Relay Selection for MIMO and Massive MIMO Two-Way Relay Networks

by

Jayamuni Mario Shashindra S. Silva

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Communications

Department of Electrical and Computer Engineering

University of Alberta

Abstract

Relay selection strategies help to improve spectral and energy efficiencies, to enhance transmission robustness, or to reduce latency in multi-relay cooperative wireless networks. Two novel relay selection strategies are proposed and analysed here for 1. Multiple-input multiple-output (MIMO) 2. Massive MIMO, amplify-and-forward (AF) two-way relay networks (TWRNs). Specifically, they are designed to minimize the overall outage probability or maximize the achievable sum rate. Interestingly, the first strategy amounts to maximizing the minimum of the eigenvalues of the Wishart matrices of the channels from the selected relay to the two user nodes. Counter-intuitively, the latter strategy amounts to maximizing the minimum of the determinant of the same Wishart matrices. The performance of these two strategies is investigated by deriving lower/upper bounds of the overall outage probability and the average sum rate approximations in closed-form. Further, the asymptotic high SNR approximations of the outage probability are derived, and thereby, the achievable diversity-multiplexing trade-off is quantified.

Our analysis shows that the transmit power at user nodes can be scaled down proportional to their number of antennas and thus the use of a large scale antennas systems results in significant power savings. Further use of large number of antennas makes relay selection process simple and deterministic. Our results reveal that relay selection indeed significantly increases the sum rate and decreases the outage probability of wireless systems and will be useful for the wireless systems.

$\sim$

Preface

This thesis is an original work conducted by Jayamuni Mario Shashindra Srimal Silva.

Chapter 3 of this thesis has been accepted for publication as S. Silva, G. Amarasuriya, C. Tellambura and M. Ardakani, "Relay Selection Strategies for MIMO Two-Way Relay Networks with Spatial Multiplexing," IEEE Trans. Commun. Also this work in part was presented as S. Silva, G. Amarasuriya, C. Tellambura, and M. Ardakani, "Relay selection for MIMO Two-Way relay networks with spatial multiplexing," in IEEE ICC 2015 - Workshop on Cooperative and Cognitive Networks (CoCoNet) (ICC'15 - Workshops 16), London, United Kingdom, Jun. 2015, pp. 943–948. I was responsible for the concept formation, technical apparatus, simulation data collection and manuscript composition with the assistance and guidance of G. Amarasuriya. C. Tellambura and M. Ardakani were the supervisory authors and was involved with concept formation and manuscript composition.

$\sim$

Acknowledgements

I would like to thank Dr. Chintha Tellambura and Dr. Masoud Ardakani for supervising this thesis and other academic work. They gave me the freedom to choose the research area. The study presented in this thesis is supported by TELUS Communications Company and Natural Sciences and Engineering Research Council of Canada (NSERC). Special thanks goes to Dr. Gayan Amarasuriya for helping and working with me during the research. Also I would like to thank my lab mates and friends for their valuable support. Last but not least I would thank my wife and parents for supporting me throughout my academic career.

$\sim$

Contents

List of Figures

List of Acronyms

Acronym	Meaning
4G	Fourth generation.
5G	Fifth generation.
AF	Amplify-and-forward.
BC	Broadcast.
CDF	Cumulative distribution function.
CSI	Channel state information.
DMT	Diversity multiplexing trade-off.
Gbps	Gigabits per second.
IoT	Internet of things.
kbps	kilobits per second.
LOS	Line of sight.
LSAS	Large scale antenna system.
LTE	Long-term evolution.
LTE-A	Long-term evolution-advanced.
M2M	Machine-to-machine.
MAC	Multiple-access.
Mbps	Megabits per second.
MMSE	Minimum mean square error.
mmWave	Millimeter wave.
MTC	Machine type communications.
MU-MIMO	Multi-user MIMO.
OWRN	One-way relay network.

Acronym	**Meaning**
PDF	Probability density function.
RF	Radio frequency.
SINR	Signal-to-interference-plus-noise ratio.
SNR	Signal-to-noise ratio.
TDD	Time division duplexing.
TWRN	Two-way relay network.
VC	Virtual channel.
Wi-Fi	Wireless Fidelity.
WiMAX	Worldwide Interoperability for Microwave Access.
ZF	Zero forcing.

List of Symbols

Elementary & Special Functions

Notation	Definition
$E_1(z)$	Exponential integral function for the positive values of the real part of z [3, Eqn. (8.211)]
$\Gamma(z)$	Gamma function [3, Eqn. (8.310.1)]
$\Gamma(a, z)$	Upper incomplete Gamma function [3, Eqn. (8.350.2)]
$\log_2(\cdot)$	logarithm to base 2

Probability & Statistics

Let X be a random variable [4, Ch. 6], and $\mathcal{A}$, an arbitrary event.

Notation	Definition
$\mathcal{E}_\Lambda\{X\}$	expected value of X over Λ
$f_X(\cdot)$	Probability density function (PDF) of X
$F_X(\cdot)$	Cumulative distribution function (CDF) of X
$\bar{F}_X(\cdot)$	Complimentary CDF of X
$\mathcal{P}[\mathcal{A}]$	probability of $\mathcal{A}$

Matrices

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ denote an $m \times n$ complex matrix.

Notation	Definition/ interpretation
$\mathbf{A}(i,j)$	i^{th} element on the j^{th} column of $\mathbf{A}$
$\|\|\mathbf{A}\|\|_F$	Frobenius norm of $\mathbf{A}$
$\mathbf{A} = \text{diag}\,(a_1, \ldots, a_n)$	$\mathbf{A}$ is rectangular diagonal; a_1 through a_n are the non-zero diagonal elements
$\mathbf{A}^{-1}$	inverse of $\mathbf{A}$ (for $m = n$)
$\mathbf{A}^*$	conjugate of $\mathbf{A}$
$\mathbf{A}^H$	conjugate transpose (Hermitian-transpose) of $\mathbf{A}$
$\mathbf{A}^T$	transpose of $\mathbf{A}$
$\text{diag}\,(\mathbf{A})$	elements on the main diagonal of $\mathbf{A}$
$\text{eig}\,(\mathbf{A})$	eigenvalues of $\mathbf{A}$ (for $m = n$) [5, Ch. 5]
$\mathbf{I}_n$	$n \times n$ identity matrix of rank n
$\mathbf{O}_{m \times n}$	$m \times n$ matrix of all zeros n
$A \otimes B$	Kronecker product of matrices A and B
$\text{trace}(\mathbf{A})$	trace of $\mathbf{A}$ (for $m = n$) [5, p.186]

Miscellaneous

Notation	Definition
$\|a\|$	absolute value of scalar a
a^*	complex conjugate of scalar a
$k!$	factorial of k [6, Eqn. (6.1.5)]
$\binom{n}{k}$	binomial coefficient, n choose k [6, Sec. 3.1]

Chapter 1

Introduction

1.1 Problem Statement

Increasing the spectral efficiency of wireless networks is essential for the future of wireless communication industry in order to support high data rates, more devices and more connectivity for generation wireless systems. Towards this goal, this thesis investigates several emerging wireless technologies.

1.2 Wireless Communication

Mobile signal receptivity has become one of the main factors considered by people when buying a new home in USA according to RootMetrics [7]. According to this survey, importance of mobile connectivity was ranked higher than the availability of schools and proximity to public transportation services. Thus the importance of wireless communication services in people's lives can not be over emphasized.

As stated by Cisco [8], mobile data traffic increased by $69\ \%$ during the year 2014. The number of mobile connected devices in the world is increased by 497 million and surpassed the world population. This trend is expected to continue for the next five years. According to [8], in 2019 the number of mobile connected devices per capita is expected to be 1.5. The average mobile data speed will surpass 2 Mbps by 2016. These forecasts impose stringent requirements on wireless communication technologies, which must thus be developed to meet the above mentioned growth requirements.

Fifth Generation (5G) wireless [9, 10] has therefore been proposed to provide

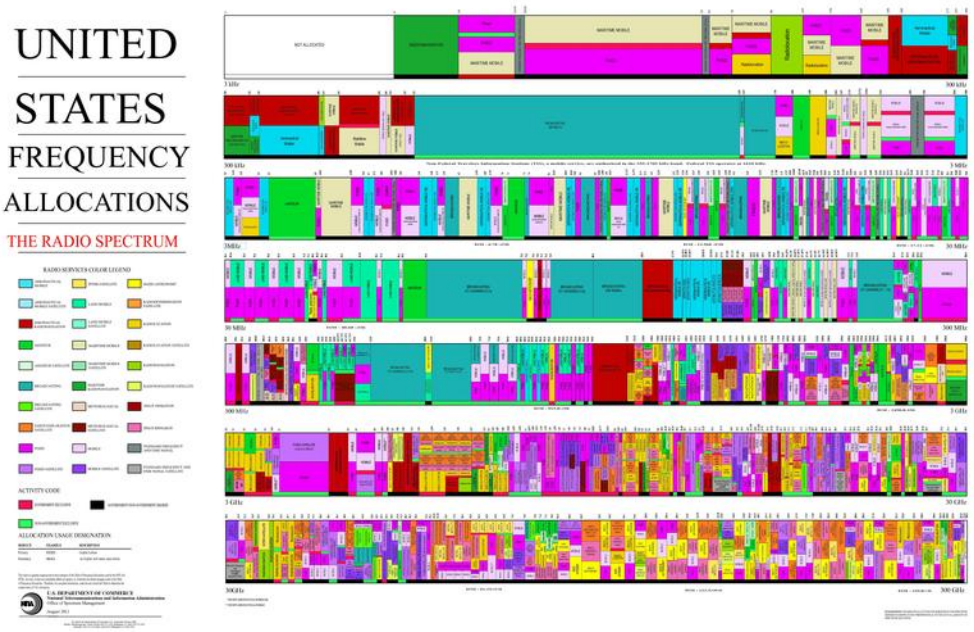


Figure 1.1: Spectrum allocation in USA. [1]

very high data rates reaching Gigabits per seconds, a 1000 fold increase over current systems. Furthermore the latency of 5G networks should be 10 times lower. Conceptual design of 5G systems has just begun and deployment is expected by year 2020.

The growth of 5G and other wireless networks depends on the availability of spectrum. However the current wireless spectrum is limited to 30 - 3000 MHz. The reason is that frequencies above this range suffers from heavy attenuation and frequencies below this range requires very large antenna dimensions. Although many portions of this desirable spectrum are already assigned, some of them are under utilized due to various reasons. For example, Fig. 1.1 shows the current wireless spectrum allocation in USA. Since this figure clearly shows that almost all the usable frequencies are allocated to different applications, to achieve high data rates, the data rate per unit bandwidth must be increased. This measure is called as spectral efficiency. As the total bandwidth available for service providers is limited and controlled by regulatory bodies, the spectral efficiency of the system must improved in order to increase the number of subscribers or the data rate of subscribers. To this end, wireless research community has developed modulation

Figure 1.2: Spectral Efficiency Gains [2]

techniques, multiple antenna systems, cooperative communications and cognitive radios.

The data rate and spectral efficiency of a wireless link are typically a logarithmic function of the signal-to-noise ratio (SNR). Thus doubling the SNR, does not double the data rate and spectral efficiency. Due to this reason, spectral efficiency gains have slowed down in recent years. For example, Fig. 1.2 clearly shows that spectral efficiencies have not increased significantly despite the introduction of many new wireless standards recently. Hence research focused on increasing the spectral efficiency is pivotal for the wireless communication industry.

Another important performance metric of wireless systems is the energy efficiency. This is typically measured as energy required to transmit or receive a unit amount of data. Energy efficiency is extremely important for mobile users as the battery life of handsets are limited. As the high data rates of 5G systems demands 100 times more energy efficiencies than 4G systems, high energy efficient wireless systems must be developed.

Unfortunately, spectral efficiency is lost due to the use of relays which are used

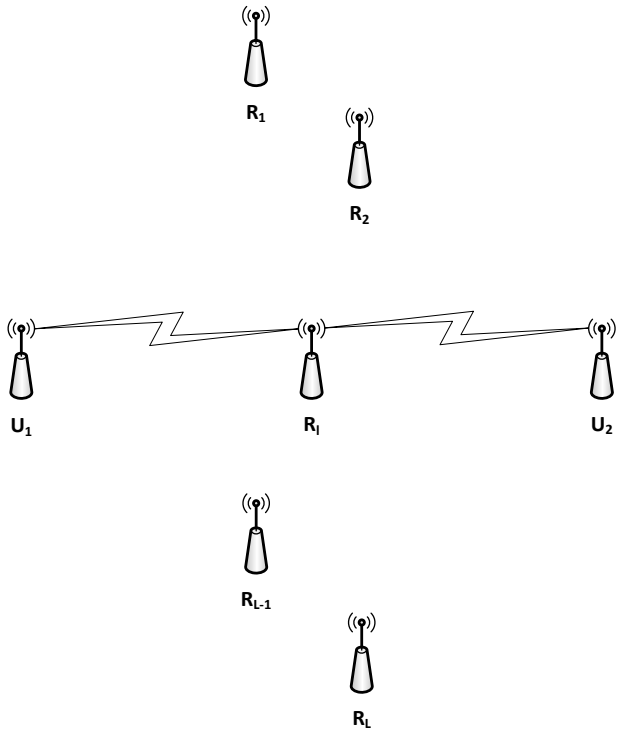


Figure 1.3: Proposed system model. R_l denotes the selected relay.

in wireless communication to increase the connection distance and the reliability of transmission. However one-way relay networks (OWRNs) have the inherent drawback of reduced capacity as it needs four time slots to transmit a message between two end nodes compared to two time slots needed in a direct communication between the users. This will reduce the spectrum efficiency in half. With the use of interference cancellation, two-way relay networks (TWRNs) are able to achieve the same data transfer using just two time slots. This will double the spectral efficiency [11] of the system compared to an OWRN.

1.3 Proposed System

In this thesis a system with two end nodes (U_1 and U_2) and L relay nodes (R_1 to R_L) is investigated (Fig. 1.3. All the nodes are MIMO enabled. The number of antennas at the user nodes should be higher than the number of antennas at each

relay node (Refer Section 3.3.1 for the explanation of this requirement). One of the available relays R_l will be selected for data transmission. For this set-up, two relay selection strategies will be investigated. The first one will minimize the overall outage probability and the second one will maximize the achievable sum-rate. The relay nodes and the end nodes perform amplify-and-forward and beamforming operations respectively.

1.4 Contributions, Significance and Outline

While MIMO TWRNs will improve the spectral efficiency, multiple relays and relay selection yield further enhancements. This thesis presents two relay selection methods to increase the spectral efficiency. The main contributions are listed as follows.

- Two relay selection methods have been proposed based on minimising the overall outage probability and maximising the achievable sum-rate of MIMO AF TWRNs. The first strategy can be simplified to maximizing the minimum of the eigenvalues of the Wishart matrices from the selected relay to the two user nodes. Further, the second strategy can be simplified to maximizing the minimum of the determinant of the same Wishart matrices.

- The performance of the two relay selection methods is characterized via the overall outage probability and the achievable sum-rates. Further, the asymptotic high SNR approximations of the outage probability are derived, and thereby, the achievable diversity-multiplexing trade-off is quantified for the proposed two relay selection schemes. This trade-off reveals that whenever the total number of relay antennas is fixed, the achievable diversity order is always a constant, and hence, the multiplexing gain can indeed be improved by equally distributing antennas among the available set of relays.

- For the second relay selection method the asymptotic performance of massive MIMO TWRNs has been analysed . The achievable asymptotic sum-rates have been derived for three power scaling methods. It is shown that signif-

icant power and spectral efficiencies can be achieved using massive MIMO for TWRN relay selection.

The thesis contributions are significant due to the following reasons.

- Although relay selection has been analysed for MIMO TWRNs with single end-to-end data streams [12, 13], relay selection for MIMO TWRNs with multiple end-to-end spatial data-streams has not been analysed yet. This thesis fills this gap in wireless research.

- Due to the increased spectral efficiency of TWRNs, they are currently being examined for next generation wireless communication standards, including Long-term evolution-advanced (LTE-A) [14, 15]. With the emergence of heterogeneous networks use of relays and relay selection will increase and the research on TWRNs will be important especially with MIMO technologies.

- As explained earlier, massive MIMO has been identified as one of the main enabling technologies for 5G systems [9]. This thesis analyses the use of massive MIMO at the user nodes of a TWRN.

- The recent research developments in millimeter wave (mmWave) wireless communications render MIMO practically viable even at the user nodes because of the extremely short wavelengths associated with the mmWave frequency bands such as 28 GHz and 38 GHz [16]. Thus mmWave frequency bands can be exploited to design MIMO transceivers and other RF elements with much smaller physical dimensions. Thus, employing multiple-antennas even at the user nodes will be practically viable in the next generation wireless standards, and hence, our system models would be useful in practice.

The outline of the thesis is as follows:

- Chapter 2

  This chapter contains the necessary background information. Wireless channel models, cooperative relay systems, 5G technologies, MIMO systems and massive MIMO systems are briefly described.

- Chapter 3

  This chapter proposes and analyses two novel relay selection strategies for MIMO AF TWRNs with spatial multiplexing. Specifically, these strategies minimize the overall outage probability or maximize the achievable sum-rate. Their performance is investigated by deriving lower/upper bounds of the overall outage probability and the average sum rate approximations in closed form. Further, the asymptotic high signal-to-noise ratio (SNR) approximations of the outage probability are derived, and thereby, the achievable diversity-multiplexing trade-off is quantified.

- Chapter 4

  This chapter investigates the asymptotic sum-rate of AF TWRNs with massive MIMO at the user nodes. To this end, three asymptotic SNR expressions are derived for three specific transmit power scaling laws, and thereby, the corresponding asymptotic sum rates are derived. Power scaling is done at the user nodes, relay nodes and both at the user and relay nodes.

- Chapter 5

  This chapter presents the conclusions of the thesis and future research directions. The impact of the thesis results on the wireless communication research is also interpreted.

Chapter 2

Background

This chapter describes wireless communication basics. First, several models for wireless path loss, shadowing and multi-path propagation are described. Second cooperative relay networks, two-way relay networks and relay selection are described. Next, 5G wireless requirements and the enabling technologies are discussed. In particular, the basics of MIMO and massive MIMO are described in detail.

2.1 Wireless Channel Modelling

In wireless channels signal propagation depends on reflection, diffraction, scattering and other phenomena. These propagation effects give rise to various imperfections. Understanding these imperfections are critical for the development and analysis of wireless communication systems. The critical wireless impairments are [17]

1. Path Loss,
2. Shadowing,
3. Multi-path Fading.

In the following sections, these impairments are described in detail.

2.1.1 Simplified Path Loss Model

Path loss is the reduction in power density (attenuation) of a signal as it propagates from the transmitter to the receiver. Apart from the distance between the receiver

and the transmitter, other factors that affect the path loss include the transmitter and the receiver heights, composition of the atmosphere, physical properties of the transceiver antennas and the signal frequency. Thus modelling the path loss is a complex task. Several path loss models including COST 231 [18], Okumura [19] and Hata [20] models are presented in [17]. The simplified path loss model, a highly simplified analytical deterministic channel model, is given as [17]

$$P_R = P_T K \left(\frac{d}{d_0}\right)^{-\eta}, \quad d \geq d_0, \tag{2.1}$$

where P_R, P_T are the received and transmitted powers and d_0 and d are the reference distance for the antenna far-field and the distance between the transmitter and the receiver respectively. This model is not valid for distances $d < d_0$ due to the scattering effects in the antenna near field. Typically d_0 is assumed to be between 1 to 10m in indoor environments and 10 to 100m in outdoor environments. K is a unit less constant which depends on the average channel attenuation and the antenna characteristics. The constant η is known as the path loss exponent and will vary depending on the environment. For example, η is between 3.7 to 6.5 for urban macro-cells and between 2.7 to 3.5 for urban micro-cells [17].

2.1.2 Shadowing

Shadowing is the received signal power fluctuations in wireless channels due to objects obstructing the propagation path between transmitter and receiver. As the location, size, and dielectric properties of the blocking objects are generally unknown, shadowing is often modelled statistically. Thus the log-normal shadowing model describes the probability density function (PDF) of ψ, which is the ratio between the transmit to received power as [17]

$$f_\psi(\psi) = \frac{\xi}{\sqrt{2\pi}\sigma_{\psi_{dB}}\psi} \exp\left[-\frac{\left(10\log_{10}(\psi) - \mu_{\psi_{dB}}\right)^2}{2\sigma^2_{\psi_{dB}}}\right], \quad 0 \leq \psi < \infty \tag{2.2}$$

where $\xi = 10/\ln 10$, $\sigma^2_{\psi_{dB}}$ is the variance of ψ in decibels and $\mu_{\psi_{dB}}$ is the average of ψ in decibels.

Often path loss and shadowing are categorized as large-scale fading, which is the signal attenuation due to movements over large areas compared to the signal wavelength, λ. Similarly small-scale fading is the signal attenuation that occurs over small distances often in the order λ. Small scale fading is often caused by multi-path fading.

2.1.3 Multi-path Fading

Mulit-path fading is the fluctuation of the received power levels at the receiver due to the radio signals reaching the receiving antenna by two or more paths. Due to multi-path propagation, the received signal is a superposition of multiple copies of the signal. Rayleigh model, Rician Model and Nakagami-m model are some widely used multi-path fading models in wireless literature. Each of these methods provide the distribution of the received power β, which depends on the square of the channel amplitude (i.e. $\beta = \frac{|h|^2}{K}$ where h is the channel amplitude and K is a constant).

Rayleigh Fading

Rayleigh fading occurs when there is no line of sight (LOS) signal between the transmitter and the receiver. Absence of LOS path results in the in-phase and the quadrature-phase components of the channel, which are the sum of many random variables. From the central limit theorem, the PDF of such a sum is a normal (Gaussian) distribution, regardless of the exact PDF of the constituent amplitudes. Thus the amplitude of the channel coefficient becomes a Rayleigh distributed random variable and consequently the received power β follows the exponential distribution [21]

$$f_\beta(x) = \frac{1}{\bar{\beta}} e^{-\frac{x}{\bar{\beta}}}, \quad 0 \leq x < \infty, \tag{2.3}$$

where $\bar{\beta}$ is the average signal power.

As mentioned earlier Rayleigh fading assumes the unavailability of a dominant path between the transmitter and the receiver. In wireless communication scenarios such as a mobile user in an urban environment communicating with a base station, the line of sight signal propagation is often blocked due to surrounding buildings.

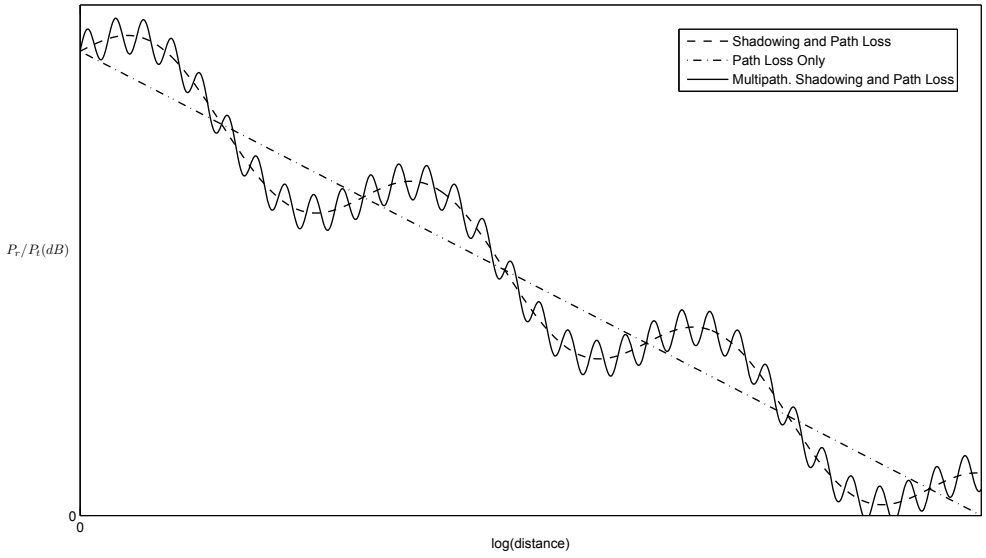


Figure 2.1: Pathloss Shadowing and Multipath

Thus the signal propagation will happen through non line of sight multi-path components. This scenario can thus accurately be approximated using Rayleigh fading. Thus, this fading model has become the standard distribution that is employed in many of wireless research studies.

