Selection diversity techniques in MIMO and cooperative relaying

Shiva Prakash Premananda

2011

Shiva Prakash Premananda. (2011). Selection diversity techniques in MIMO and cooperative relaying. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/47583>

<https://doi.org/10.32657/10356/47583>

Selection Diversity Techniques in MIMO and Cooperative Relaying

Shiva Prakash Premananda

School of Computer Science and Engineering

Nanyang Technological University

Singapore

A thesis submitted to Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2011

Acknowledgement

Firstly, my deepest gratitude to my supervisor Prof. Ian McLoughlin for being the best guide and advisor I could have ever wished for. My sincere thanks to him for admitting me into the Ph.D program and his guidance during my research studies here. I have benefited from his extensive knowledge in different domains, which at the same time gave me the confidence to explore and choose a research topic that interested me. His constant support, enthusiasm in research, kindness and humour has always been a source of positive energy and motivation, something which is very much required given the long process of Ph.D studies. I am also thankful to him for the careful revisions of my manuscripts and this thesis.

I also appreciate the financial assistance offered by Nanyang Technological University, which has supported and made my stay at Singapore a comfortable one. I am also grateful for the other excellent research facilities the institution has provided. I thank all my friends for their company and support which has made my stay at Singapore a memorable one.

Finally, I would like to thank the people closest to me, my parents and sister for their love, understanding and constant encouragement which helped make the completion of this thesis work possible.

Contents

1	Introduction	1
1.1	Overview of wireless standards	2
1.2	Thesis motivation	8
1.3	Thesis organisation	10
2	Overview of MIMO systems	14
2.1	The Multipath Fading Channel	15
2.1.1	MIMO small-scale fading model	16
2.1.2	Fading distributions	20
2.1.3	Spatial Correlation	21
2.2	Gains achievable in MIMO channels	22
2.2.1	Performance Metrics	23
2.2.2	Diversity Gain	25
2.2.3	Array Gain	30
2.2.4	Spatial multiplexing gain (SM)	31
2.3	Adaptive MIMO schemes	33
2.3.1	Adaptability with respect to CSI	35
2.3.2	Adaptive modulation and coding	37
2.4	Summary	38
3	Antenna Selection in MIMO	39
3.1	Introduction	39
3.1.1	Related work	44
3.1.2	Contributions	45
3.1.3	Summary of Results	46
3.2	System model and assumptions	47
3.2.1	TAS with feedback delay (TASD)	49
3.2.2	Wiener channel prediction	50
3.2.3	Prediction of channel power	52

3.2.4	Fading distribution of a TAS/MRC with prediction (TASP)	53
3.3	Error Rate Analysis	56
3.4	Outage Probability	62
3.5	Fading statistics of the output SNR	63
3.6	Investigation of system parameters	70
3.6.1	Influence of system parameters on channel correlation	70
3.6.2	Operating point of TASP and trade-off point with TASP	72
3.6.3	Computational complexity	75
3.7	Feedback rate reduction	75
3.7.1	Reduced feedback in an uncorrelated channel	76
3.7.2	Reduced feedback in a correlated channel	78
3.8	Summary	81
4	Analysis of Adaptive Modulation with Selective Diversity under Channel Prediction	83
4.1	Introduction	83
4.2	Overview of adaptive modulation in MIMO	84
4.2.1	Related work	85
4.2.2	Contributions	86
4.2.3	Summary of results	87
4.3	System Model and Assumptions	87
4.4	Maximum Achievable Spectral Efficiency	89
4.5	System Analysis	92
4.6	System Performance	96
4.6.1	Instantaneous BER and Constant Power	97
4.6.2	Average BER and Constant Power	98
4.6.3	Efficient threshold setting for A-BER, CP scheme.	100
4.7	Performance Analysis	101
4.8	Summary	108
5	Cooperative Relaying	111
5.1	Overview	111

5.2	Cooperative signalling strategies	114
5.2.1	Amplify and Forward	115
5.2.2	Decode and Forward	118
5.2.3	Coded Cooperation	118
5.2.4	Relay Configurations	119
5.2.5	Multi-antenna relaying	123
5.3	Summary	124
6	Multi-Antenna relay selection	125
6.0.1	Related work	125
6.0.2	Contributions	127
6.1	Dual-Hop Fixed Gain Multi-Antenna AF Relay Selection	128
6.1.1	System model	129
6.1.2	Outage Probability	131
6.1.3	Error probability	133
6.1.4	Choice of gain parameter	135
6.1.5	Statistics of the end-to-end SNR	135
6.1.6	Numerical and simulation results	136
6.1.7	Summary	139
6.2	Dual-Hop Variable Gain AF Multi-antenna Relay Selection in Nakagami- m Channels	139
6.2.1	System model and assumptions	140
6.2.2	Outage Probability	142
6.2.3	Error probability	143
6.2.4	Numerical and simulation results	144
6.2.5	Summary	145
6.3	Partial Multi-antenna Relay selection taking direct path into account	147
6.3.1	System model	147
6.3.2	Performance metrics	148
6.3.3	Channel capacity	150
6.3.4	Diversity order analysis	151
6.3.5	Numerical and simulation results	152

6.3.6	Summary	155
6.4	Predictive Dual-Hop Partial Multi-antenna Relay Selection	155
6.4.1	System model and assumptions	156
6.4.2	Performance metrics	158
6.4.3	Numerical and simulation results	160
6.4.4	Summary	164
6.5	Chapter Summary	165
7	Conclusion and Future work	167
7.1	Conclusion	167
7.2	Future Work	172
	Appendix A	174
A	List of integrals and formulas	174
A.1	Multinomial coefficients	174
A.2	Laguerre polynomial	174
A.3	Definite integration involving Bessel functions	174
A.4	Integrals involving exponential function	175
A.5	Covariance between Gaussian random variables	175
A.6	Derivation steps for Eqn. 6.14	176
A.7	Derivation steps for Eqn. 6.23	176
	Appendix B	176
B	Design and Implementation of a MIMO Hardware Test-bed	177
B.1	MIMO test-bed design issues	178
B.2	STAR (Space-time array research) Platform	179
B.3	System Architecture	180
B.4	Sub-units of the STAR Platform Transceiver	183
B.4.1	Digital Board	183
B.4.2	Mixed Signal Board	185
B.4.3	The RF boards	186
B.4.4	Synth Board	187

B.5	FPGA Design	187
B.5.1	ARM-FPGA interface	188
B.6	Implementation of the 4x4 MIMO transceiver	190
B.6.1	Loopback data flow and control:	193
B.7	MIMO test bed design considerations	194
B.7.1	Clocks and synchronization	195
B.7.2	Processing requirements	195
B.7.3	Memory Configuration and Processor Host Control	196
B.7.4	Digital IF (intermediate frequency)Architecture	197
B.7.5	Partitioning and Control plane	197
B.7.6	Existing Test-beds	198
B.8	Summary	199

List of Figures

1.1	Different wireless standards and approximate year of release into market. .	7
1.2	Transmit antenna selection in delay and rate constrained feedback paths. Based on feedback delay, a predictor is used to pick best set of antennas and also determine best rates that can be supported at any transmission instant.	10
1.3	Relay selection in dual-hop cooperative relay networks. The best set of relays forward data to the destination where their signal is combined before demodulation. When feedback delay exists a predictor can be employed at the source to improve selection.	12
2.1	Multipath propagation due to obstacles via. reflection, diffraction and scattering in a wireless MIMO channel.	17
2.2	Doppler fading for 20 Hz (in blue) and 100 Hz (in red) in Rayleigh channels with $T_s=10\mu s$	19
3.1	Block diagram showing TAS based upon information from a feedback path from MRC receive antennas and a predictor. To combat degradation due to delay, the predicted antenna index i for the future slot is sent D blocks in advance.	47
3.2	Transmission frame, showing pilot symbols at the start of each transmission block, and delay between pilot transmission and the antenna switching point.	49
3.3	BER performance comparison of BPSK for TAS/MRC (2, 1; 2) with and without prediction at different delays, showing theoretical and simulation results.	59
3.4	Feedback delay tolerance of BER in BPSK of the (2, 1; 2) scheme for different SNRs of 6, 10, 14, 18, 22 dB with and without prediction, showing the improved tolerance to feedback delay effects gained by the prediction scheme.	59
3.5	Feedback delay tolerance of BER in BPSK at $\bar{\gamma} = 10dB$ for different TAS/MRC configurations with and without prediction. (a) Influence of increasing N_r and N_t . (b) Different system arrangements with constrained total number of antennas of $N_r + N_t = 6$	61

3.6	Outage probability comparisons at $\bar{\gamma} = 10$ dB SNR between TASP/D schemes with fixed N_r and increasing N_t and vice-versa, at $R = 1 \text{ bit/s/Hz}$, showing theoretical and simulation results.	64
3.7	Comparison of normalised average SNR gain $\bar{\gamma}_{TAS/\bar{\gamma}}$ as a function of number of receive and transmit antennas at specific feedback delays of $f_d\tau = 0.25$ with and without prediction. Theoretical results are verified by simulations.	66
3.8	Normalised average SNR gains $\bar{\gamma}_{TAS/\bar{\gamma}}$ of different arrangements with and without prediction for a wide range of delays.	67
3.9	(a) Influence of filter tap spacing and E_p/N_0 at normalised feedback delays of 0.1, 0.2 and 0.4 on power correlation with a five tap filter (b) Influence of predictor length and E_p/N_0 on power correlation at fixed delays of 0.02 and 0.1 and $L_b = 100$	69
3.10	Correlation coefficient comparison for TASP, and TASP against feedback delay for E_p/N_0 of 15 dB and 25 dB with two different filter lengths of 5 and 10.	73
3.11	Equal correlation plot for TASP and TASP, where at all points $\rho_p = \rho_d$ with the corresponding E_p/N_0 (for TASP case) against delay for various L . The dashed line corresponds to the correlation value at any delay, read off from the RH axis.	73
3.12	Effect of increasing E_p/N_0 for TASP at a filter length of $L = 5$ on BER performance at $\bar{\gamma} = 10$ dB for two different configurations. The solid black curves represent TASP at different $E_p/N_0 = 15, 20, 25, 30$ dB from top to bottom, in two example systems.	74
3.13	Frame structure with Pilot Block and Data and feedback link	76
3.14	Influence of feedback rate reduction on TASP and TASP scheme, with feedback delay and reduced feedback rate. (a) 10 block delay ($\mathcal{D} = 10$) and 5 times reduction ($m = 5$) in feedback rate. (b) 20 block delay ($\mathcal{D} = 20$) and 5 times reduction in feedback rate.	77

3.15	Influence of feedback rate reduction on TASP and TASP scheme, with feedback delay and reduced feedback rate. (a) 10 block delay ($\mathcal{D} = 10$) and 20 times $m = 20$ reduction in feedback rate. (b) 20 block delay ($\mathcal{D} = 20$) and 20 times reduction in feedback rate.	80
3.16	BER Performance of TASP and TASP under reduced feedback rate with zero feedback delay and 20 times reduction in feedback rate.	81
4.1	A $(N_t, 1; N_r)$ transceiver block diagram employing both adaptive modulation and transmit antenna selection with MRC.	85
4.2	Maximum achievable spectral efficiency for different TAS configurations $(N_t, 1; N_r)$ with perfect prediction or zero delay case.	91
4.3	Maximum achievable spectral efficiency for different TAS configurations $(N_t, 1; N_r)$ with channel power correlation coefficient $\rho_p = 0.7$	91
4.4	$BER(\hat{\gamma})$ against instantaneous SNR for $\rho_p = 0.9, 0.8$ in an $(N_t, 1; 2)$ setup at $\bar{\gamma} = 20dB$	94
4.5	Plot showing improvement in spectral efficiency for A-BER scheme using threshold setting from our the proposed method compared with an existing method [45] for $\rho_p = 0.9, 0.8$ for modulation sets $K=5$ (0,4,16,64,256) and $BER_0 = 10^{-3}$ in a (2,1;2) system.	100
4.6	Average BER for I-BER at target BERs of 10^{-4} with $\rho_p = 0.999$ and $\rho_p = 0.9$ in a (2,1;2) system, shown as thick lines. Also shown are the non-adaptive average BER, BER_M for TASP/MRC M-QAM for the two values of ρ_p	101
4.7	Average BER for I-BER and A-BER schemes at target BERs of 10^{-3} and 10^{-5} with $\rho_p = 0.999$ in a (2,1;2) system, shown as thick lines. Grey lines correspond to average BER of non-adaptive M-QAM with (2,1;2) TAS with equal ρ_p	102
4.8	Spectral efficiency for I-BER and A-BER schemes for target BER of 10^{-3} and 10^{-5} with $\rho_p = 0.999$ (left) and 0.75 (right) in a (2;1,2) system; also plotted is the maximum achievable Shannon spectral efficiency for perfect prediction.	104

4.9	Effect of N_t and ρ_p on $ABER_{se}$ with $N_r = 2$ for a target BER of 10^{-4} ; also plotted for these setups is the maximum achievable Shannon spectral efficiency for perfect prediction.	105
4.10	Effect of N_t and ρ_p on $IBER_{se}$ with $N_r = 2$ for a target BER of 10^{-4} ; also plotted for these setups is the maximum achievable Shannon spectral efficiency for perfect prediction.	106
4.11	Mode selection probabilities of I-BER for $(4, 1; 2)$ and $(1, 1; 2)$ at $\rho_p = 0.99$, (corresponding to a channel MMSE of -20 dB) with $BER_0 = 10^{-4}$	107
4.12	Percentage I-BER spectral efficiency gain of $(4, 1; 2)$, $(2, 1; 2)$ and $(1, 1; 2)$ setups over the $(1, 1; 2)$ non delayed or perfectly predicted scheme at different ρ_p , corresponding to a channel MMSE of -30,-25,-20,-15,-12,-10,-8,-7,-5 dB (top to bottom) with $BER_0 = 10^{-4}$	109
5.1	Cooperative relay network	112
5.2	Cooperative relaying topologies showing two signal reinforcement schemes above, and three selection-based schemes below. Yellow circles indicate time slots.	115
5.3	Basic half-duplex three node relaying scheme employing either AF or DF	116
6.1	Dual-Hop relay selection system equipped with multiple receive antennas at source and destination.	129
6.2	Outage Probability for different relay configurations in the dual-hop multi-antenna scheme, with $\gamma_T = 5$, $\bar{\gamma}_2 = 5\bar{\gamma}_1$	137
6.3	BPSK system BER performance for different antenna configurations and relay gains.	138
6.4	A distant source communicates to a destination via a multi-antenna relay selective system. A weak direct path between $S \rightarrow D$ is also shown. . . .	140
6.5	Outage probability w.r.t $\bar{\gamma}_1$, for $\gamma_T = 5$ dB for selective relaying $L = 2$, $N_R = 2$, $K = 2$. The effect of power imbalance between hops for increasing $m_2 = 1, 2, 5, 8$ with $m_1 = 1$ is seen.	146

6.6	BER plot for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 = 2\bar{\gamma}_2$. A performance improvement is noted for increasing values of $m_2 = 1, 2, 3, 4, 5$ is noted. Sensitivity to changes in m_2 w.r.t number of transmit antennas can also be seen.	146
6.7	Capacity for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 =$ $1.25\bar{\gamma}_2$, with and without direct path.	152
6.8	Capacity for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 =$ $1.25\bar{\gamma}_2$, with and without direct path.	154
6.9	BER plot for BPSK under predictive selective relaying, for first hop temporal correlations of $\rho_p = 1, 0.9, 0.7$, comparing single antenna and multi-antenna relay switching.	161
6.10	BER degradation due to feedback delay with and without prediction at $\bar{\gamma}_1 =$ $8dB$ (top) and $\bar{\gamma}_1 = 15dB$ (bottom) for different relay schemes with $\bar{\gamma}_2 = 5\bar{\gamma}_1$. 162	
B.1	The Space time array research test-bed	179
B.2	Simplified Platform Transceiver Hardware	181
B.3	STAR Platform Architecture [182]	182
B.4	Mixed Signal Chip Architecture [184]	186
B.5	The FPGA Interface	188
B.6	A few important FPGA VHDL code blocks	189
B.7	Implementation and testing of a 4×4 MIMO transceiver	191
B.8	(a) ADC glue logic (b) Read loopback module (c) Transmitter Block . . .	192
B.9	Loopback test setup	193
B.10	Square and sine loop-backed test samples captured at 60 Msamples/s. . . .	194

List of Tables

3.1	Simulation setup for TAS/MRC with prediction	58
6.1	Comparison of existing relaying methods against proposed schemes in 6.1, 6.2, 6.3 and 6.4.	166

List of Abbreviations

AWGN	additive white Gaussian noise
MIMO	multiple-input multiple-output
SISO	single-input single-output
OSTBC	orthogonal space-time block coding
TAS	transmit antenna selection
AF	amplify and forward
DF	decode and forward
BER	bit error rate
SER	symbol error rate
LAN	local area network
BPSK	binary phase shift keying
QPSK	quadrature phase shift keying
CNA	channel noise assisted
E2E	end-to-end
BS	base station
RS	relay station
CSI	channel state information
OFDM	orthogonal frequency division multiplexing
i.i.d	independent and identically distributed
r.v	random variable
PDF	probability density function
CDF	cumulative distribution function
ADC	analogue to digital converter

DAC	digital to analogue converter
3GPP	third generation partnership project
bps	bits per second
bps/Hz	bits per second per Hertz
Mbps	megabits per second
QAM	quadrature amplitude modulation
LOS	line of sight
AMC	adaptive modulation and coding
ARQ	automatic repeat request
MRC	maximum ratio combining
MRT	maximal ratio transmission
SNR	signal to noise ratio
CDMA	code division multiple access
FDD	frequency division duplexing
FDMA	frequency division multiple access
FPGA	field programmable gate array
HiperLAN	high performance radio LAN
MMSE	minimum mean square error
MRC	maximal ratio combining
STBC	space-time block code
W-CDMA	wideband CDMA
WiMAX	worldwide inter-operability for microwave access
IEEE	Institute of electrical and electronics engineers

List of Symbols

$\mathbf{I}_n$	Identity Matrix of dimension n-by-n
$\mathbf{a}$	vector
$\mathbf{A}$	matrix
$\ \mathbf{A}\ _F$	Forbrenius norm of $\mathbf{A}$
$(\cdot)^T$	transpose of a vector
$(\mathbf{A})^H$	Hermetian transpose of $\mathbf{A}$
$\mathbb{E}[X]$	Expectation value or mean of a r.v X
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian r.v with mean μ , variance σ^2
$Q(x)$	Gaussian Q Function
$ x $	magnitude of complex number x
$f_X(\cdot)$	PDF of r.v X
$F_X(\cdot)$	CDF of r.v X
$\overline{X}$	Mean of r.v X
C_{ij}	Covariance between two r.vs X_i, X_j
$\max(x_1, x_2)$	Maximum of x_1, x_2
$\lceil x \rceil$	ceiling operator, smallest integer larger than x
$\approx$	approximately equal to
$\ln(x)$	(natural) logarithm of x to the base-e
$\log_a(x)$	logarithm of x to the base-a
$Pr(A)$	Probability of event A
$\frac{\partial f}{\partial x_i}$	partial derivative of f with respect to x_i , where f is a function on $(x_1, \dots, x_n)$
$x!$	factorial of x

Abstract

The primary objective of this thesis revolves around investigating MIMO (multiple-input multiple output) transmit antenna selection (TAS) in delay constrained networks. In this thesis, for various different scenarios of TAS arrangement, channel prediction is explored as a tool to recover some of the system gain lost due to delayed switching. In use, TAS requires at least partial channel knowledge at the transmitter in order to perform selection. In this work, performance degradation due to outdated channel knowledge is determined analytically, and related to channel characteristics. Performance gains of several different antenna placements are presented and trade-offs are analysed with respect to use of prediction. Furthermore, adaptive modulation is also explored in a TAS scenario (allied to maximal ratio combining), and again channel prediction is employed at the receiver to provide estimates of future best transmission states, which include selecting the best antenna as well as the best supported modulation scheme, optimised for the given objectives.

This thesis then extends these ideas into the interesting area of cooperative relaying. The principles of selection diversity are applied as we investigate partial relay selection in amplify-and-forward relaying for a variety of system configurations, and the effects of the multiple antennas are analysed when used either at the source, relay or destination terminal. Since selection is usually delay limited, we also consider the benefit of incorporating a predictive relay selection scheme to improve performance.

To validate and test advanced MIMO algorithms over real-world channels, an experimental architecture was designed for a reconfigurable multi-antenna test-bed. This 4×4 MIMO FPGA (field programmable gate array) base-band architecture is presented, and verified by MIMO loop-back testing. The setup is useful in validating future research work relating to MIMO and antenna selection, and also in developing efficient reconfigurable architectures that are demonstrable over physical channels.

Chapter 1

Introduction

“This is just an experiment that proves Maestro Maxwell was right - we just have these mysterious electromagnetic waves that we cannot see with the naked eye. But they are there.”

Heinrich Hertz

WIRELESS RADIO communications has its origins in the 1880s when Heinrich Hertz’s set of brilliant experiments led to the discovery of radio waves. The discovery essentially verified Maxwell’s theory of electromagnetic radiation, and marked the dawn of the wireless era. Wireless technology has made steady progress ever since Marconi’s demonstration of transatlantic radiowave communications around the turn of the 20th century. However, the past decade has seen explosive growth in the deployment of wireless communication technologies with recent advancements in cellular communications.

MIMO (multiple-input multiple-output) [1] and cooperative relaying [2] have been two significant technological breakthroughs in the present day wireless communication technology. MIMO can help to overcome the deleterious effects of fading, through the use of multiple antenna elements at both ends of the communication system. By obtaining independently faded copies of the transmitted signal, diversity benefit or improved error performance may be achieved through effective signal combining [3]. Similarly, system capacity could be greatly increased without further increase in bandwidth or transmit power, by transmitting different streams of data from different antennae and through careful signal recovery at the receiver [1], [4]. With several different degrees of freedom, a MIMO system can thus be configured in many ways, depending on the resources available and the end objective. It is seen as potential

solution for achieving high data rate demands of next generation wireless networks such as 4G (fourth generation) by offering improvements in areas such as spectral efficiency and link reliability without increasing spectrum usage or power consumption [1]. Cooperative relaying on the other hand, is similar to MIMO and mimics it in many ways. By exploiting the broadcast nature of the wireless medium, several independent collaborating nodes establish a virtual MIMO channel in forwarding a signal to the destination, which then combines the relayed transmissions and the direct transmission from the source [2], [5], [6]. This leads to increased diversity against fading at the destination. Along with diversity benefits, relaying offers advantages of increased coverage capacity, and reduced interference due to smaller transmitted power with all other factors being equal. This chapter begins with an overview of the different wireless standards and technologies deployed so far where the highlight is on increasing use of MIMO and cooperative relaying into modern day wireless communication standards in improving link throughput speeds and reliability; serving as an inspiration to conduct more research studies in these advanced technologies. Next, a section on motivation for the work in this thesis is presented, identifying key MIMO and relaying techniques for improved performance. Next, the organisation of the thesis is presented with the outline of each chapters contribution.

1.1 Overview of wireless standards

Since the inception of the cellular radio networks back in the 1980's, the wireless communication industry has witnessed an explosive and unprecedented growth. Mobile cellular subscriptions have nearly increased eight-fold from an estimated 500 million subscriptions in the year 2000 to an estimated 4.6 billion mobile cellular subscriptions at the end of 2009 [7]. Many present day cellular networks conform to the digitally encoded 2G (second generation) wireless standards such as GSM (Global System for Mobile Communications) and IS-95 (Interim Standard 95), having replaced the 1G (first generation) analogue telecommunication standards, which were introduced back in the 1980's. The digital 2G standard [8] inherently has benefits

over its analogue predecessor, through increased spectral efficiency brought about by cellular frequency reuse concept, ease of multiplexing, better power control, better compression, through the use of channel coders and enhanced privacy through encryption. While the more popular GSM standard, which accounts for over 80% of the global mobile market, uses TDMA (time division multiplex access) multiple access technique to support eight users per carrier over a 200 kHz channel, the competitor IS-95 (Interim Standard 95) uses CDMA (code division multiple access) over the entire radio band of 1.25 MHz, to increase capacity and supports 64 users which are orthogonally coded for simultaneous transmissions. However 2G mainly catered to providing voice services rather than data. The standard bandwidth for data services in GSM networks is 9.6/14.4 kbps per time slot. To cater to increased demand for data services, the evolution to 2.5G occurred through various upgrade technologies such as the High Speed Circuit Switched Data for 2.5G GSM (where multiple time slots were allowed per user), GPRS (general packet radio service) for packet switched networks optimised for conveying IP traffic, EDGE for enhanced data rates for GSM evolution, which used more efficient modulation techniques such as 8-PSK (phase shift keying) and adaptive air interface modulation schemes. EDGE paved the way for high speed data access, and accommodates a maximum of 473 kbps in eight time slots: about four times the traffic that GPRS can handle. On the other hand IS-95 evolved into the IS-95B to support increased data rates of up to 115.2 kbps, and had supported hard hand-off techniques to maintain improved link quality.

Rapid demand for increased data services and increased penetration of the mobile market, paved the way for 3G mobile wireless standards. The IMT-2000 (International Mobile Telecommunications) was responsible for standardising the new 3G W-CDMA (wide band CDMA system) or the 3GPP's (third generation partnership project) UMTS (Universal Mobile Telecommunications System), capable of delivering up to 384 kbps in outdoor environments and up to 2 Mbps in fixed indoor environments, operating in a bandwidth of 5 MHz, allowing simultaneous audio, high quality data and multimedia streaming. UMTS utilizes DS-SS (direct-sequence spread

spectrum) for channel access, and FDD (frequency-division duplexing) to achieve higher speeds and support more users compared to the TDMA schemes. The other significant 3G set of standards (by the working group of 3GPP2), is CDMA2000, which is essentially an upgrade to the earlier 2G CDMA standard IS-95 through the use of multi-carrier techniques. CDMA2000 1xEV comes in two flavours of CDMA carriers: data only (CDMA2000 1xEV-DO) or with both data and voice (CDMA2000 1xEV-DV). The CDMA2000 1xEV-DO option dedicates the radio channel strictly to data users, with greater than 2.4 Mbps of instantaneous high speed packet throughput per user while CDMA2000 1xEV-DV supports both voice and data users, and can offer data rates up to 144 kbps with about twice as many voice channels as IS-95B. Similar to UMTS, the High Speed Packet Access (HSPA) upgrade is defined for both uplink and downlink in the 3GPP Release 5 and 6 respectively, supporting peak downlink data rates of 7.2-14.6 Mbps by using adaptive modulation and coding techniques. Particularly for the HSDPA (down link) the use of receive diversity with at least two receive chains is recommended at the user terminal. Besides enhanced throughput, HSPA also significantly reduces latency. Similarly the evolved HSPA known as HSPA Evolution or HSPA+ is an upcoming wireless broadband standard defined in 3GPP Release 7 and 8 of the W-CDMA specification [9]. Evolved HSPA provides data rates up to 42 and 11 Mbps in the downlink and uplink directions respectively, per 5 MHz carrier, using MIMO technologies and higher order modulation schemes. The employment of MIMO at both UE (user equipment) and BS (base station) improves system capacity and spectral efficiency by increasing the data throughput in the downlink within the existing carrier. Transmission with MIMO is supported with configurations across the downlink with two or four transmit antennas and two or four receive antennas, which allow for multi-layer transmissions with up to four streams. HSPA+ with 64 QAM (quadrature amplitude modulation) is capable of providing peak theoretical downlink throughput rates of 21 Mbps. HSPA+ with 64 QAM and advanced antenna techniques such as 2×2 MIMO can deliver 42 Mbps theoretical capability and 11.5 Mbps on the uplink.

UMTS is gradually evolving into the Long Term Evolution (LTE) system, presently being drafted by the 3GPP group in its Release 8 and 9 specifications [9], [10]. The LTE system aims to support peak data rates of more than 100 Mbps, with increased spectral efficiency and small round trip latency (less than 10ms). LTE uses OFDMA (orthogonal frequency division multiple access) technology with different bandwidths from 1.4 MHz to 20 MHz, in discrete increments, thus very high data rates are possible along with the use of higher order modulations in combination with multi-stream MIMO transmission. OFDMA offers huge advantages in the presence of multipath propagation (over previous schemes like W-CDMA), where orthogonality between channels reduces interference, and equalisation in the frequency domain reduces complexity of receivers, along with efficient frequency domain scheduling. These are the primary reasons why LTE technology has clear advantages over WCDMA-HSPA in terms of spectral efficiency [11]. LTE adopts various MIMO technologies including transmit diversity, single user (SU)-MIMO, multi-user (MU)-MIMO, closed-loop precoding, and dedicated beamforming. Transmit diversity selection, as a form of precoding, is applied to two or four transmit antennas in the downlink, and to two transmit antennas in the uplink. Keeping bandwidth requirement constant, the only way to further increase performance is by employing MIMO techniques, thus MIMO has become an integral part of LTE. LTE-Advanced technology is based on OFDMA combined with advanced MIMO techniques [9]. Enhanced multi-antenna operation for supporting downlink SU MIMO with up to eight layer spatial multiplexing is needed to fulfil targets for LTE-Advanced peak rate spectral efficiency of 30 bps/Hz. The multi-antenna operation for supporting uplink SU-MIMO, with up to four layer spatial multiplexing is needed to fulfil requirements on the peak rate spectral efficiency of 15 bps/Hz. Both open loop and closed loop MIMO techniques are specified for different mobility conditions. Coordinated multi-point transmission (CoMP) technology is also recommended to improve the cell edge throughput/cell capacity.

On the other hand IEEE 802.16/WiMAX has emerged from a different business direction as a key advanced broadband access technology for wireless metropolitan

area networks (WMAN). The 802.16 base standards have undergone changes to incorporate multi-antenna technologies at the BS and mobile stations (MS), and OFDMA, to deliver the capacity and QoS (quality of service) demanded by high-speed digital services. Two different antenna profiles are recommended which include the Alamouti transmit diversity 2×1 technique for diversity improvement and the 2×2 spatial multiplexing scheme for increasing rate. Closed-loop techniques such as antenna selection and beamforming are also specified. Similarly, the popular IEEE 802.11g WLAN (wireless local area network) standard, employing direct-sequence spread spectrum (DSSS) and orthogonal frequency-division multiplexing (OFDM) signaling methods, has undergone amendments for upgrade to IEEE 802.11n, which provides for MIMO technology. Key MIMO techniques are the use of spatial division multiplexing, space-time block coding (STBC) and transmit beamforming. These pave the way for significant increases in the maximum raw data rate from the original 54 Mbps to 600 Mbps with the use of four spatial streams, operating 4×4 at a channel width of 40 MHz. Both 802.11n and WiMAX provide for broadband wireless access, however the key differences are in mobility and throughput. While WiMAX is a WMAN service with an approximate coverage area of 2 miles for mobile WiMAX and 5 miles for fixed WiMAX, 802.11 is a LAN technology for short range communication up to 200 feet. WiMAX offers speeds ranging from 1Mbps to 50 Mbps depending on the mobility, while 802.11n is capable of higher average throughput in excess of 200 Mbps, although it might operate at slightly lesser speeds, due to protocol overhead considerations.

In PAN (personal area networks) UWB technology offers very high data rate service by occupying a very large bandwidth (for example IEEE 802.15.3a UWB operates in a 500 MHz band), while operating at a very low transmit power. Thus, very high data rates are possible due to two reasons: a) large signal bandwidth and b) the channel being frequency selective, exploitation of multipath components using a rake receiver is possible. Due to power constraints imposed on UWB technology, the service has been restricted to short distance communication, and is therefore most applicable to PAN. MIMO technology can be applied in conjunction with UWB

Figure 1.1: Different wireless standards and approximate year of release into market.

to increase signal strength for increasing communication range or increasing data rates [12]. Other PAN technologies include Bluetooth which differs from UWB by occupying a low bandwidth of 2 MHz with a relatively larger transmit power, using a frequency hopping technique to avoid interference. The ZigBee 802.15.4 standard was developed for very low data rate applications and complexity, keeping cost and power consumption to a minimum. This operates in the ISM band and employs DSSS coding with BPSK modulation, providing maximum data rates of 250 kbps while transmitting in a 2 MHz band. Fig. 1.1 summarises the global wireless standards.

Despite the tremendous potential and advantages of MIMO technologies, the use of a large number of antennas is not always possible in many situations where form factor and power constraints are of primary concern. This is especially true in cellular mobile devices, wireless sensor and ad-hoc networks where the devices can physically house at most two transmit/receive chains.

Cooperative communications is a novel approach to harness the inherent spatial diversity, in places where the use of transmit/receive chains has to be kept low. The concept of cooperative communications is a paradigm shift from the conventional point-to-point communications network model. Cooperative communication primarily uses relaying as a key concept where the broadcast nature of the wireless channel is exploited to realize spatial diversity advantages. Cooperative communications techniques take advantage of the broadcast nature of wireless transmission, effectively creating a virtual antenna array through cooperating nodes. Relay communication technology promises significant performance gains in terms of link reliability, spectral efficiency, system capacity, and transmission range. As a result relaying and cooperation have emerged as hot research topics in wireless communication in both industry and academic circles over the past few years. The IEEE 802.16 working group, has formed a task group devoted to incorporating relay capabilities into the foundation of mobile WiMAX/IEEE 802.16e-2005. This task group is currently in the process of finishing IEEE 802.16j which is the multi-hop relay specification. This addendum to 802.16 will be fully compatible with 802.16e-2005 mobile and subscriber stations, but a base station specific to 802.16j will be required for relays to operate.

1.2 Thesis motivation

This thesis revolves around selection diversity techniques used as a low cost means to harness spatial diversity in MIMO and cooperative relaying deployments. While MIMO technology is mature and its benefits are potentially immense, the adoption of MIMO in modern hand-held wireless gadgets has been slow, mainly due to hardware and power costs of additional antennas being used, coupled with the difficulty in testing due to the lack of research test-beds. Transmit antenna selection (TAS) is a low cost mechanism used to exploit spatial diversity. In TAS, a low-rate feedback channel conveys an index identifying the best set of transmit antennas. It is well suited to downlink communications since selection diversity can be practically exploited with a large number of antenna elements and fewer RF chains at the BS in combination

with signal combining techniques with fewer RF chains at the UE equipment. Feedback information is vital in maintaining the benefits of transmit antenna selection, as outdated or corrupt channel information naturally degrades system performance. As a means of alleviating degradation, channel prediction is proposed, and is thoroughly studied through analytical characterisations for non-adaptive schemes. Link adaptation is an effective way to increase throughput in wireless systems, however the performance analysis in adaptive MIMO systems where feedback non-idealities exist has been less researched. To realise these potential benefits, the thesis also investigates the throughput gains achieved by combining TAS techniques (in outdated channels) with adaptive modulation.

On the other hand, cooperative relaying techniques tend to emulate MIMO systems in improving coverage and capacity by forming virtual antenna arrays. Cooperative diversity has gained a lot of research interest in recent times due to its innate ability to harness spatial diversity in a wireless medium, thereby helping to mitigate against the deleterious effect of fading. Relaying communications has many different degrees of freedom, depending on the availability of resources such as cooperating nodes, power and total number of communicating antennas, total number of hops, etc.. When many nodes cooperate simultaneously in relaying data and improving end-to-end communications, there is a reduction in bandwidth efficiency, along with the overhead of ensuring synchronisation. Relay selection is recognised as a bandwidth efficient technique in preserving spatial diversity, easing synchronisation requirements and reducing excessive use of transmit power. This thesis analyses different relay selection arrangements, where the principles of selection diversity are applied to the case of partial relay selection, to obtain improvement in spectral efficiency and diversity gain.

Several closed form analytical expressions such as bit error rate, outage probability, capacity, are obtained in quantifying the gains of the above schemes.

Finally the thesis details a test-bed implementation of reconfigurable MIMO hardware, which is useful in validating the performance gains of any given MIMO algo-

Figure 1.2: Transmit antenna selection in delay and rate constrained feedback paths. Based on feedback delay, a predictor is used to pick best set of antennas and also determine best rates that can be supported at any transmission instant.

rithm over realistic channels and different operating conditions. As stated, test-beds are important in verifying the true gains of MIMO algorithms over real-world wireless channels. A 4x4 MIMO transceiver is presented which is suitable for useful in conducting offline MIMO algorithm testing, and also in conducting MIMO channel measurements and modelling.

1.3 Thesis organisation

The document is organized as follows. **Chapter 2** provides a brief overview of the principles of MIMO wireless communication, different MIMO techniques and popular channel models which are utilized in the thesis. A discussion on adaptive MIMO schemes with focus on antenna selection and adaptive modulation with MIMO to increase data rates in wireless systems are discussed.

Chapter 3 details the TAS/MRC scheme where channel prediction is used to improve performance in delay and rate constrained networks. TAS is known to offer several advantages over open-loop space-time codes both in terms of increased efficiency and decoding complexity, at the cost of a low rate feedback path. Feedback is

critical in implementing TAS, as channel state information (CSI) often gets outdated because of feedback delay, causing performance degradation that may become severe. Moreover, in practical systems, the feedback channel is typically bandwidth limited. Sufficient correlation usually exists in Doppler fading channels and this is used in developing a predictive technique for improved transmit antenna selection. A predictive scheme is then developed to mitigate against delay-induced degradation. Several factors relating to TAS system performance under different channel scenarios both with and without mitigation are explored. Closed form expressions for performance metrics such as bit error rate (BER), average signal to noise ratio (SNR), gain and outage probability are derived and verified by simulations. The impact of prediction is analyzed through these measures for different TAS setups and channel prediction scenarios, as are various system design parameter considerations.

Chapter 4 deals with adaptive modulation techniques related to TAS. Here, the performance of a rate-adaptive modulation system in Rayleigh fading channels exploiting spatial diversity through TAS and maximal ratio combining at the receiver is analysed for the case of delay constrained networks. Channel prediction is employed at the receiver to provide estimates of future best transmission states, including selecting the best transmission antenna as well as the best supported M-QAM (M-Quadrature amplitude modulation) mode. A closed form expression for the Shannon capacity in such a system is derived, serving as a benchmark to evaluate the spectral efficiency (SE) of the discrete rate optimised system under different objectives, operating conditions and number of transmit/receive antennas. Closed form expressions for several performance metrics such as average SE and bit error rate are derived in order to determine optimal switching boundaries for discrete rate-adaptive M-QAM schemes, under channel prediction errors. Fig. 1.2 depicts a general TAS adaptive modulated system in a feedback rate and delay constrained channel, where the best transmission states are predicted at the receiver and sent to the transmitter several blocks ahead.

In **Chapter 5** we provide a brief literature survey and overview of co-operative relaying techniques. Cooperative diversity has gained a lot of research interest in re-

Figure 1.3: Relay selection in dual-hop cooperative relay networks. The best set of relays forward data to the destination where their signal is combined before demodulation. When feedback delay exists a predictor can be employed at the source to improve selection.

cent times due to its innate ability to harness spatial diversity in a wireless medium, thereby helping to mitigate against the deleterious effect of fading. A brief introduction to two popular relaying technologies: amplify and forward, and decode and forward relaying, is provided. Given the different degrees of freedom through which a particular signal can be relayed to the destination via several intermediate nodes, several relaying topologies are identified and presented. When many nodes cooperate simultaneously in relaying data and improving end-to-end communications, there is a reduction in bandwidth efficiency, along with the overhead of ensuring synchronisation. Relay selection is recognised as a bandwidth efficient technique in preserving spatial diversity and easing synchronisation requirements.

In **Chapter 6** we analyse the case of partial relay selection for multi-antenna relay networks. We then extend relay selection to the multi-antenna case. Capitalising on results from previous chapters, we analyse selection relaying for different relay configurations in a multi-antenna scenario—with multiple antennas located either at the source, relay or destination. Specifically, we consider a generalised relay selection

scheme for arbitrary number of relays, relay receivers, transmit antennas, different severities of fading for the first and second hops, and power imbalance between hops. We also consider predictive relay selection to improve performance in delay limited feedback channels. Fig. 1.3 depicts a partial relaying scheme where a subset of relays are selected for transmission in the second hop. Prediction may be applied to compensate for the losses due to outdated channel state information because of feedback delay or low feedback rate. Apart from deriving novel closed form solutions, we succinctly analyse system performance w.r.t different system parameters. Closed form solutions for performance metrics are derived and are verified numerically. While employing MIMO techniques to relaying is in general found to be beneficial in improving throughput and diversity, we demonstrate that this need not be so for the specific case of partial relaying; in general an intelligent resource distribution of resources is required to achieve best performance.

Chapter 7 summarizes the major contributions in each stage of work, and also contains a brief discussion on future research directions.

In **Appendix B**, an FGPA (field programmable gate array) implementation of a reconfigurable MIMO test-bed, useful in validating the performance gains of any given MIMO algorithm over realistic channels and different operating, conditions is presented. Firstly, as an overview, implementation aspects of a MIMO test-bed and design considerations for developing a flexible, scalable and modular prototype are discussed. A thorough study of the overall architecture including RF, IF, and baseband FPGA design is presented which is important as this knowledge is finally used in the development of the 4×4 MIMO baseband architecture. The final sections describe the underlying FPGA design blocks and demonstrate the set-up and validation of the 4×4 MIMO loop back test. This test-bed is useful for validating future research work in MIMO technology and also in developing efficient working architectures that are demonstrable. More importantly they allow a study of potential reconfigurable solutions for MIMO.

Chapter 2

Overview of MIMO systems

“The wireless telegraph is not difficult to understand. The ordinary telegraph is like a very long cat. You pull the tail in New York, and it meows in Los Angeles. The wireless is the same, only without the cat.”

Albert Einstein

A major obstacle to reliability across wireless channels is fading, which refers to deep attenuation of channel amplitude which may be experienced due to the mobility of users and surrounding obstacles. Traditionally, to mitigate the effects of fading, a receiver uses multiple antennas, a technology known as antenna diversity. Multiple input, multiple output (MIMO) is one mature system that utilises the multiple antennas at transmitter and receiver to improve performance of wireless communications. In fact, MIMO has already been identified as a key technology for achieving the high data rate demands of next generation wireless networks by offering increased spectral efficiency and link reliability without consequential increased spectrum usage [1]. For reliable communications, high spectral efficiency must be accompanied by low error rate. Theoretical information studies have shown that the potential channel capacity of a MIMO system can be considerably higher than that of a single-input single-output (SISO) system [13], [14].

This chapter has two goals. First it will provide a brief overview of MIMO beginning with the characterisation of a general wireless multipath channel along with some definitions of standard performance measures and channel models used in this thesis. Next a discussion about the different types of gains that can be realized by a MIMO system is presented where antenna selection is identified as an attractive

feature in bringing down hardware costs in MIMO systems. This is followed by a discussion on adaptive MIMO schemes, and a discussion on the potential benefits and challenges of combining MIMO with adaptive modulation.

2.1 The Multipath Fading Channel

A wireless channel is usually characterised using two different propagation models: to cover *large-scale* and *small-scale* fading. The large-scale fading model accounts for the average power of the wireless signal over a distance, which takes into consideration the distance d between transmitter and receiver. This model attempts to predict the average received power at any given distance d . The received intensity as per the popular *log-distance* power model is given as [15]:

$$P(d) = G(d_0) \left(\frac{d}{d_0} \right)^{-p} \quad (2.1)$$

where d_0 is the reference distance, and $G(d_0)$ is a gain depending on antenna parameters and wavelength, and p is the path loss exponent which indicates the rate at which signal strength decreases with d . The value of p is usually determined empirically and depends on the propagation environment. For free-space $p = 2$, while this can take higher values depending on the amount of shadowing the signal experiences as a result of surrounding clutter. More popular large scale outdoor propagation models include the well known Okumara and Hata empirical models [15] used for accurate urban signal power prediction for GSM frequency bands, where the height of base station is included in the analysis. Multipath fading occurs when different reflected versions of the transmitted signal interfere with each other at the receiver to cause rapid fluctuations in signal intensity over a short duration. This effect, more generally termed small-scale fading, is caused by nearby reflecting bodies or scatterers. The received wave is a superposition of the results of the phenomenon of reflection, diffraction and scattering by obstacles present in the propagating medium between the transmitter and receiver. The multipath waves create random fluctuations in the amplitude and phase of the received signal, causing signal distortion at the re-

ceiver. In other words, the small-scale fading model characterises rapid signal power fluctuations, about a local mean received power (which is itself characterised by the large-scale fading model). Several factors which influence these random fluctuations are: multipath propagation, movement of mobile receiver or relative movement of surrounding objects and the bandwidth of the transmission channel. Degradation due to multipath propagation results in multiplicative noise, and is more deleterious in nature in reducing signal quality at the receiver, when compared to effects due to receiver noise, which is additive in nature. In a MIMO system, multiple antennas at either end of the communication system are used to exploit the multipath propagation in the wireless channel, in bringing about an improvement in system performance, increasing effective throughput and signal reliability. Below we will begin by mathematically characterising a MIMO channel.

2.1.1 MIMO small-scale fading model

The local average signal power decreases significantly only when the receiver moves over large distances. Hence, in theoretical work, the mean average received power is usually taken to be a constant and most analysis is done using the small-scale signal model. It is this local average signal level that is predicted by large-scale propagation model. For a MIMO system the channel matrix at any time instant t , is given by $\mathbf{H}(t, \zeta)$ which represents a $N_r \times N_t$ channel matrix where N_r and N_t are the number of receive and transmit chains respectively. The channel matrix is represented as follows:

$$\mathbf{H}(t, \zeta) = \begin{bmatrix} h_{11}(t, \zeta) & h_{21}(t, \zeta) & h_{31}(t, \zeta) & \cdots & h_{N_t 1}(t, \zeta) \\ h_{12}(t, \zeta) & h_{22}(t, \zeta) & h_{32}(t, \zeta) & \cdots & h_{N_t 2}(t, \zeta) \\ \vdots & \vdots & \ddots & & \vdots \\ h_{1 N_r}(t, \zeta) & h_{2 N_r}(t, \zeta) & h_{3 N_r}(t, \zeta) & \cdots & h_{N_t N_r}(t, \zeta) \end{bmatrix} \quad (2.2)$$

Figure 2.1: Multipath propagation due to obstacles via. reflection, diffraction and scattering in a wireless MIMO channel.

$h_{ij}(t, \zeta)$ represents the *impulse response* of the time varying mobile radio channel between the i^{th} transmit element and the j^{th} receive element, where t denotes the time variability of the channel and ζ stands for *multipath delay* for a given t . The properties or distribution of $h_{ij}(t, \zeta)$ essentially characterise the small-scale fading model. The channel impulse response is the sum of a large number of attenuated, delayed and phase-shifted versions of the transmitted signal, interfering either constructively or destructively at the receiver. As a result, the received signal constitutes several delayed and attenuated versions of the original transmitted signal or symbol. Figure 2.1 shows a MIMO communication system undergoing multipath propagation in the presence of scattering, reflection and diffraction.

The multipath delay is termed the *delay spread* of the channel, denoted as δ , which is the delay between the first dominant multipath signal and the weakest multipath component. The delay is usually specified by the mean or the root-mean-square (r.m.s) of the delay of reflections, σ_δ . In an urban scenario, the delay is usually high due to the presence of many reflectors, such as buildings. Typical values of r.m.s *delay*

spread for 2 GHz and 6 GHz frequencies in an urban micro-cellular environment range from $100ns$ to $1\mu s$ with $400ns$ being the 80% percentile score [16]. The *coherence bandwidth* of the channel is defined as the bandwidth over which a transmitted signal can be considered frequency-flat; denoted by W_C , it bears an inverse relationship to δ or σ_δ . The bandwidth W_C with 90% correlation between frequencies can be approximated as [8]:

$$W_C \approx \frac{1}{50\sigma_\delta} \quad (2.3)$$

For narrowband fading or flat fading, we can drop the ζ dependency, as the channel becomes *frequency non-selective*. This means that the delay spread is smaller than the receiver sampling interval, and the channel is also called a flat-fading or *narrowband* MIMO channel.

Using discrete time notation, we can represent the flat-fading MIMO channel as $\mathbf{H}(k)$ at any time kT_s simply as $h_{ij}(k)$, where T_s is the symbol rate. Thus in a flat-fading channel $h_{ij}(k)$ will have constant amplitude over the symbol period or time period kT_s , so that typically $T_s > \sigma_\delta$. This is termed narrowband because the bandwidth of the transmitted signal is small compared to W_C , or $W_S < W_C$. If the antennas are well separated, then different signal paths will undergo uncorrelated or independent fading. If we assume this signal independence between each transmit-receive pair, then $h_{ij}(k)$ can be characterised as independent and identically distributed (i.i.d) complex Gaussian random variables with distribution $\mathcal{CN}(0, \sigma_h^2)$, where the mean square value, $\mathbb{E}[|h_{ij}|^2] = \sigma_h^2; \forall i, j$. The time varying nature of the channel is more clearly defined by the Doppler spread W_D or doppler shift f_d , caused by relative motion between the receiver and transmitter, or movements of objects in the fading environment. f_d is also proportional to the carrier frequency, f_c .

This accounts for how fast the channel varies over time. Fig. 2.2 depicts two cases where the Doppler frequency is 20 Hz (in blue) and 100 Hz (in red) respectively, with symbol or channel sampling time $T_s = 10\mu s$. For mobile communication in high speed vehicles, Doppler shift has a major impact on system performance. Even when mobility is relatively low and the operating frequency is high (such as WLAN with

Figure 2.2: Doppler fading for 20 Hz (in blue) and 100 Hz (in red) in Rayleigh channels with $T_s=10\mu s$.

$f_c = 5$ GHz), temporal variations caused by human movement or human induced motion on the otherwise static communication channel is enough to cause sufficient Doppler shift. In [17] Doppler spreads of up to 6 Hz were measured in a typical indoor working environment at the 900MHz GSM band; this implies that in a present day WLAN environment it is possible to have Doppler shifts in excess of 30 Hz. Related to the Doppler shift f_d , is the coherence time T_C of the channel, i.e the time over which a channel can be assumed to be constant. This parameter is useful in system design, usually in selecting the block or frame length for data transmission. This is inversely proportional to f_d , i.e for a channel with high Doppler shift, which rapidly changes, the channel remains fairly constant over a few symbol instants only. A conservative estimate for T_C to ensure negligible channel variation over the chosen block or frame interval is to choose [18]:

$$T_C \approx \frac{1}{100f_d} \quad (2.4)$$

Another rule of thumb is to take $T_C \approx 1/6f_d$, to achieve a correlation of more than 50% between channel instants separated by T_C seconds. If the channel is frequency selective, one could use an OFDM (orthogonal frequency-division multiplexing) system, to convert the channel into a set of parallel frequency-flat channels. The OFDM tech-

nique is highly advantageous since this equalisation process is simplified in frequency domain. Throughout the thesis, we will deal with the flat-fading MIMO channel. Using OFDM based subcarrier techniques, the results and techniques in this thesis can hence be applied to wideband systems.

2.1.2 Fading distributions

The small-term fading amplitudes or strengths are usually modelled by Rayleigh PDF, although several models exist which attempt to closely describe the statistical behaviour, taking into account both LOS (line of sight) and non LOS signal propagations. The fading amplitude $\alpha_{ij} = |h_{ij}|$ is typically characterised by the Rayleigh fading distribution in a non LOS condition. The PDF of $|h_{ij}|$ is distributed [19] as :

$$f_{\alpha_{ij}}(\alpha) = \frac{2\alpha}{\sigma_h^2} \exp\left(-\frac{\alpha^2}{\sigma_h^2}\right), \quad \alpha \geq 0 \quad (2.5)$$

where, as before, $\mathbb{E}[|h_{ij}|^2] = \sigma_h^2; \forall i, j$. However, one is generally more interested in the instantaneous SNR per symbol defined as $\gamma_{ij} = \frac{E_s \sigma_h^2 |h_{ij}|^2}{N_0}$, where E_s is the average power of the transmitted symbol, and N_0 is the receiver noise per receive branch. Then γ_{ij} is Rayleigh or gamma distributed [19] as¹:

$$f_{\gamma_{ij}}(\gamma) = \frac{1}{\bar{\gamma}} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right), \quad \gamma \geq 0 \quad (2.6)$$

where $\bar{\gamma} = \mathbb{E}[\gamma] = \sigma_h^2 E_s / N_0$ is the average SNR per transmit-receive branch.

