Wireless Transmission Protocols Using Relays For Broadcast and Information Exchange Channels

by

Chee Yen Leow

A thesis submitted in fulfilment of requirements for the degree of

Doctor of Philosophy of Imperial College London

Communications and Signal Processing Group

Department of Electrical and Electronic Engineering

Imperial College London

2011

Abstract

Relays have been used to overcome existing network performance bottlenecks in meeting the growing demand for large bandwidth and high quality of service (QoS) in wireless networks. This thesis proposes several wireless transmission protocols using relays in practical multi-user broadcast and information exchange channels. The main theme is to demonstrate that efficient use of relays provides an additional dimension to improve reliability, throughput, power efficiency and secrecy. First, a spectrally efficient cooperative transmission protocol is proposed for the multiple-input and single-output (MISO) broadcast channel to improve the reliability of wireless transmission. The proposed protocol mitigates co-channel interference and provides another dimension to improve the diversity gain. Analytical and simulation results show that outage probability and the diversity and multiplexing tradeoff of the proposed cooperative protocol outperforms the non-cooperative scheme. Second, a two-way relaying protocol is proposed for the multi-pair, two-way relaying channel to improve the throughput and reliability. The proposed protocol enables both the users and the relay to participate in interference cancellation. Several beamforming schemes are proposed for the multi-antenna relay. Analytical and simulation results reveal that the proposed protocol delivers significant improvements in ergodic capacity, outage probability and the diversity and multiplexing tradeoff if compared to existing schemes. Third, a joint beamforming and power management scheme is proposed for multiple-input and multiple-output (MIMO) two-way relaying channel to improve the sum-rate. Network power allocation and power control optimisation problems are formulated and solved using convex optimisation techniques. Simulation results verify that the proposed scheme delivers better sum-rate or consumes lower power when compared to existing schemes. Fourth, two-way secrecy schemes which combine one-time pad and wiretap coding are proposed for the scalar broadcast channel to improve secrecy rate. The proposed schemes utilise the channel reciprocity and employ relays to forward secret messages. Analytical and simulation results reveal that the proposed schemes are able to achieve positive secrecy rates even when the number of users is large. All of these new wireless transmission protocols help to realise better throughput, reliability, power efficiency and secrecy for wireless broadcast and information exchange channels through the efficient use of relays.

To Granny, Dad, Mum, Pei Ling, Li Ling,

Chin Han, Chee Hau, Chee Yong, and Chee Song.

Acknowledgement

First of all, I would like to express my deepest gratitude to my supervisor Prof. Kin K. Leung for his valuable guidance and unparalleled support throughout the journey of my Ph.D. studies. His dedication to work sets as a motivating example for me in completing my studies.

I would like to thank Dr. Zhiguo Ding of Newcastle University, whom I work closely with throughout my research. Dr. Ding is so generous to share his knowledge and experience with me through weekly Skype discussion. He is always there to resolve my technical questions and teaches me to be a more patient person. I would also like to thank Prof. Dennis L. Goeckel of University of Massachusetts (UMASS) for sharing his expertise in the collaboration work between Imperial College and UMASS.

The support from my family and friends has always been a driving force for me to complete my studies. I would like to thank my parents for their encouragement and blessing. I would also like to thank my girlfriend, Pei Ling for her tender loving care. Thanks to my sister, Chin Han for spending wonderful time with me in the UK during her studies. My appreciation also goes to my friends and colleagues in the EEE Department especially Zhengguo, Yun, George, Thanos, Bukhari and Chin Kim. Special thanks to my friends at the Victoria League student house especially Jessica, Siew Kee, Erh Jhen, Kent and Shou Kang who spent time cooking and hanging around with me after college hours. Thanks to the wardens Sue and Trudi for taking care of the house.

Last but not least, I would like to thank my financial sponsors, Universiti Teknologi Malaysia and Ministry of Higher Education Malaysia, for funding my studies. I would like to thank Prof. Leung for providing financial support for me to attend conferences and academic visits through the International Technology Alliance (ITA) project, which is jointly sponsored by the US Army Research Lab and the UK Ministry of Defence.

Statement of Originality

I hereby declare that the contents of this thesis are the results of my original works. Part of the contents of this thesis has been published or submitted for publication as conference or journal papers. The list of publications is as follows.

Journal Papers

1. C. Y. Leow, Z. Ding, K. K. Leung, and D. L. Goeckel, "On the study of analogue network coding for multi-pair, bidirectional relay channels," *IEEE Trans. Wireless Communications*, vol.10, no.2, pp.670-681, Feb. 2011.

2. C. Y. Leow, Z. Ding, and K. K. Leung, "Linear precoded cooperative transmission protocol for wireless broadcast channels," *IEEE Trans. Vehicular Technology*, vol.60, no.7, pp. 3509-3515, Sep. 2011.

3. C. Y. Leow, Z. Ding, and K. K. Leung, "Joint beamforming and power management for MIMO two-way relaying channels," *IEEE Trans. Vehicular Technology*, (accepted).

Conference Papers

1. C. Y. Leow, Z. Ding, and K. K. Leung, "Linear precoded cooperative transmission protocol for wireless broadcast channels," *in IEEE International Symposium on Information Theory (ISIT 2009)*, June 28 -July 3, 2009, pp. 1684-1688.

2. C. Y. Leow, Z. Ding, and K. K. Leung, "Wireless broadcast channel cooperative transmission protocol using linear precoding", *in the 3rd Annual Conference of the International Technology Alliance (ACITA 2009)*, Sep. 22-24, 2009, Maryland, USA.

3. C. Y. Leow, Z. Ding, K. K. Leung, and D. L. Goeckel, "Analogue network coding for multi-pair, bidirectional relay channels," *in the Sensor Signal Processing for Defence (SSPD 2010)*,

Sep. 29-30, 2010, London, UK.

4. C. Y. Leow, Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "On the study of analogue network coding for multi-pair, bidirectional relay channels," *in the 4th Annual Conference of the International Technology Alliance (ACITA 2010)*, Sep. 15-16, 2010, London, UK.

5. C. Y. Leow, Z. Ding, and K. K. Leung, "Joint beamforming and power allocation for MIMO two-way relaying channels," *in IEEE ICC 2011 Workshop on Planning and Optimization of Wireless Communication Networks*, June 5-9, 2011, pp. 1-5.

6. C. Y. Leow, D. L. Goeckel, and K. K. Leung, "A two-way secrecy scheme for the scalar broadcast channel with internal eavesdroppers, " *in the 45th Annual Asilomar Conference on Signals, Systems, and Computers,* Nov. 6-9, 2011, California, USA, (accepted).

7. C. Y. Leow, D. L. Goeckel, and K. K. Leung, "Two-way secrecy schemes for the broadcast channel with internal eavesdroppers, "*in the 5th Annual Conference of the International Technology Alliance (ACITA 2011)*, Sep. 27-28, 2011, Maryland, USA.

Contents

Abstract i

Acknowledgement iii

Statement of Originality v

Contents vii

List of Figures xiii

List of Tables xvii

Nomenclature xix

List of Acronyms xix

List of Mathematical Symbols xxi

1 Introduction 1

1.1 Motivations 2

1.2 Aim and Objectives 4

1.3 Contributions and Outlines of the Thesis 5

2 Background and Related Literature 7

2.1 The Wireless Channels 7

2.1.1 The Multipath Fading 8

2.1.2 Channel Models 8

2.1.3 Discrete-Time Baseband Model 9

2.1.4 CSI Acquisition and CSI Model 10

2.1.5 Diversity Techniques to Improve Reliability 11

2.1.6 Multi-User Diversity and Opportunistic Communication 12
2.1.7 Capacity of Wireless Channels 12
2.2 MIMO Communication 14
2.2.1 Spatial Multiplexing and Channel Decomposition 14
2.2.2 Linear Precoding and Beamforming 16
2.2.3 Number of Antennas and Capacity 17
2.2.4 SNR and Capacity 18
2.2.5 Diversity and Multiplexing Tradeoff 19
2.3 Cooperative Communication 22
2.4 Two-Way Relaying 24
2.5 Secrecy Communication 26

3 Cooperative Transmission Protocol with Linear Precoding for the MISO Broadcast Channel 29

3.1 Chapter Introduction 30
3.2 Description of the Proposed Protocol 33
3.2.1 Initialisation 34
3.2.2 Cooperative Transmission 34
3.2.2.1 First Time Slot 36
3.2.2.2 Second Time Slot 38
3.2.2.3 Remaining Time Slots 39
3.3 Analysis of Outage Probability and the Diversity and Multiplexing Tradeoff 40
3.4 Numerical Results 47
3.5 Chapter Conclusion 52

4 Analogue Network Coding For the Multi-Pair, Two-Way Relaying Channel 55

4.1 Chapter Introduction 56
4.2 System Model and Protocol Description 58
4.3 Joint Receive and Transmit Beamforming Design 60
4.3.1 Case I: $N = 2M - 1$ 60
4.3.2 Case II: $N \geq 2M$ 63

4.3.2.1 Null-Space Vector Selection . . . 63
4.3.2.2 Coherent Combining of Null-Space Vectors . . . 64
4.3.3 Comparable Schemes . . . 66
4.3.3.1 Pure AF Scheme . . . 66
4.3.3.2 MRR-MRT Scheme . . . 66
4.3.3.3 Zero-Forcing Scheme . . . 67
4.3.3.4 Block-Diagonalisation Based Schemes . . . 68
4.4 Ergodic Capacity Analysis . . . 70
4.5 Outage Probability and the Diversity and Multiplexing Tradeoff Analysis . . . 73
4.6 Numerical Results . . . 76
4.7 Chapter Conclusion . . . 82

5 Joint Beamforming and Power Management for the MIMO Two-Way Relaying Channel 85
5.1 Chapter Introduction . . . 86
5.2 System Model and Protocol Description . . . 89
5.2.1 Initialisation . . . 90
5.2.2 Transmission Protocol . . . 91
5.3 Beamforming Design . . . 92
5.3.1 Design of $\mathbf{F}_i$ . . . 92
5.3.2 Design of $\mathbf{W}$ . . . 93
5.3.3 Design of $\mathbf{G}_i$ . . . 94
5.4 Joint Power Allocation . . . 94
5.4.1 Subchannel SNR Derivation . . . 94
5.4.2 Sum-Rate Optimisation . . . 95
5.4.3 Proposed Power Allocation Strategies . . . 97
5.4.3.1 Proposed JPA I . . . 97
5.4.3.2 Proposed JPA II . . . 98
5.4.4 Baseline Schemes and Comparable Scheme . . . 98
5.4.4.1 Baseline Scheme: Pure AF . . . 98
5.4.4.2 Baseline Scheme: MIMO One-Way . . . 99
5.4.4.3 Comparable Scheme: Alternate Optimisation (A-Opt) . . . 100

5.5 Joint Power Control . . . 100
5.5.1 Joint Power Control Optimisation Problem . . . 101
5.5.2 Power Control Strategies and Baseline Scheme . . . 102
5.5.2.1 Joint Power Control (JPC) . . . 102
5.5.2.2 Equal Power Control (EPC) . . . 102
5.5.2.3 Baseline Scheme: MIMO One-Way . . . 103
5.6 Numerical Results . . . 104
5.6.1 Power Allocation with Fixed Total Power Constraint . . . 104
5.6.2 Power Control with Fixed Substream Rate Constraints . . . 108
5.7 Chapter Conclusion . . . 110

6 Two-Way Secrecy Schemes for the Broadcast Channel with Internal Eavesdroppers 113
6.1 Chapter Introduction . . . 114
6.2 System Model and Problem Statement . . . 116
6.3 Two-Way Secrecy Scheme for Opportunistic Broadcast . . . 118
6.3.1 Protocol Description . . . 118
6.3.2 Secrecy Rate Derivation . . . 120
6.3.3 Asymptotic Analysis When $K \to \infty$ . . . 121
6.4 Two-Way Secrecy Scheme for Non-Opportunistic Broadcast . . . 123
6.4.1 Protocol Description . . . 123
6.4.2 Secrecy Rate Derivation . . . 124
6.4.3 Asymptotic Analysis When $K \to \infty$ . . . 126
6.5 Numerical Results . . . 128
6.5.1 Opportunistic Broadcast . . . 128
6.5.2 Non-Opportunistic Broadcast . . . 130
6.6 Chapter Conclusion . . . 132

7 Conclusions and Future Works 133
7.1 Conclusions . . . 133
7.2 Future Work . . . 136

Appendix A **139**

Appendix B **149**

Appendix C **157**

Appendix D **159**

Bibliography **165**

List of Figures

2.1 The SIMO ergodic capacity versus the number of receive antennas M_r at various SNR. 18

2.2 MIMO ergodic capacity versus number of antennas ($M_r = M_t$) at $5 \leq$SNR≤ 30 with perfect CSI at the transmitter and receiver. 19

2.3 Ergodic capacity versus SNR for several antenna configurations 20

2.4 The optimal diversity and multiplexing tradeoff when $M_t = 3$ and $M_r = 4$. 21

2.5 Basic scenario in the cooperative transmission 22

2.6 Information exchange using one-way relaying 25

2.7 Information exchange using two-way relaying 26

3.1 Cooperative broadcast scenario with 2-antenna source, 2-antenna relay and 2 single-antenna users. 35

3.2 Flow chart of a CBC transmission frame of the proposed protocol. 35

3.3 Diversity and multiplexing tradeoff of the non-cooperative ZFBF scheme and the proposed CBC scheme with various number of available relays L. 49

3.4 Probability that K relays out of $L = 4$ available relays are qualified. Fixed parameters: $M = 2$ and $R = 1$ bits/s/Hz. 49

3.5 Outage probability versus SNR of the non-cooperative ZFBF scheme and the proposed CBC scheme at various target data rate R. Fixed parameters: $M = 2$ and the number of participating relays, $K = 1, 2$. 51

3.6 Outage probability versus SNR of the non-cooperative ZFBF scheme and the proposed CBC scheme with various number of participating relays K. Fixed parameters: $M = 2$ and $R = 4$ bits/s/Hz. 51

3.7 Outage probability versus target data rate of the non-cooperative scheme and the proposed CBC scheme, with various number of participating relays K. Fixed parameters: $M = 2$, SNR=30dB. 52

4.1 Scenario when two pairs of single-antenna users exchanging information with the help of a multi-antenna relay. 59
4.2 Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of the proposed protocol in comparison with the capacity bounds. Fixed parameters: $M = 2$ and $N = 3$. 77
4.3 Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of various schemes when $M = 2$ and $N = 3$. 78
4.4 Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of various scheme when $M = 2$ and $N = 4$. 79
4.5 Single-user outage probability versus SNR $(1/\sigma^2)$ when $M = 2$, $N = 4$ and $R = 2$ bits/s/Hz. 80
4.6 Single-user outage probability versus target data rate R for various schemes when $M = 2$, $N = 4$ and SNR=30dB. 81

5.1 Example of two-way relaying scenario where each node is equipped with $M = 2$ antennas. The symbols above the arrows represent the channel matrices while the directions of the arrows indicate the directions of data flows. 90
5.2 Ergodic sum-rate versus SNR $(\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2})$ for fixed $M = 4$, $P = 3$. 105
5.3 Ergodic sum-rate versus SNR $(\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2})$ for fixed $M = 4$, $P = 3$, and $\frac{1}{\sigma_r^2} = 30$dB. . 106
5.4 Ergodic sum-rate versus SNR $(\frac{1}{\sigma_r^2})$ for fixed $M = 4$, $P = 3$, and $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 5$dB. . . 107
5.5 Ergodic sum-rate versus number of antennas M, for fixed $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 15$dB, $\frac{1}{\sigma_r^2} = 30$ dB, and $P = 3$. 108
5.6 Ergodic sum-rate versus the relay location d_r for fixed $M = 2$, $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 30$ dB, $P = 3$, $\alpha = 4$. 109
5.7 Average total power consumption versus SNR $(\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2})$ when R =2 bits/s/Hz and $M = 2$. 110
5.8 Average total power consumption versus target rate R when reference SNR, $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 20$dB and $M = 2$. 111

6.1 Data flows of the proposed two-way secrecy scheme for opportunistic broadcast. . . 119
6.2 Data flows in the first phase of the proposed two-way secrecy scheme for non-opportunistic broadcast. 124

6.3 Data flows in the second phase of the proposed two-way secrecy scheme for non-opportunistic broadcast. 125
6.4 Ergodic secrecy capacity versus number of users K, when SNR=10dB. 128
6.5 Ergodic secrecy capacity versus SNR when K=10, 30, and 60. K increases in the direction of the arrow. 129
6.6 Ergodic secrecy capacity versus number of users K, when SNR=30dB. 130
6.7 Ergodic secrecy capacity versus SNR when K=15, 30, and 60. K increases in the direction of the arrow. 131

7.1 Outage probability versus SNR $(1/\sigma^2)$ under two different selection criteria in comparison with the baseline zero-forcing scheme. The fixed parameters are $M = 2$, $N = 4$ and $R = 2$ bits/s/Hz. 154

List of Tables

2.1 Classification of multipath fading 9

7.1 Complexity of key operations when $N = 2M - 1$. 157

7.2 Complexity of key operations of the null-space vector selection scheme when $N \geq 2M$. 158

7.3 Complexity of key operations of the coherent combining scheme when $N \geq 2M$. . . 158

Nomenclature

List of Acronyms

AF	amplify-and-forward
A-Opt	alternate optimisation
BD-SVD	block-diagonalisation with singular value decomposition
CBC	cooperative broadcast channel
CDF	cumulative density function
CDMA	code division multiple access
CF	compress-and-forward
CSI	channel state information
DDF	dynamic decode-and-forward
DF	decode-and-forward
DPC	dirty paper coding
EF	estimate-and-forward
EPC	equal power control
FDD	frequency division duplex
flops	floating point operations
i.i.d.	independent and identically distributed
JPA	joint power allocation
JPC	joint power control

MIMO	multiple-input and multiple-output
MISO	multiple-input and single-output
ML	maximum likelihood
MMSE	minimum mean-squared-error
MRR-MRT	maximal ratio reception and transmission
MSE	mean-squared-error
NAF	non-orthogonal amplify-and-forward
NC-SM	network coding with spatial multiplexing
OFDM	orthogonal frequency division multiplexing
PA-MF	pair-aware matched filter
PA-SDR	pair-aware with semi-definite relaxation
PDF	probability density function
QoS	quality of service
RHS	right hand side
SDMA	space division multiple access
SIMO	single-input and multiple-output
SINR	signal-to-interference-and-noise ratio
SISO	single-input and single-output
SNR	signal-to-noise ratio
SVD	singular value decomposition
TDD	time division duplex
X-OR	exclusive-or operation
ZFBF	zero-forcing beamforming

List of Mathematical Symbols

u	lower case letters denote scalars
$\mathbf{u}$	boldface lower case letters denote vectors. All vectors are column vectors by default.
$\mathbf{U}$	boldface upper case letters denote matrices
$[\mathbf{U}]_{m,m}$	the mth diagonal element of matrix $\mathbf{U}$
det($\mathbf{U}$)	determinant of matrix $\mathbf{U}$
rank($\mathbf{U}$)	rank of matrix $\mathbf{U}$
trace($\mathbf{U}$)	trace of matrix $\mathbf{U}$
null($\mathbf{U}$)	null space of matrix $\mathbf{U}$
diag($u_1, u_2, ...$)	diagonal matrix with elements $u_1, u_2, ...$
diag($\mathbf{U}_1, \mathbf{U}_2, ...$)	block-diagonal matrix with elements $\mathbf{U}_1, \mathbf{U}_2, ...$
$[.]^T$	transpose operation
$[.]^H$	Hermitian transpose operation
$[.]^*$	complex conjugate operation
$[.]^{-1}$	inverse opeartion
$[.]^\dagger$	pseudoinverse operation
$\|.\|$	absolute value
$\|\|.\|\|$	Euclidean norm
$\|\|.\|\|_F$	Frobenius norm
$\otimes$	Kronecker product
$\oplus$	exclusive-or operation
$\exp(.)$	exponent

$\log_a(.)$	logarithm with base a
$b!$	factorial for integer b
$(x)^+$	$\max\{0, x\}$
$E[.]$	statistical expected value
$P(A)$	probability of event A
$\mathcal{CN}(\mu, \sigma^2)$	complex Gaussian distribution with mean μ and variance σ^2
$\sim$	is distributed as
$\doteq$	special symbol to denote exponential equality, i.e. $f(\gamma_0) \doteq \gamma_0^n$ is used to denote $\lim_{\gamma_0 \to \infty} \frac{\log_2 f(\gamma_0)}{\log_2 \gamma_0} = n$.
$\triangleq$	equal by definition
$x \to \infty$	the value of variable x approaches positive infinity
$\{\ ,\ \}$	the set of
$\in$	is an element of
$\forall$	for all
$X \bigcup Y$	union of set X and set Y.
$X \bigcap Y$	intersection of set X and set Y.

Chapter 1

Introduction

The ever-increasing demand for mobile services range from basic voice communication to data intensive multimedia applications has stimulated countless research in the field of wireless communications. On-going research in the field is targeting to improve throughput, reliability and spectral efficiency and reduce power consumption. A multitude of techniques and solutions have been proposed to achieve such goals. One of the promising solutions is to use relays to assist wireless communication. Relay assisted communication such as cooperative communication, two-way relaying, etc, has demonstrated the advantages of employing relays in both centralised and decentralised networks.

Cooperative communication has drawn much attention in recent years. Cooperative communication invites distributed users to assist each other in order to improve reliability and throughput of wireless transmission. Due to the broadcast nature of wireless transmission, users can help to relay each other's messages which they observed through the wireless channel. Relaying, which is the key ingredient in cooperative communication, provide an additional spatial dimension to enhance the robustness and throughput of wireless transmission. The promising benefits of cooperative communication have attracted a variety of studies to investigate its potential applications in multi-user network under practical scenarios and to gain understanding of the fundamental performance limits.

The half duplex constraint where a communication node is not able to transmit and receive simultaneously in the same channel raises concern on the spectral efficiency of relaying. Particularly in a two-way information exchange scenario, conventional one-way relaying requires double the channel resources if compared to direct communication without relay. Inspired by the idea of network coding, two-way relaying has been proposed to enable more spectrally efficient relaying. Two-way relaying is

able to halve the channel resources used in conventional one-way relaying. The throughput improvement introduced by two-way relaying has motivated its application in practical multiple user pairs and multiple antennas scenarios.

On the other hand, in many wireless applications the information transmission from one point to another has to remain confidential to third party. However, due to the broadcast nature of wireless transmission, the information transmission is prone to eavesdropping by a malicious party. This introduces the study of secrecy communication to ensure that information is protected from any potential eavesdropper. Physical layer security which utilises the fluctuations of the wireless channel to achieve secrecy has been proposed to complement the application layer cryptography.

This chapter begins with the presentation of the motivations of this thesis in Section 1.1. The aim and objectives of this thesis are presented in Section 1.2. The contributions and outlines of this thesis are highlighted in Section 1.3.

1.1 Motivations

The demand for wireless services has grown tremendously in recent years. Existing wireless communication networks are evolving to meet the growing demand for large bandwidths and high quality of service (QoS). The application of relays have been proposed in various communication scenarios such as broadcast channels and information exchange channels to provide additional room to overcome bottlenecks in reliability, throughput, power efficiency or even in security.

By enabling mobile users to relay each others' messages, cooperative communication has demonstrated diversity gain improvement [1, 2]. The diversity gain helps to mitigate the adverse effect of channel fluctuations in fading environment. As a result, the QoS for delay-limited applications such as video conferencing, remote surveillance, etc, is improved. The promising benefit of cooperative communication has encouraged its application in real-world scenarios such as the multi-user broadcast channel. The broadcast channel is one of the key building blocks of wireless communications. In the broadcast channel, the source broadcasts information to multiple destinations. Initial work on the cooperative broadcast channel focuses on single antenna configuration. The dynamic decode-and-forward (DDF) relaying strategy is proposed in [3] for the broadcast channel. However, multi-antenna spatial multiplexing, i.e. simultaneous transmission of multiple messages in the same channel, has

not been considered. To bridge the gap, a cooperative broadcast channel supporting spatial multiplexing is studied in Chapter 3 of this thesis. The problem of multi-user co-channel interference is also addressed in the same chapter.

Meanwhile in the information exchange channel, two-way relaying has been proposed to address the spectral efficiency issue due to the half duplex constraint. By exploiting the broadcast nature of wireless transmission to mix multiple data streams, two-way relaying is able to reduce the channel resources used in two-way information exchange [4, 5, 6]. Two-way relaying has been studied in a practical scenario where there are multiple pairs of users that wish to exchange information with their partners. Simultaneous transmission of multiple pairs introduces the problem of co-channel interference, which is addressed by [7] in a wideband system and by [8, 9, 10] in a narrowband system. The linear precoding or transmit beamforming[1] method known as block-diagonalisation [11] is used at the multi-antenna relay in [8, 9, 10] to mitigate the interference caused by multi-pair transmission. Nevertheless, the potential benefit of block-diagonalisation in the scenario where the relay does not have enough degrees-of-freedom to spatially separate and/or decode each independent message is not covered in [8, 9, 10]. Furthermore, the diversity gain offered by block-diagonalisation in the multi-pair scenario has not been quantified. All these issues in the multi-pair, two-way relaying channel are addressed in Chapter 4 of this thesis.

Attracted by the benefits of multiple antennas in enhancing the system capacity and reliability, two-way relaying has been generalised to multi-antenna scenarios. Multi-antenna at the user and relay nodes enables the use of beamforming and power allocation to further optimise the performance. The sum-rate optimisation problem with an individual power constraint is considered in [12, 13, 14, 15]. Nonetheless, joint power allocation at all users and relays, subject to a total network power constraint is still an open problem. Although joint power allocation problem has been studied in one-way relaying channel [16, 17], the solutions are not applicable to the two-way relaying scenario due to the fundamental difference in the transmission protocol. This motivates the study of joint beamforming and power allocation in multiple-input and multiple-output (MIMO) two-way relaying channel in Chapter 5 of this thesis.

In a network with a large number of users, the probability to have one user with very good channel is high due to the multi-user diversity effect [18]. Opportunistic communication where users with

[1] Beamforming is a linear processing technique used either at the transmitter or receiver to adjust the complex weighting factors of each antenna element.

good channels are scheduled for transmission or reception, has been proposed to harvest the benefit of multi-user diversity. For instance, in single-antenna (or scalar) broadcast channels, the source transmits to the user with the largest channel gain in each time slot. Such opportunistic scheme exploiting multi-user diversity is able to achieve the optimal sum capacity of the scalar broadcast channel [18]. However, from the perspective of secrecy communication, the multi-user diversity effect can be detrimental. [19] investigates the effect of multi-user diversity to the secrecy capacity of the scalar broadcast channel adopting opportunistic transmission and Wyner's wiretap coding. It is revealed that not only the legitimate user would benefit from the multi-user diversity, but also the eavesdroppers (or idle users). In particular, when the number of users goes to infinity, the channel gain of the best eavesdropper is almost as good as the channel gain of the legitimate user. This renders the average secrecy capacity to zero. The effect of the multi-user diversity to the secrecy of the scalar broadcast channel is further examined in Chapter 6 of this thesis. The use of channel reciprocity, multi-user diversity and relays to improve the secrecy capacity in large network is studied.

1.2 Aim and Objectives

This thesis aims to demonstrate the benefits of employing relays in enhancing reliability, increasing data rate, improving power efficiency and even strengthening secrecy of wireless transmissions in practical multi-user networks such as broadcast channels and two-way information exchange channels. Focusing on this aim, several objectives are laid out.

The first objective is to propose a spectrally efficient cooperative transmission protocol for the multiple-input and single-output (MISO) broadcast channel to improve reliability. Multi-user co-channel interference is addressed. The performance of the proposed protocol is evaluated using information theoretic metrics, such as outage probability and the diversity and multiplexing tradeoff.

The second objective is to propose a spectrally efficient protocol and beamforming design to improve the data rate and reliability of multi-pair, two-way relaying channel with a multi-antenna relay and multiple pairs of single-antenna users. The performance of the protocol is assessed using information theoretic metrics such as ergodic capacity, outage probability and the diversity and multiplexing tradeoff.

The third objective is to propose a joint beamforming and power management scheme to improve

the sum-rate and power efficiency of the MIMO two-way relaying channel. Power allocation and power control problems are addressed. The performance in terms of sum-rate and total transmission power are evaluated using numerical simulation.

The fourth objective is to propose a two-way secrecy protocol for the scalar broadcast channel, under both opportunistic and non-opportunistic transmissions. The secrecy rate of the proposed scheme in a large network is assessed analytically and numerically.

1.3 Contributions and Outlines of the Thesis

The original contributions of this thesis can be summarised as follows,

- A spectrally efficient cooperative transmission protocol using linear precoding for MISO broadcast channel is proposed in Chapter 3. The proposed protocol is able to mitigate the co-channel interference among multiple destinations. Analytical and simulation results for outage probability and the diversity and multiplexing tradeoff are provided to justify the performance of the proposed protocol. Both the analytical and simulation results prove that the protocol can achieve the maximum diversity gain expressed as a sum of the number of source transmitter antennas and the number of available relays. The derived diversity and multiplexing tradeoff and the outage probability simulations show that the proposed protocol outperforms the non-cooperative scheme.

- A spectrally efficient two-way relaying protocol for the multi-pair, two-way relaying channel is proposed in Chapter 4. The protocol integrates the idea of analogue network coding in mixing two data streams originating from the same user pair, together with the spatial multiplexing of the data streams originating from different user pairs. Several beamforming schemes for the multi-antenna relay are proposed. Analytical and simulation results on ergodic capacity, outage probability and the diversity and multiplexing tradeoff are provided. Analytical and simulation results justify that the ergodic capacity, outage probability and the diversity and multiplexing tradeoff of the proposed scheme outperform comparable techniques.

- A joint beamforming and power management scheme for the MIMO two-way relaying channel is proposed in Chapter 5. Two power management issues, i.e. power allocation and power

control, are addressed. The non-convex sum-rate optimisation problem is approximated with a convex objective to enable the problem to be solved effectively using convex optimisation [20]. The power control problem is formulated as a geometric program which can also be solved efficiently using convex optimisation. Simulation results justify that the proposed scheme delivers better sum-rate performance or consumes lower transmission power, if compared to existing methods.

- Two-way secrecy schemes for the scalar broadcast channel are proposed in Chapter 6. Channel reciprocity is utilised and relays are employed to forward secret messages. Multi-user diversity is leveraged to improve the secrecy. Asymptotic results on the secrecy rate under both opportunistic and non-opportunistic transmissions are derived to justify the performance of the proposed schemes. Both analytical and simulation results reveal that positive secrecy rates can be obtained even when the number of users is large.

Overall, this thesis proposes new wireless transmission protocols which combine the use of relays, multi-antenna beamforming, power management and multi-user diversity. The proposed protocols are able to deliver significant advantages in terms of throughput, reliability, power efficiency and secrecy to the wireless broadcast and information exchange channels.

The remaining of this thesis is organised as follows. Chapter 2 gives an overview on the background of the areas of interest in this thesis. Chapter 3 presents the proposed cooperative transmission protocol for the MISO broadcast channel. In Chapter 4, the proposed two-way relaying protocol and beamforming schemes for the multi-pair, two-way relaying channel are elaborated. The proposed joint beamforming and power management scheme for the MIMO two-way relaying channel is studied in Chapter 5. Chapter 6 describes the proposed two-way secrecy schemes for the scalar broadcast channel. Chapter 7 concludes the thesis and suggests several future works related to the topics studied in this thesis.

Chapter 2

Background and Related Literature

This chapter presents some important reviews on the established results in the areas of interest. Section 2.1 covers the background knowledge of the wireless channels. Section 2.2 presents the fundamental of multiple-input and multiple output (MIMO) systems. The concept of relaying in cooperative communication and two-way relaying are discussed in Section 2.3 and 2.4 respectively. Finally, Section 2.5 gives an overview on the secrecy communication.

2.1 The Wireless Channels

Understanding the natural characteristics of the wireless channels is the key to enable better utilisation of the wireless channels. Subsection 2.1.1 describes the multipath fading in the wireless channels and the mechanisms that differentiate various types of fadings. Subsection 2.1.2 reviews the existing channel models while Subsection 2.1.3 describes the mathematical representation of the wireless channels. Subsection 2.1.4 describes how the transceiver obtains the channel state information (CSI) and the CSI models available. The diversity techniques used to improve the reliability of wireless transmission is discussed in Subsection 2.1.5 while the multi-user diversity which exploits the channel fading is covered in Subsection 2.1.6. In the final subsection, the notion of capacity is introduced.

2.1.1 The Multipath Fading

The attenuation and fluctuation of electromagnetic waves is generally known as fading. In wireless communication, the fading consists of large scale and small scale variations. Large scale fading is characterised by path loss and shadowing effect. Path loss is the average signal attenuation as a function of receiver distance from the transmitter while shadowing is the variations about the average signal attenuation due to diffraction. Small scale fading, observed as random and dramatic fluctuations, is attributed to small changes in position of transmitter, receiver or objects in between transmitter and receiver, where signal is refracted, reflected and scattered. As a result, multiple copies of signal undergoing different paths and carrying different amplitude, phase and delay impinge on the receiver. These multipath signals either add constructively or destructively at the receiver, causing severe fluctuations. Thus, small scale fading is also called multipath fading. Multipath fading is characterised by its time dispersive and time varying features [21]. Time dispersive property is defined by the multipath delay spread while the time varying property is due to the motion of the mobile user, known as Doppler shift. According to time dispersive mechanism, the multipath fading can be categorised into i. flat fading and ii. frequency selective fading. On the other hand, time varying mechanism is divided into i.slow fading and ii. fast fading. Table 2.1 summarised the classification of multipath fading [21]. Throughout this thesis, it is assumed that the channel is frequency flat and and slowly varying. The flat fading assumption is relevant not only in narrowband communication but also in wideband communication such as the orthogonal frequency division multiplexing (OFDM) system. The slow fading or quasi-static assumption is important to guarantee that channel state information (CSI) obtained from channel estimation remains valid throughout the whole data transmission period. The benefit of slow fading is that the channel estimation overhead is very small relative to the data transmission period. The small channel estimation overhead enables the system to gain the real benefits from the use of CSI in wireless transmission and reception, which will be demonstrated in Chapter 3 to Chapter 6 in this thesis.

2.1.2 Channel Models

Realistic channel model functions not only as wireless system design guideline, it is also an important tool to simulate and evaluate the performance of wireless systems. There are basically there categories

Mechanism	Category	Properties
Time Dispersive	Flat Fading	1. multipath delay spread<symbol period 2. channel coherence bandwidth>symbol rate
	Frequency Selective Fading	1. multipath delay spread>symbol period 2. channel coherence bandwidth<symbol rate
Time Varying	Slow Fading (quasi static fading, block fading)	1. channel fading rate<symbol rate 2. channel coherence time>symbol period
	Fast Fading	1. channel fading rate>symbol rate 2. channel coherence time<symbol period

Table 2.1: Classification of multipath fading

of channel model, i. the statistical model, ii. ray tracing model, and iii. measurement-based model. Statistical model characterises fading as random processes with certain long term distribution. Ray tracing model is a deterministic model utilising the geometric theory and reflection, diffraction and scattering models. It requires the site-specific details of the terrain and the building plan and thus suffers from a high computational load. In measurement-based model, empirical data collection is needed to construct the distribution of the channel. In this thesis, statistical model is preferred because it facilitates mathematical analysis and covers more general fading enviroments.

There are two major statistical models commonly used to characterise the fading channel in slow fading enviroment, namely Rayleigh distribution and Ricean distribution. When there is no line-of-sight, the multipath fading can be modelled as Rayleigh fading where each element of the channel has distribution $\mathcal{CN}(0, \sigma^2)$ with the corresponding probability density function (PDF) of $\frac{u}{\sigma^2} \exp\left(\frac{-u^2}{2\sigma^2}\right)$ for any $u \geq 0$ [18]. Such model is based on the assumption that there are large number of fading paths with random amplitude and uniform phase. In the presence of line-of-sight, Ricean distribution is used to represent both the large scale and small scale fadings. Throughout this thesis, the Rayleigh fading model is adopted in order to show the effectiveness of the proposed transmission protocols in handling the adverse channel fluctuations created by the multipath fading and also to demonstrate techniques to leverage the multipath fading. Without loss of generality, this thesis assume all the channel elements are i.i.d. with distribution $\mathcal{CN}(0, 1)$.

2.1.3 Discrete-Time Baseband Model

In the design of a wireless communication system, the baseband equivalent representation of the input and output of the communication system is used since most signal processing is done at the

baseband, before the transmit signal is up-converted to carrier frequency or after the received signal is down-converted from carrier frequency. In this thesis, the discrete-time baseband model is used. In a single-antenna point-to-point communication scenario with slowly varying frequency-flat fading, the received signal at the receiver at time instance t can be expressed as

$$y(t) = hx(t) + n(t), \tag{2.1}$$

where h is the (complex) channel coefficient, $x(t)$ is the transmitted information signal at time instance t and $n(t)$ is the circular symmetric white Gaussian noise with distribution $\mathcal{CN}(0, \sigma^2)$. Note that the channel coefficient h is a constant since the channel remains constant during the transmission period (which is less than the channel coherence time).

2.1.4 CSI Acquisition and CSI Model

The knowledge of CSI is required at the transceivers to enable to correct decoding of the information and also to enable functions such as rate control, power control, beamforming, etc. The CSI at the receiver can be estimated easily from pilot signal or training sequence from the transmitter. The CSI at the transmitter can be obtained using two methods, namely the open-loop and the close-loop methods [22].

The open-loop method assumes the channel reciprocity principle. This principle states that the forward channel is equivalent to the backward channel. Channel reciprocity holds when both forward and backward channels share the same frequency band (or frequency bands close enough together). Two conditions need to be met for the reciprocity principle to hold. First, the time delay between forward and backward channels must be much smaller than the channel coherence time. Second, the frequency difference between the two channels must be much smaller than the channel coherence bandwidth. Under channel reciprocity, the transmitter CSI can be measured at the transmitter by using the pilot signal or training sequence from the receiver. This open-loop method is suitable for time division duplex (TDD) system but not the frequency division duplex (FDD) . Throughout this thesis, it is assumed that open-loop method is used at the transmitter to acquire CSI.

On the other hand, under the close-loop method, channel estimation is done at the receiver side and the CSI is fed back to the transmitter. The only condition for this method to work is the feedback

time lag should be much smaller than channel coherence time. This method is suitable for both TDD and FDD systems. However, this method is not practical for system with large number of transmitters because the feedback overhead is linearly proportional to number of transmitters.

CSI model describes the quality of the estimated channel coefficients. Generally, it can be classified into perfect CSI model and imperfect CSI model. The most common CSI model is perfect CSI model. This model assumes the perfect knowledge of the instantaneous CSI at both receiver and transmitter. It simplifies the fading channel into a constant channel at a time instance where an immediate CSI is available. The availability of perfect CSI enables the transceivers to make the best use of the available channel to optimise the data rate and reliability. In this thesis, the perfect CSI model is used throughout Chapter 3 to Chapter 6 to demonstrate the benefits in terms of data rate, reliability, power efficiency and secrecy.

In highly mobile environment (such as fast fading), the instantaneous CSI varies too quickly and it is almost impossible for the transceivers to keep track of the current channel. In such case, imperfect CSI model can be employed. Channel statistics such as channel mean and channel variance can be used to model the CSI [23].

2.1.5 Diversity Techniques to Improve Reliability

Fading which causes fluctuations of radio-frequency signals in time, frequency and space is conventionally viewed as an inherent impairment in wireless channels. Mitigating the adverse effect of this variability in signal quality is essential to enable reliable communication through the wireless channels. Diversity techniques can be used to combat the adverse effect of multipath fading. Diversity techniques introduce redundancy in the wireless transmission by transmitting the same information symbol over multiple independent signal paths [18]. Diversity increases the signal immunity against random fading phenomenon and reduces the probability that a signal path is in deep fade. There are three main categories of diversity techniques, namely time diversity, frequency diversity and spatial diversity.

In time diversity technique, redundancy is introduced through channel coding where different parts of a codeword undergo several channel coherence periods with independent fading. In frequency diversity technique, redundancy is introduced through the frequency selective channel where independent multipath signals impinge on the receivers at different symbol times. The third technique,

spatial diversity introduces redundancy in the spatial domain using multiple transmit and/or receive antennas. The signal emanating from or impinging on different antennas undergo statistically independent fading paths, thus creating spatial diversity. Spatial diversity can also be obtained through the use of relays, which is discussed in Section 2.3. Spatial diversity is a promising solution because it provides another dimension of improvement which does not require additional time and frequency resources. The diversity benefit obtained through the above mentioned techniques is known as the diversity gain, which is characterised by the negative exponent of the SNR term of the error probability. Therefore, the steeper the slope of the error probability curve, the higher the diversity gain is. Physically, the diversity gain corresponds to the number of statistically independent paths that a message passes through before it reaches the receiver.

2.1.6 Multi-User Diversity and Opportunistic Communication

In certain situations, instead of suppressing the effect of multipath fading, one can exploit the fluctuations created by channel fading to improve the performance of wireless transmission. One distinct example can be found in the single-antenna (scalar) wireless broadcast channel with many users. Since each user undergoes independent fading, there is a high probability that one of the users will have a high quality channel. The is so-called the multi-user diversity effect. Multi-user diversity is able to provide a power gain to improve the data rate [18]. For instance, the optimal sum-rate of the scalar broadcast channel can be obtained by opportunistically transmit to the user with the largest channel gain [18]. Generally, any technique that exploit multi-user diversity in a network with many users is known as opportunistic communication. In Chapter 6, it will be shown that multi-user diversity can be leveraged to improve the secrecy capacity of the wireless broadcast channel.

2.1.7 Capacity of Wireless Channels

Capacity is an information theoretic metric to measure the maximum information that can be transferred reliably over a channel [24]. In a single-antenna point-to-point scenario, for any fixed wireless channel realisation h and assuming the use of Gaussian channel coding, the instantaneous capacity or

mutual information can be expressed as follows [18],

$$\mathcal{I} = \log_2\left(1 + \frac{P|h|^2}{\sigma^2}\right) \quad \text{bits/s/Hz}, \tag{2.2}$$

where P is the transmit power in Watts and σ^2 is the noise power in Watts. The term $\frac{P|h|^2}{\sigma^2}$ denotes the instantaneous signal-to-noise ratio (SNR) while $\frac{P}{\sigma^2}$ denotes the average SNR. The instantaneous capacity serves as an upper bound to measure the maximum data rate. Gaussian channel coding is only a mathematical construct and it is impractical for real world implementation. In practice, there is a gap between the achievable data rate and the upper bound in (2.2). In order to account for the real world performance, the following modified data rate expression is often used [25],

$$\mathrm{R} = \log_2\left(1 + \frac{P|h|^2}{\Gamma\sigma^2}\right) \quad \text{bits/s/Hz}, \tag{2.3}$$

whereΓ is the SNR gap corresponds to the target error probability and the specific channel coding scheme. The SNR gap has a typical value of $\Gamma \geq 1$ where the special case $\Gamma = 1$ corresponds to the upper bound in (2.2) where Gaussian channel coding is used.

In wireless fading channel, ergodic capacity and outage capacity are used to measure the long-term performance of the channel. Ergodic capacity is defined as the average mutual information over all channel states, i.e. $C_{ergodic} = E[\mathcal{I}]$ where the instantaneous mutual information $\mathcal{I}$ is given in (2.2). Ergodic capacity can be achieved using adaptive rate transmission and/or coding over multiple fading blocks [22]. Although ergodic capacity is commonly used for fast fading channels, it is also useful to measure the long term average data rate of slow fading channels [22].

Another metric commonly used to measure the long term performance of any fixed rate transmission is outage capacity. Outage capacity is the capacity of a fixed rate transmission before an event of outage occurred. An outage event is defined as the event when the instantaneous data rate falls below the target data rate. It is applicable to delay-constrained applications with a fixed target data rate irrespective of channel variation in a slowly varying channel. Define the outage probability as the probability that the instantaneous data rate falls below the target data rate R, i.e. $P_{out} = P(\mathcal{I} < R)$ where $\mathcal{I}$ is given in (2.2). A transmission rate that can be supported $(1 - P_{out}) \times 100\%$ of the time is known as the outage capacity, which is used to measure the real throughput.

2.2 MIMO Communication

Channel fading is not necessarily harmful to the performance of wireless communication systems. In Subsection 2.1.6 it has been shown that wireless channel fading can be exploited through opportunistic communication to improve the performance of the networks. In this section, it will be shown that channel fading can be exploited through the use of multiple antennas at the transmitter and the receiver to provide significant capacity improvement as compared to single-antenna system. The multiple antennas transceivers form the MIMO channels. Compared to single antenna, the use of multiple antennas is able to provide additional degrees of freedom in wireless systems and offers promising multiplexing and diversity gains. This section explains some important background and findings regarding the MIMO communication techniques. Subsection 2.2.1 explains the spatial multiplexing capability of a MIMO system using channel decomposition. In Subsection 2.2.2, linear precoding and beamforming are introduced. Subsection 2.2.3 discusses the asymptotic MIMO capacity as a function of number of antennas while Subsection 2.2.4 explains the asymptotic ergodic capacity as a function of SNR. In the final subsection, the tradeoff between the diversity and multiplexing gains of the MIMO channel is presented.

2.2.1 Spatial Multiplexing and Channel Decomposition

The multi-antenna transceivers provide additional spatial degrees of freedom to support the transmission of multiple independent data streams concurrently in the same channel. Such capability is known as spatial multiplexing. Spatial multiplexing is able to provide a linear increase in capacity as a function of number of transmitter and receiver antennas. To facilitate the understanding of the spatial multiplexing capability, the MIMO channel decomposition is applied to reveal the additional degrees of freedom offered by the MIMO channel.

The MIMO channel can be decomposed into non-interfering parallel channels if the instantaneous channel coefficients are known perfectly at the transmitter and receiver [26]. Consider a general discrete-time baseband equivalent MIMO system model,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.4}$$

where $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ is channel matrix with M_r receiver antennas and M_t transmitter antennas, $\mathbf{x} \in \mathbb{C}^{M_t \times 1}$ is the input symbol vector, $\mathbf{n} \in \mathbb{C}^{M_r \times 1}$ is the receiver noise vector. Using singular value decomposition (SVD) [27], the channel coefficient can be decomposed into

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{2.5}$$

Where $\mathbf{U} \in \mathbb{C}^{M_r \times M_r}$ and $\mathbf{V} \in \mathbb{C}^{M_t \times M_t}$ are unitary matrices, $\mathbf{\Lambda} \in \mathbb{R}^{M_r \times M_t}$ is the rectangular matrix with non-negative singular value λ_i (the fading amplitude) on the main diagonal and zeros for the off-diagonal elements. Substituting (2.5) into (2.4) yields,

$$\mathbf{y} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H\mathbf{x} + \mathbf{n}. \tag{2.6}$$

Define $\tilde{\mathbf{x}} = \mathbf{V}^H\mathbf{x}$, $\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y}$, and $\tilde{\mathbf{n}} = \mathbf{U}^H\mathbf{n}$, the equation above can be expressed as follows

$$\tilde{\mathbf{y}} \;\; = \;\; \mathbf{\Lambda}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \tag{2.7}$$

The matrix$\mathbf{V}$ is known as the linear precoding matrix or the transmit beamforming matrix while $\mathbf{U}$ is known as the the linear post-precessing matrix or the receive beamforming matrix. The function of the beamforming matrices is to adjust the weighting factors of the antenna elements to meet certain predefined design criterion. In this case, the beamforming matrices are designed to match the direction of the the channel. Effectively, the MIMO channel is decomposed into parallel non-interfering subchannels where each of the subchannels is able to carry one independent data stream. The parallel subchannels are also known as eigen-modes and can be expressed in scalar form

$$\tilde{y}_i = \lambda_i\tilde{x}_i + \tilde{n}_i, \quad i = 1, \ldots, M_{rank}. \tag{2.8}$$

The variable M_{rank} denotes the total number of positive λ_i, which is the rank of the matrix $\mathbf{H}$. The variable M_{rank} determines the maximum number of independent data streams that can be spatially multiplexed in the MIMO channel. From (2.8), the instantaneous capacity of the MIMO channel can

be expressed as a summation of the capacity of M_{rank} parallel channels [18],

$$\mathcal{I}_{MIMO} = \sum_{i=1}^{M_{rank}} \log_2 \left(1 + \frac{P_i \lambda_i^2}{\sigma^2}\right) \quad \text{bits/s/Hz}, \tag{2.9}$$

where the optimal power allocation factor P_i can be obtained using water-filling power allocation subjected to $\sum_{i=1}^{M_{rank}} P_i = P$. The solution of the water-filling can be expressed as [18]

$$P_i = \left(\mu - \frac{\sigma^2}{\lambda_i^2}\right)^+, \tag{2.10}$$

where the water-filling level is

$$\mu = \frac{1}{M_{rank}} \left[P + \sum_{i=1}^{M_{rank}} \frac{\sigma^2}{\lambda_i^2}\right]. \tag{2.11}$$

2.2.2 Linear Precoding and Beamforming

Precoding is generally used to describe the preprocessing of (channel coded) information bearing vector before it is transmitted using mutliple antennas through the wireless channel [22]. Depending on the linearity of the preprocessing operation, precoding can be categorised into linear precoding and non-linear precoding. Due to its lower complexity, linear precoding is desirable for practical implementation. The function of a linear precoder is the same as the transmit beamforming, where the weighting factors of the transmit antennas are adjusted according to a predefined criterion, i.e. mutual information, error probability, etc. Therefore, linear precoding and transmit beamforming are used interchangeably in this thesis. At the receiver side, linear postprocessing or receive beamforming can be performed to adjust the weighting factors of the receive antennas according to a predefined criterion. Beamforming is a more general term used to describe the linear processing at the transmitter and receiver while linear precoding is only limited to the preprocessing at the transmitter.

In order to facilitate the understanding of the role of beamforming, let us consider a point-to-point MIMO channel. Define a transmit beamforming matrix (or linear precoding matrix) $\mathbf{F}$ and a receive beamforming matrix $\mathbf{G}$. At the source, a linear precoded information vector, i.e. $\mathbf{Fx}$ is transmitted through a MIMO channel $\mathbf{H}$. The received signal at the destination can be expressed as

$$\mathbf{y} = \mathbf{HFx} + \mathbf{n}, \tag{2.12}$$

where $\mathbf{n}$ is the noise vector observed by the destination. The destination then applies the receive beamforming to the received signal, i.e. $\mathbf{Gy}$. When perfect CSI is available at both the source and destination, the rate optimal beamforming matrices are $\mathbf{F} = \mathbf{V\Sigma}$ and $\mathbf{G} = \mathbf{U}^H$, where $\mathbf{U}$ and $\mathbf{V}$ are obtained from SVD of $\mathbf{H}$ as shown in (2.5) and the diagonal matrix $\mathbf{\Sigma}$ is the power allocation matrix with its diagonal elements computed from (2.10) and (2.11).

2.2.3 Number of Antennas and Capacity

In this subsection, the effect of increasing the number of antennas at the transmitter and/or receiver to the ergodic capacity is discussed.

Consider a single-input and multiple-output (SIMO) channel with a single transmit antenna $M_t = 1$ and multiple receive antennas $M_r \geq 2$. When the CSI is known at the receiver, the optimal receive beamforming can be implemented in the SIMO channel [18]. The optimal receive beamforming is the maximal ratio combining with weighting vector $\mathbf{w} = \frac{\mathbf{h}^H}{\|\mathbf{h}\|}$ where $\mathbf{h} \in \mathbb{C}^{M_r \times 1}$ is the channel from single-antenna transmitter to the multi-antenna receiver. Figure 2.1 shows the ergodic capacity versus number of receive antennas M_r in the SIMO channel using optimal receive beamforming when $5 \leq$SNR≤ 30. From the figure, it can be observed that the increase of the number of receive antennas yields a logarithmic growth in the capacity. This means that multiple antennas at the receiver is able to deliver a power gain. Similar power gain can also be obtained in the multiple-input and single-output (MISO) channel with $M_t \geq 2$ and $M_r = 1$ provided that the transmitter has the CSI to enable maximal ratio transmit beamforming [18].

