# Diversity Multiplexing Tradeoff and Capacity Results in Relayed Wireless Networks

by

Shahab Oveis Gharan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Shahab Oveis Gharan

# Abstract

This dissertation studies the diversity multiplexing tradeoff and the capacity of wireless half-duplex multiple-relay network.

In part 1, we study the setup of the parallel Multi-Input Multi-Output (MIMO) relay network. An amplify-and-forward relaying scheme, Incremental Cooperative Beamforming, is introduced and shown to achieve the capacity of the network in the asymptotic case of the number of relays ($K$) goes to infinity with a gap scaling at most as $\frac{\log(K)}{\sqrt[4]{K}}$. This result is shown to hold as long as the power of each relay is significantly larger than $\frac{\log^3(K)\log(\log(K))}{K}$. In addition, two asymptotic Signal to Noise Ratio (SNR) regimes are studied: i) In the regime where the source power is equal to the relays power and both tend to infinity, the proposed scheme is shown to achieve the full multiplexing gain regardless of the number of relays. ii) In the regime where the source power is fixed but the power of each relay tends to infinity, the proposed scheme is shown to asymptotically achieve the network capacity.

In part 2, we study the general setup of multi-antenna multi-hop multiple-relay network. We propose a new scheme, which we call *random sequential* (RS), based on the amplify-and-forward relaying. Furthermore, we derive diversity-multiplexing tradeoff (DMT) of the proposed RS scheme for general single-antenna multiple-relay networks. It is shown that for single-antenna two-hop multiple-access multiple-relay ($K > 1$) networks (without direct link between the source(s) and the destination), the proposed RS scheme achieves the optimum DMT. However, for the case of multiple access single relay setup, we show that the RS scheme reduces to the naive amplify-and-forward relaying and is not optimal in terms of DMT, while the dynamic decode-and-forward scheme is shown to be optimal for this scenario.

In part 3, we characterize the maximum achievable diversity gain of the multi-antenna multi-hop relay network and we show that the proposed RS scheme achieves the maximum diversity gain.

In part 4, RS scheme is utilized to investigate DMT of the general multi-antenna multiple-relay networks. Here, we show that random unitary matrix multiplication at the relay nodes empowers the RS scheme to achieve a better Diversity-Multiplexing Tradeoff (DMT) as compared to the traditional AF relaying. First, we study the case of a multi-antenna full-duplex single-relay two-hop network, for which we show that the RS achieves the optimum DMT. Applying this result, we derive a new achievable DMT for the case of multi-antenna half-duplex parallel relay network. Interestingly, it turns out that the DMT of the RS scheme is optimum for the case of multi-antenna two parallel non-interfering half-duplex relays. Furthermore, we show that random unitary matrix multiplication also improves the DMT of the Non-Orthogonal AF relaying scheme of [26] in the case of a multi-antenna single relay channel. Finally, we study the general case of multi-antenna full-duplex relay networks and derive a new lower-bound on its DMT using the RS scheme.

Finally, in part 5, we study the multiplexing gain of the general multi-antenna multiple-relay networks. We prove that the traditional amplify-forward relaying achieves the maximum multiplexing gain of the network. Furthermore, we show that the maximum multiplexing gain of the network is equal to the minimum vertex cut-set of the underlying graph of the network, which can be computed in polynomial time in terms of the number of network nodes. Finally, the argument is extended to the multicast and multi-access scenarios[1].

[1]The materials of this work are also reported in [46, 47, 51, 48, 49, 50, 52, 53].

**Acknowledgements**

I would like to express my deep and sincere gratitude to my supervisor Professor Amir Khandani for his end-less support, understanding, kindness, and great supervision. During my PhD studies in Waterloo, I learnt a lot from him not only about research but also about life.

I would like to specially acknowledge my dear friend Dr. Alireza Bayesteh for his collaboration on this research and co-authorship of several papers, and also for his friendship and constant support. I also would like to thank Dr. Abolfazl Motahari, Shervan Fashandi, Dr. Hamid Farmanbar and Saeed Changiz Rezaie for the joint work on other interesting research topics such as Interference Alignment, Path Diversity in Overlay Networks, Code Design for Channels with State, and Parallel Relay Channels.

Finally, I would like to extend my deepest gratitude, love and affection to my beloved parents, Fatemeh and Mansour, and my sisters and brothers, Sheida, Shahram, Shadi, Shayan, who supported me throughout the long path of my academic endeavour and endured countless sacrifices so that I can have the absolute best education. To them, I owe all I have ever accomplished.

Now that this chapter of my life is ending, I am fortunate that this chapter ended with sweet memories of a special person. During the last one and half years of my PhD, I was so lucky and fortunate to know and fall in love with my beloved Kianoosh, and, finally, marry her.

April 2010,
Waterloo, ON

*To My Beloved Kianoosh*

# Contents

# List of Tables

# List of Figures

# List of Notations and Abbreviations

| | |
|---|---|
| MIMO | Multiple-Input Multiple-Output |
| MISO | Multiple-Input Single-Output |
| SNR | Signal to Noise Ratio |
| CSI | Channel State Information |
| CDF | Cumulative Distribution Function |
| pdf | Probability Density Function |
| AWGN | Additive White Gaussian Noise |
| Uplink (Broadcast) Channel | The channel between the source and the relays in parallel relay network |
| Downlink(Multi-Access) Channel | The channel between the relays and the destination in parallel relay network |
| DMT | Diversity-Multiplexing Tradeoff |
| CF | Compress-and-Forward |
| DF | Decode-and-Forward |
| AF | Amplify-and-forward |
| ICBS | Incremental Cooperative Beamforming Scheme |
| RS | Random Sequential |
| NAF | Non-orthogonal Amplify-and-Forward |
| SAF | Slotted Amplify-and-Forward |
| Boldface Upper-Case Letters | Matrices |
| Boldface Lower-Case Letters | Vectors |
| $\mathbf{H}_i$ | The channel between the source and the $i$'th relay in the multi-antenna parallel relay network |

| | |
|---|---|
| $\mathbf{G}_i$ | The channel between the $i$'th relay and the destination in the multi-antenna parallel relay network |
| $\mathbf{h}_i$ | The channel between the source and the $i$'th relay in the single-antenna parallel relay network |
| $\mathbf{g}_i$ | The channel between the $i$'th relay and the destination in the single-antenna parallel relay network |
| $\mathbf{H}_{i,j}$ | The channel from node $j$ to node $i$ in the general multi-antenna wireless relay network |
| $\mathbf{h}_{i,j}$ | The channel from node $j$ to node $i$ in the general single-antenna wireless relay network |
| $N_i$ | The number of antennas at the node $i$ in the general wireless multi-antenna relay network |
| $\|\mathbf{h}\|$ | Norm of vector $\mathbf{h}$ |
| $\|\mathbf{H}\|$ | Frobenius norm of matrix $\mathbf{H}$ |
| $\|\mathbf{H}\|_\star$ | Maximum absolute value among entries of $\mathbf{H}$ |
| $\lvert\mathbf{H}\rvert$ | Determinant of $\mathbf{H}$ |
| $\mathrm{Tr}(\mathbf{H})$ | Trace of $\mathbf{H}$ |
| $(\mathbf{H})^H$ | Transpose conjugate of $\mathbf{H}$ |
| $(\mathbf{H})^T$ | Transpose of $\mathbf{H}$ |
| $(\mathbf{H})^*$ | Conjugate of $\mathbf{H}$ |
| $(\mathbf{H})^\dagger$ | Pseudo inverse of $\mathbf{H}$ |
| $\mathbf{H}_\mathcal{A}$ for $\mathbf{H} = [\mathbf{H}_1 \vert \dots \vert \mathbf{H}_J]^T$ and $\mathcal{A} \subseteq \{1, 2, \dots, J\}$ | $[\mathbf{H}_k \vert\, k \in \mathcal{A}]$ |
| $\mathbf{H} \succeq 0$ | Matrix $\mathbf{H}$ is positive semi-definite |
| $\mathbf{H} \succeq \mathbf{G}$ | $\mathbf{H} - \mathbf{G} \succeq 0$ |

| | |
|---|---|
| $\mathbf{1}_n$ | $n$-dimensional vector with all entries equal to one |
| $\mathbf{I}_n$ | $n$-dimensional identity matrix |
| $\mathbf{0}$ | The vector of all zeros |
| $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ | Circularly symmetric Gaussian distribution with zero mean and unit variance |
| $G = (V, E)$ | Graph $G$ with the set of nodes $V$ and the set of edges $E$ |
| Edge cut-set | Refer to Definition 3.1 |
| Vertex cut-set | Refer to Definition 6.1 |
| $\mathcal{S}^c$ | Complement of the set $\mathcal{S}$ |
| $\lvert\mathcal{S}\rvert$ | Cardinality of the set $\mathcal{S}$ |
| $x^+$ | $\max(x, 0)$ |
| $\ln(.)$ | Natural logarithm |
| $\log(.)$ | base-2 logarithm |
| $\mathbb{E}\{.\}$ | Expectation operation |
| $\mathbb{P}\{\mathcal{B}\}$ | Probability of event $\mathcal{B}$ |
| $\mathbb{RH}(.)$ | Right hand side of the equations |
| $K$ | Number of relays |
| $P$ | Total transmit power at every node |
| $P_s$ | Total transmit power at the source (in the case that $P_s \neq P$) |
| $P_r$ | Total transmit power at the relay (in the case that $P_r \neq P$) |
| $C_{\text{ub}}$ | Upper-bound on the capacity |
| $\Gamma(.)$ | Gamma function |

| | |
|---|---|
| $f(n) = o(g(n))$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ |
| $f(n) = O(g(n))$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} < \infty$ |
| $f(n) = \omega(g(n))$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$ |
| $f(n) = \Omega(g(n))$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} > 0$ |
| $f(n) = \Theta(g(n))$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} = c$, where $0 < c < \infty$ |
| $f(n) \sim g(n)$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$ |
| $f(n) \gtrsim g(n)$ | $\lim_{n\to\infty} \frac{f(n)}{g(n)} \geq 1$ |
| $A \approx B$ | approximate equality between $A$ and $B$, such that if one replaces $A$ by $B$ in the equations, the results still hold. |
| $f(P) \doteq g(P)$ | $\lim_{P\to\infty} \frac{f(P)}{\log(P)} = \lim_{P\to\infty} \frac{g(P)}{\log(P)}$ |
| $f(P) \dot{\leq} g(P)$ | $\lim_{P\to\infty} \frac{f(P)}{\log(P)} \leq \lim_{P\to\infty} \frac{g(P)}{\log(P)}$ |

# Chapter 1

# Introduction

## 1.1 Motivation

In recent years, relay-assisted transmission, which was first introduced by van der Meulen in 1971 [15], has gained significant attention as a powerful technique to enhance the performance of wireless networks, combat the fading effect, extend the coverage, and reduce the amount of interference due to frequency reuse. The main idea is to deploy some extra nodes in the network to facilitate the communication between the end terminals. In this manner, these supplementary nodes act as spatially distributed antennas for the end terminals.

Multiple-antennas transmission is another technique that has been significantly investigated in recent years. Information-theoretic results have shown that MIMO systems have the ability to simultaneously enhance the multiplexing gain (degrees of freedom) and the diversity (reliability) in point-to-point wireless fading links [24, 17, 29]. Motivated by these results on point-to-point MIMO channels, subsequently there has been a growing interest in studying the application of

MIMO systems in more complex wireless networks. Specifically, some promising results are reported regarding MIMO multiple-access and broadcast channels in [36, 58, 42, 21, 13]. However, even in its simplest form, the capacity of a single-antenna Gaussian relay channel is still unknown. Due to this complexity, in spite of all the attention to MIMO systems, there are only a few results known concerning MIMO relay networks.

More recently, cooperative diversity techniques have been proposed as candidates to exploit the spatial diversity offered by the relay nodes (for example, see [25, 26, 34, 1]). A fundamental measure to evaluate the performance of the existing cooperative diversity schemes is the diversity-multiplexing tradeoff (DMT) which was first introduced by Zheng and Tse in the context of point-to-point MIMO fading channels [29]. Roughly speaking, the diversity-multiplexing tradeoff identifies the optimal compromise between the "transmission reliability" and the"data rate" in the high-SNR regime.

## 1.2 General Relay Networks

The classical relay channel was first introduced by Van der Meulen in 1971 [15]. In [15], a node defined as the relay enhances the transmission of information between the source and the destination. The most important subsequent results have been published by Cover and El Gamal [56]. In [56], two different coding strategies are introduced. In the first strategy, originally named "cooperation", and later known as "decode-and-forward" (DF), the relay decodes the transmitted message and cooperates with the source to send the message in the next block. In the second strategy, known as "compress-and-forward" (CF), the relay compresses the received signal and sends it to the destination. The performance of the DF strategy is limited by the quality of the source-to-relay channel, while CF's performance is mostly governed by the quality of the relay-to-destination channel [56]. The drawback of using CF strategy is that the transmitted signals by the source and the relay are independent of each other. Hence, the CF strategy is unable to exploit the power boosting advantage due to the coherent addition of the signals in the Gaussian relay channel [56].

Recently, extensions of the relay channel to the multiple relays or to multiple source/destinations have been investigated in [10, 44, 31, 32, 30, 18, 6, 8]. [18, 30] develop new coding schemes based on Decode-and-Forward and Compress-and-Forward relaying strategies for multiple-relay networks. Avestimehr *et al.* in [6] present a linear deterministic model for the wireless relay network and characterize its exact capacity. Applying the capacity-achieving scheme of the corresponding deterministic model, the authors in [6] show that the capacity of wireless single-relay channel and the diamond relay channel can be characterized within 1 bit and 2 bits, respectively, regardless of the values of the channel gains. Furthermore, in [8], the authors show that a variant of the CF relaying achieves the capacity of any general single-antenna Gaussian relay network within a constant bit number that only depends on the number of nodes in the network. The authors show in [9] that the result is still valid for both the multi-antenna Gaussian and the multi-antenna ergodic Rayleigh fading relay networks.

## 1.3 Parallel Relay Networks

A special case of the multiple-relay network is the parallel relay network, which was first introduced by Schein in [10,44]. In the set up of [10,44], the source broadcasts its data to two relays and then, the relays transmit their data in a coherent manner to the destination. Hence, the communication is performed in two hops. [44] studies the symmetric Gaussian parallel single-antenna relay network. Besides investigating the well-known "compress-and-forward" and "decode-and-forward" strategies, [44] have also studied the "amplify-and-forward" (AF) strategy. In AF relaying, the relay(s) simply amplify and transmit their received data to the destination. Despite its simplicity, in many scenarios AF relaying performs very well. Indeed, [44] shows that the time-sharing between AF and DF achieves the best lower-bound on the capacity in all SNR regimes. Furthermore, for the scenario where either the SNR value at the downlink (channel between the relays and the destination) is large or the SNR values at both the downlink and the uplink (channel between the source and the relays) are low, the time-sharing between AF and DF reduces to naive AF relaying and achieves the network capacity, asymptotically. Moreover, Gastpar in [31] proves that employing AF relaying achieves the capacity of the Gaussian parallel single-antenna relay network as the number of relays tends to infinity. AF relaying is also investigated in [43, 2, 19, 39, 20] for different Gaussian relay network scenarios.

Recently, Bolcskei *et al.* in [19, 39] extend the work of [31] to the parallel multiple-antenna parallel relay network. Unlike the parallel single-antenna relay scenario, in this case the AF multipliers are matrices rather than scalars. Hence, finding the optimum AF matrices becomes challenging. The authors in [19] study two scenarios in terms of the CSI assumption at the relays: i) coherent relaying, in which the relays have perfect CSI about their own forward channel (channel between the relay and the destination) and backward channel (channel between the source and the relay), and ii) non-coherent relaying, in which the relays have no CSI. In both scenarios, it is assumed that the source has no CSI while the destination has perfect CSI about the equivalent end-to-end channel between the source and the destination. In the coherent scenario, the authors propose a new AF scheme, called "matched filtering", and prove that the achievable rate of the proposed scheme follows the capacity of the parallel MIMO relay network with a constant gap in terms of the number of relays, $K$, in the asymptotic case of $K \to \infty$. They also show that this achievable rate grows linearly with the number of transmit antennas (reflecting the multiplexing gain [1]) and grows logarithmically in terms of the number of relays (reflecting the distributed array gain [19]). Moreover, in the non-coherent scenario, the simple AF relaying achieves a rate which grows linearly with the number of transmit antennas, but does not grow with the number of relays.

Shi *et al.* [20] present a new AF relaying scheme for the parallel MIMO relay network using the QR decomposition of the forward and backward channels in each relay. Relying on numerical results, reference [20] shows that their proposed scheme outperforms other existing AF schemes for practical number of relays and practical number of antennas.

[1]Note that, however, the multiplexing gain of the "matched filtering" scheme is zero for any arbitrary number of relays.

## 1.4 DMT in Relay Networks

The DMT of relay networks was first studied by Laneman *et al.* in [25] for half-duplex relays. In this work, the authors prove that the DMT of a network with single-antenna nodes, composed of a single source and a single destination assisted with $K$ half-duplex relays, is upper-bounded by

$$d(r) = (K+1)(1-r)^{+}. \tag{1.1}$$

This result can be established by applying either the multiple-access or the broadcast cut-set bound [12] on the achievable rate of the system. Despite its simplicity, this bound is yet the tightest on DMT[2]. The authors in [25] also suggest two protocols based on decode-and-forward (DF) and amplify-and-forward (AF) strategies for a single-relay system with single-antenna nodes. In both protocols, the relay listens to the source during the first half of the frame, and transmits during the second half. To improve the spectral efficiency, the authors propose an incremental relaying protocol in which the receiver sends a single bit feedback to the transmitter and to the relay to clarify if it has decoded the transmitter's message or needs help from the relay for this purpose. However, none of the proposed schemes are able to achieve the DMT upper-bound.

The non-orthogonal amplify-and-forward (NAF) scheme, first proposed by Nabar *et al.* in [38], has been further studied by Azarian *at al.* in [26]. In addition to analyzing DMT of the NAF scheme, reference [26] shows that NAF is the best in the class of AF strategies for single-antenna single-relay systems. The dynamic decode-and-forward (DDF) scheme has been proposed independently in [26,35,33]

[2]Indeed, as we will explain in the sequel, Pawar *et al.* in their recent work [41] have shown that (1.1) is the exact DMT for the single half-duplex relay channel.

based on the DF strategy. In DDF, the relay node listens to the sender until it can decode the message, and then re-encodes and forwards it to the receiver in the remaining time. Reference [26] analyzes the DMT of the DDF scheme and shows that it is optimal for low rates in the sense that it achieves (1.1) for the multiplexing gains satisfying $r \leq 0.5$. However, for higher rates, the relay should listen to the transmitter for most of the time, reducing the spectral efficiency. Hence, the scheme is unable to follow the upper-bound for high multiplexing gains. More importantly, the generalizations of NAF and DDF for multiple-relay systems fall far from the upper-bound, especially for high multiplexing gains.

Yuksel *et al.* in [34] apply compress-and-forward (CF) strategy and show that CF achieves the DMT upper-bound for multiple-antenna half-duplex single-relay systems. However, in their proposed scheme, the relay node needs to know the CSI of all the channels in the network which may not be practical.

More recently, Yang *et al.* in [55] propose a class of AF relaying scheme called slotted amplify-and-forward (SAF) for the case of half-duplex multiple-relay ($K > 1$) and single source/destination setup. In SAF, the transmission frame is divided into $M$ equal length slots. In each slot, each relay transmits a linear combination of the previous slots. Reference [55] presents an upper-bound on the DMT of SAF and shows that it is impossible to achieve the MISO upper-bound for finite values of $M$, even with the assumption of full-duplex relaying. However, as $M$ goes to infinity, the upper-bound meets the MISO upper-bound. Motivated by this upper-bound, the authors in [55] propose a half-duplex sequential SAF scheme. In the sequential SAF scheme, following the first slot, in each subsequent slot, one and only one of the relays is permitted to transmit an amplified version of the signal it has received in the previous slot. By doing this, the different parts of the signal are transmitted through different paths by different relays, resulting in some form of spatial diversity. However, [55] could only show that the sequential SAF achieves the MISO upper-bound for the setup of non-interfering relays, i.e. when the consecutive relays (ordered by transmission times) do not cause any interference on one another.

Yang and Belfiore in [54] study the DMT performance of the NAF scheme for the multi-antenna parallel relay setup. Moreover, based on the non-vanishing determinant criterion, the authors constructed a family of space-time codes for the NAF scheme over multi-antenna channels. However, as shown in [54], the NAF scheme falls far from the DMT upper-bound in the multiple-antenna setup, particularly for small values of multiplexing gain. Indeed, even for the case of a multi-antenna two-hop single-relay setup, the NAF scheme is unable to achieve the maximum diversity gain of the system.

Yuksel *et al.* in [34] apply CF strategy and show that CF achieves the DMT upper-bound for multi-antenna half-duplex single-relay networks. However, in their proposed scheme, the relay node needs to know the Channel State Information (CSI) of all the channels in the network, which may not be practical.

## 1.5 Achievable Rate of AF Relaying

Among the different relaying strategies, AF relaying turns out to be more suitable in practice. Indeed, in AF relaying the relays are not supposed to decode the transmitted message. Instead, they simply forward their observation of the last time-slot. Hence, the relays consume less computing power. Moreover, the end-to-end system expends a much smaller amount of delay compared with the other relaying strategies, as the relays do not need to wait a couple of time-slots in order to decode the source message or compress the received vector. Another advantage of the AF relaying is that the relay nodes do not need to have any knowledge of the codebook the source is using.

The AF relaying is mainly investigated in literature in order to exploit the cooperative diversity for the wireless relay networks (for example, see [25, 26, 55, 54, 51, 48, 49, 28]). However, the achievable rate of the AF relaying is unknown for general wireless relay networks. Indeed, [8] has shown that there exists scenarios for which the gap between the achievable rate of AF relaying and the capacity of the Gaussian relay network can be arbitrarily large.

Most recently, Avestimehr *et al.* in [8] show that a variant of the CF relaying achieves the capacity of any general single-antenna Gaussian relay network within a constant bit number that only depends on the number of nodes in the network. Furthermore, the authors show in [9] that the result is still valid for both the multi-antenna Gaussian and the multi-antenna ergodic Rayleigh fading relay networks. For the case of the relay network with nodes which are equipped with multi-antenna, the gap is only related to the summation of the number of antennas of all network nodes. Also, by relating the original problem to the linear deterministic network and applying the result of [7], the authors of [27] show that the maximum multiplexing gain of the wireless relay networks is equal to the minimum between the matrix rank corresponding to different cut-sets of the underlying graph of the network. However, the scheme of [27] also has the drawbacks of CF relaying: each relay node listens for $T$ time-slots ($T$ should approach infinity such that the argument is valid) and then multiplies the received vector by a predefined matrix of size $NT \times NT$, where $N$ is the maximum number of antennas among all nodes of the network and sends the result in the following $T$ time-slots. Hence, the scheme requires high computing power consumption at the relay nodes and imposes a large delay to the end-to-end network.

In another work [23], the authors show that cooperative communication, either between the transmitters or between the receivers, is unable to improve the multiplexing gain of a wireless $2 \times 2$ single-antenna interference channel.

## 1.6 Summary of Dissertation

In Chapter 2, we consider the parallel MIMO relay network. The network studied in this chapter consists of $K$ relays each equipped with $N$ antennas assist in data transmission between a source and a destination, each equipped with $M$ antennas ($N \geq M$). Communication takes place in two equal-time hops and the relays operate in the half-duplex mode. We propose a new AF protocol called "Incremental Cooperative Beamforming Scheme" (ICBS). We prove that the achievable rate of ICBS converges to the capacity of the parallel MIMO relay network, for asymptotically large number of relays, with a gap which vanishes to zero. Next, we study the performance of ICBS in two asymptotically high SNR regmies: i) In the regime where the power of both the source and the relays approaches infinity, we prove that ICBS achieves the full multiplexing gain; and ii) In the regime where the power of the source is fixed, but the power of each relay approaches infinity, we show that the gap between the achievable rate of ICBS and the capacity vanishes to zero. Finally, through simulation, we compare the achievable rate of ICBS against the achievable rate of "matched-filtering" scheme of [19] and the upper-bound on capacity obtained from the point-to-point capacity of the broadcast channel. Simulation results show that while the gap between "matched-filtering" and the upper-bound on capacity remain constant for different number of relays, the achievable rate of ICBS rapidly achieves the upper-bound capacity.

In Chapter 3, we study DMT in single-antenna multiple-relay networks. Here, we propose a new scheme, which we call random sequential (RS), based on the SAF relaying for general multiple-antenna multi-hop networks. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the source of the future paths on the destination of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in amplify-and-forward relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. We derive DMT of the RS scheme for general single-antenna multiple-relay networks (maximum diversity and DMT of RS scheme is investigated in the following two Chapters). Specifically, we derive: 1) the exact DMT of the RS scheme under the condition of "non-interfering relaying", and 2) a lower-bound on the DMT of the RS scheme (no conditions imposed). Finally, we prove that for single-antenna multiple-access multiple-relay networks (with $K > 1$ relays) when there is no direct link between the transmitters and the receiver and all the relays are connected to the transmitter and to the receiver, the RS scheme achieves the optimum DMT. However, for two-hop multiple-access single-relay networks, we show that the proposed scheme is unable to achieve the optimum DMT, while the DDF scheme is shown to perform optimum in this scenario.

In Chapter 4, we investigate the maximum diversity gain for the general multi-hop multi-antenna wireless relay network which is introduced in Chapter 3. We show that the proposed RS scheme achieves the maximum diversity gain of the network. Furthermore, we characterize the maximum achievable diversity gain in terms of the minimum edge cut-set of the underlying graph of the network.

In Chapter 5, we investigate DMT of AF relaying in "multi-antenna" multi-relay networks. For this purpose, we study the application of the RS scheme described in Chapter 3. First, we study the simple structure of multi-antenna full-duplex two-hop single-relay network. We show that unlike the traditional AF relaying, the RS scheme achieves the optimum DMT. Indeed, random unitary matrix multiplication empowers the RS scheme to achieve the optimum DMT. This fact will be elaborated throughout Chapter 5. Furthermore, we generalize this result to the multi-hop multi-antenna relay networks with single-relay in each hop. Next, we study the case of multi-antenna half-duplex parallel relay network and, by deriving its DMT, we show that the RS scheme improves the DMT of the traditional AF relaying scheme. Interestingly, it turns out that the DMT of the RS scheme is optimum for the multi-antenna half-duplex parallel two-relay ($K = 2$) setup with no direct link between the relays. We also show that utilizing random unitary matrix multiplication improves the DMT of the NAF relaying scheme of [26] in the case of a multi-antenna single relay channel. Finally, we study the class of general full-duplex multi-antenna relay networks whose underlying graph is *directed acyclic* and all nodes are equipped with the same number of antennas. Using the RS scheme, we derive a new lower-bound for the achievable DMT of this class of networks. It turns out that the new DMT lower-bound meets the optimum DMT at the corner points, corresponding to the maximum multiplexing gain and the maximum diversity gain of the network, respectively.

In Chapter 6, we study the achievable rate of the traditional AF relaying in the high SNR scenarios for general wireless multiple-antenna multiple-relay networks. The channel model for this Chapter is the same as the ones used in Chapters 3, 4 and 5, meaning that every two nodes are either connected through a Rayleigh fading channel or disconnected. Unlike the RS scheme which utilizes matrix multiplication and a complex scheduling for the relays transmission, in traditional AF relaying, each relay node forwards its received signal of the last time-slot in the following time-slot. No channel state knowledge is required at either the source or any of the relay nodes. However, the destination is assumed to know the end-to-end channel state. We study the pre-log coefficient of the ergodic capacity in high SNR regime, known as the multiplexing gain. We prove that the traditional AF relaying achieves the maximum multiplexing gain for any wireless multi-antenna relay network. Furthermore, we characterize the maximum multiplexing gain of the network in terms of the minimum vertex cut-set of the underlying graph of the network and show that it can be computed in polynomial-time (with respect to the number of network nodes) using the maximum-flow algorithm. Finally, we show that the argument can be easily extended to the *multicast* and *multi-access* scenarios as well.

Chapter 7 contains conclusions and directions for future research.

# Chapter 2

# Parallel MIMO Relay Networks

## 2.1 Introduction

In this chapter, we consider a parallel MIMO relay network, in which $K$ relays each equipped with $N$ antennas assist in data transmission between a source and a destination, each equipped with $M$ antennas ($N \geq M$). It is assumed that there is no direct link between the source and the destination. All the channels are assumed to be block Rayleigh fading and all the nodes in the network are assumed to be fully aware of their corresponding channels. Communication takes place in two equal-time hops and the relays operate in the half-duplex mode, i.e., the relays can not transmit and receive simultaneously.

We propose a new AF protocol called "Incremental Cooperative Beamforming Scheme" (ICBS). In this scheme, considering the uplink channel (from the source to all the relays) as a point-to-point channel, the relays cooperatively multiply the uplink channel matrix with a beamforming matrix computed from the Singular Value Decomposition (SVD) of this channel. Interestingly, to perform such an operation, each relay only needs to know its corresponding sub-matrix of the beamforming matrix. Moreover, for the outputs to be coherently added at the destination, each relay has to apply zero forcing beamforming to its corresponding downlink channel. However, the downlink noise degrades the performance of ICBS significantly if any of the relays' downlink channels is ill-conditioned. To enhance the performance of ICBS in such scenarios, a threshold parameter is introduced and the relays with ill-conditioned downlink channels are turned off. This strategy improves the overall point-to-point channel from the source to the destination. However, some diagonal and non-diagonal terms are subtracted from the equivalent end-to-end channel matrix due to turning some of the relays off. Hence, the off relays can potentially increase the *orthogonality defect* of the end-to-end channel matrix.

It is shown that for asymptotically large number of relays, one can simultaneously mitigate the downlink noise and the *orthogonality defect* due to the turned-off relays. As a result, the achievable rate of ICBS converges to the capacity of parallel MIMO relay network with a gap scaling as $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$. This result is stronger than the result of [19] in the coherent scenario, in which they show that their proposed scheme can asymptotically ($K \to \infty$) achieve the capacity up to a gap of $O(1)$. Furthermore, we prove that the asymptotic capacity of the parallel relay network remains the same as long as the power of relays scales as $\frac{\log^3(K)\log(\log(K))}{K}$, and the capacity is also achievable by ICBS. Finally, we investigate the achievable rate of ICBS in two high SNR regime scenarios. Numerical results show that the achievable rate of ICBS converges rapidly to the capacity, even for moderate number of relays.

The main contributions of the chapter are as follows:

- We derive the asymptotic capacity ($K \to \infty$) of the parallel MIMO relay network and show that the achievable rate of ICBS converges to the capacity with a gap which vanishes as $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$.
- We show that the same result can be achieved by ICBS, as long as the power of each relay scales as $\omega\left(\frac{\log^3(K)\log(\log(K))}{K}\right)$.
- ICBS is proved to achieve the full multiplexing gain, regardless of the number of relays.
- In the regime where the source power is fixed and the power of each relay tends to infinity, ICBS is shown to achieve the network capacity, asymptotically.

The rest of the chapter is organized as follows. In Section 2.2, the system model is introduced and the assumptions are stated. In Section 2.3, the proposed scheme is described. Section 2.4 is dedicated to the asymptotic analysis of the proposed schemes. Simulation results are presented in Section 2.5. Finally, Section 2.6 concludes the chapter.

## 2.2 System Model and Assumptions

### 2.2.1 System Model

The system model in this chapter, as in [19], [39] and [20], is a parallel MIMO relay network with two-hop relaying and half-dulplexing between the uplink and downlink channels. In other words, data transmission is performed in two equal-length time slots; in the first time slot, the signal is transmitted from the source to the relays, and in the second time slot, the relays transmit data to the destination. Note that there is no direct link between the source and the destination in this model. The source and the destination are equipped with $M$ antennas and each of the relays is equipped with $N$ antennas. Throughout this chapter, we assume that $N \geq M$. The channel between the source and the relays and the channel between the relays and the destination are assumed to be quasi-static Rayleigh fading. The channel from the source to the $k$th relay, $1 \leq k \leq K$, is modeled as

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k, \tag{2.1}$$

and the downlink channel is modeled as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{G}_k \mathbf{t}_k + \mathbf{z}, \tag{2.2}$$

where the channel matrices $\mathbf{H}_k$ and $\mathbf{G}_k$ are $N \times M$ and $M \times N$ i.i.d. complex Gaussian matrices with zero mean and unit variance, respectively, $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ and $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ are Additive White Gaussian Noise (AWGN) vectors, $\mathbf{r}_k$ and $\mathbf{t}_k$ are the $k$th relay's received and transmitted signal, respectively, and $\mathbf{x}$ and $\mathbf{y}$ are the source and the destination signal, respectively (figure 2.1).

The power constraints $\mathbb{E}[\mathbf{x}^H \mathbf{x}] \leq P_s$ and $\mathbb{E}_{\mathbf{x},\mathbf{n}_k}[\mathbf{t}_k^H \mathbf{t}_k] \leq P_r$ must be satisfied for the transmitted signals of the source and the relays, respectively[1]. We assume $P_r = P_s = P$ throughout this chapter, except in Theorem 2.8, where we study the case $P_r < P_s = P$, and Theorem 2.10, where we study the case $P_r \gg P_s = P$.

The task of amplify and forward (AF) relaying is to find the matrix $\mathbf{F}_k$ for each relay to be multiplied by its received signal to form the relay's output as $\mathbf{t}_k = \mathbf{F}_k \mathbf{r}_k$, such that the power constraint is satisfied for each relay. In this way, the entire

[1]Note that in the proposed ICBS, we have assumed that the power constraint for the transmitter and the relays is satisfied for *each* realization of the channels. However, for deriving the upper-bound on the capacity in Theorem 2.1, we assume an *average* power constraint (over all realizations) for the transmitter, i.e., $\mathbb{E}_{\mathbf{H}}[\mathbf{x}^H \mathbf{x}] \leq P_s$, which is a stronger result.

source-destination channel is modeled as

$$\mathbf{y} = \left( \sum_{k=1}^{K} \mathbf{G}_k \mathbf{F}_k \mathbf{H}_k \right) \mathbf{x} + \sum_{k=1}^{K} \mathbf{G}_k \mathbf{F}_k \mathbf{n}_k + \mathbf{z}. \tag{2.3}$$

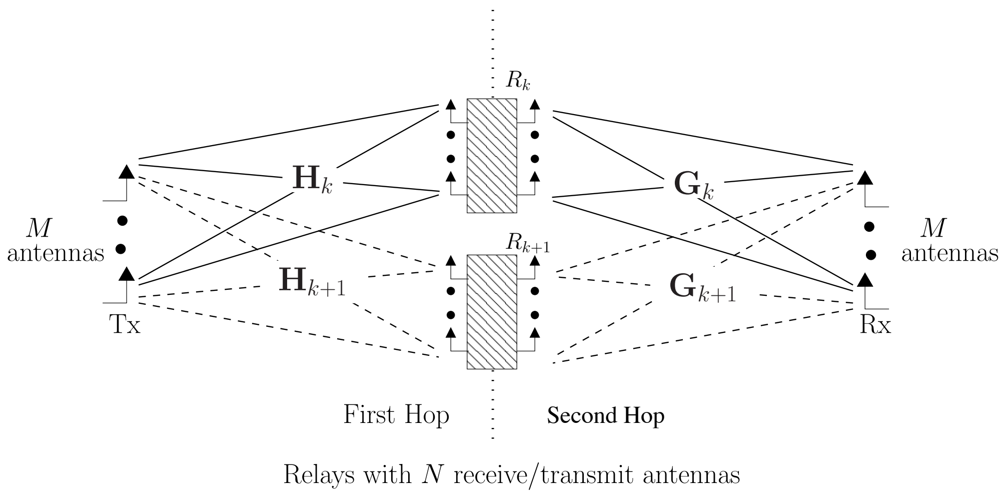


Figure 2.1: A schematic of a parallel MIMO relay network

### 2.2.2 CSI Assumptions

For the proposed ICBS, it is assumed that the source knows the uplink channel, i.e. $\mathbf{H}_1, \cdots, \mathbf{H}_K$. The source's knowledge about the uplink channel can be realized either through the feedback from the relays or by measuring the uplink channel directly assuming reciprocity of the uplink and its reverse channel. Furthermore, the source is assumed to send a $N \times M$ matrix containing the the channel information to each of the relays. This assumption is reasonable when the uplink channel variation is slow such that the transmitter has enough time and bandwidth resources to send the channel information to the relays. Furthermore, we assume that each relay knows its forward channel, i.e. $\mathbf{G}_k$. Finally, it is assumed that the destination has perfect knowledge about the equivalent point-to-point channel from the source to the destination. This information can be obtained through sending pilot signals by the source, amplified and forwarded at the relay nodes in the same manner as the information signal.

## 2.3 Proposed Method

The equivalent uplink channel can be represented as $\mathbf{H}^T \triangleq \left[\mathbf{H}_1^T|\mathbf{H}_2^T|\cdots|\mathbf{H}_K^T\right]^T$. By applying Singular Value Decomposition (SVD) to $\mathbf{H}$, we have $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$. The diagonal matrix $\boldsymbol{\Sigma}$ has at most $M$ nonzero diagonal entries corresponding to the nonzero singular values of $\mathbf{H}$. Consequently, we can rearrange the SVD such that $\mathbf{U}$ is of size $NK \times M$ while $\mathbf{V}$ and $\boldsymbol{\Sigma}$ are $M \times M$ matrices. $\mathbf{U}$ can be partitioned into $M \times N$ sub-matrices as $\mathbf{U} = \left[\mathbf{U}_1^T|\mathbf{U}_2^T|\cdots|\mathbf{U}_K^T\right]^T$. For every $1 \leq k \leq K$, let us define $\beta_k$ as $\beta_k \triangleq \left\|\mathbf{G}_k^\dagger\mathbf{U}_k^H\mathbf{r}_k\right\|^2$. In ICBS, a predefined threshold $\beta$ is coordinated between the relays. Consequently, for each realization of the channels, the relays which satisfy $\beta_k \leq \beta$ amplify and forward their received signal by the matrix $\mathbf{F}_k \triangleq \sqrt{\frac{P_r}{\beta}}\mathbf{G}_k^\dagger\mathbf{U}_k^H$ and other relays are turned off. In this way, the relays whose downlink channels are ill conditioned are turned off. It should be noted that as the active relays satisfy $\left\|\mathbf{G}_k^\dagger\mathbf{U}_k^H\mathbf{r}_k\right\|^2 \leq \beta$, the power of their transmitted signal is guaranteed to be less than or equal to $P_r$. Let us denote the set of off relays by $\mathcal{A}$, i.e. $\mathcal{A} \triangleq \{k\,|\beta_k > \beta\}$. At the destination side, we have (figure 2.2)

$$
\begin{aligned}
\mathbf{y} &= \sum_{k\in\mathcal{A}^c} \mathbf{G}_k\mathbf{F}_k\mathbf{r}_k + \mathbf{z} \\
&= \sqrt{\frac{P_r}{\beta}} \sum_{k\in\mathcal{A}^c} \mathbf{U}_k^H\mathbf{r}_k + \mathbf{z} \\
&\overset{(a)}{=} \sqrt{\frac{P_r}{\beta}} \left(\left(\mathbf{U}^H\mathbf{H} - \sum_{k\in\mathcal{A}} \mathbf{U}_k^H\mathbf{H}_k\right)\mathbf{x} + \sum_{k\in\mathcal{A}^c} \mathbf{U}_k^H\mathbf{n}_k\right) + \mathbf{z} \\
&= \sqrt{\frac{P_r}{\beta}} \left(\left(\mathbf{\Sigma}\mathbf{V}^H - \mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\right)\mathbf{x} + \mathbf{U}_{\mathcal{A}^c}^H\mathbf{n}_{\mathcal{A}^c}\right) + \mathbf{z}. \qquad (2.4)
\end{aligned}
$$

Here, $\mathbf{n} = \left[\mathbf{n}_1^T|\mathbf{n}_2^T|\cdots|\mathbf{n}_K^T\right]^T$, and $(a)$ results from the fact that $\sum_{k=1}^K \mathbf{U}_k^H\mathbf{H}_k = \mathbf{U}^H\mathbf{H}$.

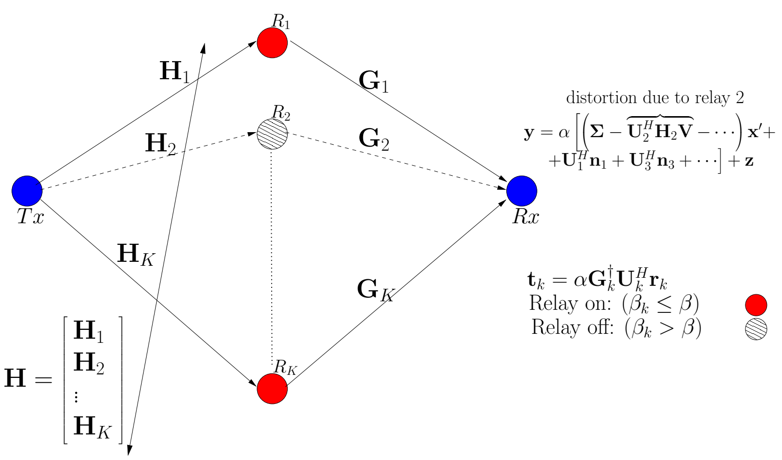


Figure 2.2: Incremental Cooperative Beamforming Scheme

As (2.4) shows, by decreasing the value of $\beta$, one can guarantee a large value for signal to the downlink noise ratio at the expense of turning off more relays. Turning off more relays results in increasing the deviation of the equivalent channel matrix from $\mathbf{\Sigma}\mathbf{V}^H$ and decreasing the determinant of the equivalent channel matrix. It will be shown in Section 2.4 that for large number of relays, it is possible to guarantee both having a large value of signal to downlink noise ratio and a small deviation from $\mathbf{\Sigma}\mathbf{V}^H$. Accordingly, we show that the achievable rate of ICBS is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the network capacity.

### 2.3.1 ICBS in Single Antenna Scenario

In the case that the network nodes are equipped with single antenna, the proposed ICBS is more easily tractable. In this case, the uplink and downlink channel matrices reduce to vectors [2], and we have

$$\mathbf{h} = [h_1 h_2 \dots h_K]^T, \mathbf{u} = \frac{\mathbf{h}}{\sqrt{\|\mathbf{h}\|}}, \sigma = \|\mathbf{h}\|, v = 1, \tag{2.5}$$

where $\mathbf{h} = \mathbf{u}\sigma v^H$ is the result of the SVD decomposition of the uplink channel. The $k$'th relay amplifies the received signal as $t_k = \alpha \frac{u_k^*}{g_k} r_k$.

In ICBS, the relays with $\beta_k = \frac{|u_k|^2}{|g_k|^2}\left(1 + |h_k|^2 P_s\right) \leq \beta$ are active and the rest are turned off. Rewriting (2.4), the received signal at the destination side can be written as

$$y = \sqrt{\frac{P_r}{\beta}}\left(\frac{\|\mathbf{h}_{\mathcal{A}^c}\|^2}{\|\mathbf{h}\|}x + \mathbf{u}_{\mathcal{A}^c}^H \mathbf{n}_{\mathcal{A}^c}\right) + z, \tag{2.6}$$

where $\mathcal{A}^c$ is the set of active relays. The above equation implies that if we can choose the value of $\beta$ such that $\sqrt{\frac{P_r}{\beta}} \gg 1$ and $\|\mathbf{h}_{\mathcal{A}^c}\| \approx \|\mathbf{h}\|$, the capacity of the uplink channel (which is an upper-bound on the capacity of the system) is achievable. The first condition requires selecting a small enough threshold value, while the second condition requires selecting a large enough threshold value to

[2]Here, we denote the problem parameters with their equivalent lowercase bold letters for the vectors or regular letters for the scalars to emphasize on their dimensions

ensure that there are enough active relays in the network. The second condition also translates to making the distortion term (defined in (2.4)) as small as possible.

## 2.4 Asymptotic Analysis

In this section, we consider the asymptotic behavior of the achievable rate of the proposed schemes. First, we show that in the asymptotic regime of $K \to \infty$, by properly choosing the value of $\beta$, the achievable rate of ICBS converges to the capacity. Next, we study the asymptotic SNR behavior of ICBS in two regimes: i) $P_r = P_s \to \infty$; and ii) $P_r \to \infty$. We show that in both SNR regimes ICBS is asymptotically optimal.

**Theorem 2.1** *Consider a parallel MIMO relay network in which $P_s = P_r = P$ is fixed and $K \to \infty$. Then, by setting the threshold as $\beta = \frac{\log(K)}{\sqrt[4]{K}}$, the achievable rate of ICBS converges to the capacity upper-bound defined on the uplink channel. More precisely,*

$$\lim_{K \to \infty} C_u(K) - R_{ICBS}(K) = 0, \tag{2.7}$$

*where $C_u(K) = \max_{\mathbf{Q}(.),\mathbb{E}[\mathrm{Tr}\{\mathbf{Q}\}] \leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \left[\log\left(\left|\mathbf{I}_{KN} + \mathbf{HQH}^H\right|\right)\right]$ is the point to point ergodic capacity of the uplink channel and $R_{ICBS}(K)$ is the achievable rate of ICBS.*

**Proof** The sequence of proof of Theorem 2.1 is as follows. In Lemma 2.2, we relate $\mathbb{P}\left[\nu > \xi\right]$, where $\nu$ denotes the norm of the distortion term defined in equation (2.4), to $\mathbb{P}[k \in \mathcal{A}]$ (the probability of turning off a relay) and $\mathbb{P}[\|\mathbf{U}_k\|^2 > \gamma]$ (the probability of having a sub-matrix with a norm greater than $\gamma$ in the unitary matrix obtained from the SVD of $\mathbf{H}$). In Lemma 2.3, we upper-bound $\mathbb{P}[\|\mathbf{U}_k\|^2 > \gamma]$. As a result, In Lemma 2.4, we show that by properly choosing the value of $\beta$, with high probability, one can simultaneously reduce the effect of the distortion to $o(K)$, while maintaining a large signal to downlink noise ratio. In Lemma 2.5, we show that with high probability, all the singular values of $\mathbf{H}^H\mathbf{H}$ scale as $NK$. Using Lemmas 2.4 and 2.5, it is concluded that with high probability, the ratio of the power of distortion to the power of signal approaches zero, while the signal to the downlink noise ratio approaches infinity at the same time. This is the key idea in the proof of Theorem 2.1.

**Lemma 2.2** *Consider a parallel MIMO relay network with $K$ relays using ICBS. We have*

$$\mathbb{P}\left[\nu > \xi\right] \;\; \leq \;\; \frac{2MNK^2}{\xi}\left(M\mathbb{P}[B_k] + \gamma\mathbb{P}[A_k] + \frac{1}{NK^2}\right), \tag{2.8}$$

*where $\nu$ is defined as $\nu \triangleq \left\|\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\right\|^2$, and $A_k$ and $B_k$ are two events defined as $A_k \equiv \{k \in \mathcal{A}\}$ and $B_k \equiv \{\|\mathbf{U}_k\|^2 > \gamma\}$, respectively.*

**Proof** See Appendix A. ∎

**Lemma 2.3** *Consider a $KN \times M$ Unitary matrix $\mathbf{U}$, where its columns $\mathbf{U}_i$, $i = 1, \cdots, M$, are isotropically distributed unit vectors in $\mathbb{C}^{NK\times 1}$. Let $\mathbf{W}$ be an arbitrary $N \times M$ sub-matrix of $\mathbf{U}$. Then, assuming $\gamma \geq \frac{MN}{K}$, we have*

$$\mathbb{P}\left[\|\mathbf{W}\|^2 \geq \gamma\right] \leq \frac{N^N}{M^{N-2}(N-1)!}(K\gamma)^{N-1}\left(1 - \frac{\gamma}{M}\right)^{N(K-1)}. \tag{2.9}$$

**Proof** See Appendix B. ∎

Next, we apply Lemmas 2.2 and 2.3 to prove that for the threshold value defined in the argument of Theorem 2.1, the distortion term scales as $o(K)$, with a high probability.

**Lemma 2.4** *Assuming $\beta \geq e\gamma$, we have*

$$\mathbb{P}\left[\nu > \xi\right] \leq \frac{2MNK^2}{\xi}\Bigg(2M\mathbb{P}[B_k]+ \\ \frac{\gamma^2}{\beta}\left(2MP_s \log\left(\frac{\beta}{\gamma}\right) + 8\frac{\gamma}{\beta}\log^{MN-1}\left(\frac{\beta}{\gamma}\right) + M\right) + \frac{1}{NK^2}\Bigg), \tag{2.10}$$

*where $\nu$ is the distortion term defined in Lemma 2.2. Moreover, for the threshold value $\beta = \frac{\log(K)}{\sqrt[4]{K}}$, we have*

$$\mathbb{P}\left[\nu > \log^2(K)\sqrt{K}\right] = O\left(\frac{1}{\sqrt[4]{K}}\right) \tag{2.11}$$

*where $\nu$ is the distortion term defined in Lemma 2.2.*

**Proof** See Appendix C. ∎

Although with the threshold value stated by Lemma 2.4, the distortion term may tend to infinity in terms of $K$, the signal term tends to infinity more rapidly. In fact, as the following Lemma shows, the singular values of the whole uplink channel matrix scale as $NK$ with probability one, as $K \to \infty$.

**Lemma 2.5** *Let $\mathbf{A}$ be an $r \times s$ matrix whose entries are i.i.d complex Gaussian random variables with zero mean and unit variance. Assume that $r$ is fixed and $s$ tends to infinity. Then, with probability one $\lambda_{\max}(\mathbf{A}\mathbf{A}^H) \sim \lambda_{\min}(\mathbf{A}\mathbf{A}^H) \sim s$, or more precisely,*

$$\mathbb{P}\left[s\left(1+4\sqrt{r\epsilon}\right) \geq \lambda_{\max}(\mathbf{A}\mathbf{A}^H) \geq \lambda_{\min}(\mathbf{A}\mathbf{A}^H) \geq s\left(1-\sqrt{6r\epsilon}\right)\right] = \\ 1 + O\left(\frac{1}{s\sqrt{\log(s)}}\right), \tag{2.12}$$

*where $\epsilon \triangleq \sqrt{\frac{2\log(s)}{s}}$.*

**Proof** See Appendix D. ∎

Next, we prove Theorem 2.1 by using the above lemmas. By applying the cut-set bound Theorem [12] on the broadcast uplink channel, it can be easily verified [19], [39] that the point-to-point capacity of the uplink channel, $C_u(K)$, is an upper-bound on the capacity of the parallel MIMO relay network. Considering the fact that each of the uplink and downlink channels are used $\frac{1}{2}$ portion of the time (due to the half-duplex assumption), we have a factor of $\frac{1}{2}$ in the expression of $C_u(K)$. Let us define

$$C_{u^\star}(K) \triangleq \frac{M}{2} \log\left(\frac{KNP_s}{M}\right) + O\left(\sqrt[4]{\frac{\log(K)}{K}}\right). \tag{2.13}$$

We first show that $C_{u^\star}(K)$ is an upper-bound for $C_u(K)$, and then prove that a lower-bound for $R_{ICBS}(K)$ converges to $C_{u^\star}(K)$. In order to prove (2.13) is an upper-bound for $C_u(K)$, we show that the amount of uplink capacity improvement obtained from waterfilling the source power between different uplink channel's realizations decays as $O\left(\sqrt[4]{\frac{\log(K)}{K}}\right)$ as $K$ grows. To upper-bound $C_u(K)$, we have

$$\begin{aligned}
C_u(K) &= \frac{1}{2} \max_{\substack{\mathbf{Q}(.)\\ \mathbb{E}[\mathrm{Tr}\{\mathbf{Q}\}]\leq P_s}} \mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_{KN} + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right|\right)\right] \\
&\overset{(a)}{=} \frac{1}{2} \max_{\substack{\mathbf{Q}(.)\\ \mathbb{E}[\mathrm{Tr}\{\mathbf{Q}\}]\leq P_s}} \mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_{M} + \mathbf{H}^H\mathbf{H}\mathbf{Q}\right|\right)\right] \\
&\overset{(b)}{\leq} \frac{M}{2} \max_{\substack{\mathbf{Q}(.)\\ \mathbb{E}[\mathrm{Tr}\{\mathbf{Q}\}]\leq P_s}} \mathbb{E}_{\mathbf{H}}\left[\log\left(1 + \frac{\mathrm{Tr}\left\{\mathbf{H}^H\mathbf{H}\mathbf{Q}\right\}}{M}\right)\right] \\
&\overset{(c)}{\leq} \frac{M}{2} \max_{\substack{\mathbf{Q}(.)\\ \mathbb{E}[\mathrm{Tr}\{\mathbf{Q}\}]\leq P_s}} \mathbb{E}_{\mathbf{H}}\left[\log\left(1 + \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})\mathrm{Tr}\{\mathbf{Q}\}}{M}\right)\right],
\end{aligned} \tag{2.14}$$

where maximization is over all functions $\mathbf{Q}(\mathbf{H})$ which satisfy the average power constraint $P_s$, i.e., $\mathbb{E}\left[\mathrm{Tr}\{\mathbf{Q}(\mathbf{H})\}\right] \leq P_s$. Here, $(a)$ follows from the matrix determinant equality[3], $(b)$ results from the fact that for any $M \times M$ positive semidefinite matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$, applying geometric inequality, we have $|\mathbf{A}| = \prod_{i=1}^{M} \lambda_i \leq \left(\frac{\sum_{i=1}^{M} \lambda_i}{M}\right)^M = \left(\frac{\mathrm{Tr}\{\mathbf{A}\}}{M}\right)^M$, and finally, $(c)$ follows from the fact that for positive semi-definite matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\mathrm{Tr}\{\mathbf{AB}\} \leq \lambda_{\max}(\mathbf{A})\mathrm{Tr}\{\mathbf{B}\}$ [22].

