Diversity and multiplexing tradeoff in general MIMO fading channels

by

Lei Zhao

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Electrical Engineering

Program of Study Committee:
Yao Ma, Major Professor
Zhengdao Wang
Huaiqing Wu

Iowa State University

Ames, Iowa

2006

Graduate College
Iowa State University

This is to certify that the master's thesis of

Lei Zhao

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratefulness to those who helped me with various aspects of conducting research and the writing of this thesis.

First and foremost, I want to thank Dr. Yao Ma for his guidance, patience and support throughout this research and the writing of this thesis. His insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education.

I would give my gratitude to Dr. Zhengdao Wang, for his teaching of advanced communications and information theory, which greatly ignited my interests in this area, also for his so many discussions with me when I came to him with my headachy problems.

I would also like to thank Dr. Huaiqing Wu for serving as my minor representative.

I owned my gratefulness to Dr. Krishna Athreya, who gave an excellent course on measure theory and probability and introduced those beautiful mathematics theorems to me, also to Dr. Aleksandar Dogandzic, from whom I learned most of my knowledge on modern signal processing.

My gratitude extends to Jinghua Jin, Jialing Liu, Peng Yu, Mo Wei, Benhong Zhang, Dongbo Zhang, for collaborating and co-authoring papers with me.

Finally, I would like to give my thanks to my parents and my wife for supporting me and making my life a diversified and colorful one.

ABSTRACT

The optimal diversity-multiplexing tradeoff (DMT) has provided a comprehensive view for multiple-input multiple-output (MIMO) systems with multiple transmit and receive antennas. It is widely used as a benchmark to evaluate different space-time schemes in high signal-to-noise ration (SNR) regime. However, previous results depend on the assumption of independent and identical distributed (i.i.d) Rayleigh fading while practical channel models may be more complicated. In this thesis, we investigate for a family of MIMO fading channels the optimal DMT , which is characterized by the near zero and near infinity behaviors of its bounded probability density function (pdf). We believe this family is quite general since it includes i.i.d Rayleigh as well as many other known models as special cases. The progress we have made is built upon our analysis of the probability of the outage set. We find the relation between the parameters of a channel model and the probability of the dominant events in the outage set in high SNR regime. Two different methods are used to obtain the asymptotic outage probability for low and high multiplexing gains. Based on the outage result, we show that optimal DMT can be characterized exactly in a simple piece-wise linear function given long enough channel coherent length while we only provide lower and upper bounds for the optimal DMT when the channel coherent length is short. In the extension section, we consider the effect of correlation on optimal DMT given that the transmit correlation and receive correlation are separable. We also discuss the effect of non-zero-mean channel and derive the optimal DMT result for some specific fading types added with a determinant channel mean.

CHAPTER 1. Introduction

1.1 Literature Review

Recently multiple antenna system has received intense research attention for its potential to combat fading environment. Multiple antennas at both transmitter and receiver ends increase the number of fading paths. The receiver can obtain multiple replicas of one information symbol for detection and therefore the detection error probability is decreased. On the other hand, we could use the extra degrees of freedom brought by the multiple fading paths to transmit independent information streams to increase the transmission bit rate of the system. In high signal-to-noise ration (SNR) regime, the error probability is quantized by diversity gain, i.e. $\lim_{\mathrm{SNR}\to\infty} -\log P_e/\log \mathrm{SNR}$. The maximum transmission rate of the system is given by the ergodic capacity of the MIMO channel.

While traditional design focused on maximizing either the diversity gain or the transmission rate, Zheng and Tse gave a novel asymptotic view of MIMO systems in their seminal work [1] by considering them jointly. Their result is essentially the tradeoff between the error probability and the transmission rate of a MIOM system. From a DMT's point of view, we can say a scheme has a spatial multiplexing gain r and a diversity advantage d, if the rate of the scheme scales like $r \log \mathrm{SNR}$ and the average error probability decays like $1/\mathrm{SNR}^{-d}$. A remarkable result of the paper is the characterization of the optimal DMT for a MIMO systems with n_T transmit antennas and n_R receive antennas under i.i.d Rayleigh fading. It was proved that the optimal diversity gain $d(r)$ given the system operates with a multiplexing gain r is a simple piecewise linear function [1, Fig. 1] shown in Fig 1.1 here.

Following similar ideas, optimal tradeoff curves have been calculated in different scenarios . For example, the DMT for the non-coherent MIMO channel is considered in [2]. The tradeoff

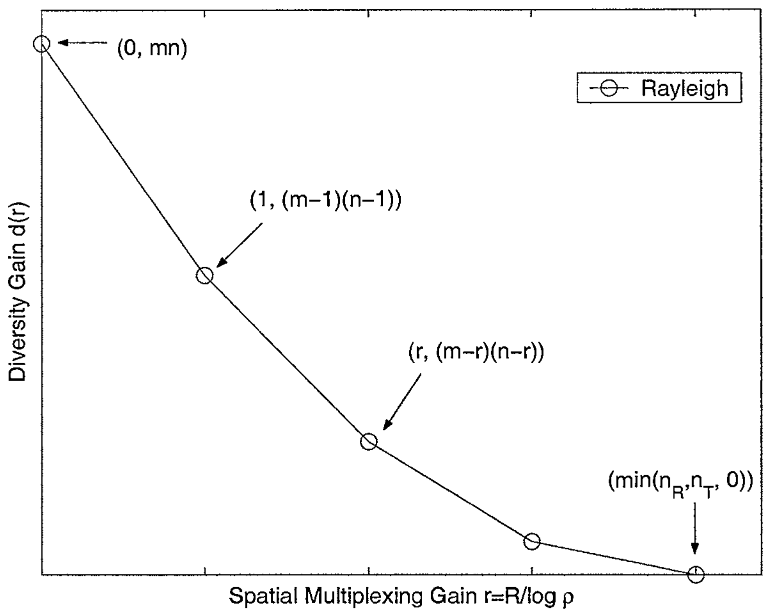

Figure 1.1 Diversity and Multiplexing Tradeoff for i.i.d Rayleigh Fading

result in multiple-access channels is obtained in [3]. In [4], the authors derived the DMT for cooperative wireless systems. the DMT result for multiaccess relay channel is provided in [5]. To determine which point of the optimal DMT curve a MIMO system should operate on, the author of [6] considered an additional end-to-end distortion constraint. The MIMO ARQ Channel is studied from a DMT point of view in [7]. Inner bounds and outer bounds are derived for the tradeoff of a 2-by-2 broadcast/multiple access channels [8].

Popular schemes such as Alamouti, V-BLAST and D-BLAST are evaluated using diversity and multiplexing tradeoff as a metric in [1]. This kind of tradeoff has been serving as a new performance benchmark to compare existing schemes and evaluate new ones [3, 9, 10]. Meanwhile, the lack of optimality of many existing schemes inspires the design of new MIMO schemes that achieve the optimal tradeoff curve [11].

1.2 Motivation

Previous works [1] etc. employed the assumption of i.i.d Rayleigh fading condition to derive the optimal DMT result. In practical MIMO communication scenarios, there exist many channel conditions that cannot be accurately modeled by Rayleigh fading. For example, the line-of-sight micro-cellular communication channels can be modelled by the Rician distribution [12, 13, 14]. A large number of indoor and outdoor mobile communication channels have been shown to be matched with the Nakagami-m distribution [15, 12, 13, 14]. The Weibull distribution also has gained popularity as a versatile channel model for both indoor and outdoor digital communications [16, 17, 18, 19]. Furthermore, due to the size-limitation at the transmit and/or receive antenna arrays the fading correlation arises [20], which substantially affects the achievable performance of MIMO and space-time coded systems [21, 22, 23]. Therefore, it is of theoretical interest and practical importance to develop a technique to enable the calculation of optimal DMT in general fading channel cases.

The various fading types discussed above motivate us to investigate the DMT problem in a general fading channel condition, which includes Rayleigh, Rician, Nakagami-m, Weibull and Nakagami-q fading distributions , and to quantify the effects of correlation, among the MIMO channel elements and the non-zero channel mean.

1.3 Thesis Outline

In Chapter 2, we first introduce the MIMO system model and review some definitions such as scheme, diversity gain, multiplexing gain and others made by Zheng and Tse in [1]. We define the regular fading conditions, which the MIMO channel we study satisfies. We also give examples of existing channel models that meet the regular fading conditions. Next, a brief mathematical prerequisite on random matrix is provided, followed by our discussion on outage diversity. We review the concept of outage and the technique that is used to obtain asymptotic result in [1]. Next several simplifications are made for the derivation of outage probability. We give two heuristic calculations for small multiplexing gain case ($r < 1$)and large multiplexing case ($r \geqslant 1$) to provide some intuitive explanation. We use the outage probability as an lower bound on the detection error probability and construct an upper bound using i.i.d Gaussain ensemble as the input codebook. Based on the coherent block length of the channel, we give either exact optimal DMT curve or lower and upper bounds for general fading. At the end of the chapter, we illustrate the difference between our new result and the previous i.i.d Rayleigh result.

In Chapter 3, the optimal DMT result obtained in Chapter 2 is extended to an even more general scenario. First, we show that Nakagami-q can be included in our general model. Next, the effect of channel correlations are discussed. We also consider the effect and non-zero mean and provide an explicit optimal DMT for some existing channel models. The joint effect of channel mean and channel correlation is investigated at the end of the chapter and an explicit optimal DMT result is summarized in a corollary.

We make the conclusions in Chapter 4 and summarize the contributions of this thesis.

The result in this thesis is also reported in [28] and [29]

1.4 Notations

Bold faced letters denote random variables, vectors, or matrices; plain letters denote the corresponding realizations or constants; I_m denotes $m \times m$ identity matrix; superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^\dagger$ denote scalar complex conjugate, vector and matrix transpose and conjugate transpose, respectively. $||\cdot||_F$ denotes the Frobenius norm. Unless otherwise indicated, the (i,j)th entry of a matrix H is denoted by h_{ij} or $[H]_{ij}$.

We also list two tables for convenience of reading this thesis:

DMT	diversity-multiplexing tradeoff
pdf	probability density function
MIMO	multiple-input multiple-output
SNR	signal-to-noise ratio
i.i.d.	independent and identical distributed

Table 1.1 Abbreviation Table.

$\boldsymbol{X}$	transmitted codeword
$\boldsymbol{H}$	channel matrix
$\boldsymbol{W}$	additive white Gaussian noise
ρ	normalized transmit power
n_T	number of transmit antennas
n_R	number of receive antennas
l	channel coherent block length
$\boldsymbol{\lambda}_i$	the ith ordered eigenvalue of $\boldsymbol{H}\boldsymbol{H}^\dagger$ or $\boldsymbol{H}^\dagger\boldsymbol{H}$
$\boldsymbol{\lambda}$	$[\boldsymbol{\lambda}_1, ..., \boldsymbol{\lambda}_{\min\{n_T,n_R\}}]$
$\boldsymbol{\alpha}_i$	$\boldsymbol{\alpha}_i = -\log \boldsymbol{\lambda}_i / \log \rho$
r	multiplexing gain
$d^*(r)$	the supreme of all the scheme achieving r
$d_{\text{out}}(r)$	outage diversity for multiplexing gain r

Table 1.2 Frequently used variables

CHAPTER 2. Diversity and Multiplexing Tradeoff for Independent Fading Channels

2.1 System Model

We consider a wireless link with n_T transmit and n_R receive antennas. The fading coefficient $\boldsymbol{h}_{ij}$ is the complex path gain from transmit antenna j to receive antenna i. We assume that the coefficients are independent random variables, and write $\boldsymbol{H} = [\boldsymbol{h}_{ij}] \in \mathcal{C}^{n_R \times n_T}$. More general cases will be explored in Section 3. $\boldsymbol{H}$ is assumed to be known at the receiver, but unknown at the transmitter. We also assume that $\boldsymbol{H}$ remains constant within a block of l symbols, then the received data within one block can be written as

$$\boldsymbol{Y} = \sqrt{\frac{\rho}{n_T}} \boldsymbol{H}\boldsymbol{X} + \boldsymbol{W}, \tag{2.1}$$

where $\boldsymbol{X} \in \mathcal{C}^{n_T \times l}$ is the transmitted codeword, the additive noise matrix $\boldsymbol{W}$ has independent circular symmetric complex Gaussian distributed entries $\boldsymbol{w}_{ij} \sim \mathcal{CN}(0,1)$, and ρ controls the signal-to-noise ratio (SNR).

We assume that $\forall\, i, j$, the pdf of channel $\boldsymbol{h}_{ij}$, i.e. $p_{\boldsymbol{h}_{ij}}(h)$, satisfies the following conditions[1]:

1. *near 0 behavior:* $|h| \to 0$, $p_{\boldsymbol{h}_{ij}}(h) = \Theta(|h|^{t_{ij}})$, $t_{ij} \geqslant 0$
2. *near ∞ behavior:* $|h| \to \infty$, $p_{\boldsymbol{h}_{ij}}(h) = O(e^{-b_{ij}|h|^{\beta_{ij}}})$, $b_{ij} > 0$ and $\beta_{ij} > 0$
3. $p_{\boldsymbol{h}_{ij}}(h)$ is bounded by a constant K.

