國 立 交 通 大 學

電信工程研究所

碩 士 論 文

記憶型多重存取萊斯衰減通道之衰減數

The Fading Number of Multiple-Access Rician Fading Channel with Memory

研究生： 周育興

指導教授： 莫詩台方 博士

中 華 民 國 一 百 零 一 年 八 月 三 十 一 日

記憶型多重存取萊斯衰減通道之衰減數

The Fading Number of Multiple-Access Rician Fading Channel with Memory

研 究 生：周育興　　Student：Chou Yu-Hsing

指導教授：莫詩台方 博士　　Advisor：Dr. Stefan M. Moser

國 立 交 通 大 學
電信工程研究所
碩 士 論 文

A Thesis

Submitted to Institute of Communication Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Communication Engineering

August 2012

Hsinchu, Taiwan, Republic of China

中華民國一百零一年八月

Information Theory Laboratory

Institute of Communications Engineering
National Chiao Tung University

Master Project

The Fading Number of Multiple-Access Rician Fading Channel with Memory

Chou Yu-Hsing

Advisor: Prof. Dr. Stefan M. Moser
National Chiao Tung University, Taiwan

Graduation Committee: Prof. Dr. Chen Po-Ning
National Chiao Tung University, Taiwan
Prof. Dr. Chen Fu-Chiarng
National Chiao Tung University, Taiwan

記憶型多重存取萊斯衰減通道之衰減數

研究生：周育興　　　　　　　　　　指導教授：莫詩台方　博士

國立交通大學電信工程研究所碩士班

中文摘要

在本篇論文中，我們分析記憶型萊斯衰減多重存取通道的總通道容量。在此通道中，衰減程序爲高斯分佈並且有一個可目視的路徑成分，而且，有一個以上的使用者在同一時間裡傳送資料。爲了簡化我們的分析，我們只考慮單傳送天線單接收天線的情況，也就是說，所有傳送端的使用者和接收端都僅使用單一天線。

在衰減通道容量的分析中，我們還不知道通道容量的精準表示式。我們使用一種稱作漸進分析的方法，在極限當可用的功率趨近無限大時得到通道容量。它顯示通道總容量在高訊號與雜訊比時會以雙指數成長達到無限大。而在高訊號與雜訊比的展開式中，第二項是一個叫做衰減數的常數。

在我們的研究中，我們找到一個單傳送天線單接收天線，m個使用者，一般記憶型萊斯衰減多重存取通道衰減數的上界。和其自然的下界－單使用者，單傳送天線單接收天線通道的衰減數結合後，我們得到精確的單傳送天線單接收天線，m個使用者，一般記憶型萊斯衰減多重存取通道衰減數。爲了達到這個衰減數，我們必須停止比較差的使用者們傳送，並且讓最好的使用者們用分享時間的方式傳送。

The Fading Number of Multiple-Access Rician Fading Channel with Memory

Student: Chou Yu-Hsing Advisor: Prof. Stefan M. Moser

Institute of Communications Engineering
National Chiao Tung University

Abstract

In this thesis we analyze the sum-rate capacity of the Rician fading multiple-access channel (MAC) with memory. The fading process of the channel is Gaussian in addition to a line-of-sight component. Moreover, there are more than one user sending data at the same time. To simplify our analysis, we consider the single-input single-output (SISO) case, *i.e.*, all the transmitters and the receiver use one antenna.

In the analysis of the fading channel capacity, the exact expression of the capacity is not yet known. A way called *asymptotic analysis* is used to derive the channel capacity in the limit when the available power tends to infinity. It is shown that at high signal-to-noise ratio (SNR), the sum-rate capacity grows to infinity doublelogarithmically. The second term in the high-SNR expansion is a constant called *fading number*.

In our work, we derive an upper bound on the fading number of the general m-user SISO Rician fading MAC with memory. Combining the natural lower bound on the fading number of the single-user SISO channel, we then obtain the exact fading number of the general m-user SISO Rician fading MAC with memory. To achieve the fading number, we have to switch off the worse users and allow the best users communicate by time-sharing.

Acknowledgments

I would like to thank my advisor, Prof. Stefan M. Moser, for providing me this project and offer. He guides me through the difficulties. When I cannot understand him or follow his brilliant ideas, he is always willing to slow down and explain again with patience. With his valuable suggestion, I could have solve the problems and complete the work. He also provides me the best environment — the Information Theory LAB with the iMACs, the printer, the comfortable sofa, and the coffee machine. I have learned much from him, not only related to the research but also the philosophy.

Next, I would to thank all the members in the Information Theory LAB. They gives me a lot of advice about my oral presentation and research problems. With their company, I had a great time during the whole master career. I also want to thank the former LAB member — Lin Gu-Rong. Although he has graduated, he concerns about my work and help me a lot.

Finally, I would like to thank my family. They always support me when I am depressed. Without them, I could not finish my research and my master degree.

Hsinchu, 31 August 2012

Chou Yu-Hsing

Contents

List of Figures

Chapter 1

Introduction

With the development of the technology, the usage of wireless communication is more and more common nowadays. Therefore, it is important to analyze the communication in this area. There is much interference in nature: we divide the influences of the interference in wireless channel model into additive Gaussian noise and multiplicative noise called *fading*. The fading impact on the signal amplitude is often destructive and the channels with fading impact usually called fading channels. Moreover, it is more complicated to design a good communication system for fading channels than the additive white Gaussian noise (AWGN) channel. In this thesis, we research a multiple-access fading channel with memory, where we restrict the fading to the special case of *Rician fading*.

In Rician fading the multiplicative noise is Gaussian distributed with a line-of-sight path between the transmitter and the receiver. Furthermore, we assume memory which implies that not only the present fading process but also the past fading process affects the output of the channel.

Multiple-access indicates that there are more than one user on the transmitter side that send data at the same time. The difference between one user with multiple antennas and multiple-access is that in the latter all users are individual and have no knowledge of the other users, *i.e.*, they are independent of each other, while in the former all antennas cooperate. A common example of a multiple-access channel (MAC) are several mobile phones in the same area that communicate to the same base station.

In order to make a communication system efficient, we have to analyze the channel capacity, which was initially introduced in the famous landmark paper of Shannon — "A Mathematical Theory of Communication" [12]. In this paper, Shannon proved that in every communication channel, we can transmit data reliably with a theoretical maximum rate denoted *capacity*, *i.e.*, for every transmission rate below the capacity, we can make the probability of transmission errors as small as wished. As a result, how good a channel is can be judged by the capacity of the channel.

In the case of MAC, the rate of one user might be affected by other users. It is unfair to decide whether the channel is good or not based on only one user. For this reason, we calculate the sum of all users' rate and use the *sum-rate capacity* to denote the maximum of the sum to replace the original capacity.

Although there are many research results about the wireless communication channel, the channel capacity of a general fading channel is still unknown. To solve the problem, researchers have tried many different approaches. A common approach is to assume that from the training sequences, the receiver can estimate the channel state *perfectly*. Unfortunately, it is impossible to measure the channel state perfectly even if we have sent a large amount of training data. Moreover, another problem is the bandwidth for these training sequences cannot be neglected.

Even though the receiver cannot have perfect knowledge of the channel state, the receiver can be assumed to have some intelligence of the channel by the received information data. We call a channel model *noncoherent*, when both the transmitter and the receiver have no idea about the real state of the channel, but they know the characteristics of the channel.

Since the exact expression of noncoherent channel's capacity is not yet known, a way called *asymptotic analysis* is used to derive the channel capacity at asymptotic high and low signal-to-noise ratio (SNR). In [3], [6] and [8], Lapidoth and Moser have derived the asymptotic high-SNR capacity of general single-user fading channels; the asymptotic low-SNR capacity of fading channel has also been derived in [11]; and the asymptotic high-SNR sum-rate capacity of the memoryless MAC is derived in [4]. In our work, we extend the result of the memoryless MAC [4] to the case with memory.

Since the evaluation of the noncoherent channel's capacity includes the problem of optimization, it is difficult to get the exact channel capacity. Instead of deriving the channel capacity directly, we find an upper and a lower bound of the channel capacity and try to make them tight. From [7], we know a natural upper bound and a natural lower bound on the sum-rate capacity of MAC. The sum-rate capacity of a MAC can be upper-bounded by the capacity of the multiple-input single-output (MISO) channel and lower-bounded by the capacity of the single-input single-output (SISO) channel. Unfortunately, the upper bound is loose. Based on the duality-based upper bound on the mutual information in [3] and [6], we obtain a tighter upper bound on the sum-rate capacity.

In [6, Theorem 6.10], the result is proved that in the regime of high-SNR, the capacity grows only double-logarithmically in the SNR such that the capacity mainly is decided by a constant called *fading number*. This result directly extends to our model, too. Consequently, we focus on the computation of the fading number in this thesis. The precise definition of the fading number is given in Section 3.4.

The main contributions of this thesis are as follows. Firstly, we find an upper bound on the fading number of MAC from the duality-based upper bound on sum-rate capacity of MAC. Next, we derive the exact fading number of the two-user and the general m-user MAC.

The structure of this thesis is as follows: In the reminder of this chapter we will shortly describe our notation. The channel model will be introduced in Chapter 2. In Chapter 3, we will give some concepts that are related to our analysis. In Chapter 4, we review some previous results that we will use in the following chapters. The main results and the derivation of the results are shown in Chapter 5 and Chapter 6. Finally, we will give the conclusion and discuss the results and future works in Chapter 7.

In order to make this thesis easier to read, we attempt to use a consistent and precise notation. For random quantities, we use upper-case letters such as X to denote scalar random variables, and their realizations are written in lower-case, *e.g.*, x. For random vectors we use bold-face capitals, *e.g.*, $\mathbf{X}$ and bold lower-case for their realization. Constant matrices are denoted by a special font of upper-case letters, *e.g.*, H and for random matrices we use another font, *e.g.*, $\mathbb{H}$. Scalars are typically denoted using Greek letters or lower-case Roman letters.

Some exceptions that are widely used in literature and therefore kept in their customary shape are as follows:

- $h(\cdot)$ denotes the differential entropy of a continue random variable.
- $I(\cdot;\cdot)$ denotes the mutual information functional.

Furthermore, we use the capitals Q and W to denote probability distribution functions:

- $Q(\cdot)$ denotes a distribution on an input of a channel.
- $W(\cdot|\cdot)$ denotes a channel law, *i.e.*, the distribution of the channel output when the channel input is given.

The letter C denotes the channel capacity for single-user or the sum-rate capacity for multiple-user. The energy per symbol is denoted by $\mathcal{E}$. Also note that we use $\log(\cdot)$ to denote the natural logarithmic function.

Chapter 2

Channel Model

In this chapter, we will introduce our channel model of the multiple-access Rician fading channel. We assume that the channel model is noncoherent in the sense that neither the transmitter nor the receiver knows the real state of the channel model. Moreover, the transmitter and the receiver only have the information about of channel characteristics, *e.g.*, the distribution of the channel state. In Section 2.1, we consider the m-user SISO channel model and give some mathematical formulas. In Section 2.2, we will describe the special cases of the two-user MAC and the memoryless MAC.

2.1 The m-User SISO Rician Fading MAC with Memory

The multiple-access channel is a channel where more than one user transmits data to the receiver at the same time. Furthermore, every user does not know the state of other users *i.e.*, all users are independent of each other. In our channel model, we assume that the channel has memory, *i.e.*, the current channel state depends on the past channel states. We assume that the channel is a discrete-time model with a common clock known to all users.

We consider a SISO multiple-access channel with m users at the transmitter side. Each user and the receiver use only one antenna. As illustrated in Figure 2.1, the total number of transmit antennas is m.

At time k, the output $Y_k \in \mathbb{C}$ is given by

$$Y_k = \mathbf{H}_k^{\mathsf{T}} \mathbf{x}_k + Z_k \tag{2.1}$$

$$= H_{1,k} x_{1,k} + \cdots + H_{m,k} x_{m,k} + Z_k. \tag{2.2}$$

Here $\mathbf{H}_k \in \mathbb{C}^m$ is a random vector that denotes the time-k random fading vector; $\mathbf{x}_k \in \mathbb{C}^m$ denotes the time-k input vector of m users; $Z_k \in \mathbb{C}$ is a random variable that denotes the time-k additive noise.

The additive noise $\{Z_k\}$ is a independent and identically distributed (IID), zero-mean, circularly symmetric, complex Gaussian random variable, *i.e.*, $\{Z_k\}$ is IID $\sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ for some $\sigma^2 > 0$.

In general, we can assume that the fading process $\{\mathbf{H}_k\}$ and the additive noise $\{Z_k\}$ are independent, and neither does depend on the input $\{\mathbf{x}_k\}$. Furthermore, we can assume the fading process $\mathbf{H}_k$ be any distribution. In our analysis, we only discuss the case of stationary Rician fading with memory. Thus, every component $\{H_{i,k}\}$ which represents the channel fading for the i-th user is Gaussian distributed, *i.e.*, at time k

$$H_{i,k} \sim \mathcal{N}_{\mathbb{C}}\big(d_i, \sigma^2\big), \quad i = 1, \ldots, m \tag{2.3}$$

for some $\sigma^2 > 0$, where $d_i \in \mathbb{C}$ is a constant called the line-of-sight component. The memory is defined by a spectral distribution function $\mathsf{F}_i(\lambda)$ such that the prediction error of $H_{i,k}$ given from its past $\{H_{i,j}\}_{j=-\infty}^{k-1}$ is

$$\epsilon_i^2 = \exp\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \log \mathsf{F}_i'(\lambda) \mathsf{d}\lambda\right), \tag{2.4}$$

which is shown in [9] and [13]. Moreover, we assume the different channels to be independent, *i.e.*,

$$\{H_{i,k}\} \perp\!\!\!\perp \{H_{j,k}\}, \quad \forall i, j = 1, \ldots, m, \quad i \neq j. \tag{2.5}$$

As for the input, because of the property of the multiple-access channel, the different users are not allowed to cooperate, *i.e.*,

$$\{X_{i,k}\} \perp\!\!\!\perp \{X_{j,k}\}, \quad \forall i, j = 1, \ldots, m, \quad i \neq j. \tag{2.6}$$

Furthermore, we use a symbol $\mathcal{E}$ to denote the total power allowed and consider one of three different constraints for the input:

- **Peak-Power Constraint:** At every time-step every user i is allowed to use a power of at most $\frac{\kappa_i}{m}\mathcal{E}$:

$$\Pr\left[|X_{i,k}|^2 > \frac{\kappa_i}{m}\mathcal{E}\right] = 0 \tag{2.7}$$

  for some fixed number $\kappa_i > 0$.

- **Average-Power Constraint:** Averaged over the length of a codeword, every user i is allowed to use a power of at most $\frac{\kappa_i}{m}\mathcal{E}$

$$\mathsf{E}\left[|X_{i,k}|^2\right] \leq \frac{\kappa_i}{m}\mathcal{E} \tag{2.8}$$

  for some fixed number $\kappa_i > 0$.

- **Power-Sharing Average-Power Constraint:** Averaged over the length of a codeword all users together are allowed to use a power of at most $\bar{\kappa}\mathcal{E}$

$$\mathsf{E}\left[\sum_{i=1}^{m} |X_{i,k}|^2\right] \leq \bar{\kappa}\mathcal{E} \tag{2.9}$$

  for some fixed number $\bar{\kappa} > 0$.

In all cases, we have the signal-to-noise ratio:

$$\text{SNR} \triangleq \frac{\mathcal{E}}{\sigma^2}. \tag{2.10}$$

Note that if $\kappa_i = 1$ for all i, we have the special case where all users have an equal power available. Also note that in (2.7) and (2.8), we have normalized the power to the number of user m. From an engineering point of view, this might be strange; however, in regard of our freedom to choose κ_i, it is irrelevant, and it simplifies our analysis since we can easily connect the power-sharing average-power constraint with other two constraints. Indeed, if we define $\bar{\kappa}$ to be the average of the constants $\{\kappa_i\}_{i=1}^m$, *i.e.*,

$$\bar{\kappa} \triangleq \frac{1}{m} \sum_{i=1}^{m} \kappa_i \tag{2.11}$$

then the three constraints are in order of strictness: the peak-power constraint is the most stringent of the three constraints, *i.e.*, if (2.7) is satisfied for all $i = 1, \ldots, m$, then the other two constraints are also satisfied; and the average-power constraint is the second most stringent in the sense that if (2.8) is satisfied for all $i = 1, \ldots, m$, then the power-sharing average-power constraint (2.9) is also satisfied. In the remainder of this thesis, we will always assume that (2.11) holds.

It is worth mentioning that the slackest constraint, *i.e.*, the power-sharing average-power constraint, implicitly allows a form of cooperation: even if the user are still assumed to be statistically independent, we do allow cooperation concerning power allocation. This is not very realistic, however, we include it anyway because it will help in deriving bounds on the sum-rate capacity. As a matter of fact, it will turn out that the asymptotic sum-rate capacity is unchanged irrespective of which constraint is assumed.

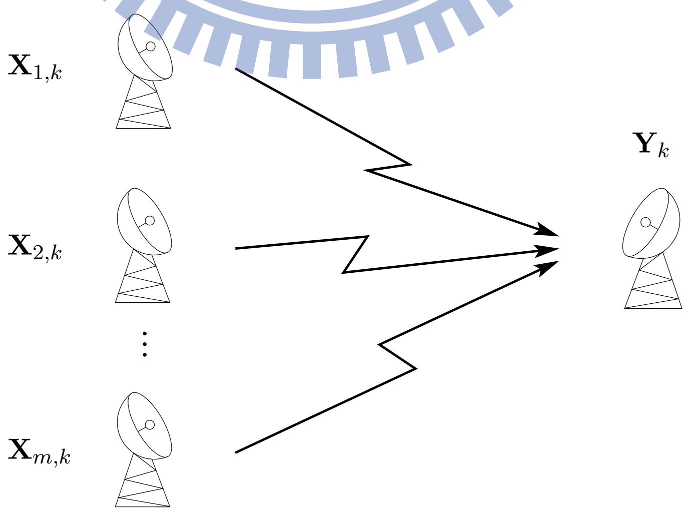

Figure 2.1: The m-user SISO MAC.