Now, we apply Lemma 2.5 in order to upper-bound RHS of (2.14). Let us define the event $\mathcal{C}$ as the event that $\lambda_{\max}(\mathbf{H}^H\mathbf{H}) \leq KN\left(1 + 4\sqrt{M}\sqrt[4]{\frac{2\log(KN)}{KN}}\right)$. Moreover, let us define $C_u(K, \mathbf{Q}(.))$ as the expression inside the max function in RHS of (2.14). We have

$$
\begin{aligned}
C_u(K, \mathbf{Q}(.)) \;\; = \;\; & \mathbb{E}_{\mathbf{H}}\left[\log\left(1 + \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})\mathrm{Tr}\{\mathbf{Q}\}}{M}\right)\middle|\, \mathcal{C}\right]\mathbb{P}\left\{\mathcal{C}\right\} + \\
& \mathbb{E}_{\mathbf{H}}\left[\log\left(1 + \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})\mathrm{Tr}\{\mathbf{Q}\}}{M}\right)\middle|\, \mathcal{C}^c\right]\mathbb{P}\left\{\mathcal{C}^c\right\} \qquad (2.15)
\end{aligned}
$$

Let us define $C_{\mathbf{Q}}(\mathbf{H}) \triangleq \log\left(1 + \frac{\lambda_{\max}(\mathbf{H}^H\mathbf{H})\mathrm{Tr}\{\mathbf{Q}\}}{M}\right)$. The first term in the RHS of (2.15) can be upper-bounded as follows.

$$
\begin{aligned}
\mathbb{E}_{\mathbf{H}}\left[C_{\mathbf{Q}}(\mathbf{H})|\,\mathcal{C}\right] \;\; & \overset{(a)}{\leq} \;\; \log\left(1 + \frac{KN\left(1 + O\left(\sqrt[4]{\frac{\log(K)}{K}}\right)\right)\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}|\,\mathcal{C}\right]}{M}\right) \\
& \overset{(b)}{\leq} \;\; \log\left(\frac{KNP_s}{M}\right) + O\left(\sqrt[4]{\frac{\log(K)}{K}}\right). \qquad (2.16)
\end{aligned}
$$

Here, $(a)$ follows from concavity of the log function and $(b)$ follows from the fact that $\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}|\,\mathcal{C}\right] \leq \mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}\right] \leq P_s$. Moreover, the second term in the RHS of

[3] Assuming $\mathbf{A}$ and $\mathbf{B}$ to be $M \times N$ and $N \times M$ matrices respectively, we have $|\mathbf{I}_M + \mathbf{AB}| = |\mathbf{I}_N + \mathbf{BA}|$ [22].

(2.15) can be upper-bounded as follows.

$$
\begin{aligned}
\mathbb{E}_{\mathbf{H}}\left[C_{\mathbf{Q}}(\mathbf{H})|\,\mathcal{C}^c\right] &\overset{(a)}{\leq} \log\left(1+\frac{\mathbb{E}_{\mathbf{H}}\left[\lambda_{\max}(\mathbf{H}^H\mathbf{H})\big|\,\mathcal{C}^c\right]}{M}\right)+\log\left(1+\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}|\,\mathcal{C}^c\right]\right) \\
&\overset{(b)}{\leq} \log\left(1+\frac{KN}{\mathbb{P}\left\{\mathcal{C}^c\right\}}\right)+\log\left(1+P_s\right) \\
&\leq \log(K)+O\left(\frac{1}{K}\right)+\log\left(1+P_s\right)-\log\left(\mathbb{P}\left\{\mathcal{C}^c\right\}\right). \qquad (2.17)
\end{aligned}
$$

Here, $(a)$ follows from concavity of the log function and $(b)$ follows from the facts that i) $\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}|\,\mathcal{C}^c\right] \leq \mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\{\mathbf{Q}\}\right] \leq P_s$, and ii) $\mathbb{E}_{\mathbf{H}}\left[\lambda_{\max}(\mathbf{H}^H\mathbf{H})\big|\,\mathcal{C}^c\right] \leq \frac{\mathbb{E}_{\mathbf{H}}\left[\lambda_{\max}(\mathbf{H}^H\mathbf{H})\right]}{\mathbb{P}\{\mathcal{C}^c\}}$.

On the other hand, from Lemma 2.5, we know $\mathbb{P}\{\mathcal{C}^c\} = O\left(\frac{1}{K\sqrt{\log(K)}}\right)$. Applying this fact and combining (2.15), (2.16), and (2.17), we have

$$
\begin{aligned}
C_u(K,\mathbf{Q}(.)) &\leq \log\left(\frac{KNP_s}{M}\right)+O\left(\sqrt[4]{\frac{\log(K)}{K}}\right)-\mathbb{P}\left\{\mathcal{C}^c\right\}\log\left(\mathbb{P}\left\{\mathcal{C}^c\right\}\right) \\
&\leq \log\left(\frac{KNP_s}{M}\right)+O\left(\sqrt[4]{\frac{\log(K)}{K}}\right)=C_{u^\star}(K) \qquad (2.18)
\end{aligned}
$$

Now, we lower-bound $R_{ICBS}(K)$. Re-arranging (2.4), we have $\mathbf{y}=\sqrt{\frac{P_r}{\beta}}\mathbf{H}^\star\mathbf{x}+\mathbf{n}^\star$ where $\mathbf{H}^\star \triangleq \mathbf{\Sigma}-\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}$ and $\mathbf{n}^\star$ is the additive colored Gaussian vector with the covariance matrix $\mathbf{P}_{\mathbf{n}^\star}=\mathbf{I}_M+\frac{P_r}{\beta}\mathbf{U}_{\mathcal{A}^c}^H\mathbf{U}_{\mathcal{A}^c}$. The achievable rate of such a system is

$$
\begin{aligned}
R_{ICBS}(K) &= \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M+\frac{P_rP_s}{M\beta}\mathbf{H}^{\star H}\mathbf{H}^\star\mathbf{P}_{\mathbf{n}^\star}^{-1}\right|\right)\right] \\
&\overset{(a)}{\geq} \frac{M}{2}\log\left(\frac{P_r}{P_r+\beta}\right)+\frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M+\frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^\star\right|\right)\right] \qquad (2.19) \\
&\overset{(b)}{\geq} \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M+\frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^\star\right|\right)\right]+O\left(\frac{\log(K)}{\sqrt[4]{K}}\right) \qquad (2.20)
\end{aligned}
$$

where $(a)$ follows from the fact that $\mathbf{P}_{\mathbf{n}^\star} \preccurlyeq (\frac{P_r}{\beta}+1)\mathbf{I}_M$ and $(b)$ follows from the assumption that $\beta = \frac{\log(K)}{\sqrt[4]{K}}$. For convenience, let us define $R_L(K)$ as

$$R_L(K) \triangleq \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M + \frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^\star\right|\right)\right]. \tag{2.21}$$

Defining the events $\mathcal{E}$ and $\mathcal{F}$ as

$$\mathcal{E} \equiv \left\{\lambda_{\min}\left(\mathbf{H}^H\mathbf{H}\right) \geq KN\left(1-\sqrt{6M}\sqrt[4]{\frac{2\log(NK)}{NK}}\right)\right\}, \tag{2.22}$$

and

$$\mathcal{F} \equiv \left\{\left\|\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\right\|^2 \leq \log^2(K)\sqrt{K}\right\}, \tag{2.23}$$

we have

$$\mathbb{P}\left[\mathcal{E},\mathcal{F}\right] \overset{(a)}{\geq} 1-\mathbb{P}[\mathcal{E}^c]-\mathbb{P}[\mathcal{F}^c] \overset{(b)}{=} 1+O\left(\frac{1}{\sqrt[4]{K}}\right). \tag{2.24}$$

Here, $(a)$ follows from union bound inequality and $(b)$ follows from Lemmas 2.4 and 2.5. Let us define $\mathbf{\Lambda} \triangleq \mathbf{\Sigma}\mathbf{\Sigma}^H$ and assume that the diagonal entries of $\mathbf{\Lambda}$ are ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$. Thus, $R_L(K)$ can be lower-bounded as

$$
\begin{aligned}
R_L(K) &\geq \frac{1}{2}\mathbb{P}\left[\mathcal{E},\mathcal{F}\right]\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^{\star}\right|\right)\Big|\mathcal{E},\mathcal{F}\right] && (2.25)\\
&\overset{(a)}{\geq} \mathbb{P}\left[\mathcal{E},\mathcal{F}\right]\mathbb{E}_{\mathbf{H}}\left[\log\left(\left(\prod_{i=1}^{M}\lambda_i^{\frac{1}{2}} - \sum_{i=1}^{M} i!\binom{M}{i}\left\|\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\right\|_{\star}^{i}\prod_{j=1}^{M-i}\lambda_j^{\frac{1}{2}}\right)\right)\Big|\mathcal{E},\mathcal{F}\right]\\
&\quad +\mathbb{P}\left[\mathcal{E},\mathcal{F}\right]\frac{M}{2}\log\left(\frac{P_s}{M}\right)\\
&\overset{(b)}{\geq} \mathbb{P}\left[\mathcal{E},\mathcal{F}\right]\Bigg\{\mathbb{E}_{\mathbf{H}}\left[\log\left(1-\sum_{i=1}^{M} i!\binom{M}{i}\left(\frac{\left\|\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\right\|^2}{\lambda_{\min}\left(\mathbf{H}^H\mathbf{H}\right)}\right)^{\frac{i}{2}}\right)\Big|\mathcal{E},\mathcal{F}\right]\\
&\quad +\frac{M}{2}\log\left(\frac{P_s}{M}\right)+\frac{1}{2}\sum_{i=1}^{M}\mathbb{E}_{\mathbf{H}}\left[\log\left(\lambda_i\right)\Big|\mathcal{E},\mathcal{F}\right]\Bigg\}\\
&\overset{(c)}{\geq} \mathbb{P}\left[\mathcal{E},\mathcal{F}\right]\Bigg\{\frac{M}{2}\log\left(\frac{P_s}{M}\right)+\frac{1}{2}\sum_{i=1}^{M}\mathbb{E}_{\mathbf{H}}\left[\log\left(\lambda_i\right)\Big|\mathcal{E},\mathcal{F}\right]-\\
&\quad \frac{M\log(K)}{\sqrt{N}\sqrt[4]{K}}\left(1+O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)\right)\Bigg\} && (2.26)\\
&\overset{(d)}{\geq} \frac{M}{2}\log\left(\frac{KNP_s}{M}\right)+O\left(\frac{\log(K)}{\sqrt[4]{K}}\right). && (2.27)
\end{aligned}
$$

Here, $(a)$ follows from an upper-bound on the determinant expansion [4] of $\mathbf{\Lambda}^{\frac{1}{2}} - \mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\mathbf{V}$, expanded over all possible set entries between $\mathbf{\Lambda}^{\frac{1}{2}}$ and $\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\mathbf{V}$, $(b)$ follows by upper-bounding the maximum absolute value entry of the matrix $\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}$ with the matrix Frobenius norm and also lower-bounding $\lambda_i$ with $\lambda_{\min}(\mathbf{H}^H\mathbf{H})$, $(c)$ results from the upper-bound on $\|\mathbf{U}_{\mathcal{A}}^H\mathbf{H}_{\mathcal{A}}\|$ and the lower-bound on $\lambda_i$ obtained from definition of the events $\mathcal{E}$ and $\mathcal{F}$, and finally, $(d)$ follows from the lower-bound on $\lambda_i$ conditioned on the event $\mathcal{E}$ and also the upper-bound on $\mathbb{P}\left[\mathcal{E},\mathcal{F}\right]$ obtained

[4] $\det\left(A\right)=\sum_{\pi}\left(-1\right)^{\sigma(\pi)}a_{1\pi_1}a_{2\pi_2}\cdots a_{n\pi_n}\leq\sum_{\pi}\left|a_{1\pi_1}a_{2\pi_2}\cdots a_{n\pi_n}\right|$, where $\sigma$ is the parity function of permutation.

in (2.24). Now, defining

$$R_S(K) = \frac{M}{2} \log\left(\frac{KNP_s}{M}\right), \tag{2.28}$$

according to (2.20), (2.21), and (2.27), we have

$$\lim_{K\to\infty} R_{ICBS}(K) - R_S(K) \geq 0. \tag{2.29}$$

Furthermore, following the definition of $C_{u^\star}(K)$ in (2.13), we observe

$$\lim_{K\to\infty} C_{u^\star}(K) - R_S(K) = 0. \tag{2.30}$$

Knowing $C_{u^\star}(K)$ is an upper-bound on the capacity completes the proof. ∎

**Corollary 2.6** *Achievable rate of ICBS is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the upper-bound on the capacity of the network, i.e. $C_u(K)$.*

**Proof** Revisiting the inequality series, (2.18), (2.20), and (2.27), it is easy to verify that the achievable rate of ICBS is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the capacity upper-bound. ∎

An interesting fact stated in the following Corollary is that as the number of relays increases, the instantaneous achievable rate of the ICBS scheme for a realization of the network channels is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the ergodic capacity of the network, with probability 1 (in probability).

**Corollary 2.7** *Consider the parallel MIMO relay network and ICBS with the threshold $\beta = \frac{\log(K)}{\sqrt[4]{K}}$. Let us define $R_{ICBS}(\mathbf{H}, \mathbf{G})$ as the instantaneous achievable rate of ICBS for a realization of the network channels. Then, we have*

$$\mathbb{P}\left[R_{ICBS}(\mathbf{H}, \mathbf{G}) \leq C_u(K) + O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)\right] = O\left(\frac{1}{\sqrt[4]{K}}\right).$$

**Proof** First, we note that $R_{ICBS}(\mathbf{H},\mathbf{G}) = \frac{1}{2}\log\left(\left|\mathbf{I}_M + \frac{P_r P_s}{M\beta}\mathbf{H}^{\star H}\mathbf{H}^{\star}\mathbf{P}_{\mathbf{n}^\star}^{-1}\right|\right)$. Revisiting the inequality series (2.20), we observe that $R_{ICBS}(\mathbf{H},\mathbf{G})$ is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the function $R_L(\mathbf{H},\mathbf{G}) \triangleq \frac{1}{2}\log\left(\left|\mathbf{I}_M + \frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^{\star}\right|\right)$. On the other hand, revisiting the inequality series (2.27), we observe that conditioned on the events $\mathcal{E}$ and $\mathcal{F}$, the function $R_L(\mathbf{H},\mathbf{G})$ is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the capacity upper-bound. As a result, conditioned on $\mathcal{E}\bigcap\mathcal{F}$, instantaneous achievable rate of ICBS is at most $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ below the capacity. This completes the proof. ∎

Another interesting result is that by increasing the number of relays, each relay can operate with much less power as compared to the source, while the scheme achieves the same rate asymptotically.

**Theorem 2.8** *As long as the power of each relay scales as* $P_r = \omega\left(\frac{\log^3(K)\log(\log(K))}{K}\right)$, *we have*

$$\lim_{K\to\infty} R_{ICBS}(K) - C_u(K) = \lim_{K\to\infty} R_{ICBS}(K) - \frac{M}{2}\log\left(\frac{KNP_s}{M}\right) = 0. \tag{2.31}$$

**Proof** First, we select $\beta$ such that $\beta = o(P_r)$ and $\beta = \omega\left(\frac{\log^3(K)\log(\log(K))}{K}\right)$. Now, We apply the argument of first part of Lemma 2.4 to upper-bound $\mathbb{P}[\nu > \xi]$. Applying Lemma (2.10) with values of $\gamma = \frac{4\log(K)}{K}$ and $\xi = o(K)$, we have

$$\mathbb{P}[\nu > \xi] = o\left(\frac{1}{\log(K)}\right) \tag{2.32}$$

Now, revisiting the step from (2.19) to (2.20) with the new value of $\beta$ and $P_r$ and noting $\beta = o(P_r)$, we conclude that in this case

$$R_{ICBS}(K) \geq R_L(K) + o(1). \tag{2.33}$$

Furthermore, revisiting the steps (2.26) and (2.27) in lower-bounding $R_L(K)$ and noting that $\xi = o(K)$ and $\mathbb{P}[\mathcal{E}, \mathcal{F}] = \mathbb{P}[\nu > \xi] = o\left(\frac{1}{\log(K)}\right)$, we have

$$R_L(K) \geq \frac{M}{2} \log\left(\frac{KNP_s}{M}\right) + o(1). \tag{2.34}$$

Combining (2.33) and (2.34) completes the proof. ∎

**Theorem 2.9** *Consider the parallel MIMO relay network with $P_s = P_r = P \to \infty$. Then, ICBS with the threshold $\beta = P\log^2(P)$ achieves the maximum multiplexing gain of the network which is equal to $\frac{M}{2}$.*

**Proof** First, we can apply the cut-set bound Theorem [12] to upper-bound the ergodic capacity with the point-to-point capacity of the broadcast uplink channel. As the broadcast uplink channel can be considered as a $(KN) \times M$ MIMO channel and this channel is active for half of the time, we conclude that $\frac{M}{2}$ is an upper-bound on the multiplexing gain of the network.

Hence, we only need to prove that ICBS with the threshold $\beta = P\log^2(P)$ achieves the multiplexing gain of $\frac{M}{2}$. To prove, we first show that with probability $1 + O\left(\frac{1}{\log(P)}\right)$ all relays are on. The probability of a relay being off can be upper-bounded by (C.3). Applying Lemma 2.3 for $\gamma = M\left(1 - \frac{1}{\log(P)}\right)$, we conclude $\mathbb{P}[B_k] = O\left(\frac{1}{\log(P)}\right)^{N(K-1)}$. Furthermore, we can apply (C.4) to upper-bound $\mathbb{P}[D_k]$. Noting $\delta = \frac{\gamma}{\beta} \leq \frac{M}{P\log^2(P)}$, we conclude $\mathbb{P}[D_k] = O\left(\frac{1}{\log(P)}\right)$. Accordingly, the probability that all relays are on can be lower-bounded as

$$\mathbb{P}\{\mathcal{A} = \varnothing\} \geq 1 - \sum_{k=1}^{K} \mathbb{P}\{A_k\} \overset{(a)}{\geq} 1 - \sum_{k=1}^{K} \mathbb{P}\{B_k\} + \mathbb{P}\{D_k\} = 1 + O\left(\frac{1}{\log(P)}\right), \tag{2.35}$$

where $(a)$ follows from (C.3).

Now, we apply the inequality (2.19) to lower-bound the achievable rate of ICBS. We have

$$
\begin{aligned}
R_{ICBS}(P) &\geq R_L(P) + \frac{M}{2}\log\left(\frac{P}{P+\beta}\right) \\
&\overset{(a)}{\geq} \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M + \frac{P}{M}\mathbf{H}^{\star H}\mathbf{H}^{\star}\right|\right)\right] - \frac{M}{2}\log\left(\log\left(P\right)\right) \\
&\geq \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M + \frac{P}{M}\mathbf{H}^{\star H}\mathbf{H}^{\star}\right|\right)\middle|\, \mathcal{A}=\varnothing\right]\mathbb{P}\left[\mathcal{A}=\varnothing\right] - \\
&\quad \frac{M}{2}\log\left(\log\left(P\right)\right) \\
&\overset{(b)}{=} \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M + \frac{P}{M}\mathbf{H}^{H}\mathbf{H}\right|\right)\right] - \frac{M}{2}\log\left(\log\left(P\right)\right) - \\
&\quad \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left(\left|\mathbf{I}_M + \frac{P}{M}\mathbf{H}^{H}\mathbf{H}\right|\right)\middle|\, \mathcal{A}\neq\varnothing\right]\mathbb{P}\left[\mathcal{A}\neq\varnothing\right] \\
&\overset{(c)}{\geq} \frac{M}{2}\log(P) + O(1) - \frac{M}{2}\log\left(1+\frac{P}{M}\mathbb{E}\left[\left\|\mathbf{H}\right\|^2\middle|\, \mathcal{A}\neq\varnothing\right]\right)\mathbb{P}\left[\mathcal{A}\neq\varnothing\right] \\
&\quad -\frac{M}{2}\log\left(\log\left(P\right)\right) \\
&\overset{(d)}{\geq} \frac{M}{2}\log(P) + O(1) - \frac{M}{2}\log\left(\log\left(P\right)\right)
\end{aligned} \tag{2.36}
$$

Here, $(a)$ follows from the fact that $\log(1+x) \leq x$. $(b)$ follows from the fact that conditioned on $\mathcal{A}=\varnothing$, we have $\mathbf{H}^{\star}=\mathbf{H}$. $(c)$ follows from the ergodic capacity formula for the MIMO point-to-point channel [24] and also concavity of the log function. Finally, $(d)$ follows from the fact that $\mathbb{E}\left[\left\|\mathbf{H}\right\|^2\middle|\, \mathcal{A}\neq\varnothing\right] \leq \frac{E\left[\|\mathbf{H}\|^2\right]}{\mathbb{P}[\mathcal{A}\neq\varnothing]} = \frac{MNK}{\mathbb{P}[\mathcal{A}\neq\varnothing]}$ and also, knowing that $\mathbb{P}\left[\mathcal{A}\neq\varnothing\right] = O\left(\frac{1}{\log(P)}\right)$. Applying (2.36), we observe $\lim_{P\to\infty}\frac{R_{ICBS}(P)}{\log(P)} \geq \frac{M}{2}$. This completes the proof. ∎

In Theorem 2.9, we have shown the optimality of ICBS in the scenario where the power of both the source and the relays tends to infinity. In the following theorem, we study the scenario where only the power of the relays tends to infinity, and show that ICBS is optimum in this scenario as well.

**Theorem 2.10** *Consider the parallel MIMO relay network in which the source power $P$ and the number of relays $K$ are fixed, but the power of each relay ($P_r$) tends to infinity. Then, the ICBS scheme with the threshold $\beta = O\left(\frac{P_r}{\log(P_r)}\right)$ achieves the capacity of parallel MIMO relay network. More precisely, the achievable rate of ICBS is at most $O\left(\frac{1}{\log(P_r)}\right)$ below the capacity.*

**Proof** The proof applies the same technique as the one that is used in the proof of Theorem 2.9. We show that the capacity of the network is asymptotically equal to the achievable rate of ICBS conditioned on $\mathcal{A} = \varnothing$.

First, Applying Lemma 2.3 for $\gamma = M\left(1 - \frac{\log(P_r)}{P_r}\right)$, we conclude $\mathbb{P}\left[B_k\right] = O\left(\frac{\log(P_r)}{P_r}\right)^{N(K-1)}$. Furthermore, we can apply (C.4) to upper-bound $\mathbb{P}[D_k]$. Noting $\delta = \frac{\gamma}{\beta} \leq \frac{M\log(P_r)}{P_r}$, we conclude $\mathbb{P}[D_k] = O\left(\frac{\log^2(P_r)}{P_r}\right)$. Accordingly, the probability that all relays are on can be lower-bounded as

$$\mathbb{P}\left\{\mathcal{A} = \varnothing\right\} = 1 + O\left(\frac{\log^2(P_r)}{P_r}\right) \tag{2.37}$$

From the cut-set bound Theorem, the point-to-point capacity of the broadcast channel is an upper-bound for the ergodic capacity of the network. Let us assume that for each channel realization $\mathbf{H}$, the input covariance matrix $\mathbf{Q}^\star(\mathbf{H})$ maximizes the broadcast channel rate, i.e. $\mathbf{Q}^\star(\mathbf{H}) = \underset{\mathbf{Q},\mathrm{tr}(\mathbf{Q})\leq P_s}{\operatorname{argmax}} \left|\mathbf{I}_M + \mathbf{H}^H\mathbf{H}\mathbf{Q}\right|$. Revisiting the inequality series of (2.36), we can lower-bound the achievable rate of ICBS as follows.

$$
\begin{aligned}
R_{ICBS}(P_r) &\overset{(a)}{\geq} \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_M + \frac{P_s}{M}\mathbf{H}^{\star H}\mathbf{H}^{\star}\mathbf{Q}^{\star}(\mathbf{H})\right|\right] - \frac{M}{2}\log\left(1+\frac{\beta}{P_r}\right) \\
&\overset{(b)}{\geq} C_{ub} - \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_M + \frac{P_s}{M}\mathbf{H}^{H}\mathbf{H}\mathbf{Q}^{\star}(\mathbf{H})\right|\middle|\mathcal{A}\neq\varnothing\right]\mathbb{P}\left[A\neq\varnothing\right] - \\
&\quad O\left(\frac{1}{\log(P_r)}\right) \\
&\overset{(c)}{\geq} C_{ub} - \frac{M}{2}\log\left(1+\frac{P_s}{M}\mathbb{E}_{\mathbf{H}}\left[\|\mathbf{H}\|^2\middle|\mathcal{A}\neq\varnothing\right]\right)\mathbb{P}\left[A\neq\varnothing\right] - \\
&\quad O\left(\frac{1}{\log(P_r)}\right) \\
&\overset{(d)}{\geq} C_{ub} - \frac{M}{2}\log\left(1+\frac{P_sNKP_r}{\log^2(P_r)}\right)\frac{\log^2(P_r)}{P_r} - O\left(\frac{1}{\log(P_r)}\right) \\
&= C_{ub} + O\left(\frac{1}{\log(P_r)}\right),
\end{aligned}
\tag{2.38}
$$

where $C_{ub} \triangleq \frac{1}{2}\mathbb{E}_{\mathbf{H}}\left[\log\left|\mathbf{I}_M + \frac{P_s}{M}\mathbf{H}^{H}\mathbf{H}\mathbf{Q}^{\star}(\mathbf{H})\right|\right]$ is the capacity upper-bound based on the point-to-point capacity of the broadcast channel. Here, $(a)$ follows from (2.19). $(b)$ follows from the fact that conditioned on $\mathcal{A}=\varnothing$, we have $\mathbf{H}^{\star}=\mathbf{H}$. $(c)$ follows from the same argument as the one used in the proof of (2.14). Finally, $(d)$ results from the fact that $\mathbb{E}_{\mathbf{H}}\left[\|\mathbf{H}\|^2\middle|\mathcal{A}\neq\varnothing\right] \leq \frac{\mathbb{E}_{\mathbf{H}}\left[\|\mathbf{H}\|^2\right]}{\mathbb{P}[\mathcal{A}\neq\varnothing]} = \frac{MNK}{\mathbb{P}[\mathcal{A}\neq\varnothing]}$. (2.38) completes the proof. ∎

## 2.5 Simulation Results

Figure 2.3 shows the simulation results for the achievable rate of ICBS, BNOP matched filtering scheme [19], and the upper-bound of the capacity based on the uplink Cut-Set for varying number of relays. The number of transmitting and receiving antennas is $M=N=2$, and the SNR is $P_s=P_r=10dB$. While both of the schemes demonstrate logarithmic scaling of rate in terms of $K$, we observe that there is a significant gap between the BNOP scheme and the scheme proposed here, reflecting the gap of $O(1)$ in the achievable rate of [19]. On the other hand, the gap between ICBS and the upper-bound approaches zero, as the number of relays increases.

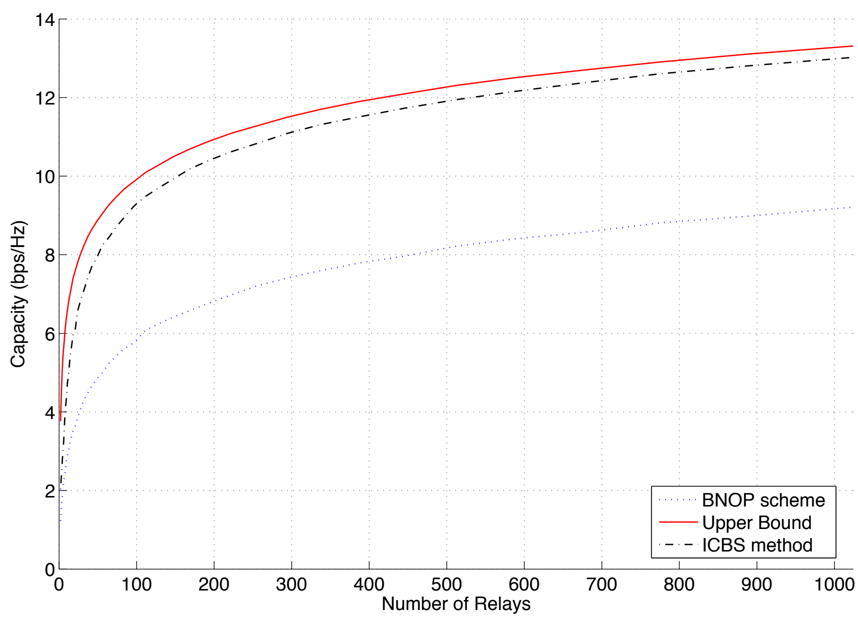


Figure 2.3: Upper-bound of the capacity and achievable rate of ICBS and BNOP matched filtering schemes vs. number of relays

## 2.6 Conclusion

A simple new scheme, Incremental Cooperative Beamforming Scheme (ICBS), based on Amplify and Forward (AF) relaying strategy is introduced in the parallel

MIMO relay network. ICBS is shown to achieve the capacity of parallel MIMO relay network for $K \to \infty$. The scheme is shown to approach the upper-bound of the capacity with a gap no more than $O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$. As a result, it is shown that the capacity of a parallel MIMO relay network is $C(K) = \frac{M}{2}\log\left(1+\frac{KNP}{M}\right)+O\left(\frac{\log(K)}{\sqrt[4]{K}}\right)$ in terms of the number of relays, $K$. Moreover, it is shown that as the number of relays increases, the relays in ICBS can operate using much less power without any performance degradation. Finally, two asymptotic SNR regimes are investigated: i) In the regime where the source power is equal to the relays power and both tend to infinity, ICBS is shown to achieve the full multiplexing gain, regardless of the number of relays; and ii) In the regime where the source power is fixed but the power of each relay tends to infinity, ICBS is shown to achieve the network capacity with a gap which vanishes as $O\left(\frac{1}{\log(P_r)}\right)$. The simulation results confirm the validity of the theoretical arguments.

# Chapter 3

# Diversity-Multiplexing Tradeoff in Single-Antenna Multiple-Relay Networks

## 3.1 Introduction

In this chapter, we consider the general multi-antenna multiple-relay networks. The model which is introduced in this chapter is going to be used in the following chapters, as well. We propose a new scheme, which we call random sequential (RS), based on the SAF relaying for general multiple-antenna multi-hop networks. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the transmitter of the future paths on the receiver of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in amplify-and-forward relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. Furthermore, each relay node knows the CSI of its corresponding backward channel, and the receiver knows the equivalent end-to-end channel. The reason behind random unitary matrix multiplication at the relays can be described as follows: Using the traditional AF relaying at multi-antenna relay nodes, there exists a chance that the eigenvectors corresponding to the large eignenvalues of the incoming channel matrix of the relay project to the eigenvectors corresponding to the small eignenvalues of the relay's outgoing channel matrix. This event degrades the performance of traditional AF relaying in the MIMO setup. However, in the RS scheme, utilizing the random unitary matrix multiplication at the relay nodes for different time-slots, such an event is much more unlikely to happen. Consequently, the RS scheme achieves better diversity gain and DMT compared to the traditional AF relaying[1].

In this chapter, we study DMT of the RS scheme in single-antenna wireless relay networks. In the following 2 chapters, the maximum achievable diversity gain and DMT of the RS scheme are investigated for multi-antenna wireless relay networks. Here, we derive the DMT of the RS scheme for general single-antenna multiple-relay networks. Specifically, we derive: 1) the exact DMT of the RS scheme under the condition of "non-interfering relaying", and 2) a lower-bound on the DMT of the RS scheme (no conditions imposed). Finally, we prove that for single-antenna multiple-access multiple-relay networks (with $K > 1$ relays) when there is no direct link between the transmitters and the receiver and all the relays

[1]This fact will be elaborated more in the following 2 chapters where we show that RS scheme achieves the optimum diversity gain in any general multi-antenna wireless relay network and, it also achieves the optimum DMT in a class of multi-antenna multiple-relay networks.

are connected to the transmitter and to the receiver, the RS scheme achieves the optimum DMT. However, for two-hop multiple-access single-relay networks, we show that the proposed scheme is unable to achieve the optimum DMT, while the DDF scheme is shown to perform optimally in this scenario.

The rest of the chapter is organized as follows. In section 3.2, the system model is introduced. In section 3.3, the proposed random sequential scheme is described. Section 3.4 is dedicated to the DMT analysis of the proposed RS scheme for the single-antenna relay networks. Finally, section 3.5 concludes the chapter.

## 3.2 System Model

The setup throughout this chapter consists of $K$ relays assisting the source and the destination in the half-duplex mode, i.e. at a given time, the relays can either transmit or receive. Each two nodes are assumed either i) to be connected by a quasi-static flat Rayleigh-fading channel, i.e. the channel gains remain constant during a block of transmission and change independently from block to block; or ii) to be disconnected, i.e. there is no direct link between them. Hence, the undirected graph $G = (V, E)$ is used to show the connected pairs in the network[2]. The node set is denoted by $V = \{0, 1, \ldots, K + 1\}$ where the $i$'th node is equipped with $N_i$ antennas. Nodes 0 and $K + 1$ correspond to the source and the destination nodes, respectively[3]. The received and the transmitted vectors at the $k$'th node are shown

[2]Notice that any where the underlying graph is directed, like in Remarks 3.4 and 3.13, the assumption will be explicitly mentioned.

[3]Throughout this chapter, it is assumed that the network consists of one source. However, in Theorems 3.16 and 3.19, we study the case of two-hop multiple sources single destination scenario.

by $\mathbf{y}_k$ and $\mathbf{x}_k$, respectively. Hence, at the destination side of the $a$'th node, we have

$$\mathbf{y}_a = \sum_{\{a,b\} \in E} \mathbf{H}_{a,b} \mathbf{x}_b + \mathbf{n}_a, \tag{3.1}$$

where $\mathbf{H}_{a,b}$ shows the $N_a \times N_b$ Rayleigh-distributed channel matrix between the $a$'th and the $b$'th nodes and $\mathbf{n}_a \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}_{N_a}\right)$ is the additive white Gaussian noise. We assume reciprocal channels between each pair. Hence, $\mathbf{H}_{a,b} = \mathbf{H}_{b,a}^T$. However, it can be easily verified that all the statements of this chapter are valid under the non-reciprocity assumption. In the scenario of single-antenna networks, the channel between nodes $a$ and $b$ is denoted by $h_{\{a,b\}}$ to emphasize both the SISO and the reciprocity assumptions. As in [26] and [55], each relay is assumed to know the state of its backward channel, and moreover, the destination knows the equivalent end-to-end channel. Hence, unlike the CF scheme in [34], no CSI feedback is needed. All nodes have the same power constraint, $P$. Finally, we assume that the topology of the network is known by the nodes such that they can perform a distributed AF strategy throughout the network.

Throughout this chapter, we make some further assumptions in order to prove our statements. First, we consider the scenario in which nodes with a single antenna are used[4]. Moreover, in Theorems 3.7, 3.9, 3.16, and 3.19, where we address DMT optimality of the RS scheme, we assume that there is no direct link between the source(s) and the destination. This assumption is reasonable when the source and the destination are far from each other and the relay nodes establish the connection between the end nodes. Moreover, we assume that all the relay nodes are connected to the source and to the destination through quasi-static flat Rayleigh-fading channels. Hence, the network graph is two-hop. Specifically, we denote the

[4]The case of the nodes equipped with multi-antenna is investigated in the following 2 chapters.

output vector at the source as $\mathbf{x}$, the input vector and the output vector at the $k$'th relay as $\mathbf{r}_k$ and $\mathbf{t}_k$, respectively, and the input at the destination as $\mathbf{y}$.

## 3.3 Proposed Random Sequential (RS) Amplify-and-Forwarding Scheme

In the proposed RS scheme, a sequence $\mathrm{P} \equiv (\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ of $L$ paths[5] originating from the source and culminating in the destination with the length $(l_1, l_2, \ldots, l_L)$ are involved in connecting the source to the destination sequentially $(\mathrm{p}_i(0) = 0, \mathrm{p}_i(l_i) = K+1)$. Note that any path p of $G$ can be selected multiple times in the sequence.

Furthermore, the entire block of transmission is divided into $S$ slots, each consisting of $T'$ symbols. Hence, the entire block consists of $T = ST'$ symbols. Let us assume the source intends to send information to the destination at a rate of $r$ bits per symbol. To transmit a message $w$, the source selects the corresponding codeword from a Gaussian random code-book consisting of $2^{ST'r}$ elements each with length $LT'$. Starting from the first slot, the source sequentially transmits the $i$'th portion $(1 \le i \le L)$ of the codeword through the sequence of relay nodes in $\mathrm{p}_i$. More precisely, a timing sequence $\{s_{i,j}\}_{i=1,j=1}^{L,l_i}$ is associated with the path sequence. The source sends the $i$'th portion of the codeword in the $s_{i,1}$'th slot. Following the transmission of the $i$'th portion of the codeword by the source, in the $s_{i,j}$'th slot,

[5] Throughout this chapter, a path p is defined as a sequence of the graph nodes $(v_0, v_1, v_2, \ldots, v_l)$ such that for any $i$, $\{v_i, v_{i+1}\} \in E$, and for all $i \neq j$, we have $v_i \neq v_j$. The length of the path is defined as the total number of edges on the path, $l$. Furthermore, $\mathrm{p}(i)$ denotes the $i$'th node that p visits, i.e. $\mathrm{p}(i) = v_i$.

$1 \leq j \leq l_i$, the node $\mathrm{p}_i(j)$ receives the transmitted signal from the node $\mathrm{p}_i(j-1)$. Assuming $\mathrm{p}_i(j)$ is not the destination node, i.e. $j < l_i$, it multiplies the received signal in the $s_{i,j}$'th slot by a $N_{\mathrm{p}_i(j)} \times N_{\mathrm{p}_i(j)}$ random, uniformly distributed unitary matrix $\mathbf{U}_{i,j}$ which is known at the destination side, amplifies the signal by the maximum possible coefficient $\alpha_{i,j}$ considering the output power constraint $P$ and $\alpha_{i,j} \leq 1$, and transmits the amplified signal in the $s_{i,j+1}$'th slot. Furthermore, the timing sequence $\{s_{i,j}\}$ should have the following properties

(1) for all $i, j$, we have $1 \leq s_{i,j} \leq S$.

(2) for $i < i'$, we have $s_{i,1} < s_{i',1}$ (the ordering assumption on the paths)

(3) for $j < j'$, we have $s_{i,j} < s_{i,j'}$ (the causality assumption)

(4) for all $i < i'$ and $s_{i,j} = s_{i',j'}$, we have $\{\mathrm{p}_i(j), \mathrm{p}_{i'}(j'-1)\} \notin E$ (no noncausal interference assumption). This assumption ensures that the signal of the future paths causes no interference on the output signal of the current path. This assumption can be realized by designing the timing of the paths such that in each time slot, the current running paths are established through disjoint hops.

At the destination side, having received the signal of all paths, the destination decodes the transmitted message $w$ based on the signal received in the time slots $\{s_{i,l_i}\}_{i=1}^{L}$. As we observe in the sequel, the fourth assumption on $\{s_{i,j}\}$ converts the equivalent end-to-end channel matrix to lower-triangular in the case of single-antenna nodes, or to block lower-triangular in the case of multiple-antenna nodes.

An example of a three-hop network consisting of $K = 4$ relays is shown in Figure (3.1). It can easily be verified that there are exactly 12 paths in the graph connecting the source to the destination. Now, consider the four paths $\mathrm{p}_1 = (0, 1, 3, 5)$, $\mathrm{p}_2 = (0, 2, 4, 5)$, $\mathrm{p}_3 = (0, 1, 4, 5)$ and $\mathrm{p}_4 = (0, 2, 3, 5)$ connecting the

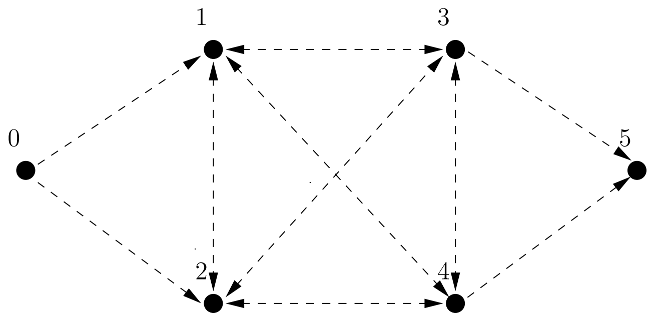


Figure 3.1: An example of a 3-hop network where $N_0 = N_5 = 2, N_1 = N_2 = N_3 = N_4 = 1$.

source to the destination. Assume the RS scheme is performed with the path sequence $\mathrm{P}_1 \equiv (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_4)$. Table 3.1 shows one possible valid timing sequence associated with RS scheme with the path sequence $\mathrm{P}_1$. As seen, the first portion of the source's codeword is sent in the 1st time slot and is received by the destination through the nodes of the path $\mathrm{P}_1(1) \equiv (0, 1, 3, 5)$ as follows: In the 1st slot, the source's signal is received by node 1. Following that, in the 2nd slot, node 1 sends the amplified signal to node 3, and finally, in the 3rd slot, the destination receives the signal from node 3. As observed, for every $1 \leq i \leq 3$, signal of the $i$'th path interferes on the output signal of the $i+1$'th path. However, no interference is caused by the signal of future paths on the outputs of the current path. The timing sequence corresponding to Table 3.1 can be expressed as $s_{i,j} = i + \lfloor \frac{i}{3} \rfloor + j - 1$ and it results in the total number of transmission slots to be equal to 7, i.e. $S = 7$.

As an another example, consider RS scheme with the path sequence $\mathrm{P}_2 \equiv (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_1, \mathrm{p}_2)$. Table 3.2 shows one possible valid timing-sequence for the RS scheme with the path sequence $\mathrm{P}_2$. Here, we observe that the signal on every path interferes on the output of the next two consecutive paths. However, like the scenario with $\mathrm{P}_1$, no interference is caused by the signal of future paths on the output signal of the current path. The timing sequence corresponding to Table 3.2 can be expressed as $s_{i,j} = i+j-1$ and it results in the total number of transmission slots equal to 6, i.e. $S = 6$.

It is worth noting that to achieve higher spectral efficiencies (corresponding to larger multiplexing gains), it is desirable to have larger values for $\frac{L}{S}$. Indeed, $\frac{L}{S} \to 1$ is the highest possible value. However, this can not be achieved in some graphs (an example is the case of two-hop single relay scenario studied in the next section where $\frac{L}{S} = 0.5$). On the other hand, to achieve higher reliability (corresponding to larger diversity gains between the end nodes), it is desirable to utilize more paths of the graph in the path sequence. It is not always possible to satisfy both of these objectives simultaneously. As an example, consider the single-antenna two-hop relay network where there is a direct link between the end nodes, i.e. $G$ is the complete graph. Here, all the nodes of the graph interfere with each other, and consequently, in each time slot only one path can transmit signal. Hence, in order to achieve $\frac{L}{S} \to 1$, only the direct path $(0, K+1)$ should be utilized for almost all the time.

As an another example, consider the 3-hop network in Figure (3.1). As we will see in the following sections, the RS scheme corresponding to the path sequence $\mathrm{P}_1$ achieves the maximum diversity gain of the network, $d = 4$. However, it can easily be verified that no valid timing-sequence can achieve fewer number of transmission slots than the one shown in Table 3.1. Hence, $\frac{L}{S} = \frac{4}{7}$ is the best RS scheme can achieve with $\mathrm{P}_1$. On the other hand, consider the RS scheme with the path sequence $\mathrm{P}_2$. Although, as seen in the sequel, the scheme achieves the diversity gain $d = 2$ which is below the maximum diversity gain of the network, it utilizes fewer number of slots compared to the case using the path sequence $\mathrm{P}_1$. Indeed, it achieves $\frac{L}{S} = \frac{4}{6}$.

In the two-hop scenario investigated in the next section, we will see that for asymptotically large values of $L$, it is possible to utilize all the paths needed to achieve the maximum diversity gain and, at the same time, devise the timing sequence such that $\frac{L}{S} \to 1$. Consequently, it will be shown that in this setup, the proposed RS scheme achieves the optimum DMT.

## 3.4 Diversity-Multiplexing Tradeoff

In this section, we analyze the performance of the RS scheme in terms of the DMT for the single-antenna multiple-relay networks. First, in subsection 3.4.1, we study the performance of the RS scheme for the case of non-interfering relays where there exists neither causal nor noncausal interference between the signals sent through different paths. In this case, as there exists no interference between different paths, we can assume that the amplification coefficients take values greater than one, i.e. the constraint $\alpha_{i,j} \leq 1$ can be omitted. Under the condition of non-interfering relays, we derive the exact DMT of the RS scheme. As a result, we show that the RS scheme achieves the optimum DMT for the setup of non-interfering two-hop multiple-relay ($K > 1$) single-source single-destination, where there exists no direct link between the relay nodes and between the source and the destination (more precisely, $E = \{\{0, k\}, \{k, K+1\}\}_{k=1}^{K}$). To prove this, we assume that the RS scheme relies on $L = BK$ paths, $S = BK + 1$ slots, where $B$ is an integer number, and the path sequence is $\mathrm{Q} \equiv (\mathrm{q}_1, \ldots, \mathrm{q}_K, \mathrm{q}_1, \ldots, \mathrm{q}_K, \ldots, \mathrm{q}_1, \ldots, \mathrm{q}_K)$ where $\mathrm{q}_k \equiv (0, k, K+1)$. In other words, every path $\mathrm{q}_k$ is used $B$ times in the sequence.

Here, each $K$ consecutive slots are called a sub-block. Hence, the entire block of transmission consists of $B+1$ sub-blocks. The timing sequence is defined as $s_{i,j} = i+j-1$. It is easy to verify that the timing sequence satisfies the requirements. Here, we observe that the spectral efficiency is $\frac{L}{S} = 1 - \frac{1}{S}$ which converges to 1 for asymptotically large values of $S$. By deriving the exact DMT of the RS scheme, we prove that the RS scheme achieves the optimum DMT for asymptotically large values of $S$.

In subsection 3.4.2, we study the performance of the RS scheme for general single-antenna multiple-relay networks. First, we study the performance of RS scheme for the setup of two-hop single-source single-destination multiple-relay ($K > 1$) networks where there exists no direct link between the source and the destination; However, no additional restriction is imposed on the graph of the interfering relay pairs. We apply the RS scheme with the same parameters used in the case of two-hop non-interfering networks. We derive a lower-bound for DMT of the RS scheme. Interestingly, it turns out that the derived lower-bound merges with the upper-bound on the DMT for asymptotic values of $B$. Next, we generalize our result and derive a lower-bound on DMT of the RS scheme for general single-antenna multiple-relay networks.

Finally, in subsection 3.4.3, we generalize our results for the scenario of single-antenna two-hop multiple-access multiple-relay ($K > 1$) networks where there exists no direct link between the sources and the destination. Here, we apply the RS scheme with the same parameters as used in the case of single-source single-destination two-hop relay networks. However, it should be noted that here, instead of sending data from the single source, all the sources send data coherently. By deriving a lower-bound on the DMT of the RS scheme, we show that in this network the RS scheme achieves the optimum DMT. However, as studied in subsection 3.4.4, for the setup of single-antenna two-hop multiple-access single-relay networks where there exists no direct link between the sources and the destination, the proposed RS scheme reduces to naive amplify-and-forward relaying and is not optimum in terms of the DMT. In this setup, we show that the DDF scheme achieves the optimum DMT.

### 3.4.1 Non-Interfering Relays

In this subsection, we study the DMT behavior of the RS scheme in general single-antenna multi-hop relay networks under the condition that there exists neither causal nor noncausal interference between the signals transmitted over different paths. More precisely, we assume the timing sequence is designed such that if $s_{i,j} = s_{i',j'}$, then we have $\{\mathrm{p}_i(j), \mathrm{p}_{i'}(j'-1)\} \notin E$. This assumption is stronger than the fourth assumption on the timing sequence (here the condition $i < i'$ is omitted). We call this the "non-interfering relaying" condition. Under this condition, as there exists no interference between signals over different paths, we can assume that the amplification coefficients take values greater than one, i.e. the constraint $\alpha_{i,j} \leq 1$ can be omitted.

First, we need the following definition.

**Definition 3.1** *For a network with the connectivity graph $G = (V, E)$, an edge cut-set on $G$ is defined as a subset $\mathcal{S} \subseteq V$ such that $0 \in \mathcal{S}, K+1 \in \mathcal{S}^c$. The weight of the edge cut-set corresponding to $\mathcal{S}$, denoted by $w(\mathcal{S})$, is defined as*

$$w_G(\mathcal{S}) \quad = \sum_{a \in \mathcal{S}, b \in \mathcal{S}^c, \{a,b\} \in E} N_a N_b. \tag{3.2}$$

**Theorem 3.2** *Consider a half-duplex single-antenna multiple-relay network with the connectivity graph $G = (V, E)$. Assuming "non-interfering relaying", the RS scheme with the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ achieves the diversity gain corresponding to the following linear programming optimization problem*

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \sum_{e \in E} \mu_e, \tag{3.3}$$

*where $\boldsymbol{\mu}$ is a vector defined on edges of $G$ and $\hat{\mathcal{R}}$ is a region of $\boldsymbol{\mu}$ defined as*

$$\hat{\mathcal{R}} \equiv \left\{ \boldsymbol{\mu} \;\middle|\; \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1}, \sum_{i=1}^{L} \max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \geq L - Sr \right\}.$$

*Furthermore, the DMT of the RS scheme can be upper-bounded as*

$$d_{RS,NI}(r) \leq (1-r)^{+} \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{3.4}$$

*where $\mathcal{S}$ is an edge cut-set on $G$. Finally, by properly selecting the path sequence, one can always achieve*

$$d_{RS,NI}(r) \geq (1 - l_G r)^{+} \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{3.5}$$

*where $\mathcal{S}$ is an edge cut-set on $G$ and $l_G$ is the maximum path length between the source and the destination.*

**Proof** See Appendix E. ∎

**Remark 3.3** *In scenarios where the minimum edge cut-set on $G$ is achieved by a cut-set of the MISO or SIMO form, i.e., the edges that cross the cut-set are either originated from or destined to the same vertex, the upper-bound on the diversity gain of the RS scheme derived in* (3.4) *meets the information-theoretic upper-bound on the diversity gain of the network. Hence, in this scenario, any RS scheme that achieves* (3.4) *indeed achieves the optimum DMT.*

**Remark 3.4** *In general, the upper-bound* (3.4) *can be achieved for various certain graph topologies by wisely designing the path sequence and the timing sequence. One example is the case of the layered network [6] in which all the paths from the source to the destination have the same length* $l_G$*. Let us assume that the relays are allowed to operate in the full-duplex manner. In this case, it easily can be observed that the timing sequence corresponding to the path sequence* $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ *used in the proof of* (3.5) *can be modified to* $s_{i,j} = i + j - 1$*. Accordingly, the number of slots is decreased to* $S = L + l_G - 1$*. Rewriting* (E.10)*, we have* $d_{RS,NI}(r) = \left(1 - r - \frac{l_G - 1}{L} r\right)^+ \min_{\mathcal{S}} w_G(\mathcal{S})$ *which achieves* $(1-r)^+ \min_{\mathcal{S}} w_G(\mathcal{S})$ *for large values of* $L$*.*

**Example 3.5** *Consider the half-duplex* 3 *hops network in Figure 3.1. Let us assume all the nodes are having single-antenna, i.e.* $N_i = 1$*. Here, the minimum edge cut-set is achieved by the MISO and SIMO cuts disconnecting the source or the destination from the other nodes. As a result,* $2(1-r)^+$ *is an upper-bound for DMT of the network. However, the RS schemes with the timing-sequence depicted in Tables 3.1 and 3.2 are interfering. Hence, the argument of Theorem 3.2 can not be applied to analyze their achievable DMT. Yet, Theorem 3.2 can be applied. Indeed, as* $l_G = 5$*, from* (3.5)*, there exists a non-interfering RS scheme with DMT greater than or equal to* $2(1 - \frac{r}{5})^+$*. Moreover, as the maximum-flow of the graph can be obtained by the union of two disjoint paths* $(0, 1, 3, 5)$ *and* $(0, 2, 4, 5)$ *each with length* 3*, the lower-bound on the DMT of non-interfering RS scheme can be improved to* $2(1 - \frac{r}{3})^+$*.*

**Example 3.6** *Consider the directed half-duplex* 3 *hops network in Figure 3.2. Here, similar to Example 3.5, the minimum edge cut-set is achieved by the MISO*

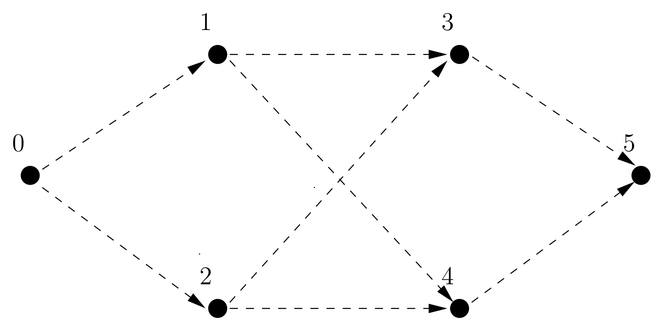


Figure 3.2: An example of a directed 3 hops network where $\forall i : N_i = 1$.