Rayleigh fading is a special case of more general Nakagami-m fading model which was designed to fit empirical data by Nakagami in early 1960s [22].

Nakagami-m Fading

The PDF of the received power β according to Nakagami-m Fading is given as [22]

$$f_{\beta}(x) = \frac{x^{m-1}}{\Gamma(m)} \left(\frac{x}{\bar{\beta}}\right)^{m} e^{-\frac{mx}{\bar{\beta}}}, \quad 0.5 \leq m < \infty, \quad 0 \leq x < \infty. \tag{2.4}$$

The parameter m describes the level of fading and as in the Rayleigh fading case $\bar{\beta}$ is the average signal power. Changing the m value gives rise to different fading scenarios. When $m = 1$ this model simplifies to Rayleigh fading and when $m = \infty$, this model denotes no fading scenario.

2.1.4 Combined Channel Models

Sections 2.1.1 to 2.1.3 models wireless channel impairments separately. But theses impairments cascade, resulting in a composite channel. Thus a composite model of path loss, shadowing and multi-path fading is described in Section 2.5.3. This composite model versus the distance between the transmitter and the receiver is illustrated in Fig. 2.1. More complex fading models has been proposed in [23].

2.2 Cooperative Relay Technologies

2.2.1 Introduction

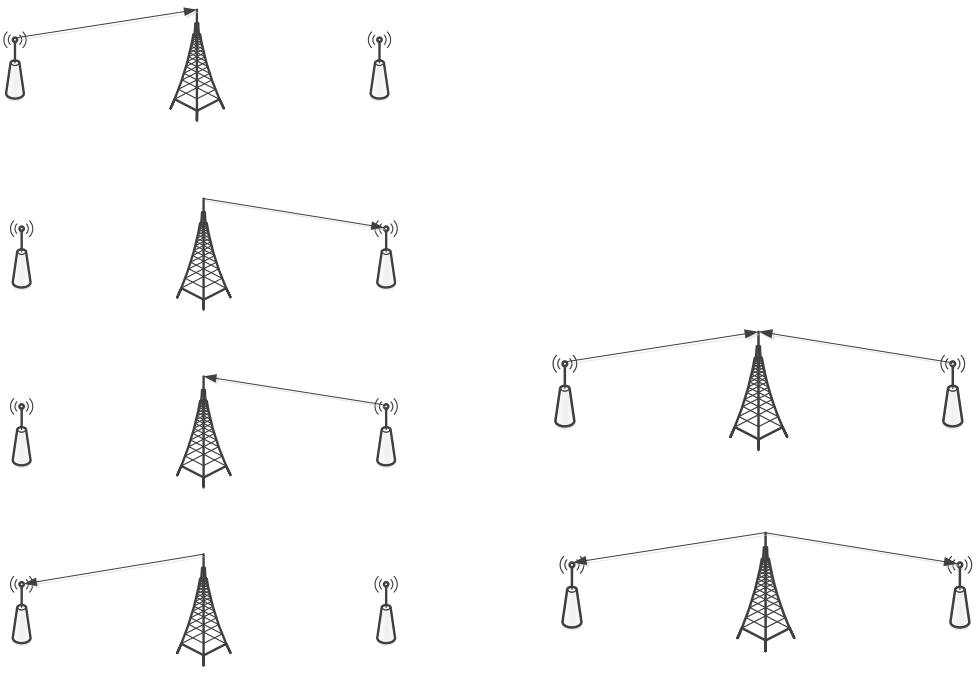
Cooperative relays increase the coverage [24] by acting as intermediate repeaters between the transmitter and receiver. The use of relays results in increase of the overall quality of service of the network. Cooperative relays have already been used in wireless standards such as IEEE 802.16 J [25] and LTE-A [14, 15].

Cooperative relay nodes can be categorized into two types according to the mode of their operation [26, 27]. Amplify and forward (AF) relays, amplify the received signal and forward it to the destination node. The relay does not attempt to decode the received signal. AF relays may amplify the added noise at the relay when forwarding to the destination. This is known as the noise amplification problem.

On the other hand, decode and forward (DF) relays will decode the received data, encode it again and retransmit it to the destination. Thus decoding errors at the DF relay may also be forwarded to the destination. Also the complexity of the DF relay nodes are higher than that of the AF relay due to the added decoding functionality of the relay.

2.2.2 One-way Relay Networks and Two-way Relay Networks

A one-way relay receives from or transmits to a single user node at a given time. But a two-way relay communicates with two user nodes simultaneously. A more general multi-way relay receives from or transmits to multiple user nodes simultaneously.



(a) One-way Relay Networks (b) Two-way Relay Networks

Figure 2.2: Comparison of OWRN and TWRN operation

In a one-way relay network (OWRN) with half-duplex nodes, four time-slots are needed for a single data transfer between two end nodes. This results in loss of spectral efficiency. However two-way relay networks (TWRNs) which uses interference cancellation or physical layer network coding requires only two time slots to transfer data between two end nodes [11, 28]. Thus a TWRN doubles the data rate compared to an OWRN. A comparison of signal transfer steps of an OWRN with a TWRN is shown in the Fig. 2.2.

In the first time slot both the end nodes transmit their respective data to the relay node. In the second time-slot the relay will amplify and forward the received signal to end nodes. End nodes receive their own transmitted signal and the signal transmitted by the other node. As end nodes know their own transmitted signal, using basic signal processing they can easily decode the transmitted message of the other node. Removal of this self interference is known as network coding.

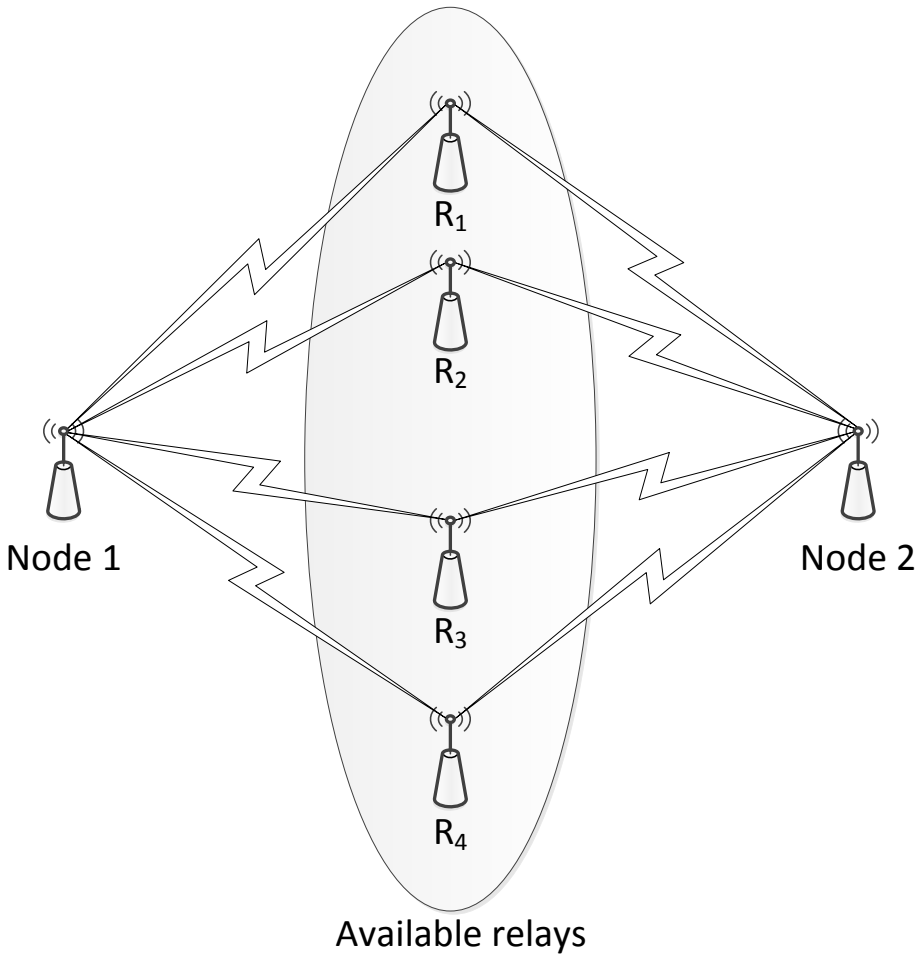


Figure 2.3: System with multiple available relays

2.2.3 Relay Selection

Wireless systems may have several nodes that can be used to relay signals between two nodes. Thus one or more nodes from this pool of nodes can be selected to act as relays between a transmitter (Tx) and a receiver (Rx) [29]. This process is known as relay selection.

This process is illustrated in the Fig. 2.3. The Node 1 and Node 2 can not communicate with each other due to the unavailability of a direct channel between them. However there are four intermediate nodes, R_1 to R_4, that can assist the data transmission between Node 1 and Node 2 by operating as a relay. Choosing the best combination of R_1 to R_4 is the relay selection problem.

The number of relays selected can be one or many. Single relay selection schemes select just one node to act as the relay. These schemes include best relay selection, nearest neighbour selection and best worse channel selection. Best relay selection selects the relay which provides the path with maximum SNR [30]. Nearest neighbour selection selects the nearest relay to the user node [31]. Relay communication involves two channels, one from the source to the relay and the one from the relay to the destination. Best worse channel selection method selects the relay which has the best channel conditions in it's worse channel [32].

Selecting multiple relays will increase the complexity of the system as well as complexity of the relay selection procedure. System complexity arises due to the phase adjustments required at the transmitter to enable the addition of multiple received signals coherently at the receiver. Although non coherent multiple relay selection methods exist [33], their performance is degraded due to non coherent combining. Relay selection complexity is due to the exponential number of possibilities to select multiple relays (If the number of potential relays are L, there are $2^L - 1$ possible ways to select multiple relays). In [34], relay ordering is presented for multiple relay selection. And in [35] several multiple relay selection methods have been analysed for TWRNs.

2.3 5G Wireless Systems

2.3.1 Introduction

Although current Fourth Generation (4G) wireless is designed for mobile broadband, with the forecast of 10 fold growth of mobile data and 100 fold increase of connected devices [8], 4G will soon be inadequate. Thus Fifth Generation (5G) wireless [9] has been proposed to provide very high data rates reaching Gigabits per seconds, a 1000 fold increase over 4G systems. 5G systems are expected to be implemented by year 2020 [9].

While initial 1G services were mainly focussed on voice services, 2G systems were evolved around improved voice and text messaging. The main driver for 3G was the integrated voice and affordable mobile internet. 4G was focussed on high

capacity mobile multimedia [36]. Yet 5G systems will support a wide range of applications. These which include but not limited to

- Mobile Broadband,
- Smart cities and smart homes,
- Smart grids,
- Health monitoring systems,
- Augmented/Virtual reality,
- Industrial monitoring systems,
- Automotive safety systems,
- Machine type communications (MTC).

Thus it is clear that mobile broadband is not the only main driving component of 5G systems. This shift in the driver represents a major paradigm shift in wireless communications. Instead of few applications, new wireless standards are expected to cater for a vast range of applications.

It should be noted that for the first time in wireless communication standards there may be no dedicated voice services in 5G. Instead voice services are expected to be provided as an application running on data connectivity [36].

2.3.2 Requirements of 5G Systems

The diverse requirements of above mentioned use cases include high throughput, low latency and the support for a large number of devices. High throughput is required by applications like high definition videos and augmented reality applications. The peak data rates for 5G systems are expected to be in the range of 10Gbps. And 95% of users should be able to get 100Mbps data rate. 10Mbps data rate is expected in everywhere including the rural areas [37].

5G systems should offer latencies less than 1 ms. These low latencies are essential for time critical applications such as automotive safety systems and also

required by machine type communications (MTC) where different devices communicate with each other without any human interactions.

5G systems should also support 10-100 times more devices than current systems. With the internet of things (IoT) where almost everything in the environment becomes a wireless device to coordinate with other objects, the number of connected devices will increase drastically. These devices may range from high bandwidth consuming devices with rechargeable batteries such as mobile phones and tablet computers to very low bandwidth consuming, battery life limited, sleep-enabled sensor nodes. Providing connectivity for all these devices will be a challenge for 5G system designers and implementers.

2.3.3 Proposed Technologies for 5G Systems

Several new and old technologies for 5G systems have already been identified and analysed [9, 10, 38]. Also several white papers have been published by wireless companies Huawei, Nokia and Ericsson [36,37,39]. Following technologies for 5G has been identified by [10]

1. Device-centric architectures,
2. Millimeter wave,
3. Massive MIMO,
4. Smarter devices,
5. Native support for machine-to-machine (M2M) communications.

According to [37] implementation of 5G systems include the use of novel radio access technologies in new spectrum regions together with the backward compatible LTE evolution systems operating in existing spectrum [37]. 5G systems are expected to operate in the mmWave region and will utilize massive MIMO. Massive MIMO systems are described in the Section 2.5.

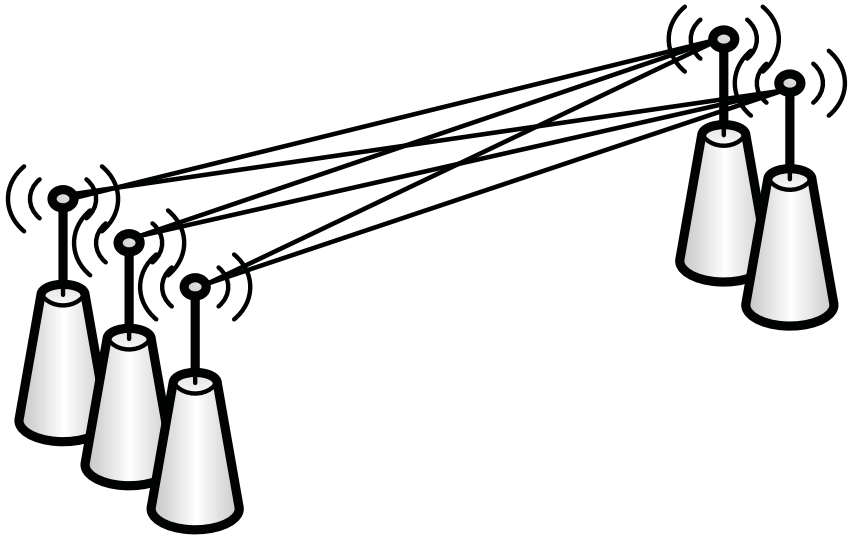


Figure 2.4: MIMO transmitters and receivers

2.4 MIMO Systems

Multiple-input multiple-output (MIMO) technology refers to the use of multiple antennas at the receiver and the transmitter. This usage results in increased spectral efficiency and/or reliability due to the increased spatial diversity. Due to these advantages, MIMO has been incorporated in wireless standards such as IEEE 802.11n (Wi-Fi), IEEE 802.11ac (Wi-Fi), HSPA+ (3G), WiMAX (4G) and Long Term Evolution (4G).

As shown in Fig 2.4, instead of a single path between the transmitter and receiver, a 2×3 MIMO system will have six paths between the transmitter and the receiver. Thus the channel between the transmitter and the receiver is represented using a 2×3 matrix of complex numbers. For a more general example with N_t transmit antennas and N_r receive antennas the channel is given as a matrix $\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$. Each (i, j) element of the matrix $\mathbf{H}$, $h_{i,j}$ represents the channel coefficient between the j^{th} transmitter antenna and the i^{th} receiver antenna for

$1 \leq j \leq N_t$ and $1 \leq i \leq N_r$. Thus the MIMO received signal may be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.5}$$

where $\mathbf{y}$ is the $N_r \times 1$ received signal vector, $\mathbf{x}$ is the $N_t \times 1$ transmit vector and $\mathbf{n}$ is the $N_r \times 1$ noise vector. Having a signal vector instead of a single symbol as the transmitting data opens up new possibilities to achieve higher performance in wireless systems. These include the following scenarios:

- Space diversity

  The transmitted data vector $\mathbf{x}$ can be selected such that the same data is repeated through the channel to the receiver. When the propagation paths of these signal streams are independent, they will undergo different amount of fading. Thus the chance of atleast one of these streams having the required signal strength is high. This repetition will increase the reliability of the system and hence the error rates will be reduced.

- Spatial multiplexing

  Instead of repeating the same data through the channel, vector $\mathbf{x}$ can be selected such that several different data symbols are transmitted to the destination. Thus the amount of information carried between the transmitter and the receiver in a MIMO system will be increased compared to a normal single-input single-output system. Thus the use of MIMO with spatial multiplexing will increase the data rate of the system.

- Reduction of interference

  MIMO can be used to support multiple users simultaneously by exploiting the spatial diversity of MIMO channel to reduce the interferences among different users. This results in multi-user MIMO systems which are described in Section 2.4.5.

Thus MIMO systems may be designed with different objectives and thus operate in different modes. For instance, seven modes of MIMO for downlink have been

defined in LTE [40]. Diversity gain/order, multiplexing gain and diversity multiplexing trade-off are some performance matrices used to compare different MIMO modes.

2.4.1 Diversity Gain/Order

The diversity gain (or diversity order) d quantifies the number of independently faded replicas of a transmitted symbol the destination receives and can be defined as follows [41]:

$$d = -\lim_{\gamma \to \infty} \frac{\log\left(P_{\text{out}}(\mathcal{R}, \gamma)\right)}{\log\left(\gamma\right)}, \tag{2.6}$$

where $P_{\text{out}}(\mathcal{R}, \gamma)$ is the information rate outage probability evaluated at SNR γ while keeping the threshold rate at $\mathcal{R}$. Diversity gain can be used to derive the asymptotic performance analysis [42]. Further improvements in asymptotic performance analysis are done in [43, 44].

2.4.2 Multiplexing Gain

The multiplexing gain r can be defined as follows [41]:

$$r = \lim_{\gamma \to \infty} \frac{\mathcal{R}(\gamma)}{\log\left(\gamma\right)}, \tag{2.7}$$

where $\mathcal{R}(\gamma)$ is the achievable data rate evaluated at the SNR γ.

2.4.3 Diversity Multiplexing Trade-off (DMT)

The seminal work of Zheng and Tse [41] showed that both diversity and multiplexing gains can be simultaneously obtained for a given multiple-antenna channel subjected to a fundamental trade-off. Thus, the multiplexing gain comes at the expense of sacrificing the diversity gain and vice versa. By using (2.6) and (2.7), the DMT can be defined as [41]

$$d(r) = -\lim_{\gamma \to \infty} \frac{\log\left(P_{\text{out}}(r\log(\gamma))\right)}{\log\left(\gamma\right)}, \tag{2.8}$$

where diversity order $d(r)$ is given as a function of the multiplexing gain r.

2.4.4 MIMO Beamforming

MIMO Beamforming is the process of separating a MIMO channel in to multiple virtual channels (VC) in the space dimension by using signal processing at the transmitter and the receiver. Beamforming typically use linear signal processing at the transceivers and heavily rely on the availability of the channel state information. Each virtual channel can be considered as a single-input single-output channel and thus modulation, coding and resource allocation schemes used for single-antenna systems can directly be used on these virtual channels.

Beamforming utilizes transmit and receiver beamforming matrices $\mathbf{W}_t$ and $\mathbf{W}_r$, so that the product $\mathbf{W}_r \times \mathbf{H} \times \mathbf{W}_t$ has the desired diagonal structure. There are several popular linear beamforming methods such as zero-forcing (ZF), maximum-ratio combining/receiving and minimum mean square error (MMSE).

In this thesis, the precoders and detectors at the user nodes are designed by using simple transmit-ZF and receive-ZF. In particular, the proposed ZF-based transmission design is significantly less complicated and more computationally efficient than the optimal precoder/detectors in [45]. Further, ZF precoders and detectors can provide a significant fraction of the performance gains achieved by the complicated optimal counterparts whenever the constraint $\min(N_1, N_2) \geq \max_{l \in \{1,\cdots,L\}} N_{R_l}$ is satisfied. Thus, our system design preserves the two of the most important trade-offs of deploying cooperative relay networks namely (i) performance versus implementation cost and (ii) performance versus computational complexity.

2.4.5 Multi-user MIMO Systems

While MIMO was initially developed for point-to-point links, in the last few years the focus shifted to multi-user MIMO (MU-MIMO) systems [46]. MU-MIMO systems contain base stations (BSs) having multiple antennas to communicate with multiple users. Users often have single antennas. The impact of this was the low complexity at user level hardware and more propagation environment tolerant systems.

2.5 Massive MIMO

2.5.1 Introduction

Massive MIMO systems, also known as Large Scale Antenna Systems (LSAS) were proposed by Marzetta [47] as an extension of MU-MIMO. These systems are expected to play a vital role in upcoming 5G systems [9]. In massive MIMO systems the number of antennas at a base station is theoretically increased to infinity. For practical purposes, this number can be in the range of 100-1000. In this section the basics of the massive MIMO systems, the channel models and the technical challenges are presented.

2.5.2 System Model for Massive MIMO

A massive MIMO system consists of K users with single antennas interacting with a base station with M number of antennas where $M \ggg K$. For instance K will be in the range of 10 to 100 while M will be in the range 100 to 1000. For massive MIMO analysis the performance of the system will be analysed as M goes to infinity. Normally there are two scenarios.

- The number of users K in the system is kept fixed as the number of antennas at the base station M is increased.
- The number of users K in the system is also increased as the number of antennas at the base station are increased such that $\frac{M}{K} = \alpha$.

This thesis considers the first scenario where the number of users are kept constant. However previous research has shown that the main results obtained for the first case are still valid for the second scenario as well [48]. Yet the number of antennas needed to achieve near asymptotic performance may be higher in the second approach.

2.5.3 Channel Model for Massive MIMO

The channel model proposed for massive MIMO is just an extension of MIMO channels with added values to capture the path-loss, shadowing and multi-path fad-

ing. The channel matrix $\mathbf{G}$ have $M \times K$ entries. This matrix $\mathbf{G}$ will model independent small-scale fading, geometric attenuation and log-normal shadow fading. The element g_{mk} which is the channel coefficient between the mth antenna of the BS and the kth user is written as

$$g_{mk} = h_{mk}\sqrt{\beta_k}, \tag{2.9}$$

for $m \in 1, 2, \cdots, M$. Here h_{mk} models the multi-path fading and $\sqrt{\beta_k}$ models the path loss and shadow fading. The path loss and shadow fading coefficient $\sqrt{\beta_k}$ is assumed to be same from any base station antenna to each user. This channel can be represented in matrix form as

$$\mathbf{G} = \mathbf{H}\mathbf{D}^{\frac{1}{2}}, \tag{2.10}$$

where $\mathbf{D}$ is the $K \times K$ diagonal matrix with $|\mathbf{D}|_{kk} = \beta_k$ and $\mathbf{H}$ is the matrix of multi-path fading coefficients with $|\mathbf{H}|_{mk} = h_{mk}$.

The performance of the massive MIMO is based on the law of large numbers. The massive MIMO systems are based on the following result presented in [47].

$$\frac{\mathbf{G}\mathbf{G}^H}{M} \xrightarrow{a.s.} \mathbf{D}, \quad \text{as} \quad M \to \infty, \tag{2.11}$$

for $M \ggg K$. The (a.s.) stands for the following definition.

Definition 2.1 Random variable X_n converges almost surely (a.s.) to X as $n \to \infty$, if $\Pr\{X_n \to X\} = 1$ as $n \to \infty$. •

This situation is known as the favourable propagation and its consequences for the wireless communication systems are explained in [49]. As the number of antennas at the base station M goes to infinity, the columns of matrix $\mathbf{G}$ becomes independent of each other. Thus the paths from a user to each antenna will be independent and the full use of diversity can be achieved for each user.

2.5.4 General Results

The use of linear precoding with a large number of antennas promises many interesting results for wireless research. Most of these results are very favourable for high performance and data rates. These include

- Mitigation of noise and small-scale fading at the receiver,

  Due to the large number of antennas at the base station, the noise and the small-scale fading effects of the channel will be averaged at the base station. Thus the system designers only need to worry about the large scale fading and shadowing when designing massive MIMO systems. Also mitigation of fading effects results in reduced latency of the system [47].