2.2 The Simplified Channel Model

Now, we consider the special case of a two-user SISO multiple-access channel with memory. We assume that there are only two users in our channel model, *i.e.*, $m = 2$. At time k, the time-k output of channel can be simplifed as

$$Y_k = H_{1,k}x_{1,k} + H_{2,k}x_{2,k} + Z_k \tag{2.12}$$

where $H_{1,k}, H_{2,k}$ and Z_k are as the same as we mentioned before. As for the input, the two users are not allowed to cooperate, *i.e.*,

$$\{X_{1,k}\} \perp\!\!\!\perp \{X_{2,k}\}. \tag{2.13}$$

Moreover, given $X_{1,k} = x_{1,k}$, $X_{2,k} = x_{2,k}$, and the past fading process $\{\mathbf{H}_j\}_{j=-\infty}^{k-1}$, the variance of the output Y_k is

$$\epsilon_1^2|x_{1,k}|^2 + \epsilon_2^2|x_{2,k}|^2 + \sigma^2 \tag{2.14}$$

where $\epsilon_i^2 > 0$ is the prediction error (2.4).

Another special case is the memoryless version of the m-user channel (2.1), the output of which is given by

$$Y = \mathbf{H}^{\mathsf{T}}\mathbf{x} + Z \tag{2.15}$$

$$= H_1x_1 + \cdots + H_mx_m + Z. \tag{2.16}$$

Here $\mathbf{H} \in \mathbb{C}^m$ is a random vector that denotes the fading vector of fading process; $\mathbf{x} \in \mathbb{C}^m$ denotes the input vector of m users; $Z \in \mathbb{C}$ is a random variable that denotes the additive noise.

The additive noise Z and the input $\mathbf{x}$ are as the same as the case with memory. The difference between the two cases is we assume that the channel is memoryless such that the present output only depends on the present input and the present fading process. Therefore, the past fading process has no effect on the present fading, and we drop the prediction errors.

Chapter 3

Mathematical Preliminaries

In this chapter, we review some important notions that help us to analyze our channel model. We consider the memoryless channel as shown in (2.15) and (2.16) with only single user, the output of which is given as

$$Y = Hx + Z. \tag{3.1}$$

In Section 3.1, we review the knowledge of the channel capacity, furthermore, we provide the idea of sum-rate capacity to compute the channel capacity of channel with more than one user. In Section 3.2, we give the concept that the power of input escapes to infinity. Moreover, we provide a lemma expressing in some conditions that the power of input may escape to infinity. In Section 3.3, we mention the stationarity of the input distribution that makes our channel model easier to analyze. In Section 3.4, we introduce the fading number which is relative to the fading channel.

3.1 The Channel Capacity

Firstly, we review the definition of channel capacity in [12] provided by Shannon. The channel capacity of a discrete memoryless channel (DMC) is defined as

$$\mathsf{C} \triangleq \max_{Q_X} I(X;Y) \tag{3.2}$$

where the maximization is taken over all possible input distributions Q_X. Next, we consider the general case of continuous channel with memory, *i.e.*, the input and output of the channel are continuous alphabet, and the channel capacity becomes:

$$\mathsf{C} \triangleq \lim_{n\to\infty} \frac{1}{n} \sup_{Q_X^n} I(X_1^n;Y_1^n) \tag{3.3}$$

where X_1^n symbols the squence $X_1, \ldots, X_n$, and the supremum is over the set of all probability distributions of X_1^n.

Furthermore, there must be a power constraint for the input of the channel:

$$\mathsf{C} \triangleq \lim_{n\to\infty} \frac{1}{n} \sup_{Q_X^n \in \mathcal{D}} I(X_1^n;Y_1^n) \tag{3.4}$$

where $\mathcal{D}$ is the set of all probability measures Q_X^n satisfying the given constraint, *i.e.*,

$$|X_k|^2 \leq \mathcal{E}, \quad \forall k \tag{3.5}$$

for the peak power constraint or

$$\frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\left[X_k^2\right] \leq \mathcal{E} \tag{3.6}$$

for the average power constraint.

When we extend the channel model to multiple-user, the capacity becomes a capacity region. This is too complicated, so we replace it by the sum-rate capacity that is given as:

$$\mathsf{C} = \sup_{Q_{\mathbf{X}}^n \in \mathcal{D}} (\mathsf{R}_1 + \cdots + \mathsf{R}_m) \tag{3.7}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sup_{Q_{\mathbf{X}}^n \in \mathcal{D}} I(\mathbf{X}_1^n; Y_1^n) \tag{3.8}$$

where R_i is the achievable rate of the i-th user; $\mathcal{D}$ is the set of all probability measures $Q_{\mathbf{X}}^n$, the m subsets of which are independent and satisfy the given constraint (2.7) for the peak-power constraint, (2.8) for the average-power constraint or (2.9) for the power-sharing average-power constraint.

3.2 Escaping to Infinity

In this section, we introduce the concept of input distributions that escape to infinity in [3], [6]. A sequence of input distributions parameterized by the allowed cost (in our case the cost of the fading channel is the available power or the SNR, respectively) is said to *escape to infinity* if it assigns to every fixed compact set a probability that tends to zero as the allowed cost tends to infinity. That is to say such a distribution does not use any finite-cost symbols in the limit that when the allowed cost tends to infinity.

Since the asymptotic capacity of the fading channels can only be achieved by input distributions that escape to infinity, this notion is very important. As a matter of fact we can show that to achieve a mutual information of only identical asymptotic *growth rate* as the capacity, the input distribution *must* escape to infinity.

Then, we give the definition of escaping to infinity for the fading channel.

Definition 3.1. Let $\{Q_{\mathcal{E}}\}_{\mathcal{E}\geq 0}$ be a family of input distribution for the memoryless fading channel (3.1), where this family is parameterized by the available average power $\mathcal{E}$ such that

$$\mathsf{E}_{Q_{\mathcal{E}}}\left[\|X\|^2\right] \leq \mathcal{E}, \quad \mathcal{E} \geq 0. \tag{3.9}$$

We say that the input distributions $\{Q_{\mathcal{E}}\}_{\mathcal{E}\geq 0}$ *escape to infinity* if for every $\mathcal{E}_0 > 0$

$$\lim_{\mathcal{E}\uparrow\infty} Q_{\mathcal{E}}(\|X\|^2 \leq \mathcal{E}_0) = 0. \tag{3.10}$$

And we have the following lemma:

Lemma 3.2. *Assume a single-user memoryless SISO fading channel as given in* (3.1) *and let* $W(\cdot|\cdot)$ *denote the corresponding conditional channel law. Let* $\{Q_{\mathcal{E}}\}_{\mathcal{E}\geq 0}$ *be a family of input distributions satisfying the power constraint* (3.9) *and the condition*

$$\lim_{\mathcal{E}\uparrow\infty} \frac{I(Q_{\mathcal{E}}, W)}{\log\log\mathcal{E}} = 1. \tag{3.11}$$

Then $\{Q_{\mathcal{E}}\}_{\mathcal{E}\geq 0}$ *escape to infinity.*

Proof. A proof can be found in [3], [6]. □

From the engineering point of view, this concept is intuitive that the input should utilize the resource (available power) completely as the available power tends to infinity, as a result, any fixed symbol is not used in the limit.

Remark 3.3. When computing the bounds on the fading number (which is part of the capacity in the limit when $\mathcal{E}$ tends to infinity), we can assume that for every $\mathcal{E}_0 > 0$

$$\Pr\left(\|X\|^2 \leq \mathcal{E}_0\right) = 0. \tag{3.12}$$

Next, we generalize escaping to infinity to multiple user, and there is a proposition which is stated in [4].

Proposition 3.4. *Let* $\{Q_{\mathcal{E}}\}_{\mathcal{E}\geq 0}$ *be a family of joint input distributions of the multiple-access fading channel given in* (2.15) *and* (2.16)*, where the family is parameterized by the available average power* $\mathcal{E}$ *such that*

$$\mathsf{E}_{Q_{\mathcal{E}}}\left[\|\mathbf{X}\|^2\right] \leq \mathcal{E}, \quad \mathcal{E} \geq 0. \tag{3.13}$$

Let $W(\cdot|\cdot)$ *be the channel law, and* $\{Q_{\mathcal{E}}\}$ *be such that*

$$\lim_{\mathcal{E}\uparrow\infty} \frac{I(Q_{\mathcal{E}}, W)}{\log\log\mathcal{E}} = 1. \tag{3.14}$$

Then at least one user's input distribution must escape to infinity, i.e., *for any* $\mathcal{E}_0 > 0$

$$\lim_{\mathcal{E}\uparrow\infty} Q_{\mathcal{E}}\left(\bigcup_{i=1}^{m}\left\{\|X_i\|^2 \geq \frac{\mathcal{E}_0}{m}\right\}\right) = 1. \tag{3.15}$$

Proof. A proof can be found in [4]. □

3.3 Stationarity

In this section, we give the idea of the capacity achieving input distributions that are stationary in [8]. One of the main assumption about our channel model is that the fading processes and the additive noises are *stationary*. Since the assumption allows us to shift random quantities in time, it is important for the results and the derivation. From an intuitive point of view, it is obvious that a stationary channel model should have a capacity achieving input distribution that is also stationary. Unfortunately, we are not aware of a rigorous proof of this claim. However, [8, Theorem 3] proves that for the MIMO fading channel, the capacity can be approached up to a $\tau > 0$ by a distribution that looks stationary apart from edge effects. We simplify the MIMO fading channel to MISO fading channel where there is only one antenna at the output.

Theorem 3.5. *Consider a MISO channel model with input $\mathbf{x}_k \in \mathbb{C}^{n_T}$ and output $Y_k \in \mathbb{C}$ which is shown as*

$$Y_k = \mathbf{H}^\mathsf{T}\mathbf{x}_k + Z_k. \tag{3.16}$$

Note that the channel is both stationary *and* unaffected by zero input vector $\mathbf{0}$ *in the following sense: for every choice of $n \in \mathbb{N}$ and $t \in \mathbb{Z}$, for some integers $\underline{n} < -|t|$ and $\overline{n} > n + |t|$, and for every distribution $Q \in \mathcal{P}(\mathbb{C}^{n_T \times n})$ we have*

$$I\left(\mathbf{0}_{\underline{n}}^{0+t}, \mathbf{X}_{1+t}^{n+t}, \mathbf{0}_{n+1+t}^{\overline{n}}; Y_{\underline{n}}^{\overline{n}}\right) = I\left(\mathbf{X}_1^n; Y_1^n\right) \tag{3.17}$$

whenever both $\mathbf{X}_{1+t}^{n+t}$ on the LHS and $\mathbf{X}_1^n$ on the RHS have the same distribution Q.

Now fix some non-negative integer κ and some power $\mathcal{E} > 0$. Then for every $\tau > 0$ there corresponds some positive integer $\eta = \eta(\mathcal{E}, \tau)$ and some distributions $Q_{\mathcal{E},\tau}^{\kappa+1} \in \mathcal{P}(\mathbb{C}^{n_T \times (\kappa+1)})$ such that for a blocklength n sufficiently large there exists some input $\mathbf{X}_1^n$ satisfying the following conditions.

1. *The input $\mathbf{X}_1^n$ nearly achieves capacity in the sense that*

$$\frac{1}{n} I(\mathbf{X}_1^n; Y_1^n) \geq \mathrm{C}(\mathcal{E}) - \tau. \tag{3.18}$$

2. *For every integer μ with $0 \leq \mu \leq \kappa$, every length-$(\mu+1)$ block of adjacent vectors*

$$(\mathbf{X}_\ell, \dots, \mathbf{X}_{\ell+\mu}) \tag{3.19}$$

take from within the sequence

$$\mathbf{X}_\eta, \mathbf{X}_{\eta+1}, \dots, \mathbf{X}_{n-2\eta+2} \tag{3.20}$$

has the same joint distribution $Q_{\mathcal{E},\tau}^{\mu+1}$, where this distribution $Q_{\mathcal{E},\tau}^{\mu+1}$ is given as the corresponding marginal distribution of $Q_{\mathcal{E},\tau}^{\kappa+1}$.

3. *In particular, all vectors in (3.20) have the same marginal distribution $Q_{\mathcal{E},\tau}^{1}$.*

4. *The marginal distribution $Q_{\mathcal{E},\tau}^{1}$ gives rise to a second moment $\mathcal{E}$*

$$\mathsf{E}\left[\|\mathbf{X}_k\|^2\right] = \mathcal{E}, \quad k \in \{\eta, \dots, n - 2\eta + 2\}. \tag{3.21}$$

5. *The first $\eta-1$ vectors and the last $2(\eta-1)$ vectors satisfy the power constraint possibly strictly*

$$\mathsf{E}\left[\|\mathbf{X}_k\|^2\right] \leq \mathcal{E}, \quad k \in \{1,\ldots,\eta-1\} \cup \{n-2\eta+3,\ldots,n\}. \tag{3.22}$$

Proof. A proof can be found in [8]. □

Remark 3.6. Neglecting the edge effects for the moment, Theorem 3.5 basically says that, for every $\mu \leq \kappa$, every block of $\mu+1$ *adjacent* vectors has the same distribution independent of the time shift. From this it immediately follows that the distribution of every subset of (not necessarily adjacent) vectors of a $\mu+1$ block does not change when the vectors are shifted in time (simply marginalize those vectors out that are not members of the subset). Therefore, Theorem 3.5 *almost* proves that the capacity achieving input distribution is stationary: the only problems are the edge effects. Note that κ can be chosen freely, but has to remain fixed until n has been loosened to infinity. That is, to get rid of the edge effects one needs to first let n tend to infinity, before one can let κ grow.

3.4 The Fading Number

When we focus on the asymptotic analysis of channel capacity at high SNR, the channel capacity grows only double-logarithmically in the SNR, which has been shown in [3], [6]. It means that at high SNR, the addition of power is inefficient since to get an additional bit improvement in capacity, we have to square the SNR. In fact, the difference between channel capacity and $\log\log \text{SNR}$ is bounded as the SNR tends to infinity, *i.e.*,

$$\overline{\lim_{\mathcal{E}\uparrow\infty}} \left\{ \mathsf{C}(\mathcal{E}) - \log\log\frac{\mathcal{E}}{\sigma^2} \right\} < \infty. \tag{3.23}$$

The bounded term is called the *fading number*. The precise definition of fading number is as follows.

Definition 3.7. The fading number $\chi(\mathbb{H}_k)$ of a fading channel with fading matrix $\mathbb{H}_k$ is defined as

$$\chi(\mathbb{H}_k) \triangleq \overline{\lim_{\mathcal{E}\uparrow\infty}} \left\{ \mathsf{C}(\mathcal{E}) - \log\log\frac{\mathcal{E}}{\sigma^2} \right\}. \tag{3.24}$$

Thus, whenever χ is finite and the limit of (3.24) exists, we can express the capacity as

$$\mathsf{C}(\mathcal{E}) = \log\log\frac{\mathcal{E}}{\sigma^2} + \chi + o(1). \tag{3.25}$$

Here when $\mathcal{E}$ tends to infinity, the $o(1)$ term tends to zero. Therefore, at high SNR, we can approximate the channel capacity of a fading channel to

$$\mathsf{C}(\mathcal{E}) \approx \log\log\frac{\mathcal{E}}{\sigma^2} + \chi. \tag{3.26}$$

Note that this approximation is not always valid. In the regime of low-SNR to medium-SNR, the capacity is dominated by the $o(1)$ term that cannot be neglected. However, we only focus on high-SNR regime in the analysis of the asymptotic capacity, furthermore, especially in the condition that the SNR tends to infinity. Hence we can use the asymptotic expression of the channel capacity shown as (3.26) which is constructed by a double-logarithmical term decided by the SNR and a constant called fading number instead of the intractable exact expression. Moreover, since the first double-logarithmically term is always the same at high SNR, we can even only be concerned about the fading number.

The most important effect of the fading number is that it symbols a criterion for the communication system. From (3.26), we know the capacity is extremely power-inefficient at high SNR, as a result, we should avoid transmission in this regime. Since the fading number is a constant, the capacity is mainly dominated by the $o(1)$ term in low-SNR regime, by fading number in medium-SNR regime and by double-logarithmical term in high-SNR regime. The fading number can provide a threshold of how high the capacity can be before entering the high-SNR regime, *i.e.*, the fading number can give a threshold SNR_0 such that once the SNR is higher than SNR_0, the enlargement of SNR is inefficient to the channel capacity. For more details see [8].

Chapter 4

Previous Results

In this chapter we review some known results that are related to our analysis. In Section 4.1, we derive natural upper and lower bounds of the multiple-access channel sum-rate capacity. Furthermore, we find bounds on the fading number of the MAC channel. The results are based on [7]. In Section 4.2, we give an upper bound on the mutual information of memoryless MAC fading channel, which is derived in [7]. In Section 4.3, we provide two equalities introduced in [5] that are needed in our analysis.

4.1 Natural Upper and Lower Bounds

In this section, we consider the channel model of a two-user SISO MAC fading channel as shown in (2.12). Note that the difference between the MAC and the MISO fading channel with two transmitters and one receiver is that the two transmitters of the latter can cooperate while the two transmitters of the former are independent. Therefore, in the MISO fading channel, we can get higher transmission rate than the MAC fading channel, *i.e.*, the sum-rate capacity of the MAC fading channel is upper-bounded by the capacity of the MISO fading channel:

$$\mathsf{C}_{\text{MAC}}(\mathcal{E}) \leq \mathsf{C}_{\text{MISO,av}}(\mathcal{E}). \tag{4.1}$$

On the other hand, we can regard the SISO fading channel as a special case of the MAC fading channel where only one user communicates. As a result, the sum rate of MAC fading channel cannot be smaller than the single-user rate that can be achieved when the weaker of the two users is switched off, *i.e.*,

$$\mathsf{C}_{\text{MAC}}(\mathcal{E}) \geq \max_{i\in\{1,2\}} \mathsf{C}_{\text{SISO},i}(\mathcal{E}). \tag{4.2}$$

Based on (3.24), (4.1), and (4.2), we can define the fading number of MAC by

$$\chi_{\text{MAC}} \triangleq \lim_{\mathcal{E}\uparrow\infty} \left\{ \mathsf{C}_{\text{MAC}}(\mathcal{E}) - \log\left(1+\log\left(1+\frac{\mathcal{E}}{\sigma^2}\right)\right)\right\}. \tag{4.3}$$

Note that asymptotically for $\mathcal{E}\uparrow\infty$, $\log\left(1+\log\left(1+\frac{\mathcal{E}}{\sigma^2}\right)\right) = \log\log\frac{\mathcal{E}}{\sigma^2} + o(1)$.