*and SIMO cuts disconnecting the source or the destination from the other nodes. As a result,* $2(1-r)^+$ *is an upper-bound for DMT of the network. It can be easily verified that, unlike for the graph of Figure 3.1, the RS schemes with the timing-sequence depicted in Tables 3.1 and 3.2 are non-interfering. Applying Theorem 3.2, DMT of the RS schemes with the timing-sequence in Tables 3.1 and 3.2 are* $2(1-\frac{7}{4}r)^+$ *and* $2(1-\frac{3}{2}r)^+$*, correspondingly. Moreover, the RS scheme with the timing sequence in Table 3.2 can be extended to* $M$ *repetitions of* $\mathrm{p}_1, \mathrm{p}_2$ *as follows. The path sequence consists of* $L = 2M$ *paths* $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_1, \mathrm{p}_2)$ *and the timing-sequence is equal to* $s_{i,j} = i + j - 1$*. As the RS scheme still remains non-interfering, the argument of Theorem 3.2 can be applied to analyze its achievable DMT. Accordingly, DMT of the above RS scheme equals* $2(1 - \frac{M+1}{M}r)^+$*. As a result, as* $M$ *goes to infinity, the RS scheme achieves the optimum DMT,* $2(1-r)^+$*.*

Next, using Theorem 3.2, we show that the RS scheme achieves the optimum DMT in the setup of single-antenna two-hop multiple-relay networks where there exists no direct link neither between the source and the destination, nor between the relay nodes.

**Theorem 3.7** *Assume a single-antenna half-duplex parallel relay scenario with* $K$

*non-interfering relays. The proposed RS scheme with $L = BK$, $S = BK + 1$, the path sequence*

$$\mathrm{Q} \equiv (\mathrm{q}_1, \dots, \mathrm{q}_K, \mathrm{q}_1, \dots, \mathrm{q}_K, \dots, \mathrm{q}_1, \dots, \mathrm{q}_K)$$

*where $\mathrm{q}_k \equiv (0, k, K+1)$ and the timing sequence $s_{i,j} = i + j - 1$ achieves the diversity gain*

$$d_{RS,NI}(r) = \max\left\{0, K(1-r) - \frac{r}{B}\right\}, \tag{3.6}$$

*which achieves the optimum DMT curve $d_{opt}(r) = K(1-r)^+$ as $B \to \infty$.*

**Proof** See Appendix F. ∎

**Remark 3.8** *Note that as long as the complement*[6] *of the induced sub-graph of $G$ on the relay nodes $\{1, 2, \dots, K\}$ includes a Hamiltonian cycle* [7]*, the result of Theorem 3.7 remains valid. However, the paths $\mathrm{q}_1, \mathrm{q}_2, \dots, \mathrm{q}_K$ should be permuted in the path sequence according to their orderings in the corresponding Hamiltonian cycle.*

According to (3.6), we observe that the RS scheme achieves the maximum multiplexing gain $1 - \frac{1}{BK+1}$ and the maximum diversity gain $K$, respectively, for the setup of non-interfering relays. Hence, it achieves the maximum diversity gain for any finite value of $B$. Also, knowing that no signal is sent to the destination in the first slot, the RS scheme achieves the maximum possible multiplexing gain. Figure (3.3) shows the DMT of the scheme for the case of non-interfering relays and various values of $K$ and $B$.

[6] For every undirected graph $G = (V, E)$, the complement of $G$ is a graph $H$ on the same vertices such that two vertices of $H$ are adjacent if and only if they are non-adjacent in $G$. [14]

[7] A Hamiltonian cycle is a simple cycle $(v_1, v_2, \cdots, v_K, v_1)$ that goes exactly one time through each vertex of the graph [14].

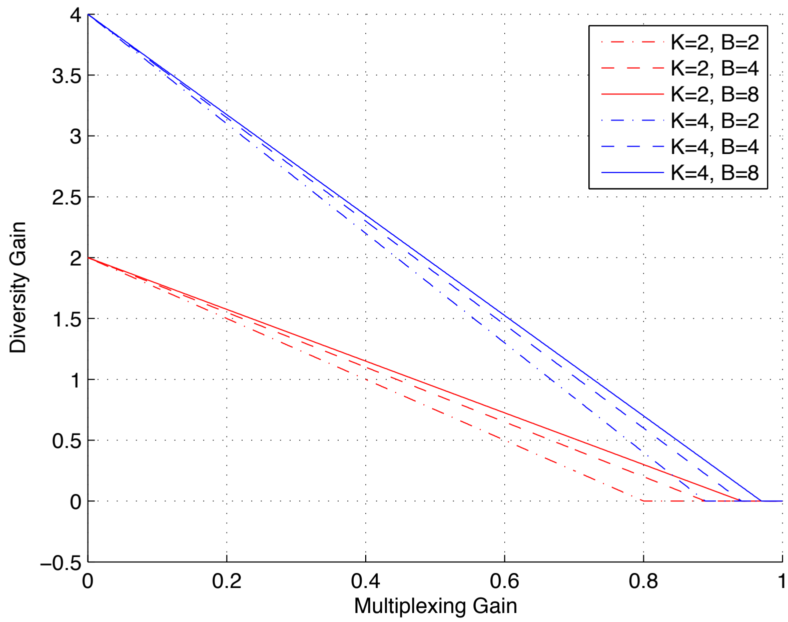


Figure 3.3: DMT of RS scheme in parallel relay network for both "interfering" and "non-interfering" relaying scenarios and for different values of $K, B$.

### 3.4.2 General Case

In this subsection, we study the performance of the RS scheme in general single-antenna multi-hop wireless networks and derive a lower bound on the corresponding DMT. First, we show that the RS scheme with the parameters defined in Theorem 3.7 achieves the optimum DMT for the single-antenna parallel-relay networks when there is no direct link between the source and the destination. Then, we generalize the statement and provide a lower-bound on the DMT of the RS scheme for the more general case.

As stated in the section "System Model", throughout the two-hop network analysis, we slightly modify our notations to simplify the derivations. Specifically, the output vector at the source, the input and the output vectors at the $k$'th relay, and the input vector at the destination are denoted as $\mathbf{x}$, $\mathbf{r}_k$, $\mathbf{t}_k$ and $\mathbf{y}$, respectively. $h_k$ and $g_k$ represent the channel gain between the source and the $k$'th relay and the channel gain between the $k$'th relay and the destination, respectively. $(k)$ and $(b)$ are defined as $(k) \equiv ((k-2) \mod K) + 1$ and $(b) \equiv b - \lfloor \frac{(k)}{K} \rfloor$. Finally, $i_{(k)}$, $\mathbf{n}_k$, $\mathbf{z}$, and $\alpha_k$ denote the channel gain between the $k$'th and the $(k)$'th relay nodes, the noise at the $k$'th relay and at the destination, and the amplification coefficient at the $k$'th relay.

Figure (3.4) shows a realization of this setup with 4 relays. As observed, the relay set $\{1, 2\}$ is disconnected from the relay set $\{3, 4\}$. In general, the output signal of any relay node $k'$ such that $\{k, k'\} \in E$ can interfere on the received signal of relay node $k$. However, in Theorem 3.9, the RS scheme is applied with the same parameters as in Theorem 3.7. Hence, when the source is sending signal to the $k$'th relay in a time-slot, just the $(k)$'th relay is simultaneously transmitting and interferes at the $k$'th relay side. As an example, for the scenario shown in Figure (3.4), we have

$$\begin{aligned} \mathbf{r}_1 &= h_1\mathbf{x} + i_4\mathbf{t}_4 + \mathbf{n}_1, \\ \mathbf{r}_2 &= h_2\mathbf{x} + \mathbf{n}_2. \end{aligned}$$

However, for the sake of simplicity, in the proof of the following Theorem, we assume that all the relays interfere with each other. Hence, at the $k$'th relay, we have

$$\mathbf{r}_k = h_k\mathbf{x} + i_{(k)}\mathbf{t}_{(k)} + \mathbf{n}_k. \tag{3.7}$$

According to the output power constraint, the amplification coefficient is bounded as $\alpha_k \leq \sqrt{\frac{P}{P\left(|h_k|^2 + \left|i_{(k)}\right|^2\right)+1}}$. However, according to the signal boosting constraint

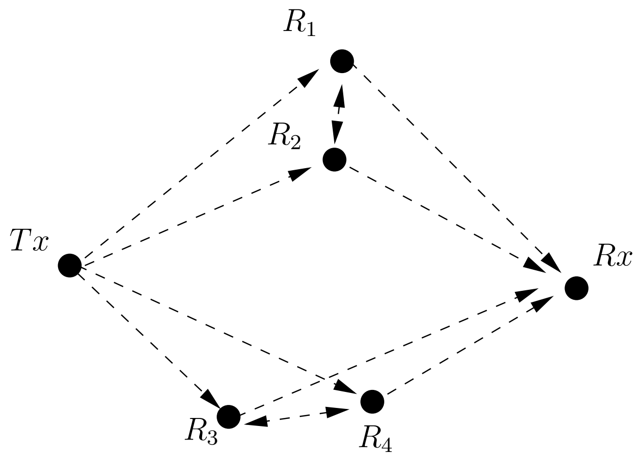


Figure 3.4: An example of the half-duplex parallel relay network setup, relay nodes $\{1, 2\}$ are disconnected from relay nodes $\{3, 4\}$.

imposed on the RS scheme, we have $|\alpha_k| \leq 1$. Hence, the amplification coefficient is equal to

$$\alpha_k = \min\left\{1, \sqrt{\frac{P}{P\left(|h_k|^2 + \left|i_{(k)}\right|^2\right) + 1}}\right\}. \tag{3.8}$$

In this manner, it is guaranteed that the noise terms of the different relays are not boosted throughout the network. This is achieved at the cost of working with the output power less than $P$. On the other hand, we know that almost surely [8] $|h_k|^2, \left|i_{(k)}\right|^2 \dot{\leq} 1$. Hence, almost surely, we have $\alpha_k \doteq 1$. This point will be elaborated further in the proof of the Theorem. Now, we prove the DMT optimality of the RS scheme for general single-antenna parallel-relay networks.

**Theorem 3.9** *Consider a single-antenna half-duplex parallel relay network with $K > 1$ interfering relays where there is no direct link between the source and the*

[8] By almost surely, we mean its probability is greater than $1 - P^{-\delta}$, for any value of $\delta > 0$.

*destination. The diversity gain of the RS scheme with the parameters defined in Theorem 3.7 is lower-bounded as*

$$d_{RS,I}(r) \geq \max\left\{0, K(1-r) - \frac{r}{B}\right\}. \tag{3.9}$$

*Furthermore, the RS scheme achieves the optimum DMT $d_{opt}(r) = K(1-r)^+$ as $B \to \infty$.*

**Proof** See Appendix G. ∎

**Remark 3.10** *The argument in Theorem 3.9 is valid no matter what the induced graph of $G$ on the relay nodes is. More precisely, the DMT of the RS scheme can be lower-bounded as* (3.9) *as long as $\{0, K+1\} \notin E$ and $\{0, k\}, \{K+1, k\} \in E$. One special case is that the complement of the induced subgraph of $G$ on the relay nodes includes a Hamiltonian cycle which is analyzed in Theorem 3.7. Here, we observe that the lower-bound on DMT derived in* (3.9) *is tight as shown in Theorem 3.7.*

Figure (3.3) shows the DMT of the RS scheme for varying number of $K$ and $B$. Noting the proof of Theorem 3.9, we can easily generalize the result of Theorem 3.9 and provide a lower-bound on the DMT of the RS scheme for general single-antenna multi-hop multiple-relay networks.

**Theorem 3.11** *Consider a half-duplex single-antenna multiple-relay network with the connectivity graph $G = (V, E)$ operated under the RS scheme with $L$ paths, $S$ slots, and the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$. Defining $\beta_e$ for each $e \in E$ as the number of paths in the path sequence that go through $e$, then the DMT of the RS* *scheme is lower-bounded as*

$$d_{RS}(r) \geq \frac{L}{\max\limits_{e \in E} \beta_e} \left(1 - \frac{S}{L} r\right)^+ . \tag{3.10}$$

**Proof** First, similar to the proof of Theorem 3.9, we show that the entire channel matrix is lower triangular. At the destination side, we have

$$\mathbf{y}_{K+1,i} = \prod_{j=1}^{l_i} h_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \alpha_{i,j} \mathbf{x}_{0,i} + \sum_{j<i} f_{i,j} \mathbf{x}_{0,j} + \sum_{j \leq i, m \leq l_j} q_{i,j,m} \mathbf{n}_{j,m}. \tag{3.11}$$

Here, $\mathbf{x}_{0,i}$ is the vector transmitted at the source side during the $s_{i,1}$'th slot as the input for the $i$'th path, $\mathbf{y}_{K+1,i}$ is the vector received at the destination side during the $s_{i,l_i}$'th slot as the output for $i$'th path, $f_{i,j}$ is the interference coefficient which relates the input of the $j$'th path ($j < i$) to the output of the $i$'th path, $\mathbf{n}_{j,m}$ is the noise vector during the $s_{j,m}$'th slot at the $\mathrm{p}_j(m)$'th node, and finally, $q_{i,k,m}$ is the coefficient which relates $\mathbf{n}_{k,m}$ to $\mathbf{y}_{K+1,i}$. Note that as the timing sequence satisfies the noncausal interference assumption, the summation terms in (3.11) do not exceed $i$. Moreover, for the sake of brevity, we define $\alpha_{i,l_i} = 1$. Defining $\mathbf{x}(s) = [x_{0,1}(s)\, x_{0,2}(s) \cdots x_{0,L}(s)]^T$, $\mathbf{y}(s) = [y_{K+1,1}(s)\, y_{K+1,2}(s) \cdots y_{K+1,L}(s)]^T$, and $\mathbf{n}(s) = [n_{1,1}(s)\, n_{1,2}(s) \cdots n_{L,l_L}(s)]^T$, we have the following equivalent lower-triangular matrix between the end nodes:

$$\mathbf{y}(s) = \mathbf{H}_T \mathbf{x}(s) + \mathbf{Q}\mathbf{n}(s). \tag{3.12}$$

Here,

$$\mathbf{H}_T = \begin{pmatrix} f_{1,1} & 0 & 0 & \dots \\ f_{2,1} & f_{2,2} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ f_{L,1} & f_{L,2} & \dots & f_{L,L} \end{pmatrix}, \tag{3.13}$$

where $f_{i,i} = \prod_{j=1}^{l_i} h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\alpha_{i,j}$, and

$$\mathbf{Q} = \begin{pmatrix} q_{1,1,1} & \cdots & q_{1,1,l_1} & 0 & 0 & 0 & \cdots \\ q_{2,1,1} & \cdots & q_{2,1,l_1} & \cdots & q_{2,2,l_2} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ q_{L,1,1} & q_{L,1,2} & \cdots & \cdots & \cdots & q_{L,L,l_L-1} & q_{L,L,l_L} \end{pmatrix}. \tag{3.14}$$

Let us define $\mu_e$ for every $e \in E$ such that $|h_e|^2 = P^{-\mu_e}$. First, we observe that similar to the proof of Theorem 3.9, it can be shown that i) $\alpha_{i,j} \doteq 1$ with probability 1[9], ii) we can restrict ourselves to the region $\mathbb{R}_+$, i.e., the region $\boldsymbol{\mu} > \mathbf{0}$. These two facts imply that $|q_{i,j,m}| \dot{\leq} 1$. This means there exists a constant $c$ which depends just on the topology of the graph $G$ and the path sequence such that $\mathbf{P}_n \triangleq \mathbf{Q}\mathbf{Q}^H \preccurlyeq c\mathbf{I}_L$ (by a similar argument as in the proof of Theorem 3.9). Hence, similar to the arguments in the equation series (G.5), the outage probability can be bounded as

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} &= \mathbb{P}\left\{\left|\mathbf{I}_L + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \leq P^{Sr}\right\} \\ &\dot{\leq} \mathbb{P}\left\{|\mathbf{H}_T|\left|\mathbf{H}_T^H\right| \leq P^{Sr-L}\right\} \\ &= \mathbb{P}\left\{\sum_{e\in E}\beta_e\mu_e \geq L - Sr\right\} \\ &\doteq \mathbb{P}\left\{\boldsymbol{\mu} \geq \mathbf{0}, \sum_{e\in E}\beta_e\mu_e \geq (L-Sr)^+\right\}, \end{aligned} \tag{3.15}$$

where $\beta_e$ is the number of paths in the path sequence that pass through $e$. Knowing that $\mathbb{P}\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0\} \doteq P^{-\mathbf{1}\cdot\boldsymbol{\mu}}$ and computing the derivative, we have $f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) = P^{-\mathbf{1}\cdot\boldsymbol{\mu}}$. Defining $\mathcal{R} = \left\{\boldsymbol{\mu} > \mathbf{0}, \sum_{e\in E}\beta_e\mu_e \geq (L-Sr)^+\right\}$ and applying the results of equa-

[9] More precisely, with probability greater than $1 - P^{-\delta}$, for any $\delta > 0$.

tion series (G.6), we obtain

$$\mathbb{P}\{\mathcal{E}\} \ \dot{\leq} \ P^{-\min\limits_{\boldsymbol{\mu}\in\mathcal{R}} \mathbf{1}\cdot\boldsymbol{\mu}} \stackrel{(a)}{\doteq} P^{-\frac{L}{\max_{e\in E}\beta_e}\left(1-\frac{S}{L}r\right)^+}, \tag{3.16}$$

where $(a)$ follows from the fact that for every $\boldsymbol{\mu} \in \mathcal{R}$, $(L-Sr)^+ \leq \sum_{e\in E}\beta_e\mu_e \leq \max_{e\in E}\beta_e \sum_{e\in E}\mu_e$ which implies that $\sum_{e\in E}\mu_e = \mathbf{1}\cdot\boldsymbol{\mu} \geq \frac{(L-Sr)^+}{\max_{e\in E}\beta_e}$, and on the other hand, defining $\boldsymbol{\mu}^\star$ such that $\boldsymbol{\mu}^\star(\hat{e}) = \frac{(L-Sr)^+}{\beta_{\hat{e}}}$ where $\hat{e} = \underset{e\in E}{\operatorname{argmax}}\ \beta_e$ and otherwise $\boldsymbol{\mu}^\star(e) = 0$, we have $\boldsymbol{\mu}^\star \in \mathcal{R}$ and $\mathbf{1}\cdot\boldsymbol{\mu}^\star = \frac{L}{\max_{e\in E}\beta_e}\left(1-\frac{S}{L}r\right)^+$. (3.16) completes the proof of Theorem 3.11. ∎

**Remark 3.12** *The lower-bound of* (3.5) *can also be proved by using the lower-bound of* (3.10) *obtained for DMT of the general RS scheme. In order to prove this, one needs to apply the RS scheme with the same path sequence and timing sequence used in the proof of* (3.5) *in Theorem 3.2. Putting $S = L_0 d_G$ and $S \leq l_G L$ in* (3.10) *and noting that for all $e \in E$, we have $\beta_e \in \{0, L_0\}$,* (3.5) *is easily obtained.*

**Remark 3.13** *It should be noted that* (3.4) *is still an upper-bound for the DMT of the RS scheme, i.e., even for the case of interfering relays. This is due to the fact that in the proof of* (3.4) *the non-interfering relaying assumption is not used. However, by employing the RS scheme with causal-interfering relaying and applying* (3.10)*, one can find a bigger family of graph topologies that can achieve* (3.4)*. Such an example is the two-hop relay network studied in Theorem 3.9. Another example is the case that $G$ is a directed acyclic graph (DAG)*[10] *and the relays are operating in the full-duplex mode. Here, the argument is similar to that of Remark 3.4. Assume*

[10]A directed acyclic graph $G$ is a directed graph that has no directed cycles.

*that each* $\hat{\mathrm{p}}_{\mathrm{i}}$ *is used* $L_0$ *times in the path sequence in the form that* $\mathrm{p}_{(i-1)L_0+j} \triangleq \hat{\mathrm{p}}_{\mathrm{i}}, 1 \leq j \leq L_0$*. Let us modify the timing sequence as* $s_{i,j} = i + j - 1 + \sum_{k=1}^{\lceil \frac{i}{L_0} \rceil - 1} \hat{l}_k$ *which results in* $S = L + \sum_{i=1}^{d_G} l_i$*. Here, it is easy to verify that only non-causal interference exists between the signals corresponding to different paths. However, by considering the paths in the reverse order or equivalently reversing the time axis, the paths can be observed with the causal interference. Hence, the result of Theorem 3.11 is still valid for such paths. Here, knowing that for all* $e \in E$*, we have* $\beta_e \in \{0, L_0\}$ *and applying* (3.10)*, we have* $d_{RS}(r) \geq d_G \left(1 - r - \frac{\sum_{i=1}^{d_G} l_i}{L_0 d_G}\right)^+$ *which achieves* (3.4) *for asymptotically large values of* $L_0$*.*

**Remark 3.14** *In this remark, we address the problem of designing a proper timing sequence for the RS scheme. Let us consider the RS scheme with a fixed path sequence. From Theorem 3.11, we observe that the objective is to maximize* $\frac{L}{\max e \in E \beta_e} \left(1 - \frac{S}{L}\right)^+$*. As the path sequence is fixed, we conclude that* $\beta_e$ *is determined for every edge of the graph. As a result, the objective is to determine a proper timing-sequence, i.e. satisfying the ordering, causality, and no noncausal interference assumptions, such that the objective function* $\frac{S}{L}$ *is minimized. In other words, the number of used time-slots,* $S$*, should be minimized. It turns out that this optimization problem can be solved by applying Dynamic Programming with a polynomial time complexity in terms of* $V$ *and* $L$*.*

**Example 3.15** *Consider the half-duplex* 3 *hops network of Figure 3.1 assuming all nodes having single-antenna, i.e.,* $\forall i : N_i = 1$*. According to the argument of Theorem 3.11, the RS schemes with the path sequences* $\mathrm{P}_1$ *and* $\mathrm{P}_2$ *and the timing sequences depicted in Tables 3.1 and 3.2 achieve DMT at least greater than or equal to* $2(1 - \frac{7}{4}r)^+$ *and* $2(1 - \frac{3}{2}r)^+$*, respectively. This is due to the fact that for* *both RS schemes, we have* $\max_{e\in E}\beta_e = \beta_{\{3,5\}} = 2$*. Similar to the argument of Example 3.6, the RS scheme with the path sequence* $\mathrm{P}_2$ *and the timing sequence of Table 3.2 can be extended to* $M$ *repetitions of* $\mathrm{p}_1, \mathrm{p}_2$*. Accordingly, the RS scheme achieves the optimum DMT of the network in Figure 3.1 which is* $2(1-r)^+$*. Notice that assuming non-interfering relaying, Example 3.5 could only show that the RS scheme achieves* $2(1-\frac{r}{3})^+$*.*

### 3.4.3 Multiple-Access Parallel Relays Scenario

In this subsection, we generalize the result of Theorem 3.9 to the multiple-access scenario aided by multiple relay nodes. Here, similar to Theorem 3.9, we assume that there is no direct link between each source and the destination. However, no restriction is imposed on the induced subgraph of $G$ on the relay nodes. Assuming having $M$ disjoint source nodes, we show that for the rate sequence $r_1\log(P), r_2\log(P), \ldots, r_M\log(P)$, in the asymptotic case of $B \to \infty$ ($B$ is the number of sub-blocks), the RS scheme achieves DMT of $d_{RS,MAC}(r_1, r_2, \ldots, r_M) = K\left(1-\sum_{m=1}^{M} r_m\right)^+$, which is shown to be optimum due to the cut-set bound on the cut-set between the relays and the destination. Here, the notations are slightly modified compared to the ones used in Theorem 3.9 to emphasize the fact that multiple signals are transmitted from multiple sources. Throughout this subsection and the next one, $\mathbf{x}_m$ and $h_{m,k}$ denote the transmitted vector at the $m$'th source and the Rayleigh channel coefficient between the $m$'th source and the $k$'th relay, respectively. Hence, at the received side of the $k$'th relay, we have

$$\mathbf{r}_k = \sum_{m=1}^{M} h_{m,k}\mathbf{x}_m + i_{(k)}\mathbf{t}_{(k)} + \mathbf{n}_k, \tag{3.17}$$

where $\mathbf{x}_m$ is the transmitted vector of the $m$'th sender. The amplification coefficient at the $k$'th relay is set to

$$\alpha_k = \min\left\{1, \sqrt{\frac{P}{P\left(\sum_{m=1}^{M} |h_{m,k}|^2 + \left|i_{(k)}\right|^2\right) + 1}}\right\}. \tag{3.18}$$

Here, the RS scheme is applied with the same path sequence and timing sequence as in the case of Theorem 3.7 and 3.9. However, it should be mentioned that in the current case, during the slots that the source is supposed to transmit the signal, i.e. in the $s_{i,1}$'th slot, all the sources send their signals coherently. Moreover, at the destination side, after receiving the $BK$ vectors corresponding to the outputs of the $BK$ paths, the destination node decodes the messages $\omega_1, \omega_2, \ldots, \omega_K$ by joint-typical decoding of the received vectors in the corresponding $BK$ slots and the transmitted signal of all the sources, i.e., in the same way that joint-typical decoding works in the multiple access setup [12]. Now, we prove the main result of this subsection.

**Theorem 3.16** *Consider a multiple-access channel consisting of $M$ transmitting nodes aided by $K > 1$ half-duplex relays. Assume there is no direct link between the sources and the destination. The RS scheme with the path sequence and timing sequence defined in Theorems 3.7 and 3.9 achieves a diversity gain of*

$$d_{RS,MAC}(r_1, r_2, \ldots, r_M) \geq \left[K\left(1 - \sum_{m=1}^{M} r_m\right) - \frac{\sum_{m=1}^{M} r_m}{B}\right]^+, \tag{3.19}$$

*where $r_1, r_2, \ldots, r_M$ are the multiplexing gains corresponding to users $1, 2, \ldots, M$. Moreover, as $B \to \infty$, it achieves the optimum DMT which is $d_{opt,MAC}(r_1, r_2, \ldots, r_M) = K\left(1 - \sum_{m=1}^{M} r_m\right)^+$.*

**Proof** At the destination side, we have

$$\begin{aligned}
\mathbf{y}_{b,k} &= g_{(k)}\mathbf{t}_{b,k} + \mathbf{z}_{b,k} \\
&= g_{(k)}\alpha_{(k)}\left(\sum_{\substack{1\le b_1\le b, 1\le k_1\le K\\ b_1K+k_1<bK+k}} p_{b-b_1,k,k_1}\left(\sum_{m=1}^{M} h_{m,k_1}\mathbf{x}_{m,b_1,k_1} + \mathbf{n}_{b_1,k_1}\right)\right) + \mathbf{z}_{b,k},
\end{aligned} \tag{3.20}$$

where $p_{b,k,k_1}$ is defined in the proof of Theorem 3.9 and $\mathbf{x}_{m,b,k}$ represents the transmitted signal of the $m$'th sender in the $k$'th slot of the $b$'th sub-block. Similar to (G.3), we have

$$\mathbf{y}(s) = \mathbf{G\Omega F}\left(\sum_{m=1}^{M} \mathbf{H}_m\mathbf{x}_m(s) + \mathbf{n}(s)\right) + \mathbf{z}(s), \tag{3.21}$$

where $\mathbf{y}_s, \mathbf{n}_s, \mathbf{z}_s, \mathbf{G}, \mathbf{\Omega}, \mathbf{F}$ are defined in the proof of Theorem 3.9, $\mathbf{H}_m = \mathbf{I}_B \otimes \mathit{diag}\{h_{m,1}, h_{m,2}, \cdots, h_{m,K}\}$ and $\mathbf{x}_m(s) = [x_{m,1,1}(s), x_{m,1,2}(s), \cdots, x_{m,B,K}(s)]^T$. Similarly, we observe that the entire channel from each of the sources to the destination acts as a MIMO channel with a lower triangular matrix of size $BK \times BK$.

Here, the outage event occurs whenever there exists a subset $\mathcal{S} \subseteq \{1, 2, \ldots, M\}$ of the sources such that

$$I\left(\mathbf{x}_{\mathcal{S}}(s); \mathbf{y}(s)|\mathbf{x}_{\mathcal{S}^c}(s)\right) \le (BK+1)\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P). \tag{3.22}$$

This event is equivalent to

$$\log\left|\mathbf{I}_{BK} + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \le (BK+1)\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P). \tag{3.23}$$

where $\mathbf{P}_n$ is defined in the proof of Theorem 3.9, $\mathbf{H}_T = \mathbf{G\Omega F H}_{\mathcal{S}}$, and

$$\mathbf{H}_{\mathcal{S}} = \mathbf{I}_B \otimes \mathit{diag}\left\{\sqrt{\sum_{m\in\mathcal{S}} |h_{m,1}|^2}, \sqrt{\sum_{m\in\mathcal{S}} |h_{m,2}|^2}, \cdots, \sqrt{\sum_{m\in\mathcal{S}} |h_{m,K}|^2}\right\}. \tag{3.24}$$

Defining such an event as $\mathcal{E}_{\mathcal{S}}$ and the outage event as $\mathcal{E}$, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}\right\} &= \mathbb{P}\left\{\bigcup_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathcal{E}_{S}\right\} \\ &\leq \sum_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\mathcal{E}_{S}\right\} \\ &\leq \left(2^{M}-1\right) \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\mathcal{E}_{S}\right\} \\ &\doteq \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\mathcal{E}_{S}\right\}. \end{aligned} \tag{3.25}$$

Hence, it is sufficient to upper-bound $\mathbb{P}\left\{\mathcal{E}_{S}\right\}$ for all $\mathcal{S}$.

Defining $\hat{\mathbf{H}}_{\mathcal{S}} = \mathbf{I}_{B} \otimes diag\left\{\max_{m\in\mathcal{S}}\left|h_{m,1}\right|, \max_{m\in\mathcal{S}}\left|h_{m,2}\right|, \cdots, \max_{m\in\mathcal{S}}\left|h_{m,K}\right|\right\}$, we have $\hat{\mathbf{H}}_{\mathcal{S}}\hat{\mathbf{H}}_{\mathcal{S}}^{H} \preccurlyeq \mathbf{H}_{\mathcal{S}}\mathbf{H}_{\mathcal{S}}^{H}$. Therefore,

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{S}\right\} &\leq \mathbb{P}\Bigg\{\log\left|\mathbf{I}_{BK} + P\mathbf{G}\boldsymbol{\Omega}\mathbf{F}\hat{\mathbf{H}}_{\mathcal{S}}\hat{\mathbf{H}}_{\mathcal{S}}^{H}\mathbf{F}^{H}\boldsymbol{\Omega}^{H}\mathbf{G}^{H}\mathbf{P}_{n}^{-1}\right| \leq \\ &\qquad (BK+1)\left(\sum_{m\in\mathcal{S}} r_{m}\right)\log\left(P\right)\Bigg\} \\ &\triangleq \mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\}. \end{aligned} \tag{3.26}$$

Assume $\max_{m\in\mathcal{S}}\left|h_{m,k}\right|^{2} = P^{-\mu_{k}}$, and $\left|g_{k}\right|^{2} = P^{-\nu_{k}}$, $\left|i_{k}\right|^{2} = P^{-\omega_{k}}$, and $\mathcal{R}$ as the region in $\mathbb{R}^{3K}$ that defines the outage event $\hat{\mathcal{E}_{\mathcal{S}}}$ in terms of the vector $[\boldsymbol{\mu}^{T}, \boldsymbol{\nu}^{T}, \boldsymbol{\omega}^{T}]^{T}$. Similar to the proof of Theorem 3.9, we have $\mathbb{P}\left\{\mathcal{R}\right\} \doteq \mathbb{P}\left\{\mathcal{R}_{+}\right\}$ where $\mathcal{R}_{+} = \mathcal{R}\bigcap\mathbb{R}_{+}^{3K}$. Rewriting the equation series of (G.5), we have

$$\begin{aligned}\mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} &\dot{\leq} \mathbb{P}\Bigg\{-BK\frac{\log\left[3\left(B^{2}K^{2}+1\right)\right]}{\log(P)} + BK\left(1-\sum_{m\in\mathcal{S}} r_{m}\right) - \sum_{m\in\mathcal{S}} r_{m} \leq \\ &\qquad B\sum_{k=1}^{K}\left(\mu_{k}+\nu_{k}\right), \mu_{k}, \nu_{k}, \omega_{k} \geq 0\Bigg\}. \end{aligned} \tag{3.27}$$

On the other hand, as $\{h_{m,k}\}$'s are independent random variables, we conclude that for $\boldsymbol{\mu}^{0}, \boldsymbol{\nu}^{0} \geq \mathbf{0}$, we have $\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^{0}, \boldsymbol{\nu} \geq \boldsymbol{\nu}^{0}\right\} \doteq P^{-\mathbf{1}\cdot\left(|\mathcal{S}|\boldsymbol{\mu}^{0}+\boldsymbol{\nu}^{0}\right)}$. Similar to the proof of Theorem 3.9, by computing the derivative with respect to $\boldsymbol{\mu}, \boldsymbol{\nu}$, we have $f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \doteq P^{-\mathbf{1}\cdot(|\mathcal{S}|\boldsymbol{\mu}+\boldsymbol{\nu})}$. Defining $l_0 \triangleq -\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + \left(1 - \sum_{m\in\mathcal{S}} r_m\right) - \frac{\sum_{m\in\mathcal{S}} r_m}{BK}$, the region $\hat{\mathcal{R}}$ as $\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\mu}, \boldsymbol{\nu} \geq \mathbf{0}, \frac{1}{K}\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu}) \geq l_0\right\}$, the cube $\mathcal{I}$ as $\mathcal{I} \triangleq [0, Kl_0]^{2K}$, and for $1 \leq i \leq 2K$, $\mathcal{I}_i^c = [0,\infty)^{i-1} \times [Kl_0, \infty) \times [0,\infty)^{2K-i}$, we have

$$\begin{aligned}
\mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} &\overset{(a)}{\dot{\leq}} \mathbb{P}\{\hat{\mathcal{R}}\} \\
&\leq \int_{\hat{\mathcal{R}}\bigcap\mathcal{I}} f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu},\boldsymbol{\nu})\, d\boldsymbol{\mu} d\boldsymbol{\nu} + \sum_{i=1}^{2K} \mathbb{P}\left\{[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T]^T \in \hat{\mathcal{R}} \cap \mathcal{I}_i^c\right\} \\
&\dot{\leq} \operatorname{vol}(\hat{\mathcal{R}} \cap \mathcal{I}) P^{-\min\limits_{[\boldsymbol{\mu},\boldsymbol{\nu}]\in\hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot(|\mathcal{S}|\,\boldsymbol{\mu}+\boldsymbol{\nu})} + 2KP^{-Kl_0} \\
&\overset{(b)}{\doteq} P^{-Kl_0} \\
&\doteq P^{-\left[K\left(1-\sum_{m\in\mathcal{S}} r_m\right) - \frac{\sum_{m\in\mathcal{S}} r_m}{B}\right]}.
\end{aligned} \tag{3.28}$$

Here, (a) follows from (3.27) and (b) follows from the fact that $\hat{\mathcal{R}}\bigcap\mathcal{I}$ is a bounded region whose volume is independent of $P$ and the fact that $\min\limits_{[\boldsymbol{\mu},\boldsymbol{\nu}]\in\hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot(|\mathcal{S}|\,\boldsymbol{\mu}+\boldsymbol{\nu}) = Kl_0$, which is achieved by having $\boldsymbol{\mu} = \mathbf{0}$. Comparing (3.25), (3.26) and (3.28), we observe

$$\mathbb{P}\{\mathcal{E}\} \dot{\leq} \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\{\mathcal{E}_S\} \dot{\leq} \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\hat{\mathcal{E}}_{\mathcal{S}}\right\} \dot{\leq} P^{-\left[K\left(1-\sum_{m=1}^M r_m\right) - \frac{\sum_{m=1}^M r_m}{B}\right]}. \tag{3.29}$$

Next, we prove that $K\left(1 - \sum_{m=1}^M r_m\right)^+$ is an upper-bound on the diversity gain of the system corresponding to the sequence of rates $r_1, r_2, \ldots, r_M$. We have

$$\mathbb{P}\{\mathcal{E}\} \geq \mathbb{P}\left\{\max_{p(\mathbf{t}_1,\mathbf{t}_2,\ldots,\mathbf{t}_K)} I(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_K; \mathbf{y}) \leq \left(\sum_{m=1}^M r_m\right)\log(P)\right\} \overset{(a)}{\doteq} P^{-K\left(1-\sum_{m=1}^M r_m\right)^+}. \tag{3.30}$$

Here, (a) follows from the DMT of the point-to-point MISO channel proved in [29]. This completes the proof. ∎

**Remark 3.17** *The argument of Theorem 3.16 is valid for the general case in which any arbitrary set of relay pairs are non-interfering.*

**Remark 3.18** *In the symmetric situation for which the multiplexing gains of all the users are equal to say $r$, the lower-bound in* (3.19) *takes a simple form. First, we observe that the maximum multiplexing gain which is simultaneously achievable by all the users is $\frac{1}{M} \cdot \frac{BK}{BK+1}$. Noting that no signal is sent to the destination in $\frac{1}{BK+1}$ portion of the time, we observe that the RS scheme achieves the maximum possible symmetric multiplexing gain for all the users. Moreover, from (3.19), we observe that the RS scheme achieves the maximum diversity gain of $K$ for any finite value of $B$, which turns out to be tight as well. Finally, the lower-bound on the DMT of the RS scheme is simplified to $\left[K\left(1 - Mr\right) - \frac{Mr}{B}\right]^+$ for the symmetric situation.*

### 3.4.4 Multiple-Access Single Relay Scenario

As we observe, the arguments of Theorems 3.7, 3.9 and 3.16 concerning DMT optimality of the RS scheme are valid for the scenario of having multiple relays ($K > 1$). Indeed, for the single relay scenario, the RS scheme is reduced to the simple amplify-and-forward relaying in which the relay listens to the source in the first half of the frame and transmits the amplified version of the received signal in the second half. However, like the case of non-interfering relays studied in [55], the DMT optimality arguments are no longer valid. On the other hand, we show that the DDF scheme achieves the optimum DMT for this scenario.

**Theorem 3.19** *Consider a multiple-access channel consisting of $M$ transmitting nodes aided by a single half-duplex relay. Assume that all the network nodes are* *equipped with a single antenna and there is no direct link between the sources and the destination. The amplify-and-forward scheme achieves the following DMT*

$$d_{AF,MAC}(r_1, r_2, \ldots, r_M) = \left(1 - 2\sum_{m=1}^{M} r_m\right)^+. \tag{3.31}$$

*However, the optimum DMT of the network is*

$$d_{MAC}(r_1, r_2, \ldots, r_M) = \left(\frac{1 - 2\sum_{m=1}^{M} r_m}{1 - \sum_{m=1}^{M} r_m}\right)^+, \tag{3.32}$$

*which is achievable by the DDF scheme of [26].*

**Proof** First, we show that the DMT of the AF scheme follows (3.31). At the destination side, we have

$$\mathbf{y} = g\alpha\left(\sum_{m=1}^{M} h_m \mathbf{x}_m + \mathbf{n}\right) + \mathbf{z}, \tag{3.33}$$

where $h_m$ is the channel gain between the $m$'th source and the relay, $g$ is the down-link channel gain, and $\alpha = \sqrt{\frac{P}{P\sum_{m=1}^{M}|h_m|^2+1}}$ is the amplification coefficient. Defining the outage event $\mathcal{E}_{\mathcal{S}}$ for a set $\mathcal{S} \subseteq \{1, 2, \ldots, M\}$, similar to the case of

Theorem 3.16, we have

$$
\begin{aligned}
\mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\} &= \mathbb{P}\left\{I\left(\mathbf{x}_{\mathcal{S}};\mathbf{y}|\mathbf{x}_{\mathcal{S}^c}\right) < 2\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)\right\} \\
&= \mathbb{P}\left\{\log\left(1+P\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)|g|^2|\alpha|^2\left(1+|g|^2|\alpha|^2\right)^{-1}\right) < \right. \\
&\qquad \left. 2\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)\right\} \\
&\doteq \mathbb{P}\left\{\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)|g|^2|\alpha|^2\min\left\{1,\frac{1}{|g|^2|\alpha|^2}\right\} \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}}|h_m|^2 \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\qquad \mathbb{P}\left\{\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)|g|^2|\alpha|^2 \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(b)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}}|h_m|^2 \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\qquad \mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}}|h_m|^2\right)\min\left\{P,\frac{1}{\sum_{m=1}^{M}|h_m|^2}\right\} \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\sum_{m\in\mathcal{S}}|h_m|^2 \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\qquad \mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}}|h_m|^2\right) \le P^{-2\left(1-\sum_{m\in\mathcal{S}} r_m\right)}\right\} + \\
&\qquad \mathbb{P}\left\{\frac{|g|^2\sum_{m\in\mathcal{S}}|h_m|^2}{\sum_{m=1}^{M}|h_m|^2} \le P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\}.
\end{aligned}
\tag{3.34}
$$

In the above equation, $\mathbb{P}\{\min(X,Y)\le z\} = \mathbb{P}\left\{(X\le z)\bigcup(Y\le z)\right\} \doteq \mathbb{P}\{X\le z\} + \mathbb{P}\{Y\le z\}$ results in $(a)$. $(b)$ follows from the fact that $|\alpha|^2$ can be asymptotically written as $\min\left\{P,\frac{1}{\sum_{m=1}^{M}|h_m|^2}\right\}$.

Since $\{|h_m|^2\}_{m=1}^M$ are i.i.d. random variables with exponential distribution, it follows that $\sum_{m\in\mathcal{S}} |h_m|^2$ has Chi-square distribution with $2|\mathcal{S}|$ degrees of freedom, which implies that

$$\mathbb{P}\left\{\sum_{m\in\mathcal{S}} |h_m|^2 \leq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}\right\} \doteq P^{-|\mathcal{S}|\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}. \tag{3.35}$$

To compute the second term in (3.34), defining $\epsilon_1 \triangleq P^{-2\left(1-\sum_{m\in\mathcal{S}} r_m\right)}$, we have

$$\begin{aligned}\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq \epsilon_1\right\} &\overset{(a)}{\dot{\geq}} \mathbb{P}\left\{|g|^2 \leq \epsilon_1\right\} \\ &\doteq \epsilon_1,\end{aligned} \tag{3.36}$$

where $(a)$ follows from the fact that as $\sum_{m\in\mathcal{S}} |h_m|^2$ has Chi-square distribution, we have $\sum_{m\in\mathcal{S}} |h_m|^2 \dot{\leq} 1$ with probability one (more precisely, with a probability greater than $1-P^{-\delta}$ for every $\delta > 0$). On the other hand, we have

$$\begin{aligned}\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq \epsilon_1\right\} &\leq \mathbb{P}\left\{|g|^2 |h_m|^2 \leq \epsilon_1\right\} \\ &\doteq \epsilon_1.\end{aligned} \tag{3.37}$$

Putting (3.36) and (3.37) together, we have

$$\mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq \epsilon_1\right\} \doteq \epsilon_1. \tag{3.38}$$

Now, to compute the third term in (3.34), defining $\epsilon_2 \triangleq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)}$, we observe

$$\epsilon_2 \doteq \mathbb{P}\left\{|g|^2 \leq \epsilon_2\right\} \leq \mathbb{P}\left\{|g|^2 \frac{\sum_{m\in\mathcal{S}} |h_m|^2}{\sum_{m=1}^M |h_m|^2} \leq \epsilon_2\right\} \overset{(a)}{\dot{\leq}} \mathbb{P}\left\{|g|^2\left(\sum_{m\in\mathcal{S}} |h_m|^2\right) \leq \epsilon_2\right\} \overset{(b)}{\doteq} \epsilon_2.$$

Here, $(a)$ follows from the fact that with probability one, we have $\sum_{m=1}^M |h_m|^2 \dot{\leq} 1$ and $(b)$ follows from (3.38). As a result

$$\mathbb{P}\left\{|g|^2 \frac{\sum_{m\in\mathcal{S}} |h_m|^2}{\sum_{m=1}^M |h_m|^2} \leq \epsilon_2\right\} \doteq \epsilon_2 \tag{3.39}$$

From (3.35), (3.38), and (3.39), we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\} &\doteq P^{-|\mathcal{S}|\left(1-2\sum_{m\in\mathcal{S}} r_m\right)^+} + P^{-2\left(1-\sum_{m\in\mathcal{S}} r_m\right)^+} + P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)^+} \\ &\doteq P^{-\left(1-2\sum_{m\in\mathcal{S}} r_m\right)^+}. \end{aligned} \tag{3.40}$$

Observing (3.40) and applying the argument of (3.25), we have

$$\mathbb{P}\left\{\mathcal{E}\right\} \doteq \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \mathbb{P}\left\{\mathcal{E}_{\mathcal{S}}\right\} \doteq P^{-\left(1-2\sum_{m=1}^{M} r_m\right)^+}. \tag{3.41}$$

This completes the proof for the AF scheme. Now, to compute the DMT of the DDF scheme, let us assume that the relay listens to the transmitted signal for the $l$ portion of the time until it can decode it perfectly. Hence, we have

$$l = \min\left\{1, \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \frac{\left(\sum_{m\in\mathcal{S}} r_m\right)\log(P)}{\log\left(1+\left(\sum_{m\in\mathcal{S}} |h_m|^2\right)P\right)}\right\}. \tag{3.42}$$

The outage event occurs whenever the relay can not transmit the re-encoded information in the remaining portion of the time. Hence, we have

$$\mathbb{P}\left\{\mathcal{E}\right\} \doteq \mathbb{P}\left\{(1-l)\log\left(1+|g|^2 P\right) < \left(\sum_{m=1}^{M} r_m\right)\log(P)\right\}. \tag{3.43}$$

Assuming $|h_m|^2 = P^{-\mu_m}$ and $|g|^2 = P^{-\nu}$, at high SNR, we have

$$l \approx \min\left\{1, \max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \frac{\sum_{m\in\mathcal{S}} r_m}{1-\min_{m\in\mathcal{S}} \mu_m}\right\}. \tag{3.44}$$

Equivalently, an outage event occurs whenever

$$\left(1-\max_{\mathcal{S}\subseteq\{1,2,\ldots,M\}} \frac{\sum_{m\in\mathcal{S}} r_m}{1-\min_{m\in\mathcal{S}} \mu_m}\right)(1-\nu) < \sum_{m=1}^{M} r_m. \tag{3.45}$$

In order to find the probability of the outage event, we first find an upper-bound on the outage probability and then, we show that this upper-bound is indeed tight. Defining $R = \sum_{m=1}^{M} r_M$ and $\mu = \sum_{m=1}^{M} \mu_m$, we have

$$R \overset{(a)}{>} \left(1-\frac{\sum_{m\in\mathcal{S}_0} r_m}{1-\min_{m\in\mathcal{S}_0} \mu_m}\right)(1-\nu) > \left(1-\frac{R}{1-\mu}\right)(1-\nu). \tag{3.46}$$

Here, $(a)$ follows from (3.45). Equivalently,

$$R \overset{(a)}{>} \frac{(1-\mu)(1-\nu)}{(1-\mu)+(1-\nu)} > \frac{1-\mu-\nu}{(1-\mu)+(1-\nu)}, \tag{3.47}$$

where $(a)$ follows from (3.46). It can be easily checked that (3.47) is equivalent to

$$R > (1-R)(1-\mu-\nu). \tag{3.48}$$

In other words, any vector point $[\mu_1, \mu_2, \ldots, \mu_M, \nu]$ in the outage region $\mathcal{R}$, i.e., the region that satisfies (3.45), also satisfies (3.48). As a result, defining $\mathcal{R}'$ as the region defined by (3.48), we have

$$\mathbb{P}\{\mathcal{E}\} \quad \leq \quad \mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\}, \tag{3.49}$$

where $\boldsymbol{\pi} \triangleq [\mu_1, \mu_2, \ldots, \mu_M, \nu]$. Similar to the approach used in the proofs of Theorems 3.9 and 3.16, $\mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\}$ can be computed as

$$\mathbb{P}\{\boldsymbol{\pi} \in \mathcal{R}'\} \doteq P^{-\frac{1-2R}{1-R}}. \tag{3.50}$$

Hence,

$$\mathbb{P}\{\mathcal{E}\} \dot{\leq} P^{-\frac{1-2R}{1-R}}. \tag{3.51}$$

For lower-bounding the outage probability, we note that all the vectors $[\mu_1, \cdots, \mu_M, \nu]$ for which $\mu_m > 0, m = 1, \cdots, M$ and $\nu > \frac{1-2R}{1-R}$, lie in the outage region defined in (3.45). In other words,

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} \quad &\geq \quad \mathbb{P}\left\{\boldsymbol{\pi} > \left[0, \cdots, 0, \frac{1-2R}{1-R}\right]\right\} \\ &\doteq \quad P^{-\frac{1-2R}{1-R}}. \end{aligned} \tag{3.52}$$

Combining (3.51) and (3.52) yields

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} \quad &\doteq \quad P^{-\frac{1-2R}{1-R}} \\ &= \quad P^{-\frac{1-2\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}}, \end{aligned} \tag{3.53}$$

which completes the proof for the DMT analysis of the DDF scheme.

Next, we prove that the DDF scheme achieves the optimum DMT. As the channel from the transmitters to the receiver is a degraded version of the channel between the transmitters and the relay, similar to the argument of [56] for the case of single-source single-relay, we can easily show that the decode-forward strategy achieves the capacity of the network for each realization of the channels. Now, consider the realization in which for all $m$ we have, $|h_m|^2 \leq \frac{1}{M}$. As we know, $\mathbb{P}\left\{\forall m : |h_m|^2 \leq \frac{1}{M}\right\} \doteq 1$. Let us assume in the optimum decode-and-forward strategy, the relay spends $l$ portion of the time for listening to the transmitter. According to the Fano's inequality [12], to make the probability of error in decoding the transmitters' message at the relay side approach zero, we should have $l \log\left(1 + \frac{P}{l}\sum_{m=1}^{M} |h_m|^2\right) \geq \left(\sum_{m=1}^{M} r_m\right)\log(P)$. Accordingly, we should have $l \geq \sum_{m=1}^{M} r_m$. On the other hand, in order that the receiver can decode the relay's message with a vanishing probability of error in the remaining portion of the time, we should have $(1-l)\log\left(1 + \frac{P}{1-l}|g|^2\right) \geq \sum_{m=1}^{M} r_m \log(P)$. Hence, we have $\mathbb{P}\{\mathcal{E}\} \geq \mathbb{P}\left\{|g|^2 \leq cP^{-\left(1-\frac{\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}\right)}, \forall m : |h_m|^2 \leq \frac{1}{M}\right\} \doteq P^{-\left(1-\frac{\sum_{m=1}^{M} r_m}{1-\sum_{m=1}^{M} r_m}\right)^{+}}$, for a constant $c$. This completes the proof.

∎

Figure 3.5 shows DMT of the AF scheme and the DDF scheme for multiple access single relay setup consisting of $M = 2$ sources assuming *symmetric* situation, i.e. $r_1 = r_2 = r$. As can be observed in this figure, although the AF scheme achieves the maximum multiplexing gain and maximum diversity gain, it does not achieve the optimum DMT in any other points of the tradeoff region.

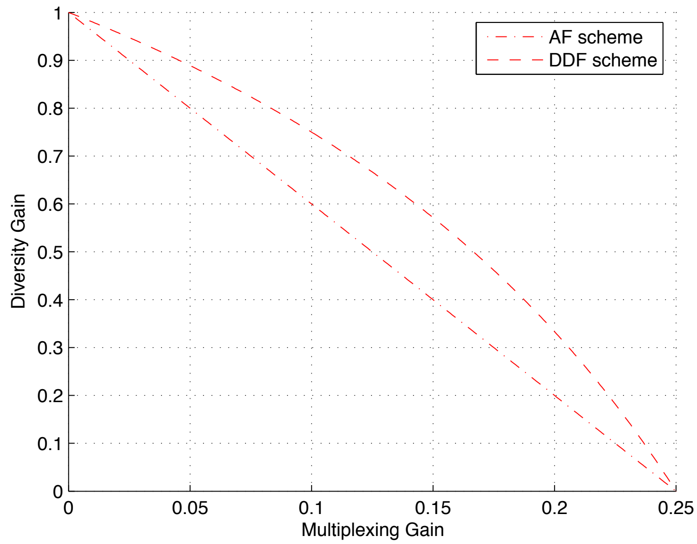


Figure 3.5: Diversity-Multiplexing Tradeoff of AF scheme versus the optimum and DDF scheme for multiple access single relay channel consisting of $M = 2$ sources assuming *symmetric* transmission, i.e. $r_1 = r_2 = r$.