- Simple multiple access layer,

  Due to the law of large numbers, each sub carrier in the massive MIMO systems will have the same channel gain. Thus each terminal will be given the full bandwidth and physical layer control signals are not required [49].

- Robustness to interferences,

  It is shown that the massive MIMO systems are robust against the interferences caused by other equipments. Thus co-channel and inter-cell interferences are not present in massive MIMO systems [47].

Massive MIMO offers very high energy efficiencies. It is shown that the system power can be reduced inversely proportional to the number of antennas at the base station M in massive MIMO systems when channel state information (CSI) is available. Even when systems only have imperfect CSI, the power can be reduced inversely proportional to the square root of the number antennas at the base station $\sqrt{M}$.

2.5.5 Massive MIMO Research Problems

Several massive MIMO research problems are mentioned.

1. Pilot Contamination

   Pilot contamination occurs due to the interferences among the users who use the same pilot sequence during the channel estimation phase. This contamination occurs due to the limitation of the number of orthogonal pilot sequences available to a system [48]. Moreover, increasing the number of antennas at the base station does not reduce this problem. The solutions include pilot scheduling algorithms and intelligent pilot sequence allocation algorithms [50, 51].

2. Channel Reciprocity

   Massive MIMO mainly operates with time division duplexing which requires uplink and downlink channel reciprocity. But due to the hardware constraints, this assumption may not be true. More research is required to solve this reciprocity mismatches in massive MIMO uplink and downlink.

3. Hardware Impairments

   Using a very large number of antennas at the base station introduces new hardware related problems. These includes high power consumption of signal processing hardware. Due to this the total power consumption at the base station will be increased significantly. New research needs to be done to overcome these impairments.

4. Deployment Scenarios

   Massive MIMO systems should be compatible with current wireless systems. One scenario suggested by [37], is the use of massive MIMO in high frequency mmWave range while the current LTE-A systems works in the current wireless spectrum, allowing backward compatibility for 4G devices. It suggests that after some time, massive MIMO can fall back into current wireless spectrum, when the support for 4G is not required. Also use of massive MIMO for backhaul links has also been considered as a possible application.

5. Antenna Arrangement and Placements

Other main research area of massive MIMO systems will be the placement of the antennas at the base station. Several antenna structures have already been analysed for performance. These include circular structures and 3D antenna arrangements. Also the practical number of antennas needed to achieve massive MIMO performance has been another research problem [52].

2.6 Conclusion

This chapter reviewed related topics, including wireless channels, cooperative relay systems, 5G technologies, MIMO systems and massive MIMO systems.

Chapter 3

Relay Selection Strategies for MIMO Two-Way Relay Networks with Spatial Multiplexing

This chapter presents and analyses relay selection strategies for MIMO TWRNs. The obtained results are verified with Monte-Carlo simulations. Further the insights provided by the obtained results are discussed.

3.1 Introduction

The advantages offered by MIMO, TWRNs and relay selection were described in the previous chapter. While this thesis develops two relay selection strategies for MIMO TWRNs with zero-forcing (ZF) transmit and receive filtering, several previous work have developed relay selection strategies for TWRNs employing MIMO-enabled user and relay nodes [12,35,53,54]. Thus this introduction section presents the related previous research and establishes the uniqueness and the significance of this thesis. The motivation of the research is also presented.

Prior related research on MIMO TWRNs: In [12] joint relay and antenna selection strategies over Rayleigh fading are investigated, with closed-form exact and high SNR approximations of outage probability. Thus, achievable diversity order is quantified. Further, [12] analyzes and quantifies the performance degradation due to feedback-delay effect and spatially correlated fading. Reference [13] develops a joint beamforming and relay selection strategy to maximize the end-to-end

signal-to-noise ratio (SNR) and hence to minimize the overall outage probability. In particular, in [13], the overall outage probability is derived in closed-form, and thereby, the achievable diversity order is quantified. Note that the relay selection strategies in [12] and [13] have been developed for a single end-to-end spatial data-stream. Thus, relay selection for MIMO TWRNs with multiple end-to-end spatial data-streams (i.e., with spatial multiplexing) has not yet been investigated.

For the sake of completeness, several important prior studies on relay selection for single-antenna TWRNs are next summarized. In [35], single and multiple relay selection schemes are developed and analyzed. Single-relay selection is based on maximizing the worst SNR of the two user nodes. Further, in the multiple relay selection scheme of [35], a subset of available relays are selected by using the concept of relay ordering proposed in [34]. Here, the available relays are first ordered in the ascending order of the end-to-end SNR, and then, the best subset of relays, which maximizes the worst SNR of the two user nodes, are selected. Further, in [55], an optimal relay selection scheme is developed with full-duplex nodes based on maximizing the effective signal-to-interference-plus-noise ratio (SINR). All aforementioned studies [35, 55] treat AF relays. But [56–58] consider decode-and-forward (DF) relays. Moreover, in [54, 59], the relay selection schemes are studied for single-antenna TWRNs with physical-layer network coding.

Motivation and our contribution: In the extensive wireless relay literature, only two studies develop relay selection strategies for MIMO TWRNs. Moreover, these strategies in [12] and [13] are limited to a single end-to-end spatial data stream only. Thus, they are more suited for practical scenarios where the corresponding wireless channels are ill-conditioned (due to lack of rich scattering) or if the transmission reliability via the diversity gains is preferred over the data rates via spatial multiplexing gains. In this context, no prior results exist on designing and analyzing of relay selection strategies for MIMO TWRNs with spatial multiplexing (i.e., with more than one end-to-end spatial data-streams). This observation thus motivates our work and fills the aforementioned gap in relay literature by proposing and analysing two novel relay selection strategies for MIMO TWRNs with spatial multiplexing. Our strategies are motivated by the practical need for boosting the throughput of

MIMO TWRNs operating in rich-scattering wireless channels, and hence prioritize the use of available degrees of freedom for spatial multiplexing. The proposed relay selection strategies and our main contributions can be next summarized as follows:

Specifically, our first relay selection strategy maximizes the SNR of the worst data substream at the weakest user node and consequently minimizing the overall outage probability. Interestingly, it amounts to selecting the relay which maximizes the minimum of the eigenvalues of the Wishart matrices of the selected relay to the two user nodes. Moreover, our second strategy maximizes the achievable sum rate. Unlike the first strategy, this one selects the relay which maximizes the minimum of the determinants of the same Wishart matrices. Intuitively, the first strategy is designed to maximize the overall diversity order by alleviating the reduction of the available degrees of freedom due to spatial multiplexing, otherwise majority of the achievable degrees of freedom would be utilized for boosting the diversity order as shown in [12, 13]. Counter-intuitively, the second one maximizes the overall spatial multiplexing gain by maximizing the effective determinant of the corresponding Wishart matrices, and thereby paving the way to transmit multiple data-streams through well-conditioned spatial sub-channels.

The performance of the two proposed relay selection strategies is investigated by deriving the lower/upper bounds of the overall outage probability and the average sum rate approximations in closed-form. Further, the asymptotic high SNR outage probability approximations are then used for deriving the achievable diversity-multiplexing trade-off. Thereby, the maximum overall diversity order and achievable spatial multiplexing gain are quantified to obtain valuable insights into practical implementation of relay selection strategies for MIMO TWRNs. Moreover, rigorous simulation results are presented to investigate/compare the performance of the proposed selection strategies and as well to verify our analysis.

Notation: $\mathbf{Z}^H$, $[\mathbf{Z}]_k$, and $\lambda_k\{\mathbf{Z}\}$ denote the Hermitian-transpose, the $k\text{th}$ diagonal element, and the kth eigenvalue of the matrix, $\mathbf{Z}$, respectively. $\mathcal{E}_\Lambda\{z\}$ is the expected value of z over Λ, and the operator $\otimes$ denotes the Kronecker product. $\mathbf{I}_M$ and $\mathbf{O}_{M\times N}$ are the $M \times M$ Identity matrix and $M \times N$ matrix of all zeros, respectively. $f(x) = o\left(g(x)\right)$, $g(x) > 0$ states that $f(x)/g(x) \to 0$ as $x \to 0$. Further, $E_1(z)$ is

the exponential integral function for the positive values of the real part of z [3, Eqn. (8.211)]. $\Gamma(z)$ is the Gamma function [3, Eqn. (8.310.1)], and $\Gamma(a, z)$ is the upper incomplete Gamma function [3, Eqn. (8.350.2)].

3.2 Application Scenarios

This section identifies application scenarios of the proposed system. The system model is briefly presented and two specific application scenarios for the system are identified.

It should be noted that the relay nodes in the proposed system has a lower complexity compared to end nodes due to the following reasons.

- The relays only perform amplify-and-forward operation. Thus the signal processing complexity at the relays will be considerably lower.
- The number of antennas at the relay nodes are less than those at the nodes U_1 and U_2. Thus the hardware complexity of the relay antenna structure will be lower than the end nodes.

The application scenarios are as follows.

1. Relaying between two user nodes

   In this case U_1 and U_2 will be user nodes and R_l can be a base station. The ability to employ multiple-antennas at the user nodes has been one of the limitations that prevents ubiquitous usage of MIMO technology in the current wireless systems. However, the recent research developments in millimeter wave (mmWave) wireless communications render MIMO practically viable even at the user nodes because of the extremely short wavelengths associated with the mmWave frequency bands such as 28 GHz and 38 GHz [16]. Consequently, mmWave frequency bands can be exploited to design MIMO transceivers and other RF elements with much smaller physical dimensions [60].

2. Backhaul links

Backhaul links are used to interconnect base stations with each other and with the base station controllers [61]. Often, these links carry a large amount of data and the reliability is very important. Currently most backhaul links are either optical fibre or point-to-point microwave links. Replacing wired backhaul links with wireless links has been suggested in [62, 63].

Thus the proposed system can be applied to a backhaul link. In this scenario, the nodes U_1 and U_2 will be base stations. Thus the assumption of multiple antennas at U_1 and U_2 is justified. Relays R_1 to R_L can be used to minimize the outage probability or to maximize the achievable sum-rate. The base stations can use either relay selection strategy depending on the requirement of the system. When the requirements of the system changes, more relay nodes can be deployed for increased performance.

3.3 MIMO ZF TWRN systems

This section presents the system, channel, and signal models. Further, the exact end-to-end SNRs at the user nodes are derived, leading to the lower and upper bounds of the SNRs.

3.3.1 System and channel model

The system model consists of two user nodes (U_1 and U_2) and L relay nodes (R_l for $l \in \{1, \cdots, L\}$). User node U_i is equipped with N_i antennas for $i \in \{1, 2\}$, and the lth relay node has N_{R_l} antennas. All nodes are assumed half-duplex terminals, and all channel amplitudes are assumed independently distributed, frequency flat-Rayleigh fading. Thus, the channel matrix from U_i to R_l can be defined as $\mathbf{F}_{i,l} \sim \mathcal{CN}_{N_{R_l} \times N_i} \left(\mathbf{0}_{N_{R_l} \times N_i}, \mathbf{I}_{N_{R_l}} \otimes \mathbf{I}_{N_i} \right)$, where $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$. The channel coefficients are assumed to be fixed during two consecutive time-slots, and hence, the channel matrix from R_l to U_i can be written as $\mathbf{F}_{l,i} = \mathbf{F}_{i,l}^T$ by using the reciprocity property of wireless channels. The additive noise at all the receivers is modelled as complex zero mean additive white Gaussian (AWGN) noise. The direct channel between U_1 and U_2 is assumed to be unavailable due to large pathloss and

heavy shadowing effects [11, 53, 64].

One complete transmission cycle of our two-way relaying model consists of two phases, namely (i) multiple-access (MAC) phase and (ii) broadcast (BC) phase. In the MAC phase, two user nodes utilize transmit-ZF precoding for transmitting their signals simultaneously towards the selected relay. Then, the relay receives a superimposed signal which is a function of both signals belonging to two users. During the BC phase, the selected relay transmits an amplified version of its received signal back to two users. Then, each user node employs its corresponding receive-ZF detector to receive the signals belonging to the other user node by using self-interference cancellation [11].

It is worth noting that for employing transmit-ZF, the number of antennas at the transmitter should be equal for higher than that of the receiver [65]. Similarly, in order to use receive-ZF, the antenna count at the receiver should be equal or higher than that of the transmitter [65]. To fulfil the aforementioned requirement, the antenna configurations at the user nodes and relays should satisfy the following antenna constraints:

$$N_1 \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}, \tag{3.1}$$

$$N_2 \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}, \tag{3.2}$$

$$\min(N_1, N_2) \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}. \tag{3.3}$$

3.3.2 Signal model

During two time-slots, U_1 and U_2 exchange their signal vectors ($\mathbf{x}_1$ and $\mathbf{x}_2$) by selecting one of the available L relays. Here, the selected relay is denoted as R_l for the sake of the exposition. In the first time-slot, U_1 and U_2 transmit $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively, towards R_l by employing transmit-ZF precoding over the multiple access channel. The received signal at R_l can then be written as

$$\mathbf{y}_{R_l} = g_{1,l}\mathbf{F}_{1,l}\mathbf{W}_{T_{1,l}}\mathbf{x}_1 + g_{2,l}\mathbf{F}_{2,l}\mathbf{W}_{T_{2,l}}\mathbf{x}_2 + \mathbf{n}_{R_l} \text{ for } l \in \{1, \cdots, L\}, \tag{3.4}$$

where the $N_{R_l} \times 1$ signal vector $\mathbf{x}_i$ satisfies $\mathcal{E}\left[\mathbf{x}_i\mathbf{x}_i^H\right] = \mathbf{I}_{N_{R_l}}$ for $i \in \{1, 2\}$. Thus, the $N_i \times 1$ precoded-transmit signal vector at U_i is given by $\mathbf{W}_{T_{i,l}}\mathbf{x}_i$. In (3.4), $\mathbf{n}_{R_l}$

is the $N_{R_l} \times 1$ zero mean AWGN vector at R_l satisfying $\mathcal{E}\left(\mathbf{n}_{R_l}\mathbf{n}_{R_l}^H\right) = \mathbf{I}_{N_{R_l}}\sigma_{R_l}^2$. In particular, in (3.4), $\mathbf{W}_{T_{i,l}}$ is the transmit-ZF precoder at U_i and can be defined as [66]

$$\mathbf{W}_{T_{i,l}} = \mathbf{F}_{i,l}^H \left(\mathbf{F}_{i,l}\mathbf{F}_{i,l}^H\right)^{-1} \quad \text{for} \;\; i \in \{1,2\} \;\; \text{and} \;\; l \in \{1,\cdots,L\}. \tag{3.5}$$

Moreover, in (3.4), $g_{i,l}$ is the power normalizing factor at U_i and is designed to constrain its long-term transmit power as follows [53]:

$$g_{i,l} = \sqrt{\frac{\mathcal{P}_i}{\mathcal{E}\left[\mathrm{Tr}\left(\mathbf{W}_{T_{i,l}}\mathbf{W}_{T_{i,l}}^H\right)\right]}} = \sqrt{\frac{\mathcal{P}_i}{\mathcal{E}\left[\mathrm{Tr}\left(\left(\mathbf{F}_{i,l}\mathbf{F}_{i,l}^H\right)^{-1}\right)\right]}}, \tag{3.6}$$

for $i \in \{1,2\}$ and $l \in \{1,\cdots,L\}$ where $\mathcal{P}_i$ is the transmit power at U_i. Next, the expected value of the trace of an inverse Wishart matrix can be evaluated as [67]

$$\mathrm{Tr}\left(\mathcal{E}\left[\left(\mathbf{F}_{i,l}\mathbf{F}_{i,l}^H\right)^{-1}\right]\right) = \frac{N_{R,l}}{N_i - N_{R,l}}. \tag{3.7}$$

Next, by substituting (3.7) into (3.6), the power normalizing factor at U_i can be derived as

$$g_{i,l} = \sqrt{\frac{\mathcal{P}_i(N_i - N_{R_l})}{N_{R_l}}}, \;\; \text{for} \;\; i \in \{1,2\} \;\; \text{and} \;\; l \in \{1,\cdots,L\}. \tag{3.8}$$

In order to evaluate (3.7) finitely and hence to obtain a finite non-zero value for $g_{i,l}$ (3.8) at each user node, the constraint $N_i \neq N_{R_l}$ needs to be satisfied. Thus, in order to employ transmit-ZF/receive-ZF at U_1 and U_2 during two transmission phases, and for satisfying the long-term transmit power constraint at R_l, the antenna configurations at the relay and user nodes should satisfy the condition $min(N_1, N_2) > \max_{l\in\{1,\cdots,L\}} N_{R_l}$.

In the second time-slot, R_l amplifies $\mathbf{y}_{R_l}$ and broadcasts this amplified-signal towards both user nodes. Each user node then receives its signal vector by using the corresponding receive-ZF detector as follows:

$$\mathbf{y}_{U_{i,l}} = \mathbf{W}_{R_{i,l}}\left(G_l\mathbf{F}_{l,i}\mathbf{y}_{R_l} + \mathbf{n}_i\right), \;\; \text{for} \;\; i \in \{1,2\} \;\; \text{and} \;\; l \in \{1,\cdots,L\}, \tag{3.9}$$

where G_l is the amplification factor at G_l and can be defined as

$$G_l = \sqrt{\frac{\mathcal{P}_{R_l}}{g_{1,l}^2 + g_{2,l}^2 + \sigma_{R_l}^2}}, \tag{3.10}$$

where $g_{l,i}$ is defined in (3.8) and $\sigma^2_{R_l}$ is the variance of the additive noise at R_l. Moreover, in (3.9), $\mathcal{P}_{R_l}$ is the transmit power at R_l, $\mathbf{F}_{l,i}$= $\mathbf{F}^T_{i,l}$, and $\mathbf{n}_i$ is the $N_i \times 1$ zero mean AWGN at U_i satisfying $\mathcal{E}\left(\mathbf{n}_i\mathbf{n}_i^H\right) = \mathbf{I}_{N_i}\sigma_i^2$ for $i \in \{1, 2\}$. Furthermore, in (3.9), $\mathbf{W}_{R_{i,l}}$ is the receive-ZF matrix at U_i, and can be given as [66]

$$\mathbf{W}_{R_{i,l}} = \left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}\mathbf{F}^H_{l,i}, \quad \text{for} \;\; i \in \{1, 2\} \;\; \text{and} \;\; l \in \{1, \cdots, L\}. \tag{3.11}$$

3.3.3 Exact end-to-end SNR

In this subsection, the exact end-to-end SNR of the kth data substream for $k \in \{1, \cdots, N_{R_l}\}$ is derived by using the signalling model presented in Section 3.3.2. To this end, by substituting (3.4), (3.5), and (3.11) into (3.9), the received signal vector at U_i can be written in an alternative form as follows:

$$\mathbf{y}_{U_{i,l}} = G_l\left(g_{i,l}\mathbf{x}_i + g_{i',l}\mathbf{x}_{i'} + \mathbf{n}_{R_l}\right) + \left[\left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}\mathbf{F}^H_{l,i}\right]\mathbf{n}_i, \tag{3.12}$$

for $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$. By carefully observing the received signal at the user nodes in (3.12), it can readily be seen that the noise term is no longer Gaussian, but the noise is non-Gaussian (coloured) due to the application of the receive-ZF detector ($\mathbf{W}_{R_{i,l}}$). It is worth noting that this coloured noise resulted due to the fact that the receive-ZF detector matrix is not a unitary matrix. Next, the filtered, coloured noise at the user nodes can explicitly be expressed as follows:

$$\tilde{\mathbf{n}}_i = \left[\left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}\mathbf{F}^H_{l,i}\right]\mathbf{n}_i. \tag{3.13}$$

Next, by using self-interference cancellation to (3.12), the signal vector of $U_{i'}$ received at U_i can be extracted as follows:

$$\tilde{\mathbf{y}}_{U_{i,l}} = G_l\left(g_{i',l}\mathbf{x}_{i'} + \mathbf{n}_{R_l}\right) + \tilde{\mathbf{n}}_i \;\; \text{for} \;\; i \in \{1, 2\}, i' \in \{1, 2\}, \;\; \text{and} \;\; i \neq i'. \tag{3.14}$$

By using (3.14), the post-processing end-to-end SNR of the kth data substream at U_i can be derived as

$$\gamma^{(k)}_{U_{i,l}} = \frac{G_l^2 g_{i',l}^2}{G_l^2\sigma^2_{R_l} + \sigma_i^2\left[\left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}\right]_{k,k}} \;\; \text{for} \;\; k \in \{1, \cdots, N_{R_l}\}. \tag{3.15}$$

By substituting G (3.10) and $g_{i',l}$ (3.8) into (3.15), the end-to-end SNR in (3.15) can be written in a more insightful form as

$$\gamma_{U_{i,l}^{(k)}} = \frac{(N_{i'} - N_{R_l})\bar{\gamma}_{i',l}\bar{\gamma}_{l,i}}{N_{R_l}\bar{\gamma}_{l,i} + ((N_{i'} - N_{R_l})\bar{\gamma}_{i',l} + (N_i - N_{R_l})\bar{\gamma}_{i,l} + N_{R_l}) \left[\left(\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right)^{-1}\right]_{k,k}}, \quad (3.16)$$

where $k \in \{1, \cdots, N_{R_l}\}$, $\bar{\gamma}_{i,l} \triangleq \mathcal{P}_i/\sigma_{R_l}^2$, $\bar{\gamma}_{l,i} \triangleq \mathcal{P}_{R_l}/\sigma_i^2$, $i, i' \in \{1, 2\}$, $l \in \{1, \cdots, L\}$ and $i \neq i'$.

Remark II.1: It is worth noting that $\gamma_{U_{1,l}^{(k)}}$ and $\gamma_{U_{2,l}^{(k)}}$ for $k \in \{1, \cdots, N_{R_l}\}$ of (3.16) are statistically independent for a given k and l. However, post-processing SNRs of multiple substreams belonging to a given user node are correlated, i.e., $\gamma_{U_{i,l}^{(k)}}$ and $\gamma_{U_{i,l}^{(k')}}$ are correlated for a given i and l. Due to this correlation effect, derivation of the probability distributions of the SNR of the smallest data substream appears mathematically intractable, and hence, the SNR bounds are derived in the next subsection.

3.3.4 Bounds on the SNR of the smallest data substream and their probability distributions

In this subsection, lower and upper bounds on the end-to-end SNR of the smallest data substream are derived. In particular, these SNR lower and upper bounds are used in deriving the important performance metrics in closed-form in the sequel.

Lower bound on the SNR of the smallest data substream

A lower bound of the SNR of the smallest data substream can be derived as follows: The maximum diagonal element of $\left(\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right)^{-1}$ can be upper bounded as [66]

$$\max_{k \in \{1 \cdots N_{R_l}\}} \left[\left(\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right)^{-1}\right]_{k,k} \leq \lambda_{\max}\left\{\left(\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right)^{-1}\right\} = \lambda_{min}^{-1}\left\{\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right\}. \quad (3.17)$$

By substituting (3.17) into (3.16), the SNR of the smallest substream of U_i for $i \in \{1, 2\}$ can then be lower bounded as follows [53]:

$$\gamma_{U_{i,l}^{min}} = \min_{k \in \{1 \cdots N_{R_l}\}} \gamma_{U_{i,l}^{(k)}} \geq \gamma_{U_{i,l}^{min}}^{\text{lb}} = \frac{\alpha_{i,l}}{\beta_{i,l} + \zeta_{i,l}\lambda_{min}^{-1}\left\{\mathbf{F}_{l,i}^H \mathbf{F}_{l,i}\right\}}, \quad (3.18)$$

where $\alpha_{i,l} = (N_{i'} - N_{R_l})\bar{\gamma}_{i',l}\bar{\gamma}_{l,i}$, $\beta_{i,l} = N_{R_l}\bar{\gamma}_{l,i}$, and $\zeta_{i,l} = (N_{i'} - N_{R_l})\bar{\gamma}_{i',l} + (N_i - N_{R_l})\bar{\gamma}_{i,l} + N_{R_l}$ for $i \in \{1,2\}$, $i' \in \{1,2\}$, $i \neq i'$, and $l \in \{1, \cdots, L\}$. It is worth noting that $\gamma^{\text{lb}}_{U^{min}_{1,l}}$ and $\gamma^{\text{lb}}_{U^{min}_{2,l}}$ are statistically independent for any l.

