MULTIPLE ANTENNA COMMUNICATION OVER FADING CHANNELS: PERFORMANCE LIMITS AND COOPERATIVE TRANSMISSION STRATEGIES

ZHU YONGLAN

NATIONAL UNIVERSITY OF SINGAPORE

2008

MULTIPLE ANTENNA COMMUNICATION OVER FADING CHANNELS: PERFORMANCE LIMITS AND COOPERATIVE TRANSMISSION STRATEGIES

ZHU YONGLAN

(*B. S., Peking University*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2008

Acknowledgment

I am glad to have the chance to thank everyone who has supported me during my PhD study. The work in this thesis could not have been accomplished without the guidance and support of many people.

First of all, I would like to express my sincere appreciation to my supervisors, Dr. Yan Xin and Prof. Pooi-Yuen Kam, for their expert and valuable guidance throughout my PhD study. I am deeply impressed by their enthusiasm, dynamism and motivation for research. They gave me a lot of advices, insights, inspirations and encouragement, from which I have benefited tremendously in both intellectual and personal growth. It has been my great pleasure to work under their supervision.

I also would like to thank my colleagues in Communications Lab and ECE-I2R Wireless Communications Lab for their warm friendship and stimulating discussions in research. Special thanks to Jinhua Jiang, Lan Zhang, Jun He, Le Cao, Seyed Hossein Seyedmehdi, Rong Li, Yan Li, Feifei Gao, Wei Cao, Jianwen Zhang, Hossain A. K. M. Mahtab, Nitthita Chirdchoo, Qi Zhang and Lokesh Bheema Thiagarajan.

I want to thank my friends who have made my stay at NUS enjoyable. Especially, Yijing Fan, Yaqiong Zhang, Hui Zhang and Alexander Shapeev for sharing many laughs and hours of conversations. My tender and sincere thanks go to Pierre Chuong for his encouragement and support in innumerable ways.

Last but not least, I would like to thank my parents for their endless love and support.

Contents

Summary

In a point-to-point communication link, multiple-input multiple-output (MIMO) technology has been used to increase spectral efficiency and reliability of transmissions, and in wireless networks, cooperative transmission techniques have been developed for striking desirable tradeoffs among power, bandwidth, and complexity. In both MIMO and cooperative relay systems, multiple antennas are used: in MIMO systems, all the transmit/receive antennas are collocated at the same terminal where all the transmit/receive antennas can fully cooperate, whereas in cooperative relay systems, the transmit/receive antennas are distributed at different terminals and thus the transmit/receive antennas cannot fully cooperate because the cooperative relays do not have a priori or complete knowledge about information being transmitted from a source, and the source and relays need to communicate with each other over fading links. Even though multiple antennas are employed in both MIMO and cooperative relay systems with different fashions, the use of multiple antennas offers great potential to improve reliability and spectral efficiency of wireless transmissions for both systems.

In this dissertation, we investigate performance limits and cooperative transmission strategies for MIMO and cooperative relay communication systems. We first obtain some bounds on the complementary cumulative distribution function (CCDF) of the mutual information (MI) of a MIMO system in a Rician fading environment. The MI distributions are obtained through the statistical distributions of

the determinant and trace of a noncentral complex Wishart matrix. These bounds on the CCDF of the MI show explicitly the effects of system parameters on the distribution of the MI. We then turn our attention to cooperative relay systems. We derive the outage probability and ergodic rate for a decode-and-forward (DF) relay system in a Rayleigh fading environment, and propose optimal transmission strategies including transmit signaling and power allocations to minimize outage probability or maximize ergodic rate. We also derive the outage probability of DF relay systems in a Rician fading environment. The information theoretic study, however, only provides performance limits for the system; it does not allude to any methods for achieving these limits. We apply differential modulation to relay systems and study the relay systems from an end-to-end error performance perspective. For DF, we design maximum likelihood and piecewise linear receivers at the destination which takes account of average error probabilities of all the source-relay transmissions. For selection relaying, we shift the computational complexity from the destination to relays which compute their own instantaneous error probabilities and use them to make decisions on whether to transmit or remain silent. We also obtain the average error probability of the relay systems for both protocols over Rayleigh fading channels.

List of Figures

Abbreviations

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CSI	Channel State Information
DBPSK	Differential Binary Phase Shift Keying
DF	Decode-and-Forward
i.i.d.	independent and identically distributed
LOS	Line-Of-Sight
LST	Layered Space Time
MAP	Maximum a posteriori Probability
MGF	Moment Generating Function
MI	Mutual Information
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MRC	Maximum Ratio Combining
PDF	Probability Distribution Function
PL	Piecewise Linear

PSK	Phase Shift Keying
RV	Random Variable
SIMO	Single Input Multiple Output
SNR	Signal-to-Noise Ratio
SR	Selection Relaying
STBC	Space Time Block Code
STTC	Space Time Trellis Code

Notation

In this dissertation, scalar variables are written as plain lower-case letters, vectors as bold-face lower-case letters, and matrices as bold-face upper-case letters. Some further used notations and commonly used acronyms are listed in the following:

a	plain lower-case to denote scalars
$\boldsymbol{a}$	boldface lower-case to denote column vectors
$\boldsymbol{A}$	boldface upper-case to denote matrices
$(\cdot)^*$	the conjugate operation
$(\cdot)^T$	the transpose operation
$(\cdot)^H$	the conjugate transpose operation
$\det(\cdot)$	the determinant of a matrix
$\mathrm{tr}(\cdot)$	the trace of a matrix
$\mathbb{E}(\cdot)$	the statistical expectation operation
$\Re(\cdot)$	the real part of the argument
$\otimes$	the Kronecker product
$\|\cdot\|_F^2$	the Frobenius norm square
$\mathrm{erfc}(\cdot)$	the complementary error function
$\Gamma(\cdot)$	the Gamma function
$\Gamma(\cdot,\cdot)$	the upper incomplete Gamma function
$I_m(\cdot)$	the m-th order modified Bessel function of the first kind

$K_m(\cdot)$	the m-th order modified Bessel function of the second kind
$Q(\cdot,\cdot)$	the first order Marcum Q-function
$Q_m(\cdot,\cdot)$	the generalized Marcum Q-function

Chapter 1

Introduction

The rapidly growing demands for wireless services such as high speed internet, mobile television, and mobile computing have stimulated tremendous research efforts towards improving reliability and throughput of communication links. Many technical challenges stem from the time varying nature of wireless medium (e.g., multi-path fading), interference, as well as power and bandwidth limitations. Among these challenges, multi-path fading (often simply called fading) significantly degrades the reliability of the transmission links, and becomes a bottleneck for enhancing data rate in the wireless systems. In addition, the available radio spectrum is limited and a significant increase in spectral efficiency is needed to meet the needs for communication capacity.

One of the most effective solutions to mitigate the detrimental effects of fading is to exploit *diversity* techniques. Roughly speaking, the basic idea is to send signals that carry the same information through different paths, and obtain multiple independently faded replicas of the data symbols at the receiver, thus achieving more reliable detection. Depending on the domain where replicas of transmitted signals are created, diversity techniques can be categorized into three types: time diversity, frequency diversity and space diversity. Either time or frequency diversity induces a loss in

bandwidth efficiency due to the redundancy introduced in time or frequency domain. Compared to time and frequency diversities, space diversity, which can be provided by sufficiently separated or differently polarized antennas, can enhance performance without causing a loss in bandwidth efficiency. Therefore, space diversity has become one of the most favored solutions in numerous wireless applications.

A conventional approach to achieving space diversity is to employ multiple transmit and/or multiple receive antennas. If the antennas are placed sufficiently far apart, the channel gains between different antenna pairs are independent. For a mobile terminal, a half to one carrier wavelength separation among antennas is sufficient to guarantee that the channel gains are independent. Through transmitting the replicas of the signal through different antennas, and/or combining the different replicas together at the receiver, space diversity can be achieved. Traditionally, space diversity is achieved by employing multiple receive antennas (single-input multiple-output or SIMO channels) at the receiver where combining, selection or switching of the received signals is performed. This is so-called *receive diversity*. By deploying multiple transmit antennas (multiple-input single-output or MISO channels) at the transmitter, *transmit diversity* techniques shift the complexity associated with realizing diversity to the transmitter. A well-known transmit diversity scheme is *space-time coding* that is capable of achieving diversity and coding gains compared to single-antenna uncoded transmissions.

Communication systems with multiple transmit and multiple receive antennas, often simply called multiple-input multiple-output (MIMO) systems, provide even greater potential. In addition to the aforementioned diversity benefits, it has been revealed that increasing the number of both transmit and receive antennas has great potential to enhance spectral efficiency of wireless transmissions. MIMO channels exhibit fading and provide additional spatial dimensions. Under certain fading

conditions, additional degrees of freedom can be provided by MIMO channels, and be exploited to increase capacity which highlights the potential spectral efficiency of MIMO channels. The capacity of the MIMO channels has been proved to grow approximately linearly with the minimum number of transmit and receive antennas. Therefore, MIMO technology not only leads an improved error performance but also enables high spectral efficiency communication.

However, in wireless networks, there are many terminals (transmitters and/or receivers), and each transmitter or receiver usually only has a single antenna because of the transmitter's and receiver's size limitation. In addition, maintaining direct reliable transmission links between widely separated terminals requires high transmission power, which not only incurs energy inefficiency but also introduces strong interference to the other nearby terminals. As an alternative, the communication between a pair of far-apart terminals can take place through several intermediate terminals. In this transmission mode, the multiple terminals can act as a virtual antenna array to achieve space diversity in a distributed fashion. This space diversity is called *cooperative diversity*. Cooperative diversity can be used in cellular, ad-hoc, and hybrid networks in order to increase coverage, throughput and capacity. In cellular networks, cooperative relay transmission has been used to extend network coverage. In sensor networks such as military battlefield communication networks, the use of wire-line infrastructure is often precluded and the radios might be substantially power constrained. For these ad-hoc or sensor networks, wireless relay transmissions have the potential to greatly improve reliability, energy efficiency, as well as throughput of wireless networks.

We will introduce the conventional multiple antenna system and the cooperative diversity system in the following two sections.

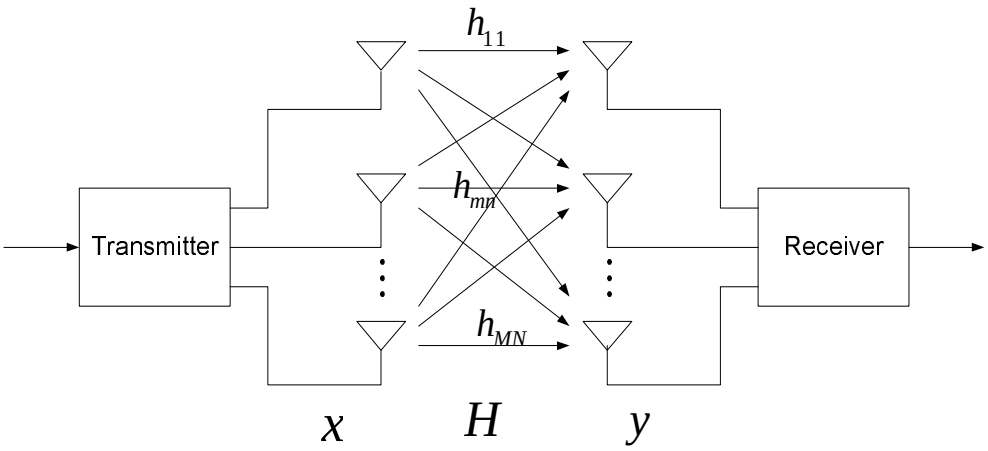

Figure 1.1: Block diagram of a MIMO system.

1.1 Multiple-Input Multiple-Output (MIMO) Systems

The conventional multiple antenna system has one transmitter and one receiver where multiple antennas are employed at the transmitter and/or receiver. The multiple transmit/receive antennas are collocated at the transmitter/receiver, and full cooperation among the transmit/receive antennas is possible.

1.1.1 MIMO System Model

A typical MIMO system with N transmit antennas and M receive antennas is shown in Fig. 1.1. The input-output relationship can be expressed in vector form as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{1.1}$$

where $\boldsymbol{H}$ is the $M \times N$ channel matrix with the mn-th entry h_{mn} describing the channel gain between the n-th transmit antenna and the m-th receive antenna, $\boldsymbol{x}$ is an N-dimensional transmitted vector with the n-th element the transmit signal from the n-th transmit antenna, $\boldsymbol{y}$ is an M-dimensional received vector with the m-th element the receive signal at the m-th receive antenna, and $\boldsymbol{n}$ is an additive white Gaussian

noise (AWGN) vector with the m-th element the AWGN at the m-th receive antenna.

There have been great interests in MIMO channels of the form shown in Fig. 1.1. It has been well documented that employing MIMO system not only improves bit error rate (BER) performance but also boosts the channel capacity. Substantial efforts have been made on characterizing the ultimate information theoretic limits of the MIMO systems, and designing practical coding and decoding algorithms that have high spectral efficiency.

1.1.2 Information Theoretic Performance Limits

Established by Claude Shannon, the notion of *capacity* is referred to the maximum rate of communication for which an arbitrarily small error probability can be achieved, and is commonly used to characterize the fundamental performance limit of reliable communication.

In analyzing the capacity of fading channels, there are two commonly used statistics: *ergodic capacity* and *outage probability*. In fading environments, since the channel $\boldsymbol{H}$ is random, the mutual information (MI) associated with the MIMO channel is a random variable (RV) and has a statistical distribution function. The ergodic capacity is defined as the ensemble average of the MI. If the fading process is ergodic, the ergodic capacity is equal to the Shannon capacity, and an arbitrary small error probability can be achieved if the transmission rate is smaller than the Shannon capacity assuming that an asymptotically optimal codebook is used. If the fading process is non-ergodic, e.g., the channel is randomly selected at the beginning of the transmission and remains constant during the transmission, the Shannon capacity of the channel can be zero because there is a nonzero probability that a particular realization of the channel is incapable of supporting arbitrarily low error rates, no matter which code is selected. In this case, the ergodic capacity is not meaningful

and the performance measure is the outage probability. The outage probability is the cumulative distribution function (CDF) of the MI, and measures the tradeoff between the transmission rate and the reliability.

Ergodic Capacity of MIMO systems

There has been substantial work on characterizing the ergodic capacity of MIMO systems under a variety of fading conditions. The ergodic capacity of the MIMO channel has been developed for several different cases which depend on the availability of channel state information (CSI) at the transmitter and/or receiver: CSI available to the receiver only [1, 2]; CSI available to both the transmitter and the receiver [2]; and CSI available at neither the transmitter nor the receiver [3–5]. The information theoretic study has revealed that MIMO systems have the potential to provide dramatic increases in capacity. For example, when CSI is available at the receiver only, the ergodic capacity of a MIMO system grows linearly with the minimum number of transmit and receive antennas [2]. The linear increase of the ergodic capacity at high signal-to-noise ratio (SNR) is due to the increase in degrees of freedom from spatial multiplexing. The linear increase at low SNRs is due to a power gain from receive beam-forming. The linear increase at intermediate SNR ranges is due to a combination of both these gains. We will focus on flat fading channels with CSI available at the receiver only.

A wireless channel is often too complex to be precisely modeled, and statistical models for characterizing the channel effects are widely employed. There is a range of relatively simple and accurate statistical models for fading channels depending on the particular propagation environment. The two commonly used fading models are Rayleigh fading and Rician fading.

Perhaps Rayleigh fading is the most commonly adopted model in the literature. If

we assume that fading is caused by the superposition of a large number of independent scattered components, then the in-phase and the quadrature components of the fading coefficient are independent zero-mean Gaussian processes. Thus, the fading coefficient is zero-mean complex Gaussian distributed. For independent Rayleigh fading among the individual links, the ergodic capacity of a MIMO system was obtained exactly in [2], where the ergodic capacity was expressed in terms of Laguerre polynomials. Since the exact expression of the ergodic capacity is quite complex and is not given in a closed form, [6] presented some tight lower bounds on the ergodic capacity in closed forms, and [7] presented the limiting value of the ergodic capacity when the numbers of transmit and receive antennas tend to infinity and the ratio of the number of transmit antennas to the number of receive antennas is held constant. The capacity derivation and the optimality of Gaussian codebooks are based on the assumption that perfect CSI is available at the receiver. When channel estimation errors exist, performance degradations in terms of the MI were investigated in [8], and some bounds on the ergodic capacity were obtained in [9].

Rayleigh fading is known to be a reasonable channel model applicable to many wireless systems. However, when line-of-sight (LOS) paths exist between transmit and receive antennas, the fading coefficients are no longer zero-mean complex Gaussian distributed, and they are nonzero-mean complex Gaussian distributed. In this case, the channels are modeled as Rician fading which is more general than Rayleigh fading, since Rician fading reduces to Rayleigh fading when the LOS component is equal to zero.

The ergodic capacity of a MIMO system in Rician fading environments has been presented in [10–14]. By making use of the joint distribution of the eigenvalues of a noncentral Wishart matrix, the exact ergodic capacity was obtained in multiple integral forms [10], which only can be evaluated by numerical integrations. Some tight upper

and lower bounds on the ergodic capacity have been derived in [11–13] in simpler forms. In [11], the MI was re-expressed as a linear sum of determinants of noncentral Wishart distributed sub-matrices. The new expression of the MI, together with the convexity and concavity of some functions, enables the ergodic capacity to be upper and lower bounded [11]. Actually, [10, 11] only considered the case that the channels between different transmit and receive antennas are statistically independent, and it is not true if the antennas are not separated enough. The ergodic capacity bounds of independent or single-correlated Rician fading MIMO channels with a rank-1 mean matrix were derived in [12] through computing the expected values of the determinant and log-determinant of a noncentral complex Wishart matrix. A more general study on the ergodic capacity of spatially correlated Rician fading MIMO channels was done in [13] through exploiting the statistical property of the noncentral matrix-variate complex quadratic forms. Some upper and lower bounds on the ergodic capacity of the single- or double- correlated channels with an arbitrary mean matrix were obtained in [13].

Statistical Distribution of the MI of the MIMO Systems

Compared to the ergodic capacity, the statistical distribution of the MI is fairly useful in either ergodic or non-ergodic fading channels, and it provides more comprehensive information on the MIMO systems than the ergodic capacity does. Therefore, the statistical distribution of the MI is of interest. However, because of the complexity of the statistical distribution of a complex Wishart matrix and the determinant operation, the study on the statistical distribution of the MI is quite involved.

There are several different studies [15–20] on the statistical distribution of the MI in a Rayleigh fading environment. The complementary cumulative distribution function (CCDF) of the MI was first studied in [15] via Monte Carlo simulations, which

empirically provides us some information on the statistical distribution of MI. But the experimental approach, e.g., Monte Carlo simulations, is quite time consuming and cannot indicate how parameters affect the MI distribution. Later, based on the central limit theorem, [16] proposed using a Gaussian distribution function to approximate the statistical distribution of the MI, for which only the mean and the variance of the MI need to be calculated. The Gaussian approximation is quite accurate for a large number of antennas. But when the number of antennas is small, which is generally the case in practice, the Gaussian approximation becomes loose. To overcome this shortcoming of the Gaussian approximation, the exact CDF of the MI of a dual antenna MIMO system was derived in [17] by using a Jacobian on the joint density function of the eigenvalues of a complex Wishart matrix. However, the complexity of the Jacobian transformation increases exponentially with the increase in the number of antennas. Thus the result in [17] cannot be generalized to a system with a larger number of antennas. A more comprehensive study on the distribution of the MI has been conducted in [18–20]. The characteristic function of the MI was obtained in [18] through rewriting the joint distribution of the eigenvalues of a Wishart matrix in a new form; while the moment generating function (MGF) of the MI has been given in [19, 20]. These results in [18–20] are valid for an arbitrary number of antennas, and the CDF of the MI can be expressed as the Fourier transform of the characteristic function or the inverse Laplace transform of the MGF.

For Rician fading channels, the noncentral properties of a complex Wishart matrix make the MI distribution much more complicated than that in the case of Rayleigh fading channels. Similar to [17] which only considers Rayleigh fading channels, the exact CDF of the MI for a dual antenna MIMO system was derived in [21] for Rician fading channels by using a Jacobian transformation. Recently, [14] obtained the MGF of the MI for independent Rician fading MIMO channels. However, to obtain the CDF

of the MI for MIMO Rician fading channels, we need resort to numerical integrations to take the inverse Laplace transform of the MGF. Based on the MGF of the MI, the mean value and the variance of the MI were also obtained in [14] for independent Rician fading channels. In addition, the mean and the variance of the MI were present in [13] for single- or double-sided correlated Rician fading channels. With the mean and variance of the MI, a Gaussian distribution has been used to approximate the distribution of the MI in [13, 14].

All the aforementioned results on the distribution of the MI rely on the joint density function of the eigenvalues of a central or noncentral Wishart matrix. Until now, there is still no explicit expression available for the CDF of the MI, which is an important performance measure of MIMO systems. Unfortunately, due to the complicated mathematical problem related to the eigenvalues of a Wishart matrix, an explicit expression for the CDF of the MI is very difficult to obtain. Instead, some upper and lower bounds in simple forms may be derived.

1.1.3 Space-Time (ST) Coding

ST coding is an effective and practical way to improve spectral efficiency of MIMO fading channels. Coding is performed in both space and time domains to introduce correlation between signals transmitted from various antennas in various time periods. ST coding can achieve transmit diversity and coding gain over uncoded systems without sacrificing bandwidths. Typical examples of ST coding schemes include space-time trellis codes (STTC), space-time block codes (STBC) and layered space-time (LST) codes.

A well known upper bound on the pairwise error probability and the corresponding STTC design criteria (rank and determinant criteria) were derived in [22]. Alamouti's scheme [23], which is for a system with two transmit antennas,

admits a particularly simple linear decoding algorithm and achieves full diversity. The key feature of the Alamouti's scheme is the orthogonality between the sequences transmitted by the two transmit antennas. Alamouti's scheme was generalized to STBC with an arbitrary number of transmit antennas [24]. A LST architecture that can attain a tight lower bound on the MIMO channel capacity was proposed in [25]. Some other forms of ST code, such as the differential ST code [26] and non-coherent unitary ST code [27], also were proposed.

1.2 Cooperative Relay Systems

As we see in the previous section, using multiple antennas has offered high reliability and spectral efficiency for point-to-point communication systems. In cooperative relay systems, multiple antennas are utilized in a distributed manner such that communications between nodes may take place through several intermediate nodes to save power, reduce interference and increase diversity. Cooperative communications in such systems are realized by exploiting the broadcast nature of the wireless medium and allowing terminals to cooperatively transmit information through relaying. The relaying system realizes many benefits over traditional MIMO systems in terms of adaptability and capacity. In fact, every additional terminal added to the network increases the overall connectivity and improves fault tolerance. In contrast to the conventional forms of space diversity with physical arrays, cooperative relay systems benefit from space diversity by using multiple terminals as a virtual antenna array. The cooperative relay systems are fundamentally different from the point-to-point MIMO systems in the sense that the cooperative relays do not have a priori or complete knowledge about information transmitted from the source and only have faded and noisy versions of the transmitted information.

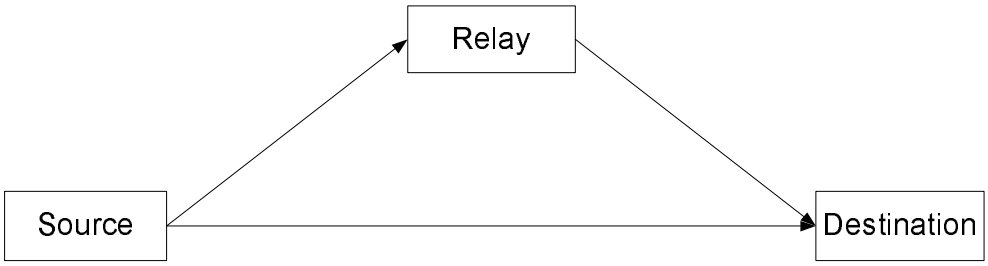

Figure 1.2: The classic relay channel model.

1.2.1 Relay System Models

A classic relay model [28–30] is shown in Fig. 1.2. The classic relay channel model comprises of three terminals: a source that transmits information, a destination that receives information, and a relay that both receives and transmits information in order to enhance communication between the source and destination. As illustrated in Fig. 1.2, the relay performs the following function: receiving signals transmitted from the source, processing those received signals, and transmitting the processed signals to the destination. Roughly speaking, the use of the relay transmission is able to increase the capacity and/or improve reliability of an end-to-end transmission from the source to destination.

For a cooperative relay system with multiple relays (L relays), there are two categories of relay systems: *parallel* relay system as shown in Fig 1.3, and *serial* relay system as shown in Fig. 1.4. As illustrated in Fig 1.3, for a parallel relay system, all the L relays receive signals transmitted from the source, and the destination receive signals transmitted from the source and all the L relays. As illustrated in Fig. 1.4, for a serial relay system, relaying can be performed in multiple stages and relays receive signals transmitted from the source and the previous several relays, such that relays as well as the destination benefits from space diversity.

Cooperative communication [31, 32] is a generalization of the relay channel to

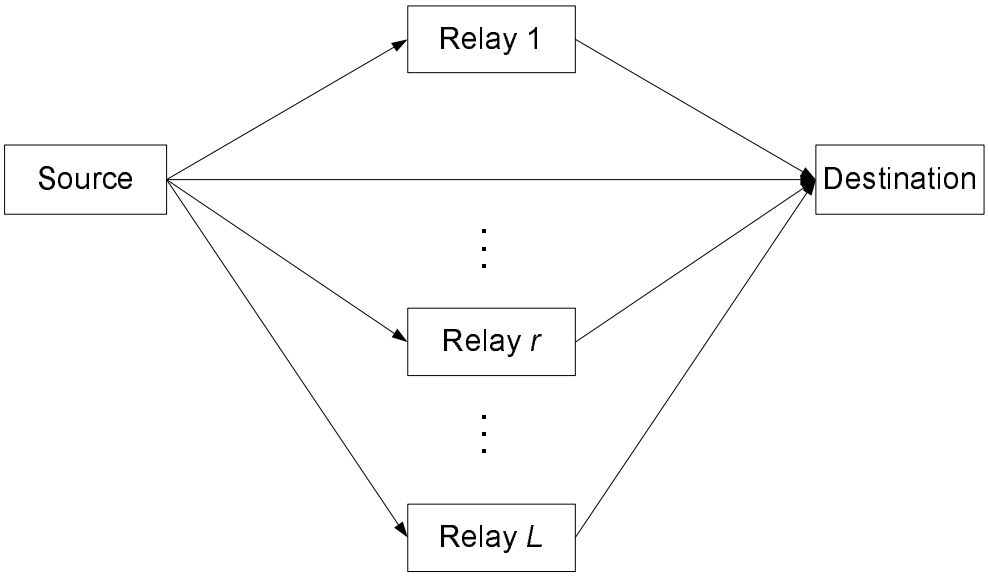

Figure 1.3: The parallel relay channel model with L relays.

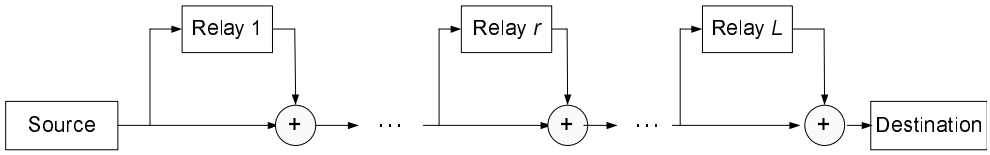

Figure 1.4: The serial relay channel model with L relays.

multiple sources which have information to be transmitted and also serve as relays for one another. In this case, the sources and relays act as partners for one another [33]. A combination of relaying and cooperation is also possible, and is often referred to generically "cooperative communications".

Cooperative relay communication leverages space diversities available when multiple transmission links are independently faded. For example, if the direct link from the source to the destination experiences a deep fade, there remains a significant probability that the signal transmitted from the source can be effectively communicated to the destination via one of the relays.

Cooperative communication in a wireless network involves issues of protocol

layering and cross-layer architectures. In physical layer, encoding and signal processing are required at the source and relay(s) [31, 32, 34], and decoding and signal processing are required at the destination and relay(s) [35–37]. Scheduling [38–40] for transmissions in time and frequency has to be addressed by protocols in the link layer and medium-access control layer in coordination with the physical layer. Symbol, frame and carrier synchronization [41–43] is also of particular importance at the physical and link layers. Placing participating terminals into cooperating groups is a cross-layer issue that can involve the physical, medium-access control, link and even network layers [34, 44, 45]. In addition, the overhead introduced by internode communications might detract from the gains obtained via cooperation. Designing an effective cooperative communication system requires insights about all of these issues. Appropriate architecture and algorithms (specific encoding, decoding and signal processing) are very important, and can depend upon the radio hardware, system complexity and application context.

Full-Duplex Mode and Half-Duplex Mode

A relay node can work either in *full-duplex* mode or *half-duplex* mode. If a relay node works in a full-duplex mode, the relay node receives and transmits at the same time on the same frequency band. Since the relay node has full knowledge of what to transmit, it can cancel the interference from its transmit antenna at its receive antenna. The full-duplex mode has high spectral efficiency, but high complexity.

In contrast to the full-duplex mode, a half-duplex mode means that the relay node cannot transmit and receive at the same time or at the same frequency band. There are two types of half-duplex modes: time-division-duplex (TDD) and frequency-division-duplex (FDD). In the TDD mode, for a given time interval D, the relay node is in receive mode for a fraction of the time aD and in transmit mode for

the rest of the time $(1-a)D$. Similarly, in the FDD mode, a given bandwidth W is divided into a bandwidth of aW over which the relay node receives, and a bandwidth $(1-a)W$ over which the relay node transmits. The destination listens over the whole bandwidth W. Clearly, from an information theoretic point of view, TDD and FDD are equivalent. Compared to the full-duplex mode, the half-duplex mode is easier to realize, but it has lower spectral efficiency.

Relay Protocols

A relay node can employ one of several protocols to process the received signals and forward to the destination. The most commonly used four protocols are given as follows.

- Amplify-and-Forward (AF) [35, 44]: The relay node transmits a linear combination of its past received signals. The main advantage of the AF protocol is its simple operation at the relay node. But the transmit signals from the relay are analog signals, and the end-to-end performance is limited since the noise at the relay is also amplified and forwarded to the destination.

- Compress-and-Forward (CF) [30, 46–49]: The relay node transmits a compressed and quantized version of its past received signals. Compression is done using Wyner-Ziv source coding. When the relay-destination link is strong, the CF cooperative relay channel behaves like a SIMO channel [49, Remark 20]. The main drawback of the CF protocol is its relatively high complexity.

- Decode-and-Forward (DF) [30, 44]: The relay node applies some form of detection and/or decoding algorithms to its received signals and re-encodes the information into its transmit signals. This decoding and re-encoding process often corresponds to a non-linear transformation of the received signals.

Although decoding at the relays has advantages of reducing the impact of receiver noise, the performance of the DF protocol is limited by the source-relay link. If the source-relay link is strong, the DF cooperative relay channel behaves like a MISO channel [49].

- Selection Relaying (SR) [44, 50, 51]: The performance of the DF protocol is constrained by the reliability of the source-relay link. In the DF protocol, if a relay makes detection errors, erroneous symbols from the relay will lead to a degradation in post-combining SNR at the destination, and a symbol error at the destination becomes highly likely. This can be viewed as a form of error propagation in the relay system. If a relay retransmits all the symbols regardless of the reliability in their decisions, the possible error propagation will limit the end-to-end performance of the DF relaying. If the systems are designed with embedded error detection codes, the relay can forward only if no error is detected. In this case, the relay also can correct some symbol errors by decoding and re-encoding the received data block. Alternatively, the instantaneous SNR of the source-relay link can be used as an indicator of the reliability of the relay detection. If the source-relay instantaneous SNR is larger than a certain threshold, the probability of an error at the relay is small and hence the relay transmits the signal. Otherwise, the relay remains silent.

There are many other relaying protocols which are extended versions of the above protocols, e.g., partial DF, dynamic DF, estimate-and-forward, etc.

It has been shown that cooperative relaying techniques offer great performance improvement from various perspectives: increased capacity or enlarged capacity region [30, 49, 52]; improved reliability in terms of diversity gain [34, 44, 53, 54], diversity-multiplexing tradeoff [55], and bit or symbol error probabilities [35, 37,

56–58]. In the following, we will introduce cooperative relay systems from two perspectives: an information theoretic perspective and an end-to-end performance perspective.

1.2.2 Information Theoretic Study on Cooperative Relay Systems

Early formulation theoretic study on cooperative relay systems was launched in early 1970's [28, 29], and continued by many others, e.g, Cover and El Gamal [30], with a primary focus on the classic relay channel. More recently, multiple relay models have been examined [49, 59–61]. For the classic relay model in Fig. 1.2, Cover and El Gamal [30] obtained the channel capacity for *physically degraded*[1] relay channels, and a capacity upper bound and an achievable rate for general relay channels. Various study on cooperative relaying has been done from the information theoretic perspective for decades, however, the capacity or capacity region, are not known in general. Nevertheless, bounds on capacity and numerous achievable rates have been obtained [49, 59–61].

The information rates achieved by various half-duplex relaying protocols and their corresponding outage probabilities were obtained in [44]. In [52], some upper and lower bounds on the capacity of classic relay channel were presented when the relay node operates in full duplex mode or half-duplex mode. With the transmitter CSI, several power allocation strategies that maximize the upper and lower bounds of ergodic capacity or minimize the outage probability were also proposed in [52]. For MIMO relay channels, [62] presented some upper and lower bounds on the capacity of the classic relay channel when the relay node operates in full duplex mode for the cases of both AWGN channels and Rayleigh fading channels. Capacity theorems and

[1]Degradedness means that the destination receives a corrupted version of what the relay receives, all conditioned on the relay transmit signal.

achievable rates for systems with multiple relays have been derived in [49, 60, 61]. It was shown in [49] that for Rayleigh fading DF relay channels, the transmit signals from source and relay, which maximize the ergodic rate, are statistically independent. In the low SNR regime, the spatial power allocation strategy that maximizes the ergodic rate was presented in [63], and the spatial power allocation strategy that maximizes the upper bound of the ergodic capacity was derived in [64]. While the work in [52, 63, 64] emphasizes on optimizing ergodic performance measures, a recent paper [65] proposed a spatial power allocation strategy that minimizes the outage probability of the information rate for AF relay channels.

For a classic relay channel employing DF protocol, the highest information rate achievable until now was derived in [30]. The corresponding outage probability and the ergodic rate in fading channels are the two important information theoretic performance measures. However, the analytical expressions of these quantities are not available in the literature for an arbitrary correlation between the transmit signals from the source and relay. When transmitters only have statistical CSI, the optimal power allocation and transmission strategies in the DF relay system to minimize the outage probability or maximize the achievable rate are also needed.

1.2.3 End-to-End Performance in Cooperative Systems

For a relay system, several encoding strategies have been proposed for forwarding the source message by the relay(s): simple uncoded or repetition coding, ST coding [34], and more idealistic block Markov coding derived from the information theoretic framework established by Cover and El Gamal [30]. The error probability analysis is usually done for repetition coding or ST coding. In general, ST coding is superior to uncoded and repetition, since ST coding provides diversity without causing a substantial loss in spectral efficiency [22].

End-to-End Performance of Uncoded Relay Systems

For a classic relay channel, the bit or symbol error probabilities have been derived in [56] for AF relay, and in [37] for DF relay. It is worth noting that the work [56] expressed the average received SNR at the destination as a harmonic mean of two exponentially distributed RVs, and then obtained the MGF of the average received SNR and the symbol error probability. The paper [37] proposed a cooperative maximum ratio combining (C-MRC) scheme for DF relay channel and obtained its corresponding symbol error probability.

The extensions of the error performance analysis from a classic relay channel to a multiple relay channel are done for a parallel relay in [35], or a serial relay in [57, 66], or a combination of these two in [58]. In [35], the maximum likelihood (ML) receiver for the coherent AF parallel relay was designed. The ML receiver for the coherent AF parallel relay is a linear operation on the received signals, and thus easy to implement. The error probability analysis for the coherent AF relay is quite difficult, since it requires averaging the performance over the fading coefficients. Some progress in the error probability analysis of the coherent AF parallel relay has been reported in [35,56,58]. In [57], four different models, which are classified according to relay processing (AF or DF) and signal reception model (from all preceding terminals or from the immediately preceding terminal only), were considered. In [66], each relay coherently combines the signals received from the source and one or more of the previous relays. In [58], symbol error probabilities for high SNR were derived for parallel relay, serial relay and combined configurations.

All the aforementioned studies assume that the receivers have perfect CSI, and utilize the CSI for coherent detection. However, channel estimation for multiple transmission links is complex and costly in fading environments, and incurs communication overheads, especially when the fading is rapid and the number of relay

nodes is large. To obviate the need for channel estimation for wireless relay systems, non-coherent or differential modulations are suitable alternatives.

The general non-coherent ML detector for the AF relay was derived in [67], and the ML detectors for on-off keying (OOK) and binary frequency shift keying (BFSK) non-coherent AF relay were derived in [68]. However, the ML detectors are not given in closed form and involve numerical integrations, which makes the ML detectors very difficult for implementation. The key challenge in deriving the ML decision rule for the non-coherent AF relay is to obtain the statistics of the combination of products and sums of Gaussian RVs. For a non-coherent DF relay system with BFSK modulation, [36] derived the nonlinear ML detector and proposed a piecewise linear (PL) detector which has performance similar to the ML detector.

For an AF relay system using differential modulations, various receivers and their corresponding error probabilities have been studied in [69–72]. Linear combiners at the destination have been proposed [69, 70, 72] based on the assumption that the effective noise at the destination is Gaussian distributed. For a DF relays using differential modulations, a nonlinear ML detector and an equal gain combiner with low complexity were derived in [71] and [73], respectively. More recently, the paper [72] derived the ML detector and the PL detector for the DF relay with differential binary phase shift keying (DBPSK) modulation. All the results on relay systems with differential modulations have made the assumption that the fading coefficients remain unchanged in two consecutive symbol intervals. This assumption may not be valid when the fading is relatively rapid.

Distributed ST Coding

In cooperative relay systems, several terminals can share their resources and cooperate to work as a virtual MIMO system. Hence, the ST coding in conventional MIMO

systems also can be extended to cooperative relay systems. The ST coding in cooperative relay systems is called *distributed ST coding* due to the fact that the encoding operation is distributed among cooperating terminals.

The use of distributed ST coding in cooperative relay channels was proposed in [34] in which orthogonal STBC were used and each node transmits a column of the STBC matrix. It is difficult to design orthogonal STBC [74] for a large number of terminals, and a rate loss is unavoidable. Thus distributed orthogonal STBC can only be applied to a network with a small total number of terminals. A new class of distributed STBC was proposed in [75] where each node is assigned a unique signature. [76] presented a regenerative relay system with the Alamouti code, and analyzed the effects of errors at the relays for the regenerative distributed ST coding. The non-regenerative relay system with Alamouti code was presented in [77]. An error-aware distributed ST code was proposed in [78] for a cooperative relay system with one or two relays. The error-aware distributed ST scheme allows feedback from destination and error probability feedforward from relays. In [79], linear dispersion ST codes were applied among relays in a parallel relay system. The pairwise error probability, diversity gain and coding gain of the distributed linear dispersion ST codes were analyzed. The distributed ST coding also has been extended to differential modulation [50, 80] and non-coherent modulation [51].

1.3 Research Objectives and Contributions

As addressed in Section 1.1.2, the statistical distribution of the MI for a MIMO system in a fading environment has great importance, but no explicit expression of the MI distribution is available. We study in Chapter 2 the CCDF of the MI of a MIMO system in a Rician fading environment. We propose a new approach that examines the

MI distribution through investigating the statistical distribution of the determinant and trace of a noncentral complex Wishart matrix, instead of using the joint distribution of the eigenvalues of the Wishart matrix, as is common in the literature. The PDF and CCDF of the determinant of a noncentral complex Wishart matrix derived in Chapter 2 are important results in their own right, and may have applications beyond the MI distribution analysis. Our result on the CDF of the MI serves as the outage probability when the fading process is non-ergodic, which is an important performance measure of the MIMO system in fading environments. The CDF also may be used to evaluate the mean value of the MI or ergodic capacity. In addition, our study should be easily reduced to the Rayleigh fading case since it is a special case of Rician fading. Furthermore, our approach to the CCDF of the MI should greatly simplify the calculation procedure, and may provide easy and accurate ways to deal with MI related calculations for MIMO systems.

For a cooperative relay system, the outage probability of the highest information rate and ergodic rate of the classic relay channel with DF protocol in a fading environment are still unknown, as addressed in Section 1.2.2. We derive the outage probability and ergodic rate of the classic relay channel employing the DF protocol in Chapter 3 for Rayleigh fading environments. For fixed channel gains, we rewrite the highest information rate in terms of a Hermitian complex Gaussian quadratic form. It enables us to evaluate the outage probability and ergodic rate. Both the outage probability and the ergodic rate are functions of the SNRs and the correlation coefficient between the transmit signals from the source and relay. Based on the derived outage probability and ergodic rate, we design optimal transmission strategies which minimize the outage probability (we call it outage-optimal) or maximize the ergodic rate (we call it ergodic-optimal). The optimal transmission strategies include the optimal correlation coefficient of the transmit signals from source and relay, and

the optimal transmit power allocation between source and relay. When no spatial power allocation is available, we show that if the target transmission rate is smaller than a certain threshold, the outage-optimal transmit signals from the source and relay are independent; if the target transmission rate is larger than the threshold, the outage-optimal transmit signals from the source and relay are not necessarily independent, and their correlation generally depends on the SNRs of the links and the target transmission rate. However, the ergodic-optimal transmit signals from the source and relay are statistically independent irrespective of the SNRs. Moreover, we further show that the ergodic rate is a monotonically decreasing function of the correlation coefficient between the transmit signals from the source and relay. When spatial power allocation between source and relay is feasible, we design optimal power allocation strategies. In a small outage scenario, which is generally of practical interest, we derive an outage-optimal spatial power allocation strategy between source and relay. Our findings suggest that it is not always beneficial to use a relay. When the channel from source to relay is weak, the relay terminal may have a large probability of making decoding errors and the transmit signals from the relay may become interference at the destination. The outage-optimal power allocation generally depends on several parameters, including the total transmit power, the target transmission rate, and the variances of the transmission links. Additionally, the ergodic-optimal spatial power allocation strategy is obtained by numerically solving certain transcendental equations, and it depends on the variances of the three links and the SNRs.

When an LOS path exists, wireless channels are subject to Rician fading which is more general than Rayleigh fading in the sense that it includes Rayleigh fading as a special case. We also extend our results in Chapter 3 for the outage probability of a classic DF relay channel in a Rayleigh fading environment to the outage probability in a Rician fading environment in Chapter 4. Due to the noncentral property of the

Rician fading, the outage probability of the highest information rate in a Rician fading environment is much more complicated than the one under Rayleigh fading. We obtain an analytical expression for the outage probability, which can be computed by using standard numerical techniques. Through a geometric interpretation of the outage probability, we also obtain an upper bound and a lower bound on the outage probability in simple closed forms. These bounds can serve as fairly good approximations to the outage probability. Moreover, when the channel statistics (Rice factors and variances) are known at the source and relay nodes, the outage probability can be minimized by choosing an appropriate correlation coefficient between the transmit signals from the source and relay. Relying on numerical methods, we obtain the optimal correlation coefficient that minimizes the outage probability. Numerical results reveal that for large values of Rice factors, the optimal correlation coefficient is not necessarily zero, but instead, depends generally on SNRs, variances and Rice factors of the channels, as well as the rate threshold.

After studying the cooperative relay systems from an information theoretic perspective, we investigate the cooperative relay systems from an end-to-end error performance perspective in Chapter 5. The error probability of a parallel relay system using differential modulation is analyzed for both DF protocol and SR protocol. For the DF protocol, the destination takes account of the average error probabilities of all the source-relay transmissions and performs ML or PL detection. The PL detector has similar performance to the nonlinear ML detector. In both the ML and PL detectors, the received signals at the destination from the source and all the relays are combined with different weights, because the transmission links in the relay system have different statistics. Then, the BER of the PL detector is derived, whereas the BER analysis for the nonlinear ML detector appears intractable. For a DF single relay system, we obtain the exact BER of the PL detector and its high SNR approximation. The

exact BER is a simple function of the SNRs, the variances and fade rates of all the transmission links. The BER approximation at high SNR explicitly shows the diversity order and the different effects of the source-relay link and the relay-destination link on the end-to-end error performance. Moreover, for a DF multiple relay system, we obtain a Chernoff upper bound on the BER and a high SNR approximation on the BER. For the SR protocol, the computational complexity is shifted from the destination to the relays. The relays need to compute their own instantaneous error probabilities in detecting source information and make decisions on whether to transmit or remain silent according to certain instantaneous error probabilities. At one frame, the relays that are transmitting form a random set which is a subset of the set containing all the potential relays. The destination performs simple MRC reception, and the end-to-end performance is analyzed at high SNR. It reveals that the SR protocol offers a space diversity order of the number of all the potential cooperating nodes, not just the number of nodes that are transmitting to the destination.