We call channels that satisfy the above three conditions *regular*. This general model is valid for many frequently used fading channels, including complex Gaussian, whose amplitude follows Rayleigh or Rician distribution, as well as Nakagami-m ($m \geqslant 1$), Weibull ($\eta \geqslant 2$), and

[1] We use $f(x) = \Theta(g(x))$ to mean there exits positive c_1, c_2, such that $c_1 g(x) \leqslant f(x) \leqslant c_2 g(x)$ and use $f(x) = O(g(x))$ to mean there exits positive c_2, such that $f(x) \leqslant c_2 g(x)$

	pdf	t	b	β
Rayleigh	$\frac{1}{\pi\Omega}e^{-\frac{\lvert h\rvert^2}{\Omega}}$	0	$\frac{1}{\Omega}-\psi$	2
Rician	$\frac{1}{\pi\Omega}e^{-\frac{\lvert h-\mu\rvert^2}{\Omega}}$	0	$\frac{1}{\Omega}-\psi$	2
Nakagami-m	$\frac{m^m\lvert h\rvert^{2m-2}}{\pi\Omega^m\Gamma(m)}e^{-\frac{m\lvert h\rvert^2}{\Omega}}$	$2m-2$	$\frac{m}{\Omega}-\psi$	2
Weibull	$\frac{\eta\Omega^{-\eta}}{2\pi}\lvert h\rvert^{\eta-2}e^{-(\frac{\lvert h\rvert}{\Omega})^\eta}$	$\eta-2$	$\Omega^{-\eta}-\psi$	η
Nakagami-q [2]	$\frac{(1+q^2)}{2\pi q\Omega}I_0(\frac{(1-q^4)\lvert h\rvert^2}{4q^2\Omega})e^{-\frac{(1+q^2)^2}{4q^2\Omega}\lvert h\rvert^2}$	0	$\frac{(1+q^2)}{2\Omega}-\psi$	2

ψ can be any positive number which makes parameter $b > 0$.
For example, for Nakagami-m, ψ can be $m/\Omega/2$.

Table 2.1 Pdf of different fading channels.

Nakagami-q ($0 < q \leqslant 1$) with the assumption that the phase is uniformly distributed over $[0, 2\pi)$ and independent from the amplitude. The corresponding parameters t, b and β are listed in Table 2.1. Ω is the channel variance, except for Weibull, for which case, the channel variance is $\Omega^2\Gamma(1+2/\eta)$, where $\Gamma(\cdot)$ is the gamma function.

Since $\boldsymbol{H}$ has independent elements, the pdf of $\boldsymbol{H}$ can be written as

$$p_{\boldsymbol{H}}(H) = \prod_{i=1}^{n_R}\prod_{j=1}^{n_T} p_{\boldsymbol{h}_{ij}}([H]_{ij}) \tag{2.2}$$

A codebook $\mathcal{C}$ of rate R with unit bits per channel use has $|\mathcal{C}| = \lfloor 2^{Rl} \rfloor$ codewords $\{X(1), ..., X(|\mathcal{C}|)\}$. We further assume a power constraint on the codebook $\mathcal{C}$:

$$\frac{1}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|} ||X(i)||_F^2 \leqslant n_T l, \tag{2.3}$$

so that ρ in (2.1) is the average transmit power, regardless of the value of n_T.

The following definitions were introduced in [1]:

Definition 1: The symbol $\doteq$ denotes *exponential equality*, i.e., we write $f(\rho) \doteq \rho^b$ to denote

$$\lim_{\rho\to\infty}\frac{\log f(\rho)}{\log \rho} = b, \tag{2.4}$$

and $\dot{\geqslant}$, $\dot{\leqslant}$ are similarly defined.

Definition 2: A *scheme* is a family of codes $\{\mathcal{C}(\rho)\}$ of block length l, one at each transmit power level. $R(\rho)$ is the rate of the code $\mathcal{C}(\rho)$.

[2] Notice that $1 \leqslant I_0(x) = \sum_{k=0}^{\infty}\frac{(x/2)^{2k}}{k!k!} \leqslant \left(\sum_{k=0}^{\infty}\frac{(x/2)^k}{k!}\right)^2 = e^x$.

Definition 3: A scheme $\{\mathcal{C}(\rho)\}$ is said to achieve *spatial multiplexing gain* r and *diversity gain* d if the data rate $R(\rho)$ and the average error probability $P_e(\rho)$ satisfy the following equalities, respectively,

$$\lim_{\rho\to\infty} \frac{R(\rho)}{\log\rho} = r, \tag{2.5}$$

$$P_e(\rho) \doteq \rho^{-d}. \tag{2.6}$$

For each r, we define $d^*(r)$ as the supremum of the diversity gain achieved over all schemes.

2.2 Mathematical Prerequisite on Random Matrices

In this section, we present some important results on random matrices without proof. Interested readers can refer to [24] and [25].

For a nonsingular matrix $\boldsymbol{H} \in \mathcal{C}^{m\times n}$ $(m \leqslant n)$ with pdf $p_{\boldsymbol{H}}(H)$, there is a unique LQ factorization

$$\boldsymbol{H} = \boldsymbol{L}\boldsymbol{Q}, \tag{2.7}$$

where $\boldsymbol{L} \in \mathcal{C}^{m\times m}$ is upper-triangular with positive diagonal elements and $\boldsymbol{Q} \in \mathcal{C}^{m\times n}$ with $\boldsymbol{Q}\boldsymbol{Q}^\dagger = I_m$. If $\boldsymbol{H}$ is random, the pdf of $\boldsymbol{L}$ can be written as

$$p_{\boldsymbol{L}}(L) = \prod_{i=1}^{m} l_{ii}^{2n-2i+1} \cdot \int_{V_{m,n}} p_{\boldsymbol{H}}(LQ)dQ, \tag{2.8}$$

where $V_{m,n}$ is the complex Stiefel manifold, i.e., a sub-manifold of m by n complex matrices $\boldsymbol{Q}$ such that $\boldsymbol{Q}\boldsymbol{Q}^\dagger = I_m$, and the dimension of $V_{m,n}$ is $2mn - m^2$.

Let

$$\boldsymbol{W} \equiv \boldsymbol{H}\boldsymbol{H}^\dagger = \boldsymbol{L}\boldsymbol{L}^\dagger, \tag{2.9}$$

then the pdf of $\boldsymbol{W}$ is

$$\begin{aligned} p_{\boldsymbol{W}}(W) &= \left.\left(2^{-m}\prod_{i=1}^{m} l_{ii}^{2i-2m-1} p_{\boldsymbol{L}}(L)\right)\right|_{LL^\dagger=W} \\ &= \left.\left(2^{-m}|\det W|^{n-m} \cdot \int_{V_{m,n}} p_H(LQ)dQ\right)\right|_{LL^\dagger=W}. \end{aligned} \tag{2.10}$$

Since $\boldsymbol{H}$ is nonsingular with probability one, $\boldsymbol{W} \in \mathcal{C}^{m\times m}$ has full rank of m. Thus, there is a unique eigenvalue decomposition of $\boldsymbol{W}$ as

$$\boldsymbol{W} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\dagger}, \tag{2.11}$$

if we assume the diagonal elements of $\boldsymbol{\Lambda}$ are ordered non-decreasingly and the first row of the unitary matrix $\boldsymbol{U}$ is real and non-negative. Then, we have the pdf of $\boldsymbol{\Lambda}$ as

$$p_{\boldsymbol{\Lambda}}(\Lambda) = \frac{1}{(2\pi)^m}\prod_{i<j}(\lambda_i - \lambda_j)^2 \cdot \int_{V_{m,m}} p_{\boldsymbol{W}}(U\Lambda U^{\dagger})dU. \tag{2.12}$$

The factor $\frac{1}{(2\pi)^m}$ comes from the fact that we assume the first row of U is non-negative and integrate U over the whole manifold $V_{m,m}$.

We know that

$$\boldsymbol{L}\boldsymbol{L}^{\dagger} = \left(\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\right)\left(\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\right)^{\dagger},$$

and $\boldsymbol{L}$ is unique for a given $\boldsymbol{H}$, hence there is a unique unitary matrix $\boldsymbol{Q}_1 \in \mathcal{C}^{m\times m}$ for the given $\boldsymbol{H}$ such that

$$\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{Q}_1. \tag{2.13}$$

Combining (2.10), (2.12) and (2.13), we obtain

$$\begin{aligned}
p_{\boldsymbol{\Lambda}}(\Lambda) &= \frac{1}{(4\pi)^m}\left(\prod_{i=1}^{m}\lambda_i^{n-m}\right)\cdot\prod_{i<j}(\lambda_i - \lambda_j)^2 \cdot \int_{V_{m,m}}\int_{V_{m,n}} p_H(U\Lambda^{1/2}Q_1Q)dQdU \\
&= \frac{1}{(4\pi)^m}\left(\prod_{i=1}^{m}\lambda_i^{n-m}\right)\cdot\prod_{i<j}(\lambda_i - \lambda_j)^2 \cdot \int_{V_{m,m}}\int_{V_{m,n}} p_H(U\Lambda^{1/2}Q)dQdU,
\end{aligned} \tag{2.14}$$

the second equality follows from the fact that the measure defined by dQ is invariant under unitary transformations.

2.3 Outage Formulation

Consider a non-ergodic fading channel model

$$\boldsymbol{y}_i = \sqrt{\frac{\rho}{n_T}}\boldsymbol{H}\boldsymbol{x}_i + \boldsymbol{w}_i, \quad \text{for } i = 1, 2, \ldots, \infty, \tag{2.15}$$

where $\boldsymbol{x}_i \in \mathcal{C}^{n_T}$, $\boldsymbol{y}_i \in \mathcal{C}^{n_R}$, and $\boldsymbol{w}_i \in \mathcal{C}^{n_R}$ are the transmitted data, received data, and additive white Gaussian noise at time i. The channel matrix $\boldsymbol{H}$ is chosen randomly but is held fixed for all time.

outage is defined as the event that the mutual information of this channel does not support a target data rate R [26]. Without loss of optimality, the input distribution can be taken to be complex Gaussian with a covariance matrix C_{xx}, then

$$I(\boldsymbol{x}_i; \boldsymbol{y}_i | \boldsymbol{H} = H) = \log\det\left(I_{n_R} + \frac{\rho}{n_T} H C_{xx} H^{\dagger}\right). \tag{2.16}$$

Optimizing over all input distributions that satisfy the average power constraint, the outage probability is

$$P_{\text{out}}(R) = \inf_{C_{xx} \geqslant 0, \text{Tr}(C_{xx}) \leqslant n_T} P\left[\log\det\left(I_{n_R} + \frac{\rho}{n_T} \boldsymbol{H} C_{xx} \boldsymbol{H}^{\dagger}\right) < R\right], \tag{2.17}$$

where the probability is taken over the random channel matrix $\boldsymbol{H}$. It is shown in [1] that

$$P_{\text{out}}(R) \doteq P\left[\log\det(I_{n_R} + \rho \boldsymbol{H}\boldsymbol{H}^{\dagger}) < R\right]. \tag{2.18}$$

We assume, without loss of generality, that $n_R \leqslant n_T$. This is because

$$\log\det\left(I_{n_R} + \frac{\rho}{n_T} \boldsymbol{H}\boldsymbol{H}^{\dagger}\right) = \log\det\left(I_{n_T} + \frac{\rho}{n_T} \boldsymbol{H}^{\dagger}\boldsymbol{H}\right).$$

hence, swapping n_R and n_T has no effect on the mutual information, except a scaling factor of n_T/n_R on ρ, which can be ignored on the scale of interest.