From [6], the fading number of the SISO Rician fading channel with memory is given as

$$\chi_{\text{SISO}} = \log(|d|^2) - \text{Ei}\big(-|d|^2\big) - 1 + \log\frac{1}{\epsilon^2}, \tag{4.4}$$

where $\text{Ei}(\cdot)$ is the exponential integral function defined as

$$\text{Ei}(-\xi) \triangleq -\int_{\xi}^{\infty} \frac{e^{-t}}{t}dt, \quad \xi > 0, \tag{4.5}$$

and where $\epsilon^2 > 0$ is the prediction error (2.4).

Hence, from (4.2), we have

$$\chi_{\text{MAC}} \geq \max_{i\in\{1,2\}} \chi_{\text{SISO},i} \tag{4.6}$$

$$= \max_{i\in\{1,2\}} \left\{ \log(|d_i|^2) - \text{Ei}\big(-|d_i|^2\big) - 1 + \log\frac{1}{\epsilon_i^2} \right\}. \tag{4.7}$$

4.2 An Upper Bound on the Mutual Information of Memoryless MAC

We know the multiple-access channel is quite similar to the MISO channel, therefore, we recall an upper bound on the capacity of MISO channel in [3], [6] to derive the sum-rate capacity of MAC channel. By choosing the output distribution as a generalized Gamma distribution, we derive the upper bound from the dual expression of mutual information, and we have the following lemma.

Lemma 4.1. *Consider a memoryless version of the MISO fading channel* (3.16)*, the output of which is shown as*

$$Y = \mathbf{H}^{\mathsf{T}}\mathbf{x} + Z. \tag{4.8}$$

Then the mutual information between input and output of the channel is upper-bounded as follows:

$$I(\mathbf{X};Y) \leq -h(Y|\mathbf{X}) + \log\pi + \alpha\log\beta + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + (1-\alpha)\mathsf{E}\left[\log(|Y|^2+\nu)\right]$$
$$+\frac{1}{\beta}\mathsf{E}\left[|Y|^2\right] + \frac{\nu}{\beta} \tag{4.9}$$

where $\alpha, \beta > 0$ and $\nu \geq 0$ are parameters that can be chosen freely, but must not depend on $\mathbf{X}$.

Proof. A proof can be found in [3], [6]. □

Next, we choose the parameters α, β and ν appropriately, (4.9) can be further simplified to an upper bound on the mutual information of MAC fading channel.

Lemma 4.2. *For the memoryless version of SISO Rician fading MAC* (2.15) *and* (2.16)*, an upper bound of the mutual information is given as follows:*

$$I(\mathbf{X};Y) \le -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right)\right] + \epsilon_\nu + \alpha(\log\beta - \log\sigma^2 + \gamma)$$
$$+ \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta}. \tag{4.10}$$

Proof. The proof can be found in Appendix A and in [7]. □

4.3 Two Equalities for the SISO Rician Fading MAC

In asymptotic analysis, we only consider the input distribution escaping to infinity. From Proposition 3.4, we know that at least one user's input distribution escapes to infinity as the input distribution escapes to infinity. With this constraint, we can show that the following two expectations are equal to zero.

Lemma 4.3. *Consider a memoryless version of SISO Rician fading MAC* (2.15) *and* (2.16)*. With the constraints of power-sharing average-power constraint and the input distribution escapes to infinity, we have*

$$\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_1||X_1||d_i||X_i|}{|X_1|^2+\cdots+|X_m|^2}\right] = 0 \tag{4.11}$$

$$\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\right] = 0 \tag{4.12}$$

for $i, j \in \{2, \ldots, m\}$, $i \neq j$, *and where* $\mathcal{A}$ *is the set defined as*

$$\mathcal{A} \triangleq \Bigg\{ \{Q_{\mathbf{X}}\}_{\mathcal{E}>0} : X_i \perp\!\!\!\perp X_j, \quad i,j = 1,\ldots,m, \quad \forall i \neq j; \quad \mathsf{E}\left[\sum_{i=1}^{m}\|\mathbf{X}_i\|^2\right] \le \bar{\kappa}\mathcal{E};$$
$$\lim_{\mathcal{E}\uparrow\infty} Q_{\mathcal{E}}\left(\bigcup_{i=1}^{m}\left\{|X_i|^2 \ge \frac{\mathcal{E}_0}{m}\right\}\right) = 1 \text{ for any fixed } \mathcal{E}_0 > 0 \Bigg\}. \tag{4.13}$$

Proof. The proof can be found in Appendix B and in [5]. □

Chapter 5

Main Results

In this chapter, we present our main results about the fading number of SISO Ricain fading MAC with memory. In Section 5.1, we find an upper bound on the fading number of the m-user SISO Rician fading MAC with memory from the duality-based bounds. In Section 5.2, we show the exact fading number of the two-user SISO Rician fading MAC with memory. In Section 5.3, we generalize the channel model to the m-user case and provide the exact fading number, furthermore, we discuss the power constraint of the input.

5.1 An Upper Bound on Fading Number of the m-User SISO Rician Fading MAC with Memory

Proposition 5.1. *Consider a SISO Rician fading multiple-access channel with m users as defined in* (2.1) *and* (2.2)*. Then the sum-rate fading number can be upper-bounded by:*

$$\chi_{\text{MAC}} \leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}\in\mathcal{A}}} \left\{ \mathsf{E}\left[\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_d\mathbf{X}}{\|\mathbf{X}\|^2} \right) - \text{Ei}\left(-\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_d\mathbf{X}}{\|\mathbf{X}\|^2} \right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2} \right) - 1 \right] \right\} \tag{5.1}$$

where $\mathcal{A}$ is the set as defined in (4.13)*; D_d is the diagonal matrix defined as*

$$\mathsf{D}_d \triangleq \text{diag}\left(|d_1|^2, \ldots, |d_m|^2\right); \tag{5.2}$$

and D_ϵ is the diagonal matrix defined as

$$\mathsf{D}_\epsilon \triangleq \text{diag}\left(\epsilon_1^2, \ldots, \epsilon_m^2\right) \tag{5.3}$$

with prediction errors ϵ_i shown in (2.4).

This proposition shows that the upper bound of the fading number depends on not only the line-of-sight components but also the prediction errors. Compared with [5] and [7], the effect of memory is the third term in (5.1) which is independent of line-of-sight components.

5.2 The Fading Number of the Two-User SISO Rician Fading MAC with Memory

Theorem 5.2. *Consider a two-user SISO Rician fading multiple-access channel as defined in* (2.12)*. Then the sum-rate fading number is given by*

$$\chi_{\text{MAC-2}} = \max_{i\in\{1,2\}} \chi_{\text{SISO},i} \tag{5.4}$$

$$= \max_{i\in\{1,2\}} \left\{\log\left(|d_i|^2\right) - \text{Ei}\big(-|d_i|^2\big) - \log\left(\epsilon_i^2\right) - 1\right\} \tag{5.5}$$

This fading number of the two-user SISO MAC holds in all three cases when the peak-power constraint (2.7)*, the average-power constraint* (2.8)*, or the power-sharing average-power constraint* (2.9) *is considered.*

This theorem shows that the lower bound in (4.7) is tight. Moreover, the sum-rate capacity of the two-user SISO MAC is decided by the better one of the two users. Note that if the fading numbers of the two users are the same, the sum-rate capacity of MAC can be achieved by time-sharing.

5.3 The Fading Number of the m-User SISO Rician Fading MAC with Memory

Theorem 5.3. *Consider a SISO Rician fading multiple-access channel with m users as defined in* (2.1) *and* (2.2)*. Then the sum-rate fading number is given by*

$$\chi_{\text{MAC}} = \max_{i\in\{1,\ldots,m\}} \chi_{\text{SISO},i} \tag{5.6}$$

$$= \max_{i\in\{1,\ldots,m\}} \left\{\log\left(|d_i|^2\right) - \text{Ei}\big(-|d_i|^2\big) - \log\left(\epsilon_i^2\right) - 1\right\} \tag{5.7}$$

This fading number of the m-user SISO MAC holds in all three cases when the peak-power constraint (2.7)*, the average-power constraint* (2.8)*, or the power-sharing average-power constraint* (2.9) *is considered.*

This theorem shows that the sum-rate capacity of the m-user SISO MAC is decided by the best one of m users. Similarly to the two-user case, if more than one user has the best channel, *i.e.*, they have the maximum fading number, the sum-rate capacity of MAC can be achieved by time-sharing between these users.

In our analysis, we have allowed three different types of power constraints: an individual peak-power constraint for each user, an individual average-power constraint for each user, and a combined power-sharing average-power constraint among all users. The power-sharing constraint does not make sense in a practical setup as it requires the users to share a common battery, while their signals still are restricted to be independent. However, the inclusion of this case helps with the analysis. Furthermore, it turns out that the pessimistic results described above even hold if we allow for such power sharing.

Within the three types of constraint, we do allow for different power settings for different users as long as the constraints scale linearly (see the constants κ_i and $\bar{\kappa}$ in (2.7)–(2.9)).

Chapter 6

Derivation of Results

In this chapter, we give the proofs of the main results shown in Chapter 5. In Section 6.1, we derive an upper bound on the fading number of SISO MAC based on the proof of the MIMO case in [8]. In Section 6.2, the fading number of the two-user SISO Rician fading MAC with memory is derived by the concepts of Section 4.1 and Section 6.1. In Section 6.3, from the result of Section 6.2, we provide the proof of the exact fading number in the m-user case.

6.1 Derivation of Proposition 5.1

To upper-bound the fading number of SISO MAC, we follow the steps of MIMO case stated in [8]. Fix some power $\mathcal{E} > 0$, and let $\tau > 0$ be an arbitrary value. From Theorem 3.5, we can fix a positive integer κ, and let $\eta = \eta(\mathcal{E}, \tau) \in \mathbb{Z}^+$ and $Q^{\kappa+1}_{\mathcal{E},\tau} \in \mathcal{P}(\mathbb{C}^{n_\mathrm{T}\times(\kappa+1)})$ which is the set of all input distributions over $\mathbb{C}^{n_\mathrm{T}\times(\kappa+1)}$ on $\mathbb{C}^{n_\mathrm{T}}$. Moreover, let blocklength n and input $\mathbf{X}_1^n$ satisfying (3.18)–(3.22) so that

$$\mathrm{C}_{\mathrm{MAC}}(\mathcal{E}) \le \frac{1}{n} I\left(\mathbf{X}_1^n; Y_1^n\right) + \tau \tag{6.1}$$

$$= \frac{1}{n}\sum_{k=1}^{n} I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) + \tau. \tag{6.2}$$

For the region of $1 \le k \le \eta + \kappa - 1$ and $n - 2\eta + 3 \le k \le n$, we use the crude bound

$$I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) \le I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\right) \tag{6.3}$$

$$\le \mathrm{C}_{\mathrm{IID}}(\mathcal{E}) + I\left(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\right) \tag{6.4}$$

where $\mathrm{C}_{\mathrm{IID}}(\mathcal{E})$ denotes the sum-rate capacity of the SISO memoryless MAC fading channel as given in (2.15) and (2.16) with an available average power of at most $\mathcal{E}$ as guaranteed in (3.21) and (3.22). Here the first inequality can be derived as follows:

$$I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) = I\left(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\right) - I\left(Y_k; Y_1^{k-1}\right) \tag{6.5}$$

$$\le I\left(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\right) \tag{6.6}$$

$$= I\left(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.7}$$
$$\leq I\left(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \mathbf{H}_1^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.8}$$
$$= I\left(\mathbf{H}_1^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.9}$$
$$= I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_1^{k-1}; Y_k \,\middle|\, \mathbf{X}_k\right) \tag{6.10}$$
$$= I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_1^{k-1}; \mathbf{X}_k, Y_k\right) \tag{6.11}$$
$$\leq I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_1^{k-1}; \mathbf{H}_k, \mathbf{X}_k, Y_k\right) \tag{6.12}$$
$$= I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_1^{k-1}; \mathbf{H}_k\right) \tag{6.13}$$
$$\leq I\left(\mathbf{X}_k; Y_k\right) + I\left(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\right). \tag{6.14}$$

Here (6.5) follows from the chain rule; (6.6) follows from the non-negativity of mutual information; (6.7) follows because we prohibit feedback; (6.8) follows from the inclusion of the additional random vectors $\mathbf{H}_1^{k-1}$ in the mutual information term; (6.9) follows from the chain rule and the fact that if the past fading processes $\mathbf{H}_1^{k-1}$ and present input $\mathbf{X}_k$ are given, the past inputs $\mathbf{X}_1^{k-1}$ and outputs Y_1^{k-1} are independent of the present output Y_k; (6.10) follows from the chain rule; (6.11) follows from the chain rule and the fact that since $\mathbf{X}_k$ and $\mathbf{H}_1^{k-1}$ are independent, $I\left(\mathbf{H}_1^{k-1}; \mathbf{X}_k\right) = 0$; the two steps (6.12), (6.13) are similar to (6.8) and (6.9); and (6.14) follows once more from the inclusion of additional random vectors in the mutual information and from stationarity.

Because (6.4) is uniformly bounded in n, we conclude that

$$\mathsf{C}_{\mathrm{MAC}}(\mathcal{E}) \leq \varliminf_{n\uparrow\infty} \frac{1}{n} I\left(\mathbf{X}_1^n; Y_1^n\right) + \tau \tag{6.15}$$
$$= \varliminf_{n\uparrow\infty} \frac{1}{n-\kappa-3(\eta-1)} \sum_{k=\eta+\kappa}^{n-2(\eta-1)} I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) + \tau \tag{6.16}$$

Since Theorem 3.5 guarantees that every $(\kappa+1)$-block $(X_{k-\kappa}, \ldots, X_k)$ has the same distribution $Q_{\mathcal{E},\tau}^{\kappa+1}$, we only have to focus on the region $\eta + \kappa \leq k \leq n - 2(\eta-1)$.

Now, we further upper-bound $I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right)$ for such k to continue:

$$I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) = I\left(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\right) - I\left(Y_k; Y_1^{k-1}\right) \tag{6.17}$$
$$\leq I\left(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\right) \tag{6.18}$$
$$= I\left(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.19}$$
$$\leq I\left(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \mathbf{H}_1^{k-\kappa-1}, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.20}$$
$$= I\left(\mathbf{X}_{k-\kappa}^{k-1}, \mathbf{H}_1^{k-\kappa-1}, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}, \mathbf{X}_k; Y_k\right) \tag{6.21}$$
$$= I\left(\mathbf{X}_k; Y_k\right) + \underbrace{I\left(\mathbf{X}_{k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_k\right)}_{=0} + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right)$$

$$+ I\left(\mathbf{H}_1^{k-\kappa-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}\right) \tag{6.22}$$

$$= I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right)$$

$$+ I\left(\mathbf{H}_1^{k-\kappa-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}\right) \tag{6.23}$$

$$\leq I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right) + \delta(\kappa). \tag{6.24}$$

Here the three steps (6.17)–(6.19) are the same as (6.5)–(6.7); (6.20) follows from the inclusion of the additional random vectors $\mathbf{H}_1^{k-\kappa-1}$ and the random variables $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}$ in the mutual information term; (6.21) follows from the chain rule and the fact that if $\mathbf{H}_1^{k-\kappa-1}$, $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}$ and $\mathbf{X}_{k-\kappa}^k$ are given, the past output Y_1^{k-1} and input $\mathbf{X}_1^{k-\kappa-1}$ are independent of the present output Y_k; (6.22) follows from the chain rule; (6.23) follows from the fact that the past inputs $\mathbf{X}_{k-\kappa}^{k-1}$ and the present output Y_k are independent as the present input $\mathbf{X}_k$ is given; and (6.24) follows from [8, Lemma 18].

Note that $\delta(\kappa)$ does neither depend on k nor on the input $\{\mathbf{X}_k\}$ and monotonically tends to zero as κ tends to infinity due to the stationarity of $\{\mathbf{H}_k\}$.

We continue (6.24) as follows:

$$I\left(\mathbf{X}_1^n; Y_k \,\middle|\, Y_1^{k-1}\right) \leq I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right) + \delta(\kappa) \tag{6.25}$$

$$\leq I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k, \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right) + \delta(\kappa) \tag{6.26}$$

$$= I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right)$$

$$+ \underbrace{I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; Y_k \,\middle|\, \mathbf{X}_{k-\kappa}^k, \mathbf{H}_k^\mathsf{T}\mathbf{X}_k\right)}_{=0} + \delta(\kappa) \tag{6.27}$$

$$= I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\middle|\, \mathbf{X}_{k-\kappa}^k\right) + \delta(\kappa) \tag{6.28}$$

$$= I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=k-\kappa}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\middle|\, \hat{\mathbf{X}}_{k-\kappa}^k, \{\|\mathbf{X}_\ell\|\}_{k-\kappa}^k\right) + \delta(\kappa) \tag{6.29}$$

$$= I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=k-\kappa}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\middle|\, \hat{\mathbf{X}}_{k-\kappa}^k\right) + \delta(\kappa). \tag{6.30}$$

Here (6.26) follows from the inclusion of the additional random variable $\mathbf{H}_k^\mathsf{T}\mathbf{X}_k$ in the mutual information; (6.27) follows from the chain rule; (6.28) follows the fact that the additive noise Z_k is independent of the fading processes $\mathbf{H}_{k-\kappa}^{k-1}$; in (6.29), we split the vectors $\mathbf{X}_\ell$ up into magnitude $\|\mathbf{X}_\ell\|$ and direction $\hat{\mathbf{X}}_\ell$that $\hat{\mathbf{X}}_\ell \triangleq \frac{\mathbf{X}_\ell}{\|\mathbf{X}_\ell\|}$; and (6.30) follows from dividing each term by the magnitude of the input vectors and from the fact that $\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=k-\kappa}^k$ is independent of $\{\|\mathbf{X}_\ell\|\}_{\ell=k-\kappa}^k$ as $\hat{\mathbf{X}}_{k-\kappa}^k$ is given.