For the MIMO case when the number of transmit antennas and the receive antennas are increased simultaneously, there is a linear increase in capacity as shown in [28] and [26]. Figure 2.2 shows the ergodic capacity versus the number of antennas ($M_r = M_t$) for $5 \leq$SNR≤ 30 when perfect CSI is available at the transmitter and receiver. It can be observed that the capacity grows linearly with the simultaneous increase of the number of antennas at the transmitter and receiver. In this case, the use of multiple antennas not only delivers a power gain but also increases the multiplexing gain [18]. The multiplexing gain determines the number of independent data streams that can be transmitted simultaneously in the same channel. The exact definition of multiplexing gain is given in Section 2.2.5.

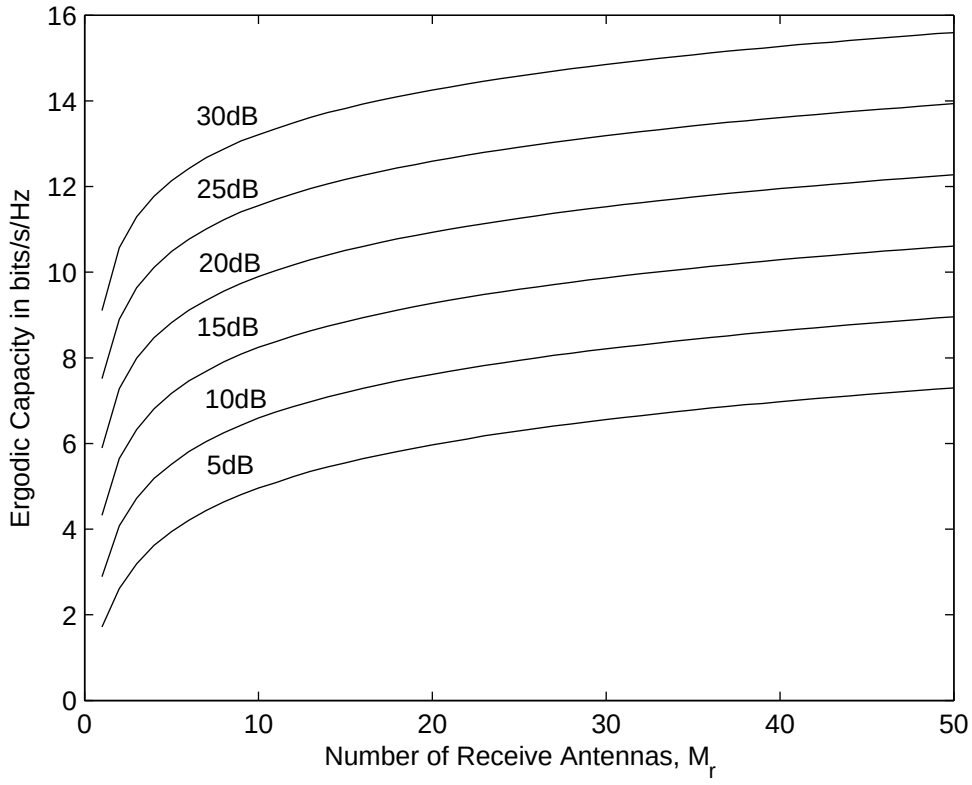


Figure 2.1: The SIMO ergodic capacity versus the number of receive antennas M_r at various SNR.

2.2.4 SNR and Capacity

What is the effect on the capacity when the SNR is increased? To obtain the high SNR asymptotic behaviour of the ergodic capacity, the number of antennas, M_r and M_t are fixed while the SNR is raised towards infinity. Define the mean SNR as $\gamma_0 = \frac{P}{\sigma^2}$. For a single user MIMO channel, the ergodic capacity when $\gamma_0 \to \infty$ can be approximated as [22]

$$E[\mathcal{I}_{MIMO}] \approx \min(M_t, M_r) \log_2 \gamma_0 + \mathcal{O}(1). \tag{2.13}$$

The pre-log factor in the above equation is known as the multiplexing gain of the MIMO channel. The actual meaning of the multiplexing gain is that each increment of 3dB of SNR increases the ergodic capacity as much as $\min(M_t, M_r)$ bits per channel use.

The multiplexing gain demonstrates the significant capacity improvement of using multi-antenna transceivers over the conventional single antenna setting. Figure 2.3 displays the ergodic capacity as a function of SNR for different number of antennas when perfect CSI is available at the transmitter and

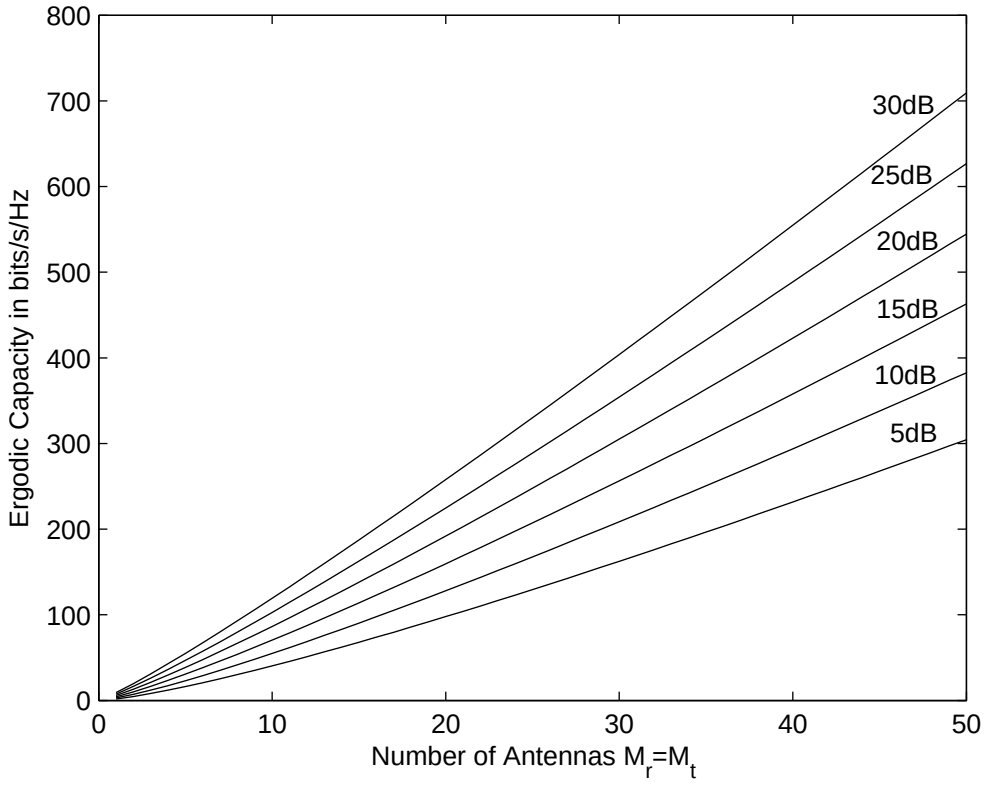


Figure 2.2: MIMO ergodic capacity versus number of antennas ($M_r = M_t$) at $5 \leq$SNR≤ 30 with perfect CSI at the transmitter and receiver.

receiver. When $M_r \times M_t$ is set to 2×2, the capacity has a significant gain if compared to the single antenna 1×1 case. As the number of transmit and receive antennas are simultaneously increased, a higher capacity can be achieved. The slope of the capacity curve increases with the multiplexing gain, i.e. $\min(M_t, M_r)$. The 4×2 and 2×2 settings have a similar capacity slope because their multiplexing gains are exactly the same. The only difference is that the 4×2 configuration has a higher power gain if compared to the 2×2 configuration.

2.2.5 Diversity and Multiplexing Tradeoff

The use of multiple antennas offers two important improvements, namely the diversity gain as discussed in Subsection 2.1.5 which improves the reliability and the multiplexing gain as presented in Subsection 2.2.4 which improves the data rate. Formally, the diversity gain d and the multiplexing gain r of a MIMO system can be defined asymptotically as follows [29]

$$d \triangleq \lim_{\gamma_0 \to \infty} -\frac{\log_2 P_e(\gamma_0)}{\log_2 \gamma_0} \text{ and } r \triangleq \lim_{\gamma_0 \to \infty} \frac{R(\gamma_0)}{\log_2 \gamma_0}, \tag{2.14}$$

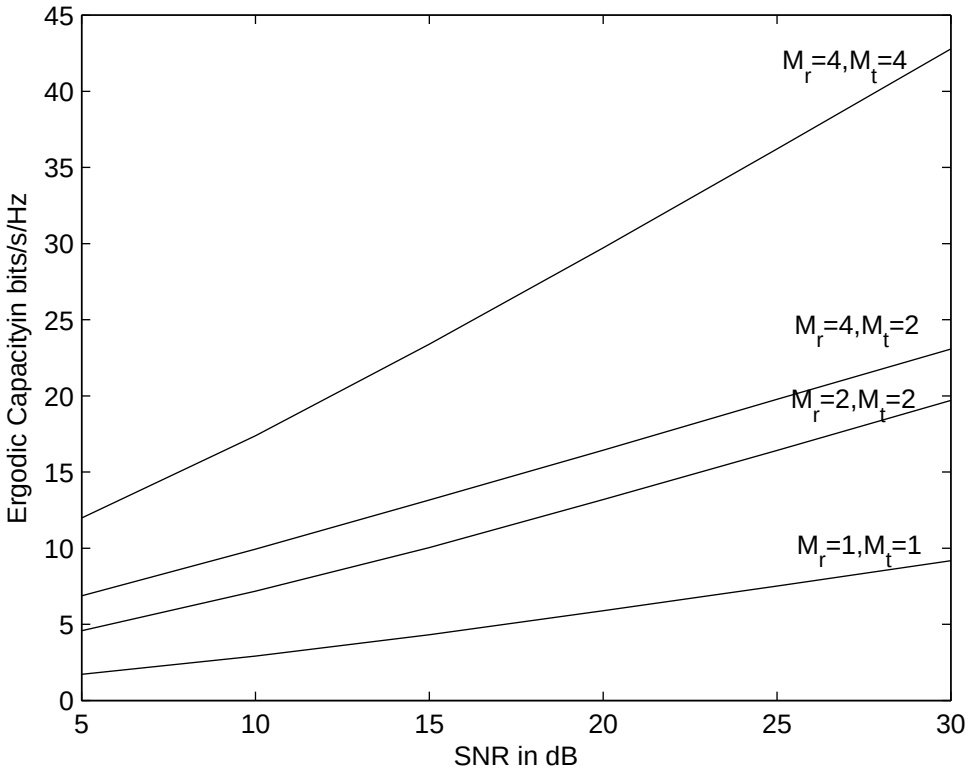


Figure 2.3: Ergodic capacity versus SNR for several antenna configurations

where P_e is maximum likelihood (ML) probability of detection error, R is the target data rate in bits/s/Hz, and γ_0 is the average SNR. Since both diversity techniques and spatial multiplexing techniques exploit the additional spatial dimension offered by the multiple antennas, there is a fundamental tradeoff between the diversity gain and multiplexing gain. The optimal diversity and multiplexing tradeoff of the MIMO channel is a piecewise linear function [29]

$$d^*(r) = (M_t - r)\,(M_r - r)\,, \tag{2.15}$$

where $r = 0, 1, \ldots, \min(M_t, M_r)$. An example of the optimal diversity and multiplexing tradeoff when $M_t = 3$ and $M_r = 4$ is shown in Figure 2.4. The maximum diversity gain, $d(r = 0)$ characterises the rate of decay of the error probability with SNR, when the data rate R is fixed. This point represents an extreme way of using the increase in the SNR to improve the reliability, for a fixed data rate. From Figure 2.4, it can be observed that the maximum diversity gain of the specific MIMO channel is $d(r = 0) = M_t M_r = 12$. On the other extreme, the maximum multiplexing gain, $d(r) = 0$, determines the rate of increase of the data rate with SNR, when the error probability P_e

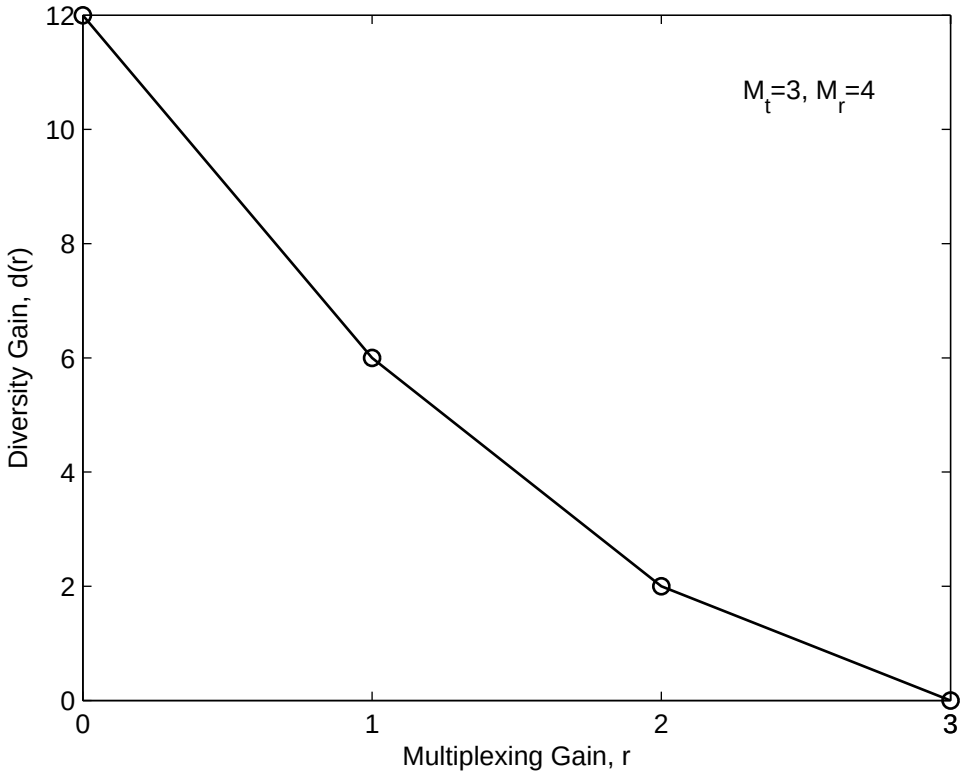


Figure 2.4: The optimal diversity and multiplexing tradeoff when $M_t = 3$ and $M_r = 4$.

is fixed. This is another extreme way of using the increase in the SNR to increase the data rate, for a fixed reliability. From Figure 2.4, it can be observed that the maximum multiplexing gain of the specific MIMO channel is $r = \min(M_t, M_r) = 3$ when $d(r) = 0$. Positive multiplexing gain $(r > 0)$ and positive diversity gain $(d > 0)$ can be obtained simultaneously, subject to the diversity and multiplexing tradeoff in (2.15). In slow fading channel, this corresponds to a system supporting a variable data rate $R(\gamma_0)$ as a function of SNR to obtain a fixed portion r of the maximum multiplexing gain of the optimal ergodic capacity. At the same time, the system will maintain a variable error probability $P_e(\gamma_0)$ as a function of SNR to obtain a fixed portion d of the maximum diversity gain of the fixed rate communication. Overall, the diversity and multiplexing tradeoff serves as an important tool to evaluate the tradeoff between rate and reliability of any centralised or distributed multi-antenna system. In Chapter 3 and Chapter 4 of this thesis, the diversity and multiplexing tradeoff is used to evaluate the performance of the proposed protocols.

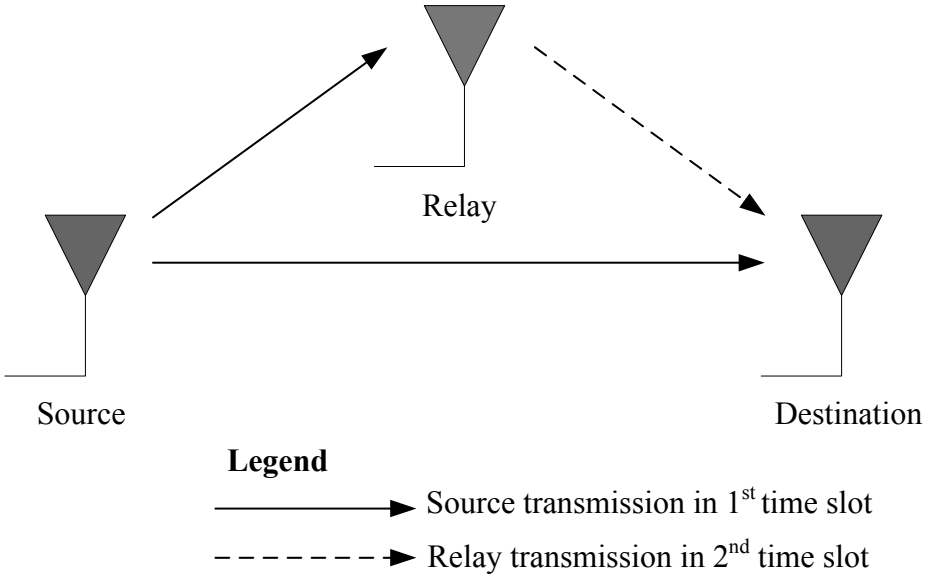


Figure 2.5: Basic scenario in the cooperative transmission

2.3 Cooperative Communication

The idea of cooperative communication dates back to the study on the capacity of the relay channel in [30] and [31] in the 70s. The research on cooperative communication in the context of contemporary communication system is pioneered by [1] in the cellular network and [2] in the ad hoc network.

Cooperative communication can be described easily in a generic relay channel consisting three nodes acting as source, relay and destination respectively, as shown in Figure 2.5. This can be seen as adding a relaying link to the direct link from source to destination. Picturing the single-antenna source and the single-antenna relay as two nodes that cooperate and share their antennas, a virtual antenna array of two antennas is formed. This virtual antenna array resembles the functions of a physical antenna array with co-located antennas, which provides additional degrees of freedom to improve throughput and reliability [32].

In practice, a cooperative transmission protocol operates in two orthogonal phases due to the half duplex constraint. Under half duplex constraint, an antenna cannot transmit and receive simultaneously in the same time and frequency channel. The half duplex constraint occurs because the transmit power is much larger than the received signal power, which will saturate the receiver amplifier if both transmission and reception occurs simultaneously. Figure 2.5 summarises the cooperative transmission protocol in two time slots. In the first time slot, the source transmits information to

the destination. Following the broadcast nature of wireless transmission, the relay is able to receive the information broadcast by the source. In the second time slot, the relay forwards the processed observation from the previous time slot to the destination. The destination combines both copies of received signal from the source and the relay to decode the desired information. Since the propagation paths of the two transmitting nodes (source and relay) have statistically independent fading, cooperative communication creates spatial diversity gain. This diversity gain averages out the adverse effect of multipath fading and increases the reliability of wireless transmission. The spatial diversity gain obtained through cooperative transmission is termed as user cooperation diversity in [1] or cooperative diversity in [2]. It is worth to mention that all nodes can become the relays for each other. Therefore, cooperative communication is a flexible and mutually benefiting technique to improve the performance of a wireless network.

The relaying strategies of cooperative communication depend very much on how the relay processes the information received from the source. The most common relaying strategies are decode-and-forward (DF) relaying and amplify-and-forward (AF) relaying. In DF relaying, the relay decodes the received signal from the source, re-encodes the source information and forwards it to the destination. In AF relaying, the relay does not attempt to decode the received signal from the source, but just forwards the power normalised observation to the destination. Both DF and AF are commonly used in the literature due to their straightforward and easy-to-understand operations, as well as their simplicity in mathematical modelling. When compared to DF relaying, AF relaying has lower computational complexity since the relay does not need to decode the source message. However, AF requires the destination to know the channel state information (CSI) of the source-to-relay channel to enable decoding at the destination, while DF does not. DF relaying is able to prevent noise-propagation by decoding the source message perfectly, while AF suffers from noise-propagation at low signal-to-noise ratios (SNR) due to the fact that the AF relay forwards noisy observation to the destination. Other existing relaying strategies are compress-and-forward (CF) and coded-cooperation. In CF relaying, the relay forwards the quantised source message to the destination. On the other hand, coded cooperation introduces redundancy in the channel coding level. Each channel codeword is partitioned into two segments where the first segment is transmitted by the source and the second segment is transmitted by the relay. A more comprehensive review on the cooperative communication can be found in [33]. One interesting finding in [33] is that all users benefit from cooperating among each

other regardless of their local channel quality. This mutual gain enjoyed by all users motivates the application of cooperative transmission in wireless network.

Nonetheless, due to the half duplex constraint, additional channel resources are needed for relaying. Recall the existing cooperative protocol described previously. Two orthogonal channel uses, e.g. two time slots as shown in Figure 2.5, are needed for cooperative transmission: one for the source transmission and the other for the relay transmission. Due to the fact that the source is only allowed to transmit new information half of the time, the spectral efficiency of such a cooperative protocol is inferior. To address the spectral efficiency issue, [3] proposes non-orthogonal relaying techniques, known as non-orthogonal amplify-and-forward (NAF) and dynamic decode-and-forward (DDF). The important feature of non-orthogonal relaying is to allow both the source and the relay to transmit simultaneously in the same channel, without violating the half duplex constraint. Both NAF and DDF allow the source to transmit message to destination all the time, utilising all the available channel resources. NAF relaying can be described in two time slots. In the first time slot, the source transmits a message while the relay listens. In the second time slot, the source transmits a new message while the relay repeats the noisy source message observed in the previous time slot. The difference between AF and NAF lies in the second time slot: NAF allows the source to continue transmitting new message while AF does not. DDF also features similar non-orthogonal relaying. However, unlike NAF, the relay in DDF does not have a fixed transmission time slot. The relay transmission happens whenever the relay gains enough information of the source message. During the relay transmission, the relay forwards the re-encoded source message to the destination. Reference [3] also demonstrates that the application of NAF and DDF in uplink and downlink respectively, yields significant performance improvement in the high spectral efficiency region.

2.4 Two-Way Relaying

In practical scenarios, information flows not only in one direction, but in both. For instance, in a cellular network there are downlink (channel from base station to mobile user) and uplink (channel from mobile user to base station) to enable information to flow in both directions. Similar examples can be found in ad-hoc networks, where nodes are exchanging information in both directions. This important communication scenario is known as the information exchange channel.

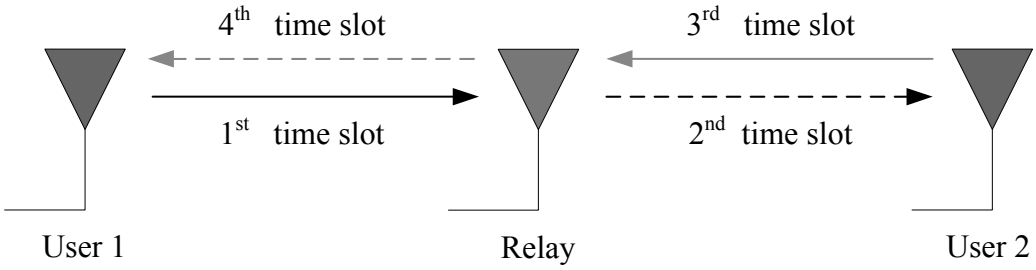


Figure 2.6: Information exchange using one-way relaying

In certain situations, the direct link between the source and destination might not be available. This happens in situations such as when source-to-destination channel is in a deep fade or undergoing severe shadowing, where the link quality is too weak to support any communication. In cellular systems, this also commonly occurs when the mobile user is located at cell edge, where the coverage of the base station is weak. In such situations, the information exchange between a pair of users depends on the relay. Using a conventional one-way relaying technique designed for uni-directional communication, the information exchange can only be completed in four channel uses due to half duplex constraint. Figure 2.6 shows the conventional one-way relaying scheme used for the information exchange. The information flows from user 1 to the relay, then from the relay to user 2 and vice versa, where a total of four time slots are used. This doubles the number of time slots used in direct communication without a relay (when direct link between source and destination exists). In such a information exchange scenario without a direct link from source to destination, one-way relaying is spectrally inefficient.

Similar two-way information exchange scenarios have been studied in wired networks, where a powerful technique known as network coding [34] is proposed. Network coding allows intermediate nodes (relays) to mix the information packets from multiple links in order to enhance the network throughput. Inspired by the idea of network coding, two-way relaying has been proposed in wireless networks. Two-way relaying makes use of the broadcast nature of wireless transmission to enable information mixing. Based on the idea of network coding, two-way relaying in wireless network is able to enhance the network throughput by reducing the channel resources used in the information exchange between a pair of users. Two-way relaying schemes such as strategies summarised in [4] based on DF relaying, analogue network coding [5] based on AF relaying, physical network coding

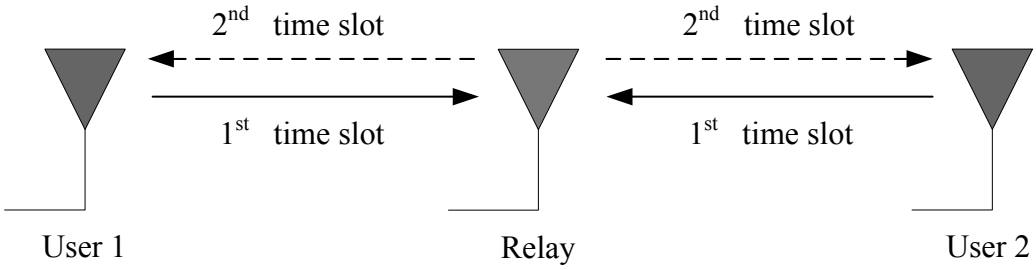


Figure 2.7: Information exchange using two-way relaying

[6] based on estimate-and-forward (EF) relaying, etc are able to complete the two way information passing in only two phases. Figure 2.7 explains the two-way relaying protocol in two time slots. In the first time slot, two users transmit simultaneously in the same channel to the relay. In the second time slot, the relay forwards the processed mixture to the users and each user uses the knowledge of his previously transmitted message, known as self-interference, to decode the new message from his partner. Since the total channel use is halved, the number of independent data streams that can be transmitted or received simultaneously per channel use in the network is doubled as compared to one-way relaying which requires four orthogonal channel uses. This promising throughput improvement offered by two-way relaying motivates its application in wireless networks.

2.5 Secrecy Communication

In the previous subsections, the broadcast nature of wireless transmission is utilised through cooperative communication and two-way relaying to deliver positive improvements such as better reliability and higher throughput. Nonetheless, the broadcast nature of wireless transmission also poses a security threat to wireless transmission. Due to the fact that the transmission over the air can be listened by other nodes, the secrecy of the information transmitted from the sender to the legitimate receiver could be compromised. Malicious nodes or adversaries with a good channel connection to the sender might be able to eavesdrop the transmission broadcast by the sender and gain access to confidential information. In commercial applications such as online shopping, internet banking, etc, the secrecy of the information transfer is top priority. In consumer applications such as video-calling, messaging, etc, the privacy of end users is of concern. These motivate the study of secrecy communication.

Traditionally, the secrecy of the wireless transmission is achieved using computation-theoretic security. Computation-theoretic security relies on the inability of the eavesdropper to solve high-complexity cryptographic problems. Computation-theoretic security assumes the physical transmission medium is error free and the eavesdropper has limited computational capability. However, perfect secrecy where the eavesdropper gains no knowledge of the secret message, cannot be guaranteed. This is due to the fact that the physical channel is imperfect and the computational capability of the eavesdropper might be underestimated.

In order to achieve perfect secrecy, Shannon [35] introduces the concept of information-theoretic security, which does not impose any limitation on the computational capability of the eavesdropper. Under the information-theoretic security, the randomness of the physical channel is exploited to prevent the eavesdropper from decoding the secret message. For this reason, information-theoretic security is also known as physical layer security. Shannon shows that perfect secrecy can be achieved by means of a one-time pad, where the secret message is combined with a random key before transmission (using an exclusive-or operation, X-OR). However, the shortcoming of a one-time pad is that it requires a new key for each new message and the key has to have the same length as the secret message. These result in large overhead in key sharing which demotivates its practical use.

Since then, the information-theoretic security has been widely studied. Wyner [36] introduces the wire-tap channel, a simple network consists of one sender, one legitimate receiver and one eavesdropper (or wiretapper). Wyner concentrates on the degraded wiretap channel where the eavesdropper channel is a noisier version of the legitimate channel. Wyner shows that a non-zero secrecy capacity can be achieved using his proposed wiretap coding scheme without the need of prior key-sharing. The wiretap coding maximises the information gained by legitimate user and minimises the information obtained by the eavesdropper. Reference [37] extends Wyner's scheme to non-degraded channels while [38] generalises it to the Gaussian channels. Wyner's wiretap coding scheme serves as an important framework for the later development of secrecy communications in wireless fading channels. Reference [39] and [40] investigate the secrecy communications in the fading channels. It is shown that by opportunistically exploiting the fluctuation of the channels, a positive secrecy capacity can be achieved using Wyner's wiretap scheme. This holds even though on the average eavesdropper channel is better than the legitimate channel [40].

Chapter 3

Cooperative Transmission Protocol with Linear Precoding for the MISO Broadcast Channel

The wireless broadcast channel is one of the most important scenarios in wireless communications. In the wireless broadcast channel, the co-channel interference among multiple users limits the achievable data rate. Practical linear precoding methods such as zero-forcing beamforming (ZFBF) have been proposed to null out co-channel interference and improve data rate. However, the achievable diversity gain of the ZFBF is limited by the number of source transmitter antennas. Cooperative transmission is able to provide another dimension of diversity gain improvement. In this chapter, a spectrally efficient cooperative transmission protocol is proposed for the multiple-input and single-output (MISO) broadcast channel. The protocol employs linear precoding, namely the ZFBF with user scheduling at the source and the ZFBF at the relays to mitigate co-channel interference. The performance of the proposed protocol is evaluated using information theoretic metrics, such as outage probability and the diversity and multiplexing tradeoff. The proposed protocol achieves the maximum diversity gain expressed as a sum of the number of source transmitter antennas and the number of available relays. The diversity and multiplexing tradeoff of the proposed protocol outperforms the comparable non-cooperative scheme, even when only a single relay is used. The use of multiple relays further improves the tradeoff between data rate and reliability. For a large number of relay

candidates, the diversity and multiplexing tradeoff of the proposed protocol completely surpasses the existing non-cooperative ZFBF scheme. Both the analytical and simulation results justify that the proposed protocol achieves better system robustness and better rate and reliability tradeoff.

3.1 Chapter Introduction

The wireless broadcast channel is one of the most important scenarios in wireless communications. In the broadcast channel, the source broadcasts information to multiple destinations. A typical example is the downlink of cellular network, where the base station broadcasts information to multiple mobile users [41]. This chapter focuses on the broadcast scenario where the base station transmits independent messages to each of the mobile users.

Multiple antenna transceivers are commonly employed in wireless networks to provide diversity and multiplexing gains [22]. In the wireless broadcast channel, the MISO configuration is particularly of interest. The single antenna constraint is due to the size and battery life limitation of the mobile terminals. Leveraging the degrees of freedom of the MISO channels, the optimal sum-rate and the optimal capacity region can be achieved using non-linear precoding technique known as dirty paper coding (DPC) [42]. With non-causal knowledge of the co-channel interference, DPC uses coding technique to precode each user's message so that the multi-user co-channel interference can be eliminated. However, high computation cost and complexity make it less attractive if compared to suboptimal techniques such as linear precoding. One of the well-known linear precoding methods is the zero-forcing beamforming (ZFBF). The ZFBF is a suboptimal precoding technique that cancels the co-channel interference among multiple users by diagonalising the channels from the source to multiple destinations [43]. It has been shown that ZFBF can achieve the optimal sum-rate when the signal-to-noise ratio (SNR) is asymptotically large [44], or when the number of users approaches infinity [45]. When a large number of users are available, the optimal sum-rate can be achieved by utilising the multiuser diversity [18] in scheduling, where the transmission is scheduled to a subset of users with preferable channel conditions. The ZFBF with semi-orthogonal user selection proposed in [46] can achieve the optimal sum-rate and perform reasonably well for system with a practical number of users, i.e less than 100 users. The semi-orthogonal user selection algorithm schedules the transmission to a subset of users whose channels are almost orthogonal to each other. The advantages in

terms of practicality and spectral efficiency (achieves the optimal sum-rate asymptotically) motivates the application of ZFBF in broadcast transmission.

The achievable diversity gain of the MISO broadcast channels using DPC or linear precoding is constrained by the number of transmitter antennas at the source node. Cooperative transmission is able to provide another dimension of diversity gain improvement, by employing distributed nodes acting as relays to form virtual antennas which forward the information broadcast by the source to the destinations [2]. In MISO broadcast channels supporting space division multiple access (SDMA), relays equipped with multi-antenna can be utilised to provide another dimension of diversity gain improvement. A multi-antenna relay can be an infrastructure relay with centralised antennas which has very good channel connection with the source [47]. The high quality link between source and relay is achieved by careful placement of the infrastructure relay. Furthermore, a multi-antenna relay can also be formed by clustering several single-antenna users who are in the idle mode and have superior channel connections with the source [48]. On the other hand, in certain military and civilian applications where the base station is mounted on tanks, trucks, buses, trains, etc [49], the number of antennas at the base station is limited due to size consideration. In these applications, cooperative transmission employing relays is an attractive solution to provide diversity gain improvement without increasing the size of the base station.

Initial work on cooperative broadcast channels focuses on single antenna configuration. A dynamic decode-and-forward (DDF) strategy is proposed in [3]. It is implemented in the broadcast channels with N destinations which act as relays for each other. Full diversity of order N is achievable by the DDF broadcast scheme. However, there is no spatial multiplexing involved. In each time slot, only one new message is transmitted. To bridge the gap, the cooperative broadcast channels supporting spatial multiplexing in the MISO broadcast channels is studied in this chapter. Introducing relays into the network might not be beneficial when the base station and the relays share common cellular bandwidth, as evident in the multi-hop relaying schemes which require orthogonal channel for each relay transmission hop [50]. However, by allowing the relay and the base station to access the channel simultaneously, i.e. non-orthogonal relaying, common bandwidth overhead due to relaying can be avoided. Simultaneous transmissions to multiple destinations introduce co-channel interference problem, which is addressed in this chapter.

In this chapter, a cooperative broadcast channel (CBC) transmission protocol using linear pre-

coding is proposed for the MISO broadcast channel consist of a source, multiple destinations and one or more relays. A relay can be either an infrastructure relay with centralised antennas or a distributed relay formed by clustering several single-antenna nodes. By using the existing ZFBF with semi-orthogonal user selection at the source [46] coupled with the ZFBF at the multi-antenna relays, co-channel interference among multiple destinations can be nullified. All qualified relays are scheduled in a round robin fashion to forward the source messages to the target destinations. At each time slot, new messages will be transmitted along with the relayed messages from the previous time slot. This non-orthogonal transmission strategy allows the source and the relays to access the shared bandwidth simultaneously, and the outcome is a spectrally efficient cooperative transmission. At each of the destinations, simple successive decoding is used to decode the mixture of the new message and the relayed message.

The diversity and multiplexing tradeoff is a tool commonly used to assess the spectral efficiency of the multiple-input and multiple-output (MIMO) system [29]. Since the MISO cooperative broadcast channels can be viewed as a special case of MIMO, the performance of the proposed CBC protocol is evaluated using the diversity and multiplexing tradeoff. The achievable diversity and multiplexing tradeoff of the proposed protocol outperforms the comparable non-cooperative ZFBF scheme [46], even when only single relay is used. Increasing the number of relays further improves the tradeoff between data rate and reliability of the proposed protocol. When a large number of relays is available, the achievable diversity and multiplexing tradeoff of the proposed CBC scheme completely surpasses the comparable non-cooperative scheme. The maximum diversity gain expressed as a sum of the number of source transmitter antennas and the number of available relays is achievable by the proposed scheme. As a comparison, the maximum diversity gain achievable by the non-cooperative ZFBF scheme is limited by the number of source transmitter antennas. Monte-Carlo simulations show that the outage probability of the proposed scheme outperforms the non-cooperative ZFBF scheme. Simulations also justify that the protocol offers better rate and reliability tradeoff.

This chapter is organised as follows. In Section 3.2, the proposed CBC protocol is presented. The analytical development of outage probability and the diversity and multiplexing tradeoff of the CBC protocol are shown in Section 3.3. In Section 3.4, the numerical results on the outage probability are provided to evaluate the performance of the proposed protocol. Section 3.5 concludes this chapter.

3.2 Description of the Proposed Protocol

Consider a broadcast scenario with one source, M active destinations[1] and L relays. The source and the relays are equipped with M_t and M_r antennas respectively, while the destination nodes are equipped with single antenna. A multi-antenna relay can be an infrastructure relay with centralised antennas, which is commonly proposed for a cellular network [47]. Otherwise, a multi-antenna relay can be formed by clustering several single-antenna users who are in the idle mode[2]. The relays use the decode-and-forward (DF) relaying strategy. Since full spatial multiplexing supporting M active destinations simultaneously is of interest, the number of antennas at the source and relays are set as $M_t = M$ and $M_r = M$ respectively. An example of the cooperative broadcast scenario with $M = 2$ and $L = 1$ is shown in Figure 3.1.

The network is assumed to be symmetrical, where each channel is independent and identically distributed (i.i.d.) with the same channel statistics and all destinations have the same target data rate, i.e. symmetrical rate. The channels are modelled as frequency non-selective, quasi-static Rayleigh fading, where all channels remain constant in the duration of a CBC transmission frame. The noise observed at each receiver is circularly symmetric complex Gaussian distributed, i.e. $n \sim \mathcal{CN}(0, \sigma_n^2)$. The baseband equivalent, discrete-time channel model is used to model the continuous-time channel. The practical half duplex constraint is imposed on all nodes, i.e. nodes cannot transmit and receive simultaneously. Time division duplex (TDD) is used to satisfy this constraint.

The proposed CBC protocol can be described in a CBC transmission frame. Each CBC transmission frame consists of two phases, namely the initialisation phase and the cooperative transmission phase. Under quasi-static fading, the channel coherence time (where the channel remains constant) is much longer than the delay requirement. As a result, the cooperative transmission phase can be repeated multiple times without going through the initialisation phase. The quasi-static fading assumption is important to guarantee that the channel state information (CSI) obtained through the initialisation phase remains valid in the cooperative transmission phase. In general, a CBC transmission frame consists of one initialisation phase and multiple cooperative transmission phases. The

[1] Active destinations are the users served by the source at a particular time. In a cellular network using SDMA, the total number of users is usually much larger than the number of antennas at the base station. Scheduling is usually performed to enable time sharing between groups of users, where the number of users in each active group is at most the number of antennas at the source (assuming full spatial multiplexing).

[2] The bandwidth cost of clustering (i.e. information exchange overhead between members in a cluster) to the cellular system can be avoided by using the secondary radio resource available at the mobile users, such as Wi-Fi, Bluetooth, etc.

duration of an initialisation phase is much more smaller than the total duration of multiple cooperative transmission phases. Figure 3.2 shows the flow chart of a CBC transmission frame when K qualified relays are used to assist the transmission. The details of the protocol are described in the following subsections.

3.2.1 Initialisation

All nodes assume perfect knowledge of the local channel state information (CSI) to enable linear precoding at the source and relays, and coherent detection at the relays and destinations. More precisely, the source has the CSI of the all source-destination channels while the relays have the CSI of the source-relay, inter-relay, and relay-destination channels. The destinations have the CSI of their respective source-destination and relay-destination channels. Using the channel reciprocity principle, the CSI can be acquired using the open-loop method [22]. Using this method, the source and each of the destinations will take turns to broadcast the training symbols. All nodes use such training information to estimate their local channels. With the knowledge of its local channels, every relay knows whether it can decode the source messages successfully. The criterion for successful decoding will be discussed in the next subsection. Say there are K qualified relays. The qualified relays take turns to broadcast the training symbols so that each destination can estimate the relay-destination link.

3.2.2 Cooperative Transmission

Before going into the details of the protocol description, the ZFBF with semi-orthogonal user scheduling proposed in [46] is reviewed. The ZFBF with user scheduling is used for the direct transmission between the source and the destinations in the proposed protocol. Using the ZFBF with scheduling, in each transmission time slot, the source schedules M active destinations which have good channel gains and channel directions matching the beam directions of the ZFBF. When the total number of users in the network (comprising all active and idle users) is asymptotically large, M perfectly orthogonal channels can be scheduled using the user scheduling algorithm proposed in [46]. In this case, the channel inversion operation in the ZFBF reduces to a rotation operation, where the beam directions match the user channel directions exactly. As a result, the power loss due to channel inversion in ZFBF (where the beamforming weights do not match the channels) can be avoided. In addition, the optimal sum-rate can be achieved asymptotically, by having a large number of users.

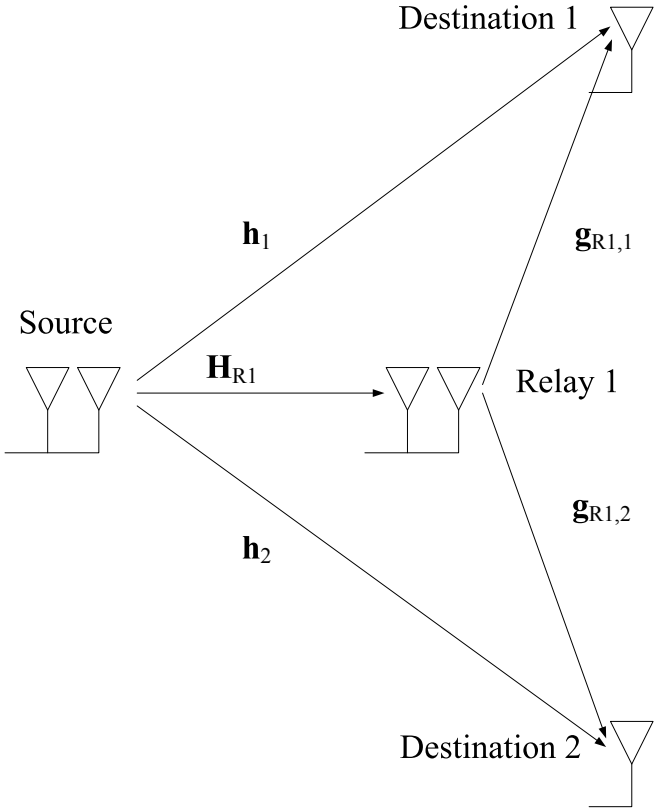


Figure 3.1: Cooperative broadcast scenario with 2-antenna source, 2-antenna relay and 2 single-antenna users.

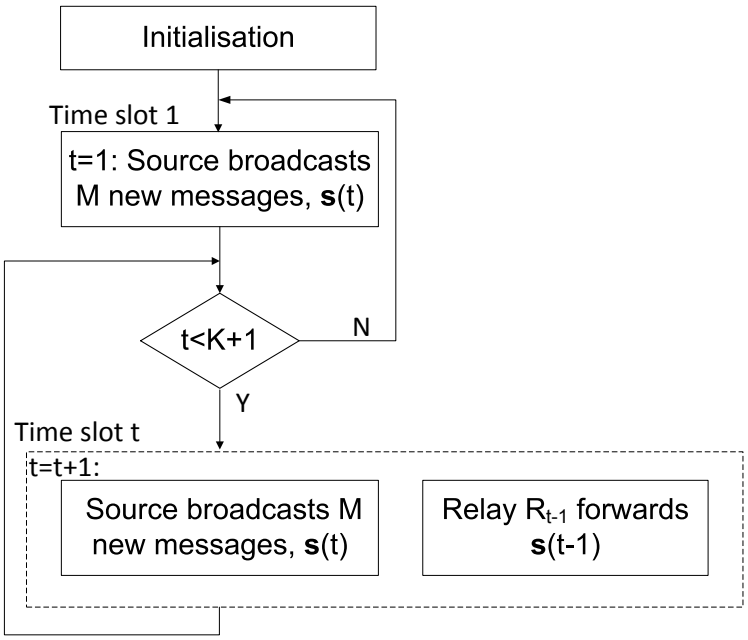


Figure 3.2: Flow chart of a CBC transmission frame of the proposed protocol.

The ZFBF with semi-orthogonal user scheduling [46] is the baseline non-cooperative scheme considered in this chapter. The same scheme is used for the direct transmission between the source and destinations, in order to isolate the benefits of using cooperative relays. For convenience of analysis, it is assumed that an asymptotically large number of users is available to achieve perfectly orthogonal source-destination channels. Note that the ZFBF with semi-orthogonal user scheduling also works for a finite number of users, and delivers a good sum-rate performance under reasonable number of users, i.e. less than 100. Refer to [46] for details. Applying the ZFBF with scheduling, M channel interference-free parallel transmission can be achieved in each transmission time slot. The parallel channels allow us to concentrate on the signal description at only one single destination. In the following, the details of the proposed protocol are described, in time slotted fashion.

3.2.2.1 First Time Slot

In the first time slot, the source transmits M messages simultaneously, such that $\mathbf{x}(1) \in \mathbb{C}^{M\times 1} = \mathbf{W}\mathbf{s}(1)$, where the column vector $\mathbf{s}(1) \in \mathbb{C}^{M\times 1} = \begin{bmatrix} s_1(1) & \cdots & s_M(1) \end{bmatrix}^T$ and the scalar $s_m(t)$ is the message transmitted at time slot t to mth destination. The channel from the multi-antenna source to the single-antenna destination m is denoted by row vector $\mathbf{h}_m^T \in \mathbb{C}^{1\times M}$ while $\mathbf{W} \in \mathbb{C}^{M\times M} = \begin{bmatrix} \mathbf{w}_1 & \ldots & \mathbf{w}_M \end{bmatrix}$ is the precoding weighting matrix with column vector $\mathbf{w}_m \in \mathbb{C}^{M\times 1}$ as the precoding weighting vector of the mth user. Since M parallel orthogonal MISO channels can be created using user scheduling in an asymptotically large network, the ZFBF with semi-orthogonal user scheduling can be modeled as a maximal ratio transmission beamformer [18] at each parallel channel, such that $\mathbf{w}_m = \frac{\mathbf{h}_m^*}{||\mathbf{h}_m||}$, where the normalising factor is to ensure unit transmission power for each stream, i.e. $\mathbf{w}_m^H\mathbf{w}_m = 1$. This corresponds to the fact that when the user channels are orthogonal to each other, the channel inversion operation in ZFBF reduces to a rotation operation, where the beam directions directly match the channel directions. Since the high SNR performance of the proposed protocol is of interest, i.e. diversity and multiplexing gains, equal power allocation among multiple users is sufficient [18]. To ensure equal power allocation among users, $\mathbf{W}^H\mathbf{W} = \mathbf{I}_M$.

The signal received by mth destination can be expressed as

$$y_m(1) = \sqrt{E_s}\mathbf{h}_m^T\mathbf{w}_m s_m(1) + n_m, \tag{3.1}$$

where E_s is the average transmit power for each message, n_m is the noise at mth destination and $\mathbf{h}_m^T \mathbf{w}_m = ||\mathbf{h}_m||$. The signal received by the kth relay is

$$\mathbf{r}_k(1) = \sqrt{E_s}\mathbf{H}_{R_k}\mathbf{W}\mathbf{s}(1) + \mathbf{n}_k, \tag{3.2}$$

where $\mathbf{H}_{R_k} \in \mathbb{C}^{M \times M}$ is the channel matrix from multi-antenna source to kth multi-antenna relay and $\mathbf{n}_k \in \mathbb{C}^{M \times 1}$ is the noise vector observed by relay k.

Out of total L available relays, there are K qualified relays that manage to decode all messages from the source successfully. Since symmetrical system is considered, each user is allocated with the same data rate, R bits/s/Hz. Assuming Gaussian coding with long codewords and zero-forcing receiver [51], the qualification criterion for the kth relay can be expressed as

$$\log_2\left(1 + \frac{\gamma_0}{\left[\left(\mathbf{W}^H\mathbf{H}_{R_k}^H\mathbf{H}_{R_k}\mathbf{W}\right)^{-1}\right]_{m,m}}\right) > R, \quad \forall m \in \{1, 2, \ldots, M_r\}, \tag{3.3}$$

where $\gamma_0 = \frac{E_s}{\sigma^2}$ is the mean SNR. Note that the zero-forcing criterion requires $M_r \geq M_t$. All K qualified relays decode the messages and store them in memory.

A linear receiver, namely the zero-forcing receiver is used at the relays, for several reasons. First, the linear receiver has lower computational complexity when compared to the optimal maximum likelihood (ML) decoder. Second, the linear receiver enables the separation of linear spatial filtering and single-input and single-output (SISO) decoding [51]. The SISO decoding (or per-stream decoding) is more suitable for distributed implementation where the relay is a cluster of idle users, since each single-antenna user can decode his/her data independently using an individual SISO decoder. For the choice of a linear receiver, zero-forcing receiver is sufficient since it achieves the same high SNR performance as the minimum mean-squared-error (MMSE) receiver [51]. The choice of using a zero-forcing receiver not only makes the analysis tractable but also fulfils the goal to study the achievable performance of the proposed protocol. Nonetheless, there is a price to pay for using zero-forcing receiver. The diversity gain achieved by zero-forcing receiver is at most $M_r - M_t + 1$ [52], as opposed to the full diversity of $M_t M_r$ offered by the channel itself. It will be shown in Section III and IV that the diversity loss of using zero-forcing receiver can be compensated by using multiple relays.

3.2.2.2 Second Time Slot

During the second time slot, the source broadcasts M new messages concurrently, such that $\mathbf{x}(2) \in \mathbb{C}^{M\times 1} = \mathbf{W}\mathbf{s}(2)$. Denote row vector $\mathbf{g}_{R_k,m}^T \in \mathbb{C}^{1\times M}$ as the channel between the relay R_k and the mth destination. All K qualified relays are scheduled to transmit in a round robin fashion. Only one relay transmits in each cooperative time slot. Specifically, the first relay forwards the precoded messages, $\mathbf{x}_{R_1} \in \mathbb{C}^{M\times 1} = \mathbf{P}_{R_1}\mathbf{s}(1)$, where the precoding matrix $\mathbf{P}_{R_1} \in \mathbb{C}^{M\times M} = [\mathbf{p}_{R_1,1} \cdots \mathbf{p}_{R_1,m} \cdots \mathbf{p}_{R_1,M}]$. Each column vector $\mathbf{p}_{R_1,m} \in \mathbb{C}^{M\times 1}$ should ensure no interference to destinations other than the target mth destination. Denote $\mathbf{A}$ as a matrix which contains all the interfering channels (all channels other than its targeted destination). In this instance, denote

$$\mathbf{A} \in \mathbb{C}^{M\times M} = \begin{bmatrix} \mathbf{g}_{R_1,1} & \cdots & \mathbf{g}_{R_1,m-1} & \mathbf{g}_{R_1,m+1} & \cdots & \mathbf{g}_{R_1,M} \end{bmatrix}^T, \tag{3.4}$$

and $\mathbf{p}_{R_1,m}$ lies in the null space of $\mathbf{A}$, such that

$$\mathbf{p}_{R_1,m} \in \text{null}(\mathbf{A}), \tag{3.5}$$

so that $\mathbf{g}_{R_1,i}^T\mathbf{p}_{R_1,j} = 0$ for any $i \neq j$. The null-space vector exists when the dimension condition, $M_r > \text{rank}(\mathbf{A})$ is satisfied, and the independence condition where $\mathbf{g}_{R1,m}^T$ is linearly independent of the rows in $\mathbf{A}$ is met [11]. The first condition is satisfied since $M_r = M$ is chosen. The latter is satisfied with high probability since the channels are i.i.d. The computation of the null-space vectors can be found in [27]. Note that $\mathbf{p}_{R_1,m}$ has unit power, i.e $\mathbf{p}_{R_1,m}^H\mathbf{p}_{R_1,m} = 1$. Similar to the source transmission, equal power allocation among multiple users is assumed at the relay, i.e. $\mathbf{P}_{R_1}^H\mathbf{P}_{R_1} = \mathbf{I}$. The ZFBF is able to multiplex M streams of messages simultaneously and eliminate the co-channel interference among destinations. At the mth destination, the observation is

$$y_m(2) = \sqrt{E_s}||\mathbf{h}_m||s_m(2) + \sqrt{E_s}\mathbf{g}_{R_1,m}^T\mathbf{p}_{R_1,m}s_m(1) + n_m(2). \tag{3.6}$$

Although the remaining relays also receive a mixture of the messages, they will be able to decode the source messages using simple successive decoding since they satisfy the qualification criterion, as stated in (3.3) and thus have knowledge of the messages transmitted from the source in the previous

time slot.