Since the elements of $\boldsymbol{H}$ are independent, $\boldsymbol{H}$ has full rank of m with probability one. Let $\boldsymbol{\lambda}_1 \leqslant \boldsymbol{\lambda}_2 \leqslant \ldots \leqslant \boldsymbol{\lambda}_{n_R}$ be the nonzero eigenvalues of $\boldsymbol{W}$, data rate be $R = r\log\rho$ (which means a multiplexing gain r), we have

$$P_{\text{out}}(r\log\rho) \doteq P\left[\log\det(I_{n_R} + \rho \boldsymbol{H}\boldsymbol{H}^{\dagger}) < R\right] = P\left[\prod_{i=1}^{n_R}(1 + \rho\lambda_i) < \rho^r\right].$$

Denotes $\boldsymbol{\alpha}_i = -\log\boldsymbol{\lambda}_i/\log\rho$ and $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{n_R}]$. It is shown in [1] that

$$P_{\text{out}}(r\log\rho) \doteq P\left[\sum_i (1 - \alpha_i)^+ < r\right] = \int_{\mathcal{A}(r)} p_{\boldsymbol{\alpha}}(\alpha)d\alpha, ^3 \tag{2.19}$$

[3]This $\doteq$ is rigorous as used in [1] since $\rho^{(1-\alpha_i)^+} \leqslant 1 + \rho^{1-\alpha_i} \leqslant 2\rho^{(1-\alpha_i)^+}$.

where $(x)^+$ denotes $\max\{0, x\}$ and $\mathcal{A}(r) = \{\alpha : \sum_i (1-\alpha_i)^+ < r\}$.

$$p_{\boldsymbol{\alpha}}(\alpha) = \left(\frac{\log\rho}{4\pi}\right)^{n_R} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i}\right) \cdot \prod_{i<j} (\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \cdot \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} p_{\boldsymbol{H}}(UDQ)dQdU, \tag{2.20}$$

Using (2.14) and change of variables, we obtain the joint pdf of $\boldsymbol{\alpha}$ as

$$p_{\boldsymbol{\alpha}}(\alpha) = \left(\frac{\log\rho}{4\pi}\right)^{n_R} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i}\right) \cdot \prod_{i<j} (\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \cdot \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} p_{\boldsymbol{H}}(UDQ)dQdU, \tag{2.21}$$

where $D = \text{diag}[\rho^{-\alpha_1/2}, \ldots, \rho^{-\alpha_{n_R}/2}]$. So, we have

$$\begin{aligned} P_{\text{out}}(r\log\rho) \doteq \int_{\mathcal{A}(r)} \left(\frac{\log\rho}{4\pi}\right)^{n_R} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i}\right) \cdot \prod_{i<j} (\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \\ \cdot \left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} p_{\boldsymbol{H}}(UDQ)dQdU\right] d\alpha. \end{aligned} \tag{2.22}$$

Since $P_{\text{out}}(r\log\rho) \to 0$ as $\rho \to \infty$, and we are only interested in the ρ exponent of P_{out}, i.e.,

$$\lim_{\rho\to\infty} \frac{\log P_{\text{out}}(r\log\rho)}{\log\rho}, \tag{2.23}$$

thus we can make some approximations to simplify the integral.

First, the term $\left(\frac{\log\rho}{4\pi}\right)^{n_R}$ has no effect on the ρ exponent, since

$$\lim_{\rho\to\infty} \frac{\log\left(\frac{\log\rho}{4\pi}\right)^{n_R}}{\log\rho} = 0. \tag{2.24}$$

Second, we can ignore the outer integral over the range with any $\alpha_i < 0$ when $\rho \to \infty$, and replace the outer integral range $\mathcal{A}(r)$ with $\mathcal{A}'(r) = \mathcal{A}(r) \bigcap \mathcal{R}^{n^+}$, where $\mathcal{R}^{n^+}$ is the set of real n-vectors with nonnegative elements. To see this, notice that we can assume that $\alpha_{n_R} < 0$ w.l.o.g., for the case with any $\alpha_i < 0$ and

$$\sum_{i,j} |h_{ij}|^2 = \text{tr}(HH^\dagger) = \sum_{i=1}^{n_R} \rho^{-\alpha_i} \geqslant \rho^{-\alpha_{n_R}}, \tag{2.25}$$

thus at least for one path $|h_{ij}| \geqslant \frac{1}{n_R n_T} \rho^{-\alpha_{n_R}/2}$, whose pdf can be upper bounded by an exponential term $e^{-\min\{b_{ij}\}\rho^{-\alpha_{n_R}\max\{\beta_{ij}\}/2}}$ when $\rho \to \infty$, which decays with ρ exponentially. A rigorous proof is provided in Appendix A.3.

As we will see later, the dominant part of the outage event is different for $0 \leqslant r < 1$ and $r \geq 1$. While a rigorous proof is provided in the appendix, we make a heuristic calculation here for some intuitive explanation.

2.3.1 Outage diversity when $0 \leq r < 1$

As stated before, we only need to consider $\mathcal{A}'(r)$. When $r \leq 1$, since outage occurs when $\sum_i (1-\boldsymbol{\alpha}_i)^+ < r$, it is obvious that $\boldsymbol{\alpha}_i > 1-r$. Thus every eigenvalue $\boldsymbol{\lambda}_i = \rho^{-\boldsymbol{\alpha}_i} < \rho^{-(1-r)}$ and goes to 0 as $\rho \to \infty$, which implies $|\boldsymbol{h}_{ij}| \to 0$ as $\rho \to \infty$. This means the outage occurs when all the elements of $\boldsymbol{H}$ is small, so we can use the near 0 behavior of the pdf of $\boldsymbol{H}$ to calculate the outage diversity:

$$p_{\boldsymbol{H}}(H)\Big|_{H=UDQ} = \Theta\left(\prod_{i,j} |h_{ij}|^{t_{ij}}\right)\Bigg|_{h_{ij}=[UDQ]_{ij}} \tag{2.26}$$

Notice that $[UDQ]_{ij}$ is a polynomial of ρ, where the coefficient are function of Q and U. The term with the highest order is $\rho^{-\alpha_{n_R}/2}$ and its corresponding coefficient is nonzero almost everywhere. We can approximate $|h_{ij}|^{t_{ij}}$ by $\rho^{-\alpha_{n_R} t_{ij}/2}$.

For the polynomial in the Jacobian part, $\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i} \cdot \prod_{i<j} (\rho^{-\alpha_i} - \rho^{-\alpha_j})^2$, we only need to consider the case that α_i's are distinct, since otherwise the integrand is zero. In this case, $|\rho^{-\alpha_i} - \rho^{-\alpha_j}|$ is dominated by $\rho^{-\alpha_j}$ for any $i<j$. Therefore, we can approximate this polynomial by $\rho^{-\sum_{i=1}^{n_R}(n_T-n_R+2i-1)\alpha_i}$ and approximate the outage probability by:

$$P_{\text{out}}(r \log \rho) \doteq \int_{\mathcal{A}'(r)} \rho^{-\sum_{i=1}^{n_R}(n_T-n_R+2i-1)\alpha_i} \rho^{-\alpha_{n_R}\sum_{i,j} t_{ij}/2} d\alpha. \tag{2.27}$$

As $\rho \to \infty$, the integral is dominated by the term with the largest ρ exponent:

$$-\inf_{\alpha \in \mathcal{A}'(r)} \sum_{i=1}^{n_R} (n_T - n_R + 2i - 1)\alpha_i + \alpha_{n_R} \sum_{i,j} t_{ij}/2$$

2.3.2 Outage diversity when $r \geq 1$

First, let us consider a simple upper bound on the outage probability:

$$\begin{aligned} P_{\text{out}}(r\log\rho) &\dot{\leqslant} \int_{\mathcal{A}(r)} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i}\right) \cdot \prod_{i<j}(\rho^{-\alpha_i}-\rho^{-\alpha_j})^2 \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} K^{n_R n_T} dQ dU d\alpha. \\ &\dot{\leqslant} \int_{\mathcal{A}'(r)} \rho^{-\sum_{i=1}^{n_R}(n_T-n_R+2i-1)\alpha_i} d\alpha. \end{aligned} \tag{2.28}$$

Compared with (2.27), it is obvious that this upper bound is not tight for $r < 1$. The reason is that using the upper bound $K^{n_R n_T}$ on the pdf of $\boldsymbol{H}$ ignores the fact the $p_{\boldsymbol{H}}(H)$ goes to 0 when $\rho \to \infty$ and $r < 1$. However, when $r \geqslant 1$, as ρ goes to infinity, the outage events will include the neighborhood of H with the structure[4] $\left(U \operatorname{diag}[0, ..., 0, \lambda_{n_R-\lfloor r\rfloor+1}^{1/2}, ..., \lambda_{n_R}^{1/2}]Q\right)$. By choosing proper U, Q and $(\lambda_{n_R-\lfloor r\rfloor+1}^{1/2}, ..., \lambda_{n_R}^{1/2})$, we can find that the pdf of this kind of H is bounded away from 0. And the upper bound (2.28) is tight in the ρ exponent sense. Again as $\rho \to \infty$, the integral is dominated by the term with the largest ρ exponent:

$$-\inf_{\alpha \in \mathcal{A}'(r)} \sum_{i=1}^{n_R}(n_T - n_R + 2i - 1)\alpha_i$$

Notice the above ρ exponent does not depends on parameters t_{ij}, b_{ij} and β_{ij}, and is only a function of r, n_R and n_T. Furthermore, it is the same as that of i.i.d Rayleigh channel considered in [1, Theorem 4], thus the outage diversity for $r \geqslant 1$ is the same for all the general fading channel models we consider.

Roughly summarizing, when r is small ($r < 1$), outage only occurs when all the elements of H are small so that parameters t_{ij} take effect; When $r \geqslant 1$, outage still occurs when all elements of H are small. But such outage events are no longer dominant, because in this case outage can also occur when all elements of H are bounded away from 0, which are actually the dominant error events because they are more probable than the case with all elements of H close to zero. The probability of such dominant error events does not depend on the behavior of the pdf of the elements of H near zero, specifically the parameters t_{ij}, provided the pdf is bounded, but only depends on the Jacobian between the elements of H and the eigenvalues

[4] $\lfloor x\rfloor$ is the largest integer no greater than x and $\lceil x\rceil$ is the smallest integer no less than x.

of H. The Jacobian in turn only depends on the system dimensions, n_T and n_R, and not the channel statistical characteristics.

It is of interest to notice that when $n_R \leqslant n_T$ [c.f. Appendix B]

$$\inf_{\alpha\in\mathcal{A}'(r)} \sum_{i=1}^{n_R}(n_T - n_R + 2i - 1)\alpha_i + \frac{\sum_{i,j} t_{ij}}{2}\alpha_{n_R} = \inf_{\alpha\in\mathcal{A}'(r)} \sum_{i=1}^{n_R}(n_T - n_R + 2i - 1)\alpha_i, \text{ when } r \geqslant 1$$

since the optimization will result in $\alpha_{n_R} = 0$ for both termes. So we can unify our result in the following theorem with a proof provided in Appendix A.1.

Theorem 1 *(Outage Probability) For the MIMO channel defined in* (2.1) *and* (2.2)*, let* $m = \min\{n_R, n_T\}$*, and the data rate be* $R = r\log\rho$*, with* $0 \leqslant r \leqslant m$*. The outage probability satisfies*

$$P_{out}(R) \doteq \rho^{-d_{out}(r)}, \tag{2.29}$$

where

$$d_{out}(r) = f(\alpha^*) = \inf_{\alpha\in\mathcal{A}'(r)} f(\alpha), \tag{2.30}$$

and

$$\begin{aligned} f(\alpha) &= \sum_{i=1}^{m}(|n_T - n_R| + 2i - 1)\alpha_i + \frac{\sum_{i=1}^{n_R}\sum_{j=1}^{n_T} t_{ij}}{2}\alpha_m, \\ \mathcal{A}'(r) &= \Big\{\alpha : \alpha_1 \geqslant \cdots \geqslant \alpha_m \geqslant 0, \sum_{i=1}^{m}(1-\alpha_i)^+ < r\Big\}. \end{aligned} \tag{2.31}$$

$d_{out}(r)$*, shown in Fig 2.1, is given by the piecewise-linear function connecting the points* $(k, d^*(k))$*, for* $k = 0, 1, \ldots, n$*, where*

$$d^*(k) = \begin{cases} (n_R - k)(n_T - k) & \text{if } k = 1, \ldots, n, \\ \sum_{i=1}^{n_R}\sum_{j=1}^{n_T}(1 + \frac{t_{ij}}{2}) & \text{if } k = 0. \end{cases} \tag{2.32}$$

2.4 Optimal Tradeoff Curve

The outage probability provides a lower bound on the average error probability for channel defined in (2.1), which is proved by [1, Lemma 5] for i.i.d Rayleigh fading case. Based on Theorem 1, this result still holds for the general channel model defined in this paper. For convenience, we restate it as follows without proof:

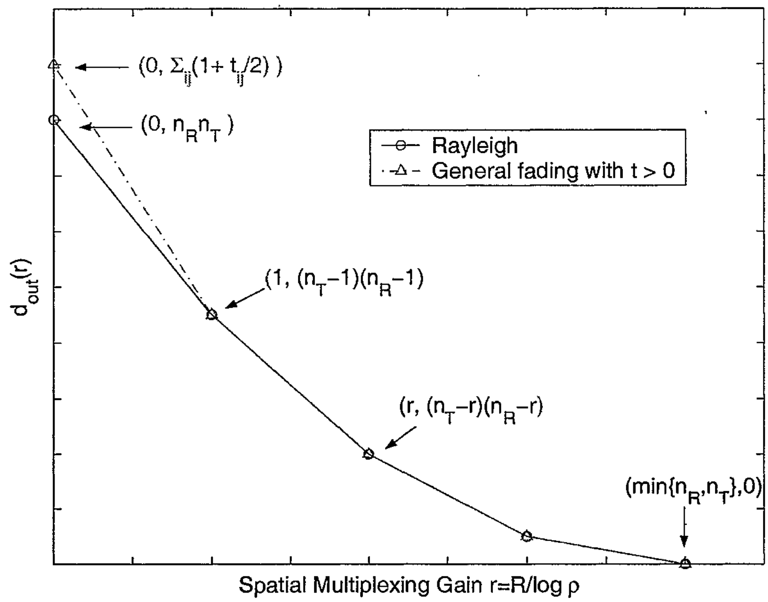

Figure 2.1 Outage diversity for Rayleigh and General Fading

Lemma 1 *(Outage Bound)[1] For the channel defined in* (2.1) *and* (2.2)*, let the data rate be scaled as* $R = r\log\rho$*. For any coding scheme, the average error probability is lower-bounded as*

$$P_e(\rho) \dot{\geqslant} \rho^{-d_{out}(r)}, \tag{2.33}$$

where $d_{out}(r)$ *is defined in* (2.30).