Nakagami- n or Rician- K Distribution

If there is one strong LOS component, then it is usually modelled by a Nakagami- n distribution or the Rice-K distribution ($K = n^2$) where n is the fading parameter and K is an indicator of the strength of the LOS component. The PDF of the SNR is given by [19]:

$$f_{\gamma}(\gamma) = \frac{(1+K)}{\bar{\gamma}} \exp\left(-\frac{\gamma(1+K)}{\bar{\gamma}} - K\right) I_0\left(2\sqrt{\frac{K(1+K)\gamma}{\bar{\gamma}}}\right); \gamma \geq 0, 0 \leq n \quad (2.7)$$

¹indices i,j can be dropped since we are assuming same distribution for all channels

Nakagami- m Distribution

The Nakagami- m channel with parameter m , has PDF is given as [19]:

$$f_{\gamma}(\gamma) = \frac{m^m \bar{\gamma}^{m-1}}{\bar{\gamma}^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right); \gamma \geq 0, m \geq 0.5 \quad (2.8)$$

where $\Gamma(\cdot)$ is the Gamma function. The Nakagami- m statistical model is known to be an accurate fit for urban radio multi-path measured fading data. This enables performance evaluation for different fading scenarios of LOS and non-LOS conditions between transmitter and receiver depending on relative distance between them. The distribution includes the Rayleigh and the one-side of the Gaussian distribution for $m = 1$ and $m = 0.5$ respectively. When $m \rightarrow \infty$, this limits to the case of no fading and only AWGN (additive white gaussian noise) effects influence performance. The Rice distribution is closely related to the Nakagami parameter via the following formula [19]:

$$m = \frac{(1 + K)^2}{1 + 2K} \quad (2.9)$$

2.1.3 Spatial Correlation

The elements of the MIMO channel matrix $\mathbf{H}$ in general might be correlated due to insufficient spacing between antenna elements, or the presence of a direct path in the propagation channel. With correlation, the effective narrowband MIMO channel can be expressed using the popular Kronecker model as :

$$\mathbf{H}_e = \mathbf{R}_r^{1/2} \mathbf{H} \mathbf{R}_t^{1/2} \quad (2.10)$$

where $\mathbf{R}_r$ and $\mathbf{R}_t$ are the receive and transmit correlation matrices respectively, $\mathbf{H}$ is the original discrete time MIMO uncorrelated matrix, with independent and identically distributed complex Gaussian entries. As such, this model assumes independence between fading processes at the transmitter and receiver. $\mathbf{R}_r$ is the $N_r \times N_r$ receive correlation matrix between rows of $\mathbf{H}$ and $\mathbf{R}_t$ is the $N_t \times N_t$ transmit covariance matrix between columns of $\mathbf{H}$. In general, correlation between antenna elements reduces diversity and capacity gains. However, it is possible to obtain low correlation

particularly in base stations, where one is able to separate the antenna elements by a distance of more than $\lambda/2$ - a spacing which is sufficient to reduce mutual coupling between co-located antenna elements, thereby lowering correlation. In small mobile receivers, space constraints usually preclude the use of multiple antennas. Usually at most two antennas can be housed. However, by exploiting different types of antenna diversity schemes (other than spatial diversity) such as polarisation (receiving on orthogonal polarisations) and pattern diversity, or a combination of the above, integrating several antennas without inducing mutual coupling is possible (with polarisation diversity one could essentially obtain the equivalent of several antennas in the same form factor). One such campaign has been undertaken in [20] where three receive antennas were housed in a mobile device to obtain a 10% outage capacity gain of 5.7 bps/Hz, with low correlation coefficients of below 0.1.

2.2 Gains achievable in MIMO channels

Consider a channel which undergoes flat fading. The transmitted signal x , is an $(N_t \times 1)$ vector as transmitted from N_t antenna elements. The receive vector y can be expressed a linear equation as follows:

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x} + \mathbf{n}(k) \quad (2.11)$$

where $\mathbf{n}$ is assumed to be the AWGN with $\mathbb{E}[\mathbf{nn}^H] = N_0\mathbf{I}_{N_r}$, with N_0 as the receiver noise power and $\mathbf{I}_{N_r}$ as the identity matrix of dimension N_r by N_r . Depending on the arrangement and antenna configuration, different types of gain can be extracted or derived from a MIMO system. Performance improvements come in the form of diversity gain, multiplexing gain and array gain or beamforming or a combination of all. Such schemes strive to increase the effective instantaneous SNR γ at the receiver through signal diversity combining or selective diversity methods. To increase net capacity C , independent stream of data are sent through different antennas at the transmitter for appropriate reception at the receiver, so that the overall capacity can be expressed as a sum of the capacities of each of the individual spatial channels.

There are a wide variety of MIMO algorithms, which exploit such gains and do so by employing either spatial division multiplexing [27], beamforming or some type of coding across space and time, which include STBCs [28, 29], space time trellis codes (STTCs) [30], and layered space time codes (STCs) [31]. In some ways, these schemes are complementary as the advantages of one are often the drawbacks of the other. Before we discuss them in more detail, we first define some of the standard and widely used performance metrics in order to best quantify the performance of any communication system. We will start with the average error probability, and outage probability or effectively outage capacity.

2.2.1 Performance Metrics

2.2.1.1 Average Symbol error probability

Assuming maximum likelihood detection, the mean symbol error rate (SER) at an average SNR is found by averaging the probability of symbol error in AWGN over the fading distribution $f_\gamma(\gamma)$ in slow fading scenarios. Average SER for coherent demodulation can be found from [21]:

$$\bar{P}_s \approx \int_0^\infty \alpha Q(\sqrt{\beta\gamma}) f_\gamma(\gamma) d\gamma \quad (2.12)$$

α and β are constellation specific constants and $Q(\cdot)$ is the Q -function [21]. The approximation is due to the Union bound approximation employed for QAM constellations. However, setting $\alpha = 1$, $\beta = 2$, an exact BER (bit error rate) solution can be found for BPSK/ QPSK, while approximate (but accurate) SER for the spectrally efficient M -QAM is obtained with $\alpha = \frac{4(\sqrt{M}-1)}{\sqrt{M}}$ and $\beta = \frac{3}{(M-1)}$. Specific methods to evaluate the above integral can be found in [19].

2.2.1.2 Outage Probability

Outage probability is defined as the probability that the received power or SNR goes below a particular threshold γ_T [21]. Mathematically this is:

$$\begin{aligned} P_{out}(\gamma_T) &= Pr(\gamma < \gamma_T) \\ &= \int_0^{\gamma_T} f_\gamma(\gamma) d\gamma \end{aligned} \quad (2.13)$$

The value of threshold is something chosen during the system design phase as a minimum SNR value for determining the QoS (i.e. requirements for acceptable performance at receiver). For example, in a voice network, the threshold SNR value may be set to obtain some minimum BER value that denotes acceptable voice quality output. $P_{out}(\gamma_T)$ would then determine the outage performance, ie. how often the output SNR γ goes below this threshold and how well the operating device meets its target.

2.2.1.3 Ergodic Capacity

The ergodic capacity of a MIMO channel with average power constraint E_s is the ensemble average of the information rate over the distribution of the channel matrix $\mathbf{H}$, given as [13], [14]:

$$C = \mathbb{E} \left[\max_{tr(\mathbf{R}_{xx}) \leq E_s} \log_2 [\det(\mathbf{I}_{N_r} + \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H)] \right] \quad (2.14)$$

where the output covariance matrix can be determined as $\mathbb{E}[\mathbf{y}\mathbf{y}^H] = \mathbf{R}_{yy} = \mathbf{H}\mathbf{R}_{xx}\mathbf{H}$. By optimising over $\mathbf{R}_{xx}$, the capacity is maximised. When there is no CSI at the transmitter end, the optimal power allocated to each of the transmit antennas is simply given by E_s/N_t where E_s is the total power available across all transmit antennas. In other words, the covariance matrix of the signal is given as $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{R}_{xx} = \frac{E_s}{N_t} \mathbf{I}_{N_t}$. The ergodic capacity of the MIMO channel can then be written as:

$$C = \mathbb{E} \left[\log_2 \det \left(\mathbf{I}_{N_r} + \frac{\bar{\gamma} \mathbf{H}\mathbf{H}^H}{N_t} \right) \right] \quad (2.15)$$

2.2.1.4 Outage capacity

Another definition similar to outage probability is the outage capacity defined as the probability that the capacity C falls below a certain required rate R [13].

$$P_{C_{out}} = Pr(C \leq R) \quad (2.16)$$

This denotes that there is a finite probability q that the channel capacity is less than the required rate R at any particular instant. In other words, we can define a $q\%$ outage capacity as the transmission rate that is guaranteed for $100 - q\%$ of channel realisations. This definition of outage capacity is more appropriate when a channel code word does not all experience all channel fade states, for example in a quasi-static fading scenario. If the channel state does not support a given rate then it is declared to be in outage. This is relevant in slowly-varying channels where the channel matrix $\mathbf{H}$ is constant over a relatively long transmission time. For any choice of rate there will be an outage probability which defines the probability that the transmitted data is not received correctly.

2.2.2 Diversity Gain

Diversity indicates the possibility of sending multiple copies of data through different independent fading paths. At least three different kinds of diversity can be exploited in multipath channels; time, frequency and space. Time diversity is exploited from the total or partial retransmission of information at different instants of time, while frequency diversity forms redundancy across different frequencies. In a fading channel, time diversity could be achieved by signal repetition with transmission instants separated by at least the coherence time, so as to achieve independent fading. In practice this is usually achieved in fast fading channels via an interleaver and error control coding scheme where the interleaver disperses different coded symbols across different transmission times. Analogous to time diversity, we can harvest the diversity gains from frequency selectivity [22], by transmitting in different frequency bands spaced apart by value equal to coherence bandwidth. For wideband transmission, the

channel response is frequency selective, and multiple copies of a transmitted symbol arrive at the receiver over several signalling intervals- these are then effectively combined at the receiver to obtain an increase in SNR. Spatial diversity, in the form of antenna diversity, is harnessed by obtaining independently faded copies of the same signal across available receiver antennas. Thus spatial diversity makes use of spatially distinct links affected by different fading paths between transmitter to receiver. Spatial diversity is a very powerful technique because it does not require an increase in time or bandwidth. At the receiver side it is possible to reconstruct the signals from the $N_t N_r$ links of the MIMO channel, obtaining a resultant signal which presents a considerably reduced variability in comparison to a SISO link. To quantify diversity gain, one usually refers to the diversity order which is in turn related to the reliability of the system.

The diversity order

This is defined as [21]:

$$G_d = - \lim_{\bar{\gamma} \rightarrow \infty} \frac{\log P_e}{\log \bar{\gamma}} \quad (2.17)$$

where P_e is the error rate or the outage probability of the system at any given value of SNR respectively. G_d will be the slope of the error curve vs. $\bar{\gamma}$ when plotted on a log – log scale. Diversity orders of different configurations are discussed below.

SIMO/MISO : In case of a single transmitter and N_r receive antennae i.e in a SIMO system, using the optimal maximum ratio combining technique (MRC) [3]. In MRC, before summing, the signals from individual branches are co-phased by obtaining channel amplitude and phase measurements. This brings about an increase in average SNR. For N_r receivers and single transmit antenna:

$$\mathbf{y}(k) = \mathbf{h}(k)\mathbf{x} + \mathbf{n}(k) \quad (2.18)$$

where $\mathbf{h}$ is a $N_r \times 1$ channel vector. Then the effective SNR after combining is given by [19]:

$$\gamma = \frac{E_s \sigma_h^2}{N_0} \sum_{j=1}^{N_r} |h_{1j}|^2 \quad (2.19)$$

SER of SIMO channel: Knowing the statistics or PDF of the SNR γ one can evaluate the performance metrics such as SER, outage probability as listed previously. If each path is Rayleigh distributed then γ will be gamma distributed as $\gamma \sim G(N_r, E(\gamma))$; assuming independence between MRC paths [19].

$$f_\gamma(\gamma) = \frac{\gamma^{N_r-1}}{[\mathbb{E}(\gamma)]^{N_r} \Gamma(N_r)} \exp\left(-\frac{\gamma}{\mathbb{E}(\gamma)}\right) \quad (2.20)$$

Specifically the error probability for the MRC scheme can be shown to decay like $1/\bar{\gamma}^{N_r}$ at high SNR, in contrast to $1/\bar{\gamma}$ in a SISO channel, i.e for the single antenna fading channel.

Capacity of a SIMO channel: For a SIMO system with single transmit and N_r receivers, (2.15) can be manipulated by Sylvester's determinant theorem, reducing to:

$$C = \mathbb{E} \left[\log_2 \left(1 + \bar{\gamma} \sum_{j=1}^{N_r} |h_{1j}|^2 \right) \right] \quad (2.21)$$

This is also the ergodic capacity of a SIMO system with MRC reception. Equivalently, if the system is configured as a MISO system, for transmit side processing, channel gains need to be fed back for transmitter preprocessing, and a MISO system can achieve the same diversity gain G_d [23] as well as the capacity in (2.21). In principle, the maximum degree of diversity that a MIMO channel can offer is given by $N_t N_r$, this can be achieved through transmit beamforming techniques, namely maximum ratio transmission or MRT, by feeding back the whole channel matrix for transmit side processing [23], [24]. The optimal transmit beamforming vector that maximises the SNR is the eigen-vector corresponding to the maximum Eigen-value of $\mathbf{H}^H \mathbf{H}$. The higher the diversity gain, the higher is the signal reliability.

Space time block codes : These are different from SIMO/MISO or MRT. It is possible to exploit spatial diversity also without knowledge of the channel at the receiver side using a suitably constructed transmit signal such as STCs [25]. Antenna diversity at the receiver is well-known, and has been studied for a long time, with maximum diversity obtained being N_r . However, achieving this diversity is not straightforward with multiple transmit antennas, when the channel is unknown at the

transmitter. Alamouti introduced a clever way to achieve the maximum diversity gain for two transmit antennas [26]. The idea is to transmit data symbols simultaneously both in space and time in such a manner so that orthogonality between data streams was maintained upon signal demodulation at the the receiver.

Capacity of the MISO channel: Using (2.15) the general capacity with single receiver and N_t transmitters reduces to:

$$C = \mathbb{E} \left[\log_2 \left(1 + \frac{\bar{\gamma} \sum_{i=1}^{N_t} |h_{i1}|^2}{N_t} \right) \right] \quad (2.22)$$

Thus from (2.22) and (2.21) it is seen that the capacity of MISO without transmitter adaption is less than that for a SIMO channel. Essentially, to achieve MISO capacity without transmitter adaptation for two transmit antennas, the Alamouti MISO system can be used. The extension of the Alamouti technique to systems with more than two transmit antennas is called Orthogonal Space-Time Block Coding (OSTBCs) [25], and presents a loss in spatial rate, since it is not able to transmit more than one symbol per time slot (averaged over the course of one block). For space-time block coding, the effective SNR as a result of orthogonality can be expressed in terms of the channel Frobenius norm, resulting in a diversity gain. Considering OSTBC, the received SNR after MRC can be given as [15] :

$$\gamma_{OSTBC} = \frac{\bar{\gamma}}{N_t} \|\mathbf{H}\|_F^2 = \frac{\bar{\gamma}}{N_t} \sum_{i=1}^{N_t} \sum_{j=1}^{N_r} |h_{ij}|^2 \quad (2.23)$$

The effective SNR is therefore a a sum of $N_t N_r$ gamma distributed random variables, scaled by a factor N_t ; therefore the PDF of γ_{OSTBC} is another Gamma distribution $\gamma \sim G(N_r N_t, E(\gamma)/N_t)$, given as:

$$f_{\gamma_{OSTBC}}(\gamma) = \frac{N_t^{N_t N_r} \gamma^{N_t N_r - 1}}{[\mathbb{E}(\gamma)]^{N_t N_r} \Gamma(N_r N_t)} \exp \left(-\frac{\gamma N_t}{\mathbb{E}(\gamma)} \right) \quad (2.24)$$

From this one could do the SER analysis. The spatial code rate is $r_s < 1$ where $r_s = N_t/T$ is the number of independent data symbols in a codeword divided by the frame length T . Although these kinds of codes are not particularly efficient in terms of achieved data-rate they have been particularly attractive for practical applications

because they are very easy to decode through linear combining owing to the absence of mutual interference due to orthogonality. To maintain orthogonality, however, OSTBC sacrifices spatial multiplexing gain.

Antenna Selection : Another method to harness diversity in a MIMO channel is through antenna selection either at the transmit or receiver end or a combination of both [27]. The idea is to pick the best set of antennas at any given instant so as to maximise the received SNR. Antenna selection is a bandwidth-efficient feedback mechanism and is a useful feature at higher speeds, when the rate of the feedback is quite high. Antenna selection has the added advantage that unlike other closed-loop MIMO modes, the number of required RF chains is equal to the number of streams. Therefore antenna selection can lower the cost of a MIMO system by replacing multiple copies of costly RF electronics by simple antenna elements and a switch. Antenna selection can be achieved in various forms by selecting a subset of best antennas at the receiver or transmitter so as to increase received SNR or equivalently minimize the BER. This means that we choose that subset of antenna elements for which the capacity of the system (or outage capacity) is the highest among all capacity values achieved by any other possible antenna subset. Another criterion that is often used for optimal antenna selection is maximization of the SNR (instantaneous or average) of the reduced-complexity system. In cases of statistical antenna selection, the criterion that is mostly used is minimization of the average symbol error rate or minimization of the average symbol error probability. Transmit antenna selection, where a single best antenna is selected amongst N_t antenna elements when combined with MRC is known to provide a diversity order of $N_t N_r$; the best antenna is selected as [28]:

$$l = \arg \max_{1 \leq i \leq N_t} (\gamma_i); \quad \gamma_i = \sum_{j=1}^{N_r} |h_{ij}|^2 \quad (2.25)$$

SER analysis: To obtain the PDF of the effective SNR one may resort to results from order statistics [29]. In TAS the best antenna SNR is referred to as the N_t^{th} order statistic. If the r.vs SNR corresponding to the different antennas are arranged as $\gamma_{(1)} \leq \gamma_{(2)} \leq \dots \leq \gamma_{(N_t)}$, the antenna corresponding to the N_t^{th} order statistic will be the one offering the maximum gain. Since γ_i are assumed to be i.i.d, then

according to order statistics, in general the PDF of γ_k , the k^{th} order statistic can be given as:

$$f_{\gamma_k}(\gamma) = \frac{N_t!}{(k-1)!(n-1)!} f_{\gamma}(\gamma) [F_{\gamma}(\gamma)]^{N_t-1} [1 - F_{\gamma}(\gamma)]^{N_t-k} \quad (2.26)$$

where $f_{\gamma}(\gamma)$ and $F_{\gamma}(\gamma)$ are the PDF and CDF of the equivalent standalone $1 \times N_r$ MRC channel.

Capacity of a Selective diversity channel: With selection diversity used at the transmit end, such as with transmit antenna selection, full transmit power can be apportioned to a single transmitter at any time, with the maximisation performed over the set of transmit antennas. Modifying 2.21, we can see that the capacity becomes:

$$C = \mathbb{E} \left[\log_2 \left(1 + \bar{\gamma} \max_{1 \leq i \leq N_t} \sum_{j=1}^{N_r} |h_{ij}|^2 \right) \right] \quad (2.27)$$

While transmit antenna selection has been less studied, historically much attention has been given to receive antenna selection. Several forms of receive antenna selection exist which are analogous to TAS alternatives, such as selection combining, switch and stay combining, and generalised switched diversity where signals N_c of N_r receive antennas are selected for effective combining [30]. A more detailed study of antenna selection techniques is presented in **Chapter3**.

2.2.3 Array Gain

This kind of gain is due to the use of an antenna array and coherent combining effect at the receiver that leads to an increase in average received SNR. Formally, this is defined as :

$$\mathbb{E}[\gamma] = \int_0^{\infty} \gamma f_{\gamma}(\gamma) d\gamma \quad (2.28)$$

where γ is the effective SNR at the receiver. This is achieved through the availability of CSI at receiver side or transmitter. In a SIMO system, array gain can be achieved via MRC signal combining (which also achieves diversity gain), while for a MISO

case, CSI has to be conveyed to the transmitter so that beamforming can achieve array gain. In both cases there is an increase in average SNR which is larger than the SISO case by the number of receive antennas (MRC) or transmitters (MISO) case. For a MIMO case, both transmitter and receiver need CSI to achieve full array gain. This requirement makes it more difficult to exploit the gain because of the inherent difficulty in obtaining CSI at the transmitter side. Knowledge of the transmitter CSI helps us to achieve diversity and spatial gains and also reduce decoding complexity. OSTBCs offer limited array gain owing to the fact that now power is equally distributed among transmit antennas. The average SNR gain is found from 2.23 as:

$$\mathbb{E}[\gamma_{OSTBC}] = \mathbb{E} \left[\frac{\bar{\gamma}}{N_t} \|\mathbf{H}\|_F^2 \right] = \mathbb{E} \left[\frac{\bar{\gamma}}{N_t} \sum_{i=1}^{N_t} \sum_{j=1}^{N_r} |h_{ij}|^2 \right] = \sigma_h^2 \bar{\gamma} N_r \quad (2.29)$$

Thus the array gain is same as that of MRC, or equivalently that of transmit side MRC. On the other hand, techniques such as transmit antenna selection can achieve about both diversity as well as array gain. Although suboptimal when compared to full blown implementations like MRT, they provide a good tradeoff between hardware and processing complexity and diversity, array, or capacity gains that can be realised.

2.2.4 Spatial multiplexing gain (SM)

Multiplexing gain is related to the data rate that the system can convey over the MIMO channel. MIMO channels can offer linearly increased capacity with respect to a SISO link without any corresponding increase in time or bandwidth. This kind of gain can be exploited by transmitting independent data signals from different antennae. In other words, the multiple transmit antennae are used to multiplex data in space offering a spatial multiplexing gain. Pioneering work by Foschini [14] and Telatar [13] showed that the capacity of MIMO can be $\min(N_t, N_r)$ times larger than that of a single-antenna alternative. This can be realised after rewriting 2.15 as will be shown in the following paragraphs.

Capacity of a MIMO channel : The MIMO channel can be converted to n SISO channels, with $r = \min(N_t, N_r)$, using singular value decomposition, and the capacity be re-written as [13]:

$$C = \mathbb{E} \left[\sum_{m=1}^r \log_2 \det \left(1 + \frac{\bar{\gamma} \lambda_m^2}{N_t} \right) \right] \quad (2.30)$$

r is the rank of the matrix and λ_m , $m = 1, \dots, r$ are the non-null eigenvalues of the $\mathbf{H}\mathbf{H}^H$ matrix when $N_r < N_t$ and $\mathbf{H}^H\mathbf{H}$ when $N_t \leq N_r$, also known as the *Wishart* matrix. A numerically integrable expression for the capacity of the MIMO channel in terms of Laguerre polynomials has been calculated by Telatar [13], after obtaining the joint density of the ordered eigenvalues.

When CSI exists at the transmitter, power can be optimally allocated via methods such as water-filling [31], so that the more power is allocated to good channels. The capacity for this case would now become [13]:

$$C = \mathbb{E} \left[\sum_{m=1}^n \log_2 \det \left(1 + \frac{p_m E_s \lambda_i^2}{N_t} \right) \right] \quad (2.31)$$

p_m is the coefficient that allocates power to the m_{th} sub channel. At high values of SNR, the channel capacity can be approximated as $C = r \log_2 \bar{\gamma}$ while at low SNR the channel capacity is approximated as $C = N_r \log_2 \bar{\gamma}$. Thus the channel capacity of a MIMO channel can be r times larger than that of a SISO channel.

Spatial multiplexing gains were originally achieved using the following two approaches:

Parallel Encoding: This is the simplest approach where a bit stream is demultiplexed over N_t different streams. Each would be independently encoded, mapped and interleaved and the association between antenna and stream fixed over time. This system can achieve an N_r^{th} order diversity and $r_s = N_t$, where $N_r \geq N_t$. In this scheme each stream can be decoded separately and so requires a simple receiver. V-BLAST (vertical Bell Labs layered space-time) achieves parallel transmission, where independent layers of data are transmitted from each antenna. It still requires joint detection through zero forcing (ZF) of the codeword from each transmit antenna, but

through successive interference cancellation, the complexity can be significantly reduced. However, this ZF-SIC system achieves a diversity order between $(N_r - N_t + 1)$ and N_r .

Serial Encoding: With this approach, the bit stream is firstly encoded using a channel encoder, mapped and interleaved, and then de-multiplexed across N_t streams to be sent on different transmit antennae. The receiver must decode jointly the entire codeword which is de-multiplexed on the different antennae. The vertical encoding, for sufficiently long codewords, obtains full diversity, but since the complexity of the receiver grows exponentially with the length of the codeword, this could easily lead to an un-implementable complexity at the receiver. D-BLAST is a trade-off between serial and parallel encoding. The data stream is firstly de-multiplexed over N_t streams with each one separately encoded, mapped and interleaved, however the association between stream and antenna is not fixed in time but decided using a rotational technique. This scheme can achieve $N_t N_r$ order diversity but has the main negative drawback that it does not employ all the space-time slots offered by the channel, and if the frame length is not sufficiently long there can be quite a big overhead.

2.3 Adaptive MIMO schemes

Future wireless systems are expected to offer optimized performance by adapting to the varying propagation and network conditions. The system should be able to switch accordingly and perform better given the nature of the channel characteristics in the form of CSI or transmission environment, which could mean low/high SNR or interference. For example, it is generally perceived that, at low SNR, STCs perform better due to their ability to exploit antenna diversity, while at high SNR, SM schemes are preferred for being able to support high rates with simpler modulation and demodulation. Thus, systems may need to be able to switch between STC and SM modes to have good performance under both low and high SNR regimes. Therefore,

a new MIMO transmission approach that adapts to the changing channel conditions based on spatial selectivity information has been proposed [32]. The system switches between different MIMO transmission schemes (beamforming, space-time TD, and spatial multiplexing) as a means of approaching the spatially correlated MIMO channel capacity with low-complexity. Since the adaptation is based on the long term spatial characteristics of the channel, it can be carried out at slow rate, avoiding excessive feedback overhead. Considerable performance gains can be obtained by transmission on the Eigen-modes of the transmit antenna correlation matrix. For example, [33] proposes an optimal linear pre-coder that assumes knowledge of the transmit antenna correlations and improves the performance of a space-time coded system by forcing transmission on the nonzero Eigen-modes of the transmit antenna correlation matrix. Linear pre-coders exploit knowledge of a channel's mean and correlation matrix to optimize performance and achieve robustness with respect to antenna correlation and mobile speed. The pre-coders exploit channel correlation matrix knowledge at the transmitter to weight the space-time block encoder so that the upper bound on pairwise error probability is minimized. Thus the system is adaptive in nature benefitting significantly from adaptation to the channel. A significant advantage of the traditional STC techniques is that they do not require any CSI knowledge at the transmitter. OSTBCs represent an important class of STC method because they achieve full diversity while enjoying simple ML decoding. However, the error rate performance can be further improved if a low rate channel state feedback is available at the transmitter. In particular, diagonally weighted OSTBCs with feedback-driven weights have been proposed in [34], while Alamouti-type wireless systems with one-bit feedback and a combination of antenna selection and power allocation have been studied in [35].

Systems could also adapt the current channel conditions by selecting the most suitable AMC (adaptive modulation and coding) scheme. AMC [36] is a proven technique in approaching the channel capacity of SISO systems. AMC has been applied to singular value decomposition (SVD)-based MIMO systems and shows a good av-

erage spectral efficiency performance [37]. Link adaptation methods combining AMC and MIMO techniques where different MIMO schemes are dynamically switched in conjunction with the appropriate AMC scheme is a challenging and promising area of research, since joint optimisation of these techniques is required to bring about increased link performance and throughput [38].

2.3.1 Adaptability with respect to CSI

CSI available at the transmitter and/or the receiver side can be exploited, either as instantaneous (short-term) information obtained through feedback or channel estimation or as statistical information (long term) based on the moments of the channel. Regarding the degree of CSI at the transmitter, signal and coding design for MIMO systems has traditionally concentrated on two extreme cases: perfect CSI, for which multi-beamforming strategies are optimal, and unavailable CSI, for which space-time codes have traditionally been proposed. MIMO designs for situations with both imperfect and incomplete CSI at the transmitter have only more recently been proposed in the literature [39]. Imperfect CSI refers to the case where the channel estimation presents some errors, while partial CSI describes a situation where the transmitter has only access to some channel information (phases, amplitude, covariance, mean, index of selected antennas etc.) instead of its actual physical realization. Imperfect channel estimation can be modelled as follows [40]:

$$\mathbf{H}(k) = \rho_e \hat{\mathbf{H}}(k) \mathbf{x} + \sqrt{1 - \rho_e^2} \mathbf{\Xi}(k) \quad (2.32)$$

where $\rho_e \in (0, 1)$ is the correlation coefficient between $\mathbf{H} \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$, the original channel matrix and the estimated matrix $\hat{\mathbf{H}} \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$ and $\mathbf{\Xi} \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$. Partial CSI situations have been far less explored, and the relationship between the quality and degree of CSI and the associated capacity-achieving architecture needs to be further studied. Clearly, scalable signal processing/coding designs that adapt themselves to the degree and quality of the available CSI will surely be the optimum choice for these situations. Partial CSI might be acquired either from a feedback

link in frequency division duplex (FDD) systems, or through TDD by exploiting reciprocity properties of the channel. To ensure correlation due to reciprocity, the duplexing distance in TDD should be far lesser than total coherence time of both forward and reverse channels, while in FDD the forward and reverse channel frequency difference should be smaller than the coherence bandwidth. While TDD appears to be the more favourable choice, processing delays can usually contribute to the effective normalised delay ($f_d\tau$), where τ is the delay incurred before acquired CSI is put to good use. For FDD, small frequency separation is usually very difficult to achieve, which makes channel reciprocity difficult to achieve in practise. Therefore CSI has to be acquired, which is in turn affected by processing delays at either end. This type of situation is more valid in a channel with high Doppler spread, where the delay associated with the return link would render any channel information completely outdated. In cases where feedback channel capacity is limited [41], only part of the instantaneous information of the channel is available. Imperfect partial CSI due to feedback delay can be modelled as follows, where the relation between actual and delayed value of the channel $\mathcal{D}$ time or symbol instants earlier is:

$$\mathbf{H}(k) = \rho_d \mathbf{H}(k - \mathcal{D}) + \sqrt{1 - \rho_d^2} \mathbf{\Xi}(k) \quad (2.33)$$

$\rho_d = J_0(2\pi f_d \mathcal{D} T_s)$ [42] is the correlation coefficient and $\mathbf{\Xi}$ is $\mathcal{CN}(0, \mathbf{I}_{N_r})$ distributed. Channel degradation due to feedback delay has a more serious impact on system performance than channel estimation errors, since estimation errors can be improved by increasing beacon-pilot power, and for receiver demodulation proper channel smoothing can be done using non-causal filtering at the receiver. Such feedback delay issues occur in adaptive systems like TAS or AMC, where a delay in feedback of antenna indices reduces diversity gains due to selection. However, even if clean channel estimates are obtained one may have to also account for feedback quantisation errors, for example in CSI reporting for code-book based precoding in spatial multiplexing schemes. To handle feedback delay, channel prediction at the receiver should be employed in delay limited feedback channels, in case the transmitter relies on feedback information [43].

2.3.2 Adaptive modulation and coding

AMC is currently being used in the 3GPP standard in order to utilize the wireless channel in the best possible way for data throughput and achieve high data rates. The factors that govern the achievable performance are channel conditions and signal-to-interference ratio (SIR) estimation accuracy. The current MIMO research focus is to find optimal AMC schemes which adapt suitably to channel conditions [44], [45]. In [46], several techniques of adaptive transmission and diversity-combining are examined, in terms of channel capacity. In particular, the first technique refers to the optimal power and rate adaptation, the second considers constant power with optimal rate adaptation and the last uses channel inversion with fixed rate. The second method offers a small increase in the capacity of the channel compared to the first one, and channel inversion suffers the largest capacity penalty, which decreases as diversity increases. The trade-off between the small complexity of channel inversion and the higher capacity obtained by the other methods is examined. The results are for Rayleigh channels in a system employing MQAM modulation. In [47], un-coded AMC MIMO systems are systematically investigated, showing that both a variable-rate variable-power and a variable-rate MIMO system can achieve full multiplexing gain, if both transmitter and receiver have perfect CSI. In [48], the design of an un-coded AMC MIMO system using imperfect CSI with channel estimation error and feedback delay was proposed. And in [49] an adaptive bit-interleaved coded modulation (BICM) MIMO system was proposed which showed excellent BER robustness against the CSI feedback delay. However, both the imperfect-CSI design and the adaptive BICM MIMO system compromised the multiplexing gain in exchange for robust BER performance. In [50], an adaptive rate-compatible punctured code BICM MIMO system is proposed to achieve near full multiplexing gain and robustness to feedback delay.

Full MIMO systems increase system hardware, and computational complexity costs, with the additional challenge in keeping the form factor small. Antenna selection is a viable approach which addresses these problems. Although antenna selection

per se has been a well researched area, its performance analysis in combination with adaptive modulation in feedback constrained channels has not been fully researched.

2.4 Summary

This chapter has provided a brief summary of some of the principles of MIMO and cooperative relaying. Under MIMO, different types of gain in the form of diversity or multiplexing, are possible through different arrangements of such systems. MIMO systems exploit multipath diversity in a rich fading environment to increase signal reliability and increase capacity of the system without additional need for bandwidth or power. Adaptive closed loop MIMO techniques such as precoding, AMC and so on are vital for harnessing full gains from a MIMO system; techniques like antenna selection are vital in reducing the extra cost of RF electronics and complexity of a full MIMO system, while at the same time maintaining diversity and capacity benefits. Combining adaptive modulation and coding together with antenna selection will therefore be beneficial in providing a good tradeoff between implementation cost and performance gains realisable from such systems. While antenna selection itself has been well investigated, its performance w.r.t imperfections in CSI feedback, and adaptive modulation has not been well researched. **Chapters 3** and **4** deal with transmit antenna selection considering the effect of degraded CSI quality and discuss channel prediction applied to TAS to improve its performance under both non-adaptive and adaptive modulation schemes.

Chapter 3

Antenna Selection in MIMO

“I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician: he is also a child placed before natural phenomena which impress him like a fairy tale.”

Marie Curie

3.1 Introduction

Although MIMO systems are often cited as a solution for achieving the high data rate demands of future wireless networks, increased spectral efficiency and link reliability comes at the cost of increased hardware complexity, power consumption and signal processing complexities. The physical antennas are themselves inexpensive, but the main driving factor is the requirement for multiple parallel RF chains at transmitter, receiver or both. Every extra transmit/receive antenna requires its own hardware chain comprising a power amplifier, low noise amplifier, ADC/DAC etc. Antenna subset selection, where transmission/reception is performed through a subset of available antenna elements, helps in reducing the implementation cost while retaining most of the benefits of MIMO technology. A selection of antenna elements, which are much cheaper than RF chains, is made available at the transmitter and/or receiver, by a dynamic selection scheme where only a subset of antennas are chosen at both ends. In this way, only the best set of antennae are used, while the remaining antennae are not employed, thus reducing the number of required RF chains (since we switch antennas, not RF chains) and receiver complexity, but preferably gaining some channel benefit from their locations as they are switched into and out of

the link dynamically. It has potentially reduced hardware cost compared to space-time or MIMO coding, due to this reduction in the amount of radio frequency (RF) hardware required, and possibly also in reduced computational complexity. Whilst receive antenna selection is perhaps more common, transmit antenna selection also offers several advantages, particularly for hardware-costly transmit schemes such as those requiring linearisation. Early theoretical works in antenna selection chiefly considered receiver diversity systems, such as [42], [51], [52], [53], [54], [55], where one or more branches with largest receive SNR at each instant, were combined to maximize the instantaneous output SNR; the selection schemes are referred to as hybrid selection/maximum ratio combining (H-S/MRC). The schemes suffered small performance loss when compared to the full blown MRC scheme, with considerable improvements in receiver complexity.

Antenna selection in a MIMO system has gained a lot of interest in recent years, where subset transmit/receive antenna selection is performed based on instantaneous or statistical channel knowledge. The best set of antennas are selected either to maximise data rate or minimize error probability [27], [56], [57], [58]. When the channel is fully known at both transmitter and receiver, water filling can be used to maximise capacity [59] and, (from Chapter 2), we have seen that the mutual information or the capacity in a MIMO channel in the absence of transmitter knowledge at any given average SNR $\bar{\gamma}$ is:

$$C_{nonsel} = \log_2 \left[\det(\mathbf{I}_{N_r} + \frac{\bar{\gamma}}{N_t} \mathbf{H} \mathbf{H}^H) \right] \quad (3.1)$$

where power is distributed equally amongst transmit antennas. With antenna selection under capacity maximisation, the effective capacity now becomes,

$$C_{sel} = \sum_{p=1}^{N_r} \log_2 \left(1 + \frac{\bar{\gamma}}{N_t} \lambda_i(\mathbf{H}_{sel} \mathbf{H}_{sel}^H) \right) \quad (3.2)$$

where λ_i are the eigen-values of $\mathbf{H}_{sel} \mathbf{H}_{sel}^H$. For example, if we consider single receive antenna selection with transmitter CSI unknown, the overall system capacity is:

$$C_{max} = \max_{1 \leq i \leq N_r} \log_2 \left(1 + \frac{\bar{\gamma}}{N_t} \sum_{j=1}^{N_t} |h_{ij}|^2 \right) \quad (3.3)$$

To perform receive antenna selection a matrix $\mathbf{H}_{sel}$ of size $N'_r \times N_t$ is selected from the full channel matrix $\mathbf{H}_{N_r \times N_t}$, such that the chosen subset created by striking $N_r - N'_r$ rows from $\mathbf{H}$ results in maximum capacity. For spatial multiplexing systems such as the well known BLAST (Bell Laboratories Layered Space-Time), capacity due to antenna selection at the receiver has been shown to be comparable to the full-complexity system as long as the number of selected chains at the receiver is greater than or equal to the number of transmit elements i.e $N'_r \geq N_t$ [60]. The important conclusion we can draw is that whether selection is done at the transmitter or receiver, full diversity order can be maintained as if all the antennas are used, although with some loss in average SNR. A similar analysis can be done for transmit side selection, with transmitter CSI knowledge, however, instead of assuming equal transmit power levels, better power allocation to selected set of transmit antennas improves performance [61].

For antenna selection in STBC, the selection algorithm chooses the antenna subsets that maximizes the channel Frobenius norm (in case CSI is available), which results in both coding and diversity gain [56]. Results indicated an increased system diversity with the improvement in outage capacity/probability being a good indicator of such an increase. For example considering OSTBC, the received SNR without selection is given as [15] :

$$\gamma_{nonsel} = \frac{\bar{\gamma}}{N_t} \|\mathbf{H}\|_F^2 = \frac{\bar{\gamma}}{N_t} \sum_{i=1}^{N_t} \sum_{j=1}^{N_r} |h_{ij}|^2 \quad (3.4)$$

If the best N'_t out of N_t are selected to maximize SNR, then the effective capacity is represented as:

$$\gamma_{max} = \frac{\bar{\gamma}}{N_t} \sum_{i=1}^{N'_t} \beta_i \quad \text{with} \quad \beta_i = \sum_{j=1}^{N_r} |h_{ij}|^2 \quad ; \beta_1 > \beta_2 \cdots \beta_{N'_t} > \cdots \beta_{N_t} \quad (3.5)$$

Thus N'_t out of N_t columns with the highest Frobenius norms are picked for transmission. Various other antenna subset selection criteria have been discussed in the literature, including power and SNR maximization [56] (which was discussed above), maximization of the ergodic capacity of the equivalent channel [58], [60], [62] and minimization of the average probability of error [28], [57], [63], [64].

The benefits of antenna selection are in general reduced due to channel impairments and non-ideal channel conditions, such as correlation between antenna elements. With only statistical information of the channel, antenna selection in spatial multiplexing with a zero forcing receiver was analysed in [65], where capacity was shown to be maximised by selecting the transmit antennas with minimum product of the diagonal elements of the transmit correlation matrix. To minimize symbol error rate, the optimal scheme was to choose the transmit antenna subset that minimized the maximum diagonal value of the corresponding inverse transmit correlation matrix. Likewise in [66], statistical channel knowledge of the rank of the channel, is used to maximise channel capacity, where a full rank subset transmit antenna selection is chosen, resulting in a fewer number of transmit chains. The optimal signalling for maximising capacity MIMO antenna selection in correlated channels was analysed in [67], [68], [69].

Secondly, channel imperfection exists in the form of channel estimation errors at the receiver, used as a decision for either receive or transmit antenna selection. Only a few papers have studied antenna selection with STC and estimation errors, such as in [28], [70], [71]. The conclusion of these papers is that channel estimation error has no impact on the diversity order when imperfect channel estimates are used in antenna selection and space-time decoding, although some performance degradation is seen. The impact of channel estimation errors can be significantly reduced if the SNR of the pilot tones is comparable to, or larger than, the SNR during the actual data transmission [60].

Another form of impairment is error in the feedback channel which can cause an antenna to be selected which is not the optimum requested by the receiver. In [72] optimal signaling for TAS in the presence of feedback errors is considered. However, feedback delay is of more serious concern to TAS, leading to outdated channel knowledge at the transmitter. Acting on degraded information to select a transmit antenna effectively reduces antenna diversity, adversely affecting symbol error probability (SEP) and BER. In fast fading channels, a return link delay might render any

channel information completely outdated by the time the transmit antennas switch. This is especially true if TAS is used in mobile systems operating at S band frequencies of around 2 to 3.7 GHz, such as in WiMAX, where Doppler shift is an important consideration. Furthermore, such adaptive antenna techniques, when combined with wireless systems that have restricted uplink feedback rates, such as W-CDMA (wide-band code-division multiple-access) or 3GPP (third-generation partnership project – with a 1.5 kbps uplink bandwidth) are more likely to experience large reverse channel feedback delays. Other factors typically contributing to switching delay include decode latency, ARQ (automatic repeat request) handling and block buffering. Therefore methods which combat feedback delay need to be devised either at the receiver or transmitter. In this work we will address the problem of feedback delay and restricted reverse link feedback rates in a TAS system with MRC employed at the receiver.

One major advantage of TAS/MRC is that the performance of downlink channels can be improved by transmitter diversity techniques at the base stations. TAS can be a cost-effective and low complexity closed loop technique in which a receiver periodically advises the transmitter to select a subset of transmit antennas, chosen to maximise receiver SNR. A scheme with a small number of antennas at the mobile set and simple receiver complexity, as well as a reduced number of RF chains is desirable for downlink transmission. At the same time, diversity and array gain benefits are maintained. The array gain is created by apportioning available power to the set of antennas that yield maximum instantaneous SNR. Additional array gain is achieved by MRC combining of the receiver antenna SNRs. Furthermore, antenna index feedback needs only a low bandwidth channel (compared to full CSI feedback).

In the following sections, we analyse a generalized TAS/MRC scheme such as [28], where a single best antenna is selected for transmission based on the knowledge of fading statistics at the receiver. MRC, as an optimal combining scheme irrespective of channel fading statistics, is suitable for most amplitude and phase modulated signals [3, 19], and has been adopted at the receiver. At high SNRs a TAS/MRC scheme such as this can achieve full diversity order, and can approach the perfor-

mance of some complex space-time codes of the same spectral efficiency [28]. In use, TAS requires at least partial channel knowledge at the transmitter in order to perform selection. This knowledge usually comes in the form of an index to the best set of antenna/antennas fed back from the receiver; which implies a delay between the channel being sampled (at the receiver) and this knowledge being acted upon (at the transmitter). In this work, performance degradation due to outdated channel knowledge is determined analytically, and related to channel characteristics. A predictive scheme is then developed to mitigate against delay-induced degradation. Several factors relating to TAS system performance under different channel scenarios both with and without mitigation are explored. Closed form expressions for performance metrics such as BER, outage probability, average signal-to-noise ratio gain and higher order moments of output SNR are derived and verified by simulations. The impact of prediction is analyzed for different TAS setups and channel prediction scenarios, as are various system design parameters.

3.1.1 Related work

Channel prediction was proposed initially for a DECT (Digital European Cordless Telecommunications) framework to improve transmitter combining and selection diversity gain [73], however the analysis was limited to a single receive antenna at the mobile. The importance of long range channel prediction beyond coherence time in mitigating the effect of deep fades was investigated by Eyceoz and Duel-Hallen [74], while others [75], [76] applied channel and power prediction to improve transmit diversity performance in frequency selective W-CDMA channels. However, the analysis was restricted to a two antenna, single receiver system, where an upper bound for BER with perfect prediction was the performance benchmark. With regard to channel prediction, an extensive analysis of linear prediction for mobile radio channel coefficients was developed by Ekman [77, 78], who proposed efficient noise reduction using Wiener smoothers, as well as use of an unbiased power predictor in SISO channels. Øien *et al.* [79] employed prediction using pilot symbol assisted modulation

(PSAM), for combatting delay in an adaptive modulation receive diversity system. Similarly Zhou and Giannakis [80] dealt with channel prediction for transmit beamforming with MRC. Linear channel prediction was applied for two transmit selective Alamouti STBC schemes [81], however this did not include a complete mathematical analysis for the prediction process. Joint transmit and receive antenna selection was considered in [18], [82], [83] where the impact of feedback delay on BER was analysed over flat fading Rayleigh channels. Cai and Giannakis [84] considered the performance of an ideal combined transmit selection diversity and receive generalized selection combining over Rayleigh fading channels. Moving beyond this, a spatial diversity analysis of an ideal TAS/MRC scheme was published by Chen *et al* [28], where performance metrics such as BER for binary phase shift keying (BPSK), outage probability and gains of the TAS/MRC scheme for flat Rayleigh fading channels were derived. In [85], [86], the performance analysis of TAS is analysed in Nakagami- m channels, again ideal non-delayed feedback is considered. The same authors proposed new methods to reduce feedback rates in TAS systems [87], while in [88] different low feedback rate arrangements of TAS are analysed in the presence of erroneous feedback. Zhang [89] carried out a general analysis for the joint selection of L_t out of N_t transmit and L_r out of N_r antennas for orthogonal space-time coded systems and the error rate performance is analysed. In [90], prediction for TAS is considered where its outage performance is analysed under prediction and channel estimation errors.

3.1.2 Contributions

Different from the above methods, in this chapter, long range prediction using a block fading model is applied to TAS /MRC and its performance is analysed w.r.t several system parameters chosen appropriately for maximum performance and minimum receiver computational complexity. Additionally, several new results for performance measures are presented. Thirdly, a trade off between predictive and non-predictive approach is analysed at different operating SNRs. Finally, we also consider the effect of low feedback rates in the presence of feedback delays and compare predictive and

non-predictive approaches in such a scenario.

3.1.3 Summary of Results

To mitigate issues related to delay between channel measurement and switching, we consider a predictive-TAS (TASP) scheme allied with MRC at the receiver, which employs a power predictor to exploit temporal channel correlation. This predictor is designed for TAS in future transmission blocks based upon current and historical channel information [91]. The benefits of using long range prediction (LRP) of channel values for TASP decisions are evaluated and compared with a non-predictive (TASD) scheme acting on outdated channel information, for a range of system configurations. Several performance metrics related to BER, outage probability and SNR gains, which will be useful for efficient system design, have been derived and analysed. A block fading model facilitating LRP is incorporated into the TAS setup, and TASD is compared with TASP. Several design trade-off possibilities, that may be of use to system implementers, will be discussed. In all cases we assume that noiseless estimates are available for receiver demodulation, so that feedback delay effects, with and without prediction, can be isolated.

Section 3.2 presents the system structure and channel models, while in Section 3.2.4, the fading probability density function (PDF) of the TASP scheme is derived. In Section 3.3, closed form BER equations are derived and the merits of using a predictor discussed with respect to various TAS/MRC configurations. Section 3.4 derives the outage probability of the TASP scheme and compares it with the delayed version. Next in Section 3.5, the fading statistic of the TASP system is derived for use in recognizing TAS/MRC systems of similar performance. Section 3.6 then discusses the investigations of TASP performance w.r.t channel power prediction performance and their relevant operating characteristics. Section 3.7 discusses feedback rate reduction strategies in such a system, then Section 3.8 will conclude the chapter.

Figure 3.1: Block diagram showing TAS based upon information from a feedback path from MRC receive antennas and a predictor. To combat degradation due to delay, the predicted antenna index i for the future slot is sent D blocks in advance.

3.2 System model and assumptions

In the following introductory paragraphs, we will first describe the TAS/MRC system and the antenna selection process in an ideal system with no feedback delay constraints. Next we briefly describe the issue of feedback delay in TAS/MRC. In Section 3.2.2 we review the well known Wiener channel predictor. In Section 3.2.3, we derive statistics for the power prediction process, the results of which will be used in Section 3.2.4, to derive TASP/MRC fading distributions.

A block flat fading SIMO (single-input multiple-output) channel with N_t possible transmit and N_r receive antennas is considered in a TAS/MRC system, as shown in Fig. 3.1. The block fading model is appropriate when the coherence time T_c is large enough so that the fading coefficients are assumed to be constant over one block. Based upon an index l , $1 \leq l \leq N_t$, fed back to the transmit end, a single best transmit antenna is selected from N_t candidates for block data transmission. This arrangement is denoted as $(N_t, 1; N_r)$. Diversity reception with maximal ratio combining and coherent demodulation, is employed at the receiver, and a block stationary channel assumed. We adopt the well known PSAM technique that was formulated in [92]. Our PSAM block structure essentially amounts to a SIMO version of the full MIMO beam-forming system [80], which extends PSAM plus diversity [79] (consisting of a single transmit and multiple receive antennas).

As shown in Fig. 3.2, periodic pilot symbols from each antenna are inserted at the beginning of every block of length L_b in a turn-by-turn fashion, so as to enable channel estimation for each transmit-receive pair. The pilot symbols constitute the first N_t entries of the block and aid in channel prediction and estimation. For data transmission, only one active transmit antenna is used for each block. The channel matrix of size $(N_r \times N_t)$ for the k^{th} block is $\mathbf{H}(k)$ with the complex fading coefficients $h_{ij}(k)$, $1 \leq i \leq N_t$, $1 \leq j \leq N_r$ as its entries.

The channels $h_{ij}(k)$ are characterised as spatially independent and identically distributed (i.i.d)¹ complex Gaussian random variables with distribution $\mathcal{CN}(0, \sigma_h^2)$ that follow Jakes model [42], [93] with Doppler spread f_d (assumed equal for all channels). Also they are assumed to change slowly so that multiple training signals can be transmitted without significant frame overhead [36], making antenna selection for block fading a practical solution. Essentially, we will assume that channel coefficients are constant over a block, and temporally correlated between blocks. For each block instant k , the transmitter receives switching information from the receiver: an integer l , denoting the index of the best transmit antenna to be used. The criterion for selection is to pick an antenna which provides the highest power or SNR gain amongst all transmitters. In an ideal case, where there is no feedback delay, the receiver picks the transmit antenna with signal power p_{max} to maximize the post processing SNR at the output of the maximal ratio combiner.

$$p_{max}(k) = \max_{1 \leq i \leq N_t} p_i(k) \quad (3.6)$$

$$p_i(k) = \sum_{j=1}^{N_r} |h_{ij}(k)|^2 \quad (3.7)$$

The index, l , of p_{max} is communicated to the transmitter through a feedback channel. Then for subsequent block k the received signal vector will be expressed as:

$$\mathbf{y}(k; m) = \sqrt{E_s} \mathbf{h}_l(k) x(k; m) + \mathbf{z}(k; m) \quad (3.8)$$

¹The assumption of i.i.d is valid since large antenna separation at the base station is feasible, while for the mobile receiver the separation distance can be made small because of a rich scattering environment.