3.2.2.3 Remaining Time Slots

The relaying process continues until all K qualified relays have been used. Due to half duplex constraint, each relay can only be used once in each cooperative transmission phase[3]. The use of multiple relays, i.e. $K > 1$, is beneficial in terms of the diversity and multiplexing tradeoff, which will be discussed in Section III and IV. Let column vector $\mathbf{s}_m \in \mathbb{C}^{(K+1)\times 1} = \begin{bmatrix} s_m(1) & \cdots & s_m(K+1) \end{bmatrix}^T$, noise vector $\mathbf{n} \in \mathbb{C}^{(K+1)\times 1} = \begin{bmatrix} n_m(1) & \cdots & n_m(K+1) \end{bmatrix}^T$ channel matrix

$$\mathbf{H}_m \in \mathbb{C}^{(K+1)\times(K+1)} = \begin{bmatrix} ||\mathbf{h}_m|| & 0 & \cdots & 0 \\ \mathbf{g}_{R_1,m}^T \mathbf{p}_{R_1,m} & ||\mathbf{h}_m|| & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \mathbf{g}_{R_K,m}^T \mathbf{p}_{R_K,m} & ||\mathbf{h}_m|| \end{bmatrix}. \tag{3.7}$$

The signal model after stacking $K+1$ time slots can be written as

$$\mathbf{y}_m = \mathbf{H}_m \mathbf{s}_m + \mathbf{n}. \tag{3.8}$$

The elements on the main diagonal of the effective channel matrix $\mathbf{H}_m$ represent the direct transmission links (source to destination) while the elements on the lower sub-diagonal represent the cooperative links (relays to destination). Assuming Gaussian coding and long codewords are used to convey messages, the mutual information at mth destination node can be expressed as

$$\mathcal{I}_K = \frac{1}{K+1} \log_2 \det \left(\mathbf{I}_{K+1} + \gamma_0 \mathbf{H}_m \mathbf{H}_m^H \right), \tag{3.9}$$

where $\gamma_0 = \frac{E_s}{\sigma_n^2}$ is the mean SNR and the subscript K represents the number of qualified relays.

The mutual information of the non-cooperative ZFBF with semi-orthogonal scheduling scheme can be derived by omitting the cooperative links (sub-diagonal elements) from the channel matrix

[3]When a relay is transmitting, it misses the new messages broadcast by the source, due to half duplex constraint. As a result, it will not be able to perform successive decoding to decode the new messages broadcast by the source in the subsequent time slot. This prohibits the relay to be reused in each cooperative transmission phase.

$\mathbf{H}_m$. It follows that the mutual information at the mth destination node of the non-cooperative ZFBF scheme is

$$\mathcal{I}_{ZFBF} = \log_2\left(1 + \gamma_0||\mathbf{h}_m||^2\right). \tag{3.10}$$

Note that the assumption of large number of users is used, where perfectly orthogonal channels can be formed by using the semi-orthogonal user scheduling. As a result, the ZFBF reduces to maximal ratio transmit beamforming, which has mutual information shown in (3.10).

3.3 Analysis of Outage Probability and the Diversity and Multiplexing Tradeoff

This section provides the performance evaluation of the proposed CBC protocol using outage probability and the diversity and multiplexing tradeoff. The diversity gain d and the multiplexing gain r of a MIMO system can be defined as [29]

$$d \triangleq \lim_{\gamma_0 \to \infty} -\frac{\log_2 P_e(\gamma_0)}{\log_2 \gamma_0} \text{ and } r \triangleq \lim_{\gamma_0 \to \infty} \frac{R(\gamma_0)}{\log_2 \gamma_0}, \tag{3.11}$$

where P_e is maximum likelihood (ML) probability of detection error, R is the target data rate in bits/s/Hz, and γ_0 is the mean SNR. The diversity and multiplexing tradeoff can be derived using outage formulation, which is usually used for non-ergodic fading channels, i.e. quasi-static channels.

The use of ZFBF at the source and the relay enables the decomposition of a MISO broadcast channel into parallel MISO channels, i.e. a set of point-to-point MISO channels orthogonal to each other. Recall the symmetrical network assumption, i.e. all channels have the same channel statistics and all destination nodes have a symmetric rate R. This results in each of the parallel MISO channels exhibiting statistically identical outage behaviour. Hence, the analysis of the individual point-to-point MISO channel is sufficient to assess the performance of the proposed protocol.

Define the outage event experienced by each user of the proposed CBC protocol as

$$\mathcal{O} \triangleq \bigcup_{\mathrm{K}} \mathcal{O}_K, \tag{3.12}$$

where $\mathcal{O}_K$ is the event that the mutual information when there are K qualified relays, is below the individual target data rate, i.e $\mathcal{I}_K \leq R$. The single-user outage probability can be derived using the following formula,

$$P(\mathcal{O}) = \sum_{K=0}^{L} P(\mathcal{O}_K)P(K), \tag{3.13}$$

where $P(K)$ is the probability that K relays are qualified to participate in the cooperation, and $P(\mathcal{O}_{\mathrm{K}})$ is the single-user outage probability when K relays are qualified.

Before addressing the question of how many relays are qualified, the probability that a relay decodes the source messages successfully is determined. In order to formalise the qualification criterion into a probability, the statistical distribution of the SNR at each relay is derived. The channel matrix from the source with M_t transmitter antennas to the relay with M_r receiver antennas can be generalised as $\mathbf{H}_{R_k}$ with dimension $M_r \times M_t$. Recall (3.2), the virtual channel matrix can be written as $\mathbf{H} = \mathbf{H}_{R_k}\mathbf{W}$, where $\mathbf{H}_{R_k} \sim \mathcal{CN}(0, \mathbf{I}_{M_t} \otimes \mathbf{\Sigma})$, $\mathbf{\Sigma}$ is the covariance matrix with dimension $M_r \times M_r$. In other words, the columns of $\mathbf{H}_{R_k}$ are independent $M_r \times 1$ random vectors, each with covariance matrix $\mathbf{\Sigma}$. The weighting matrix $\mathbf{W}$ is treated as a constant matrix with orthogonal columns , i.e $\mathbf{w}_i^H \mathbf{w}_j = 0$ when the column number $i \neq j$. This is valid when semi-orthogonal user scheduling [46] is employed and an asymptotically large network is assumed. In this case, the columns of $\mathbf{H}$ remain independent of each other, i.e $\mathbf{h}_i^H \mathbf{h}_j = 0$ when the column number $i \neq j$. As discussed in Section 3.2.2.1, the zero-forcing receiver is employed at the relay. Using the zero-forcing criterion at the receiver, the received signal is decomposed into M_r parallel streams. Denote $\mathbf{h}_k$ as the kth column of $\mathbf{H}$, $\tilde{\mathbf{H}}$ as the remaining columns of $\mathbf{H}$ after removing the kth column, and $\mathbf{w}_k$ as the kth column of $\mathbf{W}$. The instantaneous SNR of the kth stream at the relay receiver can be obtained using the following lemma.

Lemma 3.1. *Using a zero-forcing receiver, the SNR of kth stream at the relay receiver can be expressed as*

$$\gamma_k = \gamma_0 \mathbf{h}_k^H \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H \mathbf{h}_k, \tag{3.14}$$

where $\gamma_0 = \frac{E_s}{\sigma_n^2}$, $\mathbf{Q}$ is the eigen vector, $\mathbf{\Lambda}$ is the eigen value and $\mathbf{Q\Lambda Q^H} = \mathbf{I} - \tilde{\mathbf{H}} \left(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}} \right)^{-1} \tilde{\mathbf{H}}^H$.

Proof. Refer to [53] . ■

Before computing the distribution of the SNR of the kth data stream, the distribution of $\mathbf{h}_k$ is determined. The subsequent lemma provides us the distribution of $\mathbf{h}_k$.

Lemma 3.2. *The weighted sum of independent complex Gaussian random vectors is a Gaussian random vector. Given the distribution $\mathbf{H}_{R_k} \sim \mathcal{CN}(0, \mathbf{I}_{M_t} \otimes \mathbf{\Sigma})$, the vector $\mathbf{h}_k = \mathbf{H}_{R_k} \mathbf{w}_k$ is complex Gaussian distributed with zero mean and covariance $||\mathbf{w}_k||^2 \mathbf{\Sigma}$, such that $\mathbf{h}_k \sim \mathcal{CN}_{Mr}(0, ||\mathbf{w}_k||^2 \mathbf{\Sigma})$.*

Proof. Refer to [54]. ■

By using Lemma 3.1 and Lemma 3.2, the distribution of the SNR of the kth data stream at the relay receiver is obtained, as shown in the following lemma.

Lemma 3.3. *When $\mathbf{H}_{R_k} \sim \mathcal{CN}(0, \mathbf{I}_{M_t} \otimes \sigma^2 \mathbf{I}_{Mr})$, the SNR of the kth stream at the relay receiver is a chi-square variable with probability density function (PDF)*

$$f(\gamma_k) = \frac{(\gamma_k)^{M_r - M_t}}{(\gamma_0 \sigma^2 ||\mathbf{w}_k||^2)^{(M_r - M_t + 1)} (Mr - M_t)!} \exp\left(-\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2} \right). \tag{3.15}$$

The corresponding cumulative density function (CDF) is

$$F(\gamma_k) = 1 - \exp\left(-\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2} \right) \sum_{n=0}^{M_r - M_t} \frac{1}{n!} \left(\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2} \right)^n. \tag{3.16}$$

Proof. Refer to Appendix A. ■

With the knowledge of the SNR distribution of each data stream at the relay receiver, the probability that a relay is qualified can be determined. Assuming the SNR of each stream is independent, the following lemma is obtained.

Lemma 3.4. *Assuming an independent SNR distribution for each stream at the relay receiver, the probability that a relay is qualified at high SNR can be approximated as follows,*

$$P(A) \approx 1 - M_r \frac{\gamma^{M_r - M_t + 1}}{(M_r - M_t + 1)!}, \tag{3.17}$$

where A is the event that a relay is qualified, $\gamma = \frac{2^R - 1}{\gamma_0 ||\mathbf{w}_k||^2}$ *and* $R = r \log_2(\gamma_0)$.

Proof. Refer to Appendix A. Note that the SNR of each stream in a linear receiver is not strictly independent. However, this assumption does not affect the diversity of the system and is used to make the analysis tractable, as discussed in [55]. ■

Given a total of L available relays, the number of qualified relays, K, is expressed using the following lemma.

Lemma 3.5. *At high SNR, the probability that* K *relays satisfy criterion* (3.3) *can be approximated as*

$$P(K) \approx \frac{L!}{(L-K)!K!} \left(M_r \frac{\gamma^{M_r - M_t + 1}}{(M_r - M_t + 1)!} \right)^{L-K}, \tag{3.18}$$

where $\gamma = \frac{2^R - 1}{\gamma_0 ||\mathbf{w}_k||^2}$, $R = r \log_2(\gamma_0)$ *and* $K \leq L$.

Proof. Refer to Appendix A. ■

At high SNR, i.e. $\gamma_0 \to \infty$, $P(K = L) \to 1$ and $P(K \neq L) \to 0$. This implies that all relays are qualified to participate in the cooperative transmission at high SNR. The simulation of the size of the qualified relays set without the high SNR approximation will be discussed in the next section.

With the knowledge of the number of qualified relays, the outage performance of the proposed protocol can be determined. The upper bound of $P(\mathcal{O}_{\mathrm{K}})$ is expressed in the following lemma.

Lemma 3.6. *Assuming full spatial multiplexing at the relays, i.e.* $M_r = M$*, the outage probability for the event that* K *relays are qualified at high SNR can be expressed as*

$$P(\mathcal{O}_{\mathrm{K}}) \leq \frac{\left(\frac{\gamma_0^r - 1}{\gamma_0}\right)^{M_t}}{M_t! \gamma_0^K} \left((-1)^K - \gamma_0^{(K+1)r} \sum_{i=0}^{K-1} \frac{\left(\ln \gamma_0^{(K+1)r}\right)^i}{i!(-1)^{K-i}} \right), \tag{3.19}$$

where the multiplexing gain r *is as defined in* (3.11)*.*

Proof. Refer to Appendix A. ■

Recall the definition of diversity gain d and the multiplexing gain r of a MIMO system in (3.11). Since the ML error probability can be tightly bounded by the outage probability at high SNR [29], the outage probability can be used to derive the diversity and multiplexing tradeoff. Recall that channel decomposition and symmetric network assumption allow us to focus on the performance analysis of individual point-to-point MISO channel. MISO is a special case of MIMO, therefore the tradeoff formulation in [29] can be directly applied to our case.

Combining the result from Lemma 3.5 and Lemma 3.6, the single-user outage probability as defined in (3.13) can be obtained. Using the single-user outage probability formulation, the following diversity and multiplexing tradeoff experienced by each user can be established.

Theorem 3.1. *Assuming full spatial multiplexing at the source and relay, i.e.* $M_t = M$ *and* $M_r = M$*, the achievable diversity and multiplexing tradeoff of the proposed CBC protocol is*

$$d_{CBC}(r) = M(1-r) + [L - (L+1)r]^+, \tag{3.20}$$

where $0 \leq r \leq 1$*.*

Proof. Assuming full spatial multiplexing at the source and relay, i.e. $M_r = M_t = M$ and combining the results from Lemma 3.5 and Lemma 3.6, the outage probability of the proposed CBC protocol can be expressed as

$$P(\mathcal{O}) = \sum_{K=0}^{k} P(\mathcal{O}_K) P(K) \doteq {\gamma_0}^{-(L-K)(1-r)} \gamma_0^{-[(M+K)(1-r)-r]}. \tag{3.21}$$

Notice that the exponential equality is obtained only when $\gamma \to \infty$. At high SNR, the exponent of SNR of the approximation in (3.18) remains the same as the original $P(K)$. Similarly, the upper bound in (3.19) preserves the exponent of SNR as the original $P(\mathcal{O}_K)$. Therefore, the exponential equality in (3.21) can be obtained.

The diversity and multiplexing tradeoff can be obtained by combining the exponent of SNR γ_0,

$$d(r) = \begin{cases} M(1-r) + L - (L+1)\, r, & if \quad 0 \leq r \leq \frac{L}{L+1} \\ M(1-r) & if \quad \frac{L}{L+1} < r \leq 1 \end{cases}. \tag{3.22}$$

The derived diversity and multiplexing tradeoff is the best achievable tradeoff of the proposed scheme. Hence, the theorem is proved. ■

Remark for Theorem 3.1

As a comparison, the single-user diversity and multiplexing tradeoff achievable by the non-cooperative ZFBF broadcast scheme with scheduling [46] can be represented by the optimal tradeoff of the MISO channel with M source antennas, such that

$$d_{ZFBF}(r) = M(1-r), \tag{3.23}$$

where $0 \leq r \leq 1$.

The maximum diversity gain, $d(r = 0)$ characterises the rate of decay of the outage probability with SNR, when the data rate is fixed. This point represents an extreme way of using the increase in the SNR to improve the reliability, for a fixed data rate. The maximum diversity gain achievable by the non-cooperative scheme, as in (3.23), is constrained by the number of transmitter antennas at the source, M. In contrast, the proposed CBC protocol can achieve maximum diversity gain of $M + L$. The proposed protocol provides an extra L order diversity gain, which is a function of number of available relays. The additional diversity gain of L can be achieved since at high SNR, all relays are qualified to participate in the cooperative transmission. For the other extreme, the maximum multiplexing gain, $d(r) = 0$, determines the rate of increase of the data rate with SNR, when the

outage probability is fixed. This is another extreme way of using the increase in the SNR to increase the data rate, for a fixed reliability. The proposed CBC achieves the same maximum multiplexing gain, $r = 1$, as the non-cooperative scheme. This implies that no extra bandwidth is consumed by the CBC protocol for using relay to assist the communication between source and destinations. This benefit comes from the use of non-orthogonal relaying in the proposed protocol.

Positive multiplexing gain $(r > 0)$ and positive diversity gain $(d > 0)$ can be obtained simultaneously, subjected to the diversity and multiplexing tradeoff. As in [3], we say that protocol A uniformly dominates protocol B when $d_A(r) \geq d_B(r)$ for any multiplexing gain r. The diversity and multiplexing tradeoff of the CBC protocol uniformly dominates the non-cooperative ZFBF protocol, i.e. $d_{CBC}(r) \geq d_{ZFBF}(r)$ for any r. When the multiplexing gain $r < \frac{L}{L+1}$, the diversity and multiplexing tradeoff of the proposed protocol is always better than the non-cooperative ZFBF's, i.e. $d_{CBC}(r) > d_{ZFBF}(r)$. When the multiplexing gain $r > \frac{L}{L+1}$, the tradeoff the proposed protocol is identical to the non-cooperative protocol. This can be explained in the following example. Suppose there is only 1 relay participating in the cooperation, i.e. $L = 1$. In this case, 2 messages are broadcast by the source during each cooperative transmission phase, while only 1 of the messages is repeated by the relay. This indicates that cooperative links (relay to destinations) provided by the relay cannot support multiplexing gain greater than $\frac{1}{2}$. In general, the maximum multiplexing gain supportable by the cooperative links is a function of number of participating relays, i.e. $r = \frac{L}{L+1}$. The cooperative links can support higher multiplexing gain when more relays participate in the cooperative transmission. This observation suggests that the use of multiple relays is able to deliver improvement in terms of rate and reliability. When L is large enough, i.e $\frac{L}{L+1} \approx 1$, the achievable diversity and multiplexing tradeoff of the proposed CBC protocol significantly outperforms the non-cooperative protocol for all r. Furthermore, when L is large, the tradeoff achieved by the proposed CBC protocol approaches the optimal diversity and multiplexing tradeoff of the MISO channel with $M + L$ transmitter antennas, i.e. $d_{MISO}(r) = (M + L)(1 - r)$. In this case, adding a relay has the same effect as adding a physical antenna to the source transmitter. The performance improvement in terms of diversity and multiplexing tradeoff justifies the contribution of the proposed CBC protocol.

In order to better understand the tradeoff offered by the proposed protocol, the diversity and multiplexing tradeoff curves of the non-cooperative ZFBF protocol and the proposed CBC protocol are visualised in Figure 3.3. In this figure, the fixed parameters are $M_t = M_r = M = 4$. Equation

(3.20) and (3.23) are used to plot the diversity and multiplexing tradeoff curves. For the proposed protocol, the number of available relays L is varied from 1 to 8. From Figure 3.3, it can be observed that the maximum diversity gain (when the multiplexing gain is zero) of the proposed protocol is always better than the non-cooperative scheme, even when only one relay is used. The maximum diversity gain of the proposed protocol is a linear function of the number of relays: adding more relays further improves the diversity gain. Generally, the proposed CBC protocol uniformly dominates the non-cooperative protocol, for any number of relays. The tradeoff curve of the CBC protocol is a piecewise linear function of two pieces, with a crossing point between the pieces occurs at $r = \frac{L}{L+1}$. This crossing point indicates the maximum multiplexing gain supportable by the cooperative links, as discussed in the previous paragraph. As the number of relays increases, the crossing point is shifted towards the direction of $r = 1$. This indicates that the cooperative links can support higher multiplexing gain when higher number of relays are used. For instance, using reasonable number of relays, i.e. $L = 2$, the proposed protocol delivers better tradeoff in a wide multiplexing gain range, i.e. $0 \leq r \leq \frac{2}{3}$.

3.4 Numerical Results

In this section, numerical results demonstrate the performance gain of the proposed CBC scheme in comparison to the existing non-cooperative ZFBF scheme. Firstly, the size of the qualified relay set is simulated without high SNR approximation and Lemma 3.5 is verified. Subsequently, three experiments are carried out using Monte-Carlo simulation to compare the outage performance of the CBC protocol with the non-cooperative ZFBF protocol. Throughout the simulations, full spatial multiplexing is assumed, i.e. $M_t = M$ and $M_r = M$. The SNR is defined as $\gamma_0 = \frac{E_s}{\sigma^2}$ where it is assumed $E_s = 1$. The non-cooperative ZFBF scheme used as baseline for comparison is given in (3.10), where the assumption of large number of users is made. The outage probability of this non-cooperative scheme is $P(\mathcal{I}_{ZFBF} < R)$.

Lemma 3.5 states that at high SNR, all relays are qualified to participate in the cooperative transmission, such that $P(K = L) \to 1$ and $P(K \neq L) \to 0$. The simulation here provides the probability when K relays out of L available relays are qualified, $P(K)$, without using the high SNR approximation. In this simulation, (7.9) and (7.24) are used. Refer to Appendix A. The fixed parameters are

the $M = 2$, target data rate $R = 1$ bits/s/Hz and number of available relays $L = 4$. The number of qualified relays K is varied from 0 to 4. Figure 3.4 shows the curves of $P(K)$ versus SNR. At low SNR, i.e. SNR=0dB, the probability that there is no qualifying relay, i.e. $P(K = 0)$, dominates. This is due to the fact that at low SNR, the signals received by the relays are power limited, resulting the decoding of the messages difficult. As the SNR increases, the number of qualified relays increases. When the SNR is larger than 10dB, the probability that the size of the qualified relay set is equal to the size of the available relays, i.e. $P(K = L)$ dominates. At medium SNR, i.e. SNR=15dB, almost all relays are qualified to participate in the cooperative transmission. As SNR increases beyond 15dB, it can be observed that, $P(K = L) \to 1$ and $P(K \neq L) \to 0$. This agrees with the result in Lemma 3.5, stating that at high SNR, all relays are qualified. In the subsequent experiments, it is assumed that the number of qualified relays is equal to the size of the available relays, following the fact that almost all relays are qualified at reasonable SNR, i.e. SNR$\geq$15dB. The terms, "participating relay" are used to reflects that $P(K = L) = 1$.

In the first experiment, the outage probability of the proposed protocol and the non-cooperative ZFBF scheme is plotted as a function of SNR. At the same time, the target data rate R is varied from 1 to 4 bits/s/Hz. The fixed parameters are $M = 2$ and the number of participating relays, $K = 1, 2$. By observing the outage probability plot in Figure 3.5, it is obvious that the outage performance of the proposed protocol is better than the non-cooperative ZFBF scheme at all data rates shown in the figure, even by using only one relay. Recall that the diversity gain characterises the slope of the outage probability curve. The steeper the slope, the higher the diversity gain is. Generally, the result in Figure 3.5 shows that the proposed protocol is able to achieve higher diversity gain when compared to the non-cooperative scheme. This additional diversity gain is due to the use of relays which utilise the cooperative links (relay to destinations) to deliver statistically independent faded replicas of the source messages to the destinations. On the other hand, as the data rate decreases, the outage probability of both schemes improves. This fits the general understanding of the tradeoff between rate and reliability: transmitting with lower data rate enables more reliable communication. In general, the superior outage performance demonstrated by the proposed protocol is a direct result of exploiting the relays to achieve another dimension of diversity gain, which is a function of the number of participating relays. The relationship between the number of participating relays and the outage performance is studied in the next experiment.

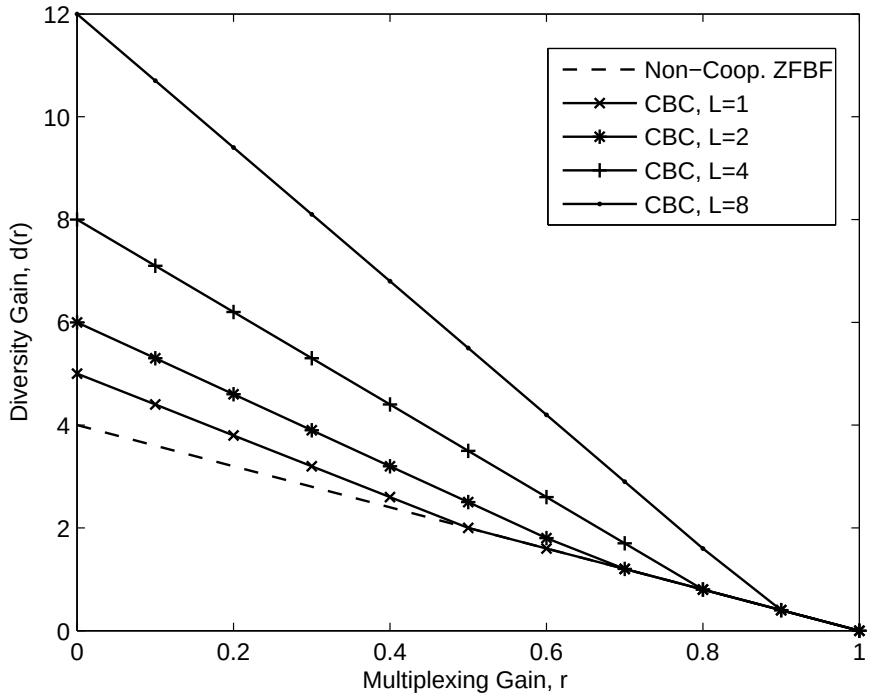


Figure 3.3: Diversity and multiplexing tradeoff of the non-cooperative ZFBF scheme and the proposed CBC scheme with various number of available relays L.

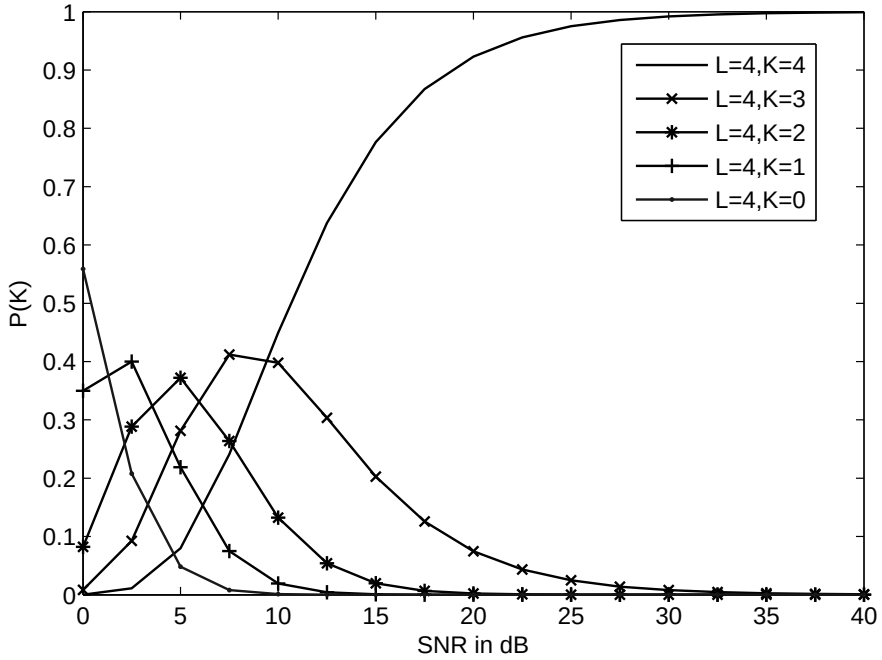


Figure 3.4: Probability that K relays out of $L = 4$ available relays are qualified. Fixed parameters: $M = 2$ and $R = 1$ bits/s/Hz.

In the second experiment, the outage probability versus SNR of the non-cooperative ZFBF scheme and the proposed CBC scheme with different number of participating relays are simulated in order to study the relationship between the number of participating relays and the outage performance. The fixed parameters are $M = 2$ and $R = 4$ bits/s/Hz. The number of participating relays K is varied from 1 to 8. Referring to Figure 3.6, with only one relay, the outage probability of the CBC protocol is smaller than the non-cooperative ZFBF scheme. As the number of participating relays increases, the outage probability of the proposed scheme decreases. This implies that by increasing the number of participating relays, the diversity gain of the proposed scheme is directly increased. The use of multiple relays is beneficial, as it is able to deliver larger performance gain in terms of reliability.

The final experiment studies the relationship between the outage probability and the target data rate of the proposed CBC protocol in comparison with the non-cooperative ZFBF protocol at fixed SNR. The fixed parameters are $M = 2$ and SNR=30dB. Figure 3.7 shows the curves of outage probability versus target data rate, R, at fixed SNR. The outage probability curves of the proposed CBC protocol with different number of participating relays are plotted in the same figure. In general, the proposed protocol delivers performance gain against the non-cooperative protocol, evident from the higher data rate supportable by the proposed protocol at a fixed outage probability. For instance, at 0.01% outage, the proposed protocol can support a data rate up to 5.5 bits/s/Hz by using 2 relays, if compared to the maximum data rate of 4 bits/s/Hz supportable by the non-cooperative scheme. When more relays participate in the cooperation, the proposed protocol can support even higher data rates. From the figure, it can be seen that the outage probability curves of the proposed protocol meet the outage probability curve of the non-cooperative protocol at some points. The locations of these crossing points depend on the number of participating relays. As the number of relays increases, the crossing point shifts towards the direction of higher data rate. Before the crossing point, the proposed protocol delivers performance gain against the non-cooperative protocol. Beyond the crossing point, both the proposed protocol and the non-cooperative protocol have identical performance. The explanation follows. The crossing point can be interpreted as the maximum data rate supportable by the cooperative links formed by the relay/relays. When the target data rate is below the maximum supportable data rate, the cooperative links are active. The opposite happens (cooperative links inactive) when the target data rate is greater than the maximum supportable data rate. The maximum data rate supportable by the cooperative links increases as more relays participate in the cooperation. This is in

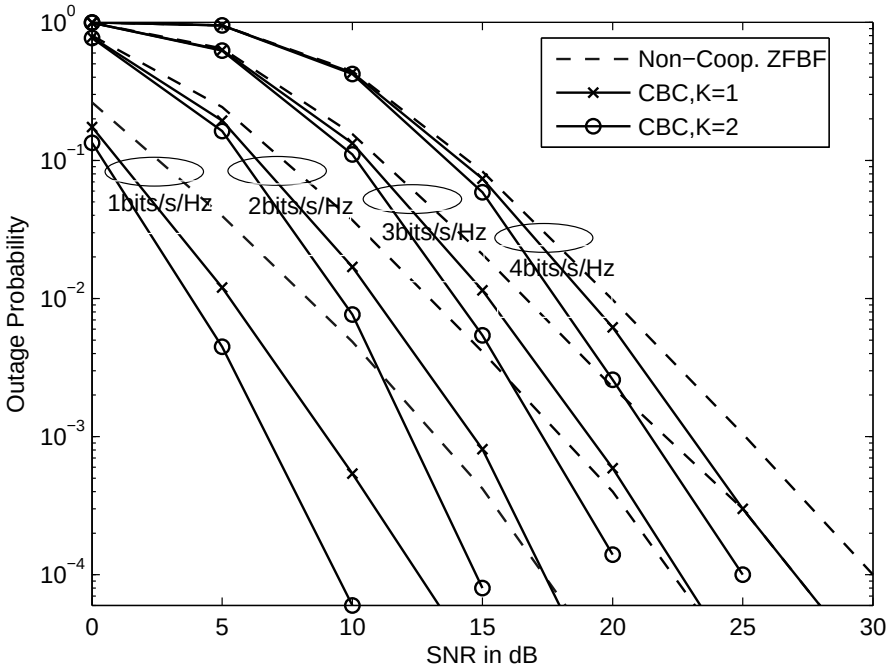


Figure 3.5: Outage probability versus SNR of the non-cooperative ZFBF scheme and the proposed CBC scheme at various target data rate R. Fixed parameters: $M = 2$ and the number of participating relays, $K = 1, 2$.

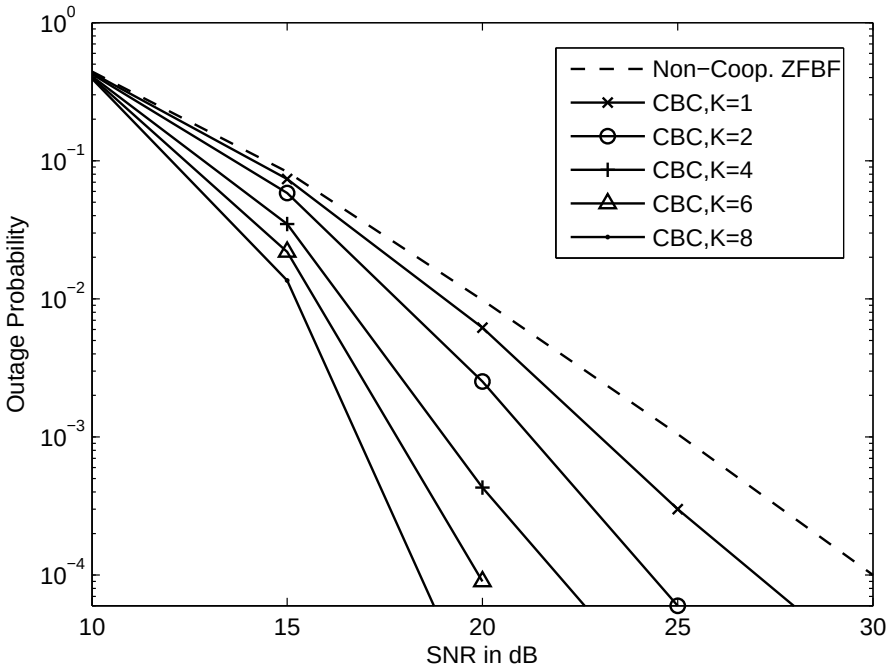


Figure 3.6: Outage probability versus SNR of the non-cooperative ZFBF scheme and the proposed CBC scheme with various number of participating relays K. Fixed parameters: $M = 2$ and $R = 4$ bits/s/Hz.

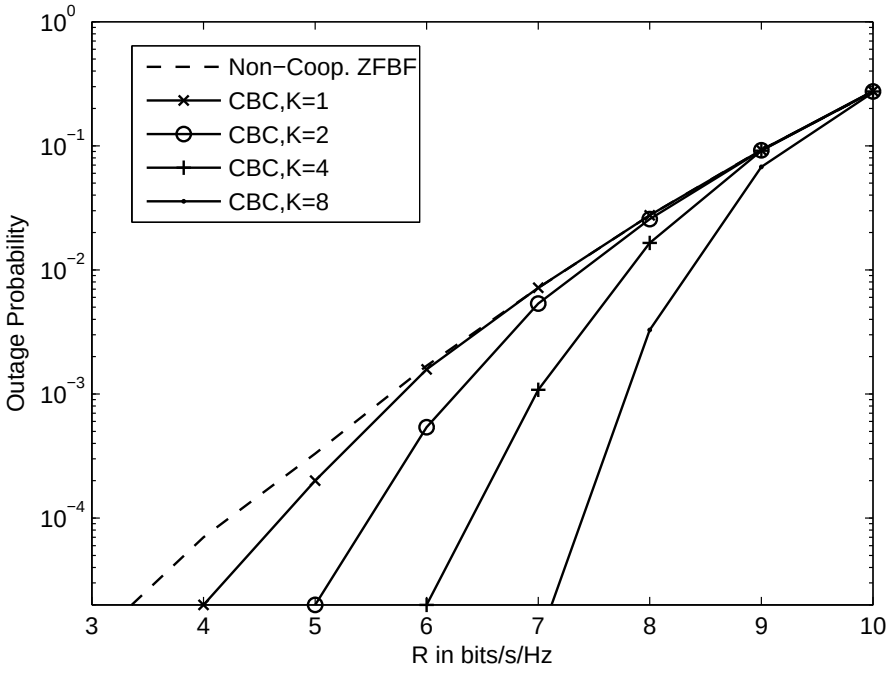


Figure 3.7: Outage probability versus target data rate of the non-cooperative scheme and the proposed CBC scheme, with various number of participating relays K. Fixed parameters: $M = 2$, SNR=30dB.

line with the result in Theorem 3.1. The maximum supportable data rate mentioned here is equivalent to the maximum multiplexing gain (supportable by the cooperative links) described in the discussion of Theorem 3.1. When more relays are used, the proportion of the unrepeated message[4] to the total messages becomes smaller. As a result, a higher data rate can be supported by the cooperative links. In general, the use of multiple relays is able to deliver better rate and reliability tradeoff.

3.5 Chapter Conclusion

In this chapter, a spectrally efficient linear precoded cooperative transmission protocol for the MISO broadcast channel has been proposed. The proposed protocol is able to avoid the co-channel interference among multiple destinations using practical linear precoding and provide another dimension for improvement. The outage behaviour and the diversity and multiplexing tradeoff of the proposed protocol are studied in order to quantify the performance of the proposed protocol. Analytical results show that the maximum diversity gain expressed as the sum of the number of source transmitter antennas and the available relays can be achieved. Analytical results also show that the proposed

[4]Recall that the message in the final time slot of each cooperative transmission phase is not repeated by the relay. Refer to Subsection 3.2.2.3.

protocol uniformly dominates the non-cooperative protocol. When the number of available relays is large, the diversity and multiplexing tradeosff of the proposed protocol completely outperforms the comparable non-cooperative protocol. Monte-Carlo simulations further justify that the proposed protocol achieves better robustness than the comparable scheme. Both the analytical and simulation results agree that the use of multiple relays is beneficial in terms of improving the tradeoff between the data rate and reliability.

Chapter 4

Analogue Network Coding For the Multi-Pair, Two-Way Relaying Channel

A two-way information exchange channel is an important scenario occuring in many practical wireless networks. For instance, in a cellular network there are downlink (channel from base station to mobile user) and uplink (channel from mobile user to base station) channels to enable information to flow in both directions. Similar examples can be found in ad-hoc networks, where nodes exchange information in both directions. This chapter studies a two-way information exchange scenario where multiple pairs of users exchange information within a pair, with the help of a dedicated multi-antenna relay. The protocol integrates the idea of analogue network coding in mixing two data streams originating from the same user pair, together with the spatial multiplexing of data streams originating from different user pairs. The key feature of the protocol is that it enables both the relay and the users to participate in interference cancellation. In this chapter, several beamforming schemes for the multi-antenna relay are proposed. The performance of the proposed protocol under several beamforming schemes are evaluated using information theoretical metrics such as ergodic capacity, outage probability and the diversity and multiplexing tradeoff. Analytical and simulation results reveal that the ergodic capacity, outage probability and the diversity and multiplexing tradeoff of the proposed beamforming schemes outperform comparable techniques.

4.1 Chapter Introduction

Network coding [34] is a powerful technique which allows intermediate nodes to mix the signals from multiple links in order to enhance the network throughput. The application of network coding in wireless networks has received growing interest in recent years. Based on the idea of network coding, two-way relaying in wireless networks is able to enhance the network throughput by reducing the channel resources used in the information exchange between nodes. Two-way relaying schemes such as the strategies summarised in [4] based on decode-and-forward (DF) relaying, analogue network coding [5] based on amplify-and-forward (AF) relaying, physical network coding [6] based on estimate-and-forward (EF) relaying, etc are able to complete the two way information passing in only two phases. Since the total number of channel uses is halved, the number of independent data streams that can be transmitted or received simultaneously per channel use in the network is doubled, as compared to the time division protocol which requires four orthogonal channel uses. Refer to Figure 2.6 and Figure 2.7.

Attracted by the benefits of multiple antennas in enhancing the system capacity and reliability, two-way relaying has been generalised to the multi-antenna case. The authors in [56] generalise the DF based two-way relaying to a multi-antenna setting using the classical multiple access capacity region and devise an optimal broadcast strategy based on point-to-point multiple-input and multiple-output (MIMO) links. On the other hand, [12] proposes an AF based protocol which uses a zero-forcing beamformer to eliminate the co-channel interference between users. However, no network coding is used to mix the data streams from two users. The researchers in [57] and [13] consider the sum rate optimising AF based beamforming design, for the case where only the relay is equipped with multi-antenna. Based on analogue network coding, the beamforming schemes proposed in [57] and [13] are able to deliver sum rate improvement when compared with [12].

The extension of the two-way relaying to multiple user pairs introduces the problem of multi-user interference. The authors in [7] propose a scheme for a code division multiple access (CDMA) system where each pair of users shares a common spreading code as a mean to reduce the multi-user interference. The demodulate-and-forward based scheme proposed in [7] uses a multi-user receiver which requires high computational complexity at the relay. The suboptimal AF based scheme proposed in [7] suffers from poor error performance when the number of users is low, due to high noise level.

The researchers in [8] propose a scheme using DF based relay with multiple antennas for narrow band system. Precoding based on X-OR is used at the relay to encode the messages from the same user pair while block-diagonalisation [11] is used to mitigate the interference caused by multiple user pairs. Since the DF based scheme requires higher complexity for decoding and encoding at the relay as compared to the AF based scheme, [9] and [10] consider the multi-pair scenario with an AF based multi-antenna relay. Similar to [8], the schemes in [9] and [10] use the idea of block-diagonalisation [11] to eliminate the multi-pair interfere

nce and to forward the mixture containing the desired message and the self-interference to the each user. Each user then uses knowledge of the previously transmitted message to subtract the self-interference from the mixture to decode the new message. The schemes in [9] and [10] have shown that a higher sum rate can be achieved in comparison with the conventional multi-user zero-forcing scheme. Nevertheless, the potential benefit of block-diagonalisation in the scenario where the relay does not have enough degrees of freedom to spatially separate and/or decode each independent message is not covered in [8, 9, 10]. Furthermore, the spatial diversity gain offered by block-diagonalisation in the multi-pair scenario has not been studied.

This chapter studies a scenario where multiple pairs of users exchange information within pair, with the help of a dedicated AF based, multi-antenna relay. The transmission protocol employed in this chapter utilises the principal concept of network coding in mixing two data streams originating from the same user pair, coupled with the spatial multiplexing of the data streams originating from different user pairs. The key feature of the protocol is that it enables both the relay and the users to participate in interference cancellation: the relay eliminates co-channel interference due to multi-pair while each user eliminates the self-interference. Several low complexity beamforming schemes are proposed based on the idea of block-diagonalisation. The performance is evaluated using information theoretical metrics, such as ergodic capacity, outage probability and the diversity and multiplexing tradeoff. Two cases have been considered in this chapter. First, the case where the number of antennas at the relay is less than the total single-antenna users. Second, the case where the number of antennas at the relay is at least the total number of single-antenna users. In the first case, simulation results show that the proposed beamforming scheme is able to deliver significant ergodic capacity improvement and higher multiplexing gain if compared with existing schemes based on time sharing between pairs. Upper and lower bounds on ergodic capacity achieved by the proposed scheme are derived to quantify

the performance. In the second case, it is shown that appropriate selection or coherent combining of null-space vectors is able to achieve all the available diversity gain offered by block-diagonalisation. The proposed beamforming with coherent combining of null-space vectors achieves the highest ergodic capacity and the lowest outage probability among all comparable schemes, while the proposed beamforming with null-space vector selection performs close to the former. The proposed beamforming schemes deliver higher diversity gain as compared with existing zero-forcing scheme [12] while several comparable schemes based on block-diagonalisation [9, 10] fail to do so due to non-coherent combining of the diversity streams. The proposed schemes also offer better rate and reliable tradeoff as compared to the zero-forcing scheme [12] and the comparable schemes based on time sharing between pairs [57]. Analytical results on outage probability and the diversity and multiplexing tradeoff are derived to quantify the achievable performance.

This chapter is organised as follows. The system model and protocol are described in Section 4.2, while the proposed beamforming schemes are discussed in Section 4.3. The analytical results on the ergodic capacity are presented in Section 4.4, followed by the analytical results on outage probability and the diversity and multiplexing tradeoff in Section 4.5. In Section 4.6, the numerical results on ergodic capacity and outage probability are discussed. Section 4.7 concludes this chapter.

4.2 System Model and Protocol Description

Consider a scenario where there are M pairs of single-antenna users who wish to exchange information with their partners, with the help of a dedicated AF based relay equipped with N antennas. An example of the scenario where $M = 2$ and $N = 3$ is shown in Figure 4.1. The symmetric case is assumed, where all nodes are subjected to unit power constraint and have the same channel statistics. All channels undergo i.i.d. quasi-static Rayleigh fading and channel reciprocity is assumed. The receiver is corrupted by circularly symmetric additive white Gaussian noise with distribution $\mathcal{CN} \sim (0, \sigma^2)$. The half duplex constraint is assumed throughout the chapter and it is realised using time division duplexing. Every user knows his and his partner's effective user-to-relay channel state information (CSI) while the relay has the CSI of all user-to-relay links.

The notation (i, j) is used to represent the pair of user i and user j who exchange information with each other, such that the mth user pair is denoted as $(2m-1, 2m)$. The channel from user i to the relay

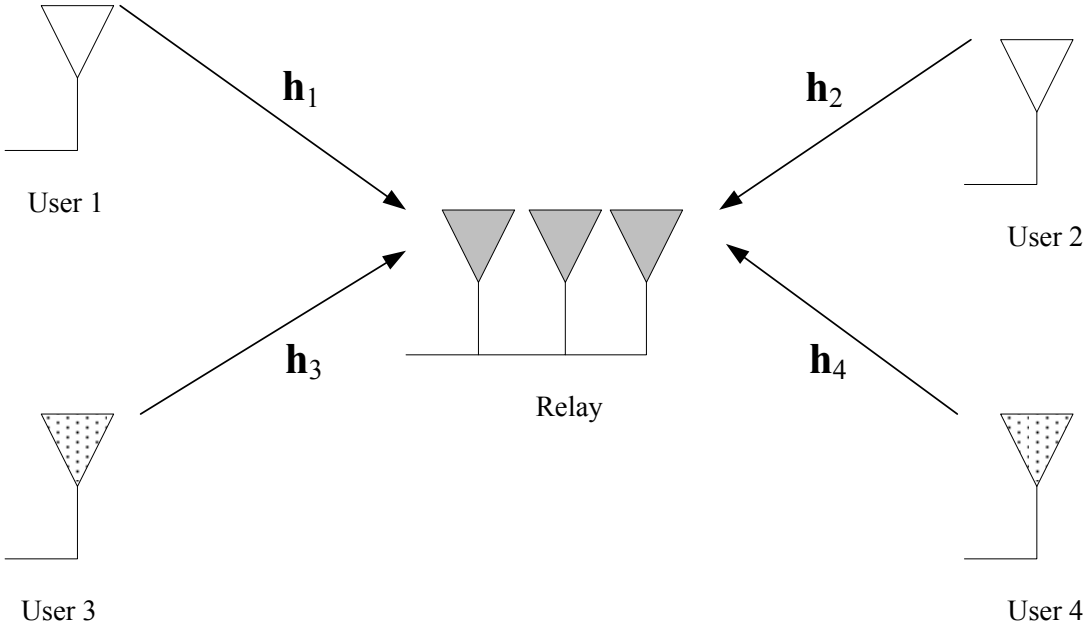


Figure 4.1: Scenario when two pairs of single-antenna users exchanging information with the help of a multi-antenna relay.

is $\mathbf{h}_i \in \mathbb{C}^{N\times 1}$, the message transmitted from user i is $x_i \in \mathbb{C}$, the combined channels of user pair (i,j) is $\mathbf{H}_{i,j} = [\ \mathbf{h}_i \quad \mathbf{h}_j\ ] \in \mathbb{C}^{N\times 2}$ and the message vector of user pair (i,j) is $\mathbf{x}_{i,j} = [\ x_i \quad x_j\ ]^T \in \mathbb{C}^{2\times 1}$. The noise observed by the relay and user i is $\mathbf{n} \in \mathbb{C}^{N\times 1}$ and $n_i \in \mathbb{C}$ respectively. Define the multi-pair interference[1] channel seen by user pair (i,j) as $\tilde{\mathbf{H}}_{i,j} \in \mathbb{C}^{N\times 2(M-1)}$ by stacking all user channels other than $\mathbf{H}_{i,j}$. Similarly, the message vector conveyed through $\tilde{\mathbf{H}}_{i,j}$ is defined as $\tilde{\mathbf{x}}_{i,j} \in \mathbb{C}^{2(M-1)\times 1}$ by stacking all messages other than $\mathbf{x}_{i,j}$.

The employed protocol combines the application of analogue network coding within one user pair and spatial multiplexing between user pairs, thus it is named as the network coding with spatial multiplexing (NC-SM) protocol. The NC-SM protocol can be described in two time slots. In the first time slot, M pairs of users transmit simultaneously in the same channel with unit power. The relay observes a mixture of all messages from the users, which can be expressed as

$$\mathbf{r} = \sum_{m=1,}^{M} \mathbf{H}_{i,j}\mathbf{x}_{i,j} + \mathbf{n}, \tag{4.1}$$

where $i = 2m-1,\ j = 2m$. In the second time slot, the AF based relay broadcasts the linearly processed observation, i.e. $\mathbf{Fr}$, where $\mathbf{F} \in \mathbb{C}^{N\times N}$ is the beamforming matrix at the relay. The signal

[1]The multi-pair interference is different from the multi-user interference defined in the literature. The multi-user interference is the interference observed by each user, not by each pair.

received by user i can be expressed as

$$y_i = \mathbf{h}_i^T \mathbf{F} \mathbf{H}_{i,j} \mathbf{x}_{i,j} + \mathbf{h}_i^T \mathbf{F} \tilde{\mathbf{H}}_{i,j} \tilde{\mathbf{x}}_{i,j} + \mathbf{h}_i^T \mathbf{F} \mathbf{n} + n_i. \tag{4.2}$$

The first term on the right hand side (RHS) of the equation contains the mixture of messages from user pair (i, j), the second term contains the multi-pair interference while the last two terms contain the relay propagated noise and the receiver noise of user i. The unique feature of the NC-SM protocol is to allow both the relay and the users to participate in the interference cancellation. The relay eliminates the multi-pair interference, $\mathbf{h}_i^T \mathbf{F} \tilde{\mathbf{H}}_{i,j} \tilde{\mathbf{x}}_{i,j}$, while user i removes the self-interference, $\mathbf{h}_i^T \mathbf{F} \mathbf{h}_i x_i$. The design of beamforming matrix $\mathbf{F}$ is discussed in the following section.

4.3 Joint Receive and Transmit Beamforming Design

In this section, a low complexity beamforming design at the relay is presented in two cases. Case I: $N = 2M - 1$, which corresponds to the case when the number of antennas at the relay is less than the total number of users; and case II: $N \geq 2M$, which corresponds to the case when the number of antennas at the relay is at least the total number of users. Subsection 7.2 explains case I while Subsection 7.2 discusses case II. Note that the proposed protocol does not operate when $N < 2M-1$, due to the zero-forcing criterion, which is discussed in detail in the following subsections. In addition, a brief review of the comparable beamforming schemes is provided in Subsection 4.3.3 to facilitate the comparison between the proposed schemes and the existing schemes.

4.3.1 Case I: $N = 2M - 1$

This case corresponds to the situation where conventional multi-antenna receiver or transmitter is not able to spatially support $2M$ independent data streams, due to the limitation of the available degrees of freedom [22], i.e. $\min(N, 2M)$. The scheme proposed in [8] does not work under this case due to insufficient number of antennas at the relay, while [9] and [10] have not explored this specific setting. It will be shown in the following paragraph that the proposed beamforming structure is able to support $2M$ independent data streams (from $2M$ users) simultaneously, given only $N = 2M - 1$, by aligning the data streams of each user pair to occupy only 1 spatial dimension. As a result, a higher

multiplexing gain can be achieved.