The bound in (3.18) is a function of the minimum eigenvalue of Wishart matrix $\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}$. The distribution of such eigenvalues is available in the existing literature [68] and using this, the cumulative distribution function (CDF) of $\gamma^{\text{lb}}_{U^{min}_{i,l}}$ can be written as follows (see Appendix A.1 for the proof):

$$\mathrm{F}_{\gamma^{\text{lb}}_{U^{min}_{i,l}}}(x) = \begin{cases} 1 - \dfrac{\left|\left[\mathbf{Q}_i\left(\frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}\right)\right]\right|}{\prod_{j=1}^{N_{R_l}}\left[\Gamma(N_i-j+1)\Gamma(N_{R_l}-j+1)\right]}, & 0 < x < \frac{\alpha_{i,l}}{\beta_{i,l}} \\ 1, & x \geq \frac{\alpha_{i,l}}{\beta_{i,l}}. \end{cases} \quad (3.19)$$

In (3.19), $\mathbf{Q}_i(x)$ is an $N_{R_l} \times N_{R_l}$ matrix with the (u,v)th element given by [68, Eq. (2.73)]

$$\left[\mathbf{Q}_i(x)\right]_{u,v} = \Gamma(N_i - N_{R_l} + u + v - 1, x)\,. \quad (3.20)$$

Upper bound on the SNR of the smallest data substream

Next, an upper bound on the SNR of the smallest data substream can be derived as follows: The maximum diagonal element of $\left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}$ can be lower bounded by any of its diagonal elements [53]. Thus, the SNR of the smallest substream of U_i for $i \in \{1,2\}$ can be upper bounded as

$$\gamma_{U^{min}_{i,l}} = \min_{k \in \{1 \cdots N_{R_l}\}} \gamma_{U^{(k)}_{i,l}} \leq \gamma^{\text{ub}}_{U^{min}_{i,l}} = \frac{\alpha_{i,l}}{\beta_{i,l} + \zeta_{i,l}\left[\left(\mathbf{F}^H_{l,i}\mathbf{F}_{l,i}\right)^{-1}\right]_{k,k}}, \quad (3.21)$$

where $\alpha_{i,l}$, $\beta_{i,l}$ and $\zeta_{i,l}$ are defined in (3.18). Again, $\gamma^{\text{ub}}_{U^{min}_{1,l}}$ and $\gamma^{\text{ub}}_{U^{min}_{2,l}}$ are independent for any l.

Using distribution of the kth diagonal element of the inverse Wishart matrix [69], the CDF of $\gamma^{\text{ub}}_{U^{min}_{i,l}}$ can then be written as follows (see Appendix A.1 for the proof):

$$F_{\gamma_{U_{i,l}^{min}}^{\text{ub}}}(x) = \begin{cases} 1 - \frac{\Gamma\left(N_i - N_{R_l}+1, \frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}\right)}{\Gamma(N_i - N_{R_l}+1)}, & 0 < x < \frac{\alpha_{i,l}}{\beta_{i,l}} \\ 1, & x \geq \frac{\alpha_{i,l}}{\beta_{i,l}}. \end{cases} \tag{3.22}$$

It is worth noting that both lower and upper SNR bounds in (3.18) and (3.21), respectively, converges to the exact SNR whenever $N_{R_l} = 1$ for $l \in \{1, \cdots, L\}$. Moreover, our analytical and simulation results clearly reveal that the outage bounds derived by using these SNR bounds are asymptotically parallel to the exact outage curves (see Fig. 3.1) in Section 3.6. Further, the tightness of them significantly improves as N_{R_l} approaches unity.

3.4 Problem Formulation

This section formulates the two relay selection strategies for MIMO ZF TWRNs. The first strategy (named SS-1) minimizes the overall outage probability. The second one (named SS-2) maximizes the achievable sum rate.

3.4.1 Relay selection based on minimizing the overall outage probability (SS-1)

In this subsection, the best relay selection strategy is formulated for maximizing the SNR of the worst data substream at the weakest user node, and thereby, for minimizing the overall outage probability. Moreover, if the user nodes are MIMO-enabled with spatial multiplexing, the overall system performance can be improved by selecting a relay which maximizes the SNR of the worst data substream at the weakest user node. Thus, our first relay selection strategy can be formulated for maximizing the effective end-to-end SNR as follows:

$$L^* = \underset{l \in \{1,\cdots,L\}}{\text{argmax}} \left[\min \left(\min_{k \in \{1,\ldots,N_{R_l}\}} \gamma_{U_{1,l}^{(k)}}, \min_{k \in \{1,\ldots,N_{R_l}\}} \gamma_{U_{2,1}^{(k)}} \right) \right], \tag{3.23}$$

where L^* is the index of the selected relay based on maximizing the SNR of the worst data substream at the weakest user node. Further, in (3.23), $\gamma_{U_{i,l}^{(k)}}$ is the end-to-end SNR of the kth data substream at U_i for $i \in \{1, 2\}$ (3.16).

Next, the overall outage probability of our system set-up is defined as the probability that the effective end-to-end SNR fall below a predefined thershold SNR γ_{th}. Thus, the overall outage probability, when the lth arbitrary relay is selected, can be written as follows [53]:

$$P_{\text{out},l} = \Pr\left[\min\left(\min_{k\in\{1\cdots N_{R_l}\}} \gamma_{U_{1,l}^{(k)}}, \min_{k\in\{1\cdots N_{R_l}\}} \gamma_{U_{2,l}^{(k)}}\right) \leq \gamma_{th}\right], \tag{3.24}$$

where $l \in \{1, \cdots, L\}$, γ_{th} is the preset threshold SNR, and $\gamma_{U_{i,l}^{(k)}}$ is the end-to-end SNR of the kth data substream at U_i for $i \in \{1, 2\}$ (3.16) when the lth relay is selected. Whenever, the best relay is selected based on the selection criterion in (3.23), it can readily be seen from (3.24) that the overall outage probability can be minimized.

If the transmit and noise powers at all the relays are the same (i.e., $\mathcal{P}_{R_l} = \mathcal{P}_R$ and $\sigma_{R_l}^2 = \sigma_R^2$ for $l \in \{1, \cdots, L\}$) and the number of antennas are equal for each relay, then by using (3.17), the relay selection strategy in (3.23) can be further simplified as follows:

$$L^* = \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}} \left[\min\left(\lambda_{min}\left\{\mathbf{F}_{l,1}^H \mathbf{F}_{l,1}\right\}, \lambda_{min}\left\{\mathbf{F}_{l,2}^H \mathbf{F}_{l,2}\right\}\right)\right]. \tag{3.25}$$

Interestingly, as per (3.25), the best relay, which minimizes the overall outage probability, maximizes the minimum of the two smallest eigenvalues of the two Wishart matrices pertinent to the corresponding relay to its two user nodes.

3.4.2 Relay selection based on maximizing the sum rate (SS-2)

This subsection formulates the best relay selection strategy based on maximizing the sum rate. In MIMO TWRNs with symmetric data traffic, each user node needs to transmit its data substreams with a common rate such that these data substreams can be decoded correctly by the intended receivers, and hence, the corresponding sum rate, when the lth relay is selected, can be defined as follows:

$$\mathcal{R}_l = 2\min\left(\mathcal{R}_{U_{1,l}}, \mathcal{R}_{U_{2,l}}\right), \tag{3.26}$$

where $\mathcal{R}_{U_{i,l}}$ is the sum of data substreams rates at U_i for $i \in \{1,2\}$, and can be written as

$$\mathcal{R}_{U_{i,l}} = \frac{1}{2}\sum_{k=1}^{N_{R_l}} \log\left(1+\gamma_{U_{i,l}^{(k)}}\right). \tag{3.27}$$

Please note that the factor of two appears in (3.26) is due to the presence of two user nodes in the TWRN of interest. Further, the pre-log factor of one-half in (3.27) is due to the two time-slots used for multiple-access and broadcast phases. By first substituting (3.27) into (3.26), and then by performing several manipulations, the sum rate can be written in an alternative form as follows:

$$\mathcal{R}_l = \log\left(\min\left(\prod_{k=1}^{N_{R_l}}\left(1+\gamma_{U_{1,l}^{(k)}}\right), \prod_{k=1}^{N_{R_l}}\left(1+\gamma_{U_{2,l}^{(k)}}\right)\right)\right). \tag{3.28}$$

By using (3.28), the best relay selection based on maximizing the sum rate can then be formulated for moderate-to-high SNR regime as follows:[1]

$$L^* = \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}}\left[\mathcal{R}_l\right] = \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}}\left[\min\left(\prod_{k=1}^{N_{R_l}}\gamma_{U_{1,l}^{(k)}}, \prod_{k=1}^{N_{R_l}}\gamma_{U_{2,l}^{(k)}}\right)\right], \tag{3.29}$$

where L^* is the index of the selected best relay. Again, if $\mathcal{P}_{R_1} = \mathcal{P}_R$ and $\sigma^2_{R_1} = \sigma^2_R$ for $l \in \{1,\cdots,L\}$ and number of antennas are equal for each relay, then the relay selection strategy given in (3.29) can further be simplified as follows:

$$L^* = \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}}\left[\min\left(\prod_{k=1}^{N_{R_l}}\left(\left[\left(\mathbf{F}_{l,1}^H\mathbf{F}_{l,1}\right)^{-1}\right]_{k,k}\right)^{-1}, \prod_{k=1}^{N_{R_l}}\left(\left[\left(\mathbf{F}_{l,2}^H\mathbf{F}_{l,2}\right)^{-1}\right]_{k,k}\right)^{-1}\right)\right], \tag{3.30}$$

Then, we recall the inequality between the product of diagonal elements a positive-definite matrix and its determinant as follows

[70, 71]:

$$|(\mathbf{A})| \leq \prod_{i=1}^{M} a_{ii}, \tag{3.31}$$

[1]Eqn. (3.29) follows due to the fact that for moderate-to-high SNR regime (i.e., for $\gamma_{U_{i,l}^{(k)}} >> 1$), the term $\left(1+\gamma_{U_{i,l}^{(k)}}\right)$ can be approximated by $\gamma_{U_{i,l}^{(k)}}$. Thus, we use the approximation $1+\gamma_{U_{i,l}^{(k)}} \approx \gamma_{U_{i,l}^{(k)}}$ for obtaining (3.29) from (3.28).

where $\mathbf{A}$ is an $M \times M$ positive-definite matrix with the ith diagonal element denoted as a_{ii} for $i \in \{1, \cdots, M\}$. By using the bound given by (3.31) in (3.30), the selection strategy of (3.29) can further be simplified as

$$L^* = \underset{l \in \{1,\cdots,L\}}{\operatorname{argmax}} \left[\min \left(\left| \left(\mathbf{F}_{l,1}^H \mathbf{F}_{l,1} \right) \right|, \left| \left(\mathbf{F}_{l,2}^H \mathbf{F}_{l,2} \right) \right| \right) \right]. \tag{3.32}$$

Thus, as per (3.32), the best relay, which maximizes the sum rate, can be selected by maximizing the minimum of the determinants of the two Wishart matrices from the relay to its two user nodes. Our extensive simulations has shown that, the selection of relays based on (3.32), instead of the more complicated criterion (3.30), does not change the maximum sum-rate achieved.

3.5 Performance Analysis

In this section, we derive the performance of our proposed relay selection strategies. For the first selection strategy (SS-1), we first derive the upper and lower bounds of the overall outage probability, and quantify the achievable diversity-multiplexing trade-off by using the corresponding high SNR outage approximations. For the second selection strategy (SS-2), a tight approximation for the achievable ergodic sum rate is derived in closed-form.

3.5.1 Overall outage probability analysis of SS-1

The overall outage probability of the optimal relay selection based on minimizing the overall outage probability (SS-1) can be written as follows:

$$P_{\text{out}} = \Pr \left[\max_{l \in \{1,\cdots,L\}} \left(\min \left(\min_{k \in \{1 \cdots N_{R_l}\}} \gamma_{U_{1,l}^{(k)}}, \min_{k \in \{1 \cdots N_{R_l}\}} \gamma_{U_{2,l}^{(k)}} \right) \right) \leq \gamma_{th} \right], \tag{3.33}$$

where $\gamma_{U_{i,l}^{(k)}}$ is the end-to-end SNR of the kh data substream at U_i received via the lth relay (3.16).

The closed-form derivation of the exact overall outage probability of (3.33) appears mathematically intractable due to the statistical dependence of the substream SNRs of a given user node (see Remark II.1). Thus, the upper and lower bounds of the overall outage probability are derived in closed-form.

Upper bound on the overall outage probability

An upper bound of the overall outage probability can be written by using the SNR lower bound in (3.18) as follows[2]:

$$P_{\text{out}}^{\text{ub}} = \Pr\left[\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma_{U_{1,l}^{min}}^{\text{lb}}, \gamma_{U_{2,l}^{min}}^{\text{lb}}\right)\right) \leq \gamma_{th}\right], \tag{3.34}$$

where $\gamma_{U_{i,l}^{min}}^{\text{lb}}$ is a lower bound for the smallest data substream SNR at U_i received via the lth relay and is defined in (3.18). Using the CDF of a maximum value of a given random variables, the outage upper bound can be derived in closed-form as follows (see Appendix A.2 for the proof):

$$P_{\text{out}}^{\text{ub}} = \prod_{l=1}^{L}\left(\sum_{i=1}^{2} F_{\gamma_{U_{i,l}^{min}}^{\text{lb}}}(\gamma_{th}) - \prod_{i=1}^{2} F_{\gamma_{U_{i,l}^{min}}^{\text{lb}}}(\gamma_{th})\right), \tag{3.35}$$

where $F_{\gamma_{U_{i,l}^{min}}^{\text{lb}}}(\gamma_{th})$ is the CDF of $\gamma_{U_{i,l}^{min}}^{\text{lb}}$ and is defined in (3.19).

Next, an asymptotic high SNR approximation of the upper bound of the overall outage probability can be derived by using (3.35) as follows (see Appendix A.3 for the proof):

$$P_{\text{out}}^{\text{ub},\infty} = \left(\prod_{l=1}^{L}\Lambda_l\right)\left(\frac{\gamma_{th}}{\bar{\gamma}_{U,R}}\right)^{d_{\text{ub}}} + o\left(\bar{\gamma}_{U,R}^{-(d_{\text{ub}}+1)}\right), \tag{3.36}$$

where $\bar{\gamma}_{1,l} = \bar{\gamma}_{2,l} = \bar{\gamma}_{U,R_l}$, $\bar{\gamma}_{l,1} = \bar{\gamma}_{l,2} = \bar{\gamma}_{R_l,U}$, $\bar{\gamma}_{R_l,U} = \nu_l\bar{\gamma}_{U,R_l}$, and $\bar{\gamma}_{U,R_l} = C_l\bar{\gamma}_{U,R}$. In (3.36), d_{ub} is an upper bound of the achievable diversity order and can be derived as

$$d_{\text{ub}} = \sum_{l=1}^{L}\left(\min(N_1, N_2) - N_{R_l} + 1\right). \tag{3.37}$$

Besides, in (3.36), Λ_l can be derived as follows [53]:

$$\Lambda_l = \begin{cases} \dfrac{\eta_{1,l}\left(\frac{N_1+N_2-2N_{R_l}}{N_2-N_{R_l}}\right)^{N_1-N_{R_l}+1}}{(N_1-N_{R_l}+1)C_l^{N_1-N_{R_l}+1}}, & N_1 < N_2 \\ \dfrac{\eta_{2,l}\left(\frac{N_1+N_2-2N_{R_l}}{N_1-N_{R_l}}\right)^{N_2-N_{R_l}+1}}{(N_2-N_{R_l}+1)C_l^{N_2-N_{R_l}+1}}, & N_2 < N_1 \\ \dfrac{\left(\eta_{1,l}+\eta_{2,l}\right)2^{N-N_{R_l}+1}}{(N-N_{R_l}+1)C_l^{N-N_{R_l}+1}}, & N_1 = N_2 = N, \end{cases} \tag{3.38}$$

[2]Our analytical and simulation results reveal that this outage upper bound is asymptotically parallel to the exact outage curves and converges to the exact outage whenever $N_{R_l} \to 1$ for $l \in \{1, \cdots, L\}$ (see Fig. 3.1) in Section 3.6.

Moreover, in (3.38), $\eta_{i,l}$ for $i \in \{1,2\}$ can be defined as follows:

$$\eta_{i,l} = \begin{cases} \frac{\left|\left|(\mathbf{B}_{i,l}\right)\right.}{\prod_{j=1}^{N_{R_l}}\left[\Gamma(N_i-j+1)\Gamma(N_{R_l}-j+1)\right]}, & N_{R_l} \neq 1 \\ 1/\Gamma(N_i), & N_{R_l} = 1, \end{cases} \tag{3.39}$$

where $\mathbf{B}_{i,l}$ is an $(N_{R_l}-1) \times (N_{R_l}-1)$ matrix with the (p,q)th element given by $[\mathbf{B}_{i,l}]_{p,q} = \Gamma(N_i - N_{R_l} + p + q + 1)$ for $i \in \{1,2\}$.

Lower bound on the overall outage probability

A lower bound of the overall outage probability can then be defined by using the SNR upper bound in (3.21) as[3]

$$P_{\text{out}}^{\text{lb}} = \Pr\left[\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma_{U_{1,l}^{min}}^{\text{ub}}, \gamma_{U_{2,l}^{min}}^{\text{ub}}\right)\right) \leq \gamma_{th}\right], \tag{3.40}$$

where $\gamma_{U_{i,l}^{min}}^{\text{ub}}$ is an upper bound for the smallest data substream SNR at U_i received via the lth relay and is defined in (3.21). Next, the outage upper bound can be derived in closed-form as (see Appendix A.2 for the proof)

$$P_{\text{out}}^{\text{lb}} = \prod_{l=1}^{L}\left(\sum_{i=1}^{2} F_{\gamma_{U_{i,l}^{min}}^{\text{ub}}}(\gamma_{th}) - \prod_{i=1}^{2} F_{\gamma_{U_{i,l}^{min}}^{\text{ub}}}(\gamma_{th})\right), \tag{3.41}$$

where the CDF of $\gamma_{U_{i,l}^{min}}^{\text{ub}}$ is defined in (3.22).

Next, the asymptotic high SNR approximation for the lower bound of the overall outage probability can be derived as follows (see Appendix A.3 for the proof):

$$P_{\text{out}}^{\text{lb},\infty} = \left(\prod_{l=1}^{L}\Theta_l\right)\left(\frac{\gamma_{th}}{\bar{\gamma}_{U,R}}\right)^{d_{\text{lb}}} + o\left(\bar{\gamma}_{U,R}^{-(d_{\text{lb}}+1)}\right), \tag{3.42}$$

where d_{lb} is a lower bound of the achievable diversity order and is given by

$$d_{\text{lb}} = \sum_{l=1}^{L}\left(\min(N_1, N_2) - N_{R_l} + 1\right). \tag{3.43}$$

[3]Again, our analytical and simulation results clearly show that this outage lower bound is asymptotically parallel to the exact outage curves and converges to the exact outage whenever $N_{R_l} \to 1$ for $l \in \{1, \cdots, L\}$ (see Fig. 3.1) in Section 3.6.

Further, in (3.42), Θ_l is defined as follows:

$$\Theta_l = \begin{cases} \frac{\left(\frac{N_1+N_2-2N_{R_l}}{N_2-N_{R_l}}\right)^{N_1-N_{R_l}+1}}{\Gamma(N_1-N_{R_l}+2)C_l^{N_1-N_{R_l}+1}}, & N_1 < N_2 \\ \frac{\left(\frac{N_1+N_2-2N_{R_l}}{N_1-N_{R_l}}\right)^{N_2-N_{R_l}+1}}{\Gamma(N_2-N_{R_l}+2)C_l^{N_2-N_{R_l}+1}}, & N_2 < N_1 \\ \frac{2^{N-N_{R_l}+2}}{\Gamma(N-N_{R_l}+2)C_l^{N-N_{R_l}+1}}, & N_1 = N_2 = N \end{cases}. \tag{3.44}$$

3.5.2 Diversity-multiplexing trade-off of SS-1

Multiple antennas can be used to increase diversity order or spatial multiplexing gains in wireless systems. The seminal work of Zheng and Tse [41] showed that both types of these gains can be simultaneously obtained for a given multiple-antenna channel subjected to a fundamental trade-off. Thus, the multiplexing gain comes at the expense of sacrificing the diversity gain and vice versa. In this section, we mainly focus on quantifying the diversity-multiplexing trade-off (DMT) of the proposed SS-1. To this end, two DMT expressions are derived by using the lower and upper bound of the asymptotic high SNR approximations of the outage probability, and thereby, the achievable DMT, maximum diversity order, and spatial multiplexing gain are quantified in closed-form.

To begin with, the diversity order and spatial multiplexing gain can be defined as follows [41]:

$$d = -\lim_{\gamma\to\infty} \frac{\log\left(P_{\text{out}}(\mathcal{R},\gamma)\right)}{\log(\gamma)} \quad \text{and} \quad r = \lim_{\gamma\to\infty} \frac{\mathcal{R}(\gamma)}{\log(\gamma)}, \tag{3.45}$$

where $P_{\text{out}}(\mathcal{R},\gamma)$ is the information rate outage probability evaluated at γ while keeping the threshold rate at $\mathcal{R}$. Further, $\mathcal{R}(\gamma)$ is the achievable data rate evaluated at the γ. Next, by using (3.45), the DMT can be defined as [41]

$$d(r) = -\lim_{\gamma\to\infty} \frac{\log\left(P_{\text{out}}(r\log(\gamma))\right)}{\log(\gamma)}. \tag{3.46}$$

In the following subsections, the achievable DMT of best relay selection for MIMO TWRNs is quantified by using (3.46).

