrum is the square-root of a raised cosine with parameter  $T$  and rolloff factor  $\beta$ . Therefore,  $g_I(t)$  has a raised-cosine spectrum, and it is given by

$$g_I(t) = \text{sinc}(\pi t/T) \frac{\cos(\pi \beta t/T)}{1 - 4\beta^2 t^2/T^2}.$$

The temporal correlation of interference vectors at times  $jT$  and  $qT$  is

$$\mathbf{\Lambda}_N(j, q) = \sum_{k=-\infty}^{\infty} \left\{ g_I(jT - kT - \tau) g_I(qT - kT - \tau) \right\}. \quad (4.29)$$

We note that  $\mathbf{\Lambda}_N(j, q)$  depends on  $j - q$ . This indicates that the sequence consisting of interference vectors is stationary. Hence, the temporal correlation matrix is symmetric Toeplitz. By appropriate truncation of the infinite series in (4.29), we can numerically calculate the temporal correlation matrix. For the case of  $\beta = 1$ ,  $T = 1$ , here are two examples of temporal correlation matrices at different interferer delays.

1.  $\tau = 0.5$ , the element of the temporal correlation matrix is

$$\Lambda_N(j, q) = \begin{cases} 0.5 & j = q \\ 0.25 & |j - q| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \leq j, q \leq N - 1. \quad (4.30)$$

2.  $\tau = 0.25$ , the element of the temporal correlation matrix is

$$\Lambda_N(j, q) = \begin{cases} 0.75 & j = q \\ 0.125 & |j - q| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \leq j, q \leq N - 1.$$

#### 4.5.2.2 Interferer at a lower data rate than desired signal

The desired and interfering users employ the transmit impulse pulses whose spectrums are square-root of raised cosines with the same rolloff factor  $\beta$  and parameters  $T$  and  $T_I$ , respectively. The waveform of  $g_I(t)$  is given by

$$g_I(t) = \mathcal{F}^{-1} \left\{ \sqrt{G_{rc, T_I}(f)} \sqrt{G_{rc, T}(f)} \right\}$$

where  $\mathcal{F}^{-1}$  denotes the inverse Fourier transform, and  $G_{rc, T}(f)$  is the raised-cosine Fourier spectrum with parameter  $T$  and rolloff factor  $\beta$ . With  $T_I = mT$ , the temporal correlation of interference vectors at times  $jT$  and  $qT$  is

$$\Lambda_N(j, q) = \sum_{k=-\infty}^{\infty} \left\{ g_I(jT - mkT - \tau) g_I(qT - mkT - \tau) \right\}. \quad (4.31)$$

Unlike in the case of same-data-rate interferer where  $\Lambda_N(j, q)$  depends on  $j - q$ , in the case of lower-data-rate interferer,  $\Lambda_N(j, q)$  depends on the values of  $j$  and  $q$ . This indicates that the sequence consisting of interference vectors is cyclostationary. A sequence is said to be cyclostationary in the wide sense if its mean and autocorrelation are periodic [40, 89]. It can be seen that  $\Lambda_N(j, q)$  is periodic with period  $m$ , i.e.,  $\Lambda_N(j, q) = \Lambda_N(j + m, q + m)$ . As

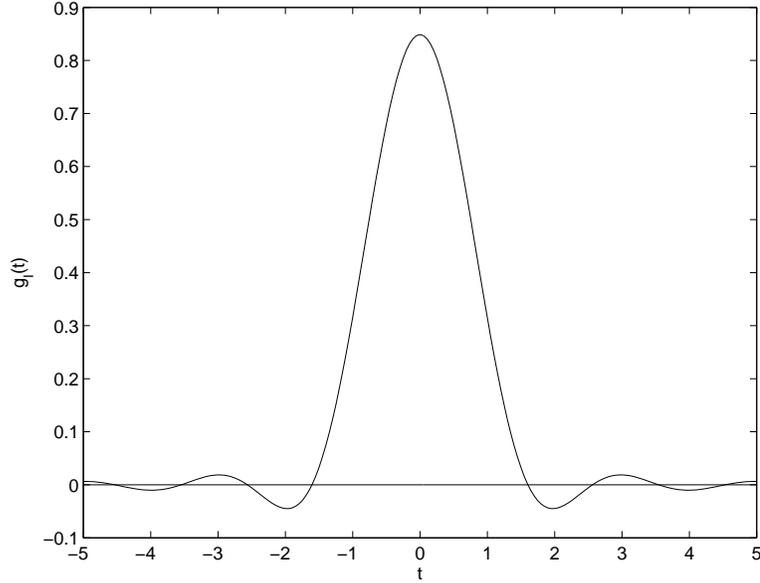


Figure 4.1. Waveform of  $g_I(t)$  for the case of lower-data-rate interferer with  $T_I = 2T$ ,  $T = 1$  and  $\beta = 1$ .

a result, the temporal correlation matrix  $\mathbf{\Lambda}_N$  is symmetric, but not Toeplitz. Furthermore, for  $N \geq m$ , the number of nontrivial eigenvalues of  $\mathbf{\Lambda}_N$  is  $\lceil N/m \rceil$  where  $\lceil \cdot \rceil$  rounds the argument to the nearest integer towards infinity [72]; therefore,  $\mathbf{\Lambda}_N$  is singular.

By appropriate truncation of the infinite series in (4.31), we can numerically calculate the temporal correlation matrix. For the case of  $T_I = 2T$ ,  $T = 1$  and  $\beta = 1$ , the waveform of  $g_I(t)$  is given by

$$g_I(t) = \int_{-1/2}^{1/2} \sqrt{[1 + \cos(2\pi f)] \cdot \frac{1}{2} [1 + \cos(\pi f)]} e^{j2\pi f t} df,$$

and it is shown in Fig. 4.1. The following are three examples of temporal correlation matrices with training length  $N = 8$  at different interferer delays:

1.  $\tau = 0$ ,

$$\Lambda_8 = \begin{bmatrix} 0.725 & 0.252 & -0.075 & -0.002 & -0.015 & -0.000 & -0.006 & -0.000 \\ 0.252 & 0.200 & 0.252 & 0.111 & -0.002 & 0.016 & -0.000 & 0.007 \\ -0.075 & 0.252 & 0.725 & 0.252 & -0.075 & -0.002 & -0.015 & -0.000 \\ -0.002 & 0.111 & 0.252 & 0.200 & 0.252 & 0.111 & -0.002 & 0.016 \\ -0.015 & -0.002 & -0.075 & 0.252 & 0.725 & 0.252 & -0.075 & -0.002 \\ -0.000 & 0.016 & -0.002 & 0.111 & 0.252 & 0.200 & 0.252 & 0.111 \\ -0.006 & -0.000 & -0.015 & -0.002 & -0.075 & 0.252 & 0.725 & 0.252 \\ -0.000 & 0.007 & -0.000 & 0.016 & -0.002 & 0.111 & 0.252 & 0.200 \end{bmatrix}.$$

2.  $\tau = 0.25$ ,

$$\Lambda_8 = \begin{bmatrix} 0.648 & 0.400 & -0.048 & -0.006 & -0.010 & -0.001 & -0.004 & -0.000 \\ 0.400 & 0.277 & 0.105 & 0.084 & 0.002 & 0.011 & 0.000 & 0.005 \\ -0.048 & 0.105 & 0.648 & 0.400 & -0.048 & -0.006 & -0.010 & -0.001 \\ -0.006 & 0.084 & 0.400 & 0.277 & 0.105 & 0.084 & 0.002 & 0.011 \\ -0.010 & 0.002 & -0.048 & 0.105 & 0.648 & 0.400 & -0.048 & -0.006 \\ -0.001 & 0.011 & -0.006 & 0.084 & 0.400 & 0.277 & 0.105 & 0.084 \\ -0.004 & 0.000 & -0.010 & 0.002 & -0.048 & 0.105 & 0.648 & 0.400 \\ -0.000 & 0.005 & -0.001 & 0.011 & -0.006 & 0.084 & 0.400 & 0.277 \end{bmatrix}.$$

(4.32)

3.  $\tau = 0.5$ ,

$$\mathbf{\Lambda}_8 = \begin{bmatrix} 0.462 & 0.461 & 0.018 & -0.008 & 0.001 & -0.001 & 0.000 & -0.000 \\ 0.461 & 0.462 & 0.044 & 0.018 & 0.004 & 0.001 & 0.001 & 0.000 \\ 0.018 & 0.044 & 0.462 & 0.461 & 0.018 & -0.008 & 0.001 & -0.001 \\ -0.008 & 0.018 & 0.461 & 0.462 & 0.044 & 0.018 & 0.004 & 0.001 \\ 0.001 & 0.004 & 0.018 & 0.044 & 0.462 & 0.461 & 0.018 & -0.008 \\ -0.001 & 0.001 & -0.008 & 0.018 & 0.461 & 0.462 & 0.044 & 0.018 \\ 0.000 & 0.001 & 0.001 & 0.004 & 0.018 & 0.044 & 0.462 & 0.461 \\ -0.000 & 0.000 & -0.001 & 0.001 & -0.008 & 0.018 & 0.461 & 0.462 \end{bmatrix}.$$

### 4.5.3 Data detection without estimating channel and interference

During a training period of  $N$  symbol vectors, instead of estimating the channel matrix and interference statistics, one can alternatively employ a least squares (LS) estimate of matrix  $\mathbf{M}$  which minimizes the average estimation error

$$f_2(\mathbf{M}) = \text{trace} \left\{ \frac{1}{N} \sum_{i=0}^{N-1} (\mathbf{x}_i - \mathbf{M}\mathbf{y}_i)(\mathbf{x}_i - \mathbf{M}\mathbf{y}_i)^\dagger \right\}.$$

By setting  $\partial f_2(\mathbf{M})/\partial \mathbf{M} = \mathbf{0}$ , we obtain

$$\mathbf{M} = \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1} \quad (4.33)$$

where sample correlation matrices  $\mathbf{R}_{xy}$  and  $\mathbf{R}_{yy}$  are defined in (4.20) and (4.19), respectively. The transmitted signal vector  $\mathbf{x}_i$  is detected as  $Q(\mathbf{M}\mathbf{y}_i)$  where  $Q(\cdot)$  is the quantization appropriate to the signal constellation. We remark that (4.33) is the well-known Direct Matrix Inversion (DMI) algorithm [121] generalized for multiple input signals. A significant loss in performance is expected for this LS detector since without estimates of channel and spatial interference correlation matrices, iterative MMSE detection cannot be performed.

#### 4.5.4 Simulation results

Monte Carlo simulations are used to assess the benefits of taking temporal and spatial interference correlation into account for channel estimation and data detection in the case of one interferer. Although it is difficult to know the temporal interference correlation in practice, it is of interest to examine the performance loss due to ignoring this correlation.

We evaluate average symbol error rates (SERs) over different realizations of the desired and interfering users' channel matrices. The channel model in Section 2.3 is used to simulate the channel matrices. Uniform linear arrays are employed. Independent Rayleigh fading is assumed for the interfering user modelling a lack of a line-of-sight component and severe fading for signals coming from other cells. Hence, the elements of channel matrix  $\mathbf{H}_I$  are i.i.d. zero-mean circularly symmetric complex Gaussian with unit variance. For the desired user, the angle of departure of the specular signal at the mobile and the angle of arrival at the base station are 0 and 60 degrees with respect to the normals of each array, respectively; these angles are the same across the array. The antenna spacings at the mobile and the base station are 0.5 and 10 wavelengths, respectively. In the simulation, both independent and correlated MIMO links are considered for the desired user.

We assume that the desired user has 5 transmitting and 5 receiving antennas, and the interfering user has 6 transmitting antennas<sup>3</sup>. Both the desired and interfering users employ uncoded QPSK modulation. The training signal vectors are taken to be columns of an FFT matrix [55]. This guarantees that the training sequences from different transmitting antennas are orthogonal. We define  $\text{SIR}(\text{dB}) = 10 \log P_s / P_I$ . Without loss of generality, we let  $P_I = 1$  in the simulation. The SERs of two cases are simulated: (1) interferer at the same data rate as the desired signal, and (2) interferer at a lower data rate than the desired

---

<sup>3</sup>To make the spatial interference correlation matrix nonsingular, we let the number of receiving antennas be no greater than the number of transmitting antennas of the interfering user, i.e.,  $N_r \leq L$ .

signal, i.e., the data rate of the desired user is twice that of the interferer.

In Figs. 4.2 to 4.7, with solid and dashed lines representing one- and two-vector-symbol data detection, respectively, we show the average SERs for the following cases:

- perfectly known channel parameters and interference statistics, with one-vector-symbol (curve 1) and two-vector-symbol (curve 2) detection;
- channel and spatial interference correlation matrices are estimated assuming known temporal interference correlation, with one-vector-symbol (curve 3) and two-vector-symbol (curve 4) detections;
- channel and spatial interference correlation matrices are estimated assuming temporally white interference, with one-vector-symbol detection (curve 5);
- only the channel matrix  $\mathbf{H}$  is estimated assuming temporally white interference; an identity spatial interference correlation matrix is used in one-vector-symbol data detection (curve 6).
- least-squares (LS) estimate of the transmitted signal vector without ordered detection (Section 4.5.3) (curve 7).

The first case is presented for reference. The fourth case corresponds to the current BLAST system [47] [55].

#### **4.5.4.1 Interferer at the same data rate as desired signal**

We examine the case of  $T = 1$ ,  $\beta = 1$ ,  $\tau = 1/2$ , and the nonsingular temporal interference correlation matrix shown in (4.30). Figs. 4.2 to 4.4 show the average SERs for training lengths  $2N_t$ ,  $4N_t$  and  $6N_t$ , respectively. Independent Rayleigh fading is assumed for the

desired user. Comparing the LS detection (curve 7) with other methods, as expected, much lower symbol error rates can be achieved by using ordered MMSE detection.

Comparing curves 5 and 6, we observe that for training lengths  $4N_t$  and  $6N_t$ , gains can be obtained by estimating spatial interference correlation. However, for a short training length, for example,  $2N_t$ , due to inaccurate estimates of spatial interference correlation, it is better to estimate only the channel matrix and assume spatially white interference. The performance gap between curves 5 and 6 increases as the training lengths increase due to better estimates of interference spatial correlation obtained with longer training lengths.

By examining curves 3 and 5, we observe that the improvement in taking temporal interference correlation into account in *channel estimation* is not significant, and this improvement decreases as the training length increases. This phenomenon can be explained as follows: in estimating channel and temporal interference correlation matrices for temporally colored interference, the received signal vectors first undergo a transformation which temporally whitens the interference vectors as discussed in Section 4.3.3. Since the temporal correlation in (4.30) drops quickly to zero after one time lag, the benefit in the temporal whitening of interference vectors is not significant, especially for long training lengths.

By comparing curves 3 and 4, there is a slight improvement of two-vector-symbol over one-vector-symbol detection. This implies that not much gain can be achieved by taking temporal interference correlation into account in *data detection* owing to the low temporal correlation. Due to better estimates of channel and interference spatial correlation matrices obtained with a longer training length, the performance gap between curves 3 and 4 increases as the training length increases.

In Fig. 4.3, for training length  $4N_t$ , by comparing curves 4 and 6, we observe a total of 1.5dB gain in SIR by estimating spatial interference correlation and taking explicit advantage of known temporal interference correlation in channel estimation and data detection.

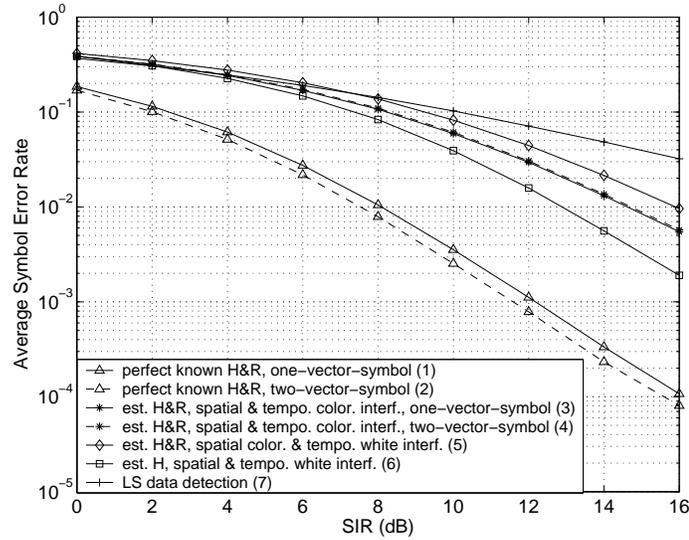


Figure 4.2. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $2N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate.

About 1dB of the gain is due to the estimation of spatial interference correlation, and the remaining 0.5dB gain is due to exploiting temporal interference correlation in channel estimation and data detection. We also observe that training length  $4N_t$  achieves most of the performance gain.

#### 4.5.4.2 Interferer at a lower data rate than desired signal

We examine the case of  $T_I = 2T$ ,  $T = 1$ ,  $\beta = 1$ ,  $\tau = 0.25$  and the temporal interference correlation matrix for training length 8 shown in (4.32). Recall that the temporal correlation matrix in the lower-data-rate-interferer case is singular. To avoid the singularity, the diagonal elements of  $\mathbf{\Lambda}_N$  are increased by a small amount; hence, the temporal correlation matrix used for channel estimation may be modified to  $\mathbf{\Lambda}_N + \delta \mathbf{I}_N$  within the proposed framework. In our simulation, we chose  $\delta = 0.01$ .

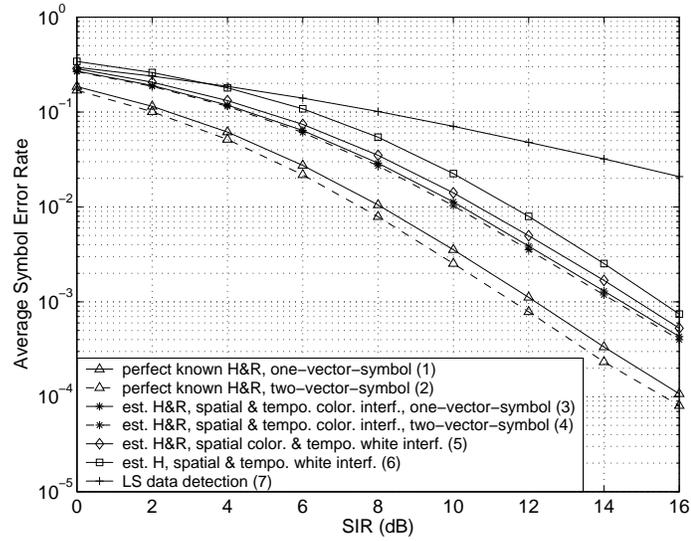


Figure 4.3. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $4N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate.

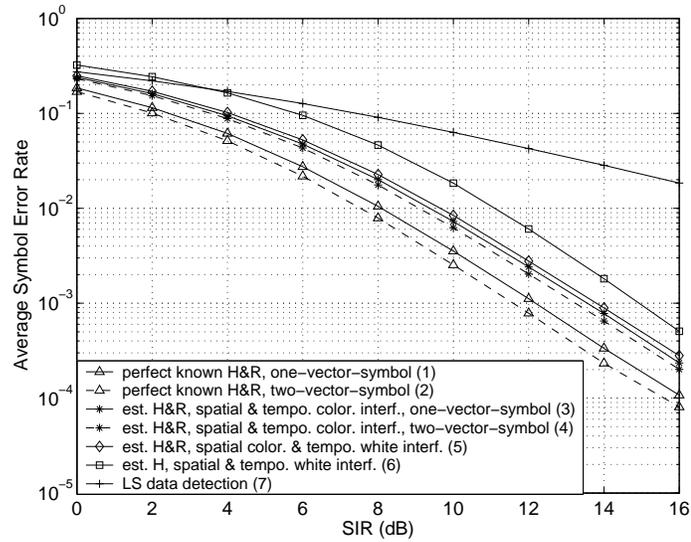


Figure 4.4. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $6N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate.

The same set of average SER curves as in the same-data-rate-interferer case are simulated. Figs. 4.5 to 4.7 show the SERs for different training lengths. Independent Rayleigh fading is assumed for the desired user. As in the case of the same-data-rate interferer, curve 7 illustrates the poor performance without ordered detection. Curves 5 and 6 suggest that for short training lengths it is better to estimate only the channel matrix and assume spatially white interference in data detection; however, for moderately long training lengths, gains can be obtained by estimating spatial interference correlation.

By examining curves 3 and 5, we observe that the improvement in taking temporal interference correlation into account in *channel estimation* is not that significant. However, this improvement is larger than that in the same-data-rate-interferer case due to the high temporal correlation in the lower-data-rate-interferer case.

In contrast to the same-data-rate-interferer case, curves 3 and 4 show that the improvement of two-vector-symbol over one-vector-symbol detection is significant due to the higher temporal interference correlation. This implies that significant gain can be achieved by taking temporal interference correlation into account in *data detection* for the lower-data-rate-interferer case.

By comparing curves 4 and 6 in Fig. 4.6, for training length  $4N_t$ , there is a total of 4dB gain in SIR by estimating spatial interference correlation and taking advantage of known temporal interference correlation in channel estimation and data detection. About 3.5dB of the gain is due to exploiting temporal interference correlation in channel estimation and data detection. Hence, when there is lower-data-rate interference, considerable performance loss occurs if we ignore the temporal interference correlation. In addition, as in the same-data-rate-interferer case, we observe that training length  $4N_t$  achieves most of the performance gain.

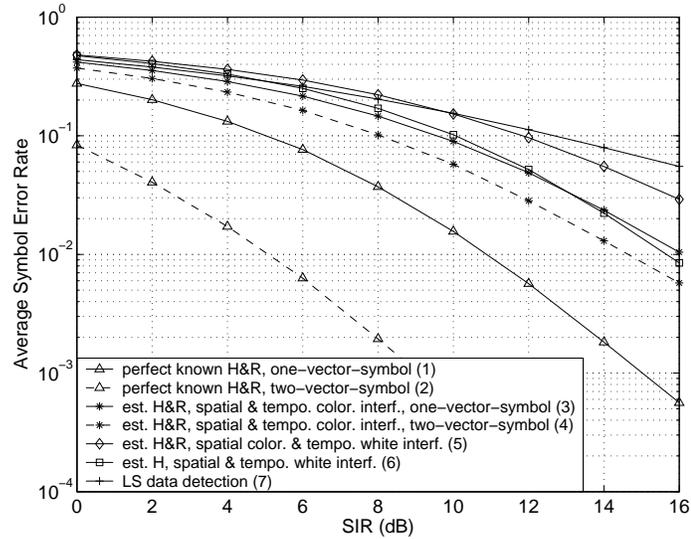


Figure 4.5. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $2N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user.