With Lemma 1 providing a lower bound on the average error probability, to obtain the ρ exponent of P_e, we only need to derive an upper bound on P_e (a lower bound on the optimal diversity gain).

Consider at data rate $R = r\log\rho$,

$$\begin{aligned} P_e(\rho) =& P_{\text{out}}(R)P(\text{error}|\text{outage}) + P(\text{error, no outage}) \\ \leqslant& P_{\text{out}}(R) + P(\text{error, no outage}). \end{aligned} \tag{2.34}$$

By choosing the input to be the random code from the i.i.d Gaussian ensemble, the second term in (2.34) can be upper-bounded via a union bound [1]:

$$P(\text{error, no outage}) \dot{\leqslant} \int_{\mathcal{A}^c(r)} p_{\alpha}(\alpha)\rho^{-l[\sum_{i=1}^{\min\{n_R,n_T\}}(1-\alpha_i)^+ - r]} d\alpha, \tag{2.35}$$

where $\mathcal{A}^c(r)$ is the complementary set of $\mathcal{A}(r)$.

Using the similar argument used in the outage probability, we can change the integral region to $\mathcal{A}^c(r)\bigcap R^{\min\{n_R,n_T\}+}$ and approximate ρ exponent of the integral, which is stated in the following lemma.

Lemma 2 *For the MIMO channel defined in* (2.1) *and* (2.2)*, let* $m = \min\{n_R, n_T\}$*, and the data rate be* $R = r\log\rho$*, with* $0 \leqslant r \leqslant n$*. The average error probability when no channel outage occurs satisfies:*

$$P(\mathit{error,\ no\ outage}) \dot{\leqslant} \rho^{-d_G(r)}, \tag{2.36}$$

where

$$d_G(r) = d_G(r,\alpha^*) = \inf_{\alpha\in\mathcal{A}^c(r)\bigcap\mathcal{R}^{m+}} d_G(r,\alpha), \tag{2.37}$$

and

$$\begin{aligned} d_G(r,\alpha) &= \sum_{i=1}^{m}(2i-1+|n_T-n_R|)\alpha_i + \frac{\sum_{i,j} t_{ij}}{2}\alpha_m + l\Big(\sum_{i=1}^{m}(1-\alpha_i)^+ - r\Big), \\ \mathcal{A}^c(r)\bigcap\mathcal{R}^{m^+} &= \left\{\alpha : \alpha_1 \geqslant \ldots \geqslant \alpha_m \geqslant 0, \sum_{i=1}^{m}(1-\alpha_i)^+ \geqslant r\right\}. \end{aligned} \tag{2.38}$$

It can be verified that [c.f. [1, Section IV-A]] when $l \geqslant n_T + n_R - 1 + \sum_{i,j} t_{ij}/2$, $d_G(r)$ agrees with $d_{\text{out}}(r)$; when $l < n_T + n_R - 1 + \sum_{i,j} t_{ij}/2$, $d_G(r) = d_{\text{out}}(r)$ for $m - k_1 < r \leqslant m$, while $d_G(r)$ is linear with slope $-l$ and is strictly below $d_{\text{out}}(r)$ for $0 \leqslant r \leqslant m - k_1$, where

$$k_1 = \min\left\{\left\lceil\frac{l-|n_R-n_T|-1}{2}\right\rceil, \min\{n_R,n_T\}-1\right\}. \tag{2.39}$$

For convenience, we call a system with n_T transmit, n_R receive antennas, and a block length l as an (n_T, n_R, l) system. Combining the results from Theorem 1, Lemma 1, and Lemma 2, we can see that given a multiplexing gain r, the optimal diversity gain is bounded by $\min\{d_G(r), d_{\text{out}}(r)\} \leqslant d^*(r) \leqslant d_{\text{out}}(r)$. Whether the optimal tradeoff curve $d^*(r)$ can be exactly characterized depends on the relation among n_T, n_R, l and t_{ij} as we conclude in Theorem 2 and Theorem 3.

Theorem 2 *For an* (n_T, n_R, l) *system with independent and regular fading pathes and* $l \geqslant n_T + n_R - 1 + \sum_{i,j} t_{ij}/2$*, the optimal tradeoff curve* $d^*(r)$ *is given by the piecewise-linear* *function connecting the points* $(k, d^*(k))$, $k = 0, 1, ..., \min\{n_R, n_T\}$, *where*

$$d^*(k) = \begin{cases} (n_R - k)(n_T - k) & \text{if } k = 1, \ldots, \min\{n_R, n_T\}, \\ \sum_{i=1}^{n_R} \sum_{j=1}^{n_T} (1 + \frac{t_{ij}}{2}) & \text{if } k = 0. \end{cases} \tag{2.40}$$

In particular, the maximum diversity gain $d^*_{\max} = \sum_{i=1}^{n_R} \sum_{j=1}^{n_T} (1 + \frac{t_{ij}}{2})$ *and the maximum multiplexing gain* $r^*_{\max} = \min\{n_R, n_T\}$.

Theorem 3 *For an* (n_T, n_R, l) *system with independent and regular fading pathes and* $l < n_T + n_R - 1 + \sum_{i,j} t_{ij}/2$, *the optimal tradeoff curve* $d^*(r)$ *is upper bounded by* $d_{out}(r)$ *and lower bounded by* $d_G(r)$. *With* k_1 *provided by eq. (2.39), for* $\min\{n_R, n_T\} - k_1 < r \leqslant \min\{n_R, n_T\}$, $d_G(r)$ *agrees with the upper bound* $d_{out}(r)$ *; For* $0 \leqslant r \leqslant \min\{n_R, n_T\} - k_1$, $d_G(r)$ *is linear with slope* $-l$ *and is strictly below* $d_{out}(r)$.

Remark: It is interesting to notice that the optimal tradeoff curve (or bounds) only depends on parameter t_{ij} and not on b_{ij}, β_{ij} and K.

As an example, in Figs. 2.2 – 2.4, tradeoff curves are plotted for systems with $n_R = n_T = 4$, i.i.d Nakagami-m fading ($m = 1.5$) and block length $l = 1, 10, 15$, compared with the curves for i.i.d Rayleigh case. As indicated by Fig. 2.2, when the block length is short for both Nakagami-m and Rayleigh cases, their lower bounds are the same but Nakagami-m case yields a higher upper bound; When block length is long enough for Rayleigh case but still short for Nakagami-m case, the upper and lower bounds for Rayleigh case are exactly the same while the lower bound of Nakagami-m case is no less than that of Rayleigh case as shown in Fig. 2.3. Fig. 2.4 illustrates the tradeoff curves when the block length is long for both cases. The upper and lower bounds for either of the two channels are identical, thus exactly characterize the channel. We observe that, although the Nakagami-m case has a higher tradeoff curve when $0 < r < 1$, it yields the same result when $r \geqslant 1$ as the Rayleigh case. In general, different fading models only affect the first segment of the optimal tradeoff curve (or bounds), compared with the results in [1].

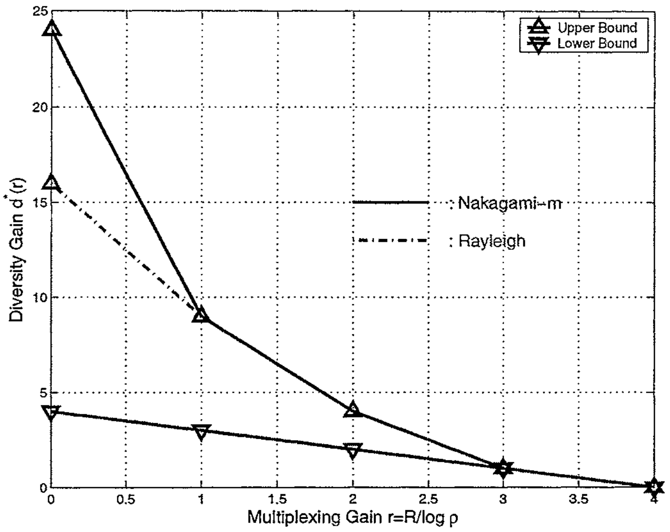

Figure 2.2 The bounds of optimal tradeoff curves for a (4, 4, 1) system over Rayleigh and Nakagami-m ($m = 1.5$) fading channels

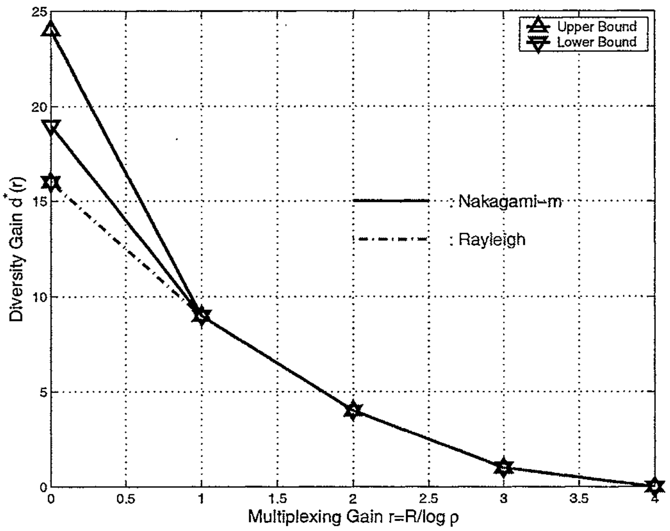

Figure 2.3 The bounds of optimal tradeoff curves for a (4, 4, 10) system over Rayleigh and Nakagami-m ($m = 1.5$) fading channels

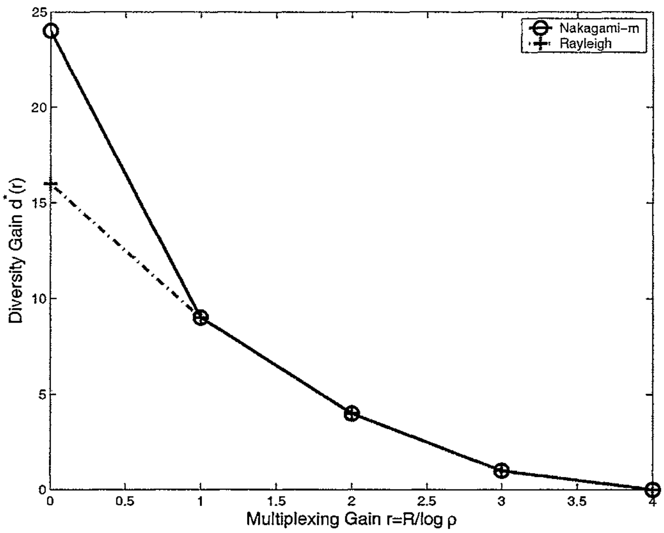

Figure 2.4 Optimal tradeoff curves for a (4, 4, 15) system over Rayleigh and Nakagami-m ($m = 1.5$) fading channels

CHAPTER 3. Extensions

In previous section, we derive the DMT result for independent channel model. Based on the techniques we have use, we are ready to extend our result to some more general and practical models.

3.1 Effect of Correlation

We will study the DMT of a correlated MIMO channel in this section. Suppose that the channel matrix $\boldsymbol{H}$ can be decoupled as $\Sigma_1 \widetilde{\boldsymbol{H}} \Sigma_2$, where $\widetilde{\boldsymbol{H}}$ has independent and regular entries, Σ_1 and Σ_2 are the covariance matrices at the receiver and the transmitter, respectively, both of full rank.

First, let us consider a special case, where Σ_1 and Σ_2 are unitary matrices. By (2.14), the distribution of the eigenvalues of $\boldsymbol{H}\boldsymbol{H}^\dagger$ and that of $\widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{H}}^\dagger$ are the same. Thus those two channels yield the same DMT.

Second, let us consider another special case where Σ_1 and Σ_2 are diagonal matrices with positive elements on the diagonal. It is easy to see that the elements of $\boldsymbol{H}$ are still independent. The pdf's of elements of $\boldsymbol{H}$ share the same parameters as those of $[\widetilde{\boldsymbol{H}}]_{ij}$ except that the variance Ω's are re-scaled, which has no effect on the tradeoff as long as it is a positive constant. Thus $\boldsymbol{H}$ and $[\widetilde{\boldsymbol{H}}]_{ij}$ yield the same tradeoff.