DMT by using outage lower bound

To begin with, the DMT is quantified by using the asymptotic outage lower bound in (3.42). In this context, the achievable mutual information can be upper bounded as

$$\mathcal{I}_{\text{ub}}= \left(\min_{l\in\{1,\cdots,L\}} (N_{R_l})\right) \log\left(1+\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma^{\text{ub}}_{U^{min}_{1,l}},\gamma^{\text{ub}}_{U^{min}_{2,l}}\right)\right)\right), \quad (3.47)$$

where $\gamma^{\text{ub}}_{U^{min}_{i,l}}$ is defined in (3.21). A lower bound of the information outage probability can be then defined by using (3.47) as follows:

$$P^{\text{lb}}_{\text{out}} = \Pr\left(\mathcal{I}_{\text{ub}} \leq \mathcal{R}_{th}\right), \quad (3.48)$$

where $\mathcal{R}_{th}$ is the pre-set threshold rate of the overall system and is defined as $\mathcal{R}_{th} = r\log\left(1+\bar{\gamma}_{U,R}\right)$ [72]. By substituting (3.47) into (3.48), the outage lower bound can be written in an alternative form as follows:

$$P^{\text{lb}}_{\text{out}} = \Pr\left(\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma^{\text{ub}}_{U^{min}_{1,l}},\gamma^{\text{ub}}_{U^{min}_{2,l}}\right)\right) \leq \left(\bar{\gamma}_{U,R}\right)^{\frac{r}{N^{min}_R}} - 1\right), \quad (3.49)$$

where $N^{min}_R = \min_{l\in\{1,\cdots,L\}} N_{R_l}$. By using (3.42), the outage lower bound in (3.49) can be asymptotically approximated when $\bar{\gamma}_{U,R}\to\infty$ as follows:

$$P^{\text{lb},\infty}_{\text{out}} \approx \bar{\gamma}_{U,R}^{-\sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right)\left(1-\frac{r}{N^{min}_R}\right)}. \quad (3.50)$$

The lower bound of the achievable DMT can then be derived by using its definition in [72] as

$$d^{\text{lb}}(r) = \left(\sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right)\right)\left(1-\frac{r}{N^{min}_R}\right). \quad (3.51)$$

DMT by using outage upper bound

A lower bound of the achievable mutual information can be written as

$$\mathcal{I}_{\text{lb}} = N^{min}_R \log\left(1+\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma^{\text{lb}}_{U^{min}_{1,l}},\gamma^{\text{lb}}_{U^{min}_{2,l}}\right)\right)\right), \quad (3.52)$$

where $\gamma^{\mathrm{lb}}_{U^{min}_{i,l}}$ is defined in (3.18). Then, the corresponding information outage probability upper bound can be defined by using (3.52) as

$$P^{\mathrm{lb}}_{\mathrm{out}} = \Pr\left(\mathcal{I}_{\mathrm{lb}} \leq \mathcal{R}_{th}\right), \tag{3.53}$$

By substituting (3.52) into (3.53), the upper bound of the outage probability can be then written as

$$P^{\mathrm{ub}}_{\mathrm{out}} = \Pr\left(\max_{l\in\{1,\cdots,L\}}\left(\min\left(\gamma^{\mathrm{lb}}_{U^{min}_{1,l}}, \gamma^{\mathrm{lb}}_{U^{min}_{2,l}}\right)\right) \leq \left(\bar{\gamma}_{U,R}\right)^{\frac{r}{N^{min}_R}} - 1\right). \tag{3.54}$$

An asymptotic high SNR approximation of the outage upper bound in (3.54) can be derived as

$$P^{\mathrm{ub},\infty}_{\mathrm{out}} \approx \bar{\gamma}_{U,R}^{-\sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right)\left(1-\frac{r}{N^{min}_R}\right)}. \tag{3.55}$$

The corresponding upper bound of the achievable DMT can then be derived as

$$d^{\mathrm{ub}}(r) = \left(\sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right)\right)\left(1-\frac{r}{N^{min}_R}\right). \tag{3.56}$$

Remark IV.1: Note that the DMT derived by using the lower and upper bounds of the overall outage probability given in (3.51) and (3.56), respectively, are the same. Thus, it can be concluded that the achievable diversity order is given by

$$d(r) = \left(\sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right)\right)\left(1-\frac{r}{\min\limits_{l\in\{1,\cdots,L\}}(N_{R_l})}\right). \tag{3.57}$$

This DMT (3.57) provides several important insights. The maximum achievable diversity order and the spatial multiplexing gain can be derived as

$$d_{max} = \sum_{l=1}^{L}\left(\min(N_1,N_2)-N_{R_l}+1\right) = L\left(\min(N_1,N_2)+1\right) - \sum_{l=1}^{L} N_{R_l}. \tag{3.58}$$

$$r_{max} = \min_{l\in\{1,\cdots,L\}} N_{R_l}. \tag{3.59}$$

Thus, the spatial multiplexing gain of SS-1 is solely governed by the number of antennas at the relay which has the smallest antenna array (N^{min}_R). However, whenever N^{min}_R is increased for a given N_1 and N_2 in order to increase the multiplexing

gain, the achievable diversity order decreases. This decrease in diversity order is more prominent when the MIMO TWRN operates without the best relay selection (i.e., $L = 1$). The best relay selection based on SS-1 alone does not increase the achievable spatial multiplexing gain as it is independent of the number of available relays (L). Nevertheless, the multiplexing gain can indeed be improved by maximizing minimum number of antennas at any relay. Counter-intuitively, the best relay selection based on SS-1 improves the diversity order by a multiplicative factor of L whenever $N_{R_l} = N_R$ for $l \in \{1, \cdots, L\}$, and thereby, compensates the diversity order degradation whenever N_R is increased in an attempt to improve the multiplexing gain.

Interestingly, the second equality of d_{max} in (3.58) reveals that whenever the sum of the relay antennas is fixed, the maximum diversity order is always a constant. In this context, the multiplexing gain can readily be improved by equally distributing antennas among the available set of relays.

3.5.3 Sum rate analysis of SS-2

In this subsection, the sum rate of the best relay selection based on maximizing the sum rate (SS-2) is investigated. Again, the exact derivation of the ergodic sum rate of SS-2 appears mathematically intractable due to statistical dependence of data substream SNRs belonging to a given user node (see remark II.1). Thus, in this context, an approximation of the ergodic sum rate can be derived for obtaining useful insights. To begin with, the sum of the rates of data substreams of U_i (3.27) can be approximated by its lower bound as

$$\mathcal{R}_{U_{i,l}} = \frac{1}{2} \sum_{k=1}^{N_{R_l}} \log\left(1+\gamma_{U_{i,l}^{(k)}}\right) \approx \frac{N_{R_l}}{2} \log\left(1+\gamma_{U_{i,l}^{min}}\right), \tag{3.60}$$

for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ where $\gamma_{U_{i,l}^{min}}$ is defined in (3.21). By first substituting (3.60) into (3.26), and then by performing several mathematical formulations, an approximation for the sum rate for the best relay selection based on SS-2

can then be written as follows[4]:

$$\mathcal{R}_{L^*} \approx \left(\max_{l\in\{1,\cdots,L\}} (N_{R_l})\right) \log\left(1+\min\left(\gamma_{U^{min}_{1,L^*}}, \gamma_{U^{min}_{2,L^*}}\right)\right). \tag{3.61}$$

By using (3.21), $\gamma_{U^{min}_{i,L^*}}$ for $i \in \{1,2\}$ in (3.61) can be approximated as

$$\gamma_{U^{min}_{i,L^*}} \approx \frac{\alpha_{i,L^*}}{\beta_{i,L^*} + \zeta_{i,L^*}\left[\left(\mathbf{F}^H_{L^*,i}\mathbf{F}_{L^*,i}\right)^{-1}\right]_{k,k}}, \tag{3.62a}$$

where $k \in \{1,\cdots,N_R^{max}\}$. Further, α_{i,L^*}, β_{i,L^*}, and ζ_{i,L^*} are defined as follows:

$$\alpha_{i,L^*} = (N_{i'} - N_R^{max})\bar{\gamma}_{i',L^*}\bar{\gamma}_{L^*,i}, \tag{3.62b}$$

$$\beta_{i,L^*} = N_R^{max}\bar{\gamma}_{L^*,i}, \tag{3.62c}$$

$$\zeta_{i,L^*} = (N_{i'} - N_R^{max})\,\bar{\gamma}_{i',L^*} + (N_i - N_R^{max})\bar{\gamma}_{i,L^*} + N_R^{max}. \tag{3.62d}$$

Here, in (3.62b), (3.62c), and (3.62d), $N_R^{max} = \max_{l\in\{1,\cdots,L\}} N_{R_l}$, $i \in \{1,2\}$, $i' \in \{1,2\}$, and $i \neq i'$. The average sum rate approximation can next be derived by taking the expectation of $\mathcal{R}_{L^*}$ in (3.61) over the corresponding SNR, $\gamma_{\text{eq}} = \min\left(\gamma_{U^{min}_{1,L^*}}, \gamma_{U^{min}_{2,L^*}}\right)$, as follows (see Appendix A.4 for the proof):

$$\begin{aligned}\bar{\mathcal{R}}_{L^*} \approx\ & \frac{N_R^{max}}{\ln(2)} \sum_{m=0}^{N-N_R^{max}} \sum_{n=0}^{N-N_R^{max}} \frac{1}{2^{m+n}\ m!\ n!} \left(\mathbb{J}(m+n, 2\zeta_{L^*}, \alpha_{L^*}+\beta_{L^*})\right. \\ &- \mathbb{J}(m+n, 2\zeta_{L^*}, \beta_{L^*}) - (m+n)\left(\mathbb{J}(m+n-1, 2\zeta_{L^*}, \alpha_{L^*}+\beta_{L^*})\right. \\ &- \left.\left.\mathbb{J}(m+n-1, 2\zeta_{L^*}, \beta_{L^*})\right)\right),\end{aligned} \tag{3.63}$$

where the function $\mathbb{J}(x,y,z)$ is given by

$$\begin{aligned}\mathbb{J}(x,y,z) &= \int_0^\infty \lambda^x \exp(-\lambda)\ln(y+z\lambda)\,\mathrm{d}\lambda \\ &= \Gamma(x+1)\left(\ln(y) + \sum_{p=0}^{x} \frac{1}{\Gamma(x-p+1)}\left(\left(\frac{-y}{z}\right)^{x-p}\exp\left(\frac{y}{z}\right)E_1\left(\frac{y}{z}\right)\right.\right. \\ &\quad \left.\left. + \sum_{q=1}^{x-p}\Gamma(q)\left(\frac{-y}{z}\right)^{x-p-q}\right)\right).\end{aligned} \tag{3.64}$$

Remark IV.2: The close-form approximation of the ergodic sum rate presented in (3.63) is valid for different number of relay antennas (N_{R_l} for $l \in \{1,\cdots,L\}$).

[4]It is worth noting that the sum rate approximation in (3.61) provides very close results as shown by our extensive Monte-Carlo simulations in Fig. 3.4 presented in Section 3.6.

Nevertheless, our numerical results reveal that this approximation slightly weakens when all available relays are equipped with the same number of antennas ($N_{R_l} = N_R$ for $i \in \{1, \cdots, L\}$). Thus, for this case, an tight ergodic sum rate approximation can next be derived explicitly as follows (see Appendix A.5 for the proof):

$$\begin{aligned}\bar{\mathcal{R}}_{L^*} &\approx \frac{N_R}{\ln(2)} \sum_{l=1}^{L} \sum_{k=1}^{2l(N-N_R)} (-1)^l \binom{l}{L} \frac{\beta_{k,2l,N-N_R+1}}{(2l)^k} \\ &\times \left(k\mathbb{J}(k-1, 2l\zeta, \alpha+\beta) + \mathbb{J}(k, 2l\zeta, \beta)\right. \\ &\quad \left. -k\mathbb{J}(k-1, 2l\zeta, \beta) - \mathbb{J}(k, 2l\zeta, \alpha+\beta)\right), \end{aligned} \tag{3.65}$$

where the function $J(x, y, z)$ is defined in (3.64). Here, in (3.65), $\alpha = \alpha_{i,l}$, $\zeta = \zeta_{i,l}$, and $\beta = \beta_{i,l}$ for all $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$. The multinomial coefficient $\beta_{k,2l,N-N_R+1}$ in (3.65) can be written as

$$\beta_{k,2l,N-N_R+1} = \sum_{i=k-N+N_R}^{k} \frac{\beta_{i,2l-1,N-N_R+1}}{(k-i)!} I_{[0,(2l-1)(N-N_R)]}(i). \tag{3.66}$$

In (3.66), $\beta_{0,0,N-N_R+1} = \beta_{0,2l,N-N_R+1} = 1$, $\beta_{k,1,N-N_R+1} = 1/\ k!$, $\beta_{1,2l,N-N_R+1} = 2l$, $I_{[a,c]}(b) = 1$ for $a \leq b \leq c$, and $I_{[a,c]}(b) = 0$ otherwise.

3.5.4 Diversity-multiplexing trade-off of SS-2

By following similar steps to those in Section 3.5.2, the DMT of SS-2 can be derived as follows[5]:

$$d(r) = \left(\sum_{l=1}^{L'} \left(\min(N_1, N_2) - \max_{l \in \{1,\cdots,L\}} (N_{R_l}) + 1\right)\right)\left(1 - \frac{r}{\max\limits_{l \in \{1,\cdots,L\}} (N_{R_l})}\right), \tag{3.67}$$

where L' is the number of relays having $N_R^{max} = \max_{l \in \{1,\cdots,L\}}(N_{R_l})$ number of antennas. Interestingly, by comparing (3.57) and (3.67), it can be noted that the DMT of SS-2 becomes identical to that of SS-1 whenever all available relays are equipped with the same number of antennas (i.e., $N_{R_l} = N_R$ for $i \in \{1, \cdots, L\}$).

[5]The proof of the DMT of SS-2 (3.67) is omitted due to its repetitive nature to the proof of (3.57).

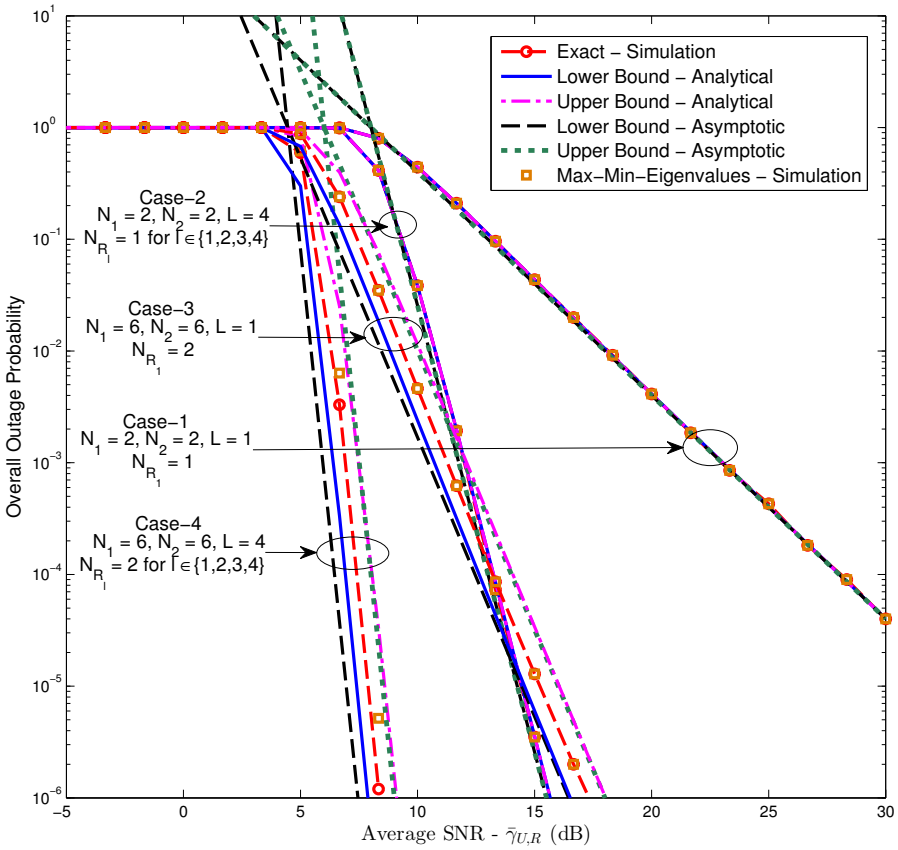


Figure 3.1: The overall outage probability of the best relay selection based on SS-1 for MIMO AF TWRNs with equal number of antennas at relays. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$.

3.6 Numerical Results

In this section, numerical and simulation results are presented to investigate the performance of our proposed selection strategies. To this end, the overall outage probability, average sum rate, and diversity-multiplexing curves are plotted. These plots are generated by using both analytical expressions and Monte-Carlo simulations. Importantly they also justify the validity of our analysis.

3.6.1 Overall outage probability of SS-1

In Fig. 3.1, the performance of the best relay selection based on minimizing the outage probability (SS-1) is plotted for four specific system set-ups[6] (see case-1 to case-4 in Fig. 3.1). To this end, the upper and lower outage bounds are plotted by using (3.35) and (3.41), respectively. The outage probability curves for $L = 1$ (i.e., no relay selection) are plotted for comparison purposes. The outage curves of case-1/case-2 and case-3/case-4 reveal the significant performance gains of best relay selection. For example, at an outage probability of 10^{-5}, selecting the best out of four relays (case-4) provides almost 7.5 dB SNR gain over no relay selection (case-3). Further, our asymptotic outage curves reveal that selecting the best out of four relays (case -2) provides four-times the diversity order that is achieved by no relay selection (case-1). Besides, Fig. 3.1 clearly reveals that our upper and lower outage bounds approach to the simulated curves for single-antenna relay TWRNs. Further, the outage curves obtained by maximizing the minimum of the eigenvalues of the Wishart matrices closely follow the simulated outage curves obtained via (3.23), and this observation justifies the validity of (3.25).

Fig. 3.2, the overall outage probability of SS-1 is plotted for three relays equipped with different number of antennas. The high SNR outage asymptotics are plotted by using (3.36) and (3.42) to investigate the achievable diversity order by three specific system configurations (see case-1, case-2, and case-3 in Fig. 3.2.). The outage curves are plotted by using Monte-Carlo simulations for diversity order comparison purposes. Fig. 3.2 shows that the achievable diversity order of case-1, case-2, and case-3 are eleven, nine, and seven respectively. These diversity orders match with those of (3.58), validating our diversity order analysis. Moreover, for a fixed number of antennas at the two user nodes and a fixed number of relays, the system set-up having the least number of aggregate relay antennas provides the highest diversity order, whereas, the system set-up having the highest number of aggregate relay antennas provides the least diversity order. Intuitively, the achievable overall

[6]These outage curves are plotted for TWRNs with relays having equal number of relay antennas for a given case. The outage curves for TWRNs with relays having different number of antennas are plotted separately in Fig. 3.2 to avoid lack of clarity due to cluttering of the curves.

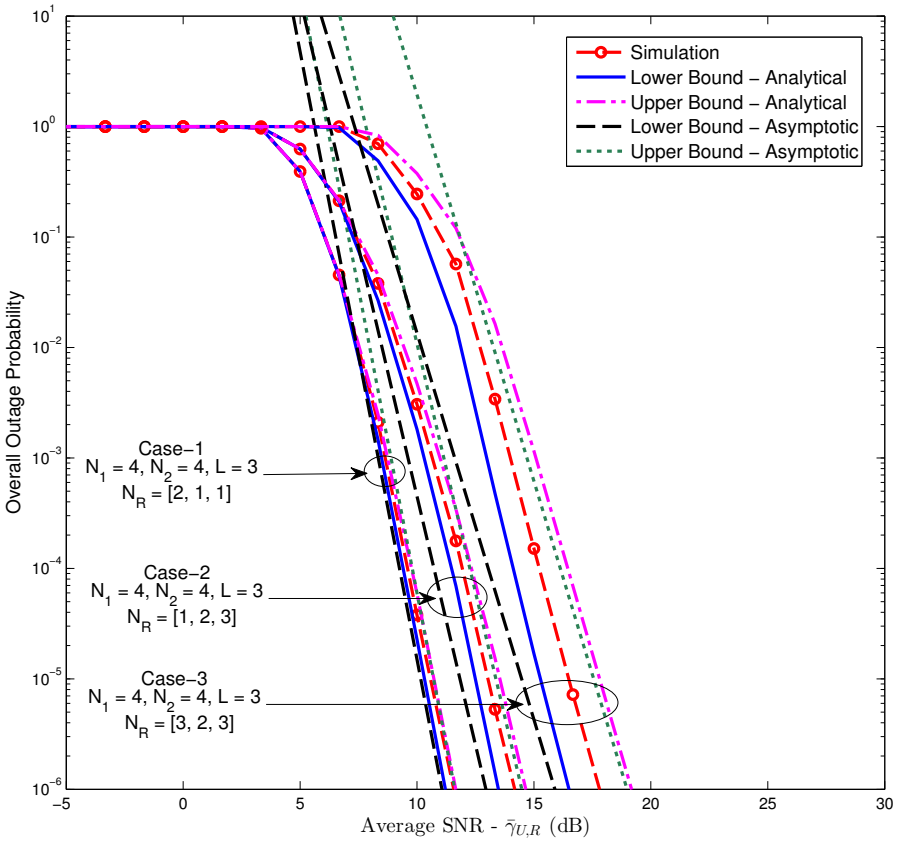


Figure 3.2: The overall outage probability of best relay selection (SS-1) for MIMO AF TWRNs with different number of antennas at relays. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$

diversity order by the best relay selection based on SS-1 is given by the aggregate of the diversity orders of each end-to-end relayed-channel.

3.6.2 Sum rate of SS-2

In Fig. 3.3, the sum rate of the best relay selection based on maximizing the sum rate (SS-2) is investigated when the number of antennas at each relay is fixed to three. Six different cases are plotted to obtain insights about the achievable sum rate performance of SS-2. Monte-Carlo simulations and the sum rate approximations from (3.65) are plotted. Fig. 3.3 clearly reveals that the sum rate heavily depends on (i) the relay antenna count, (ii) user node antenna count, and (iii) number of relays. For

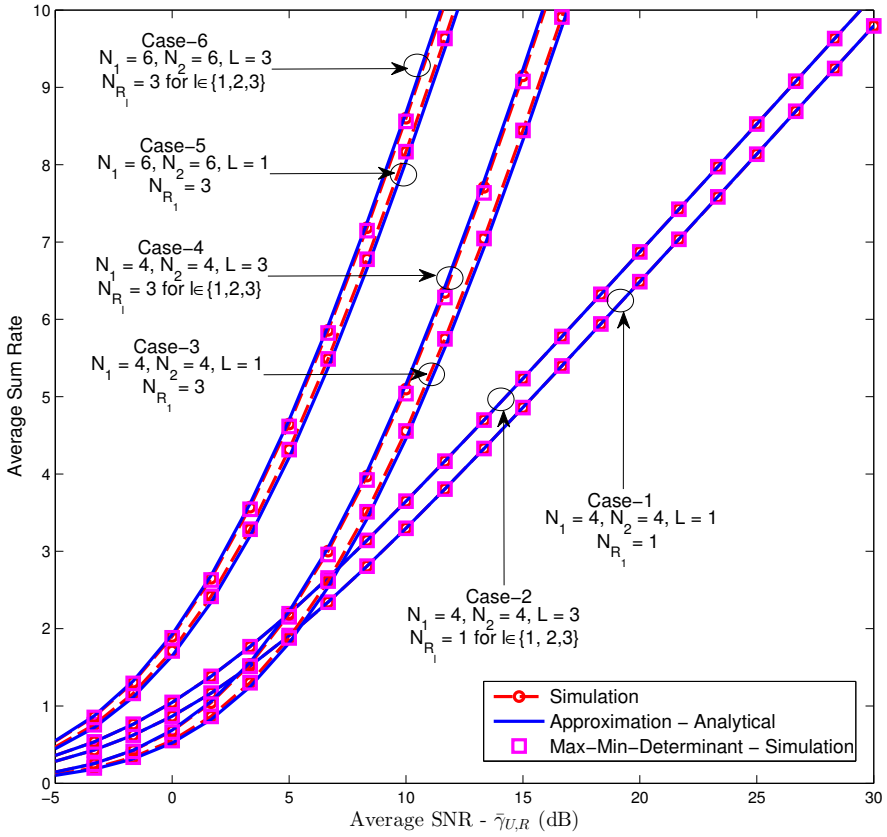


Figure 3.3: The achievable average sum rate of the best relay selection based on SS-2 for MIMO AF TWRNs with equal number of antennas at relays. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$.

example, at an SNR of 15 dB, three relays with user nodes each with four antennas (case-2/case-4) achieves a sum rate increase of four bits/channel-use/Hz by upgrading the single-antenna relays (case-2) to triple-antenna relays (case-4). Moreover, at an SNR of 10 dB, for the triple-relay/triple-antenna-relay TWRN, a sum rate gain of 3 bits/channel-use/Hz can be achieved by increasing the number of user antennas from four (case-4) to six (case-6). Besides, the triple-relay TWRN with quadruple-antenna users (case-4) provides about a sum rate gain of one bits/channel-use/Hz over the single-relay counterpart (case-3). Fig. 3.3 shows that the sum rate curves obtained via relay selection based on maximizing the minimum determinant of the Wishart matrices (3.32) coincides with the simulated sum rate curves. Further, our

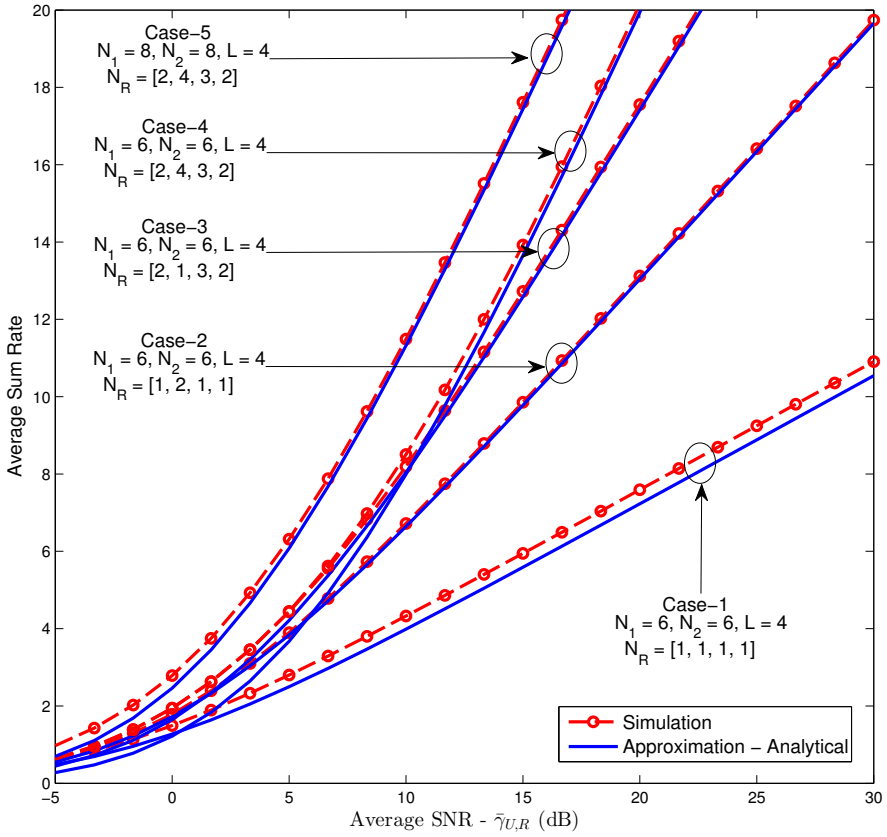


Figure 3.4: The achievable average sum rate of the best relay selection based on SS-2 for MIMO AF TWRNs with different number of antennas at relays. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$.

analytical sum rate approximation is significantly tight to the simulated sum rate curves.