Figure 3.2: Transmission frame, showing pilot symbols at the start of each transmission block, and delay between pilot transmission and the antenna switching point.

with $k \geq 0$, $N_t \leq m \leq L_b - 1$, and $x(k; m)$ represents a data symbol at time instant $(kL_bT_s + mT_s)$ transmitted from the selected antenna with E_s as the power per data symbol, in the k^{th} block. T_s is the symbol duration in seconds and block length is L_b symbols. $\mathbf{h}_l(k)$ is a $(N_r \times 1)$ channel vector, the l^{th} column of the channel matrix $\mathbf{H}(k)$, and is assumed to be perfectly estimated. $\mathbf{z}(k)$ is the AWGN vector with distribution $\mathcal{CN}(0, \sigma_z^2 \mathbf{I}_{N_r})$ with $\sigma_z^2 = N_0$, the receiver AWGN noise power for a given receive branch. $\mathbf{I}_{N_r}$ is an identity matrix of dimension N_r . Perfect channel estimation is assumed to enable coherent receiver demodulation. A common way to achieve this in practice, used in adaptive systems such as [94], [80], [79], is through noncausal channel smoothing Wiener interpolation filters [92], [95]. We also assume that index l is received error-free at the transmitter.

3.2.1 TAS with feedback delay (TASD)

Even with perfect channel estimates being used for antenna selection, the performance of TAS/MRC will degrade as a result of acting on outdated knowledge at the transmitter. Large delays render feedback information useless, effectively breaking the feedback loop. Delays are commonly caused by error coding/decoding, reverse channel blocking, ARQ, bandwidth restrictions or higher layer protocol processing. Given a $\mathcal{D}$ block delay, the transmitter relies on switching information sent by the

receiver $\mathcal{D}$ blocks earlier. Thus, (3.6) and (3.7) now become:

$$p'_{max}(k) = \max_{1 \leq i \leq N_t} p'_i(k) \quad (3.9)$$

$$p'_i(k) = \sum_{j=1}^{N_r} |h_{ij}(k - \mathcal{D})|^2 \quad (3.10)$$

The channel coefficients at the present block, k , differ from those at block $(k - \mathcal{D})$, with the relationship [96], [40]:

$$h_{ij}(k) = \rho_d \sigma_h^2 h_{ij}(k - \mathcal{D}) + \sqrt{\sigma_h^2 - \rho_d^2} n_{ij}(k) \quad (3.11)$$

where the correlation between the true and delayed or outdated fading gain is given as $\rho_d = \mathbb{E}[h_{ij}(k)h_{ij}^*(k - \tau)] = \sigma_h^2 J_0(2\pi f_d \tau)$ [42]. $\mathbb{E}[\cdot]$ is the expectation operator, $J_0(\cdot)$ is the zeroth order Bessel function of the first kind and τ is the feedback time delay length given by $\tau = \mathcal{D}L_b T_s$. $n_{ij}(k)$ is AWGN with zero mean and unit variance. Delay causes channel mismatch, hence performance degradation occurs as a result of poor antenna selection. In addition, since $f_d \propto v f_c$, where v is the vehicle speed and f_c is the operating carrier frequency, the normalised delay term $f_d \tau$ becomes larger for higher vehicle speed and higher carrier frequency. Thus the channel correlation ρ_d is dependent on feedback delay, velocity and carrier. For very small values of normalised feedback delay $f_d \tau \ll 1$, sufficient channel correlation exists so that $p'_{max}(k) \approx p_{max}(k)$. Therefore switching information is relevant, and the BER degradation may be small. For a large delay, ρ_d tends to zero; essentially resulting in the transmitter acting on completely outdated channel knowledge, causing incorrect antenna selection. Intuitively one would then expect the TASD/MRC system to behave like an open loop system: simple MRC with one transmit antenna. In general, despite using a clear estimate for antenna selection, delay degrades the BER of TAS/MRC, an effect that will be explored in Section 3.2.4 and beyond.

3.2.2 Wiener channel prediction

To combat delay, we aim to predict channel coefficients ahead of time, to improve antenna selection performance. The channel prediction filter here is strictly causal in

nature, and to enable prediction, channel estimates for each block are obtained using PSAM [92]. For a block of data, channel estimation is carried out independently for all channels, with the entries of channel matrix $\mathbf{H}(k)$ estimated as:

$$\tilde{h}_{ij}(k) = h_{ij}(k) + v_{ij}(k) \quad (3.12)$$

where $\tilde{h}_{ij}(k)$ is the channel estimate while $v_{ij}(k)$ is the AWGN channel estimation error with distribution $\mathcal{CN}(0, \sigma_v^2)$, with $\sigma_v^2 = N_0/E_p$, E_p being the power of the pilot symbol. Thus the variance of the estimated channel amplitude is given by $\sigma_h^2 = \sigma_h^2 + \sigma_v^2$, since it is assumed that $h_{ij}(k)$ and $v_{ij}(k)$ are statistically independent. Note that the same estimates will aid in channel smoothing for perfect receiver demodulation as assumed before, therefore $\tilde{h}_{ij}(k) = h_{ij}(k)$, for receiver symbol detection. For channel prediction, a vector of delayed estimates obtained from (3.12) are used. E_p/N_0 here is assumed to be at least equal to the average receive SNR, $\bar{\gamma}$. In practice, the pilot power could be made large enough to achieve channel smoothing for demodulation purposes [97].

The Wiener-Hopf equation for the $\mathcal{D}$ block ahead predicted channel is $\hat{h}_{ij}(k + \mathcal{D}) = \mathbf{w}_{opt}^H \tilde{\mathbf{h}}_{ij}$, where $\tilde{\mathbf{h}}_{ij}$ is the complex vector of estimated fading amplitudes, corresponding to prediction length L given by $\tilde{\mathbf{h}}_{ij} = [\tilde{h}_{ij}(k), \tilde{h}_{ij}(k-1), \dots, \tilde{h}_{ij}(k-(L-1))]^T$ and $\mathbf{w}_{opt}$ is the optimal complex coefficient vector given by $\mathbf{w}_{opt} = [\mathbf{R}]^{-1} \mathbf{r}$ where $[\mathbf{R}]_{\varphi\vartheta} = \sigma_h^2 J_0(2\pi f_d |\varphi - \vartheta| L_b T_s) + \sigma_v^2 \delta(\varphi - \vartheta)$, and $\mathbf{r}_\varphi = \sigma_h^2 J_0(2\pi f_d |D + \varphi - 1| L_b T_s)$ $\varphi, \vartheta = 1, 2, \dots, L$. When $\sigma_h^2 = 1$, the normalised correlation coefficient between the true and the predicted channel is given as $\hat{\rho}_{h\hat{h}} = \sqrt{\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}}$ which is bounded: $0 \leq \hat{\rho}_{h\hat{h}} \leq 1$, a value of one meaning perfect prediction and zero meaning no correlation between predicted and actual channel. The prediction error for any channel is given by $\epsilon_c(k + \mathcal{D}) = h(k + \mathcal{D}) - \hat{h}(k + \mathcal{D})$ with the mean square error (MSE) being minimized when the optimal coefficient vector $\mathbf{w} = \mathbf{w}_{opt}$ is used. Then the MSE is given by $\min_{\mathbf{w}_{opt}} \sigma_{\epsilon_c}^2 = \sigma_h^2 - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$ and is bounded by $0 \leq \sigma_{\epsilon_c}^2 \leq \sigma_h^2$. Then the true channel can be written as:

$$h_{ij}(k + \mathcal{D}) = \hat{h}_{ij}(k + \mathcal{D}) + \sqrt{\sigma_h^2 - \hat{\rho}_{h\hat{h}}^2} n_{ij}(k + \mathcal{D}) \quad (3.13)$$

where $n_{ij}(k + D)$ is AWGN with zero mean and unit variance. The predicted channel amplitude is also a Gaussian random variable with variance $\sigma_h^2 = \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$.

3.2.3 Prediction of channel power

In this section we develop a framework for the MRC power predictor, to determine fading statistics for predicted power and post-processing SNR at the receiver. This will be useful in deriving fading statistics for the TASP/MRC system discussed in Section 3.2.4. Based on the $\mathcal{D}$ block ahead predicted channel coefficients, at instant k , the receiver computes the corresponding predicted channel power for each antenna and selects a transmit antenna, l , corresponding to the maximum power gain. Mathematically this is:

$$\hat{p}_i(k + \mathcal{D}) = \sum_{j=1}^{N_r} \left| \hat{h}_{ij}(k + \mathcal{D}) \right|^2 \quad (3.14)$$

$$\hat{p}_{max}(k + \mathcal{D}) = \max_{1 \leq i \leq N_t} \hat{p}_i(k + \mathcal{D}) \quad (3.15)$$

Examining the statistics of the power predictor will naturally lead us to the derivation of the fading PDF of the post-processing SNR, given in Section 3.2.4. Note that the average value of error $\epsilon_p(k + \mathcal{D}) = p_i(k + \mathcal{D}) - \hat{p}_i(k + \mathcal{D})$, while predicting based on power, is non-zero as was the case with simple channel amplitude prediction; it is known to be biased [77], [78]. It can be seen that $\mathbb{E}[\epsilon_p(k + \mathcal{D})] = N_r(\sigma_h^2 - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r})$. Thus the power prediction is itself biased. The MSE for the biased predictor is $\sigma_{\epsilon_p}^2 = \mathbb{E}[|p_i(k + \mathcal{D}) - \hat{p}_i(k + \mathcal{D})|^2]$. With use of the average channel power gain $\mathbb{E}[p_i(k + \mathcal{D})] = N_r \sigma_h^2$ and the average predicted power gain $\mathbb{E}[\hat{p}_i(k + \mathcal{D})] = N_r \sigma_h^2 = N_r \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$, along with $\mathbb{E}[h_i(k + \mathcal{D}) \hat{h}_i^*(k + \mathcal{D})] = \hat{\rho}_{\hat{h}\hat{h}}$, and identity in (A.5), the value of the MSE for the biased power predictor can now be determined, after deriving the following:

$$\mathbb{E}[p_i(k + \mathcal{D})^2] = N_r(N_r + 1)\sigma_h^4 \quad (3.16)$$

$$\mathbb{E}[\hat{p}_i(k + \mathcal{D})^2] = N_r(N_r + 1)\sigma_h^4 \quad (3.17)$$

$$\mathbb{E}[p_i(k + \mathcal{D})\hat{p}_i(k + \mathcal{D})] = N_r^2 \sigma_h^2 \sigma_h^2 + N_r |\hat{\rho}_{\hat{h}\hat{h}}|^4 \quad (3.18)$$

The above equations are also useful in determining the normalised correlation coefficient ρ_p between the predicted and actual SNR, which is described in Section 3.2.4.

For simplicity, we shall assume $\sigma_h^2 = 1$ for the rest of the calculations. Then the instantaneous post-processing SNR $\gamma_i(k + \mathcal{D})$ for any transmit antenna is:

$$\gamma_i(k + \mathcal{D}) = \frac{E_s}{N_0} p_i(k + \mathcal{D}) \quad (3.19)$$

and the corresponding predicted SNR gain is,

$$\hat{\gamma}_i(k + \mathcal{D}) = \frac{E_s}{N_0} \hat{p}_i(k + \mathcal{D}) \quad (3.20)$$

with their means as:

$$\bar{\gamma}_i = \mathbb{E}[\gamma_i(k + \mathcal{D})] = \frac{E_s}{N_o} N_r \sigma_h^2 = \frac{E_s}{N_o} N_r \quad (3.21)$$

$$\bar{\hat{\gamma}}_i = \mathbb{E}[\hat{\gamma}_i(k + \mathcal{D})] = \frac{E_s}{N_o} N_r \sigma_h^2 = \frac{E_s}{N_o} N_r \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r} \quad (3.22)$$

Both are gamma distributed with their PDF: $\gamma_i \sim G(N_r, \bar{\gamma})$ and $\hat{\gamma}_i \sim G(N_r, \bar{\hat{\gamma}})$ where N_r is the shape factor, $\bar{\gamma} = \frac{E_s}{N_o}$ and $\bar{\hat{\gamma}} = \frac{E_s}{N_o} \sigma_h^2$ the scale factors of the gamma distributions, also equal to the average SNR per symbol for the true and predicted channel respectively. For an error rate analysis it is required to determine the fading distribution of the true maximum channel SNR $\gamma_{max}(k + \mathcal{D})$, which is the maximum of the SNRs $\gamma_i(k + \mathcal{D})$. Similarly the predicted maximum SNR is $\hat{\gamma}_{max}(k + \mathcal{D})$.

3.2.4 Fading distribution of a TAS/MRC with prediction (TASP)

The TASP/MRC symbol error probability, outage probability and fading statistics, can be derived by first obtaining an expression for the PDF $f_{\gamma_{max}}(\gamma)$, which is the distribution of the random variable γ_{max} . We will omit time indices because of the assumption of a stationary random process. The PDF of $f_{\gamma_{max}}(\gamma)$ is:

$$f_{\gamma_{max}}(\gamma) = \int_0^\infty f(\gamma|\hat{\gamma}) f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} = \int_0^\infty \frac{f(\gamma, \hat{\gamma})}{f_{\hat{\gamma}}(\hat{\gamma})} f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} \quad (3.23)$$

$f(\gamma, \hat{\gamma})$ being the joint distribution of γ and $\hat{\gamma}$, $f(\gamma|\hat{\gamma})$ the PDF of γ conditioned on $\hat{\gamma}$, and $f_{\hat{\gamma}}(\hat{\gamma})$ the PDF of the predicted power given by (we have dropped index i since

the distribution is the same for all antennas):

$$f_{\hat{\gamma}}(\hat{\gamma}) = \frac{\hat{\gamma}^{N_r-1}}{\bar{\gamma}^{N_r} (N_r - 1)!} \exp\left(-\frac{\hat{\gamma}}{\bar{\gamma}}\right) \quad (3.24)$$

$f_{\hat{\gamma}_{max}}(\hat{\gamma})$ is the PDF of $\hat{\gamma}_{max}$. To derive $f_{\hat{\gamma}_{max}}(\hat{\gamma})$, we determine the CDF of $\hat{\gamma}$:

$$F_{\hat{\gamma}}(\hat{\gamma}) = 1 - \exp(-\hat{\gamma}/\bar{\gamma}) \sum_{m=0}^{N_r-1} (1/m!) (\hat{\gamma}/\bar{\gamma})^m \quad (3.25)$$

Making use of order statistics [29], the PDF $f_{\hat{\gamma}_{max}}(\hat{\gamma})$ can be calculated as follows:

$$f_{\hat{\gamma}_{max}}(\hat{\gamma}) = N_t f(\hat{\gamma}) [F(\hat{\gamma})]^{N_t-1} \quad (3.26)$$

Both γ and $\hat{\gamma}$ are mutually correlated, determined through a bi-variate gamma distribution $G' \sim (N_r, \bar{\gamma}, \bar{\hat{\gamma}}, \rho_p)$ as follows:

$$f_{\gamma, \hat{\gamma}}(\gamma, \hat{\gamma}) = \frac{f(\gamma) f(\hat{\gamma}) (N_r - 1)!}{1 - \rho_p} \left(\frac{\rho_p \gamma \hat{\gamma}}{\bar{\gamma} \bar{\hat{\gamma}}} \right)^{-(N_r-1)/2} I_{N_r-1} \left(\frac{2\sqrt{\rho_p}}{(1 - \rho_p)} \sqrt{\frac{\gamma \hat{\gamma}}{\bar{\gamma} \bar{\hat{\gamma}}}} \right) \cdot \exp \left(-\frac{\rho_p}{1 - \rho_p} \left(\frac{\gamma}{\bar{\gamma}} + \frac{\hat{\gamma}}{\bar{\hat{\gamma}}} \right) \right) \quad (3.27)$$

$I_{N_r-1}(\cdot)$ is the modified Bessel function of the first kind with order $(N_r - 1)$ and ρ_p is the power correlation coefficient of the true and predicted fading SNRs. To proceed further, we express $\bar{\hat{\gamma}} = r\bar{\gamma}$ ($r = \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$), as seen from (3.21) and (3.22)) where $0 \leq r \leq 1$. Utilising $f_{\hat{\gamma}}(\hat{\gamma})$, the PDF of $\hat{\gamma}$, we can write $f(\gamma|\hat{\gamma})$ as:

$$f_{(\gamma|\hat{\gamma})}(\gamma|\hat{\gamma}) = \left(\frac{r\gamma}{\rho_p \hat{\gamma}} \right)^{(N_r-1)/2} I_{N_r-1} \left(\frac{2\sqrt{\rho_p \gamma \hat{\gamma}}}{(1 - \rho_p) \sqrt{r\bar{\gamma}}} \right) \cdot \exp \left(-\frac{\rho_p \hat{\gamma} + r\gamma}{r(1 - \rho_p) \bar{\gamma}} \right) \frac{1}{(1 - \rho_p) \bar{\gamma}} \quad (3.28)$$

Furthermore, to determine ρ_p we know that,

$$\rho_p = Cov(\gamma, \hat{\gamma}) [Var(\gamma) Var(\hat{\gamma})]^{-0.5} \quad (3.29)$$

Using (3.16-3.18) and (3.21-3.22), the value of ρ_p can now be determined. The power correlation coefficient can be shown to equal the square of the channel correlation coefficient so that $\rho_p = \rho_{hh}^2$ (note that this is because the coefficient vector $\mathbf{w}_{opt}$ is also

the underlying factor used in the channel power prediction calculation). Note that in this case also $r = \rho_p$. Examining (3.28) more closely, we see that as ρ_p approaches zero (meaning very poor prediction), the term in the exponential becomes independent of $\hat{\gamma}$. By re-factorising, $f_{(\gamma|\hat{\gamma})}(\gamma|\hat{\gamma})$ is found to be independent of $\hat{\gamma}$ and equates to $f_{\gamma}(\gamma)$. This means that the true instantaneous SNR γ will be independent of the predicted SNR $\hat{\gamma}$, thereby rendering TASP ineffective with the system behaving like a simple MRC system with a single transmit antenna $(1, 1; N_r)$.

By use of multi-nominal and binomial theorems, the power of the polynomial in equation (3.26) can be expanded for use with (3.28). Finally, using integration results from [98, eqn.(6.643.4)] (See A.3.1 in Appendix A), the final closed form PDF expression of (3.23) is derived as:

$$f_{\gamma_{max}}(\gamma) = \frac{N_t! \gamma^{N_r-1}}{(N_r-1)!} \cdot \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!} \exp\left(-\frac{(i+1)\gamma}{[i(1-\rho_p)+1]\bar{\gamma}}\right) \cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) j! \frac{[(1-\rho_p)\bar{\gamma}]^j}{[[i(1-\rho_p)+1]\bar{\gamma}]^{j+N_r}} L_j^{(N_r-1)}\left(-\frac{\rho_p\gamma}{[i(1-\rho_p)+1](1-\rho_p)\bar{\gamma}}\right) \quad (3.30)$$

where $L_j^{N_r-1}(\cdot)$ is the Laguerre polynomial [99] (see A.2) and $\eta_{N_r}(i, j)$ [98] are the coefficients of $z^j, j = 0, 1, \dots, i(N_r-1)$, in the expansion of $(\sum_{j=0}^{N_r-1} z^j/j!)^i$, obtained as per equations in (A.1). For the TASP case, the fading PDF is derived in a similar manner. However the important difference is that in the TASP case, as per (3.10) we will use a perfectly estimated or noiseless channel coefficient for antenna selection. This can be thought of as a highly smoothed version of the channel estimate, used for generating instantaneous antenna index. Note that for TASP, the PDF of the maximum of SNRs would result in the same form as (3.30), the difference being in the power correlation coefficient ρ_p , which has to be replaced now with $\rho_p = J_0^2(2\pi f_d \tau)$ [42]. Later we will see that power correlation ρ_p in TASP can be made to exceed ρ_d for any channel feedback delay, thus outperforming a TASP system; even one that uses a clear channel estimate for index generation. The advantage of the TASP scheme will be observed in the following sections.

3.3 Error Rate Analysis

The average SER at an average SNR $\bar{\gamma}$ per receive path is found by averaging the probability of symbol error in AWGN over the fading distribution of the TASD/P schemes, in slow fading or in quasi-static scenarios. Exact, or approximate (but accurate), values of SER for coherent demodulation can be found from:

$$\bar{P}_s(\bar{\gamma}) = \int_0^\infty P_s(\gamma) f_{\gamma_{max}}(\gamma) d\gamma \quad (3.31)$$

where $P_s(\gamma) \approx \alpha Q(\sqrt{\beta\gamma})$. α and β are determined by specific constellations and $Q(\cdot)$ is the Q -function [21]. With $\alpha = 1$, $\beta = 2$, the exact BER for BPSK is found, while an approximate SER for rectangular M -QAM (quadrature amplitude modulation) is found with $\alpha = 4(\sqrt{M} - 1)/\sqrt{M}$ and $\beta = 3/(M - 1)$. To evaluate (3.31) we first expand the Laguerre polynomial using (A.2). Next collecting the terms containing exponential and the powers of γ in (3.30) along with the Q -function. We then have the following inner integral to solve:

$$\begin{aligned} J_{N_r}(i, k, \beta) &= \int_0^\infty \alpha Q(\sqrt{\beta\gamma}) \exp\left(-\frac{i+1}{i(1-\rho_p)+1}\right) \left(\frac{\gamma}{\bar{\gamma}}\right)^{k+N_r+1} d\gamma \\ &= \alpha (k + N_r - 1)! \left[\frac{(i(1-\rho_p)+1)(1-\mu_i)}{2(i+1)} \right]^{k+N_r} \\ &\quad \cdot \sum_{l=0}^{l+N_r-1} \frac{(k + N_r - 1 + l)!}{l! (k + N_r - 1)!} \left(\frac{1+\mu_i}{2}\right)^l \end{aligned} \quad (3.32)$$

$$\mu_i = \sqrt{\frac{\beta(i(1-\rho_p)+1)\bar{\gamma}}{2(i+1) + \beta(i(1-\rho_p)+1)\bar{\gamma}}} \quad (3.33)$$

which can be solved using the result of the integral [19, (5A.5)]. Finally eqn. (3.31) is presented as:

$$\begin{aligned} \bar{P}_s(\bar{\gamma}) = & \alpha \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ & \cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) (j + N_r - 1)! \sum_{k=0}^j \frac{\rho_p^k (1 - \rho_p)^{j-k}}{[i(1 - \rho_p) + 1]^j} \left[\frac{(1 - \mu_i)}{2(i + 1)} \right]^{k+N_r} \\ & \cdot \sum_{l=0}^{k+N_r-1} \frac{(k + N_r - 1 + l)!}{l! (k + N_r - 1)!} \left(\frac{1 + \mu_i}{2} \right)^l \end{aligned} \quad (3.34)$$

As a check, we may consider the $(N_t, 1; 1)$ case with $\beta = 2$ for BPSK modulation, and $\rho_p = 1$, implying full prediction or the immediate (no delay) case, then $\eta_1(i, j) = 1$, and the BER in (3.34) becomes:

$$\bar{P}_s(\bar{\gamma}) = N_t! \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \left[\frac{(1 - \sqrt{\bar{\gamma}/(\bar{\gamma} + i + 1)})}{2(i + 1)} \right]$$

This simple case exactly matches [28, eqn.(26)]. Similarly when $\rho_p = 0$, $\bar{P}_s(\bar{\gamma}) = \left[\frac{1 - \sqrt{\bar{\gamma}/(\bar{\gamma} + 1)}}{2} \right]$ which is the trivial case of BER of BPSK in a Rayleigh fading SISO link.

Numerical example

To illustrate the benefits of TASP and further verify the result, let us consider the BER performance of TASP $(2, 1; 2)$ compared to simple TASD. For all numerical and Monte Carlo simulation setups, we shall set realistic values of $T_s = 1e - 06$, $E_p/N_0 = 30$ dB, $f_d = 100$ Hz, $L_b = 1/(100f_dT_s)$, $L = 5$ (5 tap predictor). The setup for the simulations are given in a table below. The correlation coefficient is affected by many system parameters, so that several dimensions of system performance trade-off are possible. The choice of these parameters, and how they impact system performance, will be explored in Section 3.6.

Fig. 3.3 compares the BER performance of BPSK, in the TAS/MRC system, with and without prediction, for small to severe normalised feedback delays² of 0.02, 0.1

²Note that the normalised delay is the product: $f_d\tau = f_dDL_bT_s \propto v f_c DL_bT_s$, which depends on vehicle speed, carrier frequency, block delay, block length and symbol duration.

Channel	Frequency-flat Rayleigh channel
Doppler Spectrum	Jakes Doppler Spectrum $S(f) = \frac{1}{\pi f_d \sqrt{1-(f/f_d)^2}}$
Maximum Doppler frequency	$f_d = 100Hz$
Sample period	$T = 1e - 06$
Block length	$L_b = 1/(100f_dT_s)$
Average path gain	0 dB
Normalised path gain	1
Pilot to noise power ratio	$E_p/N_0 = 30$ dB
Filter taps	5
Block delay	$\mathcal{D} = 1$ to 300
Normalised feedback delay	$f_d\tau = .01$ to 3

Table 3.1: Simulation setup for TAS/MRC with prediction

and 0.2. In order to verify eqn. (3.34) we also plot simulation results. For small $\bar{\gamma}$ and small delay, the performance of both schemes is nearly equal. For a $1e - 3$ BER, compared to TASP, TASP offers a gain of approximately $2dB$ at 0.1 delay, and $3dB$ at 0.2 delay. Similarly for a $1e - 5$ BER, the gains are about $4dB$ for 0.1 and 0.2 feedback delay; while for lower BERs the gains are found to be $7dB$ and $5dB$ respectively. At large delays, as expected, the TASP system behaves like an open loop $(1, 1; 2)$ MRC system: as seen in the graph, TASP tends towards the $(1, 1; 2)$ MRC BER curve, while the TASP system also tends to the $(1, 1; 2)$ MRC BER at a slower rate. These observations also indicate that there is an increase in degradation rate w.r.t delay for the TASP/D case.

The degradation effect is seen more clearly in Fig. 3.4 which plots BER against normalised feedback delay for five different average receive SNRs in a $(2, 1; 2)$ configuration. In comparison to TASP for a delay of 0.02, the performance of the TASP system is almost as good as the system without any feedback delay. For small SNRs and small delays, the performance of TASP and TASP are quite similar, with additive noise dominating performance. TASP degradation begins at around 0.03 and

Figure 3.3: BER performance comparison of BPSK for TAS/MRC (2,1;2) with and without prediction at different delays, showing theoretical and simulation results.

Figure 3.4: Feedback delay tolerance of BER in BPSK of the (2,1;2) scheme for different SNRs of 6, 10, 14, 18, 22 dB with and without prediction, showing the improved tolerance to feedback delay effects gained by the prediction scheme.

higher, while TASP holds BER steady until about 0.1 delay. At higher SNRs of around 22 dB, TASP experiences degradation as early as 0.01 delay, while TASP can tolerate up to about 0.04. For small SNRs, in both TASP/P, degradation is more pronounced at larger feedback delays than at smaller ones. This effect can be explained by looking at the power correlation coefficient ρ_p in the BER equations. For TASP, $\rho_p = J_0^2(2\pi f_d \tau)$ [42], which falls off rapidly for increasing values of τ . By comparison, the power correlation coefficient for the TASP case is equal to $\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$, which is a quadratic function, with shallower slope at smaller values of delay compared to the coefficient in TASP. The comparison of the correlation coefficients for both cases will be explored in Section 3.6.

Fig. 3.5 plots the degradation of BER for different setups of TAS evaluated at a receive SNR $\bar{\gamma}$ of 10 dB, chosen so as to yield a range of BER values. Even small delays starting from 0.02 in the TASP arrangement cause significant deterioration in BER. By contrast the predictive scheme sustains performance almost unchanged out to around 0.1 delay. Relating this to a real system, at a carrier frequency of 900 MHz and walking speed of 1 m/s, a time delay of not more than 6 ms can be tolerated by TASP. TASP, by contrast, can withstand up to 33 ms. Similarly for vehicles moving at 27.7 m/s (100 km/h), TASP tolerates up to 0.72 ms while TASP can tolerate about 3.6 ms of delay.

Fig. 3.5a looks at the influence of increasing N_t keeping N_r fixed and vice-versa at different delays. For increasing N_r , the BER decreases considerably at smaller delays, while at extreme delays, it gets capped to a value corresponding to its open loop MRC configuration. The BER improvement due to increasing N_t is less pronounced for higher delays in the TASP system, while the TASP system can help to harness this gain. It is also noteworthy that, for the TASP case, systems with lower BER requirement are more sensitive to feedback delay. The effects of adding extra diversity are also discussed in the following sections on outage probability and average fading gains.

Fig. 3.5b shows the BER at $\bar{\gamma} = 10$ dB for different systems of equal $N_r + N_t$. Such

Figure 3.5: Feedback delay tolerance of BER in BPSK at $\bar{\gamma} = 10dB$ for different TAS/MRC configurations with and without prediction. (a) Influence of increasing N_r and N_t . (b) Different system arrangements with constrained total number of antennas of $N_r + N_t = 6$.

a plot can aid in choosing the total number of antennas used in a communications system for maximum performance w.r.t BER, at any given delay. For example, with prediction, in the region of $0.01 < f_d\tau < 0.25$, for $\bar{\gamma} = 10 \text{ dB}$, the $(2, 1; 4)$ setup outperforms all other schemes, with the same number of antennas, in terms of BER. For $f_d\tau \geq 0.25$, the non-selective $(1, 1; 5)$ MRC setup performs best. However, when prediction is not employed, $(2, 1; 4)$ outperforms the other non-predictive setups in the region of $0.01 < f_d\tau < 0.11$.

3.4 Outage Probability

The outage probability P_{out} at any SNR $\bar{\gamma}$ is defined as the probability that the instantaneous capacity $C_{\bar{\gamma}}$ (*bits/s/Hz*) is less than a given capacity target R and is given as $P_{out} = Pr\{C_{\bar{\gamma}} < R\}$ [13], which, in a block fading channel essentially is:

$$\begin{aligned} P_{out}(\bar{\gamma}, R) &= Pr\{\log_2(1 + \gamma_{max}) < R\} \\ &= Pr\{\gamma_{max} < (2^R - 1)\} \end{aligned} \quad (3.35)$$

Setting $z = (2^R - 1)/\bar{\gamma}$, outage probability is given as:

$$\begin{aligned} P_{out}(z) &= \int_0^z f_{\gamma_{max}}(\gamma) d\gamma \\ &= \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) \\ &\quad \cdot (N_r + j - 1)! \sum_{k=0}^j \binom{j}{k} \frac{\rho_p^k (1 - \rho_p)^{j-k}}{[(i(1 - \rho_p) + 1)]^j} \\ &\quad \cdot \left\{ \frac{1}{(i + 1)^{N_r+k}} - \exp\left(-\frac{(i + 1)z}{[(i(1 - \rho_p) + 1)]}\right) \right. \\ &\quad \cdot \left. \sum_{l=0}^{k+N_r-1} \frac{[1/(N_r + k - 1 - l)!] z^{N_r+k-l-1}}{(i + 1)^{l+1} [i(1 - \rho_y) + 1]^{N_r+k-l-1}} \right\} \end{aligned} \quad (3.36)$$

where the integral in (A.4.1) has been utilized. As a check, consider the $(N_t, 1; 1)$ case: here $\eta_1(i, j) = 1$. Hence,

$$P_{out}(z) = N_t \sum_{i=0}^{N_t-1} \frac{(-1)^i}{(i + 1)} \binom{N_t - 1}{i} \left[1 - \exp\left(-\frac{(i + 1)z}{[i(1 - \rho_p) + 1]}\right) \right] \quad (3.37)$$

Using $\frac{N_t}{i+1} \binom{N_t-1}{i} = \binom{N_t}{i+1}$ and $\sum_0^{N_t} (-1)^i \binom{N_t}{i} = 0$, when $\rho_p = 1$ for full prediction, (3.37) reduces to $P_{out}(z) = (1 - \exp(-z))^{N_t}$ which, for high SNR, has a diversity order of N_t . Similarly, when $\rho_p = 0$, $P_{out}(z)$ becomes $1 - \exp(-z)$ which is the simple SISO case with diversity reducing to unity. Both cases help to verify eqn. (3.36), and to explore further Fig. (3.6) plots the outage probability for different setups of TASD/P, at $\bar{\gamma} = 10 \text{ dB}$ with $R = 1 \text{ bit/s/Hz}$. These reveal how the systems behave for different degrees of fading correlation and number of antennas. Simulation results are also given to verify the analytical expressions.

Fig. 3.6a shows outage plots for $(N_t, 1; 2)$, and the dependency of increasing N_t having 2 receive antennas for TASP/D. For TASD, increasing N_t decreases P_{out} considerably for smaller delays, however for $N_t = 4$, degradation begins earlier than 0.01 delay. At larger delays, the effect of increasing N_t offers little improvement as seen around $f_d\tau = 0.1$. The TASP setup however tolerates delays up to around $f_d\tau = 0.05$, and behaves in a similar fashion to TASD for greater delays; while at extreme delays, as expected, P_{out} tends to a limit corresponding to the open loop MRC configuration of $(1, 1; 2)$ for both cases. It is also seen that for high order schemes, the degradation rate is higher for higher delays; whereby $(4, 1; 2)$ degrades faster than a $(3, 1; 2)$ scheme. Fig. 3.6b is plotted for increasing N_r as $(2, 1; N_r)$ and behaves in a similar fashion except being more delay tolerant than $(N_t, 1; 2)$. For increasing values of N_r , P_{out} decreases considerably at smaller delays. Similar to the previous case, for extreme delays, each setup tends to its corresponding $(1, 1; N_r)$ MRC configuration. The setups here have the same number of antennas as in Fig. 3.6a, with MRC being the dominant factor in providing increased array gains. They also having the advantage of decaying at a slower rate as delay increases. Thus a $(2, 1; 4)$ setup may be better than a $(4, 1; 2)$ setup.

3.5 Fading statistics of the output SNR

We evaluate and compare the different TAS configurations based on moments of the signal amplitude γ_{max} . Closed form expressions for average SNR gain, signal

Figure 3.6: Outage probability comparisons at $\bar{\gamma} = 10 \text{ dB}$ SNR between TASP/D schemes with fixed N_r and increasing N_t and vice-versa, at $R = 1 \text{ bit/s/Hz}$, showing theoretical and simulation results.

variance and amount of fading (*AoF*) are derived. Such measures could find use in comparing system performance of the existing TASP/MRC scheme with those that use different fading models, and possibly across other MIMO architectures, where closed form solutions of common performance metrics such as BER and outage probability may not be readily available. The average SNR gain can be used to quantify and compare across different TAS/MRC schemes. Specifically, here we compare the gains of different TAS/MRC schemes in the presence and absence of the predictor.

The average SNR gain of the TASP $(N_t, 1; N_r)$ system is denoted as $\bar{\gamma}_{TAS} = \mathbb{E}[\gamma_{max}]$ and is given by:

$$\mathbb{E}[\gamma_{max}] = \int_0^\infty \gamma f_{\gamma_{max}}(\gamma) d\gamma \quad (3.38)$$

For a $(N_t, 1; N_r)$ scheme, using (3.30),

$$\begin{aligned} \mathbb{E}[\gamma_{max}] &= \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ &\cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) j! \frac{[(1 - \rho_p)\bar{\gamma}]^j}{[i(1 - \rho_p) + 1]\bar{\gamma}]^{j+N_r}} \\ &\cdot \int_0^\infty \gamma^{N_r} \exp\left(-\frac{(i+1)}{[i(1 - \rho_p) + 1]\bar{\gamma}}\right) \cdot L_j^{(N_r-1)}\left(-\frac{[\rho_p/(1 - \rho_p)]\gamma}{[i(1 - \rho_p) + 1]\bar{\gamma}}\right) d\gamma \end{aligned} \quad (3.39)$$

then expanding the Laguerre polynomial, and integrating with the help of (A.4.1), we have:

$$\begin{aligned} \mathbb{E}[\gamma_{max}] &= \frac{\bar{\gamma} N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \sum_{j=0}^{i(N_r-1)} \frac{\eta_{N_r}(i, j) j! (j + N_r - 1)!}{[i(1 - \rho_p) + 1]^{(j-1)}} \\ &\cdot \sum_{k=0}^j \frac{\rho_p^k (1 - \rho_p)^{j-k} (k + N_r)}{(j - k)! k! [(i + 1)]^{N_r+k+1}} \end{aligned} \quad (3.40)$$

Fig. 3.8 plots the gains of the predictor and delayed scheme for a fixed normalised delay of 0.25, verified by simulation results which are also plotted. Note that for the delayed scheme, the diversity benefits from adding an additional transmit antenna diminish as delay gets larger. Prediction is therefore crucial in maintaining transmit diversity. When prediction is not used, it is beneficial to invest in receive diversity. For larger N_r , the gain increase from adding an extra transmit antenna is less.

Figure 3.7: Comparison of normalised average SNR gain $\bar{\gamma}_{TAS}/\bar{\gamma}$ as a function of number of receive and transmit antennas at specific feedback delays of $f_d\tau = 0.25$ with and without prediction. Theoretical results are verified by simulations.

Figure 3.8: Normalised average SNR gains $\bar{\gamma}_{TAS}/\bar{\gamma}$ of different arrangements with and without prediction for a wide range of delays.

Fig. 3.8 compares post-processing SNR gains achieved by different TASP and TASP configurations for a wide range of delays. As before both TASP and TASP show a decrease in gains as delay increases. Depending upon the amount of gain required at a given delay, the predictor can attempt to keep gain constant. This has a direct bearing on the amount of power put into the pilot symbols. In all graphs, increasing N_r , keeping N_t fixed increases the gain, but not linearly. The gain in going from one receive antenna to two is greater than from going from two to three: in general increasing N_r gives diminishing returns of post-processing SNR gain. Looking across the graphs, at given delay, any gain improvement obtained by increasing N_t (keeping N_r constant) is less than from increasing N_r (keeping N_t constant). Depending on the required gain, a system designer can choose the most appropriate configuration based upon available resources. For example, as seen from Fig. 3.8, at a delay of 0.1, it would be wiser to invest in a (2, 1; 4) configuration with prediction than to use a (4, 1; 3) without prediction under the given conditions, since the (2, 1; 4) is better

protected through MRC than $(4, 1; 3)$.

Although this performance criterion is useful in comparing different TAS schemes, it does not capture all the diversity benefits since the fading fluctuations are also dependent on the second order statistics, namely the signal amplitude variance, or the second moment of the output SNR. The second moment of the random variable γ_{max} can be similarly obtained as:

$$\begin{aligned} \mathbb{E}[\gamma_{max}^2] &= \frac{\bar{\gamma}^2 N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ &\quad \cdot \sum_{j=0}^{i(N_r-1)} \frac{\eta_{N_r}(i, j) j! (j + N_r - 1)!}{[i(1 - \rho_p) + 1]^{j-2}} \\ &\quad \cdot \sum_{k=0}^j \frac{\rho_p^k (1 - \rho_p)^{j-k} (k + N_r)(k + N_r + 1)}{(j - k)! k! [(i + 1)]^{N_r+k+2}} \end{aligned} \quad (3.41)$$

The amount of fading (*AoF*), associated with the fading PDF of strength γ_{max} , is derived as:

$$AoF = \frac{var[\gamma_{max}]}{\mathbb{E}[\gamma_{max}]^2} = \frac{\mathbb{E}[\gamma_{max}^2] - \mathbb{E}[\gamma_{max}]^2}{\mathbb{E}[\gamma_{max}]^2} \quad (3.42)$$

This performance measure was introduced by Charash [100] as a measure of fading severity and is independent of the average SNR per symbol $\bar{\gamma}$. Thus, two systems with the same *AoF* may be considered to fade similarly. Win and Winters [101] used the square root of the ratio of the variance of the combined output SNR to its squared mean, to assess the effectiveness of a hybrid diversity combining scheme in the presence of Rayleigh fading. This performance measure is also equal to the square-root of the *AoF*. For example, in comparing the fading severity of $(2, 1; 4)$ and $(3, 1; 3)$ under no delay, *AoF* evaluated as per (3.42), is approximately -8 dB for both, and one may expect them to have a similar BER performance. As a check, Fig. 3.5a verifies this, and both schemes have nearly equal BER for small delays at $\bar{\gamma} = 10$ dB. However, *AoF* is not always a good indicator of BER performance (although may be better than average SNR). Higher order statistics of the fading SNR can also be derived, and used for further system analysis and design.

Figure 3.9: (a) Influence of filter tap spacing and E_p/N_0 at normalised feedback delays of 0.1, 0.2 and 0.4 on power correlation with a five tap filter (b) Influence of predictor length and E_p/N_0 on power correlation at fixed delays of 0.02 and 0.1 and $L_b = 100$.

3.6 Investigation of system parameters

3.6.1 Influence of system parameters on channel correlation

The correlation coefficient which governs the BER of the TASP/MRC system, generally depends on a number of parameters such as FIR (finite impulse response) predictor length, pilot SNR, symbol rate, channel sampling frequency (location of pilot symbol), training length, Doppler frequency and feedback delay. At higher bit-rates, to limit the number of filter coefficients, it is beneficial to reduce pilot insertion frequency. This sub-sampling frequency (SSF) is usually kept at multiples of the Doppler frequency ($\text{SSF} \geq 2f_d$), satisfying the sampling theorem. Having a relatively high over-sampling rate may mean clearer channel estimates but also poorer long range prediction with a fixed filter order, and vice-versa. We shall look at the influence of these parameters one by one.

3.6.1.1 Influence of E_p/N_0 and SSF

As mentioned, our simulations and numerical evaluations use a symbol period of $T = 1\mu\text{s}$, with Doppler of 100 Hz. To find the optimal sampling or adaptation rate in the presence of noise, we must determine the optimal delay spacing at any given E_p/N_0 and fixed filter length. We fix the filter length to a nominal value of $L = 5$ in our simulations, since having a small filter length reduces the computational burden of receiver end processing; it also reduces errors in estimation of the coefficients themselves. Next we plot channel correlation w.r.t. delay spacing for different E_p/N_0 values and find the delay spacing at which maximum correlation occurs, as shown in Fig. 3.9a. For small prediction ranges, we could use a low sampling rate, since sufficient channel correlation exists over longer periods. For larger prediction ranges, the adaptation rate varies with the SNR as shown. For lower SNR it is beneficial to keep a larger block length, and vice-versa. Note however, no prior noise smoothing has been done and it is possible that any out-of-band noise gets amplified, decreasing correlation. Thus higher order filters can efficiently perform better prediction by

suppressing noise. Also increasing the filter tap length reduces the optimal sampling frequency, since more predictor coefficients can provide better long term prediction.

3.6.1.2 Influence of E_p/N_0 and predictor filter length

At a given $\bar{\gamma}$, normalised feedback delay and filter delay spacing, we can determine the dependency of the correlation and filter length with increasing E_p/N_0 . For smaller delays, shorter training and prediction-filter lengths are able to assure sufficient correlation. However at increased delays, greater training length offers more correlation gain than increasing filter length, since clearer channel estimates offer greater prediction performance. This is shown in Fig. 3.9b where two sets of curves for normalised delays of 0.02 and 0.1 are plotted with different E_p/N_0 and increasing filter lengths. We can see that, for a larger delay, the effect of changing pilot power is more pronounced. Maintaining the same parameters, and improving the E_p/N_0 , naturally increases correlation again because of better channel estimates. For small filter lengths like $L = 5$, the absolute change in correlation value upon increasing pilot power is greater at larger delays than at small delays. This trend is also seen for higher order filters, but reducing slowly with increasing L . These results and observations corroborate findings in [74, 77]. Depending on the filter order and the amount of power of the pilot, the predictor can always be made to offer greater correlation. These system parameters influence the correlation for the TASP system which increases or decreases at a fixed delay. However for better prediction at greater delays, the block length may have to be increased, which may in turn account for channel variability and estimation errors during demodulation at the receiver.

3.6.1.3 Choice of block length (L_b)

As noted above the choice of block length or pilot spacing will also be determined by the capabilities of the noise smoother and interpolation that happen during channel estimation at the receiver. For Rayleigh flat fading, we need to estimate a single tap for each of the $(N_r \times N_t)$ channels. In general, for a continuous flat fading channel,

the estimation error will be contributed by the Doppler variance and noise [102]. For quasi-static or block fading, if we neglect the noise introduced by Doppler variance, then the MSE or the variance of the estimated channel will be $N_0/(L_t E_s)$ where L_t is the number of training symbols for each transmit antenna (where each is transmitted round-robin), over the entire frame. As a rule of thumb, pilot symbols are inserted every $1/(100 f_d)$ s to ensure that the channel changes negligibly over this period. As seen from Fig. 3.9a when the prediction distance ($f_d \tau$) is relatively small, for example, $f_d \tau = 0.1$, this condition can be easily met, without significantly lowering the correlation coefficient or predictor performance. Fig. 3.10 shows the comparison of correlation coefficients for the delayed and predicted cases with $L_b=100$.

3.6.2 Operating point of TASP and trade-off point with TASP

Depending on threshold output BER, BER_T for a given configuration, and knowing the normalised delay that must be tolerated, the graphs presented in the previous subsection may be used to first determine the threshold correlation, and from that, the required filter order and E_p/N_0 . This is worst-case design: when delay is less, BER will naturally improve. It may happen that TASP performs well enough to achieve the target BER at that delay, in which case prediction is not necessary. Otherwise, TASP can be chosen to improve performance. It is interesting to determine the point at which TASP becomes advantageous over TASP, through setting $\rho_p = \rho_d$ (i.e when the correlation of both schemes are equal). A few such intersecting points can be seen by following the TASP curve in Fig. 3.10.

To explore further, Fig. 3.11 plots the equal-correlation points for different delays against pilot SNR with varying filter order. Thus, knowing the feedback delay of a given system, it is possible to determine whether TASP of a given filter order outperforms TASP. If one wants to improve system performance beyond this point (i.e. in terms of increasing correlation or BER) then TASP should be used. Thus it can be seen that even when TASP has a clear, but outdated, estimate available, TASP can perform better by predicting into the future from noisy channel estimates.

Figure 3.10: Correlation coefficient comparison for TASP, and TASP against feedback delay for E_p/N_0 of 15 dB and 25 dB with two different filter lengths of 5 and 10.

Figure 3.11: Equal correlation plot for TASP and TASP, where at all points $\rho_p = \rho_d$ with the corresponding E_p/N_0 (for TASP case) against delay for various L . The dashed line corresponds to the correlation value at any delay, read off from the RH axis.

Figure 3.12: Effect of increasing E_p/N_0 for TASP at a filter length of $L = 5$ on BER performance at $\bar{\gamma} = 10$ dB for two different configurations. The solid black curves represent TASP at different $E_p/N_0 = 15, 20, 25, 30$ dB from top to bottom, in two example systems.

The setups of Fig. 3.12 further highlight design tradeoffs, for $(3, 1; 2)$ and $(2, 1; 4)$ arrangements respectively: they exhibit crossover points between the TASP and TASD curves for each arrangement. It is at these points that the performance of both TASP and TASD are equal. For example at a delay of 0.08 and $E_p/N_0 = 15$ dB (top solid curves), the BER of TASP and TASD equate. This can be seen in both $(3, 1; 2)$ and $(2, 1; 4)$, which provide BERs of 5×10^{-5} and 3×10^{-7} respectively. From Fig. 3.11 we can see that both schemes offer a correlation of almost 0.9 at the given delay, and that TASP requires a 5 tap filter, found from the intersection of $E_p/N_0 = 15$ dB and $f_d\tau = 0.08$ which lies between the $L = 4$ and $L = 5$ curves. Note that, if better BER is required, further improvement is possible with TASP through increasing L or E_p/N_0 . However from Fig. 3.12, we note that at extremes of very small ($< .02$) and very large normalised delays (≥ 0.4), the bundle of TASP BER curves approach asymptotes. This implies that in those extreme regions, the effect of increases in E_p/N_0 will be minimal.

3.6.3 Computational complexity

As described, we choose $L = 5$, with a relatively high pilot power to noise power ratio, to keep the complexity of obtaining predicted channel values for each transmit-receive link small. Since prediction is strictly causal in nature, a L step prediction for future block and channel estimate (in practice) for the present block demodulation have to be calculated separately. These are similar processes, and the quality of the MMSE channel estimate will be far better than the predicted coefficients, since the channel estimator's performance is not affected by delay and secondly, Wiener smoothing can be employed to improve quality. Compared to other methods, the additional complexity of the TAS prediction method comes from predicting the coefficients for each transmit-receive pair. In practise, the first task is in obtaining an inverse of the channel coefficient matrix $\mathbf{R}$, (computed depending on L) to obtain the optimal complex coefficient vector $\mathbf{w}_{opt} = [\mathbf{R}]^{-1}\mathbf{r}$. Since essentially $\mathbf{R}$ is a Topelitz matrix, the solution can be computed efficiently using the well known Levinson-Durbin recursion algorithm. The system may pre-compute this off-line for a range of expected feedback delays, if the channel characteristics are not changing so quickly. Secondly, in computing the channel coefficients themselves through $\hat{h}_{ij} = \mathbf{w}_{opt}^H \tilde{\mathbf{h}}_{ij}$. This requires L multiplications per transmit-receive pair, totalling LN_rN_t operations, per block. Thus it is seen that it is beneficial to maintain a small L as this helps keep the complexity low in both cases. Based on the expected delay, the receiver may also periodically monitor the channel correlation between the predictive and non-predictive approaches in order to make an intelligent decision in choosing the appropriate method.

3.7 Feedback rate reduction

In this section we analyse the impact of reduced feedback rate reduction on a $(N_t, 1; N_r)$ system where, due to feedback rate constraints, transmit antenna switching is done at a lower rate (instead of switching at the frame rate). The delay τ between the receiver outputting predicted CSI and the transmitter obtaining this to select

Figure 3.13: Frame structure with Pilot Block and Data and feedback link

the best antenna, was assumed to be a multiple of the frame length, so that $\tau = DL_bT_s$, where T is the symbol period. Furthermore, in our analysis we consider different feedback update rates $f = 1/(mL_bT_s)$, where m is an integer. Fig 3.13 shows this schedule. Here $f = 1/(mL_bT_s)$ $m = 1, 2, 3, \dots$. Reduction of feedback rates are more useful in correlated channels where even at reduced switching rates, the system still benefits from channel correlation for the time duration when antenna switching does not take place, thus performing better than in uncorrelated channels. The system can be looked upon as operating as a full TASP/D system for $1/m$ percent of the time and as partial TAS or simple MRC otherwise, depending on the temporal correlation. Partial TASP/D refers to a system when antenna switching takes place at only intervals of mL_bT_s s, meaning adaptation rate is slower than block rate. We will show that prediction can also improve performance in feedback rate constrained systems. We will analyse feedback rate reduction both with and without delay.