Note that (6.30) only depends on $\mathbf{X}_{k-\kappa}^k$ which has a distribution $Q_{\mathcal{E},\tau}^{\kappa+1}$ according to Theorem 3.5. As a result, using the stationarity and combining (6.30) with (6.16), we have

$$\mathrm{C}_{\mathrm{MAC}}(\mathcal{E}) \leq \varliminf_{n\uparrow\infty} \frac{1}{n-\kappa-3(\eta-1)} \sum_{k=\eta+\kappa}^{n-2(\eta-1)} \left(I(\mathbf{X}_k; Y_k) + I\left(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=k-\kappa}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\middle|\, \hat{\mathbf{X}}_{k-\kappa}^k\right)\right) + \delta(\kappa) + \tau \tag{6.31}$$

$$= \lim_{n\uparrow\infty} \frac{1}{n-\kappa-3(\eta-1)} \sum_{k=\eta+\kappa}^{n-2(\eta-1)} \Big(I\left(\mathbf{X}_k; \mathbf{H}_{\eta+\kappa}^{\mathsf{T}} \mathbf{X}_k + Z_{\eta+\kappa}\right)$$
$$+ I\left(\mathbf{H}_{\eta+\kappa-1}^{\mathsf{T}} \hat{\mathbf{X}}_{k-1}, \ldots, \mathbf{H}_{\eta}^{\mathsf{T}} \hat{\mathbf{X}}_{k-\kappa}; \mathbf{H}_{\eta+\kappa}^{\mathsf{T}} \hat{\mathbf{X}}_k \,\Big|\, \hat{\mathbf{X}}_{k-\kappa}^{k}\right)\Big) + \delta(\kappa) + \tau \tag{6.32}$$
$$= I\left(\mathbf{X}_{\eta+\kappa}; \mathbf{H}_{\eta+\kappa}^{\mathsf{T}} \mathbf{X}_{\eta+\kappa} + Z_{\eta+\kappa}\right) + I\left(\{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=\eta}^{\eta+\kappa-1}; \mathbf{H}_{\eta+\kappa}^{\mathsf{T}} \hat{\mathbf{X}}_{\eta+\kappa} \,\Big|\, \hat{\mathbf{X}}_{\eta}^{\eta+\kappa}\right)$$
$$+ \delta(\kappa) + \tau. \tag{6.33}$$

Here in (6.32) we shift $\mathbf{H}_k$ and Z_k to $\mathbf{H}_{\eta+\kappa}$ and $Z_{\eta+\kappa}$ due to the stationarity of $\{\mathbf{H}_k, Z_k\}$; and (6.33) follows from the fact that for all $k \in \{\eta+\kappa, \ldots, n-2(\eta-1)\}$ the distribution of $\mathbf{X}_{k-\kappa}^{k}$ is $Q_{\mathcal{E},\tau}^{\kappa+1}$ given in Theorem 3.5.

To make (6.33) easy to read, we introduce a slight misuse in notation: for pure notational convenience we will assume from now on that $\mathbf{X}_{-\kappa}^{0} \sim Q_{\mathcal{E},\tau}^{\kappa+1}$, *i.e.*, that from now on $\mathbf{X}_{-\kappa}^{0}$ is quasi-stationary. Note that there is no contradiction between this notation and the edge-effects of Theorem 3.5 since this is a notational choice. Hence, we can drop η and rewrite (6.33) as follows:

$$\mathsf{C}_{\text{MAC}}(\mathcal{E}) \le I\left(\mathbf{X}_0; \mathbf{H}_0^{\mathsf{T}} \mathbf{X}_0 + Z_0\right) + I\left(\{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}; \mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}\right) + \delta(\kappa) + \tau. \tag{6.34}$$

We can find that the first term in (6.34) is the mutual information of the SISO memoryless MAC and independent of the past. Therefore, the memory effect only comes from the second term, and we upper-bound the second term to continue:

$$I\left(\{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}; \mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}\right)$$
$$= h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}\right) - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}, \{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}\right) \tag{6.35}$$
$$= h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_0\right) - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}, \{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}\right) \tag{6.36}$$
$$= \log \pi e - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}, \{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}\right) \tag{6.37}$$
$$\le \log \pi e - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}, \mathbf{H}_{-\kappa}^{-1}, \{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}\right) \tag{6.38}$$
$$= \log \pi e - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_{-\kappa}^{0}, \mathbf{H}_{-\kappa}^{-1}\right) \tag{6.39}$$
$$= \log \pi e - h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0 \,\Big|\, \hat{\mathbf{X}}_0, \mathbf{H}_{-\kappa}^{-1}\right) \tag{6.40}$$
$$= \log \pi e - \mathsf{E}\left[h\left(\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{x}}_0 \,\Big|\, \hat{\mathbf{X}}_0 = \hat{\mathbf{x}}_0, \mathbf{H}_{-\kappa}^{-1}\right)\right] \tag{6.41}$$
$$= \mathsf{E}[\log \pi e] - \mathsf{E}\left[\log \pi e\left(\epsilon_{1,\kappa}^2 |\hat{X}_{1,0}|^2 + \cdots + \epsilon_{m,\kappa}^2 |\hat{X}_{m,0}|^2\right)\right] \tag{6.42}$$
$$= \mathsf{E}\left[\log \pi e - \log \pi e\left(\epsilon_{1,\kappa}^2 |\hat{X}_{1,0}|^2 + \cdots + \epsilon_{m,\kappa}^2 |\hat{X}_{m,0}|^2\right)\right] \tag{6.43}$$
$$= -\mathsf{E}\left[\log\left(\epsilon_{1,\kappa}^2 |\hat{X}_{1,0}|^2 + \cdots + \epsilon_{m,\kappa}^2 |\hat{X}_{m,0}|^2\right)\right]. \tag{6.44}$$

Here (6.35) follows the definition of the mutual information; (6.36) follows the fact that if the present input $\hat{\mathbf{X}}_0$ is given, $\mathbf{H}_0^{\mathsf{T}} \hat{\mathbf{X}}_0$ is independent of the past input $\hat{\mathbf{X}}_{-\kappa}^{-1}$; (6.38) follows that the conditioning reduces differential entropy; (6.39) follows the fact that when $\hat{\mathbf{X}}_{-\kappa}^{0}$ and $\mathbf{H}_{-\kappa}^{-1}$ are given, $\{\mathbf{H}_{\ell}^{\mathsf{T}} \hat{\mathbf{X}}_{\ell}\}_{\ell=-\kappa}^{-1}$ can be dropped because it is decided by these two terms;

(6.40) follows the same step as (6.36); and in (6.42), $\epsilon_{i,\kappa}$ is the prediction error for i-th user from the past fading $\{H_{i,k}\}_{k=-\kappa}^{-1}$.

Combined with (6.44), (6.34) becomes

$$
\begin{aligned}
\mathsf{C}_{\mathrm{MAC}}(\mathcal{E}) &\leq I\left(\mathbf{X}_0; \mathbf{H}_0^{\mathsf{T}}\mathbf{X}_0 + Z_0\right) - \mathsf{E}\left[\log\left(\epsilon_{1,\kappa}^2|\hat{X}_{1,0}|^2 + \cdots + \epsilon_{m,\kappa}^2|\hat{X}_{m,0}|^2\right)\right] + \delta(\kappa) + \tau && (6.45)\\
&= I\left(\mathbf{X}_0; \mathbf{H}_0^{\mathsf{T}}\mathbf{X}_0 + Z_0\right) - \mathsf{E}\left[\log\left(\epsilon_1^2|\hat{X}_{1,0}|^2 + \cdots + \epsilon_m^2|\hat{X}_{m,0}|^2\right)\right] + \tau && (6.46)\\
&= I\left(\mathbf{X}_0; \mathbf{H}_0^{\mathsf{T}}\mathbf{X}_0 + Z_0\right) - \mathsf{E}\left[\log\left(\frac{\epsilon_{1,\kappa}^2|X_{1,0}|^2 + \cdots + \epsilon_{m,\kappa}^2|X_{m,0}|^2}{|X_{1,0}|^2 + \cdots + |X_{m,0}|^2}\right)\right] + \tau && (6.47)\\
&= I\left(\mathbf{X}_0; \mathbf{H}_0^{\mathsf{T}}\mathbf{X}_0 + Z_0\right) - \mathsf{E}\left[\log\left(\frac{\mathbf{X}_0^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}_0}{\|\mathbf{X}_0\|^2}\right)\right] + \tau. && (6.48)
\end{aligned}
$$

Here in (6.46), we let κ tends to infinity which makes sure that $\delta(\kappa) \to 0$ as can be seen from [8, Lemma 18] and the prediction error $\epsilon_{i,\kappa}$ is equal to ϵ_i which is shown as (2.4); (6.47) follows the definition $\hat{\mathbf{X}}_\ell \triangleq \frac{\mathbf{X}_\ell}{\|\mathbf{X}_\ell\|}$; and (6.48) follows from the Rayleigh-Ritz Theorem [1, Theorem 4.2.2], and we have defined the matrix

$$
\mathsf{D}_\epsilon \triangleq \mathrm{diag}\left(\epsilon_1^2, \ldots, \epsilon_m^2\right). \tag{6.49}
$$

Since every term in (6.48) is independent of the time, we can drop the time parameter and rewrite (6.48) as follows

$$
\begin{aligned}
\mathsf{C}_{\mathrm{MAC}}(\mathcal{E}) &\leq I\left(\mathbf{X}; \mathbf{H}^{\mathsf{T}}\mathbf{X} + Z\right) - \mathsf{E}\left[\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right] + \tau && (6.50)\\
&= I\left(\mathbf{X}; Y\right) - \mathsf{E}\left[\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right] + \tau && (6.51)\\
&\leq -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right)\right] - \mathsf{E}\left[\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right] + \epsilon_\nu \\
&\quad + \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1 + \|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta} + \tau && (6.52)\\
&= -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right] + \epsilon_\nu \\
&\quad + \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1 + \|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta} + \tau. && (6.53)
\end{aligned}
$$

Here (6.52) follows the Lemma 4.2.

Note that this bound still depends on the distribution $Q_{\mathcal{E},\tau}^{\kappa+1}$ which is guaranteed to exist by Theorem 3.5. However, the exact form of the bound is unknown. Fortunately, the bound is independent to the time, and we can upper-bound this expression by maximizing it over all probability measures $Q_{\mathbf{X}}$:

$$
\begin{aligned}
\mathsf{C}_{\mathrm{MAC}}(\mathcal{E}) \leq \sup_{Q_{\mathcal{E}} \in \mathcal{A}} \Bigg\{ &-1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_\epsilon\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right] + \epsilon_\nu \\
&+ \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1 + \|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta}
\end{aligned}
$$

$$+\tau\Bigg\} \tag{6.54}$$

$$= \sup_{Q_{\mathcal{E}} \in \mathcal{A}} \left\{ -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right]\right\} + \epsilon_\nu$$
$$+ \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta}$$
$$+ \tau \tag{6.55}$$

Here we define $\mathcal{A}$ to be the set of all probaility measures in $Q_{\mathbf{X}}$ satisfying the constraints that all users are independent (2.6), the power-sharing average-power constraint (2.9) and that the input distribution of at least one user escapes to infinity as the available power $\mathcal{E}$ tends to infinity shown in Proposition 3.4, *i.e.*,

$$\mathcal{A} \triangleq \Bigg\{ \{Q_{\mathbf{X}}\}_{\mathcal{E}>0} : X_i \perp\!\!\!\perp X_j, \quad i,j = 1,\ldots,m, \quad \forall i \neq j; \quad \mathsf{E}\left[\sum_{i=1}^{m} \|\mathbf{X}_i\|^2\right] \leq \bar{\kappa}\mathcal{E};$$
$$\lim_{\mathcal{E}\uparrow\infty} Q_{\mathcal{E}}\left(\bigcup_{i=1}^{m}\left\{|X_i|^2 \geq \frac{\mathcal{E}_0}{m}\right\}\right) = 1 \text{ for any fixed } \mathcal{E}_0 > 0 \Bigg\}. \tag{6.56}$$

Note that we drop the time parameter k here since (6.55) is independent of the time.

From the definition of the MAC fading number (4.3) and (6.55), we can derive the following upper bound on the MAC fading number:

$$\chi_{\text{MAC}} = \varlimsup_{\mathcal{E}\uparrow\infty} \left\{ \mathsf{C}_{\text{MAC}}(\mathcal{E}) - \log\left(1 + \log\left(1 + \frac{\mathcal{E}}{\sigma^2}\right)\right)\right\} \tag{6.57}$$

$$= \varlimsup_{\mathcal{E}\uparrow\infty} \left\{ \sup_{Q_{\mathcal{E}} \in \mathcal{A}} \left\{ -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right)\right]\right\} + \epsilon_\nu \right.$$
$$+ \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta} + \tau$$
$$\left. - \log\left(1 + \log\left(1 + \frac{\mathcal{E}}{\sigma^2}\right)\right)\right\} \tag{6.58}$$

Next, we choose the free parameters α and β as follows:

$$\alpha \triangleq \alpha(\mathcal{E}) = \frac{\nu}{\log\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right)} \tag{6.59}$$

$$\beta \triangleq \beta(\mathcal{E}) = \frac{1}{\alpha(\mathcal{E})} e^{\nu/\alpha(\mathcal{E})} \tag{6.60}$$

for some constant $\nu \geq 0$, which leads to the following asymptotic behavior:

$$\varlimsup_{\mathcal{E}\uparrow\infty} \left\{ \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) - \log\frac{1}{\alpha}\right\} = \log(1 - e^{-\nu}); \tag{6.61}$$

$$\varlimsup_{\mathcal{E}\uparrow\infty} \alpha\left(\log\beta - \log\sigma^2 + \gamma\right) = \nu; \tag{6.62}$$

$$\varlimsup_{\mathcal{E}\uparrow\infty} \left\{ \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta}\right\} = 0; \tag{6.63}$$

$$\varlimsup_{\mathcal{E}\uparrow\infty} \left\{ \log\frac{1}{\alpha} - \log\left(1 + \log\left(1 + \frac{\mathcal{E}}{\sigma^2}\right)\right)\right\} = -\log\nu. \tag{6.64}$$

(Compare with [3, Appendix VII], [6, Sec. B.5.9])

As a result, (6.58) can be upper-bounded as follows:

$$\begin{aligned}\chi_{\text{MAC}} \leq \varlimsup_{\mathcal{E}\uparrow\infty} \Bigg\{ \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ -1 + \mathsf{E}\Bigg[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right)\Bigg]\Bigg\} + \epsilon_\nu \\ + \alpha(\log\beta - \log\sigma^2 + \gamma) + \log\Gamma\left(\alpha, \frac{\nu}{\beta}\right) + \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E} + \sigma^2\right) + \frac{\nu}{\beta} + \tau \\ - \log\left(1 + \log\left(1 + \frac{\mathcal{E}}{\sigma^2}\right)\right)\Bigg\} \qquad (6.65)\end{aligned}$$

$$\begin{aligned}\leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ \mathsf{E}\Bigg[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\} + \epsilon_\nu + \nu \\ + \log(1 - e^{-\nu}) - \log\nu + \tau. \qquad (6.66)\end{aligned}$$

Let ν tend to zero which makes sure that $\epsilon_\nu \to 0$ as can be seen from (A,6). Since τ is an arbitrary value, (6.66) can be rewritten as

$$\chi_{\text{MAC}} \leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ \mathsf{E}\Bigg[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\}. \qquad (6.67)$$

Furthermore, we define

$$f(\xi) \triangleq \log(\xi) - \mathrm{Ei}(-\xi) \qquad (6.68)$$

$$\mathsf{F}(\mathbf{X}) \triangleq -\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1 \qquad (6.69)$$

$$\mathsf{G}_1(\mathbf{X}) \triangleq \frac{|d_1|^2|X_1|^2 + \cdots + |d_m|^2|X_m|^2}{|X_1|^2 + \cdots + |X_m|^2} \qquad (6.70)$$

$$\mathsf{G}_2(\mathbf{X}) \triangleq \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m} \frac{|d_i||X_i||d_j||X_j|}{|X_1|^2 + \cdots + |X_m|^2}, \qquad (6.71)$$

and upper-bound the RHS in (6.67) as follows:

$$\begin{aligned}&\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ \mathsf{E}\Bigg[\log\left(\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^{\mathsf{T}}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\} \\ &= \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ \mathsf{E}\Bigg[\log\left(\frac{|d_1X_1 + \cdots + d_mX_m|^2}{|X_1|^2 + \cdots + |X_m|^2}\right) - \mathrm{Ei}\left(-\frac{|d_1X_1 + \cdots + d_mX_m|^2}{|X_1|^2 + \cdots + |X_m|^2}\right) \\ &\qquad - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\} \qquad (6.72) \\ &= \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{ \mathsf{E}\Bigg[\log\left(\frac{|d_1X_1|^2 + \cdots + |d_mX_m|^2}{|X_1|^2 + \cdots + |X_m|^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m} \frac{|d_iX_i|\cdot|d_jX_j|}{|X_1|^2 + \cdots + |X_m|^2}\right) \\ &\qquad - \mathrm{Ei}\left(-\left(\frac{|d_1X_1|^2 + \cdots + |d_mX_m|^2}{|X_1|^2 + \cdots + |X_m|^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m} \frac{|d_iX_i|\cdot|d_jX_j|}{|X_1|^2 + \cdots + |X_m|^2}\right)\right)\end{aligned}$$

$$- \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\} \tag{6.73}$$

$$\begin{aligned}
&\leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Bigg\{\mathsf{E}\Bigg[\log\Bigg(\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\Bigg) \\
&\qquad - \mathrm{Ei}\Bigg(-\Bigg(\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\Bigg)\Bigg) \\
&\qquad - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\Bigg]\Bigg\}
\end{aligned} \tag{6.74}$$

$$= \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \left\{\mathsf{E}\left[f\left(\mathsf{G}_1(\mathbf{X}) + \mathsf{G}_2(\mathbf{X})\right) + \mathsf{F}(\mathbf{X})\right]\right\} \tag{6.75}$$

$$\leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathbf{X}}\in\mathcal{A}} \sup_{Q_{\tilde{\mathbf{X}}}\in\mathcal{A}} \left\{\mathsf{E}_{\mathbf{X},\tilde{\mathbf{X}}}\left[f\left(\mathsf{G}_1(\mathbf{X}) + \mathsf{G}_2(\tilde{\mathbf{X}})\right) + \mathsf{F}(\mathbf{X})\right]\right\} \tag{6.76}$$

$$\leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathbf{X}}\in\mathcal{A}} \sup_{Q_{\tilde{\mathbf{X}}}\in\mathcal{A}} \left\{\mathsf{E}_{\mathbf{X}}\left[f\left(\mathsf{G}_1(\mathbf{X}) + \mathsf{E}_{\tilde{\mathbf{X}}}\left[\mathsf{G}_2(\tilde{\mathbf{X}})\right]\right) + \mathsf{F}(\mathbf{X})\right]\right\} \tag{6.77}$$

$$\leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathbf{X}}\in\mathcal{A}} \sup_{Q_{\tilde{\mathbf{X}}}\in\mathcal{A}} \left\{\mathsf{E}_{\mathbf{X}}\left[f\left(\mathsf{G}_1(\mathbf{X}) + \sup_{Q_{\tilde{\mathbf{X}}}\in\mathcal{A}}\mathsf{E}_{\tilde{\mathbf{X}}}\left[\mathsf{G}_2(\tilde{\mathbf{X}})\right]\right) + \mathsf{F}(\mathbf{X})\right]\right\} \tag{6.78}$$

$$= \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathbf{X}}\in\mathcal{A}} \left\{\mathsf{E}_{\mathbf{X}}\left[f\left(\mathsf{G}_1(\mathbf{X}) + \sup_{Q_{\tilde{\mathbf{X}}}\in\mathcal{A}}\mathsf{E}_{\tilde{\mathbf{X}}}\left[\mathsf{G}_2(\tilde{\mathbf{X}})\right]\right) + \mathsf{F}(\mathbf{X})\right]\right\}. \tag{6.79}$$