In Fig. 3.3, we saw how the total number of antennas and relays tends to increase the sum rate. In Fig. 3.4, we try to isolate the effect of the number of relay antennas. For this purpose, we fix the number of relays to four and plot the sum rate performance with different numbers of relay antennas. Fig. 3.4 clearly shows that the sum rate is heavily dependent on the number of relay antennas. By comparing sum rate curves of case-1, case-2, case-3, and case-4, we conclude that the achievable sum rate, and hence the spatial multiplexing gain is solely governed by the relay with the largest antenna array. For example, at an SNR of 20 dB, case-4

provides sum rate gains of 12.5, 7, and 2.5 bits/channel-use/Hz over case-1, case-2, and case-3, respectively. Intuitively, these sum rate gains are a direct consequence of having a quadruple, triple, and dual-antenna relay as the relay with the largest antenna array in case-4, case-3, and case-2, respectively. Further, TWRNs of case-4 and case-5 have the same set of relays, however, the TWRNs of the former and latter cases are equipped with sextuple-antenna and octuple-antenna user nodes, respectively. Thus, case-5 provides a sum rate gain of 3.5 bits/channel-use/Hz over case-4, nevertheless, the spatial multiplexing gains of both these cases are the same as the largest relay-antenna array is four for both the cases. Again, the sum rate approximations are plotted by using our analysis (3.63) and they are fairly tight to the Monte-Carlo plots[7].

3.6.3 Diversity-multiplexing trade-off

In Fig. 3.6.3, the achievable diversity-multiplexing trade-off (DMT) of best relay selection strategies for MIMO TWRNs with relays having same number of antennas is investigated. Thus, DMT curves in Fig. 3.6.3 valid for both SS-1 and SS-2[8]. Here, two sets of DMT curves are plotted by changing the number of relays and size of the relay antenna arrays. Fig. 3.6.3 clearly reveals that the diversity order increases whenever the number of available relays (L) is increased, however, the achievable spatial multiplexing gain does not increase with L. Counter-intuitively, the achievable diversity order decreases whenever the number of antennas at the relay (N_R) is increased; compensating this trend is the increase of the multiplexing gain. These two observations clearly reveal that the spatial multiplexing gain is solely governed by N_R. Whenever N_R is increased, much of the available degrees of freedom boosts the multiplexing gain, and consequently, the achievable diversity order decreases. Nevertheless, the best relay selection significantly alleviates this inherent reduction of diversity order by increasing the degrees of freedom available for boosting the overall diversity order. Thus, there is a fundamental DMT associ-

[7]The analytical sum rate curve corresponding to case-1 slightly deviates away from the exact sum rate curve. Thus, this observation clearly justifies remark IV-2 and hence the derivation of a tighter sum rate approximation (3.65) for the TWRNs with relays having the same number of antennas.

[8]Whenever all available relays are equipped with the equal number of antennas, the DMT of SS-1 (3.46) and DMT of SS-2 (3.67) become identical.

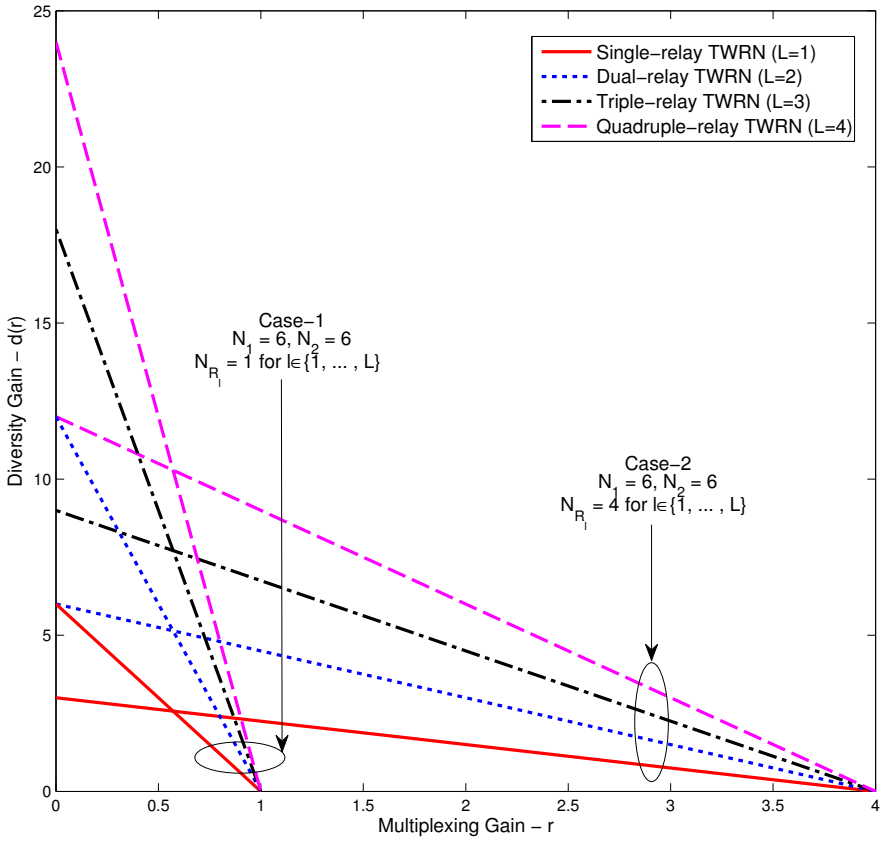


Figure 3.5: Diversity-multiplexing trade-off (DMT) comparison.

ated with TWRNs with multiple-data streams, and the best relay selection indeed improves the achievable DMT.

In Fig. 3.6.3, the achievable DMTs of both SS-1 and SS-2 are studied for MIMO TWRNs whenever the available relays are equipped with different number of antennas. DMT curves of case-1 and case-2 clearly reveal that the maximum achievable diversity order of SS-1 is always higher than that of SS-2, whereas SS-2 provides the highest multiplexing gain. This behaviour is not surprising as relay selection based on SS-1 is designed to minimize the outage probability and thereby to maximize the overall diversity order, while the relay selection based on SS-2 is designed to maximize the sum rate.

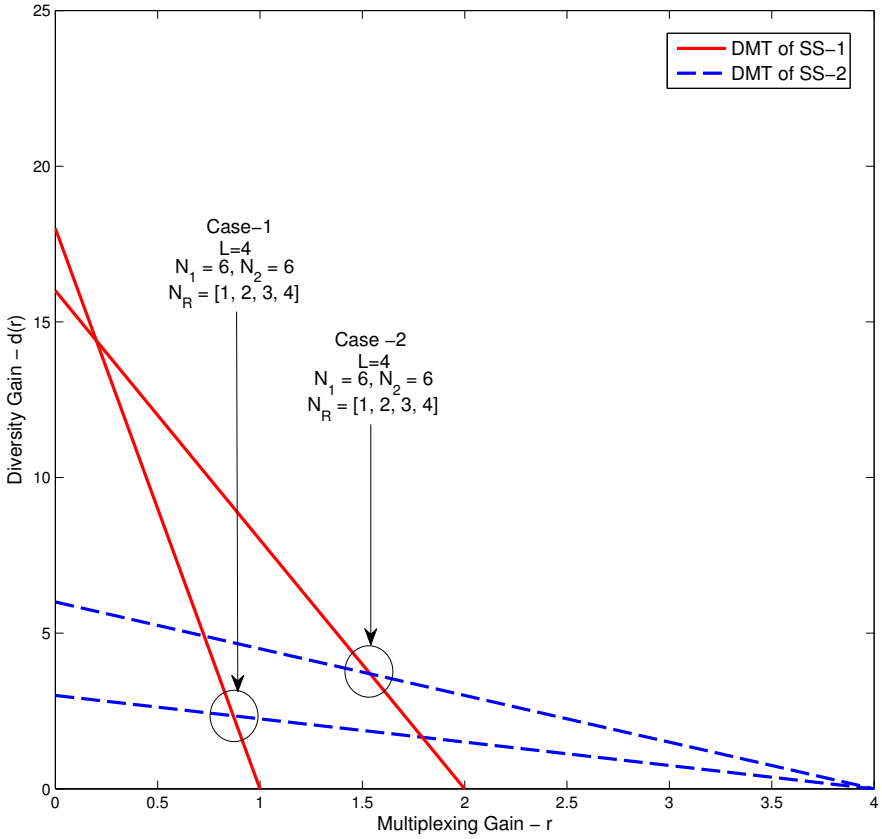


Figure 3.6: Diversity-multiplexing trade-off (DMT) comparison.

3.6.4 Outage probability and sum rate comparison between SS-1 and SS-2

In Fig. 3.6.4 and Fig. 3.6.4, the overall outage probability and sum rate performances of our proposed relay selection strategies are compared. To this end, in Fig. 3.6.4 and Fig. 3.6.4, the sum rate and outage probability, respectively, are plotted for TWRNs with best relay selection based on SS-1 and SS-2 by considering four cases of system configurations. Fig. 3.6.4 clearly shows that the sum rate performance of SS-2 is significantly better than that of SS-1. This observation is not surprising as SS-2 is explicitly designed to maximize the sum rate. However, Fig. 3.6.4 reveals that the outage probability performance of SS-1 is considerably better than that of SS-2. Again, this behaviour is not surprising, as well, because SS-1

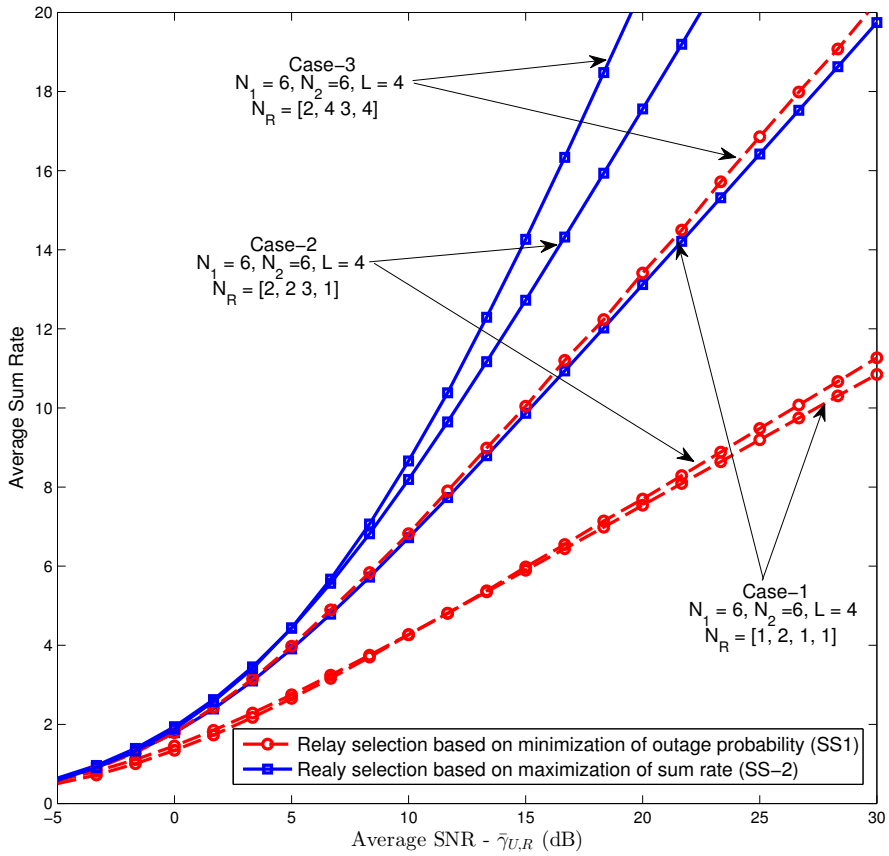


Figure 3.7: Sum rate comparison of two relay selection strategies for MIMO TWRNs with spatial multiplexing. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$.

is designed to minimize the overall outage probability. Counter-intuitively, case-1 provides the worst sum rate performance (see Fig. 3.6.4), as well, the best outage probability performance (see Fig. 3.6.4) irrespective of the relay selection strategy (i.e., SS-1 or SS-2). Similarly, case-3 provides the best sum rate performance and worst outage probability performance. The aforementioned observations are due to direct consequences of the achievable diversity-multiplexing trade-off, and notably, larger relay antenna arrays indeed increases the multiplexing gain, while reducing the overall diversity order.

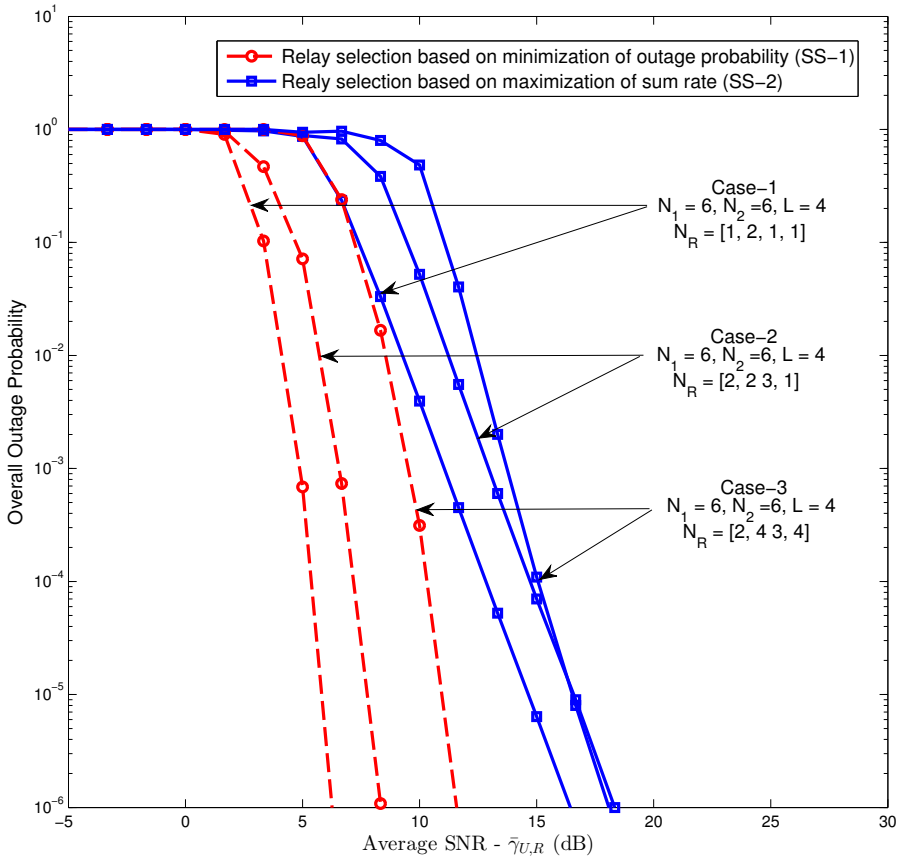


Figure 3.8: Outage probability comparison of two relay selection strategies for MIMO TWRNs with spatial multiplexing. The average transmit SNRs $\bar{\gamma}_{i,l}$ and $\bar{\gamma}_{l,i}$ for $i \in \{1, 2\}$ and $l \in \{1, \cdots, L\}$ are assumed equal and denoted as $\bar{\gamma}_{U,R}$.

3.7 Conclusion

Two novel relay selection strategies are proposed and analyzed for MIMO AF TWRNs with spatial multiplexing. The first strategy minimizes the overall outage probability of the two user nodes, and it amounts to selecting the relay which maximizes the minimum of the eigenvalues of the corresponding Wishart matrices. The second strategy maximizes the achievable sum rate by selecting the relay which maximizes the minimum of the determinant of the Wishart matrices. The performance of these strategies was studied by deriving the lower/upper bounds of the overall outage probability and average sum rate approximations in closed-form.

Further, achievable diversity order and spatial multiplexing gain were characterized by deriving the fundamental diversity-multiplexing trade. Notably, this trade-off shows that whenever the sum of relay antennas is fixed, then the achievable diversity order always becomes a constant, and hence, the overall multiplexing gain can be improved by equally distributing the antennas among the available set of relays. Several numerical results were presented to compare the performance gains of the proposed relay selection strategies and to validate our analysis. Our results reveal that the proposed relay selection strategies provide substantial improvements in outage probability, diversity order, and the average sum rate. This study thus confirms the potential use of MIMO AF TWRNs in the context of emerging, next generation wireless networks. Moreover, our proposed sum rate selection algorithm (SS2) has also been extended to massive MIMO configurations in the next chapter.

Chapter 4

Relay Selection Strategies with Massive MIMO Two-Way Relay Networks

This chapter analyses the relay selection which is designed to maximise the achievable sum-rate for massive MIMO TWRNs.

4.1 Introduction

In the previous chapter, two relay selection methods for two-way relay networks with spatial multiplexing were developed and analysed. As mentioned in Chapter 1, due to the development of millimetre wave (mmWave) communications, the physical sizes of transceiver antennas have decreased [60]. Thus the deployment of a large number of antennas at the transmitters and receivers will be possible even in user nodes. This deployment results in large scale antenna systems or massive MIMO systems. Both mmWave and massive MIMO are expected to be used in future 5G technologies [9].

As explained in Section 2.5, massive MIMO have following main properties:

- It eliminates the effects of small scale fading and noise at the receivers.
- It eliminates the inter-cell interference.
- It has a very high energy efficiency.

- Transmit power can be scaled inversely proportional to the number of transmit antennas at the base station.

Due to these advantages offered by massive MIMO, appropriate relay selection schemes are important. In this analysis of the system under massive MIMO conditions several interesting results were obtained. The end-to-end SNR of each sub channel in each relay was found to be a deterministic value. The relay selection schemes proposed here, becomes deterministic when large scale antenna systems are used in the user nodes. To maximise the sum-rate of the system, the relay with highest number of antennas have to be chosen. The system is analysed under three power scaling scenarios namely power scaling at user nodes, power scaling at relay nodes and power scaling at both user and relay nodes.

The asymptotic analysis is presented whenever the number of antennas at two user nodes grow without bound for the second relay selection strategy proposed in Chapter 3. As explained in Chapter 3, the number of antennas at relay nodes and the user nodes are constrained as follows:

$$N_1 \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}, \tag{4.1}$$

$$N_2 \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}, \tag{4.2}$$

$$\min(N_1, N_2) \geq \max_{l\in\{1,\cdots,L\}} N_{R_l}. \tag{4.3}$$

Due to this restrictions the number of antennas at the relay nodes can not be increased arbitrarily. Thus the implementation of a large scale antenna systems or massive MIMO should be done at the user nodes.

4.2 System and Channel Model

The system and channel model is same as in Section 3.3.1 in Chapter 3. Yet in this case, the channel matrix from U_i to R_l is defined as $\mathbf{F}_{i,l} = \tilde{\mathbf{F}}_{i,l}\mathbf{D}_{i,l}^{1/2}$, where $\tilde{\mathbf{F}}_{i,l} \sim \mathcal{CN}_{N_{R_l}\times N_i}\left(\mathbf{0}_{N_{R_l}\times N_i}, \mathbf{I}_{N_{R_l}} \otimes \mathbf{I}_{N_i}\right)$ captures the fast fading and $\mathbf{D}_{i,l} = \eta_{i,l}\mathbf{I}_{N_{R_l}}$ accounts for the pathloss.

Also the number of antennas at the user nodes U_1 and U_2, N_1 and N_2 grows unbounded while the number of antennas at the relay node R_l, N_l for $l \in \{1, 2, \cdots L\}$

is kept unchanged. For simplicity it is assumed that as $N_1, N_2 \rightarrow \infty$, the ratio between the number of antennas at U_1 and U_2 is kept constant. Thus

$$\alpha = \frac{N_2}{N_1}. \tag{4.4}$$

This assumption does not affect the system analysis, but only makes the results much simpler. The end-to-end SNR of the kth substream in the lth relay and ith user can also be represented as

$$\gamma_{U_{i,l}^{(k)}} = \frac{P_{R_l} P_{i'}}{P_{R_l}\sigma_{R_l}^2 \mathrm{Tr}\left(\mathcal{E}\left\{\mathbf{V}_{i',l}^{-1}\right\}\right) + \sigma_i^2 \left[\mathbf{V}_{l,i}^{-1}\right]_{k,k} \mathrm{Tr}\left(\mathcal{E}\left\{\mathbf{V}_{i',l}^{-1}\right\}\right)\left(\frac{P_i}{\mathrm{Tr}\left(\mathcal{E}\left\{\mathbf{V}_{i,l}^{-1}\right\}\right)} + \frac{P_{i'}}{\mathrm{Tr}\left(\mathcal{E}\left\{\mathbf{V}_{i',l}^{-1}\right\}\right)} + \sigma_{R_l}^2\right)}, \tag{4.5}$$

where $\mathbf{V}_{l,i} = \mathbf{F}_{l,i}^H \mathbf{F}_{l,i}$, $\mathbf{V}_{i,l} = \mathbf{F}_{i,l} \mathbf{F}_{i,l}^H$, $l \in \{1, \cdots, L\}$, $k \in \{1, \cdots, N_{R_l}\}$, $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$. As the number of antennas at the user nodes asymptotically approaches ∞, we recall the following identity from [47] as

$$\lim_{N_i \rightarrow \infty} \frac{\mathbf{F}_{i,l} \mathbf{F}_{i,l}^H}{N_i} = \lim_{N_i \rightarrow \infty} \frac{\mathbf{V}_{i,l}}{N_i} = \mathbf{D}_{i,l}. \tag{4.6}$$

4.3 Asymptotic end-to-end SNR

In this section the asymptotic end-to-end SNR will be derived for three power scaling methods. It is shown that the end-to-end SNR values are independent from fast fading when massive MIMO is used at the user nodes. And the randomness of the SNR values will disappear and deterministic values can be obtained for end-to-end SNR for each data sub-stream in each relay. This makes relay selections simple.