Finally, for arbitrary full-rank covariance matrices Σ_1 and Σ_2, we can do the eigenvalue decomposition of both matrices and get $\boldsymbol{H} = \Sigma_1 \widetilde{\boldsymbol{H}} \Sigma_2 = U_1 \Lambda_1 U_1^\dagger \widetilde{\boldsymbol{H}} U_2 \Lambda_2 U_2^\dagger$. Based on the discussions of previous two special cases, it is straightforward to prove the following corollary:

Corollary 1 *Suppose that the channel matrix $\boldsymbol{H}$ can be decoupled as $\Sigma_1 \widetilde{\boldsymbol{H}} \Sigma_2$, where $\widetilde{\boldsymbol{H}}$ has independent and regular entries, Σ_1 and Σ_2 are the covariance matrices at the receiver and the* *transmitter, respectively, both of full rank. Then $d_{out}(r)$ and $d_G(r)$ of this channel are the same as those of a MIMO system with channel $\widetilde{\boldsymbol{H}}$. The optimal DMT curve (or bounds) for channel $\boldsymbol{H}$ is the same as that for $\widetilde{\boldsymbol{H}}$, which can be characterized using Theorem 2 and Theorem 3.*

3.2 Effect of Channel Mean

In Table 2.1, the only model with a non-zero channel mean is the Rician channel. For both Rayleigh and Rician channels, the parameter t_{ij} is 0. Therefore, for a fading coefficient $\boldsymbol{h}_{ij}$ with complex Gaussian distribution, its channel mean will not affect the optimal tradeoff.

For other models where $t_{ij} \neq 0$, adding a channel mean $\overline{H}$, i.e., $\boldsymbol{H} = \widetilde{\boldsymbol{H}} + \overline{H}$, where $\widetilde{\boldsymbol{H}}$ has independent and regular entries, results in the pdf of the fading coefficient $\boldsymbol{h}_{ij}$ as

$$p_{\boldsymbol{h}_{ij}}(h) = p_{\widetilde{\boldsymbol{h}}_{ij}}(h - [\overline{H}]_{ij})$$

It is obvious that the near ∞ behavior of the pdf of $\boldsymbol{h}_{ij}$ is the same as that of $\widetilde{\boldsymbol{h}}_{ij}$, i.e. as $|h| \to \infty$, $p_{\boldsymbol{h}_{ij}}(h) = O(e^{-b_{ij}|h|^{\beta_{ij}}})$. Also $p_{\boldsymbol{h}_{ij}}(h)$ is bounded by K. But the near 0 behavior of $\boldsymbol{h}_{ij}$'s pdf is the near $-[\overline{H}]_{ij}$ behavior of $\widetilde{\boldsymbol{h}}_{ij}$'s pdf, which the DMT result of channel $\boldsymbol{H}$ depends on.

Let us associate $(p(\cdot), \bar{h})$ with a channel, where $\bar{h}$ is the channel mean and $p(\cdot)$ is the pdf of the channel subtract the channel mean. We call a channel strictly regular if

1. $p(\cdot)$ satisfies the three regular conditions,
2. if $\bar{h} \neq 0$, when $|h + \bar{h}| \to 0$, $p(h) = \Theta(1)$.

Notice when $\bar{h} \neq 0$, a strictly regular channel is equivalent to a zero-mean complex gaussian channel (recall that its paramter $t = 0$) from a DMT's point of view, which means that the non-zero will be non-beneficial to the DMT result. In summary, we have the following corollary.

Corollary 2 *Suppose that the channel matrix $\boldsymbol{H}$ has independent and strictly regular entries. Denote $\overline{\boldsymbol{H}} = E\{\boldsymbol{H}\}$ and $\widetilde{\boldsymbol{H}} = \boldsymbol{H} - \overline{\boldsymbol{H}}$. Then $d_{out}(r)$ and $d_G(r)$ of channel $\boldsymbol{H}$ are the same as those of a MIMO system with channel matrix $\boldsymbol{B}$ with independent elements such that:*

- *If $\overline{H}_{ij} \neq 0$, $\boldsymbol{B}_{ij}$ has a zero mean complex Gaussian distribution;*

- *if* $\overline{H}_{ij} = 0$, $\boldsymbol{B}_{ij}$ *has the same distribution as* $\widetilde{\boldsymbol{H}}_{ij}$.

The optimal diversity and multiplexing tradeoff (or bounds) for channel $\boldsymbol{H}$ *is the same as that for* $\boldsymbol{B}$, *which can be characterized using Theorem 2 and Theorem 3.*

We remark that for channels that are not strictly regular, the presence of a non-zero channel mean may be beneficial.

3.3 Effect of Combination of Channel Mean and Channel Correlation

In fact we can generalize the result in Section 3.1 and Section 3.2 in one corollary. The proof is similar thus omitted here.

Corollary 3 *Suppose that the channel matrix* $\boldsymbol{H}$ *has the form of* $\boldsymbol{H} = \Sigma_1 \hat{\boldsymbol{H}} \Sigma_2$, *where* Σ_1 *and* Σ_2 *are the covariance matrices at the receiver and the transmitter, respectively, both of full rank;* $\hat{\boldsymbol{H}}$ *has independent and strictly regular entries. Then* $d_{out}(r)$ *and* $d_G(r)$ *of channel* $\boldsymbol{H}$ *are the same as those of a MIMO system with channel matrix* $\hat{\boldsymbol{H}}$, *which can be calculated via Corollary 2.*

CHAPTER 4. Conclusions

In this thesis, we derived the optimal multiplexing-diversity tradeoff for general MIMO fading channels, which include different fading types as special cases. We also treated channels with non-identical fading distributions, spatial correlation, and non-zero channel means.

Our approach is based on the joint pdf of the eigenvalues of the Gram matrix of the channel which is the mathematical foot stone to determine the DMT. It is available in a simple and nice form [1, eq.(10)] under i.i.d Rayleigh channel assumption. Although a closed form is hard to derive for other fading types, we show that it is still possible to characterize this pdf in the high SNR regime to derive the DMT. Roughly speaking, if the joint pdf of the eigenvalues can be well approximated by a polynomial near zeros and as exponential near infinity, the DMT can be computed explicitly, which we believe is novel according to the literature we have. The order of this polynomial approximation depends on the near-zero behavior of the specific channel model we choose, to which different DMT results are attributed.

We showed that for a (n_R, n_T, l) system with n_R receive antennas, n_T transmit antennas, and encoding block length l, the optimal tradeoff is determined by a set of parameters t_{ij}, $i \in [1, n_R]$, $j \in [1, n_T]$, one for each fading path, describing the near-zero (or deep-fade) behavior of the probability density function of the fading path. The i.i.d Rayleigh fading case considered in [1] corresponds to $t_{ij} = 0$, $\forall i, j$. Compared with the results in [1] for i.i.d. Rayleigh channels, the optimal tradeoff in the general case may be different only on first segment, i.e. for multiplexing gain $r \in [0, 1)$, which suggests that for $r \geq 1$, the optimal tradeoff depends only on the MIMO system array structure, rather than the channel fading types. We proved that under certain full-rank assumptions spatial correlation has no effect on the optimal tradeoff. We discussed the channel mean effect from a DMT point's of view, and also argued that non-zero channel means are not beneficial for multiplexing-diversity tradeoff for some fading types.

Since we only require the near zero and near infinity behavior of the channel pdf with the bounded condition satisfied, any new practical channel model is relatively easy to check whether it falls within the category we have considered. With the assist of our result and the mathematical techniques we have developed, it is now possible to compare different space time schemes under a unified framework for not only i.i.d Rayleigh channel but a much more generalized fading models while the optimal DMT result we have derived in the paper can serve as a benchmark. Our new DMT results may also facilitate a more comprehensive understanding of the limiting performance of MIMO systems under generalized fading conditions.

APPENDIX A. Proofs

A.1 Proof of Theorem 1

Without loss of generality, we assume that $n_R \leqslant n_T$. Let $\mathcal{A}'(r) = \mathcal{A}(r) \bigcap \mathcal{R}^{n_R^+} = \{\alpha : \sum_{i=1}^{n_T}(1-\alpha_i)^+ < r\}$. Recall that,

$$P_{\text{out}}(r \log \rho) \doteq \int_{\mathcal{A}'(r)} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i}\right) \cdot \prod_{i<j}(\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \cdot \left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} p_{\boldsymbol{H}}(UDQ)dQdU\right] d\alpha. \tag{A.1}$$

A.1.1 Outage diversity when $0 \leqslant r < 1$

Notice that when $r < 1$, $\forall \alpha \in \mathcal{A}'(r)$, since $\sum_i (1-\alpha_i)^+ < r$, it is obvious that $\forall i, 1-\alpha_i < r$, thus $\lambda_i = \rho^{-\alpha_i} \leqslant \rho^{-(1-r)}$. Therefore for every element of H, $|h_{ij}|^2 \leqslant \text{tr}(HH^\dagger) = \sum_{i=1}^{n_R} \rho^{-\alpha_i} \leqslant n_R \rho^{-(1-r)}$. As $\rho \to \infty$, $|h_{ij}| \to 0$, so we can use the near 0 behavior of $p_{\boldsymbol{h}_{ij}}(\cdot)$.

$$P_{\text{out}}(r \log \rho) \doteq \int_{\mathcal{A}'(r)} \prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i} \cdot \prod_{i<j}(\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \cdot \left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} \prod_{i=1}^{n_R}\prod_{j=1}^{n_T} |\,[UDQ]_{ij}\,|^{t_{ij}} dQdU\right] d\alpha. \tag{A.2}$$

Define

$$g(\alpha, \rho, U, Q, i, j) = |[UDQ]_{ij}| \tag{A.3}$$

A.1.1.1 upper bound

Since U and Q are unitary matrices, their elements are bounded by 1. Notice $[UDQ]_{ij}$ is just a polynomial of ρ with the highest order $-\alpha_{n_R}/2$, therefore there exits a positive constant

N_1 such that $g(\alpha,\rho,U,Q,i,j) \leqslant N_1 \rho^{-\alpha_{n_R}/2}$. Thus we can upper bound the outage probability as:

$$
\begin{aligned}
&P_{\text{out}}(r\log\rho) \\
&\dot{\leqslant} \int_{\mathcal{A}'(r)} \prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i} \cdot \prod_{i<j} (\rho^{-\alpha_i}-0)^2 \cdot \left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} \prod_{i=1}^{n_R}\prod_{j=1}^{n_T} \rho^{-\alpha_{n_R}\sum_{i,j} t_{ij}/2} dQ dU \right] d\alpha \\
&\dot{\leqslant} \int_{\mathcal{A}'(r)} \rho^{-f(\alpha)} d\alpha.
\end{aligned}
\tag{A.4}
$$

where

$$
f(\alpha) = \sum_{i=1}^{n_R} (n_T - n_R + 2i - 1)\alpha_i + \alpha_{n_R} \sum_{i,j} \frac{t_{ij}}{2}. \tag{A.5}
$$

Following the same argument used to prove [1, (44)], we obtain

$$
\int_{\mathcal{A}'(r)} \rho^{-f(\alpha)} d\alpha \doteq \rho^{-f(\alpha^*)}, \tag{A.6}
$$

where $\alpha^* = \arg\inf_{\mathcal{A}'(r)} f(\alpha)$.

A.1.1.2 Lower bound

For any $\delta > 0$, define

$$
\mathcal{S}(\delta) \equiv \{\alpha : |\alpha_i - \alpha_j| > \delta, \forall i \neq j\}, \tag{A.7}
$$

For any $0 < \epsilon < \frac{1}{2\max\{n_R,n_T\}}$, define

$$
\begin{aligned}
\mathcal{Q}_{n_R,n_T}(\epsilon) &= \left\{ Q \in \mathcal{C}^{n_R\times n_T} : QQ^\dagger = I_{n_R}, \quad |[Q]_{n_R j}| \geqslant \epsilon, j = 1, ..., n_T \right\}, \\
\mathcal{U}_{n_R,n_R}(\epsilon) &= \left\{ U \in \mathcal{C}^{n_R\times n_R} : UU^\dagger = I_{n_R}, \quad |[U]_{i n_R}| \geqslant \epsilon, i = 1, ..., n_R \right\},
\end{aligned}
\tag{A.8}
$$

Given $Q \in \mathcal{Q}_{n_R,n_T}(\epsilon)$ and $U \in \mathcal{U}_{n_R,n_R}(\epsilon)$, the amplitude of the coefficient of $\rho^{-\alpha_{n_R}/2}$ in the expansion of $g(\alpha,\rho,U,Q,i,j)$ can be lower bounded by ϵ^2.

Given $\alpha \in \mathcal{S}(\delta)$, we could lower bound $g(\alpha,\rho,U,Q,i,j)$ as:

$$
g(\alpha,\rho,U,Q,i,j) \geqslant \epsilon^2 \rho^{-\alpha_{n_R}/2} - \sum_{i=1}^{n_R-1} \rho^{-\alpha_i/2} \geqslant \rho^{-\alpha_{n_R}/2}(\epsilon^2 - n_R \cdot \rho^{-\delta/2}), \text{ when } \rho > 1. \tag{A.9}
$$

It is true that $\forall \delta > 0$, $\epsilon^2 - n_R \rho^{-\delta/2} > 0$, when $\rho > (\frac{n_R}{\epsilon^2})^{2/\delta}$.