1.4 Organization of the Dissertation

The rest of this dissertation is organized as follows.

Chapter 2 investigates the statistical distribution of the MI for a MIMO system in a Rician fading environment. The PDF and CCDF of the determinant of a noncentral complex Wishart matrix are derived, and some upper and lower bounds on the CCDF of the MI are proposed. In non-ergodic fading, these bounds on capacity CCDF can provide insight of MI outage probability.

Chapter 3 presents the outage probability and ergodic rate for a classic DF relay system with full-duplex transmission in a Rayleigh fading environment. With statistical CSI, optimal transmission strategies including optimal transmit signaling

and spatial power allocation are derived to minimize the outage probability or maximize the ergodic rate.

Chapter 4 extends the results in Chapter 3 to the case of Rician fading. An analytical expression for the outage probability, and an upper bound and a lower bound on the outage probability in simple closed forms are obtained. When the channel statistics (Rice factors and variances) are known at the source and relay nodes, we obtain the optimal correlation coefficient that minimizes the outage probability using numerical methods.

Chapter 5 is devoted to receiver design and error probability analysis for parallel relay systems with differential modulation for DF or SR protocols. The DF protocol places the detection complexity at the destination while keeping relatively simple transceivers at the relays. The SR protocol, on the other hand, shifts the computational complexity to the relays and keeps simple detection at the destination.

Finally, Chapter 6 summarizes our work, and points out a number of future research directions.

Chapter 2

On the Mutual Information Distribution of MIMO Rician Fading Channels

In a fading environment, the statistical distribution of the MI of a MIMO system depends on the joint distribution of the eigenvalues of a Wishart matrix, and is quite complex in general. In this chapter, we obtain simple expressions for the statistical distributions of the determinant and the trace of a Wishart matrix. Based on these obtained expressions, we derive some simple and tight bounds on the CCDF of the MI of a MIMO system in Rician fading environments. The obtained bounds on the CCDF of MI provide further insights into the channel MI, and show the effects of the system parameters on the MI distribution explicitly. In addition, results for the Rayleigh fading channels are obtained as a special case.

2.1 Introduction

In the past decade, the information theoretic study of MIMO fading channels has attracted considerable research attention [1]. The primary focus of such a study is on *ergodic capacity*, which is defined as the ensemble average of the MI[1] over the statistical distribution of the channels [81]. In particular, the ergodic capacity has been investigated in [2, 3, 6] for Rayleigh fading, and in [10–14] for Rician fading. For ergodic channels, the ergodic capacity is viewed as an important performance measure, since it can be viewed as the *Shannon capacity* in the sense that we can transmit with an arbitrarily small error probability when the transmission rate is less than ergodic capacity. On the other hand, for non-ergodic channels, the ergodic capacity has no physical significance and instead the proper performance measure is the *outage probability* which is actually the CDF of the MI. In either ergodic or non-ergodic channels, the statistical distribution of MI is fairly useful in obtaining the corresponding performance measures, and more importantly it provides a more comprehensive view about MIMO fading channels than the ergodic capacity does. This motivates many recent studies on the distribution of MI.

However, because of the complexity of the statistical distribution of a Wishart matrix, the study on the statistical distribution of MI becomes quite involved. For Rayleigh fading channels, the CCDF of MI[2] of a MIMO system was first studied by Monte Carlo simulations in [15]. Although some heuristic discussions on the statistical distribution of the MI were provided in [15], it is quite time consuming to use Monte

[1]When the transmit signals are independent circularly symmetric Gaussian distributed with equal power at each transmit antenna, the MI is known as the capacity in [15–17], since this kind of transmit signals can achieve the ergodic capacity in independent and identically distributed Rayleigh fading channels.

[2]The MI is also called the conditional capacity in [11].

Carlo simulations and rather difficult to examine how the system parameters will affect the MI distribution through Monte Carlo simulations. Later, based on the central limit theorem, [16] proposed using the Gaussian distribution function to approximate the statistical distribution of MI, and obtained the variance of MI for MIMO Rayleigh fading channels. The Gaussian approximation is accurate for a large number of antennas. But when the number of antennas is small, which is in fact the case of practical interests, the Gaussian approximation becomes loose. To overcome this shortcoming of the Gaussian approximation, [17] and [21] derived the exact CDF of the MI for dual antenna MIMO systems in Rayleigh and Rician fading environments, respectively. Jacobian on the joint density of the eigenvalues of a Wishart matrix is applied to obtain the CDF of MI in [17] and [21]. However, the complexity of Jacobian increases exponentially with an increase in the number of antennas. This prohibits the application of the methods in [17] and [21] to a system with a larger number of antennas. More comprehensive studies on the statistical characterization of MI have been conducted in [18–20] for Rayleigh fading environments, and in [13, 14] for Rician fading environments. These results are valid for an arbitrary number of antennas. The characteristic function of MI was obtained in [18] by using a new expression for the joint distribution of the eigenvalues of a Wishart matrix. The moment generating function of MI was given in [19, 20] for Rayleigh fading environments, and in [14] for Rician fading environments. However, to obtain the CDF of MI for MIMO fading channels, one needs to take the inverse Laplace transform of the moment generating function, and need to resort to numerical integrations. In addition, the mean and variance of MI were also presented in [13] for single- and double-sided correlated Rician fading channels with an arbitrary-rank mean matrix, and in [14] for independent Rician fading channels, and then the Gaussian distribution was used to approximate the CCDF of MI [13, 14]. All of the aforementioned results on the distribution of the MI

rely on the joint distribution of the eigenvalues of a central or non-central Wishart matrix. To date, no explicit expression is obtained for the CCDF of MI for MIMO fading channels.

Instead of using the joint distribution of the eigenvalues of a Wishart matrix, as is common in literature, a new approach is proposed in this chapter to derive some upper and lower bounds on the statistical distribution, i.e., the CCDF, of the MI for MIMO Rician fading channels. We first bound the MI by the determinant or trace of a complex Wishart matrix. Then we derive the statistical distributions of the determinant and trace of a non-central complex Wishart matrix. Finally, we use the derived distributions to obtain upper and lower bounds on the CCDF of the MI in Rician fading environments. Our approach to the CCDF of the MI greatly simplifies the calculation procedure, and provides easy and accurate ways to deal with MI related calculations for MIMO fading channels. The derived bounds on the CCDF of the MI are expressed in simple forms, and show explicitly the effects of the system parameters on the distribution of the MI. Furthermore, the bounds are valid for an arbitrary number of antennas, and in particular, they are tight for small numbers of antennas, which is generally the case in practical systems. In addition, since Rician fading reduces to Rayleigh fading when LOS does not exist, our results can be readily applied to MIMO systems in Rayleigh fading environments [82]. The statistical distribution of the determinant of a complex Wishart matrix derived here is an important result in its own right, and may have applications beyond the MI distribution analysis.

The remainder of this chapter is organized as follows. In Section 2.2, we describe the system model and the assumptions. In Section 2.3, we present some bounds on the MI of the MIMO system. In Sections 2.4 and 2.5, we investigate the statistical distribution of the determinant and the trace of a non-central Wishart matrix, respectively. In Section 2.6, new and simple bounds on the CCDF of the MI

are obtained for MIMO Rician fading channels. In Section 2.7, the results for MIMO Rayleigh fading channels are presented as a special case of Rician fading channels. Comparisons of the results are shown in Section 2.8. Finally, we conclude with Section 2.9.

2.2 MIMO System Model

Consider a single user MIMO system with N transmit and M receive antennas. The M-dimensional received vector is mathematically represented as [2, 15]

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

where $\boldsymbol{H}$ is the $M \times N$ channel matrix with the mn-th entry h_{mn} describing the channel gain between the n-th transmit antenna and the m-th receive antenna, $\boldsymbol{x}$ is an N-dimensional transmitted vector, and $\boldsymbol{n}$ is AWGN. We assume $\mathbb{E}[|h_{mn}|^2] = (1+K)\sigma^2$, $\mathbb{E}[\boldsymbol{n}\boldsymbol{n}^H] = N_0\boldsymbol{I}_M$ and $\mathbb{E}[\boldsymbol{x}^H\boldsymbol{x}] = E$. For a Rician fading environment, the channel matrix $\boldsymbol{H}$ can be decomposed into the sum of a deterministic (specular) matrix and a variable (scattered) matrix, i.e.,

$$\boldsymbol{H} = \sqrt{\sigma^2 K}\,\overline{\boldsymbol{H}} + \sqrt{\sigma^2}\,\widetilde{\boldsymbol{H}}, \quad K \geq 0, \tag{2.2}$$

where $\overline{\boldsymbol{H}}$ is a deterministic matrix and $\widetilde{\boldsymbol{H}}$ is a random matrix. The elements of $\widetilde{\boldsymbol{H}}$, $\{\tilde{h}_{mn}\}$, are independent and identically distributed (i.i.d.), complex, Gaussian random variables with zero mean and unit variance, i.e., $\tilde{h}_{mn} \sim \mathcal{CN}(0,1)$. The Rice factor, $K \geq 0$, represents the ratio of deterministic energy to scattered energy; $K = 0$ corresponds to Rayleigh fading; while $K \to \infty$ corresponds to non-fading channels. The mean matrix of $\boldsymbol{H}$ is $\mathbb{E}[\boldsymbol{H}] = \sqrt{\sigma^2 K}\overline{\boldsymbol{H}}$. Thus, the channel matrix $\boldsymbol{H}$ has a matrix variate complex Gaussian distribution, i.e.,

$$\boldsymbol{H} \sim \mathcal{CN}_{M,N}\left(\sqrt{\sigma^2 K}\,\overline{\boldsymbol{H}},\ \sigma^2\boldsymbol{I}_M \otimes \boldsymbol{I}_N\right). \tag{2.3}$$

We assume that the deterministic matrix $\overline{\boldsymbol{H}}$ is of rank one [10]. By applying singular value decomposition, an arbitrary rank-1 matrix $\overline{\boldsymbol{H}}$ can be decomposed as $\overline{\boldsymbol{H}} = \boldsymbol{U}\overline{\boldsymbol{M}}\boldsymbol{V}^H$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, and $\overline{\boldsymbol{M}} = [\boldsymbol{m}, \boldsymbol{0}, \cdots, \boldsymbol{0}]$ with both $\boldsymbol{m}$ and $\boldsymbol{0}$ being $M \times 1$ column vectors.

We consider the scenario where the receiver has perfect knowledge of the CSI, and the transmitter has no channel knowledge at all (neither CSI nor fading distribution). In this case, for any realization of $\boldsymbol{H}$, the MI, $I(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H})$, is maximized when the transmit signal is circularly symmetric, zero-mean, complex, Gaussian distributed. Only the covariance matrix $\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^H]$ of the capacity-achieving transmit signal depends on the fading distributions. For i.i.d. Rayleigh fading channels, the capacity-achieving signal covariance matrix is $\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^H] = (E/N)\boldsymbol{I}_N$. This kind of transmit signals is also robust for transmission over Rician fading channels [10, 13, 14]. Hence, we still consider circularly symmetric, zero-mean, complex, Gaussian distributed transmit signals with covariance matrix $\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^H] = (E/N)\boldsymbol{I}_N$ in the Rician fading case. Therefore, the MI in this case is given by [81, Eq. (4.10)]

$$\begin{aligned} I = I(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H}) &= \log_2 \det\left(\boldsymbol{I}_N + \gamma \boldsymbol{H}^H \boldsymbol{H}\right) \quad \text{bps/Hz} \\ &= \log_2 \det\left(\boldsymbol{I}_N + \sigma^2 \gamma \boldsymbol{Z}^H \boldsymbol{Z}\right) \quad \text{bps/Hz}, \end{aligned} \tag{2.4}$$

where $\gamma = E/(N_0)$ is the average SNR at each transmit antenna, and

$$\boldsymbol{Z} = \sqrt{\frac{1}{\sigma^2}}\boldsymbol{H} = \overline{\boldsymbol{H}}\sqrt{K} + \widetilde{\boldsymbol{H}}, \qquad \boldsymbol{Z} \sim \mathcal{CN}_{N,M}(\overline{\boldsymbol{H}}\sqrt{K}, \boldsymbol{I}_N \otimes \boldsymbol{I}_M).$$

Without loss of generality, we assume $M \geq N$. Let $\lambda_n (n = 1, 2, \ldots, N)$ denote the nonzero eigenvalues of $\boldsymbol{Z}^H\boldsymbol{Z}$. The MI in (2.4) can be expressed as [2], [81, Eq. (4.13)]

$$I = \log_2\left(\prod_{n=1}^{N}\left(1 + \sigma^2\gamma\lambda_n\right)\right). \tag{2.5}$$

If $M < N$, the expression for the MI in (2.5) and our results in this chapter remain the same, except for interchanging M with N.

Note that since the channel matrix $\boldsymbol{H}$ is random, the MI, I in (2.5), is also a RV. If the fading process is ergodic, a *Shannon capacity* exists and is given by $C = \mathbb{E}[I]$, where the expectation is with respect to the channel $\boldsymbol{H}$. If the fading process is non-ergodic, a proper performance measurement is the *outage probability* $P(I \leq R)$. Actually, in ergodic channels, the MI also has a statistical distribution function, e.g., the CDF. The statistical distribution of MI exists for both ergodic and non-ergodic fading processes. We will investigate the CCDF of the MI of a Rician fading MIMO system.

The MI (2.5) of the MIMO Rician fading channels is characterized by all the eigenvalues of $\boldsymbol{Z}^H\boldsymbol{Z}$. To study the statistical properties of the MI, the traditional method is to make use of the joint distribution of the eigenvalues of $\boldsymbol{Z}^H\boldsymbol{Z}$, which is highly complex in general. We will use a new approach to studying the statistics of the MI. Firstly, we will obtain some new bounds on the MI in the next section.

2.3 New Bounds on the MI of MIMO Channels

Making use of some inequalities and properties of the determinant and the trace of a matrix, we will obtain, in this section, some bounds on the MI in terms of the determinant or trace of $\boldsymbol{Z}^H\boldsymbol{Z}$, instead of the eigenvalues. Specifically, we derive a lower bound on the MI in terms of the determinant of $\boldsymbol{Z}^H\boldsymbol{Z}$, and an upper bound and a lower bound on the MI in terms of the trace of $\boldsymbol{Z}^H\boldsymbol{Z}$.

2.3.1 Determinant Bounds

We first derive a lower bound on the MI in terms of the determinant $\det(\boldsymbol{Z}^H\boldsymbol{Z})$. Applying the inequality [83, Eq. (15)]: $\prod_{n=1}^{N}(1+x_n) \geq (1+(\prod_{n=1}^{N}x_n)^{1/N})^N$ for

$x_n > 0$, to (2.5), we obtain a lower bound on the MI as

$$I \geq N \log_2 \Big(1 + \sigma^2 \gamma \Big(\prod_{n=1}^{N} \lambda_n \Big)^{1/N} \Big). \tag{2.6}$$

Since the determinant of a matrix is equal to the product of its non-zero eigenvalues: $\prod_{n=1}^{N} \lambda_n = \det(\boldsymbol{Z}^H \boldsymbol{Z})$, we can rewrite the above lower bound on the MI as

$$I \geq I_{\text{det}} := N \log_2 \Big(1 + \sigma^2 \gamma \det(\boldsymbol{Z}^H \boldsymbol{Z})^{1/N} \Big), \tag{2.7}$$

where I_{det} is the lower bound on the MI in terms of the determinant $\det(\boldsymbol{Z}^H \boldsymbol{Z})$.

2.3.2 Trace Bounds

We next derive an upper bound and a lower bound on the MI in terms of the trace $\text{tr}(\boldsymbol{Z}^H \boldsymbol{Z})$.

Upper Bound

Applying the arithmetic and geometric means inequality: $(\prod_{n=1}^{N} x_n)^{1/N} \leq \sum_{n=1}^{N} x_n / N$ for $x_n > 0$, to (2.5), we can upper bound the MI as

$$I \leq \log_2 \Big(\frac{1}{N} \sum_{n=1}^{N} \big(1 + \sigma^2 \gamma \lambda_n \big) \Big)^N = N \log_2 \Big(1 + \frac{\sigma^2 \gamma}{N} \sum_{n=1}^{N} \lambda_n \Big). \tag{2.8}$$

Since the trace of a matrix is equal to the sum of its eigenvalues: $\sum_{n=1}^{N} \lambda_n = \text{tr}(\boldsymbol{Z}^H \boldsymbol{Z})$, we can rewrite the above upper bound on the MI as

$$I \leq I_{\text{tr\_U}} := N \log_2 \Big(1 + \frac{\sigma^2 \gamma}{N} \text{tr}(\boldsymbol{Z}^H \boldsymbol{Z}) \Big), \tag{2.9}$$

where $I_{\text{tr\_U}}$ is the upper bound on the MI in terms of the trace $\text{tr}(\boldsymbol{Z}^H \boldsymbol{Z})$.

Lower Bound

Expanding $\prod_{n=1}^{N} (1 + \sigma^2 \gamma \lambda_n)$, we obtain the inequality

$$\prod_{n=1}^{N} (1 + \sigma^2 \gamma \lambda_n) \geq 1 + \sigma^2 \gamma \Big(\sum_{n=1}^{N} \lambda_n \Big) + (\sigma^2 \gamma)^N \Big(\prod_{n=1}^{N} \lambda_n \Big) \geq 1 + \sigma^2 \gamma \text{tr}(\boldsymbol{Z}^H \boldsymbol{Z}).$$

This allows us to lower bound the MI as

$$I \geq I_{\text{tr\_L}} := \log_2 \Big(1 + \sigma^2 \gamma \text{tr}(\boldsymbol{Z}^H \boldsymbol{Z})\Big), \tag{2.10}$$

where $I_{\text{tr\_L}}$ is the lower bound on the MI in terms of the trace $\text{tr}(\boldsymbol{Z}^H \boldsymbol{Z})$. The lower bound $I_{\text{tr\_L}}$ is tight only at low SNRs, because the inequality (2.10) ignores the terms containing second and higher powers of SNRs. In the low SNR regime, it was proved in [6] that the lower bound $I_{\text{tr\_L}}$ is strictly tighter than the lower bound I_{det}.

Since the MI can be bounded by the quantities in terms of the determinant or trace of the Wishart matrix $\boldsymbol{Z}^H \boldsymbol{Z}$, in order to study the statistics of the MI, we will study the statistics of the determinant of the Wishart matrix in the next section.

2.4 Statistical Distribution of the Determinant of a Wishart Matrix

In this section, we derive the PDF and the CCDF of the determinant of a non-central Wishart matrix, which will be used to evaluate the statistical distribution of the MI in Section 2.6.

2.4.1 Determinant of a Wishart matrix

The matrix $\boldsymbol{Z}$ has a matrix variate complex Gaussian distribution, i.e., $\boldsymbol{Z} \sim \mathcal{CN}_{M,N}(\overline{\boldsymbol{H}}\sqrt{K}, \boldsymbol{I}_M \otimes \boldsymbol{I}_N)$, and $\boldsymbol{Z}^H \boldsymbol{Z}$ is called a noncentral complex Wishart matrix. Note that in Section 2.2 we have the singular value decomposition $\overline{\boldsymbol{H}} = \boldsymbol{U}\overline{\boldsymbol{M}}\boldsymbol{V}^H$. Defining $\boldsymbol{W} = \boldsymbol{U}^H \boldsymbol{Z} \boldsymbol{V}$, we have $\det(\boldsymbol{Z}^H \boldsymbol{Z}) = \det(\boldsymbol{W}^H \boldsymbol{W})$ and $\boldsymbol{W} = \overline{\boldsymbol{M}}\sqrt{K} + \boldsymbol{U}^H \widetilde{\boldsymbol{H}} \boldsymbol{V}$. Since the entries of $\widetilde{\boldsymbol{H}}$ are i.i.d. Gaussian RVs with zero mean and unit variance, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, the distribution of $\boldsymbol{U}^H \widetilde{\boldsymbol{H}} \boldsymbol{V}$ is the same as the distribution of $\widetilde{\boldsymbol{H}}$. Therefore, the matrix $\boldsymbol{W}$ also has a matrix variate complex

Gaussian distribution, i.e., $\boldsymbol{W} \sim \mathcal{CN}_{M,N}(\overline{\boldsymbol{M}}\sqrt{K}, \boldsymbol{I}_M \otimes \boldsymbol{I}_N)$, and $\boldsymbol{W}^H\boldsymbol{W}$ is also a noncentral complex Wishart matrix.

The Bartlett's decomposition of the noncentral Wishart matrix is written as $\boldsymbol{W}^H\boldsymbol{W} = \boldsymbol{T}^H\boldsymbol{T}$, where $\boldsymbol{T}$ is an upper-triangular matrix with positive diagonal elements. It has been show in [12, Section II-C], [84, Section III] and [85, Theorem 10.3.8] that the elements of $\boldsymbol{T}$, $t_{mn}(1 \leq m \leq n \leq N)$, are all independent RVs. Moreover, $|t_{11}|^2$ is non-central chi-square distributed with $2M$ degrees of freedom and non-centrality parameter $\delta = K\boldsymbol{m}^H\boldsymbol{m} = K\|\overline{\boldsymbol{M}}\|_F^2 = K\|\overline{\boldsymbol{H}}\|_F^2$, $|t_{nn}|^2$ $(n = 2, \ldots, N)$ is central chi-square distributed with $2(M - n + 1)$ degrees of freedom, and $t_{mn}(1 \leq m < n \leq N)$ is complex Gaussian distributed with zero mean and unit variance, i.e., $|t_{11}|^2 \sim \mathcal{X}^2_{2M}(\delta)$, $|t_{nn}|^2 \sim \mathcal{X}^2_{2(M-n+1)}, (n = 2, \ldots, N)$, and $t_{mn} \sim \mathcal{CN}(0, 1), (1 \leq m < n \leq N)$.

Since the determinant of a triangular matrix equals the product of its diagonal elements, we can express the determinant of the Wishart matrix $\boldsymbol{W}^H\boldsymbol{W}$ as

$$\det(\boldsymbol{W}^H\boldsymbol{W}) = \det(\boldsymbol{T}^H\boldsymbol{T}) = \det(\boldsymbol{T}^H)\det(\boldsymbol{T}) = \prod_{n=1}^{N} |t_{nn}|^2. \tag{2.11}$$

From the above equation, we can see that $\det(\boldsymbol{Z}^H\boldsymbol{Z}) = \det(\boldsymbol{W}^H\boldsymbol{W})$ is the product of N independent RVs, one of which is non-central chi-square distributed and the remaining $(N - 1)$ ones of which are central chi-square distributed. Let X_n denote $|t_{nn}|^2$, i.e., $X_n = |t_{nn}|^2, (n = 1, 2, \ldots, N)$, and thus $X_1 \sim \mathcal{X}^2_{2M}(\delta)$, $X_n \sim \mathcal{X}^2_{2(M-n+1)}, (n = 2, \ldots, N)$. The PDFs of the X_ns are given by [86, Eqs. (2.118) and (2.110)]

$$\begin{aligned} f_{X_1}(x_1) &= \left(\frac{x_1}{\delta}\right)^{(M-1)/2} e^{-(\delta+x_1)} I_{M-1}(2\sqrt{\delta x_1}), \\ f_{X_n}(x_n) &= \frac{1}{(M-n)!} x_n^{M-n} e^{-x_n}, \qquad n = 2, \ldots, N, \end{aligned} \tag{2.12}$$

where $I_\alpha(x)$ is the αth-order modified Bessel function of the first kind. By using the series representation of $I_\alpha(x)$ [86, Eq. (2.120)], the PDF of the non-central chi-square

RV X_1 can be expressed as a weighted sum of central chi-square PDFs with Poisson distributed weights [87, Corollary 1.3.5],

$$f_{X_1}(x_1) = \sum_{k=0}^{\infty} e^{-\delta} \frac{\delta^k}{k!} \cdot \frac{1}{(M-1+k)!} x_1^{M-1+k} e^{-x_1}, \tag{2.13}$$

where $e^{-\delta}\delta^k/k!$ is a Poisson probability with mean δ. When the non-centrality parameter δ equals zero, (2.13) only has the term corresponding to $k = 0$ and reduces to the PDF of a central chi-square RV with $2M$ degrees of freedom.

Let $\det(\boldsymbol{Z}^H\boldsymbol{Z}) = Y = \prod_{n=1}^{N} X_n$, and $f_Y(y)$ denote the PDF of Y. We now study the statistical distribution, i.e., PDF and CCDF, of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ through studying a product of independent RVs.

2.4.2 Distribution functions in Meijer G-functions

In this subsection, we will derive the PDF and CCDF of the determinant of a noncentral Wishart matrix in terms of the Meijer G-functions by using the Mellin integral transform.

The Mellin integral transform is typically used to study the distribution of the products or quotients of RVs [87–89], because the Mellin transform of the PDF of a product of independent RVs can be written as a product of the Mellin transforms of the PDFs of individual RVs.

A real function, $f(x)$, is *Mellin transformable* if the following conditions are satisfied: 1) $f(x)$ is defined only for $x \geq 0$ and single valued almost everywhere for $x \geq 0$; 2) the integral $\int_0^\infty x^{k-1}|f(x)|dx$ is bounded for some real value $k > 0$. The Mellin transform of $f(x)$ is given by [87, Ch. 2.8]

$$M_s\big(f(x)\big) = \int_0^\infty x^{s-1} f(x) dx, \tag{2.14}$$

and inverse Mellin transform is given by

$$f(x) = M_s^{-1}\left[M_s\big(f(x)\big)\right] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s} M_s\big(f(x)\big)ds, \tag{2.15}$$

where the path of integration is any line parallel to the imaginary axis and lying within the strip of analyticity of $M_s\big(f(x)\big)$. Eqs. (2.14) and (2.15) constitute a transform pair, and $f(x)$ is uniquely determined by its Mellin transform $M_s\big(f(x)\big)$.

It is easy to see that $f_Y(y)$ and $f_{X_n}(x_n)$ are Mellin transformable. By the definition of Mellin transform, it is clear that $M_s\big(f_Y(y)\big) = \mathbb{E}[y^{s-1}]$ and $M_s\big(f_{X_n}(x_n)\big) = \mathbb{E}[x_n^{s-1}]$. Since $\{X_n\}_{n=1}^N$ are independent RVs, according to the properties of Mellin transform [87, Ch. 4.3], we can write the Mellin transform of $f_Y(y)$ as a product of the Mellin transforms of $\{f_{X_n}(x_n)\}_{n=1}^N$, i.e.,

$$M_s\big(f_Y(y)\big) = \prod_{n=1}^{N} M_s\big(f_{X_n}(x_n)\big).$$

Based on the PDF of X_n given in (2.12) and (2.13), we can easily compute the Mellin transform of $f_{X_n}(x_n)$, and further obtain the Mellin transform of $f_Y(y)$ as

$$M_s\big(f_Y(y)\big) = \prod_{n=1}^{N} M_s\big(f_{X_n}(x_n)\big) = \frac{e^{-\delta}}{A}\left(\sum_{k=0}^{\infty}\frac{\delta^k}{k!}\frac{\Gamma(s+M-1+k)}{(M-1+k)!}\right)\cdot\prod_{n=2}^{N}\Gamma(s+M-n), \tag{2.16}$$

where $A = \prod_{n=2}^{N}(M-n)!$. Taking the inverse Mellin transform of $M_s\big(f_Y(y)\big)$, we obtain the PDF $f_Y(y)$ as

$$f_Y(y) = \frac{e^{-\delta}}{A}\sum_{k=0}^{\infty}\frac{\delta^k}{k!(M-1+k)!}G_{0,N}^{N,0}\left(y\big|_{M-1+k,M-2,M-3,\cdots,M-N}^{-}\right), \tag{2.17}$$

where $G_{p,q}^{m,n}(x\,|_{\cdots}^{\cdots})$ is the Meijer G-function. The Meijer G-function, a generalization of the generalized hypergeometric function, can be defined by the contour integral:

$$G_{p,q}^{m,n}\left(x\,\big|_{b_1,b_2,\cdots,b_q}^{a_1,a_2,\cdots,a_p}\right) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s}\frac{\prod_{j=1}^{m}\Gamma(s+b_j)\prod_{j=1}^{n}\Gamma(1-a_j-s)}{\prod_{j=n+1}^{p}\Gamma(s+a_j)\prod_{j=m+1}^{q}\Gamma(1-b_j-s)}ds. \tag{2.18}$$

The CCDF of Y can be obtained by using the Mellin transform of $f_Y(y)$, i.e., $M_s\big(P(Y>y)\big) = s^{-1}M_{s+1}\big(f_Y(y)\big)$ [87, Eq. (4.3.7)]. Thus, we can obtain the CCDF of Y as

$$P(Y>y) = M_s^{-1}\left[s^{-1}M_{s+1}\big(f_Y(y)\big)\right] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} y^{-s}s^{-1}M_{s+1}\big(f_Y(y)\big)ds$$
$$= \frac{e^{-\delta}}{A}\sum_{k=0}^{\infty}\frac{\delta^k}{k!(M-1+k)!}G_{1,N+1}^{N+1,0}\Big(y\,\Big|_{0,M+k,M-1,M-2,\cdots,M-N+1}^{1}\Big). \quad (2.19)$$

The PDF and CCDF of Y are expressed by the Meijer G-functions. Even though the Meijer G-functions can be computed in some mathematical software packages, evaluating the Meijer G-functions normally involves a heavy computational load and is time consuming. We next provide some bounds on the PDF and CCDF of Y in series form.

2.4.3 Bounds on distribution functions in series form

In this subsection, we provide some bounds on the PDF and CCDF of Y in series form instead of Meijer G-functions. Although these bounds are not the exact PDF and CCDF of Y, they only involve some simple functions, and can be evaluated faster than the Meijer G-functions.

The joint PDF of $X_1,\cdots,X_N$ is $f_{X_1\cdots X_N}(x_1,\cdots,x_N) = \prod_{n=1}^{N} f_{X_n}(x_n)$, and hence the PDF of Y can be written as

$$f_Y(y) = \int\cdots\int_D \frac{1}{\big(\prod_{n=1}^{N-1}x_n\big)}\left(\prod_{n=1}^{N-1}f_{X_n}(x_n)\right)f_{X_N}\left(\frac{y}{\prod_{n=1}^{N-1}x_n}\right)d\boldsymbol{x}, \quad (2.20)$$

where the multiple integral domain is $D = \{0 \le x_1, x_2, \cdots, x_{N-1} < \infty\}$ and $d\boldsymbol{x} = dx_1dx_2\cdots dx_{N-1}$. Substituting the PDFs of X_n's from (2.12) and (2.13) into (2.20), we obtain the PDF of Y as

$$f_Y(y) = \frac{e^{-\delta}}{A}\sum_{k=0}^{\infty}\frac{\delta^k y^{M-N}}{k!(M-1+k)!}\int\cdots\int_D x_1^k\left(\prod_{n=1}^{N-1}x_n^{N-1-n}\right)\exp\left(-\sum_{n=1}^{N-1}x_n\frac{y}{\prod_{n=1}^{N-1}x_n}\right)d\boldsymbol{x}. \quad (2.21)$$

The CCDF of Y can be calculated as $P(Y > y) = \int_y^\infty f_Y(t)dt$, and is expressed as

$$P(Y > y) = \frac{e^{-\delta}}{A} \sum_{k=0}^{\infty} \frac{\delta^k}{k!} \int \cdots \int_D x_1^k \Big(\prod_{n=1}^{N-1} x_n^{M-n} \Big) \frac{\Gamma\big(M-N+1, \frac{y}{\prod_{n=1}^{N-1} x_n}\big)}{(M-1+k)!} \exp\Big(-\sum_{n=1}^{N-1} x_n \Big) d\boldsymbol{x}$$
$$= \sum_{k=0}^{\infty} \frac{e^{-\delta}\delta^k (M-N)!}{Ak!(M-1+k)!} \sum_{r=0}^{M-N} \frac{y^r}{r!} \int \cdots \int_D x_1^k \Big(\prod_{n=1}^{N-1} x_n^{M-n-r} \Big) \exp\Big(-\sum_{n=1}^{N-1} x_n - \frac{y}{\prod_{n=1}^{N-1} x_n} \Big) d\boldsymbol{x}, \tag{2.22}$$

in which $\Gamma(\alpha, x)$ is the upper incomplete gamma function, i.e., $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t} dt$. When α is an integer n, we have $\Gamma(n, x) = (n-1)! e^{-x} \sum_{k=0}^{n-1} x^k / k!$.

When $N = 2$, by making use of the integral formula

$$\int_0^\infty x^n \exp(-x - \frac{y}{x}) dx = 2y^{\frac{1+n}{2}} K_{1+n}(2\sqrt{y}),$$

the PDF (2.21) and CCDF (2.22) of Y are reduced, respectively, to

$$f_Y(y) = \frac{e^{-\delta} y^{M-2}}{(M-2)!} \sum_{k=0}^{\infty} \frac{\delta^k}{k!} \frac{2y^{\frac{1+k}{2}}}{(M-1+k)!} K_{1+k}(2\sqrt{y}),$$
$$P(Y > y) = e^{-\delta} 2y^{M/2} \sum_{k=0}^{\infty} \frac{\delta^k}{k!} \frac{y^{k/2}}{(M-1+k)!} \sum_{r=0}^{M-2} \frac{y^{r/2}}{r!} K_{M-r+k}(2\sqrt{y}), \tag{2.23}$$

where $K_n(x)$ is the nth-order modified Bessel function of the second kind. In fact, (2.23) is the series form of (2.19) when $N = 2$. When $N > 2$, there is no closed-from expression for the multiple integral in (2.21) and (2.22). However, we derive one inequality which can be applied repeatedly to obtain lower bounds on (2.21) and (2.22). Applying the following inequality (see Appendix A for the proof):

$$\int_0^\infty x^n \exp\big(-x - \frac{y}{x}\big) dx \geq n! \exp\big(-\frac{y}{n}\big), \qquad n = 1, 2, 3, \cdots, \tag{2.24}$$

for $y \geq 0$ to (2.22) $(N-1)$ times, we can obtain the following lower bound on (2.22):

$$P(Y > y) \geq e^{-\delta} \sum_{k=0}^{\infty} \frac{\delta^k}{k!} \sum_{r=0}^{M-N} \frac{(M-1+k-r)!}{(M-1+k)!} \Big(\prod_{n=2}^{N-1} \frac{(M-n-r)!}{(M-n)!} \Big) \frac{y^r}{r!}$$
$$\exp\left(-\frac{(M-N-r)! y}{(M-2-r)!(M-1+k-r)} \right). \tag{2.25}$$

This is the lower bound on the CCDF of Y for $N \geq 2$, and (2.23) is the exact CCDF for $N = 2$. We will use these results to evaluate the CCDF of the MI of MIMO Rician fading channels.

2.5 Statistical Distribution of the Trace of a Wishart Matrix

In this section, we study the distribution of the trace of a Wishart matrix, which will be used to evaluate the CCDF of the MI in the next section. The trace of $\boldsymbol{Z}^H\boldsymbol{Z}$ can be expressed as a sum of the magnitude square of all the elements of a Gaussian variate matrix, i.e.,

$$\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z}) = \|\boldsymbol{Z}\|_F^2 = \frac{1}{\sigma^2}\|\boldsymbol{H}\|_F^2 = \frac{1}{\sigma^2}\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{mn}|^2. \tag{2.26}$$

It can be easily seen that $\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z})$ is non-central chi-square distributed with $2MN$ degrees of freedom, variance $1/2$ and non-centrality parameter $\delta = K\|\overline{\boldsymbol{H}}\|_F^2$, i.e., $\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z}) \sim \mathcal{X}^2_{2MN}(\delta)$. Thus, the CCDF of $\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z})$ is given by [86, Eq. (2.124)]

$$P\big(\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z}) > z\big) = Q_{MN}\big(\sqrt{2\delta}, \sqrt{2z}\big) = e^{-\delta-z}\sum_{k=0}^{\infty}\frac{\delta^k}{k!}\sum_{r=0}^{MN-1+k}\frac{z^r}{r!}, \tag{2.27}$$

where $Q_m(a,b)$ is the generalized Marcum Q-function of order m, i.e., $Q_m(a,b) = \int_b^\infty x(x/a)^{m-1}e^{-(x^2+a^2)/2}I_{m-1}(ax)dx$.

When the non-centrality parameter δ equals zero, the CCDF (2.27) reduces to

$$P\big(\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z}) > z\big) = \frac{1}{(MN-1)!}\Gamma(MN, z) = e^{-z}\sum_{r=0}^{MN-1}\frac{z^r}{r!}, \tag{2.28}$$

where $\Gamma(a,x)$ is the upper incomplete Gamma function, i.e., $\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}dt$.

2.6 New Approaches to CCDF of the MI of MIMO Rician fading Channels

So far, the evaluations of the CCDF or CDF of the MI are based on the joint distribution of the eigenvalues of $\boldsymbol{Z}^H\boldsymbol{Z}$. The available expressions for the CCDF or CDF of the MI either are the inverse Laplace transform of the moment generating function [14], or are approximated by Gaussian distribution with the mean and variance of the MI obtained in [14]. These expressions are quite difficult to evaluate. In this section, we will provide some new and simple approaches to investigating the CCDF of the MI based on the CCDF of the determinant or trace of the noncentral Wishart matrix. We will present some new and simple, upper and lower bounds on the CCDF of the MI of MIMO Rician fading channels.

2.6.1 Determinant Bounds

According to the lower bound (2.7) on the MI, the CCDF of the MI is lower bounded as

$$P(I > R) \geq P_{\det}(R) := P(I_{\det} > R) = P\big(\det(\boldsymbol{Z}^H\boldsymbol{Z}) > a^N(2^{R/N}-1)^N\big)\,, \quad (2.29)$$

where $a = 1/(\sigma^2\gamma)$. We will derive here some lower bounds on the CCDF of the MI using the distribution of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ given in (2.19), (2.23) and (2.25).

Bounds in Meijer G-functions

Applying (2.19) to (2.29), we can obtain a lower bound on the CCDF of the MI as

$$P_{\det}(R) = \frac{e^{-\delta}}{A}\sum_{k=0}^{\infty}\frac{\delta^k}{k!(M-1+k)!}G_{1,N+1}^{N+1,0}\left(a^N(2^{R/N}-1)^N\,\Big|_{0,M+k,M-1,M-2,\cdots,M-N+1}^{1}\right). \quad (2.30)$$

Although the lower bound $P_{\det}(R)$ in (2.30) involves infinite series, the sum of infinite series will converge very fast, and it will be illustrated later in Section 2.8. We also will illustrate later that calculating the lower bound $P_{\det}(R)$ in (2.30) is more computational efficient than using Monte Carlo simulations.

Bounds in series form

When $N = 2$, the CCDF of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ can be expressed as (2.23). Thus, applying (2.23) to (2.29), we obtain the lower bound on the CCDF of the MI for $N = 2$ as

$$P_{\det}(R) = e^{-\delta}2a^M(2^{R/N}-1)^M\sum_{k=0}^{\infty}\frac{\delta^k}{k!}\frac{a^k(2^{R/N}-1)^k}{(M-1+k)!}\times \left[\sum_{r=0}^{M-2}\frac{a^r}{r!}(2^{R/N}-1)^r K_{M-r+k}\big(2a(2^{R/N}-1)\big)\right]. \quad (2.31)$$

Since (2.25) is also a lower bound on the CCDF of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ for $N \geq 2$, we can apply (2.25) to (2.29), and obtain another lower bound on the CCDF of the mutual information for $N \geq 2$ as

$$P_{\det}(R) \geq P_{\det 1}(R) := e^{-\delta}\sum_{k=0}^{\infty}\frac{\delta^k}{k!}\sum_{r=0}^{M-N}\frac{(M-1+k-r)!}{r!(M-1+k)!}\left(\prod_{n=2}^{N-1}\frac{(M-n-r)!}{(M-n)!}\right) a^{rN}(2^{R/N}-1)^{rN}\exp\left(-\frac{(M-N-r)!a^N(2^{R/N}-1)^N}{(M-2-r)!(M-1+k-!r)}\right). \quad (2.32)$$

When $N = 2$, both (2.31) and (2.32) can serve as the lower bounds for the CCDF of the MI. However, compared with (2.32), (2.31) is tighter though more complicated, since it involves modified Bessel functions.

2.6.2 Trace Bounds

We will derive one upper bound and one lower bound on the CCDF of the MI based on the properties of $\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{Z})$.

Upper Bound

According to the upper bound (2.9) on the MI, the CCDF of the MI is upper bounded as

$$P(I > R) \leq P_{\text{tr\_U}}(R) := P(I_{\text{tr\_U}} > R) = P\big(\text{tr}(\boldsymbol{Z}^H \boldsymbol{Z}) > aN(2^{R/N} - 1)\big). \quad (2.33)$$

By applying (2.27) to (2.33), the upper bound on the CCDF of the MI in (2.33) is computed as

$$P_{\text{tr\_U}}(R) = Q_{MN}\Big(\sqrt{2\delta}, \sqrt{2aN(2^{R/N} - 1)}\Big). \quad (2.34)$$

Lower Bound

According to the lower bound (2.10) on the MI, the CCDF of the MI is lower bounded as

$$P(I > R) \leq P_{\text{tr\_L}}(R) := P(I_{\text{tr\_L}} > R) = P\big(\text{tr}(\boldsymbol{Z}^H \boldsymbol{Z}) > a(2^R - 1)\big). \quad (2.35)$$

With (2.27) and (2.35), the lower bound on the CCDF of the MI in (2.35) is computed as

$$P_{\text{tr\_L}}(R) = Q_{MN}\big(\sqrt{2\delta}, \sqrt{2a(2^R - 1)}\big). \quad (2.36)$$

The lower bound (2.36) is tight only at low SNR, because (2.10) ignores the terms containing higher powers of SNR. In the low SNR regime, the lower bound $P_{\text{tr\_L}}(R)$ is strictly tighter than the lower bound $P_{\text{det}}(R)$ since $I_{\text{tr\_L}}$ is strictly tighter than I_{det} as indicated in Section 2.3.

The upper bound $P_{\text{tr\_U}}(R)$ and lower bound $P_{\text{tr\_L}}(R)$ on the CCDF of the MI are expressed in simple closed forms.

2.7 CCDF of the MI of MIMO Rayleigh Fading Channels

Rayleigh fading is a special case of Rician fading with the Rice factor $K = 0$. Our results derived for the Rician fading case in the previous sections can be readily applied to the Rayleigh fading case. In this section, we will explicitly give the statistical properties of the MI for a MIMO system in Rayleigh fading environments.

For Rayleigh fading, the non-centrality parameter δ equals zero, and our bounds obtained in Section 2.6 on the CCDF of the MI can be simplified. The lower bound (2.30) on the CCDF of the MI is reduced to

$$P_{\det}(R) = \frac{1}{\prod_{n=1}^{N}(M-n)!} G_{1,N+1}^{N+1,0}\left(a^N(2^{R/N}-1)^N \,|_{0,M,M-1,M-2,\cdots,M-N+1}^{1}\right) \quad (2.37)$$

For $N = 2$, the lower bound (2.31) on the CCDF of the MI is also reduced to

$$P_{\det}(R) = \frac{2a^M(2^{R/N}-1)^M}{(M-1)!} \sum_{r=0}^{M-2} \frac{a^r}{r!}(2^{R/N}-1)^r K_{M-r}(2a). \quad (2.38)$$

For $N \geq 2$, another lower bound on the CCDF of the MI for $N \geq 2$ is

$$P_{\det}(R) \geq P_{\det 1}(R) = \sum_{r=0}^{M-N} \frac{(M-1-r)!}{r!(M-1)!}\left(\prod_{n=2}^{N-1}\frac{(M-n-r)!}{(M-n)!}\right) a^{rN}(2^{R/N}-1)^{rN} \exp\left(-\frac{(M-N-r)!a^N(2^{R/N}-1)^N}{(M-2-r)!(M-1-r)}\right). \quad (2.39)$$

We can see that for Rayleigh fading, the determinant bounds on the CCDF of MI are simple. It involves evaluation of one Meijer-G function or a finite exponential series.

In Rayleigh fading, the upper bound (2.34), and the lower bound (2.36) on the CCDF of the MI reduce, respectively, to

$$P_{\text{tr\_U}}(R) = \frac{1}{(MN-1)!}\Gamma\big(MN, aN(2^{R/N}-1)\big), \quad (2.40)$$

$$P_{\text{tr\_L}}(R) = \frac{1}{(MN-1)!}\Gamma\big(MN, a(2^{R}-1)\big). \quad (2.41)$$

The trace bounds on the CCDF of the MI are expressed by incomplete Gamma functions, and are easy to evaluate.