The proposed beamforming matrix $\mathbf{F}$ consists of the receive beamforming matrix $\mathbf{W}_R$ and transmit beamforming matrix $\mathbf{W}_T$, which are directly cascaded as follows,

$$\mathbf{F} = \mathbf{W}_T \mathbf{A} \mathbf{W}_R, \tag{4.3}$$

where the receive beamforming matrix $\mathbf{W}_R \in \mathbb{C}^{M \times N}$ and the transmit beamforming matrix $\mathbf{W}_T \in \mathbb{C}^{N \times M}$ while the diagonal matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$ is the power allocation matrix. Due to channel reciprocity, $\mathbf{W}_T = \mathbf{W}_R^T$. This allows us to concentrate on the design of the transmit beamforming matrix. For simplicity, the subscript of $\mathbf{W}_T$ is omitted, i.e. $\mathbf{W}_T = \mathbf{W}$. Represent $\mathbf{W} = \begin{bmatrix} \mathbf{w}_{1,2} & \dots & \mathbf{w}_{2M-1,2M} \end{bmatrix}$ where $\mathbf{w}_{i,j} \in \mathbb{C}^{N \times 1}$ is the transmit beamforming vector for user pair (i, j), the diagonal matrix $\mathbf{A} = \text{diag} \begin{pmatrix} \alpha_{1,2} & \dots & \alpha_{2M-1,2M} \end{pmatrix}$, and $\mathbf{F}_{i,j} \in \mathbb{C}^{N \times N} = \mathbf{w}_{i,j} \mathbf{w}_{i,j}^T$ as the effective beamforming matrix for pair (i, j). The equation in (4.3) can be rewritten as $\mathbf{F} = \sum_{m=1}^{M} \alpha_{i,j} \mathbf{F}_{i,j}$ where $i = 2m - 1,\ j = 2m$. The design objective of $\mathbf{w}_{i,j}$ is to ensure that each user pair is free from the multi-pair interference. In other words, the zero-forcing criterion $\tilde{\mathbf{H}}_{i,j}^T \mathbf{w}_{i,j} = \mathbf{0}$ has to be satisfied for all pairs (i, j), where $\mathbf{0}$ is a column vector of all zeros. This criterion coincides with the block-diagonalisation[2] proposed for the MIMO broadcast channels in [11]. To satisfy this criterion, vector $\mathbf{w}_{i,j}$ is designed to lie in the null-space of the multi-pair interference channel, i.e. $\mathbf{w}_{i,j} = \text{null}(\tilde{\mathbf{H}}_{i,j}^T)$, which exists as a non-zero vector when $N = 2M - 1$. Note that $\text{rank}(\mathbf{H}_{i,j}^T \mathbf{w}_{i,j}) = 1$, i.e. each user pair only occupies one spatial dimension. This enables the relay to spatially multiplex $2M$ independent streams by using only $N = 2M - 1$ antennas.

The transmission from the relay is subject to a unit power constraint. The power constraint can be expressed as

$$\sum_{m=1}^{M} \alpha_{i,j}^2 \left(||\mathbf{F}_{i,j} \mathbf{H}_{i,j}||_F^2 + \sigma^2 ||\mathbf{F}_{i,j}||_F^2 \right) \ \leq \ 1, \tag{4.4}$$

where $i = 2m - 1,\ j = 2m$. Note that the expected value, $E[\mathbf{n}\mathbf{n}^H] = \sigma^2 \mathbf{I}$. Since high SNR performance, i.e. diversity gain and multiplexing gain, is of interest, equal power allocation across

[2]Each user pair in the multi-pair scenario considered here is analogous to each multi-antenna user in the MIMO broadcast channels considered in [11]. However, different from the MIMO broadcast channels, the users in each pair are not able to cooperate with each other, i.e. linear postprocessing within user pair is not possible. Hence, the block-diagonalisation proposed in [11] cannot be directly applied in the multi-pair scenario.

M data streams from M user pairs is sufficient. Although optimal power allocation among user pairs is able to further improve the sum rate performance at finite SNR, it improves neither the diversity gain nor the multiplexing gain at infinite SNR. This is because when SNR goes to infinity, equal power allocation is asymptotically optimal [18]. Using equal power allocation, the equation above is satisfied in equality by choosing $\alpha_{i,j} = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{||\mathbf{F}_{i,j}\mathbf{H}_{i,j}||_F^2 + \sigma^2}}$. Note that $||\mathbf{F}_{i,j}||_F^2 = ||\mathbf{w}_{i,j}\mathbf{w}_{i,j}^T||_F^2 = 1$. The signal received by user i can be expressed as

$$y_i = \alpha_{i,j}\mathbf{h}_i^T\mathbf{F}_{i,j}\left(\mathbf{h}_i x_i + \mathbf{h}_j x_j + \mathbf{n}\right) + n_i, \tag{4.5}$$

while the signal received by user j is

$$y_j = \alpha_{i,j}\mathbf{h}_j^T\mathbf{F}_{i,j}\left(\mathbf{h}_i x_i + \mathbf{h}_j x_j + \mathbf{n}\right) + n_j. \tag{4.6}$$

Note that for all $p \neq i$ and $q \neq j$, we have $\mathbf{h}_i^T\mathbf{F}_{p,q}\mathbf{h}_i = 0$, $\mathbf{h}_i^T\mathbf{F}_{p,q}\mathbf{h}_j = 0$, $\mathbf{h}_j^T\mathbf{F}_{p,q}\mathbf{h}_i = 0$ and $\mathbf{h}_j^T\mathbf{F}_{p,q}\mathbf{h}_j = 0$. Since user i has the knowledge of x_i, and the knowledge of the effective channels, $\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_i$ and $\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_j$, user i can decode the desired message x_j by subtracting the self-interference $\alpha_{i,j}\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_i x_i$ from the received mixture. A similar strategy is used by user j to decode the desired message x_i. Notice that the effective channels are scalars. The effective scalar channels carrying self-interference, $\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_i$ and $\mathbf{h}_j^T\mathbf{F}_{i,j}\mathbf{h}_j$ can be fed back from the relay to user i and j respectively using orthogonal feedback channels, while the effective scalar channel carrying desired message, $\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_j$ can be fed back from the relay to user pair (i, j) simultaneously using a common feedback channel, with low overhead. Note that $\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_j = \mathbf{h}_j^T\mathbf{F}_{i,j}\mathbf{h}_i$. The user pairs (i, j) do not need to know the exact channel vectors $\mathbf{h}_i$ and $\mathbf{h}_j$.

Assuming Gaussian channel coding, the mutual information of user i can be described as

$$\mathcal{I}_i = \frac{1}{2}\log_2\left(1 + \frac{\alpha_{i,j}^2|\mathbf{h}_i^T\mathbf{F}_{i,j}\mathbf{h}_j|^2}{\sigma^2\left(\alpha_{i,j}^2||\mathbf{h}_i^T\mathbf{F}_{i,j}||^2 + 1\right)}\right). \tag{4.7}$$

The pre-log factor of 1/2 reflects the two time slots used to complete the information exchange. The mutual information of user j, $\mathcal{I}_j$ can be obtained by interchanging $\mathbf{h}_i$ and $\mathbf{h}_j$ in (4.7). Define a floating point operation as one complex multiplication or addition. The number of floating point operations

(flops) is used to measure the computational complexity. It can be easily shown that the computational complexity of the proposed scheme is at most $\mathcal{O}(MN^3)$ flops. See Appendix C. Generally, this beamforming scheme performs well when both channels $\mathbf{h}_i$ and $\mathbf{h}_j$ have good channel gain[3].

4.3.2 Case II: $N \geq 2M$

Similar to the previous case, the design objective of the beamforming matrix $\mathbf{F}$ is to ensure that multi-pair interference is nullified. However, unlike the previous case, the dimension of $\mathrm{null}(\tilde{\mathbf{H}}_{i,j}^T)$ is greater than one, indicating that the null space consists of multiple vectors. The transmit beamforming matrix for pair (i, j) is now $\mathbf{W}_{i,j} \in \mathbb{C}^{N\times(N-2(M-1))}$, where $\mathbf{W}_{i,j} = \mathrm{null}(\tilde{\mathbf{H}}_{i,j}^T)$. Multiple null-space vectors are able to improve the diversity gain by providing multiple statistically independent paths for the messages to travel through. In order to benefit from the additional diversity gain, the beamforming structure need to be carefully designed. A trivial choice of directly cascading the receive and transmit beamforming matrices as in (4.3), destroys the diversity gain offered by multiple null-space vectors. This is due to the fact that the superposition of multiple diversity streams can either add up constructively or destructively at the destinations. It will be shown that appropriate selection or coherent combining of null-space vectors is important to achieve the available diversity gain offered by block-diagonalisation based beamforming in comparison with the zero-forcing scheme [12]. The block-diagonalisation with singular value decomposition (BD-SVD) [10] and pair-aware matched filter (PA-MF) [9] fail to achieve the available diversity gain because the diversity streams are not coherently combined. In this subsection, two beamforming schemes which are able to achieve all diversity gain offered by block-diagonalisation are proposed. One is based on null-space vector selection and the other is based on coherent combining of all null-space vectors.

4.3.2.1 Null-Space Vector Selection

This subsection proposes a beamformer with null-space vector selection. The relay performs selection to determine the null-space vector that can deliver the best performance in maximising the sum rate of each user pair (i, j). The overall beamforming structure $\mathbf{F}$ is similar to (4.3). Since high SNR

[3]The performance degrades when either one of the channels is in deep fade. This can be solved by increasing the diversity gain, achieved by either using more antennas at the relay as discussed in Subsection 7.2, or implementing user selection. When the number of users in the network is large, the relay can select the user pairs with good channel quality at a particular time instance to become active users in order to harvest the benefit of multi-user diversity.

performance is of interest, equal power allocation among user pairs is sufficient. Denote the kth null-space vector for pair (i, j), obtained from the kth column of $\mathbf{W}_{i,j}$, as $\mathbf{w}_{i,j}(k)$. The null-space vector selection criterion for user pair (i, j) can be expressed as

$$\arg \max_{k=1,\ldots,N-2(M-1)} \mathcal{I}_i(k) + \mathcal{I}_j(k), \tag{4.8}$$

where $\mathcal{I}_i(k)$ and $\mathcal{I}_j(k)$ are the mutual information of user i and user j respectively, when $\mathbf{w}_{i,j}(k)$ is used. The value of $\mathcal{I}_i(k)$ can be obtained by replacing $\mathbf{F}_{i,j} = \mathbf{w}_{i,j}(k)\mathbf{w}_{i,j}^T(k)$ in (4.7) while $\mathcal{I}_j(k)$ can be derived similarly. The best null-space vector, denoted as $\mathbf{w}_{i,j}(k_{best})$ is able to maximise the sum rate of user pair (i, j) and is used as the receive and transmit beamforming vectors for pair (i, j). The sum rate of user pair (i, j) is used instead of the individual rate because the best null-space vector, $\mathbf{w}_{i,j}(k_{best})$ affects both user i and user j simultaneously. In other words, each beam carries the mixture of the messages of user pair (i, j). The received signal and the mutual information of user i can be expressed in similar way as (4.5) and (4.7) by substituting $\mathbf{F}_{i,j} = \mathbf{w}_{i,j}(k_{best})\mathbf{w}_{i,j}(k_{best})^T$. The worst case computational complexity of this selection scheme is $\mathcal{O}(MN^3)$ flops, which is the same as the previous scheme. See Appendix C. This selection scheme serves as a lower bound for the derivation of the outage probability and the diversity and multiplexing tradeoff, which are discussed in Section 4.5 .

4.3.2.2 Coherent Combining of Null-Space Vectors

In contrast to the null-space vector selection scheme, the beamformer proposed in this subsection utilises all the available null-space vectors. In order to guarantee that the superposition of multiple diversity streams at the target destination is constructive, the following beamforming structure is proposed,

$$\mathbf{F} = \sum_{m=1}^{M} \mathbf{W}_{i,j}\mathbf{B}_{i,j}\mathbf{A}_{i,j}\mathbf{P}_{\pi}\mathbf{B}_{i,j}^T\mathbf{W}_{i,j}^T, \tag{4.9}$$

where $i = 2m - 1,\ \ j = 2m$. The matrix $\mathbf{W}_{i,j} \in \mathbb{C}^{N\times(N-2(M-1))}$ is the transmit beamforming matrix for pair (i, j), the matrix $\mathbf{B}_{i,j} \in \mathbb{C}^{(N-2(M-1))\times 2}$ is the channel matching matrix for pair (i, j), the diagonal matrix $\mathbf{A}_{i,j} \in \mathbb{R}^{2\times 2}$ is the power allocation matrix for user pair (i, j) while the matrix $\mathbf{P}_{\pi} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is the permutation matrix.

The channel matching matrix for pair (i,j) is designed as $\mathbf{B}_{i,j} = \mathbf{W}_{i,j}^H \mathbf{H}_{i,j}^*$. The effective relay to user (i,j) channel can be expressed as $\mathbf{H}_{i,j}^T \mathbf{W}_{i,j} \mathbf{B}_{i,j} = \begin{bmatrix} \Phi_i & \Psi \\ \Psi^* & \Phi_j \end{bmatrix}$, where the diagonal elements $\Phi_i = \sum_{k=1}^{N-2(M-1)} \left|\mathbf{h}_i^T \mathbf{w}(k)\right|^2$, $\Phi_j = \sum_{k=1}^{N-2(M-1)} \left|\mathbf{h}_j^T \mathbf{w}(k)\right|^2$ and the off-diagonal element $\Psi = \sum_{k=1}^{N-2(M-1)} \mathbf{h}_i^T \mathbf{w}(k) \mathbf{w}^H(k) \mathbf{h}_j^*$. The channel matching matrix $\mathbf{B}_{i,j}$ ensures that main diagonal elements Φ_i and Φ_j contain the coherently combined (at zero phase) diversity streams of user i and user j respectively, while the off-diagonal element, Ψ, contains the non-coherent superposition of the correlated streams for user i and j. The permutation matrix $\mathbf{P}_\pi$ plays an important role to ensure that diversity gain (contributed by the coherently combined diversity streams) is preserved when the transmit beamforming matrix $\mathbf{W}_{i,j} \mathbf{B}_{i,j} \mathbf{A}_{i,j}$ and receive beamforming matrix $\mathbf{B}_{i,j}^T \mathbf{W}_{i,j}^T$ are cascaded.

The relay is subjected to unit power constraint. Since high SNR performance is of interest, equal power allocation among users is sufficient. Under equal power allocation, the power allocation matrix $\mathbf{A}_{i,j} = \alpha_{i,j} \mathbf{I}$. Denote $\mathbf{F}_{i,j} \in \mathbb{C}^{N \times N} = \mathbf{W}_{i,j} \mathbf{B}_{i,j} \mathbf{P}_\pi \mathbf{B}_{i,j}^T \mathbf{W}_{i,j}^T$ such that $\mathbf{F} = \sum_{m=1}^{M} \alpha_{i,j} \mathbf{F}_{i,j}$ where $i = 2M-1$ and $j = 2M$. The power constraint can be expressed similarly as in (4.4). The power constraint is satisfied in equality by choosing $\alpha_{i,j} = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{||\mathbf{F}_{i,j}\mathbf{H}_{i,j}||_F^2 + \sigma^2 ||\mathbf{F}_{i,j}||_F^2}}$. The signal received by user i can be written in the same way as in (4.5). User i is able to decode the desired message x_j by subtracting the self-interference from the observation. Expand the effective channel carrying the desired message of user i,

$$\alpha_{i,j} \mathbf{h}_i^T \mathbf{F}_{i,j} \mathbf{h}_j \;=\; \alpha_{i,j} \Phi_i \Phi_j + \alpha_{i,j} |\Psi|^2. \tag{4.10}$$

The first term on the RHS of (4.10) contains the multiplication of the coherently combined diversity streams of user i, i.e. Φ_i, and the coherently combined diversity streams of user j, i.e. Φ_j, while the last term contains the magnitude square of the non-coherent combination of the correlated streams of user i and user j, i.e. Ψ. Recall that the diversity gain is obtained when statistically independent (uncorrelated) streams are used. Hence, only the first term in (4.10) contributes to the diversity gain. The correlated streams in the last term of (4.10) are allowed to combine non-coherently as it does not contribute to the diversity gain. Note that the last term does not affect the diversity gain contributed by the first term, as it has zero phase.

The mutual information of user i can be written in similar form as in (4.7). The received signal and the mutual information of user j can be derived easily. It is worth mentioning that the beamforming

structure in (4.9) reduces the overhead needed to feed back the CSI of the effective scalar channel carrying the desired message of user i and j. Specifically, user i and user j only need to know the magnitude of scalar $|\mathbf{h}_i^T \mathbf{F}_{i,j} \mathbf{h}_j|$, which can be fed back from the relay using a common feedback channel with only half the amount of overhead needed in case I and case II with null-space vector selection. The computational complexity of this coherent combining scheme is at most $\mathcal{O}(MN^3)$ flops, which is no worse than the selection scheme discussed previously. See Appendix C.

4.3.3 Comparable Schemes

Several existing schemes with AF based relay are used for comparison, namely the pure AF scheme, maximal ratio reception and transmission (MRR-MRT) [57], zero-forcing [12] and the block-diagonalisation based schemes [9, 10]. To enable fair comparison, a relay with N antennas assumes a unit power constraint with equal power allocation among users. In general, the mutual information of the comparable schemes can be written in similar form as in (4.7), where the overall beamforming matrix for pair (i, j), $\mathbf{F}_{i,j}$, the power normalisation factor $\alpha_{i,j}$ and the pre-log factor β depend on the specific scheme. The comparable schemes are described in the following.

4.3.3.1 Pure AF Scheme

In the pure AF scheme, the relay forwards the power normalised observation to the target node pair without any linear processing at the relay. Since no beamforming is performed, the weighting matrix is just an identity matrix, i.e. $\mathbf{F}_{i,j} = \mathbf{I}$. This scheme is extended to multi-pair using time sharing between pairs. The power normalisation factor is $\alpha_{i,j} = \frac{1}{\sqrt{||\mathbf{H}_{i,j}||_F^2 + N\sigma^2}}$ while the pre-log is set as $\beta = \frac{1}{2M}$ to reflect the total number of time slots used to complete the information exchange of M pairs.

4.3.3.2 MRR-MRT Scheme

The MRR-MRT scheme [57] is used for comparison because it is able to achieve near optimal sum rate in the single user pair scenario. In this scheme, the relay broadcasts a channel matched and power normalised observation to the target destinations. The weighting matrix of user pair (i, j) is chosen to match the forward and the backward channels, such that $\mathbf{F}_{i,j} = \mathbf{H}_{i,j}^* \mathbf{P}_\pi \mathbf{H}_{i,j}^H$ where $\mathbf{P}_\pi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

This scheme is extended to multi-pair using time sharing between pairs. The power normalisation factor is $\alpha_{i,j} = \frac{1}{\sqrt{||\mathbf{F}_{i,j}\mathbf{H}_{i,j}||_F^2+\sigma^2||\mathbf{F}_{i,j}||_F^2}}$ while the pre-log $\beta = \frac{1}{2M}$ accounts for the total number of time slots used.

4.3.3.3 Zero-Forcing Scheme

In the zero-forcing scheme [12], the relay uses conventional multi-user zero-forcing to separate the received mixture from each user pair into two orthogonal data streams and uses zero-forcing to direct the orthogonal data streams to the intended destinations. The relay eliminates all the co-channel interference and each user receives only the desired message from his partner. The weighting matrix is chosen to be $\mathbf{F}_{i,j} = \left(\mathbf{H}_{i,j}^T\right)^\dagger \mathbf{P}_\pi \mathbf{H}_{i,j}^\dagger$, so that the channels are completely diagonalised. The permutation matrix $\mathbf{P}_\pi$ acts as data switching matrix. This scheme is different from the proposed NC-SM schemes because it separates the data streams originating from the same user pair while the proposed NC-SM schemes preserve the mixture from the same user pair and orthogonalise only the data streams originating from different user pairs. Two cases are considered under the existing zero-forcing scheme. In the first case, when $2 \leq N < 2M$, the zero-forcing scheme is extended to multiple pairs using time sharing between pairs because the relay does not have sufficient antennas to diagonalise the channels. The power normalisation factor is $\alpha_{i,j} = \frac{1}{\sqrt{||\mathbf{F}_{i,j}\mathbf{H}_{i,j}||_F^2+\sigma^2||\mathbf{F}_{i,j}||_F^2}}$ while the pre-log factor for this case is $\beta = \frac{1}{2M}$. The possible combination of time sharing and spatial multiplexing among multi-pair is not considered because it leads to a scheduling problem and complicates the study. In the second case, when $N \geq 2M$, the zero-forcing scheme has enough degrees of freedom to spatially support M pairs of users simultaneously. In this case, the overall beamforming matrix $\mathbf{F} = \left(\mathbf{H}^T\right)^\dagger \mathbf{D}_\pi \mathbf{H}^\dagger$ where $\mathbf{H} \in \mathbb{C}^{N\times 2M}$ is the combined channels of all $2M$ users and block-diagonal matrix $\mathbf{D}_\pi = \mathrm{diag}(\ \mathbf{P}_\pi \ \ \ldots \ \ \mathbf{P}_\pi\ ) \in \mathbb{C}^{2M\times 2M}$. The power normalisation factor becomes $\alpha = \frac{1}{\sqrt{||\mathbf{FH}||_F^2+\sigma^2||\mathbf{F}||_F^2}}$ while the pre-log becomes $\beta = \frac{1}{2}$ to indicate that only two time slots are used. Although both the zero-forcing scheme and the proposed schemes support spatial multiplexing, it will be shown in the numerical results section that the proposed schemes outperform the former, in both the ergodic capacity and outage probability.

4.3.3.4 Block-Diagonalisation Based Schemes

There are several variants of block-diagonalisation based schemes, proposed in [9] and [10] to study the case $N \geq 2M$. These schemes perform block-diagonalisation to eliminate the multi-pair interference, i.e. $\mathbf{W}_{i,j} = \text{null}(\tilde{\mathbf{H}}_{i,j}^T)$. The power allocation factor for each user pair (i,j) can be expressed as $\alpha_{i,j} = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{||\mathbf{F}_{i,j}\mathbf{H}_{i,j}||_F^2 + \sigma^2 ||\mathbf{F}_{i,j}||_F^2}}$ and the pre-log $\beta = \frac{1}{2}$. The difference between these existing schemes and the proposed schemes lies in how the effective beamformer $\mathbf{F}_{i,j}$ deal with multiple null-space vectors.

The authors in [10] adapt the beamforming structure proposed in [11] into following structure

$$\mathbf{F}_{i,j} = \mathbf{W}_{i,j}\mathbf{V}_{i,j}\mathbf{V}_{i,j}^T\mathbf{W}_{i,j}^T, \tag{4.11}$$

where $\mathbf{V}_{i,j} \in \mathbb{C}^{(N-2(M-1))\times 2}$ is the unitary matrix containing the first two right singular vectors obtained from the singular value decomposition (SVD) of the effective relay to user (i,j) channel, $\mathbf{H}_{i,j}^T\mathbf{W}_{i,j}$. This scheme is named as BD-SVD to facilitate the comparison. Unlike the broadcast scenario in [11] where linear post-processing can be performed at the multi-antenna destination to decouple the received mixture, the single-antenna user pair are unable to do so since cooperation between users is not allowed. The BD-SVD beamforming design in (4.11) results in the non-coherent combining of the diversity streams. As a consequence, the diversity gain achieved by this scheme is no better than the zero-forcing scheme [12]. This effect will be shown in the numerical results section.

The authors in [9] propose two variants of block-diagonalisation based schemes using the idea of multicasting, i.e. each user pair is treated as a multicast group. The first scheme, known as pair-aware with channel matching (PA-MF), has the following structure [9]

$$\mathbf{F}_{i,j} = \mathbf{W}_{i,j}\mathbf{B}_{i,j}\mathbf{m}\mathbf{m}^T\mathbf{B}_{i,j}^T\mathbf{W}_{i,j}^T \tag{4.12}$$

where the channel matching matrix $\mathbf{B}_{i,j} = \mathbf{W}_{i,j}^H\mathbf{H}_{i,j}^*$ and vector $\mathbf{m} = [\ 1 \quad 1\ ]^T$. The PA-MF beamforming structure in (4.12) is a direct cascading of the effective transmit and receive multicast beamforming vectors, i.e. $\mathbf{F}_{i,j} = \mathbf{m}_{multicast}\mathbf{m}_{multicast}^T$, where the transmit multicast vector, $\mathbf{m}_{multicast} = \mathbf{W}_{i,j}\mathbf{B}_{i,j}\mathbf{m}$. Although the same channel matching matrix $\mathbf{B}_{i,j}$ as in the proposed coherent combining scheme is used, this PA-MF beamforming structure is not able to achieve the

diversity gain offered by block-diagonalisation. Recall that the effective channel carrying the desired message of user i is $\alpha_{i,j}\mathbf{h}_i^T\mathbf{m}_{multicast}\mathbf{m}_{multicast}^T\mathbf{h}_j$. Expand the effective channel from the relay to user i,

$$\mathbf{h}_i^T\mathbf{m}_{multicast} = \sum_{k=1}^{N-2(M-1)} \left(|\mathbf{h}_i^T\mathbf{w}(k)|^2 + \mathbf{h}_i^T\mathbf{w}(k)\mathbf{w}^H(k)\mathbf{h}_j^*\right), \tag{4.13}$$

and the effective channel from user j to the relay,

$$\mathbf{m}_{multicast}^T\mathbf{h}_j = \sum_{k=1}^{N-2(M-1)} \left(|\mathbf{w}^T(k)\mathbf{h}_j|^2 + \mathbf{h}_i^H\mathbf{w}^*(k)\mathbf{w}^T(k)\mathbf{h}_j\right). \tag{4.14}$$

It can be easily seen that each channel-matched diversity stream, i.e. $|\mathbf{h}_i^T\mathbf{w}(k)|^2$, is adding up either constructively or destructively with the correlated stream, i.e. $\mathbf{h}_i^T\mathbf{w}(k)\mathbf{w}^H(k)\mathbf{h}_j^*$, since the correlated stream has random phase. As a result, the diversity gain offered by block-diagonalisation is destroyed. This effect will be shown in the numerical results section.

The authors in [9] also propose a scheme known as pair-aware with semi-definite relaxation (PA-SDR) which uses max-min optimisation to choose the multicast beamforming vector. The structure of the PA-SDR beamforming is [9]

$$\mathbf{F}_{i,j} = \mathbf{W}_{i,j}\mathbf{m}\mathbf{m}^T\mathbf{W}_{i,j}^T \tag{4.15}$$

where the multicast vector $\mathbf{m} \in \mathbb{C}^{(N-2(M-1))\times 1}$ is chosen by solving the following optimisation problem [9],

$$\begin{aligned} \max_{\mathbf{m}} \quad \min \quad & \left(\left|\frac{\mathbf{h}_i^T\mathbf{W}_{i,j}\mathbf{m}}{\sigma^2}\right|^2, \left|\frac{\mathbf{h}_j^T\mathbf{W}_{i,j}\mathbf{m}}{\sigma^2}\right|^2\right), \\ & s.t. \quad ||\mathbf{m}||^2 \leq 1. \end{aligned} \tag{4.16}$$

It is shown in [58] that such NP-hard optimisation problem can be solved using semi-definite programming by relaxing the rank constraint of the equivalent problem, which requires further randomisation post-processing. Although this PA-SDR scheme is able to achieve all the diversity gain offered by block-diagonalisation, it has higher computational complexity if compared to the proposed schemes in this chapter. PA-SDR has the worst case complexity of $\mathcal{O}(N^7)$ flops [58] as compared to the proposed schemes with at most $\mathcal{O}(MN^3)$ flops. See Appendix C. It will be shown in the numerical

results section that this PA-SDR scheme does not perform better than the proposed coherent combining scheme in terms of ergodic capacity and outage probability, despite the higher computational complexity required.

4.4 Ergodic Capacity Analysis

This section provides analytical results on the ergodic capacity for case I, where the proposed NC-SM scheme achieves higher multiplexing gain than existing schemes[4]. Since symmetrical channels are considered, i.e. all user nodes have the same channel statistics, it is sufficient to study the single-user ergodic capacity. Define the single-user ergodic capacity as the per-user long term data rate a system can support, which can be expressed as

$$C_{erg} = \int_0^\infty \mathcal{I} f_{\mathcal{I}}(\mathcal{I}) d\mathcal{I}, \tag{4.17}$$

where $f_{\mathcal{I}}(\mathcal{I})$ is the probability density function (PDF) of the mutual information $\mathcal{I}$. Recall the mutual information of the proposed protocol in (4.7) where Gaussian coding is assumed. Note that $\mathbf{F}_{i,j} = \mathbf{w}_{i,j}\mathbf{w}_{i,j}^T$. Omit the subscript of $\mathbf{w}$ for simplicity. First, the properties of the variables in $\mathcal{I}$ are examined. In order to find common variables to simplify the SNR expresion in (4.7), one defines the matrix $\mathbf{W} = \mathbf{w}\mathbf{w}^H$. The matrix $\mathbf{W}$ is a positive semi-definite Hermitian matrix and more importantly, $\mathbf{W}$ carries the idempotent property, such that $\mathbf{W}\mathbf{W}^H = \mathbf{W}$. Using the associative property of matrix multiplication, the numerator $|\mathbf{h}_i^T\mathbf{w}\mathbf{w}^T\mathbf{h}_j|^2$ in (4.7) can be split into $|\mathbf{h}_i^T\mathbf{w}|^2|\mathbf{w}^T\mathbf{h}_j|^2$. Each of the absolute square can be written as vector multiplication, i.e. $|\mathbf{h}_i^T\mathbf{w}|^2 = \mathbf{h}_i^T\mathbf{W}\mathbf{h}_i^*$. Applying the idempotent property of $\mathbf{W}$, one can represent $\mathbf{h}_i^T\mathbf{W}\mathbf{h}_i^* = \mathbf{h}_i^T\mathbf{W}\mathbf{W}^H\mathbf{h}_i^* = ||\mathbf{h}_i^T\mathbf{W}||^2$. Following similar approach, one can write $|\mathbf{w}^T\mathbf{h}_j|^2 = ||\mathbf{W}^T\mathbf{h}_j||^2$, where $\mathbf{W}^T = \mathbf{w}^*\mathbf{w}^T$. Define the auxiliary variables $x = \frac{1}{\sigma^2}||\mathbf{h}_i^T\mathbf{W}||^2$ and $y = \frac{1}{\sigma^2}||\mathbf{W}^T\mathbf{h}_j||^2$. After some algebraic manipulations, the mutual information in (4.7) can be simplified as follows,

$$\mathcal{I} = \frac{1}{2}\log_2\left(1 + \frac{xy}{(M+1)x + My + M}\right). \tag{4.18}$$

[4]The high SNR approximation used in deriving the ergodic capacity in this section is too coarse to capture the power gain (or array gain) due to the use of additional antennas in case II. As a result, the analytical capacity expression for case II is omitted.

Note that $||\mathbf{w}\mathbf{w}^T\mathbf{h}_i||^2 = ||\mathbf{h}_i^T\mathbf{W}||^2$. Define auxiliary variable $z = \frac{xy}{(M+1)x+My+M}$. The following lemma quantifies the cumulative distribution function (CDF) of variable z.

Lemma 4.1. *The CDF of the auxiliary variable z can be expressed as*

$$F_z(z) = 1 - 2\sigma^2 \exp\left(-(2M+1)\sigma^2 z\right) \sqrt{M(M+1)z^2 + Mz} \times K_1\left(2\sigma^2\sqrt{M(M+1)z^2 + Mz}\right), \tag{4.19}$$

where M is the number of user pairs and $K_1(z)$ is the modified Bessel function of the second kind.

Proof. Refer to Appendix B. ■

Observe that the CDF of z is expressed in terms of a modified Bessel function of second kind, which has no closed form solution. Thus, the next lemma is used to obtain the upper and lower bounds for the CDF of variable z in order to ease the difficulty in the subsequent analytical development of the ergodic capacity.

Lemma 4.2. *The CDF of the auxiliary variable z can be bounded as*

$$1 - \exp\left(-(2M+1)\sigma^2 z\right) \leq F_z(z) \leq 1 - \exp\left(-(2M+1)\sigma^2 z - 2\sigma^2\sqrt{M(M+1)z^2 + Mz}\right). \tag{4.20}$$

Proof. Refer to Appendix B. ■

In order to facilitate the analytical development of the ergodic capacity, the following high SNR approximation for the upper bound of the CDF of z is used,

$$1-\exp\left(-(2M+1)\sigma^2 z-2\sigma^2\sqrt{M(M+1)z^2+Mz}\right)$$
$$\approx 1-\exp\left(-\left(2M+1+2\sqrt{M(M+1)}\right)\sigma^2 z\right). \tag{4.21}$$

Recall that the mutual information is a function of z, such that $\mathcal{I}(z)=\frac{1}{2}\log_2(1+z)$. According to Section 7.4.9 in [59], the expected value of $\mathcal{I}$ (which is the ergodic capacity) can be derived from the following formula

$$C_{erg}=\int_0^\infty \mathcal{I}(z)f_z(z)dz. \tag{4.22}$$

The density function $f_z(z)$ can be obtained by differentiating $F_z(z)$ over z, i.e. $f_z(z)=\frac{dF_z(z)}{dz}$. The following theorem expresses the ergodic capacity of the proposed protocol.

Theorem 4.1. *The single-user ergodic capacity of the proposed NC-SM protocol can be upper and lower bounded using a high SNR approximation, as follows,*

$$c\left(2M+1+2\sqrt{M(M+1)}\right)\leq C_{erg}\leq c\left(2M+1\right), \tag{4.23}$$

where the function $c(x)=\frac{1}{2}\exp(x\sigma^2)\log_2\left(\frac{1}{x\sigma^2\exp(\gamma)}\right)$ *and* γ *is the Euler constant.*

Proof. Refer to Appendix B. ■

Remark for Theorem 4.1

For a fixed number of user pairs M, the lower bound of the ergodic capacity at high SNR, i.e. $\sigma^2\to 0$, can be written as $C_{erg,lower}\approx\frac{1}{2}\log_2\left(\frac{1}{\sigma^2}\right)-\frac{1}{2}\log_2\left(\left(2M+1+2\sqrt{M(M+1)}\right)\exp(\gamma)\right)$. The first term of the RHS is a function of SNR while the second term is a constant independent of the SNR. Therefore, for every 3dB increase in SNR, there is a 0.5 bits/s/Hz increment on the ergodic capacity. In other words, the achievable per-user multiplexing gain is $\frac{1}{2}$, which is independent of the number of user pairs. The second term captures two effects. First, the power loss due to the null-space projection operation used in the beamformer and second, the equal power sharing among M pairs of users. As the number of user pairs increases, the fixed transmission power at the relay is shared by more users.

Hence each user gets a smaller portion of the total power as M increases. This explains why the second term increases logarithmically with M. Overall, the single-user ergodic capacity decreases as the number of active user pairs increases, but the individual multiplexing gain obtainable by each user is maintained.

4.5 Outage Probability and the Diversity and Multiplexing Tradeoff Analysis

This section presents the analytical results on outage probability and the diversity and multiplexing tradeoff of the proposed NC-SM schemes. Since symmetric channels are considered, the analysis of single-user outage performance is adequate. Define the outage probability, $P(I < R)$, as the probability that the mutual information of a user falls below the individual target data rate R. The outage performance of the proposed null-space vector selection scheme is analysed, for two obvious reasons. First, it is the generalisation of the proposed protocol to the case $N \geq 2M-1$, which includes both case I and case II. For example, when $N = 2M-1$, the proposed null-space vector selection scheme reverts to the proposed scheme in case I. Second, the proposed null-space vector selection scheme is a lower bound for the proposed coherent combining scheme. This argument is verified in the numerical results section. The following theorem captures the outage probability of the proposed NC-SM schemes.

Theorem 4.2. *The single-user outage probability of the proposed NC-SM schemes is*

$$P\left(\mathcal{I} < R\right) = \left(1 - \exp\left(-(2M+1)\sigma^2\zeta\right)\omega K_1\left(\omega\right)\right)^{N-2(M-1)}, \tag{4.24}$$

where $\omega = 2\sigma^2\sqrt{M(M+1)\zeta^2 + M\zeta}$ *and* $\zeta = 2^{2R} - 1$.

Proof. Refer to Appendix B. ■

Remark for Theorem 4.2

At fixed data rate R, the auxiliary variable $w \to 0$ when $\sigma^2 \to 0$. Using the approximation for the exponent, i.e. $\underset{x\to 0}{\exp}(-x) = 1 - x$ and also the approximation for the modified Bessel function of second kind, i.e. $\underset{x\to 0}{K_1}(x) = \frac{1}{x}$, one can obtain the high SNR approximation of the outage probability

$$P\left(\mathcal{I} < R\right) \approx \left((2M+1)\sigma^2\zeta\right)^{N-2(M-1)}. \tag{4.25}$$

Representing the mean SNR as $\gamma_0 = \frac{1}{\sigma^2}$, it is readily shown that the outage probability decays as $\frac{1}{\gamma_0^{N-2(M-1)}}$. The exponent of the SNR defines the rate of decay and it is well-known as the diversity gain. When the diversity gain is higher, the outage probability decays faster with increasing SNR. The diversity gain achieved by the proposed NC-SM schemes is $N - 2M + 2$. It is better than the conventional multi-user zero-forcing scheme. Since the analytical outage probability of the zero-forcing scheme in [12] is not available, we have to make a comparison with the conventional zero-forcing receiver [52] used in MIMO channels with $2M$ transmitter antennas and N receiver antennas. Provided that $N \geq 2M$, the diversity gain achieved by each substream in zero-forcing receiver is $N - 2M + 1$. Clearly, the proposed NC-SM schemes deliver additional diversity gain as compared to the MIMO zero-forcing receiver, thanks to the beamformer design which only removes the co-channel interference caused by other user pairs. In comparison with the multiple-input and single-output (MISO) upper bound which has optimal diversity gain of N, the loss of diversity gain experienced by the proposed scheme due to the block-diagonalisation is $2M - 2$. This is lower than the loss of $2M - 1$ suffered by conventional zero-forcing scheme.

Similar to the MIMO channels, there is a fundamental tradeoff between the rate and reliability in the multi-pair, bidirectional relay channels. In order to quantify this relationship, we derive the diversity and multiplexing tradeoff for the proposed NC-SM schemes. Recall that the diversity and multiplexing gain can be defined as [29]

$$d \triangleq -\lim_{\gamma_0\to\infty} \frac{\log\left[P_e(\gamma_0)\right]}{\log \gamma_0} \text{ and } r \triangleq \lim_{\gamma_0\to\infty} \frac{R(\gamma_0)}{\log \gamma_0}, \tag{4.26}$$

where P_e is the maximum likelihood (ML) probability of detection error, R is the target data rate in

bits/Hz/s, γ_0 is the mean SNR and in our case $\gamma_0 = \frac{1}{\sigma^2}$. The outage probability is used to obtain the diversity and multiplexing tradeoff because the ML error probability is tightly bounded by the outage probability at high SNR. Using the result from Theorem 4.2, one obtains the following corollary which quantifies the diversity and multiplexing tradeoff of the proposed NC-SM schemes.

Corollary 4.1. *The achievable per-user diversity and multiplexing tradeoff of the proposed NC-SM schemes is*

$$d(r) = (1-2r)\,(N-2M+2), \quad 0 \le r \le 1, \tag{4.27}$$

provided that $N \geq 2M-1$.

Proof. As shown in the remark of Theorem 4.2, the high SNR approximation for the outage probability can be written as

$$P\left(\mathcal{I} < R\right) \approx \left((2M+1)\sigma^2\zeta\right)^{N-2(M-1)}. \tag{4.28}$$

Substituting $R = r\log_2\left(\gamma_0\right)$ into (4.28), one obtains the following

$$\begin{aligned} P\left(\mathcal{I} < R\right) &\approx \left((2M+1)\left(\frac{1}{\gamma_0}\right)\left(\gamma_0^{2r}-1\right)\right)^{N-2(M-1)} \\ &\doteq \gamma_0^{-(1-2r)(N-2(M-1))}. \end{aligned} \tag{4.29}$$

From the exponential equality, the diversity and multiplexing tradeoff of the proposed NC-SM scheme is readily shown as

$$d(r) = (1-2r)\,(N-2M+2), \tag{4.30}$$

and the corollary is proved. ■

Remark for Corollary 4.1

The maximum diversity gain (when $r = 0$), describes how fast the outage probability decays with the SNR at fixed data rate. This maximum diversity gain of $N-2M+2$ is in line with Theorem 4.2. On the other extreme, the maximum multiplexing gain (when $d = 0$), describes how the data rate

grows with the SNR at fixed outage probability. The multiplexing gain of $1/2$, achievable by each user is independent of the number of user pairs M. This agrees with the result in Theorem 4.1. The multiplexing gain of $1/2$ is indeed the best possible gain that can be achieved in the half duplex, two-way relay channels being studied. The total multiplexing gain achievable in the network is the sum of all individual multiplexing gain, which is equal to M. Both the individual and network multiplexing gains of the proposed NC-SM schemes outperform the existing schemes in [12] (when $N < 2M$) and [57], which are designed for single-pair and extended to multi-pair using time sharing between pairs. The per-user multiplexing gain achievable by comparable schemes [12, 57] with time sharing between pairs is $1/2M$, which decreases as M gets larger. The detailed comparisons are shown in the next section. The simultaneous increase of data rate (positive r) and reliability (positive d) is possible in the proposed NC-SM schemes, as long as the diversity and multiplexing tradeoff in (4.27) is satisfied. The proposed NC-SM schemes achieve better diversity and multiplexing tradeoff than the zero-forcing scheme (when $N \geq 2M$) for all data rates. Furthermore, the proposed NC-SM schemes offer better tradeoff at high data rate region as compared to the diversity gain optimal scheme, the maximal ratio reception-transmission (MRR-MRT) scheme [57]. These arguments are justified by the simulation results in the next section. Note that the analytical results on the outage probability and the diversity and multiplexing tradeoff of the existing schemes are not available in the literature, which prevent direct analytical comparison. Until this point, the optimal diversity and multiplexing tradeoff for the multi-pair, two-way relay channels with a half duplex constraint remains an open problem.

4.6 Numerical Results

This section presents several Monte Carlo simulation results on the single-user ergodic capacity and single-user outage probability, in order to validate the analytical results and assess the performance of the proposed NC-SM schemes in comparison with existing methods. All schemes assume equal power allocation to enable fair comparisons.

Figure 4.2 shows the single-user ergodic capacity versus SNR of the proposed NC-SM scheme for case I, i.e. $N = 2M - 1$. The curve generated using Monte Carlo simulation is compared with the analytical bounds in Theorem 4.1. The number of user pairs is fixed at $M = 2$ while the number

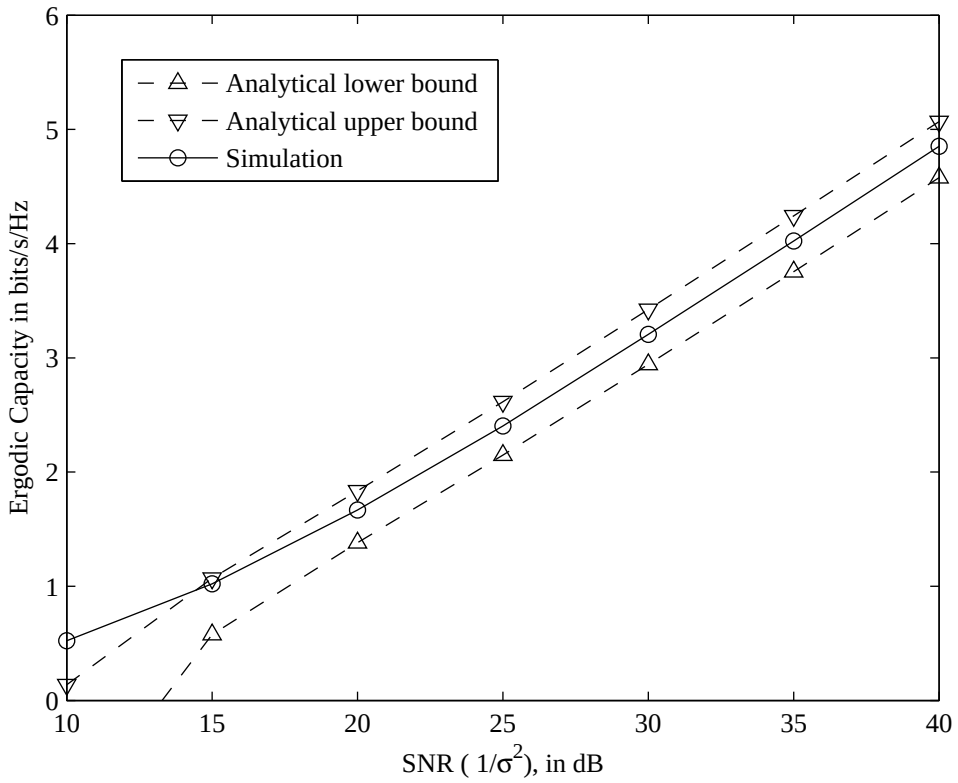


Figure 4.2: Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of the proposed protocol in comparison with the capacity bounds. Fixed parameters: $M = 2$ and $N = 3$.

of antennas at the relay is fixed at $N = 3$. The analytical upper and lower bounds approximate the ergodic capacity very well, starting from medium to high SNR, i.e. SNR>15dB. Recall that the slope of the ergodic capacity curve characterises the multiplexing gain. The simulation result agrees with the analytical bounds that the proposed NC-SM scheme can achieve per-user multiplexing gain of $1/2$ (or a total multiplexing gain of M).

Figure 4.3 compares the single-user ergodic capacity versus SNR of the proposed NC-SM scheme and three comparable AF based schemes for case I, i.e. $N = 2M-1$. The fixed parameters are $M = 2$ and $N = 3$. The baseline schemes: 1. pure AF, 2. MRR-MRT [57], and 3. zero-forcing [12]. The baseline schemes are extended to multi-pair using time sharing between pairs. Note that the zero-forcing [12] cannot support all user pairs simultaneously because the zero-forcing criterion requires $N \geq 2M$. It can be observed that from medium to high SNR, i.e. SNR>17dB, the ergodic capacity of the proposed NC-SM scheme outperforms all existing AF based schemes. At high SNR, i.e. SNR=30dB, significant capacity gains of 36%, 45% and 75% are obtained by the proposed NC-SM scheme in comparison with the MRR-MRT, zero-forcing, and the pure AF schemes respectively. The proposed protocol demonstrates that spatial multiplexing across different pairs coupled with network

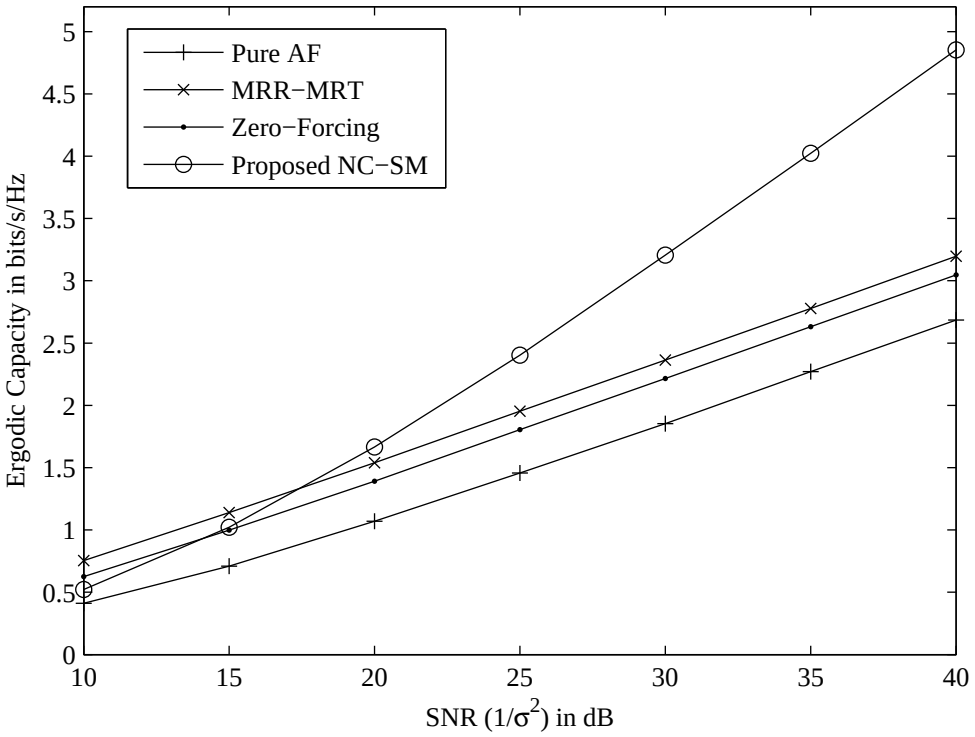


Figure 4.3: Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of various schemes when $M = 2$ and $N = 3$.

coding within same pair is spectrally more efficient. By observing the slope of the curves, it is obvious that the proposed NC-SM scheme delivers the highest multiplexing gain among all schemes, thanks to the efficient use of the available degrees of freedom to spatially support multiple user pairs. At low SNR, a system supporting only single-stream, i.e. MRR-MRT, performs better than spatial multiplexing system such as the proposed scheme. Recall that at low SNR, allocating all of the transmission power to the best subchannel (corresponds to single-stream system) is better than distributing the power among all subchannels (corresponds to spatial-multiplexing system) [18].

Figure 4.4 shows the single-user ergodic capacity versus SNR of the proposed NC-SM schemes in comparison with the existing AF based schemes, for case II, i.e. $N \geq 2M$. The fixed parameters are $M = 2$ and $N = 4$. Recall that in case II, the proposed NC-SM beamformer uses either the null-space vector selection or coherent combining of null-space vectors. The comparable schemes are MRR-MRT, zero-forcing, BD-SVD [10], PA-MF and PA-SDR [9]. All comparable schemes are able to support all user pairs simultaneously (spatial multiplexing), except for the MRR-MRT which uses time-sharing between pairs. From Figure 4.4, it can be observed that all schemes supporting spatial multiplexing achieve higher ergodic capacity and higher multiplexing gain when compared with the scheme based on time sharing between pairs (MRR-MRT scheme). All spatial multiplexing

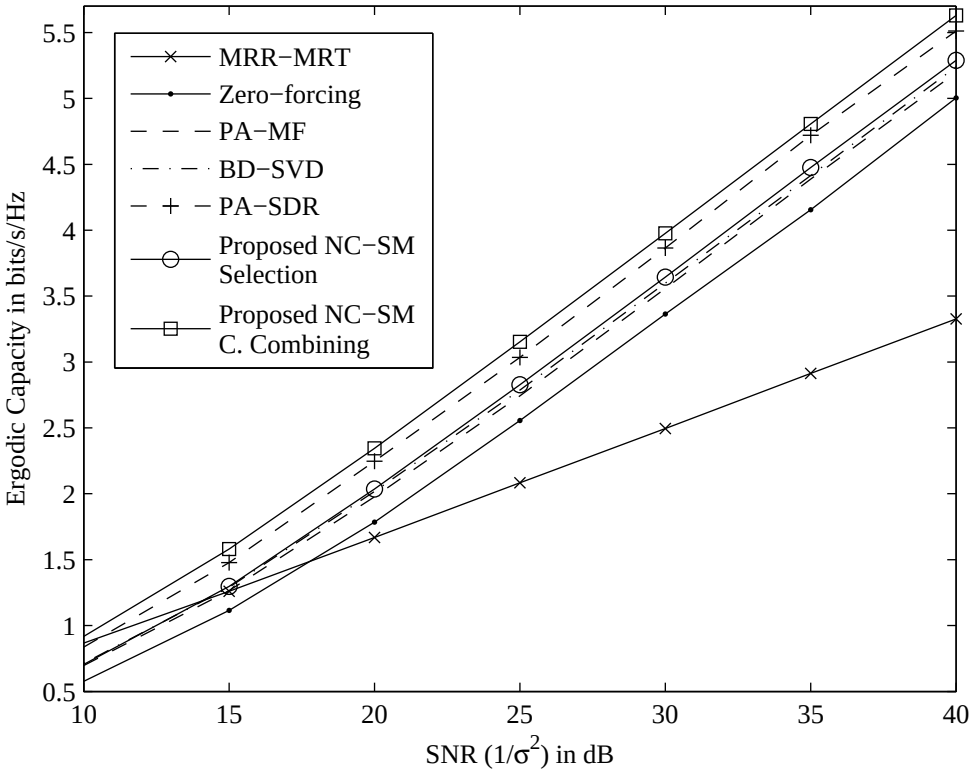


Figure 4.4: Single-user ergodic capacity versus SNR $(1/\sigma^2)$ of various scheme when $M = 2$ and $N = 4$.

schemes achieve the same multiplexing gain, evident from the slope of ergodic capacity curves. It can be observed that block-diagonalisation based schemes (including proposed NC-SM schemes, BD-SVD, PA-MF and PA-SDR) are able to achieve higher ergodic capacity for any fixed SNR, or deliver power gain for any fixed ergodic capacity, when compared with the zero-forcing scheme. Among all block-diagonalisation based schemes, the proposed NC-SM with coherent combining delivers the best performance. The PA-SDR scheme performs close to the proposed NC-SM with coherent combining, but it comes at the cost of higher computational complexity. The higher complexity of PA-SDR is due to the use of semi-definite programming in computing the multicast vectors [58]. The NC-SM with null-space vector selection, is about 2 dB away from the coherent combining scheme. Notice that the PA-MF and BD-SVD schemes do not perform better than the proposed null-space vector selection scheme. Similar to Figure 4.3, MRR-MRT dominates at low SNR.