Now, we can construct a lower bound on $P_{\text{out}}(r \log \rho)$:

$$
\begin{aligned}
&P_{\text{out}}(r \log \rho) \\
\dot{\geqslant}& \int_{\mathcal{A}'(r) \bigcap \mathcal{S}(\delta)} \prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i} \prod_{i<j} ((1-\rho^{-\delta})\rho^{-\alpha_j})^2 \cdot \rho^{-\frac{\sum_{ij} t_{ij}}{2}\alpha_{n_R}} \left(\epsilon^2 - n_R \cdot \rho^{-\delta/2}\right)^{\sum_{i,j} t_{ij}/2} \\
&\quad \cdot \left[\int_{\mathcal{U}_{n_R,n_R}(\epsilon)} dU \cdot \int_{\mathcal{Q}_{n_R,n_T}(\epsilon)} dQ \right] d\alpha \\
\doteq& (1-\rho^{-\delta})^{{n_R}^2 - n_R} \left(\epsilon^2 - n_R \cdot \rho^{-\delta/2}\right)^{\sum_{i,j} t_{ij}/2} \int_{\mathcal{A}' \bigcap \mathcal{S}(\delta)} \rho^{-f(\alpha)} d\alpha, \\
&\quad \text{[c.f. Lemma 4 in Appendix A.2]} \\
\doteq& \int_{\mathcal{A}'(r) \bigcap \mathcal{S}(\delta)} \rho^{-f(\alpha)} d\alpha \\
\doteq& \rho^{-\inf_{\mathcal{A}'(r) \bigcap \mathcal{S}(\delta)} f(\alpha)}.
\end{aligned}
\tag{A.10}
$$

By the continuity of $f(\cdot)$, $\inf_{\mathcal{A}'(r) \bigcap \mathcal{S}(\delta)} f(\alpha)$ approaches $f(\alpha^*)$ as $\delta \to 0$. Based on the upper and lower bounds we have constructed, we have

$$
P_{\text{out}}(r \log \rho) \doteq \rho^{-d_{\text{out}}(r)}, \tag{A.11}
$$

where $d_{\text{out}}(r)$ is defined in (2.30).

A.1.2 Outage diversity when $r \geqslant 1$

A.1.2.1 Upper bound

Notice that since $p_{\boldsymbol{h}_{ij}}(\cdot)$ is bounded by K. It is straightforward to obtain an upper bound on $P_{\text{out}}(r \log \rho)$:

$$
\begin{aligned}
&P_{\text{out}}(r \log \rho) \\
\dot{\leqslant}& \int_{\mathcal{A}'(r)} \left(\prod_{i=1}^{n_R} \rho^{-(n_T-n_R+1)\alpha_i} \right) \cdot \prod_{i<j} (\rho^{-\alpha_i} - \rho^{-\alpha_j})^2 \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} K^{n_R n_T} dQ dU d\alpha. \\
\dot{\leqslant}& \int_{\mathcal{A}'(r)} \rho^{-\sum_{i=1}^{n_R} (n_T - n_R + 2i - 1)\alpha_i} d\alpha \\
\doteq& \rho^{-f'(\alpha^*)},
\end{aligned}
\tag{A.12}
$$

where $f'(\alpha) = \sum_{i=1}^{n_R} (n_T - n_R + 2i - 1)\alpha_i$ and $\alpha^* = \arg\inf_{\mathcal{A}'(r)} f'(\alpha)$.

A.1.2.2 lower bound

We have assumed that the entries of $\boldsymbol{H}$ are independent and regular. For each value of $r \in [1, n_R)$, we could find a channel realization $H_0(r)$, two positive numbers ζ and ω, both depending on r, such that $H_0(r)$ is of rank $\lfloor r \rfloor$ and $\forall H$ in the neighbor hood of $H_0(r)$, i.e. $||H - H_0(r)||_F < \zeta$, its pdf is lower bounded by ω, i.e. $p_{\boldsymbol{H}}(H) > \omega$. To see this, notice that we can find a rank $\lfloor r \rfloor$ matrix $H(r)$ with non-zero elements, and $H_0(r) = aH(r)$ is also of rank $\lfloor r \rfloor$ where a is a positive number. Thus we can make a arbitrarily small so that we can use the near 0 behavior of the channel to get the lower bound ω. [1]

The singular value decomposition of $H_0(r)$ is $H_0(r) = U_0(r)D_0(r)Q_0(r)$. Since $H_0(r)$ is of rank $\lfloor r \rfloor$, $[D_0(r)]_{ii} = 0$, for $1 \leqslant i \leqslant n_R - \lfloor r \rfloor$ and $[D_0(r)]_{ii} > 0$, for $n_R - \lfloor r \rfloor + 1 \leqslant i \leqslant n_R$. Also we can consider the $H_0(r)$ with distinct positive singular values[2].

Define

$$\begin{aligned}
&\mathcal{B}_{U_0(r)}(x) = \{U : U \in V_{n_R,n_R}, ||U - U_0(r)||_F < x\}, \\
&\mathcal{B}_{Q_0(r)}(x) = \{Q : Q \in V_{n_R,n_T}, ||Q - U_0(r)||_F < x\}, \\
&\mathcal{C}(D_0(r), \rho, x) = \left\{ \alpha : \begin{cases} \rho^{-\alpha_i} \leqslant x, & 1 \leqslant i \leqslant n_R - \lfloor r \rfloor; \\ [D_0(r)]_{ii} \leqslant \rho^{-\alpha_i} \leqslant [D_0(r)]_{ii} + x, & n_R - \lfloor r \rfloor + 1 \leqslant i \leqslant n_R. \end{cases} \right\}^{3} \qquad \text{(A.13)} \\
&\mathcal{S}(\delta, r) = \{\alpha : |\alpha_i - \alpha_j| > \delta, \text{ for } 1 \leqslant i < j \leqslant n_R - \lfloor r \rfloor\}.
\end{aligned}$$

By the continuity of Frobenius norm, there exits a positive number x, which depends on ζ such that $\forall U \in \mathcal{B}_{U_0(r)}(x)$, $\forall Q \in \mathcal{B}_{Q_0(r)}(x)$ and $\forall \alpha \in \mathcal{C}(D_0(r), \rho, x)$, $||UDQ - H_0(r)||_F < \zeta$ (recall that $D = \mathrm{diag}[\rho^{-\alpha_1/2}, ..., \rho^{-\alpha_{n_R}/2}]$).

Since $[D_0(r)]_{ii}$'s are different when $i \geqslant n_R - \lfloor r \rfloor + 1$, we can make x small enough such that $\forall \alpha \in C(D_0(r), \rho, x)$, $|\rho^{-\alpha_i} - \rho^{-\alpha_j}| > x$ for $n_R - \lfloor r \rfloor + 1 \leqslant i < j \leqslant n_R$ and $x < [D_0(r)]_{jj}$, for $n_R - \lfloor r \rfloor + 1 \leqslant j \leqslant n_R$.

[1] Here we just use the near-zero behavior to find $H_0(r)$, but actually, we only require that for each integer r, we can find a channel matrix realization $H_0(r)$ such that $H_0(r)$ has rank r and a neighborhood in which the pdf of every matrix can be lower bounded by a positive number.

[2] A $H_0(r)$ with repeated singular values will also be enough for our proof. We can construct non-overlapping intervals of the largest $\lfloor r \rfloor$ eigenvalues in set $\mathcal{C}(D_0(r), \rho, x)$ to get a lower bound.

[3] It is possible that some α_i are negative, but as $\rho \to \infty$, those $\alpha_i \to 0$ so it is not contradictionay to our claim that we can focus on $\mathcal{A}'(r)$ instead of $\mathcal{A}(r)$.

So for $\alpha \in \mathcal{S}(\delta, r) \bigcap \mathcal{C}(D_0(r), \rho, x)$, we can handle the term $\prod_{i<j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2$ in the following way:

$$
\begin{aligned}
&\prod_{i<j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \\
\geqslant &\prod_{1 \leqslant i<j \leqslant n_R - \lfloor r \rfloor} \left((1-\rho^{-\delta})\rho^{-\alpha_j}\right)^2 \cdot \prod_{n_R - \lfloor r \rfloor + 1 \leqslant i<j \leqslant n_R} x^2 \cdot \prod_{i \leqslant n_R - \lfloor r \rfloor < j} ([D_0]_{jj} - x)^2 \\
\geqslant &\prod_{j=1}^{n_R - \lfloor r \rfloor} \rho^{-(2j-2)\alpha_j} \prod_{1 \leqslant i<j \leqslant n_R - \lfloor r \rfloor} \left(1-\rho^{-\delta}\right)^2 \cdot \prod_{n_R - \lfloor r \rfloor + 1 \leqslant i<j \leqslant n_R} x^2 \\
&\cdot \prod_{i \leqslant n_R - \lfloor r \rfloor < j} ([D_0]_{jj} - x)^2
\end{aligned}
\tag{A.14}
$$

Only the first term $\prod_{j=1}^{n_R - \lfloor r \rfloor} \rho^{-(2j-2)\alpha_j}$ affects the ρ exponent. The other terms can be viewed as constants and removed in the following calculations.

Notice that our choice of x and δ does not depend on ρ.

For fixed x, any $\delta > 0$ and ρ large enough, we have $|\log x / \log \rho| < \delta$, and $|\log([D_0]_{ii} + x)/\log \rho| < \delta$, for $n_R - \lfloor r \rfloor + 1 \leqslant i \leqslant n_R$.

It follows that when ρ is large enough, we have $\mathcal{A}_0(r, \delta) \subset \mathcal{A}(r) \bigcap \mathcal{C}(D_0(r), \rho, x) \bigcap \mathcal{S}(\delta, r)$, where[4]

$$
\begin{aligned}
\mathcal{A}_0(r, \delta) = &\{\alpha_1^{n_R - \lfloor r \rfloor} : \sum_{i=1}^{n_R - \lfloor r \rfloor} (1-\alpha_i)^+ < r - \lfloor r \rfloor (1+\delta)\} \\
&\bigcap \{\alpha_1^{n_R - \lfloor r \rfloor} : \alpha_1 \geqslant \cdots \geqslant \alpha_{n_R - \lfloor r \rfloor} \geqslant \delta, |\alpha_i - \alpha_j| > \delta\},
\end{aligned}
\tag{A.15}
$$

and $\alpha_1^q = \{\alpha_1, \alpha_2, ..., \alpha_q\}$.

[4]If r is an integer, we can use continuity to get $d_{\mathrm{out}}(r)$.

Then the lower bound of $P_{\text{out}}(r\log\rho)$ can be constructed as:

$$
\begin{aligned}
&P(r\log\rho)\\
&\dot{\geqslant} \int_{\mathcal{A}(r)\bigcap\mathcal{C}(D_0(r),\rho,x)} \prod_{i=1}^{n_R}\rho^{-(n_T-n_R+1)\alpha_i}\prod_{i<j}|\rho^{-\alpha_i}-\rho^{-\alpha_j}|^2\\
&\qquad\cdot\,\omega\left[\int_{\mathcal{B}_{U_0}(x)} dU\int_{\mathcal{B}_{Q_0}(x)} dQ\right]d\alpha\\
&\dot{\geqslant} \int_{\mathcal{A}(r)\bigcap\mathcal{C}(D_0(r),\rho,x)} \prod_{i=1}^{n_R}\rho^{-(n_T-n_R+1)\alpha_i}\prod_{i<j}|\rho^{-\alpha_i}-\rho^{-\alpha_j}|^2 d\alpha,\\
&\qquad\text{[c.f. Lemma 3 in Appendix A.2]}\\
&\dot{\geqslant} \int_{\mathcal{A}(r)\bigcap\mathcal{C}(D_0(r),\rho,x)\bigcap\mathcal{S}(\delta,r)} \prod_{i=1}^{n_R-\lfloor r\rfloor}\rho^{-(n_T-n_R+2i-1)\alpha_i}\cdot\prod_{i=n_R-\lfloor r\rfloor+1}^{n_R}\rho^{-(n_T-n_R+1)\alpha_i}d\alpha,\\
&\qquad\text{[c.f. eq. (A.14)]}\\
&\dot{\geqslant} \int_{\mathcal{A}_0(r,\delta)} \prod_{i=1}^{n_R-\lfloor r\rfloor}\rho^{-(n_T-n_R+1)\alpha_i}\alpha_i d\alpha_1\cdots d\alpha_{n_R-\lfloor r\rfloor}\\
&\qquad\cdot\prod_{i=n_R-\lfloor r\rfloor+1}^{n_R}\int_{-\log_\rho([D_0(r)]_{ii}+x)}^{-\log_\rho([D_0(r)]_{ii})}\rho^{-(n_T-n_R+2i)\alpha_i}d\alpha_i,\\
&\dot{\geqslant} \int_{\mathcal{A}_0(r,\delta)} \rho^{-\sum_{i=1}^{n_R-\lfloor r\rfloor}(n_T-n_R+2i-1)\alpha_i}d\alpha_1\cdots d\alpha_{n_R-\lfloor r\rfloor}\\
&\qquad\cdot\prod_{i=n_R-\lfloor r\rfloor+1}^{n_R}\log_\rho\left(\frac{[D_0(r)]_{ii}+x}{[D_0(r)]_{ii}}\right)([D_0(r)]_{ii})^{n_T-n_R+2i}\\
&\dot{\geqslant} \int_{\mathcal{A}_0(r,\delta)} \rho^{-f_r(\alpha)}d\alpha_1\cdots d\alpha_{n_R-\lfloor r\rfloor}
\end{aligned}
\tag{A.16}
$$

where

$$
f_r(\alpha)=\sum_{i=1}^{n_R-\lfloor r\rfloor}(n_T-n_R+2i-1)\alpha_i. \tag{A.17}
$$

So the lower bound of outage probability is

$$
P(r\log\rho)\geqslant\rho^{-\inf_{\mathcal{A}_0(\delta)(r)}f_r(\alpha)},\quad\forall\delta>0 \tag{A.18}
$$

By the continuity of $f_r(\cdot)$, $\inf_{\mathcal{A}_0(r,\delta)}f_r(\alpha)$ approaches $\inf_{\mathcal{A}_0(r,0)}f_r(\alpha)$ as $\delta\to 0$. The upper bound and lower bounds we have obtained for $r\geqslant 1$ case turns out to be two simple optimization problems. It is easy to show that the diversity of those two bounds are the same, which can be further shown to be a simple curve $d_{\text{out}}(r)$ defined by (2.30).