## 3.5 Conclusion

In this chapter, a general model is described for the multi-antenna multiple relay networks. In this model, each pair of nodes are assumed to be either connected through a quasi-static Rayleigh fading channel or disconnected. A new scheme called *random sequential* (RS), based on the amplify-and-forward relaying, is introduced for this setup. DMT of the RS scheme is investigated for the single-antenna multiple relay networks. Bounds on DMT of the RS scheme are derived for a general single-antenna multiple-relay network. Specifically, 1) the exact DMT of the RS scheme is derived under the assumption of "non-interfering relaying"; 2) a lower-bound is derived on the DMT of the RS scheme (no conditions imposed).

Finally, it is shown that for the single-antenna two-hop multiple-access multiple-relay network setup where there is no direct link between the transmitter(s) and the receiver, the RS scheme achieves the optimum diversity-multiplexing tradeoff. However, for the multiple access single relay scenario, we show that the RS scheme is unable to perform optimum in terms of the DMT, while the dynamic decode-and-forward scheme is shown to achieves the optimum DMT for this scenario.

| time-slot | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $P_1(1)$ | $0 \to 1$ | $1 \to 3$ | $3 \to 5$ | — | — | — | — |
| $P_1(2)$ | — | $0 \to 2$ | $2 \to 4$ | $4 \to 5$ | — | — | — |
| $P_1(3)$ | — | — | — | $0 \to 1$ | $1 \to 4$ | $4 \to 5$ | — |
| $P_1(4)$ | — | — | — | — | $0 \to 2$ | $2 \to 3$ | $3 \to 5$ |

Table 3.1: One possible valid timing for RS scheme with the path sequence $P_1 = (p_1, p_2, p_3, p_4)$.

| time-slot | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P_2(1)$ | $0 \to 1$ | $1 \to 3$ | $3 \to 5$ | — | — | — |
| $P_2(2)$ | — | $0 \to 2$ | $2 \to 4$ | $4 \to 5$ | — | — |
| $P_2(3)$ | — | — | $0 \to 1$ | $1 \to 3$ | $3 \to 5$ | — |
| $P_2(4)$ | — | — | — | $0 \to 2$ | $2 \to 4$ | $4 \to 5$ |

Table 3.2: One possible valid timing for RS scheme with the path sequence $P_2 = (p_1, p_2, p_1, p_2)$.

# Chapter 4

# Maximum Diversity Gain of Relay Networks

## 4.1 Introduction

In this chapter, we study the setup of multi-antenna multiple relay network which is introduced in Chapter 3. We investigate the maximum achievable diversity gain of the relay network. For this purpose, we utilize the RS scheme which is proposed in Chapter 3. We prove that the RS scheme achieves the maximum diversity gain of any multi-antenna wireless relay network. It turns out that the maximum achievable diversity gain of the network is equal to the minimum weight between all edge cut-sets of the underlying graph of the network. Hence, for any relay network, the maximum achievable diversity gain of the network can be characterized in polynomial time (with respect to the number of network nodes) using the maximum flow algorithm [14]. However, we show that there exists multi-antenna scenarios for which the RS scheme does not achieve the optimum DMT.

Indeed, we show that in order to achieve the optimum DMT, in some scenarios, multiple interfering nodes have to transmit together (and interfere on the receiver node of each other) during the same time-slot.

The rest of this chapter is as follows. In section 4.2, we investigate the maximum achievable diversity gain of the multi-antenna relay networks and in section 4.3,we conclude the chapter.

## 4.2 Maximum Diversity Achievability Proof in General Multi-Hop Multiple-Antenna Scenario

**Theorem 4.1** *Consider a relay network with the connectivity graph $G = (V, E)$ and $K$ relays, in which each two adjacent nodes are connected through a Rayleigh-fading channel. Assume that all the network nodes are equipped with multiple antennas. Then, by properly choosing the path sequence, the proposed RS scheme achieves the maximum diversity gain of the network which is equal to*

$$d_G = \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{4.1}$$

*where $\mathcal{S}$ is an edge cut-set on $G$. The edge cut-set and weight of a cut-set are defined in Defintion 3.1.*

**Proof** First, we show that $d_G$ is indeed an upper-bound on the diversity-gain of the network. To show this, we do not consider the half-duplex nature of the relay nodes and assume that they operate in full-duplex mode. Consider an edge cut-set

$\mathcal{S}$ on $G$. We have

$$\begin{aligned}
\mathbb{P}\{\mathcal{E}\} &\overset{(a)}{\dot{\geq}} \mathbb{P}\{I(X(\mathcal{S});Y(\mathcal{S}^c)|X(\mathcal{S}^c)) < R\} \\
&\overset{(b)}{=} \mathbb{P}\left\{\sum_{k\in\mathcal{S}^c} I(X(\mathcal{S});Y_k|Y(\mathcal{S}^c/\{1,2,\ldots,k\}),X(\mathcal{S}^c)) < R\right\} \\
&\overset{(c)}{\geq} \prod_{k\in\mathcal{S}^c} \mathbb{P}\left\{I(X(\mathcal{S});Y_k|X(\mathcal{S}^c)) < \frac{R}{|\mathcal{S}^c|}\right\} \\
&\overset{(d)}{\doteq} \prod_{k\in\mathcal{S}^c} P^{-|\{e\in E|k\in e, e\cap\mathcal{S}\neq\oslash\}|} \\
&\doteq P^{-w_G(\mathcal{S})}, \qquad (4.2)
\end{aligned}$$

where $R$ is the target rate which does not scale with $P$ (i.e., $r = 0$). Here, $(a)$ follows from the cut-set bound Theorem [12] and the fact that for the rates above the capacity, the error probability approaches one (according to Fano's inequality [12]), $(b)$ follows from the chain rule on the mutual information [12], $(c)$ follows from the facts that i) $(Y_k, X(\{0,1,\ldots,K+1\}), Y(\mathcal{S}^c/\{1,2,\ldots,k\}))$ form a Markov chain [12] and as a result, $I(X(\mathcal{S});Y_k|Y(\mathcal{S}^c/\{1,2,\ldots,k\}),X(\mathcal{S}^c)) \leq I(X(\mathcal{S});Y_k|X(\mathcal{S}^c))$, and ii) $I(X(\mathcal{S});Y_k|X(\mathcal{S}^c))$ depends only on the channel matrices between $X(\mathcal{S})$ and $Y_k$ and as all the channels in the network are independent of each other, it follows that the events

$$\left\{I(X(\mathcal{S});Y_k|X(\mathcal{S}^c)) < \frac{R}{|\mathcal{S}^c|}\right\}_{k\in\mathcal{S}^c}$$

are mutually independent, and finally $(d)$ follows from the diversity gain of the MISO channel. Considering all possible cut-sets on $G$ and using (4.2), we have

$$\mathbb{P}\{\mathcal{E}\} \dot{\geq} P^{-\min_{\mathcal{S}} w_G(\mathcal{S})}. \qquad (4.3)$$

Now, we prove that this bound is indeed achievable by the RS scheme. First, we provide the path sequence needed to achieve the maximum diversity gain. Consider the graph $\hat{G} = (V, E, w)$ with the same set of vertices and edges as the graph $G$ and the weight function $w$ on the edges as $w_{\{a,b\}} = N_a N_b$. Consider the maximum-flow algorithm [14] on $\hat{G}$ from the source node 0 to the sink node $K+1$. Since the weight function is integer over the edges, according to the Ford-Fulkerson Theorem [14], one can achieve the maximum flow which is equal to the minimum cut of $\hat{G}$ or $d_G$ by the union of elements of a sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G})$ of paths $(L = d_G)$. We show that this family of paths are sufficient to achieve the optimum diversity. Here, we do not consider the problem of selecting the path timing sequence $\{s_{i,j}\}$. We just assume that a timing sequence $\{s_{i,j}\}$ with the 4 requirements defined in section 3.3 exists. However, it should be noted that as we consider the maximum diversity throughout the Theorem, we are not concerned with $\frac{S}{L}$. Hence, we can select the path timing sequence such that no two paths cause interference on each other.

Noting that the received signal at each node is multiplied by a random isotropically distributed unitary matrix, at the destination side we have

$$\mathbf{y}_{K+1,i} = \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0}\mathbf{x}_{0,i} + \sum_{j<i}\mathbf{X}_{i,j}\mathbf{x}_{0,j} + \sum_{j\le i, m\le l_j}\mathbf{Q}_{i,j,m}\mathbf{n}_{j,m}. \quad (4.4)$$

Here, $\mathbf{x}_{0,i}$ is the vector transmitted at the source side during the $s_{i,1}$'th slot as the input for the $i$'th path, $\mathbf{y}_{K+1,i}$ is the vector received at the destination side during the $s_{i,l_i}$'th slot as the output for $i$'th path, $\mathbf{U}_{i,j}$ denotes the multiplied unitary matrix at the $\mathrm{p}_i(j)$'th node of the $i$th path, $\mathbf{X}_{i,j}$ is the interference matrix which relates the input of the $j$'th path $(j < i)$ to the output of the $i$'th path, $\mathbf{n}_{j,m}$ is the noise vector during the $s_{j,m}$'th slot at the $\mathrm{p}_j(m)$'th node of the network, and finally, $\mathbf{Q}_{i,k,m}$ is the matrix which relates $\mathbf{n}_{k,m}$ to $\mathbf{y}_{K+1,i}$. Notice that as the timing sequence satisfies the noncausal interference assumption, the summation terms in (4.4) do not exceed $i$. Defining $\mathbf{x}(s) = \left[\mathbf{x}_{0,1}^T(s)\,\mathbf{x}_{0,2}^T(s)\cdots\mathbf{x}_{0,L}^T(s)\right]^T$, $\mathbf{y}(s) = \left[\mathbf{y}_{K+1,1}^T(s)\,\mathbf{y}_{K+1,2}^T(s)\cdots\mathbf{y}_{K+1,L}^T(s)\right]^T$, and $\mathbf{n}(s) = \left[\mathbf{n}_{1,1}^T(s)\,\mathbf{n}_{1,2}^T(s)\cdots\mathbf{n}_{L,l_L}^T(s)\right]^T$, we have the following equivalent block lower-triangular matrix between the end nodes

$$\mathbf{y}(s) = \mathbf{H}_T\mathbf{x}(s) + \mathbf{Q}\mathbf{n}(s). \tag{4.5}$$

Here,

$$\mathbf{H}_T = \begin{pmatrix} \mathbf{X}_{1,1} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{X}_{2,1} & \mathbf{X}_{2,2} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{X}_{L,1} & \mathbf{X}_{L,2} & \cdots & \mathbf{X}_{L,L} \end{pmatrix}, \tag{4.6}$$

where $\mathbf{X}_{i,i} = \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0}$, and

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{1,1,1} & \cdots & \mathbf{Q}_{1,1,l_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{Q}_{2,1,1} & \cdots & \mathbf{Q}_{2,1,l_1} & \cdots & \mathbf{Q}_{2,2,l_2} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \mathbf{Q}_{L,1,1} & \mathbf{Q}_{L,1,2} & \cdots & \cdots & \cdots & \mathbf{Q}_{L,L,l_L-1} & \mathbf{Q}_{L,L,l_L} \end{pmatrix}. \tag{4.7}$$

Having (5.46), the outage probability can be written as

$$\mathbb{P}\left\{\mathcal{E}\right\} \;=\; \mathbb{P}\left\{\left|\mathbf{I}_L + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| < 2^{SR}\right\}, \tag{4.8}$$

where $\mathbf{P}_n = \mathbf{Q}\mathbf{Q}^H$. First, similar to the proof of Theorem 3.9, we can show that $\alpha_{i,j} \doteq 1$ with probability 1[1], and also show that there exists a constant $c$ which depends just on the topology of graph $G$ and the path sequence such that $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$.

[1]More precisely, with probability greater than $1 - P^{-\delta}$ for any $\delta > 0$.

Assume that for each $\{a,b\} \in E$, $\lambda_{\max}\left(\mathbf{H}_{a,b}\right) = P^{-\mu_{\{a,b\}}}$, where $\lambda_{\max}\left(\mathbf{A}\right)$ denotes the greatest eigenvalue of $\mathbf{A}\mathbf{A}^H$. Also, assume that

$$\begin{aligned}\gamma_{i,j} \triangleq & \left|\mathbf{v}_{r,\max}^H\left(\mathbf{H}_{\{\mathrm{p}_i(j+1),\mathrm{p}_i(j)\}}\right)\mathbf{U}_{i,j}\mathbf{v}_{l,\max}\left(\mathbf{H}_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\mathbf{U}_{i,j-1}\mathbf{H}_{\{\mathrm{p}_i(j-1),\mathrm{p}_i(j-2)\}}\cdots\right.\right. \\ & \left.\left.\mathbf{H}_{\{\mathrm{p}_i(1),0\}}\right)\right|^2 = P^{-\nu_{i,j}},\end{aligned} \tag{4.9}$$

where $\mathbf{v}_{l,\max}\left(\mathbf{A}\right)$ and $\mathbf{v}_{r,\max}\left(\mathbf{A}\right)$ denote the left and the right eigenvectors of $\mathbf{A}$ corresponding to $\lambda_{\max}\left(\mathbf{A}\right)$, respectively. The outage probability can be upper-bounded as

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}\right\} &\overset{(a)}{\leq} \mathbb{P}\left\{\lambda_{\max}\left(\left(\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right)^{\frac{1}{2}}\right) \leq \left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(b)}{\leq} \mathbb{P}\left\{\lambda_{\max}\left(\mathbf{H}_T\right) \leq c\left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(c)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\lambda_{\max}\left(\mathbf{X}_{i,i}\right) \leq c\left(2^{SR}-1\right)P^{-1}\right)\right\} \\ &\overset{(d)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1 - \log\frac{c\left(2^{SR}-1\right)}{P}\right)\right\} \\ &\overset{(e)}{\doteq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1\right)\right\}.\end{aligned} \tag{4.10}$$

In the above equation, $(a)$ follows from the fact that $1+\lambda_{\max}(\mathbf{A}^{\frac{1}{2}}) \leq \left|\mathbf{I}+\mathbf{A}\right|$, for a positive semi-definite matrix $\mathbf{A}$. $(b)$ results from $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$. $(c)$ follows from the fact that $\lambda_{\max}(\mathbf{H}_T) \geq \max_i \lambda_{\max}(\mathbf{X}_{i,i})$. To obtain $(d)$, we first show that

$$\lambda_{\max}(\mathbf{AUB}) \geq \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{A})\left|\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\right|^2, \tag{4.11}$$

for any matrices $\mathbf{A}$, $\mathbf{U}$ and $\mathbf{B}$. To show this, we write

$$\begin{aligned}
\lambda_{\max}(\mathbf{AUB}) &= \max_{\substack{\mathbf{x} \\ \|\mathbf{x}\|^2=1}} \left\|\mathbf{x}^H\mathbf{AUB}\right\|^2 \\
&\geq \|\mathbf{v}_{l,\max}(\mathbf{A})\mathbf{AUB}\|^2 \\
&= \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\sum_i \mathbf{v}_{l,i}(\mathbf{B})\sigma_i(\mathbf{B})\mathbf{v}_{r,i}^H(\mathbf{B})\right\|^2 \\
&\overset{(a)}{=} \sum_i \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,i}(\mathbf{B})\sigma_i(\mathbf{B})\mathbf{v}_{r,i}^H(\mathbf{B})\right\|^2 \\
&\geq \left\|\sigma_{\max}(\mathbf{A})\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\sigma_{\max}(\mathbf{B})\mathbf{v}_{r,\max}^H(\mathbf{B})\right\|^2 \\
&\overset{(b)}{=} \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{B})\left|\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\right|^2,
\end{aligned} \tag{4.12}$$

where $\sigma_i(\mathbf{A})$ denotes the $i$'th singular value of $\mathbf{A}$, and $\sigma_{\max}(\mathbf{A})$ denotes the singular value of $\mathbf{A}$ with the highest norm. Here, $(a)$ follows from the fact that as $\{\mathbf{v}_{r,i}(\mathbf{B})\}$ are orthogonal vectors, the square-norm of their summation is equal to the summation of their square-norms. $(b)$ results from the fact that $\lambda_i(\mathbf{A}) = |\sigma_i(\mathbf{A})|^2, \forall i$. By recursively applying (4.11), it follows that

$$\lambda_{\max}(\mathbf{X}_{i,i}) \;\geq\; \prod_{j=1}^{l_i} \lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(j),\mathrm{p}_i(j-1)}\right) \prod_{j=1}^{l_i-1} \gamma_{i,j}. \tag{4.13}$$

Noting the definitions of $\mu_{\{i,j\}}$ and $\nu_{i,j}$, $(d)$ easily follows. Finally, $(e)$ results from the fact that as $P \to \infty$, the term $\log \frac{c\left(2^{SR}-1\right)}{P}$ can be ignored.

Since the left and the right unitary matrices resulting from the SVD of an i.i.d. complex Gaussian matrix are independent of its singular value matrix [57] and $\mathbf{U}_{i,j}$ is an independent isotropically distributed unitary matrix, we conclude that all the random variables in the set $\left\{\{\mu_e\}_{e\in E}, \{\nu_{i,j}\}_{1\leq i\leq L, 1\leq j<l_i}\right\}$ are mutually independent. From the probability distribution analysis of the singular values of circularly symmetric Gaussian matrices in [29], we can easily prove

$\mathbb{P}\left\{\mu_e \geq \mu_e^0\right\} \doteq P^{-N_a N_b \mu_e^0} = P^{-w_e \mu_e^0}$. Similarly, as $\mathbf{U}_{i,j}$ is isotropically distributed, it can be shown that $\mathbb{P}\left\{\nu(i,j) \geq \nu_0(i,j)\right\} \doteq P^{-\nu_0(i,j)}$. Hence, defining $\boldsymbol{\mu} = [\mu_e]_{e\in E}^T$, $\boldsymbol{\nu} = [\nu_{i,j}]_{1\leq i\leq L, 1\leq j<l_i}^T$, and $\mathbf{w} = [w_e]_{e\in E}$, we have

$$\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}_0, \boldsymbol{\nu} \geq \boldsymbol{\nu}_0\right\} \doteq P^{-(\mathbf{1}\cdot\boldsymbol{\nu}+\mathbf{w}\cdot\boldsymbol{\mu})}. \tag{4.14}$$

Let us define $\mathcal{R}$ as the region in $\mathbb{R}^{|E|+\sum_{i=1}^{L} l_i - L}$ of the vectors $\left[\boldsymbol{\mu}^T \boldsymbol{\nu}^T\right]^T$ such that for all $1 \leq i \leq L$, we have $\sum_{j=1}^{l_i} \mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1} \nu_{i,j} \geq 1$. Using the same argument as in the proof of Theorem 3.9, we conclude that $\mathbb{P}\left\{\mathcal{R}\right\} = \mathbb{P}\left\{\mathcal{R} \bigcap \mathbb{R}_+^{|E|+\sum_{i=1}^{L} l_i - L}\right\}$. Hence, defining the region $\mathcal{R}_+$ as $\mathcal{R}_+ = \mathcal{R} \bigcap \mathbb{R}_+^{|E|+\sum_{i=1}^{L} l_i - L}$ and $d_0 = \min_{[\boldsymbol{\mu}^T \boldsymbol{\nu}^T]^T \in \mathcal{R}_+} \mathbf{w}\cdot\boldsymbol{\mu} + \mathbf{1}\cdot\boldsymbol{\nu}$, which can easily be verified to be bounded, and applying the same argument as in the proof of Theorem 3.9, we have

$$\mathbb{P}\left\{\mathcal{E}\right\} \dot{\leq} \mathbb{P}\left\{\mathcal{R}_+\right\} \doteq P^{-d_0}. \tag{4.15}$$

To complete the proof, we have to show that $d_0 = d_G$, or equivalently, $d_0 = L$ (note that $L = d_G$). The value of $d_0$ is obtained from the following linear programming optimization problem

$$\begin{aligned} &\min\ \mathbf{w}\cdot\boldsymbol{\mu} + \mathbf{1}\cdot\boldsymbol{\nu} \\ &s.t. \quad \boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\nu} \geq \mathbf{0}, \forall i \sum_{j=1}^{l_i} \mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1))\}} + \sum_{j=1}^{l_i-1} \nu_{i,j} \geq 1. \end{aligned} \tag{4.16}$$

According to the argument of linear programming [45], the solution of the above linear programming problem is equal to the solution of the dual problem which is

$$\begin{aligned} &\max\ \sum_{i=1}^{L} f_i \\ &s.t. \quad \mathbf{0} \leq \mathbf{f} \leq \mathbf{1}, \forall e \in E, \sum_{e\in \mathrm{p}_i} f_i \leq w_e. \end{aligned} \tag{4.17}$$

Let us consider the solution $\mathbf{f}_0 = \mathbf{1}$ for (4.17). As the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ consists of the paths that form the maximum flow in $\hat{G}$, we conclude that for every $e \in E$, we have $\sum_{e \in \mathrm{p}_i} 1 \leq w_e$. Hence, $\mathbf{f}_0$ is a feasible solution for (4.17). On the other hand, as for all feasible solutions $\mathbf{f}$ we have $\mathbf{f} \leq \mathbf{1}$, we conclude that $\mathbf{f}_0$ maximizes (4.17). Hence, we have

$$d_0 = \min \ \mathbf{w} \cdot \boldsymbol{\mu} + \mathbf{1} \cdot \boldsymbol{\nu} \overset{(a)}{=} \max \ \sum_{i=1}^{L} f_i = L = d_G. \tag{4.18}$$

Here, $(a)$ results from duality of the primal and dual linear programming problems. This completes the proof. ∎

**Remark 4.2** *It is worth noting that according to the proof of Theorem 4.1, any RS scheme achieves the maximum diversity of the wireless multi-antenna relays network as long as its corresponding path sequence includes the paths* $\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G}$ *used in the proof of Theorem 4.1.*

**Example 4.3** *Consider the half-duplex* 3 *hops network of Figure 3.1. Here, as the path sequence* $\mathrm{P}_1 \equiv (\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_4)$ *forms the maximum flow of* $\hat{G}$*, the RS scheme with the path sequence* $\mathrm{P}_1$ *and the timing sequence of Table 3.1 achieves the maximum diversity of the network which is equal to* 4*. However, the diversity of the RS scheme with the path sequence* $\mathrm{P}_2$ *and the timing sequence of Table 3.2 is equal to* 2 *as it utilizes just two SISO edges* $\{1, 3\}$ *and* $\{2, 4\}$ *of the second hop.*

Theorem 4.1 shows that the RS scheme is capable of exploiting the maximum achievable diversity gain in multiple-antenna multiple-relay wireless networks. However, as the following example shows, the RS scheme is unable to achieve the maximum multiplexing gain in a general multiple-antenna multiple-node wireless network.

**Example 4.4** *Consider a two-hop relay network consisting of $K = 4$ relay nodes. The source and the destination are equipped with $N_0 = N_5 = 2$ antennas, while each of the relays has a single receiving/transmitting antenna. There exists no direct link between the source and the destination, i.e. $\{0, 5\} \notin E$. For the sake of simplicity, assume that the relays are non-interfering, i.e. $1 \leq a \leq 4, 1 \leq b \leq 4, \{a, b\} \notin E$. Let us partition the set of relays into $\mathcal{S}_0 = \{1, 2\}, \mathcal{S}_1 = \{3, 4\}$. Consider the following amplify-and-forward strategy: In the i'th time slot, the relay nodes in $\mathcal{S}_{i \bmod 2}$ transmit what they have received in the last time slot, while the relay nodes in $\mathcal{S}_{(i+1) \bmod 2}$ receive the source's signal. It can be easily verified that this scheme achieves a maximum multiplexing gain of $r = 2$. However, the proposed RS scheme achieves a maximum multiplexing gain of $r = 1$.*

## 4.3 Conclusion

In this chapter, the general setup of multi-antenna multiple relay network which is introduced in Chapter 3, is investigated and the maximum achievable diversity gain is characterized. Furthermore, the RS scheme is shown to achieve the maximum achievable diversity gain of the relay network. The maximum achievable diversity gain is turned out to be equal to the minimum weight between all edge cut-sets of the underlying graph of the network. However, certain multi-antenna scenarios are shown in which the RS scheme does not achieve the optimum DMT. Indeed, it is proved that in order to achieve the optimum DMT, in some scenarios, multiple interfering nodes have to transmit together (and interfere on the receiver node of each other) during the same time-slot.

# Chapter 5

# Diversity-Multiplexing Tradeoff in Multi-Antenna Multiple Relay Networks

## 5.1 Introduction

In this chapter, we investigate the benefits of AF relaying in multi-antenna multi-relay networks. For description of the general multi-antenna relay network, you can refer to Chapter 3. For this purpose, we study the application of the RS scheme proposed in Chapter 3. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the transmitted signal of the future paths on the received signal of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in AF relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. We derive the DMT of the RS scheme for multi-antenna multi-relay networks. To accomplish this task, we first study a simple structure, namely the multi-antenna full-duplex two-hop single-relay network. We show that unlike the traditional AF relaying, the RS scheme achieves the optimum DMT. This fact can be justified as follows: using the traditional AF relaying, there exists a chance that the eigenvectors corresponding to the largest eignenvalues of the incoming channel matrix of the relay project to the eigenvectors corresponding to the small eignenvalues of the relay's outgoing channel matrix. This event degrades the performance of traditional AF relaying in the multi-antenna setup. However, in the RS scheme, due to the random independent unitary matrix multiplication at the relay nodes for different time-slots, such an event is much less likely to happen. This fact will be elaborated throughout the Chapter.

Next, we study the case of multi-antenna half-duplex parallel relay network and, by deriving its DMT, we show that the RS scheme improves the DMT of the traditional AF relaying scheme. Interestingly, it turns out that the DMT of the RS scheme is optimum for the multi-antenna half-duplex parallel two-relay ($K = 2$) setup with no direct link between the relays. We also show that utilizing random unitary matrix multiplication improves the DMT of the NAF relaying scheme of [26] in the case of a multi-antenna single relay channel.

Finally, we study the class of general full-duplex multi-antenna relay networks whose underlying graph is *directed acyclic* and all nodes are equipped with the same number of antennas. Using the RS scheme, we derive a new lower-bound for the achievable DMT of this class of networks. It turns out that the new DMT lower-bound meets the optimum DMT at the corner points, corresponding to the maximum multiplexing gain and the maximum diversity gain of the network, respectively. Another point worth mentioning is that the RS scheme is *robust* in the sense that it achieves all points of the DMT curve with no modification of the underlying parameters. In other words, the relay nodes of the network perform the same operation, no matter at which point of the DMT curve the scheme is operating.

For description of the system model and the RS scheme, you can refer to Chapter 3. The rest of the chapter is organized as follows. Section 5.2 is dedicated to Diversity-Multiplexing Tradeoff analysis of the RS scheme in the multi-antenna setup. This section is further divided into four subsections as follows. Subsection 5.2.1 studies the multi-antenna single-relay two-hop network and also the multi-antenna multi-hop relay network with one relay in each hop. Subsection 5.2.2 is dedicated to the multi-antenna half-duplex parallel relay network. The well-known multi-antenna single-relay channel (with direct link between the source and the destination) is investigated in subsection 5.2.3. Subsection 5.2.4 studies the achievable DMT of the RS scheme for the general multi-antenna full-duplex relay networks whose underlying graph is directed acyclic. Finally, section 5.3 concludes the chapter.

## 5.2 Diversity-Multiplexing Tradeoff

### 5.2.1 Two-Hop Single Relay Network

This setup corresponds to the network consisting of a source, a single full-duplex relay and a destination with no direct link between the source and the destination. The source, relay, and destination are equipped with $m$, $p$, and $n$ antennas, respectively (see Figure 5.1). The channel between the source and the relay is denoted by

**H** and the channel between the relay and the receiver is denoted by **G**. First, we study the achievable DMT by the traditional AF scheme in Lemmas 5.1 and 5.2 and show the achievable DMT in general does not match with the optimal value, which is achievable by the DF scheme[1]. Then, in Theorem 5.3 we prove that using the proposed RS scheme, which is a modification of the traditional AF scheme, the optimal DMT is indeed achievable. Theorem 5.7 generalizes the result of Theorem 5.3 to the case of multi-hop relay network and shows that the proposed RS scheme still achieves the optimum DMT provided that a certain relationship between the number of antennas at nodes is satisfied.

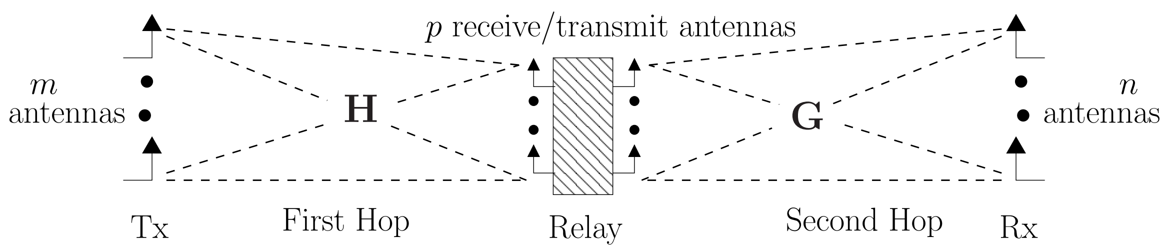


Figure 5.1: Schematic of a multi-antenna single-relay two-hop network

In the traditional AF strategy, the received signal at the relay is multiplied by a constant $\alpha$ such that the power constraint at the relay is satisfied and then it is transmitted to the destination. The corresponding received signal at the destination can be written as

$$\mathbf{y} = \alpha \mathbf{G}\mathbf{H}\mathbf{x}_t + \alpha \mathbf{G}\mathbf{n}_r + \mathbf{n}_d, \tag{5.1}$$

where $\mathbf{x}_t$ denotes the transmitted signal from the source and $\mathbf{n}_r \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_p)$ and $\mathbf{n}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$ denote the noise vectors at the relay and at the destination,

[1] In fact, this configuration is a special case of the degraded relay channel studied by [56]. In [56], the authors show that the DF scheme achieves the capacity of the degraded relay channel.

respectively.

**Lemma 5.1** *The DMT of the system given in (5.1) is upper-bounded by the DMT of the following system:*

$$\mathbf{y} = \alpha \mathbf{G}\mathbf{H}\mathbf{x}_t + \mathbf{n}_d, \tag{5.2}$$

*and is lower-bounded by the DMT of the following system:*

$$\mathbf{y} = \alpha \mathbf{G}\mathbf{H}\mathbf{x}_t + \sqrt{c\log(P)+1}\,\mathbf{n}_d, \tag{5.3}$$

*for some constant $c$.*

**Proof** See Appendix H. ∎

**Lemma 5.2** *The DMT of the systems in (5.2) and (5.3) are equal.*

**Proof** See Appendix I. ∎

A direct conclusion of the Lemmas 5.1 and 5.2 is that the DMT of the two-hop network can be expressed as the DMT of the product channel **GH** which is computed in [54]. Due to the result given in Proposition 1 in [54], assuming $m, n \geq p$, the DMT of the product channel $\mathbf{A} = \mathbf{G}\mathbf{H}$ is a piecewise-linear function connecting the points $(r, d_{\mathbf{A}}(r)), r = 0, 1, \ldots, p$, where

$$d_{\mathbf{A}}(r) = (p-r)(q-r) - \frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor, \tag{5.4}$$

$q = \min(m, n)$ and $\Delta = |m-n|$. On the other hand, the piecewise-linear function connecting the integer points $(r, (p-r)(q-r))$ can be easily derived as the upper-bound by considering each of the source-relay or the relay-destination cuts. Comparing (5.4) with the upper-bound, it follows that the traditional AF scheme achieves the optimum DMT only when $r \geq p - \Delta$. This motivates us to use a variant of AF scheme which achieves the optimum DMT in all cases. In fact, using the traditional AF scheme, there are three sources of outage: (i) the outage in the source-relay link, (ii) the outage in the relay-destination link, and (iii) the projection of the eigenmodes of $\mathbf{H}$ over the eigenmodes of $\mathbf{G}$ is very small. More precisely, the matrix $\mathbf{V}^H(\mathbf{G})\mathbf{U}(\mathbf{H})$, in which $\mathbf{V}^H(\mathbf{G})$ denotes the right eigenvector matrix from the Singular Value Decomposition (SVD) of $\mathbf{G}$ and $\mathbf{U}(\mathbf{H})$ denotes the left eigenvector matrix from the SVD of $\mathbf{H}$, has very small eigenvalues. The extra term $\frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor$ in (5.4) is due to the third source of outage. The first two outage events depend on the distribution of the eigenvalues of $\mathbf{H}$ and $\mathbf{G}$, while the third event depends solely on the *direction* of the eigenvectors of these two matrices. This suggests that in order to eliminate the extra terms in $\frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor$ one can multiply the received signal at the relay by $\alpha\mathbf{\Theta}$, for some $p \times p$ unitary matrix $\mathbf{\Theta}$ (for preserving the power constraint at the relay). However, it should be noted that if $\mathbf{\Theta}$ does not change across the transmission block, the performance of the systems does not change. Therefore, we propose that in each transmission slot an independent unitary matrix $\mathbf{\Theta}_l$ is used and at the destination side the decoding is performed across $L$ transmission slots[2]. This is exactly what is being done in the proposed RS scheme. Indeed, this setup is a simple example of the general setup of the relay network studied in Chapter 3 in which the source and the destination are connected through a single path. In this case the proposed RS scheme reduces to the following: the source's message is sent using $L$ slots through the same path; at the relay side the received signal is multiplied by a randomly inde-

[2]From practical point of view, the transmission slots are assumed to be long enough to make the probability of error solely dominated by the outage event. This fact is more elaborated in [29].

pendent (through different slots) unitary matrix, and subsequently, it is multiplied by a scalar $\alpha \leq 1$[3] such that the power constraint is satisfied, and the result is transmitted in the next slot. At the destination, following receiving the signal of the slots $2, 3, \ldots, L+1$, the source message is decoded. In the following theorem, we show that as long as $L$ is above a certain threshold, the probability of the third outage event is negligible compared to the first two outage events and hence, the optimum DMT is achievable by the RS scheme.

**Theorem 5.3** *Consider a two-hop network consisting of a source with $m$ antennas and a destination with $n$ antennas that are connected through a full-duplex relay with $p$ antennas. Let us define $q = \min(m, n)$. Providing $L$ is large enough such that $L \geq \min^2(p, q) \max(p, q)$, RS scheme achieves the optimum DMT, which is a piecewise-linear function connecting the points $(k, (p-k)(q-k))$, where $k = 0, 1, \ldots, \min(p, q)$.*

**Proof** Using Lemmas 5.1 and 5.2, the DMT of the system using the proposed RS scheme is equal to the DMT of the following system:

$$\mathbf{Y} = \alpha \mathbf{\Omega} \mathbf{X}_t + \mathbf{N}_d, \tag{5.5}$$

where $\mathbf{X}_t \triangleq [\mathbf{x}_t(1), \cdots, \mathbf{x}_t(L)]^T$, $\mathbf{Y} = [\mathbf{y}(1), \cdots, \mathbf{y}(L)]^T$, and $\mathbf{N}_d \triangleq [\mathbf{n}_d(1), \cdots, \mathbf{n}_d(L)]^T$, in which $\mathbf{x}_t(l)$, $\mathbf{y}(l)$ and $\mathbf{n}_d(l)$ denote the transmitted signal, received signal and

[3]Note that, as can be observed from the proof of Lemma 5.1 in Appendix I, the constraint $\alpha \leq 1$ does not affect the DMT of the system.

noise in the $l$th slot, respectively, and

$$\boldsymbol{\Omega} \triangleq \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_L \end{bmatrix}, \tag{5.6}$$

in which $\mathbf{A}_l \triangleq \mathbf{G}\boldsymbol{\Theta}_l\mathbf{H}$. Hence, the matrix of the end-to-end channel is a block diagonal matrix consisting of $\mathbf{A}_l$'s. Assuming that the transmitted signals in each slot are independent of each other, the mutual information between the input and the output of (5.5) can be written as

$$\mathcal{I}(\mathbf{X}_t; \mathbf{Y}) = \sum_{l=1}^{L} \log \left| \mathbf{I} + \alpha^2 \frac{P}{m} \mathbf{A}_l \mathbf{A}_l^H \right|, \tag{5.7}$$

in which it is assumed that $\mathbf{x}_t(l) \sim \mathcal{CN}(\mathbf{0}, \frac{P}{m}\mathbf{I}_m)$, $\forall l = 1, \cdots, k$. Using the above equation, the probability of outage can be written as

$$\mathbb{P}\{\mathcal{O}\} = \mathbb{P}\left\{ \sum_{l=1}^{L} \log \left| \mathbf{I} + \alpha^2 \frac{P}{m} \mathbf{A}_l \mathbf{A}_l^H \right| < Lr \log(P) \right\}, \tag{5.8}$$

or equivalently,

$$\mathbb{P}\{\mathcal{O}\} = \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} \log \left( 1 + \alpha^2 \frac{P}{m} \lambda_j(\mathbf{A}_l) \right) < Lr \log(P) \right\}, \tag{5.9}$$

where $\lambda_i(\mathbf{A})$ denotes the $i$th ordered eigenvalue of $\mathbf{A}^H\mathbf{A}$ ($\lambda_1 > \lambda_2 > \cdots > \lambda_{\min}$). Defining $\gamma_j(\mathbf{B}) \triangleq -\frac{\log(\lambda_j(\mathbf{B}))}{\log(P)}$ and $\delta \triangleq -\frac{\log(\alpha^2)}{\log(P)}$, we have

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &= \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} \log \left( 1 + \frac{1}{m} P^{1-\delta-\gamma_j(\mathbf{A}_l)} \right) < Lr \log(P) \right\} \\ &\doteq \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} (1 - \delta - \gamma_j(\mathbf{A}_l))^+ < Lr \right\}. \end{aligned} \tag{5.10}$$

First, we show that $\alpha \doteq 1$ (or $\delta \doteq 0$), with probability one[4]. For this purpose, we write $\alpha^2$ as follows:

$$\begin{aligned}
\alpha^2 &= \min\left(1, \frac{P}{\mathbb{E}_{\mathbf{x}_t,\mathbf{n}_r}\left\{\|\mathbf{H}\mathbf{x}_t + \mathbf{n}_r\|^2\right\}}\right) \\
&= \min\left(1, \frac{P}{\mathsf{Tr}\left\{\mathbf{H}\mathbf{Q}_{\mathbf{x}_t}\mathbf{H}^H + \mathbf{I}\right\}}\right) \\
&= \min\left(1, \frac{P}{\mathsf{Tr}\left\{\frac{P}{m}\mathbf{H}\mathbf{H}^H + \mathbf{I}\right\}}\right) \\
&= \min\left(1, \frac{P}{\frac{P}{m}\|\mathbf{H}\|^2 + m}\right). \qquad (5.11)
\end{aligned}$$

From the above equation, we have

$$\mathbb{P}\{\delta \geq \delta_0\} < \mathbb{P}\left\{\|\mathbf{H}\|^2 > P^{\delta_0-\epsilon}\right\}, \qquad (5.12)$$

for all $\delta_0, \epsilon > 0$. Noting that $\|\mathbf{H}\|^2$ has Chi-square distribution with $2mp$ degrees of freedom, it follows that

$$\mathbb{P}\left\{\|\mathbf{H}\|^2 > P^{\delta_0-\epsilon}\right\} \sim \frac{P^{mp(\delta_0-\epsilon)}}{(mp)!}\exp\left\{-P^{\delta_0-\epsilon}\right\}. \qquad (5.13)$$

Choosing $\epsilon = \frac{\delta_0}{2}$, the above equation implies that for $\delta_0 > 0$, the probability $\mathbb{P}\{\delta \geq \delta_0\}$ approaches to zero much faster than polynomially. More precisely, defining the event $\mathscr{F} \equiv \{\delta \dot{>} 0\}$, we have $\mathbb{P}\{\mathscr{F}\} = o(P^{-c})$ for any positive constant $c$. Since $\mathbb{P}\{\mathcal{O}\} \dot{\geq} P^{-(q-r)(p-r)}$ (lower-bound on the outage probability corresponding to the DMT upper-bound ), we can write

$$\begin{aligned}
\mathbb{P}\{\mathcal{O}\} &= \mathbb{P}\{\mathcal{O}|\mathscr{F}\}\mathbb{P}\{\mathscr{F}\} + \mathbb{P}\{\mathcal{O}|\mathscr{F}^c\}\mathbb{P}\{\mathscr{F}^c\} \\
&\overset{(a)}{\sim} \mathbb{P}\{\mathcal{O}|\mathscr{F}^c\}, \qquad (5.14)
\end{aligned}$$

[4]Note that due to the definition of $\alpha$, we always have $\delta \geq 0$.

where $(a)$ follows from the fact that $\mathbb{P}\{\mathscr{F}\} = o\left(\mathbb{P}\{\mathcal{O}\}\right)$. In other words, one can replace $\delta$ with zero in (5.10), which results in

$$\mathbb{P}\{\mathcal{O}\} \doteq \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{j=1}^{\min(p,q)}\left(1-\gamma_j(\mathbf{A}_l)\right)^{+} < Lr\right\}. \tag{5.15}$$

Moreover, we have

$$\begin{aligned}\lambda_i(\mathbf{A}_l) &\leq \|\mathbf{A}_l\|^2 \\ &\overset{(a)}{\leq} \|\mathbf{G}\|^2\|\mathbf{H}\|^2,\end{aligned} \tag{5.16}$$

where $(a)$ results from the fact that $\|\mathbf{AB}\|^2 \leq \|\mathbf{A}\|^2\|\mathbf{B}\|^2$, for any two matrices $\mathbf{A}$ and $\mathbf{B}$. Consider a negative number $\varepsilon$. From the above equation, it follows that

$$\begin{aligned}\mathbb{P}\{\gamma_i(\mathbf{A}_l) \leq \varepsilon\} &\leq \mathbb{P}\left\{\|\mathbf{G}\|^2\|\mathbf{H}\|^2 \geq P^{-\varepsilon}\right\} \\ &\leq \mathbb{P}\left\{\|\mathbf{G}\|^2 \geq P^{-\frac{\varepsilon}{2}}\right\} + \mathbb{P}\left\{\|\mathbf{H}\|^2 \geq P^{-\frac{\varepsilon}{2}}\right\} \\ &\overset{(5.13)}{\sim} \frac{P^{-\frac{np\varepsilon}{2}}}{(np)!}\exp\left\{-P^{-\frac{\varepsilon}{2}}\right\} + \frac{P^{-\frac{mp\varepsilon}{2}}}{(mp)!}\exp\left\{-P^{-\frac{\varepsilon}{2}}\right\} \\ &= o\left(P^{-(q-r)(p-r)}\right).\end{aligned} \tag{5.17}$$

As a result, following (5.14), we can assume that $\gamma_j(\mathbf{A}_l) \geq 0, \forall j = 1, \cdots, \min(p,q)$, in (5.15).

In order to compute the outage probability in (5.15), we need to find the statistical behavior of $\gamma_j(\mathbf{A}_l)$. Since we are interested in upper-bounding the outage probability of the RS scheme, finding an upper-bound for $\gamma_j(\mathbf{A}_l)$, or equivalently, a lower-bound for $\lambda_j(\mathbf{A}_l)$ would be sufficient. This is performed in the following lemma.

**Lemma 5.4** *Consider matrices* $\mathbf{G}$ *and* $\mathbf{H}$ *with the size of* $m \times p$ *and* $p \times n$*, respectively, and a* $p \times p$ *matrix* $\mathbf{\Theta}$*. Assume* $\mathbf{G}$ *and* $\mathbf{H}$ *are singular value decomposed as*

$\mathbf{G} = \mathbf{U}(\mathbf{G})\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}^H(\mathbf{G})$ *and* $\mathbf{H} = \mathbf{U}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{H})\mathbf{V}^H(\mathbf{H})$*, respectively. We have*

$$\lambda_i\left(\mathbf{G}\mathbf{\Theta}\mathbf{H}\right) \geq \lambda_i(\mathbf{G})\lambda_i(\mathbf{H})\lambda_{\min}\left(\mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta}\mathbf{U}_{(1,i)}(\mathbf{H})\right), \tag{5.18}$$

*where* $\lambda_i(\mathbf{A})$ *and* $\lambda_{\min}(\mathbf{A})$ *denote the i'th largest eigenvalue and the minimum eigenvalue of* $\mathbf{A}^H\mathbf{A}$*, respectively, and* $\mathbf{A}_{(a,b)}$ *denotes the submatrix of* $\mathbf{A}$ *consisting of the* $a, a+1, \ldots, b$*'th columns of* $\mathbf{A}$.

**Proof** See Appendix J. ∎

The above lemma relates $\lambda_i(\mathbf{A}_l)$ to $\lambda_i(\mathbf{G})$ and $\lambda_i(\mathbf{H})$, which facilitates the subsequent derivations. A direct consequence of the above lemmas is that

$$\gamma_i(\mathbf{A}_l) \leq \gamma_i(\mathbf{G}) + \gamma_i(\mathbf{H}) + \gamma_{\min}(\mathbf{\Psi}_{i,l}), \tag{5.19}$$

where $\mathbf{\Psi}_{i,l} \triangleq \mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta}_l\mathbf{U}_{(1,i)}(\mathbf{H})$. As the statistical behaviors of $\gamma_i(\mathbf{G})$ and $\gamma_i(\mathbf{H})$ are known from [29], it is sufficient to derive the asymptotic behavior of $\gamma_{\min}(\mathbf{\Psi}_{i,l})$, or equivalently, $\lambda_{\min}(\mathbf{\Psi}_{i,l})$, which is performed in the following lemma:

**Lemma 5.5** *Assuming small enough* $\varepsilon$*, we have*

$$\mathbb{P}\left\{\lambda_{\min}(\mathbf{\Psi}_{i,l}) \leq \varepsilon\right\} \leq \eta\sqrt[i]{\varepsilon}, \tag{5.20}$$

*for some constant* $\eta$.

**Proof** See Appendix K. ∎

A direct consequence of the above lemma is that

$$\mathbb{P}\left\{\gamma_{\min}(\mathbf{\Psi}_{i,l}) > \theta\right\} \dot{\leq} P^{-\frac{\theta}{i}}. \tag{5.21}$$

Defining the $L \times 1$ vector $\boldsymbol{\psi} \triangleq [\psi(1), \cdots, \psi(L)]^T$ as $\psi(l) \triangleq \max_i \gamma_{\min}(\boldsymbol{\Psi}_{i,l})$, we have

$$\begin{aligned}
\mathbb{P}\{\boldsymbol{\psi} \geq \boldsymbol{\psi}_0\} &\overset{(a)}{=} \prod_{l=1}^{L} \mathbb{P}\{\psi(l) \geq \psi_0(l)\} \\
&= \prod_{l=1}^{L} \mathbb{P}\left\{ \bigcup_{i=1}^{\min(p,q)} (\gamma_{\min}(\boldsymbol{\Psi}_{i,l}) \geq \psi_0(l)) \right\} \\
&\overset{(b)}{\dot{\leq}} P^{-\frac{\mathbf{1}\cdot\boldsymbol{\psi}_0}{\min(p,q)}}
\end{aligned} \tag{5.22}$$

As $\boldsymbol{\Theta}_l$'s are independent isotropic unitary matrices, their products with any possibly correlated set of unitary matrices constructs a set of independent isotropic unitary matrices [57]. Accordingly, $\boldsymbol{\Psi}_{i,l}$'s are independent for different values of $l$, which results in $(a)$. Also, $(b)$ follows from Lemma 5.5 and the union bound inequality.

Let us define the $1 \times \min(p, q)$ vectors

$$\begin{aligned}
\boldsymbol{\chi}(\mathbf{H}) &\triangleq \left[\gamma_{\min(p,m)}(\mathbf{H}), \gamma_{\min(p,m)-1}(\mathbf{H}), \ldots, \gamma_{1+\min(p,m)-\min(p,q)}(\mathbf{H})\right], \\
\boldsymbol{\chi}(\mathbf{G}) &\triangleq \left[\gamma_{\min(p,n)}(\mathbf{G}), \gamma_{\min(p,n)-1}(\mathbf{G}), \ldots, \gamma_{1+\min(p,n)-\min(p,q)}(\mathbf{G})\right].
\end{aligned}$$

Notice that these vectors include the log-values of the corresponding $\min(p, q)$ smallest eigenvalues of $\mathbf{H}\mathbf{H}^H$ and $\mathbf{G}\mathbf{G}^H$, respectively. Now, applying the result of Lemma 5.4 to (5.15), we can upper-bound the outage probability of the end-to-end channel as

$$
\begin{aligned}
\mathbb{P}\{\mathcal{O}\} &\dot{\leq} \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\gamma_i(\mathbf{G})-\gamma_i(\mathbf{H})-\gamma_{\min}(\mathbf{\Psi}_{i,l})\right)^{+} < Lr\right\} \\
&\dot{\leq} \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\gamma_i(\mathbf{G})-\gamma_i(\mathbf{H})-\psi(l)\right)^{+} < Lr\right\} \\
&\overset{(a)}{\dot{\leq}} \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\gamma_{1+\min(p,n)-i}(\mathbf{G})-\gamma_{1+\min(p,m)-i}(\mathbf{H})-\psi(l)\right)^{+} < Lr\right\} \\
&= \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right)^{+} < Lr\right\}, \qquad (5.23)
\end{aligned}
$$

where $(a)$ follows from the fact that the log-values ($\gamma_i$'s) corresponding to the smallest eigenvalues of $\mathbf{H}\mathbf{H}^H$ and $\mathbf{G}\mathbf{G}^H$ are greater than the log-values corresponding to the largest eigenvalues of these matrices. According to (5.22), to upper-bound the outage probability, it is sufficient to upper-bound the probability of the region of $(\boldsymbol{\psi}, \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\chi}(\mathbf{G}))$ that satisfies (5.23). The following lemma gives a general formula for computing such an upper-bound:

**Lemma 5.6** *Consider a fixed region $\mathcal{R} \subseteq [0,\infty)^n$. Assume that a uniformly continuous*[5] *non-negative function $f(\mathbf{x})$ ($f(\mathbf{x}) \geq 0$) is defined over $[0,\infty)^n$ such that for all $\mathbf{x} \in [0,\infty)^n$ we have $\mathbb{P}\{\mathbf{y} \geq \mathbf{x}\} \dot{\leq} P^{-f(\mathbf{x})}$. Then, we have*

$$\mathbb{P}\{\mathbf{x} \in \mathcal{R}\} \dot{\leq} P^{-\inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})}. \qquad (5.24)$$

**Proof** See Appendix L. ∎

[5]A *uniformly continuous* function $f : \mathcal{M} \to \mathcal{N}$ where $\mathcal{M} \subseteq \mathbb{R}^m, \mathcal{N} \subseteq \mathbb{R}^n$ is a function that has the following property: for every $\epsilon$, there exists a constant $g(\epsilon) > 0$ such that for all $\mathbf{x}, \mathbf{y} \in \mathcal{M}$, $\|\mathbf{x}-\mathbf{y}\| \leq g(\epsilon)$, we have $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq \epsilon$.