Here (6.74) follows from the Cauchy-Schwarz inequality and the fact that $\xi \mapsto f(\xi)$ is monotonically increasing; in (6.76), we replace $\mathbf{X}$ in $\mathsf{G}_2(\mathbf{X})$ by $\tilde{\mathbf{X}}$ and take the supremum over all $Q_{\tilde{\mathbf{X}}}$ without the constraint that $\tilde{\mathbf{X}} = \mathbf{X}$; (6.77) follows from the Jensen's inequality; and (6.78) follows because $f(\xi)$ is monotonically increasing.

Since $Q_{\tilde{\mathbf{X}}}$ is independent of $Q_{\mathbf{X}}$, we can regard the $\tilde{\mathbf{X}}$-term in (6.79) as a constant upper-bounded as follows:

$$\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\mathsf{G}_2(\tilde{\mathbf{X}})\right] = \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m} \mathsf{E}\left[\frac{|d_i||\tilde{X}_i||d_j||\tilde{X}_j|}{|\tilde{X}_1|^2+\cdots+|\tilde{X}_m|^2}\right] \tag{6.80}$$

$$\begin{aligned}
&\leq 2\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \sum_{j=2}^{m} \mathsf{E}\left[\frac{|d_1||\tilde{X}_1||d_j||\tilde{X}_j|}{|\tilde{X}_1|^2+\cdots+|\tilde{X}_m|^2}\right] \\
&\quad + 2\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \sum_{i=2}^{m}\sum_{j=i+1}^{m} \mathsf{E}\left[\frac{|d_i||\tilde{X}_i||d_j||\tilde{X}_j|}{|\tilde{X}_1|^2+\cdots+|\tilde{X}_m|^2}\right]
\end{aligned} \tag{6.81}$$

$$\leq 2\sum_{j=2}^{m} \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_1||\tilde{X}_1||d_j||\tilde{X}_j|}{|\tilde{X}_1|^2+\cdots+|\tilde{X}_m|^2}\right]$$

$$+ 2\sum_{i=2}^{m}\sum_{j=i+1}^{m}\overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\mathsf{E}\left[\frac{|d_i||\tilde{X}_i||d_j||\tilde{X}_j|}{|\tilde{X}_1|^2+\cdots+|\tilde{X}_m|^2}\right] \tag{6.82}$$

$$= 2\cdot 0 + 2\cdot 0 = 0. \tag{6.83}$$

Here in (6.81), we separate the expectations into two kinds; in (6.82), we split the supremum into many separate suprema; and (6.83) follows from Lemma 4.3.

Therefore, combined (6.67) with (6.79) and (6.83), we have

$$\chi_{\text{MAC}} \le \overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\left\{\mathsf{E}[f(\mathsf{G}_1(\mathbf{X}))+\mathsf{F}(\mathbf{X})]\right\} \tag{6.84}$$

$$= \overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\left\{\mathsf{E}\left[\log\left(\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2}\right) - \text{Ei}\left(-\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\right]\right\} \tag{6.85}$$

$$= \overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\left\{\mathsf{E}\left[\log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_d\mathbf{X}}{\|\mathbf{X}\|^2}\right) - \text{Ei}\left(-\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_d\mathbf{X}}{\|\mathbf{X}\|^2}\right) - \log\left(\frac{\mathbf{X}^{\mathsf{T}}\mathsf{D}_{\epsilon}\mathbf{X}}{\|\mathbf{X}\|^2}\right) - 1\right]\right\}. \tag{6.86}$$

Here (6.86) follows from the Rayleigh-Ritz Theorem [1, Theorem 4.2.2], and we have defined the matrix

$$\mathsf{D}_d \triangleq \text{diag}\left(|d_1|^2,\ldots,|d_m|^2\right). \tag{6.87}$$

6.2 Derivation of Theorem 5.2

From Proposition 5.1, we have an upper bound on fading number of the two-user SISO MAC with memory:

$$\chi_{\text{MAC-2}} \le \overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\left\{\mathsf{E}\left[\log\left(\frac{|d_1|^2|X_1|^2+|d_2|^2|X_2|^2}{|X_1|^2+|X_2|^2}\right) - \text{Ei}\left(-\frac{|d_1|^2|X_1|^2+|d_2|^2|X_2|^2}{|X_1|^2+|X_2|^2}\right) - \log\left(\frac{\epsilon_1^2|X_1|^2+\epsilon_2^2|X_2|^2}{|X_1|^2+|X_2|^2}\right) - 1\right]\right\} \tag{6.88}$$

$$= \overline{\lim_{\mathcal{E}\uparrow\infty}}\sup_{Q_{\mathcal{E}}\in\mathcal{A}}\left\{\mathsf{E}\left[\log\left(|d_1|^2|\hat{X}_1|^2+|d_2|^2|\hat{X}_2|^2\right) - \text{Ei}\left(-\left(|d_1|^2|\hat{X}_1|^2+|d_2|^2|\hat{X}_2|^2\right)\right) - \log\left(\epsilon_1^2|\hat{X}_1|^2+\epsilon_2^2|\hat{X}_2|^2\right) - 1\right]\right\}. \tag{6.89}$$

Note that we have a lower bound on fading number of the two-user SISO MAC fading channel (4.7). If the upper bound of fading number (6.89) exists in the condition that only one user communicates and the other one is switched off, Theorem 5.2 is proved.

Since $\hat{\mathbf{X}}$ is an unit vector, *i.e.*,

$$|\hat{X}_1|^2 + |\hat{X}_2|^2 = 1, \tag{6.90}$$

we can we use ϑ with $0 \le \vartheta \le 1$ to denote the power allocation of the first user such that

$$|\hat{X}_1|^2 = \vartheta, \tag{6.91}$$

$$|\hat{X}_2|^2 = 1 - \vartheta, \quad 0 \le \vartheta \le 1. \tag{6.92}$$

And we have

$$\chi_{\text{MAC-2}} \le \max_{0\le\vartheta\le 1} \big\{\log\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big) - \text{Ei}\big(-\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big)\big) \\ - \log\big(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\big) - 1\big\}. \tag{6.93}$$

Here we drop the moduli of line-of-sight components because they are given and we can assume they are always real.

To finish the proof, our main purpose is to prove that the RHS of (6.93) is always on the boundary, *i.e.*, $\vartheta = 0$ or $\vartheta = 1$. Since it has four arbitrary values, it's difficult to analyze. To make it easier, we separate the four arbitrary values into three cases:

1. $d_1 > d_2$ and $\epsilon_1 < \epsilon_2$,
2. $d_1 > d_2$, $\epsilon_1 > \epsilon_2$, and $\frac{d_1}{\epsilon_1} < \frac{d_2}{\epsilon_2}$,
3. $d_1 > d_2$, $\epsilon_1 > \epsilon_2$, but $\frac{d_1}{\epsilon_1} > \frac{d_2}{\epsilon_2}$.

In the first case, the first user has the larger line-of-sight component, and the smaller prediction error. We know that the larger the line-of-sight component is, and the smaller the prediction error is, the better the channel is. Therefore, it is obvious that the first user has the better channel.

The (6.93) becomes

$$\chi_{\text{MAC-2}} \le \max_{0\le\vartheta\le 1} \big\{\log\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big) - \text{Ei}\big(-\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big)\big) \\ - \log\big(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\big) - 1\big\} \tag{6.94}$$

$$\le \max_{0\le\vartheta\le 1} \big\{\log\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big) - \text{Ei}\big(-\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big)\big) - 1\big\} \\ + \max_{0\le\vartheta\le 1} \big\{-\log\big(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\big)\big\} \tag{6.95}$$

$$= \log\big(d_1^2\big) - \text{Ei}\big(-d_1^2\big) - 1 - \log\big(\epsilon_1^2\big). \tag{6.96}$$

Here in (6.95), we split the maximum into two maximums, and (6.96) follows form the fact that $\xi \mapsto \log(\xi) - \text{Ei}(-\xi) - 1$ is monotonically increasing (see Figure 6.1) and $\xi \mapsto -\log(\xi)$ is monotonically decreasing. In this case, the first user is switched on and the second user is switched off at all time. The maximum is achieved at $\vartheta = 1$.

Next, we look at the second case. The first user has the better line-of-sight component, but the worse prediction error. However, the impact of the prediction error is very large such that the second user has the better ratio of line-of-sight component to prediction error. We can prove that the second user has the better channel:

$$\chi_{\text{MAC-2}} \le \max_{0\le\vartheta\le 1} \big\{\log\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big) - \text{Ei}\big(-\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big)\big) \\ - \log\big(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\big) - 1\big\} \tag{6.97}$$

$$= \max_{0\le\vartheta\le 1} \left\{\log\left(\frac{d_1^2\vartheta + d_2^2(1-\vartheta)}{\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)}\right) - \text{Ei}\big(-\big(d_1^2\vartheta + d_2^2(1-\vartheta)\big)\big) - 1\right\} \tag{6.98}$$

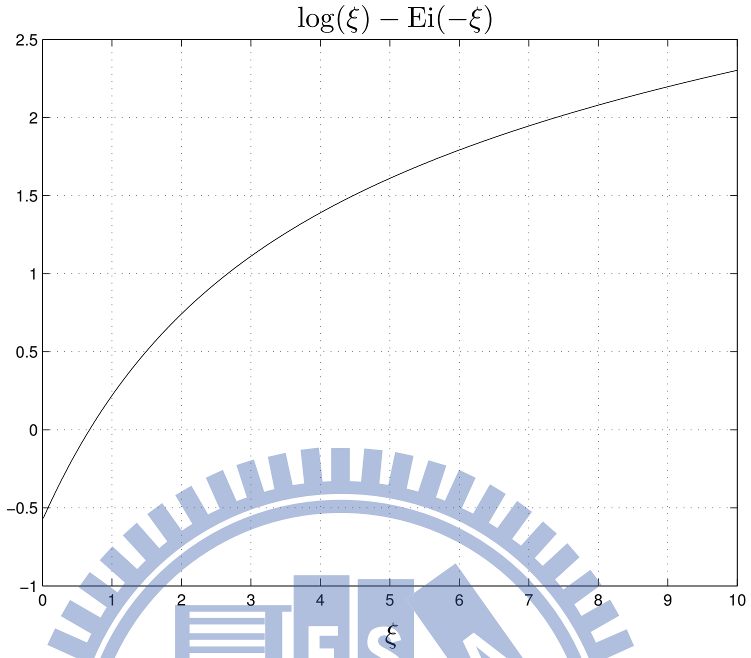

Figure 6.1: The plot of $\xi \mapsto \log(\xi) - \mathrm{Ei}(-\xi)$ for ξ from 0 to 10.

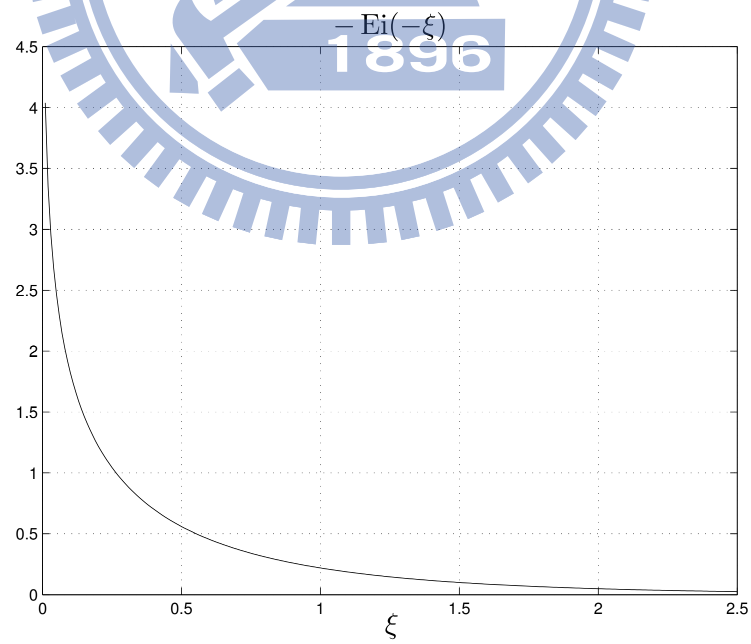

Figure 6.2: The plot of $\xi \mapsto -\mathrm{Ei}(-\xi)$ for ξ from 0 to 2.5.

$$\leq \max_{0\leq\vartheta\leq 1}\left\{\log\left(\frac{d_1^2\vartheta + d_2^2(1-\vartheta)}{\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)}\right)\right\} + \max_{0\leq\vartheta\leq 1}\left\{-\operatorname{Ei}\bigl(-\bigl(d_1^2\vartheta + d_2^2(1-\vartheta)\bigr)\bigr)\right\} - 1 \tag{6.99}$$

$$= \log\left(\frac{d_2^2}{\epsilon_2^2}\right) - \operatorname{Ei}\bigl(d_2^2\bigr) - 1. \tag{6.100}$$

Here in (6.99), we split the maximum into two maximums, and (6.100) follows from the fact that $\xi \mapsto \log(\xi)$ is monotonically increasing and $\xi \mapsto -\operatorname{Ei}(-\xi)$ is monotonically decreasing (see Figure 6.2). In this case, the first user is switched off and the second user is switched on at all time. The maximum is achieved at $\vartheta = 0$.

In the third case, the first user has the better line-of-sight, but worse prediction error. However, the impact is not large enough to make the first user worse that we are not sure which user has the better channel. Therefore, we have to try another way to analyze (6.93) in this case.

Now, we focus on the function

$$\varphi : \vartheta \mapsto \log\left(\frac{d_1^2\vartheta + d_2^2(1-\vartheta)}{\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)}\right) - \operatorname{Ei}\bigl(-\bigl(d_1^2\vartheta + d_2^2(1-\vartheta)\bigr)\bigr) - 1. \tag{6.101}$$

It consists of three terms: a log term, an exponential integral term and a constant 1. The shape of (6.101) only depends on the log term and the exponential integral term. With these two terms being monotonically concave or convex, the shape can only be one out of four cases, as shown in Figure 6.3: monotonically increasing, monotonically decreasing, convex, or concave. Of these four cases, only the concave case breaks our proof that the maximum may appear in the middle of the interval between $\vartheta = 0$ and $\vartheta = 1$. If we want to prove the maximum only exists on the boundary, we have to show the shape of (6.101) is never concave.

Note that the slope decides how the shape looks like. Hence we describe the shape of (6.101) by the sign of the slope: if the slope is always positive, the shape is monotonically increasing; similarly, if the slope is always negative, the shape is monotonically decreasing; if the slope goes from negative to positive as ϑ goes from 0 to 1, the shape is convex; and if the slope goes from positive to negative, the shape is concave.