A.2 Lemma 3 and Lemma 4

Lemma 3 *Define the topological metric/distance in the complex Stiefel manifold $V_{m,n}$ as*

$$d(Q_1, Q_2) = ||Q_1 - Q_2||_F, \quad \textit{for } Q_1, Q_2 \in V_{m,n}.$$

$\forall Q_0 \in V_{m,n}$ and $\forall r > 0$, the volume of a ball centered at Q_0 with radius r is always positive. i.e.,

$$\int_{d(Q,Q_0)\leqslant r} dQ > 0$$

Proof: First, due to the invariant property of dQ, the volume does not depend on the center of the ball. Second, when the distance is defined as the geodesic distance instead, a positive lower bound of the volume of a ball is given in [27, (42)–(46)]. Finally, the distances under these two definitions are locally equivalent [27, (21)], which completes the proof. □

Lemma 4 *For any $0 < \epsilon < \frac{1}{2\max\{n_R, n_T\}}$, define*

$$\begin{aligned} \mathcal{Q}_{n_R,n_T}(\epsilon) &= \left\{Q \in \mathcal{C}^{n_R \times n_T} : QQ^\dagger = I_{n_R}, \quad |[Q]_{n_R j}| \geqslant \epsilon, j = 1, ..., n_T \right\}, \\ \mathcal{U}_{n_R,n_R}(\epsilon) &= \left\{U \in \mathcal{C}^{n_R \times n_R} : UU^\dagger = I_{n_R}, \quad |[U]_{i n_R}| \geqslant \epsilon, i = 1, ..., n_R \right\}, \end{aligned} \tag{A.19}$$

Then

$$\int_{\mathcal{Q}_{n_R,n_T}(\epsilon)} dQ > 0, \qquad \int_{\mathcal{U}_{n_R,n_R}(\epsilon)} dU > 0. \tag{A.20}$$

Proof: We can find a $U_0 \in V_{n_R,n_R}$ whose last column is $[\frac{1}{\sqrt{n_R}}, \frac{1}{\sqrt{n_R}}, ..., \frac{1}{\sqrt{n_R}}]^T$. Thus for any $0 < \epsilon < \frac{1}{2\sqrt{\max\{n_R,n_T\}}}$, we can construct a ball $B_{U_0}(x)$ centered at U_0 with sufficient small radius $0 < x \leqslant \epsilon$, such that $B_{U_0}(x)$ is a subset of $\mathcal{U}_{n_R,n_R}(\epsilon)$. To see this, $\forall U \in B_u$, $||U_0 - U||_F \leqslant r$. Thus, for $0 < r \leqslant \epsilon$, $1 \leqslant i \leqslant n_R$

$$r \geqslant \left|[U_0]_{in_R} - [U]_{in_R}\right| \geqslant \left|[U_0]_{in_R}\right| - |[U]_{in_R}| \geqslant 2\epsilon - |[U]_{in_R}|, \tag{A.21}$$

then we have $|[U]_{in_R}| \geqslant 2\epsilon - r \geqslant \epsilon$, i.e., $U \in \mathcal{U}_{n_R,n_R}(\epsilon)$. Since $B_{U_0}(x) \subseteq \mathcal{U}_{n_R,n_R}(\epsilon)$, by Lemma 3, $\int_{\mathcal{U}_{n_R,n_R}(\epsilon)} dU > 0$. Similarly $\int_{\mathcal{Q}_{n_R,n_T}(\epsilon)} dQ > 0$. □

A.3 Discussion on ignoring negative α_i

Let us first partition the outage set $\mathcal{A}(r)$ in the following way:

$$\mathcal{A}(r) = \mathcal{A}'(r) \bigcup_{q=1}^{n_R-1} \left(\mathcal{A}(r) \bigcap \mathcal{B}_q\right), \tag{A.22}$$

where $\mathcal{B}_q = \{\alpha : \alpha_1, ..., \alpha_q \geqslant 0 > \alpha_{q+1} \geqslant \cdots \geqslant \alpha_{n_R}\}$. Notice that $\mathcal{B}_0$ is the set in which $\forall i, \alpha_i < 0$ and $\mathcal{A}(r) \bigcap \mathcal{B}_0 = \emptyset$.

To ignore the integration over the region other than $\mathcal{A}'(r)$, we need to show that:

$$\int_{\mathcal{A}(r) \bigcap \mathcal{B}_q} p_{\boldsymbol{\alpha}}(\alpha) d\alpha \dot{\leqslant} \int_{\mathcal{A}'(r)} p_{\boldsymbol{\alpha}}(\alpha) d\alpha, \tag{A.23}$$

Notice

$$\begin{aligned} \mathcal{A}(r) \bigcap \mathcal{B}_q =& \{\alpha : \sum_{i=1}^{q} (1-\alpha_i)^+ < r - (n_R - q) + \sum_{i=q+1}^{n_R} \alpha_i, \\ & \alpha_1, ..., \alpha_q \geqslant 0, \alpha : \alpha_1, ..., \alpha_q \geqslant 0 > \alpha_{q+1} \geqslant \cdots \geqslant \alpha_{n_R}\} \\ \subset& \{\alpha : \sum_{i=1}^{q} (1-\alpha_i)^+ < r - (n_R - q) \\ & \alpha_1, ..., \alpha_q \geqslant 0, \alpha : \alpha_1, ..., \alpha_q \geqslant 0 > \alpha_{q+1} \geqslant \cdots \geqslant \alpha_{n_R}\} \\ \equiv & \mathcal{C}(q, r) \end{aligned} \tag{A.24}$$

When $r < n_r - q$, $\mathcal{C}(q, r) = \emptyset$, it is trivial that $\int_{\mathcal{A}(r) \bigcap \mathcal{B}_q} p_{\boldsymbol{\alpha}}(\alpha) d\alpha \dot{\leqslant} \int_{\mathcal{A}'(r)} p_{\boldsymbol{\alpha}}(\alpha) d\alpha$,

When $r \geqslant n_r - q$, we have

$$\begin{aligned} & \int_{\mathcal{A}(r) \bigcap \mathcal{B}_q} p_{\boldsymbol{\alpha}}(\alpha) d\alpha) \\ < & \int_{\mathcal{C}(r,q)} p_{\boldsymbol{\alpha}}(\alpha) d\alpha \\ \dot{\leqslant} & \int_{\mathcal{C}(r,q)} \rho^{-\sum_{i=1}^{q}(n_T - n_R + 2i - 1)\alpha_i} \\ & \left[\int_{V_{n_R, n_R}} \int_{V_{n_R, n_T}} \rho^{-\sum_{i=q+1}^{n_R}(n_T - n_R + 2i - 1)\alpha_i - \frac{\sum_{ij} t_{ij}}{2} \alpha_{n_R}} p_{\boldsymbol{H}}(UDQ) dQ dU \right] d\alpha \end{aligned} \tag{A.25}$$

Choose a sufficient large M, such that when $|h| > M, \forall i, j$, $p_{\boldsymbol{h}_{ij}(h)} < O(e^{-b_{ij}|h|^{\beta_{ij}}})$, i.e. the near ∞ behavior takes effect.

If $\rho^{-\alpha_{n_R}} > M\sqrt{n_R n_T}$, notice $\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^\dagger) \geqslant \sum_{i=q+1}^{n_R} \rho^{-\alpha_i}$, there exist i, j such that $|h|_{ij} \geqslant \frac{1}{\sqrt{n_R n_T}} \sum_{i=q+1}^{n_R} \rho^{-\alpha_i} > M$, thus the pdf of this $\boldsymbol{h}_{ij}$ can be upper bounded by an exponential function. Define $b = \min_{ij} b_{ij}$, $\beta = \max_{ij} \beta_{ij}$. This pdf can be upper bounded by $O(e^{-b|h|^\beta})$. Since $(\sum_{i=q+1}^{n_R} \rho^{-\alpha_i})^{\beta/2} \geqslant \frac{1}{n_R} \sum_{i=q+1}^{n_R} \rho^{-\alpha_i \beta/2}$, this pdf can be further upper bounded by $O(e^{-b' \sum_{i=q+1}^{n_R} \rho^{-\alpha_i \beta'}})$, where $b' = \frac{b}{n_R^{\beta/2+1} n_T \beta/2}$ and $\beta' = \beta/2$. We can use K to upper bound the pdf of each other elements in $\boldsymbol{H}$:

$$\int_{0>\alpha_{q+1}\geqslant\cdots\geqslant\alpha_{n_R}} \left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} \rho^{-\sum_{i=q+1}^{n_R}(n_T-n_R+2i-1)\alpha_i - \frac{\sum_{ij} t_{ij}}{2}\alpha_{n_R}} p_{\boldsymbol{H}}(UDQ)dQdU\right]$$
$$d\alpha_{q+1}\cdots d\alpha_{n_R}$$
$$< \int_{0>\alpha_{q+1}\geqslant\cdots\geqslant\alpha_{n_R}} \rho^{-\sum_{i=q+1}^{n_R}(n_T-n_R+2i-1)\alpha_i - \frac{\sum_{ij} t_{ij}}{2}\alpha_{n_R}} O(e^{-b' \sum_{i=q+1}^{n_R} \rho^{-\alpha_i \beta'}}) d\alpha_{q+1}\cdots d\alpha_{n_R}$$
$$\cdot \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} dQdU$$
$$\leqslant M_1 \text{ where } M_1 \text{ is a positive number and does not depend on } \rho \tag{A.26}$$

If $\rho^{-\alpha_{n_R}} < M\sqrt{n_R n_T}$, then we use $K^{n_R n_T}$ to upper bound $p_{\boldsymbol{H}}(UDQ)$ and use $(\frac{M}{\sqrt{n_R n_T}})^y$ to upper bound $\rho^{-\sum_{i=q+1}^{n_R}(n_T-n_R+2i-1)\alpha_i - \frac{\sum_{ij} t_{ij}}{2}\alpha_{n_R}}$, where $y = \sum_{i=q+1}^{n_R}(n_T - n_R + 2i - 1) + \frac{\sum_{ij} t_{ij}}{2}$

$$\int_{0>\alpha_{q+1}\geqslant\cdots\geqslant\alpha_{n_R}>-\log(M\sqrt{n_R n_T})/\log\rho}$$
$$\left[\int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} \rho^{-\sum_{i=q+1}^{n_R}(n_T-n_R+2i-1)\alpha_i - \frac{\sum_{ij} t_{ij}}{2}\alpha_{n_R}} p_{\boldsymbol{H}}(UDQ)dQdU\right]$$
$$d\alpha_{q+1}\cdots d\alpha_{n_R}$$
$$< \left(\frac{\log(M\sqrt{n_R n_T})}{\log\rho}\right)^{n_R-q+1} (\frac{M}{\sqrt{n_R n_T}})^y K^{n_R n_T}$$
$$\cdot \int_{V_{n_R,n_R}} \int_{V_{n_R,n_T}} dQdU$$
$$\leqslant \left(\frac{\log M_2}{\log\rho}\right)^{n_R-q+1} \quad \text{where } M_2 \text{ is a positive number and does not depend on } \rho \tag{A.27}$$

Based on (A.25), we have:

$$
\begin{aligned}
&\int_{\mathcal{A}(r)\bigcap\mathcal{B}_q} p_{\boldsymbol{\alpha}}(\alpha)d\alpha) \\
&\dot{\leqslant} \left((M_1+(\frac{\log M_2}{\log\rho})^{n_R-q+1}\right)\int_{\sum_i^q(1-\alpha_i)^+<r-(n_R-q),\alpha_1\geqslant\cdots\geqslant\alpha_q\geqslant 0}\rho^{-\sum_{i=1}^q(n_T-n_R+2i-1)\alpha_i}d\alpha_1\cdots d\alpha_q \\
&\dot{\leqslant} \int_{\sum_i^q(1-\alpha_i)^+<r-(n_R-q),\alpha_1\geqslant\cdots\geqslant\alpha_q\geqslant 0}\rho^{-\sum_{i=1}^q(n_T-n_R+2i-1)\alpha_i}d\alpha_1\cdots d\alpha_q \\
&\doteq\rho^{-\inf_{\sum_i^q(1-\alpha_i)^+<r-(n_R-q),\alpha_1\geqslant\cdots\geqslant\alpha_q\geqslant 0}\sum_{i=1}^q(n_T-n_R+2i-1)\alpha_i} \\
&\dot{\leqslant}\rho^{-\inf_{\mathcal{A}'(r)}\sum_{i=1}^{n_R}(n_T-n_R+2i-1)\alpha_i+\frac{\sum_{ij}t_{ij}}{2}\alpha_{n_R}} \\
&\dot{\leqslant}\rho^{-d_{\mathrm{out}}(r)}
\end{aligned}
\tag{A.28}
$$

Based on the discussions above, we can always ignore the region $\mathcal{A}(r)\bigcap\mathcal{B}_q$, $0\leqslant q\leqslant n_R-1$ and only consider $\mathcal{A}'(r)$.