According to the upper-bound in (5.22) and the distribution of $\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H})$ derived in [29], we have

$$\begin{aligned}
&\mathbb{P}\left\{\boldsymbol{\psi} \geq \hat{\boldsymbol{\psi}}, \boldsymbol{\chi}(\mathbf{G}) \geq \boldsymbol{\chi}', \boldsymbol{\chi}(\mathbf{H}) \geq \boldsymbol{\chi}''\right\} \dot{\leq} P^{-\frac{1}{\min(p,q)}\sum_{l=1}^{L}\hat{\psi}(l) - \sum_{i=1}^{\min(p,q)} \mathsf{a}_i \chi_i'' + \mathsf{b}_i \chi_i'} \\
&\overset{(a)}{\dot{\leq}} P^{-\frac{1}{\min(p,q)}\sum_{l=1}^{L}\hat{\psi}(l) - \sum_{i=1}^{\min(p,q)} (2i-1+|p-q|)\left(\chi_i' + \chi_i''\right)}, \qquad (5.25)
\end{aligned}$$

where $\mathsf{a}_i \triangleq 2[i + \min(p,m) - \min(p,q)] - 1 + |p-m|$ and $\mathsf{b}_i \triangleq 2[i + \min(p,n) - \min(p,q)] - 1 + |p-n|$ and $(a)$ follows from the fact that

$$\begin{aligned}
\mathsf{a}_i &= 2[i + \min(p,m) - \min(p,q)] - 1 + |p-m| \\
&= 2i - 1 + m + p - 2\min(p,q) \\
&\geq 2i - 1 + q + p - 2\min(p,q) \\
&= 2i - 1 + |p-q|, \qquad (5.26)
\end{aligned}$$

and similarly, $\mathsf{b}_i \geq 2i - 1 + |p - q$. Now, we can apply the result of Lemma 5.6 to the region defined in (5.23) and the upper-bound derived in (5.25). Accordingly, we have

$$\mathbb{P}\left\{\mathcal{O}\right\} \dot{\leq} P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\psi}) \in \mathcal{R}} \frac{1}{\min(p,q)}\sum_{l=1}^{L}\psi(l) + \sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)(\chi_i(\mathbf{G}) + \chi_i(\mathbf{H}))}, \quad (5.27)$$

where the region $\mathcal{R}$ is defined as

$$\begin{aligned}
\mathcal{R} \triangleq \Bigg\{ (\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\psi}) \Bigg| \boldsymbol{\psi} \geq \mathbf{0}, \quad \sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)} \left(1 - \chi_i(\mathbf{G}) - \chi_i(\mathbf{H}) - \psi(l)\right)^{+} \leq Lr, \\
\chi_1(\mathbf{G}) \geq \cdots \geq \chi_{\min(p,q)}(\mathbf{G}) \geq 0, \quad \chi_1(\mathbf{H}) \geq \cdots \geq \chi_{\min(p,q)}(\mathbf{H}) \geq 0 \Bigg\}. \quad (5.28)
\end{aligned}$$

Let us assume $L \geq \min(p,q)\left(\sum_{i=1}^{\min(p,q)} 2i - 1 + |p-q|\right) = \min^2(p,q)\max(p,q)$. We define $\min(p,q) \times 1$ vector $\boldsymbol{\varphi} \triangleq [\varphi_1, \cdots, \varphi_{\min(p,q)}]^T$ as $\varphi_i \triangleq \chi_i(\mathbf{G}) + \chi_i(\mathbf{H}) +$

$\frac{1}{L}\sum_{l=1}^{L}\psi(l)$. For each $(\boldsymbol{\chi}(\mathbf{G}),\boldsymbol{\chi}(\mathbf{H}),\boldsymbol{\psi})\in\mathcal{R}$, we have

$$\begin{aligned} Lr &\geq \sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right)^+ \\ &= \sum_{i=1}^{\min(p,q)}\sum_{l=1}^{L}\max\left\{0,1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right\} \\ &\geq \sum_{i=1}^{\min(p,q)}\max\left\{0,\sum_{l=1}^{L}1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right\} \\ &= L\sum_{i=1}^{\min(p,q)}\left(1-\varphi_i\right)^+. \end{aligned} \tag{5.29}$$

On the other hand, according to (5.27) we have

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &\dot{\leq} P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}),\boldsymbol{\chi}(\mathbf{H}),\boldsymbol{\psi})\in\mathcal{R}}\frac{1}{\min(p,q)}\sum_{l=1}^{L}\psi(l)+\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)(\chi_i(\mathbf{G})+\chi_i(\mathbf{H}))} \\ &\dot{\leq} P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}),\boldsymbol{\chi}(\mathbf{H}),\boldsymbol{\psi})\in\mathcal{R}}\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)\varphi_i} \\ &\overset{(5.29)}{\dot{\leq}} P^{-\min_{\boldsymbol{\varphi}\in\hat{\mathcal{R}}}\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)\varphi_i}, \end{aligned} \tag{5.30}$$

where $\hat{\mathcal{R}}$ is defined as

$$\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\varphi}\,\middle|\,\varphi_1\geq\cdots\geq\varphi_{\min(p,q)}\geq 0,\ \sum_{i=1}^{\min(p,q)}\left(1-\varphi_i\right)^+\leq r\right\}, \tag{5.31}$$

noting that according to the definition of $\boldsymbol{\varphi}$ we can easily conclude that $\varphi_1\geq\cdots\geq\varphi_{\min(p,q)}\geq 0$. According to [29], (5.30) defines the probability of outage from the rate $r\log(P)$ in an equivalent $p\times q$ point-to-point multi-antenna Rayleigh fading channel. Hence, we have

$$d_{RS}(r)\geq d_{p\times q}(r). \tag{5.32}$$

On the other hand, due to the cut-set bound Theorem [12] we know that the DMT of the system is upper-bounded by the minimum of the DMT of the equivalent point-to-point $p \times m$ and $n \times p$ multi-antenna channels. Hence,

$$d_{RS}(r) \leq d_{opt}(r) = d_{p\times q}(r). \tag{5.33}$$

Comparing (5.32) and (5.33) completes the proof. ∎

The statement of Theorem 5.3 can be generalized to multi-hop networks. However, in a general multi-hop network, the RS scheme does not necessarily achieve the optimum DMT for any number of antennas at the network nodes. The following theorem gives a sufficient condition for the RS scheme to achieve the optimal DMT in a multi-hop network:

**Theorem 5.7** *Consider a multi-antenna multi-hop network consisting of a single source and destination and full-duplex relays, with exactly one relay in each hop. Assume that each relay is connected to the relays in the previous and next hop. Moreover, assume that for a fixed* $1 \leq m \leq h$*, we have* $\max(N_m, N_{m-1}) \leq \min(N_0, N_1, \ldots, N_{m-2}, N_{m+1}, \ldots, N_h)$ *where* $h$ *denotes the number of hops and* $N_i$ *denotes the number of antennas at the relay in the* $i$*'th hop. (*$N_0$ *and* $N_h$ *denote the number of antennas at the source and destination, respectively). Providing* $L$ *is large enough such that* $L \geq \min^2(N_m, N_{m-1}) \max(N_m, N_{m-1})$*, the RS scheme achieves the optimum DMT, which is a piecewise-linear function connecting the points* $(k, (N_m - k)(N_{m-1} - k)), k = 0, 1, \ldots, \min(N_m, N_{m-1})$*.*

**Proof** Using the same argument as in Theorem 5.3, we can show that the probability of outage from the rate $r\log(P)$ is equal to

$$\mathbb{P}\{\mathcal{O}\} \doteq \mathbb{P}\left\{\sum_{l=1}^{L}\sum_{j=1}^{N_{\min}} (1-\gamma_j(\mathbf{A}_l))^+ < Lr\right\}, \tag{5.34}$$

where $N_{\min} \triangleq \min\{N_m, N_{m-1}\}$, $A_l \triangleq \mathbf{G}_h \mathbf{\Theta}_{l,h-1} \mathbf{G}_{h-1} \cdots \mathbf{\Theta}_{l,1} \mathbf{G}_1$, and $\mathbf{G}_i$ denotes the channel matrix between the nodes of the $i$'th hop and $i-1$'th hop. On the other hand, applying the argument of Lemma 5.4, we have

$$\gamma_i(\mathbf{A}_l) = \gamma_i(\mathbf{G}_h) + \gamma_{\min}(\mathbf{\Psi}_{i,l,h-1}) + \gamma_i(\mathbf{G}_{h-1}) + \cdots + \gamma_{\min}(\mathbf{\Psi}_{i,l,1}) + \gamma_i(\mathbf{G}_1), \quad (5.35)$$

where $\mathbf{\Psi}_{i,l,j} \triangleq \mathbf{V}^H_{(1,i)}(\mathbf{G}_{j+1}) \mathbf{\Theta}_{l,j} \mathbf{U}_{(1,i)}(\mathbf{G}_j \mathbf{\Theta}_{l,j-1} \mathbf{G}_{j-1} \cdots \mathbf{\Theta}_{l,1} \mathbf{G}_1)$. Moreover, we can easily check that the statement of Lemma 5.5 is yet valid. Hence, similar to the proof of Theorem 5.3, we define the $L \times 1$ vector $\boldsymbol{\psi} \triangleq [\psi(1), \cdots, \psi(L)]^T$ as $\psi(l) \triangleq \max_{i,j} \gamma_{\min}(\mathbf{\Psi}_{i,l,j})$. We have

$$\begin{aligned}
\mathbb{P}\{\boldsymbol{\psi} \geq \boldsymbol{\psi}_0\} &\overset{(a)}{=} \prod_{l=1}^{L} \mathbb{P}\{\psi(l) \geq \psi_0(l)\} \\
&= \prod_{l=1}^{L} \mathbb{P}\left\{ \bigcup_{i=1}^{N_{\min}} \bigcup_{j=1}^{h} (\gamma_{\min}(\mathbf{\Psi}_{i,l,j}) \geq \psi_0(l)) \right\} \\
&\overset{(b)}{\dot{\leq}} P^{-\frac{\mathbf{1}\cdot\boldsymbol{\psi}_0}{N_{\min}}}. \qquad (5.36)
\end{aligned}$$

As $\mathbf{\Theta}_{l,j}$'s are independent isotropic unitary matrices, their products with any possibly correlated set of unitary matrices constructs a set of independent isotropic unitary matrices [57]. Accordingly, $\mathbf{\Psi}_{i,l,j}$'s are independent for different values of $l, j$, which results in $(a)$. Also, $(b)$ follows from Lemma 5.5 and the union bound inequality.

Accordingly, applying (5.35) to (5.34) and rewriting inequality series of (5.23), we can upper-bound the outage probability as

$$\mathbb{P}\{\mathcal{O}\} \dot{\leq} \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{i=1}^{N_{\min}} \left(1 - \sum_{j=1}^{h} \chi_i(\mathbf{G}_j) - \psi(l)\right)^+ < Lr \right\}, \quad (5.37)$$

where $\chi_i(\mathbf{G}_j) \triangleq \gamma_{N_{\min}+1-i}(\mathbf{G}_j)$, i.e., the reverse ordering of $\gamma_i(\mathbf{G}_j)$'s. Let us define the $N_{\min} \times 1$ vectors $\boldsymbol{\chi}(\mathbf{G}_j)$'s as $\boldsymbol{\chi}(\mathbf{G}_j) \triangleq [\chi_1(\mathbf{G}_j), \chi_2(\mathbf{G}_j), \ldots, \chi_{N_{\min}}(\mathbf{G}_j)]^T$

containing the corresponding log-values of the $N_{\min}$ smallest eigenvalues of $\mathbf{G}_j\mathbf{G}_j^H$. Notice that $\chi_1(\mathbf{G}_j) \geq \chi_2(\mathbf{G}_j) \geq \cdots \geq \chi_{N_{\min}}(\mathbf{G}_j) \geq 0$. According to the upper-bound in (5.36) and the statistical behavior of the eigenvalues of $\mathbf{G}_j\mathbf{G}_j^H$ derived in [29], we have

$$\begin{aligned}
&\mathbb{P}\left\{\boldsymbol{\psi} \geq \hat{\boldsymbol{\psi}}, \boldsymbol{\chi}(\mathbf{G}_j) \geq \hat{\boldsymbol{\chi}}(\mathbf{G}_j), j=1,\ldots,h\right\} \quad \dot{\leq} \\
&P^{-\frac{1}{N_{\min}}\sum_{l=1}^{L}\hat{\psi}(l)-\sum_{j=1}^{h}\sum_{i=1}^{N_{\min}}(2(i+\min(N_j,N_{j-1})-N_{\min})-1+|N_j-N_{j-1}|)\hat{\chi}_i(\mathbf{G}_j)} \overset{(a)}{\dot{\leq}} \\
&P^{-\frac{1}{N_{\min}}\sum_{l=1}^{L}\hat{\psi}(l)-\sum_{i=1}^{N_{\min}}(2i-1+|N_m-N_{m-1}|)\left(\sum_{j=1}^{h}\hat{\chi}_i(\mathbf{G}_j)\right)}. 
\end{aligned} \tag{5.38}$$

Here, $(a)$ results from the fact that $2\min(N_j, N_{j-1}) - 2N_{\min} + |N_j - N_{j-1}| = N_j + N_{j-1} - 2N_{\min} \geq N_m + N_{m-1} - 2N_{\min} = |N_m - N_{m-1}|$ which comes form the assumption of $\max(N_m, N_{m-1}) \leq \min(N_0, N_1, \ldots, N_{m-2}, N_{m+1}, \ldots, N_h)$. Now, we can apply the result of Lemma 5.6 to the region defined in (5.38) and the upper-bound derived in (5.37). Accordingly, we have

$$\begin{aligned}
&\mathbb{P}\{\mathcal{O}\} \dot{\leq} \\
&P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}_1),\ldots,\boldsymbol{\chi}(\mathbf{G}_h),\boldsymbol{\psi})\in\mathcal{R}}\frac{1}{N_{\min}}\sum_{l=1}^{L}\psi(l)+\sum_{i=1}^{N_{\min}}(2i-1+|N_m-N_{m-1}|)\left(\sum_{j=1}^{h}\hat{\chi}_i(\mathbf{G}_j)\right)},
\end{aligned} \tag{5.39}$$

where the region $\mathcal{R}$ is defined as

$$\begin{aligned}
\mathcal{R} \triangleq &\left\{(\boldsymbol{\chi}(\mathbf{G}_1),\ldots,\boldsymbol{\chi}(\mathbf{G}_h),\boldsymbol{\psi}) \middle| \chi_1(\mathbf{G}_j) \geq \cdots \geq \chi_{N_{\min}}(\mathbf{G}_j) \geq 0, j=1,2,\ldots,h\right. \\
&\left., \quad \boldsymbol{\psi} \geq \mathbf{0}, \quad \sum_{l=1}^{L}\sum_{i=1}^{N_{\min}}\left(1-\psi(l)-\sum_{j=1}^{h}\chi_i(\mathbf{G}_j)\right)^{+} \leq Lr\right\}.
\end{aligned} \tag{5.40}$$

Similar to the proof of Theorem 5.3, we define $N_{\min}\times 1$ vector $\boldsymbol{\varphi} \triangleq [\varphi_1, \cdots, \varphi_{N_{\min}}]^T$ as $\varphi_i \triangleq \sum_{j=1}^{h}\chi_i(\mathbf{G}_j) + \frac{1}{L}\sum_{l=1}^{L}\psi(l)$. Rewriting the inequality series in (5.29) and (5.30), we can upper-bound the outage probability as

$$\mathbb{P}\{\mathcal{O}\}\dot{\leq}P^{-\min_{\boldsymbol{\varphi}\in\hat{\mathcal{R}}}\sum_{i=1}^{N_{\min}}(2i-1+|N_m-N_{m-1}|)\varphi_i}, \tag{5.41}$$

where $\hat{\mathcal{R}}$ is defined as

$$\hat{\mathcal{R}} \triangleq \left\{ \boldsymbol{\varphi} \left| \varphi_1 \geq \cdots \geq \varphi_{N_{\min}} \geq 0, \sum_{i=1}^{\min(p,q)} (1-\varphi_i)^+ \leq r \right. \right\}. \tag{5.42}$$

According to [29], (5.41) and (5.42) define the probability of outage from the rate $r\log(P)$ in an equivalent $N_m \times N_{m-1}$ point-to-point multi-antenna Rayleigh fading channel. Hence, we have

$$d_{RS}(r) \geq d_{N_m \times N_{m-1}}(r). \tag{5.43}$$

On the other hand, due to the cut-set bound Theorem [12], we know that the DMT of the system is upper-bounded by the minimum of the DMT of the channels of different hops. Hence,

$$d_{RS}(r) \leq d_{opt}(r) = d_{N_m \times N_{m-1}}(r). \tag{5.44}$$

Comparing (5.43) and (5.44) completes the proof. ∎

**Corollary 5.8** *Consider a multi-antenna multi-hop network consisting of a single source, a single destination and full-duplex relays with exactly one relay in each hop and assume that all the nodes are equipped with $N$ antennas. Providing $L$ is large enough such that $L \geq N^3$, the RS scheme achieves the optimum DMT, which is the piecewise-linear function connecting the points $(k, (N-k)^2), k = 0, 1, \ldots, N$.*

### 5.2.2 Parallel Relay Network

In this subsection, we consider the setup of a multi-antenna parallel relay network. In specific, we consider a two-hop network consisting of $K > 1$ half-duplex relays with the assumption that there is no direct link between the source and the destination. The source and the destination are shown by nodes 0 and $K+1$, respectively, while the $K$ parallel relays are denoted by the nodes $1, 2, \ldots, K$. Earlier, the optimum DMT of the single-antenna parallel relay network is derived in Theorems 3.9 and 3.16. Indeed, it is shown in Theorem 3.16 that the RS scheme can achieve the optimum DMT of the single-antenna multiple-access parallel relay network. However, much less is known regarding the DMT of the multi-antenna parallel relay networks.

Here, we show that the RS scheme achieves a better DMT with respect to the traditional AF relaying and also with respect to the other results reported in the literature. Moreover, we show that the RS scheme achieves the optimum DMT of the 2-relay parallel relay network.

**Theorem 5.9** *Consider a multi-antenna parallel relay network consisting of a source equipped with $m$ antennas, a destination equipped with $n$ antennas and $K$ half-duplex relays each equipped with $p$ antennas. Assume that there exists no direct link between the source and the destination*[6]*. For any fixed $B \geq \min^2(p,q)\max(p,q)$, the RS scheme*[7] *with $L = BK$ number of paths, $S = BK + 1$ number of slots, the path sequence*

$$\mathrm{Q} \equiv (\mathrm{q}_1, \ldots, \mathrm{q}_K, \mathrm{q}_1, \ldots, \mathrm{q}_K, \ldots, \mathrm{q}_1, \ldots, \mathrm{q}_K),$$

*in which $\mathrm{q}_k \equiv (0, k, K+1)$, and the timing sequence $s_{i,j} = i + j - 1$, achieves the diversity gain*

$$d_{RS}(r) \geq K d_{p\times q}\left(\left(1 + \frac{1}{BK}\right) r\right), \tag{5.45}$$

*where $q \triangleq \min(m,n)$ and $d_{p\times q}(r)$ denotes the diversity gain of the point-to-point*

[6]Note that in this theorem, the relays are not assumed to be isolated from each other, i.e., there may exist some links between the relays.

[7]The reader is encouraged to have a look at the definition of RS scheme and its parameters in Chapter 3.

*$p \times q$ multi-antenna Rayleigh fading channel corresponding to the rate $r\log(P)$. Moreover, as $B \to \infty$, the RS scheme achieves the diversity gain $Kd_{p\times q}(r)$.*

**Proof** The proof steps are the same as the ones presented for Theorem 3.9 in Chapter 3 and Theorem 6.7 of [28]. Let us denote the channel between the $k$'th relay and the source and the channel between the $k$'th relay and the destination by $\mathbf{H}_k$ and $\mathbf{G}_k$, respectively. Moreover, let us define $r(i) \triangleq (i-1) \mod K+1$. Similar to the proof of Theorem 3.9, one can easily show that the end-to-end channel from the source to the destination can be shown by a block lower-triangular matrix. More precisely, we have

$$\mathbf{y} = \mathcal{F}\mathbf{x} + \mathcal{Q}\mathbf{n}_r + \mathbf{n}_d. \tag{5.46}$$

Here, $\mathbf{x}$ denotes the vector corresponding to all the paths transmitted by the source, $\mathbf{y}$ denotes the vector corresponding to all the paths received by the destination,

$$\mathcal{F} = \begin{pmatrix} \mathbf{F}_{1,1} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{F}_{2,1} & \mathbf{F}_{2,2} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{F}_{L,1} & \mathbf{F}_{L,2} & \cdots & \mathbf{F}_{L,L} \end{pmatrix}, \tag{5.47}$$

where[8] $\mathbf{F}_{i,i} = \mathbf{G}_{r(i)}\alpha_i\boldsymbol{\Theta}_i\mathbf{H}_{r(i)}$, and

$$\mathcal{Q} = \begin{pmatrix} \mathbf{Q}_{1,1} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{Q}_{L,1} & \mathbf{Q}_{L,2} & \cdots & \mathbf{Q}_{L,L} \end{pmatrix}, \tag{5.48}$$

[8]As only the value of $\mathbf{F}_{i,i}$ is needed in the proof of Theorem 3.9, we just give the value of $\mathbf{F}_{i,i}$ here. The formula for $\mathbf{F}_{i,j}$, $i < j$, is much more complicated and hence, we decide not to bring it. The same thing for $\mathbf{Q}_{i,j}$ defined right after this.

where $\mathbf{Q}_{i,i} = \mathbf{G}_{r(i)}\alpha_i\mathbf{\Theta}_i$. The DMT of the end-to-end channel is equal to

$$\begin{aligned} d_{RS}(r) &= \lim_{P\to\infty} -\frac{\log\left(\mathbb{P}\left\{I(\mathbf{x};\mathbf{y}) < (L+1)r\log(P)\right\}\right)}{\log(P)} \\ &= \lim_{P\to\infty} \frac{\log\left(\mathbb{P}\left\{\log\left(\left|\mathbf{I}_{Ln} + P\mathcal{F}\mathcal{F}^H\left(\mathbf{I}_{Ln} + \mathcal{Q}\mathcal{Q}^H\right)^{-1}\right|\right) < (L+1)r\log(P)\right\}\right)}{-\log(P)}. \end{aligned} \tag{5.49}$$

Noting $\mathcal{F}$ is block lower-diagonal and applying Theorem 3.3 in [28], we have

$$\left|\mathbf{I}_{Ln} + P\mathcal{F}\mathcal{F}^H\left(\mathbf{I}_{Ln} + \mathcal{Q}\mathcal{Q}^H\right)^{-1}\right| \geq \prod_{l=1}^{L}\left|\mathbf{I}_n + P\mathbf{F}_{l,l}\mathbf{F}_{l,l}^H\left(\mathbf{I}_n + \sum_{i=1}^{l}\mathbf{Q}_{l,i}\mathbf{Q}_{l,i}^H\right)^{-1}\right|. \tag{5.50}$$

Note that according to the constraint in the RS scheme, we have $\alpha_l \leq 1$. Hence, one can apply the same argument as in Lemma 5.2 and show that

$$\begin{aligned} &\mathbb{P}\left\{\prod_{l=1}^{L}\left|\mathbf{I}_n + P\mathbf{F}_{l,l}\mathbf{F}_{l,l}^H\left(\mathbf{I}_n + \sum_{i=1}^{l}\mathbf{Q}_{l,i}\mathbf{Q}_{l,i}^H\right)^{-1}\right| < P^{(L+1)r}\right\} \doteq \\ &\mathbb{P}\left\{\prod_{l=1}^{L}\left|\mathbf{I}_n + P\mathbf{F}_{l,l}\mathbf{F}_{l,l}^H\right| < P^{(L+1)r}\right\}. \end{aligned} \tag{5.51}$$

Moreover, using the argument in the proof of Theorem 5.3, one can show that with probability one, we have $\alpha_l \doteq 1$. Hence, defining $\mathbf{A}_{k,b} \triangleq \mathbf{G}_k\mathbf{\Theta}_{(b-1)K+k}\mathbf{H}_k$, we have

$$\mathbb{P}\left\{\prod_{l=1}^{L}\left|\mathbf{I}_n + P\mathbf{F}_{l,l}\mathbf{F}_{l,l}^H\right| < P^{(L+1)r}\right\} \doteq \mathbb{P}\left\{\prod_{k=1}^{K}\prod_{b=1}^{B}\left|\mathbf{I}_n + P\mathbf{A}_{k,b}\mathbf{A}_{k,b}^H\right| < P^{(L+1)r}\right\}. \tag{5.52}$$

Let us define the $1\times K$ random vector $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \ldots, \sigma_K]$ where $\sigma_k$ is defined as $\sigma_k \triangleq \frac{\sum_{b=1}^{B}\log\left|\mathbf{I}_n + P\mathbf{A}_{k,b}\mathbf{A}_{k,b}^H\right|}{B\log(P)}$. Notice that $\sigma_k$'s are independent of each other. As $B \geq \min^2(p,q)\max(p,q)$, we can apply Theorem 5.3 to $\sigma_k$'s. Hence, for any fixed $1\times K$ vector $\hat{\boldsymbol{\sigma}} \geq \mathbf{0}$, we have

$$\mathbb{P}\left\{\boldsymbol{\sigma} \geq \hat{\boldsymbol{\sigma}}\right\} \overset{(a)}{\doteq} P^{-\sum_{k=1}^{K} d_{p\times q}(\hat{\sigma}_k)}. \tag{5.53}$$

Here, $(a)$ results from Theorem 5.3 and the fact that $\sigma_k$'s are independent of each other. Denoting the outage event as $\mathcal{O}$, according to (5.49), (5.50), (5.51), and (5.52), we have

$$\mathbb{P}\left\{\mathcal{O}\right\} \dot{\leq} \mathbb{P}\left\{\sum_{k=1}^{K} \sigma_k \leq \left(K + \frac{1}{B}\right) r\right\}. \tag{5.54}$$

Let us define the region $\mathcal{R} \triangleq \left\{\boldsymbol{\sigma} \geq \mathbf{0} \middle| \sum_{k=1}^{K} \sigma_k \leq \left(K + \frac{1}{B}\right) r\right\}$. We have

$$\mathbb{P}\left\{\mathcal{O}\right\} \dot{\leq} \mathbb{P}\left\{\mathcal{R}\right\} \overset{(a)}{\dot{\leq}} P^{-\min_{\boldsymbol{\sigma} \in \mathcal{R}} \sum_{k=1}^{K} d_{p \times q}(\sigma_k)} \overset{(b)}{\doteq} P^{-K d_{p \times q}\left(\left(1 + \frac{1}{BK}\right) r\right)}. \tag{5.55}$$

Here, $(a)$ results from Lemma 5.6. $(b)$ results from the fact that $d_{p \times q}(r)$ is a convex decreasing function and, as a result, we have $\frac{1}{K} \sum_{k=1}^{K} d_{p \times q}(\sigma_k) \geq d_{p \times q}(\frac{1}{K} \sum_{k=1}^{K} \sigma_k) \geq d_{p \times q}\left(\left(1 + \frac{1}{BK}\right) r\right)$. (5.55) completes the proof of the Theorem. ∎

In the following Theorem, we show that the RS scheme achieves the optimum DMT for the two-relays half-duplex parallel relay network in which there exists no direct link between the relays and $m = n$, but the two parallel relays can have a different number of antennas. Figure 5.2 shows the schematic of such a network.

**Theorem 5.10** *Consider a multi-antenna parallel relay network consisting of a source and a destination each equipped with $m$ antennas, and $K = 2$ half-duplex relays equipped with $n_k, k = 1, 2$ antennas. Assume that there exists no direct link between the source and the destination and also between the two relays. Consider the RS scheme with $L = BK$, $S = BK + 1$, and the path and timing sequences defined in Theorem 5.9. As $B \to \infty$, the RS scheme achieves the optimum DMT of the network.*

**Proof** First, notice that according to the argument of Theorem 5.9, as $B \to \infty$, the RS scheme achieves the DMT $d_{RS,\infty}(r) \triangleq \min_{0 \leq \nu \leq 2r} d_{m \times n_1}(\nu) + d_{m \times n_2}(2r - \nu)$.

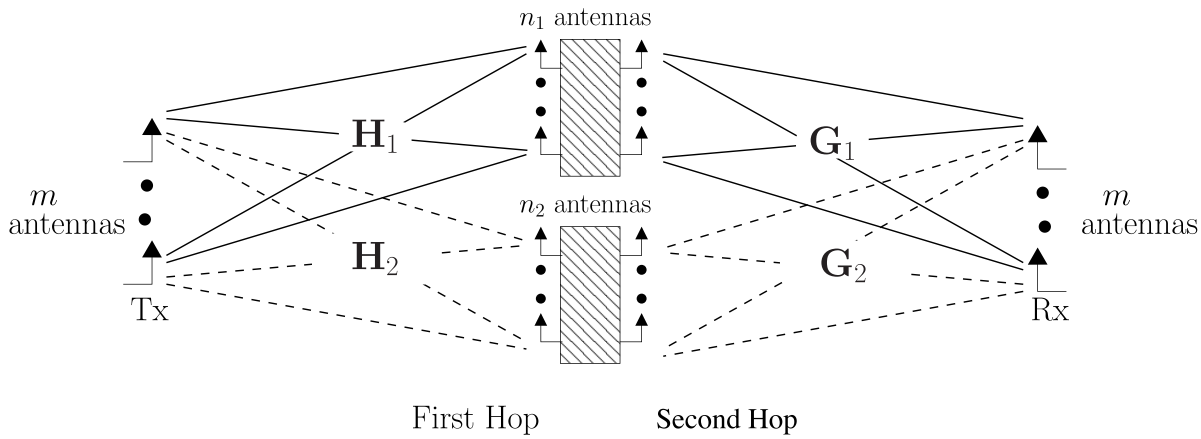


Figure 5.2: A schematic of the MIMO parallel 2 relays network with no direct link between the source and the destination and also between the relays

Now, to prove the Theorem, we just have to show that $d_{RS,\infty}(r)$ is indeed an upper-bound for the optimum DMT. According to the cut-set Theorem [12], we have an upper-bound for the capacity of the network for each channel realization. Hence, we can apply the cut-set Theorem to find an upper-bound for the optimum DMT. In general, for any arbitrary half-duplex relay network with $K$ relays and any set $\{0\} \subseteq \mathcal{S} \subseteq \{0, 1, \dots, K\}$, we say the network is in the *state* $\mathcal{S}$, if the network nodes in $\mathcal{S}$ are transmitting and the network nodes in $\mathcal{S}^c \triangleq \{0, 1, \dots, K+1\} / \mathcal{S}$ are receiving. Notice that as the source is always transmitting and the destination is always receiving, we have $0 \in \mathcal{S}, K+1 \in \mathcal{S}^c$. Accordingly, we define a $1 \times 2^K$ *state vector* $\boldsymbol{\rho}$ such that for any set $\{0\} \subseteq \mathcal{S} \subseteq \{0, 1, \dots, K\}$, $\rho_{\mathcal{S}}$ shows the portion of time that the half-duplex relay network spends in the *state* $\mathcal{S}$ ($\sum_{\{0\} \subseteq \mathcal{S} \subseteq \{0,1,\dots,K\}} \rho_{\mathcal{S}} = 1$). As the channels are assumed to be fixed for the whole transmission period and the relay nodes and the source are assumed to have no channel state knowledge about their forward channels, we can assume that a fixed state vector $\boldsymbol{\rho}$ is associated with the strategy that achieves the optimum DMT. Denoting the outage event by $\mathcal{O}$, for any general half-duplex relay network consisting of $K$ relays, we have

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{O}\right\} &\overset{(a)}{\geq} \min_{\boldsymbol{\rho}} \mathbb{P}\left\{ \bigcup_{\{0\}\subseteq\mathcal{T}\subseteq\{0,1,\ldots,K\}} \right. \\
&\left.\left( \sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,\ldots,K\}} \rho_{\mathcal{S}} I\left( X\left(\mathcal{S}\cap\mathcal{T}\right); Y\left(\mathcal{S}^c\cap\mathcal{T}^c\right) \middle| X\left(\mathcal{S}\cap\mathcal{T}^c\right)\right) < r\log(P) \right)\right\} \\
&\overset{(b)}{\doteq} \min_{\boldsymbol{\rho}} \max_{\{0\}\subseteq\mathcal{T}\subseteq\{0,1,\ldots,K\}} \\
&\mathbb{P}\left\{ \sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,\ldots,K\}} \rho_{\mathcal{S}} I\left( X\left(\mathcal{S}\cap\mathcal{T}\right); Y\left(\mathcal{S}^c\cap\mathcal{T}^c\right) \middle| X\left(\mathcal{S}\cap\mathcal{T}^c\right)\right) < r\log(P) \right\}.
\end{aligned} \tag{5.56}$$

Here, $(a)$ follows from the cut-set bound Theorem [12] and $(b)$ follows from the union bound on the probability. Now, in our two-relay parallel setup, let us define two sets $\mathcal{T}_1 \triangleq \{0,1\}$ and $\mathcal{T}_2 \triangleq \{0,2\}$ corresponding to the two cut-sets. Moreover, let us define two events $\mathcal{O}_1$ and $\mathcal{O}_2$ as

$$\begin{aligned}
\mathcal{O}_1 &\triangleq \left\{\left|\mathbf{I}_m + P\mathbf{G}_1\mathbf{G}_1^H\right| \leq \hat{\nu}\log(P), \left|\mathbf{I}_{n_2} + P\mathbf{H}_2\mathbf{H}_2^H\right| \leq (2r-\hat{\nu})\log(P)\right\}, \\
\mathcal{O}_2 &\triangleq \left\{\left|\mathbf{I}_{n_1} + P\mathbf{H}_1\mathbf{H}_1^H\right| \leq \hat{\nu}\log(P), \left|\mathbf{I}_m + P\mathbf{G}_2\mathbf{G}_2^H\right| \leq (2r-\hat{\nu})\log(P)\right\},
\end{aligned}$$

where $\hat{\nu} \triangleq \underset{0\leq\nu\leq 2r}{\operatorname{argmin}}\ d_{m\times n_1}(\nu) + d_{m\times n_2}(2r-\nu)$. Hence, in our setup, (5.56) can be simplified as

$$
\begin{aligned}
\mathbb{P}\left\{\mathcal{O}\right\} &\overset{(a)}{\dot{\geq}} \min_{\boldsymbol{\rho}} \max \\
&\left(\mathbb{P}\left\{\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,2\}} \rho_{\mathcal{S}} I\left(X\left(\mathcal{S}\cap\mathcal{T}_1\right);Y\left(\mathcal{S}^c\cap\mathcal{T}_1^c\right)\middle| X\left(\mathcal{S}\cap\mathcal{T}_1^c\right)\right)\leq r\log(P)\right\},\right. \\
&\left.\mathbb{P}\left\{\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,2\}} \rho_{\mathcal{S}} I\left(X\left(\mathcal{S}\cap\mathcal{T}_2\right);Y\left(\mathcal{S}^c\cap\mathcal{T}_2^c\right)\middle| X\left(\mathcal{S}\cap\mathcal{T}_2^c\right)\right)\leq r\log(P)\right\}\right) \\
&\geq \min_{\boldsymbol{\rho}} \max \Big( \\
&\mathbb{P}\left\{\left(\rho_{\{0,1\}}+\rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m+P\mathbf{G}_1\mathbf{G}_1^H\right|+\left(\rho_{\{0\}}+\rho_{\{0,1\}}\right)\left|\mathbf{I}_{n_2}+P\mathbf{H}_2\mathbf{H}_2^H\right|\leq r\log(P)\right\}, \\
&\mathbb{P}\left\{\left(\rho_{\{0,2\}}+\rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m+P\mathbf{G}_2\mathbf{G}_2^H\right|+\left(\rho_{\{0\}}+\rho_{\{0,2\}}\right)\left|\mathbf{I}_{n_1}+P\mathbf{H}_1\mathbf{H}_1^H\right|\leq r\log(P)\right\}\Big) \\
&\overset{(b)}{\geq} \min_{\boldsymbol{\rho}} \max \left(\mathbf{1}\left[r-\left(\rho_{\{0,1,2\}}+\rho_{\{0,1\}}\right)\hat{\nu}-\left(\rho_{\{0,1\}}+\rho_{\{0\}}\right)(2r-\hat{\nu})\right]\mathbb{P}\left\{\mathcal{O}_1\right\},\right. \\
&\left.\mathbf{1}\left[r-\left(\rho_{\{0\}}+\rho_{\{0,2\}}\right)\hat{\nu}-\left(\rho_{\{0,2\}}+\rho_{\{0,1,2\}}\right)(2r-\hat{\nu})\right]\mathbb{P}\left\{\mathcal{O}_2\right\}\right) \\
&\overset{(c)}{\doteq} P^{-d_{RS,\infty}(r)}, \qquad (5.57)
\end{aligned}
$$

where $\mathbf{1}[x]=1$ for $x\geq 0$ and is 0 otherwise. Here, $(a)$ results from taking the maximization of the right-hand side of (5.56) over $\mathcal{T}_1,\mathcal{T}_2$. $(b)$ results from the facts that i) conditioned on $\mathcal{O}_1$ and assuming $r\geq\left(\rho_{\{0,1,2\}}+\rho_{\{0,1\}}\right)\hat{\nu}+\left(\rho_{\{0,1\}}+\rho_{\{0\}}\right)(2r-\hat{\nu})$, we have $\left(\rho_{\{0,1\}}+\rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m+P\mathbf{G}_1\mathbf{G}_1^H\right|+\left(\rho_{\{0\}}+\rho_{\{0,1\}}\right)\left|\mathbf{I}_{n_2}+P\mathbf{H}_2\mathbf{H}_2^H\right|\leq r\log(P)$; and ii) conditioned on $\mathcal{O}_2$ and assuming $r\geq\left(\rho_{\{0\}}+\rho_{\{0,2\}}\right)\hat{\nu}+\left(\rho_{\{0,2\}}+\rho_{\{0,1,2\}}\right)(2r-\hat{\nu})$, we conclude that

$$
\left(\rho_{\{0,2\}}+\rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m+P\mathbf{G}_2\mathbf{G}_2^H\right|+\left(\rho_{\{0\}}+\rho_{\{0,2\}}\right)\left|\mathbf{I}_{n_1}+P\mathbf{H}_1\mathbf{H}_1^H\right|\leq r\log(P).
$$

$(c)$ results from i) $\mathbb{P}\left\{\mathcal{O}_1\right\}=\mathbb{P}\left\{\mathcal{O}_2\right\}\doteq P^{-d_{m\times n_1}(\hat{\nu})-d_{m\times n_2}(2r-\hat{\nu})}=P^{-d_{RS,\infty}(r)}$ and

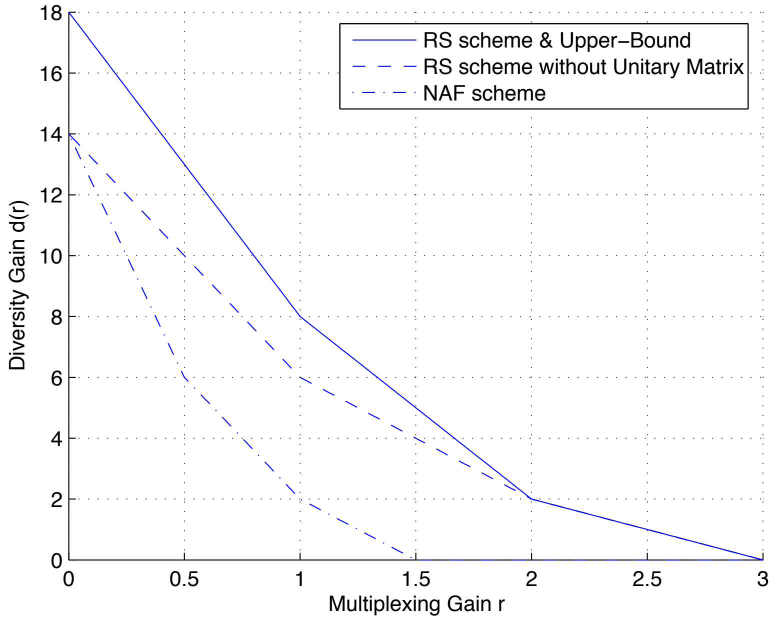


Figure 5.3: Parallel relay network with $K = 2$ relays, each node with 3 antennas and no direct link between source and destination.

ii) $\rho_0 + \rho_{0,1} + \rho_{0,2} + \rho_{0,1,2} = 1$ (due to the definition of $\boldsymbol{\rho}$) which results in having

$$r - \left(\rho_{\{0,1,2\}} + \rho_{\{0,1\}}\right)\hat{\nu} - \left(\rho_{\{0,1\}} + \rho_{\{0\}}\right)(2r - \hat{\nu}) \quad = \quad - \left[r - \left(\rho_{\{0\}} + \rho_{\{0,2\}}\right)\hat{\nu} - \left(\rho_{\{0,2\}} + \rho_{\{0,1,2\}}\right)(2r - \hat{\nu})\right]$$

and consequently,

$$\mathbf{1}\left[r - \left(\rho_{\{0,1,2\}} + \rho_{\{0,1\}}\right)\hat{\nu} - \left(\rho_{\{0,1\}} + \rho_{\{0\}}\right)(2r - \hat{\nu})\right] +$$
$$\mathbf{1}\left[r - \left(\rho_{\{0\}} + \rho_{\{0,2\}}\right)\hat{\nu} - \left(\rho_{\{0,2\}} + \rho_{\{0,1,2\}}\right)(2r - \hat{\nu})\right] = 1.$$

(5.57) completes the proof of the theorem. ∎

Figure 5.3 shows the DMT of various schemes for the parallel relay network with $K = 2$ relays and $m = n = p = 3$, i.e. 3 antennas at each node. As it is shown in Theorem 5.10, the RS scheme achieves the optimum DMT. However, if we do not apply random unitary matrix multiplication at the relay nodes, applying the steps in the proof of Theorem 5.9, one can easily show that the RS scheme achieves the DMT of $Kd_{\mathbf{GH}}(r)$, where $d_{\mathbf{GH}}(r)$ denotes the DMT of the product of the channel matrix $\mathbf{H}$ from the source to the relay and the channel matrix $\mathbf{G}$ from the relay to the destination (see (5.4)). Finally, applying the NAF scheme of [26, 38], one can easily show that the DMT $Kd_{\mathbf{GH}}(2r)$ is achievable.

### 5.2.3 Multiple-Antenna Single Relay Channel

In this subsection, we consider the most-studied scenario in the relay network, the single relay setup, in which a direct link exists between the source and the destination. The relay is assumed to be half-duplex. There have been extensive research on this particular setup toward characterization of the DMT. The authors of [26] have shown that the NAF scheme achieves the best DMT among all possible AF relaying schemes for the Single-Input Single-Output (SISO) single half-duplex relay channel. However, here we show that using independent uniformly random unitary matrices across different time-slots improves the DMT of the NAF scheme for the multi-antenna setup. In order to exploit the potential benefit from random unitary matrix multiplication, the source transmits in $2B$ consecutive time-slots. In the odd time-slots, the relay listens to the source signal. In the even time slots, the relay multiplies the received signal from the last time-slot with a uniformly random unitary matrix and then amplifies the result with the maximum possible coefficient, which is less than or equal to 1. The destination decodes the transmitted message based on the joint decoding of the signal it receives in the $2B$ time-slots.

**Theorem 5.11** *Consider a single relay channel consisting of a source, a half-duplex relay, and a destination equipped with $m$, $p$, and $n$ antennas, respectively. Let us consider a modified NAF scheme that benefits from the random unitary matrix multiplication at the relay node and the joint decoding at the destination side through $2B$ time-slots. Assuming $B \geq \min^2(p,q)\max(p,q)$ where $q \triangleq \min(m,n)$, the modified NAF scheme achieves the following DMT*

$$d_{MNAF}(r) \geq d_{m\times n}(r) + d_{p\times q}(2r). \tag{5.58}$$

**Proof** The proof is similar to the proof of Theorem 5.9. Indeed, assuming the source-destination, source-relay, and relay-destination channel matrices are denoted by $\mathbf{F}$, $\mathbf{H}$, and $\mathbf{G}$, respectively, we can show that the end-to-end channel matrix is equal to

$$\mathcal{F} = \begin{pmatrix} \mathbf{F} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\ \alpha\mathbf{G}\mathbf{\Theta}_1\mathbf{H} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{F} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \alpha\mathbf{G}\mathbf{\Theta}_2\mathbf{H} & \mathbf{F} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{5.59}$$

Here, we observe that the lower-diagonal elements are independent of the diagonal elements. Hence, we can apply Theorem 3.3 in [28]. Accordingly, the DMT corresponding to the end-to-end system is greater than or equal to the summation of the DMT of the forward channel $\mathbf{F}$ and the DMT of a two-hop channel utilized in half of the time. In other words, $d_{MNAF}(r) \geq d_{m\times n}(r) + d_{p\times q}(2r)$. Details of the proof are similar to the proof of Theorem 5.9. ∎

Figure 5.4 compares the achievable DMT of the NAF scheme with the achievable DMT of the modified NAF scheme for the single relay channel with $m = n =$

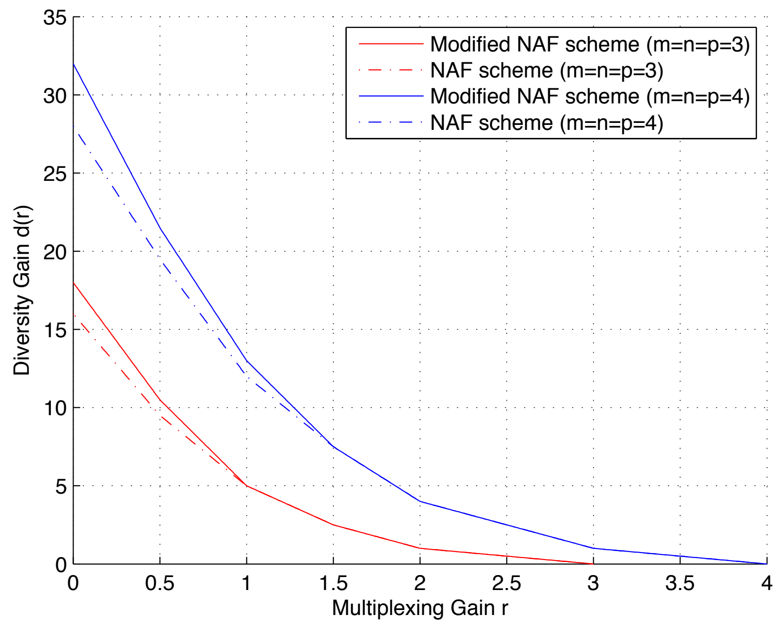


Figure 5.4: DMT of the NAF scheme versus the modified NAF scheme for Multiple-antenna single relay channel with i)3 antennas ii)4 antennas at each node.

$p = 3$ and $m = n = p = 4$ antennas. Reference [54] has shown that the NAF protocol achieves the DMT $d_{NAF}(r) \geq d_{m\times n}(r) + d_{\mathbf{GH}}(2r)$. As we observe, the modified NAF scheme outperforms the NAF scheme for small values of $r$.

### 5.2.4 General Full-Duplex Relay Networks

Here, we generalize the statement of Remark 3.13 to the multi-antenna case. Indeed, it is shown in Chapter 3 that the RS scheme achieves a linear DMT connecting the points $(0, d_{\max})$ and $(r_{\max}, 0)$, where $d_{\max}$ denotes the maximum diversity and $r_{\max}$ denotes the maximum multiplexing-gain (which is 1) for single-antenna full-duplex relay networks whose underlying graph is directed acyclic[9]. Here, we derive the multi-antenna counterpart to Remark 3.13. The reader should refer to Chapter 3 for description of the multi-antenna wireless relay network model and to Definition 3.1 for the definition of edge cut-set and the weight of a cut-set in a graph.

**Theorem 5.12** *Consider a full-duplex multi-antenna multi-relay network with the graph $G = (V, E)$ where $G$ is directed acyclic. Assume each node has $N$ antennas. The RS scheme achieves the following DMT*

$$d_{RS}(r) = \min_{\mathcal{S}} w_G(\mathcal{S}) \frac{d_{N\times N}(r)}{N^2}, \tag{5.60}$$

*where $d_{N\times N}(r)$ denotes the DMT of a $N \times N$ multi-antenna channel.*

**Proof** Using the same path sequence as the one in the proof of Remark 3.13 and applying the result of Corollary 5.8, (5.60) can be derived. The steps are the same as the steps in the proof of Theorem 5.9, noting the equivalent point-to-point channel is block lower-triangular and the function $d_{N\times N}(r)$ is convex decreasing. ∎

**Remark 5.13** *According to* (5.60)*, a specific RS scheme with fixed path and timing sequences can simultaneously achieve the maximum diversity gain, which is* $\min_{\mathcal{S}} w_G(\mathcal{S})$ *(refer to Chapter 4), and the maximum multiplexing gain, which is $N$ in a multi-antenna relay network whose underlying graph is directed acyclic. In other words, the RS scheme is robust in the sense that it achieves the corner-points of the optimum DMT with no modification of the scheme parameters.*

[9] Recall that a directed graph is called directed acyclic if it contains no directed cycle.

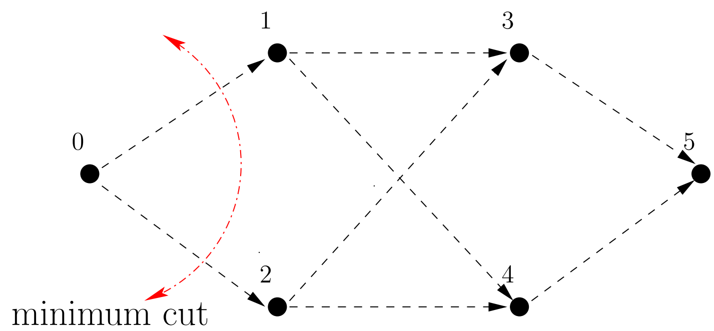


Figure 5.5: An example of a multi-antenna directed acyclic network with full-duplex relays, each node equipped with 2 antennas.

Figure 5.5 shows an example of a directed acyclic network. The relays are operating in the full-duplex mode and each node is equipped with two antennas. Here, the weight of the minimum cut-set depicted in the figure is 8. Hence, applying the argument of Theorem 5.12, the RS scheme achieves $d_{RS}(r) = 2d_{2\times2}(r)$. However, the DMT upper-bound is equal to $d_{ub}(r) = d_{2\times4}(r)$, which is obtained from the same cut-set. Although the two DMT's are equal in the corner-points, they do not coincide in between.

## 5.3 Conclusion

In this chapter, we derived new DMT results in various setups of multi-antenna relay network using AF relaying. For this purpose, the application of RS scheme proposed in Chapter 3 was studied. It was shown that random unitary matrix multiplication at the relay nodes enables the RS scheme to achieve a better DMT in comparison to the traditional AF relaying. First, the multi-antenna full-duplex single-relay two-hop network was studied for which the RS scheme was shown to achieve the optimum DMT. This result was also generalized to the multi-antenna multi-hop full-duplex relay network with one relay in each hop. Next, applying this result, a new achievable DMT was derived for multi-antenna half-duplex parallel relay networks. Interestingly, it turned out that the DMT of the RS scheme is optimum for the multi-antenna half-duplex parallel two-relay setup with no direct link between the relays. Moreover, random unitary matrix multiplication was shown to improve DMT of the NAF scheme of [26] in the setup of multi-antenna single relay channel. Finally, the general full-duplex multi-antenna relay network was studied and a new lower-bound was obtained on DMT using the RS scheme, assuming that the underlying network graph is directed acyclic.

# Chapter 6

# Multiplexing Gain of Multi-Antenna Relay Networks

In this chapter, we investigate the potential benefits of "traditional" AF relaying in the wireless multiple-antenna multiple-relay networks with Rayleigh fading channels. For description of the system model, the reader is referred to Chapter 3. In contrast to more complex AF relaying's such as RS scheme, in traditional AF relaying, each relay node forwards its received signal of the last time-slot in the following time-slot. No channel state knowledge is required at either the source or any of the relay nodes. However, the destination is assumed to know the end-to-end equivalent CSI. This CSI can be acquired through some pilot signals sent by the source through the relay network.

We study the pre-log coefficient of the ergodic capacity in high SNR regime, known as the multiplexing gain. In chapter 5, we have shown existence of certain multi-antenna wireless relay networks for which the RS scheme does not achieve the optimum multiplexing gain. However, here, we prove that the traditional AF relaying achieves the maximum multiplexing gain for any wireless multi-antenna relay network. Furthermore, we characterize the maximum multilexing gain of the network in terms of the minimum vertex cut-set of the underlying graph of the network and show that it can be computed in polynomial-time (with respect to the number of network nodes) using the maximum-flow algorithm. Finally, we show that the argument can be easily extended to the *multicast* and *multi-access* scenarios as well.

The rest of the chapter is organized as follows. Section 6.1 defines the concepts needed in the proof and explains the main results of the chapter. Section 6.2 is dedicated to the proof of the main result. Section 6.3 concludes the chapter.

## 6.1 The Main Result

Recall from Chapter 3 that the wireless relay network can be represented by a graph $G = (V, E)$ such that the adjacent nodes in $V$ are connected by a quasi-static flat Rayleigh-fading channel and non-adjacent nodes are disconnected from each other. In this chapter, we consider the networks with a directed underlying graph. However, it can be easily verified that the argument is yet valid for the undirected graphs. Also, recall that the number of antennas as node $i$ is denoted by $N_i$. Nodes 0 and $K+1$ correspond to the source and the destination nodes, respectively. Moreover, recall from Chapter 3 that the received and the transmitted vectors at the $k$'th node are shown by $\mathbf{y}_k$ and $\mathbf{x}_k$, respectively. Hence, at the receiver side of the $a$'th node, we have

$$\mathbf{y}_a = \sum_{(b,a)\in E} \mathbf{H}_{a,b}\mathbf{x}_b + \mathbf{n}_a, \tag{6.1}$$

where $\mathbf{H}_{a,b}$ shows the $N_a \times N_b$ Rayleigh-distributed channel matrix between the $a$'th and the $b$'th nodes and $\mathbf{n}_a \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{N_a})$ is the additive white Gaussian noise. All nodes have the same power constraint, $P$.

In the studied traditional AF relaying, all the relays are always active and, in each time-slot, each relay sends the amplified version of the signal it has received in the last time-slot. For definition of the edge cut-set and the weight of the edge cut-set, the reader is referred to Definition 3.1. In order to state the main argument of the chapter, we need the following definition, as well.

**Definition 6.1** *For a relay network with the directed connectivity graph $G = (V, E)$, a vertex cut-set on $G$ is defined as a subset $\mathcal{C} \subseteq V$ such that any directed path in $G$ from $0$ to $K+1$ intersects with one of the nodes in $\mathcal{C}$. In other words, in the subgraph of $G$ induced*[1] *by $V - \mathcal{C}$ the destination node $K+1$ is disconnected from the source node, $0$. The capacity of a vertex cut-set is defined as*

$$c_G(\mathcal{C}) = \sum_{v \in \mathcal{C}} N_v. \tag{6.2}$$

It should be noted that according to the above definition, the subsets $\{0\}$ and $\{K+1\}$ are vertex cut-sets on $G$.

**Theorem 6.2** *Consider a general multi-antenna full-duplex relay network with the directed connectivity graph $G = (V, E)$. The traditional AF relaying achieves the maximum multiplexing gain of the network, which is equal to*

$$m_G = \min_{\mathcal{C}} c_G(\mathcal{C}), \tag{6.3}$$

[1]For a graph $G = (V, E)$ and a subset $\mathcal{S} \subseteq V$, the subgraph of $G$ induced by $\mathcal{S}$ is defined as a graph $G_{\mathcal{S}}$ whose underlying vertex set is $\mathcal{S}$ and any two nodes in $G_{\mathcal{S}}$ are connected by an edge if and only if the similar nodes in $G$ are connected by an edge.