To find the slope, we look at the first differential of (6.101):

$$\frac{\partial\varphi}{\partial\vartheta} = \frac{d_1^2\epsilon_2^2 - d_2\epsilon_1^2 - \bigl(d_1^2 - d_2^2\bigr)\frac{\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)}{e^{d_1^2\vartheta+d_2^2(1-\vartheta)}}}{\bigl(d_1^2\vartheta + d_2^2(1-\vartheta)\bigr)\bigl(\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)\bigr)}. \tag{6.102}$$

Because in our assumption, the terms $\bigl(d_1^2\vartheta + d_2^2(1-\vartheta)\bigr)$ and $\bigl(\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)\bigr)$ in the denominator are always positive, *i.e.*, they have nothing to do with the sign of the slope. As a result, we can drop them, and the sign of the slope is decided by:

$$d_1^2\epsilon_2^2 - d_2\epsilon_1^2 - \bigl(d_1^2 - d_2^2\bigr)\frac{\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)}{e^{d_1^2\vartheta+d_2^2(1-\vartheta)}}. \tag{6.103}$$

To simplify (6.103), we define

$$c \triangleq d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2 \tag{6.104}$$

$$f(\vartheta) \triangleq \bigl(d_1^2 - d_2^2\bigr)\frac{\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)}{e^{d_1^2\vartheta+d_2^2(1-\vartheta)}}, \tag{6.105}$$

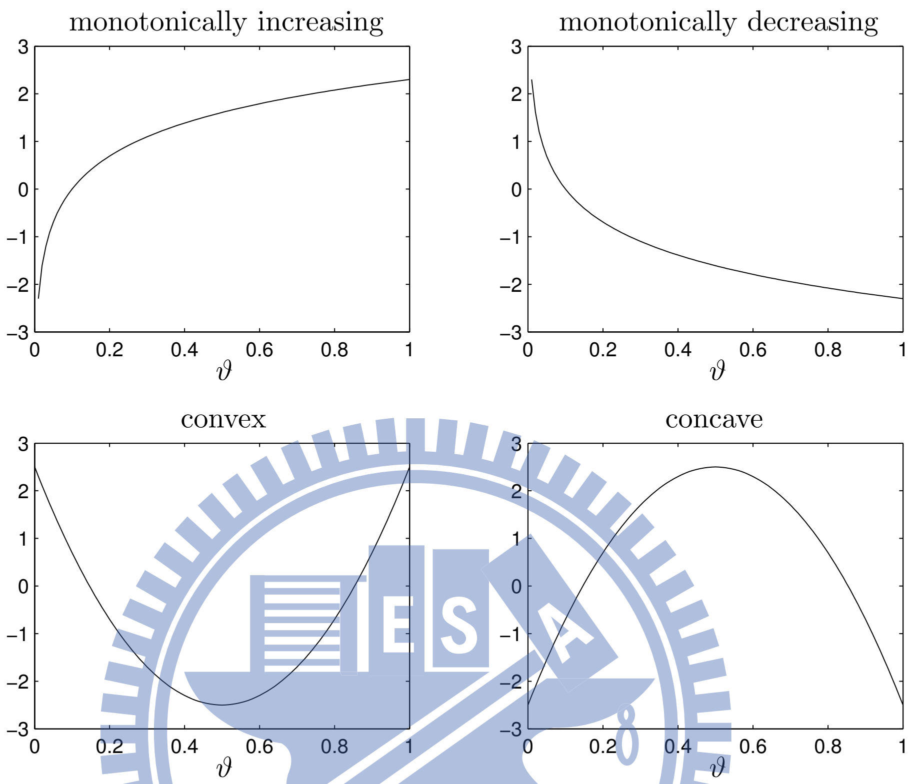

Figure 6.3: The four possible shapes of (6.101)

and rewrite (6.103) as

$$c - f(\vartheta). \tag{6.106}$$

It means that the sign of the slope is decided by the relationship between c and $f(\vartheta)$.

We consider two cases first as shown in Figure 6.4:

$$c \leq f(\vartheta), \quad 0 \leq \vartheta \leq 1 \tag{6.107}$$

and

$$c \geq f(\vartheta), \quad 0 \leq \vartheta \leq 1. \tag{6.108}$$

In the first case, $c - f(\vartheta)$ is always negative in the interval between $\vartheta = 0$ and $\vartheta = 1$, and therefore the shape is monotonically decreasing. The maximum is achieved as $\vartheta = 0$. In the other case, $c - f(\vartheta)$ is always positive in the interval between $\vartheta = 0$ and $\vartheta = 1$, and hence the shape is monotonically increasing. The maximum is achieved as $\vartheta = 1$. These two cases are in agreement with our proof. So we focus on the situation when c cuts through $f(\vartheta)$.

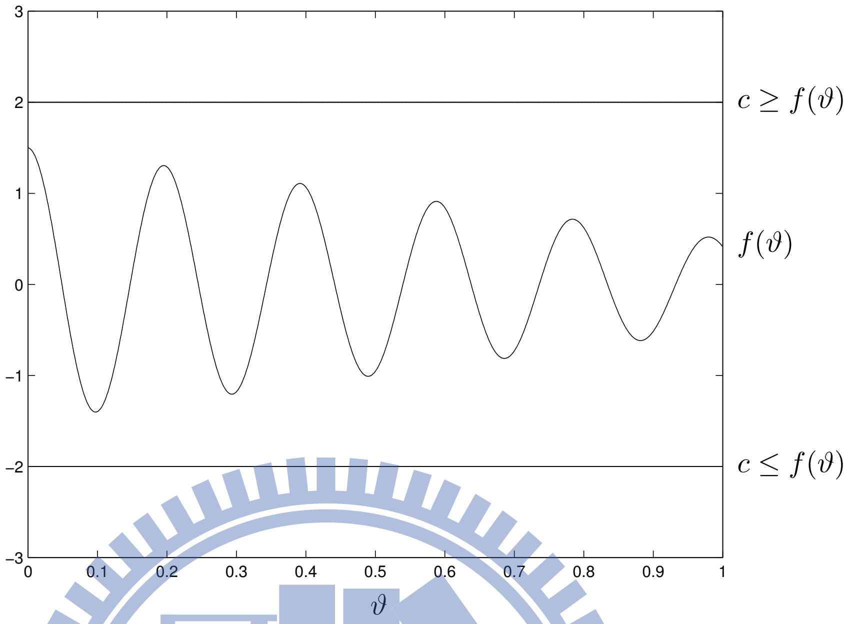

Figure 6.4: The two cases of $c > f(\vartheta)$ and $c < f(\vartheta)$.

Since c is independent of ϑ, the change of the slope only depends on $f(\vartheta)$. Therefore, we follow the same step to analyze $f(\vartheta)$ by looking at its first derivative:

$$\frac{\partial f(\vartheta)}{\partial \vartheta} = \frac{d_1^2 - d_2^2}{e^{d_1^2\vartheta + d_2^2(1-\vartheta)}} \left\{ (\epsilon_1^2 - \sigma_2^2) - (d_1^2 - d_2^2)\sigma_2^2 - (d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)\vartheta \right\}. \tag{6.109}$$

Note that since the numerator and the denominator of the fraction are always positive by assumption, we can drop the fraction. Hence the slope's sign of $f(\vartheta)$ is decided by:

$$(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 - (d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)\vartheta. \tag{6.110}$$

In (6.110), $d_1, d_2, \epsilon_1, \epsilon_2$ are given, and we have assumed that $d_1 > d_2$, $\epsilon_1 > \epsilon_2$ and $0 \leq \vartheta \leq 1$. As a result, (6.110) is a decreasing straight line as ϑ goes from 0 to 1, and we can separate this case into three subcases:

1. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \leq 0$,
2. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \geq 0$ and $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \geq (d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)$,
3. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \geq 0$ but $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \leq (d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)$.

In the first subcase, (6.110) is always negative, hence $f(\vartheta)$ is monotonically decreasing. Recall that currently we focus on the situation when $c \leq f(0)$, but $c \geq f(1)$. When ϑ goes from 0 to 1, (6.106) goes from negative to positive. It means that the shape of (6.101) is convex, and the maximum is achieved on the boundary for $\vartheta = 0$ or $\vartheta = 1$.

In the second subcase, (6.110) is positive as $\vartheta = 0$. Furthermore, the last term is too small to make the sign of (6.110) change. Even if $\vartheta = 1$, (6.110) is still positive. Before we analyze the shape of $f(\vartheta)$, we observe the following facts:

$$\text{If } (\epsilon_1^2 - \epsilon_2^2) \geq (d_1^2 - d_2^2)\epsilon_2^2, \text{ then } d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2 \leq (d_1^2 - d_2^2)\frac{\epsilon_2^2}{e^{d_2^2}}. \tag{6.111}$$

This can be proven as follows:

$$d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2 \geq (d_1^2 - d_2^2)\frac{\epsilon_2^2}{e^{d_2^2}} \iff \frac{d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2}{(d_1^2 - d_2^2)\epsilon_2^2} \geq \frac{1}{e^{d_2^2}} \tag{6.112}$$

$$\iff -\frac{d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2}{(d_1^2 - d_2^2)\epsilon_2^2} \leq -\frac{1}{e^{d_2^2}} \tag{6.113}$$

$$\iff 1 - \frac{d_1^2\epsilon_2^2 - d_2^2\epsilon_1^2}{(d_1^2 - d_2^2)\epsilon_2^2} \leq 1 - \frac{1}{e^{d_2^2}} \tag{6.114}$$

$$\iff \frac{(\epsilon_1^2 - \epsilon_2^2)d_2^2}{(d_1^2 - d_2^2)\epsilon_2^2} \leq \frac{e^{d_2^2} - 1}{e^{d_2^2}} \tag{6.115}$$

$$\iff e^{d_2^2}\frac{(\epsilon_1^2 - \epsilon_2^2)d_2^2}{(d_1^2 - d_2^2)\epsilon_2^2} \leq e^{d_2^2} - 1 \tag{6.116}$$

$$\iff e^{d_2^2}\frac{(\epsilon_1^2 - \epsilon_2^2)d_2^2}{(d_1^2 - d_2^2)\epsilon_2^2} - e^{d_2^2} + 1 \leq 0. \tag{6.117}$$

Here in (6.112), we divide the both sides by $(d_1^2 - d_2^2)\epsilon_2^2$ that is positive by assumption $d_1 > d_2$. Next, we lower-bound the LHS of (6.117) as follows:

$$e^{d_2^2}\frac{(\epsilon_1^2 - \epsilon_2^2)d_2^2}{(d_1^2 - d_2^2)\epsilon_2^2} - e^{d_2^2} + 1 \geq e^{d_2^2}\frac{(d_1^2 - d_2^2)\epsilon_2^2 d_2^2}{(d_1^2 - d_2^2)\epsilon_2^2} - e^{d_2^2} + 1 \tag{6.118}$$

$$= d_2^2 e^{d_2^2} - e^{d_2^2} + 1, \tag{6.119}$$

which follows from the assumption $(\epsilon_1^2 - \epsilon_2^2) \geq (d_1^2 - d_2^2)\epsilon_2^2$. Note that the function $\xi \mapsto \xi e^{\xi} - e^{\xi} + 1$ shown in Figure 6.5 is monotonically increasing and equals to 0 as $\xi = 0$. Since $d_2^2 \geq 0$, we have

$$d_2^2 e^{d_2^2} - e^{d_2^2} + 1 \geq 0 \tag{6.120}$$

From (6.117), (6.119) and (6.120), (6.111) is proved and can be rewritten as:

$$\text{If } (\epsilon_1^2 - \epsilon_2^2) \geq (d_1^2 - d_2^2)\epsilon_2^2, \text{ then } c \leq f(0). \tag{6.121}$$

Next, we recall that in the currently analyzed second subcase, (6.110) is always positive, hence $f(\vartheta)$ is monotonically increasing with the minimum $f(0)$ in the interval between $\vartheta = 0$ and $\vartheta = 1$. As a result, the situations that c is above $f(\vartheta)$ and c cuts through $f(\vartheta)$ do not exist. Furthermore, since (6.106) is always negative, (6.101) is monotonically decreasing. The maximum is achieved at $\vartheta = 0$.

Finally, we look at the third subcase. There (6.110) is positive at $\vartheta = 0$, however, the last term is large enough to make the sign of (6.110) change: (6.110) changes from positive

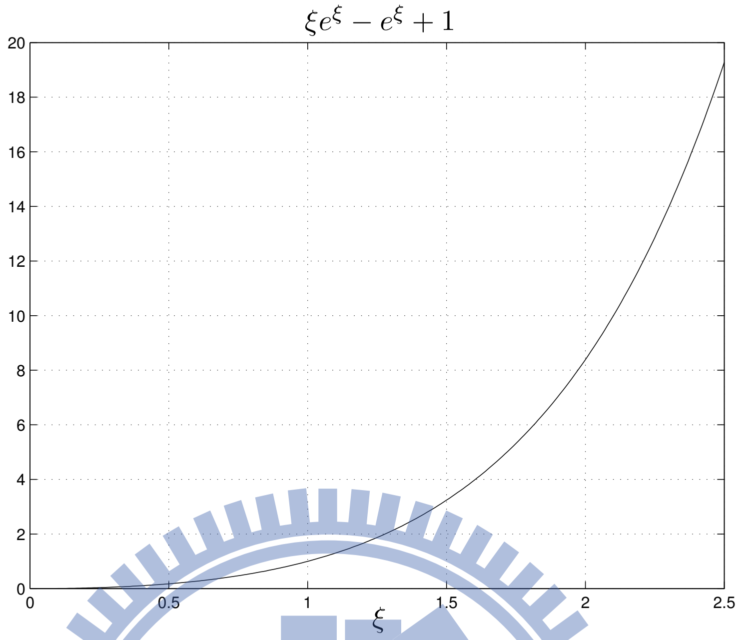

Figure 6.5: The plot of $\xi \mapsto \xi e^{\xi} - e^{\xi} + 1$ for ξ from 0 to 2.5

to negative as ϑ goes from 0 to 1, so $f(\vartheta)$ is a concave function between $\vartheta = 0$ and $\vartheta = 1$. Because in this subcase, we still have assumed $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \geq 0$, we can use the fact (6.121) as well.

Note that here the concave function $f(\vartheta)$ is one of two types as illustrated in Figure 6.6:

$$f(0) \leq f(1) \tag{6.122}$$

or

$$f(0) \geq f(1). \tag{6.123}$$

In the first type, from the fact (6.121), we have the same result as in the second subcase mentioned before. Therefore, (6.101) is monotonically decreasing and the maximum is achieved at $\vartheta = 0$. As for the second type, (6.106) changes from negative to positive as ϑ goes from 0 to 1. As a result, (6.101) is a convex function between $\vartheta = 0$ and $\vartheta = 1$. Hence the maximum exists on the boundary.

This finishes the discussion of all possible cases and proves that (6.101) cannot be concave. Therefore, the maximum is achieved for $0 \leq \vartheta \leq 1$, but it is always on the boundary $\vartheta = 0$ or $\vartheta = 1$. Furthermore, the upper bound on the fading number of the two-user SISO Rician fading MAC with memory becomes:

$$\chi_{\text{MAC-2}} \leq \max_{i \in \{1,2\}} \left\{ \log\left(|d_i|^2\right) - \text{Ei}\left(-|d_i|^2\right) - \log\left(\epsilon_i^2\right) - 1 \right\}. \tag{6.124}$$

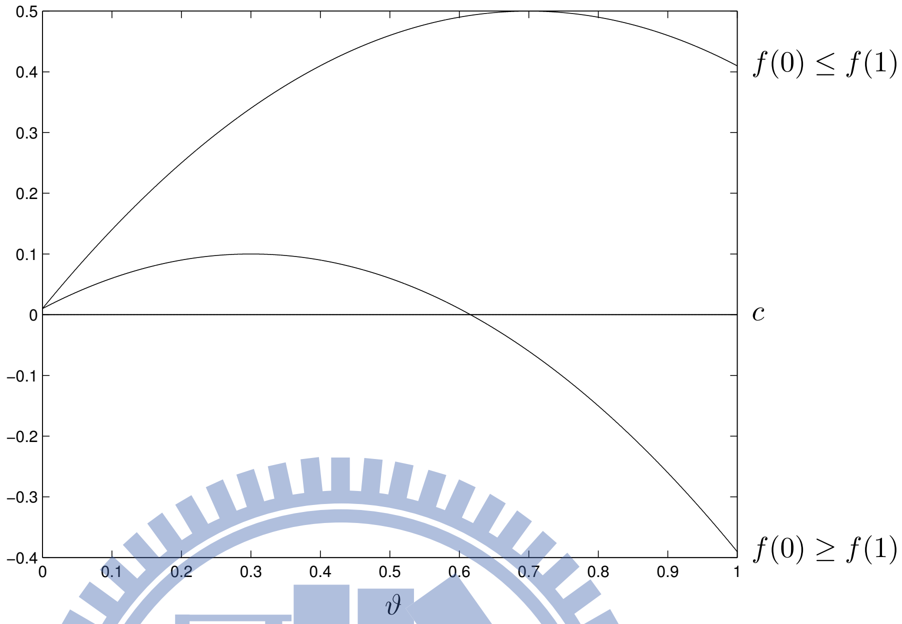

Figure 6.6: The two shapes of $f(\vartheta)$ depending on whether $f(0) \leq f(1)$ or $f(0) \geq f(1)$, $c \leq f(0)$.

From (4.7) and (6.124), we have

$$\chi_{\text{MAC-2}} = \max_{i\in\{1,2\}} \left\{\log\left(|d_i|^2\right) - \text{Ei}\left(-|d_i|^2\right) - \log\left(\epsilon_i^2\right) - 1\right\} \tag{6.125}$$

$$= \max_{i\in\{1,2\}} \chi_{\text{SISO},i}. \tag{6.126}$$

6.3 Derivation of Theorem 5.3

In Section 4.1, we have discussed the lower bound on the fading number of the two-user SISO MAC. Next, we generalize the lower bound to the m-user case. The lower bound of the m-user SISO Rician fading MAC with memory is shown as

$$\chi_{\text{MAC}} \geq \max_{i\in\{1,\ldots,m\}} \chi_{\text{SISO},i} \tag{6.127}$$

$$= \max_{i\in\{1,\ldots,m\}} \left\{\log\left(|d_i|^2\right) - \text{Ei}\left(-|d_i|^2\right) - 1 + \log\frac{1}{\epsilon_i^2}\right\}. \tag{6.128}$$

From Proposition 5.1, we have the following upper bound:

$$\chi_{\text{MAC}} \leq \varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \left\{\mathsf{E}\left[\log\left(\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2}\right)\right.\right.$$
$$- \text{Ei}\left(-\frac{|d_1|^2|X_1|^2+\cdots+|d_m|^2|X_m|^2}{|X_1|^2+\cdots+|X_m|^2}\right)$$

$$- \log\left(\frac{\epsilon_1^2|X_1|^2 + \cdots + \epsilon_m^2|X_m|^2}{|X_1|^2 + \cdots + |X_m|^2}\right) - 1\Big]\Big\} \tag{6.129}$$

$$= \overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \Big\{\mathsf{E}\Big[\log\left(|d_1|^2|\hat{X}_1|^2 + \cdots + |d_m|^2|\hat{X}_m|^2\right) - \mathrm{Ei}\Big(-\left(|d_1|^2|\hat{X}_1|^2 + \cdots + |d_m|^2|\hat{X}_m|^2\right)\Big) - \log\left(\epsilon_1^2|\hat{X}_1|^2 + \cdots + \epsilon_m^2|\hat{X}_m|^2\right) - 1\Big]\Big\}. \tag{6.130}$$

Note that the upper bound (6.130) exists in the condition that only one user communicates and the others are switched off, Theorem 5.3 is proved.