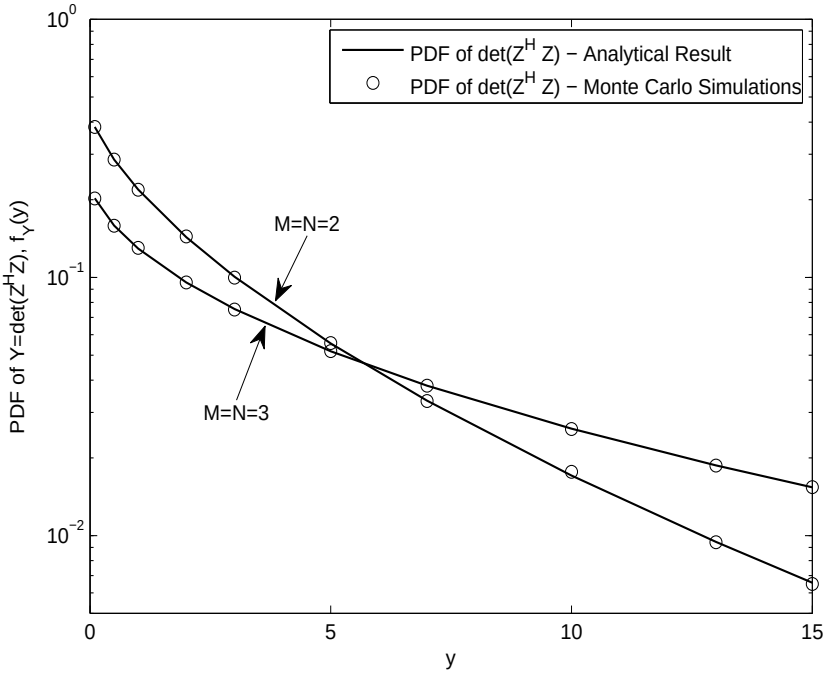

Figure 2.1: PDF of the determinant of a noncentral complex Wishart matrix with $K = 0$ dB and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

2.8 Numerical Results

In this section, we first compare the analytical results on the statistical distribution of the determinant of a non-central complex Wishart matrix with Monte Carlo simulations. Then, we compare our bounds on the CCDF of the MI with the Gaussian approximation results using the first and second moments of the MI derived in [14] , and the CCDF results obtained from Monte Carlo simulations. The Monte Carlo simulations results are obtained by generating 10^6 realizations of $\widetilde{\boldsymbol{H}}$ and then evaluating the statistics of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ or I. The average received SNR at each receive antenna is defined as $\eta = (1+K)\sigma^2 E/N_0$.

Fig. 2.1 plots the PDF of the determinant of the noncentral complex Wishart matrix for different matrix dimensions. The analytical result for the PDF of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$

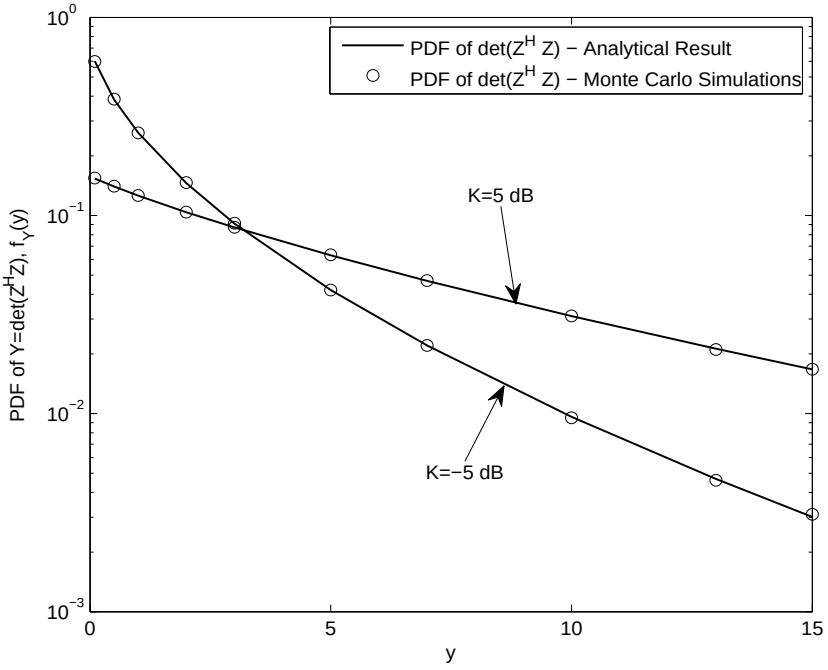

Figure 2.2: PDF of the determinant of a noncentral complex Wishart matrix with $M = N = 2$ and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

is obtained by evaluating $f_Y(y)$ in (2.17). As for the summation of infinite series in $f_Y(y)$ in (2.17), we only sum over the first 15 terms of the infinite series when $K = 0$ dB as in Fig. 2.1, and we can observe that very good accuracy is already achieved compared with Monte Carlo simulations. Fig. 2.2 plots the PDF of the determinant of the noncentral complex Wishart matrix for different Rice factors. As for the summation of infinite series in (2.17), we only sum over the first 10 terms of the infinite series when $K = -5$ dB and the first 30 terms of the infinite series when $K = +5$ dB. We plot the CCDF of the determinant of the noncentral complex Wishart matrix in Fig. 2.3 and Fig. 2.4. The analytical result for the CCDF of $\det(\boldsymbol{Z}^H\boldsymbol{Z})$ is obtained by evaluating (2.19).

Fig. 2.5 plots the CCDF of the MI versus the rate threshold R for different SNRs

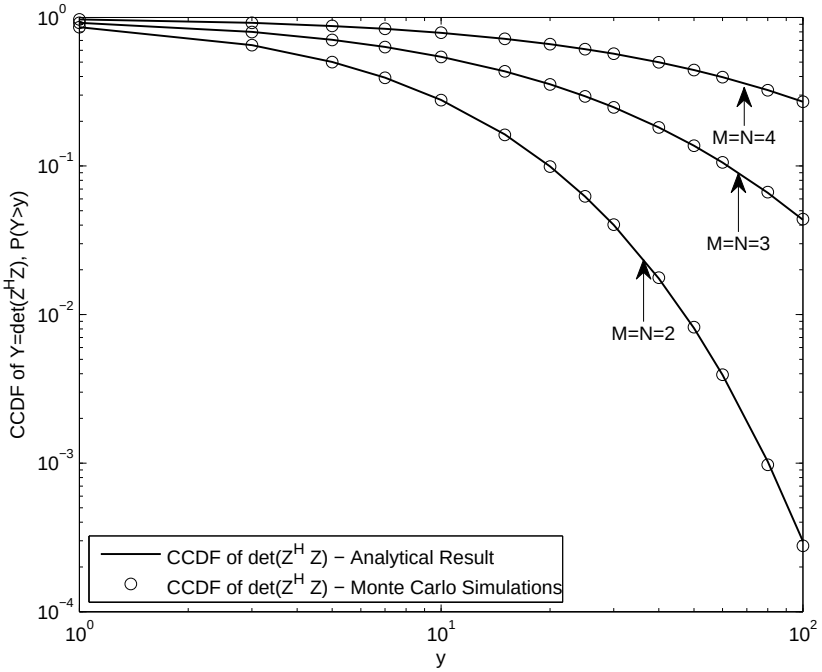

Figure 2.3: CCDF of the determinant of a noncentral complex Wishart matrix with $K = 5$ dB and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

when $M = N = 2$, $K = 0$ dB. As can be seen from the figure, the lower bound P_{det} and the upper bound $P_{\text{tr\_U}}$ are fairly tight, and they can provide a good prediction for the distribution of the MI. Although P_{det} involves summation of Meijer-G functions, the CPU time used to compute P_{det} is only $1/3$ of the CPU time spent to get the MI CCDF through Monte Carlo simulations. In Fig 2.6, we plot the trace upper bound $P_{\text{tr\_U}}$ and lower bound $P_{\text{tr\_L}}$ for a low SNR ($\eta = -15$ dB).

Fig. 2.7 plots the CCDF of the MI versus the rate threshold R for various configurations of transmit and/or receive antennas. It shows that our lower bound, P_{det}, works well for small numbers of antennas. However, with an increase in the number of antennas, our bound becomes slightly looser, while the Gaussian approximation becomes tighter. In a practical cellular communication system, the number of antennas

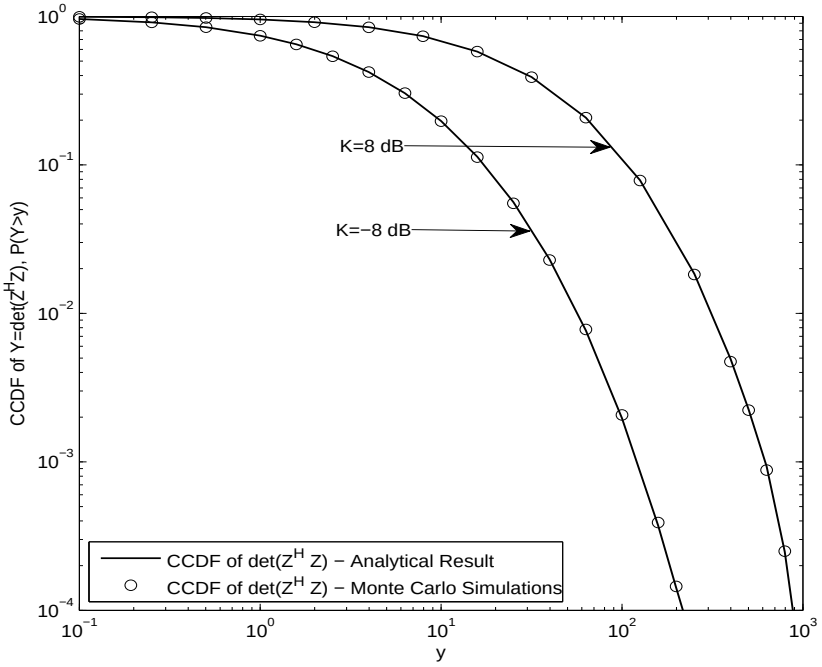

Figure 2.4: CCDF of the determinant of a noncentral complex Wishart matrix with $M = N = 3$ and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

at a mobile terminal is typically small, and hence our bound is a useful approximation to the distribution of the MI in such a situation. Comparing our bounds with the exact MI distribution in [21, Eq. (21)] which is valid only for two antennas, our results are much simpler. Therefore, when the number of antennas is small, our bounds are good approximations to the statistical distribution of the MI in Rician fading environments. We illustrate the tightness of the lower bound P_{det} at various Rice factors in Fig. 2.8. It can be observed that the lower bound P_{det} is tight for different Rice factors.

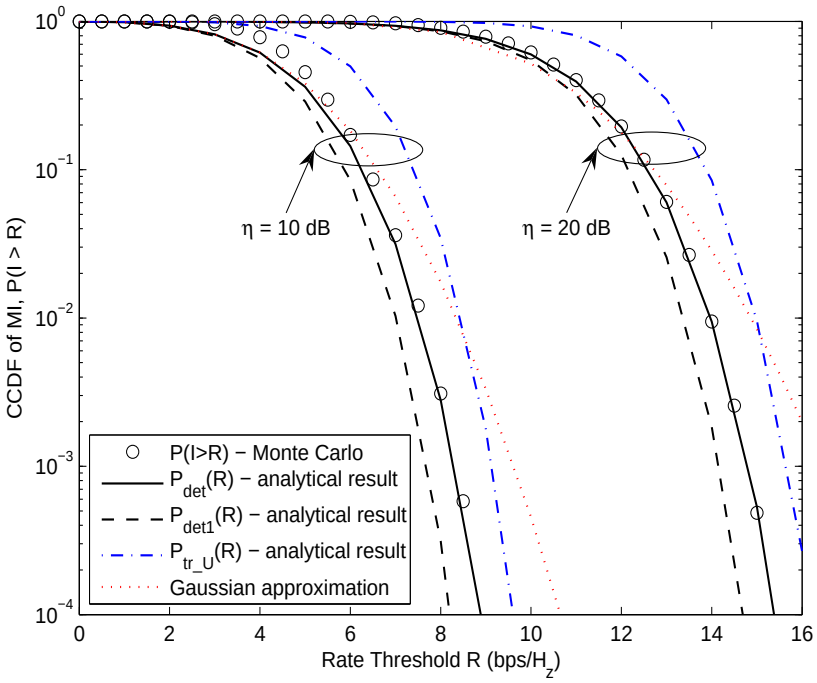

Figure 2.5: Bounds on the CCDF of the MI at various SNRs when $M = N = 2$, $K = 0$ dB , $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

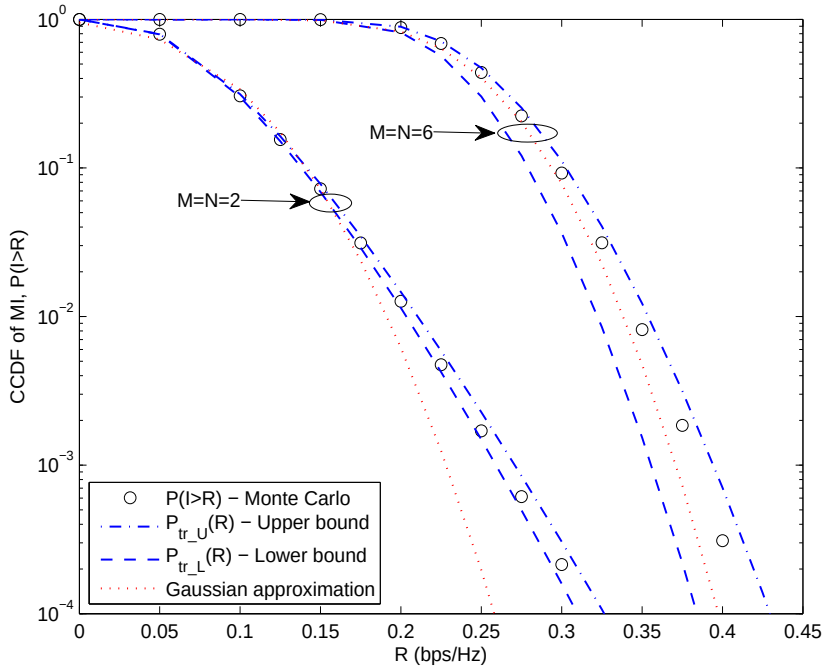

Figure 2.6: Bounds on the CCDF of the MI at a low SNR $\eta = -15$ dB with $K = -10$ dB, $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

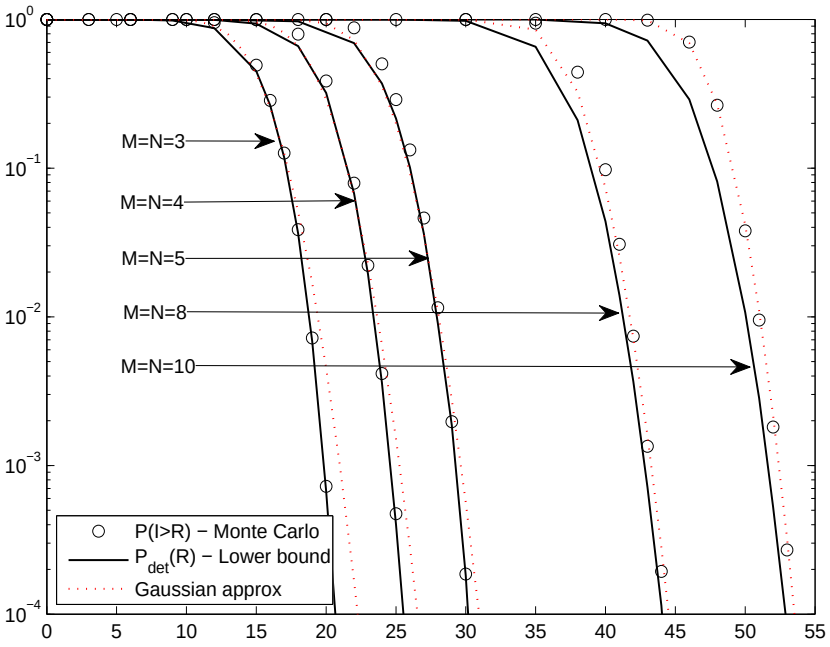

Figure 2.7: Bounds on the CCDF of the MI of MIMO channels with $K = 0$ dB, $\|\overline{\boldsymbol{H}}\|_F^2 = M$, and $\eta = 20$ dB.

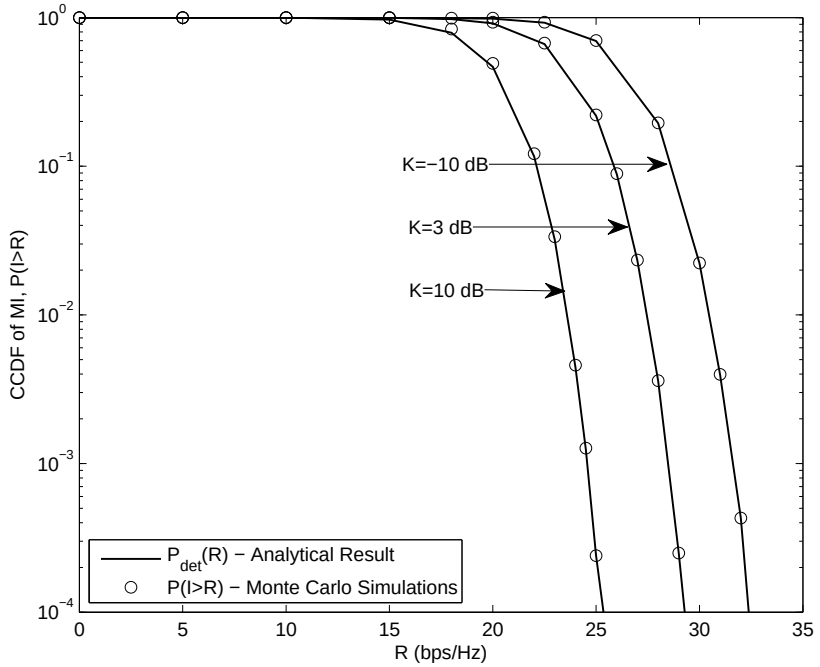

Figure 2.8: Bounds on the CCDF of the MI of MIMO channels with $M = N = 3$, $\eta = 30$ dB, and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

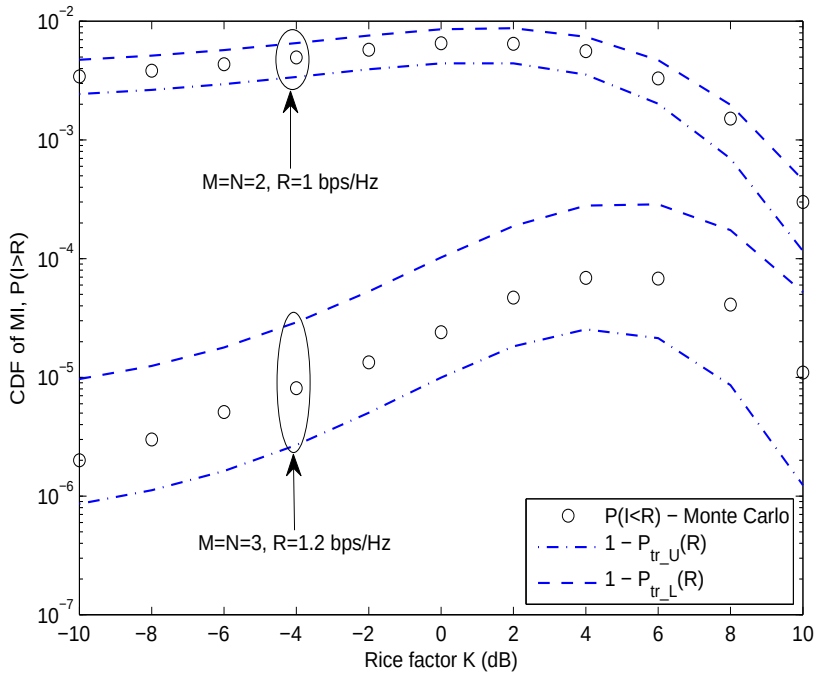

Figure 2.9: Bounds on the CDF of the MI of MIMO channels versus the Rice factor K when $\eta = 5$ dB, and $\|\overline{\boldsymbol{H}}\|_F^2 = M$.

When fading processes are non-ergodic, the CDF of MI is equal to the outage probability. In Fig. 2.9, we investigate the effects of the Rice factor on the CDF of the MI. We fix the SNR η and the rate threshold R, and plot the CDF of the MI versus the Rice factor. It can be seen from the figure that our bounds, $(1 - P_{\text{tr\_U}})$ and $(1 - P_{\text{tr\_L}})$, have the same trend as the CDF of the MI. Fig. 2.9 indicates that when the Rice factor is relatively small, the CDF of the MI increases with an increase in the Rice factor, and when the Rice factor is relatively large, the CDF of the MI decreases with an increase in the Rice factor. This can be explained as follows. When the Rice factor is relatively small, the diversity benefits dominate the line-of-sight benefits and hence the outage probability increases with an increase in Rice factor. When the Rice factor is relatively large, the line-of-sight benefits dominate the diversity benefits and hence the outage

probability decreases with an increase in Rice factor.

2.9 Conclusion

In this chapter, we developed a novel approach to bounding the CCDF of the MI of MIMO systems in Rician fading environments by exploiting the properties and statistical distributions of the determinant and trace of a noncentral complex Wishart matrix. This method avoids using the statistical distribution of the eigenvalues of a noncentral Wishart matrix which is quite complex in general. Based on the proposed approach, we derived some tight lower and upper bounds on the CCDF of the MI. The results are also readily reduced to the case of Rayleigh fading. Compared with existing results, our bounds are not only given in closed form, but also readily applicable to the evaluation of the outage probability with sufficiently high accuracy. Furthermore, the statistical distribution of the determinant of a noncentral complex Wishart matrix derived in this chapter is an important result, and may have applications beyond the MI distribution.

Chapter 3

Optimal Transmission Strategies For Rayleigh Fading Relay Channels

In this chapter, we consider a DF single relay model working in full duplex mode under a Rayleigh fading environment. The outage probability and ergodic rate for this model are derived. With the objective of either minimizing the outage probability or maximizing the ergodic rate, optimal and approximately optimal transmission strategies are developed through selecting appropriate transmit signaling and/or spatial power allocation. The derived transmission strategies only require the knowledge of the second-order statistics of the channels at the transmitters, which can be readily acquired in practice. Simulation results demonstrate that the optimal transmit signaling and/or spatial power allocation can offer considerable performance improvements over the equal power allocation strategy.

3.1 Introduction

In wireless networks, the cooperative relaying transmission is an effective technique to improve power efficiency, enhance network coverage, and mitigate detrimental

fading effects of wireless channels. The study of such a transmission technique under various settings, especially in relay channels, has received a considerable amount of research attention [30–32, 44, 49, 90]. Until now, the capacity of the general relay channels remains unknown except for some special cases [30]. Alternatively, the highest information rates achieved by various relaying protocols such as DF, AF, and CF have been extensively investigated [30,44,49]. In fading relay channels, the typical information theoretic performance measures are the *outage probability* and the *ergodic rate*, respectively, for non-ergodic and ergodic fading channels. These performance measures generally depend on the statistical correlation of the signals transmitted from source and relay, as well as the spatial power allocation between the two nodes. It was shown in [49] that for Rayleigh fading DF relay channels, the transmit signals from source and relay, which maximize the ergodic rate, are statistically independent. In a low SNR regime, a spatial power allocation strategy that maximizes the upper bound of the ergodic capacity was derived in [63,64]. In [52], several power allocation strategies that maximize lower or upper bounds of ergodic capacity were developed for various settings. Different from the previous work [52, 63, 64] which emphasizes primarily on optimizing ergodic performance measures, a recent paper [65] proposed a spatial power allocation strategy that optimizes the outage probability of the information rate for AF relay channels. In this chapter, we consider Rayleigh fading DF single relay channels with block Markov transmissions. Our focus is on the design of transmission strategies that optimize the outage probability or ergodic rate. In [52], a similar ergodic rate optimization problem was investigated. The key difference is that, at the transmitters of the source and relay, the strategies in [52] assume the availability of the complete knowledge of instantaneous CSI, whereas the strategies in this chapter only need the knowledge of variances of the channels (*statistical* CSI). Relative to the instantaneous CSI, the statistical CSI can be readily acquired at the transmitters and

does not require frequent updates.

We first derive the outage probability and erogenic rate by using a Hermitian quadratic form in complex Gaussian RVs. Based on the derived outage probability and ergodic rate, we next obtain the transmission strategies that either minimize the outage probability or maximize the ergodic rate, which we term as *outage-optimal* and *ergodic-optimal* strategies respectively. In deriving these transmission strategies, we restrict attention to two scenarios classified by the feasibility of spatial power allocation. In the scenario where no spatial power allocation is available, we show that the outage-optimal transmit signals, in contrast with the ergodic-optimal ones, are not necessarily independent, and their correlation generally depends on the SNRs of the links and the target transmission rate. Moreover, we further show that the ergodic rate is a monotonically decreasing function of the correlation coefficient between the transmit signals from the source and relay. In a small outage scenario, which is generally of practical interest, we derive an outage-optimal spatial power allocation strategy between source and relay. Our findings suggest that it is not always beneficial to use a relay, and the optimal power allocation generally depends on several parameters including the total transmit power, the target transmission rate, and the variances of the channels. Additionally, we present ergodic-optimal spatial power allocation strategies, which are obtained by numerically solving transcendental equations.

The remainder of the chapter is organized as follows. The system model of a single relay channel is introduced in Section 3.2. The outage probability and the ergodic rate for Rayleigh fading DF relay channels are obtained in Section 3.3. The outage-optimal and ergodic-optimal transmission strategies are derived in Sections 3.4 and 3.5, respectively. Finally, Section 3.6 concludes the chapter.

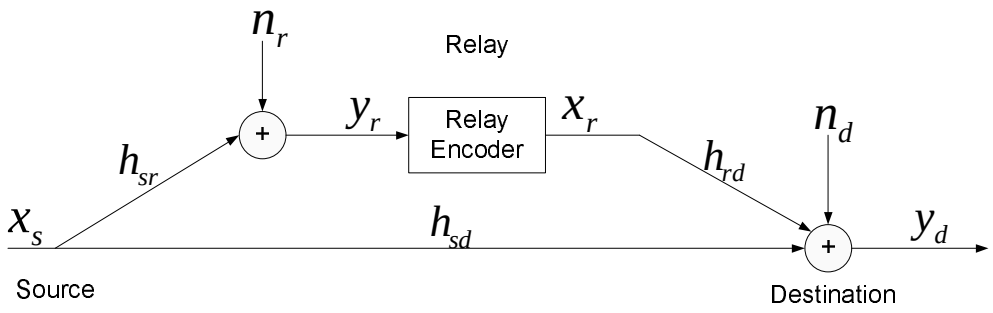

Figure 3.1: The single-relay network channel.

3.2 System Model

Consider a single relay network model as illustrated in Fig. 3.1, which consists of a source node s, a relay node r, and a destination node d. In this model, cooperating with the source, the relay facilitates the ultimate transmission from the source to destination. Specifically, the source sends a signal x_s with power E_s to both the relay and destination. The relay receives a signal y_r and performs the DF protocol where the relay fully decodes x_s, re-encodes it into a signal x_r based on the prior received signals, and sends x_r to the destination with power E_r. Mathematically, the respective received signals at the relay and destination can be expressed as

$$\begin{aligned} y_r &= \sqrt{E_s} h_{sr} x_s + n_r, \\ y_d &= \sqrt{E_s} h_{sd} x_s + \sqrt{E_r} h_{rd} x_r + n_d, \end{aligned} \tag{3.1}$$

where the AWGN terms n_r and n_d have mean zero and variance N_0, and h_{ij} $(ij \in \{sr, sd, rd\})$ denotes the channel gain between nodes i and j. By taking account of path loss, shadowing, and frequency nonselective fading [44], the channel gains h_{ij} are modeled as independent complex Gaussian RVs with mean zero and variances σ_{ij}^2, i.e., $h_{ij} \sim \mathcal{CN}(0, \sigma_{ij}^2)$. We further assume that the relay node works in the *full-duplex* mode which usually offers higher spectral efficiency than the half-duplex counterpart; perfect CSI at the corresponding receivers only; the variances of h_{sr}, h_{sd} and h_{rd} are

known at the source and relay; the transmit signals are subject to the power constraints $\mathbb{E}(|x_s|^2) = 1$ and $\mathbb{E}(|x_r|^2) = 1$; E_s and E_r satisfy a total power constraint given as $E_s + E_r \leq E$.

3.3 Outage Probability and Ergodic Rate in Rayleigh Fading Relay Channels

In this section, we begin with introducing the achievable rate for relay channels with fixed channel gains. We next derive the outage probability, and the ergodic rate for Rayleigh fading relay channels. These results will be used to derive optimal transmission strategies in later sections.

3.3.1 Achievable Rate With Fixed Channel Gains

It follows directly from [30, Theorem 1] that if the channel gains h_{sr}, h_{sd}, and h_{rd} are fixed, the following information rate is achievable:

$$I_{\text{df}} = \min\{I(x_s; y_r \mid x_r, h_{sr}), I(x_s, x_r; y_d \mid h_{sd}, h_{rd})\}.$$

Clearly, I_{df} depends on the statistical distribution of transmit signals. It was shown in [49, Proposition 2] that for any fixed h_{ij}, zero-mean jointly Gaussian distributed transmit signals, x_s and x_r, maximize I_{df}. By letting ρ be the correlation coefficient between x_s and x_r, i.e., $\rho = \mathbb{E}[x_s x_r^*]$, the corresponding information rate I_{df} can be written as [49, Eq. (106)], [52, Eq. (7)]

$$I_{\text{df}} = \min\left\{I_{\text{sir}},\ I_{\text{mac}}\right\}, \tag{3.2}$$

where

$$I_{\text{sir}} = I(x_s; y_r \mid x_r, h_{sr}) = \log_2\left(1 + |h_{sr}|^2(1-|\rho|^2)E_s/N_0\right),$$
$$I_{\text{mac}} = I(x_s, x_r; y_d \mid h_{sd}, h_{rd}) = \log_2(1+Q),$$

with

$$Q = |h_{sd}|^2 E_s/N_0 + |h_{rd}|^2 E_r/N_0 + 2\Re(\rho h_{sd} h_{rd}^*)\sqrt{E_s E_r}/N_0. \tag{3.3}$$

Obviously, I_{df} in (3.2) depends on the parameter ρ that captures the 'cooperation' between the source and relay, as well as the parameter $\nu := E_s/E$ that quantifies the power allocation between these two nodes. It should be noted that to make (3.2) non-trivial and meaningful, the parameter ν should be in the interval $(0,1)$. In the extreme case of $\nu = 1$, i.e., only the direct transmission takes place, the maximum mutual information conditional on a channel realization is $I_{\text{sd}} = \log_2(1+|h_{sd}|^2 E/N_0)$ instead of I_{df} in (3.2).

3.3.2 Outage Probability

For non-ergodic fading channels, e.g., quasi-static fading channels [91, p. 187], the outage probability is regarded as an important performance measure. In such a scenario, I_{df} is a RV depending on h_{ij}, and the outage probability of I_{df} for a target transmission rate R, is defined as

$$P_{\text{out}}(R) := P(I_{\text{df}} \leq R).$$

Note that $P_{\text{out}}(R)$ is actually the CDF of I_{df}.

Due to the mutual independence of h_{ij}, the CCDF of I_{df} with a target transmission rate R can be written as $P(I_{\text{df}} > R) = P(I_{\text{sir}} > R) \cdot P(I_{\text{mac}} > R)$. Since $|h_{sr}|^2$ is exponentially distributed and $\log(\cdot)$ is a monotonically increasing function, $P(I_{\text{sir}} >$

R) can be easily computed as follows

$$F_H(\mu,\nu,\tilde{R}) := P(I_{\text{sir}} > R) = P\big(|h_{sr}|^2 > \tilde{R}N_0/(E_s\mu)\big) = \exp\big(-\tilde{R}/(\eta_{sr}\mu)\big), \quad (3.4)$$

where $\tilde{R} := 2^R - 1$, $\mu := 1 - |\rho|^2$, and $\eta_{sr} := E_s\sigma_{sr}^2/N_0$. Note that η_{sr} represents the received SNR at the relay. In order to derive $P(I_{\text{mac}} > R)$, we need evaluate $P(Q > r)$. To do so, we express Q as a Hermitian quadratic form

$$Q = \begin{bmatrix} \sqrt{E_s/N_0}\, h_{sd}^* & \sqrt{E_r/N_0}\, h_{rd}^* \end{bmatrix} \begin{bmatrix} 1 & \rho^* \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sqrt{E_s/N_0}\, h_{sd} \\ \sqrt{E_r/N_0}\, h_{rd} \end{bmatrix}. \quad (3.5)$$

Relying on (3.5), we derive the PDF and CCDF of Q, which are collectively presented in the following lemma.

Lemma 3.1. The Hermitian quadratic form Q can be rewritten as

$$Q = \alpha|h_1|^2 + \beta|h_2|^2 = \alpha w_1 + \beta w_2, \quad (3.6)$$

where h_1 and h_2 are i.i.d. complex Gaussian RVs with zero mean and unit variance, w_1 and w_2 denote the magnitude squares of h_1 and h_2, respectively, i.e., $w_1 = |h_1|^2$, $w_2 = |h_2|^2$, α and β are given as

$$\begin{aligned} \alpha &= \frac{\eta_{sd}+\eta_{rd}}{2} + \sqrt{\Big(\frac{\eta_{sd}+\eta_{rd}}{2}\Big)^2 - \eta_{sd}\eta_{rd}\mu}, \\ \beta &= \frac{\eta_{sd}+\eta_{rd}}{2} - \sqrt{\Big(\frac{\eta_{sd}+\eta_{rd}}{2}\Big)^2 - \eta_{sd}\eta_{rd}\mu}, \end{aligned} \quad (3.7)$$

with $\eta_{sd} := E_s\sigma_{sd}^2/N_0$ and $\eta_{rd} := E_r\sigma_{rd}^2/N_0$. Moreover, the PDF $f_Q(q)$ and the CCDF $F_Q(\alpha,\beta,\tilde{R})$ of Q are, respectively, given by

$$f_Q(q) := (e^{-q/\alpha} - e^{-q/\beta})/(\alpha-\beta), \quad q \geq 0, \quad (3.8)$$

$$F_Q(\alpha,\beta,\tilde{R}) := P(Q > \tilde{R}) = \frac{\alpha e^{-\tilde{R}/\alpha} - \beta e^{-\tilde{R}/\beta}}{\alpha-\beta}. \quad (3.9)$$

Proof. Define

$$\boldsymbol{h} := [\sqrt{E_s/N_0}\, h_{sd} \;\; \sqrt{E_r/N_0}\, h_{rd}]^T \qquad \text{and} \qquad \boldsymbol{A} := \begin{bmatrix} 1 & \rho^* \\ \rho & 1 \end{bmatrix}.$$

Since $0 \leq |\rho| \leq 1$, the matrix $\boldsymbol{A}$ is Hermitian and positive semi-definite. Let $\boldsymbol{L}$ be the covariance matrix of $\boldsymbol{h}$. Since the entries of $\boldsymbol{h}$ are independently Gaussian distributed, $\boldsymbol{L}$ can be readily calculated as

$$\boldsymbol{L} := \mathbb{E}[\boldsymbol{h}\boldsymbol{h}^H] = \mathrm{diag}\{\eta_{sd}, \eta_{rd}\} = \begin{bmatrix} \eta_{sd} & 0 \\ 0 & \eta_{rd} \end{bmatrix}.$$

Applying the noise whitening technique in [92], [93, pp. 28–29] and the eigne-decomposition of $\boldsymbol{L}^{1/2}\boldsymbol{A}\boldsymbol{L}^{1/2}$, we rewrite Q as

$$Q = (\boldsymbol{h}^H\boldsymbol{L}^{-1/2})\boldsymbol{L}^{1/2}\boldsymbol{A}\boldsymbol{L}^{1/2}(\boldsymbol{L}^{-1/2}\boldsymbol{h}) = (\boldsymbol{h}^H\boldsymbol{L}^{-1/2}\boldsymbol{V})\boldsymbol{D}(\boldsymbol{V}^H\boldsymbol{L}^{-1/2}\boldsymbol{h}),$$

where $\boldsymbol{D}$ is a real diagonal matrix with diagonal elements being the eigenvalues of $\boldsymbol{L}^{1/2}\boldsymbol{A}\boldsymbol{L}^{1/2}$, α and β, and $\boldsymbol{V}$ is a unitary matrix with columns being the corresponding eigenvectors, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Defining $h_1 = \boldsymbol{v}_1^H\boldsymbol{L}^{-1/2}\boldsymbol{h}$ and $h_2 = \boldsymbol{v}_2^H\boldsymbol{L}^{-1/2}\boldsymbol{h}$, we can further simplify Q to (3.6). It is straightforward to check that h_1 and h_2 are i.i.d. complex Gaussian RVs with zero mean and unit variance. Additionally, the eigenvalues α and β can be readily calculated as given in (3.7). Following [94, Eq. (14)], the characteristic function of Q is given by

$$\Phi_Q(t) = (1 - jt\alpha)^{-1}(1 - jt\beta)^{-1}.$$

By applying the Fourier transform to $\Phi_Q(t)$, the PDF of Q can be written as

$$f_Q(q) = \begin{cases} \dfrac{e^{-q/\alpha} - e^{-q/\beta}}{\alpha - \beta}, & \alpha > \beta \geq 0, q \geq 0, \qquad (3.10a) \\ \alpha^{-2} q e^{-q/\alpha}, & \alpha = \beta > 0, q \geq 0. \qquad (3.10b) \end{cases}$$

When α equals β, (3.10a) reduces to (3.10b) by L'Hospital's rule. Thus, the PDF of Q can be simply expressed as in (3.8), and accordingly the CCDF of Q is obtained as in (3.9). □

Remark 3.1. It should be noted that Q in (3.6) is a weighted sum of i.i.d. exponential RVs. In light of this observation, we can interpret I_{mac} in the following way. The rate I_{mac}, which is achieved by sending correlated signals (x_1 and x_2) with power E_s and E_r over independent but not necessarily identically distributed fading channels (h_{sd} and h_{rd}), is the same as the one achieved by sending independent signals with respective power α and β over i.i.d. fading channels (h_1 and h_2).

Remark 3.2. Even though the quadratic form Q is a function of the complex quantity ρ, the distributions of Q including PDF (3.8) and CCDF (3.9) depend only on $|\rho|^2$ (or μ), but not the phase of ρ. This is due to the fact that in the Rayleigh fading case, the phase of ρ does not alter the distribution of the phase of $\rho h_{sd} h_{rd}^*$, since the phase of $h_{sd} h_{rd}^*$ is uniformly distributed. As $P(I_{\text{sir}} > R)$ does not depend on the phase of ρ either, so does not $P(I_{\text{df}} > R)$. Besides the parameter μ, $P(I_{\text{df}} > R)$ also depends on ν and r, and thus we denote $P(I_{\text{df}} > R)$ by $F_R(\mu, \nu, \tilde{R})$.

Following directly from Lemma 3.1, we obtain a closed-form expression for the CCDF of the rate I_{df}, which is given in the following Theorem.

Theorem 3.1. The CCDF $F_R(\mu, \nu, \tilde{R})$ of the rate I_{df} is

$$F_R(\mu, \nu, \tilde{R}) := P(I_{\text{df}} > R) = \frac{\alpha e^{-\tilde{R}/\alpha} - \beta e^{-\tilde{R}/\beta}}{\alpha - \beta} e^{-\tilde{R}/(\eta_{sr}\mu)}. \tag{3.11}$$

Furthermore, note that the outage probability of the rate I_{df}, $P_{\text{out}}(R)$, is simply $1 - F_R(\mu, \nu, \tilde{R})$, which depends on the choices of μ, ν and $\tilde{R}$.

3.3.3 Ergodic Rate

For ergodic fading channels, e.g., block fading channels [91, p. 199], the ergodic rate of the full-duplex DF relay system is given by [30, Theorem 1], [52, 62]

$$R_{\text{df}} = \min\left\{\mathbb{E}[I_{\text{sir}}],\, \mathbb{E}[I_{\text{mac}}]\right\}, \tag{3.12}$$

where $\mathbb{E}[I_{\text{sir}}]$ and $\mathbb{E}[I_{\text{mac}}]$ are the ergodic rates of the source-relay link and the multiple access links, respectively. It should be noted that the ergodic rate of the relay channel, R_{df}, is not equal to $\mathbb{E}[I_{\text{df}}]$ in general. To compute the ergodic rate R_{df} in ergodic fading channels, we need to compute the mean values of I_{sir} and I_{mac}. For $\rho = 0$, the mean value of I_{mac}, i.e., $\mathbb{E}[I_{\text{mac}}]$, was obtained in [62]. However, to the best of our knowledge, the result for ρ other than zero is not available in the literature. Based on Lemma 3.1, we derive $\mathbb{E}[I_{\text{mac}}]$ for an arbitrary ρ. This enables us to compute the ergodic rate for Gaussian input signals with an arbitrary correlation.

Theorem 3.2. For ergodic Rayleigh fading channels, the ergodic rate R_{df} of the full-duplex DF relay system is a function of μ and ν, denoted by $R_{\text{df}}(\mu, \nu)$, and is given by

$$R_{\text{df}}(\mu, \nu) = \min\left\{R_{\text{sir}}(\mu, \nu),\ R_{\text{mac}}(\mu, \nu)\right\}, \tag{3.13}$$

where $R_{\text{sir}}(\mu, \nu)$ and $R_{\text{mac}}(\mu, \nu)$ are the means of I_{sir} and I_{mac}, respectively, and are given by

$$R_{\text{sir}}(\mu, \nu) = \mathbb{E}[I_{\text{sir}}] = \exp(\mu^{-1}\eta_{sr}^{-1})\mathrm{E}_1(\mu^{-1}\eta_{sr}^{-1}) \log_2 e, \tag{3.14}$$

$$R_{\text{mac}}(\mu, \nu) = \mathbb{E}[I_{\text{mac}}] = \frac{\log_2 e}{\alpha - \beta}\left[\alpha \exp(\alpha^{-1})\mathrm{E}_1(\alpha^{-1}) - \beta \exp(\beta^{-1})\mathrm{E}_1(\beta^{-1})\right], \tag{3.15}$$

with $\mathrm{E}_1(x)$ denoting the exponential integral function, i.e., $\mathrm{E}_1(x) = \int_x^\infty e^{-t}/t\, dt$ $(x > 0)$.

Proof. The proof follows immediately from Lemma 3.1 and [62, Eq. (49)]. □

Similar to the fixed channel gain case, ν should be in the interval $(0, 1)$ to make (3.12) meaningful. In the extreme case of $\nu = 1$, the ergodic rate is $R_{\text{sd}} = \mathbb{E}[I_{\text{sd}}]$ instead of R_{df} in (3.12).

3.4 Outage-Optimal Transmission Strategies

In this section, we investigate the outage-optimal transmission strategies that minimize the outage probability of the rate I_{df}. In particular, we study the following two cases classified by whether spatial power allocation is available.