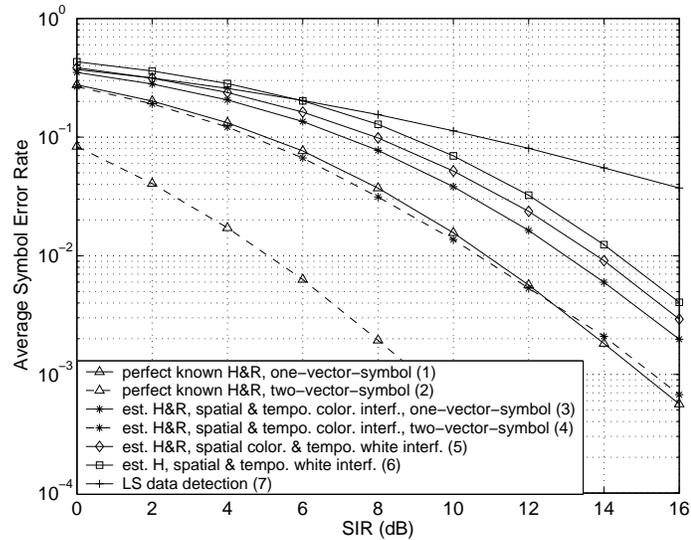


Figure 4.6. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $4N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user.

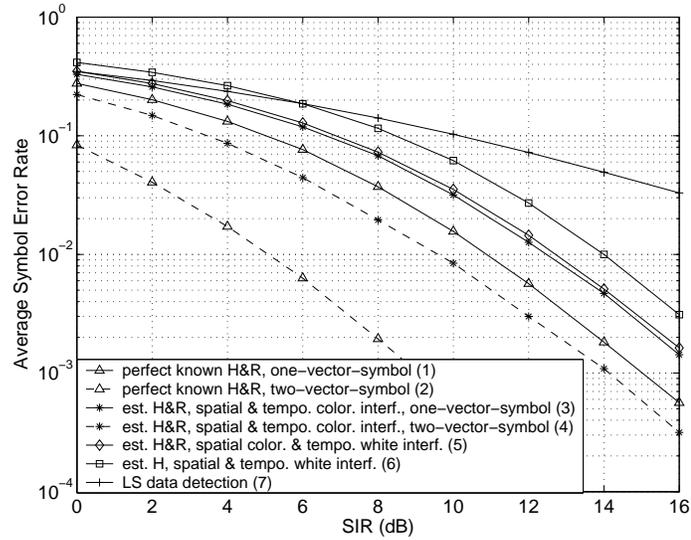


Figure 4.7. Average symbol error rate vs. SIR with  $N_t = N_r = 5$ ,  $L = 6$ , and training length  $6N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user.

#### 4.5.4.3 Correlated channels for the desired user

As mentioned in Section 2.3, MIMO links become correlated as the angle spread decreasing or the Ricean  $K$  factor increasing. Figs. 4.8 and 4.9 show that the SERs are improved as the desired user's angle spread at the base station increases for both the same- and lower-data-rate-interferer cases with SIR=10dB, training length  $4N_t$ . Correlated Rayleigh fading (Ricean factor  $K = 0$ ) is assumed for the desired user. The improvement in SER is due to the diversity gain as the channels become uncorrelated. We observe that, for angle spreads larger than 6 degrees, the performance is close to the case where MIMO links are independently faded. It is also observed that as the angle spread decreases, the SER curves obtained from channel estimates approach the curves from perfectly known channel. Therefore, the estimates of channel and spatial interference correlation matrices become more accurate as the angle spread decreases.

Figs. 4.10 and 4.11 show that SERs degrade as  $K$ -Ricean factor increases for both the same- and lower-data-rate-interferer cases with SIR=10dB, training length  $4N_t$ . The channel scattered components are independently faded for the desired user. The degradation in SERs is due to the loss of diversity gain as channels become correlated. Hence, the line-of-sight component does not improve the SERs of a MIMO system, which agrees with the degradation of information capacity observed in [30]. However, if a single data stream is transmitted simultaneously from the multiple antennas with properly selected transmit beamforming weights, we would expect that the SER would improve as the  $K$ -Ricean factor increases [31]. In Figs. 4.10 and 4.11, we also observe that as the Ricean factor increases, the SER curves obtained from channel estimates approach the curves from perfectly known channel. Therefore, the estimates of channel and spatial interference correlation matrices become more accurate as  $K$  increases.

#### 4.5.4.4 Validity of the decoupled structure of interference statistics in the presence of noise

In the presence of noise, for temporally white interference, the Kronecker product structure of interference statistics still holds; however, for temporally colored interference, the interference statistics can only be approximated by the Kronecker product. Hence, it is of interest to examine the effect of this approximation. We model the thermal noise as a zero-mean circularly symmetric complex Gaussian vector with covariance matrix  $\sigma^2 \mathbf{I}_{N_r}$ , i.e, the thermal noise is independent from antenna to antenna, and the noise power on each antenna is  $\sigma^2$ . We define interference-to-noise-power-ratio  $\text{INR} = 10 \log P_I / \sigma^2$ , where  $P_I = 1$  in the simulation. Assume independent Rayleigh fading for the desired user, training length  $4N_t$  and SIR=10dB. For the case of the same-data-rate interferer, referring to curves 3 and 5 in Fig. 4.12, we observe that for INRs lower than 17dB (noise higher than a certain

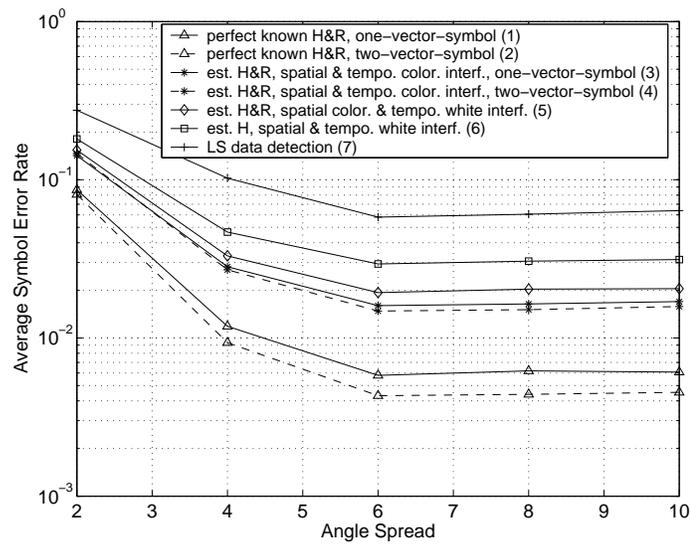


Figure 4.8. Average symbol error rate vs. angle spread with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, and training length  $4N_t$ . Correlated Rayleigh fading (Ricean factor  $K = 0$ ) is assumed for the desired user. Both the desired and interfering users are at the same data rate.

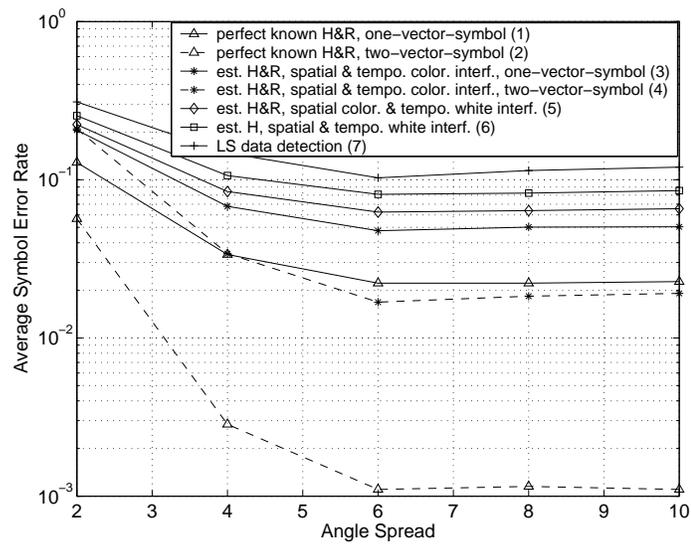


Figure 4.9. Average symbol error rate vs. angle spread with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, and training length  $4N_t$ . Correlated Rayleigh fading (Ricean factor  $K = 0$ ) is assumed for the desired user. The data rate of the desired user is twice that of interfering user.

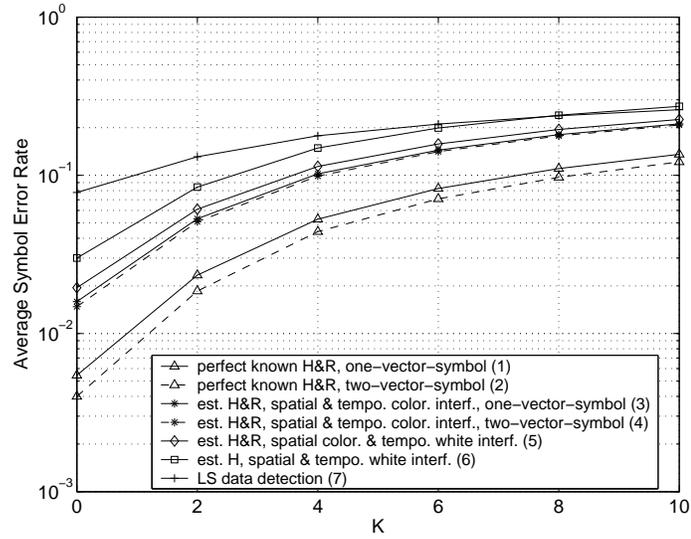


Figure 4.10. Average symbol error rate vs.  $K$ -Ricean factor with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, training length  $4N_t$ . The channel scattered components are independently faded for the desired user. Both the desired and interfering users are at the same data rate.

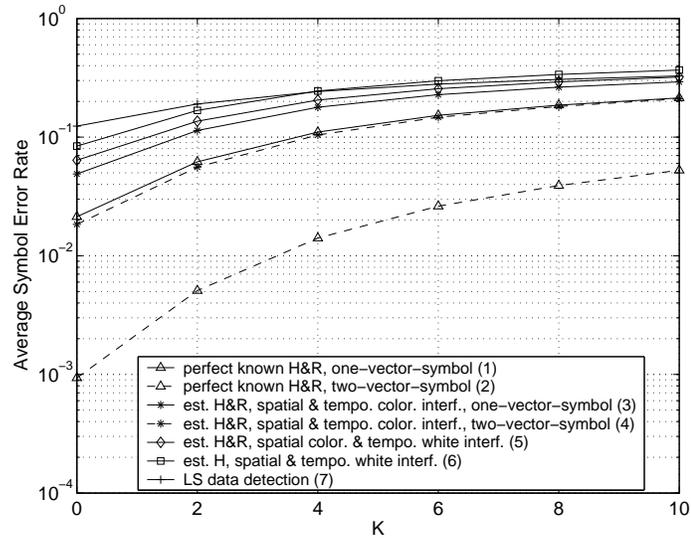


Figure 4.11. Average symbol error rate vs.  $K$ -Ricean factor with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, and training length  $4N_t$ . The channel scattered components are independently faded for the desired user. The data rate of the desired user is twice that of interfering user.

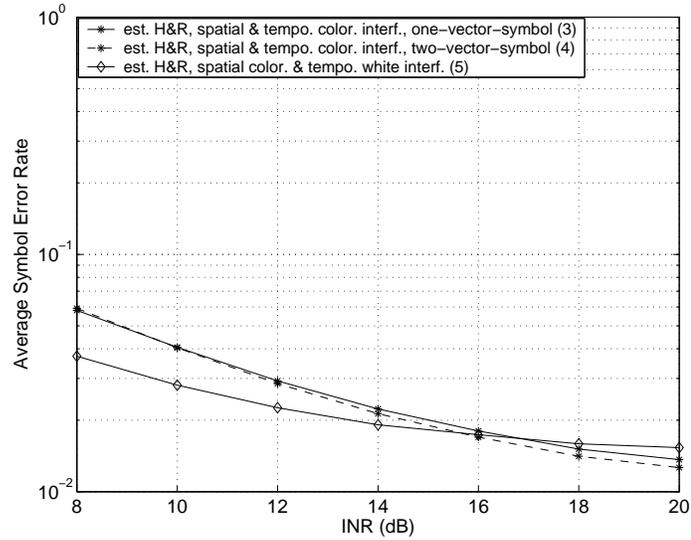


Figure 4.12. Average symbol error rate vs. INR with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, and training length  $4N_t$ . Independent Rayleigh fading is assumed for the desired user. Both the desired and interfering users are at the same data rate.

level), ignoring interference temporal correlation leads to smaller SERs compared to the case of assuming interference temporal correlation of (4.30). This suggests that the decoupled interference statistics with temporal correlation of (4.30) is an accurate approximation for INRs higher than 17dB. Similarly, Fig. 4.13 shows that, in the case of the lower-data-rate interferer, the decoupled structure of interference statistics is a valid approximation for INRs higher than 12dB.

#### 4.5.4.5 Effect of exploiting knowledge of spatial interference-plus-noise correlation

Since the temporal interference correlation is hard to know in practice, it is of interest to assess the improvement of estimating the spatial correlation of interference-plus-noise over the case of assuming the interference-plus-noise to be spatially white. With the total interference power fixed, Fig. 4.14 compares the average SER for one (solid line) and two

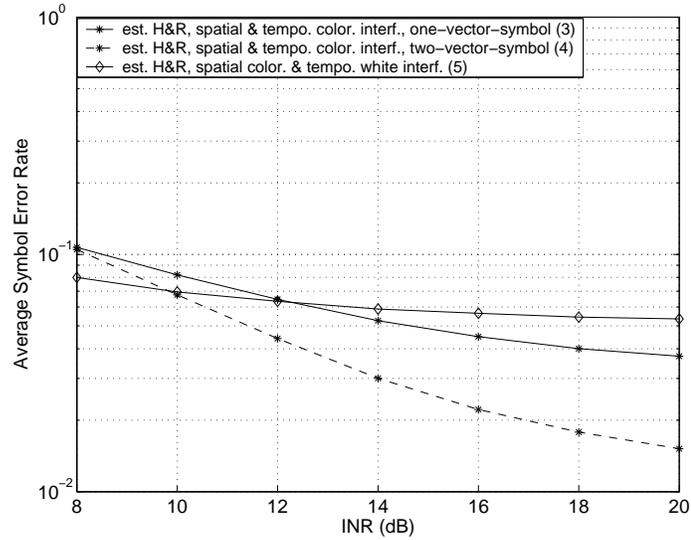


Figure 4.13. Average symbol error rate vs. INR with  $N_t = N_r = 5$ ,  $L = 6$ , SIR=10dB, and training length  $4N_t$ . Independent Rayleigh fading is assumed for the desired user. The data rate of the desired user is twice that of interfering user.

(broken line) interferers. In the case of two interferers, the interferers have the same power and are asynchronous due to random delays. Both the desired and interfering users employ a (5,5) MIMO link, the total-interference-to-noise-ratio is 12dB, and the training length is  $4N_t$ . It is assumed that both the desired and interfering users operate at the same data rate. Fig. 4.14 shows that for one interferer, there is 1.2dB gain over a wide range of SINRs by estimating the spatial correlation of interference-plus-noise; while for two interferers, the corresponding gain in SINR is negligible as the two curves are nearly superimposed. Thus, with a (5,5) MIMO link and two interferers, the interference is accurately modelled as spatially white.

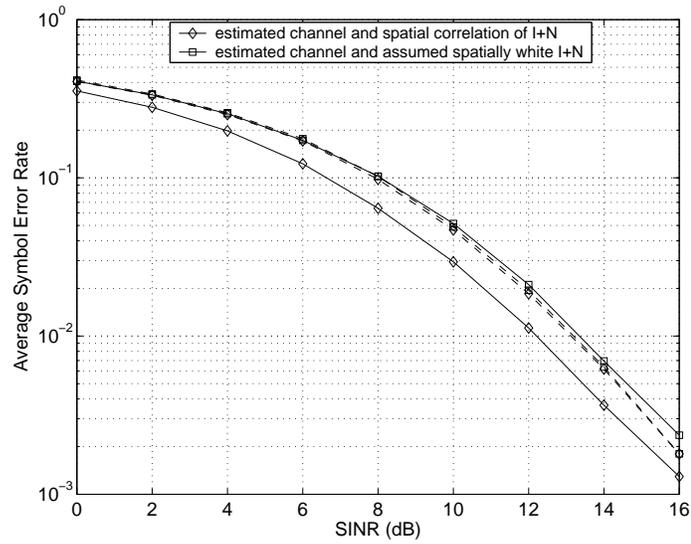


Figure 4.14. The improvement of estimating spatial correlation of interference-plus-noise in practical systems. With total interference power fixed, the solid lines are for one interferer, and the broken lines are for two interferers. Both the desired and interfering users employ a  $(5, 5)$  MIMO link, the same data rate, total-interference-to-noise-ratio of 12dB, and training length of  $4N_t$ .

#### 4.5.4.6 Confidence intervals

To obtain an accurate estimate of the actual SER in simulations, it is necessary to ensure that a sufficient number of data symbols have been processed. The number of symbols required can be determined from the confidence interval<sup>4</sup> for the estimated SER.

Let  $S_E$  and  $\hat{S}_E$  denote the true and estimated SER values and  $N_s$  the number of data symbols processed in the simulation, an approximate  $q\%$  confidence interval for  $\hat{S}_E$  can be computed as [29, p. 564]

$$\left( \hat{S}_E - Q \sqrt{\frac{\hat{S}_E(1 - \hat{S}_E)}{N_s}}, \hat{S}_E + Q \sqrt{\frac{\hat{S}_E(1 - \hat{S}_E)}{N_s}} \right) \quad (4.34)$$

where the value of  $Q$  depends on  $q$  as shown in Table 4.1. In (4.34), the width of the confidence interval is

$$W = 2Q \sqrt{\frac{\hat{S}_E(1 - \hat{S}_E)}{N_s}}. \quad (4.35)$$

From (4.35), we have

$$N_s = \frac{4\hat{S}_E(1 - \hat{S}_E)Q^2}{W^2}. \quad (4.36)$$

This gives the number of symbols required to obtain the desired confidence interval width. For example, with  $\hat{S}_E = 10^{-3}$ ,  $W = 2 \times 10^{-4}$  and a 95% confidence interval, from (4.36), we need to simulate about 384000 symbols. With five transmitter antennas in the MIMO link, 76800 vector symbols are required. In the simulations, we have simulated a sufficient number of data symbols for a 95% confidence interval and the confidence interval width equal to 20% of  $\hat{S}_E$ .

---

<sup>4</sup>For an estimation of SER, we can compute a  $q\%$  confidence interval. If we compute the  $q\%$  confidence interval for a large number of estimations of SER,  $q\%$  of the obtained confidence intervals will include the true SER.

$q$	90%	95%	99%
$Q$	1.65	1.96	2.58

Table 4.1. Corresponding values of  $Q$  for various confidence levels.

## 4.6 Conclusions

In this chapter, by modelling interference statistics as being approximately temporally and spatially separable, we have proposed an algorithm to jointly estimate channel and spatial interference correlation matrices. By exploiting the temporal interference correlation, one-vector-symbol detection has been extended to a two-vector-symbol version. In addition, multi-vector-symbol detection with more than two vector symbols is possible. In the case of one interferer, we have shown that the interference statistics is indeed temporally and spatially separable, and that the temporal interference correlation may be caused by the intersymbol interference of the interferer. The impact of temporal and spatial interference correlation on channel estimation and data detection was assessed. Our results show that much lower SERs can be achieved by estimating the channel matrix and interference statistics. For moderately long training lengths, for example, four to six times the number of transmitting antennas, gains are obtained by estimating spatial interference correlation; however, for much shorter training lengths, it is better to estimate only the channel matrix and assume spatially white interference in data detection due to poor estimates of the spatial interference correlation matrix. We have investigated the situation where high temporal interference correlation results from a cochannel interferer operating at a lower data rate. The benefit of taking temporal interference correlation into account in *channel estimation* is not significant for both same- and lower-data-rate interference cases due to the sharply reduced temporal correlation after one time lag. In the case of lower-data-rate interference, much

improvement can be achieved by taking account of the temporal interference correlation in *data detection*. For a  $(5, 5)$  MIMO link with independent Rayleigh fading, by exploiting temporal interference correlation in channel estimation and data detection, we obtain 0.5dB and 3.5dB gains in SIR for same- and lower-data-rate interference cases, respectively. Our results show that a training length equal to four times the number of transmitting antennas achieves most of the performance gain. We also observe that for angle spreads larger than 6 degrees, the SER performance is close to the case where MIMO links are independently faded.

## **Chapter 5**

# **Information Capacity of MIMO Systems with Spatially and Temporally Colored Interference**

### **5.1 Introduction**

In Chapter 4, we focused on the processing at the receiver while we assume that no knowledge of the channel matrix and interference statistics is available at the transmitter (independent data streams with uniform power allocation across the transmitting antennas). In this chapter, we consider the benefit of knowing channel matrix and interference statistics at the transmitter. We address this problem from the view point of information theory.

Current work on information capacity of MIMO systems focuses on temporally white interference. Channel capacity under both spatially and temporally white interference was studied in [35, 62, 103, 114]. In [30], the capacity was studied for spatially colored and temporally white interference; capacity expressions were derived for different degrees of knowledge of the channel matrix and interference statistics at the transmitter. As we have shown in Section 4.5, in cellular systems the interference can be not only spatially but also temporally colored. Therefore in this chapter, assuming the receiver has perfect knowledge of the channel matrix and interference statistics, we investigate the MIMO capacity

with different degrees of knowledge of the channel matrix and interference statistics at the transmitter under both spatially and temporally colored interference.

This chapter is organized as follows. The capacity in the case of full knowledge of channel matrix and interference statistics at the transmitter is derived in Section 5.2. In Section 5.3, channel capacities in the cases of partial knowledge of the channel matrix and interference statistics at the transmitter are discussed. Simulation results are presented in Section 5.4.