Besides providing ergodic capacity improvement, block-diagonalisation offers higher diversity gain as compared to the zero-forcing scheme. The diversity gain achieved by the proposed NC-SM schemes can be verified from the outage probability versus SNR curves shown in Figure 4.5. The fixed parameters are $M = 2$, $N = 4$, and R=2 bits/s/Hz. Generally, the proposed NC-SM scheme

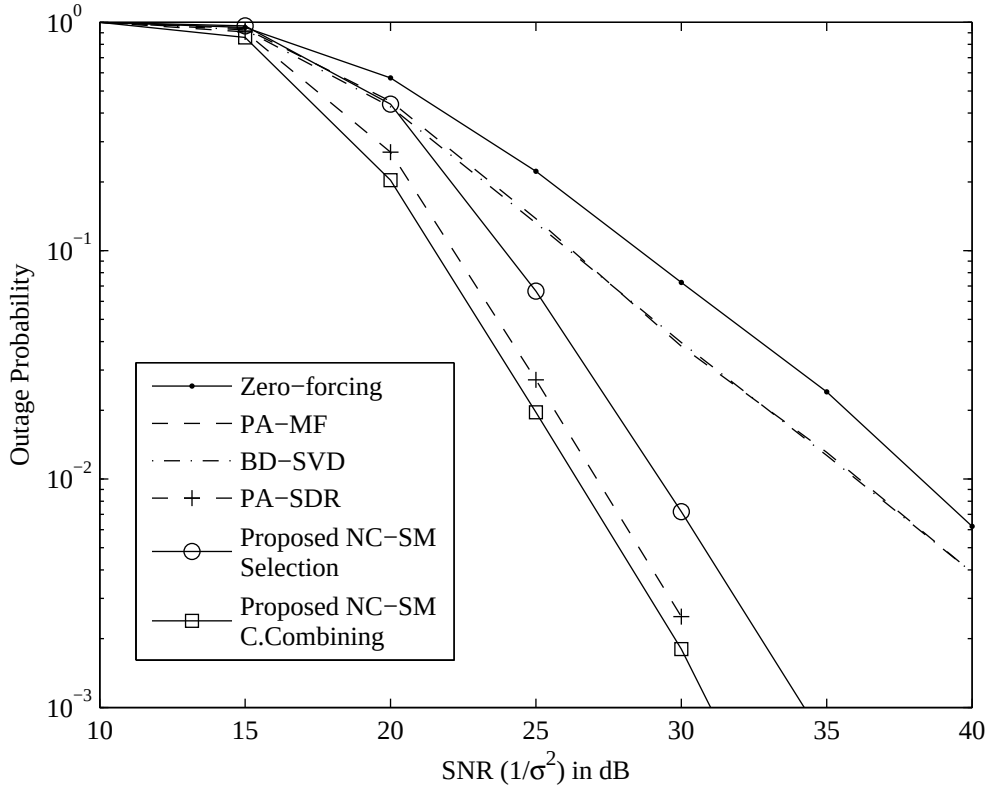


Figure 4.5: Single-user outage probability versus SNR $(1/\sigma^2)$ when $M = 2$, $N = 4$ and $R = 2$ bits/s/Hz.

with coherent combining achieves the lowest outage probability. For example, at SNR=25dB, the proposed scheme with coherent combining achieves 2% outage while the zero-forcing scheme suffers from 20% outage. In addition, for fixed outage probability, the proposed NC-SM schemes offers significant power saving if compared to the zero-forcing scheme. For instance, the proposed NC-SM with coherent combining requires 10dB less power than the zero-forcing scheme to maintain 1% outage. The PA-SDR scheme performs close to the proposed NC-SM with coherent combining while the NC-SM with selection is about 2.5dB away from the best scheme. Recall that the slope of the outage probability curve characterises the diversity gain. The steeper the outage probability curve, the higher the diversity gain is. The proposed NC-SM schemes (both selection and coherent combining schemes) and the PA-SDR are able to achieve a higher diversity gain as compared to the BD-SVD, PA-MF and zero-forcing schemes. The BD-SVD and PA-MF schemes are not able to extract the additional diversity gain offered by block-diagonalisation, due to the non-coherent combining of the diversity streams. Although the PA-SDR scheme is able to extract all the diversity gain, it suffers from high computational complexity. The proposed NC-SM schemes have lower complexity while being able to achieve all the diversity gain offered by block-diagonalisation.

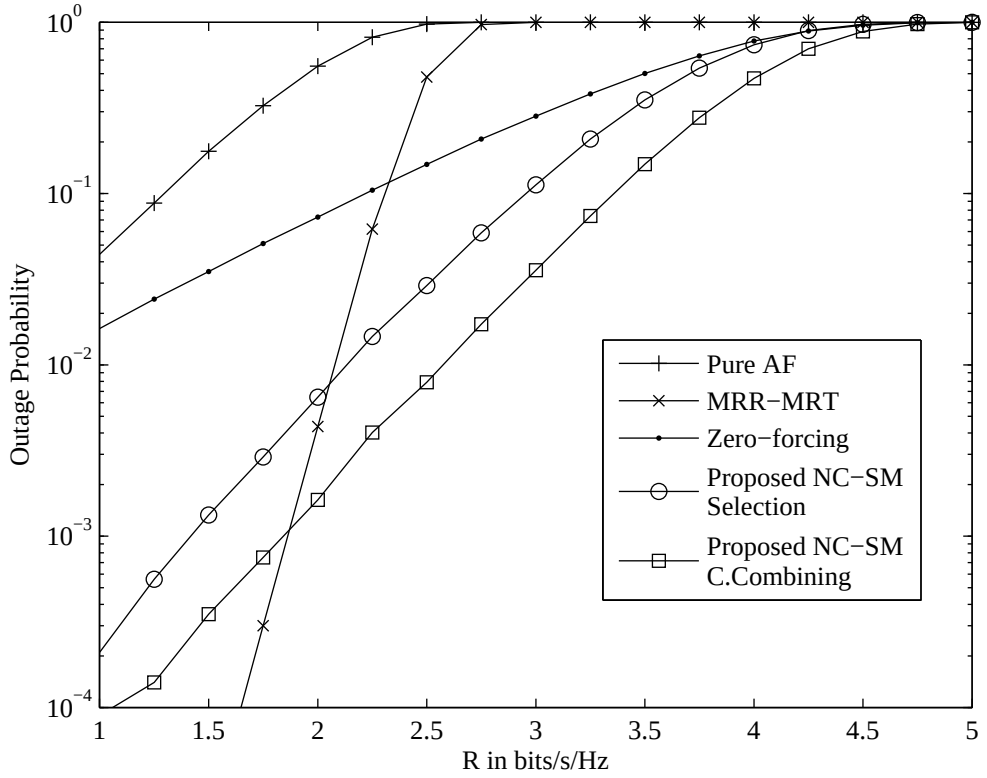


Figure 4.6: Single-user outage probability versus target data rate R for various schemes when $M = 2$, $N = 4$ and SNR=30dB.

The next simulation reveals the tradeoff between rate and reliability between several schemes. The curves for outage probability versus target data rate R, are shown in Figure 4.6. The fixed parameters are $M = 2$, $N = 4$ and SNR=30dB. The pure AF and MRR-MRT schemes use time sharing between pairs while the zero-forcing and the proposed schemes employ spatial multiplexing. Both the proposed NC-SM with selection and the proposed NC-SM with coherent combining achieve a lower outage probability when compared with the zero-forcing and the pure AF schemes, at all target data rates. For instance, when the outage probability is fixed at 5%, the proposed NC-SM with coherent combining supports a data rate up to 3bits/s/Hz, delivering data rate enhancements of 75% and 200% when compared to the zero-forcing and pure AF scheme respectively. The proposed NC-SM with selection performs close to the proposed NC-SM with coherent combining. Although the zero-forcing scheme also supports spatial multiplexing, it still underperforms the proposed schemes. This is because the zero-forcing scheme has lower diversity gain as compared to the proposed schemes. In general, the proposed schemes have better diversity and multiplexing tradeoff when compared with existing zero-forcing scheme and pure AF scheme. Besides that, the proposed schemes outperform the MRR-MRT at the high data rate region, i.e. $R > 2$ bits/s/Hz. As an example, at a fixed data rate, i.e. $R = 3$

bits/s/Hz, the proposed NC-SM with coherent combining is able to maintain 4% outage while the MRR-MRT scheme completely fails. Although the MRR-MRT scheme has the highest diversity gain as compared to all other schemes, it suffers from poor outage performance at high data rate region. In fact, its per-user multiplexing gain is only $1/2M$ due to the use of time sharing between pairs. From an implementation perspective, having a system that is able to support a higher data rate at a reasonable outage probability is certainly preferable to a system that delivers lower outage probability but operates at a very low data rate. This can be illustrated using the following expression

$$R_T = R \times (1 - P(\mathcal{I} < R)), \tag{4.31}$$

where R_T is the real throughput, R is the target data rate and $P(\mathcal{I} < R)$ is the outage probability. For a fixed $P(\mathcal{I} < R)$, the real throughput R_T produced by the proposed schemes is much better than the MRR-MRT scheme due to the ability of the proposed schemes to support higher target data rate R. Along this line, the proposed NC-SM schemes are desirable since they offer better diversity and multiplexing tradeoff at high data rate region. Although not included in the figure, it can be easily verified that the PA-SDR has similar tradeoff performance as the proposed NC-SM schemes, while PA-MF and BD-SVD have similar tradeoff performance as the zero-forcing scheme.

4.7 Chapter Conclusion

In this chapter, an analogue network coding protocol for the multi-pair, two-way relaying channel has been proposed. The proposed protocol combines analogue network coding and spatial multiplexing, which allows both the relay and the users to participate in interference cancellation. The proposed beamforming schemes yield significant improvements in terms of ergodic capacity and outage probability. The developed analytical bounds approximate the ergodic capacity closely and show that the proposed beamforming scheme achieves higher multiplexing gain than existing schemes. The analytical results for the outage probability quantifies the diversity gain and proves that the proposed beamforming schemes are able to extract all additional diversity gain offered by block-diagonalisation as compared to zero-forcing. The derived diversity and multiplexing tradeoff reveals the superior rate and reliability tradeoff offered by the proposed beamforming schemes. Simulation results agree with

the analytical results that the proposed beamforming schemes achieve higher ergodic capacity, lower outage probability and better diversity and multiplexing tradeoff than comparable schemes.

Chapter 5

Joint Beamforming and Power Management for the MIMO Two-Way Relaying Channel

In the previous chapter, it is shown that a multi-antenna relay with suitably designed receive and transmit beamforming matrices is able to improve the throughput and the reliability of the information exchange channels with multiple pairs of single-antenna users. This chapter considers the information exchange channel where not only the relay but also the user pairs are equipped with multiple antennas. The multiple antennas available at the users and the relay provide an additional dimension to improve the data rate of the two-way information exchange channel. In particular, a low-complexity joint beamforming and power management scheme is proposed in this chapter to improve the sum-rate of the network. The proposed beamformers first align the channel matrices of the user pair and then decompose the aligned channel into parallel subchannels. Two power management issues, i.e. power allocation and power control are addressed. First, a sum-rate optimising power allocation is proposed to allocate power between all subchannels and nodes. Second, a quality of service (QoS) satisfying power control is proposed to minimise the total transmission power in the network. Simulation results justify that the proposed schemes deliver better sum-rate performance or consume lower transmission power, when compared to existing schemes.

5.1 Chapter Introduction

Two-way relaying is a spectrally efficient technique to enable information exchange between two users. Two-way relaying protocols such as those based on decode-and-forward (DF) relaying [4], amplify-and-forward (AF) relaying [5], and estimate-and-forward (EF) relaying [6] are able to fulfil two-way information exchange in just two phases. Specifically, both users transmit concurrently in the same channel during the first phase while the relay broadcasts the processed mixture to both users in the second phase. Each user utilises the knowledge of the previously transmitted message known as self interference, to decode the received mixture. In comparison, conventional one-way relaying consumes four orthogonal channel uses to complete the information exchange due to half duplex constraint.

For practical considerations, non-regenerative relaying, i.e. AF relaying, is desirable when compared to other relaying methods. This is due to the fact that non-regenerative relaying has lower complexity, lower processing delay and incurs lower signal processing power if compared to regenerative relaying, i.e. DF and EF relaying. Attracted by the benefits of multiple antennas in enhancing the system capacity and reliability, subsequent works on non-regenerative two-way relaying extend to the multi-antenna scenario. Reference [57] and [13] consider the sum-rate optimising AF based beamforming and power allocation at the relay, for the case where only the relay is equipped with multiple antennas. Reference [60] studies the case where multiple relays are used to assist the information exchange and proposes a sum of mean-squared-error (MSE) minimising relay precoder subject to total power constraint at the relays. Reference [61] then generalises the scenario to include multiple pairs of single-antenna users and demonstrates that AF based beamforming at the relay is able to address co-channel interference and improve throughput and reliability.

Meanwhile, [12] and [62] look into the case with a single pair of multi-antenna users and a non-regenerative multi-antenna relay. Reference [12] proposes a sum-rate maximising beamforming design at the relay subject to a fixed power constraint. However, the number of antennas required at the relay is twice the number of antennas needed at each user, due to the zero-forcing criterion. The self interference cancellation capability at user nodes which reduces the requirement on the number of antennas at the relay is not utilised in [12]. [62] studies the joint design of beamformer at the relay and decoder at the users to minimise the sum MSE, subject to individual power constraint at each

node. The possibility of beamforming at multi-antenna users is not explored in [62].

The joint design of transmit and receive beamformers in the MIMO two-way relaying channel are studied in [14] and [15]. It is recognised in [14] and [15] that the sum-rate expression is not jointly concave with respect to the transmit and receive beamforming matrices, which complicates the optimisation of the beamforming matrices. Reference [14] proposes an iterative searching algorithm based on gradient descent method to find the locally optimal solution for the beamformers at the users and relay satisfying individual power constraints with equality. The algorithm has to be repeated extensively with different starting points in order to increase the probability of finding the best local optimal solution which corresponds to the global optimal solution. The use of multiple relays is also studied in [14]. Meanwhile, [15] proposes an alternate optimisation technique which combines searching algorithms and convex optimisation techniques to find locally optimal beamformers at the users when beamformer at the relay is fixed, and vice versa until convergence is reached. Similar to [14], the algorithms proposed in [15] to find locally optimal solutions have to be repeated multiple times to increase the probability of reaching global optimal solution. One major drawback of the algorithms proposed in [14] and [15] is the expensive computation cost involved in determining the beamforming matrices. The problem dimension of the algorithms in [14] and [15] which grows quadratically with the number of antennas increases the computational complexity significantly when the number of antennas is large. Another shortcoming is the high computation overhead involved following the fact that both [14] and [15] require an extensive repetition of a local optimisation procedure with different initial points to achieve global optimal solution. In addition, [14] and [15] only consider the case where each node is subject to fixed individual power constraints. The possible performance gain of implementing joint power allocation at all nodes subject to a total network power constraint remains unexplored.

The joint power allocation problem has been investigated in a one-way relaying scenario, e.g. in [16] and [17]. It is agreed by both [16] and [17] that the joint power allocation problem for the non-regenerative MIMO one-way relaying is neither concave nor convex. Reference [16] proposes a high signal-to-noise ratio (SNR) approximation to convert the non-concave sum-rate optimising objective into concave form which can be solved using convex optimisation techniques [20]. Reference [17] derives the convex MSE bounds for the receiver, which are used to approximate the original sum-rate optimising or MSE minimising objective function. However, the solutions cannot be directly applied

to a two-way relaying scenario due to the fundamental difference in the transmission protocol. In two-way relaying, the relay receiver needs to account for the superposition of the transmissions from both users while the transmit beamformer at the relay needs to forward information to both users simultaneously. It is remarked in [14] that channel decomposition for substream power allocation, i.e. water-filling, is not possible in MIMO two-way relaying channels[1]. Furthermore, the joint transmit and receive beamforming design in MIMO two-way relaying channels involves all three nodes in the network. In one-way relaying, only the source and relay are involved in transmit beamforming while the relay and destination participate in receive beamforming.

On the other hand, the capability of the MIMO system to support multiple parallel substreams enables spatial multiplexing of several traffic streams with various predefined QoS [63]. The QoS constraints can be target SNR, target data rates, target error rates, etc. The QoS requirement depends on the type of traffic. For instance, in a multimedia application, real-time video traffic requires higher target data rates than real-time audio traffic. In certain applications, i.e. real-time control, real-time surveillance, etc, successful information delivery defined by the QoS constraints is more important than the power constraints. All these lead to the problem of fulfilling the QoS constraints with the lowest amount of power [63]. Reference [64] studies the power control (power minimisation) problem, subject to SNR constraints for the MIMO one-way relaying channel, while [65] investigates the joint beamforming and power control problem, subject to per-user signal-to-interference-and-noise ratio (SINR) constraints for the multi-user MIMO one-way relaying channel. Nonetheless, the power control problem with QoS constraints in the MIMO two-way relaying channel remains unexplored.

This chapter considers a two-way relaying scenario consists of a pair of multi-antenna users and a non-regenerative multi-antenna relay, all equipped with M antennas. A joint transmit and receive beamforming design which enables network power allocation and power control is proposed. Firstly, the joint transmit and receive beamformers at all nodes are designed to ensure that each subchannel of user 1 and user 2 are aligned in the same signal subspace. Secondly, the singular value decomposition (SVD) is performed to decompose the aligned channel matrices of user 1 and 2 into M parallel subchannels, where each subchannel contains the mixture of user 1's and user 2's messages. The proposed beamforming design facilitates the investigation of two power management problems: 1. the joint power allocation problem which maximises the sum-rate, subject to a predefined total power

[1]Despite the negative remark in [14], this chapter demonstrates that channel decomposition is possible through the transmit and receive beamforming design detailed in Section 5.3.

constraint in the network; 2. the joint power control problem which minimises the total transmit power consumption in the network, subject to preset QoS constraints, i.e. target data rates. Problem 1 and Problem 2 are also known as rate adaptive loading and margin adaptive loading respectively. Such network power allocation and power minimisation are critical in limiting the total interference incurred to a coverage area which is often regulated by the authority.

In this chapter, the power allocation formulation which allows users and relay to jointly allocate power to each substream in order to optimise the sum-rate of the network, is shown to have non-concave utility. In order to enable the problem to be solved efficiently using convex optimisation techniques [20], a concave upper bound is derived to approximate the original non-concave objective function. On the other hand, the power control formulation which minimises the total transmission power in the network while satisfying the rate constraints is in the form of geometric program. The geometric program is transformed into convex form solvable using convex optimisation techniques. Numerical results evaluating the performance of the proposed scheme and studying the relationship between ergodic sum-rate and parameters such as SNR, number of antennas, and path loss are presented. The results show that the ergodic sum-rate of the proposed scheme significantly outperforms the baseline scheme, DF MIMO one-way relaying. Numerical results on the power control problem are also included to justify the performance of the proposed scheme and to reveal the relationship between the average total network transmission power and parameters such as SNR and target data rates. The results fully support the claim that the proposed scheme is more energy efficient in satisfying the QoS constraints, when compared to the baseline scheme.

The rest of the chapter is organised as follows. The system model and protocol description are presented in Section 5.2 while the beamforming design is discussed in Section 5.3. In Section 5.4, the joint power allocation problem is investigated while in Section 5.5, the joint power control problem is studied. Section 5.6 covers the numerical simulation results. Section 5.7 concludes the chapter.

5.2 System Model and Protocol Description

Consider a scenario where two multi-antenna users wish to exchange information with the help of a non-regenerative, multi-antenna relay. The full spatial multiplexing case where all nodes are equipped with M antennas is of interest. This configuration commonly occurs in ad hoc and sensor networks

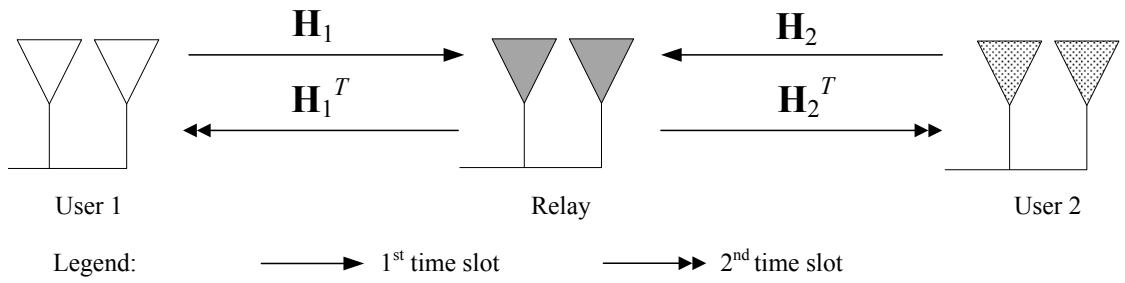


Figure 5.1: Example of two-way relaying scenario where each node is equipped with $M = 2$ antennas. The symbols above the arrows represent the channel matrices while the directions of the arrows indicate the directions of data flows.

where nodes have the same number of antennas and the relay is selected from idle users in the network. Figure 5.1 shows an example of the two-way relaying channel with $M = 2$. All channels undergo i.i.d. quasi-static Rayleigh fading and assume channel reciprocity. The receiver is corrupted by circularly symmetric additive white Gaussian noise. The half duplex constraint is assumed throughout the chapter and it is realised through time division duplexing.

5.2.1 Initialisation

Prior to the proposed information exchange protocol, the initialisation phase takes place in order to enable all nodes to estimate the channels. The proposed two-way relaying protocol requires the relay to have full knowledge of the channel state information (CSI) of both relay-to-user channels, while each user knows his and his partner's user-to-relay channels[2]. Utilising channel reciprocity, the CSI can be obtained using the open loop method [22], which can be accomplished within $3M + 1$ time slots. First, the relay uses M time slots to broadcast pilot sequences for both users to estimate their respective user-to-relay channel. Then, each user spends M time slots to broadcast pilot sequences so that the relay can estimate both relay-to-user channels. Finally, the relay consumes 1 time slot to broadcast the superposition of both users' CSI. Each user utilises the knowledge of his local CSI to decode his partner's CSI. In comparison, conventional DF MIMO one-way relaying with receiver CSI requires $3M$ time slots in order to enable all nodes to estimate their receiver CSI. The proposed

[2]Having the CSI of the partnering channel (partner's user-to-relay channel) not only enables each user to decode the information from his partner, but also facilitates each user to calculate suitable power allocation or power control factors to optimise the overall performance in the network.

two-way relaying protocol only requires one extra time slot in acquiring the desired CSI.

5.2.2 Transmission Protocol

The proposed transmission protocol can be described in two time slots. Figure 5.1 summarises the transmission flow of the proposed protocol. In the first time slot, both users transmit the linear precoded information vector to the relay, i.e. user i transmits $\mathbf{F}_i\mathbf{x}_i$ where $\mathbf{F}_i \in \mathbb{C}^{M\times M}$ is the transmit beamforming matrix of user i and $\mathbf{x}_i \in \mathbb{C}^{M\times 1}$ is the information bearing vector of user i with normalised covariance, i.e. $E\left[\mathbf{x}_i\mathbf{x}_i^H\right] = \mathbf{I}_M$. The design of $\mathbf{F}_i$ will be discussed in the following section. The signal observed by the relay can be expressed as

$$\mathbf{r} = \mathbf{H}_1\mathbf{F}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{F}_2\mathbf{x}_2 + \mathbf{n}_r, \tag{5.1}$$

where $\mathbf{H}_i \in \mathbb{C}^{M\times M}$ is the channel from user i to the relay and $\mathbf{n}_r \in \mathbb{C}^{M\times 1}$ is the noise vector observed by the relay. In the second time slot, the relay broadcasts the linear precoded observation to both users, i.e. $\mathbf{W}\mathbf{r}$ where $\mathbf{W} \in \mathbb{C}^{M\times M}$ is the joint receive and transmit beamforming matrix at the relay and $\mathbf{r} \in \mathbb{C}^{M\times 1}$ is the observation in the first time slot, expressed in (5.1). The relay beamforming matrix $\mathbf{W}$ will be discussed in the following section. The signal received by user i is

$$\mathbf{y}_i = \mathbf{H}_i^T\mathbf{W}\left(\mathbf{H}_1\mathbf{F}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{F}_2\mathbf{x}_2 + \mathbf{n}_r\right) + \mathbf{n}_i, \tag{5.2}$$

where $\mathbf{n}_i \in \mathbb{C}^{M\times 1}$ is the noise vector observed by user i. Each user performs linear post-processing to the received mixture broadcast by the relay, i.e. user i calculates $\mathbf{G}_i\mathbf{y}_i$ where $\mathbf{G}_i \in \mathbb{C}^{M\times M}$ is the receive beamforming matrix for user i. Self interference of user i contains the information transmitted by user i himself in the previous time slot, i.e. $\mathbf{G}_1\mathbf{H}_1^T\mathbf{W}\mathbf{H}_1\mathbf{F}_1\mathbf{x}_1$ is the self interference of user 1. Using the principle of analogue network coding [5], the self interference is subtracted from the mixture and the desired information vector can be decoded. Writing $\mathbf{A}_i = \mathbf{G}_i\mathbf{H}_i^T\mathbf{W}\mathbf{H}_{\bar{i}}\mathbf{F}_{\bar{i}}$ and $\mathbf{B}_i = \mathbf{G}_i\mathbf{H}_i^T\mathbf{W}$, and assume Gaussian channel coding, the mutual information of user i can be expressed as

$$R_i = \frac{1}{2}\log_2\left(\det\left(\mathbf{I}_M + \mathbf{A}_i\mathbf{A}_i^H\left(\sigma_r^2\mathbf{B}_i\mathbf{B}_i^H + \sigma_i^2\mathbf{G}_i{\mathbf{G}_i}^H\right)^{-1}\right)\right), \tag{5.3}$$

where σ_r^2 is the receiver noise power at the relay, σ_i^2 is the receiver noise power at user i and the subscript $\bar{i}$ is used to denote the complement of i, i.e. when $i = 1$, $\bar{i} = 2$. The pre-log factor reflects the two time slots used to complete the information exchange.

5.3 Beamforming Design

In this section, the proposed low-complexity design of transmit beamformer at the users $\mathbf{F}_i$, $\forall i = \{1, 2\}$, joint receive and transmit beamformer at the relay $\mathbf{W}$ and receive beamformer at the users $\mathbf{G}_i$, $\forall i = \{1, 2\}$ are described. The objective of the beamforming design is to decompose the channels into M parallel subchannels, which not only facilitates substream power allocation and power control, but also enables the use of simple SISO decoders at the users.

5.3.1 Design of $\mathbf{F}_i$

Recall that in a conventional point-to-point MIMO system, the optimal transmit and receive beamformers are designed by means of singular value decomposition (SVD), such that the channel matrix is decomposed into parallel subchannels (or eigen-modes) to enable optimal power sharing among subchannels [22]. However, this cannot be directly implemented in the two-way relaying scenario considered here, as mentioned in [14]. This is due to the fact that the relay (acts as MIMO receiver) is not able to simultaneously separate the subchannels from user 1 and user 2 which have different channel directions.

To address the above issue, subchannel alignment is proposed in the design of $\mathbf{F}_i$ to ensure that kth subchannel of user 1 and kth subchannel of user 2 occupy the same signal subspace. Specifically, the following structure for the transmit beamformer of user i is proposed

$$\mathbf{F}_i = \tilde{\mathbf{F}}_i \mathbf{V}_i \mathbf{\Sigma}_i, \tag{5.4}$$

where alignment matrix $\tilde{\mathbf{F}}_i \in \mathbb{C}^{M \times M}$ is obtained from subchannel alignment[3],

$$\mathbf{H}_1 \tilde{\mathbf{F}_1} = \mathbf{H}_2 \tilde{\mathbf{F}_2}. \tag{5.5}$$

[3]Note that $\tilde{\mathbf{F}}_i$ is not unique. However, multiplication of $\tilde{\mathbf{F}}_i$ with a unitary matrix (rotation matrix) does not change the singular values of the effective channels, i.e. $\mathbf{\Lambda}_i$ remains the same.

The subchannel alignment problem can be solved as follows,

$$\begin{bmatrix} \tilde{\mathbf{F}}_1 \\ -\tilde{\mathbf{F}}_2 \end{bmatrix} = \text{null}\left(\begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \end{bmatrix}\right), \tag{5.6}$$

where the computation of the null-space vectors can be found in [27]. Matrix $\mathbf{V}_i \in \mathbb{C}^{M\times M}$ in (5.4) is the right singular matrix obtained from the SVD of $\mathbf{H}_i\tilde{\mathbf{F}}_i$, i.e. $\mathbf{H}_i\tilde{\mathbf{F}}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$ where $\mathbf{U}_i \in \mathbb{C}^{M\times M}$ is the left singular matrix and $\mathbf{\Lambda}_i \in \mathbb{R}^{M\times M}$ is the diagonal matrix of singular values. The diagonal matrix $\mathbf{\Sigma}_i \in \mathbb{R}^{M\times M}$ in (5.4) is the transmit power allocation matrix of user i. The transmit power consumption of user i is $\|\mathbf{F}_i\|_F^2$. Due to subchannel alignment in (5.5), $\mathbf{U}_1\mathbf{\Lambda}_1\mathbf{V}_1^H = \mathbf{U}_2\mathbf{\Lambda}_2\mathbf{V}_2^H$. The subscripts of $\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$ are omitted for simplicity of notation. The received signal at the relay in (5.1) reduces to

$$\mathbf{r} = \mathbf{U}\mathbf{\Lambda}\left(\mathbf{\Sigma}_1\mathbf{x}_1 + \mathbf{\Sigma}_2\mathbf{x}_2\right) + \mathbf{n}_r. \tag{5.7}$$

5.3.2 Design of W

The design of joint receive and transmit beamformer $\mathbf{W}$ ensures that the received signal in (5.7) can be decomposed into parallel streams. To achieve the objective of subchannel decomposition, the following structure is proposed

$$\mathbf{W} = \mathbf{U}^*\mathbf{\Sigma}_r\mathbf{U}^H, \tag{5.8}$$

where $\mathbf{U}$ is the left singular matrix of $\mathbf{H}_i\tilde{\mathbf{F}}_i$, diagonal matrix $\mathbf{\Sigma}_r \in \mathbb{R}^{M\times M}$ is the transmit power allocation matrix at the relay. Notice that the use of subchannel alignment discussed in the previous subsection enables the relay to decompose the channels of user 1 and user 2 simultaneously. The total transmit power consumption at the relay is $\|\mathbf{W}\mathbf{H}_1\mathbf{F}_1\|_F^2 + \|\mathbf{W}\mathbf{H}_2\mathbf{F}_2\|_F^2 + \sigma_r^2\|\mathbf{W}\|_F^2$, where σ_r^2 is the noise power in Watts at the relay.

5.3.3 Design of $\mathbf{G}_i$

The design of the receive beamformer $\mathbf{G}_i$ is to ensure that the received signal in (5.2) can be decomposed into parallel streams. Specifically, the following structure is proposed

$$\mathbf{G}_i = \mathbf{V}^T \tilde{\mathbf{F}}_i^T, \tag{5.9}$$

which is the transposition of the transmit beamforming matrix $\mathbf{F}_i$ but without the power allocation matrix. The signal received by user i in (5.2) after the receive beamforming in (5.9) is applied simplifies to

$$\mathbf{G}_i \mathbf{y}_i = \mathbf{\Lambda}^2 \mathbf{\Sigma}_r \left(\mathbf{\Sigma}_1 \mathbf{x}_1 + \mathbf{\Sigma}_2 \mathbf{x}_2 \right) + \tilde{\mathbf{n}}_i, \tag{5.10}$$

where $\tilde{\mathbf{n}}_i = \mathbf{\Lambda}\mathbf{\Sigma}_r \mathbf{U}^H \mathbf{n}_r + \mathbf{V}^T \tilde{\mathbf{F}}_i^T \mathbf{n}_i$ is the effective noise observed by user i. As described in Section 5.2, each user is able to decode the desired information by subtracting the self interference from the mixture. For instance, user 1 is able to decode $\mathbf{x}_2$ by subtracting the self interference, $\mathbf{\Lambda}^2 \mathbf{\Sigma}_r \mathbf{\Sigma}_1 \mathbf{x}_1$, from the received mixture.

5.4 Joint Power Allocation

This section investigates the joint power allocation problem of the proposed beamforming scheme. Firstly, the SNR of each subchannel is derived. Secondly, the joint power allocation problem is formulated using the sum-rate criterion and the convexity of the optimisation problem is verified. Since the objective function is non-concave, an upper bound is derived to approximate the original objective function. The last subsections discuss the proposed power allocation strategies, baseline schemes and a comparable scheme.

5.4.1 Subchannel SNR Derivation

From (5.10), it can be observed that the channel matrices are decomposed into M parallel subchannels. In this subsection, the SNR of each subchannel is derived. Denote the transmit power allocation matrix of user 1, $\mathbf{\Sigma}_1 = \mathrm{diag}(\ \sqrt{a_1}\ \ \ldots\ \ \sqrt{a_M}\ )$, the transmit power allocation matrix of

user 2, $\mathbf{\Sigma}_2 = \text{diag}(\ \sqrt{b_1}\ \ \ldots\ \ \sqrt{b_M}\ )$, the transmit power allocation matrix of the relay, $\mathbf{\Sigma}_r = \text{diag}(\ \sqrt{c_1}\ \ \ldots\ \ \sqrt{c_M}\ )$, the diagonal matrix of singular values, $\mathbf{\Lambda} = \text{diag}(\ \sqrt{\lambda_1}\ \ \ldots\ \ \sqrt{\lambda_M}\ )$, and $\hat{\mathbf{F}}_i = \tilde{\mathbf{F}}_i \mathbf{V}_i$. The variables a_k, b_k and c_k represent the kth substream power allocation factors for user 1, user 2 and the relay respectively. Assuming a SISO decoder is used to decode each parallel stream, the SNR of the kth subchannel of user 1 can be expressed as follows,

$$\gamma_{1,k} = \frac{\lambda_k^2 b_k c_k}{\sigma_r^2 \lambda_k c_k + \sigma_1^2 \sum_{j=1}^{M} |\hat{\mathbf{F}}_1(j,k)|^2}, \tag{5.11}$$

where σ_1^2 is the noise power at user 1 receiver. Similarly, the SNR of the kth subchannel of user 2 can be written as

$$\gamma_{2,k} = \frac{\lambda_k^2 a_k c_k}{\sigma_r^2 \lambda_k c_k + \sigma_2^2 \sum_{j=1}^{M} |\hat{\mathbf{F}}_2(j,k)|^2}, \tag{5.12}$$

where σ_2^2 is the noise power at user 2 receiver. Assuming Gaussian channel coding, the instantaneous data rate (or mutual information) of user i can be expressed as

$$R_i = \frac{1}{2} \sum_{k=1}^{M} \log_2 \left(1 + \gamma_{i,k}\right), \tag{5.13}$$

where the pre-log factor reflects the two time slots used to complete the information exchange. In realistic wireless system using practical channel coding scheme, the instantaneous data rate of user i can be modified as

$$R_i' = \frac{1}{2} \sum_{k=1}^{M} \log_2 \left(1 + \frac{\gamma_{i,k}}{\Gamma}\right), \tag{5.14}$$

which includes the SNR gap Γ to account for the target error probability and the specific channel coding scheme [25]. The SNR gap has a typical value of $\Gamma \geq 1$ where the special case $\Gamma = 1$ corresponds to the upper bound in (5.13) where Gaussian coding is used. Since Γ is independent of the channel, the formulation in this chapter assumes $\Gamma = 1$ without loss of generality.

5.4.2 Sum-Rate Optimisation

Sum-rate criterion, i.e. $R_1 + R_2$, is the optimisation utility used in this chapter. All transmissions in the network are subject to a total network power constraint P in Watts, which can be written as

$$\sum_{i=1}^{2}\left(\|\mathbf{F}_i\|_F^2+\|\mathbf{W}\mathbf{H}_i\mathbf{F}_i\|_F^2\right)+\sigma_r^2\|\mathbf{W}\|_F^2\leq P. \tag{5.15}$$

The joint power constraint is the summation of the transmit power consumption at user 1, user 2 and the relay. The joint power constraint expression can be simplified. Specifically, the transmit power consumption at the relay can be simplified as $\|\mathbf{W}\mathbf{H}_1\mathbf{F}_1\|_F^2+\|\mathbf{W}\mathbf{H}_2\mathbf{F}_2\|_F^2+\sigma_r^2\|\mathbf{W}\|_F^2=\sum_{k=1}^{M}\left(\lambda_k a_k c_k+\lambda_k b_k c_k+\sigma_r^2 c_k\sum_{j=1}^{M}|\mathbf{U}(j,k)|^2\right)$. Note that $\sum_{j=1}^{M}|\mathbf{U}(j,k)|^2=1$. Similarly, one can simplify $\|\mathbf{F}_1\|_F^2=\sum_{k=1}^{M}\sum_{j=1}^{M}a_k|\hat{\mathbf{F}}_1(j,k)|^2$, and $\|\mathbf{F}_2\|_F^2=\sum_{k=1}^{M}\sum_{j=1}^{M}b_k|\hat{\mathbf{F}}_2(j,k)|^2$. To further simplify the expression, one can represent $\tilde{a}_k=a_k\sum_{j=1}^{M}|\hat{\mathbf{F}}_1(j,k)|^2$, $\tilde{b}_k=b_k\sum_{j=1}^{M}|\hat{\mathbf{F}}_2(j,k)|^2$ and $\tilde{c}_k=c_k\left(\lambda_k a_k+\lambda_k b_k+\sigma_r^2\right)$ are the effective kth substream power allocation factor for user 1, user 2 and the relay respectively.

The joint power allocation problem using sum-rate criterion[4] can be formulated as,

$$\begin{aligned}\underset{\tilde{a}_1,\ldots,\tilde{a}_M,\tilde{b}_1,\ldots,\tilde{b}_M,\tilde{c}_1,\ldots,\tilde{c}_M}{\text{maximise}}\ &\frac{1}{2}\sum_{k=1}^{M}\log_2\left(1+\frac{t_{1,k}\tilde{b}_k\tilde{c}_k}{t_{2,k}\tilde{a}_k+t_{3,k}\tilde{b}_k+t_{4,k}\tilde{c}_k+t_{5,k}}\right)\\&+\frac{1}{2}\sum_{k=1}^{M}\log_2\left(1+\frac{u_{1,k}\tilde{a}_k\tilde{c}_k}{u_{2,k}\tilde{a}_k+u_{3,k}\tilde{b}_k+u_{4,k}\tilde{c}_k+u_{5,k}}\right),\end{aligned} \tag{5.16}$$

$$\text{subject to}\quad \sum_{k=1}^{M}\left(\tilde{a}_k+\tilde{b}_k+\tilde{c}_k\right)\leq P,\tilde{a}_k\geq 0,\tilde{b}_k\geq 0,\tilde{c}_k\geq 0,\ \forall_k=\{1,\ldots,M\}, \tag{5.17}$$

where the constants $t_{1,k}=\lambda_k^2$, $t_{2,k}=\sigma_1^2\beta_k\lambda_k$, $t_{3,k}=\sigma_1^2\alpha_k\lambda_k$, $t_{4,k}=\sigma_r^2\beta_k\lambda_k$, $t_{5,k}=\sigma_1^2\sigma_r^2\alpha_k\beta_k$, $u_{1,k}=t_{1,k}=\lambda_k^2$, $u_{2,k}=\sigma_2^2\beta_k\lambda_k$, $u_{3,k}=\sigma_2^2\alpha_k\lambda_k$, $u_{4,k}=\sigma_r^2\alpha_k\lambda_k$, $u_{5,k}=\sigma_2^2\sigma_r^2\alpha_k\beta_k$, $\alpha_k=\sum_{j=1}^{M}|\hat{\mathbf{F}}_1(j,k)|^2$ and $\beta_k=\sum_{j=1}^{M}|\hat{\mathbf{F}}_2(j,k)|^2$. The optimisation problem can be solved using convex optimisation techniques [20] if the constraints are convex and the objective is concave. The inequality of the power constraint in (5.17) is affine, hence convex (and concave), with respect to all input parameters $\tilde{a}_1,\ldots,\tilde{a}_M,\tilde{b}_1,\ldots,\tilde{b}_M,\tilde{c}_1,\ldots,\tilde{c}_M$. However, it can be shown that the objective function in (5.16) is non-concave with respect to all input parameters. The detailed convexity analysis can be found in Appendix D.

It is shown that the objective function in (5.16) is non-concave with respect to the parameters $\tilde{a}_1,\ldots,\tilde{a}_M$, $\tilde{b}_1,\ldots,\tilde{b}_M$, $\tilde{c}_1,\ldots,\tilde{c}_M$. In order to ease the difficulty in solving the power allocation problem, a concave upper bound of the original objective function which can be solved efficiently

[4]The sum-rate optimisation in (5.16) with constraints in (5.17) can be easily modified to include the special case where each node has fixed power constraint. Specifically, the joint power constraint in (5.17) is separated into three individual constraints, i.e. $\sum_{k=1}^{M}\tilde{a}_k=P_1$, $\sum_{k=1}^{M}\tilde{b}_k=P_2$ and $\sum_{k=1}^{M}\tilde{c}_k=P_r$ where P_1, P_2 and P_r are the individual power constraint at user 1, user 2 and the relay respectively.

using convex optimisation techniques is derived. The following theorem summarises the concavity of the derived upper bound.

Theorem 5.1. *The following upper bound of the objective function is jointly concave with respect to input parameters* $\tilde{a}_1, \ldots, \tilde{a}_M, \tilde{b}_1, \ldots, \tilde{b}_M, \tilde{c}_1, \ldots, \tilde{c}_M$,

$$f_{upper} = \frac{1}{2}\sum_{k=1}^{M} \log_2 \left(1 + \frac{\left(\tilde{a}_k + \tilde{b}_k\right)\tilde{c}_k}{t_{a,k}\left(\tilde{a}_k + \tilde{b}_k\right) + t_{b,k}\tilde{c}_k}\right) + \frac{1}{2}\sum_{k=1}^{M} \log_2 \left(1 + \frac{\left(\tilde{a}_k + \tilde{b}_k\right)\tilde{c}_k}{u_{a,k}\left(\tilde{a}_k + \tilde{b}_k\right) + u_{b,k}\tilde{c}_k}\right), \tag{5.18}$$

where the constants $t_{a,k} = \frac{\min(t_{2,k}, t_{3,k})}{t_{1,k}}$, $t_{b,k} = \frac{t_{4,k}}{t_{1,k}}$, $u_{a,k} = \frac{\min(u_{2,k}, u_{3,k})}{u_{1,k}}$ *and* $u_{b,k} = \frac{u_{4,k}}{u_{1,k}}$.

Proof. Refer to Appendix D. ■

Remark 5.1. The power allocation factors, $\tilde{a}_1, \ldots, \tilde{a}_M, \tilde{b}_1, \ldots, \tilde{b}_M, \tilde{c}_1, \ldots, \tilde{c}_M$, obtained by solving the concave upper bound in (5.18) are sub-optimal solutions to the original problem in (5.16). The approximation in (5.18) enables the transmission power to be allocated dynamically between users and relay, while the users share identical power allocation factors. Since the positive sum of the power allocation factors of both users, i.e. $\tilde{a}_k + \tilde{b}_k$, can be represented as a single power allocation factor, the result is a combination of dynamic power sharing between users and relay, coupled with equal power sharing between users.

5.4.3 Proposed Power Allocation Strategies

In this subsection, two joint power allocation (JPA) strategies are proposed.

5.4.3.1 Proposed JPA I

The proposed JPA I computes the power allocation factors by solving the sum-rate optimisation problem in (5.16) and (5.17). As discussed in the previous subsection, the objective function in (5.16) is non-concave. In this case, the locally optimal solution does not necessarily correspond to the globally optimal solution. The globally optimal solution can be found with a certain probability by means of

randomisation based global optimisation [16]. For each channel realisation, multiple random starting vectors are generated and the local optimal solution for each starting vector is computed using convex optimisation techniques, i.e. the interior-point method [20]. The globally optimal solution for each channel realisation is the maximum of all local optimal solutions. Since this method requires the use of multiple random starting vectors, a centralised node, i.e. the relay, will compute the power allocation factors and distribute them to other nodes.

5.4.3.2 Proposed JPA II

The proposed JPA II computes the power allocation factors by solving the concave upper bound in (5.18). The power allocation factors can be calculated efficiently using the convex optimisation techniques, i.e. the interior-point method [20]. Since the upper bound objective function in (5.18) is concave, the local optimal solution obtained using convex optimisation corresponds to the global optimal solution. The computed power allocation factors correspond to the globally optimal solution of (5.18) are then substituted back into the original objective function in (5.16) to obtain the achievable sum-rate. With the CSI knowledge of the channels, each node is able to compute the power allocation factors locally, without any cooperation between nodes. Specifically, perfect knowledge of both $\mathbf{H}_1$ and $\mathbf{H}_2$ enables every node to have common knowledge of all the constants in (5.18). This allows all nodes to compute the solution to the same optimisation problem locally. The CSI training scheme to enable all nodes to have the CSI of both $\mathbf{H}_1$ and $\mathbf{H}_2$ is discussed in Subsection 5.2.1.

5.4.4 Baseline Schemes and Comparable Scheme

In this subsection, two baseline schemes, pure AF MIMO two-way relaying and DF MIMO one-way relaying schemes are presented. The best comparable scheme [15] is also discussed.

5.4.4.1 Baseline Scheme: Pure AF

In the pure AF two-way relaying scheme, the relay simply forwards the power normalised observation to the users, without any beamforming or power allocation. Assuming optimal MIMO decoders and equal power allocation, sum-rate can be computed using (5.3) with $\mathbf{F}_i = \sqrt{\frac{P_i}{M}}\mathbf{I}$, $\mathbf{G}_i = \mathbf{I}$ and $\mathbf{W} = \sqrt{\frac{P_r}{\|\mathbf{H}_1\mathbf{F}_1\|_F^2+\|\mathbf{H}_2\mathbf{F}_2\|_F^2+\sigma_r^2 M}}\mathbf{I}$ where P_i is the power constraint at user i and P_r is the power constraint at the relay. It is assumed that $P_1 = P_2 = P_r = \frac{P}{3}$ so that the total network power constraint P

is satisfied. This scheme serves as a baseline to study the contribution of power allocation to the sum-rate of the two-way relaying channel.

5.4.4.2 Baseline Scheme: MIMO One-Way

Another baseline scheme used for comparison is the DF MIMO one-way relaying. Recall that in one-way relaying, four orthogonal channel uses are consumed to complete the information exchange between user pair. In the first time slot, user 1 transmits information to the relay. After decoding the received information, the relay forwards the observation to user 2 in the second time slot. Following a similar fashion, user 2 transmits information to user 1 with the help of relay using another two time slots.

Assuming perfect transmitter and receiver CSI are available, the channel matrix from user i to relay, $\mathbf{H}_i$, can be decomposed into M parallel streams using the SVD, i.e. $\mathbf{H}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$ where $\mathbf{\Lambda}_i = \mathrm{diag}(\sqrt{\lambda_{i,1}}, \ldots, \sqrt{\lambda_{i,M}})$. Note that the $\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$ and $\lambda_{i,k}$ defined here are different from those in Subsection 5.3.1 and 5.4.1. Denote a_k and b_k as the kth substream power allocation factor of user 1 and user 2 respectively. Variable c_k is defined as the kth substream power allocation factor of the relay for transmission to user 1 while variable d_k is the kth substream power allocation factor of the relay for transmission to user 2. Represent $R_{i\to r}$ as the data rate from user i to relay while $R_{r\to i}$ as the data rate from the relay to user i. The data rate at user 1 can be expressed as

$$\begin{aligned} \mathcal{I}_1 &= \min(R_{2\to r}, R_{r\to 1}) && (5.19)\\ &\quad \min\left(\frac{1}{4}\sum_{k=1}^{M}\log_2\left(1+\frac{b_k\lambda_{2,k}}{\sigma_r^2}\right), \frac{1}{4}\sum_{k=1}^{M}\log_2\left(1+\frac{c_k\lambda_{1,k}}{\sigma_1^2}\right)\right), && (5.20)\end{aligned}$$

while the data rate at user 2 can be written as

$$\begin{aligned} \mathcal{I}_2 &= \min(R_{1\to r}, R_{r\to 2}) && (5.21)\\ &\quad \min\left(\frac{1}{4}\sum_{k=1}^{M}\log_2\left(1+\frac{a_k\lambda_{1,k}}{\sigma_r^2}\right), \frac{1}{4}\sum_{k=1}^{M}\log_2\left(1+\frac{d_k\lambda_{2,k}}{\sigma_2^2}\right)\right). && (5.22)\end{aligned}$$

Transmission power is allocated equally among nodes and data streams. Recall that in the previous subsection, the network is subject to a total power constraint of P, for every two time slots. Therefore, the power allocation factors are defined as $a_k = \frac{P}{2M}$, $b_k = \frac{P}{2M}$, $c_k = \frac{P}{2M}$, and $d_k = \frac{P}{2M}$, $\forall k =$

$\{1, \ldots, M\}$.

5.4.4.3 Comparable Scheme: Alternate Optimisation (A-Opt)

The best comparable scheme for the non-regenerative MIMO two-way relaying channel is the A-Opt scheme proposed in [15]. Due to the fact that the sum-rate expression computable from (5.3) is non-concave, [15] proposes the A-Opt scheme which alternately computes locally optimal source beamformers for fixed relay beamformer, and locally optimal relay beamformer for fixed source beamformers until convergence is reached. Several searching algorithms are proposed in [15] to determine locally optimal beamforming matrices subject to individual power constraints and assuming the use of perfect MIMO decoders at the users. Although the A-Opt scheme is able to achieve the best sum-rate under individual power constraints and symmetric SNR, it is computationally expensive to determine the beamforming matrices. Generally, the problem dimension of the searching algorithms in [15] grows quadratically with the number of antennas. This increases the computational complexity significantly when a higher number of antennas is used. In comparison, the problem dimension of the proposed JPA I and II in this chapter is linear with the number of antennas. Furthermore, due to the fact that the sum-rate is non-concave for any fixed source beamformers, the searching algorithms in [15] have to be repeated multiple times with different starting points in order to increase the probability of finding the global optimal relay beamformer. This further increases the computational overhead.

In order to obtain the simulation results for A-Opt, the weighted minimum MSE algorithm proposed in [15] is used to compute the relay beamforming matrix while semi-definite program solver in CVX toolbox [66, 67] is used to compute the user beamforming matrices. Each node is subject to individual power constraint $\frac{P}{3}$, which sums up to a joint power constraint P, to enable fair comparison with the proposed strategies in this chapter.

5.5 Joint Power Control

This section studies the power control problem in guaranteeing the predetermined QoS constraints of the two-way information transmission. Firstly, the joint power control problem is presented. Secondly, several power control strategies under the proposed scheme and the baseline scheme are

explained.

5.5.1 Joint Power Control Optimisation Problem

The objective of the power control policy is to minimise the total power constraint in the network, subject to a kth substream rate constraint of user 1, $R_{1,k}$ and a kth substream rate constraint of user 2, $R_{2,k}$, $\forall k = \{1, \ldots, M\}$. Specifically, the kth substream rate constraint of user 1 and user 2 can be expressed as $\frac{1}{2}\log_2(1+\gamma_{1,k}) \geq R_{1,k}$ and $\frac{1}{2}\log_2(1+\gamma_{2,k}) \geq R_{2,k}$ respectively, where the kth substream SNR $\gamma_{1,k}$ and $\gamma_{2,k}$ can be found in (5.11) and (5.12) respectively. The kth substream rate constraints serve as criteria to guarantee the QoS specified on the kth data stream. Notice that the kth substream rate constraints can also be expressed as kth substream SNR constraints, i.e. $\gamma_{1,k} = 2^{2R_{1,k}} - 1$, $\gamma_{2,k} = 2^{2R_{2,k}} - 1$, $\forall k = \{1, \ldots, M\}$.