3.4.1 Outage-Optimal Transmit Signals Without Spatial Power Allocation

We first consider a scenario where the source and relay can not share their power resources and each of them has its own average transmit power constraint. In other words, ν is fixed and does not need to be optimized. Designing an optimal transmission strategy amounts to finding an optimal ρ that minimizes the outage probability of I_{df} (or equivalently, finding an optimal μ that maximizes the CCDF of I_{df}). Mathematically, the optimal μ, denoted by μ_o, is selected as

$$\mu_o = \arg \min_{0 \le \mu \le 1} P_{\text{out}}(R) = \arg \max_{0 \le \mu \le 1} F_R(\mu, \nu, \tilde{R}), \tag{3.16}$$

and the optimal ρ, denoted by ρ_o, is $\sqrt{1-\mu_o}$. Correspondingly, the resulting outage probability is denoted as

$$P^I_{\text{out}}(R) := \min_{0 \le \mu \le 1} P_{\text{out}}(R) = 1 - F_R(\mu_o, \nu, \tilde{R}).$$

To tackle this problem, we need examine the monotonicity of $F_R(\mu, \nu, \tilde{R})$ with respect to μ, which is reflected by the sign of the quantity $\partial F_R(\mu, \nu, \tilde{R})/\partial \mu$ given by

$$\frac{\partial}{\partial \mu} F_R(\mu, \nu, \tilde{R}) = \frac{\alpha r + \beta \kappa \tilde{R} + \xi}{\mu(\alpha - \beta)^2} e^{-\tilde{R}/\alpha} F_H(\mu, \nu, \tilde{R}) G(\mu, \nu, \tilde{R}), \tag{3.17}$$

where $\kappa := (\alpha - \beta)/(\eta_{sr}\mu)$, $\xi := \alpha\beta(\alpha + \beta)/(\alpha - \beta)$, and

$$G(\mu, \nu, \tilde{R}) := \phi - e^{-\tilde{R}(1/\beta - 1/\alpha)} \tag{3.18}$$

with

$$\phi := \frac{(\alpha\kappa - \beta)\tilde{R} + \xi}{(\beta\kappa + \alpha)\tilde{R} + \xi}. \tag{3.19}$$

Since $\alpha \geq \beta \geq 0$ and $\mu \in [0, 1]$, the sign of $\partial F_R(\mu, \nu, \tilde{R})/\partial\mu$ in (3.17) is the same as that of $G(\mu, \nu, \tilde{R})$. Thus, the monotonicity of $F_R(\mu, \nu, \tilde{R})$ simply depends on the sign of $G(\mu, \nu, \tilde{R})$. Determining the sign of $G(\mu, \nu, \tilde{R})$ amounts to checking whether ϕ is greater than or less than $e^{-\tilde{R}(1/\beta - 1/\alpha)}$. It is obvious that if ϕ is negative or equivalently $(\alpha\kappa - \beta)\tilde{R} + \xi < 0$, then $G(\mu, \nu, \tilde{R})$ is negative. We thus only focus on the case where $(\alpha\kappa - \beta)\tilde{R} + \xi > 0$ (or $\phi > 0$). Taking logarithm of ϕ and $e^{-\tilde{R}(1/\beta - 1/\alpha)}$, we obtain an equivalent expression for $G(\mu, \nu, \tilde{R}) \gtrless 0$ as

$$g(\tilde{R}) := \ln\left[\frac{(\alpha\kappa - \beta)\tilde{R} + \xi}{(\beta\kappa + \alpha)\tilde{R} + \xi}\right] + \frac{\alpha - \beta}{\alpha\beta}\tilde{R} \gtrless 0. \tag{3.20}$$

Due to the fact that $g(0) = 0$, the sign of the derivative of $g(\tilde{R})$ may be used to indicate the sign of $g(\tilde{R})$. Hence, we calculate the derivative of $g(\tilde{R})$ as

$$g'(\tilde{R}) = \frac{dg(\tilde{R})}{d\tilde{R}} = \frac{\alpha - \beta}{\alpha\beta} \cdot \frac{(\alpha\kappa - \beta)(\beta\kappa + \alpha)\tilde{R}^2 + (\alpha + \beta)(\kappa\xi + \alpha\beta)\tilde{R} + \alpha\beta\kappa\xi}{\big[(\alpha\kappa - \beta)\tilde{R} + \xi\big]\big[(\beta\kappa + \alpha)\tilde{R} + \xi\big]}. \tag{3.21}$$

To determine the sign of $g(\tilde{R})$, we introduce the following lemmas, which will be used in our derivations of the outage-optimal transmission strategies.

Lemma 3.2. For any μ satisfying $\alpha\kappa - \beta \geq 0$, we have $g(\tilde{R}) > 0$ for any $\tilde{R} \in [0, \infty)$.

Proof. It can be easily seen from (3.21) that if μ satisfies $\alpha\kappa - \beta \geq 0$, we have $g'(\tilde{R}) > 0$ for all $\tilde{R} \geq 0$. It implies that $g(\tilde{R})$ is strictly increasing in $\tilde{R}$ for all $\tilde{R} \geq 0$. Notice that $g(0) = 0$ for all $\mu \in [0, 1]$. Hence, we can conclude that $g(\tilde{R}) > g(0) = 0$ for all $\mu \in [0, 1]$. □

Lemma 3.3. For any μ satisfying $\alpha\kappa - \beta < 0$, $g(\tilde{R}) = 0$ has a unique root $\tilde{R}_g \in \big[0, \tilde{R}_\xi\big)$, where $\tilde{R}_\xi := \xi/(\beta - \alpha\kappa)$. Furthermore, $g(\tilde{R}) \geq 0$ for $r \in [0, \tilde{R}_g]$, and $g(\tilde{R}) < 0$ for $r \in (\tilde{R}_g, \tilde{R}_\xi)$.

Proof. Let us first prove the existence of a positive root of $g(\tilde{R})$ in its domain for any μ satisfying $\alpha\kappa - \beta < 0$. Notice that $g'(0) > 0$. Due to the continuity of the right derivative of $g(\tilde{R})$, there exists a small interval $(0, \epsilon)$ such that $g(\tilde{R}) > g(0) = 0$ for any $\tilde{R} \in (0, \epsilon)$. On the other hand, we have $\lim_{\tilde{R}\to\tilde{R}_\xi^-} g(\tilde{R}) < 0$. Since $g(\tilde{R})$ is continuous and well defined for $\tilde{R} \in [0, \tilde{R}_\xi)$, $g(\tilde{R})$ has at least one root denoted by $\tilde{R}_g \in [0, \tilde{R}_\xi)$. We next prove the uniqueness of the root by contradiction. Suppose that $g(\tilde{R})$ has at least two positive roots $\tilde{R}_{g1}$ and $\tilde{R}_{g2}$, and assume $\tilde{R}_{g2} > \tilde{R}_{g1} > 0$. Thus we have $g(0) = g(\tilde{R}_{g1}) = g(\tilde{R}_{g2}) = 0$. According to Rolle's theorem, there must be at least two points $\tilde{R}_{s1} \in (0, \tilde{R}_{g1})$ and $\tilde{R}_{s2} \in (\tilde{R}_{g1}, \tilde{R}_{g2})$ at which $g'(\tilde{R}_{s1}) = g'(\tilde{R}_{s2}) = 0$. However, from (3.21), we know that $g'(\tilde{R})$ has no positive root when $\alpha\kappa - \beta > 0$; $g'(\tilde{R})$ has one negative root and one positive root when $\alpha\kappa - \beta < 0$. Thus, it is impossible that $g'(\tilde{R})$ has two positive roots $\tilde{R}_{s1}$ and $\tilde{R}_{s2}$. Hence, it is a contradiction. Together with the existence, this shows that $g(\tilde{R})$ has a unique positive root. Again, from the proof of the existence, we can readily conclude that $g(\tilde{R}) > 0$ for $\tilde{R} \in (0, \tilde{R}_g)$, and $g(\tilde{R}) < 0$ for $\tilde{R} \in (\tilde{R}_g, \tilde{R}_\xi)$. □

Applying above two lemmas (Lemma 3.2 and Lemma 3.3), we obtain the following outage-optimal transmission strategies.

Theorem 3.3. If η_{sr}, η_{sd} and η_{rd} satisfy

$$\eta_{sd}\eta_{rd}|\eta_{sd} - \eta_{rd}| \geq \min(\eta_{sd}^2, \eta_{rd}^2)\eta_{sr}, \tag{3.22}$$

then μ_o is one (or ρ_o is zero) for any target rate $R \in [0, \infty)$ (the outage-optimal signals transmitted from the source and relay are independent). Moreover, the outage probability reduces to

$$P_{\text{out}}^I(R) = 1 - \frac{1}{\eta_{sd} - \eta_{rd}} \left\{ \eta_{sd} \exp\left[-\tilde{R}(\eta_{sr}^{-1} + \eta_{sd}^{-1}) \right] - \eta_{rd} \exp\left[-\tilde{R}(\eta_{sr}^{-1} + \eta_{rd}^{-1}) \right] \right\}. \tag{3.23}$$

Proof. By the definitions of α and β in (3.7), it is easy to conclude that $\beta^2/(\alpha-\beta)$ is monotonically increasing in μ. Noting that for any $\mu \in [0,1]$, we have

$$\beta^2/(\alpha-\beta) \leq \min(\eta_{sd}^2, \eta_{rd}^2)/|\eta_{sd}-\eta_{rd}|.$$

Condition (3.22) implies $\eta_{sd}\eta_{rd}/\eta_{sr} \geq \beta^2/(\alpha-\beta)$ for $\mu \in [0,1]$, or equivalently $\alpha\kappa - \beta \geq 0$ for $\mu \in [0,1]$. By Lemma 3.2, we have $g(\tilde{R}) \geq 0$ for any $\tilde{R} \geq 0$. It implies that $G(\mu,\nu,\tilde{R}) \geq 0$ for any $\tilde{R} \geq 0$ and $\mu \in [0,1]$, or equivalently, $F_R(\mu,\nu,\tilde{R})$ is an increasing function of $\mu \in [0,1]$ for any $\tilde{R} \geq 0$. Therefore, when condition (3.22) is satisfied, μ_o is one for any target rate $R \geq 0$. □

We next define several parameters which will be used in the following Theorem.

$$\tilde{R}_0 := (\eta_{sd}+\eta_{rd})\big((\eta_{sd}+\eta_{rd})\eta_{sr}^{-1}+1\big), \tag{3.24}$$

$$\tilde{R}_1 := \frac{(\eta_{sd}+\eta_{rd})(\eta_{sd}^2+\eta_{rd}^2)(\eta_{sd}-\eta_{rd})^{-2} - 2(\eta_{sd}+\eta_{rd}+\eta_{sd}\eta_{rd})\eta_{sr}^{-1}}{1+(\eta_{sd}-\eta_{rd})^2\eta_{sr}^{-2}+(\eta_{sd}^2+\eta_{rd}^2)^2(\eta_{sd}\eta_{rd}\eta_{rd})^{-1}(\eta_{sd}+\eta_{rd})^{-1}}, \tag{3.25}$$

$$\tilde{R}_2 := \frac{\eta_{sd}+\eta_{rd}}{|\eta_{sd}-\eta_{rd}|} \cdot \frac{\eta_{sd}\eta_{rd}\min(\eta_{sd},\eta_{rd})}{\min(\eta_{sd}^2,\eta_{rd}^2)-\eta_{sd}\eta_{rd}\eta_{sr}^{-1}|\eta_{sd}-\eta_{rd}|}, \tag{3.26}$$

$$\tilde{R}_c := \max(\tilde{R}_1, \tilde{R}_2). \tag{3.27}$$

Theorem 3.4. If η_{sr}, η_{sd} and η_{rd} satisfy

$$\eta_{sd}\eta_{rd}|\eta_{sd}-\eta_{rd}| < \min(\eta_{sd}^2,\eta_{rd}^2)\eta_{sr}, \tag{3.28}$$

a) then for any rate $R \in [0, R_0]$ where $R_0 := \log_2(1+\tilde{R}_0)$, the outage-optimum μ_o is one and hence the corresponding outage probability is the same as the one given in (3.23);

b) and for any rate $R \in [R_c, \infty)$ where $R_c := \log_2(1+\tilde{R}_c)$, the outage-optimum μ_o is the unique root of $G(\mu,\nu,\tilde{R})$ in (3.18) and the outage probability is

$$P_{\text{out}}^{I}(R) = 1 - \frac{1}{\alpha_o-\beta_o}\Big\{\alpha_o \exp\big[-\tilde{R}(\eta_{sr}^{-1}\mu_o^{-1}+\alpha_o^{-1})\big] - \beta_o \exp\big[-\tilde{R}(\eta_{sr}^{-1}\mu_o^{-1}+\beta_o^{-1})\big]\Big\}, \tag{3.29}$$

where α_o and β_o are obtained by evaluating α and β in (3.7) at $\mu = \mu_o$, respectively.

Proof. See Appendix B for the proof. □

Remark 3.3. It can be observed from Theorems 3.3 and 3.4 that when the target transmission rate R is less than R_0, the outage-optimal transmitted signals from the source and relay are independent (or $\mu_o = 1$), irrespective of the transmission power.

Remark 3.4. Note that when $\eta_{sd} = \eta_{rd}$, condition (3.28) is satisfied, but R_c is infinite. In this case, the value of μ_o can be obtained by solving $G(\mu, \nu, \tilde{R}) = 0$ numerically.

We next provide an illustrative example to validate our findings. In all the examples of this chapter, we adopt the well-known path-loss model in which the path-loss exponent is 3, i.e., $\sigma_{ij}^2 \propto d_{ij}^{-3}$, where d_{ij} denotes the distance between nodes i and j. Without loss of generality, we assume that σ_{sd}^2 is normalized to one throughout the chapter, i.e., $\sigma_{sd}^2 = 1$.

Example 3.1. In this example, we select $d_{sr} = d_{sd}/3, d_{rd} = 2d_{sd}/3$, and $E_s/N_0 = E_r/N_0 = 10$ dB such that the condition given in (3.28) is satisfied. These SNRs give $R_0 = 5.63$ bps/Hz and $R_c = 6.40$ bps/Hz. By solving $G(\mu, \nu, \tilde{R}) = 0$ numerically, we find that ρ_o is zero for $R \in [0, 6.25]$ bps/Hz, and ρ_o is non-zero for the value of R beyond this interval. Fig. 3.2 plots the optimum ρ_o versus the target rate R. As can be seen from Fig. 3.2, ρ_o depends on R for given SNRs. For a relatively small R, ρ_o is always zero, while for a relatively large R, ρ_o is nonzero and increases as R increases. Fig. 3.3 compares the CCDF of I_{df} for ρ_o and $\rho = 0$. It can be observed from Fig. 3.3 that the outage probabilities achieved by ρ_o is slightly smaller than that achieved by $\rho = 0$.

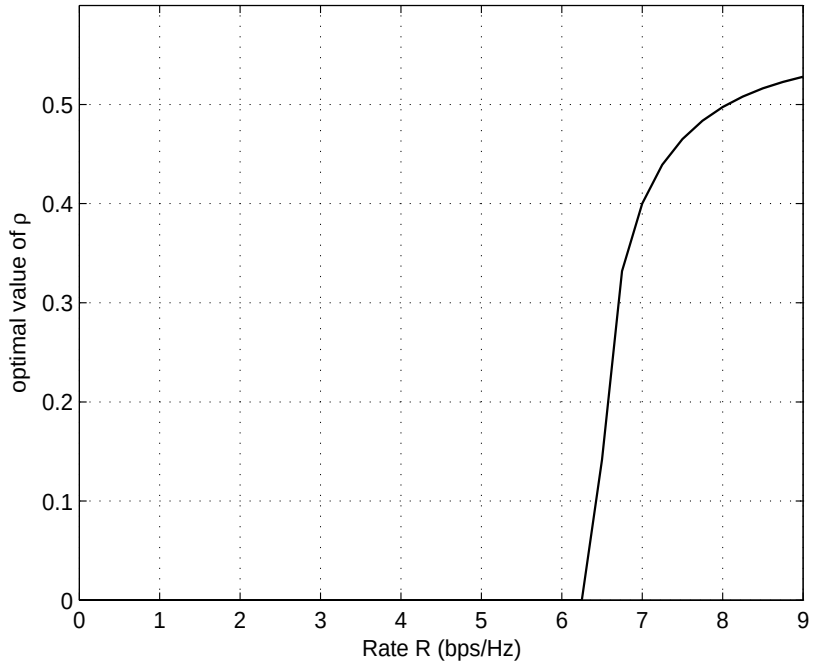

Figure 3.2: Optimal correlation coefficient ρ_o versus the target transmission rate R.

3.4.2 Outage-Optimal Spatial Power Allocation

We now consider the case where spatial power allocation between source and relay is feasible. In such a case, apart from μ, the parameter ν also needs to be optimized. For the DF relay channel, the outage-optimal transmission strategy is a pair $(\tilde{\mu}_o, \tilde{\nu}_o)$ such that

$$(\tilde{\mu}_o, \tilde{\nu}_o) := \arg \max_{\substack{0 \leq \mu \leq 1 \\ 0 < \nu < 1}} F_R(\mu, \nu, \tilde{R}). \tag{3.30}$$

It does not seem tractable to solve the problem analytically. However, we notice that, the outage probability of practical interest is relatively small (nominally $10^{-1} \sim 10^{-3}$). This motivates us to consider a *small outage* scenario, which not only describes a realistic situation but also simplifies the optimization problem (3.30).

Since $P_{\text{out}}(R) = 1 - P(I_{\text{sir}} > R) F_Q(\alpha, \beta, \tilde{R})$, we can lower bound $P_{\text{out}}(R)$ as

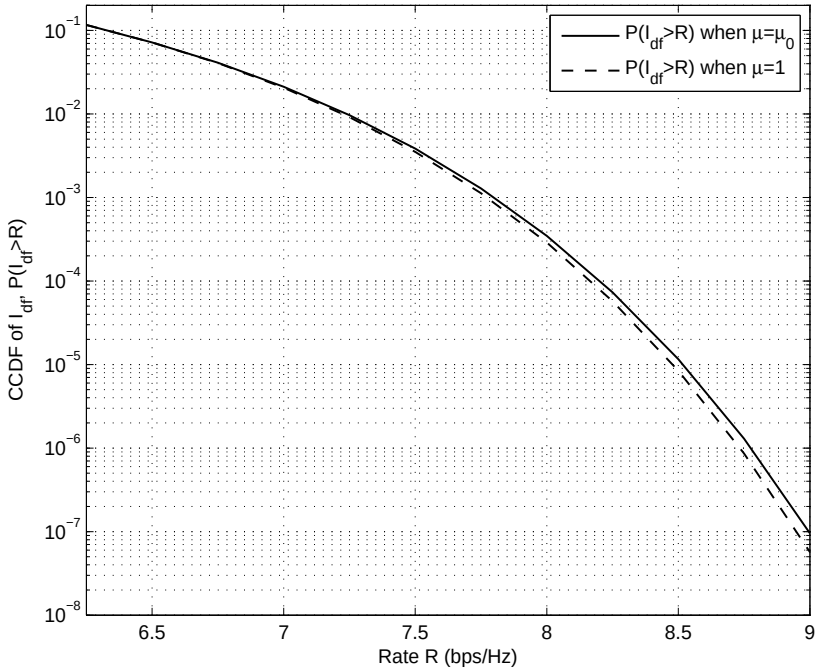

Figure 3.3: Comparisons of the CCDF of $I_{\mathbf{df}}$ for the optimal ρ_o and $\rho = 0$.

follows.

$$P_{\text{out}}(R) > 1 - F_Q(\alpha, \beta, \tilde{R}) = P(\alpha w_1 + \beta w_2 \leq \tilde{R})$$
$$\geq P(\alpha w_1 \leq \tilde{R}/2) P(\beta w_2 \leq \tilde{R}/2) > \Big[1 - \exp\Big(-\frac{\tilde{R}/2}{\alpha+\beta}\Big)\Big]^2, \quad (3.31)$$

where the last inequality follows from the fact $\alpha+\beta \geq \alpha \geq \beta$. It follows directly from (3.31) that the following inequality

$$\tilde{R}/(\alpha+\beta) \leq -2\ln\big[1 - \sqrt{P_{\text{out}}(R)}\,\big]$$

is always true. Thus, we always have $\tilde{R}/(\alpha+\beta) < 1$ if $P_{\text{out}}(R) < (1-e^{-1/2})^2 = 0.1548$. In addition, it can be easily seen from (3.7) and (3.24) that

$$\tilde{R}_0/(\alpha+\beta) = \tilde{R}_0/(\eta_{sd}+\eta_{rd}) \geq \sigma_{sd}^2/\sigma_{sr}^2 + 1$$

for all $\nu \in (0,1)$. Therefore, when outage probability $P_{\text{out}}(R)$ is less than 0.1548, the

inequality,

$$\tilde{R}/(\alpha+\beta) < 1 \leq 1 + \sigma_{sd}^2/\sigma_{rd}^2 \leq \tilde{R}_0/(\alpha+\beta),$$

holds, and hence $R < R_0$ is satisfied. In the following, we will consider the small outage scenario where $P_{\text{out}}(R) \leq 0.1548$. It is clear from Remark 3.3 that in such a scenario, $\tilde{\mu}_o = 1$. Thus the optimization problem (3.30) is simplified to

$$\tilde{\nu}_o = \arg \max_{0<\nu<1} F_R(1, \nu, \tilde{R}). \tag{3.32}$$

Moreover, in the case of $\nu = 1$, as mentioned in Section 3.3, the maximum conditional mutual information is I_{sd}, and its CCDF is given by

$$F_D(\tilde{R}) := P(I_{\text{sd}} > R) = \exp(-R_N/\sigma_{sd}^2), \tag{3.33}$$

where $R_N := \tilde{R}N_0/E = (2^R - 1)N_0/E$ and $1/R_N$ is termed as *rate normalized SNR* [44]. To determine whether the direct transmission or DF relay transmission with optimal power allocation $\tilde{\nu}_o$ should be used, we need compare $F_D(\tilde{R})$ with $F_R(1, \tilde{\nu}_o, \tilde{R})$. The outage-optimal power allocation ν_o is thus determined as

$$\nu_o = \begin{cases} \tilde{\nu}_o, & F_R(1, \tilde{\nu}_o, \tilde{R}) > F_D(\tilde{R}), \\ 1, & F_R(1, \tilde{\nu}_o, \tilde{R}) \leq F_D(\tilde{R}). \end{cases} \tag{3.34}$$

We now present optimal transmission strategies in the following two Theorems for the afore-described small outage scenario where the outage probability $P_{\text{out}}(R)$ is smaller than 0.1548.

Theorem 3.5. If $\sigma_{sr}^2 \leq \sigma_{sd}^2$, the outage-optimal transmission strategy is to use only direct transmission and allocate all the transmit power at the source, i.e., $\nu_o = 1$.

Proof. If $\sigma_{sr}^2 \leq \sigma_{sd}^2$, it is clear from (3.11) that

$$F_R(1, \nu, \tilde{R}) < \exp(-\tilde{R}\eta_{sr}^{-1}) < \exp(-R_N\sigma_{sr}^{-2}) \leq F_D(\tilde{R})$$

for all $\nu \in (0,1)$. The outage probability of using a relay is greater than that of only using the source-destination link. Thus, $\nu_o = 1$. □

This result agrees with our practical intuition: when the source-relay link is weaker than the source-destination link, the decoder at the relay has a large likelihood of making decoding errors. In such a case, it is quite likely that the signals transmitted from the relay become interferences at the destination, and hence the relay should not transmit to avoid generating any potential interferences that may reduce the reliability of the source-destination link. However, in practice, the source-relay link is generally stronger than the source-destination link ($\sigma_{sr}^2 > \sigma_{sd}^2$). In such a scenario, it does not appear mathematically tractable to obtain an exact solution to (3.34). We therefore derive the following *approximately* optimal transmission strategies.

Theorem 3.6. Consider the case where $\sigma_{sr}^2 > \sigma_{sd}^2$.

a) If $\sigma_{rd}^2/\sigma_{sd}^2 \to 0$, then $\nu_o = 1$;

b) If $\sigma_{sd}^2/\sigma_{rd}^2 \to 0$ and $\sigma_{sd}^2/\sigma_{sr}^2 \to 0$ (dual-hop), then $\nu_o = \sigma_{rd}(\sigma_{sr} + \sigma_{rd})^{-1}$;

c) If $R_N/\sigma_{ij}^2 \ll 1$, then

$$\tilde{\nu}_o = 1 - \left[(1 + 2\sigma_{sd}^2\sigma_{rd}^2\sigma_{sr}^{-2}R_N^{-1})^{1/2} + 1\right]^{-1}. \tag{3.35}$$

Proof. a) Notice that $\sigma_{rd}^2/\sigma_{sd}^2 \to 0$ implies that for any fixed $\nu \in (0,1)$, $\alpha \gg \beta$. From (3.11), we have

$$\begin{aligned} F_R(1,\nu,\tilde{R}) &\approx \exp\left[-R_N(\sigma_{sd}^2\nu)^{-1} - R_N(\sigma_{sr}^2\nu)^{-1}\right] \\ &< \exp\left[-R_N(\sigma_{sd}^2\nu)^{-1}\right] < F_D(\tilde{R}) \end{aligned}$$

for any $\nu \in (0,1)$. Thus, $\nu_o = 1$.

b) By the assumptions, the DF relay system is like a dual-hop system. We have

$$F_R(1,\nu,\tilde{R}) \approx \exp\left[-R_N\sigma_{rd}^{-2}(1-\nu)^{-1} - R_N(\sigma_{sr}^2\nu)^{-1}\right].$$

By setting $\partial F_R(1,\nu,\tilde{R})/\partial\nu$ to be zero, we obtain $\tilde{\nu}_o = \sigma_{rd}(\sigma_{sr}+\sigma_{rd})^{-1}$. Actually, $F_R(1,\nu,\tilde{R})$ is an increasing function of $\nu \in (0,\tilde{\nu}_o]$, and a decreasing function of $\nu \in [\tilde{\nu}_o, 1)$. Therefore, in this case (dual-hop), we have $\nu_o = \tilde{\nu}_o = \sigma_{rd}(\sigma_{sr}+\sigma_{rd})^{-1}$.

c) Due to the small outage probability assumption and the continuity of $1/\nu$ for $\nu > 0$, we have

$$R_N(\sigma_{sd}^2\nu)^{-1} \ll 1, \quad R_N(\sigma_{rd}^2(1-\nu))^{-1} \ll 1, \quad R_N(\sigma_{sr}^2\nu)^{-1} \ll 1$$

for any ν in a small neighborhood of $\tilde{\nu}_o$. Applying a Taylor's expansion of $F_R(1,\nu,\tilde{R})$ in (3.11) at $R_N = 0$, we obtain

$$F_R(1,\nu,\tilde{R}) \approx 1 - \frac{1}{\sigma_{sr}^2\nu}R_N + \Big[\frac{0.5}{(\sigma_{sr}^2\nu)^2} - \frac{0.5}{\sigma_{sd}^2\sigma_{rd}^2\nu(1-\nu)}\Big]R_N^2.$$

Since $R_N(\sigma_{sr}^2\nu)^{-1} \ll 1$, the higher order term $R_N^2(\sigma_{sr}^2\nu)^{-2}$ can be ignored. Thus, $F_R(1,\nu,\tilde{R})$ can be further approximated as

$$F_R(1,\nu,\tilde{R}) \approx 1 - \frac{1}{\sigma_{sr}^2\nu}R_N - \frac{0.5}{\sigma_{sd}^2\sigma_{rd}^2\nu(1-\nu)}R_N^2. \tag{3.36}$$

By solving $\partial F_R(1,\nu,\tilde{R})/\partial\nu = 0$, we obtain $\tilde{\nu}_o$ as given in (3.35), which is the globally optimal ν to maximize $F_R(1,\nu,\tilde{R})$. After finding $\tilde{\nu}_o$, we can determine ν_o according to (3.34). □

Remark 3.5. It should be pointed out that in the derivation of (3.35), certain approximations are made and thus the solutions given in Theorem 3.6 are only approximately optimal. However, at relatively high rate normalized SNRs, the impact of approximation errors on the choice of $\tilde{\nu}_o$ becomes negligible. To demonstrate this, in Fig. 3.4, we compare the outage-optimal spatial power allocation obtained by using computer search with the one given in (3.35). It can be observed from the figure that these approaches offer almost the same power allocation strategies reflected in terms of the parameter ν, especially at high rate normalized SNR.

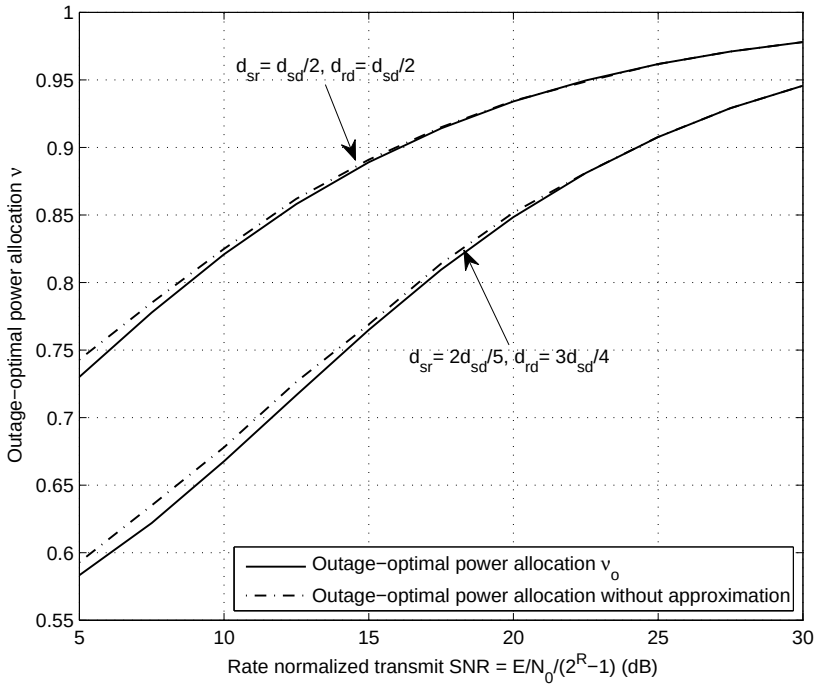

Figure 3.4: Outage-optimal spatial power allocation versus rate normalized transmit SNR.

It is easy to check that $\tilde{\nu}_o$ in (3.35) is in the interval $(0.5, 1)$, which suggests that the source should be allocated more power than the relay. This observation is consistent with what has been shown in [65, Eq. (25)] for the AF case. In addition, we note that $\tilde{\nu}_o$ is an increasing function of $\sigma_{sd}^2\sigma_{rd}^2/\sigma_{sr}^2$. Suppose that σ_{sd}^2 is fixed, e.g., the distance d_{sd} is fixed. With an increase of $\sigma_{rd}^2/\sigma_{sr}^2$, which can be realized by moving the relay toward the destination, more power should be allocated to the source, and vice versa.

Example 3.2. In Fig. 3.5, we illustrate the outage probabilities of the direct transmission, the relay transmission with equal power allocation $(\nu = 0.5)$ and the relay transmission with outage-optimal power allocation (ν_o in (3.34)). It shows that relay transmissions provide considerable outage performance improvements over the direct transmission, and the optimal power allocation strategy outperforms the equal

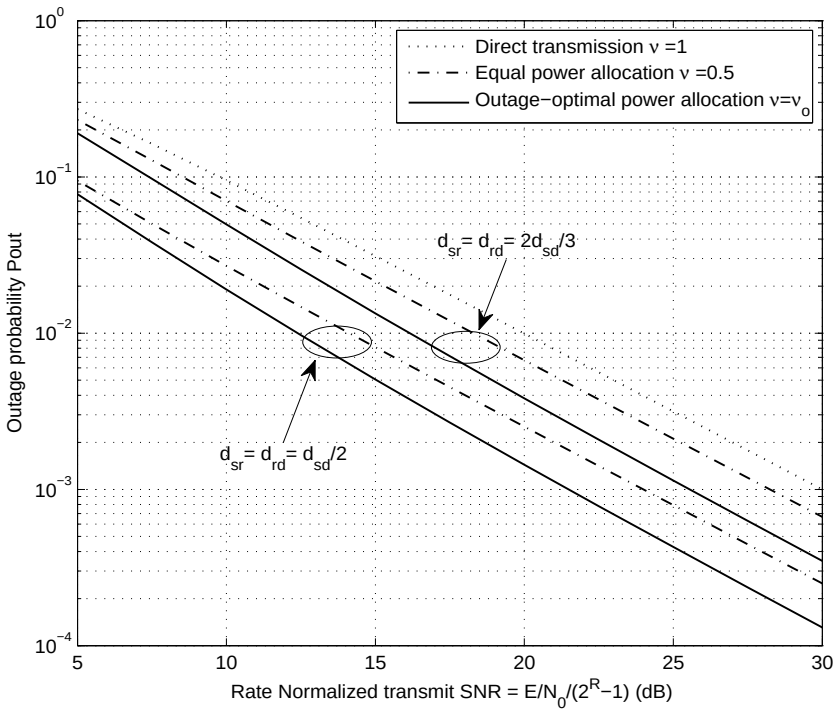

Figure 3.5: Outage probabilities versus rate normalized transmit SNR.

power one by roughly 2 dB of rate normalized SNR at an outage probability of 10^{-2}.

Example 3.3. In this example, we assume that the source, relay and destination are collinear, i.e., $d_{sr} + d_{rd} = d_{sd}$. Fig. 3.6 depicts the outage performance comparisons between the outage-optimal power allocation (ν_o) and the equal power allocation ($\nu = 0.5$) strategies. As can be seen from Fig. 3.6, the performance of the equal power allocation strategy approaches the optimal one as the relay moves toward the source, i.e., d_{sr}/d_{sd} approaches zero. This observation agrees with what we can infer from $\tilde{\nu}_o$ in (3.35).

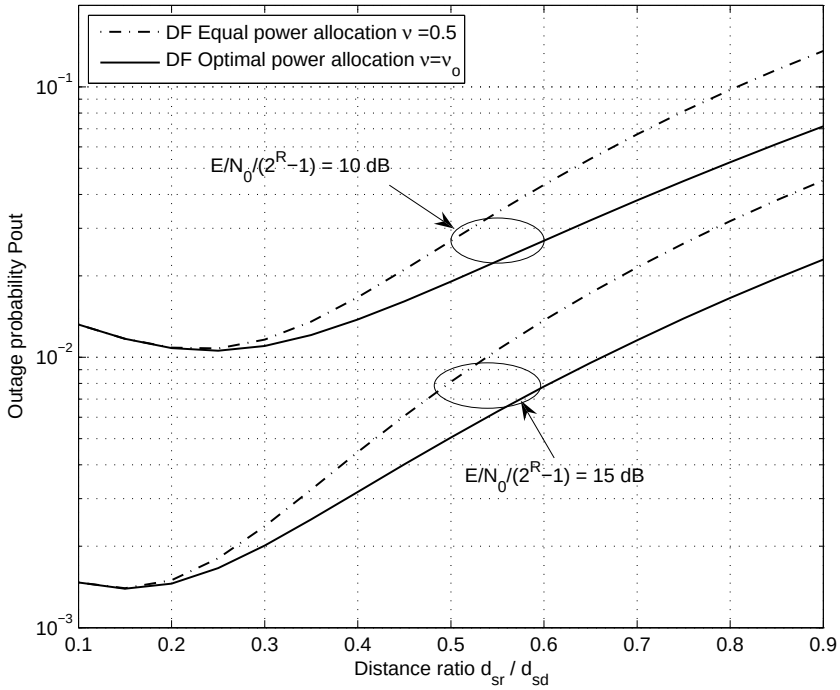

Figure 3.6: Outage probabilities versus distance ratio d_{sr}/d_{sd} when $d_{sr} + d_{rd} = d_{sd}$.

3.5 Ergodic-Optimal Transmission Strategies

In this section, we present ergodic-optimal transmission strategies which maximize the ergodic rate of ergodic fading channels for the two scenarios as discussed in the preceding section.

3.5.1 Ergodic-Optimal Transmit Signals Without Spatial Power Allocation

Following the similar discussion in Section 3.4.1, we consider the problem of designing an ergodic-optimal transmission strategy for any fixed ν as

$$\mu_e := \arg \max_{0 \le \mu \le 1} R_{\text{df}}(\mu, \nu) = \arg \max_{0 \le \mu \le 1} \min \left\{ R_{\text{sir}}(\mu, \nu), R_{\text{mac}}(\mu, \nu) \right\}. \quad (3.37)$$

It has been concluded in [49, Theorem 8] that $\mu_e = 1$ for a more general scenario, in which multiple relays are considered and the phases of the channel gains are only assumed to be uniformly distributed. We next introduce the following theorem that restates a part of the result [49, Theorem 8] for the single relay case, but complements it with our finding on the monotonicity of $R_{\text{df}}(\mu,\nu)$, and an alternative proof for the optimality of $\mu_e = 1$.

Theorem 3.7. The ergodic rate $R_{\text{df}}(\mu,\nu)$ is a monotonically increasing function of $\mu \in [0,1]$ for any fixed $\nu \in (0,1)$, and hence $\mu_e = 1$. Therefore, the maximum ergodic rate, $R^I_{\text{df}}(\nu)$, is

$$R^I_{\text{df}}(\nu) := R_{\text{df}}(1,\nu) = \min\left\{R^I_{\text{sir}}(\nu), R^I_{\text{mac}}(\nu)\right\}, \tag{3.38}$$

where

$$\begin{aligned} R^I_{\text{sir}}(\nu) &:= R_{\text{sir}}(1,\nu) = \exp(\eta_{sr}^{-1})\mathrm{E}_1(\eta_{sr}^{-1})\log_2 e, \\ R^I_{\text{mac}}(\nu) &:= R_{\text{mac}}(1,\nu) = \frac{\eta_{sd}\exp(\eta_{sd}^{-1})\mathrm{E}_1(\eta_{sd}^{-1}) - \eta_{rd}\exp(\eta_{rd}^{-1})\mathrm{E}_1(\eta_{rd}^{-1})}{\eta_{sd}-\eta_{rd}}\log_2 e. \end{aligned}$$

Proof. Clearly, $R_{\text{sir}}(\mu,\nu)$ is a monotonically increasing function of $\mu \in [0,1]$. However, the monotonicity of $R_{\text{mac}}(\mu,\nu)$ is not obvious. To demonstrate the monotonicity, we calculate

$$\begin{aligned} \frac{\partial R_{\text{mac}}(\mu,\nu)}{\partial \mu} &= -\frac{\eta_{sd}\eta_{rd}}{\alpha-\beta}\left[\frac{\partial R_{\text{mac}}(\mu,\nu)}{\partial \alpha} - \frac{\partial R_{\text{mac}}(\mu,\nu)}{\partial \beta}\right] \\ &= -\frac{\eta_{sd}\eta_{rd}}{\alpha-\beta}\,\mathbb{E}\left[\frac{w_1-w_2}{1+\alpha w_1+\beta w_2}\right], \end{aligned} \tag{3.39}$$

where the second equality follows from (3.6) and $R_{\text{mac}}(\mu,\nu) = \mathbb{E}[\log_2(1+\alpha w_1+\beta w_2)]$. It has been shown in [95, Theorem 3.1] that $\mathbb{E}\left[(w_1-w_2)/(1+\alpha w_1+\beta w_2)\right] \le 0$ when w_1 and w_2 are i.i.d. and $\alpha \ge \beta$. From (3.39), we can readily conclude that $\partial R_{\text{mac}}(\mu,\nu)/\partial\mu \ge 0$. It implies that $R_{\text{mac}}(\mu,\nu)$ is increasing in μ. Thus, $R_{\text{df}}(\mu,\nu)$ is an increasing function of μ (or a decreasing function of $|\rho|$), and it is maximized at

$\mu = 1$ (or $\rho = 0$). Lastly, we obtain $R^I_{\text{sir}}(\nu)$ and $R^I_{\text{mac}}(\nu)$ by evaluating (3.14) and (3.15) at $\mu = 1$, respectively. □

3.5.2 Ergodic-Optimal Spatial Power Allocations

We now investigate ergodic-optimal spatial power allocation strategies. Since the ergodic-optimal μ is one for any $\nu \in (0, 1)$ (see Theorem 3.7), the two-dimensional optimization problem reduces to a one-dimensional optimization problem, in which we only need determine the value of ν that maximizes the ergodic rate $R^I_{\text{df}}(\nu)$. For the DF relay channel, the ergodic-optimal power allocation strategy is $\tilde{\nu}_e$ such that

$$\tilde{\nu}_e = \arg \max_{0<\nu<1} R^I_{\text{df}}(\nu). \tag{3.40}$$

Moreover, in the case of $\nu = 1$, the ergodic rate is given by

$$R_{\text{sd}} = \mathbb{E}[I_{\text{sd}}] = \mathbb{E}[\log_2(1 + |h_{sd}|^2 E/N_0)] = \exp\Big(\frac{N_0}{\sigma^2_{sd}E}\Big)\mathrm{E}_1\Big(\frac{N_0}{\sigma^2_{sd}E}\Big)\log_2 e.$$

To determine whether to use the direct transmission or the DF relay transmission with optimal power allocation $\tilde{\nu}_e$, we need compare R_{sd} with $R^I_{\text{df}}(\tilde{\nu}_e)$. Thus, the ergodic-optimal power allocation ν_e is determined as

$$\nu_e = \begin{cases} \tilde{\nu}_e, & R^I_{\text{df}}(\tilde{\nu}_e) > R_{\text{sd}}, \\ 1, & R^I_{\text{df}}(\tilde{\nu}_e) \leq R_{\text{sd}}. \end{cases} \tag{3.41}$$

We denote the corresponding maximum ergodic DF rate as

$$R^J_{\text{df}} = \max\{R^I_{\text{df}}(\tilde{\nu}_e), R_{\text{sd}}\}. \tag{3.42}$$

Theorem 3.8. If $\sigma^2_{sr} \leq \sigma^2_{sd}$, the ergodic-optimal transmission strategy is to use only direct transmission and allocate all the transmit power at the source, i.e., $\nu_e = 1$, with $R^J_{\text{df}} = R_{\text{sd}}$.

Proof. It follows from (3.38) that $R^I_{\text{df}}(\nu) \leq R^I_{\text{sir}}(\nu)$. Since $R^I_{\text{sir}}(\nu)$ is increasing in ν and $\sigma^2_{sr} \leq \sigma^2_{sd}$, we have

$$R^I_{\text{df}}(\nu) \leq R^I_{\text{sir}}(\nu) \leq \lim_{\nu \to 1} R^I_{\text{sir}}(\nu) \leq R_{\text{sd}}, \qquad \text{for all } \nu \in (0,1).$$

It implies that to achieve the highest ergodic rate, all the power should be allocated to the source. □

As discussed in Section 3.4.2, the source-relay link is typically stronger than the source-destination link, i.e., $\sigma^2_{sr} > \sigma^2_{sd}$. In such a case, we have

$$R^I_{\text{df}}(\hat{\nu}_e) \geq \lim_{\nu \to 1} R^I_{\text{df}}(\nu) = R_{\text{sd}},$$

and thus $\nu_e = \tilde{\nu}_e$. To determine $\tilde{\nu}_e$, we need to know the monotonicity of $R^I_{\text{sir}}(\nu)$ and $R^I_{\text{mac}}(\nu)$. Clearly $R^I_{\text{sir}}(\nu)$ is an increasing function, but the monotonicity of $R^I_{\text{mac}}(\nu)$ is not evident. In the following lemma, we present the monotonic properties of $R^I_{\text{mac}}(\nu)$.

Lemma 3.4. The function $R^I_{\text{mac}}(\nu)$ is an increasing function for $\nu \in [0, \widehat{\nu})$, and a decreasing function for $\nu \in (\widehat{\nu}, 1)$, where

$$\widehat{\nu} := \arg \max_{0<\nu<1} R^I_{\text{mac}}(\nu)$$

and is determined as follows:

a) If $\sigma^2_{sd} = \sigma^2_{rd}$, then $\widehat{\nu}$ equals $1/2$;

b) If $\sigma^2_{sd} \neq \sigma^2_{rd}$, then $\widehat{\nu}$ is the unique root of the following equation

$$\begin{aligned} &\left[\frac{\sigma^2_{sd}\sigma^2_{rd}P}{\eta_{sd}-\eta_{rd}} + \frac{\sigma^2_{sd}}{\eta_{sd}}\right] \exp(\eta^{-1}_{sd})\mathrm{E}_1(\eta^{-1}_{sd}) - \sigma^2_{sd} = \\ &\left[\frac{\sigma^2_{sd}\sigma^2_{rd}P}{\eta_{sd}-\eta_{rd}} - \frac{\sigma^2_{rd}}{\eta_{rd}}\right] \exp(\eta^{-1}_{rd})\mathrm{E}_1(\eta^{-1}_{rd}) + \sigma^2_{rd}. \end{aligned} \tag{3.43}$$

Proof. Based on (3.6) and the optimality of $\mu = 1$, $R^I_{\text{mac}}(\nu)$ can be rewritten as

$$R^I_{\text{mac}}(\nu) = \mathbb{E}[\log_2(1 + \eta_{sd}w_1 + \eta_{rd}w_2)].$$

Because the second partial derivative of $R^I_{\text{mac}}(\nu)$ is always non-positive, i.e.,

$$\frac{\partial^2}{\partial \nu^2} R^I_{\text{mac}}(\nu) = -\log_2 e \cdot \mathbb{E}\left[\frac{(E\sigma^2_{sd} w_1 - P\sigma^2_{rd} w_2)^2}{(1+\eta_{sd} w_1 + \eta_{rd} w_2)^2}\right] \leq 0,$$

the optimization problem, $\max_\nu R^I_{\text{mac}}(\nu)$, is a convex problem. It is straightforward to calculate $\partial R^I_{\text{mac}}(\nu)/\partial \nu = P(\phi_1 - \phi_2)\log_2 e$ where $\phi_1 := \mathbb{E}[\sigma^2_{sd} w_1/(1+\eta_{sd} w_1 + \eta_{rd} w_2)]$ and $\phi_2 := \mathbb{E}[\sigma^2_{rd} w_2/(1+\eta_{sd} w_1 + \eta_{rd} w_2)]$. The global optimum $\widehat{\nu}$ is the unique solution of $\phi_1 = \phi_2$ as shown in (3.43). Since $\partial^2 R^I_{\text{mac}}(\nu)/\partial \nu^2$ is non-positive, $\partial R^I_{\text{mac}}(\nu)/\partial \nu$ is greater than zero and $R^I_{\text{mac}}(\nu)$ is increasing in ν for $\nu \in [0, \widehat{\nu})$; $\partial R^I_{\text{mac}}(\nu)/\partial \nu$ is less than zero and $R^I_{\text{mac}}(\nu)$ is decreasing in ν for $\nu \in (\widehat{\nu}, 1]$. In particular, when $\sigma^2_{sd} = \sigma^2_{rd}$, we have $\widehat{\nu} = 0.5$. □

We next introduce the ergodic-optimal spatial power allocation strategy for the case of $\sigma^2_{sr} > \sigma^2_{sd}$.