## 5.2 Channel Capacity: General Case

Consider a single-user link consisting of  $N_t$  transmitting and  $N_r$  receiving antennas. The  $N_r$ -dimensional received signal vector at the  $i$ th time instant is

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \mathbf{n}_i, \quad i = 0, 1, \dots, N-1 \quad (5.1)$$

where  $\mathbf{x}_i$  is the transmitted signal,  $\mathbf{n}_i$  is zero-mean circularly symmetric complex Gaussian interference vector, and  $\mathbf{H}$  is an  $N_r \times N_t$  quasi-static flat fading channel matrix. As shown in Section 4.5, the interference could be both spatially and temporally colored with separable temporal and spatial correlation. Here we assume that the interference is temporally stationary. Therefore, the interference covariance matrix can be modelled as a Kronecker product of temporal and spatial correlations where the temporal correlation matrix is symmetric Toeplitz. We remark that we are studying a channel with memory. Moreover, this channel is information-stable with additive stationary ergodic noise [28].

Concatenating  $\bar{\mathbf{y}}_N = [\mathbf{y}_0^T \cdots \mathbf{y}_{N-1}^T]^T$ ,  $\bar{\mathbf{x}}_N = [\mathbf{x}_0^T \cdots \mathbf{x}_{N-1}^T]^T$  and  $\bar{\mathbf{n}}_N = [\mathbf{n}_0^T \cdots \mathbf{n}_{N-1}^T]^T$ , we rewrite (5.1) as

$$\bar{\mathbf{y}}_N = \bar{\mathbf{H}}_N \bar{\mathbf{x}}_N + \bar{\mathbf{n}}_N \quad (5.2)$$

where  $\bar{\mathbf{H}}_N = \mathbf{I}_N \otimes \mathbf{H}$ . The covariance matrix of  $\bar{\mathbf{n}}_N$  is  $\bar{\mathbf{R}}_N = \mathbf{\Lambda}_N \otimes \mathbf{R}$  where the  $N \times N$  symmetric Toeplitz matrix  $\mathbf{\Lambda}_N$  and the  $N_r \times N_r$  matrix  $\mathbf{R}$  capture the temporal and spatial interference correlations, respectively. In addition, we assume that the correlation matrices  $\mathbf{\Lambda}_N$  and  $\mathbf{R}$  are nonsingular, and that the signal vector  $\bar{\mathbf{x}}_N$  is independent of the interference vector  $\bar{\mathbf{n}}_N$ .

Assume that the receiver has the knowledge of both channel matrix  $\mathbf{H}$  and interference statistics. Since the channel in (5.2) is information-stable, the channel capacity is expressed as [28]

$$C = \lim_{N \rightarrow \infty} \frac{1}{N} \max_{\text{tr}(\mathbf{\Sigma}_{\bar{\mathbf{x}}_N}) \leq N \cdot P_s} \mathcal{I}(\bar{\mathbf{x}}_N; \bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) \quad (5.3)$$

where  $\mathcal{I}(\cdot; \cdot)$  denotes mutual information,  $P_s$  is the constrained transmit power, and  $\mathbf{\Sigma}_{\bar{\mathbf{x}}_N} = E\{\bar{\mathbf{x}}_N \bar{\mathbf{x}}_N^\dagger\}$ . Note that we consider  $\mathbf{H}$  and  $\bar{\mathbf{R}}_N$  as channel outputs.

The mutual information term in (5.3) can be expanded as

$$\begin{aligned} & \mathcal{I}(\bar{\mathbf{x}}_N; \bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) \\ &= \mathcal{H}(\bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) - \mathcal{H}(\bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N | \bar{\mathbf{x}}_N) \\ &= \mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N) + \mathcal{H}(\mathbf{H}, \bar{\mathbf{R}}_N) - \mathcal{H}(\bar{\mathbf{y}}_N | \bar{\mathbf{x}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) - \mathcal{H}(\mathbf{H}, \bar{\mathbf{R}}_N | \bar{\mathbf{x}}_N) \\ &= \mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N) - \mathcal{H}(\bar{\mathbf{n}}_N | \bar{\mathbf{x}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) \end{aligned} \quad (5.4)$$

$$= \mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N) - \mathcal{H}(\bar{\mathbf{n}}_N | \bar{\mathbf{R}}_N) \quad (5.5)$$

$$= \mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N) - \log_2 \det(\pi e \bar{\mathbf{R}}_N) \quad (5.6)$$

where  $\mathcal{H}(\cdot)$  denotes entropy; (5.4) comes from the fact that  $\mathcal{H}(\mathbf{H}, \bar{\mathbf{R}}_N | \bar{\mathbf{x}}_N) = \mathcal{H}(\mathbf{H}, \bar{\mathbf{R}}_N)$  since  $\bar{\mathbf{x}}_N$  is independent of  $\mathbf{H}$  and  $\bar{\mathbf{R}}_N$ ; (5.5) is due to the fact that  $\bar{\mathbf{n}}_N$  is independent of  $\mathbf{H}$  and  $\bar{\mathbf{x}}_N$ ; (5.6) follows the fact that the differential entropy of a circularly symmetric complex Gaussian random vector with covariance matrix  $\mathbf{Q}$  is  $\log_2 \det(\pi e \mathbf{Q})$ .

In (5.6), to maximize the mutual information  $\mathcal{I}(\bar{\mathbf{x}}_N; \bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N)$ , we have to find the distribution of  $\bar{\mathbf{x}}_N$  which maximizes  $\mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N)$ . Note that if  $\bar{\mathbf{x}}_N$  satisfies the power

constraint  $\text{tr}(\Sigma_{\bar{\mathbf{x}}_N}) \leq N \cdot P_s$ , so does  $\bar{\mathbf{x}}_N - E\{\bar{\mathbf{x}}_N\}$ ; hence we can restrict our attention to zero-mean  $\bar{\mathbf{x}}_N$  and  $\Sigma_{\bar{\mathbf{x}}_N}$  becomes the covariance matrix of  $\bar{\mathbf{x}}_N$ . As a result,  $\bar{\mathbf{y}}_N$  has covariance matrix  $\bar{\mathbf{H}}_N \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger + \bar{\mathbf{R}}_N$ . Since circularly symmetric complex Gaussian distribution is entropy maximizer [114],  $\mathcal{H}(\bar{\mathbf{y}}_N | \mathbf{H}, \bar{\mathbf{R}}_N)$  is maximized when  $\bar{\mathbf{y}}_N$  is circularly symmetric complex Gaussian and the maximum value is  $\log_2 \det \left[ \pi e \left( \bar{\mathbf{H}}_N \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger + \bar{\mathbf{R}}_N \right) \right]$ . Clearly, zero-mean circularly symmetric complex Gaussian vector  $\bar{\mathbf{x}}_N$  maximizes the mutual information.

From (5.6), we have

$$\mathcal{I}(\bar{\mathbf{x}}_N; \bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N)$$

$$\begin{aligned} &\leq \log_2 \det \left[ \pi e \left( \bar{\mathbf{H}}_N \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger + \bar{\mathbf{R}}_N \right) \right] - \log_2 \det (\pi e \bar{\mathbf{R}}_N) \\ &= \log_2 \det \left( \bar{\mathbf{H}}_N \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger + \bar{\mathbf{R}}_N \right) + \log_2 \det (\bar{\mathbf{R}}_N^{-1}) \end{aligned} \quad (5.7)$$

$$= \log_2 \det \left( \mathbf{I}_{N \cdot N_r} + \bar{\mathbf{H}}_N \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger \bar{\mathbf{R}}_N^{-1} \right) \quad (5.8)$$

$$= \log_2 \det \left( \mathbf{I}_{N \cdot N_t} + \Sigma_{\bar{\mathbf{x}}_N} \bar{\mathbf{H}}_N^\dagger \bar{\mathbf{R}}_N^{-1} \bar{\mathbf{H}}_N \right) \quad (5.9)$$

$$\begin{aligned} &= \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \Sigma_{\bar{\mathbf{x}}_N} (\mathbf{I}_N \otimes \mathbf{H})^\dagger (\Lambda_N \otimes \mathbf{R})^{-1} (\mathbf{I}_N \otimes \mathbf{H}) \right] \\ &= \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \Sigma_{\bar{\mathbf{x}}_N} \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \end{aligned} \quad (5.10)$$

where (5.7) is due to  $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$  for a nonsingular matrix  $\mathbf{A}$  [75]; (5.8) follows from  $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$  [75]; (5.9) is due to the identity  $\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$  which is proved in Appendix D; (5.10) is due to the properties of Kronecker product in (2.18), (2.19) and (2.17). Substituting (5.10) into (5.3), the channel capacity becomes

$$C = \lim_{N \rightarrow \infty} \frac{1}{N} \max_{\text{tr}(\Sigma_{\bar{\mathbf{x}}_N}) \leq N \cdot P_s} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \Sigma_{\bar{\mathbf{x}}_N} \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \quad (5.11)$$

To find  $\Sigma_{\bar{\mathbf{x}}_N}$  that maximizes the mutual information in (5.10), we use eigenvalue decomposition and express the Hermitian matrix  $\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$  as

$$\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} = \mathbf{V}_{HR} \mathbf{D}_{HR} \mathbf{V}_{HR}^\dagger, \quad \mathbf{D}_{HR} = \text{diag}(d_{HR}^1, \dots, d_{HR}^{N_t}) \quad (5.12)$$

where  $d_{HR}^1, \dots, d_{HR}^{N_t}$  are eigenvalues of  $\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$ , and  $\mathbf{V}_{HR}$  is a unitary matrix consisting of eigenvectors of  $\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$ . Similarly, we eigenvalue-decompose the Hermitian matrix  $\Lambda_N$  as

$$\Lambda_N = \mathbf{U}_{\Lambda N} \mathbf{D}_{\Lambda N} \mathbf{U}_{\Lambda N}^\dagger, \quad \mathbf{D}_{\Lambda N} = \text{diag}(d_{\Lambda N}^1, \dots, d_{\Lambda N}^N); \quad (5.13)$$

hence

$$\Lambda_N^{-1} = \mathbf{U}_{\Lambda N} \mathbf{D}_{\Lambda N}^{-1} \mathbf{U}_{\Lambda N}^\dagger, \quad \mathbf{D}_{\Lambda N}^{-1} = \text{diag}(1/d_{\Lambda N}^1, \dots, 1/d_{\Lambda N}^N). \quad (5.14)$$

Note that  $\Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$  is Hermitian. Using eigenvalue decomposition, we have

$$\Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} = \mathbf{V}_N \mathbf{D}_N \mathbf{V}_N^\dagger, \quad \mathbf{D}_N = \text{diag}(d_N^1, \dots, d_N^{N \cdot N_t}). \quad (5.15)$$

According to the Kronecker product property (2.21), the diagonal matrix  $\mathbf{D}_N = \mathbf{D}_{\Lambda N}^{-1} \otimes \mathbf{D}_{HR}$  and the unitary matrix  $\mathbf{V}_N = \mathbf{U}_{\Lambda N} \otimes \mathbf{V}_{HR}$ . More specifically, we express the elements in  $\mathbf{D}_N$  as

$$d_N^{(i-1)N+j} = \frac{d_{HR}^i}{d_{\Lambda N}^j}, \quad i = 1, \dots, N_t, \quad j = 1, \dots, N. \quad (5.16)$$

Assuming that the transmitter knows the channel matrix and interference statistics (both temporal and spatial correlations), the mutual information in (5.10) is maximized when  $\Sigma_{\bar{x}N} = \mathbf{V}_N \mathbf{P}_N \mathbf{V}_N^\dagger$  with  $\mathbf{V}_N$  shown in (5.15) and  $\mathbf{P}_N = \text{diag}(p_N^1, \dots, p_N^{N \cdot N_t})$  [114]. The mutual information in (5.10) becomes

$$\mathcal{I}(\bar{\mathbf{x}}_N; \bar{\mathbf{y}}_N, \mathbf{H}, \bar{\mathbf{R}}_N) = \sum_{i=1}^{N \cdot N_t} \log_2(1 + p_N^i d_N^i). \quad (5.17)$$

To find  $p_N^1, \dots, p_N^{N \cdot N_t}$  that maximize (5.17) with power constraint  $\sum_{i=1}^{N \cdot N_t} p_N^i = N P_s$ , we use Lagrange multipliers. We form the Lagrangian function

$$J(p_N^1, \dots, p_N^{N \cdot N_t}, \lambda) = \sum_{i=1}^{N \cdot N_t} \log_2(1 + p_N^i d_N^i) + \lambda \left( \sum_{i=1}^{N \cdot N_t} p_N^i \right).$$

By setting  $\partial J/\partial p_i = 0$ , we have

$$\frac{d_N^i}{1 + p_N^i d_N^i} \frac{1}{\ln 2} + \lambda = 0$$

and

$$p_N^i + \frac{1}{d_N^i} = -\frac{1}{\lambda \ln 2} \triangleq \mu.$$

Since  $p_N^i$  has to be positive, we have

$$p_N^i = \left( \mu - \frac{1}{d_N^i} \right)^+$$

where  $(x)^+$  denotes the larger of 0 and  $x$ , and  $\mu$  is chosen such that

$$\sum_{i=1}^{N \cdot N_t} p_N^i = \sum_{j=1}^N \sum_{i=1}^{N_t} \left( \mu - \frac{d_{\Lambda N}^j}{d_{HR}^i} \right)^+ = N P_s.$$

Substituting (5.17) into (5.3), the capacity is expressed as

$$C_{HRA} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^{N_t} \log_2 \left\{ 1 + \left( \mu - \frac{d_{\Lambda N}^j}{d_{HR}^i} \right)^+ \frac{d_{HR}^i}{d_{\Lambda N}^j} \right\}$$

where the subscript indicates that the transmitter has the knowledge of channel matrix and both temporal and spatial interference correlations.

As  $N \rightarrow \infty$ , the eigenvalues of Toeplitz matrix  $\Lambda_N$  approach the DFT of the first row of  $\Lambda_N$ , i.e., the power spectrum of interference,  $N(f)$ ,  $0 \leq f < 1$  [25]. Therefore,

$$C_{HRA} = \sum_{i=1}^{N_t} \int_0^1 \log_2 \left[ 1 + \left( \mu - \frac{N(f)}{d_{HR}^i} \right)^+ \frac{d_{HR}^i}{N(f)} \right] df \quad (5.18)$$

where  $\mu$  is chosen such that

$$\sum_{i=1}^{N_t} \int_0^1 \left( \mu - \frac{N(f)}{d_{HR}^i} \right)^+ df = P_s. \quad (5.19)$$

This implies joint water-filling in both spatial and frequency domain as in the case of frequency selective fading [90].

In the special case of temporally white interference,  $N(f) = 1$ , the capacity in (5.18) is equivalent to the capacity in [30], i.e.,

$$C_{HR, \Lambda=I} = \sum_{i=1}^{N_t} \log_2 \left[ 1 + \left( \mu - \frac{1}{d_{HR}^i} \right)^+ d_{HR}^i \right] \quad (5.20)$$

where  $\mu$  is chosen such that

$$\sum_{i=1}^{N_t} \left( \mu - \frac{1}{d_{HR}^i} \right)^+ = P_s \quad (5.21)$$

and water-filling is applied in spatial domain only.

## 5.3 Special Cases

### 5.3.1 Transmitter knows channel and spatial interference correlation matrices

Without knowledge of temporal interference correlation, the transmitter assumes that the interference is temporally white with  $\Lambda_N = \mathbf{I}_N$ . As a result, the water-filling is applied only in spatial domain according to  $\mathbf{H}$  and  $\mathbf{R}$ . The optimal transmit covariance matrix is

$$\Sigma_{\bar{x}N} = \mathbf{I}_N \otimes \mathbf{V}_{HR} \mathbf{P} \mathbf{V}_{HR}^\dagger \quad (5.22)$$

with  $\mathbf{V}_{HR}$  in (5.12) and  $\mathbf{P} = \text{diag}(p^1, \dots, p^{N_t})$ . The elements in matrix  $\mathbf{P}$  are obtained by water-filling in spatial domain

$$p^i = \left( \mu - \frac{1}{d_{HR}^i} \right)^+ \quad (5.23)$$

where  $d_{HR}^i$  is the eigenvalue of  $\mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$  as shown in (5.12), and  $\mu$  is chosen such that

$$\sum_{i=1}^{N_t} p^i = P_s. \quad (5.24)$$

Substituting (5.22) into (5.11), we obtain

$$\begin{aligned}
C_{HR} &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \left( \mathbf{I}_N \otimes \mathbf{V}_{HR} \mathbf{P} \mathbf{V}_{HR}^\dagger \right) \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \underbrace{\Lambda_N^{-1} \otimes \mathbf{V}_{HR} \mathbf{P} \Lambda_{HR} \mathbf{V}_{HR}^\dagger}_{\mathbf{B}} \right]. \tag{5.25}
\end{aligned}$$

It can be shown that matrix  $\mathbf{B}$  is Hermitian and the eigenvalues of matrix  $\mathbf{B}$  are  $p^i d_{HR}^i / d_{\Lambda N}^j$ ,  $i = 1, \dots, N_t, j = 1, \dots, N$ . Hence, (5.25) becomes

$$\begin{aligned}
C_{HR} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^{N_t} \log_2 \left( 1 + \frac{p^i d_{HR}^i}{d_{\Lambda N}^j} \right) \\
&= \sum_{i=1}^{N_t} \int_0^1 \log_2 \left[ 1 + \frac{p^i d_{HR}^i}{N(f)} \right] df \tag{5.26}
\end{aligned}$$

where  $p^i$  is determined by (5.23) and (5.24).

### 5.3.2 Transmitter knows channel matrix

The transmitter assumes that the interference is both temporally and spatially white with  $\Lambda_N = \mathbf{I}_N$  and  $\mathbf{R} = \mathbf{I}_{N_r}$ . Hence, the water-filling is applied in spatial domain according to  $\mathbf{H}$  only. Using eigenvalue decomposition, we express

$$\mathbf{H}^\dagger \mathbf{H} = \mathbf{U}_H \mathbf{D}_H \mathbf{U}_H^\dagger, \quad \mathbf{D}_H = \text{diag}(d_H^1, \dots, d_H^{N_t}).$$

The optimal transmit covariance matrix is

$$\boldsymbol{\Sigma}_{\bar{x}N} = \mathbf{I}_N \otimes \mathbf{U}_H \mathbf{P} \mathbf{U}_H^\dagger \tag{5.27}$$

where  $\mathbf{P} = \text{diag}(p^1, \dots, p^{N_t})$ . The elements in matrix  $\mathbf{P}$  are obtained by water-filling in spatial domain according to  $\mathbf{H}$

$$p^i = \left( \mu - \frac{1}{d_H^i} \right)^+ \tag{5.28}$$

where  $\mu$  is chosen such that

$$\sum_{i=1}^{N_t} p^i = P_s. \quad (5.29)$$

Substituting (5.27) into (5.11), we have

$$\begin{aligned} C_H &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \left( \mathbf{I}_N \otimes \mathbf{U}_H \mathbf{P} \mathbf{U}_H^\dagger \right) \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \left( \mathbf{I}_N \otimes \mathbf{U}_H \mathbf{P}^{1/2} \right) \left( \mathbf{I}_N \otimes \mathbf{P}^{1/2} \mathbf{U}_H^\dagger \right) \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \left( \mathbf{I}_N \otimes \mathbf{P}^{1/2} \mathbf{U}_H^\dagger \right) \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \left( \mathbf{I}_N \otimes \mathbf{U}_H \mathbf{P}^{1/2} \right) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \underbrace{\Lambda_N^{-1} \otimes \mathbf{P}^{1/2} \mathbf{U}_H^\dagger \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \mathbf{U}_H \mathbf{P}^{1/2}}_{\mathbf{A}} \right]. \end{aligned}$$

Note that matrix  $\mathbf{A}$  is Hermitian and can be eigenvalue decomposed. Denoting the eigenvalues of matrix  $\mathbf{A}$  as  $d_A^1, \dots, d_A^{N_t}$ , together with the eigenvalues of  $\Lambda_N^{-1}$  in (5.14), we have the capacity

$$\begin{aligned} C_H &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^{N_t} \log_2 \left( 1 + \frac{d_A^i}{d_{\Lambda N}^j} \right) \\ &= \sum_{i=1}^{N_t} \int_0^1 \log_2 \left[ 1 + \frac{d_A^i}{N(f)} \right] df. \end{aligned} \quad (5.30)$$

### 5.3.3 Transmitter has no knowledge of channel and interference

Uniform power allocation is used at the transmitter, i.e., the optimal covariance matrix of the input is

$$\Sigma_{\bar{x}N} = \mathbf{I}_N \otimes \frac{P_s}{N_t} \mathbf{I}_{N_t}. \quad (5.31)$$

Substituting (5.31) into (5.11), we obtain the capacity

$$\begin{aligned} C_{\text{no info}} &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \left( \mathbf{I}_N \otimes \frac{P_s}{N_t} \mathbf{I}_{N_t} \right) \left( \Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right) \right] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \det \left[ \mathbf{I}_{N \cdot N_t} + \Lambda_N^{-1} \otimes \frac{P_s}{N_t} \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H} \right]. \end{aligned}$$

The eigenvalues of Hermitian matrix  $\Lambda_N^{-1} \otimes \mathbf{H}^\dagger \mathbf{R}^{-1} \mathbf{H}$  are shown in (5.16). Thus, we have

$$\begin{aligned} C_{\text{no info}} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^{N_t} \log_2 \left( 1 + \frac{P_s d_{HR}^i}{N_t d_{\Lambda N}^j} \right) \\ &= \sum_{i=1}^{N_t} \int_0^1 \log_2 \left( 1 + \frac{P_s d_{HR}^i}{N_t N(f)} \right) df. \end{aligned} \quad (5.32)$$

## 5.4 Numerical Results

We use the system model in Section 4.5.1 and assume that the desired and interfering users transmit at the same data rate. From (4.28) in Section 4.5.2, the spatial interference correlation matrix is  $\mathbf{R} = \frac{P_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger$  where  $P_I$  is the interference power,  $L$  is the number of interfering transmitting antennas, and  $\mathbf{H}_I$  is the interferer's channel matrix. We assume that the rolloff factor of the raised-cosine pulse shaping  $\beta = 1$ . For the interferer delay  $0.4T$  where  $1/T$  is the data rate, the elements of the temporal interference correlation matrix can be obtained by numerical calculation of (4.27) as

$$\mathbf{\Lambda}_{i,j,\tau=0.4} = \begin{cases} 0.5477 & i = j \\ 0.2261 & |i - j| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \leq i, j \leq N - 1. \quad (5.33)$$

If there is no delay between the desired user and the interferer, the elements of the temporal correlation matrix are

$$\mathbf{\Lambda}_{i,j,\tau=0} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \leq i, j \leq N - 1.$$

Hence, the interference is temporally white for a synchronized interferer. Note that the interference power is smaller in the case of an asynchronous interferer.