4.3.1 Power scaling at user nodes

In this case the transmit power at the user nodes U_1 and U_2 are scaled inversely proportional to the number of antennas. Thus

$$P_i = \frac{E_i}{N_i} \quad \text{for} \quad i \in \{1, 2\}, \tag{4.7}$$

where E_1, E_2 and P_{R_l} are fixed. By substituting (4.7) on to (4.5), we obtain

$$\gamma_{U_{i,l}^{(k)}} = \frac{P_{R_l}E_{i'}}{P_{R_l}\sigma_{R_l}^2 \mathrm{Tr}\left(\mathcal{E}\left\{\left(\frac{\mathbf{V}_{i',l}}{N_{i'}}\right)^{-1}\right\}\right) + \frac{\sigma_i^2}{N_i}\left[\left(\frac{\mathbf{V}V_{l,i}}{N_i}\right)^{-1}\right]_{k,k} \mathrm{Tr}\left(\mathcal{E}\left\{\left(\frac{\mathbf{V}_{i',l}}{N_{i'}}\right)^{-1}\right\}\right)} \times \left(\frac{E_i}{\mathrm{Tr}\left(\mathcal{E}\left\{\left(\frac{\mathbf{V}_{i,l}}{N_i}\right)^{-1}\right\}\right)} + \frac{E_{i'}}{\mathrm{Tr}\left(\mathcal{E}\left\{\left(\frac{\mathbf{V}_{i',l}}{N_{i'}}\right)^{-1}\right\}\right)} + \sigma_{R_l}^2\right)^{-1} \tag{4.8}$$

for $l \in \{1, \cdots, L\}$, $k \in \{1, \cdots, N_{R_l}\}$, $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$. By using the asymptotic values given by (4.6), we obtain

$$\gamma_{U_{i,l}^{(k)}}^{\infty} = \frac{E_{i'}\eta_{i',l}}{N_{R_l}\sigma_{R_l}^2}, \tag{4.9}$$

where $l \in \{1, \cdots, L\}$, $k \in \{1, \cdots, N_{R_l}\}$, $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$. Unlike in the previous case with finite number of antennas at the user nodes, the asymptotic end-to-end SNR is a deterministic value.

4.3.2 Power scaling at relay node

In this case the transmit power at the relay nodes R_l for $l \in \{1, \cdots, L\}$ are scaled inversely proportional to the number of antennas at the user nodes. Thus

$$P_{R_l} = \frac{E_{R_l}}{N_i}, \tag{4.10}$$

where P_1, P_2 and E_{R_l} are fixed. By substituting (4.10) to (4.5), and by using the asymptotic values given by (4.6), we obtain

$$\gamma_{U_{i,l}^{(k)}}^{\infty} = \frac{E_{R_l}P_{i'}\eta_{i,l}\eta_{i',l}}{\sigma_i^2\left(P_1\eta_{1,l}/\alpha + P_2\eta_{2,l}\right)}, \tag{4.11}$$

where $l \in \{1, \cdots, L\}$, $k \in \{1, \cdots, N_{R_l}\}$, $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$.

4.3.3 Power scaling at the user nodes and relay node

In this case the transmit power at the user nodes U_1 and U_2 and the transmit power at the relay nodes R_l for $l \in \{1, \cdots, L\}$ are scaled inversely proportional to the number of antennas at the user nodes. Thus

$$P_i = \frac{E_i}{N_i} \quad \text{for} \quad i \in \{1, 2\}, \tag{4.12}$$

$$P_{R_l} = \frac{E_{R_l}}{N_i}, \tag{4.13}$$

where E_1, E_2 and E_{R_l} are fixed. By substituting (4.12) and (4.13) on to (4.5), and by using the asymptotic values given by (4.6), we obtain

$$\gamma^{\infty}_{U^{(k)}_{i,l}} = \frac{E_{R_l} E_{i'} \eta_{i,l} \eta_{i',l}}{\sigma_i^2 (E_i \eta_{i,l} + E_{i'} \eta_{i',l}) + \sigma_{R_l}^2 N_{R_l} (E_{R_l} \eta_{i,l} + \sigma_i^2)}, \tag{4.14}$$

where $l \in \{1, \cdots, L\}$, $k \in \{1, \cdots, N_{R_l}\}$, $i \in \{1, 2\}$, $i' \in \{1, 2\}$, and $i \neq i'$.

Interestingly, the asymptotic SNRs in (4.9), (4.11), and (4.14) are independent of the fast fading component of the wireless channel, and they depend on pathloss of the channels. It is worth noting that these asymptotic SNRs are independent of the data-stream index, k, as well, and hence, we can denote $\gamma^{\infty}_{U^{(k)}_{i,l}} = \gamma^{\infty}_{U_{i,l}}$ for $k \in \{1, \cdots, N_{R_l}\}$.

4.4 Relay selection based on minimizing the overall outage probability (SS-2)

The relay selection criterion for massive MIMO-enabled TWRNs can be written by using (3.29), (4.9), (4.11), and (4.14) as

$$L^* = \underset{l \in \{1, \cdots, L\}}{\operatorname{argmax}} \left[\min \left(\left(\gamma^{\infty}_{U_{1,l}} \right)^{N_{R_l}}, \left(\gamma^{\infty}_{U_{2,l}} \right)^{N_{R_l}} \right) \right], \tag{4.15}$$

where L^* is the index of the optimal relay selected based on maximising the asymptotic sum rate. Since $\gamma^{\infty}_{U_{i,l}}$ is independent of the fast fading, the relay selection becomes significantly simple. Finally, by using (3.26), the average asymptotic sum rate of the best relay selection can be written as

$$\bar{\mathcal{R}}_{L^*} = N_{R_{L^*}} \min \left(\log \left(1 + \gamma^{\infty}_{U_{1,L^*}} \right), \log \left(1 + \gamma^{\infty}_{U_{2,L^*}} \right) \right). \tag{4.16}$$

4.5 Simulation Results

In Fig. 4.1, the asymptotic sum rate of the best relay selection for massive MIMO-enabled TWRNs is investigated. Three sum rate curves are plotted by scaling down the transmit power of user nodes and/or relay inversely proportional to the antenna

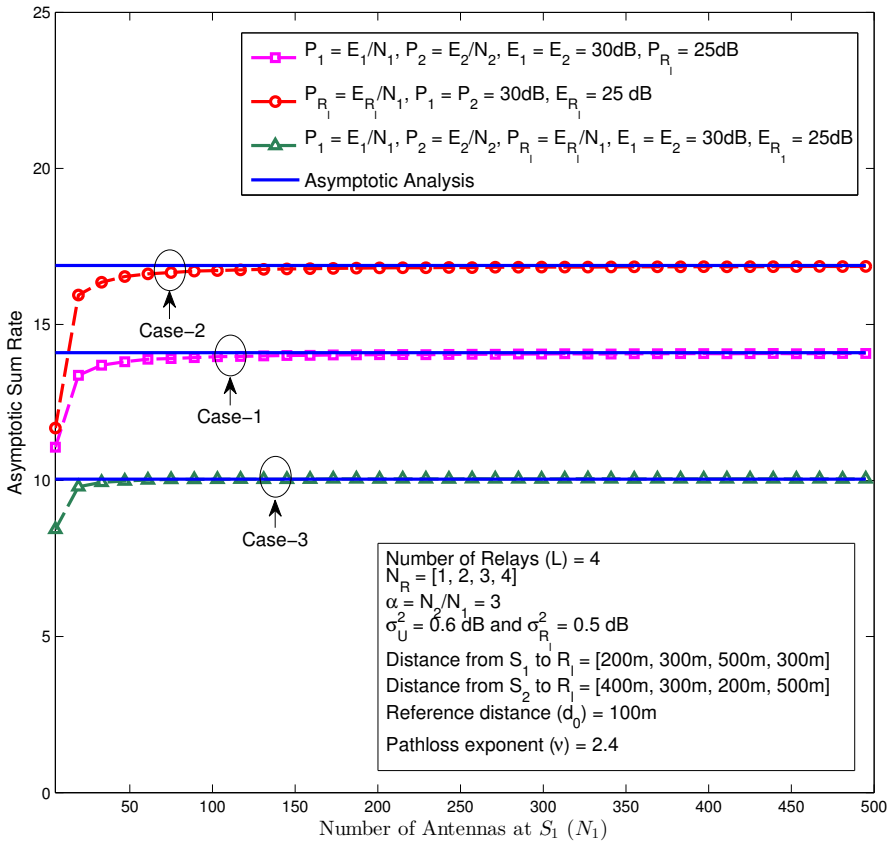


Figure 4.1: The asymptotic sum rate the best relay selection. The pathloss is modelled as $\eta_{i,l} = (d_0/d_{i,l})^{\nu}$, where ν is the pathloss exponent, d_0 is the reference distance, $d_{i,l}$ is the distance between U_i and R_l.

count at the user nodes. Our results reveal that the simulated sum rate curves approaches respective asymptotic sum rates whenever the number of antennas at the user nodes exceeds 100. An asymptotic sum rate gain of 2.7 bits/channel-use/Hz is provided by transmit power scaling only at the relays (case-2) over the transmit power scaling only at the user nodes (case-1). Moreover, the worst asymptotic sum rate is achieved when the transmit power scaling is performed at both user nodes and relays (case-3). For example, case-2 and case-1 provide asymptotic sum rate gains of 6.5 and 4 bits/channel-use/Hz, respectively, over case-3. However, case-3 provides the best energy efficiency because the transmit powers at both the users and relays can be scaled down inversely proportional to number of antennas at user nodes. Thus, massive MIMO-enabled TWRNs provide significant sum rate gains while allowing per bit energy requirement to become infinitesimal.

4.6 Conclusion

In this chapter, the system proposed in Chapter 3 was analysed for the case of large number of antennas at the user nodes by deriving the asymptotic SNR and sum rate. The resulting asymptotic SNR and sum rate expressions are independent of the fast fading effects, and consequently yield low-complexity relay selection and reduced latency in the air interface. Several numerical results compared the performance gains of the proposed relay selection strategy and validated our analysis. Our results reveal that the proposed relay selection strategy provides substantial sum-rate improvements. Further, our asymptotic analysis shows that TWRNs with massive MIMO provide substantial sum-rate gains while allowing the transmit powers at the nodes to become infinitesimal.

Chapter 5

Conclusions and Future Research Directions

5.1 Conclusions

In this thesis MIMO AF TWRNs are analysed with ZF beamforming. The contribution in each chapter is as follows.

Chapter 2 introduces the necessary wireless communication basics. Namely wireless channel models, cooperative relay systems, 5G wireless, MIMO wireless and massive MIMO have been described.

Chapter 3 proposes and analyses two novel relay selection strategies for MIMO AF TWRNs with spatial multiplexing. The first strategy minimizes the overall outage probability of the two user nodes, which amounts to selecting the relay which maximizes the minimum of the eigenvalues of the corresponding Wishart matrices. The second strategy maximizes the achievable sum rate by selecting the relay which maximizes the minimum of the determinant of the Wishart matrices. The performance of these strategies was studied by deriving the lower/upper bounds of the overall outage probability and average sum rate approximations in closed-form. Further, achievable diversity order and spatial multiplexing gain were characterized by deriving the fundamental diversity-multiplexing trade-off. Notably, this trade-off shows that whenever the sum of relay antennas is fixed, then the achievable diversity order always becomes a constant, and hence, the overall multiplexing gain can be improved by equally distributing the antennas among the available set of relays. Several numerical results were presented to compare the performance gains

of the proposed relay selection strategies and to validate our analysis. Our results reveal that the proposed relay selection strategies provide substantial improvements in outage probability, diversity order, and the average sum rate. This study thus confirms the potential use of MIMO AF TWRNs in the context of emerging, next generation wireless networks.

Chapter 4 proposes and analyses the use of massive MIMO for the TWRNs from Chapter 3 by deriving the asymptotic SNR and sum rate. These asymptotic SNR and sum rate expressions are independent of the small scale fading effects, and consequently yielding low-complexity relay selection and reduced latency in the air interface. Several numerical results were presented to validate our analysis and the performance gains of the proposed relay selection strategy. Our results reveal that it provides substantial sum rate improvements. Further, our asymptotic analysis shows that TWRNs with massive MIMO provide substantial sum rate gains while allowing the transmit powers at the nodes to become infinitesimally small.

5.2 Future Research Directions

In this section several future research directions have been proposed for the continuation of this research.

- **Analysing channel correlation among the sub-channels**

  This research made the assumption that all the channel amplitudes are independently distributed. However due to practical limitations such as the closeness of multiple antennas at the relay and the user nodes, this assumption may not be strictly true, i.e. the channel amplitudes will be correlated. Thus further research can investigate the impact of correlations.

- **Analysing imperfect CSI at the user nodes**

  This research assumed that the channel state information(CSI) is freely available at the user nodes to perform beamforming. However this assumption is unrealistic in a practical scenario and methods to acquire channel state information should be used at the user nodes. This results in imperfect chan-

nel state information at the user nodes due to practical limitations of such methods. Thus the impact of imperfect channel state information at the user nodes should be analysed for the proposed relay selection schemes. This is an essential research for the practical implementation of this relay selection schemes.

- **Effect of pilot contamination for the massive MIMO user nodes**

  Massive MIMO user nodes was analysed in Chapter 4. However this analysis ignored the effect of pilot contamination. Pilot contamination is the residual interference caused by the reuse of non-orthogonal pilot sequences in adjacent cells. It has been identified as the main performance limiting factor in massive MIMO systems [47, 50]. The effect of pilot contamination can not be mitigated by increasing the number of antennas. Thus its impact must be assessed properly before implementing massive MIMO on the proposed TWRN system. As massive MIMO holds great promise for 5G systems, this will be a great problem to analyse in future research.

Bibliography

[1] [Online]. Available: http://en.wikipedia.org/wiki/Spectrum_management

[2] Qualcomm. (2011) the visible light communications motivation. [Online]. Available: http://visiblelightcomm.com/the-visible-light-communications-motivation/

[3] I. Gradshteyn and I. Ryzhik, *Table of integrals, Series, and Products*, 7th ed. Academic Press, 2007.

[4] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill, Inc., NY, 2002.

[5] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. FL: Academic Press, Inc, 1980.

[6] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1970.

[7] (2015) Home-buying wish list: 70s appliances out, good mobile signal in. [Online]. Available: http://www.rootmetrics.com/us/blog/articles/home-buying-wish-list-70s-appliances-out-good-mobile-signal-in

[8] (2013) Cisco visual networking index: Global mobile data traffic forecast update, 20132018. CISCO. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[9] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[10] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[11] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[12] G. Amarasuriya, C. Tellambura, and M. Ardakani, "Two-way amplify-and-forward multiple-input multiple-output relay networks with antenna selection," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1513–1529, Sep. 2012.

[13] ——, "Joint beamforming and antenna selection for two-way amplify-and-forward MIMO relay networks," in *IEEE Int. Conf. on Commun. (ICC)*, Jun. 2012, pp. 4829–4834.

[14] K. Loa, C.-C. Wu, S.-T. Sheu, Y. Yuan, M. Chion, D. Huo, and L. Xu, "IMT-advanced relay standards [WiMax/LTE update]," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 40–48, Aug. 2010.

[15] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMax and LTE-advanced mobile systems," *IEEE Commun. Mag.*, vol. 47, no. 10, pp. 100–105, Oct. 2009.

[16] A. Sulyman, A. Nassar, M. Samimi, G. MacCartney, T. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 78–86, Sep. 2014.

[17] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[18] *Urban transmission loss models for mobile radio in the 900 and 1800 MHz bands*, European Cooperative in the Field of Science and Technical Research EURO-COST 231 Std., Rev. 2, Sep. 1991.

[19] E. O. T. Okumura and K. Fukuda, "Field strength and its variability in vhf and uhf land mobile service," *Review Electrical Communication Laboratory*, vol. 16, no. 9-10, pp. 825–873, Sept.-Oct. 1968.

[20] M. Hatay, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.

[21] G. L. Stüber, *Principles of Mobile Communication*, third edition ed. Springer, 2012.

[22] M. Nakagami, "The m distribution; a general formula of intensity distribution of rapid fading," *Statistical Methods in Radio Wave Propagation, W.G. Hoffman, ed.*, pp. 3–36, 1960.

[23] S. Atapattu, C. Tellambura, and H. Jiang, "A mixture gamma distribution to model the snr of wireless channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4193–4203, Dec. 2011.

[24] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[25] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16j relay-based wireless access networks: an overview," *IEEE Wireless Commun. Mag.*, vol. 15, no. 5, pp. 56–63, Oct. 2008.

[26] J. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[27] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[28] Y. Han, S. H. Ting, C. K. Ho, and W. H. Chin, "Performance bounds for two-way amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 432–439, Jan. 2009.

[29] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[30] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1429–1439, Oct. 2008.

[31] A. Sadek, Z. Han, and K. Liu, "A distributed relay-assignment algorithm for cooperative communications in wireless networks," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 4, Jun. 2006, pp. 1592–1597.

[32] V. Sreng, H. Yanikomeroglu, and D. Falconer, "Relayer selection strategies in cellular networks with peer-to-peer relaying," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, vol. 3, Oct. 2003, pp. 1949–1953 Vol.3.

[33] S. Song and K. Letaief, "System design, DMT analysis, and penalty for non-coherent relaying," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2489–2498, Sep. 2012.

[34] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1414–1423, Mar. 2009.

[35] S. Atapattu, Y. Jing, H. Jiang, and C. Tellambura, "Relay selection schemes and performance analysis approximations for two-way networks," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 987–998, Mar. 2013.

[36] "5G use cases and requirements," Nokia Networks, Tech. Rep., 2014.

[37] "5G radio access," Ericsson, Tech. Rep. Uen 284 23-3204 Rev B, Feb. 2015. [Online]. Available: http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf

[38] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.

[39] "5G: A technology vision," Huawei Technologies Co., LTD, Tech. Rep., 2013.

[40] *3GPP TS36.213: "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures"*, Std., Rev. version 8.5.0.

[41] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[42] Z. Wang and G. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[43] Y. Dhungana and C. Tellambura, "New simple approximations for error probability and outage in fading," *IEEE Wireless Commun. Lett.*, vol. 16, no. 11, pp. 1760–1763, Nov. 2012.

[44] ——, "Uniform approximations for wireless performance in fading channels," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4768–4779, Nov. 2013.

[45] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.

[46] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[47] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[48] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[49] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[50] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell tdd systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[51] T. Lee, H.-S. Kim, S. Park, and S. Bahk, "Mitigation of sounding pilot contamination in massive MIMO systems," in *Communications (ICC), 2014 IEEE International Conference on*, Jun. 2014, pp. 1191–1196.

[52] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Sep. 2011, pp. 545–550.

[53] G. Amarasuriya, C. Tellambura, and M. Ardakani, "Performance analysis of zero-forcing for two-way MIMO AF relay networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 2, pp. 53–56, Apr. 2012.

[54] Y. Li, R. H. Y. Louie, and B. Vucetic, "Relay selection with network coding in two-way relay channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4489–4499, 2010.

[55] H. Cui, M. Ma, L. Song, and B. Jiao, "Relay selection for two-way full duplex relay networks with amplify-and-forward protocol," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3768–3777, Jul. 2014.

[56] I. Krikidis, "Relay selection for two-way relay channels with MABC DF: A diversity perspective," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4620–4628, Nov. 2010.

[57] Q. Zhou, Y. Li, F. Lau, and B. Vucetic, "Decode-and-forward two-way relaying with network coding and opportunistic relay selection," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3070–3076, Nov. 2010.

[58] T. Oechtering and H. Boche, "Bidirectional regenerative half-duplex relaying using relay selection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1879–1888, May 2008.

[59] L. Song, G. Hong, B. Jiao, and M. Debbah, "Joint relay selection and analog network coding using differential modulation in two-way relay channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2932–2939, Jul. 2010.

[60] A. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[61] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan, "The evolution of cellular backhaul technologies: Current issues and future trends," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 97–113, First 2011.

[62] H. Raza, "A brief survey of radio access network backhaul evolution: part i," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 164–171, Jun. 2011.

[63] ——, "A brief survey of radio access network backhaul evolution: part ii," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 170–177, May 2013.

[64] R. H. Louie, Y. Li, and B. Vucetic, "Practical physical layer network coding for two-way relay channels: performance analysis and comparison," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 764–777, Feb. 2010.

[65] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*, 1st ed. Cambridge, UK ; New York, NY : Cambridge University Press, 2003.

[66] J. Heath, R.W., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[67] M. Antonia and S. Verdu, *Random Matrix Theory and Wireless Communications*, 1st ed., ser. Foundations and Trends in Communications and Information Theory. Now publishers inc, 2004, no. 1.

[68] L. G. Ordóñez, "Performance limits of spatial multiplexing MIMO systems," Ph.D. dissertation, Technical University of Catalonia (UPC), Barcelona, Spain, Mar. 2009.

[69] D. Gore, R. W. Heath, and A. Paulraj, "On performance of the zero forcing receiver in presence of transmit correlation," in *IEEE Intl. Symp. on Inf. Theory (ISIT)*, Lausanne, Switzerland, Jul. 2002, pp. 1–2.

[70] E. F. Beckenbach and R. E. Bellman, *Inequalities*, 1st ed. Springer Verlag, 1965.

[71] A. R. Amir-Moez and G. E. Johnston, "On the product of diagonal elements of a positive matrix," *Mathematics Magazine of Mathematical Association of America*, vol. 42, no. 1, pp. 24–26, Jan. 1969.

[72] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, Sep. 2004.

[73] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[74] A. Annamalai and C. Tellambura, "Error rates for Nakagami-m fading multichannel reception of binary and M-ary signals," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 58–68, Jan. 2001.

[75] G. Amarasuriya, C. Tellambura, and M. Ardakani, "Transmit antenna selection strategies for cooperative MIMO AF relay networks," in *IEEE Global Telecommun. Conf. (GLOBECOM 2010)*, Dec. 2010, pp. 1–5.

Appendix A

Appendices for Chapter 3

A.1 Proof of the CDFs of $\gamma_{U_{i,l}^{min}}^{\mathrm{ub}}$ and $\gamma_{U_{i,l}^{min}}^{\mathrm{lb}}$

In this appendix, the proofs of the CDFs of $\gamma_{U_{i,l}^{min}}^{\mathrm{ub}}$ and $\gamma_{U_{i,l}^{min}}^{\mathrm{lb}}$ are sketched. To this end, the corresponding SNR can be re-written in an alternative form as follows:

$$\gamma_{U_{i,l}^{min}}^{\mathrm{xb}} = \frac{\alpha_{i,l} X_{i,l}}{\beta_{i,l} X_{i,l} + \zeta_{i,l}}, \tag{A.1}$$

where $X_{i,l} = \left(\left[\left(\mathbf{F}_{l,i}^H \mathbf{F}_{l,i} \right)^{-1} \right]_{k,k} \right)^{-1}$ for $\gamma_{U_{i,l}^{min}}^{\mathrm{ub}}$ and $X_{i,l} = \lambda_{min} \left\{ \mathbf{F}_{l,i}^H \mathbf{F}_{l,i} \right\}$ for $\gamma_{U_{i,l}^{min}}^{\mathrm{lb}}$. The CDF of $\gamma_{U_{i,l}^{min}}^{\mathrm{xb}}$ can then be written as [53]

$$F_{\gamma_{U_{i,l}^{min}}^{\mathrm{xb}}}(z) = \Pr\left(X_{i,l} \leq \frac{\zeta_{i,l} z}{\alpha_{i,l} - \beta_{i,l} z} \right) = F_{X_{i,l}}\left(\frac{\zeta_{i,l} z}{\alpha_{i,l} - \beta_{i,l} z} \right), \tag{A.2}$$

for $z < \frac{\alpha_{i,l}}{\beta_{i,l}}$. Next, $F_{\gamma_{U_{i,l}^{min}}^{\mathrm{ub}}}(z)$ and $F_{\gamma_{U_{i,l}^{min}}^{\mathrm{lb}}}(z)$ can be derived by evaluating the CDFs of the kth diagonal element of the inverse Wishart matrix [69] and the smallest eigenvalue of the Wishart matrix [68] at $\zeta_{i,l} z / (\alpha_{i,l} - \beta_{i,l} z)$.