Theorem 3.9. If $\sigma^2_{sr} > \sigma^2_{sd}$, the ergodic-optimal spatial power allocation $\nu_e = \tilde{\nu}_e$ is:

a) If $R^I_{\text{sir}}(\widehat{\nu}) \geq R^I_{\text{mac}}(\widehat{\nu})$, then $\tilde{\nu}_e = \widehat{\nu}$ and $R^J_{\text{df}} = R^I_{\text{mac}}(\widehat{\nu})$.

b) If $R^I_{\text{sir}}(\widehat{\nu}) < R^I_{\text{mac}}(\widehat{\nu})$, then $\tilde{\nu}_e = \nu^* \in (\widehat{\nu}, 1)$ and $R^J_{\text{df}} = R^I_{\text{sir}}(\nu^*)$, where ν^* denotes the unique root of $R^I_{\text{sir}}(\nu) = R^I_{\text{mac}}(\nu)$ in $(\widehat{\nu}, 1)$.

Proof. a) If $R^I_{\text{sir}}(\widehat{\nu}) \geq R^I_{\text{mac}}(\widehat{\nu})$, the inequalities $R^I_{\text{df}}(\nu) \leq R^I_{\text{mac}}(\nu) \leq R^I_{\text{mac}}(\widehat{\nu}) \leq R^I_{\text{sir}}(\widehat{\nu})$ hold for any $\nu \in (0,1)$. This fact indicates that $\tilde{\nu}_e = \widehat{\nu}$ and $R^J_{\text{df}} = R^I_{\text{mac}}(\widehat{\nu})$ by the definition of R^J_{df}.

b) Define $\psi(\nu) := R^I_{\text{sir}}(\nu) - R^I_{\text{mac}}(\nu)$. Since $R^I_{\text{sir}}(\widehat{\nu}) < R^I_{\text{mac}}(\widehat{\nu})$ and $\sigma^2_{sr} > \sigma^2_{sd}$, we have $\psi(\widehat{\nu}) < 0$ and $\lim_{\nu \to 1} \psi(\nu) > 0$. This implies that $\psi(\nu)$ has at least one root $\nu^* \in (\widehat{\nu}, 1)$. Furthermore, the fact that $R^I_{\text{mac}}(\nu)$ is a decreasing function of $\nu \in (\widehat{\nu}, 1)$ (Lemma 3.4), together with the monotonicity of $R^I_{\text{sir}}(\nu)$, implies that $\psi(\nu)$ is a monotonically increasing function of $\nu \in (\widehat{\nu}, 1)$. Therefore, $\psi(\nu)$ has a unique

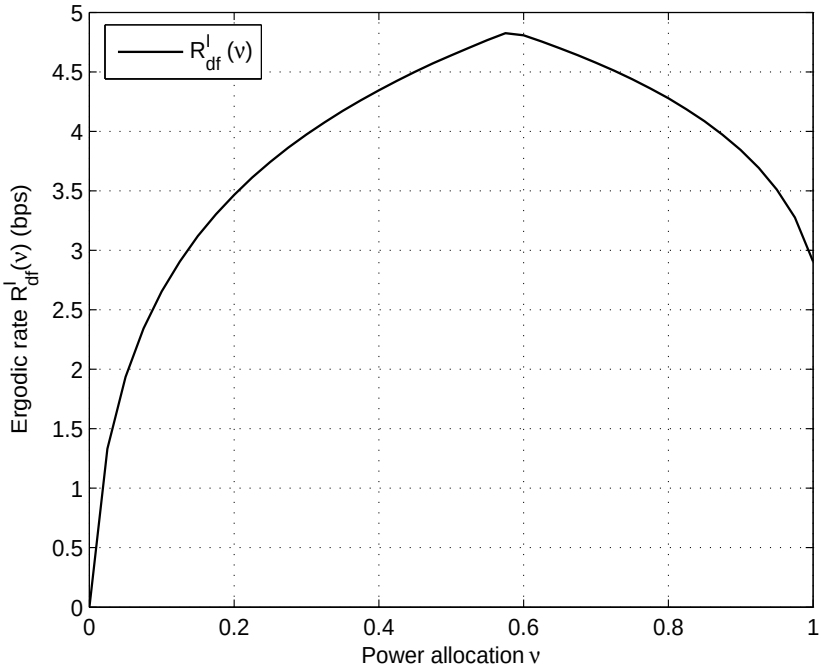

Figure 3.7: Ergodic rates $R^I_{\mathbf{df}}(\nu)$ versus power allocation ratio ν when $E/N_0 = 10$ dB and $d_{sr} = d_{rd} = d_{sd}/2$.

root $\nu^* \in (\widehat{\nu}, 1)$. With $R^I_{\text{df}}(\nu) \leq R^I_{\text{sir}}(\nu^*)$ for $\nu \in (0, \nu^*]$ and $R^I_{\text{df}}(\nu) \leq R^I_{\text{mac}}(\nu^*)$ for $\nu \in (\nu^*, 1)$, we can easily conclude that $\tilde{\nu}_e = \nu^*$ and $R^J_{\text{df}} = R^I_{\text{sir}}(\nu^*)$. □

Example 3.4. In Fig. 3.7, we plot $R^I_{\text{df}}(\nu)$ versus ν for $E/N_0 = 10$ dB and $d_{sr} = d_{rd} = d_{sd}/2$. It can be observed from Fig. 3.7 that the highest rate is achieved at roughly $\nu = 0.58$. Based on Theorem 3.9, we obtain $\widehat{\nu} = 10^{-4}$ and $\nu_e = \nu^* = 0.5825$ by numerically solving the respective equations (3.43) and $R^I_{\text{sir}}(\nu) = R^I_{\text{mac}}(\nu)$. This result empirically validates our findings in Theorem 3.9.

Example 3.5. In Fig. 3.8, we compare the ergodic rates with the equal power allocation ($\nu = 0.5$), and with the ergodic-optimal power allocation (ν_e). The ergodic-optimal power allocation strategy outperforms the equal power allocation one

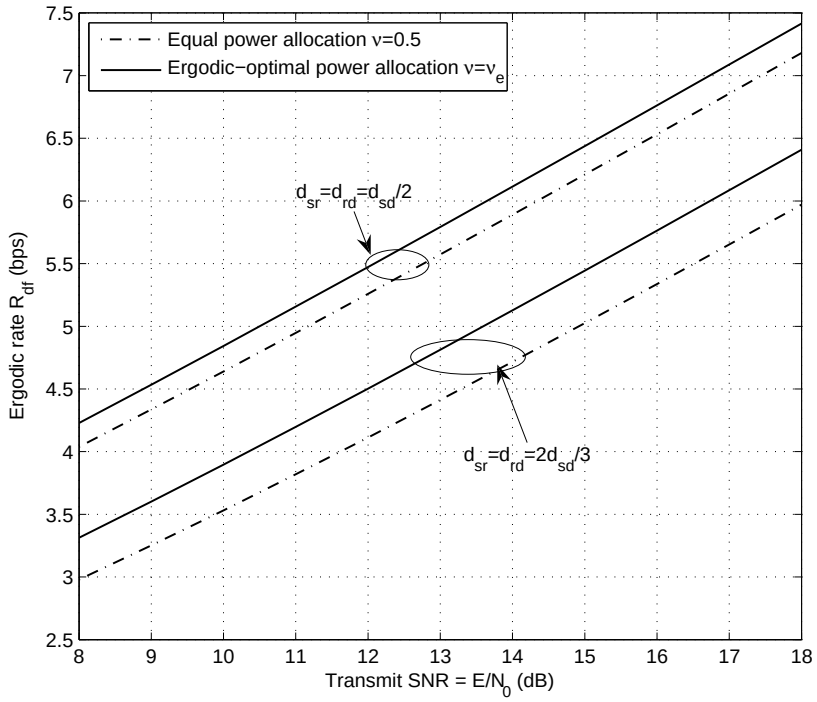

Figure 3.8: Ergodic rates versus transmit SNR.

by 0.8 dB for the case where $d_{sr} = d_{rd} = d_{sd}/2$, and by 1.5 dB for the case where $d_{sr} = d_{rd} = 2d_{sd}/3$.

Example 3.6. Assuming that the source, relay, and destination are collinear, i.e., $d_{sr} + d_{rd} = d_{sd}$, we study the impacts of the location of the relay on the ergodic rate in Fig. 3.9, for the equal power allocation ($\nu = 0.5$) and ergodic-optimal power allocation (ν_e). We observe that, when the relay is far from the source ($d_{sr}/d_{sd} > 0.5$), the ergodic-optimal power allocation strategy can greatly improve the ergodic rates. However, when the relay is close to the source ($d_{sr}/d_{sd} < 0.5$), the gain achieved by adopting the ergodic-optimal power allocation becomes insignificant. In Fig. 3.10, we plot the corresponding ergodic-optimal power allocation ν_e versus d_{sr}/d_{sd}.

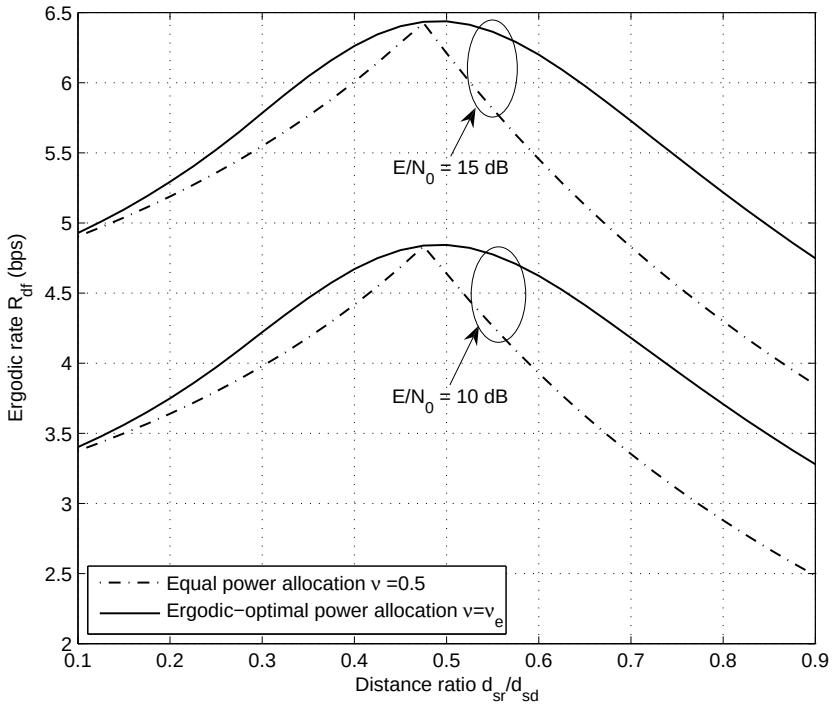

Figure 3.9: Ergodic rates versus distance ratio d_{sr}/d_{sd} when $d_{sr} + d_{rd} = d_{sd}$.

To determine $\widehat{\nu}$ and ν^*, we need solve transcendental equations involving $\mathrm{E}_1(\mathrm{x})$ function. However, at low or high SNR, the respective optimal ν can be determined as follows.

Low SNR Regime

When the SNR is sufficiently small, η_{sr}^{-1}, η_{sd}^{-1} and η_{rd}^{-1} are much greater than one. Applying $\mathrm{E}_1(z) \approx e^{-z}/z$ for any $z \gg 1$, we approximate $R_{\text{sir}}^I(\nu)$ and $R_{\text{mac}}^I(\nu)$ as

$$R_{\text{sir},l}^I(\nu) \approx \sigma_{sr}^2 P \nu \log_2 e$$

$$R_{\text{mac},l}^I(\nu) \approx [(\sigma_{sd}^2 - \sigma_{rd}^2)\nu + \sigma_{rd}^2] P \log_2 e.$$

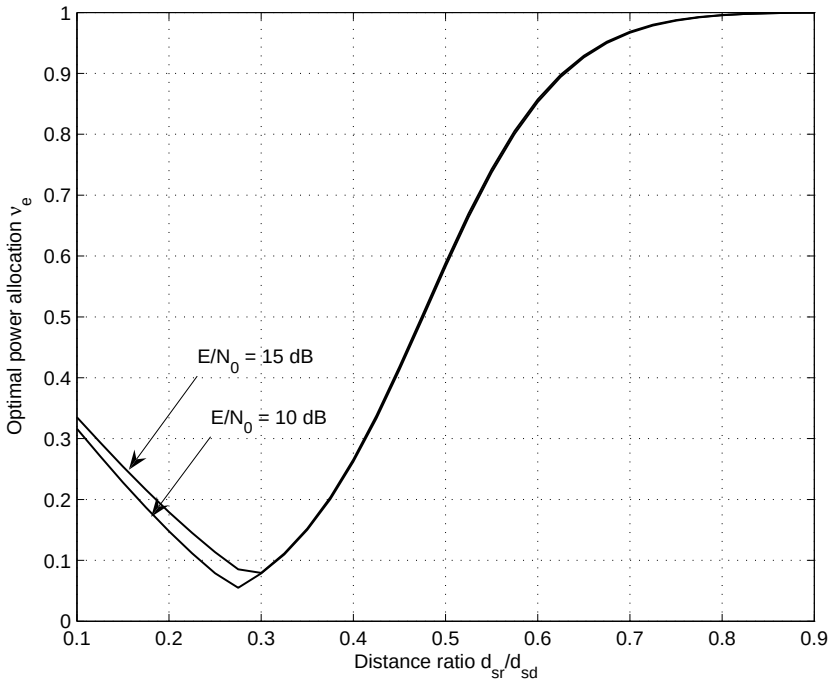

Figure 3.10: Optimal power allocation ration ν_e versus distance ratio d_{sr}/d_{sd} when $d_{sr} + d_{rd} = d_{sd}$.

If $\sigma_{sd}^2 \geq \sigma_{rd}^2$, it is obvious that $\nu = 1$ is optimal. If $\sigma_{sd}^2 < \sigma_{rd}^2$, we can obtain the optimal value of ν by setting $R_{\text{sir},l}^I(\nu) = R_{\text{mac},l}^I(\nu)$, and the optimal value of ν is

$$\nu_{e,l} := \sigma_{rd}^2(\sigma_{rd}^2 + \sigma_{sr}^2 - \sigma_{sd}^2)^{-1}.$$

High SNR Regime

When the SNR is sufficiently high, $\eta_{sr}^{-1}, \eta_{sd}^{-1}$ and η_{rd}^{-1} are much less than one. Applying the approximations, $\mathrm{E}_1(z) \approx -\gamma - \ln z$ and $e^z \approx 1$ for $z \ll 1$ with γ denoting the Euler's constant ($\gamma = 0.5772\ldots$), we approximate $R_{\text{sir}}^I(\nu)$ and $R_{\text{mac}}^I(\nu)$ as

$$\begin{aligned} R_{\text{sir},h}^I(\nu) &\approx -\gamma \log_2 e + \log_2 \eta_{sr} \\ R_{\text{mac},h}^I(\nu) &\approx -\gamma \log_2 e + \frac{\eta_{sd} \log_2 \eta_{sd} - \eta_{rd} \log_2 \eta_{rd}}{\eta_{sd} - \eta_{rd}}. \end{aligned}$$

The parameter $\widehat{\nu}_h$ that maximizes $R^I_{\text{mac},h}(\nu)$ is the solution of

$$(\eta_{sd} - \eta_{rd})/(\ln \eta_{sd} - \ln \eta_{rd}) = \sigma^2_{sd}\sigma^2_{rd}P/(\sigma^2_{sd} + \sigma^2_{rd}).$$

If $R^I_{\text{sir},h}(\widehat{\nu}_h) \geq R^I_{\text{mac},h}(\widehat{\nu}_h)$, the optimal value of ν is $\nu_{e,h} = \widehat{\nu}_h$. If $R^I_{\text{sir},h}(\widehat{\nu}_h) < R^I_{\text{mac},h}(\widehat{\nu}_h)$, the optimal value of ν is $\nu_{e,h} = \nu^*_h \in (\widehat{\nu}_h, 1)$, where ν^*_h is the unique solution of $R^I_{\text{sir},h}(\nu) = R^I_{\text{mac},h}(\nu)$.

3.6 Conclusion

We investigated optimal transmit signaling designs and spatial power allocation strategies in both outage and ergodic settings for Rayleigh fading DF relay channels. We considered two practical scenarios. In the first scenario where there is no spatial power allocation available, we showed that the outage-optimal transmit signals are not necessarily independent, and their correlation coefficient is generally determined by several system parameters, whereas the ergodic rate in ergodic fading channels is strictly decreasing in the correlation coefficient, and thus the ergodic-optimal transmit signals are always independent. In the second scenario where spatial power allocation is feasible, we derived optimal spatial power allocation strategies which depend on the second-order statistics of three links and SNRs. We further disclosed that it is not always beneficial to use a relay in this scenario.

Chapter 4

Outage Probability of Rician Fading Relay Channels

In this chapter, we extend the outage probability results in Chapter 3 for Rayleigh fading channels to the ones in Rician fading channels. We investigate the outage performance of the full-duplex DF relay system in a Rician fading environment. We derive an analytical expression for the outage probability of the highest achievable information rate of the system, which can be evaluated by employing numerical techniques. We also obtain an upper bound and a lower bound on the outage probability. These bounds are given in simple closed forms and can be easily evaluated. Furthermore, we study the impact of transmit signaling on the outage performance.

4.1 Introduction

The cooperative relaying transmission technique has attracted considerable research attention as it is capable of improving the reliability and power efficiency of wireless networks. As an important performance measure in fading scenarios, the outage probability of the achievable information rate has been investigated in Chapter 3 for

Rayleigh fading channels. However, in practice, the wireless channels are some time found to be Gaussian distributed with nonzero mean, i.e., Rician fading with finite Rice factors. The outage performance analysis of such channels is, therefore, of practical importance.

In this chapter, we consider a full-duplex DF relay system in a Rician fading environment. We first express the highest achievable information rate of the system in a Hermitian quadratic form to facilitate us to evaluate its statistical distribution. We then derive an analytical expression for the outage probability, which can be evaluated by using standard numerical techniques. Based on the geometrical representation of the outage probability, we also obtain an upper bound and a lower bound on the outage probability in simple closed forms. These bounds can work as fairly good approximations to the outage probability. Moreover, when the channel statistics (Rice factors and variances) are known at the source and relay nodes, the outage probability can be minimized by choosing an appropriate correlation coefficient between the transmit signals from the source and the relay. Relying on numerical methods, we obtain the optimal correlation coefficient that minimizes the outage probability. Our numerical results reveal that for large values of Rice factors, the optimal correlation coefficient is not necessarily zero, but instead, depends generally on SNRs, variances and Rice factors of the channels, as well as the rate threshold.

4.2 System Model

We consider the same DF single-relay system model as in Chapter 3, but operating in a Rician fading environment. The channel gains are still modeled as independent complex Gaussian random variables, but they are no longer zero mean. To differentiate with the channel gains in Chapter 3, we use $\tilde{h}_{ij}$ $(ij \in \{sr, sd, rd\})$ in this chapter to

denote the channel gain between nodes i and j in a Rician fading environment. For Rician fading channels, each $\tilde{h}_{ij}$ can be decomposed into the sum of a deterministic (or LOS) component and a variable (scattered) component, i.e.,

$$\tilde{h}_{ij} = \sqrt{K_{ij}\sigma_{ij}^2} + h_{ij}. \tag{4.1}$$

The h_{ij}s are independent zero-mean complex Gaussian RVs with variances $\mathbb{E}(|h_{ij}|^2) = \sigma_{ij}^2$. The Rice factor, $K_{ij} \geq 0$, represents the ratio of deterministic energy to random energy in the corresponding link. Thus, the channel gain $\tilde{h}_{ij}$ is complex Gaussian distributed with mean $\sqrt{K_{ij}\sigma_{ij}^2}$ and variance σ_{ij}^2, i.e., $\tilde{h}_{ij} \sim \mathcal{CN}\big(\sqrt{K_{ij}\sigma_{ij}^2}\,, \sigma_{ij}^2\big)$. We use the same notations as in Chapter 3: $\eta_{sr} = E_s\sigma_{sr}^2/N_0, \eta_{sd} = E_s\sigma_{sd}^2/N_0$, and $\eta_{rd} = E_r\sigma_{rd}^2/N_0$.

4.3 Information Rate

Similar to the information rate in Chapter 3.3.2, the highest information rate achieved by the DF single-relay systems can be written as

$$\begin{aligned} &\tilde{I}_{\text{df}} = \min\left\{\tilde{I}_{\text{sir}}, \tilde{I}_{\text{mac}}\right\}, \\ &\tilde{I}_{\text{sir}} = \log_2\left(1 + |\tilde{h}_{sr}|^2(1-|\rho|^2)E_s/N_0\right), \quad \tilde{I}_{\text{mac}} = \log_2(1+\tilde{Q}), \end{aligned} \tag{4.2}$$

where ρ is still the correlation coefficient of x_s and x_r as in Chapter 3.3.2, i.e., $\rho = \mathbb{E}[x_s x_r^*]$, and $\tilde{Q}$ is given as

$$\tilde{Q} = |\tilde{h}_{sd}|^2 E_s/N_0 + |\tilde{h}_{rd}|^2 E_r/N_0 + 2\Re(\rho\tilde{h}_{sd}\tilde{h}_{rd}^*)\sqrt{E_s E_r}/N_0. \tag{4.3}$$

The outage probability of the information rate $\tilde{I}_{\text{df}}$ at a given rate threshold R in a Rician fading environment is written as

$$P_{\text{out}}(\rho, R) = P(\tilde{I}_{\text{df}} \leq R) = 1 - P(\tilde{I}_{\text{sir}} > R)P(\tilde{I}_{\text{mac}} > R). \tag{4.4}$$

Notice that since $\tilde{h}_{sr}$ is nonzero mean Gaussian distributed, $|\tilde{h}_{sr}|^2$ is noncentral chi-square distributed with two degrees of freedom. The CCDF of $\tilde{I}_{\text{sir}}$ is computed as

$$P(\tilde{I}_{\text{sir}} > R) = Q_1\Big(\sqrt{2K_{sr}}, \sqrt{2\tilde{R}/\big[\eta_{sr}(1-|\rho|^2)\big]}\Big), \tag{4.5}$$

where $\tilde{R} = 2^R - 1$ and $Q_1(\cdot,\cdot)$ denotes the first-order Marcum Q-function [96, Eq. (4.33)]. Since $\tilde{I}_{\text{mac}} = \log_2(1+\tilde{Q})$, the CCDF of $\tilde{I}_{\text{mac}}$ becomes

$$P(\tilde{I}_{\text{mac}} > R) = P(\tilde{Q} > \tilde{R}). \tag{4.6}$$

In order to evaluate $P(\tilde{I}_{\text{mac}} > R)$, we need study the statistical properties of $\tilde{Q}$, which will be presented in the following section.

4.4 Quadratic Forms in Nonzero Mean Complex Gaussian RVs

We can see that $\tilde{Q}$ has the same form as Q in (3.3) except that $\tilde{h}_{sd}$ and $\tilde{h}_{rd}$ are nonzero mean. In this section, we first represent $\tilde{Q}$ as a Hermitian quadratic form in nonzero mean complex Gaussian RVs to facilitate us to evaluate its distribution, and we next derive the exact distribution of $\tilde{Q}$, which is given in an analytical form.

4.4.1 A Hermitian Quadratic Form Representation

Using the same technique as in Chapter 3.3.2, we can rewrite $\tilde{Q}$ as a Hermitian quadratic form in nonzero mean complex Gaussian RVs:

$$\tilde{Q} = \alpha|\tilde{h}_1|^2 + \beta|\tilde{h}_2|^2, \tag{4.7}$$

where α and β are given by (3.7) in Chapter 3.3.2, $\tilde{h}_1$ and $\tilde{h}_2$ are independent complex Gaussian distributed with nonzero mean and unit variance. Then, $|\tilde{h}_1|^2$ and $|\tilde{h}_2|^2$ are

independent noncentral chi-square random variables with two degrees of freedom and respective non-centrality parameters

$$s_1^2 = \left[\alpha K_{srd} - (\eta_{sd}K_{rd} + \eta_{rd}K_{sd}) + 2\sqrt{\eta_{sd}\eta_{rd}K_{sd}K_{rd}}\Re(\rho)\right]/(\alpha - \beta), \tag{4.8}$$

$$s_2^2 = \left[\alpha K_{srd} - (\eta_{sd}K_{sd} + \eta_{rd}K_{rd}) - 2\sqrt{\eta_{sd}\eta_{rd}K_{sd}K_{rd}}\Re(\rho)\right]/(\alpha - \beta), \tag{4.9}$$

where $K_{srd} = K_{sd} + K_{rd}$. Note that when α equals β, both the numerators and denominators in (4.8) and (4.9) are zero. In this case, $\tilde{Q}$ in (4.7) is noncentral chi-square distributed with four degrees of freedom and non-centrality parameter K_{srd}.

Comparing (4.3) with (4.7), we can see that the rate $\tilde{I}_{\text{mac}}$ achieved by sending correlated signals (x_s and x_r) with the transmit powers E_s and E_r over independent nonzero mean fading channels ($\tilde{h}_{sd}$ and $\tilde{h}_{rd}$), can be viewed as the one achieved by sending independent signals with transmit powers α and β over independent nonzero mean fading channels ($\tilde{h}_1$ and $\tilde{h}_2$). In fact, we can see from the expressions of $\tilde{Q}$ that $P(\tilde{I}_{\text{mac}} < R) = P(\tilde{Q} < \tilde{R})$ is the same as the outage probability of a MISO system [97].

4.4.2 Distribution of the Hermitian Quadratic Form

As expressed in (4.7), $\tilde{Q}$ is a linear combination (weighted sum) of independent noncentral chi-square random variables. Its distribution can be expressed as infinite series [93, Ch. 4.2]. In the following lemmas, we make use of (4.7) to derive the characteristic function and statistical distributions of $\tilde{Q}$.

Lemma 4.1. The characteristic function of $\tilde{Q}$ is given by

$$\mathbb{E}[e^{jw\tilde{Q}}] = \frac{1}{(1 - jw\alpha)(1 - jw\beta)} \exp\left(\frac{jw\alpha s_1^2}{1 - jw\alpha} + \frac{jw\beta s_2^2}{1 - jw\beta}\right). \tag{4.10}$$

Proof. Based on (4.7) and the statistical independence of $\tilde{h}_1$ and $\tilde{h}_2$, the characteristic

function of $\tilde{Q}$ can be written as

$$\mathbb{E}[e^{jw\tilde{Q}}] = \mathbb{E}[e^{jw\alpha|\tilde{h}_1|^2}] \times \mathbb{E}[e^{jw\beta|\tilde{h}_2|^2}].$$

Since $|\tilde{h}_1|^2$ and $|\tilde{h}_2|^2$ are noncentral chi-square distributed, their characteristic functions can be readily obtained as $\exp[jw\alpha s_1^2/(1 - jw\alpha)]/(1 - jw\alpha)$ and $\exp[jw\beta s_2^2/(1-jw\beta)]/(1-jw\beta)$, respectively. Thus, (4.10) follows immediately. □

Lemma 4.2. For $|\rho| = 1$ ($\rho = e^{j\theta}$), the CCDF of the quadratic form $\tilde{Q}$, $P(\tilde{Q} > \tilde{R})$, is given by

$$P(\tilde{Q} > \tilde{R}) = Q_1\Big(\sqrt{2s^2}, \sqrt{2\tilde{R}/(\eta_{sd} + \eta_{rd})}\Big), \tag{4.11}$$

where $s^2 = (\eta_{sd}K_{sd} + \eta_{rd}K_{rd} + 2\sqrt{\eta_{sd}\eta_{rd}K_{sd}K_{rd}}\cos\theta)/(\eta_{sd} + \eta_{rd})$.

Proof. In this case, we have $\alpha = \eta_{sd} + \eta_{rd}$ and $\beta = 0$. Thus, the quadratic form $\tilde{Q}$ is noncentral chi-square distributed with two degrees of freedom, and $P(\tilde{Q} > \tilde{R})$ is readily obtained as (4.11). □

Lemma 4.3. For $|\rho| \in [0, 1)$, the CCDF of the quadratic form $\tilde{Q}$, $P(\tilde{Q} > \tilde{R})$, is expressed as

$$\begin{aligned} P(\tilde{Q} > \tilde{R}) =& Q_1\left(\sqrt{2}\, s_1, \sqrt{2\frac{\tilde{R}}{\alpha}}\right) + \frac{e^{-s_1^2}}{\alpha} \times \\ & \int_0^{\tilde{R}} e^{-z/\alpha} I_0\left(2s_1\sqrt{\frac{z}{\alpha}}\right) Q_1\left(\sqrt{2}\, s_2, \sqrt{2\frac{\tilde{R}-z}{\beta}}\right) dz, \end{aligned} \tag{4.12}$$

where $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind.

Proof. When $|\rho| < 1$, we have $\alpha \geq \beta > 0$. Defining $\tilde{w}_1 = \alpha|\tilde{h}_1|^2$ and $\tilde{w}_2 = \beta|\tilde{h}_2|^2$, we can write $\tilde{Q} = \tilde{w}_1 + \tilde{w}_2$. By using Jacobian transformations, we obtain the PDFs of $\tilde{w}_1$ and $\tilde{w}_2$ as

$$f_{\tilde{w}_1}(x) = \frac{1}{\alpha} e^{-(s_1^2 + x/\alpha)} I_0\left(2s_1\sqrt{x/\alpha}\right), \quad f_{\tilde{w}_2}(x) = \frac{1}{\beta} e^{-(s_2^2 + x/\beta)} I_0\left(2s_2\sqrt{x/\beta}\right).$$

Because of the statistical independence of $\tilde{h}_1$ and $\tilde{h}_2$, the PDF of $\tilde{Q}$ is the convolution of $f_{\tilde{w}_1}(x)$ with $f_{\tilde{w}_2}(x)$, i.e.,

$$f_{\tilde{Q}}(x) = \int_0^x f_{\tilde{w}_1}(z) f_{\tilde{w}_2}(x-z) dz.$$

Once the PDF of $\tilde{Q}$ is known, the CCDF of $\tilde{Q}$ can be computed as $P(\tilde{Q} > \tilde{R}) = \int_{\tilde{R}}^{\infty} f_{\tilde{Q}}(y) dy$, and then expressed as a double integral

$$P(\tilde{Q} > \tilde{R}) = \int_{\tilde{R}}^{\infty} \int_0^y f_{\tilde{w}_1}(z) f_{\tilde{w}_2}(y-z) dz dy.$$

By interchanging the order of integrations, we can obtain $P(\tilde{Q} > \tilde{R})$ as

$$P(\tilde{Q} > \tilde{R}) = \int_0^{\tilde{R}} f_{\tilde{w}_1}(z) P(\tilde{w}_2 > \tilde{R} - z) dz + \int_{\tilde{R}}^{\infty} f_{\tilde{w}_1}(z) P(\tilde{w}_2 > 0) dz.$$

Since $P(\tilde{w}_2 > \tilde{R} - z) = Q_1\Big(\sqrt{2}\, s_2, \sqrt{2(\tilde{R} - z)/\beta}\Big)$ and $P(\tilde{w}_2 > 0) = 1$, the CCDF of $\tilde{Q}$ reduces to (4.12). □

Remark 4.1. When $\rho = 0$ and $\eta_{sd} = \eta_{rd}$, we have $\alpha = \beta = \eta_{sd}$. Thus, the CCDF in (4.12) reduces to $P(\tilde{Q}_e > \tilde{R}) = Q_2\big(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\eta_{sd}}\big)$, where $Q_2(\cdot, \cdot)$ denotes the second-order Marcum Q-function defined in [96, Eq. (4.59)], and $\tilde{Q}_e$ is obtained by evaluating $\tilde{Q}$ at $\rho = 0$ with $\eta_{sd} = \eta_{rd}$.

Remark 4.2. In a Rayleigh fading scenario, we have $K_{sd} = K_{rd} = 0$. Due to the fact that $I_0(0) = 1$ and $Q_1(0, b) = \exp(-b^2/2)$, the CCDF of $\tilde{Q}$ in (4.12) reduces to

$$P(Q > \tilde{R}) = (\alpha e^{-\tilde{R}/\alpha} - \beta e^{-\tilde{R}/\beta})/(\alpha - \beta), \tag{4.13}$$

where Q is obtained by evaluating $\tilde{Q}$ at $K_{sd} = K_{rd} = 0$, and it is equal to the Q given in Chapter 3 (3.3). This coincides with the result presented in Chapter 3 (3.9) for the Rayleigh fading relay case.

Remark 4.3. In the Rician fading case, $P(\tilde{Q} > \tilde{R})$ in (4.12) depends on both the amplitude and phase of ρ, while in the Rayleigh fading case, $P(Q > \tilde{R})$ in (4.13) only

depends on the amplitude of ρ. This is because, in the Rician fading case, the phases of h_{sd} and h_{rd} are no longer uniformly distributed, and they depend on the Rice factors K_{sd} and K_{rd}, respectively. Thus, the phase of ρ affects the distribution of $\rho h_{sd} h_{rd}^*$.

4.4.3 Geometrical Representation and Bounds on the Distribution of the Hermitian Quadratic Form

The CCDF of $\tilde{Q}$ is expressed by a single integral in (4.12). Here, we give a geometrical presentation of the CCDF of $\tilde{Q}$, and use it to obtain some simple bounds on the CCDF of $\tilde{Q}$.

Based on (4.7), we rewrite the CCDF of $\tilde{Q}$ as

$$P(\tilde{Q} > \tilde{R}) = P\left(\frac{|\tilde{h}_1|^2}{\tilde{R}/\alpha} + \frac{|\tilde{h}_2|^2}{\tilde{R}/\beta} > 1\right). \tag{4.14}$$

From the above equation, we can see that in geometrical representation, $P(\tilde{Q} > \tilde{R})$ represents the probability of the two dimensional Rician vector $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside the quarter Ellipse-$\alpha\beta$ which is centered at the origin with semi-minor axis $\sqrt{\tilde{R}/\alpha}$ and semi-major axis $\sqrt{\tilde{R}/\beta}$, as depicted in Fig. 4.1. Thus, the CCDF of $\tilde{Q}$ can be expressed as

$$P(\tilde{Q} > \tilde{R}) = P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Ellipse-}\alpha\beta\right).$$

The analytical expression for the probability of the vector $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside the quarter Ellipse-$\alpha\beta$ is complicated. We will find some simple bounds on it by making use of its geometrical representation. As illustrated in Fig. 4.1, the quarter Circle-α and quarter Circle-β are the quarter circles centered at the origin, with radius $\sqrt{\tilde{R}/\alpha}$ and $\sqrt{\tilde{R}/\beta}$, respectively. It is obvious from Fig. 4.1 that

$$\text{quarter Circle-}\alpha \subseteq \text{quarter Ellipse-}\alpha\beta \subseteq \text{quarter Circle-}\beta.$$

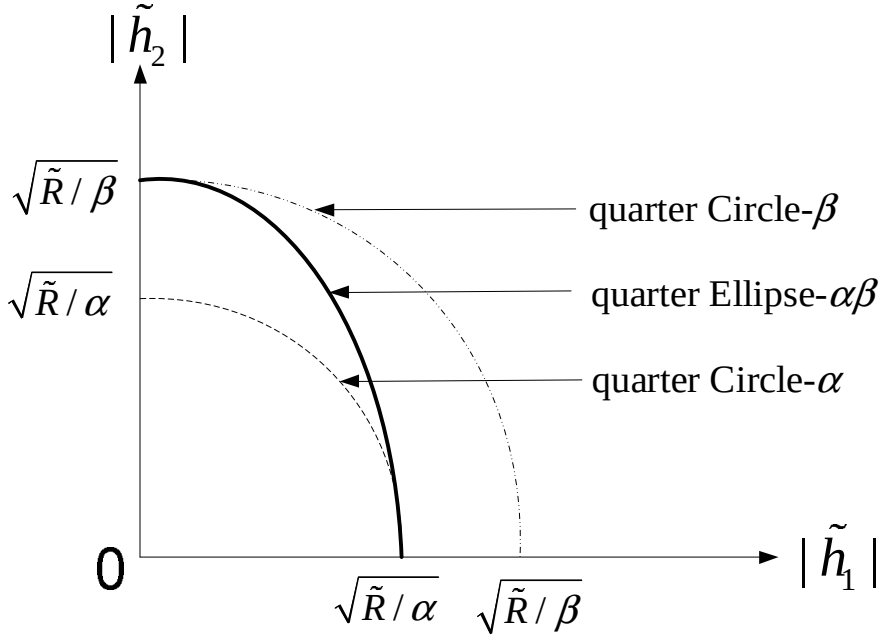

Figure 4.1: The geometrical representation of $P(\tilde{Q} > \tilde{R})$

Therefore, the probability of $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside of the quarter Ellipse-$\alpha\beta$ is less than that of $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside of the quarter Circle-α, and greater than that of $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside of the quarter Circle-β, i.e.,

$$P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Ellipse-}\alpha\beta\right) \leq P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Circle-}\alpha\right),$$
$$P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Ellipse-}\alpha\beta\right) \geq P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Circle-}\beta\right).$$

We can easily compute the probabilities of the vector $(|\tilde{h}_1|, |\tilde{h}_2|)^T$ lying outside the quarter circle-α and the quarter-circle-β, respectively, as

$$\begin{aligned} P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Circle-}\alpha\right) &= P(|\tilde{h}_1|^2 + |\tilde{h}_2|^2 > \tilde{R}/\alpha) \\ &= Q_2\left(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\alpha}\right), \\ P\left((|\tilde{h}_1|, |\tilde{h}_2|)^T \notin \text{quarter Circle-}\beta\right) &= P(|\tilde{h}_1|^2 + |\tilde{h}_2|^2 > \tilde{R}/\beta) \\ &= Q_2\left(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\beta}\right), \end{aligned}$$

since $(|\tilde{h}_1|^2 + |\tilde{h}_2|^2)$ is noncentral chi-square distributed with 4 degrees of freedom.

Therefore, the CCDF of $\tilde{Q}$ can be bounded as

$$P(\tilde{Q} > \tilde{R}) \leq Q_2\Big(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\alpha}\Big), \tag{4.15}$$

$$P(\tilde{Q} > \tilde{R}) \geq Q_2\Big(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\beta}\Big). \tag{4.16}$$

4.5 Outage Probability of the Information Rate

In this section, we derive the outage probability of the information rate, investigate the effects of the correlation coefficient ρ on the outage probability, and determine the optimal ρ which minimizes the outage probability by using numerical methods.

With (4.4), (4.5), (4.6), and Lemma 4.3, the outage probability of the information rate for any fixed $\rho \in [0, 1)$ is given by

$$P_{\text{out}}(\rho, R) = 1 - Q_1\left(\sqrt{2K_{sr}}, \sqrt{\frac{2\tilde{R}/\eta_{sr}}{1-|\rho|^2}}\right) Q_1\left(\sqrt{2}s_1, \sqrt{\frac{2\tilde{R}}{\alpha}}\right) - \frac{e^{-s_1^2}}{\alpha} Q_1\left(\sqrt{2K_{sr}}, \sqrt{\frac{2\tilde{R}/\eta_{sr}}{1-|\rho|^2}}\right) \int_0^{\tilde{R}} e^{-y/\alpha} I_0\Big(2s_1\sqrt{\frac{y}{\alpha}}\Big) Q_1\left(\sqrt{2}s_2, \sqrt{2\frac{\tilde{R}-y}{\beta}}\right) dy. \tag{4.17}$$

The outage probability in (4.17) is expressed as a single integral, which can be evaluated by standard numerical techniques. By making use of the upper and lower bounds on $P(\tilde{Q} > \tilde{R})$ in (4.15) and (4.16), the outage probability can be further bounded as

$$P_{\text{out}}(\rho, R) \geq 1 - Q_1\Big(\sqrt{2K_{sr}}, \sqrt{2\tilde{R}/[\eta_{sr}(1-|\rho|^2)]}\Big) Q_2\Big(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\alpha}\Big), \tag{4.18}$$

$$P_{\text{out}}(\rho, R) \leq 1 - Q_1\Big(\sqrt{2K_{sr}}, \sqrt{2\tilde{R}/[\eta_{sr}(1-|\rho|^2)]}\Big) Q_2\Big(\sqrt{2K_{srd}}, \sqrt{2\tilde{R}/\beta}\Big). \tag{4.19}$$

Remark 4.4. When $|\rho| = 1$, $\tilde{I}_{\text{df}}$ is zero and the outage probability is one. This can be interpreted as follows. The transmit signal, x_r, at the relay is chosen based on the prior received signals at the relay which are correlated with the prior transmitted signals

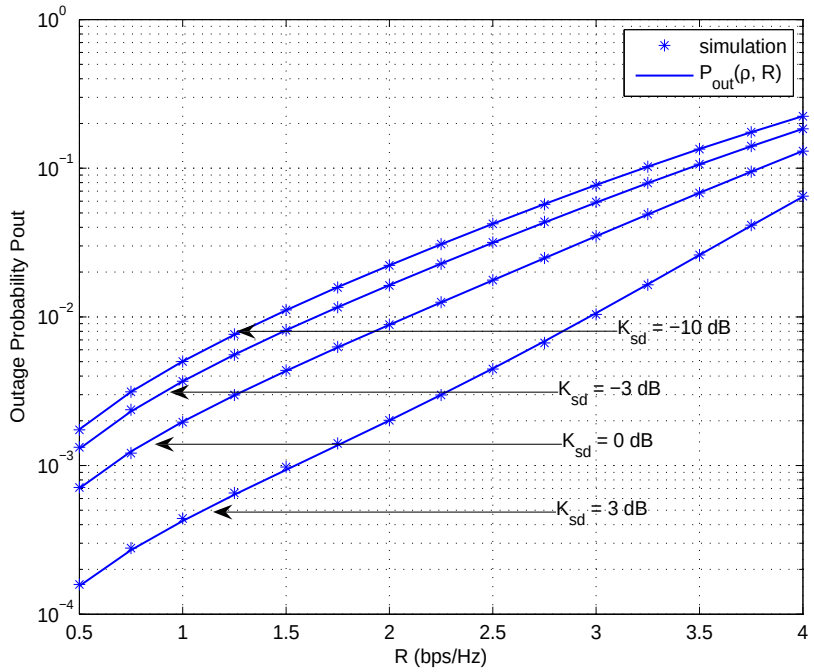

Figure 4.2: Outage probability versus R for the case that $\rho = 0$, $K_{sr} = K_{rd} = K_{sd} + 3$ (dB), and $K_{sd} = -10, -3, 0, 3$ (dB).

from the source x_s. This, together with the fact that x_s and x_r are fully correlated ($|\rho| = 1$), implies that x_s is static (x_s is a stochastic process), and hence no information is delivered from the source to the destination, i.e., $\tilde{I}_{\text{df}} = 0$.

Remark 4.5. In a Rayleigh fading environment, we have $K_{sr} = K_{sd} = K_{rd} = 0$. According to (4.13) and the fact that $Q_1(0, b) = \exp(-b^2/2)$, the outage probability (4.17) reduces to

$$P_{\text{out}}(\rho, R) = 1 - \frac{\alpha e^{-\tilde{R}/\alpha} - \beta e^{-\tilde{R}/beta}}{\alpha - \beta} \exp\left(\frac{-\tilde{R}/\eta_{sr}}{1 - |\rho|^2}\right). \tag{4.20}$$

This is consistent with the result presented in Chapter 3 (3.11) for the Rayleigh fading case.

Denote $\varepsilon_{ij}^2 = (1 + K_{ij})\sigma_{ij}^2$. Suppose that $\varepsilon_{sd}^2 = 1$ and ε_{ij}^2 is proportional to $1/d_{ij}^3$, where d_{ij} is the distance between nodes i and j. We consider the case in which

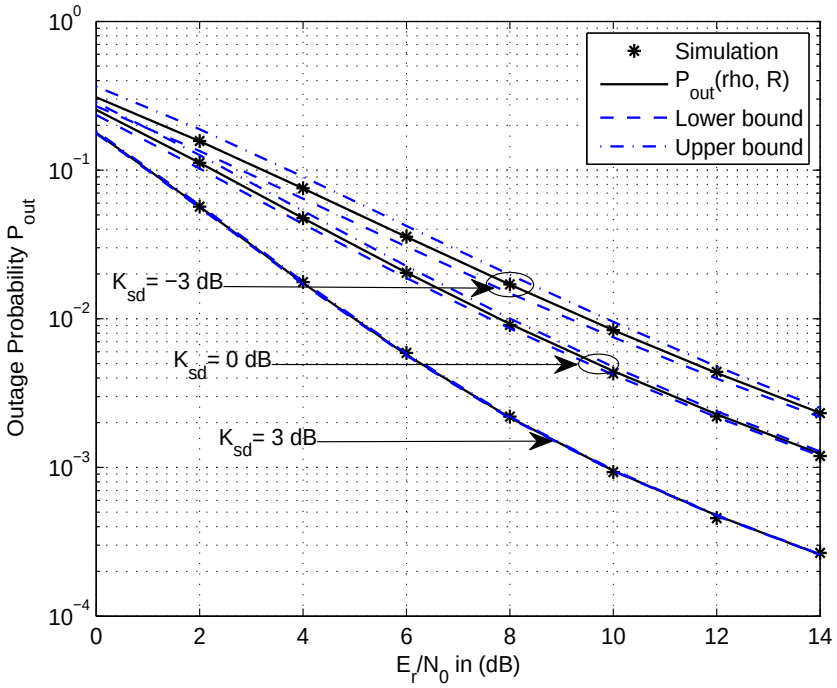

Figure 4.3: Outage probability versus SNR for the case that $R = 2$ bps/Hz, $E_s/N_0 = E_r/N_0 + 3$ (dB), $\rho = 0$, $K_{sr} = K_{rd} = K_{sd} + 3$ (dB), and $K_{sd} = -3, 0, 3$ (dB).