As in Section 4.5.4, both independent and correlated MIMO links are considered for the desired user, and independent Rayleigh fading is assumed for the interferer. The same

geometry of antenna arrays as in Section 4.5.4 is assumed for the desired user. Both the desired and interfering users are equipped with  $(4, 4)$  MIMO links, i.e.,  $N_t = N_r = L = 4$ .

For slow fading channels and applications with decoding delay constraints, an important performance measure is the capacity at a given outage probability [12, 85]. In these circumstances, the transmission duration of a codeword (this duration is long enough to accommodate a very long codeword) may be much smaller than the coherence time of the fading channel, and the channel is almost fixed during the transmission of the codeword. Hence, the channel capacity is viewed as a random entity as it depends on the instantaneous random channel parameters. The definition and operational meaning of outage capacity can be found in [18]. In this section, we consider 10% outage capacity, denoted as  $C_{0.1}$ , where  $\Pr(C < C_{0.1}) = 10\%$ .

Monte Carlo simulation is used to assess the 10% outage capacity for different assumptions of knowledge of channel matrix and interference statistics at the transmitter. We use (5.18), (5.26), (5.30) or (5.32) to calculate the capacity for a particular realization of the desired and interfering users' channel matrices. From the empirical distribution [86], we find out  $C_{0.1}$ . In the simulations, we let  $P_s = 100$  and define  $\text{SIR (dB)} \triangleq 10 \log_{10} P_s / P_I$ . We point out that all water-filling results depend not only on SIR but also on values of  $P_s$  and  $P_I$  since water-filling is a nonlinear operation.

Fig. 5.1 shows the outage capacity versus SIR for spatially and temporally colored interference (the interferer delay is  $0.4T$ ). Independent Rayleigh fading is assumed for the desired user. We observe that knowledge of the channel matrix offers a small gain in capacity, which is consistent with [104]. This is due to the independent fading of the desired user. If the elements in channel matrix  $\mathbf{H}$  are i.i.d. zero-mean complex Gaussians, water-filling power allocation does not significantly outperform uniform power allocation in that the disparity among the eigenvalues of  $\mathbf{H}^\dagger \mathbf{H}$  is not large [104]. Fig. 5.1 also shows that,

as interference becomes weak (from low to high SIR), the benefit of knowing interference spatial correlation at the transmitter diminishes. Moreover, it is observed in Fig. 5.1 that lack of knowledge of temporal correlation of the interference at the transmitter causes a small capacity reduction. This can be explained by the low interference temporal correlation in (5.33) as the correlation drops quickly to zero if the time difference is larger than one.

Fig. 5.2 shows the outage capacity for temporally white but spatially colored interference (the interferer delay is zero). Independent Rayleigh fading is assumed for the desired user. Comparing Figs. 5.1 and 5.2, we observe that temporally colored interference (asynchronous interferer) gives higher capacity. This is partially due to the smaller interference power in the case of asynchronous interferer, and partially due to the fact that temporal correlation in interference increases capacities. To see how the temporal correlation increases the capacity, let us consider an extreme case where the temporal interference correlation is unity for all time differences, i.e., interference vectors are the same at all sampling times. With the knowledge of channel matrix and interference statistics available at the receiver, the interference can be easily determined by the receiver if the transmitter sends one training vector symbol. Therefore, after the first training vector symbol, the channel can be considered as interference-free, and we can transmit unlimited amounts of information over the channel within our assumption of interference-limited environment.

Fig. 5.3 shows the effect of Ricean factor  $K$  assuming the scattered components of the desired user's channel are independently faded. As  $K$  increases, the rank of the channel matrix approaches one, and the disparity among the eigenvalues of  $\mathbf{H}^\dagger \mathbf{H}$  increases. Therefore, water-filling power allocation achieves much higher capacities than uniform power allocation does. In Fig. 5.3, as expected, the gain achieved by knowing the channel matrix increases as  $K$  grows. Due to the rank reduction of the channel matrix, we observe that

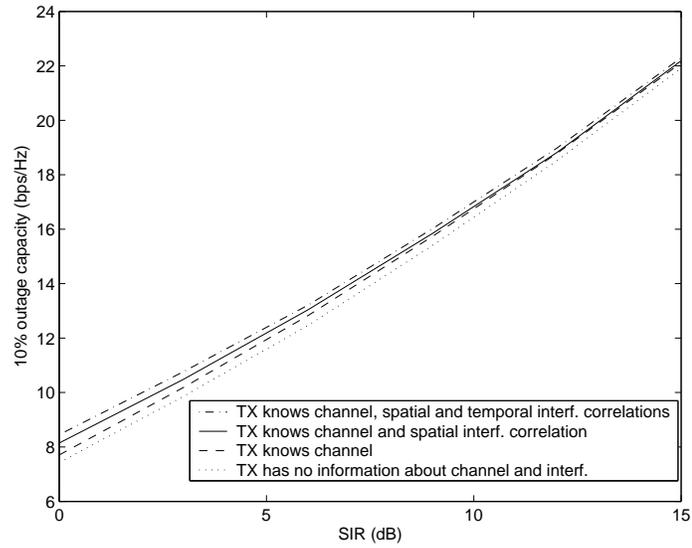


Figure 5.1. 10% outage capacity versus SIR for temporally and spatially colored interference (interferer delay  $0.4T$ ) with  $N_t = N_r = L = 4$ . Independent Rayleigh fading is assumed for the desired user.

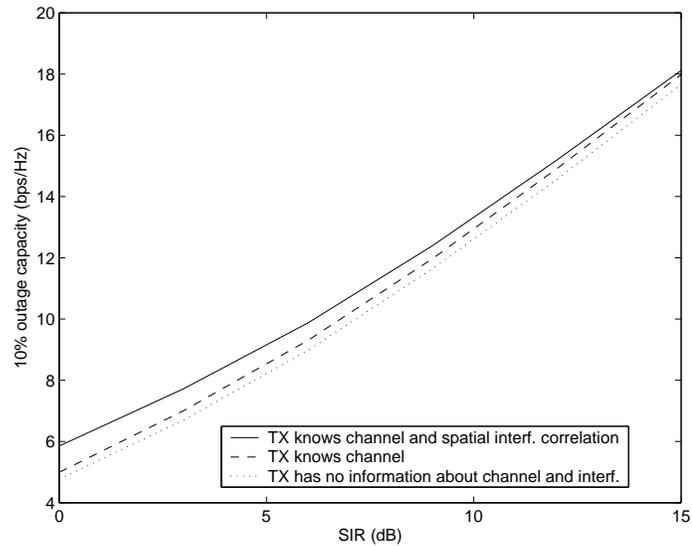


Figure 5.2. 10% outage capacity versus SIR for spatially colored interference (synchronous interferer) with  $N_t = N_r = L = 4$ . Independent Rayleigh fading is assumed for the desired user.

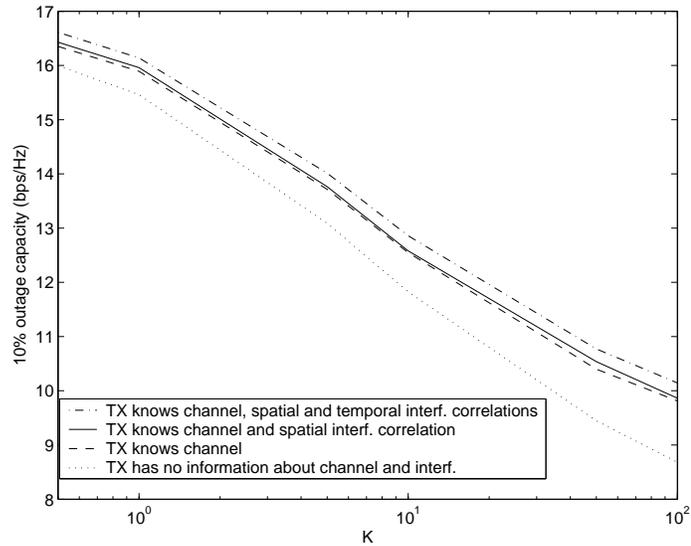


Figure 5.3. 10% outage capacity versus  $K$ -Ricean factor for temporally and spatially colored interference (interferer delay  $0.4T$ ) with  $N_t = N_r = L = 4$  and SIR=10dB. The scattered component of the desired user's channel is independently faded.

capacities decrease as  $K$  increases.

Fig. 5.4 shows the effect of angle spread at the base station antenna array assuming Ricean factor  $K = 0$ . As angle spread decreases, the MIMO links become more and more correlated, and the rank of the channel matrix is reduced to one at zero angle spread. Hence, similar to the case of increasing Ricean factor  $K$ , Fig. 5.4 shows that the capacity gain achieved by knowing channel matrix decreases as angle spread increases. As expected, it is observed that capacities increase as the angle spread increases. However, capacities are not improved significantly when angle spread is larger than 6 degrees. Hence, the MIMO channel links can be considered to be faded independently with angle spreads larger than 6 degrees. Figs. 5.3 and 5.4 also show that the knowledge of interference temporal correlation does not offer significant capacity gains.

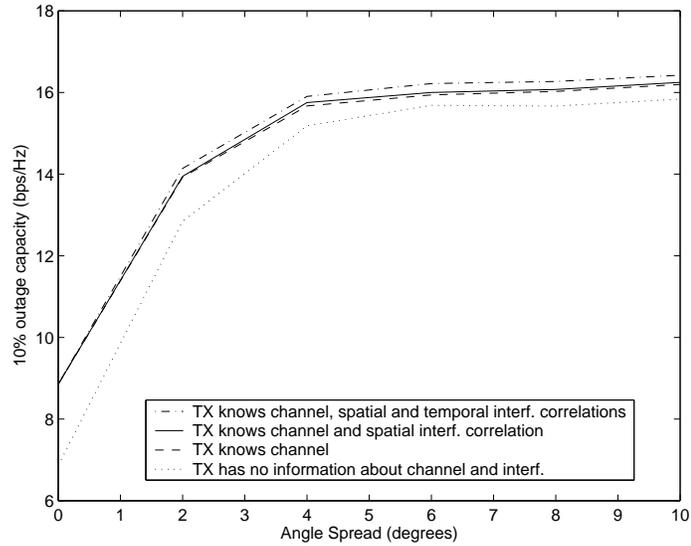


Figure 5.4. 10% outage capacity versus angle spread for temporally and spatially colored interference (interferer delay  $0.4T$ ) with  $N_t = N_r = L = 4$  and SIR=10dB. Correlated Rayleigh fading (Ricean factor  $K = 0$ ) is assumed for the desired user.

## 5.5 Conclusions

In this chapter, we have derived the channel capacities of MIMO systems under spatially and temporally correlated interference. Assuming that the receiver knows the channel matrix and interference statistics, channel capacities have been investigated for different degrees of knowledge of the channel matrix and interference statistics at the transmitter. It is shown that if the interference spatial and temporal correlation is available at the transmitter, water-filling power allocation in both spatial and frequency domain should be applied. With the temporal interference correlation of the same-data-rate interference case in Chapter 4, the results show that the knowledge of interference temporal correlation at the transmitter does not offer significant capacity improvement. As the MIMO channel links become

more correlated, knowing the channel matrix at the transmitter provides significant capacity gains. It is observed that the MIMO links can be considered to be faded independently if the angle spread at the base station antenna array is larger than 6 degrees.

# Chapter 6

## Adaptive Modulation for MIMO Systems

### 6.1 Introduction

In Chapter 4, we considered channel estimation and data detection at the receiver. In Chapter 5, we showed that, for continuous input signals, with knowledge of the channel matrix and interference statistics at the transmitter, there are gains to be achieved in information capacity via water-filling power allocation at the transmitter. In this chapter, we will consider joint processing at the transmitter and receiver. With the channel estimates in Chapter 4, we will investigate how to implement the water-filling power allocation at the transmitter in practical systems with finite alphabet input signals.

Studies for continuous input signals suggest that, if both transmitter and receiver have perfect knowledge of the channel gain matrix and interference statistics, to achieve the capacity, we can decompose the MIMO channel into independent subchannels and apply optimum water-filling on these subchannels [30], i.e., we allocate more power to subchannels with high channel gains. However, for finite alphabet input signals, power allocation

alone may not be an efficient way to maximize capacity. This is due to the different relationship between channel capacity and SNR for continuous-input signals versus discrete-input signals. Consider a single-input single-output (SISO) channel, for continuous inputs, the channel capacity increases without bound as SNR increases; while for discrete inputs, assuming the transmitted signal is drawn from an  $M$ -point constellation with equal probability, the channel capacity saturates at  $\log_2 M$  bits/channel use at high SNR. Therefore, for discrete inputs, to boost data throughput, we need to apply not only power allocation, but also other techniques, including varying the size of signal constellations. That is, for good subchannels, we allocate more power and use higher-level modulation. As a result, data rates are higher on good subchannels. In the following, joint adaptation of power and modulation level is referred to as adaptive modulation [49, 119].

Adaptive modulation for MIMO systems maximizing information rate has been considered for flat and frequency selective fading channels in [66, 90, 95, 97, 98] assuming perfect channel knowledge at both the transmitter and receiver. In practical systems, the channel transfer matrix and interference statistics are estimated by the receiver, for example, from training symbol sequences. Hence, perfect channel knowledge may not be obtainable. On the other hand, to enable adaptive modulation, channel knowledge is required at the transmitter. In frequency division duplex (FDD) systems, this requires the receiver to feed channel information back to the transmitter. In practical systems, due to limited feedback channel bandwidth, low-resolution quantized channel information is fed back [84]. Therefore, the transmitter and receiver may not have the identical channel information due to quantization error. The performance of adaptive modulation based on channel mean feedback was presented in [128]. In this chapter, we will investigate the impact of imperfect channel knowledge and feedback quantization error on adaptive modulation in MIMO systems. The channel estimates from Chapter 4 is used. Rate-distortion theory is involved in

our work to assess the achievable performance of feedback quantization. Since the results in Chapters 4 and 5 show that taking temporal interference correlation into account neither increases the channel capacity nor improves the symbol error rates, in this chapter, we will focus on temporally white interference.

To apply adaptive modulation, the MIMO channel has to be decomposed into subchannels. In Section 6.2, MIMO channel diagonalization is described for both temporally white and colored interference. The adaptive modulation algorithm is described in Section 6.3. The effects of imperfect channel estimates and feedback quantization error for temporally white interference are investigated in Sections 6.4 and 6.5, respectively.

## 6.2 Diagonalization of MIMO Channels

### 6.2.1 Temporally white interference

Fig. 6.1 shows how to decompose a MIMO channel into independent single-input single-output (SISO) subchannels for temporally white interference. Assume that the channel matrix  $\mathbf{H}$  is quasi-static, and that the interference vector  $\mathbf{n}$  is zero-mean circularly symmetric complex Gaussian with covariance matrix  $\mathbf{R}$ . In Fig. 6.1, at the transmitter side, the signal to be conveyed,  $s$ , is multiplied by a pre-processing matrix  $\mathbf{V}_{HR}$ . Note that the entries in  $s$  are independent and could have different signal constellation sizes and powers. At the receiver side, a post-processing matrix  $\mathbf{U}_{HR}^\dagger \mathbf{R}^{-1/2}$  is applied. The unitary matrices  $\mathbf{V}_{HR}$  and  $\mathbf{U}_{HR}$  are obtained via singular value decomposition (SVD) of

$$\mathbf{R}^{-1/2} \mathbf{H} = \mathbf{U}_{HR} \mathbf{\Phi}_{HR} \mathbf{V}_{HR}^\dagger. \quad (6.1)$$

Since the pre-processing matrix  $\mathbf{V}_{HR}$  is a unitary matrix, the power of the transmitted signal,  $\mathbf{x} = \mathbf{V}_{HR} s$ , is the same as that of  $s$ . As a result, the soft output in Fig. 6.1 is

expressed as

$$\begin{aligned}
\mathbf{r} &= \mathbf{U}_{HR}^\dagger \mathbf{R}^{-1/2} (\mathbf{H} \mathbf{V}_{HR} \mathbf{s} + \mathbf{n}) \\
&= \mathbf{U}_{HR}^\dagger \mathbf{R}^{-1/2} \mathbf{H} \mathbf{V}_{HR} \mathbf{s} + \underbrace{\mathbf{U}_{HR}^\dagger \mathbf{R}^{-1/2} \mathbf{n}}_{\mathbf{n}_0} \\
&= \mathbf{\Phi}_{HR} \mathbf{s} + \mathbf{n}_0
\end{aligned}$$

where it can be shown that  $E \{ \mathbf{n}_0 \mathbf{n}_0^\dagger \} = \mathbf{I}$ . Since  $\mathbf{\Phi}_{HR}$  is a diagonal matrix, the MIMO channel is transformed into independent SISO subchannels. The inputs of subchannels are elements of  $\mathbf{s}$ ; the channel gains are equal to the singular values of  $\mathbf{R}^{-1/2} \mathbf{H}$ ; the interference on each subchannel is zero-mean Gaussian with unit variance; the number of effective subchannels (subchannels with non-zero gains) is determined by the number of non-zero singular values of  $\mathbf{R}^{-1/2} \mathbf{H}$ . Hence, by linear transformations at the transmitter and receiver, the MIMO channel is decomposed into several SISO channels. The signal constellation sizes and powers of signals in  $\mathbf{s}$  will be determined by the adaptive modulation algorithm in Section 6.3. We remark that the pre-processing matrix  $\mathbf{V}_{HR}$  constitutes a form of transmit beamforming.

The post-processing matrix  $\mathbf{U}_{HR}^\dagger \mathbf{R}^{-1/2}$  can also be interpreted as follows. The received signal vector  $\mathbf{y}$  in Fig. 6.1 is first filtered by  $\mathbf{R}^{-1/2}$  to spatially whiten the interference. Then the unitary transformation  $\mathbf{U}_{HR}^\dagger$  is applied to complete the diagonalization of the composite channel  $\mathbf{R}^{-1/2} \mathbf{H}$ .

### 6.2.2 Temporally colored interference

To maximize the information transmission rate, block transmission is employed so that the temporally colored interference is taken into account. Assume that a block of  $N$  data vectors is transmitted. By stacking the  $N$  received signal vectors, the system can be modelled as (5.2) with augmented channel matrix  $\bar{\mathbf{H}}_N = \mathbf{I}_N \otimes \mathbf{H}$  and interference covariance matrix

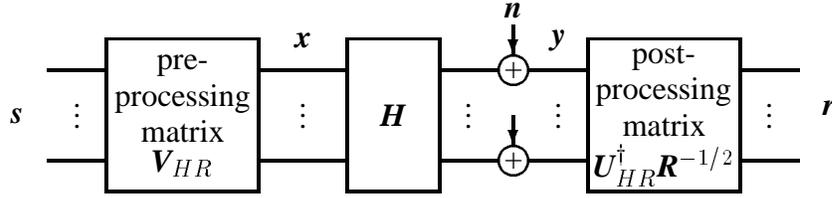


Figure 6.1. Diagonalization of an MIMO channel under temporally white interference.

$\bar{\mathbf{R}}_N = \mathbf{\Lambda}_N \otimes \mathbf{R}$ . Similar to the case of temporally white interference, we decompose the composite channel as

$$\begin{aligned} \bar{\mathbf{R}}_N^{-1/2} \bar{\mathbf{H}}_N &= (\mathbf{\Lambda}_N \otimes \mathbf{R})^{-1/2} (\mathbf{I}_N \otimes \mathbf{H}) \\ &= \left( \mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{R}^{-1/2} \right) (\mathbf{I}_N \otimes \mathbf{H}) \end{aligned} \quad (6.2)$$

$$\begin{aligned} &= \mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{R}^{-1/2} \mathbf{H} \\ &= \left( \mathbf{U}_{\Lambda N} \mathbf{D}_{\Lambda N}^{-1/2} \mathbf{U}_{\Lambda N}^{\dagger} \right) \otimes \left( \mathbf{U}_{HR} \mathbf{\Phi}_{HR} \mathbf{V}_{HR}^{\dagger} \right) \\ &= \left( \mathbf{U}_{\Lambda N} \otimes \mathbf{U}_{HR} \right) \left( \mathbf{D}_{\Lambda N}^{-1/2} \otimes \mathbf{\Phi}_{HR} \right) \left( \mathbf{U}_{\Lambda N} \otimes \mathbf{V}_{HR} \right)^{\dagger} \end{aligned} \quad (6.3)$$

where in (6.2) it can be shown that  $(\mathbf{\Lambda}_N \otimes \mathbf{R})^{-1/2} = \mathbf{\Lambda}_N^{-1/2} \otimes \mathbf{R}^{-1/2}$ ; (6.3) is due to SVDs in (5.13) and (6.1). If the SVD of

$$\bar{\mathbf{R}}_N^{-1/2} \bar{\mathbf{H}}_N = \mathbf{U}_N \mathbf{\Phi}_N \mathbf{V}_N^{\dagger}, \quad (6.4)$$

we have

$$\mathbf{U}_N = \mathbf{U}_{\Lambda N} \otimes \mathbf{U}_{HR}, \quad \mathbf{V}_N = \mathbf{U}_{\Lambda N} \otimes \mathbf{V}_{HR} \quad \text{and} \quad \mathbf{\Phi}_N = \mathbf{D}_{\Lambda N}^{-1/2} \otimes \mathbf{\Phi}_{HR}.$$

Fig. 6.2 shows the diagonalization of the MIMO channel with temporally colored interference. At the transmitter, powers and signal constellations of signals in the block  $\mathbf{s}_N = (s_{1,1}, \dots, s_{1,N}, \dots, s_{N_t,1}, \dots, s_{N_t,N})$  are determined from  $\mathbf{\Phi}_N$  in (6.4) by an adaptive modulation algorithm. At the receiver, after collecting all signals during the block, ZF detection is applied.