Since $\hat{\mathbf{X}}$ is a unit vector, *i.e.*,

$$\sum_{i=1}^{m} |\hat{X}_i|^2 = 1, \tag{6.131}$$

we can use r_i with $0 \leq r_i \leq 1$ to denote the input distribution of the i-th user such that

$$|\hat{X}_i|^2 = r_i, \tag{6.132}$$

$$\sum_{i=1}^{m} r_i = 1, \quad 0 \leq r_i \leq 1, \quad i = 1, \ldots, m, \tag{6.133}$$

and we have

$$\chi_{\mathrm{MAC}} \leq \max_{\mathbf{r}} \left\{\log\left(d_1^2 r_1 + \cdots + d_m^2 r_m\right) - \mathrm{Ei}\big(-\left(d_1^2 r_1 + \cdots + d_m^2 r_m\right)\big) - \log\left(\epsilon_1^2 r_1 + \cdots + \epsilon_m^2 r_m\right) - 1\right\} \tag{6.134}$$

Here we drop the moduli of line-of-sight components because they are given and we can assume they are always real.

To finish the proof, our main purpose is to prove that the maximum on the RHS of (6.134) is always achieved:

$$r_i = 1, \quad r_j = 0, \quad j = 1, \ldots, m, \quad j \neq i \tag{6.135}$$

for some $i \in \{1, \ldots, m\}$. As a result, we focus on the RHS of (6.134) and analyze the maximum.

From Section 6.2, we have

$$\max_{0\leq\vartheta\leq1} \left\{\log\left(d_1^2\vartheta + d_2^2(1-\vartheta)\right) - \mathrm{Ei}\big(-\left(d_1^2\vartheta + d_2^2(1-\vartheta)\right)\big) - \log\left(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\right) - 1\right\}$$
$$= \max_{\vartheta\in\{0,1\}} \left\{\log\left(d_1^2\vartheta + d_2^2(1-\vartheta)\right) - \mathrm{Ei}\big(-\left(d_1^2\vartheta + d_2^2(1-\vartheta)\right)\big) - \log\left(\epsilon_1^2\vartheta + \epsilon_2^2(1-\vartheta)\right) - 1\right\} \tag{6.136}$$

for given $d_1, d_2, \epsilon_1, \epsilon_2 > 0$. It means that the maximum of the two-user case is achieved on the boundary as $\vartheta = 0$ or $\vartheta = 1$.

Moreover, we can derive two lemmas.

Lemma 6.1. *Given $d_1, d_2, \epsilon_1, \epsilon_2 > 0$, for any $0 \le a \le 1$ and the constraint $0 \le \vartheta \le a$, we have*

$$\max_{0\le\vartheta\le a} \left\{\log\left(d_1^2\vartheta + d_2^2(a-\vartheta)\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(a-\vartheta)\right)\right) - \log\left(\epsilon_1^2\vartheta + \epsilon_2^2(a-\vartheta)\right) - 1\right\}$$
$$= \max_{\vartheta\in\{0,a\}} \left\{\log\left(d_1^2\vartheta + d_2^2(a-\vartheta)\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(a-\vartheta)\right)\right) - \log\left(\epsilon_1^2\vartheta + \epsilon_2^2(a-\vartheta)\right) -1\right\}. \tag{6.137}$$

Proof. We first separate $d_1, d_2, \epsilon_1,$ and ϵ_2 into three cases:

1. $d_1 > d_2$ and $\epsilon_1 < \epsilon_2$,
2. $d_1 > d_2, \epsilon_1 > \epsilon_2,$ and $\frac{d_1}{\epsilon_1} < \frac{d_2}{\epsilon_2}$,
3. $d_1 > d_2, \epsilon_1 > \epsilon_2,$ but $\frac{d_1}{\epsilon_1} > \frac{d_2}{\epsilon_2}$.

In the first and the second cases, we follow the same steps as in Section 6.2 and get the same result that the maximum exists on the boundary as $\vartheta = 0$ or $\vartheta = a$.

In the third condition, we define

$$d_1'^2 \triangleq d_1^2 - (1-a)d_2^2 > 0, \tag{6.138}$$
$$d_2'^2 \triangleq ad_2^2 > 0, \tag{6.139}$$
$$\epsilon_1'^2 \triangleq \epsilon_1^2 - (1-a)\epsilon_2^2 > 0, \tag{6.140}$$
$$\epsilon_2'^2 \triangleq a\epsilon_2^2 > 0, \tag{6.141}$$

and the LHS of (6.137) can be rewritten as

$$\max_{0\le\vartheta\le a} \left\{\log\left(d_1'^2\vartheta + d_2'^2(1-\vartheta)\right) - \mathrm{Ei}\left(-\left(d_1'^2\vartheta + d_2'^2(1-\vartheta)\right)\right) - \log\left(\epsilon_1'^2\vartheta + \epsilon_2'^2(1-\vartheta)\right) -1\right\}. \tag{6.142}$$

Therefore, we can follow the same step of the third case in Section 6.2, and prove that the maximum is achieved as $\vartheta = 0$ or $\vartheta = a$. Note that in our proof here, we change the constraint from $0 \le \vartheta \le 1$ to $0 \le \vartheta \le a$, as a result, the scope of the conditions (6.107), (6.108), (6.122) and (6.123) should be modified. Furthermore, (6.110) should be separated into three new subcases with the same steps to analyze:

1. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \le 0$,
2. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \ge 0$ and $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \ge a(d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)$,
3. $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \ge 0$ but $(\epsilon_1^2 - \epsilon_2^2) - (d_1^2 - d_2^2)\epsilon_2^2 \le a(d_1^2 - d_2^2)(\epsilon_1^2 - \epsilon_2^2)$.

□

Lemma 6.2. *Given $d_1, d_2, \epsilon_1, \epsilon_2 > 0$, for any $b \geq 0, c \geq 0$ and the constraint $0 \leq \vartheta \leq 1$, we have*

$$\max_{0\leq\vartheta\leq 1} \left\{\log\left(d_1^2\vartheta + d_2^2(1-\vartheta)+b\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(1-\vartheta)+b\right)\right) - \log\left(\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)+c\right) - 1\right\}$$
$$= \max_{\vartheta\in\{0,1\}} \left\{\log\left(d_1^2\vartheta + d_2^2(1-\vartheta)+b\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(1-\vartheta)+b\right)\right) - \log\left(\epsilon_1^2\vartheta+\epsilon_2^2(1-\vartheta)+c\right) - 1\right\}. \tag{6.143}$$

Proof. We define

$$d_1'^2 \triangleq d_1^2 + b > 0, \tag{6.144}$$
$$d_2'^2 \triangleq d_2^2 + b > 0, \tag{6.145}$$
$$\epsilon_1'^2 \triangleq \epsilon_1^2 + c > 0, \tag{6.146}$$
$$\epsilon_2'^2 \triangleq \epsilon_2^2 + c > 0, \tag{6.147}$$

and the LHS of (6.143) can be rewritten as (6.142). Hence from (6.136), we get the result that the maximum exists on the boundary as $\vartheta = 0$ or $\vartheta = 1$. □

From Lemma 6.1, we know that no matter how large the total power allocation of the two users is, the maximum of the two-user case still achieves on the boundary as $\vartheta = 0$ or $\vartheta = a$. Moreover,

Lemma 6.2 gives the result that even if we add two different constant in the line-of-sight components term and the prediction errors term, the maximum of the two-user case is still achieved on the boundary as $\vartheta = 0$ or $\vartheta = 1$. We combine these two lemmas: Given $d_1, d_2, \epsilon_1, \epsilon_2 > 0$, for any $0 \leq a \leq 1, b \geq 0, c \geq 0$ and the constraint $0 \leq \vartheta \leq a$, we have

$$\max_{0\leq\vartheta\leq a} \left\{\log\left(d_1^2\vartheta + d_2^2(a-\vartheta)+b\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(a-\vartheta)+b\right)\right) - \log\left(\epsilon_1^2\vartheta+\epsilon_2^2(a-\vartheta)+c\right) - 1\right\}$$
$$= \max_{\vartheta\in\{0,a\}} \left\{\log\left(d_1^2\vartheta + d_2^2(a-\vartheta)+b\right) - \mathrm{Ei}\left(-\left(d_1^2\vartheta + d_2^2(a-\vartheta)+b\right)\right) - \log\left(\epsilon_1^2\vartheta+\epsilon_2^2(a-\vartheta)+c\right) - 1\right\}. \tag{6.148}$$

Next, we define $\mathbf{s}$ to be the choice of $\mathbf{r}$ that achieves the maximum in (6.134).

$$\mathbf{s} \triangleq \operatorname*{argmax}_{\mathbf{r}} \left\{\log\left(d_1^2 r_1 + \cdots + d_m^2 r_m\right) - \mathrm{Ei}\left(-\left(d_1^2 r_1 + \cdots + d_m^2 r_m\right)\right) - \log\left(\epsilon_1^2 r_1 + \cdots + \epsilon_m^2 r_m\right) - 1\right\}. \tag{6.149}$$

Note that the maximum always exists since the expression in the maximum of (6.134) is a sum of monotonic concave or convex terms. The maximum might not be unique, though this does not affect the following arguments (simply pick one possible choice of $\mathbf{s}$). Further, we define

$$a_1 \triangleq s_1 + s_2, \tag{6.150}$$
$$b_1 \triangleq d_3^2 s_3 + \cdots + d_m^2 s_m, \tag{6.151}$$
$$c_1 \triangleq \epsilon_3^2 s_3 + \cdots + \epsilon_m^2 s_m, \tag{6.152}$$

where $0 \le a_1 \le 1, b_1 \ge 0, c_1 \ge 0$. Then we have

$$\begin{aligned}
&\log\left(d_1^2 s_1 + \cdots + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_1^2 s_1 + \cdots + d_m^2 s_m\right)\big) - \log\left(\epsilon_1^2 s_1 + \cdots + \epsilon_m^2 s_m\right) - 1 \\
&= \max_{0 \le \vartheta_1 \le a_1} \big\{\log\left(d_1^2 \vartheta_1 + d_2^2 \left(a_1 - \vartheta_1\right) + b_1\right) - \mathrm{Ei}\big(-\left(d_1^2 \vartheta_1 + d_2^2 \left(a_1 - \vartheta_1\right) + b_1\right)\big) \\
&\qquad - \log\left(\epsilon_1^2 \vartheta_1 + \epsilon_2^2 \left(a_1 - \vartheta_1\right) + c_1\right) - 1\big\} \qquad (6.153) \\
&= \max_{\vartheta_1 \in \{0, a_1\}} \big\{\log\left(d_1^2 \vartheta_1 + d_2^2 \left(a_1 - \vartheta_1\right) + b_1\right) - \mathrm{Ei}\big(-\left(d_1^2 \vartheta_1 + d_2^2 \left(a_1 - \vartheta_1\right) + b_1\right)\big) \\
&\qquad - \log\left(\epsilon_1^2 \vartheta_1 + \epsilon_2^2 \left(a_1 - \vartheta_1\right) + c_1\right) - 1\big\}, \qquad (6.154)
\end{aligned}$$

where $0 \le \vartheta_1 \le a_1$ denotes the power allocation of the first user, and where (6.154) follows from (6.148). Hence the maximum is achieved either for $\vartheta_1 = 0$, *i.e.*,

$$\log\left(d_2^2 a_1 + b_1\right) - \mathrm{Ei}\big(-\left(d_2^2 a_1 + b_1\right)\big) - \log\left(\epsilon_2^2 a_1 + c_1\right) - 1, \quad \text{as } s_1 = 0, \qquad (6.155)$$

or for $\vartheta_1 = a_1$, *i.e.*,

$$\log\left(d_1^2 a_1 + b_1\right) - \mathrm{Ei}\big(-\left(d_1^2 a_1 + b_1\right)\big) - \log\left(\epsilon_1^2 a_1 + c_1\right) - 1, \quad \text{as } s_2 = 0. \qquad (6.156)$$

Note that $a_1 = s_1 + s_2$. As a result, a_1 can be replaced by s_2 in the first case and by s_1 in the second case.

In the first case, we define

$$a_2 \triangleq s_2 + s_3, \qquad (6.157)$$

$$b_2 \triangleq d_4^2 s_4 + \cdots + d_m^2 s_m, \qquad (6.158)$$

$$c_2 \triangleq \epsilon_4^2 s_4 + \cdots + \epsilon_m^2 s_m, \qquad (6.159)$$

where $0 \le a_2 \le 1, b_2 \ge 0, c_2 \ge 0$. Then (6.155) can be rewritten as follows:

$$\begin{aligned}
&\log\left(d_2^2 s_2 + b_1\right) - \mathrm{Ei}\big(-\left(d_2^2 s_2 + b_1\right)\big) - \log\left(\epsilon_2^2 s_2 + c_1\right) - 1 \\
&= \log\left(d_2^2 s_2 + d_3^2 s_3 + \cdots + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_2^2 s_2 + d_3^2 s_3 + \cdots + d_m^2 s_m\right)\big) \\
&\quad - \log\left(\epsilon_2^2 s_2 + \epsilon_3^2 s_3 + \cdots + \epsilon_m^2 s_m\right) - 1 \qquad (6.160) \\
&= \max_{0 \le \vartheta_2 \le a_2} \big\{\log\left(d_2^2 \vartheta_2 + d_3^2 \left(a_2 - \vartheta_2\right) + b_2\right) - \mathrm{Ei}\big(-\left(d_2^2 \vartheta_2 + d_3^2 \left(a_2 - \vartheta_2\right) + b_2\right)\big) \\
&\qquad - \log\left(\epsilon_2^2 \vartheta_2 + \epsilon_3^2 \left(a_2 - \vartheta_2\right) + c_2\right) - 1\big\} \qquad (6.161) \\
&= \max_{\vartheta_2 \in \{0, a_2\}} \big\{\log\left(d_2^2 \vartheta_2 + d_3^2 \left(a_2 - \vartheta_2\right) + b_2\right) - \mathrm{Ei}\big(-\left(d_2^2 \vartheta_2 + d_3^2 \left(a_2 - \vartheta_2\right) + b_2\right)\big) \\
&\qquad - \log\left(\epsilon_2^2 \vartheta_2 + \epsilon_3^2 \left(a_2 - \vartheta_2\right) + c_2\right) - 1\big\}, \qquad (6.162)
\end{aligned}$$

where $0 \le \vartheta_2 \le a_2$ denotes the power allocation of the second user, and where (6.162) follows from (6.148). Hence the maximum is achieved either for $\vartheta_2 = 0$, *i.e.*,

$$\log\left(d_3^2 a_2 + b_2\right) - \mathrm{Ei}\big(-\left(d_3^2 a_2 + b_2\right)\big) - \log\left(\epsilon_3^2 a_2 + c_2\right) - 1, \quad \text{as } s_2 = 0, \qquad (6.163)$$

or for $\vartheta_2 = a_2$, *i.e*,

$$\log\left(d_2^2 a_2 + b_2\right) - \mathrm{Ei}\big(-\left(d_2^2 a_2 + b_2\right)\big) - \log\left(\epsilon_2^2 a_2 + c_2\right) - 1, \quad \text{as } s_3 = 0. \qquad (6.164)$$

Note that $a_2 = s_2 + s_3$. As a result, a_2 can be replaced by s_3 in the first case and by s_2 in the second case.