$d_{sr} = d_{sd}/3$ and $d_{rd} = 2d_{sd}/3$. In Fig. 4.2, we compare the analytical expression of the outage probability (4.17) with the ones obtained by Monte Carlo simulations for $E_s/N_0 = E_r/N_0 = 10$ dB. The Monte Carlo simulations are carried out by generating 8×10^5 realizations of h_{sr}, h_{sd} and h_{rd}, and by evaluating the statistics of $\tilde{I}_{\text{df}}$ in (4.2). Our analytical results match well with the Monte Carlo simulations.

Fig. 4.3 shows the outage probability and the upper and lower bounds given in (4.15) and (4.16) versus E_r/N_0 for a fixed rate $R = 2$ bps/Hz and $E_s/N_0 = E_r/N_0 + 3$ (dB). As can be easily seen from Fig. 4.3, the upper and lower bounds are quite tight in this case. Fig. 4.4 illustrates the effects of ρ and Rice factors on the outage probability. It can be observed from the figure that the outage probability decreases as the Rice factors increase. When the Rice factors are small, $\rho = 0$ yields the smallest outage

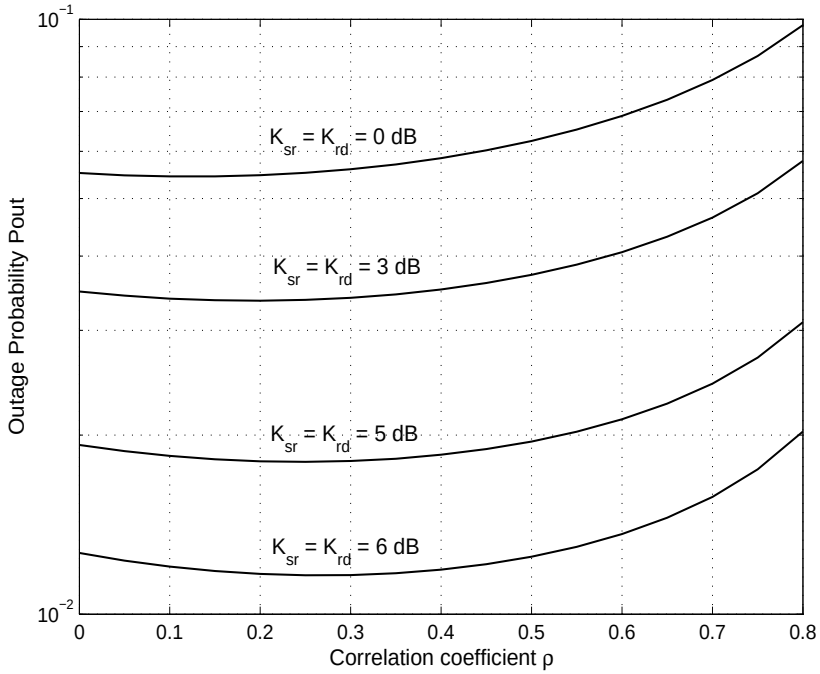

Figure 4.4: Outage probability versus ρ for the case that the phase of ρ is zero, $R = 3$ bps/Hz and $K_{sd} = 0$ dB.

probability. While for relatively large Rice factors, the value of ρ which minimizes the outage probability is not zero.

When the source and the relay know the means and variances of the three links, the outage-optimal transmit signaling in the sense of minimizing the outage probability, can be determined through choosing an optimal ρ as

$$\rho^{\text{opt}} = \arg \min_{0 \leq |\rho| \leq 1} P_{\text{out}}(\rho, R), \tag{4.21}$$

and the corresponding minimum outage probability is denoted by $P_{\text{out}}^{\text{opt}}(R)$. It appears intractable to analytically solve this optimization problem. Instead, we use the *Optimization Toolbox* in MATLAB (function: *fmincon*) to find ρ^{opt} numerically. In Fig. 4.5, we compare the outage probability for ρ^{opt} and $\rho = 0$. We observe that the optimal choice of ρ leads to a performance improvement, and it is determined by

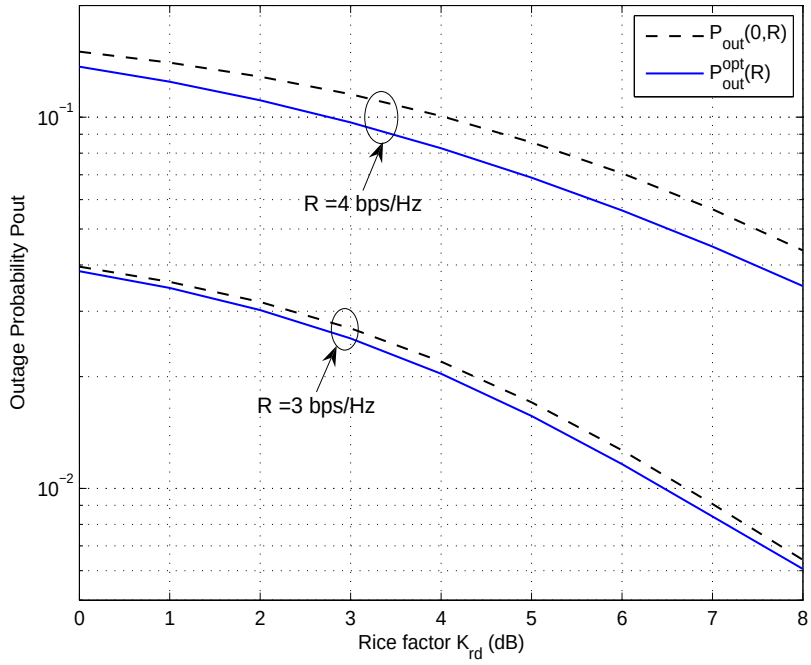

Figure 4.5: Outage probability versus K_{rd} when $K_{sr} = 6$ dB and $K_{sd} = 0$ dB.

several factors such as SNRs, rate threshold R, and the channel statistics.

4.6 Conclusion

In this chapter, we derived an exact expression of the outage probability of the full-duplex single relay DF system in a Rician fading environment. We also obtained an upper bound and a lower bound on the outage probability. These obtained bounds are expressed in simple closed forms, and hence can be easily evaluated. Based on the derived results, we determined the optimal correlation coefficient between the transmitted signals from the source and the relay by using numerical approaches, and revealed that the optimal correlation coefficient generally depends on SNRs, the rate threshold, the channel variances and the Rice factors.

Chapter 5

Differential Modulation for Decode-and-Forward Multiple Relay Systems

We consider differential modulation and demodulation in a multiple relay system where CSI is unknown to any node. Both the DF protocol and the SR protocol are analyzed in this chapter. For the DF protocol, this chapter first derives the ML detector and proposes a low complexity PL detector to closely approximate the nonlinear ML detector. Both the ML and PL detectors take the average BERs of all the source-relay transmissions into account. Then, the BER for the PL detector is analyzed. For a DF single relay system, the exact BER is obtained as a simple function of SNRs, the variances, and the fade rates of the transmission links. Based on the exact BER, a BER approximation at high SNR is derived, and it explicitly shows the diversity order and the different effects of the source-relay link and the relay-destination link on the end-to-end error performance. For a DF multiple relay system, a Chernoff upper bound on the BER and a high SNR approximation for the BER are obtained. For the SR protocol, some computational complexity is shifted from the destination to all the

relays. Each relay computes the instantaneous BERs of the source-relay transmission and use the instantaneous BERs to decide whether to transmit or remain silent. The destination performs simple MRC reception whose error performance is analyzed at high SNR. It shows from an error probability prospective that the SR protocol offers a space diversity of the number of all the potential cooperating nodes.

5.1 Introduction

In wireless communication systems, cooperative relay transmission is recognized as an effective technique to mitigate fading effects and enhance network coverage [44]. Most of the existing relay transmission schemes assume that the receivers have perfect CSI, and utilize the CSI for coherent detection [35, 37, 56, 58]. However, it is rather costly or even infeasible to obtain accurate channel estimates of multiple transmission links especially in rapid fading environments.

To obviate the need for channel estimation for wireless relay systems, non-coherent or differential modulations are proposed. General non-coherent ML detectors for AF and DF relay systems were derived in [67]. For AF relay systems with differential modulations, some simple receiver structures along with the error probability analysis had been obtained in [69, 70]. However in DF relay systems, due to the nonlinear processing at the relay nodes, the ML receiver is quite complex. For a non-coherent DF relay system with BFSK modulation, [36] derived the nonlinear ML detector and proposed a PL detector which has similar performance to the ML detector. In addition, the exact BER of the PL detector for single relay systems and the diversity order for multiple relay systems were also derived in [36]. For DF relay systems with differential modulations, the non-linear ML detector was derived in [71]. However, the error probability analysis of the ML detector seems intractable. The paper [73]

proposed a reduced complexity equal-gain combiner, and obtained its error probability. Recently, a ML detector and a PL detector were derived in [72] for the DF single relay system using DBPSK modulation. The paper [72] also computed the BER of the PL detector in terms of several infinite series. All the aforementioned studies assume that the fading coefficients remain unchanged in two consecutive symbol intervals. This assumption may not hold in rapid fading environments.

Since relay nodes have some possibilities of making decoding errors, the SR protocol was proposed to reduce error propagations at the relay nodes. From the outage probability perspective, it has been shown in [34, 44] that SR protocols can offer full space diversity. In [34, 44], the CSI of the source-relay link is used as the indicator of the reliability of the relay detection. The error probability of two-user cooperation working under the SR protocol for DBPSK modulation was analyzed in [50] where the relay node uses cyclic redundance check (CRC) after differential decoding to decide whether the relay makes correct detection.

In this chapter, we focus on a multiple relay system using DBPSK modulation. The fading processes assumed here have an arbitrary Doppler spectrum with an arbitrary Doppler bandwidth. We consider two protocols at the relay node: the DF protocol and the SR protocol. For the DF protocol, we derive the nonlinear ML detector and propose the PL detector which has similar performance to the nonlinear ML detector. Both the ML and PL detectors take the average BERs of all the source-relay transmissions into account, and the received signals at the destination from the source and all the relays are combined with different weights, because the transmission links in the relaying system have different statistics. Then, we derive the BER of the PL detector, whereas the BER analysis for the nonlinear ML detector appears intractable. For a DF single relay system, we obtain the exact BER and its high SNR approximation. The exact BER is a simple function of the SNRs, the variances

and fade rates of all the transmission links. The BER approximation at high SNR shows explicitly the diversity order and the different roles of the source-relay link and the relay-destination link in determining the end-to-end error performance. Moreover, a Chernoff upper bound on the BER and a high SNR approximation for the BER are obtained for a DF multiple relay system. For the SR protocol, each relay computes the instantaneous BERs of the source-relay transmission and use the instantaneous BERs to decide whether to transmit or remain silent. When the instantaneous BERs at a relay satisfy certain criteria, the relay belongs to the coding set which is a random set and a subset of the set containing all the potential relays. In this case, the ML receiver at the destination is an MRC receiver. We analyze the error performance of the multiple relay system working with the SR protocol at high SNR, and show from an error probability prospective that the SR protocol offers full space diversity. More specifically, the SR protocol offers a space diversity order of the number of all the potential cooperating nodes, not just the number of nodes that are transmitting to the destination. The analytical results match with simulations. Given the same total power constraint, the relay system with DBPSK modulation outperforms a non-cooperative system which only has a direct link and used also DBPSK modulation and ML detection.

5.2 Multiple-Relay System Model

We consider a relay network as illustrated in Fig. 5.1, which consists of a source node s, a destination node d, and L relay nodes $\{r\}_{r=1}^{L}$. In this model, the relays facilitate the ultimate transmission from the source to the destination by cooperating with the source. All the relays work in half-duplex mode in which they cannot transmit and receive at the same time on the same frequency band. We assume that there are in total $(L+1)$ orthogonal channels available in the network, and they can be realized in time

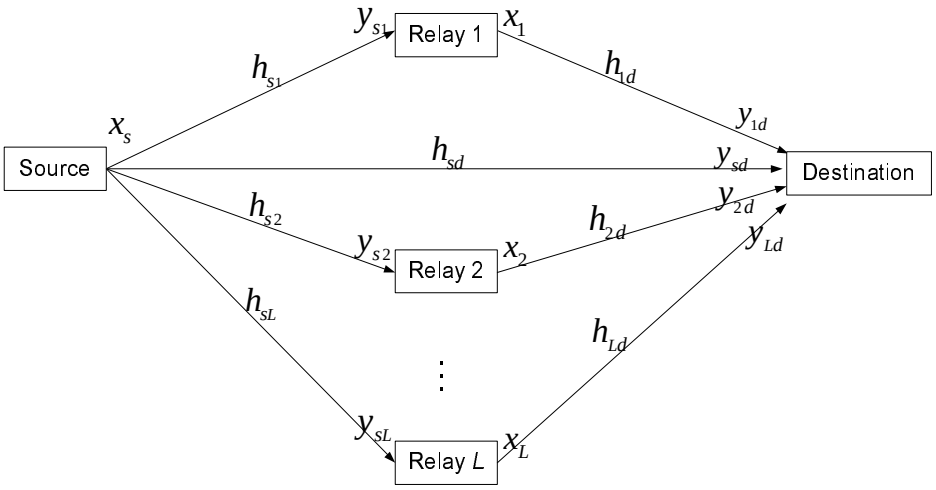

Figure 5.1: The relay system with L relays.

division, frequency division, or code division. The source broadcasts its messages to all the relays and the destination in one channel, and the relays transmit to the destination in the remaining L orthogonal channels.

In the differential transmission scheme, the information $\tilde{s}$ is differentially encoded in the transmit signal x_s. By denoting the time index as k, the k-th symbol transmitted from the source is given by

$$x_s[k] = x_s[k-1]\tilde{s}[k], \quad \tilde{s}[k] \in M_s, \tag{5.1}$$

where $M_s = \{+1, -1\}$. We assume that $\tilde{s}[k]$ has equal probabilities of being $+1$ and -1, and there are N_f information symbols in one frame. Without loss of generality, we assume that $x_s[0] = 1$ is the initial reference signal.

The source transmits the signal x_s with energy E_s to all the relays and the destination. The corresponding received signals at the relays and the destination in the kth symbol interval are

$$\begin{aligned} y_{sr}[k] &= \sqrt{E_s} h_{sr}[k] x_s[k] + n_{sr}[k], \\ y_{sd}[k] &= \sqrt{E_s} h_{sd}[k] x_s[k] + n_{sd}[k], \end{aligned} \tag{5.2}$$

where $r = 1, 2, \cdots, L$. The r-th relay transmits $x_r[k]$ to the destination with energy E_r. The received signal at the destination corresponding to the r-th relay transmission is given by

$$y_{rd}[k] = \sqrt{E_r} h_{rd}[k] x_r[k] + n_{rd}[k]. \tag{5.3}$$

Each $h_{ij}[k]$, $ij \in \{sd, sr, rd\}_{r=1}^{L}$, as illustrated in Fig. 5.1, denotes the fading coefficient between nodes i and j at the k-th symbol duration. The fading coefficients take account of path loss, shadowing and frequency nonselective fading as in [44]. Each sequence, $\{h_{ij}[k]\}_k$, consists of samples of a zero-mean, complex, Gaussian stochastic process with autocorrelation function $\mathbb{E}\big[h_{ij}[k]h_{ij}^*[k-l]\big] = \sigma_{ij}^2 R_{ij}(l)$, where σ_{ij}^2 is the average power of the fading process, and $R_{ij}(l)$ is the normalized auto-correlation function evaluated at a time difference l. We denote the correlation coefficient of the fading gains between two adjacent symbols by $\rho_{ij} = R_{ij}(1)$, and hence ρ_{ij} is a measure of the fluctuation rate of the channel fading process. We assume that the channel fading processes in different transmission links are independent, i.e., $\forall i \neq m$ or $j \neq n$, we have $\mathbb{E}\big[h_{ij}[k]h_{mn}^*[k]\big] = 0$. The sequence $\{n_{ij}[k]\}_k$ is complex AWGN with mean zero and variance N_0, i.e., $\mathbb{E}\big[|n_{ij}[k]|^2\big] = N_0$. The AWGNs are mutually independent and are independent of all channel fading processes. The transmit signals are subject to power constraints $\mathbb{E}[|x_s|^2] = 1$ and $\mathbb{E}[|x_r|^2] = 1$ with $\mathbb{E}[\cdot]$ denoting the expectation operator.

5.3 Differential Detection and Error Probability at Relay Nodes

5.3.1 Differential Detection at the r-th Relay

At the r-th relay, we will use the two adjacent received signals, $y_{sr}[k]$ and $y_{sr}[k-1]$, to make differential detection. We write the two adjacent received signals as a vector $\boldsymbol{y}_{sr} = [y_{sr}[k] \;\; y_{sr}[k-1]]^T$. Conditioned on $\tilde{s}[k]$, $y_{sr}[k]$ and $y_{sr}[k-1]$ are jointly complex Gaussian distributed with mean zero and covariance matrix $\boldsymbol{\Sigma}_{sr}$ given as

$$\boldsymbol{\Sigma}_{sr} = N_0 \begin{bmatrix} \eta_{sr}+1 & \eta_{sr}\rho_{sr}\tilde{s}[k] \\ \eta_{sr}\rho_{sr}\tilde{s}^*[k] & \eta_{sr}+1 \end{bmatrix}, \tag{5.4}$$

where $\eta_{sr} = E_s\sigma_{sr}^2/N_0$ represents the average received SNR per bit on the link from the source to the r-th relay. The joint PDF of $y_{sr}[k]$ and $y_{sr}[k-1]$ conditioned on $\tilde{s}[k]$ is

$$p\big(y_{sr}[k], y_{sr}[k-1]\,|\tilde{s}[k]\big) = p\big(\boldsymbol{y}_{sr}\,|\tilde{s}[k]\big) = \frac{1}{\pi^2|\boldsymbol{\Sigma}_{sr}|}\exp(-\boldsymbol{y}_{sr}^H\boldsymbol{\Sigma}_{sr}^{-1}\boldsymbol{y}_{sr}).$$

Taking natural logarithm of the joint PDF $p\big(y_{sr}[k], y_{sr}[k-1]\,|\tilde{s}[k]\big)$, and ignoring the terms that are independent of $\tilde{s}[k]$, we can obtain the differential ML detector at the r-th relay node as

$$\tilde{s}_r[k] = \text{sign}\left(\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big)\right) = \text{sign}\big(\boldsymbol{y}_{sr}^H\boldsymbol{B}\boldsymbol{y}_{sr}\big), \tag{5.5}$$

where $\text{sign}(\cdot)$ is the sign function and $\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

5.3.2 Average BER at the r-th Relay

Define $\bar{\epsilon}_r$ as the average BER of detecting $\tilde{s}[k]$ at the r-th relay. The average BER at the r-th relay is equal to the average BER of a SISO system, and it is given by [98]

$$\bar{\epsilon}_r = P\left(\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big) < 0\,|\,\tilde{s}[k]=+1\right) = \frac{1}{2(\eta_{sr}+1)} + \frac{(1-\rho_{sr})\eta_{sr}}{2(\eta_{sr}+1)}. \tag{5.6}$$

5.3.3 Instantaneous BER at the r-th Relay

Although the CSI of the source-relay link is not available at the relay, the relay knows its past received signal and can use it to compute instantaneous BERs.

Conditioned on $\tilde{s}[k]=+1$ and $y_{sr}[k-1]$, the PDF of $y_{sr}[k]$ can be computed as

$$p\big(y_{sr}[k]\,\big|\tilde{s}[k]=+1, y_{sr}[k-1]\big) = \frac{p\big(y_{sr}[k], y_{sr}[k-1]\,|\tilde{s}[k]=+1\big)}{p\big(y_{sr}[k-1]\big)},$$

and hence we can see that conditioned on $\tilde{s}[k]=+1$ and $y_{sr}[k-1]$, the quantity $y_{sr}[k]$ is complex Gaussian distributed with mean $y_{sr}[k-1]\eta_{sr}\rho_{sr}/(\eta_{sr}+1)$ and variance $N_0[(\eta_{sr}+1)^2-\eta_{sr}^2\rho_{sr}^2]/(\eta_{sr}+1)$. Therefore, $\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big)$ is conditionally complex Gaussian distributed with mean

$$\mathbb{E}\Big(\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big)\,\Big|\tilde{s}[k]=+1, y_{sr}[k-1]\Big) = \frac{\eta_{sr}\rho_{sr}}{(\eta_{sr}+1)}\,|y_{sr}[k-1]|^2$$

and variance

$$\mathrm{Var}\Big(\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big)\,\Big|\tilde{s}[k]=+1, y_{sr}[k-1]\Big) = \frac{N_0[(\eta_{sr}+1)^2-\eta_{sr}^2\rho_{sr}^2]}{2(\eta_{sr}+1)}\,|y_{sr}[k-1]|^2\,.$$

At the r-th relay node, the instantaneous BER of detecting $\tilde{s}[k]$ conditioned on $y_{sr}[k-1]$ can be computed as [98, Appendix]

$$\begin{aligned}\epsilon_r[k] &= P\Big(\Re\big(y_{sr}^*[k]y_{sr}[k-1]\big) < 0\,\Big|\tilde{s}[k]=+1, y_{sr}[k-1]\Big)\\ &= \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\eta_{sr}^2\rho_{sr}^2\,|y_{sr}[k-1]|^2}{N_0(\eta_{sr}+1)[(\eta_{sr}+1)^2-\eta_{sr}^2\rho_{sr}^2]}}\right)\\ &< \frac{1}{2}\exp\left(-\frac{\eta_{sr}^2\rho_{sr}^2\,|y_{sr}[k-1]|^2}{N_0(\eta_{sr}+1)[(\eta_{sr}+1)^2-\eta_{sr}^2\rho_{sr}^2]}\right),\end{aligned} \tag{5.7}$$

where the last inequality is obtained by using $\mathrm{erfc}(x) < \exp(-x^2)$, and $\mathrm{erfc}(\cdot)$ is the complementary error function, i.e., $\mathrm{erfc}(x) = 2\pi^{-1/2}\int_x^\infty e^{-t^2}dt$. The quantity $\epsilon_r[k]$ is the instantaneous BER at the r-th relay during the k-th symbol interval.

For a given average received SNR η_{sr}, noise power N_0 and fade rate ρ_{sr}, the bit errors between adjacent symbol intervals are correlated for DBPSK modulation, and thus the instantaneous BERs, $\{\epsilon_r[k]\}_{k=1}^{N_f}$, are also correlated.

5.4 Relaying Protocols

At the relay nodes, various protocols can be employed. We will consider the DF protocol and the SR protocol, respectively.

5.4.1 DF Protocol

In the DF protocol, each relay differentially demodulates its received signal, and generates an estimate $\tilde{s}_r[k]$ of $\tilde{s}[k]$ according to (5.5). Then, the r-th relay differentially remodulates $\tilde{s}_r[k]$ into $x_r[k]$ as

$$x_r[k] = x_r[k-1]\tilde{s}_r[k], \tag{5.8}$$

with $x_r[0] = 1$ as the initial reference signal, and transmits $x_r[k]$ to the destination with energy E_r.

5.4.2 SR Protocol

When some relay nodes make errors in decoding the source message, the signals transmitted from these relay nodes become interferences at the destination, and result in performance degradation. Therefore, when a relay has a large error probability in decoding the source message, this relay should remain silent instead of transmitting to the destination.

We assume that the r-th relay knows its average received SNR η_{sr}, the noise power N_0, and the fade rate ρ_{sr}. During the k-th symbol interval, since $y_{sr}[k-1]$ is available at the r-th relay node, the instantaneous BER, $\epsilon_r[k]$, can be computed at the r-th relay node according to (5.7). We can use the instantaneous BERs $\{\epsilon_r[k]\}_k$ as a selection criterion at the r-th relay to decide whether to transmit or remain silent.

We define the decoding set $\mathcal{D}$ to be the set containing the relay nodes at which

every instantaneous BER in a whole frame is smaller than a threshold ϵ, i.e.,

$$r \in \mathcal{D} \quad \Longleftrightarrow \quad \epsilon_r[k] < \epsilon, \ \forall k = 1, 2, \cdots, N_f. \tag{5.9}$$

In one frame, if there exists an instantaneous BER greater than the threshold ϵ at the r-th relay node, then the r-th relay node does not belong to the decoding set $\mathcal{D}$ in this frame.

If the r-th relay node belongs to the decoding set $\mathcal{D}$, then the r-th relay transmits to the destination using the DF protocol for the whole frame; if the r-th relay node does not belong to the decoding set $\mathcal{D}$, then the r-th relay remains silent for the whole frame, i.e.,

$$x_r[k] = \begin{cases} x_r[k-1]\tilde{s}_r[k], & r \in \mathcal{D} \\ 0, & r \notin \mathcal{D} \end{cases}, \qquad \forall k = 1, 2, \cdots, N_f. \tag{5.10}$$

The decoding set $\mathcal{D}$ is a subset of the set which contains all the relay nodes, and $\mathcal{D}$ is a random set.

5.5 Differential Detection at the Destination

5.5.1 Differential Detection at the Destination for a DF Relay System

We express the received signals at the destination from the source over two adjacent intervals as a vector $\boldsymbol{y}_{sd} = [y_{sd}[k] \ \ y_{sd}[k-1]]^T$, and the received signals at the destination from the r-th relay over two adjacent intervals as a vector $\boldsymbol{y}_{rd} = [y_{rd}[k] \ \ y_{rd}[k-1]]^T$, $r = 1, 2, \cdots, L$. We further express all the received signals at the destination over two adjacent intervals as a $2(L+1) \times 1$ vector $\boldsymbol{y}_d^{\mathrm{DF}} = \left[\boldsymbol{y}_{sd}^T \ \ \boldsymbol{y}_{1d}^T \ \ \boldsymbol{y}_{2d}^T \ \ \cdots \ \ \boldsymbol{y}_{Ld}^T\right]^T$. The detection of $\tilde{s}[k]$ at the destination is made based on

$\boldsymbol{y}_d^{\text{DF}}$. Given $\tilde{s}[k]$, the PDF of $\boldsymbol{y}_d^{\text{DF}}$ is

$$p(\boldsymbol{y}_d^{\text{DF}}|\tilde{s}[k]) = \sum_{\tilde{s}_1[k],\cdots,\tilde{s}_L[k]} p(\boldsymbol{y}_d|\tilde{s}[k],\tilde{s}_1[k],\tilde{s}_2[k],\cdots,\tilde{s}_L[k]) \prod_{r=1}^{L} p(\tilde{s}_r[k]|\tilde{s}[k]). \quad (5.11)$$

Conditioned on $\tilde{s}[k]$ and $\{\tilde{s}_r[k]\}_{r=1}^L$, the vectors, $\boldsymbol{y}_{sd}$ and $\{\boldsymbol{y}_{rd}\}_{r=1}^L$, are independent because of the mutual independence of the fading coefficients of different transmission links. Thus, the conditional PDF of $\boldsymbol{y}_d^{\text{DF}}$ given $\tilde{s}[k]$ can be rewritten as

$$p(\boldsymbol{y}_d^{\text{DF}}|\tilde{s}[k]) = p(\boldsymbol{y}_{sd}|\tilde{s}[k]) \prod_{r=1}^{L} \sum_{\tilde{s}_r[k]\in M_s} p(\boldsymbol{y}_{rd}|\tilde{s}_r[k]) p(\tilde{s}_r[k]|\tilde{s}[k]). \quad (5.12)$$

It is obvious that conditioned on $\tilde{s}[k]$ and $\{\tilde{s}_r[k]\}_{r=1}^L$, the vectors, $\boldsymbol{y}_{sd}$ and $\{\boldsymbol{y}_{rd}\}_{r=1}^L$, are zero-mean, complex, Gaussian distributed with covariance matrices $\boldsymbol{\Sigma}_{sd}$ and $\{\boldsymbol{\Sigma}_{rd}\}_{r=1}^L$, respectively, i.e.,

$$\boldsymbol{\Sigma}_{sd} = N_0 \begin{bmatrix} \eta_{sd}+1 & \eta_{sd}\rho_{sd}\tilde{s}[k] \\ \eta_{sd}\rho_{sd}\tilde{s}^*[k] & \eta_{sd}+1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{rd} = N_0 \begin{bmatrix} \eta_{rd}+1 & \eta_{rd}\rho_{rd}\tilde{s}_r[k] \\ \eta_{rd}\rho_{rd}\tilde{s}_r^*[k] & \eta_{rd}+1 \end{bmatrix}, \quad (5.13)$$

where $\eta_{sd} = E_s\sigma_{sd}^2/N_0$ and $\eta_{rd} = E_r\sigma_{rd}^2/N_0$ represent the average received SNRs per bit in the corresponding links. Taking the log-likelihood ratio of $\boldsymbol{y}_d^{\text{DF}}$ conditioned on each value of $\tilde{s}[k]$, substituting the expressions for $y_{sd}[k]$ in (5.2) and $y_{rd}[k]$ in (5.3) into $p(\boldsymbol{y}_d^{\text{DF}}|\tilde{s}[k])$, and ignoring the terms that are independent of $\tilde{s}[k]$, we obtain the ML detector for $\tilde{s}[k]$ at the destination as

$$\tilde{s}_d^{\text{DF-ML}}[k] = \text{sign}\Big(t_s + \sum_{r=1}^{L} f_{\text{ML}}(\bar{\epsilon}_r, t_r)\Big), \quad (5.14)$$

where $\bar{\epsilon}_r$ is given in (5.6) as the average BER at the r-th relay node, the function $f_{\text{ML}}(\bar{\epsilon}_r, t_r)$ is given by

$$f_{\text{ML}}(\bar{\epsilon}_r, t_r) = \ln \frac{(1-\bar{\epsilon}_r)\exp(t_r) + \bar{\epsilon}_r}{(1-\bar{\epsilon}_r) + \bar{\epsilon}_r \exp(t_r)}, \qquad \bar{\epsilon}_r \in (0, 1/2), t_r \in (-\infty, \infty), \quad (5.15)$$

and

$$
\begin{aligned}
t_s &= 2\frac{w_{sd}}{N_0}\boldsymbol{y}_{sd}^H\boldsymbol{B}\boldsymbol{y}_{sd} = 2\frac{w_{sd}}{N_0}\Re\big(y_{sd}[k]^* y_{sd}[k-1]\big),\\
t_r &= 2\frac{w_{rd}}{N_0}\boldsymbol{y}_{rd}^H\boldsymbol{B}\boldsymbol{y}_{rd} = 2\frac{w_{rd}}{N_0}\Re\big(y_{rd}[k]^* y_{rd}[k-1]\big),\\
w_{ij} &= \frac{\eta_{ij}\rho_{ij}}{(\eta_{ij}+1)^2-\eta_{ij}^2\rho_{ij}^2},\ \ ij \in \{sd, rd\}_{r=1}^{L}.
\end{aligned} \tag{5.16}
$$

The w_{ij} is the weight for the output of each transmission link according to its channel statistical information. It is easy to see that $f_{\text{ML}}(\bar{\epsilon}_r, t_r)$ is an increasing function of t_r for any fixed $\bar{\epsilon}_r$, and that $f_{\text{ML}}(\bar{\epsilon}_r, t_r) \in (-T_r, T_r)$, where

$$
T_r = \ln[(1-\bar{\epsilon}_r)/\bar{\epsilon}_r] > 0. \tag{5.17}
$$

However, the nonlinearity of $f_{\text{ML}}(\bar{\epsilon}_r, t_r)$ makes the ML detector complicated and the error probability analysis intractable. It was shown in [36] that $f_{\text{ML}}(\bar{\epsilon}_r, t_r)$ can be well approximated by a PL function $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$, i.e.,

$$
f_{\text{ML}}(\bar{\epsilon}_r, t_r) \approx f_{\text{PL}}(\bar{\epsilon}_r, t_r) = \begin{cases} -T_r, & t_r \le -T_r \\ t_r, & -T_r \le t_r \le T_r \\ T_r, & t_r \ge T_r \end{cases}. \tag{5.18}
$$

This leads to the following PL detector at the destination:

$$
\tilde{s}_d^{\text{DF-PL}}[k] = \text{sign}\Big(t_s + \sum_{r=1}^{L} f_{\text{PL}}(\bar{\epsilon}_r, t_r)\Big). \tag{5.19}
$$

Furthermore, as we will see later, the performance of this PL detector provides a very tight upper bound on the error probability of the ML detector.

Remark 5.1. For a DF single relay system, i.e., $L = 1$, when the source-destination and relay-destination links have the same statistics and the fading coefficients remain constant over two consecutive intervals, i.e., $w_{sd} = w_{rd}$ and $\rho_{ij} = 1, \forall ij \in \{sd, sr, rd\}$, the ML detector (5.14) and the PL detector (5.19) reduce to the ML detector [72, Eq. (21)] and the PL detector [72, Eq. (26)], respectively.

5.5.2 Differential Detection at the Destination for a SR Relay System

In the SR protocol, only the relay nodes which belong to the decoding set $\mathcal{D}$ will transmit to the destination. Conditioned on $\mathcal{D}$, the detection of $\tilde{s}[k]$ at the destination is made based on $\boldsymbol{y}_{sd}$ and $\{\boldsymbol{y}_{rd}\}_{r\in\mathcal{D}}$. The received signals at the destination over two adjacent intervals can be written as a $2(|\mathcal{D}|+1)\times 1$ vector $\boldsymbol{y}_d^{\text{SR}} = [\boldsymbol{y}_{sd}^T,\ \{\boldsymbol{y}_{rd}^T\}_{r\in\mathcal{D}}]$ with $|\mathcal{D}|$ representing the cardinality of the set $\mathcal{D}$. Similar to the case of DF in the last subsection, we can obtain the PDF of $\boldsymbol{y}_d^{\text{SR}}$ conditioned on $\tilde{s}[k]$ and $\mathcal{D}$ as

$$p(\boldsymbol{y}_d^{\text{SR}} \mid \tilde{s}[k], \mathcal{D}) = p(\boldsymbol{y}_{sd}|\tilde{s}[k]) \prod_{r\in\mathcal{D}} \sum_{\tilde{s}_r[k]\in M_s} p(\boldsymbol{y}_{rd}|\tilde{s}_r[k]) p(\tilde{s}_r[k]|\tilde{s}[k]). \tag{5.20}$$

When the r-th relay belongs to the decoding set $\mathcal{D}$, we have $p(\tilde{s}_r[k] \neq \tilde{s}[k]|\tilde{s}[k]) < \epsilon$. Usually, the threshold ϵ is very small to make sure the relays that belong to the decoding set have very small decoding error probabilities. Thus, we can have $p(\tilde{s}_r[k] \neq \tilde{s}[k]|\tilde{s}[k]) \approx 0$ and $p(\tilde{s}_r[k] = \tilde{s}[k]|\tilde{s}[k]) \approx 1$. Therefore, the conditional PDF of $\boldsymbol{y}_d^{\text{SR}}$ in the above equation reduces to

$$p(\boldsymbol{y}_d^{\text{SR}} \mid \tilde{s}[k], \mathcal{D}) = p(\boldsymbol{y}_{sd}|\tilde{s}[k]) \prod_{r\in\mathcal{D}} p(\boldsymbol{y}_{rd}|\tilde{s}_r[k] = \tilde{s}[k]). \tag{5.21}$$

Taking the log-likelihood ratio of $\boldsymbol{y}_d^{\text{SR}}$ conditioned on $\mathcal{D}$ and each value of $\tilde{s}[k]$, and ignoring the terms that are independent of $\mathcal{D}$ and $\tilde{s}[k]$, we obtain the ML detector for $\tilde{s}[k]$ at the destination conditioned on $\mathcal{D}$ as

$$\tilde{s}_d^{\text{SR}}[k] = \text{sign}\Big(t_s + \sum_{r\in\mathcal{D}} t_r\Big), \tag{5.22}$$

where t_s and t_r are given in (5.16).

Remark 5.2. For the SR protocol, we can see from (5.22) that the ML receiver at the destination is an MRC receiver. Compared to the ML and PL receivers for the DF protocol, the MRC receiver for the SR protocol is relatively simple, because some

computational complexity is shifted from the destination to the relay nodes which need to compute their instantaneous BERs and make decisions on whether to transmit or remain silent for the SR protocol.

Remark 5.3. For the DF protocol, the destination needs to know the average BER for all the relay nodes, i.e., $\{\bar{\epsilon}_r\}_{r=1}^{L}$. However, for the SR protocol, the destination only needs to know the decoding set $\mathcal{D}$ for each frame.

5.5.3 Statistics of the Destination Decision Metrics

From (5.14), (5.19) and (5.22), we can see that the destination decision metrics are functions of t_s and t_r. To evaluate the error performance of the detectors derived in the previous two subsections, we need to know the statistics of t_s, $\{t_r\}_{r=1}^{L}$, $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$, and $(t_s + \sum_{r\in\mathcal{D}} t_r)$. It can be seen from (5.16) that t_s and t_r involve Hermitian quadratic forms $\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij}$, $ij \in \{sd, rd\}_{r=1}^{L}$, of zero-mean, complex, Gaussian vectors. Hence, we will first study the statistics of the quadratic form $\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij}$. Based on this, we will derive the PDF of t_s, t_r, the CDF of $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$, and the CDF of $(t_s + \sum_{r\in\mathcal{D}} t_r)$. These results will be used later to evaluate the error performance of the relay system working under the DF protocol or the SR protocol.

With the noise whitening and eigne-decomposition techniques [93, Section 3.1a.1], we can express the quadratic form $\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij}$ as

$$\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij} = \begin{cases} N_0(\alpha_{ij}|u_{ij}|^2 - \beta_{ij}|v_{ij}|^2), & \tilde{s}_i[k] = +1 \\ N_0(\beta_{ij}|\tilde{u}_{ij}|^2 - \alpha_{ij}|\tilde{v}_{ij}|^2), & \tilde{s}_i[k] = -1 \end{cases}, \tag{5.23}$$

where α_{ij} and β_{ij} are given as

$$\alpha_{ij} = 1 + \eta_{ij} + \eta_{ij}\rho_{ij}, \qquad \beta_{ij} = 1 + \eta_{ij} - \eta_{ij}\rho_{ij}, \tag{5.24}$$

while u_{ij} and v_{ij} are i.i.d. complex Gaussian RVs with zero mean and unit variance; $\tilde{u}_{ij}$ and $\tilde{v}_{ij}$ are also i.i.d. complex Gaussian RVs with zero mean and unit variance. We

define $z_{ij} = \alpha_{ij}|u_{ij}|^2 - \beta_{ij}|v_{ij}|^2$ and $\tilde{z}_{ij} = \beta_{ij}|\tilde{u}_{ij}|^2 - \alpha_{ij}|\tilde{v}_{ij}|^2$. Notice that both z_{ij} and $\tilde{z}_{ij}$ are the weighted differences between two independent exponential RVs. The PDF and CDF of z_{ij} can be obtained respectively as

$$f_{z_{ij}}(z) = \begin{cases} \frac{1}{\alpha_{ij}+\beta_{ij}} e^{z/\beta_{ij}}, & z \leq 0 \\ \frac{1}{\alpha_{ij}+\beta_{ij}} e^{-z/\alpha_{ij}}, & z \geq 0 \end{cases}, \tag{5.25}$$

$$P(z_{ij} \leq z) = \begin{cases} \frac{\beta_{ij}}{\alpha_{ij}+\beta_{ij}} e^{z/\beta_{ij}}, & z \leq 0 \\ 1 - \frac{\alpha_{ij}}{\alpha_{ij}+\beta_{ij}} e^{-z/\alpha_{ij}}, & z \geq 0 \end{cases}. \tag{5.26}$$

The PDF and CDF of $\tilde{z}_{ij}$ are the same as the ones for z_{ij} except for interchanging α_{ij} with β_{ij}. With the statistics of z_{ij} and $\tilde{z}_{ij}$, the statistics of $\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij}$ are obvious since $\boldsymbol{y}_{ij}^H \boldsymbol{B} \boldsymbol{y}_{ij}$ equals either $N_0 z_{ij}$ or $N_0 \tilde{z}_{ij}$.

From (5.16), we have $t_s = 2w_{sd}z_{sd}$ conditioned on $\tilde{s}[k] = +1$. Therefore, the PDF of t_s conditioned on $\tilde{s}[k] = +1$ is given by

$$f_{t_s|\tilde{s}[k]=+1}(t) = \begin{cases} \frac{1}{2w_{sd}(\alpha_{sd}+\beta_{sd})} \exp(\frac{t}{2w_{sd}\beta_{sd}}), & t \leq 0 \\ \frac{1}{2w_{sd}(\alpha_{sd}+\beta_{sd})} \exp(\frac{-t}{2w_{sd}\alpha_{sd}}), & t \geq 0 \end{cases}. \tag{5.27}$$

We also have $t_r = 2w_{rd}z_{rd}$ if $\tilde{s}_r[k] = +1$, and $t_r = 2w_{rd}\tilde{z}_{rd}$ if $\tilde{s}_r[k] = -1$. With the PL function $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$ in (5.18) and the CDF of z_{ij}, we can obtain the CDF of $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$ conditioned on $\bar{\epsilon}_r$ and $\tilde{s}_r[k] = +1$ as

$$P\big(f_{\text{PL}}(\bar{\epsilon}_r, t_r) \leq t \,|\, \bar{\epsilon}_r, \tilde{s}_r[k]=+1\big) = \begin{cases} 0, & t \in [-\infty, -T_r) \\ \frac{\beta_{rd}}{\alpha_{rd}+\beta_{rd}} e^{\frac{t}{2w_{rd}\beta_{rd}}}, & t \in [-T_r, 0) \\ 1 - \frac{\alpha_{rd}}{\alpha_{rd}+\beta_{rd}} e^{\frac{-t}{2w_{rd}\alpha_{rd}}}, & t \in [0, T_r) \\ 1, & t \in [T_r, \infty) \end{cases}. \tag{5.28}$$

In addition, the CDF, $P\big(f_{\text{PL}}(\bar{\epsilon}_r, t_r) \leq t \,|\, \bar{\epsilon}_r, \tilde{s}_r[k]=-1\big)$, is the same as $P\big(f_{\text{PL}}(\bar{\epsilon}_r, t_r) \leq t \,|\, \bar{\epsilon}_r, \tilde{s}_r[k]=+1\big)$ except for interchanging α_{rd} with β_{rd}.

Based on the expressions of t_s and t_r in (5.16), the term $(t_s + \sum_{r \in \mathcal{D}} t_r)$ in (5.22) can

be expressed as a quadratic form:

$$t_s + \sum_{r\in\mathcal{D}} t_r = \frac{2}{N_0}\Big(w_{sd}^{1/2}\boldsymbol{y}_{sd} + \sum_{r\in\mathcal{D}} w_{rd}^{1/2}\boldsymbol{y}_{rd}\Big)^H \boldsymbol{B}\Big(w_{sd}^{1/2}\boldsymbol{y}_{sd} + \sum_{r\in\mathcal{D}} w_{rd}^{1/2}\boldsymbol{y}_{rd}\Big). \quad (5.29)$$

It is clear that $(w_{sd}^{1/2}\boldsymbol{y}_{sd} + \sum_{r\in\mathcal{D}} w_{rd}^{1/2}\boldsymbol{y}_{rd})$ is a zero-mean, complex, Gaussian vector since $\boldsymbol{y}_{sd}$ and $\boldsymbol{y}_{rd}$ are independent, zero-mean, complex, Gaussian vectors. Similar to the case of $\boldsymbol{y}_{ij}^H\boldsymbol{B}\boldsymbol{y}_{ij}$ in (5.23), conditioned on $\tilde{s}[k] = +1$, we can express $(t_s + \sum_{r\in\mathcal{D}} t_r)$ as a weighted difference between two independent exponential RVs, i.e.,

$$t_s + \sum_{r\in\mathcal{D}} t_r = 2(\alpha_{\mathcal{D}}|u_{\mathcal{D}}|^2 - \beta_{\mathcal{D}}|v_{\mathcal{D}}|^2), \quad (5.30)$$

where $u_{\mathcal{D}}$ and $v_{\mathcal{D}}$ are i.i.d. complex Gaussian RVs with zero mean and unit variance, and $\alpha_{\mathcal{D}}$ and $\beta_{\mathcal{D}}$ are given by

$$\alpha_{\mathcal{D}} = w_{sd}\alpha_{sd} + \sum_{r\in\mathcal{D}} w_{rd}\alpha_{rd}, \qquad \beta_{\mathcal{D}} = w_{sd}\beta_{sd} + \sum_{r\in\mathcal{D}} w_{rd}\beta_{rd}. \quad (5.31)$$

Similar to (5.26), we can obtain the CDF of $(t_s + \sum_{r\in\mathcal{D}} t_r)$ conditioned on $\mathcal{D}$ and $\tilde{s}[k] = +1$ as

$$P\Big(t_s + \sum_{r\in\mathcal{D}} t_r \le t \,\Big|\, \tilde{s}[k] = +1, \mathcal{D}\Big) = \begin{cases} \frac{\beta_{\mathcal{D}}}{\alpha_{\mathcal{D}}+\beta_{\mathcal{D}}}\exp(\frac{z}{2\beta_{\mathcal{D}}}), & t \le 0 \\ 1 - \frac{\alpha_{\mathcal{D}}}{\alpha_{\mathcal{D}}+\beta_{\mathcal{D}}}\exp(\frac{-z}{2\alpha_{\mathcal{D}}}), & t \ge 0 \end{cases}. \quad (5.32)$$

The PDF (5.27), and the CDFs (5.28) and (5.32) will be used in the next two sections to evaluate the error performance of the relay systems under the DF protocol and the SR protocol.