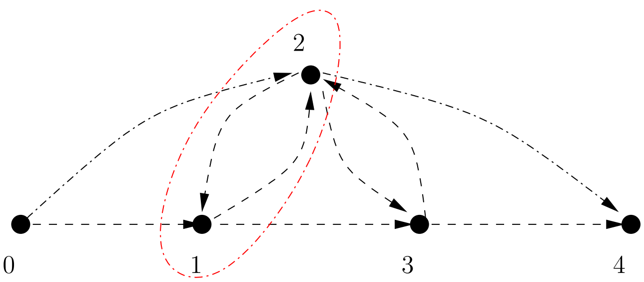


Figure 6.1: A wireless multi-antenna relay network. $N_0 = N_4 = 6, N_1 = 3, N_2 = 2, N_3 = 4$. The minimum vertex cut-set is depicted.

*where $\mathcal{C}$ is a vertex cut-set on $G$.*

**Remark 6.3** *It is worth noting that the maximum multiplexing gain value of every multi-antenna network is computable in polynomial time. Indeed, as it is shown in the proof of Theorem 6.2, the maximum multiplexing gain of the network is equal to the minimum vertex cut-set of the network graph $G$ or equivalently, the minimum cut of the graph $\hat{G}$ defined in the proof of the Theorem. Noting constructing $\hat{G}$ is feasible in polynomial time, its vertex size is linear with $V$ and also the minimum cut is computable in polynomial time from the Ford-Fulkerson Theorem, we conclude that the maximum multiplexing gain of the network is computable in polynomial time.*

Figure 6.1 shows an example of a wireless multi-antenna relay network. In this network, $N_0 = N_4 = 6, N_1 = 3, N_2 = 2, N_3 = 4$. The vertex cut-set which has the minimum capacity is $\mathcal{C} = \{1, 2\}$ and its associated capacity is equal to $c_G(\mathcal{C}) = 5$. Hence, the maximum multiplexing gain of the network is 5, which is achievable by the traditional AF relaying.

The argument of Theorem 6.2 can be easily generalized to the *multicast* and *multi-access* scenarios as well. In the multicast scenario, the source aims to send a common message to multiple destinations. In contrast, in the multi-access scenario, multiple source nodes attempt to send their independent messages to the common destination node.

**Theorem 6.4** *(Multicast Scenario) Consider a general multi-antenna full-duplex relay network with the directed connectivity graph $G = (V, E)$. The source node $s \in V$ aims to send a common message to multiple destinations $t_1, t_2, \ldots, t_M \in V$. The traditional AF relaying achieves the maximum multiplexing gain of the system, which is equal to*

$$m_G^{mc} = \min_{1 \leq i \leq M} m_G(s, t_i), \tag{6.4}$$

*where $m_G(s, t)$ is the minimum vertex cut-set between $s$ and $t$. In other words, $m_G(s, t) \triangleq \min_{\mathcal{C}} c_G(\mathcal{C})$ over all vertex cut-sets $\mathcal{C}$ between $s$ and $t$.*

**Proof** The proof is straightforward. First, it should be noted that the ergodic capacity of the multicast problem is less than or equal to the minimum value of the network ergodic capacities between the source and each of the destination nodes. As a result, $m_G^{mc} \leq \min_{1 \leq i \leq M} m_G(s, t_i)$. On the other hand, in the traditional AF relaying investigated in Theorem 6.2, the relay nodes and the source perform the same operation no matter which node the message is being sent to or what the network connectivity graph is. Hence, the argument of Theorem 6.2 can be applied for the network between $s$ and each $t_i$. Therefore, the traditional AF relaying achieves the multiplexing gain $m_{AF}^{mc} \geq \min_{1 \leq i \leq sM} m_G(s, t_i)$. This proves the argument of the Theorem. ∎

The following Theorem generalizes the argument of Theorem 6.2 to the multi-access scenario.

**Theorem 6.5** *(Multi-Access Scenario) Consider a general multi-antenna full-duplex relay network with the directed connectivity graph $G = (V, E)$. Multiple sender nodes $s_1, s_2, \ldots, s_M \in V$ aim to send independent messages $w_1, w_2, \ldots, w_M$ with the rates $r_1 \log(P), r_2 \log(P), \ldots, r_M \log(P)$ to a common destination node $t \in V$. Let us define the "multiplexing gain region" of the network as the set of all possible $M$-tuples $(r_1, r_2, \ldots, r_M)$ for which the destination can almost surely decode the message of all senders. Then, the traditional AF relaying achieves the optimum multiplexing gain region of the network. Furthermore, the optimum multiplexing gain region of the network is equal to*

$$\mathcal{M}_G^{ma} = \left\{ (r_1, r_2, \ldots, r_M) \middle| \forall \mathcal{S} \subseteq \{1, 2, \ldots, M\}, \ \sum_{m \in \mathcal{S}} r_m \leq m_G(\mathcal{S}, t) \right\}, \qquad (6.5)$$

*where $m_G(\mathcal{S}, t)$ is the minimum vertex cut-set between $\{s_i \,|i \in \mathcal{S}\}$ and $t$. In other words, $m_G(\mathcal{S}, t) \triangleq \min_{\mathcal{C}} c_G(\mathcal{C})$ over all vertex cut-sets $\mathcal{C}$ between $\{s_i \,|i \in \mathcal{S}\}$ and $t$.*

**Proof** First, we prove that the optimum multiplexing gain region of the network is a subregion of $\mathcal{M}_G^{ma}$. Next, we prove that the traditional AF relaying achieves all the points that lie in $\mathcal{M}_G^{ma}$. For any subset $\mathcal{S} \subseteq \{1, 2, \ldots, M\}$, we assume that the sender nodes in $\{s_i \,|i \in \mathcal{S}\}$ are multiple distributed antennas of a super-node $\hat{s}$ and other sender nodes, i.e. $\{s_i \,|i \notin \mathcal{S}\}$, do not interfere on the signals corresponding to $\hat{s}$. Hence, we can apply the argument of Theorem 6.2 for the multiplexing gain of the network between $\hat{s}$ and $t$. Accordingly, for any $M$-tuples that lies in the optimum multiplexing gain region of the network we have $\sum_{m \in \mathcal{S}} r_m \leq m_G(\mathcal{S}, t)$. Now, we prove that the traditional AF relaying achieves all points that lie in the region $\mathcal{M}_G^{ma}$. Let us consider an arbitrary point $(r_1, r_2, \ldots, r_M) \in$

$\mathcal{M}_G^{ma}$. Let us assume the senders are transmitting independent codewords from independent gaussian codebooks of size $P^{r_1}, P^{r_2}, \ldots, P^{r_M}$, respectively. Each relay node amplifies its received signal of the current time-slot and forwards it in the next time-slot. Let us denote the vectors transmitted by $s_1, s_2, \ldots, s_M$ as $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$, respectively, and the vector received by $t$ as $\mathbf{y}$. Going through the same steps as in the proof of Theorem 6.2, one can show that the multiplexing gain region of AF relaying is equal to the multiplexing gain region of a multiple-access channel with the equation

$$\mathbf{y} = \sum_{i=1}^{M} \boldsymbol{\mathcal{H}}_i \mathbf{x}_i + \mathbf{n}, \tag{6.6}$$

where $\boldsymbol{\mathcal{H}}_i$ is a matrix of size $N_t \times N_{s_i}$, corresponding to the end-to-end channel from $s_i$ to $t$, its entries are multivariate polynomials of the channel gains of the network and $\mathbf{n}$ is the white gaussian noise vector of variance 1. The destination performs the jointly typical decoding [12] in order to decide on the transmitted messages. The destination can decode with the error probability approaching 0 iff for any subset $\mathcal{S} \subseteq \{1, 2, \ldots, M\}$, we have

$$\left( \sum_{i \in \mathcal{S}} r_i \right) \log(P) \leq I\left(\mathbf{x}_{\mathcal{S}}; \mathbf{y} \middle| \mathbf{x}_{\mathcal{S}^c}\right), \tag{6.7}$$

where $\mathbf{x}_{\mathcal{S}} \triangleq \{\mathbf{x}_i \,| i \in \mathcal{S}\}$ and $\mathcal{S}^c \triangleq \{1, 2, \ldots M\} - \mathcal{S}$. Furthermore, from (6.6), we have

$$I\left(\mathbf{x}_{\mathcal{S}}; \mathbf{y} \middle| \mathbf{x}_{\mathcal{S}^c}\right) = \mathbb{E}\left\{ \log \left| \mathbf{I}_{N_t} + P \sum_{i \in \mathcal{S}} \boldsymbol{\mathcal{H}}_i \boldsymbol{\mathcal{H}}_i^H \right| \right\}. \tag{6.8}$$

Let us consider the network between the super-node $\hat{s}$ consisting of all nodes $\{s_i | i \in \mathcal{S}\}$ as the sender and $t$ as the destination. Revisiting equations (6.14) and (6.33) for the network between $\hat{s}$ and $t$, we conclude[2]

$$\lim_{P \to \infty} \frac{\mathbb{E}\left\{\log\left|\mathbf{I}_{N_t} + P\sum_{i\in\mathcal{S}} \boldsymbol{\mathcal{H}}_i \boldsymbol{\mathcal{H}}_i^H\right|\right\}}{\log P} = m_G(\mathcal{S}, t). \tag{6.9}$$

Therefore, in the high SNR regime, the constraint in (6.7) is equivalent to the constraint $\sum_{i\in\mathcal{S}} r_i \leq m_G(\mathcal{S}, t)$. However, this constraint is satisfied as the $M$-tuples $(r_1, r_2, \ldots, r_M)$ lies in the region $\mathcal{M}_G^{ma}$. Hence, the destination can decode the transmitted messages with an error probability approaching 0 for any $M$-tuples that lies in $\mathcal{M}_G^{ma}$. This completes the proof. ∎

## 6.2 Proof of Theorem 6.2

First, we prove the argument for the layered graphs. A graph is called layered if all the paths from the source node to the destination node have the same length. Next, we generalize the argument to any directed graph.

The traditional AF relaying scheme can be described as follows. The source node generates a gaussian codebook with codewords of length $TN_0$ where $N_0$ is the number of antennas at the source. In each time-slot, the source node transmits the corresponding $N_0$ symbols of the codeword. Following that, each relay node observes the power of its received signal in every time-slot. If the power of the received signal of the relay is less than or equal to $P\log(P)$, it amplifies the received signal by $\frac{1}{\sqrt{\log(P)}}$ and transmits the amplified signal in the next time-slot. Denoting the path length from the source to the destination by $l_G$, the destination node $K+1$ receives the transmitted symbol of the source node after $l_G - 1$ time-slots. First,

[2]Here, we used the assumption of layered network in the proof of Theorem 6.2. However, the argument is yet valid for the general case.

we find a lower-bound on the probability that all the relay nodes are active. Let us consider a relay node $i$. Defining $\mathcal{D}_i$ as the event that the relay node $i$ is active, $\mathbb{P}\left\{\mathcal{D}_i\right\}$ can be lower-bounded as

$$\begin{aligned}\mathbb{P}\left\{\mathcal{D}_i\right\} &= \mathbb{P}\left\{\mathbb{E}\left\{\|\mathbf{y}_i\|^2\right\} \leq P\log(P)\right\} \\ &\geq \mathbb{P}\left\{P\sum_{(j,i)\in E}\|\mathbf{H}_{i,j}\|^2 + 1 \leq P\log(P)\right\} \qquad (6.10)\end{aligned}$$

Here, $\mathbf{y}_i$ denotes the received vector of size $N_i$ at the node $i$ and $\mathbf{H}_{i,j}$ denotes the channel from node $j$ to node $i$. Let us define $m_i$ as $m_i \triangleq N_i \sum_{(j,i)\in E} N_j$. Noting that $\sum_{(j,i)\in E}\|\mathbf{H}_{i,j}\|^2$ is a Chi-square random variable with $2m_i$ degree of freedom, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{D}_i\right\} &\geq 1 - \sum_{k=0}^{m_i-1}\frac{\left(\log(P)-P^{-1}\right)^k}{k!}e^{P^{-1}-\log(P)} \\ &\geq 1 - c_i\frac{\left(\log(P)\right)^{m_i-1}}{P}, \qquad (6.11)\end{aligned}$$

where $c_i \triangleq e\sum_{k=0}^{m_i-1}\frac{1}{k!}$. In deriving (6.11), it is assumed $P$ is large enough such that $P \geq 1$. Now, defining $\mathcal{D}$ as the event that all the relay nodes of the network are active, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{D}\right\} &= \mathbb{P}\left\{\cap_{i=1}^{K}\mathcal{D}_i\right\} \\ &\overset{(a)}{\geq} \mathbb{P}\left\{\bigcap_{i=1}^{K}\left\{P\sum_{(j,i)\in E}\|\mathbf{H}_{i,j}\|^2 + 1 \leq P\log(P)\right\}\right\} \\ &\overset{(b)}{\geq} 1 - c\frac{\log^d(P)}{P}, \qquad (6.12)\end{aligned}$$

where $c, d \geq 0$ are constants that depend only on the characteristics of the graph $G$. Here, $(a)$ follows from (6.10) and $(b)$ follows from (6.11) and the fact that the events $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_K$ where $\mathcal{A}_i \triangleq \left\{P\sum_{(j,i)\in E}\|\mathbf{H}_{i,j}\|^2 + 1 \leq P\log(P)\right\}$ are independent. From (6.12), we observe that $\mathbb{P}\left\{\mathcal{D}\right\} \sim 1$. Hence, without any loss of generality, we can assume that with probability 1, all the relay nodes are active. In other words, the multiplexing gain of this system is equal to the system in which all the relay nodes are always active and transmit. On the other hand, from the above argument, we know that for all the channels $\mathbf{H}_{i,j}$ with probability 1 we have $\|\mathbf{H}_{i,j}\|^2 \leq \log(P)$. Knowing that for all relay nodes the amplification coefficient is equal to $\frac{1}{\sqrt{\log(P)}}$, we conclude that with probability 1 the power of the equivalent noise at the destination side is less than or equal to a constant that depends only on the topology of the network graph. As a result, the multiplexing gain of the AF relaying is equal to the multiplexing gain of a point-to-point channel whose matrix is equal to the equivalent matrix from the source to the destination. Let us denote the equivalent $N_{K+1} \times N_0$ channel matrix, the source transmitted vector, and the destination received vector by $\boldsymbol{\mathcal{H}}$, $\mathbf{x}$, and $\mathbf{y}$, respectively. Accordingly, the multiplexing gain of the AF relaying is equal to the multiplexing gain of the following channel model

$$\mathbf{y} = \boldsymbol{\mathcal{H}}\mathbf{x} + \mathbf{n} \tag{6.13}$$

where $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_{N_{K+1}}\right)$. In other words, denoting the multiplexing gain of the AF relaying by $m_{AF}$, we have

$$m_{AF} = \lim_{P\to\infty} \frac{\mathbb{E}\left\{\log\left|\mathbf{I}_{N_0} + P\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right|\right\}}{\log(P)}. \tag{6.14}$$

It should be noted that the entries of $\boldsymbol{\mathcal{H}}$ are multivariate polynomials of the entries of $\{\mathbf{H}_{i,j}\}_{(j,i)\in E}$.

Now, let us construct a graph $\hat{G} = (\hat{V}, \hat{E})$ as follows. Corresponding to each relay node $1 \leq i \leq K$ of the original graph $G$, we add $2N_i$ nodes in $\hat{G}$ and denote them by $a_{i,1}, a_{i,2}, \ldots, a_{i,N_i}$ and $b_{i,1}, b_{i,2}, \ldots, b_{i,N_i}$, respectively. Moreover, for every $1 \leq i \leq K, 1 \leq j \leq N_i$, we add an edge from $a_{i,j}$ to $b_{i,j}$. In other words, $(a_{i,j}, b_{i,j}) \in \hat{E}$. Also, corresponding to the source and destination nodes of $G$, we add $N_0 + N_{K+1} + 2$ nodes to $\hat{G}$ and denote them by $b_{0,1}, b_{0,2}, \ldots, b_{0,N_0}$ and $s$ (corresponding to the source node) and $a_{K+1,1}, a_{K+1,2}, \ldots, a_{K+1,N_{K+1}}$ and $t$ (corresponding to the destination node), respectively. $s$ is connected to $b_{0,j}$'s and also $a_{K+1,j'}$'s are connected to $t$. In other words, $(s, b_{0,j}), (a_{K+1,j'}, t) \in \hat{E}$, for $1 \leq j \leq N_0, 1 \leq j' \leq N_{K+1}$. Finally, corresponding to each pair $(i_1, i_2) \in E$ we have $(b_{i_1,j_1}, a_{i_2,j_2}) \in \hat{E}$ for all possible values of $1 \leq j_1 \leq N_{i_1}$ and $1 \leq j_2 \leq N_{i_2}$.

According to the Ford-Fulkerson Theorem [14], there exists a family of $\nu$ edge-disjoint paths $\mathrm{P} \equiv \{\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_\nu\}$ in $\hat{G}$ from $s$ to $t$ where $\nu$ is the min-cut value on $\hat{G}$ from $s$ to $t$. Considering the topology of $\hat{G}$, it is easy to verify that $\mathrm{p}_i$'s are also vertex disjoint. To show this fact, it should be noted that for every node $v, v \neq s, t$, we have either $\delta_I(v) \leq 1$ or $\delta_O(v) \leq 1$ where $\delta_I(v)$ and $\delta_O(v)$ denote the incoming and outgoing degree of $v$.

Let us consider the network channels realization in which, for every pair $(i_1, i_2) \in E$, the $(j_2, j_1)$'th entry of the matrix $\mathbf{H}_{i_2,i_1}$ is equal to 1 if one of the paths in P passes through the edge $(b_{i_1,j_1}, a_{i_2,j_2})$. Otherwise, the corresponding entry is equal to 0. More precisely, we have

$$\mathbf{H}_{i_2,i_1}(j_2, j_1) = \begin{cases} 1 & \exists\, 1 \leq v \leq \nu : (b_{i_1,j_1}, a_{i_2,j_2}) \in \mathrm{p}_v \\ 0 & \text{otherwise} \end{cases} \tag{6.15}$$

For each $1 \leq v \leq \nu$, let us denote the first node after $s$ and the last node before $t$ that the path $\mathrm{p}_v$ passes through by $b_{0,\beta_v}$ and $a_{K+1,\gamma_v}$, respectively. Since the paths are vertex disjoint, we have $\beta_v \neq \beta_{v'}$ and $\gamma_v \neq \gamma_{v'}$ for every $v \neq v'$. Moreover, as the paths are vertex disjoint, the equivalent end-to-end channel matrix corresponding to this channel's realization is equal to

$$\boldsymbol{\mathcal{H}}(i_2, i_1) = \begin{cases} 1 & \exists\, v : i_2 = \gamma_v, i_1 = \beta_v \\ 0 & \text{otherwise} \end{cases} \tag{6.16}$$

From (6.16) and knowing that $\gamma_v$'s and $\beta_v$'s are different for different values of $v$ imply that for this realization of network channels, we have

$$\mathrm{Rank}\left(\boldsymbol{\mathcal{H}}\right) = \nu. \tag{6.17}$$

Having (6.17) and applying Theorem 2.11 of [27], we conclude

$$\lim_{P\to\infty} \frac{\mathbb{E}\left\{\log\left|\mathbf{I}_{N_0} + P\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right|\right\}}{\log(P)} \geq \nu. \tag{6.18}$$

Combining (6.14) and (6.18), we have

$$m_{AF} \geq \nu. \tag{6.19}$$

Now, we prove that $\nu$ is indeed the maximum multiplexing gain of the network. If $\nu = \min(N_0, N_{K+1})$, the argument is valid as the maximum multiplexing gain of the network is less than or equal to the number of antennas at either the source or the destination side. Hence, we only have to prove the argument for the case in which $\nu < \min(N_0, N_{K+1})$.

**Lemma 6.6** *Consider the graph $\hat{G} = (\hat{V}, \hat{E})$. Assume $\nu < \min(N_0, N_{K+1})$ where $\nu$ is the minimum-cut value over $\hat{G}$ from $s$ to $t$. There exists a cut-set $\mathcal{S} \subseteq \hat{V} - \{t\}$ over $\hat{G}$ of minimum weight ($w_{\hat{G}}(\mathcal{S}) = \nu$) and a vertex cutset $\mathcal{C} \subseteq V - \{0, K+1\}$ of minimum capacity over $G$ such that*

$$\hat{E}_{\mathcal{S}} = \bigcup_{v\in\mathcal{C}} \bigcup_{i=1}^{N_v} \left\{(a_{v,i}, b_{v,i})\right\}, \tag{6.20}$$

*where $\hat{E}_{\mathcal{S}}$ denotes the edges that cross the cut-set, i.e.*

$$\hat{E}_{\mathcal{S}} \triangleq \left\{(u,v)\,\middle|(u,v) \in \hat{E}, u \in \mathcal{S}, v \in \mathcal{S}^c\right\}$$

.

**Proof** Let us consider a cut-set $\mathcal{S} \subseteq \hat{V} - \{t\}$ over $\hat{G}$ of minimum-value. For every $v \in \hat{V}$, let us define $\Delta_O(v) \triangleq \left\{(v,u)\,\middle|(v,u) \in \hat{E}\right\}$ and $\Delta_I(v) \triangleq \left\{(u,v)\,\middle|(u,v) \in \hat{E}\right\}$. It is easy to verify that we have $|\Delta_O(a_{i,j})| = |\Delta_I(b_{i,j})| = 1$ for all possible values of $i$ and $j$. Furthermore, for a subset $\mathcal{S} \subseteq \hat{V}$, let us define $\mathcal{A}_{\mathcal{S}}$ and $\mathcal{B}_{\mathcal{S}}$ as $\mathcal{A}_{\mathcal{S}} \triangleq \{a_{i,j}\,|a_{i,j} \in \mathcal{S}\}$ and $\mathcal{B}_{\mathcal{S}} \triangleq \{b_{i,j}\,|b_{i,j} \in \mathcal{S}\}$, respectively. Let us define a new cut-set $\mathcal{T}$ as $\mathcal{T} = \{s\} \cup \mathcal{A}_{\mathcal{T}} \cup \mathcal{B}_{\mathcal{T}}$ where

$$\begin{aligned}
\mathcal{A}_{\mathcal{T}} &\triangleq \mathcal{A}_{\mathcal{S}} \cup \left\{v\,\middle|v \in \mathcal{A}_{\mathcal{S}^c}, \left|\Delta_I(v) \cap \hat{E}_{\mathcal{S}}\right| \geq 1\right\}, \\
\mathcal{B}_{\mathcal{T}} &\triangleq \left\{v\,\middle|v \in \mathcal{B}_{\mathcal{S}}, \left|\Delta_O(v) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| = 0\right\}. \qquad (6.21)
\end{aligned}$$

We prove that $\mathcal{T}$ is also a cut-set of minimum weight. According to the definition of $\mathcal{T}$, we have $\mathcal{A}_{\mathcal{S}} \subseteq \mathcal{A}_{\mathcal{T}}$ and $\mathcal{B}_{\mathcal{T}} \subseteq \mathcal{B}_{\mathcal{S}}$. Now, we have

$$\begin{aligned}
w_{\hat{G}}(\mathcal{T}) - w_{\hat{G}}(\mathcal{S}) &= \left|\hat{E}_{\mathcal{T}}\right| - \left|\hat{E}_{\mathcal{S}}\right| \\
&= \left(\left|\hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| - \left|\hat{E}_{\mathcal{S}}\right|\right) + \left(\left|\hat{E}_{\mathcal{T}}\right| - \left|\hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right|\right) \\
&\overset{(a)}{=} \sum_{v \in \mathcal{A}_{\mathcal{T}} - \mathcal{A}_{\mathcal{S}}} \left(\left|\Delta_O(v) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| - \left|\Delta_I(v) \cap \hat{E}_{\mathcal{S}}\right|\right) + \\
&\quad \sum_{v \in \mathcal{B}_{\mathcal{S}} - \mathcal{B}_{\mathcal{T}}} \left(\left|\Delta_I(v) \cap \hat{E}_{\mathcal{T}}\right| - \left|\Delta_O(v) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right|\right) \\
&\overset{(b)}{\leq} \sum_{v \in \mathcal{A}_{\mathcal{T}} - \mathcal{A}_{\mathcal{S}}} \left(\left|\Delta_O(v) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| - 1\right) + \sum_{v \in \mathcal{B}_{\mathcal{S}} - \mathcal{B}_{\mathcal{T}}} \left(\left|\Delta_I(v) \cap \hat{E}_{\mathcal{T}}\right| - 1\right) \\
&\overset{(c)}{\leq} 0. \qquad (6.22)
\end{aligned}$$

Here, $(a)$ follows from the fact that $\mathcal{A}_{\mathcal{S}} \subseteq \mathcal{A}_{\mathcal{T}}$ and $\mathcal{B}_{\mathcal{T}} \subseteq \mathcal{B}_{\mathcal{S}}$ and using the basic arguments of Graph Theory [14] in counting the number of edges of a directed graph. $(b)$ follows from the fact that for every $v \in \mathcal{A}_{\mathcal{T}} - \mathcal{A}_{\mathcal{S}}$ we have $\left|\Delta_I(v) \cap \hat{E}_{\mathcal{S}}\right| \geq 1$, and also for every $v \in \mathcal{B}_{\mathcal{S}} - \mathcal{B}_{\mathcal{T}}$ we have $\left|\Delta_O(v) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| \geq 1$. Finally, $(c)$ follows from the facts that i) since $\mathcal{A}_{\mathcal{T}} - \mathcal{A}_{\mathcal{S}} \subseteq \mathcal{A}$ for every $v \in \mathcal{A}_{\mathcal{T}} - \mathcal{A}_{\mathcal{S}}$ we have $|\Delta_O(v)| = 1$, and ii) since $\mathcal{B}_{\mathcal{S}} - \mathcal{B}_{\mathcal{T}} \subseteq \mathcal{B}$ for every $v \in \mathcal{B}_{\mathcal{S}} - \mathcal{B}_{\mathcal{T}}$ we have $|\Delta_I(v)| = 1$. (6.22) proves that $\mathcal{T}$ is also a cut-set of minimum weight over $\hat{G}$.

Now, we prove that there exists a subset $\mathcal{C} \subseteq V - \{0, K+1\}$ such that $\hat{E}_{\mathcal{T}} = \bigcup_{v \in \mathcal{C}} \bigcup_{i=1}^{N_v} \{(a_{v,i}, b_{v,i})\}$. In order to prove, we first show that for every possible value of $j$ we have $b_{0,j} \in \mathcal{T}$ and $a_{K+1,j} \in \mathcal{T}^c$. Since $w_{\hat{G}}(\mathcal{T}) = \nu < N_0$, there exists a value of $j$ such that $b_{0,j} \in \mathcal{T}$. According to the definition of $\mathcal{T}$, we conclude that $\left|\Delta_O(b_{0,j}) \cap \hat{E}_{\mathcal{T}}\right| \overset{(a)}{=} \left|\Delta_O(b_{0,j}) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| = 0$ where $(a)$ follows from the fact that there exists no edge in $\hat{G}$ between the nodes in the subset $\mathcal{B}$, i.e. $(\mathcal{B} \times \mathcal{B}) \cap \hat{E} = \emptyset$. Now, let us assume there exists a value $j'$ such that $b_{0,j'} \in \mathcal{T}^c$. Since $s \in \mathcal{T}$, we have $\left|\Delta_I(b_{0,j'}) \cap \hat{E}_{\mathcal{T}}\right| = 1$. Hence, considering the cutset $\hat{\mathcal{T}} = \mathcal{T} \cup \{b_{0,j'}\}$, we have

$$\begin{aligned}
w_{\hat{G}}(\hat{\mathcal{T}}) - w_{\hat{G}}(\mathcal{T}) &= \left|\Delta_O(b_{0,j'}) \cap \hat{E}_{\hat{\mathcal{T}}}\right| - \left|\Delta_I(b_{0,j'}) \cap \hat{E}_{\mathcal{T}}\right| \\
&= \left|\Delta_O(b_{0,j}) \cap \hat{E}_{\hat{\mathcal{T}}}\right| - 1 \\
&= \left|\Delta_O(b_{0,j}) \cap \hat{E}_{\mathcal{T}}\right| - 1 \\
&= -1. \qquad (6.23)
\end{aligned}$$

(6.23) contradicts with the assumption that $\mathcal{T}$ is a cut-set of minimum value. Hence, for all possible values of $j$ we have $b_{0,j} \in \mathcal{T}$. Using the same argument, for all possible values of $j$ we have $a_{K+1,j} \in \mathcal{T}^c$. Hence, the edges that cross the cutset $\mathcal{T}$ are either of type $(a_{i,j}, b_{i,j})$, which we call *inner edges*, or of type $(b_{i_1,j_1}, a_{i_2,j_2})$, which we call *outer edges*. Now, we prove that all edges that cross the cutset $\mathcal{T}$ are inner edges. Let us assume an outer edge $(b_{i_1,j_1}, a_{i_2,j_2}) \in \hat{E}_{\mathcal{T}}$. We have $\left|\Delta_O(b_{i_1,j_1}) \cap \hat{E}_{\mathcal{A}_{\mathcal{T}} \cup \mathcal{S}}\right| = \left|\Delta_O(b_{i_1,j_1}) \cap \hat{E}_{\mathcal{T}}\right| \overset{(a)}{>} 0$ where $(a)$ follows from the fact that $(b_{i_1,j_1}, a_{i_2,j_2}) \in \hat{E}_{\mathcal{T}}$. This inequality contradicts with the assumption that $b_{i_1,j_1} \in \mathcal{B}_{\mathcal{T}}$. Hence, all the edges that cross $\mathcal{T}$ are inner edges.

Finally, we prove the argument of Lemma. Let us define a subset $\mathcal{C} \subseteq V - \{0, K+1\}$ as

$$\mathcal{C} \triangleq \left\{ v \middle| v \in V - \{0, K+1\}, \exists\, i : (a_{v,i}, b_{v,i}) \in \hat{E}_{\mathcal{T}} \right\}. \tag{6.24}$$

We prove that $\hat{E}_{\mathcal{T}} = \bigcup_{v \in \mathcal{C}} \bigcup_{i=1}^{N_v} \{(a_{v,i}, b_{v,i})\}$ ,. First, it should be noted that since all the edges that cross the cutset are inner edges, we have $\hat{E}_{\mathcal{T}} \subseteq \bigcup_{v \in \mathcal{C}} \bigcup_{i=1}^{N_v} \{(a_{v,i}, b_{v,i})\}$. Now, let us assume that $(a_{v,i}, b_{v,i}) \in \hat{E}_{\mathcal{T}}$ for some $v \in \mathcal{C}$. Accordingly, we have $\left|\Delta_I(b_{v,i}) \cap \hat{E}_{\mathcal{T}}\right| = 1$. Since $\mathcal{T}$ is a cutset of minimum weight, we conclude that $\left|\Delta_O(b_{v,i}) \cap \hat{E}_{\mathcal{T} \cup \{b_{v,i}\}}\right| > 0$. Hence, for every $1 \leq i' \leq N_v$ we have

$$\left|\Delta_O(b_{v,i'}) \cap \hat{E}_{\mathcal{T} \cup \left\{b_{v,i'}\right\}}\right| \overset{(a)}{=} \left|\Delta_O(b_{v,i}) \cap \hat{E}_{\mathcal{T} \cup \{b_{v,i}\}}\right| > 0. \tag{6.25}$$

Here, $(a)$ results from the fact that i) $\Delta_O(b_{v,i}) = \Delta_O(b_{v,i'})$ and ii) there exists no edge between the nodes in $\mathcal{B}$. From (6.25) and the definition of $\mathcal{T}$, we conclude that for all $1 \leq i' \leq N_v$, we have $b_{v,i'} \in \mathcal{T}^c$. Using the same argument, we conclude that for all $1 \leq i' \leq N_v$ we have $a_{v,i'} \in \mathcal{T}$. As a result, $\bigcup_{i'=1}^{N_v} \{(a_{v,i'}, b_{v,i'})\} \subseteq E_{\mathcal{T}}$. This proves

$$\hat{E}_{\mathcal{T}} = \bigcup_{v \in \mathcal{C}} \bigcup_{i=1}^{N_v} \{(a_{v,i}, b_{v,i})\}. \tag{6.26}$$

Now, we show that $\mathcal{C}$ is a vertex cutset over $G$. Let us assume $\mathcal{C}$ is not a vertex cut-set. Hence, there exists a path $(0, v_1, v_2, \ldots, v_l, K+1)$ where $v_i \in V - \mathcal{C}$ for all possible $v_i$'s. Accordingly, we construct a path P from $s$ to $t$ in $\hat{G}$ as

$$\mathrm{P} \equiv (s, b_{0,1}, a_{v_1,1}, b_{v_1,1}, a_{v_2,1}, b_{v_2,1}, \ldots, a_{v_l,1}, b_{v_l,1}, a_{K+1,1}, t)\,. \tag{6.27}$$

It is easy to verify that P is a valid path over $\hat{G}$. Furthermore, as for all $v_i$'s we have $v_i \in V - \mathcal{C}$, we conclude $(a_{v_i,1}, b_{v_i,1}) \notin E_{\mathcal{T}}$. Also, since $(b_{v_i,1}, a_{v_{i+1},1})$ is an outer edge, we have $(b_{v_i,1}, a_{v_{i+1},1}) \notin E_{\mathcal{T}}$. Hence, P does not cross $\mathcal{T}$. This contradicts with the assumption that $\mathcal{T}$ is a valid cut-set over $\hat{G}$. As a result, $\mathcal{C}$ is a vertex cut-set over $G$.

Finally, we prove that $\mathcal{C}$ is a minimum vertex cut-set over $G$. Let us consider any arbitrary cut-set $\mathcal{C}' \subseteq V - \{0, K+1\}$ over $G$. Let us consider the subgraph of $G$ induced by $V - \mathcal{C}'$ and denote the set of all vertices to whom the source has a directed path by $\mathcal{Q}$. Clearly, since $\mathcal{C}$ is a vertex cut-set, we have $\{K+1\} \notin \mathcal{Q}$. Now, let us define a cut-set $\mathcal{T}'$ over $\hat{G}$ as

$$\mathcal{T}' \triangleq \{0\} \cup (\cup_{1 \leq i \leq N_0} \{b_{0,i}\}) \cup (\cup_{v \in \mathcal{Q}} \cup_{1 \leq i \leq N_v} \{a_{v,i}, b_{v,i}\}) \cup (\cup_{v \in \mathcal{C}'} \cup_{1 \leq i \leq N_v} \{a_{v,i}\})\,.$$

As there exists no directed path from $s$ to $V - (\mathcal{C}' \cup \mathcal{Q} \cup \{s\})$ in the subgraph of $G$ induced by $V - \mathcal{C}'$, we conclude that there exists no edge from the nodes in $\{0\} \cup \mathcal{Q}$ to the nodes in $V - (\mathcal{C}' \cup \mathcal{Q} \cup \{s\})$. As a result, we have

$$\hat{E}_{\mathcal{T}'} = \bigcup_{v \in \mathcal{C}'} \bigcup_{i=1}^{N_v} \{(a_{v,i}, b_{v,i})\}\,. \tag{6.28}$$

Hence, for any vertex cut-set $\mathcal{C}'$ over $G$ we have $\nu \leq c_G(\mathcal{C}')$. Knowing that there exists a vertex cut-set $\mathcal{C}$ such that $\nu = c_G(\mathcal{C})$, we conclude that $\mathcal{C}$ is the minimum vertex cut-set over $G$. This completes the proof of the Lemma. ∎

Applying Lemma 6.6, for $\nu < \min(N_0, N_{K+1})$ we have $\nu = \min_{\mathcal{C}} c_G(\mathcal{C})$ where $\mathcal{C}$ is a vertex cut-set on $G$. On the other hand, when $\nu = \min(N_0, N_{K+1})$, we have

$\nu \geq \min_{\mathcal{C}} c_G(\mathcal{C})$. Hence, applying (6.19) we have

$$m_{AF} \geq \min_{\mathcal{C}} c_G(\mathcal{C}). \tag{6.29}$$

Finally, we upper-bound the maximum multiplexing gain of the network. Let us denote the maximum multiplexing gain of the network by $m_G$. Let us consider the vertex cut-set $\mathcal{C}$ with minimum capacity on $G$. In the cases where $\mathcal{C} = \{0\}$ or $\mathcal{C} = \{K+1\}$, we have $m_G \leq \min\{N_0, N_{K+1}\} = c_G(\mathcal{C})$. Let us assume the network is operating during $T$ time-slots. Let us denote the vector that the source transmits from time-slot 1 upto $\tau$ and the vector that the source transmits during the time-slot $\tau$ by $\mathbf{x}^{\tau}$ and $\mathbf{x}^{(\tau)}$, respectively. Similarly, $\mathbf{y}^{\tau}$ and $\mathbf{y}^{(\tau)}$ are defined. Furthermore, let us define $\mathbf{x}_{\mathcal{C}}$ and $\mathbf{y}_{\mathcal{C}}$ as the vectors that the nodes in $\mathcal{C}$ transmit and receive, respectively. Since $\mathcal{C}$ has the minimum capacity between the vertex cut-sets, the situation where $\mathcal{C} \neq \{0\}$ and $\mathcal{C} \neq \{K+1\}$ implies $\{0, K+1\} \cap \mathcal{C} = \emptyset$. As $\mathcal{C}$ is a vertex cut-set, $(\mathbf{x}, \mathbf{x}_{\mathcal{C}}, \mathbf{y})$ form a Markov chain. Hence, we have

$$C \overset{(a)}{=} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left\{I\left(\mathbf{x}^T; \mathbf{y}^T\right)\right\} \overset{(b)}{\leq} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left\{I\left(\mathbf{x}^T; \mathbf{x}_{\mathcal{C}}^T\right)\right\} \overset{(c)}{\leq} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left\{I\left(\mathbf{x}^T; \mathbf{y}_{\mathcal{C}}^T\right)\right\}, \tag{6.30}$$

where $C$ is the ergodic capacity of the network and the operator $\mathbb{E}$ is performed over all channels' realizations. Here, $(a)$ follows from the Fano inequality [12], $(b)$ follows from the fact that $(\mathbf{x}, \mathbf{x}_{\mathcal{C}}, \mathbf{y})$ form a Markov chain, and $(c)$ follows from the fact that $(\mathbf{x}, \mathbf{y}_{\mathcal{C}}, \mathbf{x}_{\mathcal{C}})$ form a Markov chain. Now, $\mathbb{E}\left\{I\left(\mathbf{x}^T; \mathbf{y}_{\mathcal{C}}^T\right)\right\}$ can be upper-bounded as

$$
\begin{aligned}
\mathbb{E}\left\{I\left(\mathbf{x}^T;\mathbf{y}_{\mathcal{C}}^T\right)\right\} &\overset{(a)}{\leq} \sum_{v\in\mathcal{C}}\mathbb{E}\left\{h\left(\mathbf{y}_v^T\right)\right\}-\mathbb{E}\left\{h\left(\mathbf{y}_{\mathcal{C}}^T\left|\mathbf{x}^T\right.\right)\right\} \\
&= \sum_{v\in\mathcal{C}}\mathbb{E}\left\{h\left(\mathbf{y}_v^T\right)\right\}-\sum_{\tau=1}^{T}\mathbb{E}\left\{h\left(\mathbf{y}_{\mathcal{C}}^{(\tau)}\left|\mathbf{y}_{\mathcal{C}}^{\tau-1},\mathbf{x}^T\right.\right)\right\} \\
&\overset{(b)}{\leq} \sum_{v\in\mathcal{C}}\mathbb{E}\left\{h\left(\mathbf{y}_v^T\right)\right\}-\sum_{\tau=1}^{T}\mathbb{E}\left\{h\left(\mathbf{n}_{\mathcal{C}}^{(\tau)}\right)\right\} \\
&\overset{(c)}{\leq} \sum_{v\in\mathcal{C}}\sum_{\tau=1}^{T}\mathbb{E}\left\{h\left(\mathbf{y}_v^{(\tau)}\right)-h\left(\mathbf{n}_v^{(\tau)}\right)\right\} \\
&\overset{(d)}{\leq} T\left(\sum_{v\in\mathcal{C}}N_v\right)\log(P)+TO(1) \\
&= Tc_G(\mathcal{C})\log(P)+TO(1).
\end{aligned} \tag{6.31}
$$

Here, $(a)$ follows from the fact that $h\left(\mathbf{y}_{\mathcal{C}}^T\right)\leq\sum_{v\in\mathcal{C}}h\left(\mathbf{y}_v^T\right)$, $(b)$ follows from the fact that $\mathbf{n}_{\mathcal{C}}^{(\tau)}$ is independent from $\left(\mathbf{x}^T,\mathbf{y}_{\mathcal{C}}^{\tau-1},\mathbf{y}_{\mathcal{C}}^{(\tau)}-\mathbf{n}_{\mathcal{C}}^{(\tau)}\right)$ and applying entropy power inequality[3] [12], and $(c)$ follows from the fact that $h\left(\mathbf{n}_{\mathcal{C}}^{(\tau)}\right)=\sum_{v\in\mathcal{C}}h\left(\mathbf{n}_v^{(\tau)}\right)$ and $h\left(\mathbf{y}_v^T\right)\leq\sum_{\tau=1}^{T}h\left(\mathbf{y}_v^{(\tau)}\right)$. In order to prove $(d)$, let us define $v^-$ as the set of vertices from whom there exsits an edge to $v$, i.e. $v^-\triangleq\{u\,|(u,v)\in E\}$. We have $\mathbb{E}\left\{h\left(\mathbf{y}_v^{(\tau)}\right)-h\left(\mathbf{n}_v^{(\tau)}\right)\right\}=\mathbb{E}\left\{I\left(\mathbf{x}_{v^-}^{(\tau)};\mathbf{y}_v^{(\tau)}\right)\right\}$, which is equal to the ergodic capacity of a $c_G(v^-)\times N_v$ MIMO system. As a result $\mathbb{E}\left\{h\left(\mathbf{y}_v^{(\tau)}\right)-h\left(\mathbf{n}_v^{(\tau)}\right)\right\}=\min\left(c_G(v^-),N_v\right)\log(P)+O(1)\leq N_v\log(P)+O(1)$, which results in $(d)$. Combining (6.30) and (6.31), we have

$$
m_G\leq\min_{\mathcal{C}}c_G(\mathcal{C}). \tag{6.32}
$$

[3] According to the entropy power inequality, for any independent random vectors $\mathbf{a}$ and $\mathbf{b}$ of size $n$ we have $2^{\frac{2}{n}h(\mathbf{a}+\mathbf{b})}\geq 2^{\frac{2}{n}h(\mathbf{a})}+2^{\frac{2}{n}h(\mathbf{b})}$. As a result $h(\mathbf{a}+\mathbf{b})\geq h(\mathbf{a})$.

Comparing (6.29) and (6.32), we conclude

$$m_G = m_{AF} = \min_{\mathcal{C}} c_G(\mathcal{C}). \tag{6.33}$$

(6.33) completes the proof of the Theorem for the case of the layered networks.

Now, we prove the argument of the Theorem for the case of any arbitrary networks. First, it should be noted that the inequality series (6.30) and (6.31) are still valid for any arbitrary network. As a result, (6.32) is still valid. Hence, we just need to prove $m_{AF} \geq \min_{\mathcal{C}} c_G(\mathcal{C})$.

In the traditional AF relaying, the network is operated through time-slots $t = 1, 2, \ldots, T$ as follows. The source sends a codeword of length $T$ from its gaussian codebook. Each relay node amplifies its received signal from the last time-slot and forwards it in the next time-slot with the possible amplification coefficient $\frac{1}{\sqrt{\log(P)}}$, similar to what explained for the layered network. The destination decodes the transmitted message using the joint decoding of its received vector from all of its antennas during the time-slots $t = 1, 2, \ldots, T$. Noting the destination has $N_{K+1}$ antennas, its received vector is of size $TN_{K+1}$. Let us denote the transmitted vector at the source and the received vector at the destination by $\mathbf{x} = \mathbf{x}_{t,n_1}$ and $\mathbf{y} = \mathbf{y}_{t,n_2}$, respectively, where $1 \leq t \leq T$, $1 \leq n_1 \leq N_0$ and $1 \leq n_2 \leq N_{K+1}$. Using the same argument we applied for the layered network, we conclude that the multiplexing gain of the AF relaying is equal to the multiplexing gain of a point-to-point MIMO channel whose matrix is of size $TN_{K+1} \times TN_0$ and its entries are multivariate polynomials of the entries of the network channels matrices $\{\mathbf{H}_{i,j}\}_{(j,i)\in E}$. Let us denote this channel matrix by $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}((t_2, n_2), (t_1, n_1))$ where $1 \leq t_1, t_2 \leq T$, $1 \leq n_1 \leq N_0$, and $1 \leq n_2 \leq N_{K+1}$. In other words, we have

$$m_{AF} = \frac{1}{T} \lim_{P\to\infty} \frac{\mathbb{E}\left\{\log\left|\mathbf{I}_{N_0} + P\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right|\right\}}{\log(P)}. \tag{6.34}$$

Here, the expectation is performed over all network channels realization. Now, let us consider the corresponding graph $\hat{G} = (\hat{V}, \hat{E})$, which was previously defined for the unlayered network. It can be shown that the entries of $\boldsymbol{\mathcal{H}}$ are related to the weight of paths in $\hat{G}$ as follows.

$$\begin{aligned} &\boldsymbol{\mathcal{H}}((t_2, n_2), (t_1, n_1)) = \sum_{\mathrm{p}} w(\mathrm{p}), \\ &\text{s.t.} \;\; \mathrm{p}(1) = b_{0,n_1} \;\; \& \;\; \mathrm{p}(l(\mathrm{p}) - 1) = a_{K+1,n_2} \;\; \& \;\; l(\mathrm{p}) = 3 + 2(t_2 - t_1). \end{aligned} \tag{6.35}$$

Here, the summation is over the weight of all paths p of length[4] $3 + 2(t_2 - t_1)$ in $\hat{G}$ from $s$ to $t$ such that $\mathrm{p}(1) = b_{0,n_1}$ and $\mathrm{p}(l(\mathrm{p}) - 1) = a_{K+1,n_2}$. Furthermore, the weight of a path p is defined as

$$\begin{aligned} &w(\mathrm{p}) \triangleq \prod \mathbf{H}_{i_2,i_1}(j_2, j_1), \\ &\text{s.t.} \;\; (b_{i_1,j_1}, a_{i_2,j_2}) \in \hat{E} \;\; \& \;\; \text{p passes through} \;\; (b_{i_1,j_1}, a_{i_2,j_2}). \end{aligned} \tag{6.36}$$

Applying the same argument as for the layered network, there exists a family of $\nu$ vertex-disjoint paths $\mathrm{P} \equiv \{\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_\nu\}$ in $\hat{G}$ from $s$ to $t$ where $\nu$ is the min-cut value on $\hat{G}$ from $s$ to $t$. Now, let us consider the network channels realization in which for every pair $(i_1, i_2) \in E$, the $(j_2, j_1)$'th entry of the matrix $\mathbf{H}_{i_2,i_1}$ is equal to 1 if one of the paths in P passes through the edge $(b_{i_1,j_1}, a_{i_2,j_2})$, and otherwise the corresponding entry is equal to 0. More precisely, we have

$$\mathbf{H}_{i_2,i_1}(j_2, j_1) = \begin{cases} 1 & \exists\, v : \mathrm{p}_v \text{ passes through } (b_{i_1,j_1}, a_{i_2,j_2}) \\ 0 & \text{otherwise} \end{cases} \tag{6.37}$$

From (6.36) and (6.37) and knowing the fact that the paths are vertex disjoint, we

[4]The length of a path p, which is denoted by $l(\mathrm{p})$, is defined as the number of edges that the path goes through.

conclude that for every path p in $\hat{G}$ from $s$ to $t$, we have

$$w(\mathrm{p}) = \begin{cases} 1 & \exists\, 1 \leq v \leq \nu : \mathrm{p} = \mathrm{p}_v \\ 0 & \text{otherwise} \end{cases} \tag{6.38}$$

For each $1 \leq v \leq \nu$, let us denote the first node after $s$ and the last node before $t$ that the path $\mathrm{p}_v$ passes through by $b_{0,\beta_v}$ and $a_{K+1,\gamma_v}$, respectively. Since the paths are vertex disjoint, we have $\beta_v \neq \beta_{v'}$ and $\gamma_v \neq \gamma_{v'}$ for every $v \neq v'$. Applying this fact, (6.35), and (6.38), we conclude that the equivalent end-to-end channel matrix corresponding to this specific realization for the network channels is equal to

$$\boldsymbol{\mathcal{H}}((t_2, n_2), (t_1, n_1)) = \begin{cases} 1 & \exists\, v : n_1 = \beta_v, n_2 = \gamma_v, l(\mathrm{p}_v) = 2(t_2 - t_1) + 3 \\ 0 & \text{otherwise} \end{cases} . \tag{6.39}$$

From (6.39) and knowing that $\gamma_v \neq \gamma_{v'}$ and $\beta_v \neq \beta_{v'}$ for every $v \neq v'$, we have

$$\mathrm{Rank}\,(\boldsymbol{\mathcal{H}}) = \sum_{v=1}^{\nu} \left( T - \frac{l\,(\mathrm{p}_v) - 3}{2} \right) \geq \nu \left( T - \frac{l_{\hat{G}} - 3}{2} \right) = \nu\,(T - l_G + 1)\,, \tag{6.40}$$

where $l_{\hat{G}}$ and $l_G$ denote the maximum length of a simple path[5] connecting the source to the destination in $G$ and $\hat{G}$, respectively. Having (6.40) and applying Theorem 2.11 of [27], we conclude

$$\lim_{P \to \infty} \frac{\mathbb{E}\left\{\log \left| \mathbf{I}_{N_0} + P \boldsymbol{\mathcal{H}}^H \boldsymbol{\mathcal{H}} \right| \right\}}{\log(P)} \geq \nu\,(T - l_G + 1)\,. \tag{6.41}$$

Combining (6.34) and (6.41), we have

$$m_{AF} \geq \nu - \frac{\nu\,(l_G - 1)}{T}. \tag{6.42}$$

[5]A path p in a graph $G = (V, E)$ is called simple, if it passes through each vertex of $V$ just once.

Applying Lemma 6.6, for $\nu < \min(N_0, N_{K+1})$, we have $\nu = \min_{\mathcal{C}} c_G(\mathcal{C})$ where $\mathcal{C}$ is a vertex cut-set on $G$. On the other hand, when $\nu = \min(N_0, N_{K+1})$, we have $\nu \geq \min_{\mathcal{C}} c_G(\mathcal{C})$. Hence, applying (6.42), we have

$$m_{AF} \geq \min_{\mathcal{C}} c_G(\mathcal{C}) - \frac{\upsilon\,(l_G - 1)}{T}, \tag{6.43}$$

where $\mathcal{C}$ is a vertex cut-set on $G$. Having $T \to \infty$ completes the proof of the Theorem.

## 6.3 Conclusion

The general wireless multi-antenna multiple-relay network, introduced earlier in Chapter 3, is investigated in high SNR regime. The pre-log coefficient of the ergodic capacity of the network is studied in the high SNR regime, known as multiplexing gain. It is shown that the "traditional" AF relaying achieves the maximum multiplexing gain of the network. Furthermore, the maximum multiplexing gain of the network is proved to be equal to the minimum vertex cut-set of the underlying graph of the network, which can be computed in polynomial time in terms of the number of network nodes. Finally, the argument is extended to the muticast and multi-access scenarios.

# Chapter 7

# Conclusion and Future Research

This dissertation focuses on Diversity Multiplexing Tradeoff and capacity results in relayed wireless networks.

In Chapter 2, we consider the parallel MIMO relay network. The network studied in this chapter consists of $K$ relays each equipped with $N$ antennas assist in data transmission between a source and a destination, each equipped with $M$ antennas ($N \geq M$). Communication takes place in two equal-time hops and the relays operate in the half-duplex mode. We propose a new AF protocol called "Incremental Cooperative Beamforming Scheme" (ICBS). We prove that the achievable rate of ICBS converges to the capacity of the parallel MIMO relay network, for asymptotically large number of relays, with a gap which vanishes to zero. Next, we study the performance of ICBS in two asymptotically high SNR regmies: i) In the regime where the power of both the source and the relays approaches infinity, we prove that ICBS achieves the full multiplexing gain; and ii) In the regime where the power of the source is fixed, but the power of each relay approaches infinity, we show that the gap between the achievable rate of ICBS and the capacity vanishes to zero. Finally, through simulation, we compare the achievable rate of ICBS against the achievable rate of "matched-filtering" scheme of [19] and the upper-bound on capacity obtained from the point-to-point capacity of the broadcast channel. Simulation results show that while the gap between "matched-filtering" and the upper-bound on capacity remain constant for different number of relays, the achievable rate of ICBS rapidly achieves the upper-bound capacity.