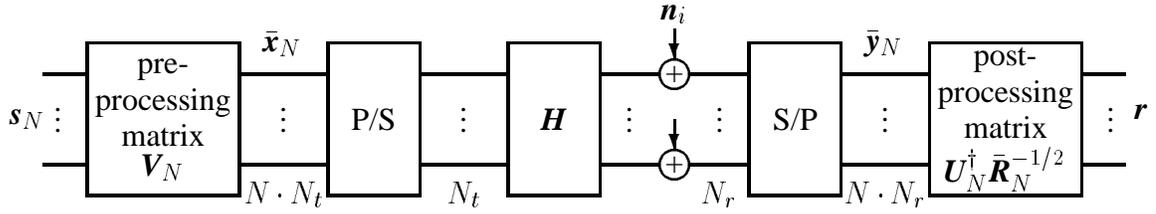


Figure 6.2. Diagonization of an MIMO channel under temporally colored interference.

### 6.3 Adaptive Modulation

Similar adaptive modulation algorithms that are used for orthogonal frequency division multiplexing (OFDM) systems can be adapted and applied to MIMO systems. Two algorithms are of practical importance: the Hughes-Hartogs algorithm which maximizes data throughput with a target bit error rate (BER) [13], and Chow-Cioffi-Bingham algorithm [24] along with its improved version [33] which transmits a fixed data rate with lowest error rate possible. Note that the latter algorithm will use up all the available transmit power, while the former one may not since the power increment from a low level modulation to a higher level is discrete. In our work, we choose the Hughes-Hartogs algorithm due to two reasons. First, most applications, such as voice and video transmission, have target BERs. Second, in wireless systems, we may want to use as little transmit power as possible to minimize the interference to other users. In addition, in our work, we restrict the modulation schemes to BPSK, QPSK, 16-square QAM and 64-square QAM.

The spectral efficiency of a MIMO system is expressed in terms of data rate per unit bandwidth in bits/sec/Hz. If we send  $k$  bits/vector symbol, the data rate is  $k/T_s$  bits/sec where  $T_s$  is the symbol duration. With Nyquist pulses, we can assume that the signal bandwidth is  $1/T_s$ . Therefore, the data rate per unit bandwidth is  $k$  bits/sec/Hz.

### 6.3.1 Adaptive modulation algorithm

We define an ordered list of possible modulation schemes as  $Q = \{\text{BPSK}, \text{QPSK}, \text{16QAM}, \text{64QAM}\}$ , i.e.,  $Q_1 = \text{BPSK}$  and so on. To facilitate adaptive modulation, given subchannel gains, we first calculate an incremental power matrix  $\mathbf{P}_{\text{incr}}$ , where the  $(m, n)$ th element  $P_{\text{incr},m,n} = P_{m,n} - P_{m-1,n}$ ,  $1 \leq m \leq 4$ , and  $P_{m,n}$  denotes the required transmit power for the  $n$ th subchannel to support modulation scheme  $Q_m$  at some target BER. Clearly,  $P_{0,n} = 0$ .

To maximize the data throughput with a transmit power constraint  $P_{\text{total}}$ , each time we choose the subchannel that requires the least incremental power. The algorithm is summarized as follows [13]:

Step 1  $i = 1$ ,  $P_{\text{used}} = 0$ , and  $\mathbf{P}_{\text{incr}}^{(i)} = \mathbf{P}_{\text{incr}}$ .

Step 2 Search the first row of  $\mathbf{P}_{\text{incr}}^{(i)}$  for the smallest element. If the  $(1, n)$ th element is the smallest and  $\mathbf{P}_{\text{incr},1,n}^{(i)} = P_{\text{incr},j,n}$ , the  $n$ th subchannel requires the smallest incremental power from modulation level  $Q_{j-1}$  to  $Q_j$ .

Step 3 If  $P_{\text{used}} + P_{\text{incr},j,n} \leq P_{\text{total}}$ , the modulation level on the  $n$ th subchannel is increased to  $Q_j$ , and  $P_{\text{used}} \leftarrow P_{\text{used}} + P_{\text{incr},j,n}$ ; otherwise, the adaptive modulation is completed and exit the algorithm.

Step 4 Obtain  $\mathbf{P}_{\text{incr}}^{(i+1)}$  by  $\mathbf{P}_{\text{incr},l,n}^{(i+1)} = \mathbf{P}_{\text{incr},l+1,n}^{(i)}$ ,  $l = 1, 2, 3$ , and  $\mathbf{P}_{\text{incr},4,n}^{(i+1)} = P_{\text{total}}$ . That is, we move the elements in the  $n$ th column of  $\mathbf{P}_{\text{incr}}^{(i)}$  up one place, and set the last element as  $P_{\text{total}}$ .

Step 5  $i = i + 1$ , and go to Step 2.

To make sure that the modulation levels on subchannels are upper-bounded by 64QAM, in updating  $\mathbf{P}_{\text{incr}}^{(i)}$ , we set the last element as  $P_{\text{total}}$ .

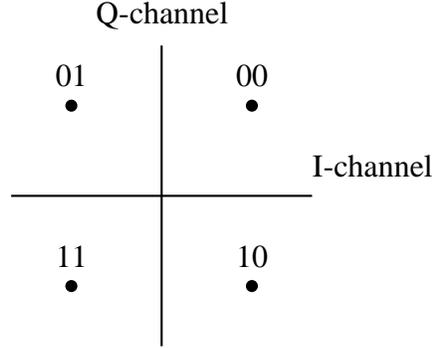


Figure 6.3. QPSK signal constellation with Gray code mapping.

### 6.3.2 Bit error rate of modulation schemes

To build up the incremental power matrix  $\mathbf{P}_{\text{incr}}$ , we need to know the required SNR for a modulation scheme to achieve the target BER. With Gray code bit mapping for QPSK, 16QAM and 64QAM as shown in Figs. 6.3 to 6.5, the BERs in AWGN can be expressed as [23, 89, 126]

$$P_{b,\text{BPSK}} = \frac{1}{2} \text{erfc}(\sqrt{\gamma}), \quad (6.5)$$

$$P_{b,\text{QPSK}} = \frac{1}{2} \text{erfc}\left(\sqrt{\frac{\gamma}{2}}\right), \quad (6.6)$$

$$P_{b,16\text{QAM}} = \frac{1}{8} \left\{ 3 \text{erfc}\left(\sqrt{\frac{\gamma}{10}}\right) + 2 \text{erfc}\left(3\sqrt{\frac{\gamma}{10}}\right) - \text{erfc}\left(5\sqrt{\frac{\gamma}{10}}\right) \right\}, \quad (6.7)$$

and

$$P_{b,64\text{QAM}} = \frac{1}{24} \left\{ 7 \text{erfc}\left(\sqrt{\frac{\gamma}{42}}\right) + 6 \text{erfc}\left(3\sqrt{\frac{\gamma}{42}}\right) - \text{erfc}\left(5\sqrt{\frac{\gamma}{42}}\right) \right. \\ \left. + \text{erfc}\left(9\sqrt{\frac{\gamma}{42}}\right) - \text{erfc}\left(13\sqrt{\frac{\gamma}{42}}\right) \right\} \quad (6.8)$$

where  $\gamma$  is SNR per symbol, and the complementary error function  $\text{erfc}(\cdot)$  is defined as  $\text{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2} dx$ . The BERs of BPSK, QPSK, 16-square QAM and 64-square

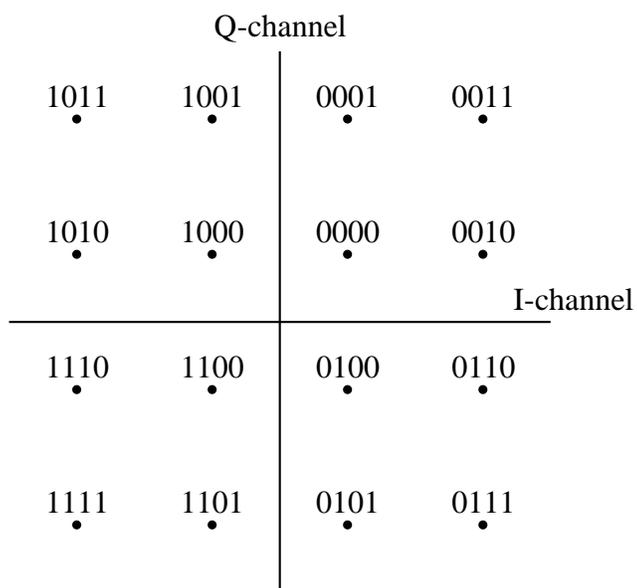


Figure 6.4. 16-square QAM signal constellation with Gray code mapping.

QAM are shown in Fig. 6.6. We use Gray mapping due to the closed-form BER expressions. According to Fig. 6.6, required SNRs to achieve a predefined target BER are summarized in Table 6.1.

Q-channel				I-channel			
101111	101101	100101	100111	000111	000101	001101	001111
101110	101100	100100	100110	000110	000100	001100	001110
101010	101000	100000	100010	000010	000000	001000	001010
101011	101001	100001	100011	000011	000001	001001	001011
111011	111001	110001	110011	010011	010001	011001	011011
111010	111000	110000	110010	010010	010000	011000	011010
111110	111100	110100	110110	010110	010100	011100	011110
111111	111101	110101	110111	010111	010101	011101	011111

Figure 6.5. 64-square QAM signal constellation with Gray code mapping.

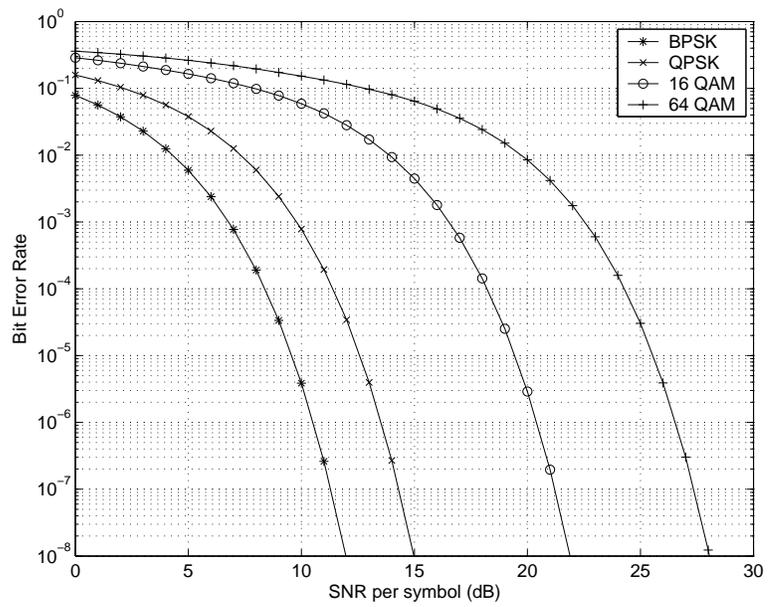


Figure 6.6. Bit error rates vs. SNR per symbol in AWGN channel for BPSK, QPSK, 16-square QAM and 64-square QAM.

target BER	BPSK	QPSK	16QAM	64QAM
$10^{-2}$	4.3203	7.2235	13.9286	19.8041
$10^{-3}$	6.7396	9.7811	16.4862	22.5000
$5 \times 10^{-6}$	10.0922	13.0415	19.7696	25.8525
$10^{-8}$	12.1198	14.9770	21.8894	28.1567

Table 6.1. Required SNRs in dB for a predefined target BER.

## 6.4 Effect of Imperfect Channel Estimates

In practical systems, the receiver estimates the channel matrix and interference statistics and feeds them back to the transmitter so that adaptive modulation can be applied at the transmitter. In practice, we usually have neither perfect channel estimates nor perfect feedback paths. In this section, to focus on the effect of imperfect channel estimates from training sequences, we assume a perfect feedback path from the receiver to transmitter and temporally white interference.

### 6.4.1 System model

Consider a MIMO link with  $N_t$  transmitting and  $N_r$  receiving antennas. The received signal is

$$y_i = \mathbf{H}\mathbf{x}_i + \underbrace{\sqrt{\frac{P_I}{L}}\mathbf{H}_I\mathbf{b}_i + \mathbf{w}_i}_{\mathbf{n}_i}$$

where  $\mathbf{x}_i$  is the transmitted vector symbol with power constraint  $P_s$ ,  $\mathbf{H}$  is the  $N_r \times N_t$  quasi-static channel matrix of the desired user, and  $\mathbf{n}_i$  is the interference-plus-noise vector. We assume that interference comes from  $L$  transmitters which could belong to one or more interferers. The total interference power is  $P_I$ , the  $N_r \times L$  matrix  $\mathbf{H}_I$  consists of quasi-static

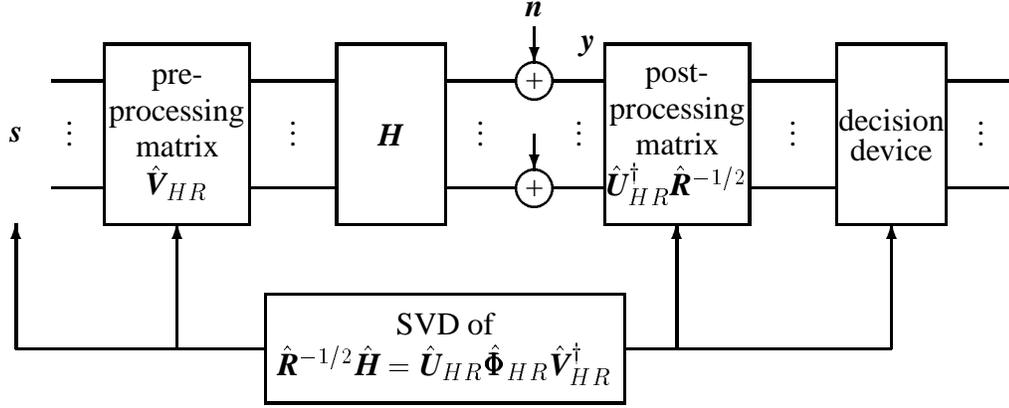


Figure 6.7. Adaptive modulation with imperfect channel estimates.

interfering users' channel gains,  $\mathbf{b}_i$  consists of  $L$  i.i.d. zero-mean interfering signals each with unit variance and  $\mathbf{w}_i$  is the additive white Gaussian noise vector with zero-mean and covariance matrix  $\sigma^2 \mathbf{I}_{N_r}$ . Note that each interfering transmitter has the same transmitted power. It can be shown that the covariance matrix of interference plus noise is

$$\mathbf{R} = \frac{P_I}{L} \mathbf{H}_I \mathbf{H}_I^\dagger + \sigma^2 \mathbf{I}_{N_r}.$$

With temporally white interference and noise, during the training period of  $N$  vector symbols, the estimates of the desired user's channel matrix and the spatial interference correlation matrix,  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{R}}$ , can be obtained by (4.17) and (4.18), respectively.

During the data transmission period, as shown in Fig. 6.7, the transmitter and receiver calculate the pre- and post-processing matrices  $\hat{\mathbf{V}}_{HR}$  and  $\hat{\mathbf{U}}_{HR}^\dagger \hat{\mathbf{R}}^{-1/2}$ , respectively, via SVD of  $\hat{\mathbf{R}}^{-1/2} \hat{\mathbf{H}} = \hat{\mathbf{U}}_{HR} \hat{\mathbf{\Phi}}_{HR} \hat{\mathbf{V}}_{HR}^\dagger$ . The power and modulation level for each transmitted signal in  $s$  are determined from  $\hat{\mathbf{\Phi}}_{HR}$  by the adaptive modulation algorithm. To demodulate the transmitted signals, the receiver scales the appropriate constellations by the power coefficient obtained from the adaptive modulation algorithm.

## 6.4.2 Simulation results

We consider a desired user with a  $(4, 4)$  MIMO link and  $L = 5$  interfering transmitters. The interfering users are assumed to experience i.i.d. Rayleigh fading. For the desired user, both independent and correlated MIMO links are considered, and the same antenna array geometry as in Section 4.5.4 is assumed. Monte Carlo simulation is used to evaluate the average system performance over different channel realizations. In the simulation, we set the total interference power  $P_I = 1$ , and interference-to-noise power ratio INR=20 dB.

In Figs. 6.8 and 6.9, we compare the performance of adaptive modulation with non-adaptive case assuming perfect knowledge of channel and spatial interference correlation matrices at the transmitter. The desired user is assumed to experience i.i.d. Rayleigh fading. For the adaptive modulation, the target bit error rate is  $10^{-3}$ . For the non-adaptive case, QPSK is assumed for each transmitter and uniform power allocation is applied among transmitting antennas; ordered MMSE detection is used at the receiver to estimate the transmitted symbols. As expected, for the adaptive case, the system operates at the target BER for all values of SIR. The data throughput increases significantly with SIR and eventually saturates at 24 bits/sec/Hz since the modulation on each subchannel is restricted to 64 QAM. For the non-adaptive case, the data throughput is fixed at 8 bits/sec/Hz for all values of SIR, and the BER improves significantly as SIR increases. At low SIR, the adaptive modulation sacrifices spectral efficiency to obtain low BERs; while at high SIR, it sacrifices the unnecessarily low BERs to achieve high data rate. We point out that in the non-adaptive case, the transmitters use up all the available power, while in the adaptive case, at high SIR, only part of the transmit power is consumed. Therefore, for applications with target BERs, adaptive modulation achieves a higher spectral efficiency.

For data throughput 8 bits/sec/Hz and BER  $10^{-3}$ , we observe that the adaptive scheme requires SIR 9 dB from Fig. 6.8, and the non-adaptive scheme requires SIR 15 dB from

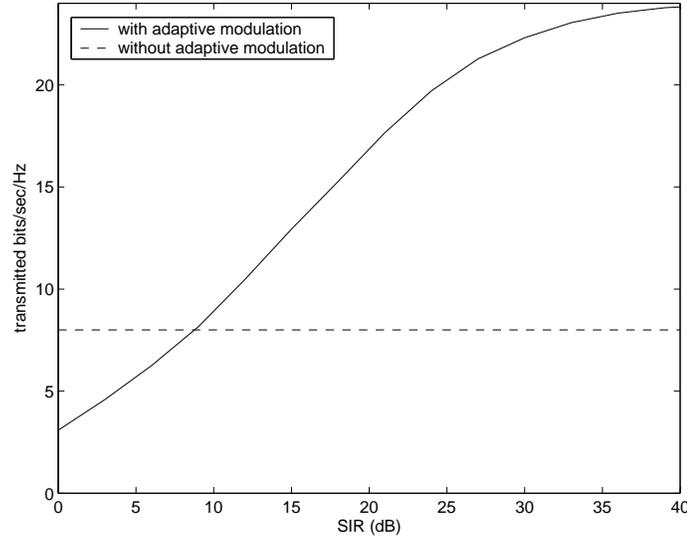


Figure 6.8. The number of bits transmitted per Hz per sec for adaptive modulation and non-adaptive scheme assuming perfect channel knowledge at the transmitter. The target BER in adaptive modulation is  $10^{-3}$ ,  $N_t = N_r = 4$ , and  $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading.

Fig. 6.9. Therefore, there are 6 dB SIR saving by using adaptive modulation.

When channel and spatial interference correlation matrices are estimated from training sequences, with target BER  $10^{-3}$ , Fig. 6.10 shows the actual achieved bit error rate versus training length. We observe that even for moderately long training lengths, e.g., six times the number of transmitting antennas, the actual achieved BER is much higher than the target BER. Hence, with estimated channel and spatial interference correlation matrices, we have to set the target BER much lower than the actual BER. After experiments, it is found that to make the actual BER be  $10^{-3}$  for training length  $6N_t$  and  $4N_t$ , we have to set the target BER to be  $5 \times 10^{-6}$  and  $10^{-8}$ , respectively. The required SNRs for target BERs  $5 \times 10^{-6}$  and  $10^{-8}$  are shown in Table 6.1.

Fig. 6.11 shows the spectral efficiency for adaptive modulation with channel estimates

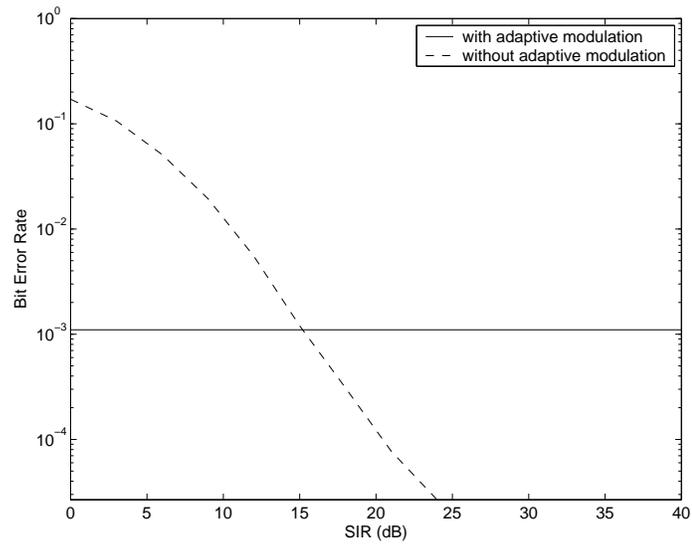


Figure 6.9. Average bit error rate for adaptive modulation and non-adaptive scheme assuming perfect channel knowledge at the transmitter. The target BER in adaptive modulation is  $10^{-3}$ ,  $N_t = N_r = 4$ , and  $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading.

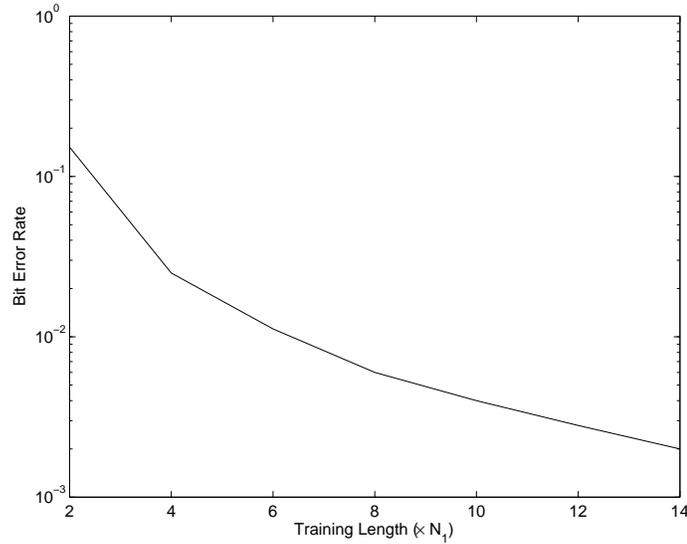


Figure 6.10. The achieved bit error rate versus training length of adaptive modulation with target BER  $10^{-3}$ . It is assumed that  $N_t = N_r = 4$ , and  $L = 5$ . The desired user is assumed to experience i.i.d. Rayleigh fading.

where the actual BERs achieve  $10^{-3}$ . As expected, the throughput improves as channel estimates become accurate. Compared to the case of perfectly known channel, about 3 bits/sec/Hz degradation occurs with training length  $4N_t$ . For example, at SIR 15 dB, with perfectly known channel we can transmit 13 bits/vector symbol, and with training length  $4N_t$  we transmit 10 bits/vector symbol.