5.6 Error Performance for a Relay System With the DF Protocol

In this section, we consider the error probability of a relay system working with the DF protocol. We obtain a closed-form expression for the BER and a high SNR

approximation for the BER of a single relay system, and a Chernoff upper bound on the BER of a multiple relay system.

5.6.1 Error Probability for a DF Single Relay System

In this subsection, we consider the error probability of a single relay system, i.e., $L = 1$. We derive the exact BER of the PL receiver (5.19), and a high SNR approximation for the BER.

Exact BER of a DF Single Relay System

According to the PL decision metric (5.19) and the equal probabilities of $\tilde{s}[k]$ being $+1$ and -1, the BER of the PL receiver at the destination is computed as

$$\begin{aligned} P_e^{\text{DF-PL}} =& P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1\big) \\ =& (1-\bar{\epsilon}_1) P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = +1\big) \\ &+ \bar{\epsilon}_1 P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = -1\big). \end{aligned} \tag{5.33}$$

Now, we compute the term $P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = +1\big)$ in (5.33) as follows:

$$\begin{aligned} &P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = +1\big) \\ &\quad = \mathbb{E}_{t_s|\tilde{s}[k]=+1}\Big[P\big(f_{\text{PL}}(\bar{\epsilon}_1, t_1) < -t \,|\, \bar{\epsilon}_1, t_s = t, \tilde{s}_1[k] = +1\big)\Big] \\ &\quad = \int_{-\infty}^{\infty} P\big(f_{\text{PL}}(\bar{\epsilon}_1, t_1) < -t \,|\, \bar{\epsilon}_1, \tilde{s}_1[k] = +1\big) f_{t_s|\tilde{s}[k]=+1}(t) dt. \end{aligned}$$

With the PDF of t_s given in (5.27) and the CDF of $f_{\text{PL}}(\bar{\epsilon}_1, t_1)$ given in (5.28), we can compute $P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = +1\big)$ in the above equation as

$$P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = +1\big) = \frac{\beta_{sd}}{\alpha_{sd}+\beta_{sd}} - P_{e1}^{\text{DF-PL}}, \tag{5.34}$$

and similarly, we can obtain

$$P\big(t_s + f_{\text{PL}}(\bar{\epsilon}_1, t_1) < 0 \,|\, \bar{\epsilon}_1, \tilde{s}[k] = +1, \tilde{s}_1[k] = -1\big) = \frac{\beta_{sd}}{\alpha_{sd}+\beta_{sd}} - P_{e2}^{\text{DF-PL}}, \quad (5.35)$$

where $P_{e1}^{\text{DF-PL}}$ and $P_{e2}^{\text{DF-PL}}$ are given as

$$P_{e1}^{\text{DF-PL}} = \varphi_{\text{PL}}(\alpha_{rd}, \beta_{rd}, \alpha_{sd}, \beta_{sd}, T_r) - \varphi_{\text{PL}}(\beta_{rd}, \alpha_{rd}, \beta_{sd}, \alpha_{sd}, T_r), \quad (5.36)$$

$$P_{e2}^{\text{DF-PL}} = \varphi_{\text{PL}}(\beta_{rd}, \alpha_{rd}, \alpha_{sd}, \beta_{sd}, T_r) - \varphi_{\text{PL}}(\alpha_{rd}, \beta_{rd}, \beta_{sd}, \alpha_{sd}, T_r), \quad (5.37)$$

with $r = 1$ and the function $\varphi_{\text{PL}}(r_1, r_2, r_3, r_4, T)$ is defined as

$$\varphi_{\text{PL}}(r_1, r_2, r_3, r_4, T) = \frac{r_1}{w_{sd}(r_1+r_2)(r_3+r_4)}\Big(\frac{1}{w_{rd}r_1}+\frac{1}{w_{sd}r_4}\Big)^{-1} \times \Big\{1 - \exp\Big[-\frac{T}{2}\Big(\frac{1}{w_{rd}r_1}+\frac{1}{w_{sd}r_4}\Big)\Big]\Big\}. \quad (5.38)$$

Therefore, with (5.33), (5.34) and (5.35), we can obtain a closed-form expression for the BER of a DF single relay system as

$$P_e^{\text{DF-PL}} = \frac{\beta_{sd}}{\alpha_{sd}+\beta_{sd}} - (1-\bar{\epsilon}_1)P_{e1}^{\text{DF-PL}} - \bar{\epsilon}_1 P_{e2}^{\text{DF-PL}}. \quad (5.39)$$

This BER expression for the PL receiver is a simple function of the SNRs, the variances and fade rates of all the transmission links. Actually, the exact BER of the PL receiver is an upper bound on the BER of the nonlinear ML receiver. As we will see later, the BER (5.39) of the PL receiver is very close to the BER of the nonlinear ML receiver in all cases for a single relay system working with DF protocol. We will verify this by simulations later.

The BER of the DF single relay system with DBPSK modulation was also obtained in [72, Eq.(46)] under the assumption that the fading coefficients remain constant over two consecutive symbol intervals. This is a special case of the BER expression (5.39) when $\rho_{ij} = 1$. Moreover, the BER expression in [72, Eq.(46)] is expressed as the sum of several infinite series, whereas the BER result (5.39) is expressed in a much simpler form.

BER Approximation at High SNR for a DF Single Relay System

We next develop a high SNR approximation of $P_e^{\text{DF-PL}}$ (5.39). The approximation will clearly show the effects of SNR, network geometry, and power allocation on the BER.

We assume that all the average received SNRs, η_{ij}, $ij \in \{sd, sr, rd\}$, become large, but remain in fixed proportion with one another. Specifically, we assume that

$$\eta_{ij} = k_{ij}\eta, \quad ij \in \{sd, sir, rd\}, \tag{5.40}$$

where η is the average single-hop SNR per bit, k_{ij}'s are constants related to network geometry and power allocation. We substitute (5.40) into (5.39), and express $P_e^{\text{DF-PL}}$ (5.39) as a function of $1/\eta$. The high SNR approximation can be obtained by applying a Taylor expansion at the point $1/\eta = 0$.

For simplicity, we assume that the fading coefficients of all the transmission links vary very slowly, and remain approximately unchanged in two consecutive symbol intervals, i.e., $\rho_{ij} = 1$, $ij \in \{sd, sr, rd\}$. With this assumption and (5.40), we can express $\alpha_{ij}, \beta_{ij}, w_{ij}$ and $\bar{\epsilon}_1$ as functions of k_{ij} and $1/\eta$ as follows

$$\alpha_{ij} = \frac{2k_{ij}+1/\eta}{1/\eta}, \quad \beta_{ij} = 1, \quad w_{ij} = \frac{k_{ij}}{2k_{ij}+1/\eta}, \quad \bar{\epsilon}_1 = \frac{1/\eta}{2(k_{sr}+1/\eta)}.$$

Using a Taylor expansion at the point $1/\eta = 0$, and ignoring the terms containing the 3rd and higher powers of $1/\eta$, we can approximate the three terms of $P_e^{\text{DF-PL}}$ in (5.39) as

$$\begin{aligned}
&\frac{\beta_{sd}}{\alpha_{sd}+\beta_{sd}} \approx \frac{1}{2k_{sd}}\frac{1}{\eta} - \frac{1}{2k_{sd}^2}\frac{1}{\eta^2}, \\
&(1-\bar{\epsilon}_1)P_{e1}^{\text{DF-PL}} \approx \frac{1}{2k_{sd}}\frac{1}{\eta} - \frac{1}{2k_{sd}}\Big(\frac{3}{2k_{rd}} + \frac{1}{k_{sd}} + \frac{1}{k_{sr}}\Big)\frac{1}{\eta^2}, \\
&\bar{\epsilon}_1 P_{e_2}^{\text{DF-PL}} \approx -\frac{4}{k_{sd}k_{sr}}\frac{1}{\eta^2}\ln(1+2k_{sr}\eta),
\end{aligned} \tag{5.41}$$

where the last approximation also makes use of the following approximation:

$$\exp\Big[ax\ln\big(\frac{x}{x+b}\big)\Big] \approx 1 - ax\ln\big(\frac{x+b}{x}\big), \text{ when } x \to 0.$$

Combining the three terms in (5.41), we have the BER approximation of $P_e^{\text{DF-PL}}$ at high SNR as

$$\begin{aligned} P_e^{\text{DF-PL}} &\approx \frac{1}{2k_{sd}}\Big[\frac{3}{2k_{rd}} + \frac{2+\ln(1+2k_{sr}\eta)}{2k_{sr}}\Big]\frac{1}{\eta^2} \\ &= \frac{1}{2\eta_{sd}}\Big[\frac{3}{2\eta_{rd}} + \frac{2+\ln(1+2\eta_{sr})}{2\eta_{sr}}\Big]. \end{aligned} \tag{5.42}$$

Remark 5.4. From (5.42), we can explicitly see that the BER of the PL receiver for a single relay system decreases as $(\eta^{-2}\ln\eta)$ when η is large. According to the definition of diversity order[1], the DF single relay system with the PL receiver has a diversity order of 2, i.e.,

$$-\lim_{\eta\to\infty}\frac{\ln(\eta^{-2}\ln\eta)}{\ln\eta} = 2 - \lim_{\eta\to\infty}\frac{\ln(\ln\eta)}{\ln\eta} = 2.$$

Although the BER of the DF single relay systems is greater by a factor of $(\ln\eta)$ than the BER of a traditional diversity system where BER decreases as η^{-2} for a large η, the diversity order of the single relay system is still 2 since the diversity order only captures the exponent of η^{-1} factor and ignores the smaller order terms.

Remark 5.5. From (5.42), we can see that η_{sr} and η_{rd} enter differently in the BER expression for the differential DF relay system. This means that the source-relay link and the relay-destination link play different roles in determining the end-to-end error performance. At a high SNR, i.e., a large η, usually we have $\ln(1+2k_{sr}\eta) = \ln(1+2\eta_{sr}) > 1$, and hence η_{sr} has greater effects on the BER than η_{rd}. This could be interpreted as follows: the reliability of the source-relay link determines the amount of possible error propagation, and hence has greater effects on the BER.

Note that η_{sr} and η_{rd} are related to both network geometry and power allocation between the source and relay, i.e., $\eta_{sr} = E_s\sigma_{sr}^2/N_0$ and $\eta_{rd} = E_r\sigma_{rd}^2/N_0$. We will illustrate how η_{sr} and η_{rd} will affect the BER in the numerical results.

[1]The diversity order is defined as $-\lim_{\text{SNR}\to\infty}(\log \text{BER})/(\log \text{SNR})$.

Furthermore, the BER approximation (5.42) for the differential DF relay reduces to [36, Eq. (19)] which is the BER for the non-coherent DF relay case except for replacing $2k_{ij}$ in (5.42) by k_{ij}. This is could be explained as follows. BFSK modulation is used in [36], while DBPSK modulation is used in this chapter. Asymptotically, DBPSK modulation is 3 dB more power-efficient than BFSK modulation. Thus, it is $2k_{ij}$ in (5.42), and k_{ij} in [36, Eq. (19)].

5.6.2 Error Probability for a DF Multiple Relay System

In this subsection, we consider the error probability of a multiple relay system, i.e., $L \geq 1$. The complexity of the error probability analysis of the PL detector increases exponentially in the number of relays. We develop a Chernoff upper bound on the BER and a high SNR approximation for the BER of a multiple relay system with the DF protocol.

The BER of a multiple relay system is defined as

$$P_e^{\text{DF-PL}} = P\Big(t_s + \sum_{r=1}^{L} f_{\text{PL}}(\bar{\epsilon}_r, t_r) < 0 \,|\, \{\bar{\epsilon}_r\}_{r=1}^{L}, \tilde{s}[k] = +1\Big).$$

By applying the Chernoff bound [86, Section 2.1.5] and the improved Chernoff bound by a factor of two [99, Section 4.2.4], the BER in the above equation can be upper bounded as

$$\begin{aligned} P_e^{\text{DF-PL}} &\leq \frac{1}{2}\mathbb{E}\bigg[\exp\Big(\delta t_s + \delta\sum_{r=1}^{L} f_{\text{PL}}(\bar{\epsilon}_r, t_r)\Big) \,|\, \{\bar{\epsilon}_r\}_{r=1}^{L}, \tilde{s}[k] = +1\bigg] \\ &= \frac{1}{2}\prod_{r=1}^{L}\Big\{(1-\bar{\epsilon}_r)\mathbb{E}\big[e^{\delta f_{\text{PL}}(\bar{\epsilon}_r, t_r)} \,|\, \bar{\epsilon}_r, \tilde{s}_r[k] = +1\big] + \bar{\epsilon}_r\mathbb{E}\big[e^{\delta f_{\text{PL}}(\bar{\epsilon}_r, t_r)} \,|\, \bar{\epsilon}_r, \tilde{s}_r[k] = -1\big]\Big\} \\ &\quad \times \mathbb{E}\big[e^{\delta t_s} | \tilde{s}[k] = +1\big], \end{aligned} \tag{5.43}$$

where $\delta \leq 0$ is the parameter to be optimized.

Based on the PDF of t_s conditioned on $\tilde{s}[k] = +1$, and t_r conditioned on $\tilde{s}_r[k] = +1$ or $\tilde{s}_r[k] = -1$, we can easily compute the characteristic functions of

t_s and $f_{\text{PL}}(\bar{\epsilon}_r, t_r)$ as

$$\begin{aligned}
&\mathbb{E}\big[e^{\delta t_s}|\tilde{s}[k]=+1\big] = \phi_1(\alpha_{sd},\beta_{sd},w_{sd},\delta),\\
&\mathbb{E}\big[e^{\delta f_{\text{PL}}(\bar{\epsilon}_r,t_r)}\,|\,\bar{\epsilon}_r,\tilde{s}_r[k]=+1\big] = \phi_1(\alpha_{rd},\beta_{rd},w_{rd},\delta) + \frac{2w_{rd}\delta}{\alpha_{rd}+\beta_{rd}}\times\\
&\qquad\qquad [\phi_2(\beta_{rd},w_{rd},T_r,\delta)-\phi_2(\alpha_{rd},w_{rd},T_r,-\delta)],\\
&\mathbb{E}\big[e^{\delta f_{\text{PL}}(\bar{\epsilon}_r,t_r)}\,|\,\bar{\epsilon}_r,\tilde{s}_r[k]=-1\big] = \phi_1(\alpha_{rd},\beta_{rd},w_{rd},-\delta) + \frac{2w_{rd}\delta}{\alpha_{rd}+\beta_{rd}}\times\\
&\qquad\qquad [\phi_2(\alpha_{rd},w_{rd},T_r,\delta)-\phi_2(\beta_{rd},w_{rd},T_r,-\delta)],
\end{aligned} \tag{5.44}$$

where $-1/(2w_{sd}\beta_{sd}) < \delta \le 0$, and the functions $\phi_1(\alpha,\beta,w,\delta)$ and $\phi_2(\alpha,w,T,\delta)$ are defined as

$$\begin{aligned}
\phi_1(\alpha,\beta,w,\delta) &= \frac{1}{(1-2w\alpha\delta)(1+2w\beta\delta)},\\
\phi_2(\alpha,w,T,\delta) &= \frac{\alpha^2}{1+2w\alpha\delta}\exp\big[-T(\frac{1}{2w\alpha}+\delta)\big].
\end{aligned}$$

Substituting (5.44) into (5.43), we can obtain the Chernoff bound on the BER. The tightest bound has to be obtained by selecting δ that minimizes the right hand side of (5.43). By solving the equation $d\phi_1(\alpha_{ij},\beta_{ij},w_{ij},\delta)/d\delta = 0$, it is easy to check that the value of δ that minimizes $\phi_1(\alpha_{ij},\beta_{ij},w_{ij},\delta)$ is $-1/2$, regardless of ij. However, it is very difficult to find the δ which minimizes the Chernoff bound (5.43) because (5.43) and (5.44) are fairly complex. Therefore, we choose $\delta = -1/2$ to evaluate the Chernoff bound, and thus obtain the BER upper bound

$$P_e^{\text{DF-PL}} \le \frac{1}{2}\left[1-\Big(\frac{\eta_{sd}\rho_{sd}}{1+\eta_{sd}}\Big)^2\right]\prod_{r=1}^{L}\big[(1-\bar{\epsilon}_r)C_r+\bar{\epsilon}_r D_r\big] \tag{5.45}$$

where C_r and D_r are given by

$$\begin{aligned}
C_r &= 1-\Big(\frac{\eta_{rd}\rho_{rd}}{1+\eta_{rd}}\Big)^2\left\{1-\exp\Big[-\frac{T_r}{2}\Big(\frac{1+\eta_{rd}}{\eta_{rd}\rho_{rd}}\Big)\Big]\right\} = 1-\Big(\frac{\eta_{rd}\rho_{rd}}{1+\eta_{rd}}\Big)^2\left[1-\Big(\frac{\bar{\epsilon}_r}{1-\bar{\epsilon}_r}\Big)^{\frac{1+\eta_{rd}}{2\eta_{rd}\rho_{rd}}}\right],\\
D_r &= \frac{4\alpha_{rd}\beta_{rd}}{4\alpha_{rd}\beta_{rd}-3(1+\eta_{rd})^2}+\frac{\eta_{rd}\rho_{rd}}{1+\eta_{rd}}\left(\frac{\alpha_{rd}e^{T_r}}{\alpha_{rd}-3\beta_{rd}}+\frac{\beta_{rd}e^{-T_r}}{3\alpha_{rd}-\beta_{rd}}\right)\exp\Big[-\frac{T_r}{2}\Big(\frac{1+\eta_{rd}}{\eta_{rd}\rho_{rd}}\Big)\Big].
\end{aligned}$$

The Chernoff upper bound on the BER is valid for a DF relay system with an arbitrary number of relay nodes.

If all the relays have the same average BER $\bar{\epsilon}$ when detecting $\tilde{s}[k]$, i.e., $\bar{\epsilon} = \bar{\epsilon}_r$ and $T = T_r$ for any $r = 1, 2, \cdots, L$, and all the transmission links from relays to destination have the fade rate and the same average received SNR per bit, i.e., $\rho_{rd} = \rho_d$ and $\eta_{rd} = \eta_d$ for any $r = 1, 2, \cdots, L$, the Chernoff bound (5.45) on BER is reduced to

$$P_e^{\text{DF-PL}} \leq \frac{1}{2}\left[1 - \left(\frac{\eta_{sd}\rho_{sd}}{1+\eta_{sd}}\right)^2\right]\left[(1-\bar{\epsilon})C + \bar{\epsilon}D\right]^L, \tag{5.46}$$

where $C = C_r$ and $D = D_r$ for any $r = 1, 2, \cdots, L$.

We next develop a high SNR approximation for the BER Chernoff bound given in (5.45). For simplicity, we assume that the fading processes are stationary, i.e., $\rho_{ij} = 1$, $ij \in \{sd, sr, rd\}_{r=1}^{L}$. In the same way as in (5.40), we assume that the average received SNRs remain in fixed proportion. Substituting $\eta_{sr} = k_{sr}\eta$ into the expression of $\bar{\epsilon}_r$ in (5.6) and applying a Taylor expansion at the point $1/\eta = 0$, we can approximate $\bar{\epsilon}_r$ and $(1 - \bar{\epsilon}_r)$ as

$$\bar{\epsilon}_r \approx \frac{1}{2k_{sr}}\eta^{-1}, \qquad 1 - \bar{\epsilon}_r \approx 1 - \frac{1}{2k_{sr}}\eta^{-1}. \tag{5.47}$$

We also substitute (5.40) into the expressions of C_r and D_r, apply a Taylor expansion at the point $1/\eta = 0$, and then obtain high SNR approximations for C_r and D_r as

$$C_r \approx \frac{2}{k_{rd}}\eta^{-1} + \left(1 - \frac{2}{k_{rd}}\eta^{-1}\right)\sqrt{\frac{\bar{\epsilon}_r}{1-\bar{\epsilon}_r}} \approx \frac{1}{\sqrt{2k_{sr}}}\eta^{-1/2} + \frac{2}{k_{rd}}\eta^{-1} \tag{5.48}$$

$$D_r \approx -\frac{1}{3k_{rd}}\eta^{-1} + \left(1 - \frac{1}{k_{rd}}\eta^{-1}\right)\sqrt{\frac{1-\bar{\epsilon}_r}{\bar{\epsilon}_r}} \approx \sqrt{2k_{sr}}\,\eta^{1/2} - \frac{\sqrt{2k_{sr}}}{k_{rd}}\eta^{-1/2}. \tag{5.49}$$

With the high SNR approximations for $\bar{\epsilon}_r$, C_r and D_r, we can have the term $\left[(1 - \bar{\epsilon}_r)C_r + \bar{\epsilon}_r D_r\right]$ be approximated as

$$(1-\bar{\epsilon}_r)C_r + \bar{\epsilon}_r D_r \approx \frac{2}{\sqrt{2k_{sr}}}\eta^{-1/2} + \frac{2}{k_{rd}}\eta^{-1} = \sqrt{\frac{2}{\eta_{sr}}} + \frac{2}{\eta_{rd}}, \tag{5.50}$$

and then obtain a high SNR approximation for the Chernoff bound on the BER as

$$P_e^{\text{DF-PL}} \sim \frac{1}{\eta^{L/2+1}} \frac{1}{k_{sd}} \prod_{r=1}^{L} \Big(\sqrt{\frac{2}{k_{sr}}} + \frac{2}{k_{rd}} \eta^{-1/2} \Big) = \frac{1}{\eta_{sd}} \prod_{r=1}^{L} \Big(\sqrt{\frac{2}{\eta_{sr}}} + \frac{2}{\eta_{rd}} \Big). \quad (5.51)$$

Remark 5.6. From the high SNR approximation (5.51), we can see that the DF relay system achieves a diversity order greater than $(L/2+1)$. This result is consistent with [36, Theorem 1] for a DF relay system with BFSK modulation. In [36, Theorem 1], the diversity order lower bound $(L/2+1)$ is obtained through using the Bhattacharyya upper bound on BER.

Remark 5.7. For a DF multiple relay system, we also can see from (5.51) that η_{sr} dominates the source-relay-destination transmission link, and η_{sr}s have greater effects on the end-to-end BER than η_{rd}s. This matches with the result in Remark 5.5 for a DF single relay system.

5.7 Error Performance for a Relay System With the SR Protocol

For a relay system with the SR protocol, the destination receiver is an MRC receiver as shown in (5.22). In this section, we will analyze the error probability of the MRC receiver for the relay system working with the SR protocol.

According to the destination decision metric in (5.22) and the fact that the decoding set $\mathcal{D}$ is a random set, the BER of the relay system with the SR protocol is computed as

$$\begin{aligned} P_e^{\text{SR}} &= P\Big(t_s + \sum_{r \in \mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k] = +1 \Big) \\ &= \sum_{\mathcal{D}} P\Big(t_s + \sum_{r \in \mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k] = +1, \mathcal{D} \Big) P(\mathcal{D}), \end{aligned} \quad (5.52)$$

where the second equality follows from the total probability theorem, $P(\mathcal{D})$ is the probability of a particular decoding set $\mathcal{D}$, and $\sum_{\mathcal{D}}(\cdot)$ represents the summation over all the possible decoding sets.

5.7.1 Error Probability Conditioned on the Decoding Set

Conditioned on $\mathcal{D}$ being the decoding set, the CDF of $(t_s + \sum_{r\in\mathcal{D}} t_r)$ is given in (5.32). Thus, we can obtain the BER conditioned on the decoding set $\mathcal{D}$ as

$$P\Big(t_s + \sum_{r\in\mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k] = +1, \mathcal{D}\Big) = \frac{\beta_{\mathcal{D}}}{\alpha_{\mathcal{D}} + \beta_{\mathcal{D}}}. \tag{5.53}$$

Although (5.53) is an exact expression of the conditional BER, we cannot see the effects of decoding set and various system parameters on the conditional BER. In the following, we develop a Chernoff bound on the conditional BER, which shows explicitly the effects of decoding set and various system parameters.

By applying the Chernoff upper bound [86, Section 2.1.5] and the improved Chernoff bound by a factor of two [99, Section 4.2.4], the conditional BER in (5.53) can be upper bounded as

$$P\Big(t_s+\sum_{r\in\mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k]=+1, \mathcal{D}\Big) \le \frac{1}{2}\mathbb{E}\big[e^{\delta^{\mathrm{SR}} t_s}|\tilde{s}[k]=+1\big] \prod_{r\in\mathcal{D}} \mathbb{E}\big[e^{\delta^{\mathrm{SR}} t_r}|\tilde{s}[k]=+1, \mathcal{D}\big].$$

The characteristic function of t_s is already given in (5.44), and the characteristic function of t_r takes the similar form as the one of t_s. The tightest upper bound on the conditional BER can be obtained by selecting $\delta^{\mathrm{SR}} = -1/2$, and given by

$$P\Big(t_s+\sum_{r\in\mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k]=+1, \mathcal{D}\Big) \le \frac{1}{2}\Big[1 - \Big(\frac{\eta_{sd}\rho_{sd}}{1+\eta_{sd}}\Big)^2\Big] \prod_{r\in\mathcal{D}} \Big[1 - \Big(\frac{\eta_{rd}\rho_{rd}}{1+\eta_{rd}}\Big)^2\Big]. \tag{5.54}$$

In fact, (5.54) is equal to the BER Chernoff bound for nonidentical MISO channels [100, Eq. (33)].

In the same way as in Eq. (5.40) in Section 5.6.1, we assume that the average received SNRs remain in fixed proportion. By applying a Taylor expansion of the right

hand side of (5.54) at the point $1/\eta = 0$, a high SNR approximation of the conditional BER can be written as

$$P\Big(t_s+\sum_{r\in\mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k]=+1, \mathcal{D}\Big) \sim \frac{1}{2}\Big[(1-\rho_{sd}^2)+\frac{2}{\eta k_{sd}}\Big]\prod_{r\in\mathcal{D}}\Big[(1-\rho_{rd}^2)+\frac{2}{\eta k_{rd}}\Big].$$

We further assume that the fading coefficients of all the transmission links change very slowly, i.e., $\rho_{ij} = 1, ij \in \{sd, rd\}$. Then, the above high SNR approximation for the conditional BER reduces to

$$P\Big(t_s+\sum_{r\in\mathcal{D}} t_r < 0 \,\Big|\, \tilde{s}[k]=+1, \mathcal{D}\Big) \sim \frac{1}{\eta^{(1+|\mathcal{D}|)}} 2^{|\mathcal{D}|}\frac{1}{k_{sd}}\prod_{r\in\mathcal{D}}\frac{1}{k_{rd}}. \tag{5.55}$$

From above equation, we can see that the conditional BER has a diversity order of $(1 + |\mathcal{D}|)$.

5.7.2 Decoding Set Probability

We now consider $P(\mathcal{D})$, the probability of a particular decoding set. Since all the potential relays make their decisions independently, and the fading coefficients of the links from the source to relays are also independent, we have

$$\begin{aligned} P(\mathcal{D}) &= \prod_{r\in\mathcal{D}} P(r\in\mathcal{D}) \times \prod_{r\notin\mathcal{D}} P(r\notin\mathcal{D}) \\ &= \prod_{r\in\mathcal{D}} P(r\in\mathcal{D}) \times \prod_{r\notin\mathcal{D}} \big[1-P(r\in\mathcal{D})\big] \end{aligned} \tag{5.56}$$

According to the relay selection criterion in (5.9) for the SR protocol, and the instantaneous BER (5.7) at the r-th relay, the probability of the r-th relay belonging to the decoding set $\mathcal{D}$ is computed as

$$\begin{aligned} P(r\in\mathcal{D}) &= P\left(\epsilon_r[1]<\epsilon, \epsilon_r[2]<\epsilon, \cdots, \epsilon_r[N_f]<\epsilon\right) \\ &= P(|y_{sr}[0]|>y_{sr}, |y_{sr}[1]|>y_{sr}, \cdots, |y_{sr}[N_f-1]|>y_{sr}) \\ &= P(|y_{sr}[0]|>y_{sr}) \prod_{k=1}^{N_f-1} P\big(|y_{sr}[k]|>y_{sr} \,\big|\, |y_{sr}[k-1]|>y_{sr}\big), \end{aligned} \tag{5.57}$$

where $y_{sr} = \sqrt{N_0(\eta_{sr}+1)/(w_{sr}\eta_{sr}\rho_{sr})}\,\text{erfc}^{-1}(2\epsilon)$, and $\text{erfc}^{-1}(\cdot)$ represents the inverse function of $\text{erfc}(\cdot)$.

Conditioned on $x_s[k]$ and $x_s[k-1]$, $y_{sr}[k]$ and $y_{sr}[k-1]$ are correlated, zero-mean, complex, Gaussian RVs. Then, $|y_{sr}[k]|$ and $|y_{sr}[k–1]|$ are bivariate Rayleigh distributed [96, Ch. 6.1]. The power correlation coefficient, ϱ_{sr}, of the two Rayleigh distributed amplitudes,$|y_{sr}[k]|$ and $|y_{sr}[k-1]|$, is given by [96, Section 6.1, Page 170]

$$\varrho_{sr} = \frac{\left|\mathbb{E}\big(y_{sr}[k]y_{sr}^*[k-1]\big)\right|^2}{\mathbb{E}\big(|y_{sr}[k]|^2\big)\cdot\mathbb{E}\big(|y_{sr}[k-1]|^2\big)} = \frac{\eta_{sr}^2\rho_{sr}^2}{(\eta_{sr}+1)^2}.$$

The bivariate Rayleigh CCDF is given by [96, Section 6.1], [101, Eq.(4)]

$$P\big(|y_{sr}[k]| > y_{sr},\, |y_{sr}[k-1]| > y_{sr}\big) = \exp\Big(-\frac{y_{sr}^2}{N_0(\eta_{sr}+1)}\Big)\times \\ \big[1 - Q_1(a_{sr}y_{sr}, b_{sr}y_{sr}) + Q_1(b_{sr}y_{sr}, a_{sr}y_{sr})\big],$$

where $Q_1(\cdot,\cdot)$ is the first order Marcum Q-function, and

$$a_{sr} = \sqrt{\frac{2/(1-\varrho_{sr})}{N_0(\eta_{sr}+1)}} = \sqrt{\frac{2w_{sr}(\eta_{sr}+1)}{N_0\eta_{sr}\rho_{sr}}}, \tag{5.58}$$

$$b_{sr} = a_{sr}\sqrt{\varrho_{sr}} = \sqrt{\frac{2w_{sr}\eta_{sr}\rho_{sr}}{N_0(\eta_{sr}+1)}}. \tag{5.59}$$

Then, the conditional probability in (5.57) can be computed as

$$P\big(|y_{sr}[k]| > y_{sr} \,\big|\, |y_{sr}[k-1]| > y_{sr}\big) = \frac{P\big(|y_{sr}[k]| > y_{sr},\, |y_{sr}[k-1]| > y_{sr}\big)}{P\big(|y_{sr}[k-1]| > y_{sr}\big)} \\ = 1 - Q_1(a_{sr}y_{sr}, b_{sr}y_{sr}) + Q_1(b_{sr}y_{sr}, a_{sr}y_{sr}). \tag{5.60}$$

By substituting the conditional probability (5.60) into (5.57), the probability of the r-th relay belonging to the decoding set $\mathcal{D}$ is expressed as

$$P(r \in \mathcal{D}) = e^{-y_{sr}^2/[N_0(\eta_{sr}+1)]}\big[1-Q_1(a_{sr}y_{sr}, b_{sr}y_{sr})+Q_1(b_{sr}y_{sr}, a_{sr}y_{sr})\big]^{N_f-1}. \tag{5.61}$$

The Marcum Q-functions in the above equation do not indicate how the systems parameters affect $P(r \in \mathcal{D})$. Hence, we derive an upper bound on $P(r \in \mathcal{D})$ as (see

Appendix C for details)

$$P(r \in \mathcal{D}) \leq (2\epsilon)^{\left[\frac{1}{w_{sr}\eta_{sr}\rho_{sr}}+4(N_f-1)\frac{\beta_{sr}^2}{(\alpha_{sr}^2+\beta_{sr}^2)}\right]}. \tag{5.62}$$

When the fading coefficients change very slowly in time, $\rho_{sr} = 1$, and the average single-hop SNR η is very large, we can obtain a high SNR approximation for $P(r \in \mathcal{D})$ by a Taylor expansion of the right hand side of (5.62) at the point $1/\eta = 0$, i.e.,

$$P(r \in \mathcal{D}) \sim 1 - \frac{\ln(\frac{1}{2\epsilon})}{k_{sr}\eta}\Big(2 + \frac{N_f - 1}{k_{sr}\eta + 1}\Big). \tag{5.63}$$

The above Taylor expansion ignores the terms containing higher orders of $1/\eta$. Since the largest term which has been ignored is positive, the upper bound (5.62) becomes the approximation (5.62). Substituting (5.63) in to (5.56), we can obtain a high SNR approximation for the decoding set probability as

$$\begin{aligned} P(\mathcal{D}) &\sim \prod_{r\in\mathcal{D}}\left\{1 - \frac{\ln(\frac{1}{2\epsilon})}{\eta k_{sr}}\Big[2 + \frac{N_f - 1}{\eta k_{sr}+1}\Big]\right\} \times \prod_{r\notin\mathcal{D}}\left\{\frac{\ln(\frac{1}{2\epsilon})}{\eta k_{sr}}\Big(2 + \frac{N_f - 1}{\eta k_{sr}+1}\Big)\right\} \\ &\sim \frac{1}{\eta^{L-|\mathcal{D}|}}\Big[\ln\big(\frac{1}{2\epsilon}\big)\Big]^{L-|\mathcal{D}|}\prod_{r\notin\mathcal{D}}\frac{1}{k_{sr}}\Big(2 + \frac{N_f - 1}{\eta k_{sr}+1}\Big). \end{aligned} \tag{5.64}$$

Combining (5.55) and (5.64) into (5.52), we can obtain a high SNR approximation for the BER of the relay system with SR protocol as

$$P_e^{\text{SR}} \sim \frac{\big[\ln(\frac{1}{2\epsilon})\big]^L}{\eta^{L+1}}\frac{1}{k_{sd}} \times \sum_{\mathcal{D}}\Big[\ln\big(\frac{1}{2\epsilon}\big)\Big]^{-|\mathcal{D}|}\prod_{r\in\mathcal{D}}\frac{1}{k_{rd}}\prod_{r\notin\mathcal{D}}\frac{1}{k_{sr}}\Big(2 + \frac{N_f - 1}{\eta k_{sr}+1}\Big). \tag{5.65}$$

For a multiple relay system with L relays, there are 2^L possible decoding sets. Hence, it is difficult to obtain the exact BER for an arbitrary L. However, from (5.65), we can explicitly see that the relay system with the SR protocol achieves a diversity order of $(L + 1)$ from the error probability perspective.

Remark 5.8. Although high SNR is assumed in (5.65), the instantaneous channel gain still can be very small, and instantanous error probability also can be quite high. The diversity order of $(L + 1)$ means that the relay system with the SR protocol offers full

diversity in the number of all the potential cooperating nodes, not just the number of nodes that are transmitting to the destination. This result is consistent with the diversity order shown from the outage probability perspective in [34, 44].

5.8 Numerical Results

In this section, we present simulation results and numerical results for relay systems that use DBPSK modulation and employ either the DF or SR protocol. The variances of the fading coefficients are assigned using a path-loss model of the form $\sigma_{ij}^2 \propto d_{ij}^{-4}$, where d_{ij} denotes the distance between nodes i and j. The quantity σ_{sd}^2 is normalized to one. The autocorrelation functions of the fading processes are assigned according to Jake's model, i.e., $\rho_{ij} = J_0(2\pi f_D T_s), \forall ij \in \{sd, sr, rd\}_{r=1}^{L}$, where f_D is the maximum Doppler frequency and T_s is the sampling period which is equal to the bit interval. Suppose that the total transmit energy is E, i.e., $E = E_s + \sum_{r=1}^{L} E_r$. For simplicity, we also assume that the source and all the relays have the same transmit energy per bit, i.e., $E_s = E_r = E/(L+1), \forall r = 1, 2, \cdots L$. We compare the BER of a relay system with that of a non-cooperative system which only has a direct link from the source to the destination. For fair comparison, the transmitted energy per bit for the non-cooperative system is selected to be E. In all the numerical results, we define SNR as E/N_0.

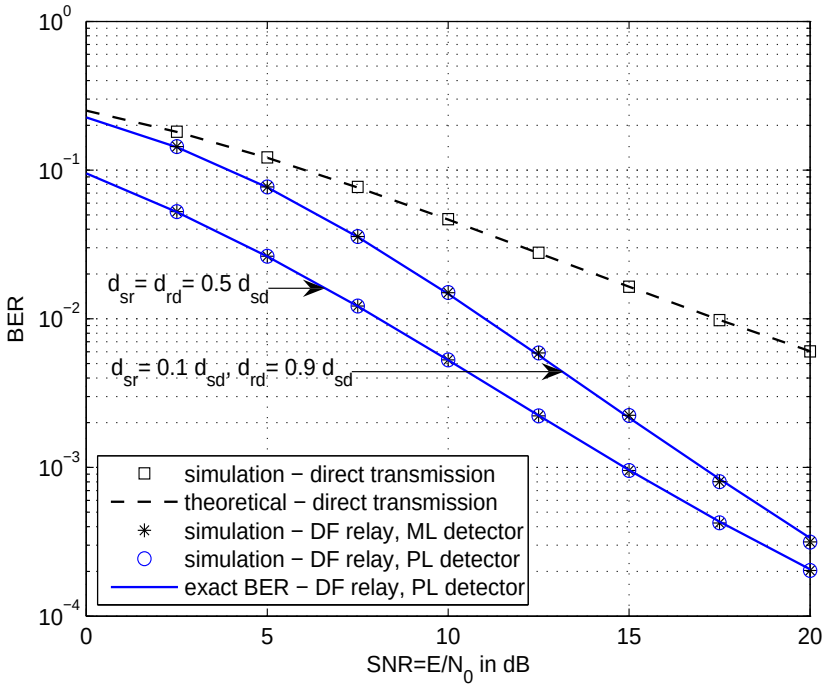

Figure 5.2: Average BER of a DF single relay system for the case of $f_D = 75$ Hz and $T_s = 2 \times 10^{-4}$s ($\rho_{ij} = 0.9978$).

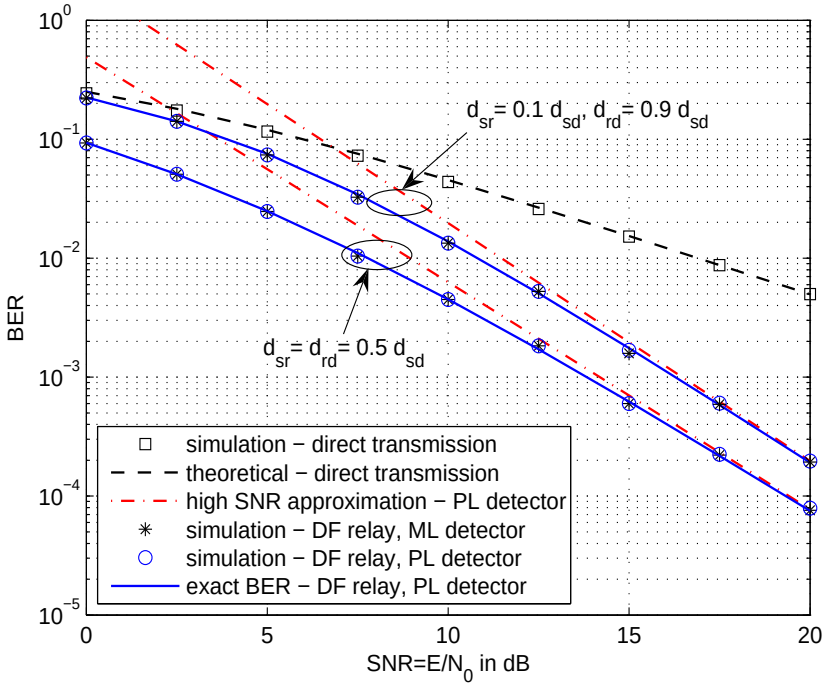

Figure 5.3: Average BER of a DF single relay system for the case of $f_D = 75$ Hz and $T_s = 2 \times 10^{-5}$s ($\rho_{ij} = 0.999978$).

In Figs. 5.2 and 5.3, we show the BER of single relay systems with the DF protocol for various network geometries and fade rates. From the slopes of the BER curves in Figs. 5.2 and 5.3, we can see clearly the diversity gains achieved by the relay systems over the non-cooperative systems. It also can be observed that there are certain shifts between the BER curves of the relay systems and that of the non-cooperative system. These shifts correspond to power gains which combat path loss and depend on the location of the relay. Furthermore, the exact BER curves of the PL detector coincide with the simulation results of the nonlinear ML detector. Since the exact BER (5.39) of the PL detector is in closed form and consists of only elementary functions, it is convenient to estimate the BER of the ML detector by using the exact BER (5.39) of the PL detector, rather than resorting to Monte Carlo simulations. The curves for the BER approximation (5.42) at high SNR are also shown in Fig. 5.3. The BER approximation (5.42) is very close to the exact BER when the SNR is greater than 12 dB.

To study the impacts of the location of the relay on the end-to-end BER for a DF single relay system, we assume that the source, relay and destination are collinear, i.e., $d_{sr} + d_{rd} = d_{sd}$, and show the BER versus the distance ratio d_{sr}/d_{sd} in Fig. 5.4. From Fig. 5.4, we can see that the end-to-end BER is decreasing with d_{sr}/d_{sd} when d_{sr}/d_{sd} is small, and is increasing with d_{sr}/d_{sd} when d_{sr}/d_{sd} is large. It also can be observed that, the end-to-end BER is minimized when the relay is closer to the source than the destination, i.e., $d_{sr}/d_{sd} < 0.5$. This means that to minimize the end-to-end BER, the source-relay link received SNR η_{sr} should be greater than the relay-destination link received SNR η_{rd}.

Fig. 5.5 presents the simulated BERs and the Chernoff bounds (5.45) for DF multiple relay systems. We can see that the Chernoff bounds (5.45) have the same slopes as the simulation results, although the Chernoff bounds are not very tight.

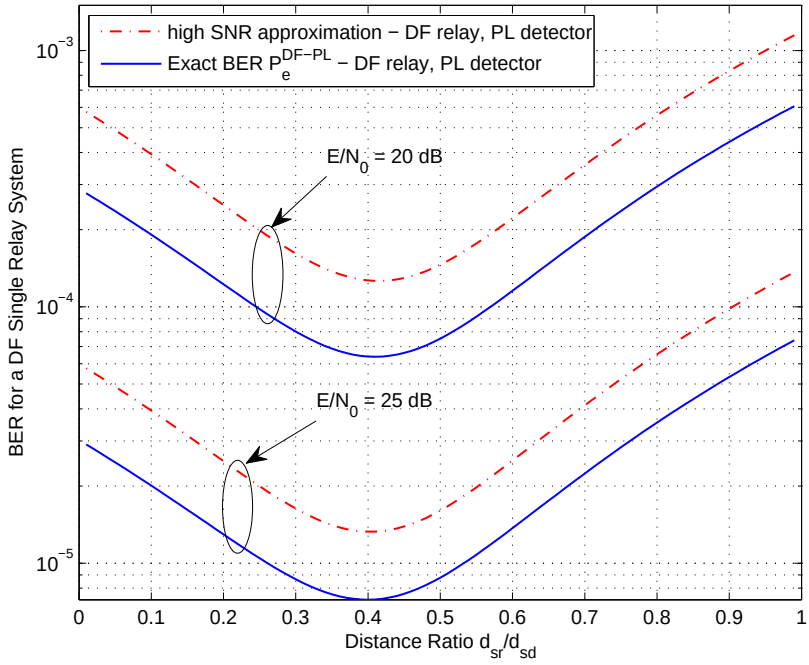

Figure 5.4: Average BER of a DF single relay system versus distance ratio d_{sr}/d_{sd} when $d_{sr} + d_{rd} = d_{sd}$, for the case of $f_D = 75$ Hz and $T_s = 2 \times 10^{-5}$s ($\rho_{ij} = 0.999978$).