A.2 Proof of the overall outage probability bounds

In this appendix, the proof of the overall outage probability bounds is sketched. To begin with, the definition of the overall outage upper bound is re-written as

$$P_{\mathrm{out}}^{\mathrm{ub}} = \Pr\left[\gamma_{\mathrm{eq}}^{\mathrm{lb}} \leq \gamma_{th} \right] = F_{\gamma_{\mathrm{eq}}^{\mathrm{lb}}}(\gamma_{th}), \tag{A.3}$$

where $\gamma_{\mathrm{eq}}^{\mathrm{lb}}$ is the equivalent SNR and is defined as

$$\gamma_{\mathrm{eq}}^{\mathrm{lb}} = \max_{l \in \{1, \cdots, L\}} \left(\gamma_l^{\mathbf{lb}} \right), \quad \text{where} \quad \gamma_l^{\mathbf{lb}} = \min\left(\gamma_{U_{1,l}^{min}}^{\mathrm{lb}}, \gamma_{U_{2,l}^{min}}^{\mathrm{lb}} \right), \tag{A.4}$$

for $l \in \{1, \cdots, L\}$. Next, the CDF of $\gamma_l^{\mathbf{lb}}$ can be derived as [4]

$$F_{\gamma_l^{\mathbf{lb}}}(x) = 1 - \left(1 - F_{\gamma^{\mathrm{lb}}_{U^{min}_{1,l}}}(x)\right)\left(1 - F_{\gamma^{\mathrm{lb}}_{U^{min}_{2,l}}}(x)\right). \tag{A.5}$$

Then, the CDF of $\gamma_{\mathrm{eq}}^{\mathrm{lb}}$ can be derived as follows [4]:

$$F_{\gamma_{\mathrm{eq}}^{\mathrm{lb}}}(x) = \Pr\left(\gamma_1^{\mathbf{lb}} \leq x, \gamma_2^{\mathbf{lb}} \leq x, \cdots, \gamma_L^{\mathbf{lb}} \leq x\right) = \prod_{l=1}^{L}\left(F_{\gamma_l^{\mathbf{lb}}}(x)\right). \tag{A.6}$$

Finally, by substituting (A.5) into (A.4), the upper bounds of the overall outage probability can be derived as shown in (3.35). The derivation of the lower bound of the overall outage probability follows the same techniques in (A.4), (A.5), and (A.6), and hence its proof is omitted.

A.3 Proof of the asymptotic high SNR outage approximations

In this appendix, the proof of the asymptotic outage lower bound is first sketched, and thereby, the asymptotic outage upper bound is deduced: To begin with, the PDF of $\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}$ for $i \in \{1, 2\}$ is derived by differentiating (3.22) by using the Leibniz integral rule as follows:

$$\begin{aligned} f_{\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}}(x) &= \frac{d}{\mathrm{d}x}\left[\frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}\right]\left(\frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}\right)^{N_i-N_{R_l}} \frac{\mathrm{e}^{-\frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}}}{\Gamma(N_i-N_{R_l}+1)} \quad \text{for} \quad 0 \leq x < \frac{\alpha_{i,l}}{\beta_{i,l}} \\ &= \frac{\beta_{i,l}(\zeta_{i,l})^{N_i-N_{R_l}+1}x^{N_i-N_{R_l}}\mathrm{e}^{-\frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}}}{\Gamma(N_i-N_{R_l}+1)(\alpha_{i,l}-\beta_{i,l}x)^{N_i-N_{R_l}+2}}, \quad \text{for} \quad 0 \leq x < \frac{\alpha_{i,l}}{\beta_{i,l}}. \end{aligned} \tag{A.7}$$

By substituting $\zeta_{i,l}$, $\alpha_{i,l}$, $\beta_{i,l}$, defined in (3.18) into (A.7), and then by taking the Taylor series expansion around $x=0$, the first order expansion[1] of $f_{\gamma^{\mathrm{ub}}_{S_i,\min}}(x)$ when $x \to 0$ can be derived as

$$f^{x\to 0}_{\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}}(x) = \frac{\phi_i^{N_i-N_{R_l}+1}x^{N_i-N_{R_l}}}{\Gamma(N_i-N_{R_l}+1)\left(C_l\bar{\gamma}_{U,R}\right)^{N_i-N_{R_l}+1}} + o\left(x^{N_i-N_{R_l}+1}\right), \tag{A.8}$$

[1]The first order expansion of $f(x)$ is the single-term polynomial approximation of $f(x)$ consisting the lowest power of x [73].

for $i \in \{1,2\}$ where $\bar{\gamma}_{1,l} = \bar{\gamma}_{2,l} = \bar{\gamma}_{U,R_l}$, $\bar{\gamma}_{l,1} = \bar{\gamma}_{l,2} = \bar{\gamma}_{R_l,U}$, $\bar{\gamma}_{R_l,U} = \nu_l \bar{\gamma}_{U,R_l}$, and $\bar{\gamma}_{U,R_l} = C_l \bar{\gamma}_{U,R}$. Further, $\phi_i = (N_i + N_{i'} - 2N_{R_l})/(N_{i'} - N_{R_l})$ for $i \in \{1,2\}$, $i' \in \{1,2\}$, $i \neq i'$. The first order expansion of the CDF of $\gamma^{\text{ub}}_{U^{min}_{i,l}}$ when $x \to 0$ can be derived by using (A.8) as follows:

$$F^{x\to 0}_{\gamma^{\text{ub}}_{U^{min}_{i,l}}}(x) = \frac{\phi_i^{N_i - N_{R_l}+1}}{\Gamma(N_i - N_{R_l}+2) C_l^{N_i - N_{R_l}+1}} \left(\frac{x}{\bar{\gamma}_{U,R}}\right)^{N_i - N_{R_l}+1} + o\left(x^{N_i - N_{R_l}+2}\right), \quad \text{(A.9)}$$

for $i \in \{1,2\}$. Next, the CDF of $\gamma_l^{\text{ub}} = min\left(\gamma^{\text{ub}}_{U^{min}_{1,l}}, \gamma^{\text{ub}}_{U^{min}_{2,l}}\right)$ can be written as

$$F_{\gamma_l^{\text{ub}}}(x) = \sum_{i=1}^{2} F_{\gamma^{\text{ub}}_{U^{min}_{i,l}}}(x) - \prod_{i=1}^{2} F_{\gamma^{\text{ub}}_{U^{min}_{i,l}}}(x). \quad \text{(A.10)}$$

Then, a polynomial approximation of $F_{\gamma_l^{\text{ub}}}(x)$ can be derived by substituting (A.9) into (A.10) as

$$\begin{aligned} F^{x\to 0}_{\gamma_l^{\text{ub}}}(x) &= \frac{\phi_1^{N_1 - N_{R_l}+1}}{\Gamma(N_1 - N_{R_l}+2) C_l^{N_1 - N_{R_l}+1}} \left(\frac{x}{\bar{\gamma}_{U,R}}\right)^{N_1 - N_{R_l}+1} \\ &+ \frac{\phi_2^{N_1 - N_{R_l}+1}}{\Gamma(N_2 - N_{R_l}+2) C_l^{N_2 - N_{R_l}+1}} \left(\frac{x}{\bar{\gamma}_{U,R}}\right)^{N_2 - N_{R_l}+1} \\ &+ \frac{\phi_1^{N_1 - N_{R_l}+1} \phi_2^{N_2 - N_{R_l}+1}}{\Gamma(N_1 - N_{R_l}+2)\Gamma(N_2 - N_{R_l}+2) C_l^{(N_1 - N_{R_l}+1)(N_2 - N_{R_l}+1)}} \\ &\quad \times \left(\frac{x}{\bar{\gamma}_{U,R}}\right)^{(N_1 - N_{R_l}+1)(N_2 - N_{R_l}+1)} \\ &+ o\left(x^{(N_1 - N_{R_l}+1)(N_2 - N_{R_l}+1)}\right). \end{aligned} \quad \text{(A.11)}$$

By collecting the lowest powers of x in (A.11), the first order expansion of $F_{\gamma_l^{\text{ub}}}(x)$ can be derived as

$$F^{x\to 0}_{\gamma_l^{\text{ub}}}(x) = \Theta_l \left(\frac{x}{\bar{\gamma}_{U,R}}\right)^{min(N_1,N_2) - N_{R_l}+1} + o\left(\bar{\gamma}_{U,R}^{-(min(N_1,N_2) - N_{R_l}+2)}\right), \quad \text{(A.12)}$$

where Θ_l is defined in (3.44). The first order expansion of the CDF of $\gamma^{\text{ub}}_{\text{eq}} = \max_{l\in\{1,\cdots,L\}}\left(\gamma_l^{\text{lb}}\right)$ can then be derived by first substituting (A.12) into $F_{\gamma^{\text{ub}}_{\text{eq}}}(x) = \prod_{l=1}^{L}\left(F_{\gamma_l^{\text{lb}}}(x)\right)$ and then by taking the lowest power of x as shown in (3.42).

The asymptotic outage upper bound can be derived by substituting the first order expansion of the CDF of γ_l^{lb} in [53, Eqn. (35)] into (A.6) and then taking the lowest power of x as shown in (3.36).

A.4 Proof of the average sum rate for relays with different antenna array sizes

In this appendix, the proof of the average sum rate of SS-2 is sketched when the relays are equipped with different number of antennas. To begin with, we recall the definition of the average sum rate approximation as follows:

$$\bar{\mathcal{R}}_{L^*} \approx \mathcal{E}_{\gamma_{\rm eq}}\left\{\left(\max_{l\in\{1,\cdots,L\}}(N_{R_l})\right)\log\left(1+\gamma_{\rm eq}\right)\right\}, \tag{A.13}$$

where $\gamma_{\rm eq}$ is the equivalent SNR and is given by

$$\gamma_{\rm eq} = \min\left(\gamma_{U^{min}_{1,L^*}}, \gamma_{U^{min}_{2,L^*}}\right). \tag{A.14}$$

Next, an approximation of the CDF of $\gamma_{\rm eq}$ can be derived by using (3.22) as follows:

$$F_{\gamma_{\rm eq}}(x) \approx 1 - \prod_{i=1}^{2}\left(\frac{\Gamma\left(N_i - N_R^{max}+1, \frac{\zeta_{i,L^*}x}{\alpha_{i,L^*}-\beta_{i,L^*}x}\right)}{\Gamma(N_i - N_R^{max}+1)}\right), \tag{A.15}$$

for $x \leq \min\left(\frac{\alpha_{2,L^*}}{\beta_{2,L^*}}, \frac{\alpha_{2,L^*}}{\beta_{2,L^*}}\right)$. Further, $F_{\gamma_{\rm eq}}(x) = 1$ for $x > \min\left(\alpha_{1,L^*}/\beta_{1,L^*}, \alpha_{2,L^*}/\beta_{2,L^*}\right)$. Next, by using [3, Eq. (8.352.2)], (A.15) can be further expanded as

$$\begin{aligned} F_{\gamma_{\rm eq}}(x) \;&\approx\; 1 - \exp\left(-\frac{\zeta_{1,L^*}x}{\alpha_{1,L^*}-\beta_{1,L^*}x} - \frac{\zeta_{2,L^*}x}{\alpha_{2,L^*}-\beta_{2,L^*}x}\right) \\ &\quad\times \sum_{m=1}^{N_1-N_R^{max}}\sum_{n=0}^{N_2-N_R^{max}} \frac{1}{m!\,n!}\left(\frac{\zeta_{1,L^*}x}{\alpha_{1,L^*}-\beta_{1,L^*}x}\right)^m\left(\frac{\zeta_{2,L^*}x}{\alpha_{2,L^*}-\beta_{2,L^*}x}\right)^n. \end{aligned} \tag{A.16}$$

The CDF in (A.16) can further be simplified, whenever both the user nodes are equipped with the same number of antennas (i.e., $N_1 = N_2 = N$), and all transmit and noise powers at each relay are the same, as follows:

$$F_{\gamma_{\rm eq}}(x) \approx 1 - \exp\left(-\frac{2\zeta_{L^*}x}{\alpha_{L^*}-\beta_{L^*}x}\right)\sum_{m=0}^{N-N_R^{max}}\sum_{n=0}^{N-N_R^{max}}\frac{1}{m!\,n!}\left(\frac{\zeta_{L^*}x}{\alpha_{L^*}-\beta_{L^*}x}\right)^{m+n}, \tag{A.17}$$

where $\alpha_{L^*} = \alpha_{i,L^*}$, $\beta_{L^*} = \beta_{i,L^*}$, and $\zeta_{L^*} = \zeta_{i,L^*}$ for $i \in \{1,2\}$, and $x < \alpha_{L^*}/\beta_{L^*}$. By differentiating (A.17), an approximation of the PDF of $\gamma_{\rm eq}$ can be derived as follows:

$$\begin{aligned} f_{\gamma_{\rm eq}}(x) \;&\approx\; \sum_{m=0}^{N-N_R^{max}}\sum_{n=0}^{N-N_R^{max}}\frac{\alpha_{L^*}\zeta_{L^*}^{m+n}x^{m+n-1}}{m!\,n!(\alpha_{L^*}-\beta_{L^*}x)^{m+n+1}}\left(\frac{2\zeta_{L^*}x}{\alpha_{L^*}-\beta_{L^*}x} - (m+n)\right) \\ &\quad\times\; \exp\left(-\frac{2\zeta_{L^*}x}{\alpha_{L^*}-\beta_{L^*}x}\right), \end{aligned} \tag{A.18}$$

where $x \leq \alpha_{L^*}/\beta_{L^*}$. Further, $f_{\gamma_{\rm eq}}(x) = 0$ for $x \geq \alpha_{L^*}/\beta_{L^*}$. Next, an approximation of the ergodic sum rate of SS-2 can be derived by averaging the sum rate in (3.61) over the PDF of $\gamma_{\rm eq}$ in (A.18) as

$$\bar{\mathcal{R}}_{L^*} \approx \mathcal{E}\{\mathcal{R}_{L^*}\} = \frac{N_R^{max}}{\ln(2)} \int_0^{\infty} \ln(1+x) f_{\gamma_{\rm eq}}(x)\, \mathrm{d}x. \tag{A.19}$$

By substituting (A.18) into (A.19), the sum rate lower bound can be written in an integral form as

$$\bar{\mathcal{R}}_{L^*} \approx \frac{N_R^{max}}{\ln(2)} \sum_{m=0}^{N-N_R^{max}} \sum_{n=0}^{N-N_R^{max}} \frac{1}{m!\, n!} \left(\mathbb{I}_1 - \mathbb{I}_2\right), \tag{A.20a}$$

where $\mathbb{I}_1$ and $\mathbb{I}_2$ can be defined as follows:

$$\mathbb{I}_1 = 2\alpha_{L^*}\zeta_{L^*}^{m+n+1} \int_0^{\frac{\alpha_{L^*}}{\beta_{L^*}}} \frac{x^{m+n}}{(\alpha_{L^*} - \beta_{L^*}x)^{m+n+2}} \exp\left(-\frac{2\zeta_{L^*}x}{\alpha_{L^*} - \beta_{L^*}x}\right) \times \ln(1+x)\, \mathrm{d}x, \tag{A.20b}$$

$$\mathbb{I}_2 = (m+n)\alpha_{L^*}\zeta_{L^*}^{m+n} \int_0^{\frac{\alpha_{L^*}}{\beta_{L^*}}} \frac{x^{m+n-1}}{(\alpha_{L^*} - \beta_{L^*}x)^{m+n+1}} \exp\left(-\frac{2\zeta_{L^*}x}{\alpha_{L^*} - \beta_{L^*}x}\right) \times \ln(1+x) \mathrm{d}x. \tag{A.20c}$$

By substituting the dummy variable $t = 2\zeta_{L^*}x/(\alpha_{L^*} - \beta_{L^*}x)$ into (A.20b) and (A.20c), the integrals $\mathbb{I}_1$ and $\mathbb{I}_2$ can be simplified as

$$\mathbb{I}_1 = \frac{1}{2^{m+n}} \int_0^{\infty} t^{m+n} \exp(-t) \ln\left(\frac{2\zeta_{L^*} + (\alpha_{L^*} + \beta_{L^*})t}{2\zeta_{L^*} + \beta_{L^*}t}\right) \mathrm{d}t, \tag{A.21a}$$

$$\mathbb{I}_2 = \frac{m+n}{2^{m+n}} \int_0^{\infty} t^{m+n-1} \exp(-t) \ln\left(\frac{2\zeta_{L^*} + (\alpha_{L^*} + \beta_{L^*})t}{2\zeta_{L^*} + \beta_{L^*}t}\right) \mathrm{d}t. \tag{A.21b}$$

Next, $\mathbb{I}_1$ and $\mathbb{I}_2$ in (A.21a) and (A.21b), respectively, can be solved in closed-form as follows:

$$\mathbb{I}_1 = \frac{1}{2^{m+n}} \left(\mathbb{J}\left(m+n, 2\zeta_{L^*}, \alpha_{L^*} + \beta_{L^*}\right) - \mathbb{J}\left(m+n, 2\zeta_{L^*}, \beta_{L^*}\right)\right), \tag{A.22a}$$

$$\mathbb{I}_2 = \frac{m+n}{2^{m+n}} \left(\mathbb{J}\left(m+n-1, 2\zeta_{L^*}, \alpha_{L^*} + \beta_{L^*}\right) - \mathbb{J}_2\left(m+n-1, 2\zeta_{L^*}, \beta_{L^*}\right)\right). \tag{A.22b}$$

where the function $\mathbb{J}(x, y, z)$ is defined in (3.64). By substituting (A.22a) and (A.22b) into (A.20a), an approximation of the ergodic sum rate of SS-2 can be derived in closed-form as in (3.63).

A.5 Proof of the average sum rate for relays with same antenna array size

In this appendix, the proof of the average sum rate approximation for the case of equal number of antennas at each relay is sketched. To begin with, the effective SNR (A.14) can be alternatively approximated as follows:

$$\gamma_{\mathrm{eq}} \approx \gamma_{\mathrm{eq}_1} = \max_{l\in\{1,\cdots,L\}} \left(\min\left(\gamma^{\mathrm{ub}}_{U^{min}_{1,l}}, \gamma^{\mathrm{ub}}_{U^{min}_{2,l}}\right)\right), \tag{A.23}$$

where $\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}$ for $i \in \{1,2\}$ is defined in (3.21). The CDF of γ_{eq} can be derived as follows:

$$F_{\gamma_{\mathrm{eq}_1}}(x) = \prod_{l=1}^{L}\left(1 - \bar{F}_{\gamma^{\mathrm{ub}}_{U^{min}_{1,l}}}(x)\,\bar{F}_{\gamma^{\mathrm{ub}}_{U^{min}_{2,l}}}(x)\right). \tag{A.24}$$

where $\bar{F}_{\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}}(x)$ is the complimentary CDF of $\gamma^{\mathrm{ub}}_{U^{min}_{i,l}}$. By using (3.22), $F_{\gamma_{\mathrm{eq}}}(x)$ in (A.24) is expanded as

$$F_{\gamma_{\mathrm{eq}_1}}(x) = \prod_{l=1}^{L}\left(1 - \prod_{i=1}^{2}\left(\frac{\Gamma\left(N_i - N_R + 1, \frac{\zeta_{i,l}x}{\alpha_{i,l}-\beta_{i,l}x}\right)}{\Gamma(N_i - N_R + 1)}\right)\right). \tag{A.25}$$

Next, by using [3, Eq. (8.352.2)], (A.25) can be further expanded as

$$\begin{aligned} F_{\gamma_{\mathrm{eq}_1}}(x) \;=\; & \prod_{l=1}^{L}\left(1-\exp\left(-\frac{\zeta_{1,l}x}{\alpha_{1,l}-\beta_{1,l}x} - \frac{\zeta_{2,l}x}{\alpha_{2,l}-\beta_{2,l}x}\right)\sum_{m=1}^{N_1-N_R}\sum_{n=0}^{N_2-N_R}\frac{1}{m!\;n!}\right. \\ & \times \left.\left(\frac{\zeta_{1,l}x}{\alpha_{1,l}-\beta_{1,l}x}\right)^{m}\left(\frac{\zeta_{2,l}x}{\alpha_{2,l}-\beta_{2,l}x}\right)^{n}\right), \end{aligned} \tag{A.26}$$

where $x < \min\left(\alpha_{1,l}/\beta_{1,l}, \alpha_{2,l}/\beta_{2,l}\right)$, and $F_{\gamma_{\mathrm{eq}}}(x) = 1$ for $x \geq \min\left(\alpha_{1,l}/\beta_{1,l}, \alpha_{2,l}/\beta_{2,l}\right)$. The CDF in (A.26) can further be simplified whenever both the user nodes are equipped with the same number of antennas (i.e., $N_1 = N_2 = N$) as follows:

$$F_{\gamma_{\mathrm{eq}_1}}(x) = \prod_{l=1}^{L}\left(1-\exp\left(-\frac{2\zeta_l x}{\alpha_l-\beta_l x}\right)\sum_{m=1}^{N-N_R}\sum_{n=0}^{N-N_R}\frac{1}{m!\;n!}\left(\frac{\zeta_l x}{\alpha_l-\beta_l x}\right)^{m+n}\right), \tag{A.27}$$

where $x < \alpha_l/\beta_l$, $\alpha_l = \alpha_{1,l} = \alpha_{2,l}$, $\beta_l = \beta_{1,l} = \beta_{2,l}$, and $F_{\gamma_{\mathrm{eq}_1}}(x) = 1$ for $x \geq \alpha_l/\beta_l$. The close-form derivation of the ergodic sum rate of SS-2 by using (A.27) appears mathematical intractable. Nevertheless, in the case of all transmit and noise powers at all relays are the same, (A.27) can be written as

$$F_{\gamma_{\mathrm{eq}_1}}(x)=\left(1-\exp\left(-\frac{2\zeta x}{\alpha-\beta x}\right)\sum_{m=0}^{N-N_R}\sum_{n=0}^{N-N_R}\frac{1}{m!\,n!}\left(\frac{\zeta x}{\alpha-\beta x}\right)^{m+n}\right)^{L}, \quad \text{(A.28)}$$

where $\alpha_l = \alpha$, $\zeta_l = \zeta$, $\beta_l = \beta$ for $l \in \{1, 2, \ldots L\}$, and $x < \alpha/\beta$. By using the binomial expansion, (A.17) can be further expanded as follows:

$$F_{\gamma_{\mathrm{eq}_1}}(x)=\sum_{l=0}^{L}(-1)^l\binom{L}{l}\exp\left(-\frac{2l\zeta x}{\alpha-\beta x}\right)\left(\sum_{m=0}^{N-N_R}\frac{1}{m!}\left(\frac{\zeta x}{\alpha-\beta x}\right)^{m}\right)^{2l}. \quad \text{(A.29)}$$

Next, by using [74, Eqn. (44)] and [75, Eqn. (6)], the CDF of γ_{eq_1} can be finally written as

$$F_{\gamma_{\mathrm{eq}_1}}(x)=\sum_{l=0}^{L}\sum_{k=0}^{2l(N-N_R)}(-1)^l\binom{L}{l}\beta_{k,2l,N-N_R+1}\left(\frac{\zeta x}{\alpha-\beta x}\right)^{k}\exp\left(-\frac{2l\zeta x}{\alpha-\beta x}\right), \quad \text{(A.30)}$$

for $x < \frac{\alpha}{\beta}$ where the multinomial coefficient $\beta_{k,2l,N-N_R+1}$ is defined in (3.66). By differentiating (A.30), the PDF of γ_{eq_1} can be derived as follows:

$$f_{\gamma_{\mathrm{eq}_1}}(x)=\sum_{l=1}^{L}\sum_{k=0}^{2l(N-N_R)}(-1)^l\binom{L}{l}\beta_{k,2l,N-N_R+1}\left(k-\frac{2l\zeta x}{\alpha-\beta x}\right)\times\frac{\alpha\zeta^k x^{k-1}}{(\alpha-\beta x)^{k+1}}\exp\left(-\frac{2l\zeta x}{\alpha-\beta x}\right), \quad \text{(A.31)}$$

where $x < \alpha/\beta$. Further, $f_{\gamma_{\mathrm{eq}}}(x) = 0$ for $x \geq \alpha/\beta$. Next, the approximation of the ergodic sum rate of SS-2 can be derived by averaging the sum rate in (3.61) over the PDF of γ_{eq_1} in (A.31) as

$$\bar{\mathcal{R}}_{L^*} \approx \mathcal{E}\{\mathcal{R}_{L^*}\} = \frac{N_R}{\ln(2)}\int_0^{\infty}\ln(1+x)f_{\gamma_{\mathrm{eq}_1}}(x)\,\mathrm{d}x. \quad \text{(A.32)}$$

Using similar techniques to those used in (3.63), a tight approximation to the ergodic sum rate of SS-2, whenever all relays are equipped with the same number of antennas, can be derived as in (3.65).