Recall that the power consumption at user 1 is a function of the power allocation factors at user 1, i.e. $P_1\left(\tilde{a}_1, \ldots, \tilde{a}_M\right) = \sum_{k=1}^{M} \tilde{a}_k$ where $\tilde{a}_k = a_k \sum_{j=1}^{M} |\hat{\mathbf{F}}_1(j,k)|^2$. Similarly, the power consumption at user 2 is $P_2\left(\tilde{b}_1, \ldots, \tilde{b}_M\right) = \sum_{k=1}^{M} \tilde{b}_k$ where $\tilde{b}_k = b_k \sum_{j=1}^{M} |\hat{\mathbf{F}}_2(j,k)|^2$, while at the relay is $P_r\left(\tilde{c}_1, \ldots, \tilde{c}_M\right) = \sum_{k=1}^{M} \tilde{c}_k$ where $\tilde{c}_k = c_k\left(\lambda_k a_k + \lambda_k b_k + \sigma_r^2\right)$. After some algebraic manipulations, the power control problem which minimises the total power consumption in the network, subject to the substream rate constraints can be formulated as follows,

$$\underset{\tilde{a}_1,\ldots,\tilde{a}_M,\tilde{b}_1,\ldots,\tilde{b}_M,\tilde{c}_1,\ldots,\tilde{c}_M}{\text{minimise}} \quad P_1\left(\tilde{a}_1, \ldots, \tilde{a}_M\right) + P_2\left(\tilde{b}_1, \ldots, \tilde{b}_M\right) + P_r\left(\tilde{c}_1, \ldots, \tilde{c}_M\right) \tag{5.23}$$

$$\text{subject to} \quad \gamma_{1,k}\left(r_{1,k}\tilde{a}_k\tilde{b}_k^{-1}\tilde{c}_k^{-1} + r_{2,k}\tilde{c}_k^{-1} + r_{3,k}\tilde{b}_k^{-1} + r_{4,k}\tilde{b}_k^{-1}\tilde{c}_k^{-1}\right) \leq 1, \tag{5.24}$$

$$\gamma_{2,k}\left(s_{1,k}\tilde{c}_k^{-1} + s_{2,k}\tilde{a}_k^{-1}\tilde{b}_k\tilde{c}_k^{-1} + s_{3,k}\tilde{a}_k^{-1} + s_{4,k}\tilde{a}_k^{-1}\tilde{c}_k^{-1}\right) \leq 1, \tag{5.25}$$

$$\forall_k = \{1, \ldots, M\},$$

where the constants $r_{1,k} = \sigma_1^2\beta_k\lambda_k^{-1}$, $r_{2,k} = \sigma_1^2\alpha_k\lambda_k^{-1}$, $r_{3,k} = \sigma_r^2\beta_k\lambda_k^{-1}$, $r_{4,k} = \lambda_k^{-2}\sigma_1^2\sigma_r^2\alpha_k\beta_k$, $s_{1,k} = \sigma_2^2\beta_k\lambda_k^{-1}$, $s_{2,k} = \sigma_2^2\alpha_k\lambda_k^{-1}$, $s_{3,k} = \sigma_r^2\alpha_k\lambda_k^{-1}$, $s_{4,k} = \lambda_k^{-2}\sigma_2^2\sigma_r^2\alpha_k\beta_k$, $\alpha_k = \sum_{j=1}^{M} |\hat{\mathbf{F}}_1(j,k)|^2$ and $\beta_k = \sum_{j=1}^{M} |\hat{\mathbf{F}}_2(j,k)|^2$.

The optimisation problem is in the form of geometric program, since the objective is an affine function (which can be generalised as a posynomial function [20]) and the constraints are posynomial functions. Although the original problem is not convex, the geometric program can be transformed

into equivalent convex form by means of change of variables and logarithmic transformation of the constraint functions. See Appendix D for details of the transformation.

5.5.2 Power Control Strategies and Baseline Scheme

In this subsection, two power control strategies using the proposed beamforming scheme are presented, namely the joint power control (JPC) and equal power control (EPC) . The baseline scheme for comparison is also explained.

5.5.2.1 Joint Power Control (JPC)

The proposed JPC corresponds to solving the power minimisation problem in (5.23) subject to rate constraints in (5.24) and (5.25). As discussed in the previous subsection, the equivalent power control problem in convex form can be solved efficiently using convex optimisation techniques, i.e. interior point method [20]. The locally optimal solution obtained by convex optimisation then corresponds to the globally optimal solution. The computation of the power allocation factors are performed locally at each node.

5.5.2.2 Equal Power Control (EPC)

In the proposed EPC, each stream from every node is allocated with a common power allocation factor, subject to the rate constraints in (5.24) and (5.25). Under this sub-optimal strategy, the power allocation factors have the following relationship, $\tilde{a}_k = \tilde{b}_k = \tilde{c}_k = \theta^*$, $\forall_k = \{1, \ldots, M\}$, where θ^* is the common power allocation factor. The common power allocation factor θ^* can be obtained analytically by solving the following quadratic equations,

$$-\theta^2 + \gamma_{1,k}\left(r_{1,k} + r_{2,k} + r_{3,k}\right)\theta + \gamma_{1,k} r_{4,k} = 0, \forall_k = \{1, \ldots, M\}, \tag{5.26}$$

$$-\theta^2 + \gamma_{2,k}\left(s_{1,k} + s_{2,k} + s_{3,k}\right)\theta + \gamma_{2,k} s_{4,k} = 0, \forall_k = \{1, \ldots, M\}, \tag{5.27}$$

From (5.26), the discriminant of the quadratic equation is $\triangle = \gamma_{1,k}^2\left(r_{1,k} + r_{2,k} + r_{3,k}\right)^2 + 4\gamma_{1,k} r_{4,k} > 0$ since all constants ($\gamma_{1,k}$,$r_{i,k}$ $\forall_1 = 1, 2, 3, 4$) are positive. When $\triangle > 0$, the real roots are $\theta_1 = \frac{\gamma_{1,k}\left(r_{1,k}+r_{2,k}+r_{3,k}\right)}{2} - \frac{\sqrt{\triangle}}{2} < 0$ and $\theta_2 = \frac{\gamma_{1,k}\left(r_{1,k}+r_{2,k}+r_{3,k}\right)+\sqrt{\triangle}}{2} > 0$. We know that the power allocation

factor is always positive, i.e. $\theta > 0$, therefore we choose $\theta = \theta_2$. Following similar steps, the real root for (5.27) is $\theta = \frac{\gamma_{2,k}\left(s_{1,k}+s_{2,k}+s_{3,k}\right)+\sqrt{\triangle}}{2}$ where $\triangle = \gamma_{2,k}^2\left(s_{1,k}+s_{2,k}+s_{3,k}\right)^2 + 4\gamma_{2,k}s_{4,k}$. The common power allocation factor, θ^* is the maximum of all real roots of $2M$ quadratic equations stated above, in order to ensure that all rate constraints are satisfied. The total power consumption in the network is $3M\theta^*$.

5.5.2.3 Baseline Scheme: MIMO One-Way

Similar to the previous section, the DF MIMO one-way relaying is considered as the baseline scheme for comparison. The details of the DF one-way relaying protocol can be found in subsection 5.4.4.2. For fair comparison, the power control policy for the one-way relaying scheme shall minimise the total power constraint in the network as well, subject to kth substream rate constraint of user 1, $R_{1,k}$ and kth substream rate constraint of user 2, $R_{2,k}$, $\forall k = \{1, \ldots, M\}$. Recall that the power consumption at user 1 is $P(a_1, \ldots, a_M) = \sum_{k=1}^{M} a_k$, the power consumption at user 2 is $P(b_1, \ldots, b_M) = \sum_{k=1}^{M} b_k$ while the power consumption at the relay is $P_r(c_1, \ldots, c_M, d_1, \ldots, d_M) = \sum_{k=1}^{M} c_k + \sum_{k=1}^{M} d_k$. After some algebraic manipulations, the power control problem under one-way relaying can be formulated as follows,

$$\begin{aligned}
\underset{a_1,\ldots,a_M,b_1,\ldots,b_M,c_1,\ldots,c_M,d_1,\ldots,d_M}{\text{minimise}} \quad & P(a_1, \ldots, a_M) + P(b_1, \ldots, b_M) \\
& + P_r(c_1, \ldots, c_M, d_1, \ldots, d_M), && (5.28) \\
\text{subject to} \quad & \psi_{1,k}\sigma_r^2\lambda_{2,k}^{-1}b_k^{-1} \leq 1, \forall k = \{1, \ldots, M\}, && (5.29) \\
& \psi_{1,k}\sigma_1^2\lambda_{1,k}^{-1}c_k^{-1} \leq 1, \forall k = \{1, \ldots, M\}, && (5.30) \\
& \psi_{2,k}\sigma_r^2\lambda_{1,k}^{-1}a_k^{-1} \leq 1, \forall k = \{1, \ldots, M\}, && (5.31) \\
& \psi_{2,k}\sigma_2^2\lambda_{2,k}^{-1}d_k^{-1} \leq 1, \forall k = \{1, \ldots, M\}, && (5.32)
\end{aligned}$$

where the constants $\psi_{1,k} = 2^{4R_{1,k}} - 1$ and $\psi_{2,k} = 2^{4R_{2,k}} - 1$. The optimisation problem above is in the form of geometric program. Using a change of variables and logarithmic transformation of the constraints, the geometric program can be converted into convex form, which can be solved using convex optimisation technique, i.e. the interior-point method.

5.6 Numerical Results

This section presents the numerical results of the proposed joint beamforming and power management scheme in comparison with existing schemes. The optimisation problems discussed in the previous sections are solved using non-linear optimisation toolbox in Matlab i.e. using the function $fmincon$ and the interior-point method. The numerical results are organised into two subsections. In the first subsection, the ergodic sum-rates of various schemes with fixed total power constraint are simulated using the Monte-Carlo method. Refer to Section 5.4 for details of the joint power allocation formulation. In the second subsection, various power control schemes with fixed rate constraints are simulated using the Monte-Carlo method. See Section 5.5 for details of the joint power control formulation.

5.6.1 Power Allocation with Fixed Total Power Constraint

In this subsection, the simulation results of the proposed JPA schemes, baseline schemes and comparable scheme are generated to study the relationship between the ergodic sum-rate and parameters such as SNR, number of antennas and path loss.

Figure 5.2 shows the ergodic sum-rate versus reference SNR ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2}$) of the proposed JPA schemes in comparison with existing schemes. The reference SNR is defined as the inverse of the noise power. In the subsequent discussion, the term SNR is used to imply the reference SNR defined here. In this simulation, the SNR at all nodes are assumed to be symmetrical (equal noise power), i.e. $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2}$. The fixed parameters are the number of antennas, $M = 4$ and the total power constraints, $P = 3$ Watts. From the figure, it can be observed that the proposed JPA I and II perform close to the A-Opt scheme at high SNR. In the range of low to medium SNR, the performance gain contributed by the proposed JPA schemes over baseline pure AF scheme is limited if compared to the A-Opt scheme. This is due to the fact that the choice of beamforming directions in the proposed JPA schemes is sub-optimal if compared to the A-Opt scheme. However, at high SNR, the suboptimal beamforming directions do not prevent the proposed JPA schemes from delivering significant performance gain against the pure AF scheme through dynamic allocation of power among substreams and nodes. It can be observed also that the JPA II performs close to the proposed JPA I. This indicates that the upper bound in Theorem 5.1 is a good approximation of the original problem in (5.16). The performance gaps between two-way relaying schemes (Pure AF, A-Opt, proposed JPA

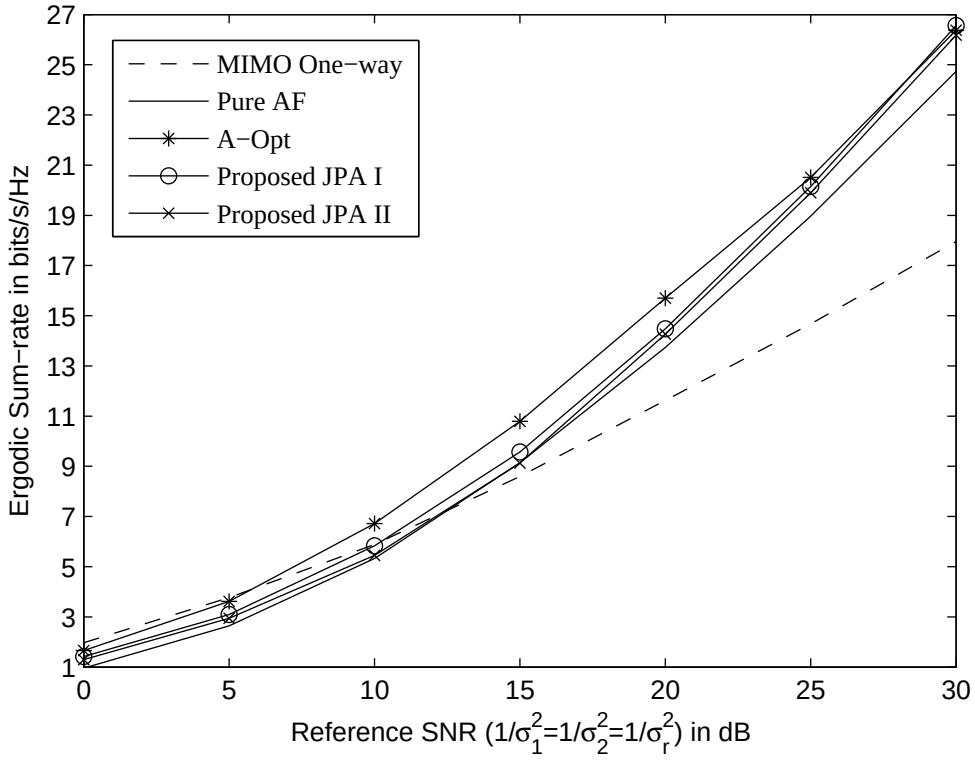


Figure 5.2: Ergodic sum-rate versus SNR ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2}$) for fixed $M = 4$, $P = 3$.

I and II,) and one-way relaying scheme enlarge with the increase of SNR. It is evident from the slope of the sum-rate curves, two-way relaying schemes are able to achieve higher multiplexing gain due to more efficient use of bandwidth.

In the following simulations, the ergodic sum-rates of the various schemes when the SNR at users and relay are asymmetrical (unequal noise power) are simulated. Figure 5.3 shows the ergodic sum-rate versus SNR at the users ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2}$) when SNR at the relay is fixed at $\frac{1}{\sigma_r^2} = 30$dB. Other fixed parameters are $M = 4$ and $P = 3$ Watts. From the figure, it can be observed that the proposed JPA I achieves the best ergodic sum-rate, followed closely by the proposed JPA II. The A-Opt scheme does not perform better than the proposed JPA schemes. This is due to the fact that in A-Opt scheme, each node is allocated with fixed amount of power that does not correlate with the asymmetric SNR. In comparison, the proposed JPA schemes respond to the asymmetric SNR by allocating power dynamically among nodes and substreams. The joint power allocation between nodes provides another dimension of improvement. Figure 5.4 shows the ergodic sum-rate versus SNR at the relay ($\frac{1}{\sigma_r^2}$) when the SNR at the users are fixed at $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 5$dB. Other fixed parameters are $M = 4$ and $P = 3$ Watts. From the figure, it is clear that the proposed JPA schemes deliver significant performance gain over the pure AF scheme. The proposed JPA I and JPA II deliver higher ergodic sum-rate than the A-Opt

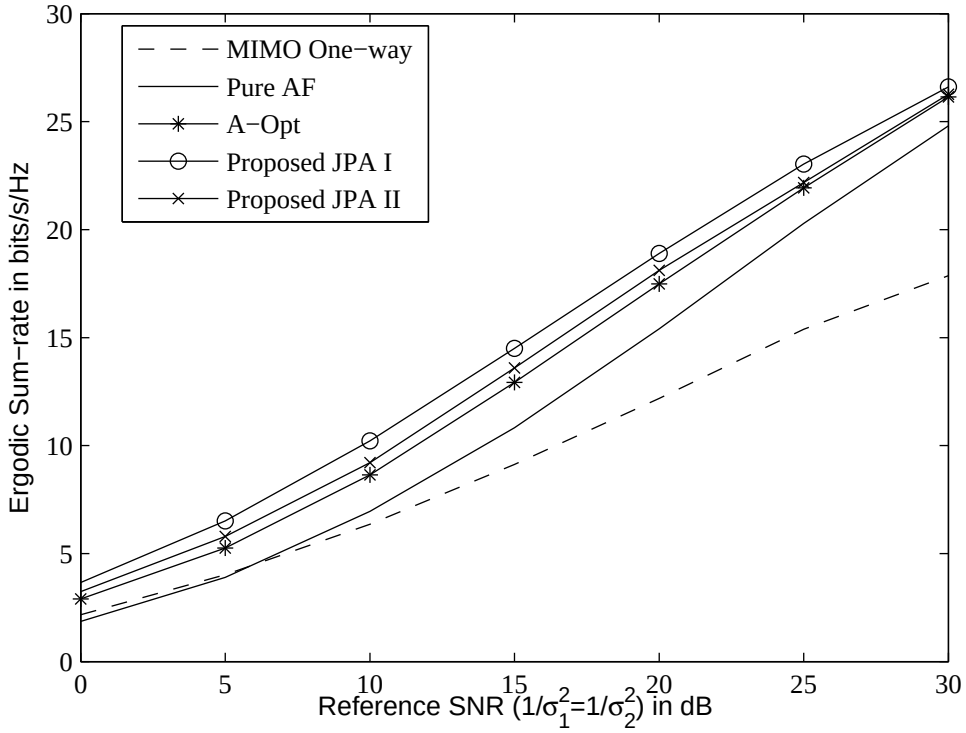


Figure 5.3: Ergodic sum-rate versus SNR ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2}$) for fixed $M = 4$, $P = 3$, and $\frac{1}{\sigma_r^2} = 30$dB.

scheme when SNR is greater than 15dB and 20dB respectively. This supports that dynamic power allocation between users and relay is able to utilise the asymmetric SNR between nodes to obtain better sum-rate performance. At low SNR, the proposed schemes do not perform as well as the A-Opt scheme due to the use of sub-optimal beamforming directions. It is interesting to observe that at low SNR, the baseline MIMO one-way relaying performs as well as the A-Opt scheme. At low SNR, the performance of two-way relaying schemes (Pure AF, A-Opt, proposed JPA I and II) is limited by not only the noise at the users but also the propagated noise from the relay. As a result, two-way relaying schemes do not perform better than one-way relaying schemes at low SNR. This effect can also be observed in Figure 5.2.

Figure 5.5 shows the ergodic sum-rate versus number of antennas M of various schemes. The fixed parameters are $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 15$dB, $\frac{1}{\sigma_r^2} = 30$ dB and $P = 3$ Watts. Generally, the ergodic sum-rates of all schemes increase linearly with the number of antennas, M. As the numbers of antennas at all nodes are increased simultaneously, the number of independent data streams supportable in the network increases. In other words, the multiplexing gain grows linearly with M. Among all power allocation schemes, the proposed JPA I achieves the best sum-rate performance, followed closely by the proposed JPA II. The proposed JPA schemes outperform the A-Opt scheme, thanks to the joint

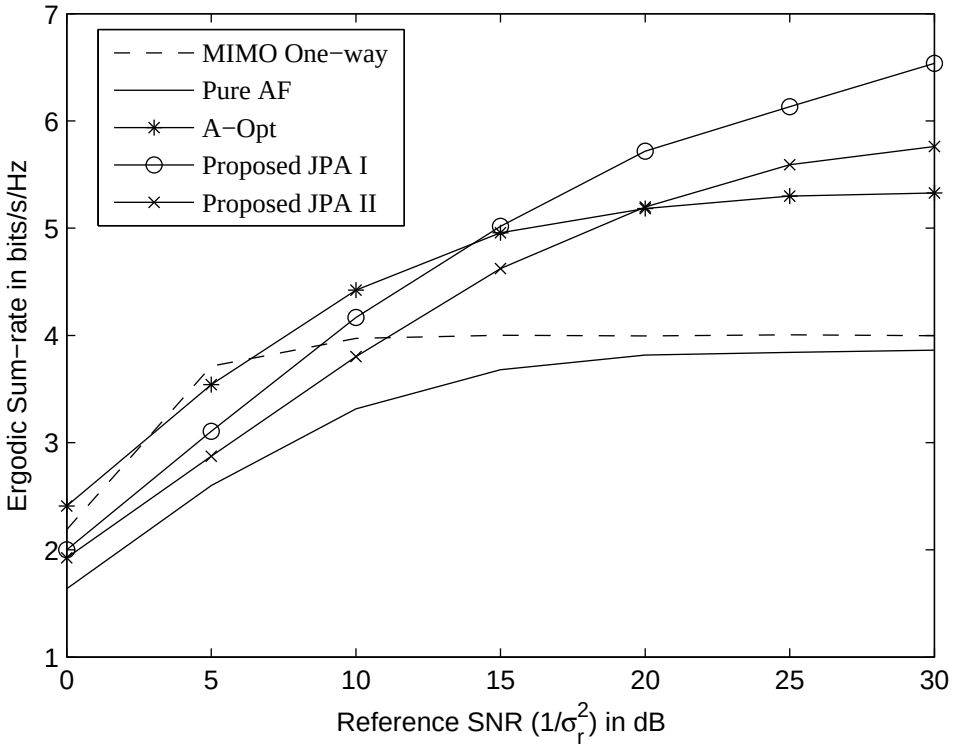


Figure 5.4: Ergodic sum-rate versus SNR ($\frac{1}{\sigma_r^2}$) for fixed $M = 4$, $P = 3$, and $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 5$dB.

power allocation between nodes. From the figure, it can be observed that the gaps between proposed schemes and pure AF scheme enlarge for increasing M. This shows that joint power allocation is vital in delivering better data rates in system with high multiplexing gain.

In the next simulation, the effect of large scale path loss to the ergodic sum-rate is investigated. The path loss is integrated in the channel model as $\frac{1}{\sqrt{d_i^\alpha}}\mathbf{H}_i$, where d_i is the distance between user i and the relay, α is the path loss exponent and $\mathbf{H}_i$ is the channel matrix between user i and the relay (as shown in Section 5.2) where its entries are i.i.d. Rayleigh distributed. A simple line network is considered where the relay is placed in between the users. Figure 5.6 shows the ergodic sum-rate versus the relay location d_r of various schemes. The distance between user 1 and the relay is $d_1 = 1 + d_r$ while the distance between user 2 and the relay is $d_2 = 2 - d_r$. The constant offset in d_i is introduced in order to ensure that the received power does not exceed the transmitted power. The fixed parameters are $M = 2$, $\alpha = 4$, $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 30$ dB, and $P = 3$ Watts. In general, all schemes achieve their best sum-rate when the relay is located in the middle of the users. The proposed JPA I delivers the best sum-rate performance and has the least sensitivity towards the variation of the location of the relay. The proposed JPA II performs close to JPA I but it is more sensitive to unequal path loss due to the fact that both users are allocated equal amount of power. Refer to the remark

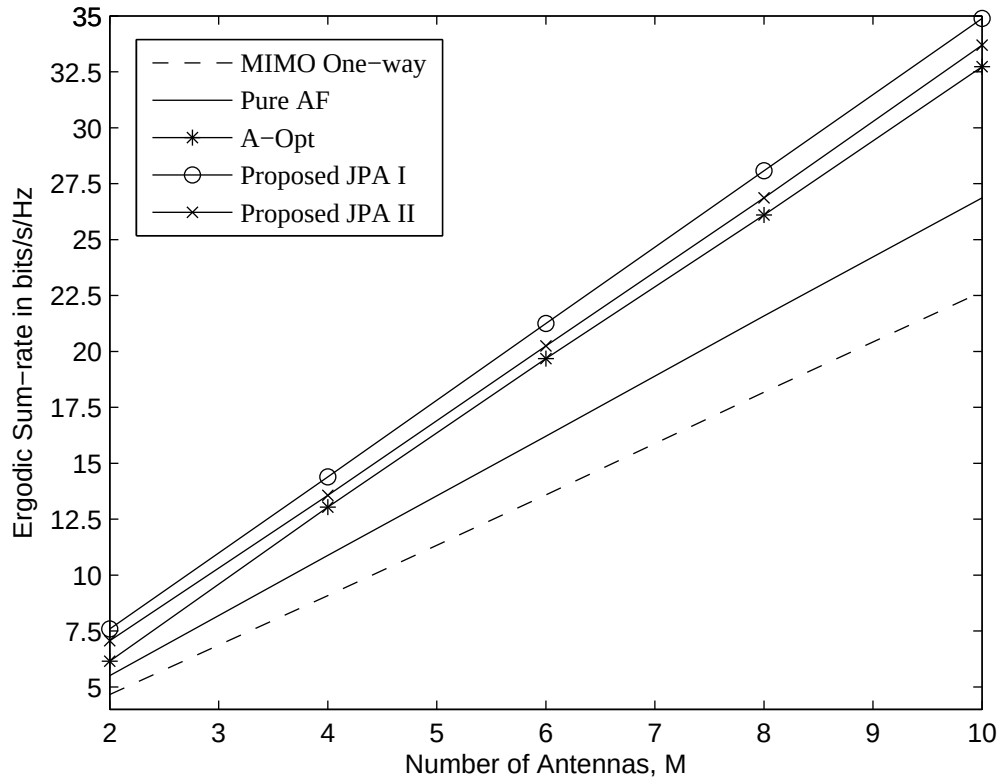


Figure 5.5: Ergodic sum-rate versus number of antennas M, for fixed $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = 15$dB, $\frac{1}{\sigma_r^2} = 30$ dB, and $P = 3$.

of Theorem 5.1. The A-Opt scheme has similar performance as the proposed JPA II. The baseline MIMO one-way relaying scheme displays the worst sum-rate performance and the highest sensitivity towards unequal path loss.

5.6.2 Power Control with Fixed Substream Rate Constraints

In this subsection, the simulation results of various schemes are presented to illustrate the relationship between the average total transmit power consumption (in Watts) and parameters such as SNR and target data rate.

Figure 5.7 shows the average total transmit power consumption in the network versus reference SNR ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2}$) of various schemes. The substream target data rate is assumed to be symmetrical, i.e. $R_{1,k} = R_{2,k} = R$, $\forall k = \{1, \ldots, M\}$. The fixed parameters are the target data rate, $R = 2$ bits/s/Hz and the number of antennas, $M = 2$. From the figure, it is obvious that the proposed OPC is the most energy efficient scheme while the baseline MIMO one-way relaying scheme is the most energy consuming scheme. When the SNR increases, the average total power of all schemes

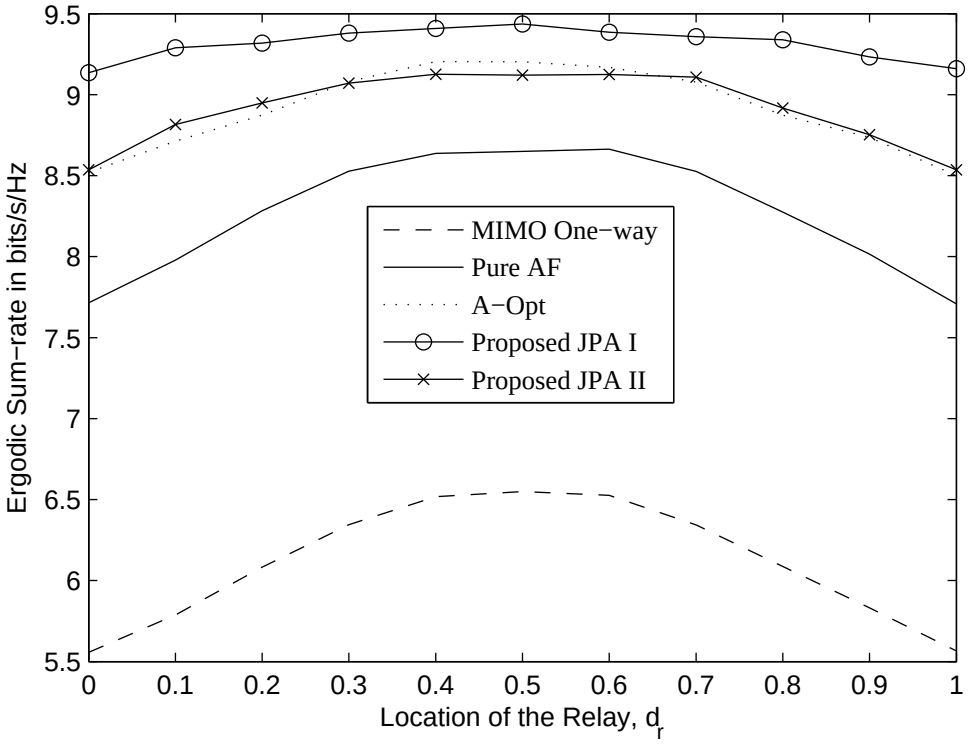


Figure 5.6: Ergodic sum-rate versus the relay location d_r for fixed $M = 2$, $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 30$ dB, $P = 3$, $\alpha = 4$.

decreases exponentially. Note that the y-axis is in logarithmic scale. At higher SNR, the noise power at each substream is lower, therefore the target data rate can be fulfilled easily with lower amount of power.

Figure 5.8 shows the average total transmit power consumption in the network versus target data rate R. Similar to the previous figure, symmetrical substream target data rate is assumed, i.e. $R_{1,k} = R_{2,k} = R$, $\forall k = \{1, \ldots, M\}$. The fixed parameters are $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 20$dB and $M = 2$. From the figure, it can be seen that the proposed OPC scheme consumes the lowest amount of power, followed closely by the proposed EPC scheme. In general, the average total power consumption of all schemes increases with the data rate R. Recall that the data rate is a logarithmic function of signal power, for any fixed noise power. The baseline MIMO one-way relaying scheme displays the most drastic increase of total power as a function of data rate. In comparison, the proposed OPC and EPC demonstrate a subtle increase of power as a function of data rate. These observations verify that the proposed schemes are able to deliver significant power savings particularly in demanding high data rate system.

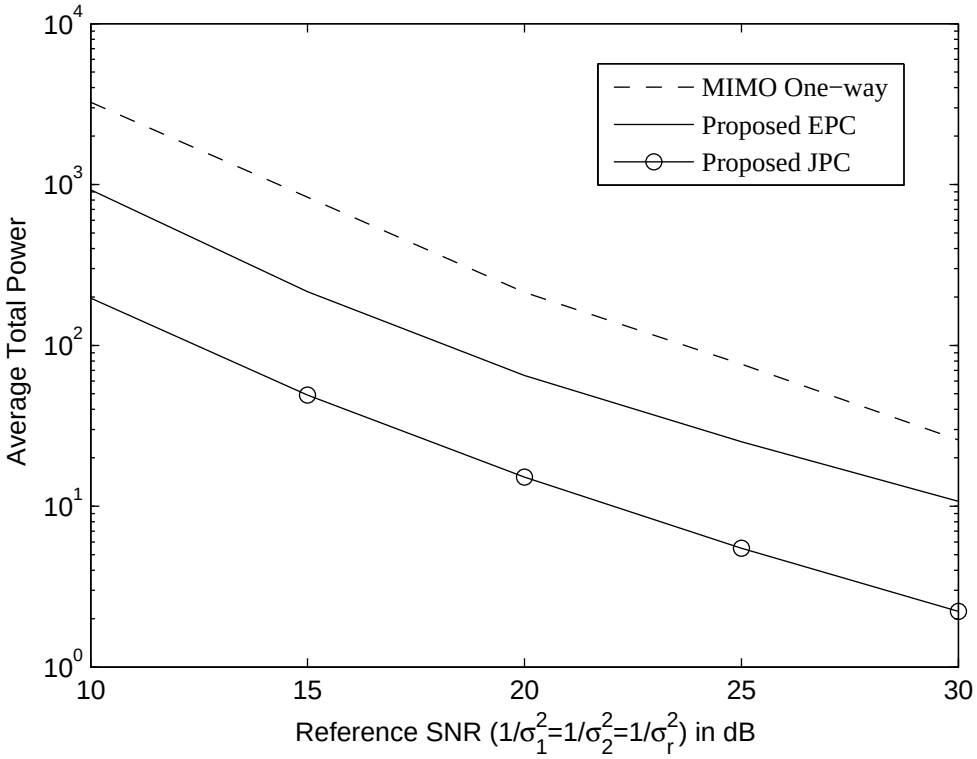


Figure 5.7: Average total power consumption versus SNR ($\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2}$) when R =2 bits/s/Hz and $M = 2$.

5.7 Chapter Conclusion

Joint beamforming and power management in the MIMO two-way relaying channel with a non-regenerative relay has been studied in this chapter. Based on the idea of subchannel alignment, transmit and receive beamformers are designed such that user channel pair can be decomposed into parallel subchannels to enable joint power allocation and joint power control. Joint power allocation dynamically allocates power to all substreams and nodes in order to maximise the sum-rate, subject to a total power constraint. On the other hand, joint power control minimises the total transmission power while satisfying the predefined target data rates. The convexity of the power allocation and power control formulations are determined. The non-concave power allocation utility function is approximated by a concave upper bound while the power control formulation is transformed from geometric program into equivalent convex form, in order to facilitate the use of efficient convex optimisation techniques in solving the optimisation problems. Numerical simulation results demonstrate that the proposed joint beamforming and power management scheme is able to deliver significant sum-rate improvement or achieve substantial transmission power saving, when compared to existing schemes.

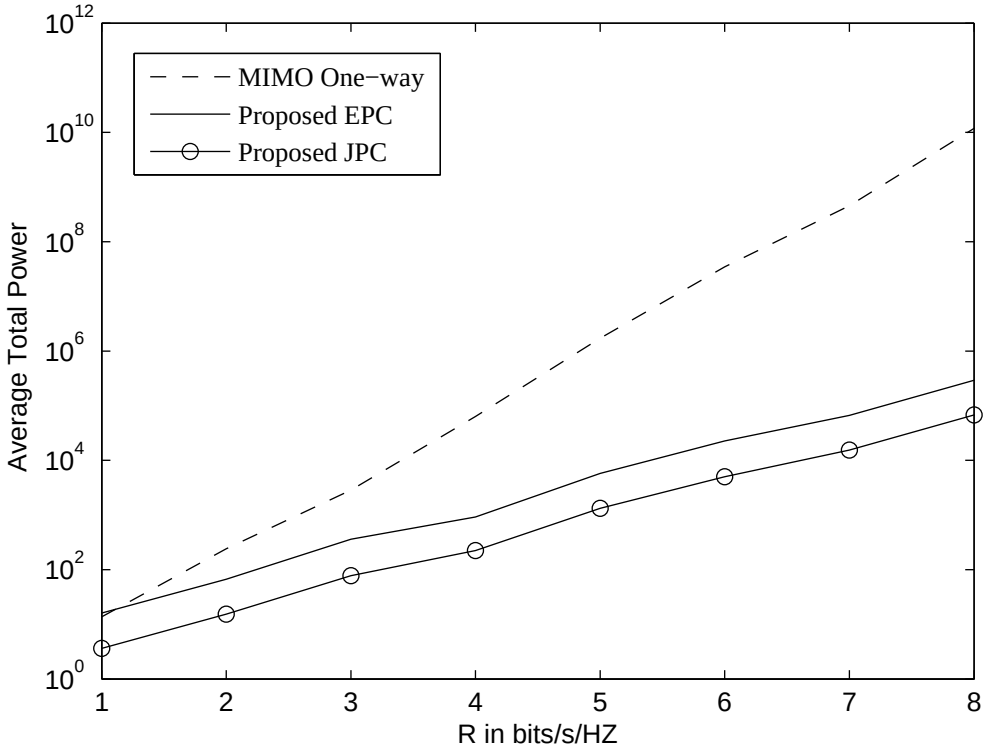


Figure 5.8: Average total power consumption versus target rate R when reference SNR, $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_2^2} = \frac{1}{\sigma_r^2} = 20$dB and $M = 2$.

Chapter 6

Two-Way Secrecy Schemes for the Broadcast Channel with Internal Eavesdroppers

In the previous chapters, it has been shown that the broadcast nature of wireless transmission enables the use of relays to deliver improvements such as better reliability, higher throughput and improved power efficiency in wireless broadcast and information exchange channels. However, the broadcast nature of wireless transmission also poses a security threat to wireless information transfer. Since transmission over wireless channels can be eavesdropped by malicious nodes, the security of information transfer could be compromised. This chapter considers the transmission of secret messages to users in a downlink cellular scenario, where the message to a given user must be kept secret from all of the other users. Multi-user diversity [18] suggests an opportunistic approach that sends a message secretly to the user with the current best channel; however, the secrecy rate goes to zero in the limit of a large number of users. Here, channel reciprocity is exploited via a two-way secrecy scheme to provide a constant positive secrecy rate to the user with the best channel. Next, motivated by the desire to transmit to a given user (rather than opportunistically to the user with the best channel), a second scheme is developed that employs relaying from other users from whom the message is still kept secret. The secrecy rates of the proposed schemes are analysed analytically in the limit of a large number of system users. In addition, for a finite number of users, simulation results demonstrate that the proposed two-way secrecy schemes are able to provide significant gains in the secrecy rates in both the opportunistic and non-opportunistic scenarios over a large range of user numbers.

6.1 Chapter Introduction

Information transmission in the wireless channel is prone to eavesdropping by unauthorised parties due to the broadcast nature of the wireless transmission. Traditionally, the secrecy of wireless transmissions is achieved using computation-theoretic security which relies on the inability of an eavesdropper to solve a "hard" problem. Such computation-theoretic approaches have the advantage that they assume the physical transmission medium to the eavesdropper is error free, and thus the eavesdropper is given a message identical to what is transmitted. However, these approaches rely on computational assumptions of the adversary that, if possible, it would be advantageous to avoid. Perfect secrecy where the eavesdropper gains no knowledge of the transmitted information cannot be guaranteed since in real life the physical channel is imperfect and the computational complexity of the eavesdropper might be underestimated.

In order to achieve security in the face of an adversary with infinite computational power, Shannon [35] introduced the concept of information-theoretic security. Shannon showed that perfect secrecy, where the mutual information between the secret message and what is observed at the eavesdropper is zero, can be achieved by means of a one-time pad [35], where the exclusive-or (X-OR) of a random key and the secret message is transmitted over the wireless channel. This is largely a negative result, as the condition to achieve perfect secrecy is that the entropy of the random key used to encrypt the secret message has at least the same entropy as the secret message. In other words, this scheme requires a new key for each new message and the key length has to be as long as the message length, which results in large overhead in key sharing.

However, Wyner [36] recognised that information-theoretic security could still be of interest if the channel from the transmitter to the eavesdropper is noisier than the channel from the transmitter to the legitimate receiver. In particular, Wyner showed that a non-zero secrecy capacity can be achieved using the wiretap coding scheme without the need of prior key-sharing. This motivated significant further work: [37] extends Wyner's scheme to non-degraded channels, while [38] generalises it to Gaussian channels. More recently, Wyner's wiretap coding scheme has served as an important framework for the development of secret communication in wireless fading channels, where variations of the wireless channel have been exploited to obtain the conditions for which a positive secrecy capacity can be employed; for example , [39] and [40] investigate secrecy communication in fading channels

and reveal that by opportunistically exploiting the fluctuation of the channels, a positive secrecy capacity can be achieved using Wyner's wiretap scheme even though on average the channel from the transmitter to the eavesdropper is less noisy than the channel from the transmitter to the legitimate user.

The study of secrecy communication extends to multi-user networks with more than the classical three-node-setup (source, legitimate user and eavesdropper). The wireless broadcast channel is of great interest, since it constitutes an important building block of modern communication networks. In [68], a scalar broadcast channel with K users and 1 external eavesdropper, which is not from the user set, is considered. Leveraging the benefit of multi-user diversity, the opportunistic secrecy transmission scheme proposed by [68] is able to achieve a positive secrecy capacity which improves with increasing number of users K. Of more interest is the work in [19], which looks into the secrecy of opportunistic transmission in a scalar broadcast channel with K users, where the transmitted message is meant to be kept secret from all but the intended recipient. Under opportunistic broadcast, only the active user with the best current channel gain is allowed to decode the transmission from the source at a particular time, and the $K-1$ idle users are potential eavesdroppers. In this scenario, not only the legitimate user but also the internal eavesdroppers are able to take advantage of the multi-user diversity, and it is shown in [19] that, as K goes to infinity, the secrecy outage probability goes to one while the average secrecy capacity reduces to zero. This is a straightforward application of order statistics: the channel gain of the best eavesdropper (user with the second largest channel gain from the transmitter) approaches very closely the channel gain of the legitimate user (user with the largest channel gain from the transmitter) when K approaches infinity. The problem worsens for the case of non-opportunistic broadcast transmission, where the legitimate user does not necessarily have the best channel among all users. In this case, only eavesdroppers benefit from the multi-user diversity.

This chapter considers how to provide security and/or privacy for wireless information transmission in a cellular downlink architecture with a large number of users, which corresponds to the scalar broadcast channel. As in [19], the goal here is to have the message intended for a single user kept secret from the remainder of the users. For the first scenario considered, as in [19], this chapter assumes that it is sufficient to transfer the secret message of any user in a given time slot, and hence proceed opportunistically by transmitting to the user with the most favorable conditions. For this opportunistic case, the proposed two-way secrecy scheme exploits the reciprocity of the forward and

backward channels between the source and the legitimate user to communicate a secret key and a one-time pad encoded secret message. In contrast to the scheme of [19], the proposed scheme is shown to achieve a positive secrecy rate in the network, even when the number of users is large. The proposed scheme is then generalised to non-opportunistic broadcast, where, in each time slot, a secret message must be transmitted to a specific user. In this case, the proposed scheme invites a pair of idle users to act as relays and utilises forward and backward channels of each of the source-to-relay and relay-to-destination links to forward secret keys and one-time pad encoded secret messages. In both cases, the proposed schemes are able to achieve positive secrecy rates when all of the nodes except the desired recipient (including the relaying nodes) are treated as eavesdroppers. Analytical and simulation results verify that the proposed schemes are able to achieve non-zero secrecy rates and non-zero ergodic secrecy capacity under opportunistic and non-opportunistic broadcast, for a large range of user numbers, and in the limit as the number of users goes to infinity.

The rest of the chapter is organised as follows. Section 6.2 describes the system model and problem statement. The proposed two-way secrecy scheme for opportunistic broadcast and its asymptotic analysis are presented in Section 6.3. Section 6.4 generalises the two-way secrecy scheme to non-opportunistic broadcast. Numerical simulation results for the ergodic secrecy capacity of finite networks employing the proposed schemes are presented in Section 6.5, and Section 6.6 concludes the chapter.

6.2 System Model and Problem Statement

Consider a scalar broadcast channel with a source and $K + 1$ users who are potential destinations. All nodes are equipped with a single antenna. It will be assumed that all channels undergo i.i.d. and slowly-varying Rayleigh fading. Given the slow fading assumption, the block fading model is adopted: the fading is fixed for a block of symbols and changes independently to another realisation for the next block [18]. Throughout, the standard half-duplex constraint is enforced: nodes can either transmit or receive at a given time but they cannot do both. Finally, it is assumed that the uplink and downlink channels from the source to the users employ the same frequency bands (or frequency bands close enough together) so that channel reciprocity holds: for a given block, the fading gain when the source transmits and the user receives is identical to that when the user transmits and the

source receives. Since it is sufficient to deal with a single block throughout the chapter, the block index is suppressed to simplify the notation; hence, the received signal at receiver i will be denoted by:

$$y_i = h_{S,i}x + n_i, \tag{6.1}$$

where $h_{S,i}$ is the (complex) channel gain from the source S to user i, x is the transmitted signal from the source, and n_i is the circularly symmetric (complex) Gaussian noise. In practice, of course, the transmission during each block would consist of a large number of symbols and a corresponding set of received samples and noise samples, making (6.1) a vector equation, but the slight abuse of notation employed in (6.1) maintains the precision of the results while hiding physical layer details that obfuscate the presentation.

Under channel fading, the opportunistic broadcast transmission where the source transmits to the user with the strongest channel at a particular time is able to produce the highest total non-secure throughput, i.e. optimal sum-rate [18]. In opportunistic broadcast transmission, the source S transmits to the user M with best instantaneous channel gain, i.e. $|h_{S,M}|^2 = \max_i |h_{S,i}|^2$,$\forall i \in \{1, \ldots, K+1\}$. The remaining K idle users $\{1, \ldots, M-1, M+1, \ldots, K+1\}$ are considered the potential eavesdroppers (internal eavesdroppers). If the eavesdropper with the best connection to the source is not able to decode the secret message successfully, all other eavesdroppers will also not succeed. Hence, the system only needs to ensure that the eavesdropper with the best connection to the source is not able to decode the secret message intended for the legitimate user. Denote the best eavesdropper channel gain as $|w_E|^2 = \max_j |h_{S,j}|^2$, $\forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$. In order to secure the transmission between the source and the legitimate user from the eavesdroppers, Wyner's wiretap coding scheme [36] can be used. The source transmits the secret message to the legitimate user using the following secrecy rate,

$$r_{one-way} = \begin{cases} \log_2\left(\frac{1+\gamma_0|h_{S,M}|^2}{1+\gamma_0|w_E|^2}\right), & \text{when } |h_{S,M}|^2 > |w_E|^2, \\ 0, & \text{when } |h_{S,M}|^2 \leq |w_E|^2, \end{cases} \tag{6.2}$$

where γ_0 is the mean signal-to-noise ratio (SNR). The derivation of $r_{one-way}$ is obtained using similar steps in [40] and serves as a baseline scheme for comparison. A positive secrecy rate can be achieved when $|h_{S,M}|^2 > |w_E|^2$. However, as the number of users gets large, i.e. $K \to \infty$, $\frac{1+\gamma_0|h_{S,M}|^2}{1+\gamma_0|w_E|^2}$ is

arbitrarily close to one with high probability. In other words, the best eavesdropper channel is nearly as good as the legitimate user channel. As a result, the ergodic secrecy capacity and the outage secrecy capacity go to zero, as shown in [19]. This result suggests that standard multi-user diversity is not effective for secrecy transmission in the broadcast channel.

The problem extends to the broadcast channel with non-opportunistic transmission. Under non-opportunistic broadcast, the legitimate user does not necessarily have the best channel connection to the source, since no user scheduling is performed. Any of the K idle users with better channels are able to eavesdrop the active transmission and thus the eavesdroppers benefit from multi-user diversity but the legitimate user does not. Using Wyner's wiretap secrecy scheme, the secrecy rate is readily shown to be zero when K is large.

In this chapter, two-way secrecy schemes are proposed to provide a positive secrecy capacity for the scalar broadcast channel with a large number of users. Two cases are considered. First, the scalar broadcast channel with opportunistic transmission is discussed in Section 6.3 and second, the scalar broadcast channel with non-opportunistic transmission is presented in Section 6.4.

6.3 Two-Way Secrecy Scheme for Opportunistic Broadcast

In this section, a scalar broadcast channel with opportunistic transmission is considered. A two-way secrecy scheme which utilises both the forward and backward channels between the source and the legitimate user is proposed. The protocol description is presented in the first subsection while the secrecy rate is derived in the second subsection. The asymptotic analysis of the secrecy rate is described in the third subsection.

6.3.1 Protocol Description

The proposed two-way secrecy scheme can be described in two time slots as shown in Figure 6.1. As per Section 6.2, the source S wishes to convey a secret message d to the destination M with the best channel using the proposed scheme. In the first time slot, the legitimate user M transmits x, a secret key encoded with wiretap codes [36], to the source S at a secrecy rate r_s. The signal received by the

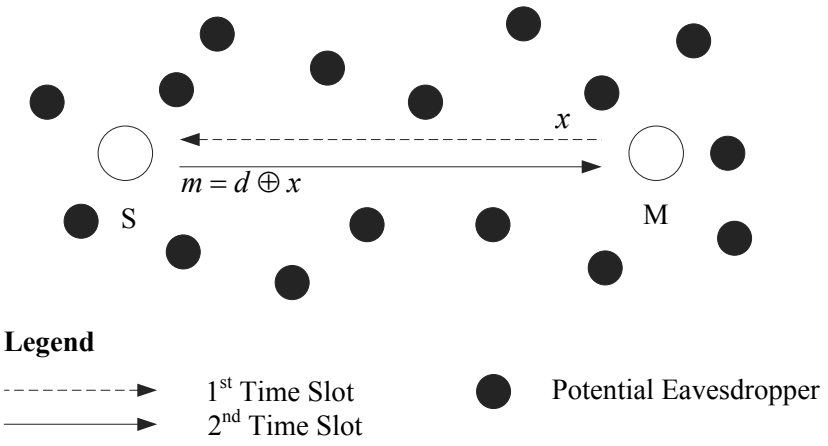


Figure 6.1: Data flows of the proposed two-way secrecy scheme for opportunistic broadcast.

source node during the first time slot can be expressed as

$$y_S(1) = h_{M,S}x + n_S(1), \tag{6.3}$$

where $n_S(t)$ is the noise observed by the source node in time slot t. Note that the same notation is used to represent the secret message or key and the transmitted (wiretap encoded) signal, i.e. symbol x is used to represent both the secret key and the transmitted signal. This slight abuse of notation is only meant for convenience and does not affect the precision of the signal model.

At the same time, the signal received by idle user j (an internal eavesdropper) is

$$y_j(1) = g_{M,j}x + n_j(1), \tag{6.4}$$

$\forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$. The secrecy rate r_s will be determined in the following subsection.

The source node decodes the secret key and uses it as a one-time pad inside a wiretap submission for the secret message intended for the legitimate user. Denote the one-time pad encoded secret message as $m = d \oplus x$.

In the second time slot, the source broadcasts the secret message m using wiretap coding at a rate

r_s. The signal received by the legitimate user can be written as

$$y_M(2) = h_{S,M}m + n_M(2). \tag{6.5}$$

Note that, under channel reciprocity, the forward channel between the source and the legitimate user is exactly the same as the backward channel shown in (6.3), i.e. $h_{M,S} = h_{S,M}$. The signal observed by any potential eavesdropper j is

$$y_j(2) = h_{S,j}m + n_j(2), \tag{6.6}$$

$\forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$. Since the legitimate user now has the knowledge of the secret key x and the one-time pad encoded message m, the secret message can be decoded successfully, i.e. $d = m \oplus x$. The eavesdroppers are not able to decode the secret message unless they have knowledge of both x and m.

6.3.2 Secrecy Rate Derivation

In this subsection, the secrecy rate r_s is determined. As per above, the proposed two-way secrecy protocol only needs to ensure that the eavesdroppers will not be able to decode either x or m. Hence, the worst channel gain among the backward and forward links of any eavesdropper j is of interest. Define the worst channel gain of an eavesdropper j as follows

$$|w_j|^2 = \min\left(|h_{S,j}|^2, |g_{M,j}|^2\right), \tag{6.7}$$

$\forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$. If the eavesdropper with the largest worst-eavesdropper-channel-gain is not able to decode the secret message, no other eavesdropper will be able to decode the secret message either. Therefore, the proposed protocol only needs to consider the best eavesdropper (the eavesdropper with the largest worst-eavesdropper-channel-gain) in the design of the secrecy scheme. Define the largest worst-eavesdropper-channel-gain as

$$|w_E|^2 = \max_j |w_j|^2, \tag{6.8}$$

$\forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$. Using the conventional wiretap coding and assuming unit transmission power, the secrecy rate r_s of the proposed two-way secrecy scheme can be expressed as

$$r_s = \begin{cases} \log_2 \left(\frac{1+\gamma_0 \left|h_{S,M}\right|^2}{1+\gamma_0 |w_E|^2} \right), & \text{when } |h_{S,M}|^2 > |w_E|^2, \\ 0, & \text{when } |h_{S,M}|^2 \leq |w_E|^2. \end{cases} \tag{6.9}$$

With equal time allocation between the two transmission time slots of the proposed protocol, the overall secrecy rate is $r_{two-way} = \frac{1}{2} r_s$.

6.3.3 Asymptotic Analysis When $K \to \infty$

In this subsection, the asymptotic secrecy rate r_s in the event of $K \to \infty$ is analysed. The following result is a critical tool.

Lemma 6.1. *Let u_i, $i = 1, 2, \ldots, K$ be a sequence of exponentially distributed variables, each with cumulative distribution function (CDF), $F_{u_i}(u) = 1 - \exp(-\lambda u)$, the extremal, $\max_i u_i$ converges almost surely as,*

$$\lim_{K \to \infty} \frac{\max_i u_i}{\ln K} = \frac{1}{\lambda} \quad \text{a.s.} \tag{6.10}$$

Proof. See [69]. ■

The asymptotic behaviour of the secrecy rate r_s when $K \to \infty$ is summarised in the following theorem.