APPENDIX B. Linear Optimization Problem

Though the optimization result for i.i.d Rayleigh case is given in [1] which is similar to the general case, we make a self-contained derivation here for completeness. As we have discussed before, we only need to consider the case $n_R \leqslant n_T$ since the result of the other case is just an exchange in n_T and n_R. Therefore, the optimization problem we face in the thesis is of the form:

$$\begin{aligned} \inf\ & f(\alpha) = \sum_{i=1}^{n_R} c_i \alpha_i \\ \text{s.t.}\ & \sum_{i=1}^{n_R} (1-\alpha_i)^+ < r \\ & \alpha_1 \geqslant \alpha_2 \geqslant \alpha_{n_R} \geqslant 0 \end{aligned} \tag{B.1}$$

where

$$c_i = \begin{cases} n_T - n_R + 2i - 1 & 1 \leqslant i \leqslant n_R - 1 \\ n_T + n_R - 1 + \frac{\sum_{i=1}^{n_R}\sum_{j=1}^{n_T} t_{ij}}{2} & i = n_R \end{cases} \tag{B.2}$$

Notice we assume $t_{ij} > 0$, we have $0 < c_i < \cdots < c_{n_R}$. And there is no benefit to make any $\alpha_i > 1$, since replace this with $\alpha_i > 1$ will always yield as smaller $f(\alpha)$. Thus the optimization problem is equivalent to:

$$\begin{aligned} \min\ & f(\alpha) = \sum_{i=1}^{n_R} c_i \alpha_i \\ \text{s.t.}\ & \sum_{i=1}^{n_R} \alpha_i \geqslant n - r \\ & 1 \geqslant \alpha_1 \geqslant \alpha_2 \geqslant \alpha_{n_R} \geqslant 0 \end{aligned} \tag{B.3}$$

Do the following change of variables:

$$\beta_i = \begin{cases} \alpha_i - \alpha_{i+1} & 1 \leqslant i \leqslant n_R - 1 \\ \alpha_{n_R} & i = n_R \end{cases} \tag{B.4}$$

$$b_i = \sum_{i=1}^{i} c_i$$

The optimization problem now changes to:

$$\begin{aligned} \min \; & \sum_{i=1}^{n_R} b_i \beta_i \\ \text{s.t.} \; & \sum_{i=1}^{n_R} i\beta_i \geqslant n - r \\ & \sum_{i=1}^{n_R} \beta_i \leqslant 1 \\ & \beta_i \geqslant 0 \end{aligned} \tag{B.5}$$

The benefit of changing α_i to β_i is that there is no ordering among β_i's, which makes the optimization problem easier to solve. After some algebra, we get:

$$b_i = \begin{cases} (n_T - n_R)i + i^2 & 1 \leqslant i \leqslant n_R - 1 \\ n_T n_R + T & i = n_R \end{cases} \tag{B.6}$$

where $T = \frac{\sum_{ij} t_{ij}}{2}$.

One useful properties of b_i's is:

$$0 < \frac{b_1}{1} < \cdots \frac{b_i}{i} < \cdots < \frac{b_{n_R}}{n_R} \tag{B.7}$$

The constrain area of β_i's is a polyhedron in n_R-dimension vector space. And the objective function is a plane in this vector space. Thus the optimal point should be one of the vertexes of this polyhedron.

The vertexes are the points obtained by the following procedure:

1. *First kind*: For each $1 \leqslant q \leqslant n_R$, the intersection point of line $\beta_i = 0, i \neq q$ and plane $q\beta_q = n_R - r$,if $\beta_q = \frac{n_R - r}{q} \leqslant 1$, i.e. $q \geqslant n_R - r$ which gives the value of the objective function as: $f = \frac{b_q}{q}(n_R - r)$

2. *Second kind*: For each $1 \leqslant q \leqslant n_R$, the intersection point of line $\beta_i = 0, i \neq q$ and plane $\beta_q = 1$, if $q \geqslant n_R - r$ which give the value of the objective function as: $f = b_q$

3. *Third kind*: For each pair of p and q such that $1 \leqslant p < q \leqslant n_R$, the intersection point of plane $\beta_i = 0, i \neq p$ and $i \neq q$, plane $p\beta_p + p\beta_p = n - r$ and plane $\beta_p + \beta_q = 1$, if $\beta_p = \frac{q-(n_R-r)}{q-p} \geqslant 0$ and $\beta_q = \frac{(n_R-r)-p}{q-p} \geqslant 0$, i.e. $p \leqslant n_R - r$ and $q \geqslant n_R - r$, which give the value of the objective function as: $f = b_p \frac{q-(n_R-r)}{q-p} + b_q \frac{(n_R-r)-p}{q-p}$

Since if $q \geqslant n_R - r$, $\frac{b_q}{q}(n_R - r) \leqslant b_q$, the vertexes of the second kind will always not be the solution.

Let us first ignore the case when r is an integer, which can be obtained by continuity. If we could find $p < n_R - r < q$, notice.

$$\begin{aligned}
&b_p \frac{q-(n_R-r)}{q-p} + b_q \frac{(n_R-r)-p}{q-p} \\
=&\frac{b_q}{q}(n_R - r) + \frac{(\frac{b_p}{p} - \frac{b_q}{q})(q - (n_R - r))}{(q-p)(pq^2} \\
\leqslant&\frac{b_q}{q}(n_R - r) \text{ since } \frac{b_p}{p} < \frac{b_q}{q}
\end{aligned} \tag{B.8}$$

Thus, if there exits $p < n_R - r < q$, then the vertexes of the first kind will always not be the solution.

Also, define $g(p,q) = b_p \frac{q-(n_R-r)}{q-p} + b_q \frac{(n_R-r)-p}{q-p}$, it is easy to check that $\frac{\partial g(p,q)}{\partial p} \leqslant 0$ and $\frac{\partial g(p,q)}{\partial p} \geqslant 0$, which means we should choose both p and q as close to $n_R - r$ as possible to minimize the objective function.

When $n_R - 1 < r < n_R$, $0 < n_R - r < 1$, we only need to consider the vertexes of the first kind. Therefore, the minimum objective function value is $b_1(n_R - 1) = (n_T - n_R + 1)(n - r)$. Thus the corresponding α and $d_{\text{out}}(r)$ is:

$$\alpha_i = \begin{cases} n_R - r & i = 1 \\ 0 & i > 1 \end{cases} \tag{B.9}$$

$$d_{\text{out}}(r) = (n_T - n_R + 1)(n - r)$$

When $m < r < m + 1$, where m is an integer and $0 \leqslant m < n_R - 1$, in this case we have $n_R - 1 - m < n_R - r < n_R - m$, the vertex we should choose is of the third kind with

$p = n_R - 1 - m$ and $q = n_R - m$ and the minimum objective function is $b_p(r-m) + b_q(m+1-r)) = b_q - (b_q - b_p)(r-m) = b_q - c_q(r-m)$. Thus the corresponding α and $d_{\text{out}}(r)$ is:

$$\alpha_i = \begin{cases} 1 & i \leqslant n_R - m - 1 \\ m+1-r & i = n_R - m \\ 0 & i > n_R - m \end{cases} \tag{B.10}$$

$$d_{\text{out}}(r) = b_{n_R - m} - c_{n_R - m}(r - m)$$

A key observation here is that $\alpha_i = 0$ when $i \geqslant n_R - \lfloor r \rfloor + 1$, which means:

$$\inf \sum_{i=1}^{n_R} c_i \alpha_i = \inf \sum_{i=1}^{n_R - \lfloor r \rfloor} c_i \alpha_i \tag{B.11}$$

under the constrain

$\sum_{i=1}^{n_R} (1-\alpha_i)^+ < r, \quad \alpha_1 \geqslant \cdots \geqslant \alpha_{n_R} \geqslant 0$

and

$\sum_{i=1}^{n_R - \lfloor r \rfloor} (1-\alpha_i)^+ < r - \lfloor r \rfloor, \quad \alpha_1 \geqslant \cdots \geqslant \alpha_{n_R - \lfloor r \rfloor}, \alpha_{n_R - \lfloor r \rfloor + 1} = \cdots = \alpha_{n_T} = 0$

respectively.

BIBLIOGRAPHY

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

[2] L. Zheng and D. N. Tse, "The diversity-multiplexing tradeoff for non-coherent multiple antenna channels," *Proc. of Allerton Conference on Communication, Control, and Computing*, Oct. 2002.

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

[4] M. Yuksel, and E. Erkip, " Diversity-Multiplexing Tradeoff in Cooperative Wireless Systems," *Conference on Information Science and Systems*, in proceeding, Jan. 2006.

[5] D. Chen, and J. Laneman, " The Diversity-Multiplexing Tradeoff for the Multiaccess Relay Channel," *Conference on Information Science and Systems*, in proceeding, Jan. 2006.

[6] T. Holliday, and A. Goldsmith, "Cross-Layer Design for MIMO Systems", *Conference on Information Science and Systems*, in proceeding, Jan. 2006.

[7] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ Channel: Diversity-Multiplexing-Delay Tradeoff", *IEEE Trans. Inform. Theory* submitted

[8] L. Weng, A. Anastasopoulos, and S. S. Pradhan, "Diversity gain regions for MIMO fading broadcast and multiple access channels," *IEEE Trans. Inform. Theory*, submitted April 2005.

[9] S. Sfar, L. Dai, and K. B. Letaief, "Optimal diversity-multiplexing tradeoff with group detection for MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1074 – 1074, July 2005.

[10] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity–multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152 – 4172, Dec. 2005.

[11] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 968–985, June 2004.

[12] G. L. Stuber, *Principles of Mobile Communications.* Norwell, MA: Kluwer Academic Publishers, 1996.

[13] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[14] M. K. Simon and M.-S. Alouini, *Digital Communications over Fading Channels: A Unified Approach to Performance Analysis.* John Wiley & Sons, 2000.

[15] M. Nakagami, "The m-distribution — A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. G. Hoffman, Ed. Oxford: Pergamon Press, 1960, pp. 3–36.

[16] N. H. Shepherd, "Radio wave loss deviation and shadow loss at 900 MHz," *IEEE Trans. Veh. Technol.*, vol. 26, pp. 309–313, Nov. 1977.

[17] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, pp. 943–968, July 1993.

[18] G. Tzeremes and C. G. Christodoulou, "Use of Weibull distribution for describing outdoor multipath fading," in *Proc. IEEE Antennas, Propagation Soc. Int. Symp.*, San Antonio, TX, June. 2002, pp. 232–235.

[19] J. Cheng, C. Tellambura, and N. C. Beaulieu, "Performance of digital linear modulations on Weibull slow-fading channels," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1265–1268, Aug. 2004.

[20] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE. J. Sel. Areas Commun.*, no. 3, pp. 550–560, April 2002.

[21] C.-N. Chuah, D. N. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE. Trans. Info. Theory*, no. 3, pp. 637–650, Mar. 2002.

[22] M. T. Ivrlac, W. Utschick, and J. A. Nossek, "Fading correlations in wireless MIMO communication systems," *IEEE. J. Sel. Areas Commun.*, no. 5, pp. 819–828, June 2003.

[23] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, no. 5, pp. 684–702, June 2003.

[24] R. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[25] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Institute of Technology, 1989.

[26] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov.–Dec. 1999.

[27] O. Henkel, "Sphere packing bounds in the Grassmann and Stiefel manifolds," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3445–3456, Oct. 2005.

[28] L. Zhao, W. Mo, Y. Ma, Z. Wang, "Diveristy and multiplexing tradeoff in general fading channels," *IEEE Trans. Inform. Theory*, submitted, Nov. 2005.

[29] L. Zhao, W. Mo, Y. Ma, Z. Wang, "Diveristy and multiplexing tradeoff in general fading channels," *Conference on Information Science and Systems*, in proceeding, Jan. 2006.