3.7.1 Reduced feedback in an uncorrelated channel

If the delay in feedback is zero, and the channel is uncorrelated, the system acts as an ideal TAS/MRC scheme for $1/m$ fraction of the time and as a simple MRC for the rest of the time. Using 3.34, BER for this system will be:

$$\bar{P}_s'(\bar{\gamma}, \rho_1, \rho_2) = (1/m)\bar{P}_s(\bar{\gamma}, \rho_1) + ((m-1)/m)\bar{P}_s(\bar{\gamma}, \rho_2) \quad (3.43)$$

with $\rho_1 = 1$ and $\rho_2 = 0$. However, if a finite delay exists, then it is easy to see that in the absence of correlation, the system reduces to the corresponding plain MRC, case i.e. a (1×2) or $(1, 1; 2)$ system.

Figure 3.14: Influence of feedback rate reduction on TASD and TASP scheme, with feedback delay and reduced feedback rate. (a) 10 block delay ($\mathcal{D} = 10$) and 5 times reduction ($m = 5$) in feedback rate. (b) 20 block delay ($\mathcal{D} = 20$) and 5 times reduction in feedback rate.

3.7.2 Reduced feedback in a correlated channel

This is a more realistic case where the transmitter benefits from residual channel correlation when switching does not happen. When $\tau = 0$, the system acts as in Section 3.7.1 (ideal TAS/MRC) for $1/m$ fraction of the time and partial TASP/D for the rest of the time. With a finite $\tau = \mathcal{D}L_bT_s$ the system acts like TASP/D system for $1/m$ fraction of the time and partial TASP/D for the rest of the time. As delay increases, this system finally tends to a simple MRC. The BER for the TASD in this scenario will be given by the following BER equation:

$$\bar{P}'_{TASD}(\bar{\gamma}, m) = (1/m)\bar{P}_s(\bar{\gamma}, \rho_1) + (1/m) \sum_{i=1}^{m-1} \bar{P}_s(\bar{\gamma}, \rho_2(i)) \quad (3.44)$$

For the non-predicted case, $\rho_1 = J_0^2(2\pi f_d \tau)$ and $\rho_2(i) = J_0^2(2\pi f_d (\mathcal{D} + i)L_bT_s)$. This is because consecutive m blocks of data use the same antenna index which is estimated and intended for the first of the m blocks. Therefore the first block operates with a correlation equal to ρ_1 , while the rest of $m - 1$ blocks operate with $\rho_2(i)$ $i = 1, \dots, m - 1$. Specifically the normalised correlation coefficient $\mathbb{E}[h_i(k + \mathcal{D} + i)h_i^*(k)] = \rho_2(i)$. Subsequent blocks receive further delayed or outdated information as a function of the block length (multiples of L_b). For TASP, similarly, the first block $\rho_1 = r_w^H R_w^{-1} r_w$ with delay $\mathcal{D}$ used in the correlation vector r_w (and r_w^H), and for the rest of the blocks, the correlation term can be approximated as $\rho_2(i) \approx J_0^2(2\pi f_d i L_b T_s) \cdot r_w^H R_w^{-1} r_w$. This is the value of the correlation coefficient $\mathbb{E}[h_i(k + \mathcal{D} + i)\hat{h}_i^*(k + \mathcal{D})]$ between the predicted channel at time $k + \mathcal{D}$ and for the subsequent blocks which do not receive updated switching information, but still benefit partially from the predicted index for block $k + \mathcal{D}$.

When delay exists, with reduced feedback rate, TASP can be made to perform better than TASD. With TASD, degradation is due to the delay not being compensated plus reduced feedback rate. With TASP, prediction improves performance for blocks which receive updated and current predicted indices, while the rest of the $(m - 1)$ blocks benefit from residual correlation existing due to prediction done at the first of the m blocks. Naturally the m th block which is farthest from the signalling

slot, is on the average worst affected in both TASP and TASP and will experience the highest instantaneous BER given m consecutive slots.

When $\mathcal{D} = 0$, with reduced rate, assuming perfect prediction for TASP, the performances of both schemes will be almost same, i.e $P'_{TASD}(\bar{\gamma}, m, \tau=0) = P'_{TASP}(\bar{\gamma}, m, \tau=0)$. Note that in the simplistic case, where no delay and no correlation exists and the feedback rate decreases, the BER tends to a limit set by (3.43). To compare the predictive and non-predictive cases, we consider a (4×1) TAS/MRC setup, under four different scenarios with different feedback delays and signalling rates: (a) moderate finite delay $\mathcal{D} = 10$, moderate signalling rate $m = 5$ (b) high delay $\mathcal{D} = 20$, moderate signalling rate $m = 5$ (c) moderate delay $\mathcal{D} = 10$, low signalling rate $m = 20$ (d) high delay $\mathcal{D} = 20$, low signalling rate $m = 20$. Fig. 3.14a compares the performance of TASP and TASP under case (a) with finite feedback delay set as $f_d\tau = .10$ and a reduction of the feedback rate set to $1/5$ of the block rate, i.e $m = 5$. Both schemes degrade in the presence of both feedback delay and reduced feedback rate. However, it is clear that TASP performs better than TASP. When delay is larger as in Fig. 3.14b, TASP is significantly able to overtake TASP over a wide range of SNRs. This is because, TASP performs more worse with larger delays, while TASP is able to recover losses at high delays.

Fig. 3.15a shows the case where signaling rate is further decreased with moderate finite delay. Compared to Fig. 3.14a with same delay, both schemes significantly degrade in performance, however TASP has a gain of about $7dB$ asymptotically. With larger delay, as in Fig. 3.15b, TASP is completely degraded and almost is similar to a 1×1 SISO, while TASP still offers about $5dB$ gain. Comparing Fig. 3.15b this to Fig. 3.14b, we see that the gap between TASP and TASP is widened when feedback rate is increased. As prediction occurs more frequently with increased feedback rate, large improvement is seen in Fig. 3.15b with $m = 5$, than compared to the $m = 20$ case leading to overall improvement.

Lastly, in Fig. 3.16 considers the case of presence of zero delay with reduced feedback rate $m = 20$, as stated before the performance of both schemes are nearly

Figure 3.15: Influence of feedback rate reduction on TASD and TASP scheme, with feedback delay and reduced feedback rate. (a) 10 block delay ($\mathcal{D} = 10$) and 20 times $m = 20$ reduction in feedback rate. (b) 20 block delay ($\mathcal{D} = 20$) and 20 times reduction in feedback rate.

Figure 3.16: BER Performance of TASD and TASP under reduced feedback rate with zero feedback delay and 20 times reduction in feedback rate.

same. This naturally performs far better than the case in Fig. 3.15b where $\mathcal{D} = 20$.

3.8 Summary

This chapter has explored the issue of transmit antenna selection based on an index fed back from an MRC receiver. The system has been shown to be sensitive to normalised switching delay (thus reducing channel correlation), which may arise from processing time in the reverse communications path and mobility between the transmitter and the receiver. In an effort to mitigate against the effects of normalised delay, a power predictor based on Wiener filtering of past outdated CSI, is introduced at the receiver, to predict a best transmit antenna for future transmission slots. The effects and benefits of channel prediction on several important performance metrics were studied and closed form expressions derived for each of them, all verified by simulation. Prediction was found useful in maintaining transmit diversity and improving both BER and outage performance. The inter-relation between system parameters

such as predictor length, pilot SNR and block length on long range prediction has been explored, and their influence on the channel correlation coefficient noted (where pilot power can improve predictor performance more than increasing the filter order). The system has been shown capable of alleviating much of the performance loss associated with outdated transmitter selection knowledge, even for delays which would have caused non-predictive TAS to be ineffective. At any given delay, a minimum threshold pilot SNR and filter order required for the predictive system to just overtake the non-predictive case was presented graphically. The predictive approach is found beneficial in combating diversity loss in systems, which would otherwise substantially degrade for normalised delays greater than 0.02. For most configurations, it was shown that the use of only a 5 tap filter, per transmit-receive path, was effective in sustaining BER out to about 0.1 normalised delay. Secondly, prediction was also found very beneficial in recovering most of the losses in a scenario where feedback delay and reduced feedback rate exist, where otherwise much of the diversity from adding transmit antennas would have been lost.

Chapter 4

Analysis of Adaptive Modulation with Selective Diversity under Channel Prediction

“I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.”

Isaac Newton

4.1 Introduction

In the previous chapter, transmit antenna selection was studied. The rationale for selecting among transmit antennas was to better apportion available transmit power, or to achieve an effective spatial diversity that is greater than would normally be achievable with the given number of active transmit antennas. Specifically, the performance of a point to point TAS/MRC closed loop MIMO system was analysed in a delay and rate constrained feedback path. The beneficial impact of channel prediction to counter the effects of degradation due to outdated CSI, was analysed. Several system performance metrics were derived and the advantages of predictive TAS scheme over the conventional TAS scheme was explored in several configurations. In this chapter, we extend the analysis of the predictive TAS to adaptive modulated schemes and study its performance for different system requirements.

4.2 Overview of adaptive modulation in MIMO

Achieving both high spectral efficiency (SE) and high data rates is one key goal of many modern communication systems, made more difficult by the problem of multi-path fading in the wireless channel from transmitter to receiver. As we have seen, the use of multiple antennas, combats fading by evening out channel fluctuations and MIMO technology has been shown to provide diversity and capacity gain without incurring power or bandwidth cost. Another method to mitigate against the effects of multi-path fading is to adapt closely to the inherent time-varying signal conditions of the channel, rather than design to either worst-case or average conditions. An example is the agile adjustment of parameters such as constellation size, transmit power and code rate to match changes in the channel. By varying some of these many degrees of freedom, spectral efficiency can be improved [103], [104], [105], [46].

Future communications systems are likely to employ a combination of the techniques mentioned: for example MIMO systems that switch between antennas whilst adjusting transmission parameters in line with changing channel characteristics. Ignoring adaptation based upon data traffic needs, these systems are sensitive to CSI or similar measures of transmission environment, including low/high SNR levels or interference. However for practical systems, adapting parameters at the transmitter (transmit-directed schemes), it is necessary to recognise that the CSI being acted upon may be incorrect: normally CSI must be communicated back from the receiver to the transmitter¹. Even in the absence of quantization and bit errors, the time between determining CSI and acting upon it causes the CSI information to be outdated. This has led to an analysis of the use of imperfect channel estimates for adaptive modulation (AM) systems for SISO channels [36]. In [96], taking channel variation into account, a robust AM system is designed using a single outdated fading estimate in

¹The transmitter usually determines CSI either via a feedback link from the receiver in frequency division duplex (FDD) systems, or by exploiting channel reciprocity in time division duplex (TDD) systems. The latter approach is not considered further in this work: we assume the use of a reverse channel from receiver to transmitter.

Figure 4.1: A $(N_t, 1; N_r)$ transceiver block diagram employing both adaptive modulation and transmit antenna selection with MRC.

a SISO fading channel. Also its effect in AM systems over a Nakagami- m fading channel with feedback delay has been considered in [94] showing the degradation of BER with feedback delay.

4.2.1 Related work

Adaptive modulation has been considered in relation to long range prediction of fading signals [106] for mobile channels. AM with channel prediction has been employed for use in SISO channels [45] where optimum solutions for rate and transmit power are derived based on prediction error variance. In [79] and [80], the channel prediction is employed for AM in multi-antenna systems, however the rate boundaries are not optimised for predicted CSI; only perfect CSI is used in calculating the switching thresholds. Also a rate adaptive modulator coupled with STBC has been utilised to combat fading by Duong and Oien [107], who argued that the scheme is practically better than transmitter eigen-beam forming systems such as in [80], since in the latter the whole channel gain matrix has to be fed back to the transmitter. By contrast, only channel signal to noise ratio (CSNR) required feeding back in the former. However, the throughput was significantly reduced when the number of transmit antennas was greater than two, as a result of rate loss contributed by STBC.

As stated before, TAS alone has been shown to achieve full diversity as if all the transmit antennas were used (at high SNR), and to outperform STCs of the same spectral efficiency [28]. Although orthogonal STBCs offer the advantage of simple

maximum likelihood (ML) decoding, high diversity order STBCs are only possible with a large number of transmit antennas where the number of receive antennas is limited. Moreover, for space time-trellis codes (STTCs), code design for a large number of transmit antennas is difficult, and demodulation becomes very complex.

4.2.2 Contributions

In this work, an adaptive modulated system with multiple transmit and receive antennas is investigated. In addition to modulation adaptation, the transmitter employs TAS, and MRC is used at the receiver. Both transmitter selection and modulation selection require CSI, which is considered degraded due to feedback delays, thereby limiting system performance. In fact, large delays render the feedback information ineffective, effectively creating an open-loop system. In the system described, to improve performance, predicted future CSI is used at the receiver to select the best parameters for transmission. Although the use of a feedback channel is necessary, in practice only a low bandwidth channel is required since it is simply the indices of the best antenna and rate that need to be fed back. In the previous chapter, channel prediction has been shown capable of mitigating issues caused by such delays in a non-adaptive (fixed mode) TAS/MRC system.

The work in this chapter investigates the issues relating to degraded CSI for a TAS/MRC system with prediction that also adaptively adjusts modulation format based upon outdated CSI, in order to maintain a target BER or spectral efficiency. We consider adapting rate rather than power, motivated by the fact that rate adaption is key to achieving high spectral efficiency, and there is little improvement in efficiency by further varying power [46], [36]. Furthermore, 3GPP HSDPA has replaced the power control mechanism (a previous feature of W-CDMA) by AMC techniques to remove power control overheads, thereby yielding power efficiency gains [108]. We also determine the best modulation switching thresholds for the rate- adaptive M-QAM schemes in the TASP/MRC system, taking prediction variance into account. This is a highly desirable component of any practical implementation, given the otherwise

exhaustive computations required to make the determination.

4.2.3 Summary of results

The organisation of the chapter is as follows: In Section 4.3, we describe the AM TASP/MRC system model and assumptions for channel prediction. Next in Section 4.4 we derive the closed form expression for the Shannon capacity bound for such TAS schemes under prediction, and in Section 4.5 we determine the closed form BER relationship for the adaptive schemes. In Section 4.6, we consider two adaptive policies a) Average BER (A-BER) with constant power policy and b) Instantaneous (I-BER) with constant power, and derive optimal switching boundaries for a TAS optimised system. Finally we optimise the TASP/MRC modulated scheme with respect to the constellation set. In Section 4.7, the performance of different TAS/MRC schemes in terms of spectral efficiency and BER are analysed for different target BERs for different degrees of channel correlation; and compared to nonadaptive predictive TAS/MRC schemes. Section 4.8 summarises this chapter.

4.3 System Model and Assumptions

We consider a point-to-point MIMO adaptive system as shown in Fig. 4.1, equipped with N_t transmit and N_r receive antennas. The transmitter is capable of switching between different M-QAM constellations as well as selecting a single best antenna based on information feedback from the receiver. This system is represented as $(N_t, 1; N_r)$. As in the previous chapter, the receiver channel estimates are assumed to be error-free for demodulation— since noncausal channel smoothing with high accuracy can be performed using Wiener interpolator filters (which are non-causal) near perfect demodulation can be assumed [79], [94], [80]. Even with perfect channel estimates at the receiver, the performance of an adaptive system such as TAS/MRC employing rate control can be degraded through delayed or outdated estimates of the channel at the transmitter. In an MIMO adaptive TAS/MRC system, degradation

from using outdated channel estimates can occur in two ways: from a) not selecting a good antenna in the present slot b) choosing an incorrect modulation level or rate. Prediction is therefore necessary to alleviate the losses that occur because of using outdated channel information. Note that as before, the channel predictor is strictly causal. In addition we need to determine optimal modulation mode switching boundaries taking into account the prediction error variance.

The basic system model and the PSAM channel prediction and estimation model are similar to the one in the previous chapter. A block stationary and ergodic channel is assumed. Known pilot symbols are transmitted [92] from each antenna in turn at different time slots into a fixed block of length L_b , and channel estimation for a block of data is carried out independently for all channels. The elements of $\mathbf{H}$, the $N_r \times N_t$ channel matrix, are hence assumed to be constant over a frame, however they are temporally correlated across blocks. $h_{ij}(m)$ is an entry of the channel matrix $\mathbf{H}(m)$, $1 \leq j \leq N_r$, $1 \leq i \leq N_t$, and is the true fading coefficient of the channel between the i th transmit and the j th receive antenna corresponding to the m^{th} transmitted block. The channels $h_{ij}(m)$ are characterised as independent and identically distributed (i.i.d) complex Gaussian random variables $\mathcal{CN}(0, \sigma_h^2)$ that follow Jakes model [42] with Doppler spread f_d and channel power σ_h^2 . The CSI is predicted by Wiener-Hopf equations so that the $\mathcal{D}$ block ahead predicted channel can be written as $\hat{h}_{ij}(m + \mathcal{D})$ i.e $\tau = \mathcal{D}L_bT_s$. The predicted channel SNR for any transmit antenna i at time $(m + \mathcal{D})$ is given by:

$$\hat{\gamma}_i = \frac{E_s}{N_0} \sum_{j=1}^{N_r} |\hat{h}_{ij}|^2 \quad (4.1)$$

E_s is the average transmit power of the symbol, N_0 is the receiver AWGN variance and E_s/N_0 is the average received SNR, depicted as $\bar{\gamma}$. The receiver is now able to take decisions in selecting the appropriate good antenna, as well as the correct rate or power for the next transmission slot. The receiver picks a single good transmit antenna from N_t candidates, based upon the maximum SNR gain:

$$\hat{\gamma}_{max} = \max_{1 \leq i \leq N_t} [\hat{\gamma}_i] \quad (4.2)$$

Besides the true maximum SNR γ_{max} , at time corresponding to predicted maximum SNR, $\hat{\gamma}_{max}$ is given as

$$\gamma_{max} = \max_{1 \leq i \leq N_t} [\gamma_i] \quad (4.3)$$

with the true SNR for antenna i as:

$$\gamma_i = \frac{E_s}{N_0} \sum_{j=1}^{N_r} |h_{ij}|^2 \quad (4.4)$$

The index of the predicted maximum SNR in (4.2) denoted as l , is fed back to the transmitter, through a feedback channel (shown in Fig. 4.1). Based on the index, a single best transmitter is chosen for transmission at time instant q , and the received signal vector can be expressed as:

$$\mathbf{y}(q) = \mathbf{h}_l(q)x(q) + \mathbf{z}(q) \quad (4.5)$$

where $x(q)$ represents the un-coded symbol transmitted from the single selected antenna, $\mathbf{z}(q)$ is the AWGN vector with distribution $\mathcal{CN}(0, \sigma_z^2 \mathbf{I}_{N_r})$, and $\mathbf{I}_{N_r}$ the identity matrix. $\mathbf{h}_l(q)$ is the $N_r \times 1$ chosen received channel vector corresponding to the maximum power gain, which is the l th column of the channel matrix $\mathbf{H}(q)$, assumed to be perfectly estimated at the receiver for demodulation with MRC. In Chapter 3, an expression is found (eqn.(3.30)) for the PDF: $f_{\gamma_{TASP}}(\gamma)$, which is equivalent to the PDF of the true maximum SNR. This was obtained by averaging out the conditional PDF $f(\gamma|\hat{\gamma})$ (eqn.(3.28)) with PDF of $f_{\hat{\gamma}_{max}}(\hat{\gamma})$. To proceed with the mathematical analysis of the AM TASP/MRC system, the PDF $f_{\gamma_{TASP}}(\gamma)$ will be made use of in the following sections.

4.4 Maximum Achievable Spectral Efficiency

The average spectral efficiency of a wireless system denotes the average information rate or capacity per unit bandwidth, given in bits per second per Hertz (bps/Hz). To achieve capacity, adaptive coding in combination with AM or adaptive coded modulation (ACM) [96] can be employed. However, in this chapter, we consider prediction benefits with rate adaptation and un-coded M-QAM modulation, although

such results can easily be extended to coded schemes [44]. Alouni and Goldsmith [94] found that at high SNR, the optimal simultaneous power and rate adaptation and the optimal rate adaptation with constant power provide roughly the same capacity. Moreover, for adaptive transmission with MRC, increasing the number of branches further decreases the gap between these two schemes [109], [110]. Hence, to analyse adaptive transmission in a transmit antenna selection-with-prediction (TASP) system, we are motivated to derive the Shannon capacity denoted as $\langle C_{oratas} \rangle$ [bits/s], with optimal rate adaptation under constant power. This provides us an upper bound for system capacity which will serve as a benchmark to compare our different transmission schemes. The maximum average spectral efficiency in bps/Hz in a channel of bandwidth W of the TASP system can be obtained using [94]:

$$\langle C_{oratas} \rangle / W = \int_0^\infty \log_2(1 + \gamma) f_{\gamma_{TASP}}(\gamma) d\gamma \quad (4.6)$$

Using eqn.(3.30) with [110, eqn. 32] and the expansion for the incomplete gamma function,

$$\Gamma(-n, x) = \frac{-1^n}{n!} \left[E_i(x) - e^x \sum_{s=0}^{n-1} (-1)^s \frac{s!}{x^{s+1}} \right] \quad (4.7)$$

where $E_i(x)$ is the exponential integral of first-order function defined as:

$$E_i(x) = \int_x^\infty \frac{\exp(-t)}{t} dt \quad x > 0, \quad (4.8)$$

the expression for $\langle C_{oratas} \rangle / W$ under imperfect prediction is derived as:

$$\begin{aligned} \langle C_{oratas} \rangle / W &= \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \exp(-\mu) \\ &\cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) (j + N_r - 1)! \sum_{k=0}^j \left(\frac{1}{\bar{\gamma}} \right)^{j+N_r} \frac{\rho_p^k (1 - \rho_p)^{j-k}}{[i(1 - \rho_p) + 1]^{j+k+N_r}} \\ &\cdot \sum_{l=1}^{k+L_r} \Gamma[-k - N_r + l, \mu] / \mu! \end{aligned} \quad (4.9)$$

where $\mu = (i + 1) / (\bar{\gamma} i(1 - \rho_p) + 1)$.

By setting $\rho_p = 1$, the maximum achievable spectral efficiency can be obtained for different TASP/MRC configurations. Figs. 4.2 and 4.3 reveal how the system

Figure 4.2: Maximum achievable spectral efficiency for different TAS configurations $(N_t, 1; N_r)$ with perfect prediction or zero delay case.

Figure 4.3: Maximum achievable spectral efficiency for different TAS configurations $(N_t, 1; N_r)$ with channel power correlation coefficient $\rho_p = 0.7$.

behaves for different degrees of fading correlation and number of antennas. Fig. 4.2 shows two sets of curves for $\rho_p = 1$, where the number of transmit antennas are increased while the number of receive antennas are kept constant. It is observed that increasing the number of transmit antennas yields diminishing capacity gains when the number of receive antennas is large. Fig. 4.3, plotted for $\rho_p = 0.7$, shows that the capacity gains, when compared to Fig. 4.2 are reduced. Also the effect of increasing N_t is further reduced when compared to the perfectly predicted case. The capacity of the $(1, 1; N_r)$ case, i.e with only receive diversity, acts as a lower bound to the capacity of the TASP/MRC schemes, since when CSI is completely outdated, i.e in a hypothetical case when $\rho_p = 0$, the TASP/MRC tends to a simple MRC system.

4.5 System Analysis

In an adaptive discrete rate (ADR) modulated scheme, the modulation mode is changed based on the instantaneous channel quality measure, which in this case is the predicted SNR value for the best transmit antenna selected for transmission in a future slot. In a K -mode modulated scheme, the transmit mode $k \in \{0, 1, \dots, K-1\}$ with throughput $R_k = \log_2(M_k)$ and constellation size $M_k \neq 0$, is chosen when $\hat{\gamma}_k \leq \hat{\gamma} < \hat{\gamma}_{k+1}$, where the switching level $\hat{\gamma}_k$ belongs to the set $\mathbf{v} = \{\hat{\gamma}_k \mid k = 0, 1, \dots, K\}$. $R_0 = 0$ when there is no transmission [111]. $\hat{\gamma}_1$ will be the cut off SNR below which there will be no transmission with $\hat{\gamma}_0 = 0$ and $\hat{\gamma}_K = \infty$. Here we will discuss the derivations of important performance measures namely average throughput, conditional BER, average BER and mode selection probability; these will be used to obtain optimal operating conditions for maximum performance under given channel prediction errors. The average throughput B is given as:

$$B = \sum_{k=0}^{K-1} R_k P_k \quad (4.10)$$

where P_k is the probability that the k th mode is selected, given by:

$$P_k = \int_{\hat{\gamma}_k}^{\hat{\gamma}_{k+1}} f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} \quad (4.11)$$

In order to derive $f_{\hat{\gamma}_{max}}(\hat{\gamma})$ we use $f_{\hat{\gamma}}(\hat{\gamma})$ the PDF of the predicted SNR $\hat{\gamma}$ for any transmit antenna (as discussed in Section 4.3) given by:

$$f_{\hat{\gamma}}(\hat{\gamma}) = \frac{\hat{\gamma}^{N_r-1}}{\bar{\gamma}^{N_r} (N_r - 1)!} \exp\left(-\frac{\hat{\gamma}}{\bar{\gamma}}\right) \quad (4.12)$$

$f_{\hat{\gamma}_{max}}(\hat{\gamma})$ is the PDF of $\hat{\gamma}_{max}$. The CDF of $\hat{\gamma}$ is then:

$$F_{\hat{\gamma}}(\hat{\gamma}) = 1 - \exp(-\hat{\gamma}/\bar{\gamma}) \sum_{m=0}^{N_r-1} (1/m!) (\hat{\gamma}/\bar{\gamma})^m \quad (4.13)$$

Making use of order statistics, the PDF $f_{\hat{\gamma}_{max}}(\hat{\gamma})$ can be calculated as follows:

$$f_{\hat{\gamma}_{max}}(\hat{\gamma}) = N_t f(\hat{\gamma}) [F(\hat{\gamma})]^{N_t-1} \quad (4.14)$$

The PDF $f_{\hat{\gamma}_{max}}(\hat{\gamma})$ can be obtained by expanding (4.14) using (A.1). Using the lower incomplete gamma function, eqn. (4.11) is evaluated as:

$$P_k = F_{\hat{\gamma}_{max}}(\hat{\gamma}_{k+1}) - F_{\hat{\gamma}_{max}}(\hat{\gamma}_k) \quad (4.15)$$

to obtain closed form solutions, where $F_{\hat{\gamma}_{max}}(\hat{\gamma})$ is the CDF of $f_{\hat{\gamma}_{max}}(\hat{\gamma})$.

Let us assume initially that we employ continuous rate M-QAM and that the modulation scheme M is chosen based upon predicted SNR $\hat{\gamma}$. In fact, thanks to Chung and Goldsmith [112], we have an approximate expression that is easily invertible given as $M(\hat{\gamma}) = 1 + \frac{1.6}{Q_0} \hat{\gamma}$, where $Q_0 = -\ln(5BER_0)$ and BER_0 is set as the desired instantaneous target BER for SNR $\hat{\gamma}$. We can therefore find an approximate expression for the instantaneous BER over the selected antenna as a function of γ , $\hat{\gamma}$ and transmit power $P(\hat{\gamma})$, that holds good when $M(\hat{\gamma}) \geq 4$ and $BER \leq 10^{-2}$ [94]:

$$BER(\gamma|\hat{\gamma}) \approx 0.2 \exp\left(-Q_0 \frac{\gamma}{\hat{\gamma}} \frac{P(\hat{\gamma})}{\bar{P}}\right). \quad (4.16)$$

Note that $\bar{P}$ is the same as E_s , the average transmit power of the symbol. This will be adapted for the case of constant transmit power: $P(\hat{\gamma}) = P$, with the average power constraint $\bar{P} = 1$. The transmit power is dependent on $\hat{\gamma}_1$ to be optimised later in Section 4.6 with the requirement $\bar{P} = 1$. The average SNR $\bar{\gamma}$ will be adjusted by varying N_0 . Note that instantaneous BER, $BER(\gamma|\hat{\gamma})$, in (4.16) evidently differs

Figure 4.4: $BER(\hat{\gamma})$ against instantaneous SNR for $\rho_p = 0.9, 0.8$ in an $(N_t, 1; 2)$ setup at $\bar{\gamma} = 20dB$.

from Q_0 , since γ differs from $\hat{\gamma}$ due to prediction errors. Considering that the receiver has no knowledge of γ , we can then find the average BER conditioned on $\hat{\gamma}$ (which can be thought of as a short term average) by averaging (4.16) over all values of the γ :

$$BER(\hat{\gamma}) = \int_0^\infty BER(\gamma|\hat{\gamma})f(\gamma|\hat{\gamma})d\gamma \quad (4.17)$$

With the aid of eqn. (4.16) and eqn. (3.28) we can solve eqn. (4.17) as:

$$BER(\hat{\gamma}) = 0.2 \left[\frac{\hat{\gamma}}{\hat{\gamma} + Q_0 P (1 - \rho_p) \bar{\gamma}} \right]^{L_r} \exp \left[- \frac{Q_0 P \hat{\gamma}}{\hat{\gamma} + Q_0 P (1 - \rho_p) \bar{\gamma}} \right] \quad (4.18)$$

The above equation is applicable to any $(N_t, 1; N_r)$ configuration and as such it does not depend on N_t . With ρ_p in the picture, in general there is an increase in instantaneous BER as a result of prediction errors, causing it to be different from the target BER, BER_o . Different system parameters such as $\rho_p, N_r, \bar{\gamma}$ influence the non-linear BER curve.

Since we restrict ourselves to discrete M-QAM, we consider $M_k(\hat{\gamma})$ taking only

four modes : 4, 16, 64, 256, so that we can substitute $Q_0 = \frac{1.6 \hat{\gamma}}{M_k(\hat{\gamma})-1}$ in (4.18). Thus (4.18) becomes:

$$BER(\hat{\gamma}) = 0.2 (1 + C_k)^{-L_r} \exp \left(- \frac{1.6 P \hat{\gamma} / (M_k - 1)}{1 + C_k P} \right) \quad (4.19)$$

where $C_k = 1.6 \bar{\gamma} (1 - \rho_p) / (M_k - 1)$. Note that this is specific to a given M_k . Fig. 4.4 shows the plot of $BER(\hat{\gamma})$ for different modulation modes for $N_r = 2$ at a received SNR of $\bar{\gamma} = 20dB$ and for $\rho_p = 0.9, 0.8$. Since at this point we are not dealing with the cut-off SNR $\hat{\gamma}_1$, we take $P = 1$ in these curves; which also is quite a valid assumption at high $\bar{\gamma}$ when $\hat{\gamma}_1$ is taken into consideration, as seen later. (4.19) was derived using (4.16) which is an approximated result, so as a check we also plot simulated BER values of (4.17) using numerical integration, using a reportedly precise approximation for $BER(\gamma|\hat{\gamma})$ [113]:

$$BER(\gamma|\hat{\gamma}) \approx 4 \frac{\sqrt{M_k(\hat{\gamma})} - 1}{R_k \sqrt{M_k(\hat{\gamma})}} \sum_{b=1}^{\sqrt{(M_k(\hat{\gamma})/2)}} Q \left((2b-1) \sqrt{3\gamma/M_k(\hat{\gamma}) - 1} \right) \quad (4.20)$$

with $Q(x) = 1/\sqrt{2\pi} \int_x^\infty \exp(-x^2/2) dx$. As seen in Fig. 4.4, there is a close match between the simulated accurate and approximated $BER(\hat{\gamma})$ for smaller M_k ; the gap narrowing for high SNRs. Interestingly when the predicted value is low, $BER(\hat{\gamma})$ decreases with decreasing ρ_p . To explain this we look at the first and second order statistics of the conditional PDF $f(\gamma|\hat{\gamma})$ in eqn. 3.28. The mean and variance of γ conditioned on $\hat{\gamma}$ are derived using (A.3.1) as follows:

$$E[\gamma|\hat{\gamma}] = \hat{\gamma} + N_r(1 - \rho_p)\bar{\gamma} \quad (4.21a)$$

$$E[(\gamma^2|\hat{\gamma})] - E[(\gamma|\hat{\gamma})]^2 = 2(N_r + 1)(1 - \rho_p)\hat{\gamma}\bar{\gamma} + (N_r(1 - \rho_p)\bar{\gamma})^2 \quad (4.21b)$$

As interpreted from (4.21a), the difference between the predicted SNR and $E[\gamma|\hat{\gamma}]$, is more when correlation is less, and lesser when correlation is higher. Thus the actual true SNR is likely to be higher than the predicted value, on the average. The variance is also smaller for lower predicted values as seen from (4.21b). Therefore, for low predicted SNRs, the expected $BER(\hat{\gamma})$ will tend to reduce for lower ρ_p . In general, for larger predicted $\hat{\gamma}$, $BER(\hat{\gamma})$ increases with decreasing ρ_p . This is because

an increase in variance about the mean $E[\gamma|\hat{\gamma}]$ is seen. Because of the asymmetric nature of the PDF $f(\gamma|\hat{\gamma})$ about its mean, its likely that lower values of γ are assigned higher probabilities, thus contributing to the overall higher BER.

The average BER of ADR M-QAM system is the ratio between the average number of bits in error and the total number of transmitted bits [36], [94] given as:

$$BER_{av} = \frac{\sum_{k=1}^{K-1} BER_k R_k}{\sum_{k=1}^{K-1} P_k R_k} \quad (4.22)$$

where the mode specific bit error rate (contribution from a specific modulation scheme M_k) is given by [45]:

$$BER_k = \int_{\hat{\gamma}_k}^{\hat{\gamma}_{k+1}} \int_0^\infty BER(\gamma|\hat{\gamma}) f(\gamma|\hat{\gamma}) f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\gamma d\hat{\gamma} \quad (4.23)$$

$$= \int_{\hat{\gamma}_k}^{\hat{\gamma}_{k+1}} BER(\hat{\gamma}) f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} \quad (4.24)$$

Combining (4.19) and the PDF in (3.26), and using (A.4.1), a closed form solution of (4.24) can be obtained:

$$BER_k = \frac{0.2N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) j! \frac{[1 + \bar{\gamma}g(1 - \rho_p)\bar{\gamma}]^j}{[(1 + \bar{\gamma}g(1 - \rho_p))(i + 1) + g\rho_p\bar{\gamma}]^{j+N_r}} \cdot \left[\exp(-\omega\hat{\gamma}_k) \sum_{l=0}^{j+N_r-1} \frac{(\omega\hat{\gamma}_k)^l}{l!} - \exp(-\omega\hat{\gamma}_{k+1}) \sum_{l=0}^{j+N_r-1} \frac{(\omega\hat{\gamma}_{k+1})^l}{l!} \right]$$

$$\text{with } \omega = \frac{(1 + \bar{\gamma}(1 - \rho_p)g)(i + 1) + g\rho_p\bar{\gamma}}{\rho_p\bar{\gamma}[(1 + \bar{\gamma})(1 - \rho_p)g]} \text{ and } g = 1.6P/(M_k - 1). \quad (4.25)$$

Finally BER_{av} (4.22) can be computed using (4.10) and (4.25).

4.6 System Performance

In this section, we consider the following setups with two types of constraint:

(a) Instantaneous BER and Constant Power (I-BER, CP), (b) Average BER and

Constant Power (A-BER, CP). The goal is to maximize SE, which involves finding optimal switching boundaries between modulation modes in each case. The first scheme requires that the instantaneous error probability remains below a specified error-rate level at every $\bar{\gamma}$ so that (4.19) is satisfied as $BER(\hat{\gamma}; \bar{\gamma}) \leq BER_0$ along with the constant power requirement. The second scheme insists on a constant average BER over a range of $\bar{\gamma}$ so that (4.22) is fulfilled as $BER_{av}(\mathbf{s}; \bar{\gamma}) = BER_0$ where $\mathbf{s} = \{\hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_K\}$, a set of threshold levels which maximise SE.

In [45], where channel prediction is considered for ADR M-QAM modulation for a SISO Rayleigh channel, the authors have derived optimum solutions for rate boundaries for variable rate-variable power, as well as variable rate-constant power schemes. When prediction variance was taken into account, and un-optimised switching thresholds (obtained from perfect CSI) were used for mode switching, their system was shown to degrade. Therefore it can be directly inferred that the TASP-MRC system will also show performance degradation if such un-optimised threshold values are used. In [44], ACM is considered with receive antenna diversity and channel prediction as well as channel estimation errors, however the system is optimised only for I-BER using a numerical approach.

Starting from the closed form expressions of average spectral efficiency, instantaneous and average BER in the previous sections, we need to determine optimal switching SNR thresholds to maximize the spectral efficiency subject to meeting the BER constraints of these two schemes, taking prediction errors into account.

4.6.1 Instantaneous BER and Constant Power

The average spectral efficiency of the TASP/MRC scheme given in (4.10) is expanded as:

$$B(\mathbf{s}; \bar{\gamma}) = \sum_{k=1}^{K-1} R_k \int_{\hat{\gamma}_k}^{\hat{\gamma}_{k+1}} f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} \quad (4.26)$$

As per the definition, a set of SNR thresholds $\mathbf{s} = \{\hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_K\}$ have to be determined at every $\bar{\gamma}$ to maximise (4.26) as follows:

$$\max_{\mathbf{s}} B(\mathbf{s}; \bar{\gamma}) \quad (4.27a)$$

$$\text{subject to } BER(\hat{\gamma}; \bar{\gamma}) \leq BER_0, \bar{P} = 1 \quad (4.27b)$$

The equality condition in (4.27b) should be satisfied at the rate region boundaries so that $BER(\hat{\gamma}_k; \bar{\gamma}) = BER_0$. The optimal switching boundaries can thus be found which also satisfy the average transmit power constraint: $\bar{P} = 1$, by inverting (4.19). In so doing we need to solve for P which is a one time adjustment of transmit power at each $\bar{\gamma}$. To solve the two unknowns we utilize the following relation between the optimal cut off SNR level, below which data transmission is suspended (can not sustain BER), and P which is:

$$P = 1/[1 - F_{\hat{\gamma}_{max}}(\hat{\gamma}_1)] \quad (4.28)$$

a value which gets close to 1, for high $\bar{\gamma}$. Inverting (4.19) at the threshold boundaries, we obtain a relationship to P :

$$\hat{\gamma}_k = -\frac{5}{8} \ln((BER_0/0.2)(1 + \mathcal{C}_k.P)^{L_r})P(1 + \mathcal{C}_k.P)(M_k - 1) \quad (4.29)$$

Finally from (4.28) and (4.29), we determine the first threshold boundary, after which P itself is determined from (4.28). Once this is done other boundary crossover values are obtained in turn from (4.29).

4.6.2 Average BER and Constant Power

In this case the requirements can be written as follows:

$$\max_{\mathbf{s}} B(\mathbf{s}; \bar{\gamma}) \quad (4.30a)$$

$$\text{subject to } BER_{av}(\mathbf{s}; \bar{\gamma}) = BER_0, \bar{P} = 1 \quad (4.30b)$$

The optimal solutions for $\mathbf{s}$ are obtained by using Lagrangian multipliers which lead to an unconstrained optimisation problem as follows:

$$J(\mathbf{s}; \bar{\gamma}) = B(\mathbf{s}; \bar{\gamma}) + \lambda\{BER_{av}(\mathbf{s}; \bar{\gamma}) - BER_0\} \quad (4.31)$$

The optimal set of switching levels should satisfy

$$\frac{\partial J}{\partial \mathbf{s}} = \frac{\partial J}{\partial \gamma_k} = 0 \quad 0 \leq k \leq K-1 \quad (4.32)$$

$$BER_{av}(\mathbf{s}; \bar{\gamma}) - BER_0 = 0 \quad (4.33)$$

By expanding and solving (4.32), we arrive at the unconstrained equation:

$$BER(\hat{\gamma}_k) = BER_0 - 1/\lambda \quad (4.34)$$

This is similar to the condition as in (4.29) and thus the cross-over threshold SNRs which satisfy (4.32) can be obtained using as:

$$\hat{\gamma}_k = -\frac{5}{8} \ln \left[((BER_0 - \lambda^{-1})/0.2)(1 + \mathcal{C}_k P)^{L_r} \right] P(1 + \mathcal{C}_k P)(M_k - 1) \quad (4.35)$$

By exploring λ in the range $\lambda < 0$, optimal boundaries that fulfil the BER constraint, can be found by extending the analysis as in [45], with the condition that $BER_0 < 0.2/(1 + \mathcal{C}_1 P)^{L_r}$. In other words, since we aim at $BER_{av}(\mathbf{s}; \bar{\gamma}) = BER_0$, considering that the instantaneous $BER(\hat{\gamma})$ decreases as $\hat{\gamma}$ increases, we expect that $BER(\hat{\gamma}_k) > BER_0$. Looking at (4.34) and (4.35) we see that this happens for negative values of λ in the region $\lambda \in \left(-\infty, \frac{1}{BER_0 - 0.2(1 + \mathcal{C}_1 P)^{-L_r}} \right]$.

On the other hand, when $BER_0 > 0.2/(1 + \mathcal{C}_1 P)^{L_r}$ we can choose a value of λ close to $\frac{1}{BER_0 - 0.2(1 + \mathcal{C}_1 P)^{-L_r}}$ to obtain the threshold set $\mathbf{s}$ closest value to the average BER. However, a set of solutions $\mathbf{s}$ found, would result in a drop in the $BER_{av}(\mathbf{s}; \bar{\gamma})$ than required, along with a reduction in spectral efficiency. This condition arises, depending on the target BER, or in instances of poor correlation or high average SNR (or a combination of both), also being dependant upon the number of receive antennas.

To improve SE whenever $BER_0 > 0.2/(1 + \mathcal{C}_1 P)^{L_r}$, optimal solutions need to be found. This is not trivial apart from by brute force numerical search over signal boundaries of $BER(\hat{\gamma}_k; \bar{\gamma})$ for every given channel SNR and modulation scheme. One improvement is to obtain the initial un-optimised Lagrangian solution, as discussed above, and then perform a constrained numerical search around this solution until the

Figure 4.5: Plot showing improvement in spectral efficiency for A-BER scheme using threshold setting from our the proposed method compared with an existing method [45] for $\rho_p = 0.9, 0.8$ for modulation sets $K=5$ (0,4,16,64,256) and $BER_0 = 10^{-3}$ in a (2,1;2) system.

required condition is met. Because of the way BER_{av} drops, we simply need to look for smaller values of threshold for each modulation scheme in order to match up to BER_{av} to increase spectral efficiency. However a search may still be time consuming, especially when the modulation set size is large. An efficient method of finding lower thresholds when $BER_0 > 0.2/(1 + \mathcal{C}_1.P)^{L_r}$ is discussed below.

4.6.3 Efficient threshold setting for A-BER, CP scheme.

An efficient heuristic proposed here is to generate a set of thresholds for a range of λ_ρ values using (4.35) with ρ_p replaced by $\rho > \rho_p$. A value of λ_ρ is found to satisfy (4.33). The idea of using $\rho > \rho_p$ is to obtain lower thresholds. In essence, this is:

$$\mathbf{v}_{opt} = \min_{\mathbf{v}} |BER_{av}(\mathbf{v}; \bar{\gamma}, \rho_p) - BER_0 B(\mathbf{v}; \bar{\gamma}, \rho_p)| \quad (4.36)$$

$$v_{\lambda_\rho}(k) = \hat{\gamma}_k \text{ for all } \lambda_\rho < 0 \text{ and } \rho > \rho_p, k = 1, ..K - 1 \quad (4.37)$$

Figure 4.6: Average BER for I-BER at target BERs of 10^{-4} with $\rho_p = 0.999$ and $\rho_p = 0.9$ in a (2,1;2) system, shown as thick lines. Also shown are the non-adaptive average BER, BER_M for TASP/MRC M-QAM for the two values of ρ_p .

Lower threshold values $\mathbf{v}_{\lambda_\rho}$ are thus generated over the range of λ_ρ , which are used to satisfy (4.36). The value of the fixed transmitter gain P and the first threshold value $\mathbf{v}_{\lambda_\rho}(1)$ can be obtained in a similar way using (4.28) with $\rho = \rho_p$. This method is used for the subsequent performance analysis and plots for the A-BER case.

Fig. 4.5 indicates the advantage of the proposed method over the method of [45] for different correlation coefficients for five mode M-QAM (note also that [45] considered only SISO channels, whereas we extend the analysis to MRC). The gains from the new method are larger when the channel correlation is low and when the number of modulation modes is high.

4.7 Performance Analysis

We choose a modulation set consisting of four square M-QAM constellations corresponding to 2 (4-QAM), 4 (16-QAM), 6 (64-QAM) and 8 (256-QAM) bits per symbol.

Figure 4.7: Average BER for I-BER and A-BER schemes at target BERs of 10^{-3} and 10^{-5} with $\rho_p = 0.999$ in a (2,1;2) system, shown as thick lines. Grey lines correspond to average BER of non-adaptive M-QAM with (2,1;2) TAS with equal ρ_p .

The correlation coefficient is affected by many system parameters, so that several dimensions of system performance trade-off are possible. Depending on the filter order and the amount of power of the pilot, the predictor can always be made to offer greater correlation at fixed feedback delay. For more details on how system parameters effect ρ_p refer to [91], [106], [77]. We analyse the system performance at different values of correlation ranging from very good to very poor prediction: $\rho_p = 0.999 \leftrightarrow 0.7$ for different TASP setups. In Fig. 4.6 we first demonstrate and analyse the performance of a (2, 1; 2) TASP/MRC system with I-BER constraints and study its average BER ($IBER_{avb}$) as well its SE. The Shannon SE with $\rho_p = 1$ is also plotted to serve as a benchmark. Also are plotted the non-adaptive TASP/MRC schemes for the two values of ρ_p . The BER curves of the fixed-mode TASP/MRC for arbitrary N_t , N_r , ρ_p for M-QAM with constellation size M are obtained as:

$$BER_M = \int_0^\infty \int_0^\infty BER(M, \gamma) f(\gamma|\hat{\gamma}) f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\gamma d\hat{\gamma} \quad (4.38)$$

where for square M-QAM [112],

$$BER(M, \gamma) = 0.2 \exp\left(\frac{-1.6 \gamma}{M-1}\right) \quad (4.39)$$

(4.38) can be reduced to:

$$BER_M = \int_0^\infty BER(\hat{\gamma}) f_{\hat{\gamma}_{max}}(\hat{\gamma}) d\hat{\gamma} \quad (4.40)$$

which is solved using (4.25). Fig. 3 compares the performance of $IBER_{avb}$ for the four mode adaptive scheme calculated using 4.25, using the optimised thresholds of IBER calculated from 4.29. While target BER, BER_0 , is maintained at the switching thresholds, the average BER $IBER_{avb}$ with I-BER constraint is always lower than BER_0 . At high average SNR, when full capacity is reached, $IBER_{avb}$ aligns itself with the fixed mode (non-adaptive) M-QAM BER curve of the highest modulation order, which in our case is the 256QAM. Next we analyse the performance of the (2;1,2) system with both A-BER and I-BER constraints and compare their average BER ($ABER_{avb}$ and $IBER_{avb}$ respectively) as well as their spectral efficiencies (denoted by $ABER_{se}$ and $IBER_{se}$ respectively). The Shannon spectral efficiency (using eqn. (4.9) with $\rho_p = 1$) is also plotted to serve as a benchmark.

Fig. 4.7 compares the performance of $ABER_{avb}$ and $IBER_{avb}$ for the five mode ($K = 5$) adaptive scheme calculated using (4.22), using the optimised thresholds of A-BER and I-BER for target BERs of 10^{-3} and 10^{-5} and channel power correlation of 0.999. As seen, $ABER_{avb}$ (the solid thick line), is maintained constant for a wide range of SNRs. However, when the full capacity of the ADR A-BER scheme is attained, as expected the BER curve aligns itself with the fixed mode TASP/MRC (non-adaptive) M-QAM BER curve of the highest modulation order, which in our case is the 256- QAM. The average BER $IBER_{avb}$ curve is always lower than the target BER and at high SNR appears parallel to the higher order 256 M-QAM BER curve, when full capacity is reached.

Fig.4.8 plots the spectral efficiencies of the schemes for $\rho_p = 0.999$ and 0.75. As expected $IBER_{se}$ is always lower than $ABER_{se}$ for the same target BER, since on the average more lower rate modes are activated in order to maintain an instantaneous

Figure 4.8: Spectral efficiency for I-BER and A-BER schemes for target BER of 10^{-3} and 10^{-5} with $\rho_p = 0.999$ (left) and 0.75 (right) in a (2;1,2) system; also plotted is the maximum achievable Shannon spectral efficiency for perfect prediction.

BER target $BER(\hat{\gamma}; \bar{\gamma}) \leq BER_0$. Also the rate boundaries for the IBER case are greater than that of ABER. It is observed that the effect of decreasing ρ_p was to reduce the capacity of both $IBER_{se}$ and $ABER_{se}$. For $\rho_p = 0.99$, it was found that the reduction in spectral efficiency compared to $\rho_p = 0.999$ in either case was very small. However, at $\rho_p = 0.75$, and at $\bar{\gamma} = 20dB$ the reductions in $ABER_{se}$ and $IBER_{se}$ were found to be 6% and 20% respectively at $BER_0 = 10^{-3}$, while at $BER_0 = 10^{-5}$ these values were 15% and 45% respectively, indicating that the $IBER_{se}$ performs worse at lower BERs than $ABER_{se}$. This effect is due to two reasons: first the cut off SNR for I-BER is increased with lower BER_0 thereby increasing outage, resulting in a decrease in SE by disallowing transmission for a certain portion of time. As a result switching occurs for higher values of $\hat{\gamma}$, which now occurs with low probability (because of the shape of $f_{\hat{\gamma}_{max}}(\hat{\gamma})$) for small $\bar{\gamma}$. Secondly, for higher $\bar{\gamma}$, threshold boundaries for lower rate modes are significantly increased, and hence are also selected with higher probability, thus slowing down overall increase in SE.

Figure 4.9: Effect of N_t and ρ_p on $ABER_{se}$ with $N_r = 2$ for a target BER of 10^{-4} ; also plotted for these setups is the maximum achievable Shannon spectral efficiency for perfect prediction.

Consequently, for lower target BERs the gap between $IBER_{se}$ and $ABER_{se}$ is increased at lower channel correlations. However, at very low correlations, this gap is expected to reduce because the capacity is affected largely by erroneous antenna selection rather than from poor rate selection.

Next we look at how SE improves with respect to increasing transmit antennas as well as ρ_p . We plot $ABER_{se}$ for three setups of $(1, 1; 2)$, $(2, 1; 2)$ and $(4, 1; 2)$ with ρ_p values of 0.999, 0.99, 0.93 in each case. The target BER is chosen as 10^{-4} for all comparisons as shown in Fig. 4.9. Higher order setups increase system SE, indicating substantial power saving. We can see that the percentage drop at a specific SNR $\bar{\gamma}$ for decreasing correlation is less in higher diversity setups. However, for very poor prediction, the drop is faster in higher order setups because of poor antenna selection. A similar trend is seen with $IBER_{se}$ for higher order setups, although the results are not plotted here. As seen before the $ABER_{se}$ will be more tolerant to prediction errors than $IBER_{se}$.

Figure 4.10: Effect of N_t and ρ_p on $IBER_{se}$ with $N_r = 2$ for a target BER of 10^{-4} ; also plotted for these setups is the maximum achievable Shannon spectral efficiency for perfect prediction.