The spectral efficiency shown in Fig. 6.11 does not take the training overhead into account. Although a longer training length yields a higher throughput for data transmission period (after training period), it requires more overhead. Therefore, we are interested in overall spectral efficiency which takes the training overhead into account. Assume that the data is transmitted frame by frame, and that the channel is estimated at the beginning of each frame. For a certain frame length, we would like to find out a good training length which maximizes the overall spectral efficiency. For example, at SIR 15 dB, Fig. 6.11

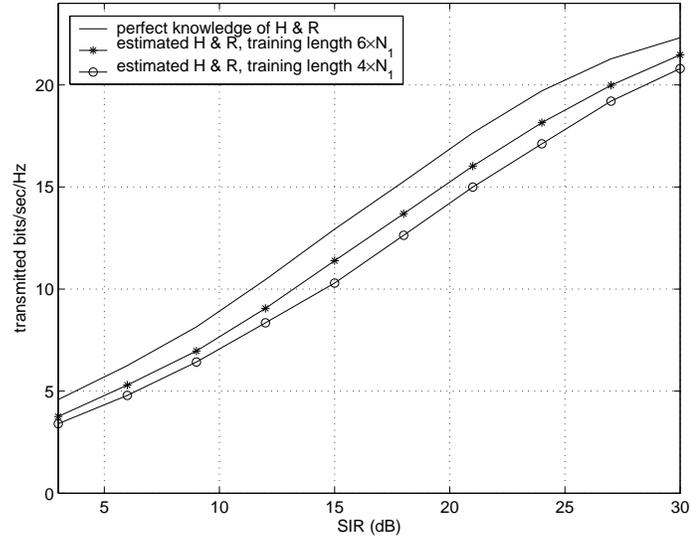


Figure 6.11. The number of bits transmitted per Hz per sec versus SIR for adaptive modulation with  $N_t = N_r = 4$ ,  $L = 5$ . The actual achieved BER is  $10^{-3}$ . The desired user is assumed to experience i.i.d. Rayleigh fading.

shows that for training length  $4N_t$  and  $6N_t$ , we can transmit 10 bits/vector symbol and 11 bits/vector symbol, respectively. If we have  $N_t = 4$  transmitting antennas and a frame length  $M$ , for training length  $4N_t$  and  $6N_t$ , the overall spectral efficiency is  $(M - 4 \times N_t) \times 10/M$  bits/sec/Hz and  $(M - 6 \times N_t) \times 11/M$  bits/sec/Hz, respectively. It can be shown that if the frame length is more than 104 vector symbols, it is worth using training length  $6N_t$ .

To investigate the effect of a LOS component in the desired user's channel, Fig. 6.12 shows the spectral efficiency of adaptive modulation for different Ricean factors where the actual achieved BER is  $10^{-3}$ . We fix SIR to be 10 dB. The scattered components of the desired user's MIMO links are assumed to be independently faded. Recall that as  $K$  increases, the rank of the MIMO channel matrix reduces. In Fig. 6.12, as the LOS component becomes more prominent, the number of subchannels is reduced, hence data throughput decreases. We observe that as  $K$  increases, the loss in spectral efficiency due

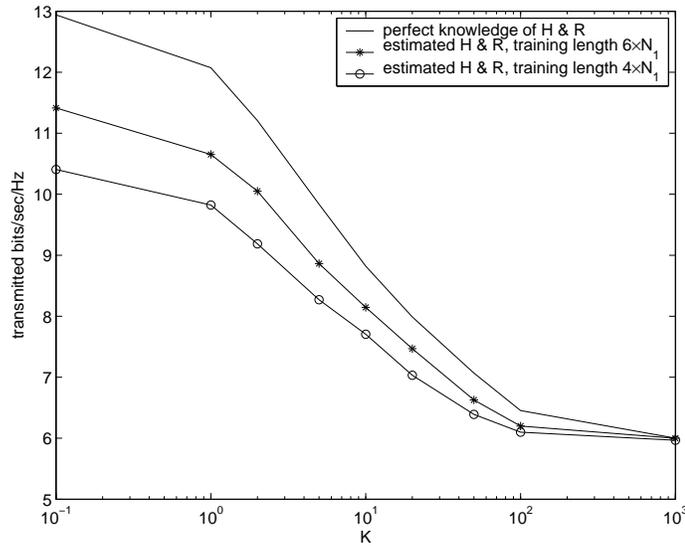


Figure 6.12. The number of bits transmitted per Hz per sec versus Ricean factor  $K$  for adaptive modulation with  $N_t = N_r = 4$ ,  $L = 5$ , SIR=10 dB. The actual achieved BER is  $10^{-3}$ . The scattered components of the desired user's channel are assumed to be independently faded.

to imperfect channel estimates diminishes. This can be explained by the fact that for a fixed training length, the channel estimates become accurate at large values of  $K$ , as we mentioned in Section 4.5.4.3.

To investigate the effect of correlated fading among the desired user's channel links, Fig. 6.13 shows data throughputs for different angle spreads at the base station antenna array with SIR 10 dB. For training length  $4N_t$  and  $6N_t$ , the actual achieved BER is  $10^{-3}$ . We assume correlated Rayleigh fading ( $K = 0$ ) for the desired user. Recall that as the angle spread increases, MIMO links become more and more uncorrelated. Fig. 6.13 shows improvements in data throughput as angle spread increases due to the increased rank of the MIMO channel matrix. It is shown that as the angle spread increases, the loss in spectral efficiency due to imperfect channel estimates slightly increases. This can be explained by

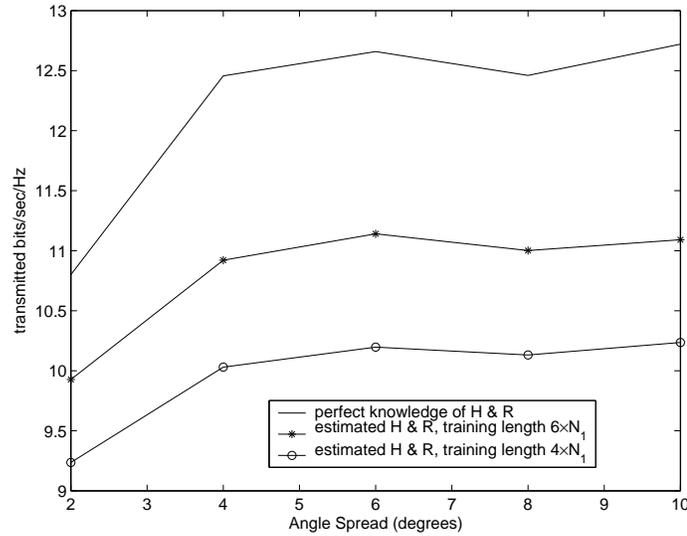


Figure 6.13. The number of bits transmitted per Hz per sec versus angle spread for adaptive modulation with  $N_t = N_r = 4$ ,  $L = 5$ , SIR=10 dB. The actual achieved BER is  $10^{-3}$ . The desired user is assumed to experience correlated Rayleigh fading ( $K = 0$ ).

the fact that for a fixed training length, the channel estimates become less accurate for a large angle spread, as we mentioned in Section 4.5.4.3. We also observe that when the angle spread is larger than 4 degrees, system performance is invariant to angle spread.

## 6.5 Effect of Feedback Quantization

To enable adaptive modulation at the transmitter, the receiver has to feed the channel information back to the transmitter. In some practical systems, the information that needs to be fed back is represented by an  $N_q$ -bit description via some quantization process. This scenario arises when a digital channel with limited bandwidth is allocated for feedback. As a result, transmitter and receiver may not have the same channel information due to quantization error. In practice, we have both imperfect feedback and imperfect channel estimates.

To separate the impact of imperfect feedback from that of imperfect channel estimates and for the purpose of analysis, in this section, we assume perfect channel knowledge at the receiver and focus on the effect of feedback quantization. Investigation of the combined impact of imperfect channel estimates and imperfect feedback is left for future work.

### 6.5.1 Analysis of quantization error

We assess the achievable quantization error by rate distortion theory [84]. To enable the analysis, we consider a noise-limited environment with negligible interference and spatially white Gaussian noise. Independent Rayleigh fading is assumed for the desired user where channel matrix  $\mathbf{H}$  has  $N_r \times N_t$  i.i.d. zero-mean circularly symmetric complex Gaussian components with unit variance. Equivalently, with real and imaginary parts,  $\mathbf{H}$  has  $2N_r N_t$  i.i.d. zero-mean real Gaussian random variables each with variance  $\sigma_H^2 = 1/2$ .

Vector quantization is employed to transmit knowledge of  $\mathbf{H}$  from the receiver to transmitter. We use  $N_q$  bits to describe the  $2N_r N_t$  real elements of  $\mathbf{H}$ . Letting  $\tilde{\mathbf{H}}$  be the reconstruction of the quantized  $\mathbf{H}$ , we consider vector quantization that minimizes mean squared error [84]

$$\sigma_\epsilon^2 = \frac{1}{2N_r N_t} \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} \sum_{k \in \{\text{real, imag}\}} E \left[ (H_{ijk} - \tilde{H}_{ijk})^2 \right]$$

where  $H_{ijk}$  and  $\tilde{H}_{ijk}$  denote the real-valued elements of matrix  $\mathbf{H}$  and  $\tilde{\mathbf{H}}$ , respectively. Since the real and imaginary components in  $\mathbf{H}$  are i.i.d. Gaussian, the mean squared error can be lower-bounded by the corresponding distortion-rate function [25], i.e.,

$$\sigma_\epsilon^2 \geq D(R_q) = \sigma_H^2 2^{-2R_q} = \frac{1}{2} 2^{-2R_q} \quad (6.9)$$

where  $R_q = N_q / (2N_r N_t)$  is the number of descriptive bits per real component of  $\mathbf{H}$ . For a fixed  $R_q$ , this lower bound can be approached arbitrarily closely as  $2N_r N_t$  goes to infinity.

For a very large  $2N_r N_t$ , this lower bound is achieved when the quantization error  $\epsilon = \mathbf{H} - \tilde{\mathbf{H}}$  is a zero-mean complex Gaussian random matrix, independent of  $\tilde{\mathbf{H}}$ , with i.i.d. real and imaginary parts each having variance  $D(R_q)$ . More specifically, for  $\epsilon_{ijk} = H_{ijk} - \tilde{H}_{ijk}$ , the lower bound of quantization error is achieved when  $\tilde{H}_{ijk} \sim \mathcal{N}(0, \sigma_H^2 - D(R_q))$ ,  $\epsilon_{ij} \sim \mathcal{N}(0, D(R_q))$ , and  $\tilde{H}_{ijk}$  is independent of  $\epsilon_{ijk}$ . Since  $H_{ijk} = \tilde{H}_{ijk} + \epsilon_{ijk}$ ,  $H_{ijk}$  and  $\tilde{H}_{ijk}$  are jointly Gaussian random variables with covariance

$$\begin{aligned} \text{cov}(H_{ijk}, \tilde{H}_{ijk}) &= E[H_{ijk}\tilde{H}_{ijk}] \\ &= E[(\tilde{H}_{ijk} + \epsilon_{ijk})\tilde{H}_{ijk}] \\ &= \sigma_H^2 - D(R_q) \end{aligned}$$

where the last equality comes from the fact that  $\epsilon_{ijk}$  and  $\tilde{H}_{ijk}$  are zero-mean and independent.

Fig. 6.14 shows the adaptive modulation with quantized feedback. The transmitter determines the unitary pre-processing matrix and adaptive modulation according to the SVD of the quantized channel matrix  $\tilde{\mathbf{H}} = \tilde{\mathbf{U}}_H \tilde{\mathbf{\Phi}}_H \tilde{\mathbf{V}}_H^\dagger$ . The receiver uses the SVD of the perfectly known channel matrix  $\mathbf{H} = \mathbf{U}_H \mathbf{\Phi}_H \mathbf{V}_H^\dagger$  to determine the post-processing matrix and scale the signal constellations for data detection.

## 6.5.2 Simulation results

Monte Carlo simulation is used to evaluate the effect of feedback quantization on the performance of adaptive modulation. To simulate matrices  $\mathbf{H}$  and  $\tilde{\mathbf{H}}$ , i.i.d. pairs of  $(H_{ijk}, \tilde{H}_{ijk})$ ,  $i = 1, \dots, N_r, j = 1, \dots, N_t, k \in \{\text{real}, \text{imag}\}$  are generated. For each pair,  $H_{ijk}$  and  $\tilde{H}_{ijk}$  are jointly real Gaussian random variables with

$$\begin{pmatrix} H_{ijk} \\ \tilde{H}_{ijk} \end{pmatrix} \sim \mathcal{N} \left( \mathbf{0}, \begin{bmatrix} 0.5 & 0.5 - D(R_q) \\ 0.5 - D(R_q) & 0.5 - D(R_q) \end{bmatrix} \right)$$

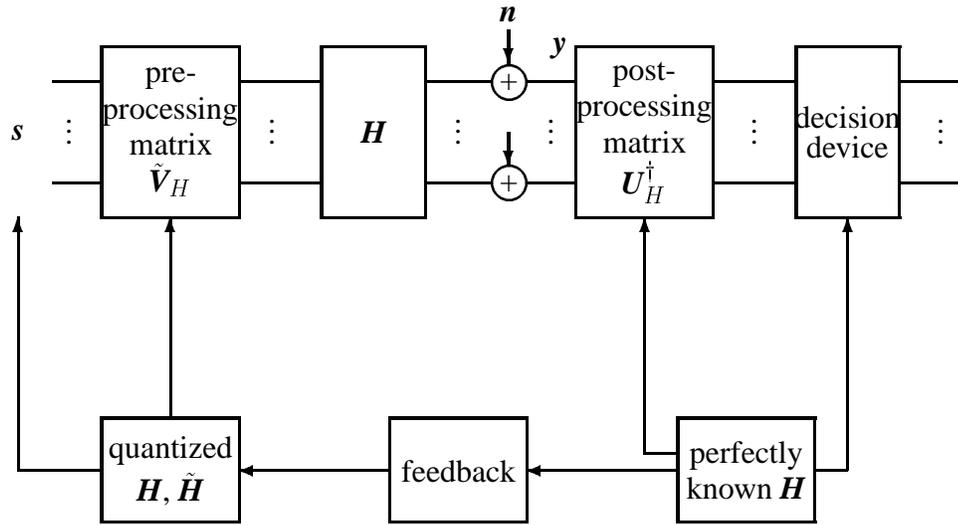


Figure 6.14. Adaptive modulation with feedback quantization.

where  $D(R_q)$  is shown in (6.9).

We consider a  $(4, 4)$  MIMO system with SNR 15 dB. Figs. 6.15 and 6.16 show the upper bounds of system performance (best possible performance) versus quantization level with target BER  $10^{-2}$ . In Fig. 6.15, we observe that the BER improves significantly as the quantization level increases, and eventually reaches the target BER. In Fig. 6.16, we observe that the spectral efficiency increases as the transmitter obtains more accurate channel information. Both figures suggest that, to achieve target performance, 15 bits/complex channel gain should be used for the feedback quantization. Figs. 6.17 and 6.18 show the performance for target BER  $10^{-3}$ . We observe that 22 bits/complex channel gain should be used for the feedback quantization. Hence, for a  $(4, 4)$  MIMO link, with target BER  $10^{-3}$ , we need  $22 \times 16 = 352$  bits to quantize the channel matrix. Since the receiver has four antennas, ignoring channel coding and assuming that each antenna uses QPSK, we need  $352/(4 \times 2) = 44$  vector symbols for the feedback of channel matrix.

To see how much overhead requires due to the feedback of channel matrix, we consider

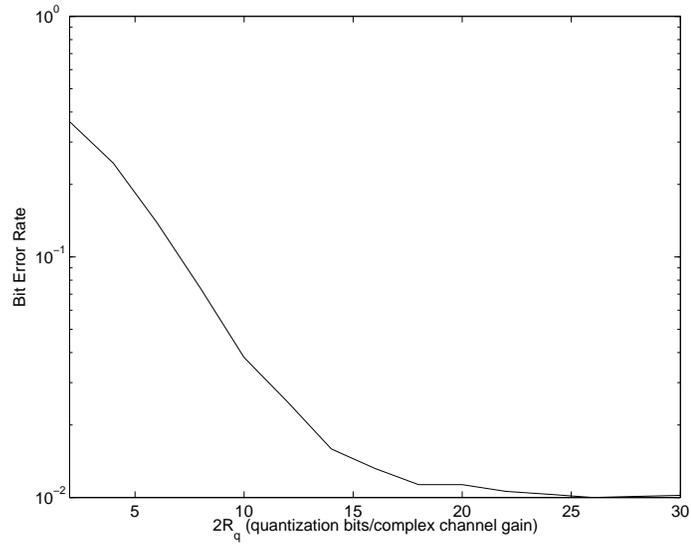


Figure 6.15. Average bit error rate versus quantization level of feedback with  $N_t = N_r = 4$ , SNR=15 dB and target BER  $10^{-2}$ .

a system operating at a carrier frequency of 1.9 GHz and symbol rate 24.3 ksymbol/sec in a bandwidth of 30 kHz [47]. If the mobile is moving at  $v = 5$  m/sec (18 km/hour), the maximum Doppler shift  $f_m = v/\lambda = 31.7$  Hz where  $\lambda$  is the wavelength at the carrier frequency. The normalized Doppler spread  $f_m/\text{symbol rate} = 0.13\%$ . The normalized Doppler spread can also be interpreted as the ratio of symbol duration to channel coherence time since the coherence time is inversely proportional to the maximum Doppler shift [91]. Coherence time is the time duration over which the channel is essentially invariant. Assuming that the frame length is 200 vector symbols (the assumption of quasi-static channel holds for normalized Doppler spread 0.13%), for the (4, 4) MIMO link and target BER  $10^{-3}$  discussed above, the overhead for feedback is  $44/200 = 22\%$ .

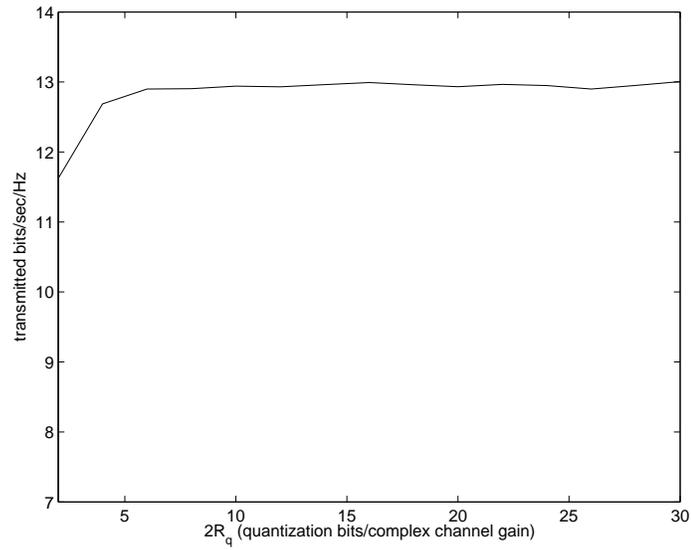


Figure 6.16. The number of bits transmitted per Hz per sec versus quantization level of feedback with  $N_t = N_r = 4$ , SNR=15dB and target BER  $10^{-2}$ .

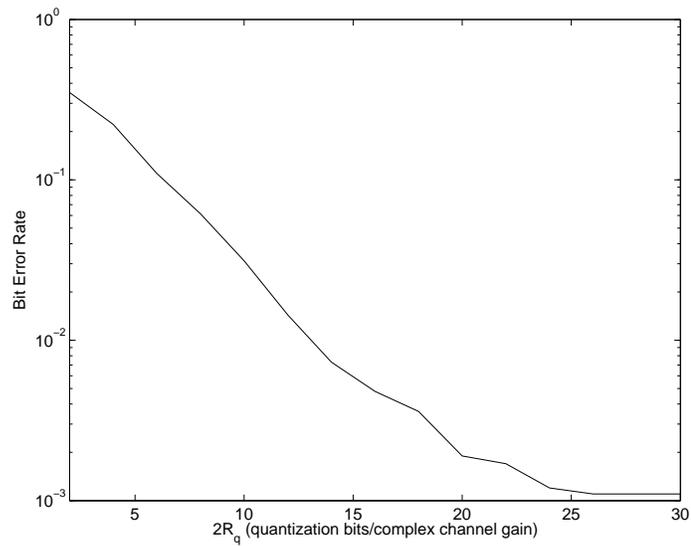


Figure 6.17. Average bit error rate versus quantization level of feedback with  $N_t = N_r = 4$ , SNR=15 dB and target BER  $10^{-3}$ .

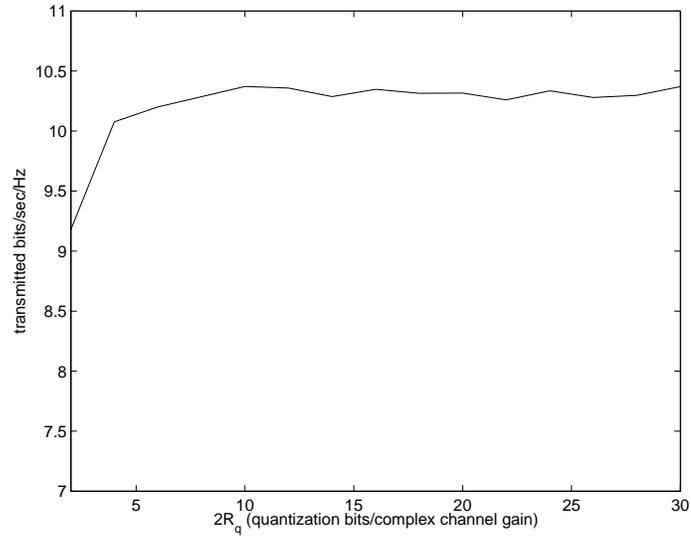


Figure 6.18. The number of bits transmitted per Hz per sec versus quantization level of feedback with  $N_t = N_r = 4$ , SNR=15dB and target BER  $10^{-3}$ .