In Chapter 3, we study DMT in single-antenna multiple-relay networks. Here, we propose a new scheme, which we call random sequential (RS), based on the SAF relaying for general multiple-antenna multi-hop networks. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the source of the future paths on the destination of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in amplify-and-forward relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. We derive DMT of the RS scheme for general single-antenna multiple-relay networks (maximum diversity and DMT of RS scheme is investigated in the following two Chapters). Specifically, we derive: 1) the exact DMT of the RS scheme under the condition of "non-interfering relaying", and 2) a lower-bound on the DMT of the RS scheme (no conditions imposed). Finally, we prove that for single-antenna multiple-access multiple-relay networks (with $K > 1$ relays) when there is no direct link between the transmitters and the receiver and all the relays are connected to the transmitter and to the receiver, the RS scheme achieves the optimum DMT. However, for two-hop multiple-access single-relay networks, we show that the proposed scheme is unable to achieve the optimum DMT, while the DDF scheme is shown to perform optimum in this scenario.

In Chapter 4, we investigate the maximum diversity gain for the general multi-hop multi-antenna wireless relay network which is introduced in Chapter 3. We show that the proposed RS scheme achieves the maximum diversity gain of the network. Furthermore, we characterize the maximum achievable diversity gain in terms of the minimum edge cut-set of the underlying graph of the network.

In Chapter 5, we investigate DMT of AF relaying in "multi-antenna" multi-relay networks. For this purpose, we study the application of the RS scheme described in Chapter 3. First, we study the simple structure of multi-antenna full-duplex two-hop single-relay network. We show that unlike the traditional AF relaying, the RS scheme achieves the optimum DMT. Indeed, random unitary matrix multiplication empowers the RS scheme to achieve the optimum DMT. This fact will be elaborated throughout the Chapter 5. Furthermore, we generalize this result to the multi-hop multi-antenna relay networks with single-relay in each hop. Next, we study the case of multi-antenna half-duplex parallel relay network and, by deriving its DMT, we show that the RS scheme improves the DMT of the traditional AF relaying scheme. Interestingly, it turns out that the DMT of the RS scheme is optimum for the multi-antenna half-duplex parallel two-relay ($K = 2$) setup with no direct link between the relays. We also show that utilizing random unitary matrix multiplication improves the DMT of the NAF relaying scheme of [26] in the case of a multi-antenna single relay channel. Finally, we study the class of general full-duplex multi-antenna relay networks whose underlying graph is *directed acyclic* and all nodes are equipped with the same number of antennas. Using the RS scheme, we derive a new lower-bound for the achievable DMT of this class of networks. It turns out that the new DMT lower-bound meets the optimum

DMT at the corner points, corresponding to the maximum multiplexing gain and the maximum diversity gain of the network, respectively.

In Chapter 6, we study the achievable rate of the traditional AF relaying in the high SNR scenarios for general wireless multiple-antenna multiple-relay networks. The channel model for this Chapter is the same as the ones used in Chapters 3, 4 and 5, meaning that every two nodes are either connected through a Rayleigh fading channel or disconnected. Unlike the RS scheme which utilizes matrix multiplication and a complex scheduling for the relays transmission, in traditional AF relaying, each relay node forwards its received signal of the last time-slot in the following time-slot. No channel state knowledge is required at either the source or any of the relay nodes. However, the destination is assumed to know the end-to-end channel state. We study the pre-log coefficient of the ergodic capacity in high SNR regime, known as the multiplexing gain. We prove that the traditional AF relaying achieves the maximum multiplexing gain for any wireless multi-antenna relay network. Furthermore, we characterize the maximum multiplexing gain of the network in terms of the minimum vertex cut-set of the underlying graph of the network and show that it can be computed in polynomial-time (with respect to the number of network nodes) using the maximum-flow algorithm. Finally, we show that the argument can be easily extended to the *multicast* and *multi-access* scenarios as well.

## 7.1 Future Research Directions

The dissertation can be continued in several directions as briefly explained in what follows.

Here, we investigated wireless relay networks in the asymptotic scenarios, either large number of relays or high SNR regime. We derived the asymptotic capacity, multiplexing gain, diversity gain, and DMT in many scenarios using the AF relaying. Another interesting scenario in which AF relaying is potentially useful is very low SNR regime (as an example, see [5, 40]). A good direction for the future research is to consider the wireless relay networks in very low SNR regime and investigate scaling of the ergodic capacity and the outage capacity.

In another research work [3, 4], we considered the $K$ user Gaussian interference channel in high SNR regime and derived its degree of freedom (DOF) for almost all channels coefficients values. Here, we showed that traditional AF relaying achieves the optimum multiplexing gain in wireless multi-antenna relay networks. As an extension, we can study multiplexing gain (or equivalently, DOF) of AF relaying in wireless relay network with multiple disjoint source-destination pairs interfering on each other.

# APPENDICES

# Appendix A

# Proof of Lemma 2.2

Applying the Markov inequality, we have

$$\begin{aligned}
\mathbb{P}\left[\nu > \xi\right] &\leq \frac{\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}^{H}\mathbf{H}_{\mathcal{A}}\|^2\right]}{\xi} \\
&\overset{(a)}{\leq} \frac{\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\|\mathbf{H}_{\mathcal{A}}\|^2\right]}{\xi} \\
&\overset{(b)}{\leq} \frac{\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\|\mathbf{H}\|^2\right]}{\xi}.
\end{aligned} \tag{A.1}$$

Here, $(a)$ is obtained by applying the norm product inequality on matrices [22], and $(b)$ results as $\mathbf{H}_{\mathcal{A}}$ is a submatrix of $\mathbf{H}$. Now, let us define the event $\mathcal{Z} \equiv \left\{\|\mathbf{H}\|^2 < 2s\right\}$ where $s \triangleq MNK$. We can write

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\|\mathbf{H}\|^2\right] &= \mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\|\mathbf{H}\|^2 \middle| \mathcal{Z}\right]\mathbb{P}[\mathcal{Z}] + \mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\|\mathbf{H}\|^2 \middle| \mathcal{Z}^c\right]\mathbb{P}[\mathcal{Z}^c] \\
&\overset{(a)}{\leq} 2s\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\right] + M\mathbb{E}\left[\|\mathbf{H}\|^2 \middle| \mathcal{Z}^c\right]\mathbb{P}[\mathcal{Z}^c],
\end{aligned} \tag{A.2}$$

where $(a)$ follows from the facts that i) conditioned on $\mathcal{Z}$, $\|\mathbf{H}\|^2$ can be upper-bounded by $2s$, and ii) $\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2|\mathcal{Z}\right]\mathbb{P}[\mathcal{Z}] \leq \mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\right] \leq \|\mathbf{U}\|^2 = M$. Hence, it is sufficient to upper-bound the term $\mathbb{E}\left[\|\mathbf{H}\|^2|\, \mathcal{Z}^c\right]\mathbb{P}[\mathcal{Z}^c]$.

Since $\|\mathbf{H}\|^2$ is the sum of $s$ i.i.d. random variables with unit mean and unit variance, we have $\mathbb{E}\left[\|\mathbf{H}\|^2\right] = s$ and $\mathrm{Var}\left[\|\mathbf{H}\|^2\right] = s$. Therefore, using the Chebyshev inequality [22], we have

$$\mathbb{P}[\mathcal{Z}^c] = \mathbb{P}\left[\|\mathbf{H}\|^2 > 2s\right] \leq \frac{1}{s}. \tag{A.3}$$

Now, we can upper-bound $\mathbb{E}\left[\|\mathbf{H}\|^2 \middle| \mathcal{Z}^c\right] \mathbb{P}[\mathcal{Z}^c]$ as follows.

$$\begin{aligned}
\left(\mathbb{E}\left[\|\mathbf{H}\|^2 \middle| \mathcal{Z}^c\right] \mathbb{P}[\mathcal{Z}^c] - s\mathbb{P}[\mathcal{Z}^c]\right)^2 &= \left(\int_{x=2s}^{\infty} (x-s)\, f_{\|\mathbf{H}\|^2}(x)dx\right)^2 \\
&\overset{(a)}{\leq} \int_{x=2s}^{\infty} f_{\|\mathbf{H}\|^2}(x)dx \cdot \int_{x=2s}^{\infty} (x-s)^2 f_{\|\mathbf{H}\|^2}(x)dx \\
&\overset{(b)}{\leq} \frac{1}{s} \cdot s = 1.
\end{aligned} \tag{A.4}$$

Here, $(a)$ follows from the Cauchy-Schwarz inequality and $(b)$ results from (A.3) and the fact that $\mathrm{Var}\left[\|\mathbf{H}\|^2\right] = s$. From inequalities (A.3) and (A.4), we conclude $\mathbb{E}\left[\|\mathbf{H}\|^2 \middle| \mathcal{Z}^c\right] \mathbb{P}[\mathcal{Z}^c] \leq 2$. Combining this fact with (A.1) and (A.2), we have

$$\mathbb{P}\left[\nu > \xi\right] \leq \frac{2MNK\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\right] + 2M}{\xi}. \tag{A.5}$$

To upper-bound $\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\right]$, we have

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{U}_{\mathcal{A}}\|^2\right] &= K\mathbb{E}\left[\|\mathbf{U}_k\|^2 | A_k\right] \mathbb{P}[A_k] \\
&= K\left(\mathbb{E}\left[\|\mathbf{U}_k\|^2 | A_k, B_k\right] \mathbb{P}[A_k, B_k] + \mathbb{E}\left[\|\mathbf{U}_k\|^2 | A_k, B_k^c\right] \mathbb{P}[A_k, B_k^c]\right) \\
&\overset{(a)}{\leq} K\left(M\mathbb{P}[B_k] + \gamma\mathbb{P}[A_k]\right).
\end{aligned} \tag{A.6}$$

Here, $(a)$ follows from the facts that i) the norm of $\mathbf{U}_k$ is upper-bounded by $\|\mathbf{U}\|^2 = M$, and ii) conditioned on the event $B_k^c$, it is upper-bounded by $\gamma$. Combining (A.5) and (A.6) completes the proof.

# Appendix B

# Proof of Lemma 2.3

Let us denote $\mathbf{W}_i$ as the $i$th column of $\mathbf{W}$. In [11], it has been shown that

$$f_{\|\mathbf{W}_i\|^2}(x) = \frac{\Gamma(NK)}{\Gamma(N)\Gamma(NK-N)} x^{N-1}(1-x)^{NK-N-1}, \quad i = 1, \cdots, M, \tag{B.1}$$

which corresponds to the Beta distribution with parameters $N$ and $NK - N$. Therefore, we have

$$\mathbb{P}\left[\|\mathbf{W}\|^2 \geq \gamma\right] = \mathbb{P}\left[\sum_{i=1}^{M} \|\mathbf{W}_i\|^2 \geq \gamma\right] \leq \mathbb{P}\left[\max_i \|\mathbf{W}_i\|^2 \geq \frac{\gamma}{M}\right] \overset{(a)}{\leq} M\mathbb{P}\left[\mathcal{F}_i\right], \tag{B.2}$$

where $(a)$ results from the Union bound on the probability, and $\mathcal{F}_i \equiv \left\{\|\mathbf{W}_i\|^2 \geq \frac{\gamma}{M}\right\}$. Defining $\gamma' \triangleq \frac{\gamma}{M}$, and using (B.1), we obtain

$$\begin{aligned}
\mathbb{P}\left[\|\mathbf{W}\|^2 \geq \gamma\right] &\leq M\left(1 - F_{\|\mathbf{W}_i\|^2}(\gamma')\right) \\
&= M\frac{\Gamma(NK)}{\Gamma(N)\Gamma(NK-N)}\int_{\gamma'}^{1} x^{N-1}(1-x)^{NK-N-1}dx \\
&\stackrel{(a)}{=} M\sum_{n=1}^{N}\frac{(NK-1)!}{(N-n)!(NK-N+n-1)!}\gamma'^{N-n}(1-\gamma')^{NK-N+n-1} \\
&\leq M\sum_{n=1}^{N}\frac{(NK\gamma')^{N-n}(1-\gamma')^{NK-N}}{(N-n)!} \\
&= \frac{M(NK\gamma')^{N-1}(1-\gamma')^{NK-N}}{(N-1)!}\times \\
&\quad \left[1+\sum_{n=1}^{N-1}\frac{(N-1)(N-2)\ldots(N-n)}{(K\gamma')^{n}}\right] \\
&\stackrel{(b)}{\leq} \frac{N^N}{M^{N-2}(N-1)!}(K\gamma)^{N-1}\left(1-\frac{\gamma}{M}\right)^{N(K-1)},
\end{aligned} \tag{B.3}$$

where $(a)$ follows from the integration by part, and $(b)$ follows from the fact that $\gamma \geq \frac{MN}{K}$.

# Appendix C

# Proof of Lemma 2.4

Applying Lemma 2.2, we have

$$\mathbb{P}\left[\nu > \xi\right] \leq \frac{2MNK^2}{\xi}\left(M\mathbb{P}[B_k] + \gamma\mathbb{P}[A_k] + \frac{1}{NK^2}\right). \tag{C.1}$$

Now, we upper-bound $\mathbb{P}[A_k]$ as follows. The event $A_k$ occurs whenever the event $\mathbb{E}_{\mathbf{x},\mathbf{n}_k}\left[\left\|\mathbf{G}_k^{\dagger}\mathbf{U}_k^H\mathbf{r}_k\right\|^2\right] > \beta$ occurs. On the other hand, conditioned on occuring $B_k^c$, we have

$$\mathbb{E}_{\mathbf{x},\mathbf{n}_k}\left[\left\|\mathbf{G}_k^{\dagger}\mathbf{U}_k^H\mathbf{r}_k\right\|^2\right] \leq \gamma\lambda_{\min}^{-1}(\mathbf{G}_k^H\mathbf{G}_k)\left(1 + P_s\|\mathbf{H}_k\|^2\right), \tag{C.2}$$

where (C.2) follows from the product norm inequality of matrices. Hence, conditioned on occuring $A_k$ and $B_k^c$, we have $\lambda_{\min}(\mathbf{G}_k^H\mathbf{G}_k) < \frac{\gamma}{\beta}(1 + P_s\|\mathbf{H}_k\|^2)$. Let us define $\delta \triangleq \frac{\gamma}{\beta}$ and $D_k$ as the event in which $\lambda_{\min}(\mathbf{G}_k^H\mathbf{G}_k) < \delta(1 + P_s\|\mathbf{H}_k\|^2)$. Accordingly, we have $\mathbb{P}[A_k, B_k^c] \leq \mathbb{P}[D_k]$. Hence, we can upper-bound $\mathbb{P}[A_k]$ as follows.

$$\mathbb{P}[A_k] = \mathbb{P}[A_k, B_k] + \mathbb{P}[A_k, B_k^c] \leq \mathbb{P}[B_k] + \mathbb{P}[D_k] \tag{C.3}$$

Moreover, $\mathbb{P}[D_k]$ can be upper-bounded as follows.

$$
\begin{aligned}
\mathbb{P}[D_k] &\le \mathbb{P}\left[\left(\lambda_{\min}(\mathbf{G}_k^H\mathbf{G}_k) \le 2P_s\delta\log\left(\frac{1}{\delta}\right)+\delta\right)\bigcup\right.\\
&\quad \left.\left(1+P_s\|\mathbf{H}_k\|^2 \ge 2P_s\log\left(\frac{1}{\delta}\right)+1\right)\right]\\
&\le \mathbb{P}\left[\lambda_{\min}(\mathbf{G}_k^H\mathbf{G}_k) \le 2P_s\delta\log\left(\frac{1}{\delta}\right)+\delta\right]+\mathbb{P}\left[\|\mathbf{H}_k\|^2 \ge 2\log\left(\frac{1}{\delta}\right)\right]\\
&\overset{(a)}{=} \int_{x=0}^{2P_s\delta\log\left(\frac{1}{\delta}\right)+\delta} Me^{-Mx}dx+\frac{1}{\Gamma(MN)}\int_{x=2\log\left(\frac{1}{\delta}\right)}^{\infty} x^{MN-1}e^{-x}dx\\
&\le 2MP_s\delta\log\left(\frac{1}{\delta}\right)+M\delta+\left[\sum_{m=0}^{MN-1}\frac{x^m e^{-x}}{m!}\right]_{x=2\log\left(\frac{1}{\delta}\right)}\\
&\overset{(b)}{\le} \delta\left(2MP_s\log\left(\frac{1}{\delta}\right)+8\delta\log^{MN-1}\left(\frac{1}{\delta}\right)+M\right).
\end{aligned}
\tag{C.4}
$$

Here, $(a)$ results from the fact that the minimum eigenvalue of $\mathbf{G}_k^H\mathbf{G}_k$ can be lower-bounded by the minimum eigenvalue of $\mathbf{W}_k^H\mathbf{W}_k$ where $\mathbf{W}_k$ is any arbitrary $M \times M$ submatrix of $\mathbf{G}_k$ [22] and also, from the probability density function of the minimum singular value of a square i.i.d. complex Gaussian matrix, derived in [16]. $(b)$ follows from $\delta \le e^{-1}$ and the fact that $\sum_{m=0}^{\infty}\frac{2^m}{m!} < 8$.

Combining (C.1), (C.3), and (C.4), we have

$$
\begin{aligned}
&\mathbb{P}\left[\nu > \xi\right] \le \frac{2MNK^2}{\xi}\times\\
&\left(2M\mathbb{P}[B_k]+\frac{\gamma^2}{\beta}\left(2MP_s\log\left(\frac{\beta}{\gamma}\right)+8\frac{\gamma}{\beta}\log^{MN-1}\left(\frac{\beta}{\gamma}\right)+M\right)+\frac{1}{NK^2}\right)
\end{aligned}
\tag{C.5}
$$

This proves the first part of Lemma. Now, assuming $\xi = \log^2(K)\sqrt{K}$, $\beta = \frac{\log(K)}{\sqrt[4]{K}}$, and $\gamma = \frac{4\log(K)}{K}$, we know $\gamma \le \frac{\beta}{e}$ for large enough values of $K$. Hence, we can apply (C.5). We have

$$
\mathbb{P}\left[\nu > \xi\right] \le \frac{2MNK\sqrt{K}}{\log^2(K)}\left(2M\mathbb{P}[B_k]+O\left(\log^2(K)K^{-\frac{7}{4}}\right)\right) \tag{C.6}
$$

Now, since we have $\gamma \leq \frac{MN}{K}$ for large enough number of relays, we can apply the argument of Lemma 2.3 to upper-bound $\mathbb{P}[B_k]$. Applying Lemma 2.3 and noting $(1-\frac{\gamma}{M})^{N(K-1)} \leq e^{-\frac{\gamma}{M}N(K-1)}$, we have

$$\mathbb{P}[B_k] = O\left(\frac{(\log(K))^{N-1}}{K^4}\right) \tag{C.7}$$

Combining (C.6) and (C.7) completes the proof of the second part of Lemma.

# Appendix D

# Proof of Lemma 2.5

The $(i, j)$th entry of $\mathbf{AA}^H$, denoted as $[\mathbf{AA}^H]_{i,j}$, can be written as

$$[\mathbf{AA}^H]_{i,j} = \mathbf{a}_i \mathbf{a}_j^H, \tag{D.1}$$

where $\mathbf{a}_i$ is the vector representing the $i$th row of $\mathbf{A}$. Let us define $\mathbf{B}$ as

$$\mathbf{B} \triangleq [\mathbf{b}_1^T | \cdots | \mathbf{b}_r^T]^T, \tag{D.2}$$

where $\mathbf{b}_i = \dfrac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$, $i = 1, \cdots, r$. We have

$$[\mathbf{BB}^H]_{i,j} = \left[ \begin{array}{ll} 1 & i = j \\ \gamma(i,j) & i \neq j \end{array} \right., \tag{D.3}$$

where $\gamma(i, j) \triangleq \mathbf{b}_i \mathbf{b}_j^H = \dfrac{\mathbf{a}_i \mathbf{a}_j^H}{\|\mathbf{a}_i\| \|\mathbf{a}_j\|}$. The pdf of $z(i, j) = |\gamma(i, j)|^2$ has been computed in [11], Lemma 3, as

$$p_{z(i,j)}(z) = (s - 1)(1 - z)^{s-2}. \tag{D.4}$$

Let us define $\mathcal{C}$ as the event that $z(i,j) < \frac{1}{\sqrt{s}}$ for all $i \neq j$. Using (D.4), we have

$$\begin{aligned}\mathbb{P}[\mathcal{C}] &= \mathbb{P}\left[\bigcap_{i \neq j}\left(z(i,j) < \frac{1}{\sqrt{s}}\right)\right] \\ &\overset{(a)}{\geq} 1 - \frac{r(r-1)}{2}\left(1 - \frac{1}{\sqrt{s}}\right)^{s-1} \\ &= 1 + O(e^{-\sqrt{s}}), \end{aligned} \tag{D.5}$$

where $(a)$ results from the Union bound on the probability, noting that $z(i,j) = z(j,i), \forall i,j$. Conditioned on $\mathcal{C}$, the orthogonality defect of $\mathbf{B}$, defined as $\frac{\prod_{i=1}^{r} \|\mathbf{b}_i\|^2}{|\mathbf{B}\mathbf{B}^H|}$, can be written as

$$\begin{aligned}\delta_{\mathcal{C}}(\mathbf{B}) &= \frac{1}{|\mathbf{B}\mathbf{B}^H|} \\ &\leq \frac{1}{1 - \frac{r(r-1)}{2\sqrt{s}} - \frac{r! - \frac{r(r-1)}{2} - 1}{s}} \\ &\overset{(a)}{\leq} \frac{1}{1 - \frac{r^2}{2\sqrt{s}}}, \end{aligned} \tag{D.6}$$

where $\delta_{\mathcal{C}}(\mathbf{B})$ denotes the orthogonality defect of $\mathbf{B}$, conditioned on $\mathcal{C}$, and $(a)$ follows from the assumption that $s$ is large enough such that $\frac{2\left(r! - \frac{r(r-1)}{2} - 1\right)}{r} < \sqrt{s}$ which results in $\frac{r! - \frac{r(r-1)}{2} - 1}{s} < \frac{r}{2\sqrt{s}}$. Hence, using the fact that the orthogonality defect of $\mathbf{A}$ and $\mathbf{B}$ are equal, conditioned on $\mathcal{C}$ we can write

$$\begin{aligned}\prod_{i=1}^{r} \lambda_i &= |\mathbf{A}\mathbf{A}^H| \\ &= \prod_{i=1}^{r} \|\mathbf{a}_i\|^2 \left[1 - \frac{r^2}{2\sqrt{s}}\right], \end{aligned} \tag{D.7}$$

where $\lambda_i$'s denote the eigenvalues of $\mathbf{A}\mathbf{A}^H$. Moreover,

$$\begin{aligned}\sum_{i=1}^{r}\lambda_i &= \mathrm{Tr}\{\mathbf{A}\mathbf{A}^H\}\\ &= \sum_{i=1}^{r}\|\mathbf{a}_i\|^2. \end{aligned} \tag{D.8}$$

Now, let us define events $\mathcal{D}_i$ as follows:

$$\mathcal{D}_i \equiv \left\{s(1-\epsilon) < \|\mathbf{a}_i\|^2 < s(1+\epsilon)\right\}, \quad i = 1, \cdots, r, \tag{D.9}$$

where $\epsilon \triangleq \sqrt{\frac{2\log(s)}{s}}$. Since $\|\mathbf{a}_i\| = \sum_{j=1}^{s}|a_{i,j}|^2$, where $a_{i,j}$ denotes the $(i,j)$th entry of $\mathbf{A}$, and having the fact that $|a_{i,j}|^2$ are i.i.d. random variables with unit mean and unit variance, using Central Limit Theorem (CLT), $\frac{1}{s}\|\mathbf{a}_i\|^2$ approaches, in probability, to a Gaussian distribution with unit mean and variance $\frac{1}{s}$, as $s$ tends to infinity. More precisely, defining $X \triangleq \frac{\frac{1}{s}\|\mathbf{a}_i\|^2}{\sqrt{\frac{1}{s}}}$ and using Theorem 5.24 in [37], we can write $q \triangleq \mathbb{P}\left[-\sqrt{2\log(s)} < X - \sqrt{s} < \sqrt{2\log(s)}\right]$ as follows:

$$\begin{aligned} q =& 1 - \left[1 - \Phi\left(\sqrt{2\log(s)}\right)\right]\exp\left\{\frac{\gamma_3\sqrt{2}\sqrt{\log^3(s)}}{3\sigma^3\sqrt{s}}\right\} - \\ & \Phi\left(-\sqrt{2\log(s)}\right)\exp\left\{-\frac{\gamma_3\sqrt{2}\sqrt{\log^3(s)}}{3\sigma^3\sqrt{s}}\right\} + O\left(s^{-1/2}e^{-\log(s)}\right) \\ \overset{(a)}{=}& 1 - \frac{1}{2\sqrt{\pi\log(s)}}e^{-\log(s)}\left[1 + \rho\sqrt{\frac{\log^3(s)}{s}} + O\left(\frac{\log^3(s)}{s}\right)\right] - \\ & \frac{1}{2\sqrt{\pi\log(s)}}e^{-\log(s)}\left[1 - \rho\sqrt{\frac{\log^3(s)}{s}} + O\left(\frac{\log^3(s)}{s}\right)\right] + O\left(\frac{1}{s\sqrt{s}}\right) \\ =& 1 - \frac{1}{s\sqrt{\pi\log(s)}}\left[1 + O\left(\frac{\log^3(s)}{s}\right)\right] + O\left(\frac{1}{s\sqrt{s}}\right), \end{aligned} \tag{D.10}$$

where $\Phi(.)$ denotes the CDF of the normal distribution, $\sigma^2$ and $\gamma_3$ denote the second and third moments of $|a_{i,j}|^2$, respectively, and $\rho \triangleq \frac{\gamma_3\sqrt{2}}{3\sigma^3}$. $(a)$ follows from i) the approximation of $\Phi(x)$ for large $x$ by $1 - \frac{1}{\sqrt{2\pi x}}e^{-\frac{x^2}{2}}$, ii) $e^x = 1 + x + O(x^2)$, for $x \to 0$, and iii) the fact that $\sigma$ and $\gamma_3$ are constants which incurs that $\rho$ is constant. From the above equation, $\mathbb{P}[\mathcal{D}_i]$ can be computed as

$$\begin{aligned}
\mathbb{P}[\mathcal{D}_i] &= \mathbb{P}\left[1 - \epsilon < \frac{1}{s}\|\mathbf{a}_i\|^2 < 1 + \epsilon\right] \\
&= \mathbb{P}\left[\sqrt{s} - \sqrt{2\log(s)} < \frac{1}{\sqrt{s}}\|\mathbf{a}_i\|^2 < \sqrt{s} + \sqrt{2\log(s)}\right] \\
&= \mathbb{P}\left[-\sqrt{2\log(s)} < X - \sqrt{s} < \sqrt{2\log(s)}\right] \\
&= 1 + O\left(\frac{1}{s\sqrt{\log(s)}}\right),
\end{aligned} \tag{D.11}$$

in which we have used the definitions of $\epsilon$ and $X$, which are $\sqrt{\frac{2\log(s)}{s}}$ and $\frac{1}{\sqrt{s}}\|\mathbf{a}_i\|^2$, respectively. Conditioned on $\mathcal{C}$ and $\mathcal{D}$, where $\mathcal{D} \triangleq \bigcap_{i=1}^r \mathcal{D}_i$, and using (D.7) and (D.8), we can write

$$\begin{aligned}
\eta &\triangleq \frac{\prod_{i=1}^r \lambda_i}{\overline{\lambda}^r} \\
&\geq \frac{\prod_{i=1}^r [s(1-\epsilon)]\left[1 - \frac{r^2}{2\sqrt{s}}\right]}{\left[\frac{1}{r}\sum_{i=1}^r s(1+\epsilon)\right]^r} \\
&\geq (1-2\epsilon)^r\left[1 - \frac{r^2}{2\sqrt{s}}\right] \\
&\overset{(a)}{\geq} 1 - 3r\epsilon,
\end{aligned} \tag{D.12}$$

where $\overline{\lambda} \triangleq \frac{1}{r}\sum_{i=1}^r \lambda_i$ and $(a)$ follows form i) $(1-2\epsilon)^r \geq 1 - 2r\epsilon$ and ii) $s$ is large enough such that $\frac{r^2}{2\sqrt{s}} < r\epsilon$. Suppose that $\lambda_{\min}(\mathbf{A}\mathbf{A}^H) = \alpha\overline{\lambda}$ ($\alpha < 1$) and

$\lambda_{\max}(\mathbf{A}\mathbf{A}^H) = \beta\overline{\lambda}$ $(\beta > 1)$. We have

$$\begin{aligned}
\eta &\overset{(a)}{\leq} \frac{\alpha\overline{\lambda}\left[\frac{1}{r-1}(r\overline{\lambda} - \alpha\overline{\lambda})\right]^{r-1}}{\overline{\lambda}^r} \\
&= \frac{\alpha(r-\alpha)^{r-1}}{(r-1)^{r-1}},
\end{aligned} \tag{D.13}$$

where $(a)$ follows from the fact that knowing $\lambda_{\min}(\mathbf{A}\mathbf{A}^H)$, the product of the rest of the singular values is maximized when they are all equal. Hence, having the sum constraint of $r\overline{\lambda}$ yields $\prod_{i=1}^r \lambda_i < \alpha\overline{\lambda}\left[\frac{1}{r-1}(r\overline{\lambda} - \alpha\overline{\lambda})\right]^{r-1}$. Using (D.12) and (D.13), and noting that $f(\alpha) \triangleq 1 - \frac{\alpha(r-\alpha)^{r-1}}{(r-1)^{r-1}}$ can be lower-bounded in the interval $[0, 1]$ by $\frac{r(1-\alpha)^2}{2(r-1)}$, we have

$$\begin{aligned}
1 - \frac{\alpha(r-\alpha)^{r-1}}{(r-1)^{r-1}} &\leq 3r\epsilon \\
\Rightarrow \frac{r(1-\alpha)^2}{2(r-1)} &\leq 3r\epsilon \\
\Rightarrow \alpha &\geq 1 - \sqrt{6(r-1)\epsilon}.
\end{aligned} \tag{D.14}$$

Similarly, we have

$$\beta \leq 1 + 3\sqrt{r\epsilon} \tag{D.15}$$

Moreover, conditioned on $\mathcal{D}$, we have $s(1-\epsilon) \leq \overline{\lambda} \leq s(1+\epsilon)$. Consequently, conditioned on $\mathcal{C}$ and $\mathcal{D}$ we have

$$\begin{aligned}
s\left(1 + 4\sqrt{r\epsilon}\right) \overset{(a)}{\geq} s(1+\epsilon)\left(1 + 3\sqrt{r\epsilon}\right) &\geq \lambda_{\max}(\mathbf{A}\mathbf{A}^H), \\
s\left(1 - \sqrt{6r\epsilon}\right) \overset{(a)}{\leq} s(1-\epsilon)\left(1 - \sqrt{6(r-1)\epsilon}\right) &\leq \lambda_{\min}(\mathbf{A}\mathbf{A}^H),
\end{aligned} \tag{D.16}$$

where $(a)$ follows assuming $\epsilon$ is small enough. As a result,

$$\begin{aligned}
&\mathbb{P}\left[s\left(1+4\sqrt{r\epsilon}\right) \geq \lambda_{\max}(\mathbf{A}\mathbf{A}^H) \geq \lambda_{\min}(\mathbf{A}\mathbf{A}^H) \geq s\left(1-\sqrt{6r\epsilon}\right)\right] \\
&\geq \mathbb{P}[\mathcal{C} \cap \mathcal{D}] \\
&\overset{(a)}{=} \mathbb{P}[\mathcal{C}]\mathbb{P}[\mathcal{D}] \\
&\overset{(b)}{=} \mathbb{P}[\mathcal{C}]\left(\mathbb{P}[\mathcal{D}_i]\right)^r \\
&\overset{(\mathrm{D.5}),(\mathrm{D.11})}{=} \left[1+O(e^{-\sqrt{s}})\right]\left[1+O\left(\frac{1}{s\sqrt{\log(s)}}\right)\right]^r \\
&= 1+O\Big(\frac{1}{s\sqrt{\log(s)}}\Big),
\end{aligned} \tag{D.17}$$

where $(a)$ follows from the fact that the norm and direction of a Gaussian vector are independent of each other [57], i.e., $\{\|\mathbf{a}_i\|\}_{i=1}^M$ and $\{\mathbf{b}_j\}_{j=1}^M$ are independent of each other, and as a result, the event $\mathcal{C}$ which is defined on $\{\mathbf{b}_j\}_{j=1}^M$ and $\mathcal{D}$ which is defined on $\{\|\mathbf{a}_i\|\}_{i=1}^M$ are independent of each other. $(b)$ follows from the fact that $\mathcal{D}_i$'s are independent and have the same probability.

# Appendix E

# Proof of Theorem 3.2

Since the relay nodes are non-interfering, the achievable rate of the RS scheme for a realization of the channels is equal to

$$R_{RS,NI}\left(\{h_e\}_{e\in E}\right) = \frac{1}{S}\sum_{i=1}^{L}$$
$$\log\left(1 + P\prod_{j=1}^{l_i}|\alpha_{i,j}|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2\left(1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1}\right), \tag{E.1}$$

where $\forall j < l_i : \alpha_{i,j} = \sqrt{\frac{P}{1+\left|h_{\{\mathrm{p}_i(j-1),\mathrm{p}_i(j)\}}\right|^2 P}}$ and $\alpha_{i,l_i} = 1$ (since $\mathrm{p}_i(l_i) = K+1$). In deriving the above equation, we have used the fact that as the paths are non-interfering, the achievable rate can be written as the sum of the rates over the paths. Note that $P\prod_{j=1}^{l_i}|\alpha_{i,j}|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2$ and $1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}|\alpha_{i,k}|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2$ represent the effective signal power and the noise power over the $i$th path, respectively. Hence, the probability of outage equals

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{E}\right\} &= \mathbb{P}\left\{R_{RS,NI}\left(\left\{h_e\right\}_{e\in E}\right) \leq r\log\left(P\right)\right\} \\
&\overset{(a)}{\doteq} \mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1}, \min\left\{\left|h_{\{0,\mathrm{p}_i(1)\}}\right|^2\prod_{k=1}^{j}\left|\alpha_{i,k}\right|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right\}_{j=0}^{l_i-1}\right\}\right. \\
&\qquad \left.\leq P^{Sr-L}\right\} \\
&\overset{(b)}{\doteq} \max_{\substack{t_1,t_2,\ldots,t_L\\ 1\leq t_i\leq l_i}}\mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1}, \left|h_{\{0,\mathrm{p}_i(1)\}}\right|^2\prod_{k=1}^{t_i-1}\left|\alpha_{i,k}\right|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right\}\right. \\
&\qquad \left.\leq P^{Sr-L}\right\} \\
&\overset{(c)}{\doteq} \max_{\substack{\mathcal{S}_1,\mathcal{S}_2,\ldots,\mathcal{S}_L\\ \mathcal{S}_i\subseteq\{1,2,\ldots,l_i-1\}}}\ \max_{\substack{t_1,t_2,\ldots,t_L\\ \max\{x\in\mathcal{S}_i\}<t_i\leq l_i}} \\
&\qquad \mathbb{P}\left\{\prod_{i=1}^{L}\max\left\{P^{-1}, P^{|\mathcal{S}_i|}\left|h_{\{\mathrm{p}_i(t_i),\mathrm{p}_i(t_i-1)\}}\right|^2\prod_{k\in\mathcal{S}_i}\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}}\right|^2\right\}\right. \\
&\qquad \left.\leq P^{Sr-L}\right\}. \qquad\qquad \text{(E.2)}
\end{aligned}$$

Here, $(a)$ follows from the facts that i) $\forall x\geq 0: \max\{1,x\}\leq 1+x\leq 2\max\{1,x\}$, which implies that $1+P\Theta\approx\max(1,P\Theta)$, where

$$\Theta \triangleq \prod_{j=1}^{l_i}\left|\alpha_{i,j}\right|^2\left|h_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\right|^2\left(1+\sum_{j=1}^{l_i-1}\prod_{k=j}^{l_i-1}\left|\alpha_{i,k}\right|^2\left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2\right)^{-1},$$

and ii) for all $x_i \geq 0$, $\frac{1}{M} \min \left\{ \frac{1}{x_i} \right\}_{i=1}^{M} \leq \left( \sum_{i=1}^{M} x_i \right)^{-1} \leq \min \left\{ \frac{1}{x_i} \right\}_{i=1}^{M}$, which implies that

$$\begin{aligned} &\left(1 + \sum_{j=1}^{l_i-1} \prod_{k=j}^{l_i-1} \left|\alpha_{i,k}\right|^2 \left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2 \right)^{-1} \approx \\ &\min \left( 1, \left\{ \left( \prod_{k=j}^{l_i-1} \left|\alpha_{i,k}\right|^2 \left|h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k+1)\}}\right|^2 \right)^{-1} \right\}_{j=1}^{l_i-1} \right). \end{aligned}$$

$(b)$ follows from the fact that for any increasing function $f(.)$ , we have

$$\max_{1 \leq i \leq M} \mathbb{P}\left\{ f\left(x_i\right) \leq y \right\} \leq \mathbb{P}\left\{ f\left( \min_{1 \leq i \leq M} x_i \right) \leq y \right\} \leq M \max_{1 \leq i \leq M} \mathbb{P}\left\{ f\left(x_i\right) \leq y \right\}.$$

$(c)$ follows from the fact that

$$\begin{aligned} &0.5 \min \left\{ 1, P \left| h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \right\} \leq \left| \alpha_{i,k} h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \leq \\ &\min \left\{ 1, P \left| h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \right\}, \end{aligned}$$

which implies that $\left| \alpha_{i,k} h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \leq \min \left\{ 1, P \left| h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \right\}$. In the last line of (E.2), $\mathcal{S}_i$ denotes the subset of $\{1, 2, \cdots, t_i - 1\}$ for which $P \left| h_{\{\mathrm{p}_i(k),\mathrm{p}_i(k-1)\}} \right|^2 \leq 1$.

Assuming $\left|h_e\right|^2 = P^{-\mu_e}$, we define the region $\mathcal{R} \subseteq \mathbb{R}^{|E|}$ as the set of points $\boldsymbol{\mu} = [\mu_e]_{e \in E}$ that the outage event occurs. Let us define $\mathcal{R}_+ = \mathcal{R} \cap \left(\mathbb{R}_+ \cup \{0\}\right)^{|E|}$. As the probability density function diminishes exponentially as $e^{-P^{\mu_e}}$ for positive values of $\mu_e$, we have $\mathbb{P}\{\mathcal{R}_+\} \doteq \mathbb{P}\{\mathcal{R}\}$. Hence, we have

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} &\doteq \mathbb{P}\{\mathcal{R}_+\} \\ &\overset{(a)}{\doteq} \max_{\substack{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_L \\ \mathcal{S}_i \subseteq \{1,2,\ldots,l_i-1\}}} \max_{\substack{t_1, t_2, \ldots, t_L \\ \max\{x \in \mathcal{S}_i\} < t_i \leq l_i}} \mathbb{P}\left\{\mathcal{R}\left(\boldsymbol{\mathcal{S}}, \mathbf{t}\right)\right\} \\ &\overset{(b)}{\doteq} \max_{\substack{\mathbf{t} \\ 1 \leq t_i \leq l_i}} \mathbb{P}\left\{\mathcal{R}_0\left(\mathbf{t}\right)\right\}, \end{aligned} \tag{E.3}$$

where

$$\mathcal{R}\left(\boldsymbol{\mathcal{S}}, \mathbf{t}\right) \equiv \Bigg\{ \boldsymbol{\mu} \in (\mathbb{R}_+ \cup \{0\})^{|E|} \,\Bigg|\, L - Sr \leq \sum_{i=1}^{L} \min \left\{ 1, \mu_{\{\mathrm{p}_i(t_i), \mathrm{p}_i(t_i - 1)\}} + \sum_{k \in \mathcal{S}_i} \mu_{\{\mathrm{p}_i(k), \mathrm{p}_i(k-1)\}} - |\mathcal{S}_i| \right\} \Bigg\},$$

$\mathbf{t} = [t_1, t_2, \ldots, t_L]$, $\boldsymbol{\mathcal{S}} = [\mathcal{S}_1, \cdots, \mathcal{S}_L]$, and $\mathcal{R}_0\left(\mathbf{t}\right) \equiv \mathcal{R}\left(\oslash, \oslash, \ldots, \oslash, t_1, t_2, \ldots, t_L\right)$, in which $\oslash$ denotes the null set. Here, $(a)$ follows from (E.2). In order to prove $(b)$, we first show that

$$\min \left\{ 1, \mu_{\{\mathrm{p}_i(t_i), \mathrm{p}_i(t_i - 1)\}} + \sum_{k \in \mathcal{S}_i} \mu_{\{\mathrm{p}_i(k), \mathrm{p}_i(k-1)\}} - |\mathcal{S}_i| \right\} \leq \max_{t'_i \in \mathcal{S}_i \cup \{t_i\}} \min \left\{ 1, \mu_{\left\{\mathrm{p}_i(t'_i), \mathrm{p}_i(t'_i - 1)\right\}} \right\}. \tag{E.4}$$

In order to verify (E.4), consider two possible scenarios: i) for all $t'_i \in \mathcal{S}_i \cup \{t_i\}$, we have $\mu_{\left\{\mathrm{p}_i(t'_i), \mathrm{p}_i(t'_i - 1)\right\}} \leq 1$. In this scenario, as in the left hand side of the inequality, we have the summation of $|\mathcal{S}_i| + 1$ positive parameters with value less than or equal to 1 subtracted by $|\mathcal{S}_i|$, we conclude that the left hand side of the inequality is less than or equal to $\mu_{\left\{\mathrm{p}_i(t'_i), \mathrm{p}_i(t'_i - 1)\right\}}$ for any $t' \in \mathcal{S}_i \cup \{t_i\}$. Hence, (E.4) is valid; ii) At least for one $t' \in \mathcal{S}_i \cup \{t_i\}$, we have $\mu_{\left\{\mathrm{p}_i(t'_i), \mathrm{p}_i(t'_i - 1)\right\}} > 1$. In this scenario, the right hand side of the inequality is equal to 1 and accordingly, (E.4) is valid. According to (E.4), we have $\mathcal{R}\left(\boldsymbol{\mathcal{S}}_1, \mathbf{t}\right) \subseteq \bigcup_{\substack{\mathbf{t}' \\ t'_i \in \mathcal{S}_i \cup \{t_i\}}} \mathcal{R}_0\left(\mathbf{t}'\right)$, which results in $(b)$ of (E.3).

On the other hand, we know that for $\boldsymbol{\mu}^0 \geq \mathbf{0}$, we have $\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0\right\} \doteq P^{-\mathbf{1} \cdot \boldsymbol{\mu}^0}$. By taking derivative with respect to $\boldsymbol{\mu}$, we have $f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) \doteq P^{-\mathbf{1} \cdot \boldsymbol{\mu}}$. Let us define $l_0 \triangleq \min_{\boldsymbol{\mu} \in \mathcal{R}_0(\mathbf{t})} \mathbf{1} \cdot \boldsymbol{\mu}$ and $\boldsymbol{\mu}_0 \triangleq \arg \min_{\boldsymbol{\mu} \in \mathcal{R}_0(\mathbf{t})} \mathbf{1} \cdot \boldsymbol{\mu}$, $\mathcal{I} \triangleq [0, l_0]^{2K}$, $\mathcal{I}_0^c \triangleq [\mu_0(1), \infty) \times [\mu_0(2), \infty) \times \cdots \times [\mu_0(L), \infty)$ and for $1 \leq i \leq L$, $\mathcal{I}_i^c \triangleq [0, \infty)^{i-1} \times [l_0, \infty) \times [0, \infty)^{L-i}$.

It is easy to verify that $\mathcal{I}_0^c \subseteq \mathcal{R}_0(\mathbf{t})$. Hence, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{R}_0(\mathbf{t})\right\} &\overset{(a)}{\doteq} \mathbb{P}\left\{\mathcal{I}_0^c\right\} + \int_{R_0(\mathbf{t})\bigcap\mathcal{I}} f_{\boldsymbol{\mu}}\left(\boldsymbol{\mu}\right) d\boldsymbol{\mu} + \sum_{i=1}^{L} \mathbb{P}\left\{\mathcal{R}_0(\mathbf{t}) \cap \mathcal{I}_i^c\right\} \\ &\overset{(b)}{\doteq} P^{-l_0}. \end{aligned} \tag{E.5}$$

Here, $(a)$ follows from the facts that i) $\mathbb{P}\left\{\bigcup_{i=1}^{M} \mathcal{A}_i\right\} \doteq \sum_{i=1}^{M} \mathbb{P}\left\{\mathcal{A}_i\right\}$, and ii) $\mathcal{I}_0^c \subseteq \mathcal{R}_0(\mathbf{t})$ and $\mathbb{R}_+^L = \mathcal{I} \bigcup \left(\bigcup_{i=1}^{L} \mathcal{I}_i^c\right)$ which imply that

$$\mathcal{R}_0(\mathbf{t}) = \mathcal{I}_0^c \bigcup \left(\mathcal{R}_0(\mathbf{t}) \bigcap \mathcal{I}\right) \bigcup \left[\bigcup_{i=1}^{M} \left(\mathcal{R}_0(\mathbf{t}) \bigcap \mathcal{I}_i^c\right)\right].$$

$(b)$ follows from the facts that $\int_{R_0(\mathbf{t})\bigcap\mathcal{I}} f_{\boldsymbol{\mu}}\left(\boldsymbol{\mu}\right) d\boldsymbol{\mu} \dot{\leq} \mathrm{vol}\left(R_0(\mathbf{t}) \bigcap \mathcal{I}\right) P^{-l_0}$ and $\mathbb{P}\left\{\mathcal{I}_0^c\right\} = \mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}_0\right\} \doteq P^{-l_0}$, noting that $\mathrm{vol}\left(R_0(\mathbf{t}) \bigcap \mathcal{I}\right)$ is a constant number independent of $P$, and $\mathbb{P}\left\{\mathcal{R}_0(\mathbf{t}) \cap \mathcal{I}_i^c\right\} \leq \mathbb{P}\left\{\mathcal{I}_i^c\right\} = P^{-l_0}$. Now, defining $\hat{\boldsymbol{\mu}} = \left[\min\left\{\mu_e, 1\right\}\right]_{e \in E}$ and $g_{\mathbf{t}}(\boldsymbol{\mu}) = \sum_{i=1}^{L} \min\left\{1, \mu_{\{\mathrm{p}_i(t_i), \mathrm{p}_i(t_i-1)\}}\right\}$, it is easy to verify that $g_{\mathbf{t}}(\hat{\boldsymbol{\mu}}) = g_{\mathbf{t}}(\boldsymbol{\mu})$ and at the same time $\mathbf{1} \cdot \hat{\boldsymbol{\mu}} < \mathbf{1} \cdot \boldsymbol{\mu}$ unless $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}$. Hence, defining $\hat{g}_{\mathbf{t}}(\boldsymbol{\mu}) = \sum_{i=1}^{L} \mu_{\{\mathrm{p}_i(t_i), \mathrm{p}_i(t_i-1)\}}$, we have

$$d_{RS,NI}(r) = \min_{\substack{\mathbf{t} \\ 1 \leq t_i \leq l_i}} \min_{\substack{\boldsymbol{\mu} \geq \mathbf{0} \\ g_{\mathbf{t}}(\boldsymbol{\mu}) \geq L - Sr}} \mathbf{1} \cdot \boldsymbol{\mu} = \min_{\substack{\mathbf{t} \\ 1 \leq t_i \leq l_i}} \min_{\substack{\mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1} \\ \hat{g}_{\mathbf{t}}(\boldsymbol{\mu}) \geq L - Sr}} \mathbf{1} \cdot \boldsymbol{\mu} = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \mathbf{1} \cdot \boldsymbol{\mu}, \tag{E.6}$$

where $\hat{\mathcal{R}} = \left\{\boldsymbol{\mu} \,\middle|\, \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1}, \sum_{i=1}^{L} \max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \geq L - Sr\right\}$. This proves the first part of the Theorem.

Now, let us define $G_{\mathrm{P}} = (V, E_{\mathrm{P}})$ as the subgraph of $G$ consisting of the edges in the path sequence, i.e. $E_{\mathrm{P}} = \left\{\left\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\right\}, \forall i, j : 1 \leq i \leq L, 1 \leq j \leq l_i\right\}$. Assume $\hat{\mathcal{S}} = \underset{\mathcal{S}}{\operatorname{argmin}}\, w_{G_{\mathrm{P}}}(\mathcal{S})$, where $\mathcal{S}$ is a cut-set on $G_{\mathrm{P}}$. We define $\hat{\boldsymbol{\mu}}$ as $\hat{\mu}_e = \frac{(L-Sr)^+}{L}$ for all $e \in E_{\mathrm{P}}$ such that $|e \cap \hat{\mathcal{S}}| = |e \cap \hat{\mathcal{S}}^c| = 1$ and $\hat{\mu}_e = 0$ for the other edges $e \in E$. As all the paths cross the cutset $\hat{\mathcal{S}}$ at least once, it follows that $\max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} = \frac{(L-Sr)^+}{L}$, which implies that $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{R}}$. Hence, we have

$$\begin{aligned} d_{RS,NI}(r) &\leq \mathbf{1} \cdot \hat{\boldsymbol{\mu}} = \frac{(L-Sr)^+}{L} \min_{\mathcal{S}} w_{G_{\mathrm{P}}}(\mathcal{S}) \overset{(a)}{\leq} \frac{(L-Sr)^+}{L} \min_{\mathcal{S}} w_G(\mathcal{S}) \\ &\overset{(b)}{\leq} (1-r)^+ \min_{\mathcal{S}} w_G(\mathcal{S}), \end{aligned} \tag{E.7}$$

where $(a)$ follows from the fact that as $G_{\mathrm{P}}$ is a sub-graph of $G$, we have $\min_{\mathcal{S}} w_{G_{\mathrm{P}}}(\mathcal{S}) \leq \min_{\mathcal{S}} w_G(\mathcal{S})$ and $(b)$ results from $S \geq L$. This proves the second part of the Theorem.

Finally, we prove the lower-bound on the DMT of the RS scheme. Let us define $d_G = \min_{\mathcal{S}} w_G(\mathcal{S})$. Consider the maximum flow algorithm [14] on $G$ from the source node 0 to the sink node $K+1$. According to the Ford-Fulkerson Theorem [14], one can achieve the maximum flow which is equal to the minimum cut of $G$ by the union of elements of a sequence $(\hat{\mathrm{p}}_1, \hat{\mathrm{p}}_2, \ldots, \hat{\mathrm{p}}_{d_G})$ of paths with the lengths $\left(\hat{l}_1, \hat{l}_2, \ldots, \hat{l}_{d_G}\right)$. Now, consider the RS scheme with $L = L_0 d_G$ paths and the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ consisting of the paths that achieve the maximum flow of $G$ such that any path $\hat{\mathrm{p}}_i$ occurs exactly $L_0$ times in the sequence. Considering $(l_1, l_2, \ldots, l_L)$ as the length sequence, we select the timing sequence as $s_{i,j} = \sum_{k=1}^{i-1} l_k + j$. It is easy to verify that, not only the timing sequence satisfies the 4 requirements needed for the RS scheme, but also the active relays with the timing sequence are non-interfering. Hence, the assumptions of the first part of the Theorem are valid. Moreover, we have $S \leq l_G L$. According to (3.3), the diversity gain of the RS scheme equals

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \sum_{e \in E} \mu_e. \tag{E.8}$$

As $\boldsymbol{\mu} \in \hat{\mathcal{R}}$, we have

$$(L - Sr)^+ \leq \sum_{i=1}^{L} \max_{1 \leq j \leq l_i} \mu_{\{\mathrm{p}_i(j), \mathrm{p}_i(j-1)\}} \overset{(a)}{\leq} L_0 \sum_{e \in E} \mu_e, \tag{E.9}$$

where $(a)$ results from the fact that as $(\hat{\mathrm{p}}_1, \hat{\mathrm{p}}_2, \ldots, \hat{\mathrm{p}}_{d_G})$ form a valid flow on $G$ (they are non-intersecting over $E$), every $e \in E$ occurs in at most one $\hat{\mathrm{p}}_i$, or equivalently, in at most $L_0$ number of $\mathrm{p}_i$'s. Combining (E.8) and (E.9), we have

$$d_{RS,NI}(r) \geq \frac{(L - Sr)^+}{L_0} \geq (1 - l_G r)^+ \, d_G = (1 - l_G r)^+ \min_{\mathcal{S}} w_G(\mathcal{S}). \tag{E.10}$$

This proves the third part of the Theorem.

# Appendix F

# Proof of Theorem 3.7

First, according to the cut-set bound Theorem [12], the point-to-point capacity of the uplink channel (the channel from the source to the relays) is an upper-bound on the achievable rate of the network. Accordingly, the diversity-multiplexing curve of a $1 \times K$ SIMO system which is a straight line (from the multiplexing gain 1 to the diversity gain $K$, i.e. $d_{opt}(r) = K(1-r)^{+}$) is an upper-bound on the DMT of the network. Now, we prove that the proposed RS scheme achieves the upper-bound on the DMT for asymptotically large values of $S$.