Theorem 6.1. *Assuming wiretap coding and i.i.d. Rayleigh fading, the rate r_s employed by the proposed two-way secrecy scheme under opportunistic broadcast converges to*

$$r_s = 1 \text{ bits/s/Hz}, \tag{6.11}$$

as $K \to \infty$ and the overall secrecy rate of the proposed scheme is $r_{two-way} = \frac{1}{2} r_s$.

Proof. First, consider $|w_j|^2, \forall j \in \{1, \ldots, M-1, M+1, \ldots, K+1\}$ as in (6.7). The minimum of two independently distributed exponential random variables with parameters λ_1 and λ_2, respectively, is easily shown to be exponential with parameter $\lambda_1 + \lambda_2$ (see [54], for example). Therefore, $|w_j|^2$ is exponentially distributed with parameter $\lambda = 2$. By Lemma 6.1, this implies that the ratio of $|w_E|^2$ in (6.8) and $\frac{\ln K}{2}$ goes to 1 almost surely as $K \to \infty$. Likewise, the ratio of $|h_{S,M}|^2$ and $\ln K$ goes to 1 almost surely. Hence, from (6.9), noting $|h_{S,M}|^2 > |w_E|^2$ almost surely, one obtains the following

$$\lim_{K\to\infty} r_s = \lim_{K\to\infty} \log_2 \left(\frac{1+\gamma_0 |h_{S,M}|^2}{1+\gamma_0 |w_E|^2} \right), \tag{6.12}$$

$$= \lim_{K\to\infty} \log_2 \left(\frac{1+\gamma_0 \ln K}{1+\gamma_0 \frac{\ln K}{2}} \right). \tag{6.13}$$

Since both the numerator and denominator go to infinity, L'Hôpital's rule can be invoked. By differentiating the numerator and denominator independently, one has the following convergence

$$\lim_{K\to\infty} r_s = \lim_{K\to\infty} \log_2 \left(\frac{\gamma_0 \left(\frac{1}{K}\right)}{\gamma_0 \left(\frac{1}{2}\right) \left(\frac{1}{K}\right)} \right), \tag{6.14}$$

$$= \lim_{K\to\infty} \log_2(2), \tag{6.15}$$

$$= 1, \tag{6.16}$$

almost surely and the theorem is proved. ■

Remark 6.1. Following similar steps as in the proof of Theorem 6.1, the asymptotic secrecy rate of the baseline scheme described in Section 6.2 when $K \to \infty$ is readily shown as $r_{one-way} = 0$ bits/s/Hz [19]. In comparison, the proposed two-way secrecy scheme is able to achieve an overall secrecy rate, $r_{two-way} = 0.5$ bits/s/Hz.

6.4 Two-Way Secrecy Scheme for Non-Opportunistic Broadcast

In this section, the proposed two-way secrecy scheme is generalised to a scalar broadcast channel with non-opportunistic transmission. In other words, a secret message must be transmitted to a specific user M. The proposed protocol makes use of the two-way secrecy scheme developed in the previous section and employs two idle users to relay the secret messages from the source to the destination. The proposed protocol ensures that the secret messages are hidden from all idle users (internal eavesdroppers), including the relays which help to forward the source messages. The proposed protocol for non-opportunistic broadcast is described in the first subsection while the secrecy rate is derived in the second subsection. The secrecy rate in large networks is analysed in the third subsection.

6.4.1 Protocol Description

The protocol starts by selecting two idle users with good channel connections to the source and the destination as relays. The selection criterion for the best relay is

$$R_1 = \arg\max_{j \in K_1} \min \left(\ |h_{S,j}|^2 , \ \ |h_{M,j}|^2 \ \right), \tag{6.17}$$

and the selection criterion for the second best relay is

$$R_2 = \arg\max_{k \in K_2} \min \left(\ |h_{S,k}|^2 , \ \ |h_{M,k}|^2 \ \right), \tag{6.18}$$

where $K_1 = \{1, \ldots, K\}$ and $K_2 = \{1, \ldots, R_1 - 1, R_1 + 1, \ldots, K + 1\}$.

Suppose the source S wishes to convey a secret message d to destination M using the proposed protocol. The source first generates a random binary string b and then sets a string c such that $d = b \oplus c$. The proposed protocol can be described in two phases.

In the first phase, three time slots are used. See Figure 6.2. In the first time slot, the first relay R_1 transmits a wiretap encoded random binary string (secret key) $\tilde{b}$ at a secrecy rate r_s. The source then transmits the one-time pad encoded message, $m_1 = b \oplus \tilde{b}$ using wiretap coding at secrecy rate r_s. In the third time slot, the relay R_1 transmits m_1 using wiretap coding at a secrecy rate r_s.

A similar process is repeated in the second phase, for the source to communicate message c to the

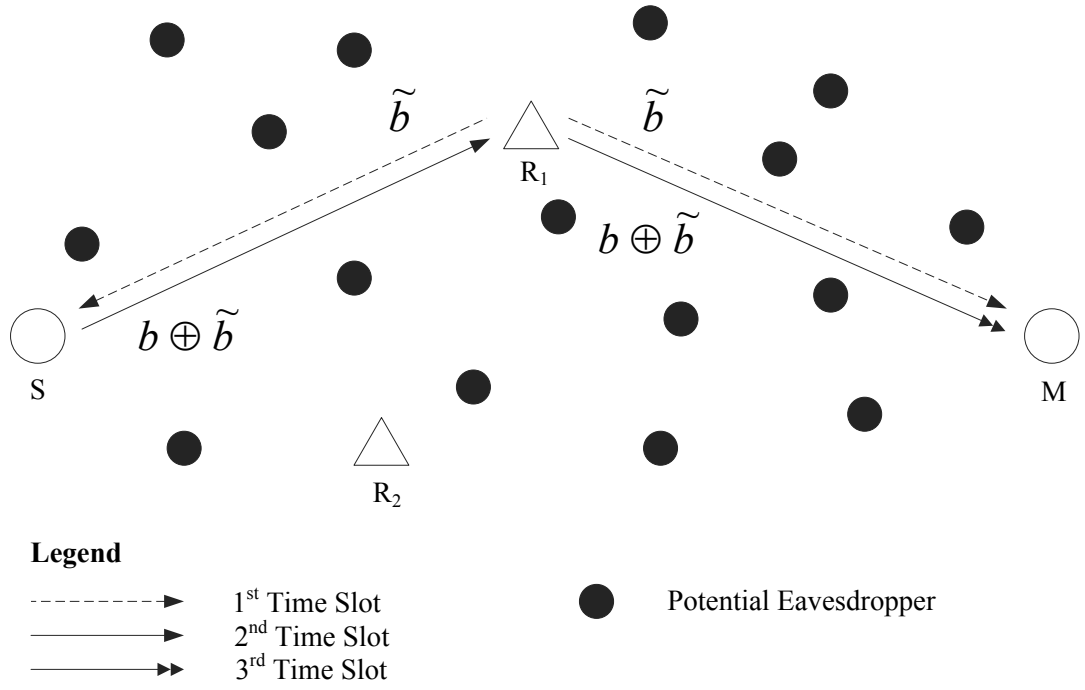


Figure 6.2: Data flows in the first phase of the proposed two-way secrecy scheme for non-opportunistic broadcast.

destination with the help of the relay R_2. See Figure 6.3. Specifically, in the first time slot, the relay R_2 transmits a wiretap encoded random binary string (secret key) $\tilde{c}$ at a secrecy rate r_s, while in the second time slot, the source transmits the one-time pad encoded message, $m_2 = c \oplus \tilde{c}$ to the relay R_2 at a secrecy rate r_s. In the third time slot, the relay transmits the message m_2 at a secrecy rate r_s. The destination is able to recover the secret message d using the knowledge of m_1, m_2, $\tilde{b}$ and $\tilde{c}$, since $d = m_1 \oplus \tilde{b} \oplus m_2 \oplus \tilde{c}$. The proposed protocol only needs to ensure that eavesdroppers and relays will miss at least one of the required messages, m_1, m_2, $\tilde{b}$ and $\tilde{c}$. The secrecy rate r_s is determined in the following subsection.

6.4.2 Secrecy Rate Derivation

In this subsection, the secrecy rate r_s for non-opportunistic broadcast is determined. As per above, each eavesdropper needs the full knowledge of m_1, m_2, $\tilde{b}$ and $\tilde{c}$ before the secret message d can be decoded. One needs to know the worst eavesdropper-channel-gain of every potential eavesdropper,

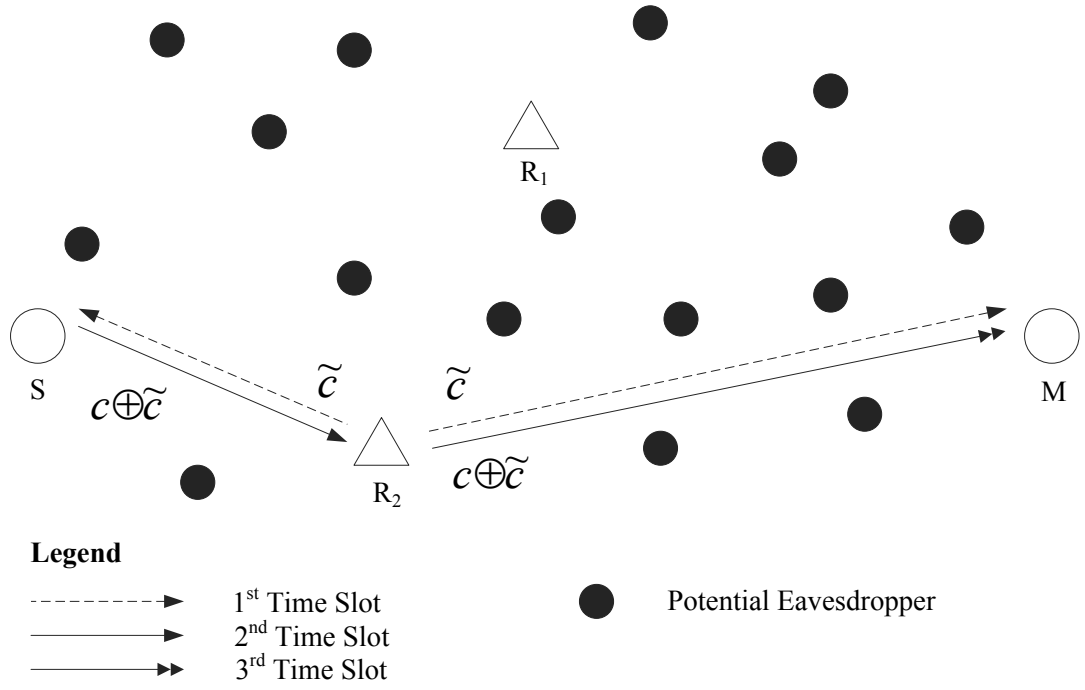


Figure 6.3: Data flows in the second phase of the proposed two-way secrecy scheme for non-opportunistic broadcast.

which is defined as follows

$$|w_l|^2 = \min\left(|h_{S,l}|^2, |h_{R_1,l}|^2, |h_{R_2,l}|^2\right), \tag{6.19}$$

$\forall l \in \{1, \ldots, R_1 - 1, R_1 + 1, \ldots, R_2 - 1, R_2 + 1, \ldots, K + 1\}$ where $h_{R_i,l}$ is the channel from relay R_i to idle user l (potential eavesdropper). It is known that if the best eavesdropper is not able to decode the secret message successfully, other eavesdroppers will not succeed either. Therefore, the proposed protocol only needs to consider the largest worst-eavesdropper-channel-gain, which can be defined as

$$|w_E|^2 = \max_l |w_l|^2, \tag{6.20}$$

$\forall l \in \{1, \ldots, R_1 - 1, R_1 + 1, \ldots, R_2 - 1, R_2 + 1, \ldots, K + 1\}$. In addition, the proposed protocol needs to prevent both the relays, R_1 and R_2 from decoding the secret message d. This can be achieved by keeping R_1 ignorant of the secret key $\tilde{c}$ and R_2 ignorant of the secret key $\tilde{b}$. For this reason, the proposed protocol accounts for the inter-relay channel between R_1 and R_2, h_{R_1,R_2} in determining the

secrecy rate. Define the overall maximum-eavesdropper-channel-gain as

$$|w_O|^2 \;=\; \max\left(|w_E|^2, |h_{R_1,R_2}|^2\right). \tag{6.21}$$

In the absence of eavesdroppers, the overall maximum-legitimate-channel-gain available for information transmission from source to destination is

$$|h_O|^2 = \max_k \min\left(\; |h_{S,k}|^2, \;\; |h_{M,k}|^2 \; \right), \tag{6.22}$$

$\forall k \in \{1, \ldots, R_1 - 1, R_1 + 1, \ldots, K + 1\}$. Using wiretap coding and assuming unit transmission power, the secrecy rate r_s of the proposed two-way secrecy scheme can be expressed as

$$r_s = \begin{cases} \log_2\left(\frac{1+\gamma_0|h_O|^2}{1+\gamma_0|w_O|^2}\right), & \text{when } |h_O|^2 > |w_O|^2, \\ 0, & \text{when } |h_O|^2 \leq |w_O|^2. \end{cases} \tag{6.23}$$

Accounting for the number of channel uses, the overall secrecy rate is $r_{two-way} = \frac{1}{6} r_s$.

6.4.3 Asymptotic Analysis When $K \to \infty$

In this subsection, the asymptotic behaviour of the secrecy rate r_s when $K \to \infty$ is analysed. The achievable secrecy rate of the proposed two-way secrecy scheme under non-opportunistic broadcast is summarised in the following theorem.

Theorem 6.2. *Assuming wiretap coding and i.i.d. Rayleigh fading, the rate r_s employed by the proposed two-way secrecy scheme under non-opportunistic broadcast converges to*

$$r_s = \log_2\left(\frac{3}{2}\right) \text{ bits/s/Hz}, \tag{6.24}$$

as $K \to \infty$ and the overall secrecy rate of the proposed scheme is $r_{two-way} = \frac{1}{6} r_s$.

Proof. First, consider $|w_l|^2$, $\forall l \in \{1, \ldots, R_1 - 1, R_1 + 1, \ldots, R_2 - 1, R_2 + 1, \ldots, K + 1\}$ as in (6.19).

From [54], the minimum of three independently distributed exponential variables with λ_1, λ_2 and λ_3 respectively, has exponential distribution with $\lambda = \lambda_1 + \lambda_2 + \lambda_3$. Therefore, $|w_l|^2$ is exponentially distributed with parameter $\lambda = 3$. Following Lemma 6.1, the ratio of $|w_E|^2$ in (6.20) and $\frac{\ln K}{3}$ goes to 1 almost surely as $K \to \infty$. Consider $|w_O|^2$ in (6.21). Since $|h_{R_1,R_2}|^2$ is just an average channel gain, one obtains $|w_O|^2 = |w_E|^2$. Recall the overall maximum-legitimate-channel-gain, $|h_O|^2$ in (6.22). One knows that $\min\left(|h_{S,k}|^2, \ |h_{M,k}|^2 \right)$ is exponential distributed with parameter $\lambda = 2$. By Lemma 6.1, the ratio of $|h_O|^2$ and $\frac{\ln K}{2}$ goes to 1 almost surely. Hence, from (6.23), noting that $|h_O|^2 > |w_O|^2$ almost surely, one obtains the following

$$\lim_{K\to\infty} r_s = \lim_{K\to\infty} \log_2\left(\frac{1+\gamma_0|h_O|^2}{1+\gamma_0|w_O|^2}\right), \tag{6.25}$$

$$= \lim_{K\to\infty} \log_2\left(\frac{1+\gamma_0\frac{\ln K}{2}}{1+\gamma_0\frac{\ln K}{3}}\right). \tag{6.26}$$

Invoking L'Hôpital's rule, one has the following convergence

$$\lim_{K\to\infty} r_s = \lim_{K\to\infty} \log_2\left(\frac{\gamma_0(\frac{1}{2})\left(\frac{1}{K}\right)}{\gamma_0(\frac{1}{3})\left(\frac{1}{K}\right)}\right), \tag{6.27}$$

$$= \lim_{K\to\infty} \log_2\left(\frac{3}{2}\right), \tag{6.28}$$

almost surely and the theorem is proved. ■

Remark 6.2. It can be easily shown that the overall secrecy rate of the baseline one-way secrecy scheme described in Section 6.2 goes to zero when $K \to \infty$. The proposed two-way secrecy scheme is able to achieve a positive overall secrecy rate, $r_{two-way} = \frac{1}{6}\log_2\left(\frac{3}{2}\right)$ bits/s/Hz. By introducing relays and two-way secrecy, the proposed protocol is able to average out the channel gain of the eavesdroppers. In other words, the proposed protocol is able to make sure that only the legitimate user is able to benefit from the multi-user diversity while the eavesdroppers fail to gain an advantage from the multi-user diversity.

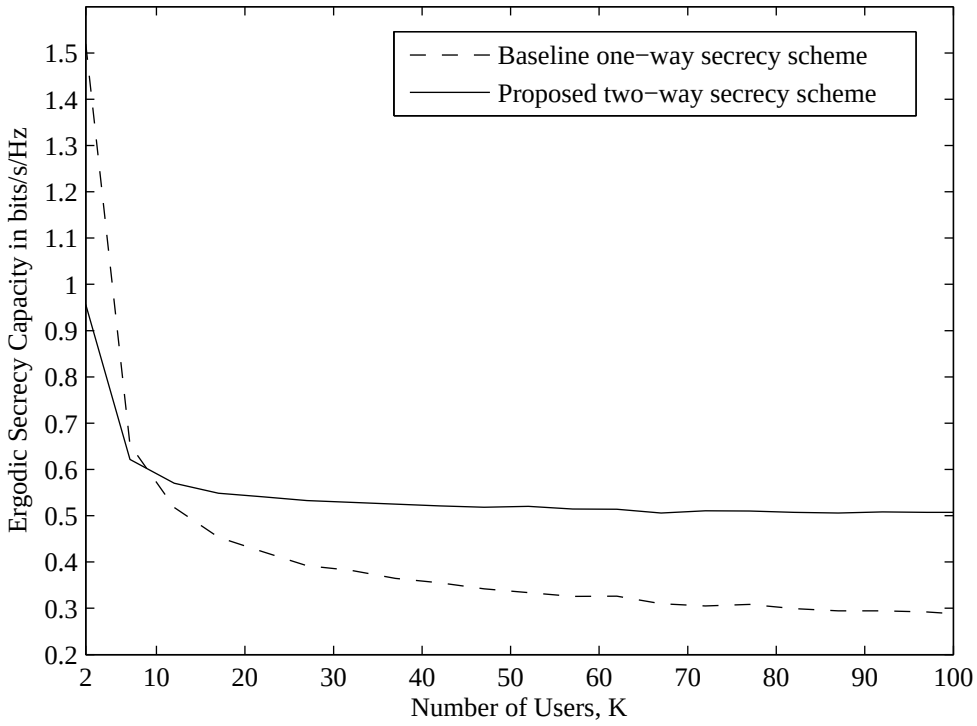


Figure 6.4: Ergodic secrecy capacity versus number of users K, when SNR=10dB.

6.5 Numerical Results

In this section, the ergodic secrecy capacity of the proposed two-way secrecy scheme and baseline one-way secrecy scheme are simulated using a Monte Carlo method to consider the gain obtained for a finite number of users. The results are organised into two subsections. The first subsection discusses the simulation results for opportunistic broadcast while the second subsection describes the simulation results for non-opportunistic broadcast.

6.5.1 Opportunistic Broadcast

In this subsection, the performance of the proposed two-way secrecy scheme under opportunistic broadcast is examined. The ergodic secrecy capacity of the proposed scheme and the baseline one-way secrecy scheme are simulated. The ergodic secrecy capacity can be obtained by averaging the secrecy rate over all channel states

Figure 6.4 shows the ergodic secrecy capacity versus the number of users K of the proposed two-way secrecy scheme and the baseline one-way secrecy scheme when SNR=10dB. When K is

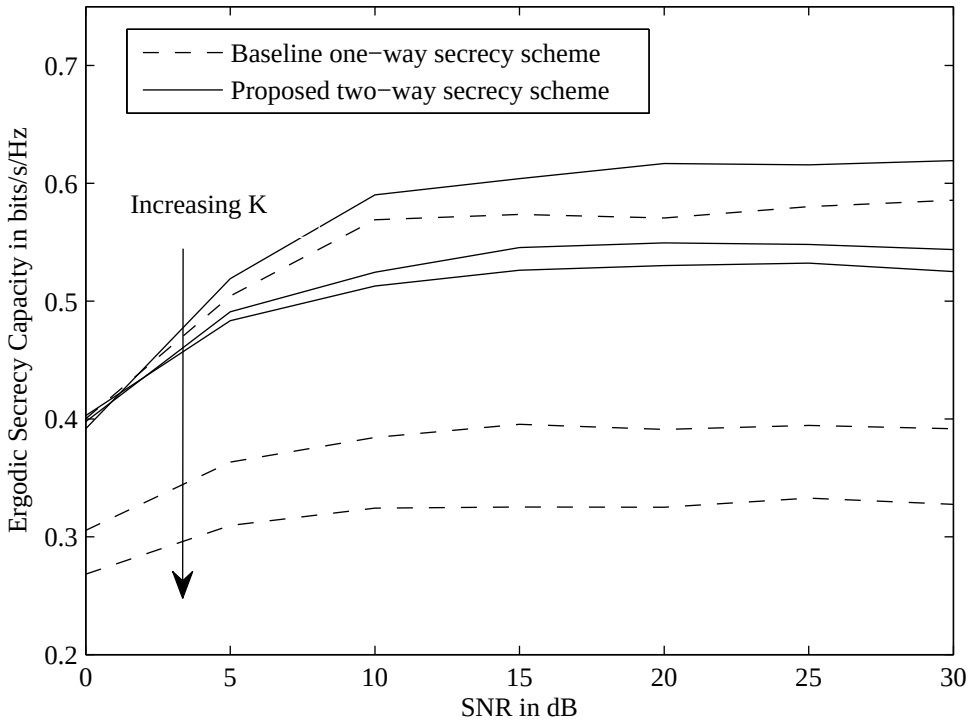


Figure 6.5: Ergodic secrecy capacity versus SNR when K=10, 30, and 60. K increases in the direction of the arrow.

increased from 2 to 10, the ergodic secrecy capacity of the baseline scheme decreases drastically. This is due to the fact that the best eavesdropper also benefits from the multi-user diversity as K increases. When $2 \leq K \leq 10$, the proposed two-way secrecy scheme does not perform better than the baseline scheme. This is because the proposed two-way secrecy scheme uses extra channel resources to communicate a secret key from the legitimate user to the source. However, when $K > 10$, the proposed scheme outperforms the baseline scheme. The proposed scheme is less sensitive to the increase of K, as is evident from the curve which decreases gradually with K and remains constant for large K. This agrees with the result in Theorem 6.1 that the secrecy rate of the proposed scheme converges to a positive constant when $K \to \infty$. In contrast, the ergodic secrecy capacity of the baseline scheme declines with K. When $K \to \infty$, the ergodic secrecy capacity of the baseline scheme will eventually go to zero, as shown in [19]. In general, the proposed two-way secrecy scheme is able to cope with the multi-user diversity through the use of the forward and the backward channels and deliver a positive secrecy capacity even when K is large.

Figure 6.5 shows the ergodic secrecy capacity versus SNR when $K = 10$, 30, and 60. When $0 \leq$SNR≤ 10dB, the ergodic secrecy capacity of both schemes increases with the SNR. When SNR>10 dB, the performance of both schemes quickly saturates because at high SNR, (6.2) and

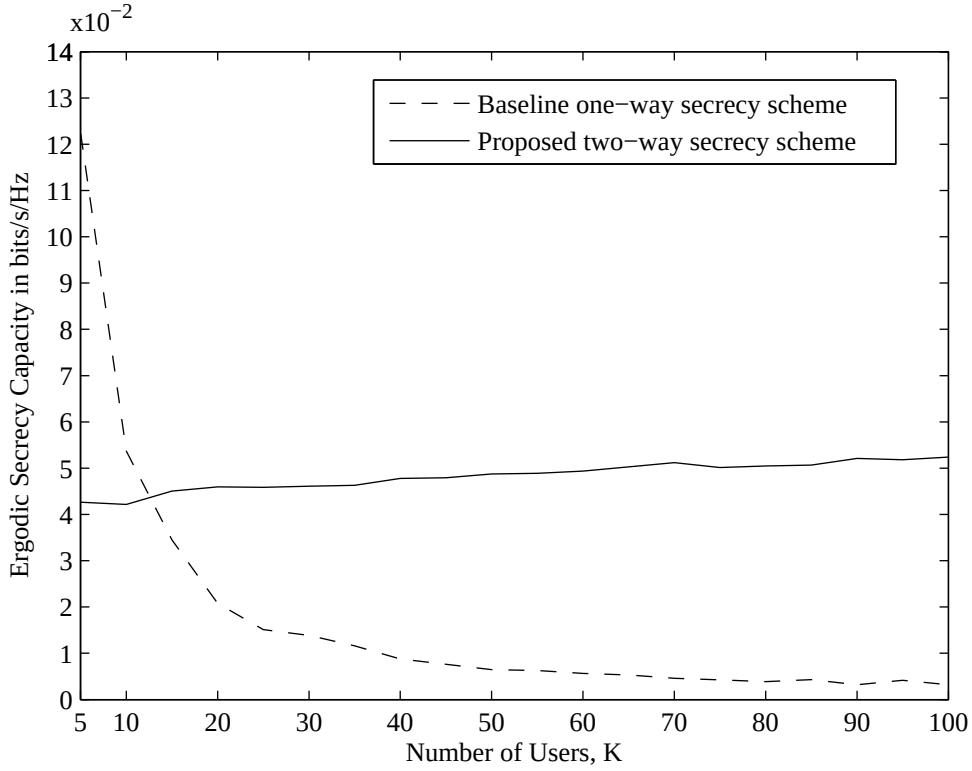


Figure 6.6: Ergodic secrecy capacity versus number of users K, when SNR=30dB.

(6.9) reveal that further increases in SNR benefit the legitimate user and eavesdroppers equally. This suggests that a medium operating SNR is sufficient for both schemes. Generally, the proposed scheme outperforms the baseline scheme. When K increases, the performance of the baseline scheme deteriorates significantly. For instance, when K is increased from 30 to 60, an 18% drop in the ergodic secrecy capacity is recorded. In comparison, the proposed two-way secrecy scheme only experiences a 4% drop in the ergodic secrecy capacity. Hence, the proposed scheme is less sensitive to the increase of K, as compared to the baseline scheme.

6.5.2 Non-Opportunistic Broadcast

In this subsection, the performance of the proposed two-way secrecy scheme for the non-opportunistic broadcast is presented. The ergodic secrecy capacity of the proposed scheme and the baseline one-way secrecy scheme are simulated.

Figure 6.6 shows the ergodic secrecy capacity versus number of users K of the proposed two-way secrecy scheme in comparison with the baseline one-way secrecy scheme when SNR = 30dB.

It can be observed that the ergodic secrecy capacity of the baseline scheme decreases with K, and approaches zero when K is large. This is due to the fact that only eavesdroppers benefit from

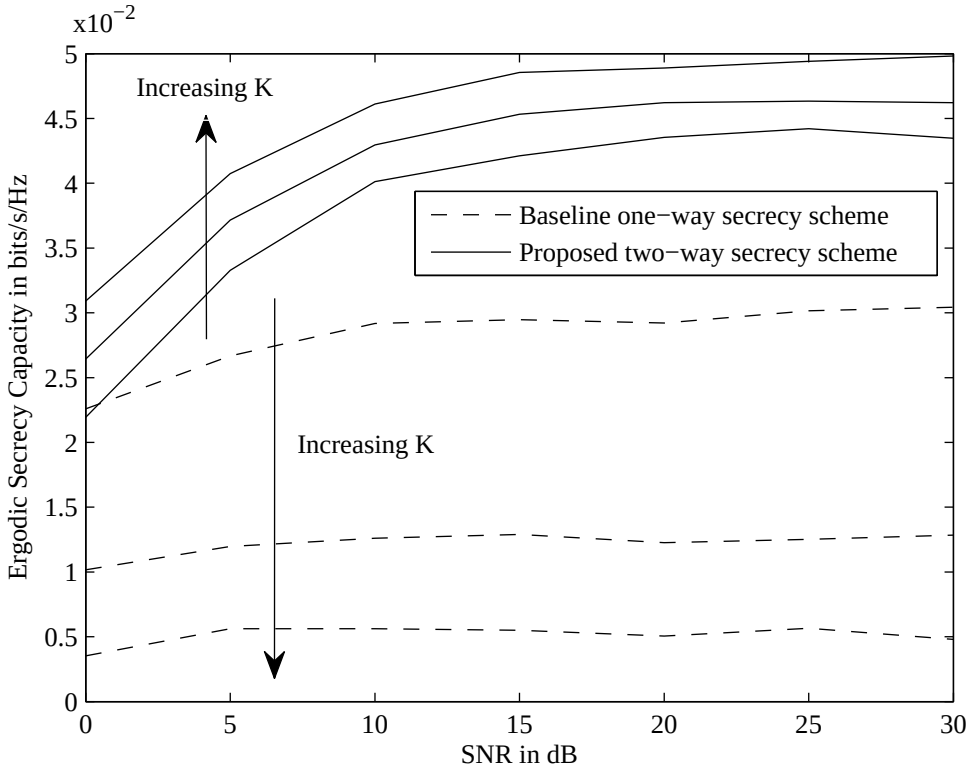


Figure 6.7: Ergodic secrecy capacity versus SNR when K=15, 30, and 60. K increases in the direction of the arrow.

multi-user diversity following the increase of K whereas the legitimate user does not benefit from the increase of K. In contrast, the ergodic secrecy capacity of the proposed scheme increases gradually with K and converges to a positive constant when K is large. This agrees with the result in Theorem 6.2. By using the proposed scheme, the legitimate user is able to benefit from the multi-user diversity through the use of relays and the forward and backward channels, while preventing the eavesdroppers from gaining a similar benefit.

Figure 6.7 shows the ergodic secrecy capacity versus SNR when $K = 15$, 30, and 60. When the SNR is increased from 0 to 10dB, both the proposed scheme and the baseline scheme display improvement in the ergodic secrecy capacity. The proposed scheme and the baseline scheme saturate when the SNR is beyond 15dB and 10dB respectively. This shows that a low to medium operating SNR is sufficient for both schemes. The ergodic secrecy capacity of the baseline scheme degrades when K increases. This is due to the fact that only the eavesdroppers benefit from the multiuser diversity through the increase of K. In contrast, the ergodic secrecy capacity of the proposed scheme increases with K. This verifies that the proposed scheme is able to leverage the multi-user diversity from increasing K.

6.6 Chapter Conclusion

Previously-proposed schemes have demonstrated the difficulty in providing information-theoretic secrecy for the wireless broadcast of a message from a base station to one of K wireless users; in particular, for standard multi-user diversity approaches, the secrecy capacity goes to zero in the limit of a large number of users. This chapter has exploited wiretap coding and reciprocity of the fading channel between the base station and the wireless users to address this problem. The proposed methods provide a positive secrecy capacity in the limit of a large number of users for both the opportunistic and non-opportunistic scenarios. Numerical results demonstrate that the proposed methods are also desirable in the finite case for more than a moderate number of users, i.e. 10-20 users, at SNR of interest.

Chapter 7

Conclusions and Future Works

In the first section of this chapter, conclusions are presented to provide a concise summary of the important findings in this thesis. The possible directions of future work are discussed in the second section.

7.1 Conclusions

This thesis has proposed several wireless transmission protocols using relays in the multi-user broadcast channels and the information exchange channels. It has been demonstrated that the use of relays is able to provide an additional dimension to improve the performance of wireless transmission in terms of reliability, throughput, power efficiency and secrecy. The findings of this thesis are summarised in the following paragraphs.

First, a spectrally efficient cooperative transmission protocol using linear precoding for the MISO broadcast channel has been proposed in Chapter 3, to improve the reliability of the broadcast transmission. Using practical zero-forcing beamforming at the source and relays, the proposed protocol is able to avoid the co-channel interference among multiple destinations and provide another dimension to improve the reliability of wireless transmission. The non-orthogonal relaying strategy used in the proposed protocol allows the source and the relays to access the shared bandwidth simultaneously, which results in a spectrally efficient cooperative transmission. The outage behaviour and the diversity and multiplexing tradeoff of the proposed protocol are analysed. Analytical results show that the maximum diversity gain expressed as the summation of the number of source transmitter antennas

and the available relays can be achieved. Analytical results also reveal that the proposed protocol uniformly dominates the non-cooperative protocol. In the event that the number of available relays is large, the diversity and multiplexing tradeoff of the proposed protocol significantly outperforms the comparable non-cooperative protocol. Numerical results further verify that the proposed protocol achieves better robustness than the comparable scheme. Both the analytical and simulation results agree that the use of multiple relays is beneficial in terms of improving the tradeoff between data rate and reliability.

Second, an analogue network coding protocol has been proposed for the information exchange channel with multiple user pairs and a multi-antenna relay in Chapter 4, to improve the data rate and reliability of the information exchange. The proposed protocol integrates analogue network coding and spatial multiplexing to enable both the relay and the users to participate in interference cancellation. Several low-complexity beamforming schemes are proposed for the multi-antenna relay to exploit the multiplexing and diversity gains. The proposed beamforming schemes deliver significant improvements in terms of ergodic capacity and outage probability. The developed analytical bounds approximate the ergodic capacity closely and show that the proposed beamforming scheme achieves higher multiplexing gain than existing schemes when the number of antennas at the relay is less than the total number of users. The analytical result on the outage probability is developed to quantify the diversity gain. The analytical result proves that the proposed beamforming schemes are able to extract all additional diversity gain offered by block-diagonalisation when the number of antennas at the relay is at least the total number of users. The diversity and multiplexing tradeoff of the proposed two-way relaying protocol is also derived. It reveals that the rate and reliability tradeoff offered by the proposed protocol outperforms existing schemes. Simulation results agree with the analytical results that the proposed protocol achieves higher ergodic capacity, lower outage probability and better diversity and multiplexing tradeoff than comparable schemes.

Third, the study on the information exchange channel extends to the scenario where both the users and the relay are equipped with multiple antennas. Specifically in Chapter 5, a low-complexity joint beamforming and power management scheme has been proposed to exploit the additional dimension offered by the use of multiple antennas to improve the sum-rate of the network. Based on the idea of subchannel alignment, transmit and receive beamformers at all nodes are designed such that user channel pair can be decomposed into parallel subchannels. The channel decomposition not only fa-

cilitates substream power allocation and power control, but also enables the use of low-complexity SISO decoders at the users. Joint power allocation dynamically allocates power to all substreams and nodes in order to maximise the sum-rate, subject to a total power constraint. On the other hand, joint power control minimises the total transmission power while satisfying the predefined target data rates. The convexity of the power allocation and power control formulations are determined. The non-concave power allocation utility function is approximated by a concave upper bound while the power control formulation are transformed from geometric program into equivalent convex form, in order to facilitate the use of efficient convex optimisation techniques in solving the optimisation problems. Numerical simulation results demonstrate that the proposed joint beamforming and power management scheme is able to deliver significant sum-rate improvement or achieve substantial transmission power saving, if compared to existing schemes.

Last but not least, the secrecy of the wireless scalar broadcast channel has been studied in Chapter 6. Two-way secrecy schemes which combine one-time pad encryption and wiretap coding are proposed for the scalar broadcast channel. The proposed schemes prevent idle users from exploiting the multi-user diversity and abusing the broadcast nature of wireless transmission to eavesdrop the information transmitted to the legitimate user. Two cases, opportunistic and non-opportunistic broadcast are considered. The proposed schemes utilise the forward and backward channels of the legitimate user to average out the channel gain of the eavesdroppers. In non-opportunistic broadcast, relays are selected from idle users to forward secret messages and to prevent the eavesdroppers from gaining an advantage of the multi-user diversity. The asymptotic secrecy rate of the proposed schemes in a large network is analysed. Analytical results reveal that the proposed schemes are able to achieve positive secrecy rates in both opportunistic and non-opportunistic broadcast even when the number of users goes to infinity. Simulations results verify that the proposed schemes are able to achieve positive ergodic secrecy capacity and significantly outperform the conventional scheme when the number of users is large.

Overall, the results in this thesis provide new insights into the potential capability of relays in the wireless broadcast and information exchange channels. All of the proposed wireless transmission protocols have demonstrated that the efficient use of relays, multi-antenna beamforming, power management and multi-user diversity are able to provide significant improvement in various aspects including throughput, reliability, power efficiency and secrecy.

7.2 Future Work

Following the studies in this thesis, several future directions are suggested to further improve the transmission protocols proposed in this thesis and to extend the application of the proposed protocols to more complicated network scenarios.

First, it would be interesting to investigate the sensitivity of the proposed protocols towards imperfect CSI and time-varying channels. Beamforming and interference cancellation procedures in the proposed CBC protocol and the proposed two-way relaying protocols might need to be adjusted to increase the immunity from imperfect CSI and time-varying channels. Dynamic rate and power management can be performed to exploit the degrees of freedom benefit available in time-varying channels.

Second, developing techniques to enable reusable relays in the proposed CBC protocol is another research direction. Recall the proposed CBC protocol in Chapter 3. A relay can only be used once in each CBC transmission time frame due to the half duplex constraint, which results in the relay missing the current source transmission when it is relaying the previous source message. Without knowledge of the source message in the kth time slot, the relay is not able to decode the new source message in the subsequent $(k+1)$th time slot where the received signal is a mixture of new source message and the previous source message (broadcast by other relay). A potential solution is to use alternating relay pair which has channels orthogonal to each other. The channel orthogonality can be achieved by means of beamforming (using more antennas) or relay scheduling. The channel orthogonality prevents the relays from interfering each other. As a result, the relays only hear the transmission from the source. In this way, the relay pair can be used alternately for many times until the channels change. The performance of such a strategy still remains an open question.

Third, the study of more complicated two-way relaying scenarios such as the multi-pair MIMO two-way relaying channels is another challenging research direction. In such scenarios, multiple pairs of users attempt to engage in information exchange simultaneously in the same channel with the help of relays. This scenario is more complicated if compared to the MIMO interference channels where information only flows in one direction. Although it has been shown that interference alignment [70] is able to achieve the best degrees of freedom in the MIMO interference channels, the applicability of interference alignment in the multi-pair MIMO two-way relaying channels remains unexplored.

Fourth, the effect of path loss to the proposed two-way secrecy protocol is another area worth investigation. Under large scale path loss, eavesdroppers located in proximity of the source have significant channel advantage over the legitimate user. This leads to the "near eavesdropper" problem. The proposed two-way secrecy protocol has the potential to overcome the "near eavesdropper" problem. The key lies in properly selecting relays which can minimise the channel advantage obtained by the near eavesdroppers. The investigation of the proposed protocol under both path loss and fading in large network is valuable to give insight on the asymptotic performance of the proposed two-way secrecy protocol.

Appendix A

Proof of Lemma 3.3

From Lemma 3.1, $\mathbf{h_k^H Q \Lambda Q^H h_k} = \mathbf{h}_k^H \mathbf{Q \Lambda Q}^H \mathbf{Q \Lambda Q}^H \mathbf{h}_k$ since $\mathbf{Q \Lambda Q}^H$ is idempotent. Due to the fact that the columns of $\mathbf{H}$ are independent of each other, the unitary matrix $\mathbf{Q}$ does not alter the distribution of $\mathbf{h}_k$. Using Lemma 3.2, it can be shown that $\mathbf{h}_k^H \mathbf{Q \Lambda Q}^H \sim \mathcal{CN}_{Mr}(0, ||\mathbf{w}_k||^2 \Sigma)$. According to [71], the sum-of-square of standardised normal variables $z = \frac{\mathbf{h}_k^H \mathbf{Q \Lambda Q}^H \mathbf{\Sigma}^{-1} \mathbf{Q \Lambda Q}^H \mathbf{h}_k}{||\mathbf{w}_k||^2}$ is chi-square distributed with $2(M_r - M_t + 1)$ degrees of freedom since the rank of $\mathbf{\Lambda}$ is $M_r - M_t + 1$. Assuming that $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_{Mr}$ (independent and identical distributed with variance σ^2), the PDF of the chi-square variable z is

$$f(z) = \frac{z^{M_r - M_t}}{\Gamma(Mr - M_t + 1)} \exp(-z). \tag{7.1}$$

Let $c = \mathbf{h}_k^H \mathbf{Q \Lambda Q}^H \mathbf{h}_k$, applying the change of variable, $z = \frac{c}{\sigma^2 ||\mathbf{w}_k||^2}$, $\mathrm{d}z = \frac{1}{\sigma^2 ||\mathbf{w}_k||^2} \mathrm{d}c$ and therefore $\int f(z) \mathrm{d}z = \int f(c) \frac{1}{\sigma^2 ||\mathbf{w}_k||^2} \mathrm{d}c$. Thus, the statistical distribution of c can be expressed as

$$f(c) = \frac{c^{Mr - M_t}}{\left(\sigma^2 ||\mathbf{w}_k||^2\right)^{(M_r - M_t + 1)} \Gamma(M_r - M_t + 1)} \exp\left(-\frac{c}{\sigma^2 ||\mathbf{w}_k||^2}\right), \tag{7.2}$$

where $\Gamma(M_r - M_t + 1) = (M_r - M_t)!$. From Lemma 3.1, applying the change of variable, $c = \frac{\gamma_k}{\gamma_0}$, $\mathrm{d}c = \frac{1}{\gamma_0} \mathrm{d}\gamma_k$ and therefore $\int f(c) \mathrm{d}c = \int f(\gamma_k) \frac{1}{\gamma_0} \mathrm{d}\gamma_k$, the PDF of the instantaneous SNR at the kth stream can be written as

$$f(\gamma_k) = \frac{\left(\gamma_k\right)^{M_r - M_t}}{\left(\gamma_0 \sigma^2 ||\mathbf{w}_k||^2\right)^{(M_r - M_t + 1)} (M_r - M_t)!} \exp\left(-\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}\right). \tag{7.3}$$

By combining the result from the table of integral, Section 3.351.1 in [72], the closed form equation for the CDF of the kth stream's instantaneous SNR is as following,

$$F(\gamma_k) = 1 - \exp\left(-\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}\right) \sum_{n=0}^{M_r - M_t} \frac{1}{n!} \left(\frac{\gamma_k}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}\right)^n, \tag{7.4}$$

and the lemma is proved. ■

Proof of Lemma 3.4

From Lemma 3.3, the probability that the SNR at the kth stream of the relay receiver is less than or equal to γ_k is expressed in (3.16). In order for the relay to decode each stream successfully, the instantaneous SNR at the kth stream must satisfy the condition $\gamma_k > 2^R - 1$, where R is the target data rate of each stream in bit/s/Hz. Formally, the probability that the kth stream can be successfully decoded can be expressed as

$$P(\gamma_k > 2^R - 1) = 1 - P(\gamma_k \leq 2^R - 1). \tag{7.5}$$

Substituting (3.16) into the equation above, it follows that

$$P(\gamma_k > 2^R - 1) = \exp\left(-\frac{2^R - 1}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}\right) \sum_{n=0}^{M_r - M_t} \frac{1}{n!} \left(\frac{2^R - 1}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}\right)^n. \tag{7.6}$$

Define the variable $\gamma = \frac{2^R - 1}{\gamma_0 \sigma^2 ||\mathbf{w}_k||^2}$ for simplicity of notation. The previous equation for the probability that the kth stream can be decoded by the relay is simplified into the following expression,

$$P(\gamma_k) = \exp(-\gamma) \sum_{n=0}^{M_r - M_t} \frac{1}{n!} (\gamma)^n, \tag{7.7}$$

where the notation $P(\gamma_k)$ implies $P(\gamma_k > 2^R - 1)$. Due to symmetrical channel assumption, the worst case SNR requirement is identical for each stream, i.e $P(\gamma_1) = P(\gamma_2) .... = P(\gamma_{M_r})$ for a total of M_r receiver antennas at each relay node. The SNR at each stream is also assumed to be independent of each other. A relay is qualified if and only if all M_r streams of the received messages can be

decoded successfully. In other words, a relay, R_k is qualified when the condition (3.3) is satisfied $\forall m \in \{1, 2, \ldots, M_r\}$. Therefore, the probability that a relay is qualified can be represented by the joint probability of the $P(\gamma_k)$ of all M_r streams. Define A as the event when a relay is qualified, the probability that a relay is qualified can be expressed as

$$P(A) = P(\gamma_1 \cap \gamma_2 \cap \cdots \cap \gamma_{M_r}) = P(\gamma_1)P(\gamma_2)\ldots P(\gamma_{M_r}), \tag{7.8}$$

since the SNR of each stream is independent of each other. Substitute (7.7) into (7.8), the probability that a relay is qualified,

$$P(A) = \left[\exp(-\gamma) \sum_{n=0}^{M_r - M_t} \frac{1}{n!}(\gamma)^n\right]^{M_r}. \tag{7.9}$$

By expanding the Taylor series of the exponential function, i.e. $\exp(-\gamma) = 1 - \gamma + \frac{\gamma^2}{2!} - \frac{\gamma^3}{3!} + \ldots$, the $P(A)$ can be written as

$$P(A) = \left[\left(1 - \gamma + \frac{\gamma^2}{2!} - \frac{\gamma^3}{3!}\ldots\right)\left(1 + \gamma + \frac{\gamma^2}{2!} + \ldots + \frac{\gamma^{M_r - M_t}}{(M_r - M_t)!}\right)\right]^{M_r}. \tag{7.10}$$

It can be observed that the right hand side of the equation above is a multiplication of an infinite series and an $(M_r - M_t)$th-order polynomial. Represent the infinite series $\left(1 - \gamma + \frac{\gamma^2}{2!} - \frac{\gamma^3}{3!}\ldots\right)$ as $p(\gamma) = p_0 + p_1\gamma + p_2\gamma^2 + \cdots$, and the polynomial $\left(1 + \gamma + \frac{\gamma^2}{2!} + \ldots + \frac{\gamma^{M_r - M_t}}{(M_r - M_t)!}\right)$ as $q(\gamma) = q_0 + q_1\gamma + \ldots + q_{M_r - M_t}\gamma^{M_r - M_t}$, the multiplication of $p(\gamma)$ and $q(\gamma)$ can be written as $r(\gamma) = p(\gamma)q(\gamma) = r_0 + r_1\gamma + r_2\gamma^2 + \ldots$, which is an infinite series. The ith coefficient of the infinite series $r(\gamma)$ can be expressed in the form of summation,

$$r_i = \sum_{j=0}^{i} p_j q_{i-j}, \tag{7.11}$$

where $p_i = \frac{(-1)^i}{i!}$ and $q_i = \begin{cases} \frac{1}{i!} & \text{when} \quad 0 \leq i \leq M_r - M_t \\ 0 & \text{when} \quad i > M_r - M_t \end{cases}$. For the case when $1 \leq i \leq M_r - M_t$, the ith coefficient is

$$r_i = \frac{(-1)^0}{0!i!} + \frac{(-1)^1}{1!(i-1)!} + \ldots + \frac{(-1)^i}{i!0!} = \frac{1}{i!}\sum_{k=0}^{i} \frac{i!}{k!(k-i)!}(-1)^k. \tag{7.12}$$

Using the binomial theorem, i.e. $(1+x)^i = \sum_{k=0}^{i} \begin{pmatrix} i \\ k \end{pmatrix} x^k$, $x = -1$, the above equation is simplified into

$$r_i = \frac{1}{i!}(1-1)^i = 0. \tag{7.13}$$

Interestingly, the ith coefficient $r_i = 0$ when $1 \leq i \leq M_r - M_t$. As a result, the equation of $r(\gamma)$ is reduced to the following infinite series

$$r(\gamma) = 1 + r_{M_r - M_t + 1}\gamma^{M_r - M_t + 1} + r_{M_r - M_t + 2}\gamma^{M_r - M_t + 2} + \ldots. \tag{7.14}$$

Using high SNR approximation, the previous equation can be further simplified into a finite series with two elements,

$$r(\gamma) \approx 1 + r_{M_r - M_t + 1}\gamma^{M_r - M_t + 1}. \tag{7.15}$$

Next, the value of the coefficient $r_{M_r - M_t + 1}$ is determined. Recall (7.11), $r_{M_r - M_t + 1}$ can be expanded into a finite series

$$\begin{aligned} r_{M_r - M_t + 1} &= p_0 q_{M_r - M_t + 1} + p_1 q_{M_r - M_t} + \\ &\quad \ldots + p_{M_r - M_t} q_1 + p_{M_r - M_t + 1} q_0. \end{aligned} \tag{7.16}$$

Following that coefficient $q_{M_r - M_t + 1} = 0$ and using (7.11),

$$\begin{aligned} r_{M_r - M_t + 1} &= \frac{(-1)^1}{1!\,(M_r - M_t)!} + \frac{(-1)^2}{2!(M_r - M_t - 1)!} + \\ &\quad \ldots + \frac{(-1)^{M_r - M_t}}{(M_r - M_t)!1!} + \frac{(-1)^{M_r - M_t + 1}}{(M_r - M_t + 1)!0!}. \end{aligned} \tag{7.17}$$

Using the binomial theorem, the above equation is reduced to

$$r_{M_r-M_t+1} = \frac{1}{(M_r-M_t+1)!}\left[\sum_{k=0}^{M_r-M_t+1}\left(\frac{(M_r-M_t+1)!}{k!\,(M_r-M_t+1-k)!}(-1)^k\right)-1\right], \quad (7.18)$$

$$= \frac{1}{(M_r-M_t+1)!}\left[(1-1)^{M_r-M_t+1}-1\right], \quad (7.19)$$

$$= \frac{-1}{(M_r-M_t+1)!}. \quad (7.20)$$

Hence, at high SNR, the coefficient $r(\gamma)$ can be approximated as a simple two terms summation,

$$r(\gamma) \approx 1 - \frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}. \quad (7.21)$$

Therefore, the probability that a relay is qualified can be written as a $r(\gamma)$ with power of M_r, such that

$$P(A) \approx \left[1 - \frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}\right]^{M_r}. \quad (7.22)$$

Applying the binomial theorem to reach the approximation, $(1-x)^n \approx 1-nx$. The probability that a relay is qualified can be expressed as

$$P(A) \approx 1 - M_r\frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}, \quad (7.23)$$

and the lemma is proved. Note that under Rayleigh fading, $\gamma = \frac{2^R-1}{\gamma_0||\mathbf{w}_k||^2}$ since $\sigma^2 = 1$. ■

Proof of Lemma 3.5

Say in a broadcast scenario, there are L available relays, K qualified relays and $K \leq L$. The probability that K relays can decode the source message successfully has a binomial distribution which is expressed as

$$P(K) = \begin{pmatrix} L \\ K \end{pmatrix} P(A)^K(1-P(A))^{L-K}. \quad (7.24)$$

Using the result of the probability that a relay is qualified, $P(A)$ in Lemma 3.4, the following high SNR approximation is obtained,

$$P(K) \approx \frac{L!}{(L-K)!K!}\left(1 - M_r\frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}\right)^K \left(M_r\frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}\right)^{L-K}. \tag{7.25}$$

Since $M_r\frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!} \ll 1$ as $\gamma_0 \to \infty$, the previous equation can be simplified into

$$P(K) \approx \frac{L!}{(L-K)!K!}\left(M_r\frac{\gamma^{M_r-M_t+1}}{(M_r-M_t+1)!}\right)^{L-K} \doteq \gamma_0^{-(1-r)(M_r-M_t+1)(L-K)}. \tag{7.26}$$

As the SNR $\gamma_0 \to \infty$, $P(K = L) \to 1$ and the probability for $K \neq L$ is very small such that $P(K \neq L) \to 0$. Thus, at high SNR, all available relays can decode the M_r streams of message successfully and the lemma is proved. ■

Proof of Lemma 3.6

Recall the mutual information of the proposed CBC protocol in (3.9), the outage probability for the event that K relays are selected can be expressed as

$$P(\mathcal{O}_{\mathrm{K}}) = P\left(\frac{1}{K+1}\log_2\det\left(\mathbf{I}_{K+1} + \gamma_{\mathbf{0}}\mathbf{H}_m\mathbf{H}_m^H\right) \leq R\right). \tag{7.27}$$