## 6.6 Conclusion

In this chapter, we have investigated the effects of imperfect channel estimates and feedback quantization error on adaptive modulation for MIMO systems. We show that for applications with target BERs, adaptive modulation achieves higher spectral efficiency compared to the non-adaptive scheme. In studying the effect of imperfect channel estimates, we assume a perfect feedback path from the receiver to transmitter. The ML estimates of channel and spatial interference correlation matrices in Chapter 4 are employed. Compared to the case of perfectly known channel, for a  $(4, 4)$  MIMO system, 3 bits/sec/Hz degradation in spectral efficiency occurs for a training length equal to four times the number of transmitting antennas and the actual achieved BER  $10^{-3}$ . When the angle spread at base station antenna array is larger than 4 degrees, the system performance is close to the case where

MIMO links are independently faded. In analyzing the effect of feedback quantization error, we assume perfect channel knowledge at the receiver and a noise-limited environment. Rate distortion theory is used to assess the achievable quantization error. It is shown that for a  $(4, 4)$  MIMO link with independent Rayleigh fading, to achieve target BER  $10^{-3}$ , 22 bits/complex channel gain should be used for the feedback quantization, which may be translated into the required overhead for feedback.

## Chapter 7

### Conclusions and Future Work

In this chapter, we summarize the major contributions in this thesis and suggest several topics for future research.

#### 7.1 Conclusions

In this thesis, we have studied the impact of spatially and temporally colored interference on the performance of MIMO systems.

In Chapter 3, outage performances of several diversity schemes have been analytically compared for an interference-limited environment in a Rayleigh fading channel. The three diversity schemes are equal gain combining (EGC), selection combining (SC) and channel-matched combining (CMC), a practical variation of maximal ratio combining (MRC) that does not require SNRs at different antennas. An exact outage probability expression was derived for EGC by accurately calculating the interference power at the output of the combiner. With this exact analysis, we show that the existing method, which calculates the interference power approximately, may lead to optimistic outage probability estimates. With four receiving antennas, the existing method may overestimate the output SIR of EGC combiner by as much as 1.5dB. The comparisons for diversity schemes show that CMC has a

lower outage probability than that of EGC, and that CMC has no greater outage probability than that of SC. The relative outage performance between EGC and SC, however, depends on the number of interferers and interferer power distribution.

In Chapter 4, we investigated algorithms of channel estimation and data detection for MIMO systems under spatially and temporally colored interference. By modelling interference statistics as being approximately temporally and spatially separable, we have proposed an algorithm to jointly estimate channel and spatial interference correlation matrices. By exploiting the temporal interference correlation, one-vector-symbol detection has been extended to a multi-vector-symbol version. In the case of one interferer, we have shown that the interference statistics is indeed temporally and spatially separable, and that the temporal interference correlation may be caused by the intersymbol interference of the interferer. The impact of temporal and spatial interference correlation on channel estimation and data detection was assessed. We also investigated the situation where high temporal interference correlation results from a cochannel interferer operating at a lower data rate. The results show that the benefit of taking temporal interference correlation into account in channel estimation is not significant for both same- and lower-data-rate interference cases. In the case of lower-data-rate interference, much improvement can be achieved by taking account of the temporal interference correlation in data detection. For a  $(5, 5)$  MIMO link with independent Rayleigh fading, by exploiting temporal interference correlation in channel estimation and data detection, we obtain 0.5dB and 3.5dB gains in SIR for same- and lower-data-rate interference cases, respectively.

In Chapter 5, with spatially and temporally colored interference, we assessed the benefit of knowing channel matrix and interference statistics at the transmitter from the view point of information theory. Assuming that the receiver knows the channel matrix and interference statistics, we derived the channel capacities of MIMO systems with different

assumptions of knowledge of channel matrix and interference statistics at the transmitter. It is shown that, with interference spatial and temporal correlation available at the transmitter, water-filling power allocation in both spatial and frequency domains should be applied. With the temporal interference correlation of the same-data-rate interference case in Chapter 4, the results show that knowledge of interference temporal correlation at the transmitter does not offer significant capacity improvement.

In Chapter 6, we proposed an adaptive modulation scheme using the channel matrix and interference statistics estimated by the receiver. We investigated the effects of imperfect channel estimates and feedback quantization error on adaptive modulation for MIMO system. In studying the effect of imperfect channel estimates, we assume a perfect feedback path from the receiver to transmitter and use the ML estimates of channel and spatial interference correlation matrices presented in Chapter 4. Compared to the case of perfectly known channel, for a  $(4, 4)$  MIMO system, it is shown that 3 bits/sec/Hz degradation in spectral efficiency occurs for a training length equal to four times the number of transmitting antennas and the actual achieved BER  $10^{-3}$ . In analyzing the effect of feedback quantization error, we assume perfect channel knowledge at the receiver and a noise-limited environment. Rate distortion theory was used to assess the achievable quantization error for independently Rayleigh faded MIMO links. It is shown that, for a  $(4, 4)$  link, to achieve the target BER  $10^{-3}$ , 22 bits/complex channel gain should be used for the feedback quantization, which may be used to calculate the required overhead for feedback.

The results in this thesis also show that a training length equal to four times the number of transmitting antennas achieves most of the performance gain. It is observed that for angle spreads larger than 5 degrees, the system performance is close to the case where MIMO links are independently faded.

## 7.2 Future Work

Here are some suggestions for future work.

- In Chapter 3, we analytically compared the performance of diversity systems in a Rayleigh fading environment. One may compare the performance of these diversity systems in Nakagami fading. However, it may not be easy to conduct an analytical comparison.
- The results in Chapter 4 show that in the case of interference from lower-data-rate users, much improvement can be achieved by taking account of the temporal interference correlation in multi-vector-symbol data detection. It is therefore of interest to investigate how to estimate an interferer's delay so that the temporal interference correlation can be estimated.
- In Chapter 6, we considered the effects of imperfect channel estimates and feedback error on adaptive modulation separately. It would be useful to evaluate the performance degradation caused by both impairments. An alternative model of feedback error as in [84] may be used.
- The adaptive modulation scheme in Chapter 6 requires the whole channel matrix to be fed back from receiver to transmitter. When the number of antennas in the system becomes large, the amount of feedback information would be impractically large. It is of interest to investigate transmission schemes which requires a reduced amount of feedback information.

## Appendix A

### Variance of $z_i[n]$

In this appendix, we show that the variance of  $z_i[n]$  in (3.3) is  $1 - \rho/4$ . We have

$$\begin{aligned} E \{ |z_i[n]|^2 \} &= E \{ z_i[n] z_i^*[n] \} \\ &= \sum_{m=-\infty}^{\infty} E \{ g^2(nT - mT - \tau_i) \} \end{aligned} \quad (\text{A.1})$$

$$= \sum_{k=-\infty}^{\infty} E \{ g^2(kT - \tau_i) \} \quad (\text{A.2})$$

$$\begin{aligned} &= \frac{1}{T} \sum_{k=-\infty}^{\infty} \int_0^T g^2(kT - \tau) d\tau \\ &= \frac{1}{T} \sum_{k=-\infty}^{\infty} \int_{(k-1)T}^{kT} g^2(t) dt \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} &= \frac{1}{T} \int_{-\infty}^{\infty} g^2(t) dt \\ &= \frac{1}{T} \int_{-\infty}^{\infty} |G(f)|^2 df \end{aligned} \quad (\text{A.4})$$

where (A.1) follows from the assumption that the data symbols of the  $i$ th interferer are independent and with unit variance; (A.2) follows from the substitution  $k = n - m$ ; (A.3) follows from the substitution  $t = kT - \tau$ ; (A.4) follows from Parseval's theorem where  $G(f)$  is the Fourier transform of  $g(t)$ . Since  $g(t)$  is a Nyquist pulse with a raised cosine

spectrum and rolloff factor  $\rho$ , we have [89]

$$G(f) = \begin{cases} T & 0 \leq |f| \leq \frac{1-\rho}{2T} \\ \frac{T}{2} \left\{ 1 + \cos \left[ \frac{\pi T}{\rho} \left( |f| - \frac{1-\rho}{2T} \right) \right] \right\} & \frac{1-\rho}{2T} \leq |f| \leq \frac{1+\rho}{2T} \\ 0 & |f| > \frac{1+\rho}{2T} \end{cases}$$

Hence, we have

$$\begin{aligned} \frac{1}{T} \int_{-\infty}^{\infty} |G(f)|^2 df &= 1 - \rho + \frac{T}{2} \int_{\frac{1-\rho}{2T}}^{\frac{1+\rho}{2T}} \left\{ 1 + \cos \left[ \frac{\pi T}{\rho} \left( f - \frac{1-\rho}{2T} \right) \right] \right\}^2 df \\ &= 1 - \rho + \frac{\rho}{2} \int_0^1 (1 + \cos \pi u)^2 du \\ &= 1 - \frac{\rho}{4} \end{aligned}$$

Hence, the variance of  $z_i[n]$  is  $1 - \rho/4$ .

## Appendix B

### Circular Symmetry of $\frac{\mathbf{c}_s^\dagger \mathbf{c}_i}{|\mathbf{c}_s|}$

In this appendix, we show that  $v_i \triangleq \frac{\mathbf{c}_s^\dagger \mathbf{c}_i}{|\mathbf{c}_s|}$  in (3.23) is a circularly symmetric complex Gaussian RV with zero-mean and unit variance. Note that  $\mathbf{c}_s$  and  $\mathbf{c}_i$  are independent zero-mean circularly symmetric complex Gaussian random vectors each with covariance matrix  $\mathbf{I}_{N_r}$ .

Let  $\mathbf{c}_s = \mathbf{c}_{s,R} + j\mathbf{c}_{s,I}$  and  $\mathbf{c}_i = \mathbf{c}_{i,R} + j\mathbf{c}_{i,I}$  where  $\mathbf{c}_{s,R}$ ,  $\mathbf{c}_{s,I}$ ,  $\mathbf{c}_{i,R}$  and  $\mathbf{c}_{i,I}$  are real Gaussian vectors. Hence,  $v_i = v_{i,R} + jv_{i,I}$  where the real and imaginary parts are

$$v_{i,R} = \frac{\mathbf{c}_{s,R}^T \mathbf{c}_{i,R} + \mathbf{c}_{s,I}^T \mathbf{c}_{i,I}}{|\mathbf{c}_s|}$$

and

$$v_{i,I} = \frac{\mathbf{c}_{s,R}^T \mathbf{c}_{i,I} - \mathbf{c}_{s,I}^T \mathbf{c}_{i,R}}{|\mathbf{c}_s|}.$$

For a given  $\mathbf{c}_s$ ,  $v_{i,R}$  and  $v_{i,I}$  are zero-mean joint Gaussians since they are linear transformation of zero-mean joint Gaussians  $[\mathbf{c}_{i,R}^T \ \mathbf{c}_{i,I}^T]^T$ . The variance of  $v_{i,R}$  conditioned on  $\mathbf{c}_s$  is

$$\begin{aligned} \text{var} [v_{i,R} | \mathbf{c}_s] &= E [v_{i,R}^2 | \mathbf{c}_s] \\ &= E [v_{i,R} v_{i,R}^T | \mathbf{c}_s] \end{aligned}$$

$$\begin{aligned}
&= E \left[ \frac{\left( \mathbf{c}_{s,R}^T \mathbf{c}_{i,R} + \mathbf{c}_{s,I}^T \mathbf{c}_{i,I} \right) \left( \mathbf{c}_{i,R}^T \mathbf{c}_{s,R} + \mathbf{c}_{i,I}^T \mathbf{c}_{s,I} \right)}{|\mathbf{c}_s|^2} \middle| \mathbf{c}_s \right] \\
&= \frac{\mathbf{c}_{s,R}^T E \left[ \mathbf{c}_{i,R} \mathbf{c}_{i,R}^T \right] \mathbf{c}_{s,R} + \mathbf{c}_{s,I}^T E \left[ \mathbf{c}_{i,I} \mathbf{c}_{i,R}^T \right] \mathbf{c}_{s,R} + \mathbf{c}_{s,R}^T E \left[ \mathbf{c}_{i,R} \mathbf{c}_{i,I}^T \right] \mathbf{c}_{s,I} + \mathbf{c}_{s,I}^T E \left[ \mathbf{c}_{i,I} \mathbf{c}_{i,I}^T \right] \mathbf{c}_{s,I}}{|\mathbf{c}_s|^2} \\
&= \frac{1}{2} \frac{\mathbf{c}_{s,R}^T \mathbf{c}_{s,R} + \mathbf{c}_{s,I}^T \mathbf{c}_{s,I}}{|\mathbf{c}_s|^2} \\
&= \frac{1}{2}
\end{aligned}$$

where  $E \left[ \mathbf{c}_{i,R} \mathbf{c}_{i,R}^T \right] = E \left[ \mathbf{c}_{i,I} \mathbf{c}_{i,I}^T \right] = \frac{1}{2} \mathbf{I}_{N_r}$  and  $E \left[ \mathbf{c}_{i,I} \mathbf{c}_{i,R}^T \right] = E \left[ \mathbf{c}_{i,R} \mathbf{c}_{i,I}^T \right] = \mathbf{0}$ . Similarly, the variance of  $v_{i,I}$  conditioned on  $\mathbf{c}_s$  is

$$\text{var} [v_{i,I} | \mathbf{c}_s] = E [v_{i,I}^2 | \mathbf{c}_s] = \frac{1}{2}.$$

The covariance of  $v_{i,R}$  and  $v_{i,I}$  conditioned on  $\mathbf{c}_s$  is

$$\begin{aligned}
&\text{cov} (v_{i,R}, v_{i,I} | \mathbf{c}_s) \\
&= E [v_{i,R} v_{i,I} | \mathbf{c}_s] \\
&= E \left[ v_{i,R} v_{i,I}^T | \mathbf{c}_s \right] \\
&= E \left[ \frac{\left( \mathbf{c}_{s,R}^T \mathbf{c}_{i,R} + \mathbf{c}_{s,I}^T \mathbf{c}_{i,I} \right) \left( \mathbf{c}_{i,I}^T \mathbf{c}_{s,R} - \mathbf{c}_{i,R}^T \mathbf{c}_{s,I} \right)}{|\mathbf{c}_s|^2} \middle| \mathbf{c}_s \right] \\
&= \frac{1}{2} \frac{\mathbf{c}_{s,I}^T \mathbf{c}_{s,R} - \mathbf{c}_{s,R}^T \mathbf{c}_{s,I}}{|\mathbf{c}_s|^2} \\
&= 0.
\end{aligned}$$

Hence, given  $\mathbf{c}_s$ ,  $v_i$  is a circularly symmetric zero-mean complex Gaussian with unit variance, and its PDF is

$$f_{v_i}(v_i | \mathbf{c}_s) = \pi^{-1} e^{-|v_i|^2}.$$

Since  $f_{v_i}(v_i | \mathbf{c}_s)$  is independent of  $\mathbf{c}_s$ , it is clear that  $v_i$  is independent of  $\mathbf{c}_s$ . Hence,  $v_i$  is a circularly symmetric zero-mean complex Gaussian with unit variance (without conditioning on  $\mathbf{c}_s$ ).

## Appendix C

### Alternative derivation of $\hat{\mathbf{H}}$ and $\hat{\mathbf{R}}$

To find the value of  $(\mathbf{H}, \mathbf{R})$  that minimizes  $f(\mathbf{H}, \mathbf{R})$  in (4.7), we set  $\partial f(\mathbf{H}, \mathbf{R})/\partial \mathbf{R} = \mathbf{0}$ . Using the identities of matrix derivative in (4.13) and (4.14), we obtain

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i) (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)^\dagger. \quad (\text{C.1})$$

Substituting  $\hat{\mathbf{R}}$  into (4.7), we note that the estimate of  $\mathbf{H}$  is determined by minimizing the determinant

$$f_1(\mathbf{H}) = \det \left\{ \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i) (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)^\dagger \right\}. \quad (\text{C.2})$$

With the weighted sample correlation matrices defined in (4.9)-(4.11), the matrix in (C.2), denoted as  $\mathbf{F}$ , can be expressed as

$$\mathbf{F} = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{i,j} (\mathbf{y}_i - \mathbf{H}\mathbf{x}_i) (\mathbf{y}_j - \mathbf{H}\mathbf{x}_j)^\dagger \quad (\text{C.3})$$

$$= \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right) \tilde{\mathbf{R}}_{xx} \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right)^\dagger + \tilde{\mathbf{R}}_{yy} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \tilde{\mathbf{R}}_{xy}. \quad (\text{C.4})$$

Now we show that both  $\left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right) \tilde{\mathbf{R}}_{xx} \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right)^\dagger$  and  $\tilde{\mathbf{R}}_{yy} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \tilde{\mathbf{R}}_{xy}$  are positive semidefinite.

**Lemma 3** *Matrix  $\mathbf{B}_1 = \tilde{\mathbf{R}}_{yy} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \tilde{\mathbf{R}}_{xy}$  is positive semidefinite.*

*Proof:* Notice that matrix  $\mathbf{B}_1$  is equal to matrix  $\mathbf{F}$  in (C.3) if  $\mathbf{H} = \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1}$ . We define matrix  $\mathbf{A}$  as

$$\mathbf{A} = \begin{bmatrix} \mathbf{y}_0 - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \mathbf{x}_0 & \cdots & \mathbf{y}_{N-1} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \mathbf{x}_{N-1} \end{bmatrix},$$

and matrix  $\mathbf{B}_1$  can be re-expressed as

$$\mathbf{B}_1 = \frac{1}{N} \mathbf{A} \Lambda_N^{-1} \mathbf{A}^\dagger$$

where  $\Lambda_N^{-1}$  is defined in (4.5). For any  $N_r \times 1$  complex vector  $\mathbf{w}$ ,

$$\mathbf{w}^\dagger \mathbf{B}_1 \mathbf{w} = \frac{1}{N} \mathbf{w}^\dagger \mathbf{A} \Lambda_N^{-1} \mathbf{A}^\dagger \mathbf{w} = \frac{1}{N} \mathbf{b}_1^\dagger \Lambda_N^{-1} \mathbf{b}_1$$

where the  $N \times 1$  vector  $\mathbf{b}_1 = \mathbf{A}^\dagger \mathbf{w}$ . Since  $\Lambda_N$  is positive definite,  $\Lambda_N^{-1}$  is positive definite as well. Hence,  $\mathbf{w}^\dagger \mathbf{B}_1 \mathbf{w} \geq 0$  for any  $\mathbf{w}$ , and matrix  $\mathbf{B}_1$  is positive semidefinite. ■

**Lemma 4** Matrix  $\mathbf{B}_2 = \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right) \tilde{\mathbf{R}}_{xx} \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right)^\dagger$  is positive semidefinite.

*Proof:* Define matrix

$$\mathbf{X} = [\mathbf{x}_0 \cdots \mathbf{x}_{N-1}]$$

and rewrite  $\tilde{\mathbf{R}}_{xx}$  in (4.11) as

$$\tilde{\mathbf{R}}_{xx} = \frac{1}{N} \mathbf{X} \Lambda_N^{-1} \mathbf{X}^\dagger.$$

For any  $N_r \times 1$  complex vector  $\mathbf{w}$ , we have

$$\begin{aligned} & \mathbf{w}^\dagger \mathbf{B}_2 \mathbf{w} \\ &= \frac{1}{N} \mathbf{w}^\dagger \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right) \mathbf{X} \Lambda_N^{-1} \mathbf{X}^\dagger \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right)^\dagger \mathbf{w} \\ &= \frac{1}{N} \mathbf{b}_2^\dagger \Lambda_N^{-1} \mathbf{b}_2 \end{aligned}$$

where the  $N \times 1$  vector  $\mathbf{b}_2 = \mathbf{X}^\dagger \left( \mathbf{H} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \right)^\dagger \mathbf{w}$ . Again, since  $\Lambda_N^{-1}$  is positive definite,  $\mathbf{w}^\dagger \mathbf{B}_2 \mathbf{w} \geq 0$  for any  $\mathbf{w}$ , and matrix  $\mathbf{B}_2$  is positive semidefinite. ■

To find the value of  $\mathbf{H}$  that maximizes  $\det(\mathbf{F})$  in (C.4), we introduce the following lemma [75].

**Lemma 5** *If matrices  $\mathbf{A}$  ( $m \times m$ ) and  $\mathbf{B}$  ( $m \times m$ ) are positive semidefinite, then*

$$\det(\mathbf{A} + \mathbf{B}) \geq \det(\mathbf{A}) + \det(\mathbf{B}).$$

Notice that, in (C.4),  $\tilde{\mathbf{R}}_{yy} - \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1} \tilde{\mathbf{R}}_{xy}$  is independent of  $\mathbf{H}$ . By applying Lemma 3, it is easy to see that  $\det(\mathbf{F})$  is minimized by choosing

$$\hat{\mathbf{H}} = \tilde{\mathbf{R}}_{xy}^\dagger \tilde{\mathbf{R}}_{xx}^{-1}. \quad (\text{C.5})$$

Substituting (C.5) into (C.1), we obtain

$$\hat{\mathbf{R}} = \tilde{\mathbf{R}}_{yy} - \hat{\mathbf{H}} \tilde{\mathbf{R}}_{xy}. \quad (\text{C.6})$$

## Appendix D

### Proof of $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$

The matrix identity  $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$  has been used extensively in the work of channel capacity of MIMO systems [30, 104, 114]. However, the author has not seen a proof of this identity in the literature. In this appendix, we will prove this identity.