Therefore, the Chernoff bounds still retain the diversity order in the BERs. In addition, it is observed that the diversity benefit is increasing as the number of relays increases, but the spectral efficiency is decreasing as the number of relays increases, since the spectral efficiency is $1/(L+1)$.

We compare the BER of the DF relay system with that of the SR relay system in Fig 5.6. We can observe that for a single relay system, both the DF and SR protocols achieve the same diversity order. However for a relay system with two relays, the DF protocol only achieves the diversity order of a DF single relay system, whereas the SR protocol achieves a diversity order of $(L+1)$. This can be interpreted as follows. For the DF relay system, the ML and PL detectors only make use of the average BERs of the source-relay transmission, while for the SR relay system, each relay makes use of

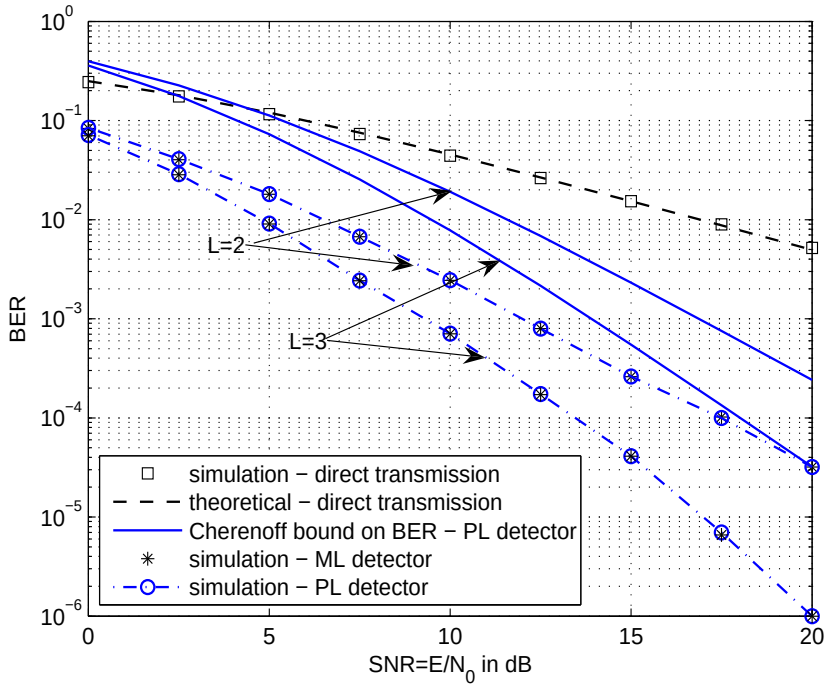

Figure 5.5: Average BER for DBPSK modulation, with $d_{sr} = d_{rd} = 0.5d_{sd}$ for the case of $f_D = 75$ Hz and $T_s = 2 \times 10^{-5}$s.

the instantaneous BERs instead of the average BER of the source-relay transmission to perform selection. The average BER of the source-relay transmission cannot fully characterize the reliability of the source-relay transmission. Therefore, the ML and PL detectors in the DF relay system cannot achieve full diversity when the number of relay nodes is greater than one.

5.9 Conclusion

In this chapter, we propose receiver designs and analyze error probabilities for DF and SR multiple relay systems with DBPSK modulation. For a DF relay system, the ML and PL detectors are derived, and they take account of the average BERs of all the

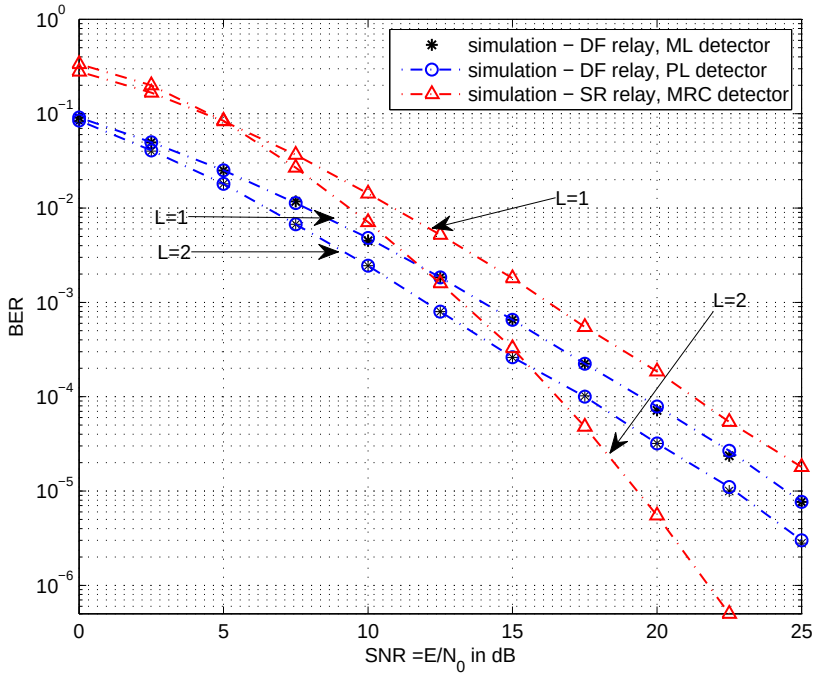

Figure 5.6: Average BER for relay systems with DF protocol or SR protocol, when $d_{sr} = d_{rd} = 0.5d_{sd}$ **for the case of** $f_D = 75$ **Hz and** $T_s = 2 \times 10^{-5}$**s.**

source-relay transmissions. The BER of the PL detector is analyzed. For a DF single relay system, the exact BER of a PL detector is obtained as a simple function of SNRs, faded rates and channel statistics. Based on the exact BER, a BER approximation at high SNR is derived. The BER approximation shows explicitly the diversity order and the different effects of the source-relay link and the relay-destination link on the end-to-end error probability. For a DF multiple relay system, a Chernoff upper bound on the BER and a high SNR approximation for the BER are obtained. The PL detector has lower complexity and can achieve a BER performance similar to that of the ML detector. For a SR relay system, each relay computes the instantaneous BERs of the source-relay transmission and use them to decide whether transmit to the destination or remain silent. In this case, the ML receiver at the destination is the MRC receiver. We

have also analyzed the error probability of the SR multiple relay system at high SNR, and have shown that the diversity order is the number of all the potential cooperating nodes.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

A multiple antenna system, because of its space diversity and spectral efficiency, is one of the most favorable solutions in numerous wireless applications. Multiple antennas have been used in both point-to-point MIMO systems and cooperative relay systems. In a point-to-point MIMO system, all the transmit antennas are collocated at the same transmitter, and they have perfect knowledge of what the other transmit antennas transmit, so they can fully collaborate to improve transmissions; Similarly, the receive antennas are collocated at the same receiver, and they have perfect knowledge of what the other receive antennas receive, so they can fully collaborate to improve receptions. Whereas a relay system benefits from space diversity by using multiple terminals as a virtual antenna array. The cooperative relay systems are fundamentally different from the point-to-point MIMO systems, because information is not a priori known to the cooperating relays and needs to be communicated over fading and noisy links. In this dissertation, we have studied performance limits of point-to-point MIMO systems and cooperative transmission strategies for cooperative relay system.

The MI of a MIMO system in a fading environment is an important information

theoretical performance measure. We have studied the statistical distribution of the MI under a Rician fading environment through studying the PDF and CCDF of the determinant and trace of a noncentral complex Wishart matrix. While the trace of a noncentral complex Wishart matrix is distributed as a noncentral chi-square RV, the determinant is distributed as a product of independent noncentral chi-square RVs. Then, by making use of the Mellin transform, we have obtained the PDF and CCDF of the determinant of a noncentral complex Wishart matrix. Finally, some tight lower and upper bounds on the CCDF of the MI are derived. Compared with existing results, the bounds derived in this dissertation are not only given in closed forms, but are also readily applicable to the evaluation of the outage probability with sufficiently high accuracy. In addition, all the results can be reduced to the case of Rayleigh fading.

For a cooperative relay system, one objective of this dissertation is to derive the outage probability and the ergodic rate of the DF relay system in a Rayleigh fading environment. Based on the outage probability and ergodic rate, this dissertation proposes an outage-optimal transmission strategy in order to minimize the outage probability, and an ergodic-optimal transmission strategy in order to maximize the ergodic rate. To the best of our knowledge, the outage probability or the ergodic rate of the DF relay system in a Rayleigh fading environment for an arbitrary correlation coefficient between the transmit signals from the source and relay are not available in the literature, because of the complexity of the information rate. Using a new approach that rewrites the information rate in terms of a Hermitian quadratic form, we have derived the outage probability and ergodic rate for an arbitrary correlation coefficient of the transmit signals from the source and relay. When the fading process is non-ergodic, with the outage probability we obtained, we can measure the tradeoff between transmission rate and reliability. When the fading process is ergodic, one can transmit with vanishing errors at a rate equal to the ergodic rate we obtained. We

have discovered a quite interesting finding that even though the information rate is a function of the correlation coefficient of the transmit signals from the source and relay, the outage probability and the ergodic rate only depend on the magnitude of the correlation coefficient, but are independent of the phase of the correlation coefficient. The independence is due to the uniform distribution of the phases of the channel gains in Rayleigh fading environments. This finding is significant because it can guide the design of transmission strategies at the source and relay. When one designs the correlation coefficient of the transmit signals from the source and relay, one only needs to design the magnitude of the correlation coefficient, and do not need to take the phase of the correlation coefficient into account, since the phase does not have any effects on the outage probability or the ergodic rate. We have also derived the outage-optimal transmission strategy in which the outage-optimal correlation coefficient generally depends on the second order statistics of the three transmission links, the transmit power and the rate threshold. When the rate threshold is less than a certain value, we have found that the outage-optimal correlation coefficient is zero irrespective of the statistics of the transmission links and the transmit power. The outage-optimal power allocation depends on the second order statistics of the three transmission links and the rate threshold. We have also found that the source should always be allocated more power than the relay. Furthermore, it has been shown for the first time that the ergodic rate is a monotonically decreasing function of the correlation coefficient of the transmit signals from the source and relay, and hence the ergodic rate is maximized when the correlation coefficient is zero, regardless of the channel statistics and the transmit power. We have also shown that it is not always beneficial to use the relay, and the ergodic-optimal power allocation depends on the second order statistics of the three transmission links. These transmission strategies can greatly improve the ergodic rate.

Moreover, we have also extended the outage probability of the DF relay system in a Rayleigh fading environment to the ones in a Rician fading environment. The extension to the Rician fading environment is not straightforward due to the noncentral property of Rician fading. We have derived an analytical expression for the outage probability in terms of a single integral. Moreover, we have provided a geometric interpretation of the outage probability, and obtain an upper bound and a lower bound on the outage probability based on the geometric interpretation. These bounds can work fairly well to approximate the outage probability. When the channel statistics are known to the source and relay, we used numerical methods to obtain the optimal correlation coefficient of the signals transmitted from the source and relay that minimizes the outage probability. For large values of Rice factors, the optimal correlation coefficient is not necessarily zero, but instead, depends on SNRs, Rice factors and variances of the channels, and the target transmission rate.

This dissertation also studies cooperative relay systems from an end-to-end performance perspective. A parallel relay system with differential modulation is considered for the DF protocol or the SR protocol. For the DF protocol, the destination performs ML or PL detections which take account of the average error probabilities at relays and the channel statistics. For a DF single relay system, we have derived the exact BER and its approximation at high SNRs. The exact BER is a simple function of the SNRs, the variances and fade rates of all the transmission links. The BER approximation at high SNR explicitly shows the diversity order and the different effects of the source-relay link and the relay-destination link on the end-to-end error performance. Moreover, for a DF multiple relay system, a Chernoff upper bound on the BER and a high SNR approximation for the BER are obtained. With the same total power constraint, the DF relay with differential modulation outperforms a non-cooperative system which only has a direct link and performs ML detection with

the same differential modulation. For the SR protocol, the complexity is shifted from the destination to the relays, and the relays make decisions on whether to transmit or remain silent according to certain instantaneous error probabilities that also need to be computed at each relay. The ML receiver at the destination is an simple MRC receiver. It is revealed that the SR protocol offers a space diversity order of the number of all the potential nodes, not just the number of nodes that are transmitting to the destination.

6.2 Future Work

6.2.1 Effects of Channel Estimation Errors on Relay Systems

In Chapters 2-4 of this dissertation, it is assumed that the CSI is perfectly known at the corresponding receivers. Usually, the CSI is obtained by channel estimation. In practice, the estimation of CSI at the receivers may not be accurate, although the estimation error can be very small when a large number of pilot signals are sent. In the classic relay system, there are three transmission links, each of which may have a channel estimation error. The effects of these channel estimation errors on the outage probability and the ergodic rate need to be studied. To employ outage-optimal or ergodic-optimal spatial power allocations proposed in Chapter 3, the second-order statistics of all the three transmission links need to be known at the transmitters at the source and relay. It is assumed in this dissertation that the second-order statistics are perfectly known at the transmitters. However, in practice, the statistics of the fading links are estimated by receivers, and then fed back to transmitters. The feedback from receivers to transmitters may have some errors, although the feedback errors can be very small since the statistics of the transmission links vary quite slowly. We can take the both channel estimation errors and the feedback errors into account in designing the outage-optimal and ergodic-optimal transmission strategies.

6.2.2 Extension of the Classic Relay System to a Multiple Relay System

One possible application of the outage-optimal and ergodic-optimal transmission strategies proposed in this study is to use them as design criteria for the transmit signals from the source and relay. The optimal correlation coefficient could be used to find optimal codebooks with this correlation coefficient. Further study is needed to extend the system with one relay to the system with multiple relays by using the same approach in this study.

6.2.3 Decision-Feedback Channel Estimation Receiver

In Chapter 5, we have employed differential modulation in the parallel relay system. A decision-feedback channel estimation receiver [102] may be employed at all the relays and the destination. In this case, the error probability performance may be improved compared to differential modulation.

6.2.4 Optimal Receiver and Error Probability Analysis for Non-coherent AF Relay systems

Because of the multiplicative operation at the relay node in a AF relay system, the noise at the destination is no longer Gaussian distributed for a non-coherent AF relay system. A near-ML detector and a diversity combining detector were proposed in our previous work [103]. The error probability for these receivers in [103] needs to be analyzed.

Bibliography

[1] J. H. Winters, "On the capacity of radio communications with diversity in a Rayleigh fading environment," vol. 05, no. 5, pp. 871–5878, Jun. 1987.

[2] E. Telatar, "Capacity of multi-antennas gaussian channels," *Europ. Trans. on Communications*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1995.

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antanna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[4] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of a discrete-time memoryless Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1290–1301, May. 2001.

[5] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[6] Ö. Oyman, R. U. Nabar, H. Bölskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.

[7] P. B. Rapajic and D. Popesu, "Information capacity of a random signature multiple-input mulitple-output channel," *IEEE Transactions on Communications*, vol. 48, no. 8, pp. 1245–1248, Aug. 2000.

[8] A. Lapidoth and S. Shamai, "Fading channels: How perfect need perfect side information be?" *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1118–1134, May. 2002.

[9] T. Yoo and A. Goldsmith, "Capacity of fading MIMO channels with channel estimation error," in *Proc. of IEEE International Communications Conf. (ICC)*, vol. 2, Jun. 2004, pp. 808–813.

[10] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1102–1111, May. 2005.

[11] X. W. Cui, Q. T. Zhang, and Z. M. Feng, "Generic procedure for tightly bounding the capacity of MIMO correlated Rician fading channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 890–898, May. 2005.

[12] S. Jin, X. Gao, and X. You, "On the ergodic capacity of rank-1 Ricean-fading MIMO channels," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 502–517, Feb. 2007.

[13] M. R. Mckay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.

[14] M. Kang and M. S. Alouini, "Capacity of MIMO Rician channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[15] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading enviornment when using multiple antennas," *Wireless Personal Communnications*, vol. 6, pp. 311–335, 1998.

[16] P. J. Smith and M. Shafi, "On a gaussian approximation to the capacity of wireless MIMO systems," in *Proc. of IEEE International Communications Conf. (ICC)*, New York, NY, Apr.28-May 2 2002, pp. 406–410.

[17] P. J. Smith, L. E. Garth, and S. Loyka, "Exact capacity distributions for MIMO systems with small number of antennas," *IEEE Communications Letters*, vol. 7, no. 10, pp. 481–483, Oct. 2003.

[18] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[19] A. M. Tulino and S. Verdú, *Foundations and Trends in Communications and Information Theory, Random Matrix Theory and Wireless Communications*. USA: now publisher, 2004.

[20] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, no. 4, pp. 657–662, Apr. 2004.

[21] P. J. Smith and L. E. Garth, "Exact capacity distributions for dual MIMO systems in Ricean fading," *IEEE Communications Letters*, vol. 8, no. 1, pp. 18–20, Jan. 2004.

[22] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: performance criterion and code construction,"

IEEE Transactions on Information Theory, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[23] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[24] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[25] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, pp. 41–59, Autumn 1996.

[26] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[27] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[28] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[29] ——, "A survey of multi-way channels in information theory: 1961-1976," *IEEE Transactions on Information Theory*, vol. IT-23, pp. 1–37, Jan. 1977.

[30] T. M. Cover and A. A. E. Gamal, "Capacity theorems for relay channel," *IEEE Transactions on Information Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[31] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, part I: System description," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[32] ——, "User cooperation diversity, part II: Implementation aspects and performance analysis," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[33] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Transactions on Communications*, vol. 52, no. 9, pp. 1470–1476, Sep. 2004.

[34] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[35] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a Rayleigh-fading environment," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1416–1421, Sep. 2004.

[36] D. Chen and J. N. Laneman, "Modulation and demodulation for cooperative diversity in wireless systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 7, pp. 1785–1794, Jul. 2006.

[37] T. Wang, A. Cano, G. B. Giannakis, and J. N. Laneman, "High-performance cooperative demodulation with decode-and-forward relays," *IEEE Transactions on Communications*, vol. 55, no. 7, pp. 1427–1438, Jul. 2007.

[38] I. Hammerström, M. Kuhn, and A. Wittneben, "Channel adapative scheduling for cooperative relay netowrks," in *Proc. of Vehicular Technology Conf. (VTC)*, Los Angeles, CA, Sep. 2004, pp. 2784–2788.

[39] A. G. Agyei and S. L. Kim, "Cross-layer multiservice opportunistic scheduling for wireless networks," *IEEE Commun. Mag.*, vol. 44, no. 6, pp. 50–57, Jun. 2006.

[40] I. Krikidis and J. C. Belfiore, "Scheduling for amplify-and-forward cooperative networks," *IEEE Trans. on Vehicular Tech.*, vol. 56, no. 6, pp. 3780–3790, Nov. 2007.

[41] M. K. Oh, X. Ma, G. B. Giannakis, and D. J. Park, "Cooperative synchronization and channel estimation in wireless sensor networks," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Asilomar, CA, Nov. 2003, pp. 238–242.

[42] D. R. Brown, G. B. Prince, and J. A. McNeill, "A method for carrier frequency and phase synchronization of two autonomous cooperative transmitters," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, New York, NY, Jun. 2005, pp. 260–264.

[43] X. Li, M. Chen, and W. Liu, "Application of STBC-encoded cooperative transmission in wireless sensor networks," *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 134–137, Feb. 2005.

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

[45] A. Bletsas, A. Khisti., D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[46] B. E. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *Proc. of ISIT*, Serrento, Italy, Jun. 2000, p. 22.

[47] M. Gastpar, G. Krammer, and P. Gupta, "The multiple relay channel: Coding and antenna-clustering capacity," in *Proc. of ISIT*, Lausanne, Switzerland, Jun. 2002, p. 136.

[48] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.

[49] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[50] P. Tarasak, H. Minn, and V. K. Bhargava, "Differential modulation for two-user cooperative diversity systems," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1891–1900, Sep. 2005.

[51] T. Wang, Y. Yao, and G. B. Giannakis, "Non-coherent distributed space-time processing for multisuer cooperative transmissions," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3339–3343, Dec. 2006.

[52] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[53] R. U. Nabar, H. Bölcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limite and space-time signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[54] P. Mitran, H. Ochia, and V. Tarokh, "Space-time diversity enhancement using cooperative communications," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2041–2057, Jun. 2005.

[55] K. Azarian, E. Gamal, and P. Philip, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4152–4172, Jun. 2005.

[56] M. O. Hasna and M. S. Alouini, "End-to-end preformance of transmission systems with relays over Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1126–1132, Nov. 2003.

[57] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1820–1830, Oct. 2004.

[58] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1264–1273, May. 2005.

[59] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[60] A. Reznik, S. R. Kulkarni, and S. Verdú, "Degraded gaussian multirelay channel: Capacity and optimal power allocation," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3037–3046, Dec. 2004.

[61] L. L. Xie and P. R. Kumar, "An achievable rate for multipe-level relay channel," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[62] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[63] X. Cai, Y. Yao, and G. B. Giannakis, "Achievable rates in low-power relay links over fading channels," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 184–194, Jan. 2005.

[64] Y. Yao, X. Cai, and G. B. Giannakis, "On energy efficiency and optimum resource allocation in wireless relay transmissions in the low-power regime," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2917–2927, Nov. 2005.

[65] N. Ahmed, M. A. Khojastepour, A. Sabharwal, , and B. Aazhang, "Outage minimization with limitted feedback for the fading relay channel," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 659–669, Apr. 2006.

[66] A. K. Sadek, W. Su, and K. J. R. Liu, "Multinode cooperative communications in wirelss networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 341–355, Jan. 2007.

[67] D. Chen and J. N. Laneman, "Cooperative diversity for wireless fading channels without channel state information," in *Proc. of Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1994.

[68] R. Annavajjala, P. C. Cosman, and L. B. Milstein, "On the performance of optimum noncoherent amplify-and-forward reception for cooperative diversity," in *Proc. of Military Commun. Conf.*, Oct. 2005.

[69] Q. Zhao and H. Li, "Performance of differential modulation with wireless relays in Rayleigh fading channels," *IEEE Communications Letters*, vol. 9, no. 4, pp. 343–345, Apr. 2005.

[70] T. Himsoon, W. Su, and K. J. R. Liu, "Differential transmission for amplify-and-forward cooperative communications," *IEEE Signal Processing Letters*, vol. 12, no. 9, pp. 597–600, Sep. 2005.

[71] W. Cho and L. Yang, "Distributed differential schemes for cooperative wireless networks," in *Proc. of Intl. Conf. on ASSP*, Toulouse, France, May. 2006.

[72] Q. Zhao and H. Li, "Differential modulation for cooperative wireless systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2273–2283, May. 2007.

[73] W. Cho and L. Yang, "Optimum energy allocation for cooperative networks with differential modulation," in *Proc. of Military Commun. Conf. (MILCOM)*, Washington, D. C., Oct. 2006.

[74] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4106–4118, Nov. 2007.

[75] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1195–1206, Jul. 2006.

[76] G. Scutari and S. Barbarossa, "Distributed space-time coding for regenerative relays networks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2387–2399, Sep. 2005.

[77] P. A. Anghel and M. Kaveh, "On the performance of distributed space-time coding systems with one or two non-regenerative relays," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 682–692, Mar. 2006.

[78] P. A. Anghel, G. Leus, and M. Kaveh, "Distributed space-time cooperative systems with regenerative relays," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3130–3141, Nov. 2006.

[79] Y. Jing and B. Hassibi, "Distributed space-time block coding in wireless relay networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[80] G. Wang, Y. Zhang, and M. Amin, "Differential distributed space-time modulation for cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3097–3108, Nov. 2006.

[81] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[82] Y. Zhu, P. Y. Kam, and Y. Xin, "A new approach to the capacity distribution of MIMO Rayleigh fading channel," in *Proc. of GLOBECOM conf.*, San Francisco, CA, USA, Nov. 27 - Dec. 1, 2006.

[83] M. K. Byun and B. G. Lee, "New bounds of pairwise error probability for space-time cdoes in Rayleigh fading," in *Proc. of Wireless Communications and Networking Conf.*, vol. 1, Mar. 2002, pp. 89–93.

[84] S. K. Jayaweera and H. V. Poor, "Performance analysis of decorrelating decision feedback detection for MIMO systems in Rician fading," in *Proc. 6th Intl. Symp. On Wilress Personal Multimedia Coomun. (WPMC)*, vol. 3, Yokusuka, Japan, Oct. 2003, pp. 29–33.

[85] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* John Wiley & Sons, Inc, 1982.

[86] J. Proakis, *Digital Communications.*, 3rd ed. New York: McGraw-Hill, 1995.

[87] M. D. Springer, *The Algebra of Random Variables.* New York: John Wiley & Sons, Inc, 1979.

[88] M. D. Springer and W. E. Thompson, "The distribution of products of Beta, Gamma and Gaussian random variables," *SIAM J. Appl. Math.*, vol. 18, no. 4, pp. 721–737, Jun. 1970.

[89] B. D. Carter and M. D. Springer, "The distribution of products, quotients and powers of independent h-function variates," *SIAM J. Appl. Math.*, vol. 33, no. 4, pp. 542–558, Dec. 1977.

[90] M. O. Hasna and M. S. Alouini, "Optimal power allocation for relayed transmissions over Reyleigh-fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1999–2004, Nov. 2004.

[91] D. Tse and P. Viswanath, *Fundamentals of Wireless Communicaiton*, Cambridge University Press, San Diego, CA, 2005.

[92] G. G. Tziritas, "On the distribution of positive-definite gaussian quadratic forms," *IEEE Transactions on Information Theory*, vol. IT-33, no. 6, pp. 895–906, Nov. 1987.

[93] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*, New York: Marcel Dekker, 1992.

[94] G. L. Turin, "The characteristic function of Hermitian quadratic form in complex normal random variables," *Biometrika*, pp. 199–201, 1960.

[95] H. Böche and E. A. Jorswieck, "On schur-convexity of expectation of weighted sum of random variables with applications," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 5, no. 2, Article 46, 2004.

[96] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, 2nd ed. John Wiley & Sons, Inc., 2004.

[97] Y. Xie, C. N. Georghiades, and A. Arapostathis, "Minimum outage probability transmission with imperfect feedback for MISO fading channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1084–1091, May. 2005.

[98] P. Y. Kam, "Bit error probabilities of MDPSK over the nonselective Rayleigh fading channel with diversity reception," *IEEE Transactions on Communications*, vol. 39, no. 2, pp. 220–224, Feb. 1991.

[99] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication.* Reading, Mass: Addison-Wesley Pub. Co., 1995.

[100] H. Fu and P. Y. Kam, "Performance of optimum and suboptimum combining diversity reception for binary DPSK over independent, nonidentical Rayleigh fading channels," in *Proc. of IEEE International Communications Conf. (ICC)*, vol. 4, Seoul, South Korea, May. 2005, pp. 2367–2371.

[101] M. K. Simon and M. S. Alouini, "A simple single integral representation of the bivariate Rayleigh distribution," *IEEE Communications Letters*, vol. 2, no. 5, pp. 128–130, May. 1998.

[102] Y. Zhu, P. Y. Kam, and Y. Xin, "A decision-feedback channel estimation receiver for independent nonidentical Rayleigh fading channels," in *Proc. of Vehicular Technology Conf. (VTC)*, Singapore, May. 2008.

[103] ——, "Non-coherent detection for amplify-and-forward relay systems in a Rayleigh fading environment," in *Proc. of GLOBECOM conf.*, Washington D. C., Nov. 2007.

[104] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey, *Table of integrals, series, and products.*, 5th ed.

[105] P. Y. Kam, "Tight bounds on Rician-type error probabilities and some applications," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3119–3128, Dec. 1994.

Appendix A

Proof of the Inequality (2.24)

In this appendix, we present the proof of the inequality (2.24)

$$\int_0^\infty x^n \exp\big(-x-\frac{y}{x}\big)dx \geq n!\exp\big(-\frac{y}{n}\big)$$

for $y \geq 0$ and $n = 1, 2, 3, \cdots$.

Proof. Consider the function

$$g(y) = \int_0^\infty x^n \exp\big(-x-\frac{y}{x}+\frac{y}{n}\big)dx. \tag{A.1}$$

The first derivative of $g(y)$ is

$$\begin{aligned} g'(y) &= \frac{1}{n}\exp\big(\frac{y}{n}\big)\Big[\int_0^\infty x^n \exp\big(-x-\frac{y}{x}\big)dx - n\int_0^\infty x^{n-1}\exp\big(-x-\frac{y}{x}\big)dx\Big] \\ &= \frac{1}{n}y^{\frac{n}{2}}e^{\frac{y}{n}}\Big[2\sqrt{y}K_{1+n}(2\sqrt{y}) - 2nK_n(2\sqrt{y})\Big] \end{aligned} \tag{A.2}$$

where $K_n(x)$ is the modified Bessel function of the second kind. By applying the recursion formula for $K_n(x)$ [104, p.982] $xK_{n-1}(x) - xK_{n+1}(x) = -2nK_n(x)$ to (A.2), the first derivative of $g(y)$ can expressed as

$$g'(y) = \frac{2}{n}y^{\frac{n+1}{2}}e^{\frac{y}{n}}K_{n-1}(2\sqrt{y}) \geq 0. \tag{A.3}$$

Because $g'(y) \geq 0$ and $g(0) = n!$, we have $g(y) \geq n!$. Therefore $\exp(-y/n)g(y) \geq n!\exp(-y/n)$, that is

$$\int_0^\infty x^n \exp\big(-x-\frac{y}{x}\big)dx \geq n!\exp\big(-\frac{y}{n}\big)$$

for $y \geq 0$ and $n \in \mathbb{N}$. The inequality holds. □

Appendix B

Proof of Theorem 3.4

Proof. If condition (3.28) is satisfied, we have $\min(\eta_{sd}^2, \eta_{rd}^2)/|\eta_{sd} - \eta_{rd}| > \eta_{sd}\eta_{rd}/\eta_{sr}$. Since $\beta^2/(\alpha - \beta)$ is monotonically increasing in $\mu \in [0, 1]$ and is in the range $\big[0, \min(\eta_{sd}^2, \eta_{rd}^2)/|\eta_{sd} - \eta_{rd}|\big]$, there exists a unique $\mu \in [0, 1]$, namely μ_1, such that $\eta_{sd}\eta_{rd}/\eta_{sr} - \beta^2/(\alpha - \beta) = 0$ (which is equivalent to $\alpha\kappa - \beta = 0$). By solving $\alpha\kappa - \beta = 0$, we obtain μ_1 as

$$\mu_1 = -2\Big(\frac{\eta_{sd} + \eta_{rd}}{\eta_{sr}} + \frac{\eta_{sd}\eta_{rd}}{\eta_{sr}^2}\Big) + \Big(\eta_{sd} + \eta_{rd} + \frac{2\eta_{sd}\eta_{rd}}{\eta_{sr}}\Big) \times \sqrt{\frac{1}{\eta_{sd}\eta_{rd}}\Big(\frac{\eta_{sd} + \eta_{rd}}{\eta_{sr}} + \frac{\eta_{sd}\eta_{rd}}{\eta_{sr}^2}\Big)}. \tag{B.1}$$

It is clear that $\alpha\kappa - \beta \geq 0$ for $\mu \in [0, \mu_1]$, and $\alpha\kappa - \beta < 0$ for $\mu \in (\mu_1, 1]$. According to Lemma 3.2, for any $\mu \in [0, \mu_1]$, $g(\tilde{R}) \geq 0$ when $\tilde{R} \geq 0$. It is thus easy to see that $F_R(\mu, \nu, \tilde{R})$ is an increasing function of $\mu \in [0, \mu_1]$ for any $\tilde{R} \geq 0$. Thus, we now focus on the case of $\mu \in (\mu_1, 1]$.

a) For $\mu \in (\mu_1, 1]$, $\alpha\kappa - \beta < 0$. According to Lemma 3.3, $g(\tilde{R})$ has a unique positive root $\tilde{R}_g$, i.e., $g(\tilde{R}_g) = 0$, but it is not mathematically tractable to obtain an analytical expression of $\tilde{R}_g$. It also should be noted that $\tilde{R}_g$ depends on μ which is actually the parameter we attempt to optimize. Hence, we turn to derive a lower bound of $\tilde{R}_g$, which is independent of μ. To do that, we first check the quadratic equation

$g'(\tilde{R}) = 0$, whose roots can be easily calculated. It can be observed from (3.21) that in this case $g'(\tilde{R})$ has only one positive root $\tilde{R}_s$, such that $g'(\tilde{R}) \geq 0$ for $\tilde{R} \in [0, \tilde{R}_s]$, and otherwise $g'(\tilde{R}) < 0$. It implies $g(\tilde{R}) > g(0) = 0$ for $\tilde{R} \in (0, \tilde{R}_s)$ and $\tilde{R}_s < \tilde{R}_g$. Although $\tilde{R}_s$ still depends on μ, we can readily find a μ-independent lower bound for $\tilde{R}_s$ as follows:

$$\begin{aligned}\tilde{R}_s &> -\frac{(\alpha+\beta)(\xi\kappa+\alpha\beta)}{(\alpha\kappa-\beta)(\beta\kappa+\alpha)} \geq \frac{(\alpha+\beta)(\xi\kappa+\alpha\beta)}{\alpha\beta} \\ &\geq (\alpha+\beta)\Big(\frac{\alpha+\beta}{\eta_{sr}}+1\Big) =: \tilde{R}_0, \end{aligned} \tag{B.2}$$

where the first inequality follows from the fact that $g'(\tilde{R})$ has one negative root, and its positive root $\tilde{R}_s$ must be greater than the sum of its two roots. Thus, we conclude that $\tilde{R}_g > \tilde{R}_s > \tilde{R}_0$. It implies that $F_R(\mu, \nu, \tilde{R})$ is an increasing function of $\mu \in (\mu_1, 1]$ for any $\tilde{R} \in [0, \tilde{R}_0]$. This, along with the previous claim that $F_R(\mu, \nu, \tilde{R})$ is an increasing function of $\mu \in [0, \mu_1]$ for any $\tilde{R} \in [0, \infty)$, concludes that $F_R(\mu, \nu, \tilde{R})$ is an increasing function of $\mu \in [0, 1]$ for all $\tilde{R} \in [0, \tilde{R}_0]$. The optimal value of μ is one for any $R \in [0, R_0]$, where $R_0 = \log_2(1 + \tilde{R}_0)$.

b) For $\mu \in (\mu_1, 1]$ and a fixed $\tilde{R}$, we first investigate the monotonicity of ϕ in (3.18) with respect to μ. Through some long but straightforward calculations, we obtain the derivative of ϕ as

$$\frac{\partial \phi}{\partial \mu} = -\frac{\eta_{sd}\eta_{rd}(\eta_{sd}+\eta_{rd})}{(\alpha-\beta)\big[(\beta\kappa+\alpha)\tilde{R}+\xi\big]^2}\tilde{R}[a(\mu)\tilde{R}+b(\mu)], \tag{B.3}$$

$$\begin{aligned} a(\mu) &= 1 + \frac{(\alpha-\beta)^2}{(\eta_{sr}\mu)^2} + \frac{(\alpha^2+\beta^2)^2}{\eta_{sr}\eta_{sd}\eta_{rd}(\eta_{sd}+\eta_{rd})\mu^2}, \\ b(\mu) &= \frac{(\eta_{sd}+\eta_{rd})^2+\alpha^2+\beta^2}{\eta_{sr}\mu} - \frac{(\eta_{sd}+\eta_{rd})(\alpha^2+\beta^2)}{(\alpha-\beta)^2}. \end{aligned}$$

It is straightforward to check that both $a(\mu)$ and $b(\mu)$ are decreasing functions. Since $a(\mu) > 0$, the ratio $-b(\mu)/a(\mu)$ is increasing in μ, i.e., $-b(\mu)/a(\mu) \leq -b(1)/a(1)$. Define $\tilde{R}_1 := -b(1)/a(1)$, which can be written as (3.25). It is clear from (B.3) that

$\partial\phi/\partial\mu < 0$ for any $\tilde{R} \in [\max\{0, \tilde{R}_1\}, \infty)$. Thus, ϕ is decreasing in $\mu \in (\mu_1, 1]$ for a fixed $r \geq \max\{0, \tilde{R}_1\}$. This, along with the fact that $e^{-\tilde{R}(1/\beta - 1/\alpha)}$ is increasing in μ, implies that $G(\mu, \nu, \tilde{R})$ is a decreasing function of $\mu \in (\mu_1, 1]$ for any $r \geq \max\{0, \tilde{R}_1\}$. Define $\tilde{R}_2$ as in (3.26). Since (3.28) is satisfied, $\tilde{R}_2 > 0$. It is easy to check that $G(1, \nu, \tilde{R}) < 0$ for $\tilde{R} \geq \tilde{R}_2$. Define $\tilde{R}_c := \max\{0, \tilde{R}_1, \tilde{R}_2\} = \max\{\tilde{R}_1, \tilde{R}_2\}$. Since $G(\mu_1, \nu, \tilde{R}) > 0$, $G(1, \nu, \tilde{R}) < 0$, and $G(\mu, \nu, \tilde{R})$ is a decreasing function of $\mu \in (\mu_1, 1]$ for any $\tilde{R} \in [\tilde{R}_c, \infty)$, there exists a unique $\mu_o \in (\mu_1, 1)$ such that $G(\mu_o, \nu, \tilde{R}) = 0$ for a fixed $\tilde{R} \in [\tilde{R}_c, \infty)$. It implies that for a fixed $\tilde{R} \in [\tilde{R}_c, \infty)$, $F_R(\mu, \nu, \tilde{R})$ is increasing in $\mu \in (0, \mu_o)$, and decreasing in $\mu \in (\mu_o, 1]$. Hence, the optimal μ is μ_o. □

Appendix C

Approximation of the Decoding Set Probability $P(r \in \mathcal{D})$

Since the first order Marcum Q-function satisfies the identity $Q_1(a,b) + Q_1(b,a) = 1 + \exp[-(a^2+b^2)/2]I_0(ab)$, we can obtain

$$1 - Q_1(a_{sr}y_{sr}, b_{sr}y_{sr}) + Q_1(b_{sr}y_{sr}, a_{sr}y_{sr}) \\ = 2Q_1(b_{sr}y_{sr}, a_{sr}y_{sr}) - \exp\Big(-\frac{a_{sr}^2+b_{sr}^2}{2}y_{sr}^2\Big)I_0(a_{sr}b_{sr}y_{sr}^2). \quad \text{(C.1)}$$

According to the inequalities [105, Eq.(B.1) and (29a)], we can upper bound the right hand side of (C.1) as

$$2Q_1(b_{sr}y_{sr}, a_{sr}y_{sr}) - \exp\Big(-\frac{a_{sr}^2+b_{sr}^2}{2}y_{sr}^2\Big)I_0(a_{sr}b_{sr}y_{sr}^2) \\ \leq \exp\left\{-2(1-\sin|\theta|)[\text{erfc}^{-1}(2\epsilon)]^2\right\}, \quad \text{(C.2)}$$

where θ satisfies the following equation

$$\frac{1-\cos\theta}{1+\cos\theta} = \frac{a_{sr}^2y_{sr}^2}{b_{sr}^2y_{sr}^2} = \frac{\eta_{sr}^2\rho_{sr}^2}{(\eta_{sr}+1)^2}, \quad \text{(C.3)}$$

and hence we can compute $\sin|\theta|$ as

$$\sin|\theta| = \frac{2(\eta_{sr}+1)\eta_{sr}\rho_{sr}}{(\eta_{sr}+1)^2+\eta_{sr}^2\rho_{sr}^2}. \quad \text{(C.4)}$$

Combining (C.2) and (C.4) into (C.1), we can obtain

$$1 - Q_1(a_{sr}y_{sr}, b_{sr}y_{sr}) + Q_1(b_{sr}y_{sr}, a_{sr}y_{sr}) \\ \leq \exp\left\{-2\frac{(\eta_{sr}+1-\eta_{sr}\rho_{sr})^2}{(\eta_{sr}+1)^2+\eta_{sr}^2\rho_{sr}^2}[\text{erfc}^{-1}(2\epsilon)]^2\right\}. \quad \text{(C.5)}$$

With (C.5) and (5.61), we can develop an upper bound on $P(r \in \mathcal{D})$ as

$$\begin{aligned} P(r \in \mathcal{D}) &\leq \exp\left\{-\left[\frac{1}{w_{sr}\eta_{sr}\rho_{sr}} + 2(N_f-1)\frac{(\eta_{sr}+1-\eta_{sr}\rho_{sr})^2}{(\eta_{sr}+1)^2+\eta_{sr}^2\rho_{sr}^2}\right][\text{erfc}^{-1}(2\epsilon)]^2\right\} \\ &= \left\{-\Big[\frac{1}{w_{sr}\eta_{sr}\rho_{sr}} + 4(N_f-1)\frac{\beta_{sr}^2}{(\alpha_{sr}^2+\beta_{sr}^2)}\Big][\text{erfc}^{-1}(2\epsilon)]^2\right\}. \end{aligned} \tag{C.6}$$

Since $\text{erfc}(x) \leq \exp(-x^2)$, we have $\text{erfc}^{-1}(x) \geq \sqrt{-\ln x}$, and hence, we can the above upper bound be simplified to

$$\begin{aligned} P(r \in \mathcal{D}) &\leq \exp\left\{\Big[\frac{1}{w_{sr}\eta_{sr}\rho_{sr}} + 4(N_f-1)\frac{\beta_{sr}^2}{(\alpha_{sr}^2+\beta_{sr}^2)}\Big]\ln(2\epsilon)\right\} \\ &= (2\epsilon)^{\left[\frac{1}{w_{sr}\eta_{sr}\rho_{sr}} + 4(N_f-1)\frac{\beta_{sr}^2}{(\alpha_{sr}^2+\beta_{sr}^2)}\right]}. \end{aligned} \tag{C.7}$$

List of Publications

1. Y. Zhu, P. Y. Kam, and Y. Xin,"Differential modulation for decode-and-forward multiple relay systems," *IEEE Transaction on Communications*, accepted with minor changes, Aug. 2008.

2. Y. Zhu, P. Y. Kam, and Y. Xin,"On the mutual information distribution of MIMO Rician fading channels," *IEEE Transaction on Communications*, accepted to publish, May. 2008.

3. Y. Zhu, Y. Xin, and P. Y. Kam, "Outage probability of Rician fading relay channels," *IEEE Transaction on Vehicular Technology*, vol. 57, no. 4, pp. 2648–2652, Jul. 2008

4. Y. Zhu, Y. Xin, and P. Y. Kam, "Optimal transmission strategies for Rayleigh fading relay channels," *IEEE Transaction on Wireless Communications*, vol.7, no.2, pp. 618–628, Feb. 2008.

5. Y. Zhu, P. Y. Kam, and Y. Xin, "A decision-feedback channel estimation receiver for independent nonidentical Rayleigh fading channels," in *Proc. IEEE Vehicular Technology Conference (VTC08)*, Singapore, May. 11–14, 2008.

6. Y. Zhu, P. Y. Kam, and Y. Xin, "Non-coherent detection for amplify-and-forward relay systems in a Rayleigh fading environment," in *Proc. IEEE Global Communications Conference (GLOBECOM07)*, Washington D. C., Nov. 26–30, 2007.

7. Y. Zhu, Y. Xin, and P. Y. Kam, "Differential modulation and demodulation for decode-and-forward multiple relay systems," in *Proc. Military Communications Conference (MILCOM07)*, Orlando, F.L., Oct. 29–31, 2007.

8. Y. Zhu, Y. Xin, and P. Y. Kam, "Optimal spatial power allocation for Rayleigh fading relay channels," in *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP07)*, Honolulu, H.A., Apr. 15–20, 2007.

9. Y. Zhu, P. Y. Kam, and Y. Xin, "A new approach to the capacity distribution of MIMO Rayleigh fading channels," in *Proc. IEEE Global Communications Conference (GLOBECOM06)*, San Francisco, C.A., Nov.27–Dec.1, 2006.

10. Y. Zhu, Y. Xin, and P. Y. Kam, "Outage-optimal transmission strategies for Rayleigh fading relay channels," in *Proc. 40th Asilomar Conference on Signals, Systems and Computer*, Pacific Grove, C.A., Oct. 29–Nov. 1, 2006.

11. Y. Zhu, Y. Xin, and P. Y. Kam, "On the outage performance of Rician fading relay channels," in *Proc. IEEE Military Communications Conference (MILCOM06)*, Washington D.C., Oct. 23–25, 2006.

12. Y. Zhu, Y. Xin, and P. Y. Kam, "On the outage performance of single-relay transmissions over quasi-static fading channels," in *Proc. 40th Annual Conference on Information Sciences and Systems (CISS06)*, Princeton University, N.J., Mar. 22–24, 2006.