Equivalently, the outage probability can be written as

$$P(\mathcal{O}_{\mathrm{K}}) = P\left(\det\left(\mathbf{I}_{K+1} + \gamma_0\mathbf{H}_m\mathbf{H}_m^H\right) \leq 2^{(K+1)R}\right). \tag{7.28}$$

It can be observed that the multiplication of the bi-diagonal matrix $\mathbf{H}_m$, i.e $\mathbf{H}_m\mathbf{H}_m^H$, is a tri-digonal matrix. As a result, $\mathbf{I}_{K+1} + \gamma_0\mathbf{H}_m\mathbf{H}_m^H$ is also a tri-diagonal matrix. According to [73], the determinant of a tri-diagonal matrix can be computed recursively using the following equation,

$$D_n = \left[1 + \gamma_0 x + \gamma_0 z_{n-1}\right] D_{n-1} - \gamma_0^2 x z_{n-1} D_{n-2}, \tag{7.29}$$

where $D_n = \det\left[\mathbf{I}_n + \gamma_0 \mathbf{H}_{n-1}\mathbf{H}_{n-1}^H\right]$, $\mathbf{H}_n$ is the top left submatrix of $\mathbf{H}_m$ with dimension $n \times n$, the variable $x = ||\mathbf{h}_m||^2$ is the norm square of the element on the principal diagonal of $\mathbf{H}_m$ and the variable $z_{n-1} = |\mathbf{g}_{R_{(n-1)},m}^T \mathbf{p}_{R_{(n-1)},m}|^2$ is the absolute square of the sub-diagonal element of $\mathbf{H}_m$. Applying such a property, the following inequality is obtained,

$$\det\left(\mathbf{I}_{K+1} + \gamma_0 \mathbf{H}_m \mathbf{H}_m^H\right) \geq (1+\gamma_0 x)^{K+1} + \prod_{i=1}^{K} (\gamma_0 z_i)\,. \tag{7.30}$$

The proof can be achieved by following same steps as in [74]. Such lower bound facilitates the subsequent analytical development, by simplifying the determinant as a sum of the variable x and the product $\prod_i^K \gamma_0 z_i$. Using the inequality, the following upper bound expression for outage probability is obtained,

$$\begin{aligned} P(\mathcal{O}_{\mathrm{K}}) &\leq P((1+\gamma_0 x)^{K+1} + \prod_{i=1}^{K} \gamma_0 z_i) \leq 2^{(K+1)R}), & (7.31)\\ &\leq P((1+\gamma_0 x) \leq 2^R) P(\prod_{i=1}^{K} \gamma_0 z_i \leq 2^{(K+1)R}). & (7.32) \end{aligned}$$

Note that the upper bound implies the worst case outage probability. Next, the distribution of x is determined in order to obtain the closed form equation for the first term of the (7.32). Since the distribution of $\mathbf{h}_m \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{M_t})$, the variable $x = ||\mathbf{h}_m||^2$ has a chi-square distribution with $2M_t$ degrees of freedom. Following similar steps as in the proof of Lemma 3.3, the PDF of x can be expressed as

$$f(x) = \frac{x^{M_t - 1} \exp(\frac{-x}{\sigma^2})}{\sigma^{2M_t}\Gamma(M_t)}. \tag{7.33}$$

By rewriting $P((1+\gamma_0 x) \leq 2^R) = P(x \leq \frac{2^R - 1}{\gamma_0})$, the CDF of x is thus

$$P(x \leq \frac{2^R - 1}{\gamma_0}) = 1 - \exp\left(-\frac{2^R - 1}{\sigma^2 \gamma_0}\right) \sum_{n=0}^{M_t - 1} \frac{1}{n!} \left(\frac{2^R - 1}{\sigma^2 \gamma_0}\right)^n. \tag{7.34}$$

Represent $\alpha = \frac{2^R-1}{\sigma^2\gamma_0}$ for sake of simplicity and using the Taylor series of the exponent, the following equation is obtained,

$$P(x \leq \alpha) = 1 - \left(1 - \alpha + \frac{\alpha^2}{2!} - \frac{\alpha^3}{3!} + \cdots\right)\left(1 + \alpha + \frac{\alpha^2}{2!} + \ldots + \frac{\alpha^{M_t-1}}{(M_t-1)!}\right), \tag{7.35}$$

where the second term in the right hand side of the equation is a multiplication of a infinite series and a polynomial of order $(M_t - 1)$. Following similar steps as in the proof of Lemma 3.4, the following high SNR approximation, $\left(1 - \alpha + \frac{\alpha^2}{2!} - \frac{\alpha^3}{3!} + \cdots\right)\left(1 + \alpha + \frac{\alpha^2}{2!} + \ldots + \frac{\alpha^{M_t-1}}{(M_t-1)!}\right) \approx 1 - \frac{\alpha^{M_t}}{M_t!}$ can be used. Hence, by replacing $\alpha = \frac{2^R-1}{\sigma^2\gamma_0}$, one obtains the following equation,

$$P(x \leq \frac{2^R - 1}{\gamma_0}) = \frac{\left(\frac{2^R-1}{\sigma^2\gamma_0}\right)^{M_t}}{M_t!}. \tag{7.36}$$

By substituting $R = r \log_2 \gamma_0$, one obtains the following expression,

$$P(x \leq \frac{\gamma_0^r - 1}{\gamma_0}) = \frac{\left(\frac{\gamma_0^r-1}{\sigma^2\gamma_0}\right)^{M_t}}{M_t!} \doteq \gamma_0^{-M_t(1-r)}, \tag{7.37}$$

which solves the first term of (7.32). Next, the closed form for $P(\prod_{i=1}^{K} \gamma_0 z_i \leq 2^{(K+1)R})$ in (7.32) is computed. Denote the mutual information $\mathcal{I} = \ln(\prod_i^K \gamma_0 z_i)$ and $f(z_1, \ldots, z_K)$ as the joint density function of $(z_1, \ldots, z_K)$. Following similar steps in [75], one obtains the following expression for the PDF of $\mathcal{I}$,

$$q(\mathcal{I}) = \frac{\exp(\mathcal{I})}{\gamma_0^K} \int_1^{\exp(\mathcal{I})} \int_1^{\psi_1} \cdots \int_1^{\psi_{K-2}} \Psi\left(\psi_1, \cdots, \psi_{K-1}, \gamma_0\right) \prod_{i=1}^{K-1} \frac{1}{\psi_i} d\psi_i, \tag{7.38}$$

where $\Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0) = f\left(\frac{\exp(\mathcal{I})}{\gamma_0\psi_1}, \frac{\psi_1}{\gamma_0\psi_2}, \cdots, \frac{\psi_{K-2}}{\gamma_0\psi_{K-1}}, \frac{\psi_{K-1}}{\gamma_0}\right)$. In order to determine the value of the density function of $\Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0)$, the joint density function of $f(z_1, \ldots, z_K)$ shall be derived. For this purpose, the individual PDF of z_k for $k = 1, \cdots, K$ needs to be determined. Note that the subscripts of $\mathbf{g}_{R_{(k)},m}$ and $\mathbf{p}_{R_{(k)},m}$ are omitted for simplicity. It is known that $\mathbf{g} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_M)$ or equivalently $g_i \sim \mathcal{CN}(0, \sigma^2)$ for $i = 1, \cdots, M$ where g_i is the ith element of vector $\mathbf{g}$. Note that we assume full spatial multiplexing at the relay, $M_r = M$. Express z_k as the norm square of weighted sum of normal variables, i.e. $z_k = |\sum_{i=1}^{M} g_i p_i|^2$. Each z_k represents the effective channel

gain from relay K to user m. According to [54], $\sum_{i=1}^{M} g_i p_i \sim \mathcal{CN}(\sum_{i=1}^{M} p_i \mu_i, \sum_{i=1}^{M} |p_i|^2 \sigma_i^2)$ since g_i is independently distributed. As a result, $\sum_{i=1}^{M} g_i p_i \sim \mathcal{CN}(0, ||\mathbf{p}||^2 \sigma^2)$. Interestingly, since $||\mathbf{p}||^2 = 1$, $\sum_{i=1}^{M} g_i p_i \sim \mathcal{CN}(0, \sigma^2)$. Thus, z_k is exponentially distributed with PDF, $f(z) = \frac{1}{\sigma^2} \exp\left(\frac{-z}{\sigma^2}\right)$. Since $\mathbf{p}$ is orthogonal to all channels other than its desirable destination, z_k is independently distributed. The joint density function $f(z_1, \ldots, z_K) = \prod_{k=1}^{K} (f(z))$. Using this property, the density function of $\Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0)$, can be expressed as

$$\begin{aligned} \Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0) &= \left(\frac{1}{\sigma^2}\right)^K \exp\left(-\frac{\exp(\mathcal{I})}{\sigma^2 \gamma_0 \psi_1}\right) \exp\left(-\frac{\psi_1}{\sigma^2 \gamma_0 \psi_2}\right) \times \\ &\quad \cdots \times \exp\left(-\frac{\psi_{K-2}}{\sigma^2 \gamma_0 \psi_{K-1}}\right) \exp\left(-\frac{\psi_{K-1}}{\sigma^2 \gamma_0}\right), \end{aligned} \tag{7.39}$$

where $\psi_i \in [1, \exp(\mathcal{I})]$ for $i = 1, \cdots, K-1$. At high SNR, the density function of the function $\Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0)$ can be approximated as following

$$\lim_{\gamma_0 \to \infty} \Psi(\psi_1, \cdots, \psi_{K-1}, \gamma_0) = \left(\frac{1}{\sigma^2}\right)^K. \tag{7.40}$$

Hence, the high SNR approximation of the density function of the mutual information $\mathcal{I}$ is

$$q(\mathcal{I}) \approx \frac{\exp(\mathcal{I})}{(\sigma^2 \gamma_0)^K} \int_1^{\exp(\mathcal{I})} \int_1^{\psi_1} \cdots \int_1^{\psi_{K-2}} \prod_{i=1}^{K-1} \frac{1}{\psi_i} d\psi_i. \tag{7.41}$$

Applying change of variable, $\phi_i = \ln \psi_i$, $d\phi_i = \frac{1}{\psi_i} d\psi_i$, one gets the following

$$q(\mathcal{I}) \approx \frac{\exp(\mathcal{I})}{(\sigma^2 \gamma_0)^K} \int_1^{\exp(\mathcal{I})} \int_1^{\phi_1} \cdots \int_1^{\phi_{K-2}} \prod_{i=1}^{K-1} d\phi_i = \frac{\exp(\mathcal{I}) \mathcal{I}^{K-1}}{(\sigma^2 \gamma_0)^K (K-1)!}. \tag{7.42}$$

Given the density of the mutual information $\mathcal{I}$, $P(\prod_{i=1}^{K} \gamma_0 z_i \leq 2^{(K+1)R})$ can be expressed as

$$P(q(\mathcal{I}) \leq (K+1) R \ln 2) = \int_0^{(K+1)R \ln 2} \frac{\exp(\mathcal{I}) \mathcal{I}^{K-1}}{(\sigma^2 \gamma_0)^K (K-1)!} d\mathcal{I}. \tag{7.43}$$

Combining the results from the table of integrals in Section 3.351.1 of [72], the closed form equation for the CDF of $q(\mathcal{I})$ is

$$P(q(\mathcal{I}) \leq (K+1)\, R \ln 2) = \frac{1}{(\sigma^2 \gamma_0)^K} \times \left(\frac{1}{(-1)^K} - \exp((K+1)\, R \ln 2) \sum_{i=0}^{K-1} \frac{((K+1)\, R \ln 2)^i}{i!(-1)^{K-i}} \right). \quad (7.44)$$

Substituting $R = r \log_2 \gamma_0$ into the previous equation, one obtains the following expression

$$P(q(\mathcal{I}) \leq (K+1)\, r \ln \gamma_0) = \frac{1}{(\sigma^2 \gamma_0)^K} \times \left(\frac{1}{(-1)^K} - \exp((K+1)\, r \ln \gamma_0) \sum_{i=0}^{K-1} \frac{((K+1)\, r \ln \gamma_0)^i}{i!(-1)^{K-i}} \right). \quad (7.45)$$

Following similar steps in [75], one obtains the closed form expression for the CDF of the $q(\mathcal{I})$,

$$\begin{aligned} P(q(\mathcal{I}) \leq (K+1)\, r \ln \gamma_0) &= \frac{1}{(\sigma^2 \gamma_0)^K} \left(\frac{1}{(-1)^K} - \gamma_0^{(K+1)r} \sum_{i=0}^{K-1} \frac{\left(\ln \gamma_0^{(K+1)r}\right)^i}{i!(-1)^{K-i}} \right), & (7.46) \\ &\doteq \gamma_0^{-(K(1-r)-r)}. & (7.47) \end{aligned}$$

Combining equations (7.37) and (7.46), the outage probability when K relays are qualified can be expressed as the following inequality,

$$P(\mathcal{O}_{\mathrm{K}}) \leq \frac{\left(\frac{\gamma_0^r - 1}{\gamma_0}\right)^{M_t}}{M_t!\,(\sigma^2 \gamma_0)^K} \left((-1)^K - \gamma_0^{(K+1)r} \sum_{i=0}^{K-1} \frac{\left(\ln \gamma_0^{(K+1)r}\right)^i}{i!(-1)^{K-i}} \right) \doteq \gamma_0^{-[(M_t+K)(1-r)-r]}. \quad (7.48)$$

and the lemma is proved. ■

Appendix B

Proof of Lemma 4.1

First, the distribution of the auxiliary variable x is determined. With the knowledge of $\mathbf{h}_i^T \sim \mathcal{CN}(0, \mathbf{I}_N)$ and applying the idempotent property of $\mathbf{W}$, the quadratic form $\mathbf{h}_i^T \mathbf{W} \mathbf{h}_i^*$ is shown in [71] to have a chi-square distribution with 2×trace($\mathbf{W}$) degrees of freedom. In this case, trace($\mathbf{W}$)=1. Denote $u = \mathbf{h}_i^T \mathbf{W} \mathbf{h}_i^*$, the PDF of u reduces to $f_u(u) = \exp(-u)$, which is an exponential distribution. Applying change of variables, $u = \sigma^2 x$, the PDF of variable x can be expressed as $f_x(x) = \sigma^2 \exp(-\sigma^2 x)$. Using similar approach, the PDF of the variable y is $f_y(y) = \sigma^2 \exp(-\sigma^2 y)$. Next, the distribution of variable $z = \frac{xy}{(M+1)x+My+M}$ is derived. Knowing that x and y are independently distributed, the CDF of z can be written as

$$F_z(z) = \int\int_{\frac{xy}{(M+1)x+My+M}} f_x(x) f_y(y) dx dy. \tag{7.49}$$

Since $x \geq 0$ and $y \geq 0$,$F_z(z)$ can be expressed with the following equation,

$$\begin{aligned} F_z(z) &= P(0 \leq y \leq (M+1)z, 0 \leq x \leq \infty) \\ &\quad + P\left((M+1)z \leq y \leq \infty, 0 \leq x \leq \frac{Myz + Mz}{y - (M+1)z}\right), \end{aligned} \tag{7.50}$$

which is then written in integral form as follows

$$F_z(z) = \int_0^{(M+1)z} \sigma^2 \exp\left(-\sigma^2 y\right) \int_0^{\infty} \sigma^2 \exp\left(-\sigma^2 x\right) dxdy + \int_{(M+1)z}^{\infty} \sigma^2 \exp\left(-\sigma^2 y\right) \int_0^{\frac{Myz+Mz}{y-(M+1)z}} \sigma^2 \exp\left(-\sigma^2 x\right) dxdy. \tag{7.51}$$

The first pair of the integrals in (7.51) can be calculated easily,

$$\int_0^{(M+1)z} \sigma^2 \exp\left(-\sigma^2 y\right) \int_0^{\infty} \sigma^2 \exp\left(-\sigma^2 x\right) dxdy = 1 - \exp(-(M+1)\sigma^2 z), \tag{7.52}$$

while the second pair of integrals in (7.51) is

$$\begin{aligned} &\int_{(M+1)z}^{\infty} \sigma^2 \exp\left(-\sigma^2 y\right) \int_0^{\frac{Myz+Mz}{y-(M+1)z}} \sigma^2 \exp\left(-\sigma^2 x\right) dxdy \\ &= \int_{(M+1)z}^{\infty} \sigma^2 \exp\left(-\sigma^2 y\right) \left[1 - \exp\left(-\sigma^2 \frac{Myz+Mz}{y-(M+1)z}\right)\right] dy, \\ &\underset{u=y-(M+1)z}{=} \exp(-(M+1)\sigma^2 z) - \sigma^2 \exp\left(-(2M+1)\sigma^2 z\right) \int_0^{\infty} \exp\left(-\alpha u - \frac{\beta}{u}\right) du, \end{aligned} \tag{7.53}$$

where $\alpha = \sigma^2$ and $\beta = (M(M+1)z^2 + Mz)\,\sigma^2$. There is no closed-form solution available for the integral in (7.53). However, one can represent the integral according to equation 3.471.9 in [72], which is shown here for convenience,

$$\int_0^{\infty} \exp\left(-\alpha u - \frac{\beta}{u}\right) du = 2\sqrt{\frac{\beta}{\alpha}} K_1\left(2\sqrt{\alpha\beta}\right), \tag{7.54}$$

where $K_1(z)$ is the modified Bessel function of second kind. Substituting (7.54) into (7.53), one obtains the following

$$\begin{aligned} &\int_{(M+1)z}^{\infty} \sigma^2 \exp\left(-\sigma^2 y\right) \int_0^{\frac{Myz+Mz}{y-(M+1)z}} \sigma^2 \exp\left(-\sigma^2 x\right) dxdy \\ &= \exp(-(M+1)\sigma^2 z) - 2\sigma^2 \exp(-(2M+1)\sigma^2 z) \\ &\times \sqrt{M(M+1)z^2 + Mz}\, K_1\left(2\sigma^2 \sqrt{M(M+1)z^2 + Mz}\right). \end{aligned} \tag{7.55}$$

Combining (7.52) and (7.55), the CDF of z can be obtained and the lemma is proved. ■

Proof of Lemma 4.2

Since the CDF of z is related to the Bessel function, it can be bounded by first determining the upper and lower bounds for the modified Bessel function of second kind. According to equation 8.432.3 in [72], one can express $K_1(z)$ as following

$$K_1(z) = \frac{z\Gamma(\frac{1}{2})}{2\Gamma(\frac{3}{2})} \int_1^\infty \exp(-zt)\sqrt{t^2-1}dt, \qquad z \geq 0, \tag{7.56}$$

where $\Gamma(x)$ is the Gamma function. Knowing that $t \geq 1$, the following lower bound is defined,

$$\begin{aligned} K_1(z) &\geq \frac{z\Gamma(\frac{1}{2})}{2\Gamma(\frac{3}{2})} \int_1^\infty \exp(-zt)\,(t-1)\,dt \\ &= \frac{z\sqrt{\pi}}{2\sqrt{\pi}/2} \int_1^\infty \exp(-zt)\,(t-1)\,dt \\ &= z\int_1^\infty \exp(-zt)\,(t-1)\,dt \\ &\underset{y=t-1}{=} z\exp(-z)\int_0^\infty \exp(-zy)\,(y)\,dt \\ &= z\exp(-z)\left(\frac{1}{z^2} - \frac{1}{z^2}\exp(-\infty)(1+\infty)\right) \\ &= \frac{\exp(-z)}{z}. \end{aligned} \tag{7.57}$$

where equation 3.351.7 in [72] is used in to obtain the equality in (7.57). Besides, using the other integral representation of the modified bessel function as in equation 8.432.6 in [72],

$$K_1(z) = \frac{z}{4}\int_0^\infty \frac{\exp(-t-\frac{z^2}{4t})}{t^2}dt \tag{7.58}$$

Since $t \geq 0$ which leads to $\exp(-t) \leq 1$, one can have the following upper bound

$$\begin{aligned} K_1(z) &\leq \frac{z}{4}\int_0^\infty \frac{\exp(-\frac{z^2}{4t})}{t^2}dt \\ &= \frac{1}{z} \end{aligned} \tag{7.59}$$

where equation 3.471.1 in [72] is used to obtain the equality in (7.59). Using the bounded $K_1(x)$, the upper and lower bounds for the CDF of z can be obtained and thus the lemma is proved. ■

Proof of Theorem 4.1

Since the distribution function $F_z(z)$ is bounded, the density function $f_z(z)$ is bounded as well. It is shown in [76] that the inequalities between the expected value of two exponential density functions of x is such $\int_0^\infty x f_1(x)dx < \int_0^\infty x f_2(x)dy$ happens when $F_1(x) \geq F_2(x)$, and $\int_0^\infty x f_1(x)dx > \int_0^\infty x f_2(x)dy$ happens when $F_1(x) \leq F_2(x)$. Using this relationship, and let the function $g(u) = \int_0^\infty \frac{1}{2} u \exp(-uz) \log_2(1+z)dz$, one can have the following bounds for the ergodic capacity,

$$g((2M+1)\sigma^2 + 2\sqrt{M(M+1)}\sigma^2) \leq C_{erg} \leq g((2M+1)\,\sigma^2). \tag{7.60}$$

The function $g(u)$ is solved as follows

$$\begin{aligned} g(u) &= \int_0^\infty \frac{1}{2} u \exp(-uz) \log_2(1+z)dz \\ &= \frac{1}{2} u \log_2(e) \int_0^\infty \exp(-uz) \ln(1+z)dz \\ &= \frac{1}{2} \log_2(e) \exp(u) E_1(u), \end{aligned} \tag{7.61}$$

where E_1 is the exponential integral and the equality (7.61) is obtained using equation 4.337.2 in [72]. The exponential integral, $E_1(u)$ can be represented as series expansion which is found in Section 5.1.11 in [77],

$$E_1(u) = -\gamma - \ln u - \sum_{n=1}^{\infty} \frac{(-1)^n u^n}{n n!}, \tag{7.62}$$

where γ is the Euler constant. When $u \to 0$, the following approximation can be made

$$E_1(u) \approx -\gamma - \ln u, \quad \text{when } u \to 0. \tag{7.63}$$

At high SNR, the noise power $\sigma^2 \to 0$, therefore one obtains the following approximation

$$g(u) \approx \frac{1}{2}\log_2(e)\exp(u)\left(-\gamma - \ln u\right). \tag{7.64}$$

Substituting (7.64) into (7.60) and performing some algebraic manipulations, the upper and lower bounds for the ergodic capacity are obtained and the theorem is proved. ■

Proof of Theorem 4.2

Define the variables $x_k = \frac{1}{\sigma^2}||\mathbf{h}_i^T\mathbf{W}(k)||^2$ and $y_k = \frac{1}{\sigma^2}||\mathbf{W}(k)^T\mathbf{h}_j||^2$, where $\mathbf{W}(k) = \mathbf{w}(k)\mathbf{w}^H(k)$, one can rewrite the beamforming selection criterion in (4.8) as follows,

$$\begin{aligned}\arg \max_{k=1,\ldots,N-2(M-1)} &\frac{1}{2}\log_2\left(1 + \frac{x_k y_k}{(M+1)\,x_k + M y_k + M}\right) \\ &+\frac{1}{2}\log_2\left(1 + \frac{x_k y_k}{M x_k + (M+1)\,y_k + M}\right).\end{aligned} \tag{7.65}$$

It can be observed that the first term and the second term of the objective function are almost identical, except at the denominator, where the multiplier of variable x_k and the multiplier of variable y_k interchange. The minor difference between the first term and the second term does not result in any statistical difference. It can be verified that the CDF of the first term and the second term are the same. In order to simplify the analytical development, the sum rate criterion in (7.65) is reduced into

$$\arg \max_{k=1,\ldots,N-2(M-1)} \left(\frac{x_k y_k}{(M+1)\,x_k + M y_k + M}\right), \tag{7.66}$$

where only the instantaneous SNR of user i is maximised. This is named as the individual rate criterion. Figure 7.1 shows the outage probability versus SNR of the proposed null-space vector selection scheme under two different selection criteria in comparison with the baseline zero-forcing scheme. In the simulation, the fixed parameters are $N = 4$, $M = 2$ and $R = 2$ bits/s/Hz. From Figure 7.1, it can be easily seen that the outage probability obtained using sum rate criterion is almost the same as the outage probability obtained using individual rate criterion. This verifies that the individual rate criterion is a very accurate approximation for the sum rate criterion.

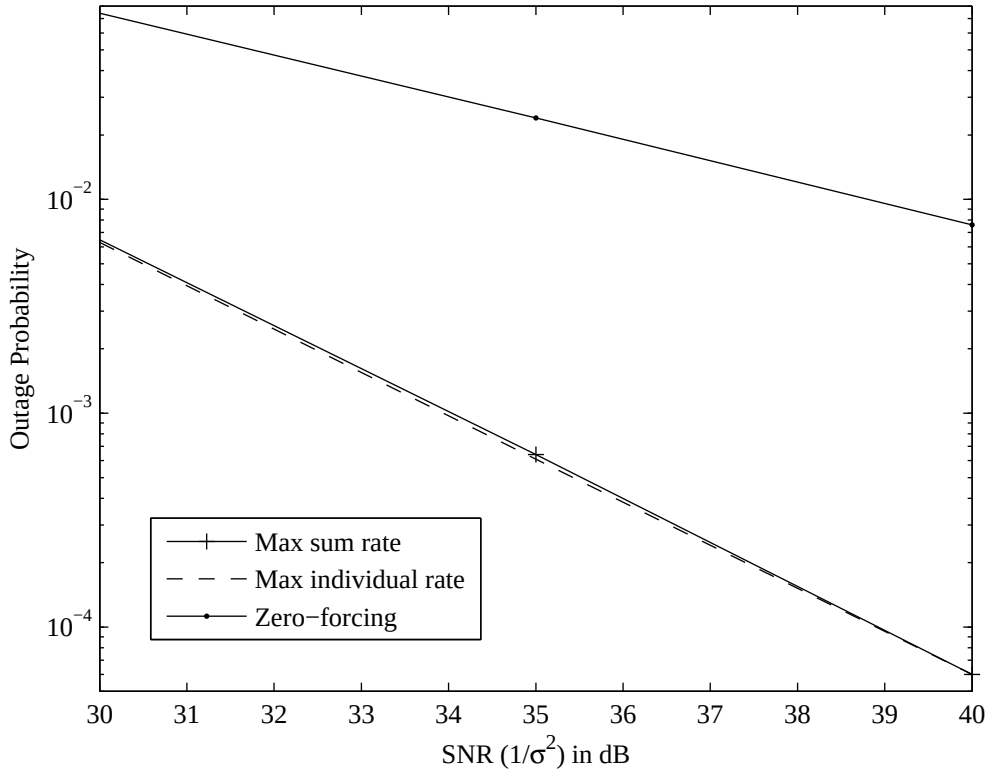


Figure 7.1: Outage probability versus SNR $(1/\sigma^2)$ under two different selection criteria in comparison with the baseline zero-forcing scheme. The fixed parameters are $M = 2$, $N = 4$ and $R = 2$ bits/s/Hz.

Using the simplified selection criterion in (7.66), one can arrange the mutual information of a user, in ascending order as follows,

$$\mathcal{I}_{(1)} \leq \mathcal{I}_{(2)} \leq \ldots \leq \mathcal{I}_{(N-2(M-1))}. \tag{7.67}$$

The null-space vector producing the maximum mutual information, $\mathcal{I}_{(N-2(M-1))}$ is selected. When the kth null-space vector is randomly chosen, the outage probability can be described as

$$P(\mathcal{I}(k) < R) = P\left(\frac{x_k y_k}{(M+1)\,x_k + M y_k + M} < 2^{2R} - 1\right). \tag{7.68}$$

Introduce the auxiliary variable $z_k = \frac{x_k y_k}{(M+1)x_k + M y_k + M}$. Following the result in Lemma 4.1, one can obtain the CDF of z_k as

$$P(z_k < R) = 1 - \exp\left(-(2M+1)\sigma^2\zeta\right)\omega K_1(\omega), \tag{7.69}$$

where $\omega = 2\sigma^2\sqrt{M(M+1)\zeta^2 + M\zeta}$ and $\zeta = 2^{2R} - 1$. In order to determine the outage probability when the best null-space vector is selected, order statistics [78] is used. Specifically, the outage probability is

$$P\left(\mathcal{I} < R\right) \;\; = \;\; \left(P(z_k < R)\right)^{N-2(M-1)}, \tag{7.70}$$

where $\mathcal{I} = \mathcal{I}_{(N-2(M-1))}$ is the mutual information when the best null-space vector is used. Substituting (7.69) into (7.70), the theorem is proved. ■

Appendix C

Complexity Analysis

This section provides the proof of the worst case computational complexity of the beamforming schemes proposed in Chapter 4. Define a floating point operation as one complex multiplication or addition. The number of floating point operations (flops) to measure the computation complexity. The complexity of the proposed schemes for two cases, case I: $N = 2M - 1$ and case II: $N \geq 2M$ are studied. The computation of the null space vectors can be obtained using Golub-Reinsch SVD technique with complexity $4mn^2 + 8n^3$ flops for any matrix with m rows and n columns [79].

Case I: $N = 2M - 1$

Refer to table 7.1, the overall complexity is at most $\mathcal{O}(MN^3)$ flops.

Key Operations	Complexity
$\mathbf{w}_{i,j} = \text{null}(\tilde{\mathbf{H}}_{i,j}^T)$, where $\tilde{\mathbf{H}}_{i,j}^T \in \mathbb{C}^{N-1\times N}$	$\mathcal{O}\left(4(N-1)N^2 + 8N^3\right) = \mathcal{O}(N^3)$
$\mathbf{F}_{i,j} \in \mathbb{C}^{N\times N} = \mathbf{w}_{i,j}\mathbf{w}_{i,j}^T$	$\mathcal{O}(N^2)$
$\|\|\mathbf{F}_{i,j}\mathbf{h}_i\|\|^2$	$\mathcal{O}(N^2)$
Overall for M user pairs	$\mathcal{O}(MN^3)$

Table 7.1: Complexity of key operations when $N = 2M - 1$.

Case II: $N \geq 2M$

Null-Space Vector Selection

Refer to table 7.2, the overall complexity is at most $\mathcal{O}(MN^3)$ flops.

Key Operations	Complexity
$\mathbf{w}_{i,j} = \text{null}(\tilde{\mathbf{H}}_{i,j}^T)$, where $\tilde{\mathbf{H}}_{i,j}^T \in \mathbb{C}^{2(M-1)\times N}$	$\mathcal{O}\left(4(2M-2)N^2 + 8N^3\right) = \mathcal{O}(N^3)$
$\mathbf{F}_{i,j}(k) \in \mathbb{C}^{N\times N} = \mathbf{w}_{i,j}(k)\mathbf{w}_{i,j}^T(k)$ and repeated for $N - 2(M-1)$ times.	$\mathcal{O}(N^3 - 2MN^2 + 2N^2) = \mathcal{O}(N^3)$
$\|\|\mathbf{F}_{i,j}(k)\mathbf{h}_i\|\|^2$ and repeated for $N - 2(M-1)$ times.	$\mathcal{O}(N^3 - 2MN^2 + 2N^2) = \mathcal{O}(N^3)$
Exhaustive search on a sequence of $N - 2M + 1$ elements	$\mathcal{O}(N - 2M + 1)) = \mathcal{O}(N)$
Overall for M user pairs	$\mathcal{O}(M \times \max(N^3, MN^2) = \mathcal{O}(MN^3)$

Table 7.2: Complexity of key operations of the null-space vector selection scheme when $N \geq 2M$.

Coherent Combining of Null-space Vectors

Refer to table 7.3, the overall complexity is at most $\mathcal{O}(MN^3)$ flops.

Key Operations	Complexity
$\mathbf{W}_{i,j} = \text{null}(\tilde{\mathbf{H}}_{i,j}^T)$, where $\tilde{\mathbf{H}}_{i,j}^T \in \mathbb{C}^{2(M-1)\times N}$	$\mathcal{O}\left(4(2M-2)N^2 + 8N^3\right) = \mathcal{O}(N^3)$
$\Phi_i = \sum_{k=1}^{N-2(M-1)} \|\mathbf{h}_i^T\mathbf{w}(k)\|^2$	$\mathcal{O}(N^2 - 2MN + N) = \mathcal{O}(N^2)$
$\mathbf{F}_{i,j} \in \mathbb{C}^{N\times N} = \mathbf{W}_{i,j}\mathbf{B}_{i,j}\mathbf{P}_\pi\mathbf{B}_{i,j}^T\mathbf{W}_{i,j}^T$	$\mathcal{O}\left(2N^2 - 4MN + 2N\right) = \mathcal{O}(N^2)$
$\|\|\mathbf{F}_{i,j}\|\|_F^2$	$\mathcal{O}(N^3)$
Overall for M user pairs	$\mathcal{O}(MN^3)$

Table 7.3: Complexity of key operations of the coherent combining scheme when $N \geq 2M$.

Pair-aware with semi-definite relaxation (PA-SDR) [9]

The complexity of the PA-SDR is dominated by the complexity of the semi-definite programming used to determine the multicast vectors. From [58], it can be easily shown that the complexity of PA-SDR is at most $\mathcal{O}\left(\left(2 + (N - 2(M-1))^2\right)^{3.5}\right) = \mathcal{O}(N^7)$ flops.

Appendix D

Convexity Analysis of (5.16)

The convexity of the objective function in (5.16) is analysed. In order to determine the convexity of the objective function, the composition rule [20] is used. Since positive weighted sum of any convex (or concave) functions preserves convexity (or concavity), one only needs to check the concavity of the kth substream rate of user 1, $\log_2\left(1+\frac{t_{1,k}\tilde{b}_k\tilde{c}_k}{t_{2,k}\tilde{a}_k+t_{3,k}\tilde{b}_k+t_{4,k}\tilde{c}_k+t_{5,k}}\right)$ and the kth substream rate of user 2, $\log_2\left(1+\frac{u_{1,k}\tilde{a}_k\tilde{c}_k}{u_{2,k}\tilde{a}_k+u_{3,k}\tilde{b}_k+u_{4,k}\tilde{c}_k+u_{5,k}}\right)$. Representing variable $x=\tilde{a}_k$, $y=\tilde{b}_k$ and $z=\tilde{c}_k$. Define function $f(x,y,z)=h(g(x,y,z))$ where $h(g)=\log_2 g$ and $g(x,y,z)=1+\frac{t_1yz}{t_2x+t_3y+t_4z+t_5}$. Note that the subscript k of the constants $t_{j,k}$ is omitted for simplicity. Recall the composition rule [20], function f is concave if

1. h is concave, $\tilde{h}$ is non-decreasing, and g is concave; or
2. h is concave, $\tilde{h}$ non-increasing, and g is convex,

where $\tilde{h}$ is the extended-value extension of function h. It can be easily verified that h is concave and $\tilde{h}$ is non-decreasing. Next, the concavity of function g is determined using Hessian matrix. The Hessian matrix can be calculated as follows,

$$\mathbf{H}(g)=\begin{bmatrix}\frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial x\partial y} & \frac{\partial^2 g}{\partial x\partial z}\\ \frac{\partial^2 g}{\partial y\partial x} & \frac{\partial^2 g}{\partial y^2} & \frac{\partial^2 g}{\partial y\partial z}\\ \frac{\partial^2 g}{\partial z\partial x} & \frac{\partial^2 g}{\partial z\partial y} & \frac{\partial^2 g}{\partial z^2}\end{bmatrix}. \tag{7.71}$$

If function g is jointly concave with respect to to parameter set (x, y, z), then g must be concave with respect to each of the variables. Note that the converse is not true. In other words, the diagonal elements of the Hessian matrix needs to be non-positive. Since the constants $t_j \geq 0, \forall j = \{1, \ldots, M\}$, and the variables $x \geq 0$, $y \geq 0$ and $z \geq 0$, the main diagonal elements can be computed as follows,

$$\frac{\partial^2 g}{\partial x^2} = \frac{2t_1 t_2^2 yz}{(t_2x + t_3y + t_4z + t_5)^3} \geq 0, \tag{7.72}$$

$$\frac{\partial^2 g}{\partial y^2} = -\frac{2t_1 t_3 z(t_2x + t_4z + t_5)}{(t_2x + t_3y + t_4z + t_5)^3} \leq 0, \tag{7.73}$$

$$\frac{\partial^2 g}{\partial z^2} = -\frac{2t_1 t_4 y(t_2x + t_3y + t_5)}{(t_2x + t_3y + t_4z + t_5)^3} \leq 0. \tag{7.74}$$

From the results, one can conclude that the function g is concave with respect to y, g is concave with respect to z but the function g is non-concave with respect to x. Therefore, function g is not jointly concave with respect to the parameter set (x, y, z). Following the composition rule, it is obvious that function f, which corresponds to the kth substream rate function of user 1, is non-concave. Applying similar steps, it can be shown that the kth substream rate function of user 2 is also non-concave. Since the kth substream rate function of both users are non-concave, it can be concluded that the sum-rate objective function in (5.16) is non-concave with respect to variables $\tilde{a}_1, \ldots, \tilde{a}_M, \tilde{b}_1, \ldots, \tilde{b}_M, \tilde{c}_1, \ldots, \tilde{c}_M$. ■

Proof of Theorem 5.1

The upper bound of the objective function can be derived from the inequality of the kth substream SNR of user 1 given by $\gamma_{1,k} = \frac{t_{1,k}\tilde{b}_k\tilde{c}_k}{t_{2,k}\tilde{a}_k + t_{3,k}\tilde{b}_k + t_{4,k}\tilde{c}_k + t_{5,k}}$ and the kth substream SNR of user 2 given by $\gamma_{2,k} = \frac{u_{1,k}\tilde{a}_k\tilde{c}_k}{u_{2,k}\tilde{a}_k + u_{3,k}\tilde{b}_k + u_{4,k}\tilde{c}_k + u_{5,k}}$. Represent $x = \tilde{a}_k$, $y = \tilde{b}_k$ and $z = \tilde{c}_k$, and omit subscript k of the constants $t_{j,k}$ for simplicity, the kth substream SNR of user 1 can be expressed as $\gamma_{1,k} = \frac{t_1 yz}{t_2x + t_3y + t_4z + t_5}$. Using the fact that $x \geq 0$, the kth substream SNR of user 1 can be upper bounded as follows,

$$\frac{t_1 yz}{t_2 x + t_3 y + t_4 z + t_5} \leq \frac{t_1 (x+y) z}{t_2 x + t_3 y + t_4 z + t_5}, \tag{7.75}$$

$$\leq \frac{t_1 (x+y) z}{t_n (x+y) + t_4 z + t_5}, \tag{7.76}$$

$$\leq \frac{(x+y) z}{t_a (x+y) + t_b z}, \tag{7.77}$$

where the second inequality is obtained by choosing $t_n = \min(t_2, t_3)$, the third inequality is obtained by omitting t_5, and normalising $t_a = \frac{t_n}{t_1}$ and $t_b = \frac{t_4}{t_1}$. Since $y \geq 0$, similar steps can be applied to the kth substream SNR of user 2 to get the following upper bound,

$$\frac{u_1 xz}{u_2 x + u_3 y + u_4 z + u_5} \leq \frac{(x+y) z}{u_a (x+y) + u_b z}, \tag{7.78}$$

where $u_a = \frac{\min(u_2, u_3)}{u_1}$ and $u_b = \frac{u_4}{u_1}$.

Denote f_{upper} as the upper bound objective function obtained using the derived kth substream SNR inequality. The proof of the concavity of f_{upper} is described in the following. Since the positive weighted sum of any convex (or concave) functions preserves convexity (or concavity), one only needs to check the concavity of kth substream rate upper bound of user 1, $\log_2 \left(1 + \frac{(x+y)z}{t_a(x+y)+t_b z}\right)$ and kth substream rate upper bound of user 2, $\log_2 \left(1 + \frac{(x+y)z}{u_a(x+y)+u_b z}\right)$. This enables one to focus on developing the proof of the concavity of the rate function of user 1. Define function $h(w, z) = \log_2 \left(1 + \frac{wz}{t_a w + t_b z}\right)$ and $w(x, y) = x + y$. From the Hessian matrix, $\mathbf{H}(h) = \begin{bmatrix} \frac{\partial^2 h}{\partial w^2} & \frac{\partial^2 h}{\partial w \partial z} \\ \frac{\partial^2 h}{\partial z \partial w} & \frac{\partial^2 h}{\partial z^2} \end{bmatrix}$, one obtains the following inequalities,

$$\frac{\partial^2 h}{\partial w^2} = -\frac{t_b z^2 (2t_a^2 w + 2t_a wz + 2t_a t_b z + t_b z^2)}{\ln 2 (t_a w + t_b z)^2 (t_a w + t_b z + wz)^2} \leq 0, \tag{7.79}$$

$$\frac{\partial^2 h}{\partial z^2} = -\frac{t_a w^2 (t_a w^2 + 2t_a t_b w + 2t_b wz + 2t_b^2 z)}{\ln 2 (t_a w + t_b z)^2 (t_a w + t_b z + wz)^2} \leq 0, \tag{7.80}$$

while the determinant of the Hessian matrix can be represented by the following inequality,

$$\det(\mathbf{H}) = \frac{2 t_a t_b w^2 z^2}{(\ln 2)^2 (t_a w + t_b z)^2 (t_a w + t_b z + wz)^3} \geq 0. \tag{7.81}$$

since the constants $t_a \geq 0$, $t_b \geq 0$, and the variables $w \geq 0$ and $z \geq 0$. Recall that the determinant of the Hessian matrix corresponds to the product of two eigen values. For concave functions, each eigen value is non-positive. By observing the inequalities from (7.79) to (7.81), matrix $\mathbf{H}$ is proved to be negative semi-definite matrix. This indicates that function $h(w, z)$ is concave with respect to (w, z). Recall that composition with affine function preserves concavity (or convexity) [20]. Since the function $w(x, y)$ is affine, one can conclude that function $h(w(x, y), z) = \log_2\left(1 + \frac{(x+y)z}{t_a(x+y)+t_b z}\right)$ is concave with respect to (x, y, z). Similar steps can be used to prove the concavity of the rate function of user 2, $\log_2\left(1 + \frac{(x+y)z}{u_a(x+y)+u_b z}\right)$. Since concavity is closed under positive summation, it can be concluded that the upper bound objective function f_{upper} is jointly concave with respect to all input parameters $\tilde{a}_1, \ldots, \tilde{a}_M$, $\tilde{b}_1, \ldots, \tilde{b}_M$, $\tilde{c}_1, \ldots, \tilde{c}_M$ and the theorem is proved. ■

Transformation of (5.24) and (5.25) to Equivalent Convex Forms

The power control problem can be transformed into equivalent convex form by performing change of variables and logarithmic transformation on constraint functions in (5.24) and (5.25). Represent $x_k = \log \tilde{a}_k$, $y_k = \log \tilde{b}_k$ and $z_k = \log \tilde{c}_k$. By taking the logarithm of the constraint in (5.24), the posynomial function can be transformed into the following log-sum-exponent function,

$$\begin{aligned} &\log\{\exp(x_k - y_k - z_k + p_{1,k}) + \exp(-z_k + p_{2,k}) \\ &\quad + \exp(-y_k + p_{3,k}) + \exp(-y_k - z_k + p_{4,k})\} \leq 0, \end{aligned} \tag{7.82}$$

where constants $p_{m,k} = \log(\gamma_{1,k} r_{m,k})$, $\forall m = \{1, \ldots, 4\}$. Recall the composition rule: composition with affine function preserves convexity [20], therefore every exponent in (7.82) remains convex. It is shown in [20] that the Hessian of the log-sum-exponent function is positive semi-definite, therefore one can conclude that the constraint in (7.82) is convex. The constraint in (5.25) can be transformed to convex form using similar approach, which is written as

$$\begin{aligned} &\log\{\exp(-z_k + q_{1,k}) + \exp(-x_k + y_k - z_k + q_{2,k}) \\ &\quad + \exp(-x_k + q_{3,k}) + \exp(-x_k - z_k + q_{4,k})\} \leq 0 \end{aligned} \tag{7.83}$$

where $q_{m,k} = \log\left(\gamma_{2,k} s_{m,k}\right), \forall m = \{1, \ldots, 4\}$. The equivalent problem can be solved efficiently using convex optimisation technique, i.e. interior point method. ■

Bibliography

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[2] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[3] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Information Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[4] R. F. Wyrembelski, T. J. Oechtering, and H. Boche, "Decode-and-forward strategies for bidirectional relaying," in *IEEE PIMRC '08*, Sep. 15–18, 2008, pp. 1–6.

[5] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *SIGCOMM '07*, Aug. 27-31, 2007, pp. 397–408.

[6] S. Zhang, S. Liew, and P. Lam, "Hot topic: Physical layer network coding," in *ACM MobiCom '06*, Sep. 23-26, 2006, pp. 358–365.

[7] M. Chen and A. Yener, "Multiuser two-way relaying for interference limited systems," in *IEEE ICC '08*, May 19–23, 2008, pp. 3883–3887.

[8] C. Esli and A. Wittneben, "One- and two-way decode-and-forward relaying for wireless multiuser MIMO networks," in *IEEE GLOBECOM 2008*, Nov. 2008, pp. 1–6.

[9] A. U. T. Amah and A. Klein, "Pair-aware transceive beamforming for non-regenerative multi-user two-way relaying," in *IEEE ICASSP 2010*, Mar. 14-19, 2010, pp. 2506–2509.

[10] E. Yilmaz, R. Zakhour, D. Gesbert, and R. Knopp, "Multi-pair two-way relay channel with multiple antenna relay station," in *IEEE ICC 2010*, May 23-27, 2010, pp. 1–5.

[11] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[12] T. Unger and A. Klein, "On the performance of two-way relaying with multiple-antenna relay stations," in *Proc. 16th IST Mobile and Wireless Communications Summit*, Jul. 1–5, 2007, pp. 1–5.

[13] N. Lee, H. J. Yang, and J. Chun, "Achievable sum-rate maximizing AF relay beamforming scheme in two-way relay channels," in *IEEE ICC Workshops '08*, May 19-23, 2008, pp. 300–305.

[14] K.-J. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Communications,*, vol. 9, no. 12, pp. 3671 –3681, Dec. 2010.

[15] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Trans. Wireless Communications*, vol. 10, no. 5, pp. 1645 – 1655, Mar. 2011.

[16] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Communications*, vol. 6, no. 8, pp. 2798–2802, Aug. 2007.

[17] C. Li, X. Wang, L. Yang, and W.-P. Zhu, "Joint source and relay power allocation scheme for a class of MIMO relay systems," *IEEE Trans. Signal Processing*, vol. 57, no. 12, pp. 4852–4860, Dec. 2009.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York: Cambridge University Press, 2005.

[19] H. Jeon, N. Kim, J. Choi, H. Lee, and J. Ha, "On multiuser secrecy rate in flat fading channel," in *Proc. IEEE MILCOM 2009*, Oct. 18-21, 2009, pp. 1–7.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. i. characterization," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 90–100, 1997.

[22] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge: University Press, Cambridge, 2007.

[23] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, 2001.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey: Wiley and Sons Inc., 2006.

[25] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding. II. coding results," *IEEE Trans. Communications*, vol. 43, no. 10, pp. 2595–2604, Oct. 1995.

[26] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, 1999.

[27] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. New York: Thomson Learning, 1988.

[28] G. FOSCHINI and M. GANS, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[29] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[30] E. C. V. D. Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[31] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Information Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[32] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge: University Press, Cambridge, 2007.

[33] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[34] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Information Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[35] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, Oct. 1949.

[36] A.D.Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[37] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[38] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Information Theory*, vol. 24, no. 4, pp. 451–456, July 1978.

[39] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Information Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[40] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Information Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[41] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. New Jersey: Prentice Hall, 2002.

[42] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[43] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Processing*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[44] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO broadcast channels," *IEEE Trans. Information Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[45] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. Communications*, vol. 55, no. 1, pp. 11–15, Jan. 2007.

[46] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Trans. Information Theory*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[47] J. Sydir and R. Taori, "An evolved cellular system architecture incorporating relay stations," *IEEE Communications Magazine*, vol. 47, no. 6, pp. 115–121, June 2009.

[48] C. Esli and A. Wittneben, "A cluster-based multiuser cooperative network," in *IEEE GLOBECOM '07*, Nov. 26-30, 2007, pp. 3926–3931.

[49] S. Nesargi and R. Prakash, "Distributed wireless channel allocation in networks with mobile base stations," *IEEE Trans. Vehicular Technology*, vol. 51, no. 6, pp. 1407–1421, Nov. 2002.

[50] L. Le and E. Hossain, "Multihop cellular networks: Potential gains, research challenges, and a resource allocation framework," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 66–73, Sep. 2007.

[51] K. Kumar, G. Caire, and A. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Information Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.

[52] D. Gore, J. R.W. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Communications Letters*, vol. 6, no. 11, pp. 491–493, Nov. 2002.

[53] M. Rupp, C. Mecklenbrauker, and G. Gritsch, "High diversity with simple space time block-codes and linear receivers," in *Proc. IEEE GLOBECOM '03*, vol. 1, Dec. 1-5, 2003, pp. 302–306.

[54] N. Balakrishnan and V. B. Nevzorov, *A Primer on Statistical Distributions*. New Jersey: John Wiley & Sons, 2003.

[55] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Signal Processing*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007.

[56] T. J. Oechtering and H. Boche, "Optimal transmit strategies in multi-antenna bidirectional relaying," in *IEEE ICASSP 2007*, vol. 3, Apr. 15–20, 2007, pp. III–145–III–148.

[57] Y.-C. Liang and R. Zhang, "Optimal analogue relaying with multi-antennas for physical layer network coding," in *IEEE ICC '08*, May 19–23, 2008, pp. 3893–3897.

[58] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2239–2251, June 2006.

[59] D. Stirzaker, *Elementary Probability*, 2nd ed. Cambridge University Press, 2003.

[60] C. Li, L. Yang, and W.-P. Zhu, "Two-way MIMO relay precoder design with channel state information," *IEEE Trans. Communications*, vol. 56, no. 12, pp. 3358 –3363, Dec 2010.

[61] C. Y. Leow, Z. Ding, K. K. Leung, and D. L. Goeckel, "On the study of analogue network coding for multi-pair, bidirectional relay channels," *IEEE Trans. Wireless Communications*, vol. 10, no. 2, pp. pp.670–681, Feb. 2011.

[62] N. Lee, H. Park, and J. Chun, "Linear precoder and decoder design for two-way AF MIMO relaying system," in *IEEE VTC 2008*, May 11-14, 2008, pp. pp.1221–1225.

[63] D. Palomar, M. Lagunas, and J. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1179 – 1197, May 2004.

[64] W. Guan, H. Luo, and W. Chen, "Linear relaying scheme for MIMO relay system with QoS requirements," *IEEE Signal Processing Letters*, vol. 15, pp. 697 –700, 2008.

[65] R. Zhang, C. C. Chai, and Y.-C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 726 –737, Feb. 2009.

[66] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Apr. 2011.

[67] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.

[68] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Information Theory*, vol. 54, no. 6, pp. 2453–2469, June 2008.

[69] P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, 1997.

[70] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the k user interference channel," *IEEE Trans. Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[71] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, ser. Wiley Series in Probability and Mathematical Statistics. New Jersey: John Wiley & Sons, 1982.

[72] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed., A. Jeffrey and D. Zwillinger, Eds. New York: Academic Press, 2000.

[73] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: University Press, Cambridge, 1985.

[74] S. Yang and J.-C. Belfiore, "Towards the optimal amplify-and-forward cooperative diversity scheme," *IEEE Trans. Information Theory*, vol. 53, no. 9, pp. 3114–3126, Sep. 2007.

[75] Z. Ding, T. Ratnarajah, and C. Cowan, "On the diversity-multiplexing tradeoff for wireless cooperative multiple access systems," *IEEE Trans. Signal Processing*, vol. 55, no. 9, pp. 4627–4638, Sep. 2007.

[76] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "On the study of network coding with diversity," *IEEE Trans. Wireless Communications*, vol. 8, no. 3, pp. 1247–1259, Mar. 2009.

[77] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. Dover, 1970.

[78] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Wiley Interscience, 2003.

[79] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Maryland: The John Hopkins Univerity Press, 1996.