**Lemma 6** *If  $\mathbf{A}$  is an  $m \times n$  matrix,  $\mathbf{B}$  is an  $n \times m$  matrix, and  $m \leq n$ , we have*

$$\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA}). \quad (\text{D.1})$$

*Proof:* Theorem 1.3.20 in [75] states that  $\mathbf{BA}$  has the same eigenvalues as  $\mathbf{AB}$ , together with an additional  $n - m$  zero eigenvalues; that is,  $p_{BA}(t) = t^{n-m} p_{AB}(t)$ , where the characteristic polynomial of  $\mathbf{AB}$  is defined as  $p_{AB}(t) = \det(t\mathbf{I} - \mathbf{AB})$ . Therefore, we have

$$\begin{aligned} p_{BA}(t) &= t^{n-m} p_{AB}(t) \\ \implies \det(t\mathbf{I}_n - \mathbf{BA}) &= t^{n-m} \det(t\mathbf{I}_m - \mathbf{AB}) \\ \implies (-1)^n \det(\mathbf{BA} - t\mathbf{I}_n) &= t^{n-m} (-1)^m \det(\mathbf{AB} - t\mathbf{I}_m). \end{aligned} \quad (\text{D.2})$$

Letting  $t = -1$  in (D.2) yields (D.1). ■

## Bibliography

- [1] V. A. Aalo and C. Chayawan. Systems using maximal ratio combining in Rayleigh fading channel with multiple interferers. *Electron. Lett.*, 36(15):1314–1315, July 2000.
- [2] V. A. Aalo and J. Zhang. Performance of antenna array systems with optimum combining in a Rayleigh fading environment. *IEEE Commun. Lett.*, pages 125–127, Apr. 2000.
- [3] V. A. Aalo and J. Zhang. Performance analysis of maximal ratio combining in the presence of multiple equal-power cochannel interferers in a Nakagami fading channel. *IEEE Trans. Veh. Technol.*, pages 497–503, Mar. 2001.
- [4] A. Abdi and M. Kaveh. A space-time correlation model for multielement antenna systems in mobile fading channels. *IEEE J. Select. Areas Commun.*, pages 550–560, Apr. 2002.
- [5] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover, 1971.
- [6] A. A. Abu-Dayya and N. C. Beaulieu. Outage probabilities of diversity cellular systems with cochannel interference in Nakagami fading. *IEEE Trans. Veh. Technol.*, pages 343–355, Nov. 1992.

- [7] F. Adachi, M. Sawahashi, and H. Suda. Wideband DS-CDMA for next-generation mobile communication systems. *IEEE Commun. Mag.*, pages 56–69, Sept. 1998.
- [8] A. Agrawal, G. Ginis, and J. M. Cioffi. Channel diagonalization through orthogonal space-time coding. In *Proc. IEEE Int. Conf. Communications*, 2002.
- [9] S. B aro, G. Bauch, A. Pavilic, and A. Semmler. Improving BLAST performance using space-time block codes and turbo decoding. In *Proc. IEEE Global Telecommun. Conf.*, 2000.
- [10] N. C. Beaulieu. An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of Rayleigh random variables. *IEEE Trans. Commun.*, pages 1463–1474, Sept. 1990.
- [11] N. C. Beaulieu and A. A. Abu-Dayya. Bandwidth efficient QPSK in cochannel interference and fading. *IEEE Trans. Commun.*, pages 2464–2474, Sept. 1995.
- [12] E. Biglieri, J. Proakis, and S. Shamai. Fading channels: information-theoretic and communications aspects. *IEEE Trans. Inform. Theory*, pages 2619–2692, Oct. 1998.
- [13] J. A. C. Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Commun. Mag.*, pages 5–14, May 1990.
- [14] R. S. Blum. Analysis of MIMO capacity with interference. In *Proc. IEEE Int. Conf. Communications*, 2003.
- [15] R. S. Blum, J. H. Winters, and N. R. Sollenberger. On the capacity of cellular systems with MIMO. *IEEE Commun. Lett.*, pages 242–244, June 2002.

- [16] V. M. Bogachev and I. G. Kiselev. Optimum combining of signals in space-diversity reception. *Telecommunication Radio Engineering*, pages 83–85, Oct. 1980.
- [17] D. G. Brennan. Linear diversity combining techniques. *Proc. IRE*, pages 1075–1102, Feb. 1959.
- [18] G. Caire, G. Taricco, and E. Biglieri. Optimum power control over fading channels. *IEEE Trans. Inform. Theory*, pages 1468–1489, July 1999.
- [19] S. Catreux, P. F. Driessen, and L. J. Greenstein. Simulation results for an interference-limited multiple-input multiple-output cellular system. *IEEE Commun. Lett.*, pages 334–336, Nov. 2000.
- [20] S. Catreux, P. F. Driessen, and L. J. Greenstein. Attainable throughput of an interference-limited multiple-input multiple-output (MIMO) cellular system. *IEEE Trans. Commun.*, pages 1307–1311, Aug. 2001.
- [21] C. Chayawan and V. A. Aalo. On the outage probability of optimum combining and maximal ratio combining schemes in an interference-limited Rice fading channel. *IEEE Trans. Commun.*, pages 532–535, Apr. 2002.
- [22] D. Chizhik, J. Ling, P. W. Wolniansky, R. A. Valenzuela, N. Costa, and K. Huber. Multiple-input-multiple-output measurements and modeling in Manhattan. *IEEE J. Select. Areas Commun.*, pages 321–331, Apr. 2003.
- [23] K. Cho and D. Yoon. On the general BER expression of one- and two-dimensional amplitude modulations. *IEEE Trans. Commun.*, pages 1074–1080, July 2002.
- [24] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham. A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels. *IEEE Trans. Commun.*, pages 773–775, Feb./Mar./Apr. 1995.

- [25] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 1990.
- [26] J. Cui and A. Sheikh. Outage probability of cellular radio systems using maximal ratio combining in the presence of multiple interferers. *IEEE Trans. Commun.*, pages 1121–1124, Aug. 1999.
- [27] O. Damen, A. Chkeif, and J. Belfiore. Lattice code decoder for space-time codes. *IEEE Commun. Lett.*, pages 161–163, May 2000.
- [28] R. L. Dobrushin. General formulation of Shannon's main theorem in information theory. *Amer. Math. Soc. Trans.*, 33:323–438, 1963.
- [29] E. J. Dudewicz and S. N. Mishra. *Modern Mathematical Statistics*. John Wiley & Sons, 1988.
- [30] F. R. Farrokhi and G. J. Foschini. Link-optimal space-time processing with multiple transmit and receive antennas. *IEEE Commun. Lett.*, pages 85–87, Mar. 2001.
- [31] F. R. Farrokhi, A. Lozano, G. J. Foschini, and R. A. Valenzuela. Spectral efficiency of FDMA/TDMA wireless systems with transmit and receive antenna arrays. *IEEE Trans. Wireless Commun.*, pages 591–599, Oct. 2002.
- [32] C. Farsakh and J. A. Nossek. Spatial covariance based downlink beamforming in an SDMA mobile radio system. *IEEE Trans. Commun.*, pages 1497–1506, Nov. 1998.
- [33] R. F. H. Fischer and J. B. Huber. A new loading algorithm for discrete multitone transmission. In *Proc. IEEE Global Telecommun. Conf.*, pages 724–728, 1996.

- [34] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, pages 41–59, Autumn 1996.
- [35] G. J. Foschini and M. J. Gans. On the limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Commun.*, 6:315–335, 1998.
- [36] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky. Simplified processing for wireless communication at high spectral efficiency. *IEEE J. Select. Areas Commun.*, pages 1841–1852, Nov. 1999.
- [37] H. E. Gamal and A. Arunachalam. Space-time coding techniques for MIMO block fading channels with co-channel interference. In *Proc. IEEE Int. Conf. Communications*, 2003.
- [38] H. Gao and P. J. Smith. Exact SINR calculations for optimum linear combining in wireless systems. *Probability in the Engineering and Information Sciences*, 12:261–281, 1998.
- [39] H. Gao, P. J. Smith, and M. V. Clark. Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels. *IEEE Trans. Commun.*, pages 666–672, May 1998.
- [40] W. A. Gardner. *Introduction to Random Processes*. McGraw Hill, 1990.
- [41] D. Gerlach and A. Paulraj. Base station transmitting antenna arrays for multipath environments. *Signal Processing*, 54:59–73, 1996.

- [42] D. Gesbert, M. Shafi, D. 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.*, pages 281–302, Apr. 2003.
- [43] J. D. Gibson. *The Mobile Communications Handbook*. CRC/IEEE, 1999.
- [44] G. Ginis and J. M. Cioffi. On the relation between V-BLAST and the GDFE. *IEEE Commun. Lett.*, pages 364–366, Sept. 2001.
- [45] L. C. Godara. Application of antenna arrays to mobile communications, part i: Performance improvement, feasibility, and system considerations. *Proc. IEEE*, pages 1031–1060, July 1997.
- [46] L. C. Godara. Application of antenna arrays to mobile communications, part ii: Beamforming and direction-of-arrival considerations. *Proc. IEEE*, pages 1195–1245, Aug. 1997.
- [47] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky. Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture. *Electron. Lett.*, 35(1):14–15, Jan. 1999.
- [48] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE J. Select. Areas Commun.*, pages 684–702, June 2003.
- [49] A. J. Goldsmith and S. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Trans. Commun.*, pages 1218–1230, Oct. 1997.
- [50] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 1983.

- [51] I. Gradshteyn and I. Ryzhik. *Tables of Integrals, Series, and Products*. Academic, 1994.
- [52] J. Guey, M. P. Fitz, M. R. Bell, and W. Kuo. Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. *IEEE Trans. Commun.*, pages 527–537, Apr. 1999.
- [53] S. W. Halpern. The effect of having unequal branch gains in practical predetection diversity systems for mobile radio. *IEEE Trans. Veh. Technol.*, pages 94–105, Feb. 1977.
- [54] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1952.
- [55] B. Hassibi. A fast square-root implementation for BLAST. In *Proc. Asilomar Conf. on Signals, Systems, and Computers*, 2000.
- [56] S. Haykin. *Digital Communications*. John Wiley & Sons, 1988.
- [57] B. M. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. Commun.*, pages 2041–2052, Dec. 2000.
- [58] B. L. Hughes. Differential space-time modulation. *IEEE Trans. Inform. Theory*, pages 2567–2578, Nov. 2000.
- [59] W. C. Jakes. *Microwave Mobile Communications*. John Wiley & Sons, 1974.
- [60] M. Kang and M. Alouini. Performance analysis of MIMO systems with co-channel interference over Rayleigh fading channels. In *Proc. IEEE Int. Conf. Communications*, 2002.

- [61] M. Kang and M. Alouini. A comparative study on the performance of MIMO MRC systems with and without co-channel interference. In *Proc. IEEE Int. Conf. Communications*, 2003.
- [62] M. Kang and M. Alouini. Impact of correlation on the capacity of MIMO channels. In *Proc. IEEE Int. Conf. Communications*, 2003.
- [63] M. Kang, M. Alouini, and L. Yang. Outage probability and spectrum efficiency of cellular mobile radio systems with smart antennas. *IEEE Trans. Commun.*, pages 1871–1877, Dec. 2002.
- [64] W. Kaplan. *Advanced Calculus*. Addison-Wesley, 1959.
- [65] S. M. Kay. *Fundamental Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998.
- [66] M. Kiessling, I. Viering, M. Reinhardt, and J. Speidel. Short-term and long-term diagonalization of correlated MIMO channels with adaptive modulation. In *Proc. IEEE Int. Symposium on Personal, Indoor and Mobile Radio Commun.*, pages 593–597, 2002.
- [67] J. S. Kwak and J. H. Lee. Performance analysis of optimum combining for dual-antenna diversity with multiple interferers in a Rayleigh fading channel. *IEEE Commun. Lett.*, pages 541–543, Dec. 2002.
- [68] J. Li, B. Halder, P. Stoica, and M. Viberg. Computationally efficient angle estimation for signals with known waveforms. *IEEE Trans. Signal Processing*, pages 2154–2163, Sept. 1995.

- [69] X. Li, H. C. Huang, A. Lozano, and G. J. Foschini. Reduced-complexity detection algorithms for systems using multi-element arrays. In *Proc. IEEE Global Telecommun. Conf.*, pages 1072–1076, 2000.
- [70] J. C. Liberti and T. S. Rappaport. *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Prentice-Hall, 1999.
- [71] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione. Transmit antennae space-time block coding for generalized OFDM in the presence of unknown multipath. *IEEE J. Select. Areas Commun.*, pages 1352–1364, July 2001.
- [72] G. Long, F. Ling, and J. G. Proakis. Fractionally-spaced equalizers based on singular values decomposition. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1988.
- [73] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela. Lifting the limits on high-speed wireless data access using antenna arrays. *IEEE Commun. Mag.*, pages 156–162, Sept. 2001.
- [74] A. Lozano and A. M. Tulino. Capacity of multiple-transmit multiple-receive antenna architectures. *IEEE Trans. Inform. Theory*, pages 3117–3128, Dec. 2002.
- [75] H. Lütkepohl. *Handbook of Matrices*. John Wiley & Sons, 1996.
- [76] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and Adaptive Signal Processing*. McGraw Hill, 2000.
- [77] T. L. Marzetta. BLAST training: estimating channel characteristics for high capacity space-time wireless. In *Proc. Allerton Conf. on Communication, Control, and Computing*, pages 22–24, 1999.

- [78] R. A. Monzingo and T. W. Miller. *Introduction to Adaptive Arrays*. John Wiley & Sons, 1980.
- [79] T. K. Moon and W. C. Stirling. *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, 2000.
- [80] R. D. Murch and K. B. Letaief. Antenna systems for broadband wireless access. *IEEE Commun. Mag.*, pages 76–83, Apr. 2002.
- [81] A. F. Naguib. *Adaptive Antennas for CDMA Wireless Networks*. PhD thesis, Dept. of Electrical Engineering, Stanford University, 1995.
- [82] A. F. Naguib and A. Paulraj. Performance of wireless CDMA with M-ary orthogonal modulation and cell site antenna arrays. *IEEE J. Select. Areas Commun.*, pages 1770–1783, Dec. 1996.
- [83] A. F. Naguib, A. Paulraj, and T. Kailath. Capacity improvement with base-station antenna arrays in cellular CDMA. *IEEE Trans. Veh. Technol.*, pages 691–698, Aug. 1994.
- [84] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE J. Select. Areas Commun.*, pages 1423–1436, Oct. 1998.
- [85] L. H. Ozarow, S. Shamai, and A. D. Wyner. Information theoretic considerations for cellular mobile radio. *IEEE Trans. Veh. Technol.*, pages 359–378, May 1994.
- [86] A. Papoulis and S. U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 2002.

- [87] T. D. Pham and K. G. Balmain. Multipath performance of adaptive antennas with multiple interferers and correlated fadings. *IEEE Trans. Veh. Technol.*, pages 342–351, Mar. 1999.
- [88] H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.
- [89] J. Proakis. *Digital Communications*. McGraw Hill, 1995.
- [90] G. G. Raleigh and J. M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Trans. Commun.*, pages 357–366, Mar. 1998.
- [91] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice-Hall, 1996.
- [92] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE J. Select. Areas Commun.*, pages 1437–1449, Oct. 1998.
- [93] S. Rice. Mathematical analysis of random noise. *Bell Syst. Tech. J.*, 23, 1944.
- [94] J. Salz and J. H. Winters. Effect of fading correlation on adaptive arrays in digital mobile radio. *IEEE Trans. Veh. Technol.*, pages 1049–1057, Nov. 1994.
- [95] H. Sampath, P. Stoica, and A. Paulraj. Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion. *IEEE Trans. Commun.*, pages 2198–2206, Dec. 2001.
- [96] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj. A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results. *IEEE Commun. Mag.*, pages 143–149, Sept. 2002.

- [97] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath. Optimal designs for space-time linear precoders and decoders. *IEEE Trans. Signal Processing*, pages 1051–1064, May 2002.
- [98] A. Scaglione, P. Stoica, S. Barbarossa, and H. Sampath. Linear precoders and decoders designs for MIMO frequency selective channels. In *Proceedings IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pages 2273–2276, 2002.
- [99] M. Sellathurai and S. Haykin. Turbo-BLAST for wireless communications: theory and experiments. *IEEE Trans. Signal Processing*, pages 2538–2546, Oct. 2002.
- [100] N. Seshadri, A. F. Naguib, V. Tarokh, and A. R. Calderband. A space-time coding modem for high-data-rate wireless communications. *IEEE J. Select. Areas Commun.*, pages 1459–1478, Oct. 1998.
- [101] A. Shah and A. M. Haimovich. Performance analysis of optimum combining in wireless communications with Rayleigh fading and cochannel interference. *IEEE Trans. Commun.*, pages 473–479, Apr. 1998.
- [102] A. Shah and A. M. Haimovich. Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference. *IEEE Trans. Veh. Technol.*, pages 1454–1463, July 2000.
- [103] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.*, pages 502–513, Mar. 2000.
- [104] D. S. Shiu. *Wireless Communication Using Dual Antenna Arrays*. Kluwer Academic Publishers, 2000.

- [105] Y. Song and S. D. Blostein. MIMO channel capacity in co-channel interference. In *Proc. 21st Biennial Symposium on Commun.*, pages 220–224, 2002.
- [106] K. W. Sowerby and A. G. Williamson. Selection diversity in multiple interferer mobile radio systems. *Electron. Lett.*, 24(24):1511–1513, Nov. 1988.
- [107] G. L. Stüber. *Principle of Mobile Communication*. Kluwer Academic Publishers, 1996.
- [108] B. Suard, A. F. Naguib, G. Xu, and A. Paulraj. Performance of CDMA mobile communication systems using antenna arrays. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1993.
- [109] A. L. Swindlehurst and G. Leus. Blind and semi-blind equalization of generalized space-time block codes. *IEEE Trans. Signal Processing*, pages 2489–2498, Oct. 2002.
- [110] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. Inform. Theory*, pages 1456–1467, July 1999.
- [111] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block coding for wireless communications: performance results. *IEEE J. Select. Areas Commun.*, pages 451–460, Mar. 1999.
- [112] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion in the presence of channel estimation errors, mobility, and multiple paths. *IEEE Trans. Commun.*, pages 199–207, Feb. 1999.

- [113] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Inform. Theory*, pages 744–765, Mar. 1998.
- [114] Í. E. Telatar. Capacity of multi-antenna gaussian channels. *European Trans. Telecommun.*, pages 585–595, Nov./Dec. 1999.
- [115] C. Tellambura and A. Annamalai. Further results on the Beaulieu series. *IEEE Trans. Commun.*, pages 1774–1777, Nov. 2000.
- [116] A. Tucker. *Applied Combinatorics*. John Wiley & Sons, 1984.
- [117] S. Verdú. *Multiuser Detection*. Cambridge University Press, 1998.
- [118] E. Villier. Performance analysis of optimum combining with multiple interferers in flat Rayleigh fading. *IEEE Trans. Commun.*, pages 1503–1510, Oct. 1999.
- [119] W. T. Webb and R. Steele. Variable rate QAM for mobile radio. *IEEE Trans. Commun.*, pages 2223–2230, July 1995.
- [120] J. H. Winters. Optimum combining in digital mobile radio with cochannel interference. *IEEE Trans. Veh. Technol.*, pages 144–155, Aug. 1984.
- [121] J. H. Winters. Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading. *IEEE Trans. Veh. Technol.*, pages 377–384, Nov. 1993.
- [122] J. H. Winters. Smart antennas for wireless systems. *IEEE Personal Commun.*, pages 23–27, Feb. 1998.

- [123] A. Wittneben. A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation. In *Proc. IEEE Int. Conf. Communications*, pages 1630–1634, 1993.
- [124] Q. Yan and R. S. Blum. Rate adaptive space-time modulation techniques for combating cochannel interference. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pages 2469–2472, 2001.
- [125] Y. Yao and A. U. H. Sheikh. Investigations into cochannel interference in microcellular mobile radio systems. *IEEE Trans. Veh. Technol.*, pages 114–123, May 1992.
- [126] D. Yoon, K. Cho, and J. Lee. Bit error probability of M-ary quadrature amplitude modulation. In *Proc. IEEE Vehicular Technology Conf.*, pages 2422–2427, 2000.
- [127] W. Zha and S. D. Blostein. Modified decorrelating decision-feedback detection of BLAST space-time system. In *Proc. IEEE Int. Conf. Communications*, 2002.
- [128] S. Zhou and G. B. Giannakis. Adaptive modulation for multi-antenna transmissions with channel mean feedback. In *Proc. IEEE Int. Conf. Communications*, 2003.

# Vita

## Yi Song

### EDUCATION

- Ph.D. Candidate    Queen's University                      Sept. 1998–Present  
Electrical and Computer Engineering
- M.Sc.(Eng.)        Queen's University                      Jan. 1996–Feb. 1998  
Electrical and Computer Engineering
- B.Eng.              Shanghai Jiao Tong University    Sept. 1991–July 1995  
Electrical and Computer Engineering

### EXPERIENCE

- Research Assistant    Sept. 1998–Present  
Electrical and Computer Engineering, Queen's University
- Co-op Student        Jan. 1998–Aug. 1998  
Advanced Wireless Technology, Nortel Networks
- Research Assistant    Jan. 1996–Dec. 1997  
Electrical and Computer Engineering, Queen's University
- Teaching Assistant    Jan. 1996–Dec. 1997, Jan. 1999–Apr. 2001  
Electrical and Computer Engineering, Queen's University

## **AWARDS**

Ontario Graduate Scholarships, Queen's University, 2001-2002.

Ontario Graduate Scholarships in Science and Technology, Queen's University, 1998-2001.

Queen's Graduate Awards, Queen's University, 1996-2002.

## **PUBLICATIONS**

Yi Song, Steven D. Blostein and Julian Cheng, "Exact Outage Probability for Equal Gain Combining with Cochannel Interference in Rayleigh Fading," *IEEE Transactions on Wireless Communications*, Sept. 2003.

Yi Song, Steven D. Blostein and Julian Cheng, "Outage Probability Comparisons for Diversity Systems with Cochannel Interference in Rayleigh Fading," Submitted to *IEEE Transactions on Wireless Communications*.

Yi Song and Steven D. Blostein, "Channel Estimation and Data Detection for MIMO Systems under Spatially and Temporally Colored Interference," Submitted to *EURASIP Journal of Applied Signal Processing MIMO Communications and Signal Processing Special Issue* (under revision).

Yi Song and Steven D. Blostein, "Data Detection in MIMO Systems with Co-Channel Interference," *IEEE Vehicular Technology Conference*, Sept. 2002, Vancouver, Canada.

Yi Song and Steven D. Blostein, "MIMO Channel Capacity in Co-Channel Interference,"

*Biennial Symposium on Communications*, May 2002, Kingston, Canada.

Yi Song and Steven D. Blostein, "Unknown Target Motion Compensation for spotlight-mode Synthetic Aperture Radar Imagery," *IEEE International Conference on Image Processing*, Sept. 2000, Vancouver, Canada.