Fig. 4.10 shows the $IBER_{se}$ graph which is similar to $ABER_{se}$ in Fig. 4.9, however, compared to $ABER_{se}$ there is a reduction of SE. As seen, the average BPS throughput curves of the adaptive square QAM schemes do not have a linear rise and is evident when the number of antennas increase. The undulations in the rise of the SE in both graphs can be explained by looking at the mode selection probabilities. The effect is explained by observing the mode selection probability curves against average SNR, which are overlapping ‘bell’ shaped curves. Fig. 4.11 plots the mode selection probabilities plotted for the $IBER$ scheme for two different setups. At any SNR, any curve will represents the probability of that mode getting selected for transmission. The ripples in the spectral efficiency graph is due to two reasons: a) use of multiple antennas, b) the the absence of 3, 5, and 7-BPS square QAM modes for discrete-rate adaptive modulated scheme. When the number of antennas is increased, the PDF of the received SNR becomes more peakier and becomes more ‘selective’. Hence, there is a chance that for a range of SNRs, a particular mode gets selected for a larger portion

Figure 4.11: Mode selection probabilities of I-BER for $(4, 1; 2)$ and $(1, 1; 2)$ at $\rho_p = 0.99$, (corresponding to a channel MMSE of -20 dB) with $BER_0 = 10^{-4}$.

of time compared to the rest for transmission. Also because of the gaps in the rates, the narrow pdf of the SNR tends to pick a certain mode with higher probability for a certain range of SNRs. Because of this, the growth in SE is slowest for the SNRs corresponding to top rounded portion for any particular mode and fastest for other regions. The ripples in the SE growth are an immediate consequence because of the above phenomenon. Another reason is because of the absence of intermediate QAM modulation modes, causing peakier selection probabilities. The effect becomes less pronounced as number of antennas decreases, since now the PDF of the channel widens, causing it to be less selective, and at any given SNR the randomness of the channel gives a chance for most of the modes to get selected for transmission. Thus it is a weighted contribution towards increase in SE from most of the modes. Notice also as the correlation decreases, the undulations decrease. This is again explained by the nature of the channel PDF, which becomes less selective. A remedy to remove these undulations is by introducing 8QAM and 32 QAM, causing a smooth rise in the SE growth, however, this may not be desirable from an implementation point of

view.

Fig. 4.12 shows the $IBER_{se}$ spectral efficiency gains of the three setups over the non delayed (1;1;2) schemes for $BER_0 = 10^{-4}$, at different power correlation values. These values correspond to a channel MMSE $10 \log_{10}(1 - \rho_p)$ and are chosen to range from $-30dB$ MSE to $-5dB$.

All arrangements can be seen to degrade with decreasing correlation. For higher order setups and at higher correlation, the ADR scheme appears less sensitive to CSI imperfection (particularly when MSE is between -35 to $-20dB$). For decreasing correlation, the fall in gain is more rapid in (4, 1; 2) than in (2, 1; 2), which is expected since more loss occurs from poor antenna selection. Such graphs can be used in system optimisation, for trade-off between computation needs (for prediction), feedback bandwidth, and hardware resource (i.e. transmitter costs), also for choice between open loop and lower order closed loop systems. For example, as seen for an MMSE of $-7dB$, the (4, 1; 2) offers almost no improvement over the (1, 1; 2) system with perfect prediction. Thus in this case if the desired SE is met, it would probably be more economical to spend resources on enhancing channel prediction, rather than on feedback schemes or costly additional hardware at the transmitter end.

4.8 Summary

This chapter has studied a multi-antenna adaptive modulation diversity system employing predictive transmit antennas and MRC at the receiver, under conditions of degraded channel knowledge. The system closed form Shannon capacity gain was first derived for optimal rate adaptation at constant power for arbitrary numbers of transmit and receive antennas. This was then used as a benchmark for two adaptive schemes which maintain constant overall transmit power, but aim to constrain instantaneous BER and average BER respectively. Closed form solutions were derived to obtain threshold boundaries (using Lagrangian multiplier techniques) in the adaption schemes to maintain system performance. An efficient heuristic to obtain these thresholds was proposed to improve system performance when optimality was

Figure 4.12: Percentage I-BER spectral efficiency gain of $(4, 1; 2)$, $(2, 1; 2)$ and $(1, 1; 2)$ setups over the $(1, 1; 2)$ non delayed or perfectly predicted scheme at different ρ_p , corresponding to a channel MMSE of -30,-25,-20,-15,-12,-10,-8,-7,-5 dB (top to bottom) with $BER_0 = 10^{-4}$.

not achieved. These solutions have been verified against numerical simulation, and applied for several designs to explore system optimization. A design space of several different degrees of channel correlation, a range of antenna configurations, and different target BERs, has been explored.

While it is evident that using adaptive modulation in combination with multi-antenna systems, such as TAS-MRC, can substantially increase system performance, it becomes altogether more important to preserve these gains by ensuring good quality feedback channel information – since this conveys important switching information which is largely responsible for those gains. Incorporating channel prediction into a TAS-MRC system is shown as being pivotal in improving system performance and in maintaining that performance in the face of CSI degradation. At the same time, it is also crucial to determine optimal operating conditions for modulation mode switching which take these prediction errors into account. In doing so the system is doubly rewarded, with improved antenna selection as well as optimal mode switching.

In summary, TAS-MRC is seen to be a viable approach for system implementation, even in the real-world situation of outdated CSI. We have shown that both average and instantaneous BER requirements can be fulfilled with system power to be constant.

Chapter 5

Cooperative Relaying

“Little by little, one travels far.”

J. R. R. Tolkien

This chapter begins with a brief overview of cooperative communications. It introduces the concept of relaying as a key technique to improve quality of service in wireless networks. Specifically, this chapter will serve as an introduction to relaying, using different relay architectures as a means to achieve cooperative diversity, and give an overview of the main categories of relaying techniques with a basic mathematical framework for standard schemes. Amongst different relaying schemes, relay selection is particularly recognised as an attractive solution for improving bandwidth usage in relay networks, while maintaining spatial diversity benefits.

5.1 Overview

Although MIMO is a promising standard to improve the performance of radio systems, implementation is often limited by cost factors, since MIMO introduces the need for extra antennas, plus increased power consumption. To improve efficiency, cooperative communication is being envisioned as an efficient way of improving system performance by exploiting the inherent spatial diversity in channels. The spatial diversity advantage of the broadcast nature of wireless transmission, results from co-operating nodes at different geographic locations being responsible for creating many independently faded versions of the signal at the receiver, thus emulating a virtual antenna array. For a given amount of transmit power, higher signal coverage and/or bit-rates can be achieved through cooperation among nodes. Direct transmissions

Figure 5.1: Cooperative relay network

can be expensive in terms of utilised power when the separation distance between transmitter and receiver is large implying that the path loss is high. Relayed transmissions can overcome this problem by sequential conveying information through two or more intermediate cooperating nodes. Cooperating relaying terminals may be base-stations, mobile hand-held devices, or fixed infrastructure-based terminals placed at strategic locations, aimed to extend signal coverage. Cooperative communication can be traced back to the pioneering work of Cover and Gamal on the information theoretic properties of the relay channel [31], where the capacity of a three-node network consisting of a source, a destination, and a relay was analysed in an AWGN discrete memory-less channel. Motivated by this concept, cooperative diversity applied to wireless fading channels has been an active research area in the past decade, and has shown to achieve significant performance gains in terms of link reliability, spectral efficiency, system capacity, and transmission range. Fig. 5.1 illustrates the concept of cooperative relaying where a base station (BS) connects to a destination via many intermediate cooperating relay nodes. As seen, relaying can be used for

coverage extension, to improve performance in severely shadowed regions, and also increase capacity and diversity. Several different relaying strategies are possible given a source, destination and number of relay nodes, depending on implementation complexity. A variety of cooperative diversity protocols that can be utilized in the network exist such as fixed, selection, and incremental relaying. These protocols employ different types and degrees of processing by relay terminals, as well as different types of combining at the destination node. For fixed relaying, the relays either amplify their received signals subject to their power constraint, or decode, re-encode, and retransmit messages. Another class of relays called compress and forward, quantize the channel output and forward it to the destination. Selection relaying builds upon fixed relaying by allowing transmitting terminals to select a suitable cooperative (or noncooperative) action based upon the measured SNR between them. Incremental relaying improves upon the spectral efficiency of both fixed and selection relaying by exploiting limited feedback from the destination and relaying only when necessary. These cooperative relaying protocols were first introduced in [2], and an outage analysis of the cooperative diversity concept was also conducted, showing performance benefits as compared to the non-relayed or direct transmission case.

A practical implementation concern of relaying is whether relays operate in full or half-duplex mode. To overcome practical implementation difficulties in realising a full-duplex relay, whereby it can transmit and receive simultaneously in the same frequency band, half-duplex relays were suggested in [2] where transmission and reception takes place in orthogonal channels. Here, in the first time slot, a source transmits to relay, while in the second slot the relay transmits to its destination. Half duplex therefore leads to a loss in capacity, and the overall capacity is scaled by a factor based on the number of orthogonal channels required for communication between source and destination, which for a three node network has a value of 2. To overcome the loss in spectral efficiency, incremental relaying was also proposed based on source feedback. However, this assumes direct path between source and destination. Two-way relaying was also introduced in [114], where source and destination

transmit simultaneously via orthogonal channels to the relay in the first phase, while the second phase relay broadcasts the combined signal to both source and destination. With the assumption that source and destination are aware of their corresponding channels to the relay, they can then subtract the back-propagating self-interference. Although two-way relaying is able to significantly recover the spectral efficiency loss from the half-duplex constraint, this scheme had the unfair advantage in utilising 50% more total transmit power when compared to one-way relaying - with a fair power constraint, two-way relaying outperformed one-way relaying when the SNR is only above a certain threshold [115]. In this thesis, we will focus more on dual-hop half-duplex relaying.

5.2 Cooperative signalling strategies

Given the extra degrees of freedom that exist in a relaying network (compared to direct link) or channel, several different relaying topologies are possible apart from the basic three-terminal communication model originally analysed by [31]. Multi-hop wireless systems have the potential to offer improved capacity and coverage over single hop radio access systems [116]. Fig. 5.2 shows five different relay network structures. Fig. 5.2a is a two-hop parallel relay harnessing cooperative diversity through many distributed terminals operating simultaneously during a particular hop. A dual-hop mode is mainly used to increase capacity of a link within the coverage area of BS and is of lower implementation cost. In this mode, the relay does not forward framing information, but operates in a centralised scheduling mode. A natural extension is to consider multi-hop parallel relaying of Fig. 5.2b. In multi-hop relays, the relay stations (RSs) generate their own framing packet or forward those provided by the BS depending on the scheduling approach. This mode is used for coverage extension. Multi-hop mode bears a higher complexity at the RS and is also subject to inter RS cell interference. The three relay structures at the bottom of Fig. 5.2 are other variants of dual hop-parallel relaying, which have certain advantages over the previous schemes, and will be discussed shortly.

Figure 5.2: Cooperative relaying topologies showing two signal reinforcement schemes above, and three selection-based schemes below. Yellow circles indicate time slots.

Depending on how the RS handles the received BS data, the relays can be classified into regenerative or non-regenerative systems. In the former, each relay fully decodes the received signal from the first hop and retransmits a decoded version into the second hop. This is more commonly referred to as decode-and-forward (DF) relaying [6]. Non-regenerative systems are low complexity relays that only amplify the incoming signal without performing any sort of decoding. They are better known as amplify-and-forward (AF) relaying [6]. Several variants of these two methods exist, such as hybrid AF/DF relaying. In general we will first look at AF and DF strategies briefly and then discuss the different relaying structures in more detail later.

5.2.1 Amplify and Forward

As the name implies, the relays amplify the incoming signal and forward it to the destination by retransmitting a noisy but amplified version. Non-regenerative systems

Figure 5.3: Basic half-duplex three node relaying scheme employing either AF or DF

can be classified into two subcategories: *CSI assisted relays* and *fixed gain relays*. CSI-assisted relays use instantaneous CSI knowledge from the first hop to calculate a second hop gain as a function of the instantaneous channel gain. CSI assisted relays are also known as *channel noise assisted amplify-and-forward* (CNA-AF) relays. Fixed gain relays are called *blind-relays* and do not need instantaneous CSI knowledge from the first hop at the relay. They perform sub-optimally when compared to CSI-assisted relays, although they are attractive from a practical aspect for their low complexity and ease of deployment. Another category of fixed gain relays, called *semi-blind* relays, use statistical knowledge from the first hop to determine second hop gain. The performance of semi-blind relays become comparable to CSI-assisted relays, and sometimes even better in the low SNR regime [117]. For analysing AF relaying, let us consider the basic three node relay wireless half-duplex system with one source, destination and intermediate relay, as shown in Fig. 5.3. Here, terminal S is communicating with terminal D through an intermediate relay and a direct path. Assume that terminal S is transmitting a signal s with an average power of $\mathcal{E}_1$ and let $\mathcal{E}_2$ be the power of the transmitted signal at the relay. Let the instantaneous channel gain of the first hop path $S \rightarrow R$, be denoted as h_{SR} and h_{RD} be the instantaneous path gain of the second hop channel $R \rightarrow D$, characterised as $\mathcal{CN}(0, \sigma_{h_1}^2)$ and $\mathcal{CN}(0, \sigma_{h_2}^2)$ respectively. The direct path $S \rightarrow D$ is given as h_{SD} , which is $\mathcal{CN}(0, \sigma_{h_0}^2)$. Then for an isolated symbol transmission the received signal at the destination in two

time instants is given as:

$$\begin{aligned} y_{SD}(n) &= h_{SD}s(n) + n_D(n) \\ y_{RD}(n+1) &= Gh_{RD}(h_{SR}s(n) + n_R(n)) + n'_D(n+1) \end{aligned} \quad (5.1)$$

where n_D , n'_D and n_R are the AWGN noise samples at the destination and relay respectively. Assuming equal variance N_0 , the effective instantaneous SNR γ_{eq} at the destination can be written as :

$$\gamma_{eq} = \gamma_0 + \frac{\gamma_1 + \gamma_2}{\gamma_1 + \gamma_2 + 1} \quad (5.2)$$

with $G^2 = \frac{1}{\mathcal{E}_1|h_{SD}|^2 + N_0}$ and $\gamma_0 = \frac{\mathcal{E}_1|h_{SR}|^2}{N_0}$, $\gamma_1 = \frac{\mathcal{E}_1|h_{SD}|^2}{N_0}$ and $\gamma_2 = \frac{\mathcal{E}_2|h_{RD}|^2}{N_0}$ as the instantaneous SNRs of each hop respectively. Average receive SNR per symbol per path for the direct, first and second hop is $\bar{\gamma}_0 = \sigma_{h_0}^2 \mathcal{E}_1 / N_0$, $\bar{\gamma}_1 = \sigma_{h_1}^2 \mathcal{E}_1 / N_0$ and $\bar{\gamma}_2 = \sigma_{h_2}^2 \mathcal{E}_2 / N_0$ respectively. The choice of gain G aims to invert the fading effect of the first channel while limiting the output power of the relay if the fading amplitude of the first hop becomes low.

However, CSI-assisted relaying requires a continuous estimate of the channel fading amplitude, which may make this choice of gain not always feasible from a practical point of view. By contrast, blind relays introduce fixed gains to the received signal regardless of the fading amplitude over the first hop. If we let $C = \mathcal{E}_2 / G^2 N_0$, then the effective end-to-end SNR can be re-written as:

$$\gamma_{eq} = \gamma_0 + \frac{\gamma_1 \gamma_2}{\gamma_2 + C} \quad (5.3)$$

Presently a lot of work has been published based on AF relaying. It can also achieve a full diversity order of two because the destination receives two independent signals from the direct-path and the relay-path [6]. Hasna and Alouini analyzed the end-to-end performance of two-hop wireless communication systems [117]. Anghel and Kaveh derived the exact average SER [118]. In [119], Tsiftsis *et al.* derived outage probability and MGF (moment generating function) in a dual-hop scheme with three nodes for both variable and fixed gain relaying in Nakagami- m channels. Ribeiro *et al.* derived SERs for general multi-hop relay networks [120].

5.2.2 Decode and Forward

In DF based relaying, each relay terminal detects the incoming signal and retransmits the detected symbol. To analyse, consider the simple case of a two-hop DF network in Fig. 5.3. Following [121], an expression for the end to end SNR, accurate in the high SNR regime, is given by:

$$\gamma_{eq} = \gamma_0 + \gamma_{SRD} = \gamma_0 + \min(\gamma_1, \gamma_2) \quad (5.4)$$

The two-hop $S \rightarrow R \rightarrow D$ channel has end to end BER given by $P_{SRD}(\gamma_{SR}, \gamma_{RD}) = P_{SR} + P_{RD} - 2P_{RD}P_{SR}$, given by the fact that errors at the destination occur either when the $S - R$ transmission is received correctly and the $R \rightarrow D$ transmission is received in error; or, when the $S \rightarrow R$ transmission is received in error and the $R \rightarrow D$ transmission is received correctly. One can think of the BER contribution of the $S \rightarrow R \rightarrow D$ link as the error probability at the receiver of an equivalent one-hop AWGN link whose output SNR:

$$\gamma_{SRD} = \frac{1}{\alpha} [Q^{-1}(P_{SRD})]^2 \quad (5.5)$$

and α is a constant which depends on the underlying constellation. Since the pioneering work of Sendonaris *et al.* [122], [123], many studies on DF relaying have been done [121], [124], [125], [126], [127].

5.2.3 Coded Cooperation

Coded cooperation was proposed by Nosratinia and Hunter [128], [129] where cooperation integrates channel coding. Other notable recent works include [130], [131], [132]. Coded cooperation works by sending different portions of each users code word via two independent fading paths. The basic idea is that each user tries to transmit incremental redundancy to its partner. Whenever that is not possible, the users automatically revert to a non-cooperative mode. The key to the efficiency of coded cooperation is that all this is managed automatically through code design, with no feedback between the users. For the coded cooperation method, a natural issue is the

possibility of designing a better coding scheme. In their seminal work [128], Hunter and Nosratinia use RCPC (rate-compatible punctured convolutional codes), while in [131], Janani *et al.* applied turbo codes to the coded cooperation framework.

5.2.4 Relay Configurations

Here we will discuss the different relay schemes as shown in Fig. 5.2, paying particular attention to their performance benefits and limitations.

5.2.4.1 Parallel relaying

In regular cooperative diversity networks with N relaying nodes, such as Fig.5.2a, $N + 1$ orthogonal channels (or time slots) are employed. These are otherwise known as all-participate or AP schemes. With MRC performed at the destination, the total SNR γ_{eq} at the destination with AF strategy can be written as:

$$\gamma_{eq} = \gamma_0 + \sum_{i=1}^N \frac{\gamma_{SR_i} \gamma_{R_i D}}{\gamma_{SR_i} + \gamma_{R_i D} + 1} \quad (5.6)$$

where γ_{SR_i} is the instantaneous SNR between S and relay R_i , and $\gamma_{R_i D}$ is the instantaneous SNR between R_i and D . Tsiftsis *et al.* [133] and others [134] analyse the BER and outage probability analysis for such systems. Ribeiro *et. al* [120] studied such a multi-branch cooperative network, and showed that a diversity order of $N + 1$ can be achieved.

An adaptive DF scheme was proposed by Laneman and Wornell [135], where relays that successfully decode information in the first hop are allowed to transmit in the second hop. A decoding set $\mathcal{D}(s)$ with a transmitting node s , is defined as the relays which participate in second hop transmission which have capacity above the required rate R . (For the AF transmission, $\mathcal{D}(s)$ is simply the entire set of cooperating relays). Let γ_i represent the random variable for the SNR on the i^{th} link, which takes into account fading on the source to i^{th} relay link and fading on the i^{th} relay to destination.

Then the PDF of γ_i is given by [136]:

$$f_{\gamma_i}(\gamma) = f_{\gamma_i|\text{link down}}(\gamma)Pr[\text{link down}] + f_{\gamma_i|\text{link up}}(\gamma)Pr[\text{link active}] \quad (5.7)$$

If the i^{th} link is down, $f_{\gamma_i|\text{link down}}(\gamma)$ is a delta function, and $Pr[\text{link down}]$ is $Pr[\gamma_{SR_i} < \gamma_{th}]$, with γ_{th} set as the threshold value. Similarly the i^{th} link is up with a probability of $1 - Pr[\gamma_i < \gamma_{th}]$ with the corresponding $f_{\gamma_i|\text{link up}}(\gamma)$ determined by the PDF of $R_i - D$ path. The effective PDF formulation is useful when the channels are assumed to be independent but not identically distributed. It was earlier shown [135], that such a DF relaying scheme achieves a diversity order of $N + 1$, equal to the number of cooperating nodes. However, for finite SNR levels this did not necessarily mean better performance [135], [137], with SNR loss due to bandwidth inefficiency being exponential in N . A space-time diversity protocol applied to the DF scheme [135], in which all relays transmit simultaneously over the same sub-channel using a suitable space-time code, experiences SNR loss which is only linear in N while maintaining same diversity benefits. In fixed DF schemes, the assumption of perfect capability of decoding the CRC at the relay is relaxed, and hence the relays always decode and re-encode before transmission. The end-to-end instantaneous SNR can be characterised in a similar manner as in the analysis in Section 5.2.2. ([138], [139] analyse fixed DF systems).

5.2.4.2 Multi-hop relaying

In multi-hop relayed transmission, several cooperating intermediate nodes act to process the received signal from the preceding transmitting terminal and then forward it to the next terminal. Particularly, for a single branch multi-hop network, it was shown in [120] that although diversity order of one is achieved, the scheme is attractive in reducing BER when the harmonic mean of the average multi-hop SNRs is larger than the single hop SNR divided by the number of hops. The performance of such multi-hop relaying schemes has been comprehensively studied in [140], [141], [142], [143], [144], [145], [146].

5.2.4.3 Relay selection

Recently, relay selection as a form of opportunistic transmission, has been considered as an attractive technique to exploit inherent spatial diversity in cooperative networks. There is presently a lot of research interest in cooperative relay selection schemes, where relay nodes are selected based on system parameters such as maximum power or SNR [5]. Relay selection can overcome synchronisation issues and limit excess bandwidth usage while preserving diversity benefits. However distributed and centralised relay selection schemes have drawbacks in that they are time sensitive and require high channel feedback from all links involved [126]. The advantages of the regular cooperative diversity come at the expense of spectral efficiency since the source and all relays must transmit on orthogonal channels in order to avoid interfering with each other [6]. With $N + 1$ orthogonal channels, a bandwidth penalty is incurred. While a space-time relaying scheme can be more bandwidth efficient, such architectures are perhaps not viable to immediate practical implementation since they assume the availability of distributed space-time codes across the relay nodes. The inefficient utilization of the channel resources in regular cooperative diversity networks can also be enhanced by using the best-relay selection scheme. Such a scheme is shown in Fig. 5.2c and only the best-relay is selected to retransmit to the destination [5]. Hence, two channels or two time slots (regardless of the number of relays) only are required in this case. The best-relay selection scheme for cooperative networks has been introduced by Bletasas *et al.* [5], they called it *opportunistic relaying* (OR). According to opportunistic relaying, a single relay among a set of relay nodes is selected, depending on which relay provides for the best end-to-end path between source and destination. Using outage probability analysis, the same authors showed that this scheme has the same diversity order as the cooperative diversity using space-time coding for both DF and AF schemes. Here the best relay is chosen prior to transmission, utilizing prior information about instantaneous signal powers at each relay across both hops. [147], the SER at high SNR, for OR is analysed using outage probability and verified that it achieves full diversity gain of the cooperative system. The equivalent received SNR

under the best relay scheme maximised over both hops is formulated as:

$$\gamma_{eq} = \gamma_0 + \max_i \frac{\gamma_{SR_i} \gamma_{R_i D}}{\gamma_{SR_i} + \gamma_{R_i D} + 1} \quad (5.8)$$

Interestingly, a coding gain improvement of OR over AP was achieved when equal power division among transmitting nodes were assumed in both systems. Zhao *et al.* [148], analyse the same system, and derive tight lower bounds for performance metrics such as BER, outage, moments of SNR, and verify the diversity order of $N + 1$ in such as system. Similarly, Duong *et al.* [139], derived outage and BER expressions for fixed DF relay selection over Nakagami- m channels.

Another relay selection scheme consisting of a source, a destination, and multiple relays is called *selection relaying* (SC), wherein the first time slot, the source transmits a symbol to all the relays, and relay selection is employed for the second hop. In threshold-based SC DF, a set of relays that correctly decode the symbol is determined in the first slot. In the second time slot, among the relays in the determined set, only a single best relay associated with the maximum SNR in the second hop retransmits the decoded symbol to the destination. Beres and Adve [149] analyzed the capacity outage probability of the best-relay selection scheme with DF, and showed that it outperforms distributed space-time codes for network with more than three relaying nodes. This gain is due to the efficient use of power by the best-relay selection scheme networks. Some other relevant work in the SC field can be found [150], [151]. In particular Bletsas *et al.* obtained interesting results that showed SC and OR having the same outage performance under the adaptive DF scheme, and that the two schemes were outage-optimal under the aggregate power constraint at the relays [5]. Michalopoulos *et al.* derived the outage probability and average BER of a threshold-based SC DF relaying when there was a direct path [152] and showed that SC performs slightly better in terms of outage probability, while in terms of BER both systems outperformed each other depending on the SNR threshold which determines the set of relays participating in the forwarding process. Ikki and Ahmed [153], extending the work of Beres and Adve [149], use a single decoding best relay out of the decoding set in the first hop, to forward the signal on to the destination in the second hop. The

authors derive closed form expressions for outage probability and channel capacity and show that the best relay scheme maintains a full diversity order. A similar network is analysed by Xu *et al.* [154] for Nakagami- m channels, and the diversity order found to equal sum of all diversity orders between each source-relay pair, as $\sum_{i=0}^N \min(m_{sd}^i, m_{rd}^i)$, where m_{sd}^i, m_{rd}^i are the Nakagami- m parameters in the first and second hop paths. Duong *et al.* [139], also used a best relay which is selected over both hops for fixed DF, (and as mentioned previously their system is presented for Nakagami- m channels). However, when AF is employed it has been shown by Costa and Aissa [155] that, for clustered fixed gain relaying, OPR always outperform SC.

Partial relay selection, another form of relay selection similar to selection relaying, as proposed by Krikidis *et al.* [156], alleviates the overhead of complete CSI information being transmitted across both hops. This has advantages in prolonging network lifetime. Here the selection is performed in the source to relay link, where in a best relay is selected for transmission in the second hop. Recently, Suraweera *et al.* [157], and DaCosta [158], [159] studied fixed gain AF with first hop relay selection. The effective end-to-end SNR for AF partial relay selection is formulated as:

$$\gamma_{eq} = \gamma_0 + \frac{\max_i(\gamma_{SR_i})\gamma_{R_bD}}{\max_i(\gamma_{SR_i}) + \gamma_{R_bD} + 1} \quad (5.9)$$

Here R_b is the best relay selected over the first hop.

5.2.5 Multi-antenna relaying

Relaying under single antenna equipped nodes has been researched intensely. Relay networks with multi-antenna terminals have recently received considerable attention [160], [161]. By deploying multiple antennas at relays or at terminals, additional capacity or diversity improvement can be achieved. Diversity combining techniques through multiple antennas can be employed for signal re-enforcement at the relays, brought about by diversity and array gains. Similarly capacity improvement is possible through spatial multiplexing gain, without additional power and bandwidth expenditure. Deployment of multi-antenna or infrastructure relay nodes is particularly

beneficial for destination terminals with small number of receive antennas. Further, when the number of cooperating nodes are less, use of multiple antennas can be beneficial. Yuksel *et al.* [162] presented diversity-multiplexing trade offs for a multi-antenna relay network. Lee *et al.* derived the outage probability of the AF protocol in a multi-antenna relay network adopting transmit antenna selection [27]. Finally Adinoyi and Yanikomeroglu [163], considered multi-antenna relaying with threshold based selection combining for DF is analysed in Nakagami- m channels.

5.3 Summary

Relaying has developed as a key technology in improving throughput and coverage, and in better utilising resources such as bandwidth and power. This chapter gave a brief overview of standard relaying strategies, and discussed common relaying schemes such as AF and DF. Given a set of cooperating nodes, it is possible to deploy different types of relaying schemes between the source and destination, each of them well suited for a certain application, and at the same time having certain drawbacks. An all-participate scheme does not require any feedback between source and relays or relays to destination, however it is bandwidth inefficient. Co-operative relay selection is found attractive in preserving diversity benefits of a cooperative network, while saving bandwidth with minimal implementation overhead. Opportunistic relay selection requires CSI across both hops, thereby needing some means of central coordination, while partial relay selection chooses the best relay over the first or second hop, alleviating the CSI monitoring overhead. To date, most work in partial relaying concerns single antenna devices and far less work has been done on multi-antenna systems. In the next section, we will combine the advantages of relay selection with the advantages of multi-antenna technology to consider the merged techniques of partial selection relaying with multi-antenna nodes and/or destination nodes.

Chapter 6

Multi-Antenna relay selection

“Equations are more important to me, because politics is for the present, but an equation is something for eternity.”

Albert Einstein

The previous chapter discussed the advantages and simplicity of relay selection in comparison to other relaying schemes. In this chapter, we will apply concepts of selection relaying for multi-antenna systems employing the less complex amplify and forward relaying [117] (rather than a regenerative system which would fully decode the data packet at the selected relay prior to transmission over the second hop). The principles of transmit selective diversity discussed in Chapters 3 and 4 are applied to the problem of partial relay selection (PRS) in dual hop amplify and forward networks. We aim to study the performance gain in PRS schemes brought about by exploiting other degrees of freedom, such as the use of multiple antennas, varying transmitter gains across nodes, or the presence of a strong channel in a hop.

6.0.1 Related work

Relay selection based on instantaneous channel measurement is employed— a low complexity and efficient method of harnessing distributed spatial diversity across a cooperative virtual antenna array [5]. Previous systems [164], studied the performance benefits of best-relay selection (BRS) offering maximum SNR gain across both hops, however, the study was limited to single antenna nodes. Although BRS in single antenna systems offers full diversity gain equal to the number of relays used, the system is dependent on a centralised controller to obtain CSI across both hops. Partial relay selection, as proposed by Krikidis *et al.* [156], alleviates the overhead of complete

CSI information being transmitted across both hops (this assumed variable relay gain). This has advantages in prolonging network lifetime. Recently, Suraweera *et al.* [157], and DaCosta and Aissa [158,165] analysed fixed gain AF with first hop relay selection. However, both systems were restricted to single antenna terminals at the relay nodes and destination. Interestingly, performance of the single antenna PRS schemes was not found to increase appreciably when the number of relay terminals increases. Instead, gains were limited by the quality of the non-selective second hop from relay to destination.

Little work has been published concerning performance analysis of relaying in multi-antenna systems. However, recent studies on MIMO relays, prompted by standardization discussions for 3GPP LTE Advanced systems and IEEE802.16j, have demonstrated improvements in link reliability and spectral efficiency [160,161] over SISO alternatives. Infrastructure-based (fixed) multi-antenna relays have also demonstrated diversity and capacity benefits for wireless terminals with limited number of antennas [163], [166]. Source beamforming has been studied [167] for single antenna relay (with variable gain) and single antenna destination, in which it was observed that performance gains quickly deteriorate when the number of transmit antennas exceeds two. In [159] Costa and Aissa studied analysed a relay network with multiple destination receive antennas in conjunction with source beamforming and the moment generating function approach was used to obtain BER performance. However, increasing the transmit or receive antennas at either source or destination end to boost performance may not be feasible in many situations. Moreover, in some cases increasing the number of receivers might prove detrimental to overall improvement. When the number of transmitters or receivers at the end terminals is to be kept small, the use of multiple intermediate relays (where in one or more ‘good’ relays are selected to convey information) to improve performance is a potential solution.

6.0.2 Contributions

This chapter investigates methods to improve the performance of previously proposed PRS schemes by exploiting different degrees of freedom that may be available to the system. We investigate the performance of PRS in a multi-antenna system, where either the source, destination or relay may be employed with more than one antenna. Depending upon the scenario, a particular type of configuration may be employed to suit the situation. The use of the versatile Nakagami- m statistical model, known to be an accurate fit for urban radio multi-path propagation, enables the investigation of different fading scenarios (from source to relay and relay to destination respectively) used to bring about additional gain.

Different arrangements are possible given the extra degrees of freedom and fading scenarios. Each arrangement will have its own advantages and will suit a particular application. In this chapter we consider four different scenarios.

- In Section 6.1, we cover single antenna relay selection when combined with multiple antenna BSs. In this case relays may be assumed to be blind or semi-blind so that only fixed amplification is employed, obviating the need for channel estimation at the relay; this has several advantages, especially when mobile devices are used as relays themselves.
- In cases where the end terminals are single antenna devices, fixed infrastructure multi-antenna relays employing MRC and MRT may be deployed for effective source to destination communication. Section 6.2 deals with this scenario, where CNA-AF systems in Nakagami fading channels are analysed as applied to multi-antenna relay selection schemes.
- When the source is not completely shadowed from the destination, a weak path between the source and destination might exist. This can be exploited to gain additional performance benefits. In 6.3 we discuss this scenario in a multi-antenna relay selection scenario.

- In Section 6.4, we consider the problem of outdated feedback in relay selection in CNA-AF systems. This scheme is advantageous in modelling a high Doppler fading scenario, where a mobile source communicates to a remote station via several intermediate nodes. Channel prediction is then applied to improve performance in relay selection systems. Specifically, we derive closed form solutions for performance metrics (verified through simulations) for all the above scenarios. We also discuss in detail, the effect of multiple-antennas at the relay and destination as well as power imbalance between hops.

Although it might be obvious that multiple antennas or increased transmit powers should increase system performance, we show that a careful choice of system parameters are vital in achieving performance improvement over other PRS schemes. Significant insights into the performance, as a function of the fading and power imbalances between hops are also presented.

6.1 Dual-Hop Fixed Gain Multi-Antenna AF Relay Selection

Introduction

This section explores a system in which a multi antenna source and destination communicate with the aid of an intermediate, single antenna, relay which is selected from several candidates by the source transmitter. In the following sections we formulate the multi-antenna AF problem and consider the effects of multiple antennas on the E2E performance. We discuss blind, semi-blind AF multi-antenna relaying in Rayleigh fading.

We analyse a beamforming system for Rayleigh fading channels with multiple relay nodes, each equipped with a single antenna, as in Fig. 6.1. Best relay selection, based on instantaneous SNR, is made at the transmitter and, to prevent performance saturation over the first hop, the second hop link is strengthened via multiple receive

Figure 6.1: Dual-Hop relay selection system equipped with multiple receive antennas at source and destination.

antennas at the destination. We study this system in terms of outage probability (P_{out}) and SER, and derive the moment generating function (MGF), statistical moments of the end-to-end SNR. Since the relays are assumed to be blind or semi-blind to channel state information (CSI), they offer only fixed amplification to the incoming signal—easier to implement than a variable gain AF system, at the expense of some small performance loss. We also derive the amplification gain for the semi-blind scenario. Performance improvement over previous fixed gain PRS schemes such as [157], [158], [165] are realised.

6.1.1 System model

We consider a relay network, as in Fig. 6.1, where source S equipped with N_{ts} transmit antennas, communicates with relatively closely spaced N_t single antenna relay nodes $R_1, R_2, \dots, R_{N_t}$. In the absence of a direct path, a single relay selected by the transmitter, relays data to destination D which is equipped with N_{rd} receive antennas. Based on pilot data broadcast by the relays, instantaneous CSI is measured at the source itself, thus alleviating the need for a feedback channel. The channel matrix at S during initial channel estimation is denoted as $\mathbf{H}$, of dimension $N_{ts} \times N_t$ with complex fading coefficients h_{ij} , $1 \leq i \leq N_{ts}$, $1 \leq j \leq N_t$ as its entries. The channels h_{ij} are characterised as spatially independent and identically distributed (i.i.d) complex Gaussian random variables with distribution $\mathcal{CN}(0, 1)$. Also, they change slowly enough that multiple training signals can be transmitted without significant

frame overhead. The source computes the instantaneous channel power for each relay as $p_j = \sum_{i=0}^{N_{ts}} |h_{ij}|^2$ and picks a single best relay $1 \leq l \leq N_t$, corresponding to the maximum power: $p_{max} = \max_{1 \leq j \leq N_t} p_j$. The first transmission hop to this relay, denoted as $S \rightarrow R_l$, has channel $\mathbf{h}_{SR_l}$. This is the $(N_{ts} \times 1)$ channel vector being the l^{th} column of channel matrix $\mathbf{H}$, assumed to be correctly estimated at S . To perform transmit diversity beamforming, S computes an optimal beamforming vector [23] $\mathbf{w}_{SR_l} = \frac{\mathbf{h}_{SR_l}^H}{\sqrt{\mathbf{h}_{SR_l}^H \mathbf{h}_{SR_l}}}$ after which data is transmitted on the $S \rightarrow R_l$ channel. $(\cdot)^H$ is the Hermitian transpose. The signal received at the relay is amplified by fixed gain G , and transmitted to D via the $R_l \rightarrow D$ channel $\mathbf{h}_{R_l D}$. This is a $(N_{rd} \times 1)$ channel vector which is $\mathcal{CN}(0, \mathbf{I}_{N_{rd}})$ distributed, also assumed to be estimated correctly at D ($\mathbf{I}$ is an identity matrix). Here, the signal is finally demodulated, employing MRC [3] at the receiver with $\mathbf{w}_{R_l D} = \frac{\mathbf{h}_{R_l D}^H}{\sqrt{\mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}}$ as the optimal MRC weight vectors at the destination. Then the final combined output at the destination is:

$$y_{R_l D} = G \mathbf{w}_{R_l D} \mathbf{h}_{R_l D} (s \mathbf{w}_{SR_l} \mathbf{h}_{SR_l} + n_{R_l}) + \mathbf{w}_{R_l D} \mathbf{n}_D \quad (6.1)$$

where s is the transmitted symbol, n_{R_l} is the AWGN noise sample at the relay characterised as $\mathcal{CN}(0, N_0)$, while $\mathbf{n}_D$ is the noise vector modelled as $\mathcal{CN}(0, N_0 \mathbf{I}_{N_{rd}})$ where N_0 is the variance of AWGN. The equivalent SNR at the destination is then calculated as:

$$\gamma_{eq} = \frac{\frac{\mathcal{E}_1 \mathbf{h}_{SR_l}^H \mathbf{h}_{SR_l}}{N_0} \frac{\mathcal{E}_2 \mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}{N_0}}{\frac{\mathcal{E}_2}{N_0 G^2} + \frac{\mathcal{E}_2 \mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}{N_0}} \quad (6.2)$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ are the powers of the transmitted signal at the source and relay. If we keep $C = \frac{\mathcal{E}_2}{N_0 G^2}$ constant for fixed gain G , then (6.2) can be simplified as:

$$\gamma_{eq} = \gamma_1 \gamma_2 / (\gamma_2 + C) \quad (6.3)$$

where $\gamma_1 = \frac{\mathcal{E}_1 \mathbf{h}_{SR_l}^H \mathbf{h}_{SR_l}}{N_0}$ and $\gamma_2 = \frac{\mathcal{E}_2 \mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}{N_0}$ represent the instantaneous SNRs of the first and second hops respectively. The average receive SNR for each diversity path in the first and second hop are $\bar{\gamma}_1 = \frac{\mathcal{E}_1}{N_0}$ and $\bar{\gamma}_2 = \frac{\mathcal{E}_2}{N_0}$ respectively. In the following section we will formulate the statistics of each hop and use these to derive precise expressions for P_{out} , SER, MGF, moments of end-to-end SNR γ_{eq} , and the gain G in a semi-blind relay scenario.

6.1.2 Outage Probability

The outage probability P_{out} at any given average received SNR is defined as the probability that the instantaneous SNR is less than a threshold, γ_T . In the relay system under consideration this is expressed as [117]:

$$\begin{aligned} P_{out}(\gamma_T) &= P_{out}(\gamma_{eq} < \gamma_T) \\ &= \int_0^\infty Pr\left[\frac{\gamma_1\gamma_2}{\gamma_2 + C} < \gamma_T | \gamma_2\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \end{aligned} \quad (6.4)$$

where $f_{\gamma_2}(\gamma_2)$ is the PDF of the $R_l \rightarrow D$ channel. To solve (6.4) we first need $F_{\gamma_1}(\cdot)$, the CDF of γ_1 , so that firstly, the inner probability term is solved. Making use of order statistics [168, p.246], the PDF $f_{\gamma_1}(\gamma)$ can be expressed as follows:

$$f_{\gamma_1}(\gamma) = N_t f_{\gamma_0}(\gamma) [F_{\gamma_0}(\gamma)]^{N_t-1} \quad (6.5)$$

where $f_{\gamma_0}(\gamma)$ is the PDF of the SNRs of the individual and independent $S \rightarrow R_{i=1,\dots,N_t}$ links, characterised by a gamma PDF $\mathcal{G} \sim (N_{ts}, \bar{\gamma}_1)$ given by

$$f_{\gamma_0}(\gamma) = \frac{\gamma^{N_{ts}-1}}{\bar{\gamma}_1^{N_{ts}} (N_{ts} - 1)!} \exp\left(-\frac{\gamma}{\bar{\gamma}_1}\right) \quad (6.6)$$

with their CDF given as:

$$F_{\gamma_0}(\gamma) = 1 - \exp(-\gamma/\bar{\gamma}_1) \sum_{j=0}^{N_{ts}-1} (1/j!)(\gamma/\bar{\gamma}_1)^j \quad (6.7)$$

To obtain a series expansion of (6.5), $[F_{\gamma_0}(\gamma)]^{N_t-1}$ is expanded first via the binomial theorem followed by multinomial theorem, and later combined with (6.6). Then, grouping exponential terms and powers of γ , the definite integral of (6.5) is then solved to obtain $F_{\gamma_1}(\cdot)$ as follows:

$$\begin{aligned} F_{\gamma_1}(z) &= \int_0^z f_{\gamma_1}(\gamma) d\gamma \\ &= \frac{N_t!}{\bar{\gamma}_1^{N_{ts}} (N_{ts} - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i! (N_t - 1 - i)!} \\ &\quad \cdot \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i, j) (1/\bar{\gamma}_1)^j \int_0^z \gamma^{j+N_{ts}-1} \exp(-(i+1)\gamma/\bar{\gamma}_1) d\gamma \end{aligned} \quad (6.8)$$

where z is a constant and $\eta_{N_{ts}}(i, j)$ [98, 0.314] (or see A.1) are the coefficients of $\varphi^j, j = 0, 1, \dots, i(N_{ts} - 1)$ in the expansion of $(\sum_{j=0}^{N_{ts}-1} \varphi^j / j!)^i$, used in the expansion of $[F_{\gamma_0}(\gamma)]^{N_t-1}$. Finally, using [98, eqn.(3.381.1)] (6.8) is solved and $F_{\gamma_1}(z)$ is cast in a complimentary CDF form as follows:

$$F_{\gamma_1}(z) = 1 - \frac{N_t!}{(N_{ts} - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i \exp(-(i+1)z/\bar{\gamma}_1)}{i!(N_t - 1 - i)!} \cdot \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i, j)(N_{ts} + j - 1)! \sum_{l=0}^{j+N_{ts}-1} \frac{z^l}{\bar{\gamma}_1^l l! (i+1)^{j+N_{ts}-l}} \quad (6.9)$$

Moving our focus back to solving (6.4), we first substitute $z = \gamma_T \left(1 + \frac{C}{\gamma_2}\right)$ in (6.9) to obtain $Pr\left[\frac{\gamma_1 \gamma_2}{\gamma_2 + C} < \gamma_T | \gamma_2\right]$. Then a binomial expansion of $(1 + C/\gamma_2)^l$ is done. Together, this is then coupled with $f_{\gamma_2}(\gamma_2)$, the PDF of the $R_l \rightarrow D$ which is $\mathcal{G} \sim (N_{rd}, \bar{\gamma}_2)$:

$$f_{\gamma_2}(\gamma_2) = \frac{\gamma_2^{N_{rd}-1}}{\bar{\gamma}_2^{N_{rd}} (N_{rd} - 1)!} \exp\left(-\frac{\gamma_2}{\bar{\gamma}_2}\right) \quad (6.10)$$

Collecting the power and exponential terms of γ_2 , the integration of (6.4) is facilitated using [98, eqn.(3.471.9)] (see A.4.2). Thus P_{out} (or equivalently $F_{\gamma_{eq}}(\gamma_T)$): the CDF of γ_{eq} is obtained as:

$$\begin{aligned} P_{out}(\bar{\gamma}_1, \bar{\gamma}_2, N_t, N_{ts}, N_{rd}, \gamma_T) &= \int_0^\infty F_{\gamma_1}\left[\gamma_T(1 + C/\gamma_2)\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \\ &= 1 - \frac{N_t!}{(N_{ts} - 1)!(N_{rd} - 1)! \bar{\gamma}_2^{N_{rd}}} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i, j)(N_{ts} + j - 1)! \\ &\quad \cdot \sum_{l=0}^{j+N_{ts}-1} \frac{2}{(i+1)^{j+N_{ts}-l} l!} \left(\frac{\gamma_T}{\bar{\gamma}_1}\right)^l \exp\left(-\frac{\gamma_T(i+1)}{\bar{\gamma}_1}\right) \sum_{p=0}^l \binom{l}{p} C^p \\ &\quad \cdot \left[\frac{C\gamma_T(i+1)\bar{\gamma}_2}{\bar{\gamma}_1}\right]^{\frac{(N_{rd}-p)}{2}} \mathbf{K}_{N_{rd}-p}\left(2\sqrt{\frac{\gamma_T C(i+1)}{\bar{\gamma}_1 \bar{\gamma}_2}}\right) \end{aligned} \quad (6.11)$$

where $\mathbf{K}_v(\cdot)$ is the modified Bessel function of the second kind of order v [98]. As a useful check to our derivation, we can reduce the expression to a simplified case that matches previously published expressions. To do this set $N_{ts} = N_{rd} = N_t = 1$, and then (6.11) reduces to [117, eqn.(9)].

6.1.3 Error probability

The mean symbol error rate (SER) at an average SNR is found by averaging the probability of symbol error in AWGN over the fading distribution $f_{\gamma_{eq}}(\gamma)$ of the relay scheme, in slow fading scenarios (as explained in Section 2.21). SER for coherent demodulation can be found from:

$$\bar{P}_s(\bar{\gamma}_1, \bar{\gamma}_2, N_t, N_{ts}, N_{rd}) \approx \int_0^\infty P_s(\gamma) f_{\gamma_{eq}}(\gamma) d\gamma \quad (6.12)$$

where $P_s(\gamma) \approx \alpha Q(\sqrt{\beta\gamma})$ where α and β are constellation specific constants and $Q(\cdot)$ is the Q -function [21]. Eqn. (6.12) can also be expressed using $F_{\gamma_{eq}}(\cdot)$ [164], so that $\bar{P}_s$ is derived without having to obtain the PDF $f_{\gamma_{eq}}(\cdot)$:

$$\bar{P}_s(\bar{\gamma}_1, \bar{\gamma}_2, N_{ts}, N_t, N_{rd}, \alpha, \beta) \approx \frac{\alpha}{\sqrt{(2\pi)}} \int_0^\infty F_{\gamma_{eq}}(x^2/\beta) \exp(-x^2/2) dx \quad (6.13)$$

Using $\gamma_T = x^2/\beta$ in (6.11) and grouping powers of x , and the exponential term with $\exp(-x^2/2)$ in (6.13), an integral containing a combination of power, exponential, and Bessel functions is solved using [98, eqn.(6.631.3)](see A.3.2). (an intermediate step is shown in Appendix eqn. A.6.1). Thus the SER of the desired modulation scheme under relaying can be evaluated as :

$$\begin{aligned} & \bar{P}_s(\bar{\gamma}_1, \bar{\gamma}_2, N_t, N_{ts}, N_{rd}, \alpha, \beta) \\ & \approx \frac{\alpha}{2} \left[1 - \frac{N_t! \bar{\gamma}_2^{-N_{rd}}}{\sqrt{2\pi} (N_{ts} - 1)! (N_{rd} - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i! (N_t - 1 - i)!} \right. \\ & \quad \cdot \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i, j) (N_{ts} + j - 1)! \sum_{l=0}^{j+N_{ts}-1} \frac{1}{(i+1)^{j+N_{ts}-l} l!} \\ & \quad \cdot \left(\frac{1}{\beta \bar{\gamma}_1} \right)^l \sum_{p=0}^l \binom{l}{p} C^p (a_i C \bar{\gamma}_2)^{(N_{rd}-p)/2} \Gamma[l + 1/2] \Gamma[l + N_{rd} - p + 1/2] \\ & \quad \cdot \exp \left(\frac{b_i}{2a_i + 1} \right) \frac{(a_i + 1/2)^{(p-2l-N_{rd})/2}}{\sqrt{b_i}} \mathbf{W}_{(p-2l-N_{rd})/2, (N_{rd}-p)/2} \left(\frac{b_i}{a_i + 1/2} \right) \left. \right] \end{aligned} \quad (6.14)$$

where $b_i = \frac{C(i+1)}{\bar{\gamma}_1 \bar{\gamma}_2 \beta}$, $a_i = \frac{(i+1)}{\bar{\gamma}_1 \beta}$, $\Gamma[\cdot]$ is the gamma function. $\mathbf{W}_{k,\mu}(\cdot)$ is the Whittaker function [98] which can be evaluated in terms of the Bessel or hypergeometric ${}_2F_0(a, b; z)$ function.

Again as a check we can reduce our derivation to a simplified case (that has been published elsewhere) where source and destination are single antenna devices. In this case, we set $N_{ts} = N_{rd} = 1$; simplifying (6.14) and expanding $\mathbf{W}_{k,\mu}(\cdot)$, now as a series of Bessel $\mathbf{K}_v(\cdot)$ functions [169], where $k = -1/2$ and $\mu = 1/2$. This allows us to reduce (6.14) to [157, eqn.(12)]: the case for N_t relay nodes but source and destinations with only single antennas.

Moment Generating Function

The MGF approach [19] is a powerful tool for evaluating the SER for a wide variety of M -ary modulations such as the M -ary phase shift keying (M -PSK), M -ary differential phase shift keying (M -DPSK) and M -QAM. The MGF of γ_{eq} is given by $\mathcal{M}_{\gamma_{eq}}(s) = \mathbb{E}[\exp(-s\gamma_{eq})]$, where $\mathbb{E}[\cdot]$ is the expectation operator. For example, it is well known that the BER for binary DPSK is $\bar{P}_s = 0.5\mathcal{M}_{\gamma_{eq}}(1)$. It is convenient to express $\mathcal{M}_{\gamma_{eq}}(s)$ as a integral containing the complimentary CDF, $1 - F_{\gamma_{eq}}(\gamma)$, so that the tedious differentiation step to obtain the PDF $f_{\gamma_{eq}}(\gamma)$ from (6.11) can be skipped. This is because using integration by parts, it can be shown that:

$$\mathcal{M}_{\gamma_{eq}}(s) = 1 - s \int_0^\infty \exp(-\gamma_{eq}s)(1 - F_{\gamma_{eq}}(\gamma)) d\gamma \quad (6.15)$$

Using $F_{\gamma_{eq}}(\cdot)$ from (6.11) in the above equation and [98, eqn.(6.643.3)](see A.3.3), and after simplifications the closed form solution for the MGF can be written as:

$$\begin{aligned} \mathcal{M}_{\gamma_{eq}}(s) = & 1 - \frac{sN_t!\bar{\gamma}_1}{(N_{ts}-1)!(N_{rd}-1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i \exp\left(\frac{C(i+1)}{2\bar{\gamma}_2(\bar{\gamma}_1s+i+1)}\right)}{i!(N_t-1-i)!} \cdot \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i,j) \\ & \cdot \sum_{l=0}^{j+N_{ts}-1} \frac{(N_{ts}+j-1)!}{l(i+1)^{j+N_{ts}-l}} \sum_{p=0}^l \binom{l}{p} C^{\frac{N_{rd}+p-1}{2}} \frac{(i+1)^{\frac{N_{rd}-p-1}{2}}}{(\bar{\gamma}_1s+i+1)^{\frac{2l+N_{rd}-p-1}{2}}} \\ & \cdot \frac{\Gamma[N_{rd}-p+l+1]}{\bar{\gamma}_2^{(N_{rd}+p-1)/2}} \mathbf{W}_{-(2l+N_{rd}-p+1)/2, (N_{rd}-p)/2} \left(\frac{C(i+1)}{\bar{\gamma}_2(\bar{\gamma}_1s+i+1)} \right) \end{aligned} \quad (6.16)$$

As a check, consider selection among N_t single antenna terminals so that $N_{ts} = N_{rd} = 1$. Knowing that $W_{-\mu-1/2, \mu}(z) = \exp(z/2)z^{\frac{1}{2}-\mu}\Gamma(-2\mu, z) = \exp(z/2)z^{\frac{1}{2}-\mu}E_{2\mu+1}(z)$, where $E_n(z)$ is the exponential integral [98, p.xxxv], we can reduce our MGF to

exactly [165, eqn.(22)]. While for $N_{ts} = N_t = N_{rd} = 1$, using the recurrence relation of $E_n(z)$, this simplifies even more to [117, eqn.(12)].