As the relay pairs are non-interfering ($1 \leq k \leq K : \{k, (k \mod K) + 1\} \notin E$), the result of Theorem 3.2 can be applied. As a result

$$d_{RS,NI}(r) = \min_{\boldsymbol{\mu} \in \hat{\mathcal{R}}} \sum_{e \in E} \mu_e, \tag{F.1}$$

where

$$\hat{\mathcal{R}} = \left\{ \boldsymbol{\mu} \,\middle|\, \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1}, \sum_{i=1}^{BK} \max_{1 \leq j \leq 2} \mu_{\left\{\mathrm{q}_{(i-1) \mod K+1}(j), \mathrm{q}_{(i-1) \mod K+1}(j-1)\right\}} \geq BK - (BK+1)r \right\}. \tag{F.2}$$

Hence, we have

$$BK\left(1-r-\frac{1}{BK}r\right)^{+} \overset{(a)}{\leq} B\sum_{k=1}^{K}\max\left\{\mu_{\{0,k\}},\mu_{\{K+1,k\}}\right\} \leq B\sum_{e\in E}\mu_e, \tag{F.3}$$

where $(a)$ results from the fact that every path $\mathrm{q}_k$ is used $B$ times in the path sequence. Hence, DMT can be lower-bounded as

$$d_{RS,NI}(r) \geq K\left(1-r-\frac{1}{BK}r\right)^{+}. \tag{F.4}$$

On the other hand, considering the vector $\hat{\boldsymbol{\mu}} = [\hat{\mu}_e]_{e\in E}$ where $\forall 1 \leq k \leq K : \hat{\mu}_{\{0,k\}} = \left(1-r-\frac{1}{BK}r\right)^{+}$ and $\forall k, k' \neq 0 : \hat{\mu}_{\{k,k'\}} = 0$, it is easy to verify that $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{R}}$. Hence,

$$d_{RS,NI}(r) \leq \sum_{e\in E}\hat{\mu}_e = K\left(1-r-\frac{1}{BK}r\right)^{+}. \tag{F.5}$$

Combining (F.4) and (F.5) completes the proof.

# Appendix G

# Proof of Theorem 3.9

First, we show that the entire channel matrix is equivalent to a lower triangular matrix. Let us define $\mathbf{x}_{b,k}, \mathbf{n}_{b,k}, \mathbf{r}_{b,k}, \mathbf{t}_{b,k}, \mathbf{z}_{b,k}, \mathbf{y}_{b,k}$ as the portion of signals that is sent or received in the $k$'th slot of the $b$'th sub-block. At the destination side, we have

$$\begin{aligned} \mathbf{y}_{b,k} &= g_{(k)}\mathbf{t}_{b,k} + \mathbf{z}_{b,k} \\ &= g_{(k)}\alpha_{(k)} \left( \sum_{\substack{1\le b_1 \le b, 1 \le k_1 \le K \\ b_1 K + k_1 < bK + k}} p_{b-b_1,k,k_1} \left( h_{k_1} \mathbf{x}_{b_1,k_1} + \mathbf{n}_{b_1,k_1} \right) \right) + \mathbf{z}_{b,k}. \end{aligned} \quad \text{(G.1)}$$

Here, $p_{b,k,k_1}$ has the following recursive formula $p_{0,k,k} = 1, p_{b,k,k_1} = i_{((k))}\alpha_{((k))}p_{(b),(k),k_1}$. Defining the square $BK \times BK$ matrices $\mathbf{G} = \mathbf{I}_B \otimes diag\{g_1, g_2, \cdots, g_K\}$, $\mathbf{H} =$

$\mathbf{I}_B \otimes diag\{h_1, h_2, \cdots, h_K\}$, $\mathbf{\Omega} = \mathbf{I}_B \otimes diag\{\alpha_1, \alpha_2, \cdots, \alpha_K\}$, and

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ p_{0,2,1} & 1 & 0 & 0 & \dots \\ p_{0,3,1} & p_{0,3,2} & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ p_{B-1,K,1} & p_{B-1,K,2} & \dots & p_{0,K,K-1} & 1 \end{pmatrix}, \tag{G.2}$$

where $\otimes$ is the Kronecker product [22] of matrices and $\mathbf{I}_B$ is the $B \times B$ identity matrix, and the $BK \times 1$ vectors $\mathbf{x}(s) = [x_{1,1}(s), x_{1,2}(s), \cdots, x_{B,K}(s)]^T$, $\mathbf{n}(s) = [n_{1,1}(s), n_{1,2}(s), \cdots, n_{B,K}(s)]^T$, $\mathbf{z}(s) = [z_{1,2}(s), z_{1,3}(s), \cdots, z_{B+1,1}(s)]^T$, and $\mathbf{y}(s) = [y_{1,2}(s), y_{1,3}(s), \cdots, y_{B+1,1}(s)]^T$, we have

$$\mathbf{y}(s) = \mathbf{G\Omega F}(\mathbf{Hx}(s) + \mathbf{n}(s)) + \mathbf{z}(s). \tag{G.3}$$

Here, we observe that the matrix of the entire channel is equivalent to a lower triangular matrix of size $BK \times BK$ for a MIMO system with a colored noise. The probability of outage of such a channel for the multiplexing gain $r$ ($r \leq 1$) is defined as

$$\mathbb{P}\{\mathcal{E}\} = \mathbb{P}\left\{\log\left|\mathbf{I}_{BK} + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| \leq (BK+1)r\log(P)\right\}, \tag{G.4}$$

where $\mathbf{P}_n = \mathbf{I}_{BK} + \mathbf{G\Omega FF}^H\mathbf{\Omega}^H\mathbf{G}^H$, and $\mathbf{H}_T = \mathbf{G\Omega FH}$. Assume $|h_k|^2 = P^{-\mu_k}$, $|g_k|^2 = P^{-\nu_k}$, $|i_k|^2 = P^{-\omega_k}$, and $\mathcal{R}$ as the region in $\mathbb{R}^{3K}$ that defines the outage event $\mathcal{E}$ in terms of the vector $[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T, \boldsymbol{\omega}^T]^T$, where $\boldsymbol{\mu} = [\mu_1\mu_2\cdots\mu_K]^T, \boldsymbol{\nu} = [\nu_1\nu_2\cdots\nu_K]^T, \boldsymbol{\omega} = [\omega_1\omega_2\cdots\omega_K]^T$. The probability distribution function (and also the complement of the cumulative distribution function) decays exponentially as $P^{-P^{-\delta}}$ for positive values of $\delta$. Hence, the outage region $\mathcal{R}$ is almost surely equal to $\mathcal{R}_+ = \mathcal{R} \bigcap \mathbb{R}_+^{3K}$. Now, we have

$$
\begin{aligned}
\mathbb{P}\{\mathcal{E}\} &\overset{(a)}{\leq} \mathbb{P}\left\{ |\mathbf{H}_T|^2 |\mathbf{P}_n|^{-1} \leq P^{-BK(1-r)+r} \right\} \\
&\overset{(b)}{\leq} \mathbb{P}\Bigg\{ -B \sum_{k=1}^{K} \left( \mu_k + \nu_k - \min\left\{0, \mu_k, \omega_{(k)}\right\} \right) - \frac{BK\log(3) + \log|\mathbf{P}_n|}{\log(P)} \leq \\
&\qquad -BK(1-r) + r \Bigg\} \\
&\overset{(c)}{\dot{\leq}} \mathbb{P}\Bigg\{ -BK \frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + BK(1-r) - r \leq B \sum_{k=1}^{K} (\mu_k + \nu_k), \\
&\qquad \mu_k, \nu_k, \omega_k \geq 0 \Bigg\}. \qquad \text{(G.5)}
\end{aligned}
$$

Here, $(a)$ follows from the fact that for a positive semidefinite matrix $\mathbf{A}$, we have $|\mathbf{I} + \mathbf{A}| \geq |\mathbf{A}|$ and $(b)$ follows from the fact that

$$
|\alpha_k|^2 = \min\left\{1, \frac{P}{P^{1-\mu_k} + P^{1-\omega_{(k)}} + 1}\right\} \geq \frac{1}{3}\min\left\{1, P, P^{\mu_k}, P^{\omega_{(k)}}\right\}
$$

and assuming $P$ is large enough such that $P \geq 1$. Finally, $(c)$ is proved as follows:

As $|\alpha_k| \leq 1$, we conclude $p_{n,k,k_1} \leq 1$. Hence, the sum of the entries of each row in $\mathbf{F}\mathbf{F}^H$ is less than $B^2K^2$. Now, consider the matrix $\mathbf{A} \triangleq B^2K^2\mathbf{I} - \mathbf{F}\mathbf{F}^H$. From the above discussion, it follows that for every $i$, we have $A_{i,i} \geq \sum_{i \neq j} |A_{i,j}|$. Hence, for every vector $\mathbf{x}$, we have $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq \sum_{i<j} |A_{i,j}| x_i^2 + |A_{i,j}| x_j^2 \pm 2|A_{i,j}| x_i x_j = \sum_{i<j} |A_{i,j}| \left(x_i \pm x_j\right)^2 \geq 0$, and as a result $\mathbf{A}$ is positive semidefinite, which implies that $\mathbf{F}\mathbf{F}^H \preccurlyeq B^2K^2\mathbf{I}_{BK}$. Consequently, we have $\mathbf{P}_n \preccurlyeq \mathbf{I}_{BK} + B^2K^2\mathbf{G}\mathbf{\Omega}\mathbf{\Omega}^H\mathbf{G}^H$. Moreover, Knowing the fact that $\mathbb{P}\{\mathcal{R}\} \doteq \mathbb{P}\{\mathcal{R}_+\}$, and conditioned on $\mathcal{R}_+$, one has $|g_k|^2 \leq 1$, which implies that $\mathbf{G}\mathbf{G}^H \preccurlyeq \mathbf{I}$. Combining this with the fact that $\mathbf{\Omega}\mathbf{\Omega}^H \preccurlyeq \mathbf{I}$ (as $|\alpha_k|^2 \leq 1,\ \forall k$) yields $\mathbf{P}_n \preccurlyeq \mathbf{I}_{BK} + B^2K^2\mathbf{G}\mathbf{\Omega}\mathbf{\Omega}^H\mathbf{G}^H \preccurlyeq \left(B^2K^2+1\right)\mathbf{I}_{BK}$. Moreover, conditioned on $\mathcal{R}_+$, we have $\min\left\{0, \mu_k, \omega_{(k)}\right\} = 0$. This completes the proof of $(c)$.

On the other hand, we have $\mathbb{P}\left\{\boldsymbol{\mu} \geq \boldsymbol{\mu}^0, \boldsymbol{\nu} \geq \boldsymbol{\nu}^0, \boldsymbol{\omega} \geq \boldsymbol{\omega}^0\right\} \doteq P^{-\mathbf{1}\cdot\left(\boldsymbol{\mu}^0+\boldsymbol{\nu}^0+\boldsymbol{\omega}^0\right)}$, for any vectors $\boldsymbol{\mu}^0, \boldsymbol{\nu}^0, \boldsymbol{\omega}^0 \geq \mathbf{0}$. Similar to the proof of Theorem 3.2, by taking derivative with respect to $\boldsymbol{\mu}, \boldsymbol{\nu}$, we have $f_{\boldsymbol{\mu},\boldsymbol{\nu}}(\boldsymbol{\mu}, \boldsymbol{\nu}) \doteq P^{-\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu})}$. Defining $l_0 \triangleq -\frac{\log\left[3\left(B^2K^2+1\right)\right]}{\log(P)} + (1-r) - \frac{r}{BK}$, $\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\mu}, \boldsymbol{\nu} \geq \mathbf{0}, \frac{1}{K}\mathbf{1}\cdot(\boldsymbol{\mu}+\boldsymbol{\nu}) \geq l_0\right\}$, the cube $\mathcal{I}$ as $\mathcal{I} \triangleq [0, Kl_0]^{2K}$, and for $1 \leq i \leq 2K$, $\mathcal{I}_i^c \triangleq [0,\infty)^{i-1} \times [Kl_0, \infty) \times [0,\infty)^{2K-i}$, we observe

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{E}\right\} &\overset{(a)}{\dot{\leq}} \mathbb{P}\{\hat{\mathcal{R}}\} \\
&\overset{(b)}{\leq} \int_{\hat{\mathcal{R}}\bigcap\mathcal{I}} f_{\boldsymbol{\mu},\boldsymbol{\nu}}\left(\boldsymbol{\mu}, \boldsymbol{\nu}\right) d\boldsymbol{\mu} d\boldsymbol{\nu} + \sum_{i=1}^{2K} \mathbb{P}\left\{[\boldsymbol{\mu}^T, \boldsymbol{\nu}^T]^T \in \hat{\mathcal{R}} \cap \mathcal{I}_i^c\right\} \\
&\dot{\leq} \operatorname{vol}(\hat{\mathcal{R}} \cap \mathcal{I}) P^{-\min_{\left[\boldsymbol{\mu}_0^T, \boldsymbol{\nu}_0^T\right]^T \in \hat{\mathcal{R}}\bigcap\mathcal{I}} \mathbf{1}\cdot(\boldsymbol{\mu}_0+\boldsymbol{\nu}_0)} + 2KP^{-Kl_0} \\
&\overset{(c)}{\doteq} P^{-Kl_0} \\
&\doteq P^{-\left[K(1-r)-\frac{r}{B}\right]}.
\end{aligned} \tag{G.6}$$

Here, $(a)$ follows from (G.5), knowing $\hat{\mathcal{R}} = \left(\hat{\mathcal{R}}\bigcap\mathcal{I}\right) \bigcup \left[\bigcup_{i=1}^M \left(\hat{\mathcal{R}}\bigcap\mathcal{I}_i^c\right)\right]$ and using the union bound on the probability results in $(b)$, and $(c)$ follows from the fact that $\hat{\mathcal{R}}\bigcap\mathcal{I}$ is a bounded region whose volume is independent of $P$. (G.6) completes the proof of Theorem 3.9.

# Appendix H

# Proof of Lemma 5.1

Assuming a multiplexing gain of $r$, the diversity of the original system can be written as

$$d(r) = \lim_{P\to\infty} -\frac{\log\left(\mathbb{P}\left\{\log\left|\mathbf{I}+\alpha^2 P\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\left(\mathbf{I}+\alpha^2\mathbf{G}\mathbf{G}^H\right)^{-1}\right| < r\log(P)\right\}\right)}{\log(P)} \tag{H.1}$$

Since $\alpha^2\mathbf{G}\mathbf{G}^H \succeq \mathbf{0}$, it follows that

$$\left|\mathbf{I}+\alpha^2 P\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\left(\mathbf{I}+\alpha^2\mathbf{G}\mathbf{G}^H\right)^{-1}\right| < \left|\mathbf{I}+\alpha^2 P\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\right|. \tag{H.2}$$

This implies that

$$d(r) \le d_u(r) \triangleq \lim_{P\to\infty} -\frac{\log\left(\mathbb{P}\left\{\log\left|\mathbf{I}+\alpha^2 P\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\right| < r\log(P)\right\}\right)}{\log(P)}, \tag{H.3}$$

which is the DMT of the system in (5.2). Moreover, from the output power constraint $P$ for the relay[1], we have $\alpha = \min\left(1, \frac{P}{\mathbb{E}\{\|\mathbf{H}\mathbf{x}_t+\mathbf{n}_r\|^2\}}\right)$. As a result, we have

$$\alpha^2\mathbf{G}\mathbf{G}^H \preceq \|\mathbf{G}\|^2\mathbf{I}. \tag{H.4}$$

[1]Note that as we are looking for a lower-bound for the DMT, it is fine to select any arbitrary value for $\alpha$ which satisfies the output power constraint.

Defining the event $\mathscr{C} \equiv \{\|\mathbf{G}\|^2 > c\log(P)\}$, and noting that $\|\mathbf{G}\|^2$ is a Chi-square random variable with $2pn$ degrees of freedom, we have

$$\begin{aligned} \mathbb{P}\{\mathscr{C}\} &= \sum_{k=0}^{pn-1} \frac{(c\log(P))^k}{k!} e^{-c\log(P)} \\ &\sim d(\log(P))^{pn-1} P^{-c}, \end{aligned} \tag{H.5}$$

where $d = \frac{c^k}{k!}$. Defining the outage event of the system in (5.1) as $\mathcal{O}$, we have

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &= \mathbb{P}\{\mathcal{O}|\mathscr{C}\}\mathbb{P}\{\mathscr{C}\} + \mathbb{P}\{\mathcal{O}|\mathscr{C}^c\}\mathbb{P}\{\mathscr{C}^c\} \\ &\leq \mathbb{P}\{\mathscr{C}\} + \mathbb{P}\{\mathcal{O}|\mathscr{C}^c\}\mathbb{P}\{\mathscr{C}^c\}. \end{aligned} \tag{H.6}$$

Conditioned on $\mathscr{C}^c$, the probability of the outage event can be upper-bounded as

$$\mathbb{P}\{\mathcal{O}|\mathscr{C}^c\} \leq \mathbb{P}\left\{\log\left|\mathbf{I} + \alpha^2 \frac{P}{c\log P + 1}\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\right| < r\log(P) \middle| \mathscr{C}^c\right\}, \tag{H.7}$$

which is equal to the probability of the outage event of the system in (5.3), denoted as $\mathcal{O}_l$, conditioned on $\mathscr{C}^c$. In other words, $\mathbb{P}\{\mathcal{O}|\mathscr{C}^c\} \leq \mathbb{P}\{\mathcal{O}_l|\mathscr{C}^c\}$. Substituting in (H.6) yields

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &\leq \mathbb{P}\{\mathscr{C}\} + \mathbb{P}\{\mathcal{O}_l|\mathscr{C}^c\}\mathbb{P}\{\mathscr{C}^c\}. \\ &\leq \mathbb{P}\{\mathscr{C}\} + \mathbb{P}\{\mathcal{O}_l\}. \end{aligned} \tag{H.8}$$

Since the capacity of the two-hop network is equal to the capacity of both the source-relay and the relay-destination links, it follows that the outage event $\mathcal{O}$ includes $\mathcal{O}_{sr}$, the event of outage in the source-relay link, and $\mathcal{O}_{rd}$, the event of outage in the relay-destination link. As a result,

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &\geq \max\{\mathbb{P}\{\mathcal{O}_{sr}\}, \mathbb{P}\{\mathcal{O}_{rd}\}\} \\ &\overset{(a)}{\dot{\geq}} \max\left\{P^{-mp}, P^{-pn}\right\} \\ &\dot{\geq} P^{-pn}. \end{aligned} \tag{H.9}$$

$(a)$ results from the fact that the outage probability corresponding to the multiplexing gain $r$ is greater than or equal to the outage probability corresponding to the multiplexing gain 0. Setting $c = 2pn - 1$, from (H.5), it is concluded that

$$\frac{\mathbb{P}\{\mathscr{C}\}}{\mathbb{P}\{\mathcal{O}\}} = \Theta\left(\left(\frac{\log(P)}{P}\right)^{pn-1}\right). \tag{H.10}$$

From the above equation and (H.8), it follows that

$$\mathbb{P}\{\mathcal{O}\} \dot{\leq} \mathbb{P}\{\mathcal{O}_l\}, \tag{H.11}$$

which incurs that the DMT of the original system is lower-bounded by the DMT of the system in (5.3). This completes the proof of Lemma 5.1.

# Appendix I

# Proof of Lemma 5.2

Defining $P' = \frac{P}{c\log(P)+1}$, the DMT of the system in (5.3) can be written as

$$\begin{aligned}
d_l(r) &= \lim_{P\to\infty} -\frac{\log\left(\mathbb{P}\left\{\log\left|\mathbf{I}+\alpha^2 P'\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\right| < r\log(P)\right\}\right)}{\log(P)} \\
&\overset{(a)}{=} \lim_{P'\to\infty} -\frac{\log\left(\mathbb{P}\left\{\log\left|\mathbf{I}+\alpha^2 P'\mathbf{G}\mathbf{H}\mathbf{H}^H\mathbf{G}^H\right| < r'\log(P')\right\}\right)}{\log(P')} \\
&\overset{(b)}{=} d_u(r') \\
&\overset{(c)}{=} d_u(r),
\end{aligned} \tag{I.1}$$

where $r' \triangleq r\frac{\log(P)}{\log(P')}$ and $(a)$ follows from the fact that as $P' = \frac{P}{c\log(P)+1}$, we have $P = P'\left[c\log(P') + O(\log\log(P'))\right]$, which implies that $\lim_{P\to\infty}\frac{\log(P')}{\log(P)} = 1$. In other words, in the first line of the preceding equation, one can substitute $\log(P)$ by $\log(P')$. $(b)$ results from the fact that the second line in the right hand side of the preceding equation is exactly the DMT of the system in (5.2) at $r'$, which is denoted by $d_u(r')$. Finally, $(c)$ follows from the facts that $\lim_{P\to\infty}\frac{r'}{r} = 1$ and the continuity of the DMT curve which implies that $d_u(r') = d_u(r)$. This completes the proof of Lemma.

# Appendix J

# Proof of Lemma 5.4

First, notice that for values of $i > \min\{m, p\}$ or $i > \min\{p, n\}$, $\lambda_i(\mathbf{G})$ or $\lambda_i(\mathbf{H})$ are defined as zero. Hence, the argument of this lemma is obvious in these cases. Now, we prove the argument for $i \leq \min\{m, n, p\}$. According to the Courant-Fischer-Weyl Theorem [22], we have

$$\lambda_i\left(\mathbf{G\Theta H}\right) = \max_{\mathcal{S}, \dim(\mathcal{S})=i} \min_{\mathbf{x}\in\mathcal{S}} \frac{\|\mathbf{G\Theta Hx}\|^2}{\|\mathbf{x}\|^2}. \tag{J.1}$$

Now, let us define $\mathcal{S}_0 = \langle \mathbf{v}_1(\mathbf{H}), \mathbf{v}_2(\mathbf{H}), \ldots, \mathbf{v}_i(\mathbf{H}) \rangle$ where $\mathbf{v}_j(\mathbf{H})$ denotes the $j$'th column of $\mathbf{V}(\mathbf{H})$ and $\langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \rangle$ denotes the span of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$. We have

$$\begin{aligned}
\lambda_i(\mathbf{G\Theta H}) &\geq \min_{\mathbf{x}\in\mathcal{S}_0} \frac{\|\mathbf{G\Theta Hx}\|^2}{\|\mathbf{x}\|^2} \\
&\stackrel{(a)}{=} \min_{\mathbf{x}'\in\mathbb{C}^i} \frac{\left\|\mathbf{G\Theta HV}_{(1,i)}(\mathbf{H})\mathbf{x}'\right\|^2}{\|\mathbf{x}'\|^2} \\
&= \min_{\mathbf{x}'\in\mathbb{C}^i} \frac{\left\|\mathbf{G\Theta U}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{H})\mathbf{I}_{n\times i}\mathbf{x}'\right\|^2}{\|\mathbf{x}'\|^2} \\
&= \min_{\mathbf{x}'\in\mathbb{C}^i} \frac{\left\|\mathbf{G\Theta U}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}_{(1,i)}(\mathbf{H})\mathbf{x}'\right\|^2}{\|\mathbf{x}'\|^2} \\
&= \min_{\substack{\mathbf{x}'\in\mathbb{C}^i \\ \|\mathbf{x}'\|^2\geq 1}} \left\|\mathbf{G\Theta U}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}_{(1,i)}(\mathbf{H})\mathbf{x}'\right\|^2 \\
&= \min_{\substack{\mathbf{x}'\in\mathbb{C}^i \\ \|\mathbf{x}'\|^2\geq 1}} \left\|\mathbf{G\Theta U}_{(1,i)}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}_{i}(\mathbf{H})\mathbf{x}'\right\|^2 \\
&\stackrel{(b)}{\geq} \min_{\substack{\mathbf{y}'\in\mathbb{C}^i \\ \|\mathbf{y}'\|^2\geq 1}} \lambda_i(\mathbf{H}) \left\|\mathbf{G\Theta U}_{(1,i)}(\mathbf{H})\mathbf{y}'\right\|^2 \\
&\stackrel{(c)}{=} \min_{\substack{\mathbf{y}'\in\mathbb{C}^i \\ \|\mathbf{y}'\|^2\geq 1}} \lambda_i(\mathbf{H}) \left\|\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}^H(\mathbf{G})\mathbf{\Theta U}_{(1,i)}(\mathbf{H})\mathbf{y}'\right\|^2 \\
&\stackrel{(d)}{\geq} \min_{\substack{\mathbf{y}'\in\mathbb{C}^i \\ \|\mathbf{y}'\|^2\geq 1}} \lambda_i(\mathbf{H}) \left\|\mathbf{\Lambda}^{\frac{1}{2}}_{(1,i)}(\mathbf{G})\mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta U}_{(1,i)}(\mathbf{H})\mathbf{y}'\right\|^2 \\
&\stackrel{(e)}{\geq} \min_{\substack{\mathbf{y}'\in\mathbb{C}^i \\ \|\mathbf{y}'\|^2\geq 1}} \lambda_i(\mathbf{G})\lambda_i(\mathbf{H}) \left\|\mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta U}_{(1,i)}(\mathbf{H})\mathbf{y}'\right\|^2 \\
&\stackrel{(f)}{=} \lambda_i(\mathbf{G})\lambda_i(\mathbf{H})\lambda_{\min}\left(\mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta U}_{(1,i)}(\mathbf{H})\right),
\end{aligned} \tag{J.2}$$

where $\mathbf{I}_{n\times i}$ denotes the diagonal identity $n \times i$ matrix and $\Lambda_i^{\frac{1}{2}}(\mathbf{H})$ denotes the square submatrix of $\Lambda^{\frac{1}{2}}(\mathbf{H})$ consisting of its first $i$ rows and first $i$ columns. Here,

$(a)$ follows from the fact that i) $\mathbf{x} \in \mathcal{S}_0$ is equivalent to $\mathbf{x} = \mathbf{V}_{(1,i)}\mathbf{x}'$ for some $\mathbf{x}' \in \mathbb{C}^i$; and ii) $\|\mathbf{x}\|^2 = \|\mathbf{x}'\|^2$. $(b)$ follows from the fact that for any $\mathbf{x}' \in \mathbb{C}^i$, $\|\mathbf{x}'\|^2 \geq 1$, defining $\mathbf{y}' = \frac{1}{\sqrt{\lambda_i(\mathbf{H})}}\mathbf{\Lambda}_i^{\frac{1}{2}}\mathbf{x}'$, we have $\|\mathbf{y}'\|^2 \geq \|\mathbf{x}'\|^2 \geq 1$. $(c)$ follows from the fact that for any unitary matrix $\mathbf{P}$, we have $\|\mathbf{PA}\|^2 = \|\mathbf{A}\|^2$. $(d)$ follows from the fact that defining $\mathbf{z}' = \mathbf{\Theta}\mathbf{U}_{(1,i)}(\mathbf{H})\mathbf{y}'$, we have

$$\left\|\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}^H(\mathbf{G})\mathbf{z}'\right\|^2 = \left\|\mathbf{\Lambda}_{(1,i)}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}_{(1,i)}^H(\mathbf{G})\mathbf{z}'\right\|^2 + \left\|\mathbf{\Lambda}_{(i+1,m)}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}_{(i+1,m)}^H(\mathbf{G})\mathbf{z}'\right\|^2.$$

$(e)$ follows from the fact that defining $\mathbf{w}' = \mathbf{V}_{(1,i)}^H(\mathbf{G})\mathbf{\Theta}\mathbf{U}_{(1,i)}(\mathbf{H})\mathbf{y}'$, we have $\left\|\mathbf{\Lambda}_{(1,i)}^{\frac{1}{2}}\mathbf{w}'\right\|^2 \geq \lambda_i(\mathbf{G})\|\mathbf{w}'\|^2$. Finally, $(f)$ follows from the Courant-Fischer-Weyl Theorem [22]. (J.2) completes the proof of Lemma.

# Appendix K

# Proof of Lemma 5.5

We have

$$
\begin{aligned}
\left|\boldsymbol{\Psi}_{i,l}\boldsymbol{\Psi}_{i,l}^H\right| &= \lambda_{\min}(\boldsymbol{\Psi}_{i,l})\prod_{j=1}^{i-1}\lambda_i(\boldsymbol{\Psi}_{i,l}) \\
&\overset{(a)}{\leq} \lambda_{\min}(\boldsymbol{\Psi}_{i,l})\left(\frac{\mathsf{Tr}\left\{\boldsymbol{\Psi}_{i,l}\boldsymbol{\Psi}_{i,l}^H\right\}}{i-1}\right)^{i-1}, \qquad \text{(K.1)}
\end{aligned}
$$

where $(a)$ follows from the *Geometric Inequality* and the fact that $\sum_{j=1}^{i-1}\lambda_i(\boldsymbol{\Psi}_{i,l}) \leq \mathsf{Tr}\left\{\boldsymbol{\Psi}_{i,l}\boldsymbol{\Psi}_{i,l}^H\right\}$. $\mathsf{Tr}\left\{\boldsymbol{\Psi}_{i,l}\boldsymbol{\Psi}_{i,l}^H\right\}$ can be upper-bounded as follows:

$$
\begin{aligned}
\mathsf{Tr}\left\{\boldsymbol{\Psi}_{i,l}\boldsymbol{\Psi}_{i,l}^H\right\} &= \left\|\mathbf{V}_{(1,i)}^H(\mathbf{G})\boldsymbol{\Theta}_l\mathbf{U}_{(1,i)}(\mathbf{H})\right\|^2 \\
&\overset{(a)}{\leq} \left\|\mathbf{V}_{(1,i)}(\mathbf{G})\right\|^2\left\|\boldsymbol{\Theta}_l\mathbf{U}_{(1,i)}(\mathbf{H})\right\|^2 \\
&\overset{(b)}{\leq} \left\|\mathbf{V}_{(1,i)}(\mathbf{G})\right\|^2\left\|\mathbf{U}_{(1,i)}(\mathbf{H})\right\|^2 \\
&\overset{(c)}{=} i^2. \qquad \text{(K.2)}
\end{aligned}
$$

In the preceding equation, $(a)$ follows from the the fact that $\|\mathbf{AB}\|^2 \leq \|\mathbf{A}\|^2\|\mathbf{B}\|^2$ for any two matrices $\mathbf{A}$ and $\mathbf{B}$. $(b)$ results from the fact that $\|\boldsymbol{\Theta}\mathbf{A}\| = \|\mathbf{A}\|$, for any matrix $\mathbf{A}$ and any unitary matrix $\boldsymbol{\Theta}$. Finally, $(c)$ follows from the fact that as $\mathbf{U}(\mathbf{H})$ and $\mathbf{V}(\mathbf{G})$ are unitary matrices, each of their columns has unit norm. Combining (K.1) and (K.2) yields

$$\left|\mathbf{\Psi}_{i,l}\mathbf{\Psi}_{i,l}^{H}\right| \;\; \leq \;\; \lambda_{\min}(\mathbf{\Psi}_{i,l})\left(\frac{i^2}{i-1}\right)^{i-1}, \tag{K.3}$$

which implies that

$$\lambda_{\min}(\mathbf{\Psi}_{i,l}) \geq c\left|\mathbf{\Psi}_{i,l}\mathbf{\Psi}_{i,l}^{H}\right|, \tag{K.4}$$

for some constant $c$. Denoting the $j$th column of $\mathbf{\Psi}_{i,l}$ as $\mathbf{\Psi}_{i,l}^{(j)}$, we write the determinant of $\mathbf{\Psi}_{i,l}\mathbf{\Psi}_{i,l}^{H}$ as

$$\left|\mathbf{\Psi}_{i,l}\mathbf{\Psi}_{i,l}^{H}\right| = \prod_{j=1}^{i}\beta_j, \tag{K.5}$$

where $\beta_j$ denotes the square norm of the projection of $\mathbf{\Psi}_{i,l}^{(j)}$ over the null-space of the subspace spanned by $\left\{\mathbf{\Psi}_{i,l}^{(s)}\right\}_{s=1}^{j-1}$. Combining (K.4) and (K.5) yields

$$\begin{aligned}
\mathbb{P}\left\{\lambda_{\min}(\mathbf{\Psi}_{i,l}) \leq \varepsilon\right\} \;\; &\leq \;\; \mathbb{P}\left\{\prod_{j=1}^{i}\beta_j \leq \kappa\varepsilon\right\} \\
&\leq \;\; \mathbb{P}\left\{\bigcup_{j=1}^{i}\left\{\beta_j \leq \sqrt[i]{\kappa\varepsilon}\right\}\right\} \\
&\overset{(a)}{\leq} \;\; \sum_{j=1}^{i}\mathbb{P}\left\{\beta_j \leq \sqrt[i]{\kappa\varepsilon}\right\},
\end{aligned} \tag{K.6}$$

where $\kappa = \frac{1}{c}$ and $(a)$ follow from the union bound on the probability. $\mathbf{\Psi}_{i,l}^{(s)}$ can be considered as the projection of the $s$th column of the matrix $\mathbf{R}_l \triangleq \mathbf{\Theta}_l\mathbf{U}_{(1,i)}(\mathbf{H})$, denoted by $\mathbf{R}_l^{(j)}$, over the $i$-dimensional subspace spanned by $\{\mathbf{V}_t(\mathbf{G})\}_{t=1}^{i}$, which is denoted by $\boldsymbol{\mathcal{P}}$. Now, consider a random unit vector $\mathbf{w}$ in $\boldsymbol{\mathcal{P}}$, which is orthogonal to the first $j-1$ columns of the matrix $\mathbf{R}_l$. Since $\{\mathbf{\Psi}_{i,l}^{(s)}\}_{s=1}^{j-1}$ are the projections of the first $j-1$ columns of $\mathbf{R}_l$ over $\boldsymbol{\mathcal{P}}$, it follows that $\mathbf{w}$ is also orthogonal to $\{\mathbf{\Psi}_{i,l}^{(s)}\}_{s=1}^{j-1}$. In other words, $\mathbf{w}$ belongs to $\boldsymbol{\mathcal{Q}}^{\perp}$, the null-space of the subspace spanned by $\left\{\mathbf{\Psi}_{i,l}^{(s)}\right\}_{s=1}^{j-1}$[1]. Hence, we have

$$\begin{aligned}\beta_j &= \left\|{\mathbf{\Psi}_{i,l}^{(j)}}^H * \boldsymbol{\mathcal{Q}}^{\perp}\right\|^2 \\ &\geq \left|{\mathbf{\Psi}_{i,l}^{(j)}}^H \mathbf{w}\right|^2, \end{aligned} \tag{K.7}$$

where $\mathbf{a}^H * \boldsymbol{\mathcal{T}}$ denotes the projection of the vector $\mathbf{a}$ over the subspace $\boldsymbol{\mathcal{T}}$. The second line in the preceding equation follows from the fact that the norm of the projection of a vector over a subspace is greater than the norm of the projection of that vector over any arbitrary unit vector in that subspace. Since $\mathbf{\Psi}_{i,l}^{(j)}$ is the projection of $\mathbf{R}_l^{(j)}$ over $\boldsymbol{\mathcal{P}}$, $\mathbf{R}_l^{(j)}$ can be written as

$$\mathbf{R}_l^{(j)} = \mathbf{\Psi}_{i,l}^{(j)} + {\mathbf{\Psi}_{i,l}^{(j)}}^{\perp}, \tag{K.8}$$

where ${\mathbf{\Psi}_{i,l}^{(j)}}^{\perp}$ denotes the projection of $\mathbf{R}_l^{(j)}$ over $\boldsymbol{\mathcal{P}}^{\perp}$, the null-space of $\boldsymbol{\mathcal{P}}$. Since $\mathbf{w} \in \boldsymbol{\mathcal{P}}$, it follows that $\mathbf{w}^H {\mathbf{\Psi}_{i,l}^{(j)}}^{\perp} = 0$. This means that ${\mathbf{\Psi}_{i,l}^{(j)}}^H \mathbf{w} = {\mathbf{R}_l^{(j)}}^H \mathbf{w}$. Combining the above with (K.7) yields

$$\beta_j \geq \left|{\mathbf{R}_l^{(j)}}^H \mathbf{w}\right|^2. \tag{K.9}$$

Since $\mathbf{\Theta}_l$ and $\mathbf{U}(\mathbf{H})$ are unitary matrices and $\mathbf{U}(\mathbf{H})$ is isotropically distributed [57], it follows that $\mathbf{\Theta}_l \mathbf{U}(\mathbf{H})$ is an isotropic unitary matrix. This means i) $\left\{\mathbf{R}_l^{(s)}\right\}_{s=1}^{j}$ is an orthonormal set, which implies that $\mathbf{R}_l^{(j)}$ is orthogonal to $\mathbf{R}_l^{(s)}$, $s = 1, \cdots, j-1$, and ii) $\mathbf{R}_l^{(j)}$ is an isotropic unit vector. As a result, $\mathbf{R}_l^{(j)}$ is an isotropic unit vector in the $(p-j+1)$-dimensional subspace perpendicular to $\left\{\mathbf{R}_l^{(s)}\right\}_{s=1}^{j-1}$. Noting that

[1] Note that as $j \leq i$, $\boldsymbol{\mathcal{Q}}^{\perp}$ has at least dimension of 1.

$\mathbf{w}$ is also in this subspace and $\mathbf{R}_l^{(j)}$ and $\mathbf{w}$ are independent of each other, from [11], Lemma 3, the CDF of $Z_j \triangleq \left|\mathbf{R}_l^{(j)^H}\mathbf{w}\right|^2$ can be computed as

$$F_{Z_j}(z) = 1 - (1-z)^{p-j}. \tag{K.10}$$

Combining (K.6), (K.9), and (K.10), it follows that

$$\begin{aligned}
\mathbb{P}\left\{\lambda_{\min}(\mathbf{\Psi}_{i,l}) \leq \varepsilon\right\} &\leq \sum_{j=1}^{i} F_{Z_j}(\sqrt[i]{\kappa\varepsilon}) \\
&= \sum_{j=1}^{i}\left[1 - \left(1 - \sqrt[i]{\kappa\varepsilon}\right)^{p-j}\right] \\
&\stackrel{(a)}{\leq} \sum_{j=1}^{i}(p-j)\sqrt[i]{\kappa\varepsilon} \\
&= \eta\sqrt[i]{\varepsilon},
\end{aligned} \tag{K.11}$$

where $\eta = i(p - \frac{i+1}{2})\sqrt[i]{\kappa}$. In the above equation, $(a)$ follows from the fact that $(1-x)^n \geq 1 - nx$, $\forall n \geq 0, 0 \leq x \leq 1$. This completes the proof of Lemma 5.5.

# Appendix L

# Proof of Lemma 5.6

We can assume that $f(\mathbf{x})$ is defined such that $\mathbf{x} \leq \mathbf{y} \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y})$; otherwise, we can redefine $f(\mathbf{x})$ as $f(\mathbf{x}) = \inf_{\mathbf{y} \geq \mathbf{x}} f(\mathbf{y})$. Let us define the set $\mathcal{S}(\epsilon)$ for every $\epsilon > 0$ as $\mathcal{S}(\epsilon) \triangleq \left\{\mathbf{x} \in [0, \infty)^n \,\middle|\, d(\mathbf{x}, \mathcal{R}) \leq \sqrt{n}\epsilon, \frac{\mathbf{x}}{\epsilon} \in \mathbb{Z}^n\right\}$ where $d(\mathbf{x}, \mathcal{R}) = \inf_{\mathbf{y} \in \mathcal{R}} \|\mathbf{x} - \mathbf{y}\|$. Also, let us define the partial order $\leq$ for two sets $\mathcal{A}, \mathcal{B} \subseteq [0, \infty)^n$ as follows: $\mathcal{A} \leq \mathcal{B}$ iff for every $\mathbf{x} \in \mathcal{B}$ there exists a $y \in \mathcal{A}$ such that $\mathbf{y} \leq \mathbf{x}$[1]. One can easily verify that $\mathcal{S}(\epsilon) \leq \mathcal{R}$. Now, let us define the set $\mathcal{L}(\epsilon)$ as $\mathcal{L}(\epsilon) \triangleq \{\mathbf{x} \in \mathcal{S}(\epsilon) \,|\, \nexists \mathbf{y} \in \mathcal{S}(\epsilon), \mathbf{y} < \mathbf{x}\}$, i.e. $\mathcal{L}(\epsilon)$ consists of the minimal members of $\mathcal{S}(\epsilon)$. Notice that the elements of $\mathcal{S}(\epsilon)$ can be mapped to the elements of $\mathbb{Z}_+^n$ such that the partial order $<$ between the real vectors is kept between the corresponding integer vectors[2]. For every subset $\mathcal{A} \subseteq \mathbb{Z}_+^n$ and every $\mathbf{x} \in \mathcal{A}$, there exists a minimal member $\mathbf{y} \in \mathcal{A}$ such that $\mathbf{y} \leq \mathbf{x}$. Accordingly, we have such a property for $\mathcal{S}(\epsilon)$. This means $\mathcal{L}(\epsilon) \leq \mathcal{S}(\epsilon)$. Noticing $\mathcal{S}(\epsilon) \leq \mathcal{R}$, we conclude that $\mathcal{L}(\epsilon) \leq \mathcal{R}$. On the other hand, it is easy to check that

[1]It should be noted that $\leq$ has the properties that i) $\mathcal{A} \leq \mathcal{B}$ and $\mathcal{B} \leq \mathcal{C}$, results in $\mathcal{A} \leq \mathcal{C}$; and ii) $\mathcal{A} \leq \mathcal{A}$.

[2]Such a mapping could be as follows: $m(\mathbf{x}) = \frac{\mathbf{x}}{\epsilon}$.

$\mathcal{L}(\epsilon)$ is a finite set, i.e. $|\mathcal{L}(\epsilon)| < \infty$. Hence, we have

$$\mathbb{P}\left\{\mathcal{R}\right\} \overset{(a)}{\leq} \mathbb{P}\left\{\bigcup_{\mathbf{x}'\in\mathcal{L}(\epsilon)} (\mathbf{x} \geq \mathbf{x}')\right\} \overset{(b)}{\leq} |\mathcal{L}(\epsilon)| \max_{\mathbf{x}'\in\mathcal{L}(\epsilon)} \mathbb{P}\left\{\mathbf{x} \geq \mathbf{x}'\right\} \overset{(b)}{\dot{\leq}} P^{-\min_{\mathbf{x}\in\mathcal{L}(\epsilon)} f(\mathbf{x})}, \quad \text{(L.1)}$$

where $(a)$ follows from $\mathcal{L}(\epsilon) \leq \mathcal{R}$ and $(b)$ follows from the fact that $|\mathcal{L}(\epsilon)| < \infty$. Now, let us define $h(\epsilon) = \min_{\mathbf{x}\in\mathcal{L}(\epsilon)} f(\mathbf{x})$. For every $\epsilon$, we have $h(\epsilon) < \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})$ (otherwise, according to (L.1), the statement of the Lemma is proved). Now, we prove that $\lim_{\epsilon\to 0} h(\epsilon)$ exists and in fact, we have $\lim_{\epsilon\to 0} h(\epsilon) = \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})$. As $f(\mathbf{x})$ is *uniformly continuous*, there exists a positive function $g(\epsilon)$ ($g(\epsilon) > 0$) such that for all $\mathbf{x}, \mathbf{y}$, $\|\mathbf{x} - \mathbf{y}\| \leq g(\epsilon)$, we have $|f(\mathbf{x}) - f(\mathbf{y})| \leq \epsilon$. Consider any positive constant $\delta > 0$ and any $\epsilon \leq \frac{g(\delta)}{\sqrt{n}}$. According to the definition, for any $\mathbf{x} \in \mathcal{L}(\epsilon)$, there exists a $\mathbf{y} \in \mathcal{R}$ such that $\|\mathbf{x} - \mathbf{y}\| \leq \sqrt{n}\epsilon \leq g(\delta)$. Accordingly, $|f(\mathbf{x}) - f(\mathbf{y})| \leq \delta$. Hence, we have $h(\epsilon) = \min_{\mathbf{x}\in\mathcal{L}(\epsilon)} f(\mathbf{x}) \geq \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x}) - \delta$. On the other hand, we know $h(\epsilon) \leq \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})$. This proves $\lim_{\epsilon\to 0} h(\epsilon) = \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})$. Noticing $\lim_{\epsilon\to 0} h(\epsilon) = \inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})$ and applying (L.1) proves the Lemma.

# Bibliography

[1] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman. A simple cooperative diversity method based on network path selection. *IEEE J. Select. Areas Commun.*, 24(3):659–672, March 2006.

[2] A. El Gamal, M. Mohseni, and S. Zahedi. Bounds on capacity and minimum energy-per-bit for AWGN relay channels. *Information Theory, IEEE Transactions on*, 52:1545–1561, April 2006.

[3] A. Motahari, Sh. Oveis Gharan, and A. K. Khandani. Real interference alignment with real numbers. *Information Theory, IEEE Transactions on*, August 2009. submitted for Publication, available online at http://arxiv.org/pdf/0908.1208/.

[4] A. Motahari, Sh. Oveis Gharan, M. Maddah-Ali, and A. K. Khandani. Real interference alignment: Exploiting the potential of single antenna systems. *Information Theory, IEEE Transactions on*, October 2009. submitted for Publication, available online at http://arxiv.org/pdf/0908.2282/.

[5] A. S. Avestimehr and D. N. C. Tse. Outage capacity of the fading relay channel in the low-SNR regime. *Information Theory, IEEE Transactions on*,

53:1401–1415, April 2007.

[6] A. S. Avestimehr, S. N. Diggavi and D. N. C. Tse. A deterministic approach to wireless relay networks. In *45th Allerton Conference on Communication, Control, and Computing*, 2007.

[7] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse. Wireless network information flow. In *45th Allerton Conference on Communication, Control, and Computing*, 2007.

[8] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse. Approximate capacity of Gaussian relay networks. In *IEEE Int. Symp. Information Theory*, 2008.

[9] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse. Wireless network information flow: a deterministic approach. *Information Theory, IEEE Transactions on*, June 2009. submitted for Publication.

[10] B. Schein and R. G. Gallager. The Gaussian parallel relay network. In *IEEE Int. Symp. Information Theory*, page 22, 2000.

[11] A. Bayesteh and A. K. Khandani. On the user selection for MIMO broadcast channels. *Information Theory, IEEE Transactions on*, 54(3):1086–1107, 2008.

[12] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, New york, 1991.

[13] D. N. C. Tse, P. Viswanath, and L. Zheng. Diversity-multiplexing tradeoff in multiple-access channels. *Information Theory, IEEE Transactions on*, 50(7):1859– 1874, Sept. 2004.

[14] R. Diestel. *Graph Theory*. Springer, 2006.

[15] E. C. van der Meulen. Three-terminal communication channels. *Adv. Appl. Prob.*, 3:120–154, 1971.

[16] A. Edelman. *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, MIT, 1989.

[17] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Pres. Comm.*, 6(3):311–335, March 1998.

[18] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *Information Theory, IEEE Transactions on*, 51:3037–3063, Sept. 2005.

[19] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj. Capacity scaling laws in MIMO relay networks. *IEEE Trans. on Wireless Communications*, 5(6):1433–1444, June 2006.

[20] H. Shi, T. Abe, T. Asai, and H. Yoshino. A relaying scheme using QR decomposition with phase control for MIMO wireless networks. In *IEEE International Conf. on Comm.*, volume 4, pages 2705–2711, 2005.

[21] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *Information Theory, IEEE Transactions on*, 52(9):3936–3964, 2006.

[22] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[23] A. Host-Madsen and A. Nosratinia. The multiplexing gain of wireless networks. In *IEEE International Symposium on Information Theory*, pages 2065–2069, 2005.

[24] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Trans. Tel.*, 10(6):585–595, Nov./Dec. 1999.

[25] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *Information Theory, IEEE Transactions on*, 50(12):3062–3080, Dec. 2004.

[26] K. Azarian, H. El Gamal, and Ph. Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *Information Theory, IEEE Transactions on*, 51(12):4152–4172, Dec. 2005.

[27] K. Sreeram, S. Birenjith and P. Vijay Kumar. DMT of Multi-hop Cooperative Networks - Part I: Basic Results. Available online at http://arxiv.org/pdf/0808.0234, August 2008.

[28] K. Sreeram, S. Birenjith and P. Vijay Kumar. Multi-hop cooperative wireless networks: diversity multiplexing tradeoff and optimal code design. Available online at http://arxiv.org/pdf/0802.1888, February 2008.

[29] L. Zheng and D. N. C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *Information Theory, IEEE Transactions on*, 49:1073– 1096, May 2003.

[30] LL Xie and PR Kumar. An achievable rate for the multiple-level relay channel. *Information Theory, IEEE Transactions on*, 51:1348– 1358, April 2005.

[31] M. Gastpar and M. Vetterli. On the capacity of large Gaussian relay networks. *Information Theory, IEEE Transactions on*, 51:765–779, March 2005.

[32] M. Gastpar, G. Kramer, and P. Gupta. The multiple-relay channel: Coding and antenna-clustering capacity. In *IEEE Int. Symp. Information Theory*, page 136, 2002.

[33] M. Katz and S. Shamai (Shitz). Transmitting to colocated users in wireless ad hoc and sensor networks. *Information Theory, IEEE Transactions on*, 51(10):3540–3563, Oct. 2005.

[34] M. Yuksel and E. Erkip. Cooperative wireless systems: a diversity-multiplexing tradeoff perspective. *Information Theory, IEEE Transactions on*, Aug. 2006. Under Review.

[35] P. Mitran, H. Ochiai, and V. Tarokh. Space-time diversity enhancements using collaborative communications. *Information Theory, IEEE Transactions on*, 51(6):2041–2057, June 2005.

[36] P. Viswanath, D. N. C. Tse, and V. A. Anantharam. Asymptotically optimal water-filling in vector multiple-access channels. *Information Theory, IEEE Transactions on*, 47(1):241–267, Jan. 2001.

[37] Valentin V. Petrov. *Limit Theorems of Probability Theory: Sequences of Indpendent Random Variables.* Oxford University Press, 1995. Page 183.

[38] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler. Fading relay channels: performance limits and space-time signal design. *IEEE J. Select. Areas Commun.*, 22(6):1099–1109, Aug. 2004.

[39] R. U. Nabar, O. Oyman, H. Bolcskei, and A. J. Paulraj. Capacity scaling laws in MIMO wireless networks. In *Allerton Conference on Communication, Control, and Computing*, pages 378–389, 2003.

[40] S. Changiz Rezaei, Sh. Oveis Gharan, and A. K. Khandani. A new achievable rate for the Gaussian parallel relay channel. In *IEEE international conference on Symposium on Information Theory*, pages 194–198, 2009.

[41] S. Pawar, A. S. Avestimehr, and D. N. C. Tse. Diversity-multiplexing tradeoff of the half-duplex relay channel. In *46th Annual Allerton Conference on Communication, Control, and Computing*, pages 27–33, 2008.

[42] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *Information Theory, IEEE Transactions on*, 49(10):2658–2668, Oct. 2003.

[43] S. Zahedi, M. Mohseni, and A. El Gamal. On the capacity of AWGN relay channels with linear relaying functions. In *IEEE Int. Symp. Information Theory*, page 399, 2004.

[44] B. Schein. *Distributed Coordination in Network Information Theory.* PhD thesis, MIT, Sept. 2001.

[45] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley, 1998.

[46] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani. Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network. *Information Theory, IEEE Transactions on*, 2006. revised, available online at http://cst.uwaterloo.ca/.

[47] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani. Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network. In *45th Allerton Conference on Communication, Control, and Computing*, 2007.

[48] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani. Diversity-multiplexing tradeoff in multiple-relay network-part I: proposed scheme and single-antenna networks. In *46th Allerton Conference on Communication, Control, and Computing*, September 2008.

[49] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani. Diversity-multiplexing tradeoff in multiple-relay network-part II: multiple-antenna networks. In *46th Allerton Conference on Communication, Control, and Computing*, September 2008.

[50] Sh. Oveis Gharan, A. Bayesteh and A. K. Khandani. Diversity-Multiplexing Tradeoff in Multi-Antenna Multi-Relay Networks: Improvements and Some Optimality Results. *Information Theory, IEEE Transactions on*, December 2009. submitted for Publication, available online at http://arxiv.org/pdf/0912.3978/.

[51] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani. On the diversity-multiplexing tradeoff in multiple-relay network. *Information Theory, IEEE Transactions on*, 55:5423–5444, December 2009.

[52] Sh. Oveis Gharan and A. K. Khandani. Multiplexing Gain of Amplify-Forward Relaying in Wireless Multi-Antenna Relay Networks. *Information Theory, IEEE Transactions on*, December 2009. submitted for Publication, available online at http://arxiv.org/pdf/0912.3981/.

[53] Sh. Oveis Gharan and A. K. Khandani. Multiplexing Gain of Amplify-Forward Relaying in Wireless Multi-Antenna Relay Networks. In *IEEE International Symposium on Information Theory*, 2010. Accepted for presentation.

[54] Sh. Yang and J.-C. Belfiore. Optimal space-time codes for the MIMO amplify-and-forward cooperative channel. *Information Theory, IEEE Transactions on*, 53(2):647–663, 2007.

[55] Sh. Yang and J.-C. Belfiore. Towards the optimal amplify-and-forward cooperative diversity scheme. *Information Theory, IEEE Transactions on*, 53:3114–3126, Sept. 2007.

[56] T. M. Cover and A. El Gamal. Capacity theorems for the relay channel. *Information Theory, IEEE Transactions on*, 25(5):572–584, Sept. 1979.

[57] A. M. Tulino and S. Verdu. *Random Matrix Theory and Wireless Communications*. Now Publishers Inc, 2004.

[58] W. Yu and J. M. Cioffi. Sum capacity of Gaussian vector broadcast channels. *Information Theory, IEEE Transactions on*, 50:1875–1892, Sept. 2004.