6.1.4 Choice of gain parameter

If the relays have statistical knowledge about first hop fading, it is possible to set G to the average of the CSI-assisted variable relay gain [117] as follows:

$$G^2 = \mathbb{E} \left[\frac{\mathcal{E}_2}{\mathcal{E}_1 \mathbf{h}_{SR_i}^H \mathbf{h}_{SR_i} + N_0} \right] = \int_0^\infty \frac{\mathcal{E}_2}{N_0(\gamma_1 + 1)} f_{\gamma_1}(\gamma_1) d\gamma_2 \quad (6.17)$$

Using (6.5) and [98, eqn.(3.383.10)](see A.4.3) the gain parameter can then be solved as :

$$G^2 = \frac{\bar{\gamma}_2 N_t!}{(N_{ts} - 1)! \bar{\gamma}_1^{N_{ts}}} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \exp \left(\frac{(i+1)}{\bar{\gamma}_1} \right) \cdot \sum_{j=0}^{i(N_{ts}-1)} \frac{\eta_{N_{ts}}(i, j)}{\bar{\gamma}_1^j} \Gamma[j + N_{ts}] \Gamma \left[1 - j - N_{ts}, \frac{i+1}{\bar{\gamma}_1} \right] \quad (6.18)$$

where $\Gamma[\cdot, \cdot]$ represents the incomplete Gamma function [98]. From (6.18), C is obtained as $C = \frac{\mathcal{E}_2}{N_0 G^2}$. If this constant gain was implemented in a system, this would make the performance of semi-blind relays comparable to that of CSI-assisted (variable gain) relays, while at the same time keeping implementation complexity relatively low. In such a system, the gain parameter may be periodically advised to the relays as and when channel statistics changes.

6.1.5 Statistics of the end-to-end SNR

Here the closed form expression for μ_n , the n^{th} moment of the output SNR is derived. Important system measures such as average output SNR, variance or higher order central moments such as skewness and kurtosis parameters, can hence be obtained. These aid in comparison with other systems using a different fading model or where closed form solutions of BER or outage probability are not readily available. Higher order moments are also useful for further system analysis such as in SNR estimation

[170]. In terms of the CDF, μ_n is given as:

$$\mu_n = \mathbb{E}[\gamma_{eq}^n] = n \int_0^\infty \gamma_{eq}^{n-1} (1 - F_{\gamma_{eq}}(\gamma)) d\gamma \quad (6.19)$$

From (6.11), (6.19) is derived¹ using [98, eqn.(6.631.3)] (see A.3.2) and after simplification results in:

$$\begin{aligned} \mu_n = & \frac{n N_t! \exp(C/2\bar{\gamma}_2) \bar{\gamma}_1^n}{(N_{ts} - 1)!(N_{rd} - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ & \cdot \sum_{j=0}^{i(N_{ts}-1)} \frac{\eta_{N_{ts}}(i, j)(N_{ts} + j - 1)!}{(i + 1)^{N_{ts}+n+j}} \sum_{l=0}^{j+N_{ts}-1} \frac{\Gamma[l+n]}{l!} \sum_{p=0}^l \binom{l}{p} \left(\frac{C}{\bar{\gamma}_2}\right)^{\frac{N_{rd}+p-1}{2}} \\ & \cdot \Gamma[N_{rd} - p + l + n] \mathbf{W}_{-(2(l+n)+N_{rd}-p-1)/2, (N_{rd}-p)/2} \left(\frac{C}{\bar{\gamma}_2}\right) \end{aligned} \quad (6.20)$$

As a check, with $N_{ts} = N_{rd} = 1$ and N_t relays, (6.20) reduces to [157, eqn.(9)]. Next, we can derive the n^{th} central moments about the mean as $\mu'_n = \mathbb{E}[(\gamma_{eq} - \mathbb{E}[\gamma_{eq}])^n]$ which is expanded as [168]: $\mu'_n = \sum_{j=0}^n \frac{n!(-1)^{n-j} \mu_j \mu_1^{n-j}}{j!(n-j)!}$. Important third and fourth standardised moments such as the skewness (indicative of PDF asymmetry) and kurtosis (PDF peakiness) can be derived from this to gain further insight into the nature of the output SNR: the n^{th} standardised moment is $\frac{\mu_n}{\mu_2^{n/2}}$.

6.1.6 Numerical and simulation results

Next, outage probability and BER for BPSK are evaluated from these equations, and confirmed through simulation. For clarity, $C = \bar{\gamma}_1$ in all plots, so $G^2 = \mathcal{E}_2/\mathcal{E}_1$. Fig. 6.2 plots the performance of the single and multi-antenna relay selection schemes, with the dual-hop plot systems represented as (N_t, N_{ts}, N_{rd}) . The figure compares outage performance of several single antenna nodes and multi-antenna systems, with $\bar{\gamma}_2 = 5\bar{\gamma}_1$ and a specific value of threshold $\gamma_T = 5$.

Clearly, a steep improvement can be obtained through the use of multiple receive antennas. We see that the performance of $(N_t, 1, 1)$ is relatively flat, and the effect of adding relays quickly diminishes; the gains being limited largely by the second

¹The n^{th} moment can be obtained from (6.16) as $\mathbb{E}[\gamma_{eq}^n] = (-1)^n \frac{d^n}{ds^n} \mathcal{M}_{\gamma_{eq}}(s)|_{s=0}$. However, it is easier to derive from (6.19).

Figure 6.2: Outage Probability for different relay configurations in the dual-hop multi-antenna scheme, with $\gamma_T = 5$, $\bar{\gamma}_2 = 5\bar{\gamma}_1$.

hop channel, as noted in [157], [165] (single antenna case). Large gains are observed even with a small number of relay terminals. Thus the use of multiple antennas is particularly beneficial when the number of relay nodes is limited.

Another advantage is that when reverse channel bandwidth is limited it helps reduce CSI feedback load to the source as the relay nodes used for selection increase in number. Following information theoretic results in [162], the diversity order of this system will be $\min(N_t \times N_{ts}, N_{rd})$, the minimum number of first hop and second hop diversity paths. With optimal relay selection, the effective number of forward diversity paths created are $N_t \times N_{ts}$, suggestive of the diversity order of the first path, whereas, for the second hop, the diversity order is N_{rd} . Thus in the multi-antenna cases shown, the curves are found to have similar slopes at high SNR, and for the single antenna case of $(N_t, 1, 1)$ the diversity order of the second hop channel determines performance at high SNR. It is also seen that if the quality of the first hop is better than that of the second hop (realised by multiple relays and transmit antennas) then

Figure 6.3: BPSK system BER performance for different antenna configurations and relay gains.

large performance gains are possible, i.e., there is room for performance improvement by increasing N_{rd} or G . However, if the opposite is true, or if the second hop link is comparable to the first, then performance saturates quickly, as seen in the $(1, 2, 4)$ setup: increasing N_{rd} produces almost no improvement. Similarly, little improvement is seen by increasing G . Fig. 6.3 plots the BER of $(5, 1, 1)$ and $(2, 2, 2)$ setups with different gain ratios between first and second hop. Note the total number of antennas in $(5, 1, 1)$ is one more than in $(2, 2, 2)$. Performance improvement is evident in the $(2, 2, 2)$ case over and above $(5, 1, 1)$, although we note that in reality, large values of G may not be profitable. This effect is again explainable noting that once second hop is sufficiently stronger through increased G , saturation occurs and increasing G further does not contribute to any performance improvement.

6.1.7 Summary

Closed form performance metric solutions have been derived for a wireless two-hop system with multi-antenna source and destination communicating via a single antenna AF best-relay selected from a set of candidates. The source and destination employ beamforming and MRC respectively. The analytical solutions, along with simulations, have been used in one possible design space exploration: that of determining how best to allocate the system resource of additional antennas: in more relays, in more antennas at source, or more at destination. It was seen that careful selection of system parameters is important to realise the best improvement. It is worth noting two variant subsets from within the possible set of systems encompassed by these generalised solutions: (a) setting $N_{rd} = N_{ts}$ constitutes a symmetrical system of multi-antenna devices communicating with the assistance of a set of several simple single-antenna relay nodes, (b) that the equations presented also apply for the case of single antenna source/destination with multiple antenna relays (having MRC for receive and beamforming for transmit); a very different deployment scenario, but nevertheless also one of potential practical relevance.

6.2 Dual-Hop Variable Gain AF Multi-antenna Relay Selection in Nakagami- m Channels

Introduction

We analyse a dual-hop amplify and forward multi-antenna relay selection in Nakagami- m channels. In this scheme, a source-selected best relay chosen out of several relatively closely spaced candidates based on maximum instantaneous SNR (signal to noise ratio), performs receive maximal ratio combining (MRC) on received data, applies variable gain, and then uses beamforming to transmit on to a destination. Novel closed form solutions for outage probability and symbol error (SER) rate are derived for arbitrary number of relays, receive/transmit antennas and fading parameters, and

Figure 6.4: A distant source communicates to a destination via a multi-antenna relay selective system. A weak direct path between $S \rightarrow D$ is also shown.

verified by simulations. In particular, the benefit of multiple antennas is observed, and the system studied for different severities of fading, power imbalance between hops, as well as influence of number of antennas. Since it is useful to explore the effect of different antenna configurations in various channel environments, generalised multi-antenna relay selection for Nakagami- m fading channels is analysed. The amplifying gain is set to invert fading at the selected relay [6]. The closed form solutions are verified through Monte Carlo simulation, with accurate channel estimates assumed for demodulation and combining. The effect of fading severity is analysed with respect to a power imbalance between hops, as well as number and location of antennas. Many degrees of freedom in this proposed PRS system can be adjusted for performance gain over such schemes as [156, 171] which report performance saturation with increasing number of relays.

6.2.1 System model and assumptions

Consider a relay network in a flat fading Nakagami- m channel as shown in Fig. 6.4 comprising source S , L relay nodes $R_1, R_2, \dots, R_L$ each equipped with N_R receive and K transmit antennas, and single destination, D . S selects a best relay R_l , l : $1 \leq l \leq L$, which forwards amplified messages to D . Half-duplex transmission occurs in two phases, first as $S \rightarrow R_l$, followed by $R_l \rightarrow D$. A weak path is assumed to

be absent in this scenario. $\mathcal{E}_1$ and $\mathcal{E}_2$ are the power of the transmitted signal at S and R_l and N_0 the power of the additive white Gaussian noise at R_l and D . The channel matrix between S and $R_{v=1,2,\dots,L}$ during initial channel estimation is denoted as $\mathbf{H} = [h_{ij}]_{N_R \times L}$ where h_{ij} s are characterised as spatially independent and identically distributed (i.i.d) with distribution $\mathcal{CN}(0, 1)$. The source picks a single best relay with index l , based on highest instantaneous channel power computed at each relay as $p_{max} = \max_{1 \leq j \leq L} p_j$ with $p_j = \sum_{i=0}^{N_R} |h_{ij}|^2$. The first transmission hop to this relay, denoted as $S \rightarrow R_l$, has channel $[\mathbf{h}_{SR_l}]_{N_R \times 1}$, being the l^{th} column of $\mathbf{H}$. The selected relay performs MRC using $\mathbf{w}_{\mathbf{SR}_l} = \frac{\mathbf{h}_{SR_l}^H}{\sqrt{\mathbf{h}_{SR_l}^H \mathbf{h}_{SR_l}}}$. Before transmitting the signal to D the signal received at the relay is amplified by a gain G . R_l obtains the $R_l \rightarrow D$ channel $[\mathbf{h}_{R_l D}]_{N_{rd} \times 1}$ channel vector which is $\mathcal{CN}(0, \mathbf{I}_{N_K})$ distributed. To perform transmit diversity beamforming over the selected hop, R_l computes an optimal beamforming vector [23] $\mathbf{w}_{R_l D} = \frac{\mathbf{h}_{R_l D}^H}{\sqrt{\mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}}$ after which data is transmitted to D . The final output at D is therefore is written as:

$$y_{R_l D} = G \mathbf{w}_{R_l D}^H \mathbf{h}_{R_l D} (\mathbf{w}_{SR_l}^H \mathbf{h}_{SR_l} s + \mathbf{n}_{R_l}) + n_D \quad (6.21)$$

where n_D are the noise samples at D characterised as $\mathcal{CN}(0, N_0)$ and respectively with N_0 as the AWGN variance. Then the equivalent SNR at the destination is given as $\gamma_{eq} = \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + c}$ [6, 117], with $\gamma_1 = \frac{\mathcal{E}_1 \mathbf{h}_{SR_l}^H \mathbf{h}_{SR_l}}{N_0}$ and $\gamma_2 = \frac{\mathcal{E}_2 \mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}{N_0}$ represent the instantaneous SNRs of the direct path, first and second hops respectively. The same equations hold if we take the relay to have a single transmit antenna and D to be equipped with K receive antennas. The CDF of the random variable γ_1 (which is the instantaneous SNR at the best relay), $F_{\gamma_1}(\cdot)$, is obtained using order statistics, found as $F_{\gamma_1}(\cdot) = [F(\gamma)]^L$, where $F(\gamma)$ is the CDF of the statistically identical $S - R_{i=1,\dots,L}$ links, characterised by independent and identical Nakagami- m random variables. Essentially, γ is the random variable corresponding to the MRC combiner output SNR at each relay, characterised by the gamma distribution $\gamma \sim \mathcal{G}(m_1 N_R, \bar{\gamma}_1/m_1)$ so that its PDF $f_\gamma(\gamma) = \frac{\gamma^{(m_1 N_R - 1)} \exp(-\frac{m_1 \gamma}{\bar{\gamma}_1})}{(\bar{\gamma}_1/m_1)^{m_1 N_R} (m_1 N_R - 1)!}$ [163], with $\bar{\gamma}_1 = \mathbb{E}(\gamma) = \mathcal{E}_1/N_0$; $\mathbb{E}$ as the expectation operator. For the second hop $R_l \rightarrow D$, with transmit beamforming, its PDF $f_{\gamma_2}(\gamma_2)$, is given by $\gamma_2 \sim \mathcal{G}(m_2 K, \bar{\gamma}_2/m_2)$ with $\bar{\gamma}_2 = \mathbb{E}(\gamma_2) = \mathcal{E}_2/N_0$. m_1, m_2

(≥ 0.5) are the Nakagami- m fading parameters for the first and second hop channels respectively.

6.2.2 Outage Probability

The outage probability P_{out} which is the probability that the instantaneous output SNR drops below a threshold γ_T , i.e, $P_{out}(\gamma_T) = P_{out}(\gamma_{eq} < \gamma_T)$. In variable gain relaying this is [117]:

$$P_{out}(\gamma_T) = \int_0^\infty Pr\left[\frac{\gamma_1\gamma_2}{\gamma_1 + \gamma_2 + 1} < \gamma_T | \gamma_2\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \quad (6.22)$$

With some rearrangement (see A.7.1 in Appendix for derivation), the above can be expressed more conveniently as

$$P_{out}(\gamma_T) = 1 - \int_{\gamma_T}^\infty \left\{ 1 - F_{\gamma_1}\left[\frac{\gamma_T(\gamma_2 + 1)}{\gamma_2 - \gamma_T}\right] \right\} f_{\gamma_2}(\gamma_2) d\gamma_2. \quad (6.23)$$

To obtain $F_{\gamma_1}(\cdot)$ we first need the CDF of $\gamma : F_\gamma(\gamma)$, which can be easily obtained as:

$$F_\gamma(\gamma) = 1 - \exp(-m_1\gamma/\bar{\gamma}_1) \sum_{k=0}^{m_1N_R-1} \left(\frac{m_1\gamma}{k!\bar{\gamma}_1}\right)^k \quad (6.24)$$

Then an expansion for $F_{\gamma_1}(\cdot)$, which is $[F_\gamma(\gamma)]^L$, is found. It is convenient to express it as:

$$\begin{aligned} F_{\gamma_1}(\gamma) &= 1 - L! \sum_{i=0}^{L-1} \frac{(-1)^i \exp\left(-\frac{m_1(i+1)\gamma}{\bar{\gamma}_1}\right)}{(i+1)!(L-i-1)!} \\ &\quad \cdot \sum_{j=0}^{(i+1)(m_1N_R-1)} \eta_{m_1N_R}(i+1, j) \left(\frac{m_1}{\bar{\gamma}_1}\right)^j \gamma^j \end{aligned} \quad (6.25)$$

with $\eta_{m_1N_R}(i+1, j)$ [98, eqn.(0.314)] as the coefficients of $(m_1\gamma/\bar{\gamma}_1)^j, j = 0, 1, \dots, (i+1)(m_1N_R-1)$, in the expansion of $(\sum_{j=0}^{m_1N_R-1} (m_1\gamma/\bar{\gamma}_1)^j/j!)^{(i+1)}$. A change of variable allows resetting limits from 0 to ∞ (details in Appendix A.7.1); combining $1 - F_{\gamma_1}(\cdot)$ with $f_{\gamma_2}(\cdot)$ and using the binomial expansion for the terms containing γ_T , we can finally evaluate (6.22) using [98, eqn.(3.471.9)](see A.4.2). An exact closed form

equation for P_{out} is then found:

$$\begin{aligned}
 P_{out}(\gamma_T) = & 1 - \frac{2L!}{(\bar{\gamma}_2/m_2)^{m_2K}} \sum_{i=0}^{L-1} \frac{(-1)^i}{(i+1)!(L-1-i)!} \exp\left(-\frac{m_1(i+1)\gamma_T}{\bar{\gamma}_1} - \frac{\gamma_T m_2}{\bar{\gamma}_2}\right) \\
 & \cdot \sum_{j=0}^{(i+1)(m_1N_R-1)} \eta_{N_R}(i+1, j) \left(\frac{m_1}{\bar{\gamma}_1}\right)^j \gamma_T^j \sum_{p=0}^j \frac{j!(\gamma_T+1)^{j-p}}{(j-p)!p!} \sum_{q=0}^{m_2K-1} \frac{\gamma_T^{m_2K-1-q}}{(m_2K-1-q)!q!} \quad (6.26) \\
 & \cdot \left[\frac{\bar{\gamma}_2(i+1)m_1\gamma_T(1+\gamma_T)}{\bar{\gamma}_1 m_2} \right]^{\frac{(q+p-j+1)}{2}} K_{(q+p-j+1)} \left(2 \sqrt{\frac{m_1 m_2 (i+1) \gamma_T (1+\gamma_T)}{\bar{\gamma}_1 \bar{\gamma}_2}} \right)
 \end{aligned}$$

where $K_v(\cdot)$ is the modified Bessel function of the second kind of order v [98].

6.2.3 Error probability

The average SER for coherent demodulation of various modulation schemes in a slow fading channel is [147] : $\bar{P}_s = \frac{\alpha}{\sqrt{(2\pi)}} \int_0^\infty F_{\gamma_{eq}}(x^2/\beta) \exp(-x^2/2) dx$, where α and β determine specific constellations, and $F_{\gamma_{eq}}(\cdot)$ is the CDF of γ_{eq} . To facilitate finding a closed form solution of the above integral using [98, eqn.(6.621.3)](see A.3.4), an approximation of $\gamma_T + 1 \approx \gamma_T$ is made in (6.26), then substituting $\gamma_T = x^2/\beta$ we obtain $P_{out}(x^2/\beta)$ or equivalently $F_{\gamma_{eq}}(x^2/\beta)$. Combining this with $\exp(-x^2/2)$ we then solve the integral to obtain $\bar{P}_s$ as follows:

$$\begin{aligned}
 \bar{P}_s \approx & \frac{\alpha}{\sqrt{2\pi}} \left[\frac{\sqrt{2\pi}}{2} - \frac{\sqrt{\pi} L!}{(\bar{\gamma}_2/m_2)^{m_2K}} \sum_{i=0}^{L-1} \frac{(-1)^i}{(i+1)!(L-1-i)!} \right. \\
 & \cdot \sum_{j=0}^{(i+1)(m_1N_R-1)} \frac{\eta_{N_R}(i+1, j)}{\Gamma[j + m_2K + 1]} \left(\frac{m_1}{\bar{\gamma}_1}\right)^j \\
 & \cdot \sum_{p=0}^j \frac{j!}{(j-p)!p!} \sum_{q=0}^{m_2K-1} \frac{\Gamma[m_2K - q - p - 0.5]}{(m_2K - 1 - q)! q! \beta^{m_2K+q+p+1}} \quad (6.27) \\
 & \cdot \frac{\Gamma[m_2K + q + p + 1.5] \left[\frac{4(i+1)m_1}{\bar{\gamma}_1} \right]^{(q+p-j+1)}}{\left[\frac{m_1(i+1)}{\beta \bar{\gamma}_1} + \frac{m_2}{\beta \bar{\gamma}_2} + \frac{1}{2} + \frac{2}{\beta} \sqrt{\frac{(i+1)m_1 m_2}{\bar{\gamma}_1 \bar{\gamma}_2}} \right]^{m_2K+q+p+1.5}} \\
 & \left. \cdot {}_2\mathbf{F}_1(m_2K + q + p + 1.5, q + p - j + 1.5; m_2K + j + 1; \zeta) \right]
 \end{aligned}$$

where $\zeta = \left[\frac{m_1(i+1)}{\beta\bar{\gamma}_1} + \frac{m_2}{\beta\bar{\gamma}_2} + \frac{1}{2} - \frac{2}{\beta} \sqrt{\frac{m_1 m_2(i+1)}{\bar{\gamma}_1 \bar{\gamma}_2}} \right]$ and ${}_2F_1(\cdot, \cdot, \cdot; \cdot)$ is the Gauss hypergeometric function [98]. The approximation made in solving the integral holds good in the medium to high SNR region, as will be seen in the graphs plotted for BER. Note that we have assumed integer values for m_1, m_2 , although a similar analysis can be done for non-integer values after obtaining infinite summations for $[F_\gamma(\gamma)]^L$.

6.2.4 Numerical and simulation results

In the following, the shorthand (L, N_R, K) identifies the system arrangement. All derived equations are verified as being accurate through Monte Carlo simulation (and this evidence is plotted on each graph). Firstly, Fig. 6.5 shows $P_{out}(\gamma_T)$ for a $(2, 2, 2)$ system for different ratios of $\bar{\gamma}_1$ to $\bar{\gamma}_2$, and several values of m_2 . If we assume that the distance between S and a cluster of relays will generally be larger than from the relays to D , we can take $m_1 = 1$, amounting to Rayleigh fading between S to R_l . (for example, a long distance communications system using local relay selection to handle last-mile linking in a scattering urban environment). Strengthening first hop diversity through relay selection is thus beneficial, while for the second hop, proximity to relays (increase in m_2) improves second hop diversity. In the evaluation we set $\gamma_T = 5dB$ as a reasonable outage probability threshold. As seen in Section 6.1.6, the diversity order of this system will be $\min(m_1 L N_R, m_2 K)$. Therefore, if the first hop is strong, either through more relays, relay receive antennas, greater m_1 , or greater first hop transmit power or lower path loss, the system will be more sensitive to changes in m_2 and K . This also implies that increasing first hop resources will cause saturation of system performance beyond a certain point, and that performance can then be improved through strengthening the second hop (either by increasing gain, or through increasing K and m_2). As seen in Fig. 6.5, when $\bar{\gamma}_2 < \bar{\gamma}_1$, the system is more sensitive to the fading parameter m_2 (i.e. proximity of D and R_l). When the ratio improves or when $\bar{\gamma}_1 = \bar{\gamma}_2$, performance naturally increases by having a stronger second hop, then sensitivity to m_2 decreases. The opposite effect is expected to occur when m_1 is

concerned i.e when m_2 is kept constant and when $\bar{\gamma}_2 > \bar{\gamma}_1$, the system would be more sensitive to changes in m_1 and less sensitive as the ratio $\bar{\gamma}_2 : \bar{\gamma}_1$ gets smaller. However, as stated above, the degree of sensitivity to changes in m_1 and m_2 also depends on the diversity gains over the two hops. Fig. 6.6 plots BER for single and multiple antenna relay selection in setups of $(4, 1, 1)$ and $(4, 2, 2)$ with $\bar{\gamma}_1 = 2\bar{\gamma}_2$. In general, a large performance improvement is evident from the use of multiple antennas. As noted, the fading parameter plays an important role in realising large performance gains in the relay system. The $(4, 1, 1)$ system is evidently more sensitive to m_2 , while a higher m_2 is beneficial in providing a balance between diversity gains over the first and second hops. A similar effect is seen in $(4, 2, 2)$, which performs much better than through the use of an additional antenna. When the diversity gain of the first hop is high, any increase in the number of relay transmit antennas will bring about improvement in the BER; this is beneficial especially when m_2 is small. However, the system is expected to be less sensitive to variations in m_2 and performance saturates quickly for increasing values of m_2 . These observations are seen in the $(4, 2, 4)$ setup.

6.2.5 Summary

The performance of a source-selected relaying system in a Nakagami- m fading environment has been analysed. A system was studied in which one of several MRC-equipped relays are chosen based upon highest instantaneous SNR over the first hop. In this system, the selected relay employs MRC, amplifies and forwards the message using multi-antenna beamforming to the destination. Generalised closed-form expressions for outage probability and BER were derived, and used to explore the beneficial effect of employing multiple antennas in the system. The effect of power imbalance between hops, different Nakagami- m fading parameters, as well as the influence of number of transmit antennas was also explored.

Figure 6.5: Outage probability w.r.t $\bar{\gamma}_1$, for $\gamma_T = 5$ dB for selective relaying $L = 2, N_R = 2, K = 2$. The effect of power imbalance between hops for increasing $m_2 = 1, 2, 5, 8$ with $m_1 = 1$ is seen.

Figure 6.6: BER plot for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 = 2\bar{\gamma}_2$. A performance improvement is noted for increasing values of $m_2 = 1, 2, 3, 4, 5$ is noted. Sensitivity to changes in m_2 w.r.t number of transmit antennas can also be seen.

6.3 Partial Multi-antenna Relay selection taking direct path into account

The previous analysis assumed the complete absence of any direct communication channel path from S to D . This section will, by contrast, consider a case where the direct path ($S \rightarrow D$) is included in the transmission process. For mathematical tractability, approximations of the equivalent E2E SNR will be used. Specifically via an MGF analysis, closed form solutions for end-to-end outage probability and MGF, and system capacity will be derived.

6.3.1 System model

The system model is similar to the one in Section 6.2.1, except that this time we consider the direct path between source and destination $S \rightarrow D$ channel, as shown in Fig. 6.4. For the $S \rightarrow D$ direct path, let h_{SD} be the channel gain. The final output at D due to direct and relayed paths are therefore written as:

$$y_{SD} = \mathbf{h}_{SD}s + n_{D_1} \quad (6.28)$$

$$y_{R_l D} = G \mathbf{w}_{R_l D} \mathbf{h}_{R_l D} (\mathbf{w}_{S R_l} \mathbf{h}_{S R_l} s + \mathbf{n}_{R_l}) + n_{D_2} \quad (6.29)$$

where n_{D_1} and n_{D_2} are the noise samples at D characterised as $\mathcal{CN}(0, N_0)$ and $\mathcal{CN}(0, \mathbf{I}_{N_R})$ respectively with N_0 as the AWGN variance and $\mathbf{I}$ is the identity matrix. Then the equivalent SNR at the destination is given as $\gamma_{eq} = \gamma_0 + \gamma_{SRD}$ where $\gamma_{SRD} = \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + c}$ [6, 117], with $\gamma_0 = \frac{\varepsilon_1 |h_{SD}|^2}{N_0}$, $\gamma_1 = \frac{\varepsilon_1 \mathbf{h}_{S R_l}^H \mathbf{h}_{S R_l}}{N_0}$ and $\gamma_2 = \frac{\varepsilon_2 \mathbf{h}_{R_l D}^H \mathbf{h}_{R_l D}}{N_0}$ representing the instantaneous SNRs of the direct path, first and second hops respectively.

With $c = 1$, the channel noise assisted (CNA) instantaneous variable relay gain will be chosen so as to invert the fading effect of the first hop such that $G = \frac{\varepsilon_2}{N_0 + \varepsilon_1 \mathbf{h}_{S R_l}^H \mathbf{h}_{S R_l}}$ i.e while limiting the output power of the relay, while $c = 0$ results in a high SNR approximation $\gamma'_{SRD} = \gamma_{SRD}|_{c=0} = \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2}$. This can be also thought of as providing a gain of $G = \frac{\varepsilon_2}{\varepsilon_1 \mathbf{h}_{S R_l}^H \mathbf{h}_{S R_l}}$ (i.e ignoring the noise term at R_l). γ'_{SRD} is upper bounded as $\gamma'_{SRD} < \gamma_{up} \triangleq \min(\gamma_1, \gamma_2)$ [172].

6.3.2 Performance metrics

In this section we derive useful closed form expressions for the moment generating function (MGF) $\mathcal{M}_{\gamma_{eq}}(\cdot)$, the CDF $F_{\gamma_{eq}}(\cdot)$, and the channel capacity of the relaying system $\mathcal{C}$.

The MGF is defined as $\mathcal{M}_{\gamma_{eq}}(s) = \mathbb{E}[\exp(-s\gamma_{eq})]$ [19]. Evidently, for the relaying system, the independence of the relayed and direct path enables us to write $\mathcal{M}_{\gamma_{eq}}(s) = \mathcal{M}_{\gamma_0}(s)\mathcal{M}_{\gamma_{SRD}}(s)$. To find $\mathcal{M}_{\gamma_{SRD}}(s)$, we find $F_{\gamma_{SRD}}(\gamma)$ first, while it is known that $\mathcal{M}_{\gamma_0}(s) = \left(\frac{m_0}{s\bar{\gamma}_1 + m_0}\right)^{m_0}$ [19]. Using the high SNR approximation γ_{up} , for γ_{SRD} , its CDF therefore is found as:

$$F_{\gamma_{SRD}}(\gamma) = 1 - F_{\gamma_1}^c(\gamma)F_{\gamma_2}^c(\gamma) \quad (6.30)$$

where $F_X^c(x)$ is the complimentary CDF (CCDF) of r.v X . First the CDF of the random variable γ_1 (which is the instantaneous SNR at the best relay), $F_{\gamma_1}(\cdot)$, is obtained using order statistics and is found as the CDF of the largest order statistic $F_{\zeta(L)} = F_{\gamma_1}(\cdot) = [F_{\zeta}(\gamma)]^L$, where $F_{\zeta}(\gamma)$ is given by $F_{\zeta}(\gamma) = 1 - \exp(-m_1\gamma/\bar{\gamma}_1) \sum_{k=0}^{m_1N_R-1} \left(\frac{m_1\gamma}{k!\bar{\gamma}_1}\right)^k$. An expansion for $F_{\gamma_1}(\cdot)$, which is $[F_{\zeta}(\gamma)]^L$, is found. It is thus advantageous to express it in CCDF form, and was given previously in eqn. (6.25).

The CDF of the second hop is : $F_{\gamma_2}(\gamma) = 1 - \exp(-m_2\gamma/\bar{\gamma}_2) \sum_{k=0}^{m_2K-1} \frac{1}{k!} \left(\frac{m_2\gamma}{\bar{\gamma}_2}\right)^k$ while for the direct $S \rightarrow D$ link the CDF is $F_{\gamma_0}(\gamma) = 1 - \exp(-m_0\gamma/\bar{\gamma}_1) \sum_{p=0}^{m_0-1} \frac{1}{p!} \left(\frac{m_0\gamma}{\bar{\gamma}_1}\right)^p$. Combining the above results we obtain:

$$\begin{aligned} F_{\gamma_{SRD}}(\gamma) &= 1 - L! \sum_{i=0}^{L-1} \frac{(-1)^i \exp\left(-\frac{m_1(i+1)\gamma}{\bar{\gamma}_1} - \frac{m_2\gamma}{\bar{\gamma}_2}\right)}{(i+1)!(L-i-1)!} \\ &\quad \cdot \sum_{j=0}^{(i+1)(m_1N_R-1)} \eta_{m_1N_R}(i+1, j) \left(\frac{m_1\gamma}{\bar{\gamma}_1}\right)^j \sum_{k=0}^{m_2K-1} \frac{1}{k!} \left(\frac{m_2\gamma}{\bar{\gamma}_2}\right)^k \end{aligned} \quad (6.31)$$

Using integration by parts it can be shown that

$$\mathcal{M}_{\gamma_{SRD}}(s) = 1 - s \int_0^\infty \exp(-\gamma_{SRD}s) F_{\gamma_{SRD}}^c(\gamma) d\gamma \quad (6.32)$$

Using (6.32), (6.31), [98, eqn.(3.383.1)] and $\mathcal{M}_{\gamma_0}(s)$ we derive:

$$\begin{aligned} \mathcal{M}_{\gamma_{eq}}(s) &= \left(\frac{m_0}{s\bar{\gamma}_1 + m_0} \right)^{m_0} s \left(\frac{m_0}{\bar{\gamma}_1} \right)^{m_0} \sum_{i=0}^{L-1} \frac{L!(-1)^i}{(i+1)!(L-i-1)!} \\ &\quad \cdot \sum_{j=0}^{(i+1)(m_1 N_R - 1)} \eta_{m_1 N_R}(i+1, j) \left(\frac{m_1}{\bar{\gamma}_1} \right)^j \sum_{k=0}^{m_2 K - 1} \frac{(j+k)!}{k!} \left(\frac{m_2}{\bar{\gamma}_2} \right)^k \\ &\quad \cdot \left(s + m_2/\bar{\gamma}_2 + \frac{m_1(i+1)}{\bar{\gamma}_1} \right)^{-(j+k+1)} (s + m_0/\bar{\gamma}_1)^{-m_0} \end{aligned} \quad (6.33)$$

To finally derive $F_{\gamma_{eq}}(\gamma)$ we utilize the relation between MGF and CDF, i.e $F_{\gamma_{eq}}(\gamma) = s^{-1} \mathbb{E}[\exp(s\gamma_{eq}) \mathcal{M}_{\gamma_{eq}}(s)]$ or,

$$F_{\gamma_{eq}}(\gamma) = \frac{1}{s} \int_0^\infty \exp(\gamma_{eq} s) \mathcal{M}_{\gamma_{eq}}(s) d\gamma \quad (6.34)$$

The $\mathcal{M}_{\gamma_{eq}}(s)$ is first expanded via partial fractions before plugging into eqn. (6.34).

Finally $F_{\gamma_{eq}}(\gamma)$ is obtained as:

$$\begin{aligned} F_{\gamma_{eq}}(\gamma) &= 1 - \underbrace{\exp\left(-\frac{m_0\gamma}{\bar{\gamma}_1}\right) \sum_{p=0}^{m_0-1} \frac{1}{p!} \left(\frac{m_0\gamma}{\bar{\gamma}_1}\right)^p}_{\text{Direct path}} - \left(\frac{m_0}{\bar{\gamma}_1}\right)^{m_0} \\ &\quad \cdot L! \sum_{i=0}^{L-1} \frac{(-1)^i}{(i+1)!(L-i-1)!} \sum_{j=0}^{(i+1)(m_1 N_R - 1)} \eta_{m_1 N_R}(i+1, j) \left(\frac{m_1}{\bar{\gamma}_1}\right)^j \\ &\quad \cdot \sum_{k=0}^{m_2 K - 1} \frac{(j+k)!}{k!} \left(\frac{m_2}{\bar{\gamma}_2}\right)^k \\ &\quad \cdot \left(\sum_{l=1+I(V+m_0)}^{1+V+I \cdot m_0} \frac{\lambda_{li} \gamma^{l-1}}{(l-1)!} \exp(-a_i \gamma) + \sum_{t=1}^{m_0} \frac{\mu_t \gamma^{t-1}}{(t-1)!} \exp(-b \gamma) \right) \end{aligned} \quad (6.35)$$

where $V = j + k$, $a_i = \frac{m_2}{\bar{\gamma}_2} + \frac{m_1(i+1)}{\bar{\gamma}_1}$, $b = \frac{m_0}{\bar{\gamma}_1}$ and I is the indicator variable defined as $I = 0$ if $a_i \neq b$ and $I = 1$ if $a_i = b$ and λ_{li} and μ_t are coefficients obtained during the

partial fraction expansion given as:

$$\begin{aligned}\lambda_{li} &= \begin{cases} \frac{\mathbf{d}^{j+k+1-l}}{(j+r+1-l)!\mathbf{d}s^{j+k+1-l}} \left(\frac{1}{s+b}\right)^{m_o} \Big|_{s=-a_i} & \text{if } a_i \neq b \\ 1, & \text{if } a_i = b \end{cases} \\ \mu_t &= \begin{cases} \frac{\mathbf{d}^{m_o-t}}{(m_o-t)!\mathbf{d}s^{m_o-t}} \left(\frac{1}{s+a_i}\right)^{j+k+1} \Big|_{s=-b} & \text{if } a_i \neq b \\ 0 & \text{if } a_i = b \end{cases}\end{aligned}\quad (6.36)$$

The result is intuitively satisfying: the first two terms denote the outage probability due to the direct path $S \rightarrow D$, while the third term contributes to the reduction in overall probability obtained through relaying.

6.3.3 Channel capacity

The maximum average spectral efficiency in bits/sec/Hz in a channel of bandwidth $\mathcal{W}$ utilised per hop in the dual-hop relaying system accounting for the extra bandwidth is:

$$\mathcal{C}/\mathcal{W} = \frac{1}{2 \log_e 2} \int_0^\infty \log_e(1 + \gamma_{eq}) f_{\gamma_{eq}}(\gamma) d\gamma \quad (6.37)$$

To obtain this, usually the PDF $f_{\gamma_{eq}}(\gamma)$ needs to be derived. However, a useful formula is given below, where the capacity can be obtained in terms of the CDF alone as:

$$\mathcal{C}/\mathcal{W} = \frac{1}{2 \log_e 2} \int_0^\infty \frac{1 - F_{\gamma_{eq}}(\gamma)}{\gamma + 1} d\gamma \quad (6.38)$$

This can be obtained through integration by parts from (6.37) and the use of the squeezing theorem plus the Markov inequality theorem. Using [98, eqn.(3.383.10)]

(see A.4.3) and (6.35), (6.38) is thus derived in closed form as:

$$\begin{aligned}
 \mathcal{C}/\mathcal{W} = & \frac{1}{2\log_e 2} \left[\exp\left(\frac{m_0}{\bar{\gamma}_1}\right) \sum_{p=0}^{m_0-1} \frac{1}{p!} \left(\frac{m_0}{\bar{\gamma}_1}\right)^p \right. \\
 & \cdot \Gamma[p+1] \Gamma\left[-p, \frac{m_0}{\bar{\gamma}_1}\right] + \left(\frac{m_0}{\bar{\gamma}_1}\right)^{m_0} \sum_{i=0}^{L-1} \frac{L!(-1)^i}{(i+1)!(L-i-1)!} \\
 & \cdot \sum_{j=0}^{(i+1)(m_1 N_R - 1)} \eta_{m_1 N_R}(i+1, j) \left(\frac{m_1}{\bar{\gamma}_1}\right)^j \sum_{k=0}^{m_2 K - 1} \frac{(j+k)!}{k!} \left(\frac{m_2}{\bar{\gamma}_2}\right)^k \\
 & \cdot \left(\sum_{l=1+I(V+m_0)}^{1+V+I \cdot m_0} \frac{\lambda_{li} \exp(a_i) \Gamma[l] \Gamma[1-l, a_i]}{(l-1)!} + \right. \\
 & \quad \left. \sum_{t=1}^{m_0} \frac{\mu_t \exp(b) \Gamma[t] \Gamma[1-t, b]}{(t-1)!} \right) \left. \right] \quad (6.39)
 \end{aligned}$$

$\Gamma[\cdot] \Gamma[\cdot, \cdot]$, as before represent the Gamma and the incomplete Gamma function respectively [98]. Note we have assumed integer values for m_1, m_2 , although a similar analysis could be performed for non-integer values after obtaining infinite summations for $[F_\gamma(\gamma)]^L$.

6.3.4 Diversity order analysis

Based on the independence of the r.vs γ_{SRD} , γ_0 , intuitively the effective diversity order of the system will be given by the sum of the diversity order of the direct path and the relayed path. Alternatively, a similar result is proved differently in [173], using an upper bound for the error probability and the definition of diversity order. It is well known that the diversity order of the direct path can be found as $G_d = m_0$ [174]. To determine the diversity order G_r of the relayed path, we utilize γ'_{SRD} . Using [162, Lemma 3], we find that G_r is determined by the minimum of the diversity order of the first and second hops, i.e $\min(G_1, G_2)$. Thus $G_1 = m_1 L N_R$; intuitively this is suggestive of the total number of diversity paths in the first hop. Similarly for the second hop $G_2 = m_2 K$. Thus $G_r = m_0 + \min(m_0 L N_R, m_2 K)$.

Figure 6.7: Capacity for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 = 1.25\bar{\gamma}_2$, with and without direct path.

6.3.5 Numerical and simulation results

As before, we use the abbreviation (L, N_R, K) to denote a relay system arrangement. Eqns. (6.35) and (6.39) are verified as being accurate through Monte Carlo simulation (and this is plotted on each graph). Firstly, 6.7 shows $P_{out}(\gamma_T)$ and compares the performance of the single antenna $(4, 1, 1)$ to a $(2, 2, 2)$ system for $\bar{\gamma}_1 = \bar{\gamma}_2$, and different values of m_2 with $m_0 = 1, m_1 = 1$. Note that the chosen setups use the same number of antennas, deployed in different arrangements. The first hop is assumed to undergo Rayleigh fading, so $m_1 = 1$ (between S to R_l) and m_2 is varied to account for proximity of the relays to the destination. We set $\gamma_T = 5dB$. As seen for $m_2 = 1, 2, 4$, the outage curves of $(4, 1, 1)$ setup will have diversity orders of 2, 3, 5, while for $(2, 2, 2)$ these will be 3, 5, 5 respectively. $(2, 2, 2)$ will be less sensitive to changes in m_2 , since saturation occurs over the second hop with increasing m_2 . With $(3, 2, 2)$ the diversity order for $m_2 = 2, 4$ will be 5 and 7 respectively. As before, the

degree of sensitivity to changes in m_1 and m_2 also depends on the diversity gains over the two hops. Similarly when $\bar{\gamma}_2 < \bar{\gamma}_1$, the system will be more sensitive to the fading parameter m_2 . When the ratio improves or when $\bar{\gamma}_1 = \bar{\gamma}_2$, outage performance improves with a stronger second hop, however sensitivity to m_2 decreases. The opposite effect is expected to occur where m_1 is concerned. In other words, when m_2 is kept constant and when $\bar{\gamma}_2 > \bar{\gamma}_1$, the system would be more sensitive to changes in m_1 and less sensitive as the ratio $\bar{\gamma}_2 : \bar{\gamma}_1$ gets smaller. As noted, the fading parameter plays an important role in realising large performance gains in the relay system. $(4, 1, 1)$ is evidently more sensitive to m_2 , where a higher m_2 is beneficial in providing a balance between diversity gains of the first and second hops. A similar effect is seen in the $(2, 2, 2)$ and $(3, 2, 2)$ setups, which performs much better than through the use additional antennas. Importantly, the direct path adds to the overall diversity gain. For example, at $\bar{\gamma}_1 = 10dB$, Fig. 6.7 shows that $(2, 2, 2)$ fares much better if a direct path is present. Also notice that the $(1, 4, 1)$ setup is always better than $(4, 1, 1)$, implying that a single relay employing receive diversity is better than a single antenna selective relaying system with the same number of antennas. However, as observed before, a proper distribution of resources or relay placement is required to ensure maximum performance, since benefits are limited by the strength of the weakest link.

The symbol error probability of the system for a variety of modulation schemes can be derived using the MGF equation 6.33 as:

$$P_s = \frac{1}{\pi} \int_0^{\pi/2} \mathcal{M}_{eq}(-a^2 / \sin^2 \phi) d\phi \quad (6.40)$$

where a is a constant that depends on the specific type of modulation used. Fig. 6.8 plots the capacity for single and multiple antenna relay selection with $\bar{\gamma}_1 = 1.25\bar{\gamma}$, both with and without a direct path taken into consideration, and specific values of m_1, m_2 . In both cases, the inclusion of the direct path yields considerable improvement in capacity. As seen from the lower graph, the capacity gap between $(2, 1, 1)$ and $(7, 1, 1)$ is very small, indicating that more than two relays may not yield substantial gains. Similar observations were made w.r.t outage probability performance in [165], [156].

Figure 6.8: Capacity for single and multi-antenna relay selection scenarios w.r.t $\bar{\gamma}_1$. $\bar{\gamma}_1 = 1.25\bar{\gamma}_2$, with and without direct path.

With multiple antennas, as seen in the upper figure, improvement in capacity can be seen. Hence $(1, 2, 1)$ has slightly more capacity than $(2, 1, 1)$ brought about by signal combining. Another advantage is that it does not require any feedback, since only one relay is used. In general a substantial performance improvement is evident from the use of multiple antennas.

6.3.6 Summary

When a direct path between the source and destination is present, the system performance of the multi-antenna relay selection setup improves considerably. The diversity order of the system was determined as $G_r = m_0 + \min(m_0 L N_R, m_2 K)$. Closed form solutions were determined for the outage probability, MGF and capacity.

6.4 Predictive Dual-Hop Partial Multi-antenna Relay Selection

In this section, we analyse a setup where all nodes in the system multiple receive antennas and a single transmit antenna. We consider outdated relay selection due to the presence of feedback delay. Since relay switching is delay-limited, power prediction is employed to mitigate against the effect of outdated CSI. Recently, the impact of outdated feedback information on relay selection has been considered [175]. Non-selective beamforming for single antenna relay/destination is an alternative strategy [167], while the impact of delayed feedback has been analysed in [176] utilising results from [177]. Da Costa [159] considered multi-antenna source and destination communication with a single antenna relay. In general, partial selection performance did not improve much when more than two relay nodes were available for selection, and was limited by the gain of the second hop.

This section extends the analysis to source, destination, and relay nodes having multiple receive, but single transmit antennas. During transmission, a source selects a best relay on the basis of predicted SNR over that link. The relay employs

maximal ratio combining at its receiver, and applies a variable gain to the received signal before forwarding to the destination, which in turn employs MRC for signal combining. This is a source-directed arrangement with no direct path from source to destination. CSI, used for relay selection, becomes outdated due to Doppler changes, thus impacting selection performance. The main aim is to (a) investigate performance improvement for multiple antennas over a single antenna, (b) investigate the system impact of feedback delay, (c) exploit temporal channel correlation through a predictor, to mitigate against item (b). First hop quality will be shown to improve with additional relays (or antennas), while second hop improvement comes primarily through receive diversity. Variable gain AF naturally makes use of CSI, estimated at the relays, and accurate channel estimates are assumed for MRC combining and demodulation. The first hop statistics are similar to those of Chapter 3, and from our previous work [91], on channel prediction for point-to-point (PtP) linking, so we build upon this foundation below. Closed form outage probability and bit error rate solutions are found for arbitrary numbers of relays and receive antennas, and used to explore trade offs between number of relays and number of antennas compared to single antenna alternatives. To assess predictor performance in combatting switching delay, comparison is made to the non-predictive system.

6.4.1 System model and assumptions

The AF relay network consists of nodes equipped with single transmitter and multiple receivers. Mobile source M_S communicates to N_t relay nodes $R_1, R_2, \dots, R_{N_t}$ each equipped with N_{rr} receive antennas. A selected best relay R_l , $l: 1 \leq l \leq N_t$, chosen based on received SNR, forwards the amplified signal to mobile destination M_D with N_{rt} receive antennas. No direct link exists between M_S and M_D , and information is conveyed half duplex with transmission in two phases, $M_S \rightarrow R_l \rightarrow M_D$. Block fading, with block length L_b and symbol rate T_s , is assumed. Thus the channel matrix for the first hop at instant k is $\mathbf{H}(k) = [h_{ij}(k)]_{N_{rr} \times N_t}$ $1 \leq j \leq N_{rr}, 1 \leq i \leq N_t$, with channels h_{ij} characterised as spatially independent and identically distributed

(i.i.d) complex Gaussian random variables with distribution $\mathcal{CN}(0, 1)$. Channels are temporally correlated and change slowly over transmission blocks - the well known Jakes fading model is used, with f_d as Doppler frequency. To obtain feedback information about the SNRs at the relays, M_S broadcasts pilot channel information as well as data using pilot symbol aided modulation (PSAM). The estimated channel matrix between M_S and relays is then $\tilde{\mathbf{H}}(k) = [\tilde{h}_{ij}(k)]_{N_{rr} \times N_t}$ where $\tilde{h}_{ij}(k)$ is the channel estimate distribution $\mathcal{CN}(0, 1 + \sigma_v^2)$, (assuming estimation and channel noise to be independent) with $\sigma_v^2 = \frac{N_0}{E_p}$ as variance of estimation error, E_p being the power of the pilot symbol and N_0 the AWGN variance, assumed to be the same for each transmit-receive chain. To compensate for known feedback delay $\mathcal{D}$, each relay R_i independently predicts combined branch power for a future transmission slot using Wiener-Hopf filtering (based on noisy estimated CSI at its receiver), and conveys this to the source via a feedback link. The predicted SNR at the i^{th} relay is then given as: $\hat{\gamma}_i(k + \mathcal{D}) = \frac{\mathcal{E}_1}{N_0} \sum_{j=1}^{N_{rr}} \left| \hat{h}_{ij}(k + \mathcal{D}) \right|^2$ where $\mathcal{E}_1$ is the power of the transmitted signal at M_S and $\hat{h}_{ij}(k + \mathcal{D}) = \mathbf{w}_{opt}^H \tilde{\mathbf{h}}_{ij}$ is the $\mathcal{D}$ block ahead predicted channel coefficient. $\tilde{\mathbf{h}}_{ij}$ is the complex vector of estimated fading amplitudes corresponding to prediction length L , defined as:

$$\tilde{\mathbf{h}}_{ij} = [\tilde{h}_{ij}(k), \tilde{h}_{ij}(k-1), \dots, \tilde{h}_{ij}(k-(L-1))]^T \quad (6.41)$$

and $\mathbf{w}_{opt}$ is the Wiener optimal complex coefficient vector $\mathbf{w}_{opt} = [\mathbf{R}]^{-1} \mathbf{r}$ where

$$[\mathbf{R}]_{\varphi\vartheta} = J_0(2\pi f_d |\varphi - \vartheta| L_b T_s) + \sigma_v^2 \delta(\varphi - \vartheta) \quad (6.42)$$

$$\mathbf{r}_{\varphi} = J_0(2\pi f_d |\mathcal{D} + \varphi - 1| L_b T_s); \quad \varphi, \vartheta = 1, 2, \dots, L.$$

$J_0(\cdot)$ is the zeroth order Bessel function of the first kind and $(\cdot)^H$ is the Hermetian transpose. Based upon computed $\hat{\gamma}_i$ fed back to M_S , a single best relay offering best predicted instantaneous SNR, with index l , is selected for the current interval: $\hat{\gamma}_l(k + \mathcal{D}) = \max_{1 \leq i \leq N_t} \hat{\gamma}_i(k + \mathcal{D})$. This relay combines the received signal using MRC. The same initial PSAM estimates are used for channel smoothing, but will yield a cleaner channel estimate, so for MRC combining at the selected relay we assume $\tilde{h}_{lj}(k) = h_{lj}(k)$, $1 \leq j \leq N_{rr}$ denoted as $[\mathbf{h}_{M_S R_l}]_{N_{rr} \times 1}$. This will be the first

hop path $M_S \rightarrow R_l$, with instantaneous MRC weight vector

$$\mathbf{w}_{M_S R_l} = \frac{\mathbf{h}_{M_S R_l}^H}{\sqrt{\mathbf{h}_{M_S R_l}^H \mathbf{h}_{M_S R_l}}} \quad (6.43)$$

The signal is then boosted by variable gain G and conveyed to the destination over the second hop $R_l \rightarrow M_D$ channel $[\mathbf{h}_{R_l D}]_{N_{rd} \times 1}$, which is $\mathcal{CN}(0, \mathbf{I}_{N_{rd}})$ distributed, also assumed correctly estimated for signal demodulation at M_D ($\mathbf{I}$ is an identity matrix) with

$$\mathbf{w}_{R_l M_D} = \frac{\mathbf{h}_{R_l M_D}^H}{\sqrt{\mathbf{h}_{R_l M_D}^H \mathbf{h}_{R_l M_D}}} \quad (6.44)$$

as the MRC weight vector. The final combined output SNR at the M_D is then:

$$y_{R_l M_D} = G \mathbf{w}_{R_l M_D}^H \mathbf{h}_{R_l M_D} (\mathbf{w}_{M_S R_l}^H \mathbf{h}_{M_S R_l} s + \mathbf{w}_{M_S R_l}^H \mathbf{n}_{R_l}) + \mathbf{w}_{R_l M_D}^H \mathbf{n}_{M_D} \quad (6.45)$$

where s is the transmitted symbol, $\mathbf{n}_{R_l}$ and $\mathbf{n}_{M_D}$ are noise vectors at relay and destination, characterised as $\mathcal{CN}(0, N_0 \mathbf{I}_{N_{rr}})$ and $\mathcal{CN}(0, N_0 \mathbf{I}_{N_{rd}})$. After simplification of (6.45), the equivalent SNR at M_D is then calculated as [117]: $\gamma_{eq} = \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + 1}$ with $G^2 = \frac{\mathcal{E}_2}{\mathcal{E}_1 (\mathbf{h}_{M_S R_l}^H \mathbf{h}_{M_S R_l}) + N_0}$ and $\gamma_1 = \frac{\mathcal{E}_1 \mathbf{h}_{M_S R_l}^H \mathbf{h}_{M_S R_l}}{N_0}$ and $\gamma_2 = \frac{\mathcal{E}_2 \mathbf{h}_{R_l M_D}^H \mathbf{h}_{R_l M_D}}{N_0}$ as the instantaneous SNRs of each hop respectively. Average receive SNR per symbol per path for first and second hop is $\bar{\gamma}_1 = \frac{\mathcal{E}_1}{N_0}$ and $\bar{\gamma}_2 = \frac{\mathcal{E}_2}{N_0}$ respectively, with $\mathcal{E}_2$ being the power of the transmitted signal at R_l . In the next section we will formulate fading statistics for the two hops in order to derive outage probability and SER.

6.4.2 Performance metrics

As before, the outage probability P_{out} for a given threshold γ_T in the variable AF relay system is generalised as [117]:

$$\begin{aligned} P_{out}(\gamma_T) &= P_{out}(\gamma_{eq} < \gamma_T) \\ &= \int_0^\infty Pr \left[\frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2 + 1} < \gamma_T | \gamma_2 \right] f_{\gamma_2}(\gamma_2) d\gamma_2 \end{aligned} \quad (6.46)$$

After manipulation the above can also be expressed as:

$$P_{out}(\gamma_T) = 1 - \int_{\gamma_T}^\infty \left[1 - F_{\gamma_1} \left(\frac{\gamma_T (\gamma_2 + 1)}{\gamma_2 - \gamma_T} \right) \right] f_{\gamma_2}(\gamma_2) d\gamma_2 \quad (6.47)$$

where $F_{\gamma_1}(\cdot)$ is the CDF of γ_1 . We now obtain an expression for $F_{\gamma_1}(\cdot)$. Briefly, we recap the derivation of the PDF; this is same as for TASP/MRC in Chapter 3, however the final CDF is cast in a alternative form so as to facilitate integration. To derive its PDF $f_{\gamma_1}(\gamma)$, let $f_{\hat{\gamma}_i}(\hat{\gamma})$ be the PDF of the i.i.d r.v.s corresponding to the predicted SNRs of $S \rightarrow R_{i=1\dots, N_t}$ links. Each $f_{\hat{\gamma}_i}(\hat{\gamma})$ is a gamma PDF [91] $\hat{\gamma}_i \sim \mathcal{G}(N_{rr}, \bar{\gamma})$ with $\bar{\gamma} = \bar{\gamma}_1 \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$. Arranging the r.v.s as $\hat{\gamma}_{(1)} \leq \hat{\gamma}_{(2)} \leq \dots \leq \hat{\gamma}_{(N_t)}$, the relay corresponding to the N_t^{th} order statistic will be selected. Since $\hat{\gamma}_i$ are assumed to be i.i.d, then according to order statistics the PDF of the predicted maximum SNR or $\hat{\gamma}_l$, is² $f_{\hat{\gamma}_l}(\hat{\gamma}) = N_t f_{\hat{\gamma}}(\hat{\gamma}) [F_{\hat{\gamma}}(\hat{\gamma})]^{N_t-1}$ [91]. $F_{\hat{\gamma}}(\hat{\gamma})$ is the CDF of $\hat{\gamma}$ and can easily be shown to be:

$$F_{\hat{\gamma}}(\hat{\gamma}) = 1 - \exp(-\hat{\gamma}/\bar{\gamma}_1) \sum_{m=0}^{N_{rr}-1} (1/m!) (\hat{\gamma}/\bar{\gamma}_1)^m \quad (6.48)$$

The PDF of γ_1 is given as: $f_{\gamma_1}(\gamma) = \int_0^\infty f_{\gamma|\hat{\gamma}}(\gamma|\hat{\gamma}) f_{\hat{\gamma}_l}(\hat{\gamma}) d\hat{\gamma}$, where the corresponding true SNR γ has PDF $\gamma \sim \mathcal{G}(N_{rr}, \bar{\gamma}_1)$ and the conditional density $f_{\gamma|\hat{\gamma}}(\gamma|\hat{\gamma}) = \frac{f_{\gamma, \hat{\gamma}}(\gamma, \hat{\gamma})}{f_{\hat{\gamma}}(\hat{\gamma})}$ where $f_{\gamma|\hat{\gamma}}(\gamma, \hat{\gamma})$ (see eqn. 3.28) is a bi-variate gamma PDF $\mathcal{G}_B(N_{rr}, \bar{\gamma}_1, \bar{\gamma}, \rho_p)$ [91] with $\rho_p = \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$. Using multinomial and binomial expansions along with [98, eqn.(6.643.4)] (A.3.1), a closed form solution of $f_{\gamma_1}(\gamma)$ can be obtained in terms of Laguerre polynomials [98](A.2). $F_{\gamma_1}(z)$ is then found as $F_{\gamma_1}(z) = \int_0^z f_{\gamma_1}(\gamma) d\gamma$ using [98, eqn.(8.350)] and (A.4.1). After rearrangement, it is then advantageous to cast $F_{\gamma_1}(z)$ in a complementary CDF form as:

$$\begin{aligned} F_{\gamma_1}(z) = & 1 - \frac{N_t!}{(N_{rr}-1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i \exp\left(-\frac{(i+1)(z/\bar{\gamma}_1)}{[(i(1-\rho_p)+1)]}\right)}{i!(N_t-1-i)!} \\ & \cdot \sum_{j=0}^{i(N_{rr}-1)} \eta_{N_{rr}}(i, j) (N_{rr}+j-1)! \sum_{k=0}^j \binom{j}{k} \frac{\rho_p^k (1-\rho_p)^{j-k}}{[(i(1-\rho_p)+1)]^j} \\ & \cdot \sum_{l=0}^{k+N_{rr}-1} \frac{(z/\bar{\gamma}_1)^l}{l! (i+1)^{k+N_{rr}-l} [i(1-\rho_p)+1]^l} \end{aligned} \quad (6.49)$$

where z is a constant and $\eta_{N_{rr}}(i, j)$ [98, eqn.(0.314)] is the coefficient of $\lambda^j, j = 0, 1, \dots, i(N_{rr}-1)$, in the expansion of $(\sum_{j=0}^{N_{rr}-1} \lambda^j/j!)^i$. The PDF of the second hop

²Index i is dropped since they have identical distributions, and time indices are omitted assuming a stationary random process.

$R_l \rightarrow D$, $f_{\gamma_2}(\gamma)$, is a well known gamma PDF, which is $\mathcal{G}(N_{rd}, \bar{\gamma}_2)$: $f_{\gamma_2}(\gamma_2) = \frac{\gamma_2^{N_{rd}-1}}{\bar{\gamma}_2^{N_{rd}}(N_{rd}-1)!} \exp(-\frac{\gamma_2}{\bar{\gamma}_2})$ Changing variables in (6.46) enables changing limits from 0 to ∞ . Using $f_{\gamma_2}(\gamma_2)$ and (6.49) in (6.46), we expand out inner binomial terms containing γ_T . Simplifying, and evaluating the integral using [98, eqn.(3.471.9)](see A.4.2), an exact expression for P_{out} becomes a function of $\mathbf{K}_v(\cdot)$, the v^{th} order modified Bessel function of the second kind [98].

$$\begin{aligned}
 P_{out}(\gamma_T) = & 1 - \frac{2N_t!}{\bar{\gamma}_2^{N_{rt}}(N_{rr}-1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!} \exp\left(-\frac{(i+1)\gamma_{th}}{[(i(1-\rho_p)+1)\bar{\gamma}_1]} - \frac{\gamma_{th}}{\bar{\gamma}_2}\right) \\
 & \cdot \sum_{j=0}^{i(N_{rr}-1)} \eta_{N_{rr}}(i, j)(N_{rr}+j-1)! \sum_{k=0}^j \binom{j}{k} \frac{\rho_p^k(1-\rho_p)^{j-k}}{[(i(1-\rho_p)+1)]^j} \\
 & \cdot \sum_{l=0}^{k+N_{rr}-1} \frac{[\gamma_T/(i(1-\rho_p)+1)\bar{\gamma}_1]^l}{l!(i+1)^{k+N_{rr}-l}} \sum_{p=0}^l \frac{l!(\gamma_T+1)^{l-p}}{(l-p)!p!} \sum_{q=0}^{N_{rt}-1} \frac{\gamma_T^{N_{rt}-q}}{(N_{rt}-1-q)!q!} \\
 & \left[\frac{\bar{\gamma}_2(i+1)\gamma_T(1+\gamma_T)}{(i(1-\rho_p)+1)\bar{\gamma}_1} \right]^{\frac{(q+p-l+1)}{2}} \mathbf{K}_{(q+p-l+1)} \left(2\sqrt{\frac{(i+1)\gamma_T(1+\gamma_T)}{(i(1-\rho_p)+1)\bar{\gamma}_1\bar{\gamma}_2}} \right)
 \end{aligned} \tag{6.50}$$

Next, the average SER for coherent demodulation in a slow fading channel is found utilizing the CDF $F_{\gamma_{eq}}(\cdot)$ (or $P_{out}(\cdot)$) [147]: $\bar{P}_s = \frac{\alpha}{\sqrt{(2\pi)}} \int_0^\infty F_{\gamma_{eq}}(\frac{x^2}{\beta}) \exp(\frac{-x^2}{2}) dx$, (α, β are determined by specific constellations). By approximating $\gamma_T \approx \gamma_T + 1$ in (6.50) setting $\gamma_T = \frac{x^2}{\beta}$, the above integral is solved using [98, eqn.(6.621.3)] to obtain a closed form solution for $\bar{P}_s$ (eqn.(6.51)) (this holds good in the moderate to high SNR regime).

6.4.3 Numerical and simulation results

BER performance is evaluated and verified through simulation. $P_{out}(\gamma_T)$ in (6.27) is directly verified by this (since BER is derived using it), but not plotted due to space constraints. In Fig. 6.9, BER for BPSK ($\alpha = 1, \beta = 2$) in different relay configurations, denoted as (N_t, N_{rr}, N_{rd}) are shown against $\bar{\gamma}_1$ for correlations³ of

³Note ρ_p is a function of several variables, finding an appropriate set for any given $\mathcal{D}$ is a general optimisation problem (Chapter 3).

Figure 6.9: BER plot for BPSK under predictive selective relaying, for first hop temporal correlations of $\rho_p = 1, 0.9, 0.7$, comparing single antenna and multi-antenna relay switching.

$\rho_p = 1, 0.9, 0.7$, where $\rho_p = 1$ implies no delay or near perfect prediction. We set $\bar{\gamma}_1 = 5\bar{\gamma}_2$. In general, multiple antennas/ relays can yield large performance improvement, however not always in specific cases, as performance may be limited by channel strength over either the first or second hop (discussed below). As seen for any given $\bar{\gamma}_1$ in the selection plots, BER degrades, as ρ_p decreases. For large delays, correlation tends to zero and diversity gains from relay switching drop to zero, just like a single relay system. For example the $(3, 1, 1)$ and $(2, 2, 2)$ systems tend to $(1, 1, 1)$ and $(1, 2, 2)$ respectively as $\rho_p \rightarrow 0$. The $(2, 2, 2)$ scheme has a diversity order⁴ of 2 for large SNRs, determined by the minimum of the diversity order over the first and second hops [162]. In other words, if the first hop channel strength is greater than the second, increasing N_{rd} is beneficial, while if the second hop diversity gain is strong, increasing N_{rr} or N_t improves performance. The $(1, 3, 1)$ single relay system

⁴The first hop is analogous to PtP un-coded best-antenna TAS/MRC. With delay free switching of T transmit antennas/ R receive chains, it is known that diversity order $G_d = TR$. In the delayed case, it can be proven that $G_d = R$.

Figure 6.10: BER degradation due to feedback delay with and without prediction at $\bar{\gamma}_1 = 8dB$ (top) and $\bar{\gamma}_1 = 15dB$ (bottom) for different relay schemes with $\bar{\gamma}_2 = 5\bar{\gamma}_1$.

$$\begin{aligned}
 \bar{P}_s \approx & \frac{\alpha}{\sqrt{2\pi}} \left[\frac{\sqrt{2\pi}}{2} - \frac{\sqrt{\pi}\beta(N_{rt}-1)!N_t!}{\bar{\gamma}_2(N_{rr}-1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!} \sum_{j=0}^{i(N_{rr}-1)} \eta_{N_{rr}}(i,j)(N_{rr}+j-1)! \right. \\
 & \cdot \sum_{k=0}^j \binom{j}{k} \frac{\rho_p^k(1-\rho_p)^{j-k}}{[(i(1-\rho_p)+1)]^j} \cdot \sum_{l=0}^{k+N_{rr}-1} \frac{(\Gamma[l+N_{rt}+1])^{-1}}{l!(i+1)^{k+N_{rr}-l}[i(1-\rho_p)+1]^l \bar{\gamma}_1^l} \\
 & \sum_{p=0}^l \frac{l!}{(l-p)!p!} \sum_{q=0}^{N_{rt}-1} \frac{\Gamma[N_{rt}-q-p+2l-0.5]}{(N_{rt}-1-q)!q!} \cdot \left[\frac{4(i+1)}{(i(1-\rho_p)+1)\bar{\gamma}_1\bar{\gamma}_2} \right]^{(q+p-l+1)} \\
 & \cdot \left. \frac{\Gamma[N_{rt}+q+p+1.5] {}_2F_1(N_{rt}+q+p+1.5, q+p-l+1.5; l+N_{rt}+1; \xi)}{\left[\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1} + \frac{1}{\bar{\gamma}_2} + \frac{\beta}{2} + 2\sqrt{\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1\bar{\gamma}_2}} \right]^{N_{rt}+q+p+1.5}} \right]
 \end{aligned} \tag{6.51}$$

$\Gamma[\cdot]$: Gamma Function, ${}_2F_1(\cdot, \cdot, \cdot; \cdot)$: Gauss hypergeometric function [98], and $\xi = \frac{\left[\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1} + \frac{1}{\bar{\gamma}_2} + \frac{\beta}{2} - 2\sqrt{\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1\bar{\gamma}_2}} \right]}{\left[\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1} + \frac{1}{\bar{\gamma}_2} + \frac{\beta}{2} + 2\sqrt{\frac{i+1}{(i(1-\rho_p)+1)\bar{\gamma}_1\bar{\gamma}_2}} \right]}$

is better than the delay limited (3, 1, 1) at low SNRs but both converge at high SNRs. Also, note that the feedback bandwidth of the (3, 1, 1) scheme would be 3 times that of (1, 3, 1) and 1.5 times more than either (2, 2, 2) or (2, 2, 3). Similarly, the (2, 2, 3) scheme will tend to (1, 2, 3) at a faster rate, since its diversity order now reduces from 3 for no delay to 2 when delayed. Fig. 6.10 shows this loss due to delay more clearly, where the BER for various setups with $\bar{\gamma}_1 = 8, 15 \text{ dB}$ is plotted. Normalised delay $f_d\tau$, with $\tau = \mathcal{D}L_bT_s$, ranges from a very low 10^{-2} to a large value of 3. The plot shows the effect of this on each setup. Note that, for the non-predictive case, it can be shown that eqn. (6.50) and (6.51) hold, with $\rho_p = J_0^2(2\pi f_d\tau)$, where noise-free, but delayed channel estimates at the receiver branches are used in computing instantaneous channel power at each relay, for feedback to the source. Firstly, the predicted case is evidently superior, tolerating larger values of $f_d\tau$. For very large $f_d\tau$, both predicted and non-predictive schemes converge to their corresponding configuration, as if one relay is used. For example, (2, 1, 1) and (7, 1, 1) tend to (1, 1, 1) while (2, 2, 2), (3, 2, 2) and (4, 2, 2) tend to (1, 2, 2). We can thus say that prediction

is beneficial in preserving spatial diversity gains realised through relay selection. As seen at lower $f_d\tau$, not much improvement is seen going from $(2, 1, 1)$ to $(7, 1, 1)$ with most gains coming from just two relays, however, the former is more delay sensitive than the latter. Importantly, adding relays improves delay tolerance in both predictive and non-predictive cases. Interestingly, increasing N_{rr} is also effective in lowering BER and tolerating delay (see $(2, 3, 2)$ and $(2, 4, 2)$ at high SNR: little degradation is seen). Note that, at low $f_d\tau$ or no delay, and $\bar{\gamma}_1$, while increasing N_t, N_{rr} lowered BER in general, saturation occurred quickly at high SNRs: performance can then be improved only through increasing N_{rd} or G (to strengthen the second hop). At high $f_d\tau$, increasing N_t, N_{rr} counters degradation due to delay - increasing N_{rd} improves BER, but makes the system more delay sensitive. Fig. 6.9 showed this for $(2, 2, 2)$ and $(2, 2, 3)$. Also, Fig. 6.10 shows that BER does not improve when N_{rd} exceeds the first hop diversity order: in delay limited cases, saturation occurs earlier. However, sensitivity to increasing resources, such as N_t and N_{rr} or N_{rd} , depends on the ratio: $\bar{\gamma}_1 : \bar{\gamma}_2$, i.e. the relative hop strengths. A higher ratio decreases sensitivity to N_t, N_{rr} but increases sensitivity to N_{rr} . Careful balancing of resources such as antennas, relays and transmit power, is necessary to maximise performance in a real system.

6.4.4 Summary

Closed form BER and outage probability expressions were derived for a delay limited dual-hop multi-antenna relay selection system. In general, adding a small number of antennas or relays into the system can improve performance, however a careful distribution of antenna resources is needed to realise substantial gains over a single antenna relay system. In terms of switching delay, increasing the number of relays or relay receive antennas can overcome some of the loss due to moderately outdated switching information. Prediction was shown to extend these benefits out to much greater delays.

6.5 Chapter Summary

This chapter has explored the performance of various partial relay selection systems which employed multiple antennas either at the source, relay, or at the destination. The performance metrics of the above cases were thoroughly analysed mainly w.r.t system outage probability and BER. While partial relay selection (PRS) has advantages over opportunistic relay selection (where CSI over both hops are considered for selection), importantly, it was noticed that system performance in PRS heavily depended on the weakest of the two links, thereby careful selection of system resources was required to realise maximum gain. Specifically this could be achieved by balancing the strength over the two hops through the use of multiple antennas, transmit gain, or improving the fading parameter of the weaker hop. Lastly, it was seen that in case of delay limited relay switching, prediction was useful to combat feedback delay; also the use of multiple relays or relay receive antennas was very beneficial in combatting delay. The following table compares and clearly points out the flexibility and improved performance of the proposed methods with some of the recent works in literature.

Paper	Hops/ Relay Gain	Relay Seln.	Multi Ant.	Naka- gami Chl.	Diversity Order Analy.	Analysis w.r.t power imbalance	Closed form expr.	Adv./ Disadv.
[156]	Dual/ Variable	✓ PRS	✗	✗	✗	✗	Outage	Perf. saturation w.r.t relays
[157]	Dual/ Fixed gain	✓ PRS	✗	✗	✗	✗	BER, Outage	Perf. saturation w.r.t relays
[159]	Dual/ Fixed Gain	✗	✓	✓	✗	Minimal w.r.t power imbalance	BER, Outage	Flexible, depends on multiple ants. for perf. improvement.
[178]	Cascaded Multi Hop. Variable	✗	✗	✗	✓ Inferred	✓ w.r.t Chl. Variance	SER Outage	High perf., Highly relies on sync. betw. relay, high b.w
[179]	Multi/ Variable	✗	✗	✓	✓ Graphical analysis	✗ Limited.	MGF SER	High latency, Perf. decreases with hop count
[180]	Dual/ Variable	✗	✓	✓	✓	✗	Outage SER	Flexible but depends on mult. ants. for gains
[164]	Dual/ Variable Gain	✓ Best Relay	✗	✗	✓	✗	SER	Full diversity, Global controller for CSI acquisition needed
[134]	Dual/ Variable Gain	✗ All Participate	✗	✓	✗	✗	MGF	Full diversity, Sync. between nodes needed
[175]	Dual/ Variable Gain	✓ PRS	✗	✗	✗	✓	Outage BER	Degrades with delay Saturation seen with increase in relays
[167]	Dual/ Variable Gain	✓ PRS	✓ At Source	✗	✗	✗	Outage BER	Degrades with delay Saturation seen with increase in relays
Sec. 6.1	Dual Fixed	✓ PRS	✓ S, D	✗	✓ Inferred	✓	Outage, BER, etc	High perf. : careful placement needed
Sec. 6.2	Dual/ Variable Gain	✓ PRS	✓ Relay Rx/Tx	✓	✓	✓	Outage BER	Same as above Model very flexible
Sec. 6.3	Dual Variable Gain	✓ PRS	✓ Relay Rx/Tx	✓	✓	✓	Outage MGF Capacity	Same as above, Model highly flexible
Sec. 6.4	Dual Variable Gain	✓ PRS	✓ R,D	✗	✓	✓	Outage BER	Prediction employed to get maximum improvement

Table 6.1: Comparison of existing relaying methods against proposed schemes in 6.1, 6.2, 6.3 and 6.4.

Chapter 7

Conclusion and Future work

In this chapter, we conclude the dissertation by highlighting the contribution of the work and discuss some future research directions.

7.1 Conclusion

The main focus of this thesis concerned the analysis of selection diversity systems which employed point-to-point MIMO and relaying techniques. Selection diversity in wireless systems has long known to be a low cost method to reduce complexity while providing diversity, array and capacity gains.

Phase I: Firstly, transmit antenna selection with MRC was studied in *delay and rate constrained* feedback links, where channel prediction was applied to combat performance degradation due to outdated feedback. The motivation was that transmit antenna selection has several advantages over other space-time codes in terms of implementation cost and complexity of design, and is well suited for use in down link transmissions. Such closed loop techniques are already part of the 3GPP LTE and IEEE 802.16e standards. The findings in the first phase of work are summarised in a point-by-point fashion as follows:

- To preserve the benefits of TAS, use of a long range channel prediction scheme was proposed and applied in predicting future best antenna states several coherence intervals ahead in order to maximise instantaneous SNR, and the system was shown to outperform the conventional scheme which operated on outdated or delayed feedback.
- Specifically, the system was analytically characterised and several performance

metrics were derived (such as BER, outage probability, SNR gains) in quantifying the superior performance of the proposed scheme.

- While the conventional scheme degraded for normalised delays beyond 0.02, the predictive system sustained delays up to 0.1 in most cases. The predictive scheme proposed, was thus found very beneficial in preserving diversity benefits of the selective systems experiencing high Doppler spreads which are expected in next generation systems where mobility and operating frequencies are on the higher end.
- Tradeoffs were recognised between use of pilot power and prediction filter order in obtaining a superior performance over the conventional method. Specifically the exact minimum requirements of pilot power and filter length for the predictive case to just overtake the non-predictive case were determined in a semi-analytical fashion.
- Lastly, the case of rate constrained feedback with delay was considered. With higher Doppler frequencies feedback rate also increases. The TAS/MRC system was analytically characterised for the case where feedback rate was much slower than frame rate. Results indicate that the system degraded much faster with both feedback delay and slow feedback rate. However, the predictive scheme performed significantly better than the uncompensated system operating on outdated feedback.

Phase II: In the second phase of work, predictive TAS was analysed in the context of adaptive modulation and coding schemes, which are key in increasing system throughput. Two factors can degrade an adaptive modulated TAS system: (a) as in the previous case, delayed feedback information causes errors in selecting the best antenna, (b) operating on delay-free calculated switching boundaries between modulation states also proves to be sub-optimal. The key contributions in this part are as follows:

- For bench-marking the performance of the adaptive modulated system, a closed form solution for the ergodic capacity under prediction errors was first determined analytically.
- In order to offset the combined degradation effect due to channel mismatch, prediction was applied and the future best transmission states were jointly determined to maximise system throughput. Based on the predicted SNR, the system was optimised in finding best transmission parameters with constraints on instantaneous BER, average BER and average transmitted power. Several closed form performance metrics were derived for such a system for arbitrary number of transmit, receive antennas and prediction error variance. The performance of an un-coded adaptive modulated scheme was bench-marked with respect to the ergodic channel capacity.
- Results indicated improvement in rate or spectral efficiency with adaptive modulation and use of multiple antennas. Systems operating on instantaneous BER constraints were more sensitive to prediction error variance compared to those using average BER constraints.
- The derived performance metrics generalised the case for receive diversity.
- Several tradeoff schemes were recognised that were made possible between hardware complexity and quality of CSI feedback.

Phase III: In the third stage of work, the principles of selection diversity were applied to the interesting field of cooperative relay communications concerning partial relay selection in dual hop transmissions. Opportunistic relaying, while advantageous in preserving bandwidth and harvesting spatial diversity gains, is difficult to achieve in practice due to the need for a central controller in obtaining CSI feedback over both hops. Specifically in this thesis, *four* different schemes of partial relay selection were studied in different arrangements where multi-antennas were used either at source, relay or destination. The use of multiple antennas was justified within the scope of the

LTE standard where at least two diversity antennas are recommended for improved downlink performance. The best relay offering highest instantaneous SNR over the first hop forwards data to the destination. Several closed form performance metrics such as BER, MGF, outage probability, average SNR gains, and capacity were derived. The results are summarised:

- In the first scheme, single antenna relays were considered with source beam-forming, with multiple-antenna destination in Rayleigh channels; where a BS communicating with several single antenna relays conveys data to destination. A best relay is selected over the first hop. In such schemes, an important conclusion derived was that in partial relaying, the weakest link determined system performance, with the diversity order of the system determined as $\min(N_t N_r, N_{rd})$. A weak link can be bettered through use of multiple antennas, increasing transmit power or number of relays (in case of first hop) over that hop. A balance between links is therefore necessary to achieve maximum performance.
- The second scheme assumed a single antenna source, and destination terminals, communicating over a selected best multi-antenna relay, in Nakagami channels. The system was analysed for different fading scenarios over the two hops. Use of multiple antennas was found beneficial in improving performance, while maintaining a balance between links. Power imbalance between two hops was also studied, where a weaker link was found to be more sensitive to variations in fading parameter.
- The third scheme extended the previous scheme in multi-antenna relaying by considering a direct weak path communication between source and destination. Importantly it was found that a diversity reception of the direct path, when present, was very beneficial in improving overall diversity and throughput gains. This system was also analysed w.r.t capacity gains. A diversity order of this system was determined as $m_0 + \min(m_1 N_t N_r, m_2 N_{rd})$.

- Lastly, selection multi-antenna relaying was applied to the case of outdated source selection. Capitalizing on results from previous analysis of the TAS system, closed form performance results were derived. The diversity order of this system in the outdated or predicted scenario was determined as $\min(N_r, N_{rd})$. As before, both first and second links had to be balanced enough to obtain best performance. Specifically it was observed that, under large prediction variance or normalised feedback delays and in the low SNR regime, use of extra relays or multiple receive antennas was beneficial in improving performance, while for high SNR and low delay, use of extra resources had minimal effect in improving performance further.

Phase IV: Finally, the design and implementation of a flexible FPGA frame work for researching MIMO algorithms over real channels was presented. Evaluation of MIMO algorithms over real world channels is important in realising the true potential of MIMO in real world channels in the presence of implementation impairments. The implementation was carried out on the advanced STAR platform, capable of handling twelve real-time bi-directional channels. Some key contributions made are listed:

- A flexible FGPA framework in conjunction with an independent control plane defined by the ARM interface for efficient reconfigurability and control.
- Capabilities of uploading arbitrary user generated data transmit patterns onto the hardware, preferably generated using MATLAB, and uploaded into the transmitter block via the ARM interface.
- A parameterized transceiver module is devised so that it is capable of transmitting arbitrary modulated signals of different signal bandwidth in conjunction with other RF interfacing modules.
- A loopback interface triggered by the host processor subsequently allows for capture of channel data into files on the server for offline processing, research analysis and investigation. The setup was thus useful in conducting and validating the performance of new MIMO algorithms over arbitrary wireless channels.

7.2 Future Work

- The performance of the channel prediction to predict best antenna can be improved by use of a unbiased predictor. It was mentioned in Section 3.2.3 that a biased power prediction is employed where the average value of prediction error is not zero. It would be interesting to analyze the gain provided by a bias-compensated or unbiased predictor for the considered TAS/MRC system.
- A cross layer technique combining AMC with ARQ (automatic repeat request) is a well known efficient method to offer further rate improvements. The performance of TAS under adaptive modulation scheme, dealt with in Chapter 4 could be further improved by using ARQ. A truncated ARQ limiting the maximum number of re-transmissions may be employed. It would be interesting to analyse the tradeoff between number of re-transmissions to antenna elements and prediction error variance in such a system.
- The antenna selection scenario can be extended to the $(N_t, K; N_r)$ STBC scheme where K transmit chains out of N_t antennas are selected to maximize received SNR. However, the scheme requires more feedback bits, i.e. $\lceil \log_2 \binom{N_t}{K} \rceil$ bits, depending on the values of N_t and K , with some rate loss if the Alamouti scheme is not used. To lower feedback rate, the scheme can be altered to a restricted selection scheme by dividing the transmit elements into N_t/K subsets and choosing the subset which offers the highest receive gain. Specifically, in addition to using only $\lceil \log_2(N_t/K) \rceil$ feedback bits, this arrangement is expected to be more robust to feedback delays, compared to $(N_t, 1; N_r)$, due to use of K transmit chains in the forward direction, although this will suffer some loss in array gain since transmit power is divided amongst K transmitting chains. It is compelling to use the Alamouti coding scheme, ($K = 2$) so that unity code rate is maintained along with a diversity order of $N_t N_r$ that is extracted from antenna selection for no delay.

- The above Alamouti style arrangement can be applied to the case of a relay selection network where each node is equipped with two transmit ($K = 2$) and N_r receive antennas. For the first hop a STBC source would first select a best relay, amongst N_t candidates, and transmit data to, whereupon data is amplified and sent to the destination. With ML decoding, this scheme can achieve a diversity order of $N_t N_r$ over the first hop and $2N_r$ for the second hop. Overall diversity would be $\min(N_t N_r, 2N_r)$ in the absence of any feedback delay, or for perfect prediction. When delay tends to infinity, the diversity order would be akin to a system with $N_t = 1$, so will be $\min(2N_r, 2N_r)$ or $2N_r$.
- More functionality can be added to the existing testbed with the ultimate aim of developing a reconfigurable MIMO platform which can be configured for different transmission or reception schemes, by efficient task partitioning between FPGA and DSPs. In other words, a flexible MIMO system can be designed in which the MIMO transmission scheme that is best adapted to the current conditions can be selected for analysis. Based on input specifications or channel conditions, the device should be able to effectively reconfigure or task partition different computing blocks. The first step in realising this would be to design a modular and parameterized structure for reconfigurable baseband blocks.

Appendix A

List of integrals and formulas

A.1 Multinomial coefficients

The coefficients $\eta_{N_r}(i, j)$ that appear in eqn. (3.30) were obtained from the following expansion [181]:

$$\left[\sum_{j=0}^{N_r-1} z^j / j! \right]^i = \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) z^j \quad (1)$$

$$\eta_{N_r}(i, 0) = 1, \eta_{N_r}(i, 1) = i; \quad \eta_{N_r}(i, i(N_r - 1)) = 1/((N_r - 1)!)^i$$

$$\eta_{N_r}(i, j) = (1/j) \sum_{l=1}^{J_0} (l(i+1) - j)/l! \eta_{N_r}(i, j-l)$$

$$J_0 = \min(j, N_r - 1), \quad 2 \leq j \leq i(N_r - 1) - 1$$

A.2 Laguerre polynomial

The Laguerre polynomial is given by [99]:

$$L_n^b(x) = \sum_{k=0}^n (-1)^k \binom{n+b}{n-k} \frac{x^k}{k!} \quad (\text{where } \binom{n}{k} \text{ is a binomial coefficient})$$

A.3 Definite integration involving Bessel functions

$$\int_0^\infty x^{n+\frac{1}{2}} e^{-\alpha x} J_v(2\beta\sqrt{x}) dx = n! \beta^v e^{-\frac{\beta^2}{\alpha}} \alpha^{(-n-v-1)} L_n^v \left(\frac{\beta^2}{\alpha} \right) \quad [n+v > -1] \quad (1)$$

where $J_v(\cdot)$ is the Bessel function of the first kind of order v .

$$\int_0^\infty x^\mu e^{-\alpha x^2} K_v(\beta x) dx = \frac{e^{-\frac{\beta^2}{\alpha}}}{\alpha^{\mu/2} 2\beta} \Gamma \left(\frac{1+v+\mu}{2} \right) \Gamma \left(\frac{1-v+\mu}{2} \right) W_{-\mu/2, v/2} \left(\frac{\beta^2}{4\alpha} \right) \quad (2)$$

$[Re(\mu) > |Re(v)| > -1]$

or equivalently:

$$\int_0^\infty x^{\mu-\frac{1}{2}} e^{-\alpha x} K_{2v}(2\sqrt{\beta}x) dx = \frac{e^{\frac{\beta^2}{\alpha}}}{\alpha^\mu 2\beta} \Gamma(1/2 + v + \mu) \Gamma(\mu - v + 1/2) W_{-\mu, v}\left(\frac{\beta^2}{\alpha}\right) \quad (3)$$

$$[Re(\mu + v + 1/2) > 0]$$

where K_v is the Bessel function of second kind of order v , W is the Whittaker function.

$$\int_0^\infty x^{\mu-1} e^{-\alpha x} K_v(\beta x) dx = \frac{\sqrt{\pi}(2\beta)^v \Gamma(v + \mu) \Gamma(\mu - v)}{(\alpha + \beta)^{\mu+v} \Gamma(\mu + 0.5)} \quad (4)$$

$$\cdot {}_2F_1\left(\mu + v; v + 0.5; \mu + 0.5; \frac{\alpha - \beta}{\alpha + \beta}\right)$$

$$[Re(\mu) > |Re(v)|, Re(\alpha + \beta) > 0]$$

where ${}_2F_1$ is the confluent hypergeometric function.

A.4 Integrals involving exponential function

The lower incomplete gamma function is given by:

$$\int_0^u x^{v-1} e^{-\mu x} dx = \frac{(v-1)!}{\mu^v} \left(1 - e^{-\mu u} \sum_{l=0}^{v-1} \frac{(\mu u)^l}{l!}\right) \quad [Re(v) > 0] \quad (1)$$

$$\int_0^\infty x^{v-1} e^{-\frac{\beta}{x} - \alpha x} dx = 2 \left(\frac{\beta}{\alpha}\right)^{v/2} K_v(2\sqrt{\beta\alpha}) \quad [Re(\beta) > 0, Re(\alpha) > 0] \quad (2)$$

$$\int_0^u \frac{x^{v-1} e^{-\mu x}}{x + \beta} dx = \beta^{v-1} \exp(\beta\mu) \Gamma(v) \Gamma(1-v, \beta\mu) \quad (3)$$

$$[|arg\beta| < \pi, Re(\mu) > 0, Re(v) > 0]$$

A.5 Covariance between Gaussian random variables

If x_i are i complex normal random variables with zero mean, with $\mathbb{E}[x_i x_j^*] = C_{ij}$ then [168]:

$$\mathbb{E}[x_1 x_2 x_3 x_4] = C_{12} C_{34} + C_{13} C_{24} + C_{14} C_{23}$$

A.6 Derivation steps for Eqn. 6.14

The integral in eqn. 5.13 is evaluated using

$$\begin{aligned}
 & \int_0^\infty F_{\gamma_{eq}}(x^2/\beta) \exp(-x^2/2) dx \\
 &= \int_0^\infty \exp(-x^2/2) dx - \frac{N_t!}{(N_{ts}-1)!(N_{rd}-1)!\bar{\gamma}_2^{N_{rd}}} \\
 & \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!} \sum_{j=0}^{i(N_{ts}-1)} \eta_{N_{ts}}(i,j)(N_{ts}+j-1)! \\
 & \cdot \sum_{l=0}^{j+N_{ts}-1} \frac{2}{(i+1)^{j+N_{ts}-l} l!} \left(\frac{1}{\bar{\gamma}_1}\right)^l \sum_{p=0}^l \binom{l}{p} C^p \int_0^\infty (x^2/\beta)^{l+(N_{rd}-p)/2} \exp\left(-\frac{x^2(i+1)}{\beta\bar{\gamma}_1}\right) \\
 & \cdot \left[\frac{C(i+1)\bar{\gamma}_2}{\bar{\gamma}_1}\right]^{\frac{(N_{rd}-p)}{2}} \mathbf{K}_{N_{rd}-p} \left(2\sqrt{\frac{x^2 C(i+1)}{\beta\bar{\gamma}_1\bar{\gamma}_2}}\right) \exp(-x^2/2) dx
 \end{aligned} \tag{1}$$

Regrouping the power and exponential terms along with the Bessel function, the integral is derived using A.3.2 or equivalently A.3.3 after identifying the constants that need to be plugged into to arrive at the final solution.

A.7 Derivation steps for Eqn. 6.23

$$\begin{aligned}
 & P_{out}(\gamma_T) \\
 &= \int_0^\infty Pr\left[\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2+1} < \gamma_T | \gamma_2\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \\
 &= \int_0^{\gamma_T} Pr\left[\gamma_1 > \frac{\gamma_T(\gamma_2+1)}{\gamma_2-\gamma_T}\right] f_{\gamma_2}(\gamma_2) d\gamma_2 + \int_{\gamma_T}^\infty Pr\left[\gamma_1 < \frac{\gamma_T(\gamma_2+1)}{\gamma_2-\gamma_T}\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \\
 &= \int_0^{\gamma_T} f_{\gamma_2}(\gamma_2) d\gamma_2 + \int_{\gamma_T}^\infty Pr\left[\gamma_1 < \frac{\gamma_T(\gamma_2+1)}{\gamma_2-\gamma_T}\right] f_{\gamma_2}(\gamma_2) d\gamma_2 \\
 &= 1 - \int_{\gamma_T}^\infty \left\{1 - F_{\gamma_1}\left[\frac{\gamma_T(\gamma_2+1)}{\gamma_2-\gamma_T}\right]\right\} f_{\gamma_2}(\gamma_2) d\gamma_2 \\
 &= 1 - \int_0^\infty \left\{1 - F_{\gamma_1}\left[\frac{\gamma_T(\varphi+\gamma_T+1)}{\varphi}\right]\right\} f_{\gamma_2}(\varphi+\gamma_T) d\varphi
 \end{aligned} \tag{1}$$

Appendix B

Design and Implementation of a MIMO Hardware Test-bed

“Observations always involve theory.”

Edwin Hubble

The past decade has seen distinct advances in the theory of MIMO techniques for wireless communications systems. Investigating the performance of highly sophisticated wireless systems, in particular with multiple transmit and receive antennas, is a difficult task. In particular, the true physical behaviour of wireless channels is quite complex and not too well understood despite the best efforts of research groups and researchers worldwide - including us. A mathematical analysis is however, the first important step in obtaining an understanding of the behaviour and forming a reasonable estimate of the performance limits of a wireless channel. However, in many cases, to maintain mathematical tractability, the uncertainties are simplified via assumptions such as reduced complexity fading model, known constant noise levels, pure additive Gaussian receiver noise, perfect synchronization, linear power amplification, and exact knowledge of channels that change only on packet boundaries. Simplifications such as these often leads to over optimistic receiver performance in mathematical analysis or simulation. Hardware test-beds and platforms have therefore become essential in validating the performance gains over real channels and in the presence of implementation impairments. In addition to the more realistic channel conditions and operating environment, test-beds have the advantage of providing a higher number of scenarios and a more exhaustive data-set.

This chapter discusses the design of a flexible and reconfigurable FPGA (field programmable gate array) architecture for efficient testing of any MIMO algorithm over

a wireless channel. This will utilize the existing advanced STAR (Space-Time Array Research) platform research [182] hardware to investigate aspects of space-time and MIMO implementations in FPGA. The research described in this chapter involves the design of multipurpose firmware for adaptive MIMO system development and testing, based upon reconfigurable FPGA architectures. The analysis and the consequential development of a flexible FPGA-based processing system is presented. To validate the functionality and feasibility of MIMO technologies, we setup a four-transmitter four-receiver MIMO test-bed in a typical indoor environment. The system has been built for researching both implementation issues and testing of MIMO algorithms over real channels. The architecture that distinguishes the test bed as a flexible and reconfigurable system is presented. This includes both the hardware and software aspects, and is followed by a discussion of implementation methods and evaluation of system research capabilities. We first briefly discuss some of the key design issues and considerations to be taken into account while building a test-bed.

B.1 MIMO test-bed design issues

Three key requirements hold while designing a successful MIMO test bed: scalability, flexibility and accessibility. The use of MIMO systems will play a key role in satisfying the ever increasing demand for high data rate services. Thus there is a need for an extensible platform in order to add additional signal processing power (either in terms of DSPs or FPGAs). The system needs to be flexible in balancing the signal processing operations amongst the processing units. While high bandwidth dominated parallel operations can be carried out in FPGAs, control and signal processing detection algorithms may be better carried out in DSPs using High Level Languages (HLLs). A combination of a bus architecture with high-speed point-to-point connections based on a standard bus is most desirable for speed, ease of interfacing and operability of the hardware for debugging or performance analysis. [183]. Lastly, accessibility to code changes and the test bed hardware, preferably to the host PC, should be easy. Designing an efficient wireless MIMO test-bed is a daunting task,

Figure B.1: The Space time array research test-bed

and requires an interdisciplinary knowledge integration, where one must approach it in a most meticulous way due to the complexity and interdependencies between the underlying processes. These processes, each equally important, must succeed independently, yet still form a harmonious union, because this mixture determines the success of the test-bed. A discussion and highlight on some key design considerations that go into the development of an efficient and successful test-bed can be found in Section B.7.

B.2 STAR (Space-time array research) Platform

The STAR platform [182], designed by Tait Electronics Ltd, is a flexible test-bed capable of delivering 12 channels of simultaneous and continuous transmit and receive data. In addition it has baseband signal processing facilities capable of executing MIMO algorithms in real time. The transceiver is connected to a PC (Personal Computer) server through an Ethernet-based data acquisition interface, and integrates an FPGA, DSP and CPU. The platform operates in the 2.4 GHz ISM band with a config-

urable RF bandwidth of either 3.84 MHz or 17 MHz and allows arbitrary modulation formats, space-time codes and MIMO algorithms to be implemented. Section B.3 and B.4 contain a description of the various modules in the platform and how they interface with each other. The structure of the VHDL (very high speed integrated circuits hardware description language code) architecture of the FPGA block and its role in system control and interfacing to the RF front end is then described. Section B.6 deals with the implementation and testing of the 4x4 baseband MIMO transceiver built onto the FPGA. Implementation methods and how the hardware-in-loop concept is used to carry out an offline baseband testing of MIMO signal processing algorithms is discussed.

B.3 System Architecture

As shown in Fig. B.1, the MIMO test-bed consists of three main blocks: The users PCs, PC Server and two platform transceivers with unique IP addresses. The FTP (file transfer protocol) interface and NFS (Network file system) interfaces are Ethernet-based software running on the PC server, mainly used for data acquisition. The web page interface, RS232 interface and telnet interface are lower bandwidth logical connections merely used for control, monitoring and debugging.

The transceiver is a reconfigurable integrated unit capable of providing up to 12 transmit and 12 receive channels by the parallel combination of RF and digital processor sub-systems. The whole platform consists of four sub-units: the digital processor sub-system (digital board), the mixed signal front end (mixed signal board), the RF sub-system and the synthesizer board. An outline of the interfacing details among four sub-units are shown in Fig. B.2. Additional details of these sub-units are provided in Section B.4. The digital board is designed for interfacing to the outside world and for processing base band signals. By closely interfacing an ARM processor, an FPGA, and a DSP, the digital sub-system can be scaled to provide more processing capability and channels. The ARM processor acts as the control CPU on the digital board. The embedding of the Linux operating system enables all the system control

and monitoring functions. The FPGA unit is considered as both the data hub and the main processing resource for low-level, high speed parallel processing. The main role of the DSP unit is to run certain additional algorithms or to provide extra signal processing capability when the FPGA resources are unavailable. On the mixed signal board, four ADCs and DACs (AD9862) operate at a rate of 60 Msamples/s providing four parallel 15MHz digital IF (intermediate frequency) interfaces. The RF boards execute two-step frequency conversions to/from the RF operating frequency within the 2.4GHz to 2.5GHz ISM band. The RF board provides four differential 15MHz IF analogue interfaces to the mixed signal board. The IF-stage consists of switchable SAW filters that provide either a 3.84 MHz or 17 MHz pass-band response.

Figure B.2: Simplified Platform Transceiver Hardware

All signal sources related to RF up/down conversion and the mixed signal interface are derived from the synthesizer board as a common high stability reference. The synthesizer board generates the RF local oscillator (LO), IF LO and ADC/DAC sample clock signals necessary to support the MIMO transceiver. Fig. B.3 shows the detailed STAR architecture.

A typical implementation cycle on the MIMO test bed is as follows. The control scripts, implementation files or simulation data are generated on user's PCs and sent to the server through either telnet or a web page interface. The server simply stores them as files in different dedicated directories. The platform transceivers are programmed to keep searching those dedicated directories on the server through the Linux NFS interface. If any new instructions are found, they will update themselves automatically by downloading and executing control scripts written in certain formats. Following the control scripts, the platform transceiver is able to be reconfigured, read simulation data from the server, modulate and transmit data into a real world wireless channel, then capture the signal, sample, process and store them back as files into the server. Therefore the received data, which could be either over-sampled digital IF waveforms or demodulated information bits, is available on the server for users to download and analyze.

B.4 Sub-units of the STAR Platform Transceiver

B.4.1 Digital Board

The digital board provides the heart of the processing capabilities. On board it has:

- Altera EP1S25B672C6 FPGA clocked at 120 MHz
- Samsung ARM S3C2410 processor clocked at 200 MHz
- TMS320C6416GLZ DSP clocked at 600 MHz

B.4.1.1 FPGA

The FGPA is a Stratix EP1S25B672C6 which provides a central interface to the ARM, DSP and LAN. In addition to this, the majority of the base band signal processing is performed here. The Stratix can be programmed via the JTAG header with a *.sof (serial out file). In practice the user will use the JTAG for quick iterations of FPGA code during debug and integration phase.

B.4.1.2 ARM

The ARM provides the interface to the outside world via serial and Ethernet ports. The CPU on the STAR digital board is a Samsung S3C2410 ARM9 variant running at 200MHz. On power up it executes a boot-loader resident in flash called GRAB (Group Research ARM boot-loader). Embedded Linux is the normal operating environment on the ARM CPU. It is from this that all system-control and monitoring occurs. The ARM connects to the rest of the STAR system through a 32 bit wide parallel memory-mapped interface. The ARM is the master and controls the FPGA as if it were a standard microprocessor peripheral.

Choice of OS: The Linux and ARM combination has been an embedded choice for about 8 years now. Linux was chosen as the embedded OS because of its stability, flexibility and easy to port features. Linux is also the main interface between the system and the outside world in the following ways:

- 32 bit memory mapped interface to the FPGA
- Serial interface to the FPGA
- 10 Mbps Ethernet interface providing telnet and NFS client access outside the board, httpd web server, TCP/IP application support

Due to limited storage capacity available to the system (16 Mbytes local SDRAM), use of NFS mounts makes the storage space accessible by the ARM almost limitless. This is especially useful in offline data recording of captured baseband samples from the FPGA. NFS allows a system to share directories and files with others over a network. By using NFS, users and programs can access files from the ARM almost as if they were local files.

Interfacing detail between the ARM and FPGA: The FPGA is mapped at physical address 0x20000000 in the ARM's memory map. Typically the ARM controls the STAR system by writing predefined binary patterns to predefined locations inside the FPGA. The ARM determines the status by reading from predefined

registers. For bulk data transfers, the ARM writes a pointer value to one register and then repeatedly reads or writes to another address, while the FPGA steps through a memory buffer outputting or inputs words as appropriate.

B.4.1.3 DSP

The DSP provides additional signal processing capabilities. Upon reset the DSP boots from code stored in the associated flash. Reset and clock signals are sourced from the FPGA. The DSP is capable of data capture at a rate of 600 Mbps and an internal RAM of size 1Mb. The main interface for data transfer between the DSP and the rest of the system is through the FPGA. The interface is a 16 bit data bus plus 5 bit address bus and control signals. A serial port debug interface and a JTAG (Joint Test Action Group) connection for boot loading is provided to the host PC for control, real time debugging and programming of the DSP. Although this is a capable device, it has not been used in our implementation work, which requires only ARM scripts and FPGA programming.

B.4.2 Mixed Signal Board

Each Mixed Signal (MS) board provides digital-analogue conversion for 4 transmit channels and analogue-digital conversion for 4 receive channels. Two mixed-signal Analog Devices AD9862 provide the ADC and DAC functions, with 2 channels each. Each transmit DAC is followed by a 3-pole elliptical low-pass filter, attenuating sampling components above 25 MHz. Each receive ADC is preceded by a 5-pole elliptical low-pass anti-alias filter with cut-off above 25 MHz. Both filter types have low delay spread within the IF pass-band. A single 60 MHz clock is fed to the board and drives both chips, which each generate internal clocks from this reference using a delay-locked loop technique. The digital interface is effected with a separate multiplexed bus system for each of transmit and receive pairs, requiring 120 MHz word rate on the bus. Transmit DAC resolution is 14 bits and receive ADC resolution is 12 bits. Control of the ADC and DAC functions is via a serial interface to the digital

Figure B.4: Mixed Signal Chip Architecture [184]

processing board. Also there are auxiliary ADCs and DACs on each chip for system expansion, available at the RF subsystem interface.

B.4.3 The RF boards

The RF boards provide up/down conversion of the signals to/from the RF operating frequency within the 2.4 GHz to 2.5 GHz ISM band. Each board provides two RF channels with differential 15MHz IF analogue interface to the mixed signal board. The transmit up-converter board provides 25dB of gain with saturated output power of +17 dBm. The receive down converter provides 43dB gain with an input 3rd order intercept point of -20 dBm. The noise figure of the receiver up to the input to the A/D is 8dB. The IF stages include SAW filters which provide either 3.84MHz or 17MHz passband response. The receiver board includes an RF switch on the input which attenuates the input signal by 40dB.

B.4.4 Synth Board

The synth board contains the RF LOs, IF LOs and ADC/DAC sample clock signals to necessary to support a 4 to 12 channel MIMO transceiver. The unit generates the following clocks:

- 60MHz ADC/DAC clocks
- 365 MHz 2nd IF LO (380 MHz IF)
- 1620-2320 MHz RF LO (2-2.7 GHz band)

B.5 FPGA Design

As shown in Fig. B.5, the FPGA provides the central interface to the ARM, DSP and the Mixed Signal Board and the Synthesizer board. The VHDL software blocks fall into two categories: the core signal processing block and the interfacing blocks. The core signal processing block is developed to deal with low-level, high speed and parallel baseband signals, though it remains an open issue regarding the digital processing partitioning between FPGA and DSP. The interfacing blocks are designed to interface other components on the platform, and because FPGA has accesses to almost all other configurable components, the ARM is able to control and configure the whole platform simply by reading or writing corresponding control registers embedded in the FPGA interfacing blocks.

The FPGA as such is divided into two blocks depending upon where it derives its clock source. The Top_10M module clocks at 10MHz and is mainly responsible for programming the registers in the Synthesizer and the Mixed Signal chip via a Serial Peripheral Interface. Apart from this it has an address decode and a memory controller unit to maintain status of various configuration registers. The Top_120M unit provides interface to all components associated with the 120MHz clock on board that is fed from the 60 MHz synthesizer clock. The ‘Mem_if_120’ component comprises an address decoder and a memory controller. Addresses are decoded and data written

Figure B.5: The FPGA Interface

to and read from the corresponding registers. The memory controller handles read and write requests from the ARM. The MS component provides glue logic for the main ADCs and DACs on the MS Board and synchronizes 4×4 channel data to the system clock through the use of internal FIFO jitter buffers.

B.5.1 ARM-FPGA interface

There are two ways to write/read to the FPGA via the arm:

1. Using a stand alone port I/O program operated from the ARM; usually done by telnetting into ARM where Linux is running. Two C programs allow us to read and write words to the FPGA.

(a) *writeport 0x < hex address > 0x < hex val >*

(b) *readport 0x < hex address >*

2. The above uses a direct memory mapping to the FPGA memory, where the FPGA is assigned a fixed virtual address in the ARM's memory space. The following

Figure B.6: A few important FPGA VHDL code blocks

can be used to read and write access a given memory location when used in a C application:

```
result = *((unsigned long *) (virt_addr + 0x< hex address >))
*((unsigned long *) virt_addr + 0x< hex address >)=< 0x hex val >
```

The address decode logic unit is used to access FPGA's internal memory and write or read data on it. When the address decode detects a valid chip select, read/write signal from the ARM, it will toggle the appropriate outputs for one clock cycle, and this will enable data to be latched onto the data bus. The FPGA has 2 Mb of memory split into various blocks. When a memory block is defined, an associated *controller component* is created. This allows the user to read/write data to one address and the controller will increment the memory block address. The input data can be multiplexed so that an FPGA internal block and the data bus can both write to the memory block, likewise the input data in can be multiplexed to write to the data bus or an internal FPGA bus. Thus by this process, a user can populate specific transmit test patterns into the FPGA internal transmit memory buffers.

B.6 Implementation of the 4x4 MIMO transceiver

Scenario:

In a completely offline implementation, the receiver directly forwards the data generated by the ADCs to a user's PC. The transmit data is usually a test pattern that is transferred by the users PC to the FPGA internal transmit memory block buffers via the ARM-FPGA memory interface. Then, all the digital signal processing functions, including digital IF demodulation and baseband signal processing, are executed in a PC-based simulation environment using MATLAB. This method evaluates the system performance over the real channels while taking advantages of powerful simulation tools. However, running at a rate of 60 Mbps and 14bits per sample, each single ADC generates a data stream with a bit rate of up to 840 Mbit/s. The data needs to be stored in on-platform-memory or simultaneously forwarded to the PC through a high bandwidth interface. In a narrow band implementation, decimation can be applied after ADCs to reduce the equivalent throughput. In the current configuration, offline implementations are limited by the maximum 2 Mb buffer inside FPGA and the 10 Mbit/s Ethernet bandwidth between the platform and PC server.

Transceiver Design:

Fig. B.9 shows the proposed transceiver design for implementing the 4×4 MIMO base band blocks. Transmit and receive memory blocks are first reserved in the FPGA. The memory controller handles read and write commands from the ARM to effect to and fro transfer of data from these blocks. Apart from this the memory controller is used to read and write status and configuration registers. The transmitter block diagram is also shown along with the VHDL I/O module in Fig. B.8. The transmitter has an internal memory, the size of which is configurable through an configuration register. Each block has a unique ID and data can be populated into a desired block using this tag. Data output rate can be controlled via the 'sample_en' input pin which

Figure B.7: Implementation and testing of a 4×4 MIMO transceiver

is derived from the master clock.

A loopback transfer test is performed to read and verify transmitted data from the transmit blocks into the ARM. To begin, the transmitter is first directly interfaced to the receive modules. A memory transfer is affected by the ARM through setting appropriate bits in the system configuration register, which is read by the 'Read_loopback' module, Fig. B.8(b), which in turn initiates the memory controller to begin reading stored samples from memory. Data sample or read rate is controlled using the 'read_en' pin. Essentially, the 'sample_en' pin in the transmitter block of Fig. B.8(c) and this pin are tied to the same clock pin. Once this test is done, the transmitter and receiver modules are decoupled for actual test over air. The MS component provides glue logic for the main ADCs and DACs on the MS Board and synchronizes 4 transmit and 4 receive channel data to the system clock through the use of internal FIFO jitter buffers. Fig. B.8(a) shows the schematic of the glue logic for interfacing with the receive ADCs. Four channel data is received through ports

Figure B.8: (a) ADC glue logic (b) Read loopback module (c) Transmitter Block

Figure B.9: Loopback test setup

MSD_RX_1.2.D[13:0] and MSD_RX_3.4.D[13:0]. A signal strobe is derived out of output 'Rx.clk.en' and is fed to the transmit module Fig. B.8(b) used to control signal output rate data bandwidth.

B.6.1 Loopback data flow and control:

In order to validate the 4×4 MIMO transceiver, a loopback test has to be performed which is also a part of the debugging process (Fig. B.9). The loopback test validates that the Digital board, Mixed Signal board, Synthesizer and RF Tx, Rx are functioning correctly. This test first configures the MS and Synthesizer boards. This is initiated by first telnetting into the server and running the utility programs for configuring the system.

15 MHz Sine wave test: In order to test out the 4x4 MIMO configurations, a Sine wave generator VHDL code block was introduced into the FPGA design which is attached to the four transmitters. This module was programmed to continuously transmit a 15 MHz carrier through each transmitter to each of the receiver one at a time. For initial validation of the circuit, the transmitters were shorted to the receiver. This test had a two and a four channel option.

Figure B.10: Square and sine loop-backed test samples captured at 60 Msamples/s.

Two channel mode: $Tx1 \rightarrow Rx1, Tx1 \rightarrow Rx2, Tx2 \rightarrow Rx1, Tx2 \rightarrow Rx2$

Four channel mode: $Tx1 \rightarrow Rx1, Tx1 \rightarrow Rx2, Tx2 \rightarrow Rx1, Tx2 \rightarrow Rx2, Tx2 \rightarrow Rx3, Tx3 \rightarrow Rx3, Tx3 \rightarrow Rx4, Tx4 \rightarrow Rx3, Tx4 \rightarrow Rx4$

Testing with arbitrary base band Square Wave data: The above loop-back test was repeated for a 6 MHz square wave test-data with 10 samples per period. These are 14 bit 2s complement baseband samples that can be generated by MATLAB, and then uploaded into the ARM server which in turn populates the transmitter block memory in FPGA. The transmitters and receivers are switched on one by one and the received samples are collected in the receive buffer and collected by the ARM for bit exact comparison. Fig. B.10 shows the received square and sine wave samples, collected by the ARM and transferred to a PC for plotting.

B.7 MIMO test bed design considerations

Designing and implementing a MIMO test bed can be very challenging task. The test-bed presents itself as a very complex system to be built and subsequently debugged. Further evaluation and validation can be an extremely time consuming and

cumbersome. Important design considerations in developing a successful test bed are presented and critical implementation issues are pointed out. We discuss in brief some key design points to be kept in mind while designing a test bed.

B.7.1 Clocks and synchronization

Several high-rate future generation wireless access standards rely on accurate timing for transmission and reception intervals and therefore large timing or frequency offsets can often not be tolerated. Coordinated transmission schemes such as MIMO requires a synchronous clock over different transmit or receive antenna modules. As the number of antennas grow, it often becomes necessary to distribute D/A and A/D conversion to multiple boards. In particular, in a carrier-based system with multiple modules, clock distribution and synchronization becomes an important issue. Further memories, interpolators, filters, and digital mixers may not allow for this [184], especially if instead of FPGAs, dedicated hardware is used. Re-synchronization measures have to be taken to effectively design a system with globally asynchronous components. This can be tedious and error-prone and adds additional hardware and latency especially in systems that use stream based processing (e.g., CDMA) [183].

B.7.2 Processing requirements

The complexity of MIMO systems generally scales linearly or quadratically with the number of antennas [185]. However, the complexity of MIMO decoding itself (typically due to matrix inversions or maximum likelihood -detection) can quickly become the dominant part [186]. A quantification of the processing requirements, however, is certainly difficult to make as it strongly depends on the implemented system and algorithms. Dedicated DSP processors have traditionally been used for wireless baseband processing. A survey of available devices as per [187], updated here, reveals clock rates of up to 1 GHz. Leading edge DSPs such as those from Texas Instruments Inc. and Analog devices Inc. contain multiple independent multiply-accumulate (MAC) cores, with capable of several thousand 16-bit MMACS (million multiply accumulates

per second). However the MMAC and other figures are upper bound peak values: whether these are achievable depends very much on software structure, other concurrent operations, and the requirements for external memory. A very rough estimate of complexity was given for a 2-channel Alamouti [26] implementation of 3 billion multiply-accumulate calculations per second [188]. It is evident that both devices are capable of a peak processing speed of the approximately required 3 billion calculations per second but do not sufficiently exceed this. A more detailed analysis reveals problems of memory bandwidth and input-output bus bandwidths that would effectively prevent the devices from handling the large data throughput required, without careful design of supporting hardware. For larger systems the computational capabilities are required to significantly exceed this if such modules were expected to be able to perform meaningful processing. In conclusion, a DSP only is appropriate when the data rate of in and outgoing data is much smaller than the clock speed the DSP performs its operations.

Such supporting hardware would probably be best achieved using a reprogrammable device such as an FPGA. Focusing on FPGA devices revealed the potential for performing all calculations in FPGA. A brief survey of contemporary FPGA devices reinforces this conclusion. Both Altera and Xilinx have bigger and faster FPGA devices currently, which include the Stratix II EP2S180 FPGA from Altera with 179400 logic elements (LEs) and 96 DSP blocks each capable of 4MACs at up to 420 MHz when paired to support 18-bit operation. In this device, use of the DSP blocks alone delivers up to 161280 MMACS even when none of the built-in logic element resources are reserved for processing. The largest Xilinx FPGA, the Virtex-4 series XC4VSX55, has 55295 logic cells, 512 embedded XtremeDSP slices each capable of a single 18x18 multiply, and operates at up to 500MHz (256000MMACS).

B.7.3 Memory Configuration and Processor Host Control

Connecting modular bulk memory to an FPGA allows many types and configurations of memory to be supported by changing the parameters of the memory module to

suit the style of the data bus width. Block processing schemes have proven to be significantly more efficient than stream processing. However, on the other hand this gain often comes along with large memory requirements that may exceed the already significant amounts of storage on today's FPGAs. It is therefore advantageous to ensure that sufficient additional memory is or can be made available on a system. Having an independent control plane allows the host to directly access and control each resource without impacting data bandwidths, needed for overall control and synchronization. This is crucial in the development of a testbed and allows the host microprocessor to have control over parameters of reconfigurable blocks and allows for easy programming of the computational blocks. Some of these points are in Chapter 7 with regard to the development of test beds and the description of the STAR platform.

B.7.4 Digital IF (intermediate frequency) Architecture

Predictability, reliability and repeatability have made digital filters preferable to traditional analogue filters for both high and low frequency applications. For maximum flexibility it is advisable to push samples as close to the antennae as possible. The reduced number of converters needed for a digital IF realization becomes another strong argument in a MIMO scenario [183] .

B.7.5 Partitioning and Control plane

A large complexity of signal processing algorithms and control functions need to be supported, requiring partitioning algorithms potentially over several FPGAs and DSPs. This currently requires engineering intuition and optimization by hand. Several interface drivers to support mixed DSP-FPGA modes have to be supported. Intelligent partitioning between these two device types gives wireless systems the best combination of features and cost-effectiveness. Most architectures implement the system control, configuration, and the signal-processing data path using a combination of microcontrollers (MCUs), FPGAs and programmable DSPs. The MCU controls the system, while the FPGA and DSP handle the data-flow processing. Sys-

tems with light processing demands and control-oriented tasks are implemented in software on a DSP; heavier loads may best be implemented in FPGAs that provide significant parallel processing benefits. Bit-level functions like randomization, interleaving/de-interleaving, bit-loading, mapping to signal constellation, are relatively straightforward and not computationally intensive and easily done in a DSP while computationally intensive algorithms can be processed in an FPGA.

B.7.6 Existing Test-beds

The Vienna MIMO test bed [187] is a flexible test-bed developed to examine MIMO algorithms and channel models described in literature by transmitting data at 2.45 GHz through real, physical channels, supporting simultaneously four transmit and four receive antennas. Combining the advantages of MATLAB and DSP/FPGA environment, the MIMO test-bed allows for rapid verification of baseband algorithms and their critical parts with minimum effort.

The Berkeley Emulation Engine (BEE2) is a generic multi-purpose FPGA based emulation platform consisting of 5 Vertex-II Pro70 FPGAs, each with an embedded PowerPC 405 core capable of supporting the most computationally intensive real-time applications [189]. Each BEE2 supports up to 16 single antenna radios operating on the 2.4 GHz band with a 14-bit 128 MHz DAC, 12-bit 64 MHz ADC, and 20 MHz bandpass filters.

Like the BEE2, the UCLA Hybrid Network Testbed is a high density FPGA system but also incorporates custom ASICs that provide real-time signal processing on the order of 100s of GOPS [190]. The platform is scalable up to 4 Tx. by 4 Rx. antennas and can support signals of up to 25 MHz in bandwidth operating on the 5.2 GHz band.

The Rice Wireless Open-Access Research Platform (WARP), on the other hand, is a single Xilinx Virtex-II Pro XC2VP70 based real-time MIMO testbed, also capable of supporting up to four single antenna radios, however, operates on either the 2.4 GHz or 5.2 GHz band [191]. Each radio has a signal bandwidth of 20 MHz and is

equipped with a 14 bit 65 Msamples/s ADC and 16-bit 160 Msamples/s DAC.

Lastly, the Software Radio Laboratory at Georgia Tech has designed a 3 by 3 MIMO transceiver testbed that combines three synchronized Agilent ESG4438C signal generators to form the multiple antenna transmitter [192]. The receiver operates on the 2.4 GHz band and is composed of three synchronized receive chains each with an RF down-converter, 12-bit 25 Msamples/s ADC, digital down converter (DDC), and MPC7410 32-bit PowerPC for realtime baseband signal processing. These testbeds provide tremendous processing capabilities intended to be used in the analysis of important interactions of higher network layer protocols under constantly changing, real-time network conditions. The Tait STAR platform [182] described in this chapter is another testbed capable of handling 12x12 simultaneous receive and transmit channels.

B.8 Summary

The chapter has described the MIMO STAR platform and its usefulness in validating a framework for the FPGA design of a 4x4 MIMO baseband transceiver. The 4x4 MIMO framework has been demonstrated and validated through loop-back tests using both in-built sine, square and arbitrary waveforms generated offline using MATLAB. Thus the test-bed is useful in validating future research work in MIMO technology and also in developing efficient working reconfigurable architectures that are demonstrable. The chapter also presented design issues that need to be considered while designing a MIMO test bed.

Publications

Journals

1. Shiva Prakash, Ian McLoughlin, “Effects of Channel Prediction for Transmit Antenna Selection with Maximal Ratio Combining in Rayleigh Fading,” *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2555-2568, July. 2011.
2. Shiva Prakash, Ian McLoughlin, “Performance of Dual-Hop Multi-Antenna Systems with Fixed Gain Amplify-and-Forward Relay Selection,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1709-1714, June. 2011.
3. Shiva Prakash, Ian McLoughlin, “Performance Analysis of Adaptive Modulation with Selective Transmitter Antenna Diversity under Channel Prediction Errors,” *IEEE Transactions on Vehicular Technology*, *Under Revision*.
4. Shiva Prakash, Ian McLoughlin, “Predictive Non-Regenerative Dual-Hop Partial Multi-antenna Relay Selection,” *IEEE Signal Processing Letters*, *Under Revision*.
5. Shiva Prakash, Ian McLoughlin, “Analysis of Dual-Hop Amplify and Forward Multiantenna Relay Selection in Nakagami- m Channels,” *IEEE Communication Letters*, *Under Review*.
6. Shiva Prakash, Ian McLoughlin, “Dual-Hop Amplify and Forward Multi-antenna Partial Relay Selection in Nakagami- m Channels” *IEEE Transactions on Vehicular Technology*, *Under Review*.

Conferences

7. Shiva Prakash, Ian McLoughlin, “Analysis of Adaptive Modulation with Antenna Selection under Channel Prediction Errors”, in the Proceedings of IEEE International Conference on Communications and Signal Processing Conference (ICCSP), Calicut, India, 10-12 February 2011.

8. Shiva Prakash, Ian McLoughlin, “Predictive Transmit Antenna Selection with Maximal Ratio Combining”, in the Proceedings of IEEE Global Communications Conference (GLOBECOM), Honolulu, Hawaii, 30 November-4 December 2009.
9. Shiva Prakash, Ian McLoughlin, “Channel prediction for mitigating feedback link issues in transmit antenna selection systems,” in the Proceedings of 20th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Tokyo, Japan, 13-16 September 2009.
10. Ian McLoughlin, Shiva Prakash, “Predictive Receive-Directed Antenna Selection for Quasi-Static Rayleigh Fading Channel” in Proceedings of ACM International Conference on Advanced Infocomm Technology, (ICAIT), Shenzhen, China, 28-31 July 2008.

References

- [1] A. Paulraj, D. Gore, R. Nabar, and H. Bolcskei, “An overview of MIMO communications - A key to gigabit wireless,” *Proc. IEEE*, vol. 92, no. 2, pp. 198 – 218, Feb. 2004.
- [2] J. Laneman, D. Tse, and G. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062 – 3080, Dec. 2004.
- [3] D. Brennan, “Linear diversity combining techniques,” *Proc. IEEE*, vol. 91, no. 2, pp. 331 – 356, Feb. 2003.
- [4] J. Winters, “On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment,” *IEEE J. Sel. Areas Commun.*, vol. 5, no. 5, pp. 871 – 878, Jun. 1987.
- [5] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, “A simple cooperative diversity method based on network path selection,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659 – 672, Mar. 2006.
- [6] J. Laneman and G. Wornell, “Energy-efficient antenna sharing and relaying for wireless networks,” in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Oct. 2000, vol.1, pp. 7 –12.
- [7] “Measuring the Information Society 2010,” <http://www.itu.int/ITU-D/ict/statistics//>, 2010.
- [8] T. Rappaport, *Wireless Communications: Principles and Practice*,. Prentice Hall, 2002.
- [9] 3GPP, TR 36.913, “Requirements for Further Advancements for E-UTRA (LTE-Advanced) (Release 8),” <http://www.3gpp.org>.

- [10] 3GPP TS 36.213, “3GPP TS 36.213: Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures,” <http://www.3gpp.org>.
- [11] H. Holma and A. Toskala, *WCDMA for UMTS: HSPA Evolution and LTE*. Wiley, 2010.
- [12] C. Zhou, N. Guo, B. Sadler, and R. Qiu, “Performance study on time reversed impulse MIMO for UWB communications based on measured spatial UWB channels,” in *Proc. IEEE Military Commun. Conf. (MILCOM)*, 2008.
- [13] I. Telatar, “Capacity of multi-antenna Gaussian channels,” *Eur. Trans. on Commun.*, vol. 10, no. 6, pp. 585–596, Nov./Dec 1999.
- [14] G. Foschini and M. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [15] M. Jankiraman, *Space-time codes and MIMO systems*. Artech House Publishers, 2004.
- [16] R. Bultitude, T. Schenk, N. den Kamp, and N. Adnani, “A Propagation-Measurement-Based Evaluation of Channel Characteristics and Models Pertinent to the Expansion of Mobile Radio Systems to Frequencies Beyond 2 GHz,” *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 382–388, Mar. 2007.
- [17] S. Howard and K. Pahlavan, “Doppler spread measurements of indoor radio channel,” *IET Electron. Lett.*, vol. 26, no. 2, pp. 107–109, 2008.
- [18] S. Thoen, L. Van der Perre, B. Gyselinckx, and M. Engels, “Performance analysis of combined transmit-SC/receive-MRC,” *IEEE Trans. Commun. Technol.*, vol. 49, no. 1, pp. 5–8, Jan. 2001.
- [19] M. Simon and M. Alouini, *Digital communication over fading channels: A Unified Approach to Performance Analysis*, 1st ed. New York, NY: John Wiley and Sons.

- [20] C. Waldschmidt, C. Kuhnert, M. Pauli, and W. Wiesbeck, "Handy MIMO," *IEE Commun. Engr.*, vol. 3, pp. 22–25, 2005.
- [21] A. Goldsmith, *Wireless communications*. Cambridge University press, 2005.
- [22] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects ," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619 –2692, Oct. 1998.
- [23] T. Lo, "Maximum ratio transmission," *IEEE Trans. Commun. Technol.*, vol. 47, no. 10, pp. 1458 –1461, Oct. 1999.
- [24] P. Dighe, R. Mallik, and S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun. Technol.*, vol. 51, no. 4, pp. 694–703, 2003.
- [25] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [26] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451 –1458, Oct. 1998.
- [27] A. Molisch and M. Win, "MIMO systems with antenna selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46 – 56, Mar. 2004.
- [28] Z. Chen, J. Yuan, and B. Vucetic, "Analysis of Transmit Antenna Selection/Maximal-Ratio Combining in Rayleigh Fading Channels," *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1312 – 1321, July 2005.
- [29] H. David and H. Nagaraja, *Order statistics*. Wiley-Interscience, 2003.
- [30] M. Alouini and M. Simon, "Performance of coherent receivers with hybrid SC/MRC over Nakagami- m fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1155–1164, 1999.

- [31] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572 – 584, Sep. 1979.
- [32] A. Forenza *et al.*, "Adaptive mimo transmission for exploiting the capacity of spatially correlated channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 619 –630, 2007.
- [33] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 239–241, 2002.
- [34] G. Ganesan, P. Stoica, and E. Larsson, "Diagonally weighted orthogonal space-time block codes," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, 2003, vol.2, pp. 1147–1151.
- [35] Y. Xue and A. Gershman, "Alamouti-type wireless communication systems with one-bit feedback," in *Proc. Sensor Array and Multichannel Signal Process. Workshop*, Jul. 2004, pp. 143 – 147.
- [36] A. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun. Technol.*, vol. 45, no. 10, pp. 1218 –1230, Oct. 1997.
- [37] A. Forenza *et al.*, "Adaptive MIMO transmission for exploiting the capacity of spatially correlated channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 619–630, 2007.
- [38] C.-B. Chae *et al.*, "Adaptive MIMO transmission techniques for broadband wireless communication systems," *Communications Magazine, IEEE*, vol. 48, no. 5, pp. 112 –118, May 2010.
- [39] D. Love, R. Heath Jr, W. Santipach, and M. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, 2004.

- [40] S. Zhou and G. Giannakis, "Adaptive modulation for multi-antenna transmissions with channel mean feedback," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2003, vol.4, pp. 2281 – 2285.
- [41] J. Akhtar and D. Gesbert, "Partial feedback based orthogonal block coding," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, 2003, vol.1, pp. 287–291.
- [42] W. Jakes, *Microwave mobile communications*. Wiley-IEEE Press, 1994.
- [43] S. Hu, H. Hallen, and A. Duel-Hallen, "Physical channel modeling, adaptive prediction and transmitter diversity for flat fading mobile channel," in *Proc. IEEE Int. Workshop on Signal Process. Advances for Wireless Commun.(SPAWC)*, 1999.
- [44] D. Duong, G. Oien, and K. Hole, "Adaptive coded modulation with receive antenna diversity and imperfect channel knowledge at receiver and transmitter," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 458 – 465, Mar. 2006.
- [45] S. Falahati *et al.*, "Adaptive modulation systems for predicted wireless channels," *IEEE Trans. Commun. Technol.*, vol. 52, no. 2, pp. 307–316, 2004.
- [46] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986 –1992, Nov. 1997.
- [47] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li, "MIMO systems with adaptive modulation," *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1828–1842, 2005.
- [48] Z. Zhou and B. Vucetic, "Design of adaptive modulation using imperfect CSI in MIMO systems," *IET Electron. Lett.*, vol. 40, no. 17, pp. 1073–1075, 2004.
- [49] Z. Zhou, B. Vucetic, Z. Chen, and Y. Li, "Design of adaptive modulation in MIMO systems using outdated CSI," in *Proc. IEEE Symp. Personal Indoor and Mobile Radio Commun. (PIMRC)*, 2006, pp. 1101–1105.

- [50] Z. Zhou and B. Vucetic, "Adaptive coded mimo systems with near full multiplexing gain using outdated csi," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 294–302, Jan. 2011.
- [51] T. Eng, N. Kong, and L. Milstein, "Comparison of diversity combining techniques for Rayleigh-fading channels," *IEEE Trans. Commun. Technol.*, vol. 44, no. 9, pp. 1117–1129, Sep. 1996.
- [52] G.-T. Chyi, J. Proakis, and C. Keller, "On the symbol error probability of maximum-selection diversity reception schemes over a Rayleigh fading channel," *IEEE Trans. Commun. Technol.*, vol. 37, no. 1, pp. 79–83, Jan. 1989.
- [53] N. Kong and L. Milstein, "Average SNR of a generalized diversity selection combining scheme," *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 57–59, Mar. 1999.
- [54] M.-S. Alouini and M. Simon, "An MGF-based performance analysis of generalized selection combining over Rayleigh fading channels," *IEEE Trans. Commun. Technol.*, vol. 48, no. 3, pp. 401–415, Mar. 2000.
- [55] M. Win and J. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun. Technol.*, vol. 49, no. 11, pp. 1926–1934, Nov. 2001.
- [56] D. Gore and A. Paulraj, "Mimo antenna subset selection with space-time coding," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2580–2588, Oct. 2002.
- [57] A. Ghrayeb and T. Duman, "Performance analysis of MIMO systems with antenna selection over quasi-static fading channels," *IEEE Trans. Veh. Technol.*, vol. 52, no. 2, pp. 281–288, Mar. 2003.
- [58] H. Shen and A. Ghrayeb, "Analysis of the outage probability for MIMO systems with receive antenna selection," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1435–1440, Jul. 2006.

- [59] T. Cover, J. Thomas, and J. Wiley, *Elements of information theory*. Wiley Online Library, 1991, vol. 1.
- [60] A. Molisch, M. Win, and J. Winters, "Capacity of mimo systems with antenna selection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2001, vol.2, pp. 570–574.
- [61] P. Smith, T. King, L. Garth, and M. Dohler, "A power scaling analysis of norm-based antenna selection techniques," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3140 –3149, Aug. 2008.
- [62] A. Gorokhov, D. Gore, and A. Paulraj, "Receive antenna selection for mimo flat-fading channels: theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2687 – 2696, Oct. 2003.
- [63] J. Heath, R.W. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2001, vol.7, pp. 2276 –2280.
- [64] Z. Chen, J. Yuan, B. Vucetic, and Z. Zhou, "Performance of Alamouti scheme with transmit antenna selection," *IET Electron. Lett.*, vol. 39, no. 23, pp. 1666 – 1668, Nov. 2003.
- [65] D. Gore, J. Heath, R.W., and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491 – 493, Nov. 2002.
- [66] D. Gore, R. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, 2000, vol.5, pp. 2785 –2788.
- [67] R. Blum and J. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 322 – 324, Aug. 2002.

- [68] Z. Xu, S. Sfar, and R. Blum, "Analysis of MIMO Systems With Receive Antenna Selection in Spatially Correlated Rayleigh Fading Channels," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 251–262, Jan. 2009.
- [69] L. Dai, S. Sfar, and K. Letaief, "Optimal antenna selection based on capacity maximization for mimo systems in correlated channels," *IEEE Trans. Commun. Technol.*, vol. 54, no. 3, pp. 563–573, Mar. 2006.
- [70] T. Gucluoglu and E. Panayirci, "Performance of transmit and receive antenna selection in the presence of channel estimation errors," *IEEE Commun. Lett.*, vol. 12, no. 5, pp. 371–373, May. 2008.
- [71] Q. Ma and C. Tepedelenlioglu, "Antenna selection for space-time coded systems with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 710–719, Feb. 2007.
- [72] Y. Li, N. Mehta, A. Molisch, and J. Zhang, "Optimal signaling and selection verification for single transmit-antenna selection," *IEEE Trans. Commun. Technol.*, vol. 55, no. 4, pp. 778–789, Apr. 2007.
- [73] A. Wittneben, "Analysis and comparison of optimal predictive transmitter selection and combining diversity for DECT," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 1995, vol.2, pp. 1527–1531.
- [74] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Commun. Lett.*, vol. 2, no. 9, pp. 254–256, Sep. 1998.
- [75] S. Hu, T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Transmitter antenna diversity and adaptive signaling using long range prediction for fast fading DS/CDMA mobile radio channels," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, 1999, vol.2, pp. 824–828.

- [76] S. Guncavdi and A. Duel-Hallen, "Performance analysis of space-time transmitter diversity techniques for WCDMA using long range prediction," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 40 – 45, Jan. 2005.
- [77] T. Ekman, "Prediction of mobile radio channels, modeling and design," Ph.D. dissertation, Uppsala University, Sweden, 2002.
- [78] T. Ekman, M. Sternad, and A. Ahlén, "Unbiased power prediction of Rayleigh fading channels," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, 2002, vol.1, pp. 280 – 284.
- [79] G. Øien, H. Holm, and K. Hole, "Impact of channel prediction on adaptive coded modulation performance in Rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 758 – 769, May 2004.
- [80] S. Zhou and G. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?" *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1285 – 1294, July 2004.
- [81] K. Baddour, C. Squires, and T. Willink, "Mobile Channel Prediction with Application to Transmitter Antenna Selection for Alamouti Systems," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Sept. 2006, pp. 1–6.
- [82] J. Tang and X. Zhang, "Error probability analysis of TAS/MRC-based scheme for wireless networks [point-to-point link example]," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Mar. 2005, vol. 2, pp. 877 – 882.
- [83] R. Radaydeh, "Impact of delayed arbitrary transmit antenna selection on the performance of rectangular qam with receive mrc in fading channels," *Communications Letters, IEEE*, vol. 13, no. 6, pp. 390 – 392, Jun. 2009.
- [84] X. Cai and G. Giannakis, "Performance analysis of combined transmit selection diversity and receive generalized selection combining in Rayleigh fading chan-

- nels,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1980 – 1983, Nov. 2004.
- [85] Z. Chen, Z. Chi, Y. Li, and B. Vucetic, “Error performance of maximal-ratio combining with transmit antenna selection in flat Nakagami-m fading channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 424 –431, Jan. 2009.
- [86] J. Romero-Jerez and A. Goldsmith, “Performance of multichannel reception with transmit antenna selection in arbitrarily distributed Nakagami fading channels,” *IEEE Trans. Wireless Commun.*
- [87] Z. Chen, I. Collings, Z. Zhou, and B. Vucetic, “Transmit antenna selection schemes with reduced feedback rate,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 1006 –1016, Feb. 2009.
- [88] L. Li, S. Vorobyov, and A. Gershman, “Transmit antenna selection based strategies in miso communication systems with low-rate channel state feedback,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1660 –1666, april 2009.
- [89] W. Zhang and C. Tellambura, “Performance Analysis of Joint Transmit and Receive Antenna Selection With Orthogonal Space-Time Coding,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2631 –2635, Jun. 2010.
- [90] T. Ramya and S. Bhashyam, “Using delayed feedback for antenna selection in MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6059 –6067, Dec. 2009.
- [91] S. Prakash and I. McLoughlin, “Predictive Transmit Antenna Selection with Maximal Ratio Combining,” in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.
- [92] J. Cavers, “An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio],” *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686 –693, Nov. 1991.

- [93] G. Stüber, *Principles of mobile communication*. Kluwer Academic Publishers, 2001.
- [94] M. Alouini and A. Goldsmith, “Adaptive modulation over Nakagami fading channels,” *Wireless Personal Communications*, vol. 13, no. 1, pp. 119–143, May 2000.
- [95] X. Tang, M.-S. Alouini, and A. Goldsmith, “Effect of channel estimation error on M-QAM BER performance in Rayleigh fading,” *IEEE Trans. Commun. Technol.*, vol. 47, no. 12, pp. 1856 –1864, Dec. 1999.
- [96] D. Goeckel, “Adaptive coding for time-varying channels using outdated fading estimates,” *IEEE Trans. Commun. Technol.*, vol. 47, no. 6, pp. 844 –855, Jun. 1999.
- [97] X. Cai and G. Giannakis, “Adaptive PSAM accounting for channel estimation and prediction errors,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 246 – 256, Jan. 2005.
- [98] I. Gradshteyn and I. Ryzhik, *Table of integrals , Series, and Products*, vol. 22.
- [99] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, 1972.
- [100] U. Charash, “Reception Through Nakagami Fading Multipath Channels with Random Delays,” *IEEE Trans. Commun. Technol.*, vol. 27, no. 4, pp. 657 – 670, Apr. 1979.
- [101] M. Win and J. Winters, “Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading,” *IEEE Trans. Commun. Technol.*, vol. 47, no. 12, pp. 1773 –1776, Dec. 1999.
- [102] Q. Sun, D. Cox, H. Huang, and A. Lozano, “Estimation of continuous flat fading MIMO channels,” in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, Mar. 2002, vol.1, pp. 189 – 193.

- [103] J. Hayes, “Adaptive Feedback Communications,” *IEEE Transactions on Communication Technology*, vol. 16, no. 1, pp. 29–34, Feb. 1968.
- [104] J. Cavers, “Variable-Rate Transmission for Rayleigh Fading Channels,” *IEEE Trans. Commun. Technol.*, vol. 20, no. 1, pp. 15–22, Feb. 1972.
- [105] W. Webb and R. Steele, “Variable rate QAM for mobile radio,” *IEEE Trans. Commun. Technol.*, vol. 43, no. 7, pp. 2223–2230, Jul. 1995.
- [106] A. Duel-Hallen, S. Hu, and H. Hallen, “Long-range prediction of fading signals,” *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.
- [107] D. Duong and G. Oien, “Optimal Pilot Spacing and Power in Rate-Adaptive MIMO Diversity Systems with Imperfect CSI,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 845–851, Mar. 2007.
- [108] T. Kolding *et al.*, “High speed downlink packet access: WCDMA evolution,” *IEEE Veh. Technol. Soc. News*, vol. 50, no. 1, pp. 4–10, 2003.
- [109] R. Mallik *et al.*, “Channel capacity of adaptive transmission with maximal ratio combining in correlated Rayleigh fading,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1124–1133, July 2004.
- [110] M.-S. Alouini and A. Goldsmith, “Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques,” *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1165–1181, Jul. 1999.
- [111] B. Choi and L. Hanzo, “Optimum mode-switching-assisted constant-power single and multicarrier adaptive modulation,” *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 536–560, May 2003.
- [112] S. T. Chung and A. Goldsmith, “Degrees of freedom in adaptive modulation: a unified view,” *IEEE Trans. Commun. Technol.*, vol. 49, no. 9, pp. 1561–1571, Sep. 2001.

- [113] J. Lu, K. Letaief, J.-I. Chuang, and M. Liou, "M-PSK and M-QAM BER computation using signal-space concepts," *IEEE Trans. Commun. Technol.*, vol. 47, no. 2, pp. 181–184, Feb. 1999.
- [114] B. Rankov and A. Wittneben, "Spectral Efficient Signaling for Half-duplex Relay Channels," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, 2005.
- [115] J. Ping and S. Ting, "Rate performance of AF two-way relaying in low SNR region," *IEEE Commun. Lett.*, vol. 13, no. 4, pp. 233–235, 2009.
- [116] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16J relay-based wireless access networks: an overview," *IEEE Wireless Commun. Mag.*, vol. 15, no. 5, pp. 56–63, 2008.
- [117] M. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1963 – 1968, Nov. 2004.
- [118] P. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a rayleigh-fading environment," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1416 – 1421, Sep. 2004.
- [119] T. A. Tsiftsis, G. K. Karagiannidis, P. T. Mathiopoulos, and S. A. Kotsopoulos, "Nonregenerative dual-hop cooperative links with selection diversity," *EURASIP J. Wirel. Commun. Netw.*, vol. 2006, no. 2, pp. 34–34, 2006.
- [120] A. Ribeiro, X. Cai, and G. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1264 – 1273, May. 2005.
- [121] T. Wang, A. Cano, G. Giannakis, and J. Laneman, "High-performance cooperative demodulation with decode-and-forward relays," *IEEE Trans. Commun. Technol.*, vol. 55, no. 7, pp. 1427 –1438, Jul. 2007.

- [122] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity. part i. system description,” *IEEE Trans. Commun. Technol.*, vol. 51, no. 11, pp. 1927 – 1938, Nov. 2003.
- [123] —, “User cooperation diversity. part ii. implementation aspects and performance analysis,” *IEEE Trans. Commun. Technol.*, vol. 51, no. 11, pp. 1939 – 1948, Nov. 2003.
- [124] M. D. Selvaraj and R. K. Mallik, “Scaled selection combining based cooperative diversity system with decode and forward relaying,” *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1 –1, 2010.
- [125] K.-S. Hwang, Y.-C. Ko, and M.-S. Alouini, “Outage probability of cooperative diversity systems with opportunistic relaying based on decode-and-forward,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5100 –5107, Dec. 2008.
- [126] D. Chen and J. Laneman, “Modulation and demodulation for cooperative diversity in wireless systems,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1785 –1794, Jul. 2006.
- [127] Y. Lee, M. hung Tsai, and S. lan Sou, “Performance of decode-and-forward cooperative communications with multiple dual-hop relays over Nakagami- m fading channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2853 – 2859, Jun. 2009.
- [128] T. Hunter and A. Nosratinia, “Diversity through coded cooperation,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 283 – 289, Feb. 2006.
- [129] T. Hunter, S. Sanayei, and A. Nosratinia, “Outage analysis of coded cooperation,” *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 375 – 391, Feb. 2006.
- [130] Z. Lin, E. Erkip, and A. Stefanov, “Cooperative regions and partner choice in coded cooperative systems,” *IEEE Trans. Commun. Technol.*, vol. 54, no. 4, pp. 760 – 760, Apr. 2006.

- [131] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 362 – 371, Feb. 2004.
- [132] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun. Technol.*, vol. 52, no. 9, pp. 1470 – 1476, Sep. 2004.
- [133] T. Tsiftsis, G. Karagiannidis, S. Kotsopoulos, and F.-N. Pavlidou, "Ber analysis of collaborative dual-hop wireless transmissions," *IET Electron. Lett.*, vol. 40, no. 11, pp. 679 – 681, May. 2004.
- [134] S. Ikki and M. Ahmed, "Performance Analysis of Cooperative Diversity Wireless Networks over Nakagami- m Fading Channel," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 334 –336, Apr. 2007.
- [135] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415 – 2425, Oct. 2003.
- [136] N. C. Beaulieu and J. Hu, "A closed-form expression for the outage probability of decode-and-forward relaying in dissimilar rayleigh fading channels," *IEEE Commun. Lett.*, vol. 10, no. 12, pp. 813 –815, Dec. 2006.
- [137] Y. Zhao, R. Adve, and T. J. Lim, "Outage probability at arbitrary snr with cooperative diversity," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 700 – 702, Aug. 2005.
- [138] Z. Yi and I.-M. Kim, "Diversity order analysis of the decode-and-forward cooperative networks with relay selection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1792 –1799, May 2008.
- [139] T. Duong, V. N. Q. Bao, and H.-j. Zepernick, "On the performance of selection decode-and-forward relay networks over Nakagami- m fading channels," *IEEE Commun. Lett.*, vol. 13, no. 3, pp. 172 –174, Mar. 2009.

- [140] M. Hasna and M.-S. Alouini, "Outage probability of multihop transmission over nakagami fading channels," *IEEE Commun. Lett.*, vol. 7, no. 5, pp. 216 – 218, May. 2003.
- [141] M. Yuksel and E. Erkip, "Diversity in relaying protocols with amplify and forward," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*. IEEE, 2004, vol.4, pp. 2025–2029.
- [142] G. Karagiannidis, T. Tsiftsis, and R. Mallik, "Bounds for multihop relayed communications in nakagami-m fading," *IEEE Trans. Commun. Technol.*, vol. 54, no. 1, pp. 18 – 22, Jan. 2006.
- [143] G. Farhadi and N. Beaulieu, "Capacity of amplify-and-forward multi-hop relaying systems under adaptive transmission," *IEEE Trans. Commun. Technol.*, vol. 58, no. 3, pp. 758 –763, Mar. 2010.
- [144] —, "Fixed relaying versus selective relaying in multi-hop diversity transmission systems," *IEEE Trans. Commun. Technol.*, vol. 58, no. 3, pp. 956 –965, Mar. 2010.
- [145] W. Zeng, C. Xiao, Y. Wang, and J. Lu, "Opportunistic cooperation for multi-antenna multi-relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3189 –3199, Oct. 2010.
- [146] C. Conne and I.-M. Kim, "Outage probability of multi-hop amplify-and-forward relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1139 –1149, Mar. 2010.
- [147] Y. Zhao, R. Adve, and T. J. Lim, "Symbol error rate of selection amplify-and-forward relay systems," *IEEE Commun. Lett.*, vol. 10, no. 11, pp. 757 –759, Nov. 2006.
- [148] A. M. H. Ikki, Salama S, "Performance of multiple-relay cooperative diversity

- systems with best relay selection over rayleigh fading channels,” *EURASIP J. on Adv. in Signal Process.*, vol. 2008, no. 2, 2008.
- [149] E. Beres and R. Adve, “Selection cooperation in multi-source cooperative networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 118–127, Jan. 2008.
- [150] —, “Optimal relay-subset selection and time-allocation in decode-and-forward cooperative networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2145–2155, Jul. 2010.
- [151] —, “Outage probability of selection cooperation in the low to medium snr regime,” *IEEE Commun. Lett.*, vol. 11, no. 7, pp. 589–597, Jul. 2007.
- [152] D. Michalopoulos and G. Karagiannidis, “Performance analysis of single relay selection in rayleigh fading,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3718–3724, Oct. 2008.
- [153] S. Ikki and M. Ahmed, “Performance analysis of adaptive decode-and-forward cooperative diversity networks with best-relay selection,” *IEEE Trans. Commun. Technol.*, vol. 58, no. 1, pp. 68–72, Jan. 2010.
- [154] F. Xu, F. Lau, Q. Zhou, and D.-W. Yue, “Outage performance of cooperative communication systems using opportunistic relaying and selection combining receiver,” *IEEE Signal Process. Lett.*, vol. 16, no. 2, pp. 113–116, Feb. 2009.
- [155] D. da Costa and S. Aissa, “Performance analysis of relay selection techniques with clustered fixed-gain relays,” *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 201–204, Feb. 2010.
- [156] I. Krikidis, J. Thompson, S. McLaughlin, and N. Goertz, “Amplify-and-forward with partial relay selection,” *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 235–237, Apr. 2008.

- [157] H. Suraweera, D. Michalopoulos, and G. Karagiannidis, “Semi-blind amplify-and-forward with partial relay selection,” *IET Electron. Lett.*, vol. 45, no. 6, pp. 317–319, Mar. 2009.
- [158] D. da Costa and S. Aissa, “Performance analysis of relay selection techniques with clustered fixed-gain relays,” *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 201–204, Feb. 2010.
- [159] —, “Cooperative dual-hop relaying systems with beamforming over Nakagami-m fading channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3950–3954, Aug. 2009.
- [160] H. Bolcskei, R. Nabar, O. Oyman, and A. Paulraj, “Capacity scaling laws in MIMO relay networks,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.
- [161] Y. Fan and J. Thompson, “MIMO Configurations for Relay Channels: Theory and Practice,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1774–1786, May. 2007.
- [162] M. Yuksel and E. Erkip, “Multiple-antenna cooperative wireless systems: A diversity;multiplexing tradeoff perspective,” *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3371–3393, Oct. 2007.
- [163] H. Adinoyi and H. Yanikomeroglu, “Cooperative relaying in multi-antenna fixed relay networks,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 533–544, Feb. 2007.
- [164] Y. Zhao, R. Adve, and T. J. Lim, “Symbol error rate of selection amplify-and-forward relay systems,” *IEEE Commun. Lett.*, vol. 10, no. 11, pp. 757–759, Nov. 2006.
- [165] D. da Costa and S. Aissa, “End-to-end performance of dual-hop semi-blind

- relaying systems with partial relay selection,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4306 –4315, Aug. 2009.
- [166] R. Pabst *et al.*, “Relay-based deployment concepts for wireless and mobile broadband radio,” *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80 – 89, Sept. 2004.
- [167] H. Min, S. Lee, K. Kwak, and D. Hong, “Effect of multiple antennas at the source on outage probability for amplify-and-forward relaying systems,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 633 –637, Feb. 2009.
- [168] A. Papoulis and S. Pillai, *Probability, Random variables and Stochastic Processes*, 4th ed., 2002.
- [169] I. Wolfram Research, <http://functions.wolfram.com/07.45.03.0009.01> .
- [170] R. Matzner and F. Englberger, “An SNR estimation algorithm using fourth-order moments,” in *Proc. IEEE Int. Symp. on Inf. Theory*, 1994, p. 119.
- [171] J.-B. Kim and D. Kim, “Comparison of tightly power-constrained performances for opportunistic amplify-and-forward relaying with partial or full channel information,” *IEEE Commun. Lett.*, vol. 13, no. 2, pp. 100 –102, Feb. 2009.
- [172] P. Anghel and M. Kaveh, “Exact symbol error probability of a cooperative network in a rayleigh-fading environment,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1416 – 1421, 2004.
- [173] B. Maham and A. Hjørungnes, “Performance analysis of repetition-based cooperative networks with partial statistical CSI at relays,” *IEEE Commun. Lett.*, vol. 12, no. 11, pp. 828 –830, Nov. 2008.
- [174] Z. Wang and G. Giannakis, “A simple and general parameterization quantifying performance in fading channels,” *IEEE Trans. Commun. Technol.*, vol. 51, no. 8, pp. 1389 – 1398, Aug. 2003.

- [175] H. Suraweera, M. Soysa, C. Tellambura, and H. Garg, "Performance analysis of partial relay selection with feedback delay," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 531–534, Jun. 2010.
- [176] H. Suraweera, T. Tsiftsis, G. Karagiannidis, and M. Faulkner, "Effect of feedback delay on downlink amplify-and-forward relaying with beamforming," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2009.
- [177] Y. Ma, D. Zhang, A. Leith, and Z. Wang, "Error performance of transmit beamforming with delayed and limited feedback," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1164–1170, 2009.
- [178] C. Conne *et al.*, "SER Analysis and PDF Derivation for Multi-Hop Amplify-and-Forward Relay Systems," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2413–2424, Aug. 2010.
- [179] V. Asghari, A. Maaref, and S. Aissa, "Symbol error probability analysis for multihop relaying over nakagami fading channels," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, april 2010, pp. 1–6.
- [180] R. Louie, Y. Li, and B. Vucetic, "Performance analysis of beamforming in two hop amplify and forward relay networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2008, pp. 4311–4315.
- [181] G. Fedele, "N-branch diversity reception of many DPSK signals in slow and nonselective Nakagami fading," *Eur. Trans. on Telecommun.*, vol. 7, no. 2, Apr. 1996.
- [182] J. Dowle, S. Kuo, K. Mehrotra, and I. McLoughlin, "An FPGA-based MIMO and space-time processing platform," *EURASIP J. App. Signal Proc.*, vol. 2006, p. 137, 2006.
- [183] T. Kaiser, *Smart antennas: state of the art*. Hindawi Pub. Corp., 2005.

- [184] M. Rupp, S. Caban, and C. Mehlführer, “Challenges in building MIMO testbeds,” in *Proc. 13th Eur. Signal Process. Conf.*, 2007.
- [185] Q. Wang, J. Eilert, and D. Liu, “Cost Analysis of Channel Estimation in MIMO-OFDM for Software Defined Radio,” in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*. IEEE, 2008, pp. 935–939.
- [186] T. Haustein *et al.*, “Real-time signal processing for multi-antenna systems: algorithms, optimization, and implementation on an experimental test-bed,” *EURASIP J. App. Signal Proc.*, vol. 2006, p. 136, 2006.
- [187] S. Caban *et al.*, “Vienna MIMO testbed,” *EURASIP J. App. Signal Proc.*, vol. 2006, p. 142, 2006.
- [188] A. Baghaie, S. Kuo, and I. McLoughlin, “FPGA implementation of space-time block coding systems,” in *Proc. IEEE Circuits and Systems Symp.on Emerging Technologies*, vol. 2. IEEE, 2004, vol.2, pp. 591–594.
- [189] D. Cabric, I. O’Donnell, M. Chen, and R. Brodersen, “Spectrum sharing radios,” *IEEE Circuits Syst. Mag.*, vol. 6, no. 2, pp. 30–45, 2006.
- [190] M. Takai *et al.*, “Scalable testbed for next-generation wireless networking technologies,” in *Int. Conf. on Testbeds and Research Infrastructures for the Development of Networks and Communities, (TRIDENTCOM)*, 2005, pp. 162 – 171.
- [191] P. Murphy *et al.*, “VALID: Custom ASIC verification and FPGA education platform,” in *Proc. of IEEE Int. Conf. on Microelectronic Systems Edu.*, 2003, pp. 64–65.
- [192] W. Xiang *et al.*, “Implementation and experimental results of a three-transmitter three-receiver OFDM/BLAST testbed,” *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 88–95, 2004.