In the case of (6.156), we define

$$a_2 \triangleq s_1 + s_3, \tag{6.165}$$

where $0 \le a_2 \le 1$, and $b_2 \ge 0, c_2 \ge 0$ are the same as given in (6.158) and (6.159). Then (6.156) can be rewritten as follows:

$$\begin{aligned}
&\log\left(d_1^2 s_1 + b_1\right) - \mathrm{Ei}\big(-\left(d_1^2 s_1 + b_1\right)\big) - \log\left(\epsilon_1^2 s_1 + c_1\right) - 1 \\
&= \log\left(d_1^2 s_1 + d_3^2 s_3 + \cdots + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_1^2 s_1 + d_3^2 s_3 + \cdots + d_m^2 s_m\right)\big) \\
&\quad - \log\left(\epsilon_1^2 s_1 + \epsilon_3^2 s_3 + \cdots + \epsilon_m^2 s_m\right) - 1 \qquad (6.166) \\
&= \max_{0 \le \vartheta_2 \le a_2} \Big\{\log\left(d_1^2 \vartheta_2 + d_3^2\left(a_2 - \vartheta_2\right) + b_2\right) - \mathrm{Ei}\big(-\left(d_1^2 \vartheta_2 + d_3^2\left(a_2 - \vartheta_2\right) + b_2\right)\big) \\
&\qquad\qquad - \log\left(\epsilon_1^2 \vartheta_2 + \epsilon_3^2\left(a_2 - \vartheta_2\right) + c_2\right) - 1\Big\} \qquad (6.167) \\
&= \max_{\vartheta_2 \in \{0, a_2\}} \Big\{\log\left(d_1^2 \vartheta_2 + d_3^2\left(a_2 - \vartheta_2\right) + b_2\right) - \mathrm{Ei}\big(-\left(d_1^2 \vartheta_2 + d_3^2\left(a_2 - \vartheta_2\right) + b_2\right)\big) \\
&\qquad\qquad - \log\left(\epsilon_1^2 \vartheta_2 + \epsilon_3^2\left(a_2 - \vartheta_2\right) + c_2\right) - 1\Big\}, \qquad (6.168)
\end{aligned}$$

where $0 \le \vartheta_2 \le a_2$ denotes the power allocation of the first user, and where (6.168) follows from (6.148). Hence the maximum is achieved either for $\vartheta = 0$, *i.e.*,

$$\log\left(d_3^2 a_2 + b_2\right) - \mathrm{Ei}\big(-\left(d_3^2 a_2 + b_2\right)\big) - \log\left(\epsilon_3^2 a_2 + c_2\right) - 1, \quad \text{as } s_1 = 0, \tag{6.169}$$

or for $\vartheta_2 = a_2$, *i.e.*,

$$\log\left(d_1^2 a_2 + b_2\right) - \mathrm{Ei}\big(-\left(d_1^2 a_2 + b_2\right)\big) - \log\left(\epsilon_1^2 a_2 + c_2\right) - 1, \quad \text{as } s_3 = 0, \tag{6.170}$$

Note that $a_2 = s_1 + s_3$. As a result, a_2 can be replaced by s_3 in the first case and by s_1 in the second case. Combining the cases (6.163), (6.164), (6.169) and (6.170), we can find that the maximum is achieved either for $s_2 = 0, s_3 = 0$, *i.e.*,

$$\log\left(d_1^2 s_1 + b_2\right) - \mathrm{Ei}\big(-\left(d_1^2 s_1 + b_2\right)\big) - \log\left(\epsilon_1^2 s_1 + c_2\right) - 1, \tag{6.171}$$

or for $s_1 = 0, s_3 = 0$, *i.e.*,

$$\log\left(d_2^2 s_2 + b_2\right) - \mathrm{Ei}\big(-\left(d_2^2 s_2 + b_2\right)\big) - \log\left(\epsilon_2^2 s_2 + c_2\right) - 1, \tag{6.172}$$

or for $s_1 = 0, s_2 = 0$, *i.e.*,

$$\log\left(d_3^2 s_3 + b_2\right) - \mathrm{Ei}\big(-\left(d_3^2 s_3 + b_2\right)\big) - \log\left(\epsilon_3^2 s_3 + c_2\right) - 1. \tag{6.173}$$

We continue to apply the same steps from (6.155) to (6.173), we define

$$a_k \triangleq s_j + s_{k+1}, \tag{6.174}$$

$$b_k \triangleq d_{k+2}^2 s_{k+2} + \cdots + d_m^2 s_m, \tag{6.175}$$

$$c_k \triangleq \epsilon_{k+2}^2 s_{k+2} + \cdots + \epsilon_m^2 s_m, \tag{6.176}$$

where $j = 1, \ldots, k$ and $k = 2, \ldots, m-2$. Then

$$
\begin{aligned}
&\log\left(d_j^2 s_j + b_{k-1}\right) - \mathrm{Ei}\big(-\left(d_j^2 s_j + b_{k-1}\right)\big) - \log\left(\epsilon_j^2 s_j + c_{k-1}\right) - 1 \\
&= \log\left(d_j^2 s_j + d_{k+1}^2 s_{k+1} + \cdots + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_j^2 s_j + d_{k+1}^2 s_{k+1} + \cdots + d_m^2 s_m\right)\big) \\
&\quad - \log\left(\epsilon_j^2 s_j + \epsilon_{k+1}^2 s_{k+1} + \cdots + \epsilon_m^2 s_m\right) - 1 \qquad (6.177) \\
&= \max_{0 \le \vartheta_k \le a_k} \big\{\log\left(d_j^2 \vartheta_k + d_{k+1}^2 (a_k - \vartheta_k) + b_k\right) - \mathrm{Ei}\big(-\left(d_j^2 \vartheta_k + d_{k+1}^2 (a_k - \vartheta_k) + b_k\right)\big) \\
&\qquad - \log\left(\epsilon_j^2 \vartheta_k + \epsilon_{k+1}^2 (a_k - \vartheta_k) + c_k\right) - 1\big\} \qquad (6.178) \\
&= \max_{\vartheta_k \in \{0, a_k\}} \big\{\log\left(d_j^2 \vartheta_k + d_{k+1}^2 (a_k - \vartheta_k) + b_k\right) - \mathrm{Ei}\big(-\left(d_j^2 \vartheta_k + d_{k+1}^2 (a_k - \vartheta_k) + b_k\right)\big) \\
&\qquad - \log\left(\epsilon_j^2 \vartheta_k + \epsilon_{k+1}^2 (a_k - \vartheta_k) + c_k\right) - 1\big\}, \qquad (6.179)
\end{aligned}
$$

where $0 \le \vartheta_k \le a_k$ denotes the power allocation of the j-th user. Here in (6.178), everything is fixed except the j-th and the $(k+1)$-th users, and (6.179) follows from (6.148). Hence the maximum is achieved either for $\vartheta_k = 0$, *i.e.*,

$$
\log\left(d_{k+1}^2 a_k + b_k\right) - \mathrm{Ei}\big(-\left(d_{k+1}^2 a_k + b_k\right)\big) - \log\left(\epsilon_{k+1}^2 a_k + c_k\right) - 1, \quad \text{as } s_j = 0, \qquad (6.180)
$$

or for $\vartheta_k = a_k$, *i.e.*,

$$
\log\left(d_j^2 a_k + b_k\right) - \mathrm{Ei}\big(-\left(d_j^2 a_k + b_k\right)\big) - \log\left(\epsilon_j^2 a_k + c_k\right) - 1, \quad \text{as } s_{k+1} = 0, \qquad (6.181)
$$

for $j = 1, \ldots, k$. Note that $a_k = s_j + s_{k+1}$. As a result, a_k can be replaced by s_{k+1} in the first case and by s_j in the second case. Combining the k cases, we can find that the maximum is achieved for one of $k+1$ cases:

$$
\log\left(d_j^2 s_j + b_k\right) - \mathrm{Ei}\big(-\left(d_j^2 s_j + b_k\right)\big) - \log\left(\epsilon_j^2 s_j + c_k\right) - 1 \qquad (6.182)
$$

for $j = 1, \ldots, k+1$ as $s_l = 0, l = 1, \ldots, k+1, l \neq j$.

In the end, we have the result that the maximum is achieved for one of $m-1$ cases:

$$
\log\left(d_j^2 s_j + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_j^2 s_j + d_m^2 s_m\right)\big) - \log\left(\epsilon_j^2 s_j + \epsilon_m^2 s_m\right) - 1 \qquad (6.183)
$$

for $j = 1, \ldots, m-1$. From (6.133), we know the sum of s_i is 1, therefore, (6.183) can be rewritten as follows: for $j = 1, \ldots, m-1$

$$
\begin{aligned}
&\log\left(d_j^2 s_j + d_m^2 s_m\right) - \mathrm{Ei}\big(-\left(d_j^2 s_j + d_m^2 s_m\right)\big) - \log\left(\epsilon_j^2 s_j + \epsilon_m^2 s_m\right) - 1 \\
&= \max_{0 \le \vartheta_m \le 1} \big\{\log\left(d_j^2 \vartheta_m + d_m^2 (1 - \vartheta_m)\right) - \mathrm{Ei}\big(-\left(d_j^2 \vartheta_m + d_m^2 (1 - \vartheta_m)\right)\big) \\
&\qquad - \log\left(\epsilon_j^2 \vartheta_m + \epsilon_m^2 (1 - \vartheta_m)\right) - 1\big\} \qquad (6.184) \\
&= \max_{\vartheta_m \in \{0, 1\}} \big\{\log\left(d_j^2 \vartheta_m + d_m^2 (1 - \vartheta_m)\right) - \mathrm{Ei}\big(-\left(d_j^2 \vartheta_m + d_m^2 (1 - \vartheta_m)\right)\big) \\
&\qquad - \log\left(\epsilon_j^2 \vartheta_m + \epsilon_m^2 (1 - \vartheta_m)\right) - 1\big\}, \qquad (6.185)
\end{aligned}
$$

where $0 \le \vartheta_m \le 1$ denotes the power allocation of the j-th user. Here (6.185) follows from (6.136). Hence the maximum is achieved either for $\vartheta_m = 0$, *i.e.*,

$$
\log\left(d_m^2\right) - \mathrm{Ei}\big(-\left(d_m^2\right)\big) - \log\left(\epsilon_m^2\right) - 1, \quad \text{as } s_j = 0, \qquad (6.186)
$$

or for $\vartheta_m = 1$, *i.e.*,

$$\log\left(d_j^2\right) - \mathrm{Ei}\left(-\left(d_j^2\right)\right) - \log\left(\epsilon_j^2\right) - 1, \quad \text{as } s_m = 0 \tag{6.187}$$

for $j = 1, \ldots, m-1$. Combining these $m-1$ cases, we derive the fact that the maximum must be of the form

$$\log\left(d_j^2\right) - \mathrm{Ei}\left(-d_j^2\right) - \log\left(\epsilon_j^2\right) - 1 \tag{6.188}$$

for some $j \in \{1, \ldots, m\}$ as $s_l = 0, l = 1, \ldots, m, l \neq j$. Consequently, the RHS of (6.134) is always in the form of (6.135). Furthermore, the upper bound of the m-user SISO Rician fading MAC with memory becomes:

$$\chi_{\mathrm{MAC}} \leq \max_{i \in \{1,\ldots,m\}} \left\{\log\left(|d_i|^2\right) - \mathrm{Ei}\left(-|d_i|^2\right) - \log\left(\epsilon_i^2\right) - 1\right\}. \tag{6.189}$$

From (6.128) and (6.189), we have

$$\chi_{\mathrm{MAC}} = \max_{i \in \{1,\ldots,m\}} \left\{\log\left(|d_i|^2\right) - \mathrm{Ei}\left(-|d_i|^2\right) - \log\left(\epsilon_i^2\right) - 1\right\} \tag{6.190}$$

$$= \max_{i \in \{1,\ldots,m\}} \chi_{\mathrm{SISO},i}. \tag{6.191}$$

Chapter 7

Discussion and Conclusion

In this thesis, we have succeeded in deriving the exact fading number (*i.e.*,the exact asymptotic capacity region) for an m-user SISO Rician fading MAC with memory. We have shown that the fading number of SISO MAC is exactly equivalent to the fading number of a single-user SISO channel. To achieve the sum-rate capacity, the optimal strategy is switching off all the users except the one who has the best fading number.

In [4], we have the fading number of memoryless case:

$$\chi_{\text{MAC-IID}} = \log\left(d_{\text{MAC-IID}}^2\right) - \text{Ei}\left(-d_{\text{MAC-IID}}^2\right) - 1 \tag{7.1}$$

where

$$d_{\text{MAC-IID}} = \max\left\{|d_1|, \ldots, |d_m|\right\}. \tag{7.2}$$

The fading number of the Rician fading memoryless SISO MAC only depends on the line-of-sight components, *i.e.*, the better the line-of-sight component is, the better the fading number is. However, in our thesis, since we assume channels with memory where we can predict the current fading from the past fading, the fading number is also influenced by the prediction errors. Even if the channel has the better line-of-sight component, it might has the worse fading number.

In Section 6.2, we have proved the two-user case with one user has the better line-of-sight component and the better prediction error. From the upper bound of the fading number in Proposition 5.1, we can derive the fading number of the two-user case with one user has the better ratio of line-of-sight component to prediction errors and the worse line-of-sight component. But in other cases, we can not distinguish which one has the better fading number by the line-of-sight components and the prediction errors directly. We only prove that the fading number cannot exist in the situation of two users working together. In general, the fading number only exist as only one user communicates.

A possible reason for this pessimistic result might be that all the users are independent, *i.e.*, they don't know the states of other users and can't cooperate with other users. Therefore, the signals transmitted from other users can only be interference and reduce the performance.

Note that in our thesis, we assume the channel model is noncoherent. Since neither the transmitter and the receiver knows the real state of channel model, some methods of lowering the interference can not be utilized. Moreover, we only consider the asymptotic capacity that the system works in the regime of high-SNR, however, real systems usually operate at low SNR. As a result, it's not necessary to degrade the multiple-access channel to single-user channel for designing a system.

Possible future works for the multiple-access fading channel might be as follows:

- Generalizing the SISO case to the MIMO case that the users and the receiver use the multiple antennas. A possible approach could be considered first for the MISO case.
- Considering the case with side-information.
- Considering the case with feedback.
- Loosening the restriction of Rician fading and considering the cases of general fading.
- Deriving bounds on the nonasymptotic capacity when the system does not operate in the regime of high-SNR.

Appendix A

Derivation of Lemma 4.2

We follows the steps in [7, Section 4.2] to derive the upper bound in Lemma 4.2. From the mutual information of MISO fading channel shown in Lemma 4.1, we have

$$\begin{aligned}
I(\mathbf{X};Y) &\leq -h(Y|\mathbf{X}) + \log\pi + \alpha\log\beta + \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) + (1-\alpha)\mathsf{E}\left[\log(|Y|^2+\nu)\right] \\
&\quad + \frac{1}{\beta}\mathsf{E}\left[|Y|^2\right] + \frac{\nu}{\beta} && \text{(A.1)}\\
&\leq -h(Y|\mathbf{X}) + \log\pi + \alpha\log\beta + \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) + (1-\alpha)\mathsf{E}\left[\log|Y|^2\right] + \epsilon_\nu \\
&\quad + \frac{1}{\beta}\mathsf{E}\left[|Y|^2\right] + \frac{\nu}{\beta} && \text{(A.2)}\\
&= -\mathsf{E}\left[\log\pi e(\|\mathbf{X}\|^2+\sigma^2)\right] + \log\pi + \alpha\log\beta + \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) \\
&\quad + (1-\alpha)\mathsf{E}\left[\mathsf{E}\left[\log|Y|^2 \,\middle|\, \mathbf{X}=\mathbf{x}\right]\right] + \epsilon_\nu + \frac{1}{\beta}\mathsf{E}\left[\|\mathbf{X}\|^2+\sigma^2+|\mathbf{d}^\mathsf{T}\mathbf{X}|^2\right] + \frac{\nu}{\beta} && \text{(A.3)}\\
&= -\mathsf{E}\left[\log(\|\mathbf{X}\|^2+\sigma^2)\right] - 1 + \alpha\log\beta + \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) + (1-\alpha)\mathsf{E}\left[\log(\|\mathbf{X}\|^2+\sigma^2)\right] \\
&\quad + (1-\alpha)\mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right)\right] + \epsilon_\nu \\
&\quad + \frac{1}{\beta}\mathsf{E}\left[\|\mathbf{X}\|^2+\sigma^2+|\mathbf{d}^\mathsf{T}\mathbf{X}|^2\right] + \frac{\nu}{\beta} && \text{(A.4)}\\
&= -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right)\right] \\
&\quad + \alpha\left(\log\beta - \mathsf{E}\left[\log(\|\mathbf{X}\|^2+\sigma^2)\right] - \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right)\right]\right) \\
&\quad + \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) + \epsilon_\nu + \frac{1}{\beta}\mathsf{E}\left[\|\mathbf{X}\|^2+\sigma^2+|\mathbf{d}^\mathsf{T}\mathbf{X}|^2\right] + \frac{\nu}{\beta}. && \text{(A.5)}
\end{aligned}$$

Here (A.1) follows Lemma 4.1; in (A.2), we assume $0 < \alpha < 1$ such that $1-\alpha > 0$ and define

$$\epsilon_\nu \triangleq \sup_{\mathbf{x}}\left\{\mathsf{E}\left[\log\left(|Y|^2+\nu\right) \,\middle|\, \mathbf{X}=\mathbf{x}\right] - \mathsf{E}\left[\log|Y|^2 \,\middle|\, \mathbf{X}=\mathbf{x}\right]\right\}, \tag{A.6}$$

such that

$$
\begin{aligned}
&(1-\alpha)\mathsf{E}\left[\log(|Y|^2+\nu)\right] \\
&\quad = (1-\alpha)\mathsf{E}\left[\log|Y|^2\right] + (1-\alpha)\left(\mathsf{E}\left[\log(|Y|^2+\nu)\right] - \mathsf{E}\left[\log|Y|^2\right]\right) && \text{(A.7)}\\
&\quad \le (1-\alpha)\mathsf{E}\left[\log|Y|^2\right] \\
&\qquad + (1-\alpha)\sup_{\mathbf{x}}\left\{\mathsf{E}\left[\log(|Y|^2+\nu)\,\middle|\,\mathbf{X}=\mathbf{x}\right] - \mathsf{E}\left[\log|Y|^2\,\middle|\,\mathbf{X}=\mathbf{x}\right]\right\} && \text{(A.8)}\\
&\quad = (1-\alpha)\mathsf{E}\left[\log|Y|^2\right] + (1-\alpha)\epsilon_\nu && \text{(A.9)}\\
&\quad \le (1-\alpha)\mathsf{E}\left[\log|Y|^2\right] + \epsilon_\nu; && \text{(A.10)}
\end{aligned}
$$

(A.3) follows the fact that the channel output is Gaussian distributed when $\mathbf{X}=\mathbf{x}$ is given; in (A.4), we evaluate the expected logarithm of a noncentral of a noncentral chi-square random as derived in [2], [3, Lemma 10.1], [6, Lemma A.6]; and (A.5) follows from simple algebraic rearrangements.

Next, we lower-bound some terms in (A.5) as follows:

$$
\mathsf{E}\left[\log(\|\mathbf{X}\|^2+\sigma^2)\right] \ge \log\sigma^2; \tag{A.11}
$$

$$
\mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right)\right] \ge -\gamma; \tag{A.12}
$$

and upper-bound another term as follows:

$$
\begin{aligned}
\mathsf{E}\left[\|\mathbf{X}\|^2+\sigma^2+|\mathbf{d}^\mathsf{T}\mathbf{X}|^2\right] &\le \mathcal{E}+\sigma^2+\mathsf{E}\left[\|\mathbf{d}\|^2\|\mathbf{X}\|^2\right] && \text{(A.13)}\\
&= \mathcal{E}+\sigma^2+\|\mathbf{d}\|^2\mathsf{E}\left[\|\mathbf{X}\|^2\right] && \text{(A.14)}\\
&\le \mathcal{E}+\sigma^2+\|\mathbf{d}\|^2\mathcal{E} && \text{(A.15)}\\
&= (1+\|\mathbf{d}\|^2)\mathcal{E}+\sigma^2. && \text{(A.16)}
\end{aligned}
$$

Here, (A.11) follows from dropping some nonnegative terms; (A.12) follows because $\log\xi - \mathrm{Ei}(-\xi) \ge -\gamma$ where $\gamma \approx 0.57$ denotes Euler's constant; and (A.13) and (A.15) follow from the Cauchy-Schwarz inequality and the fact that the input needs to satisfy the average-power constraint.

Furthermore, we bound

$$
\mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2+\sigma^2}\right)\right] \le \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right)\right], \tag{A.17}
$$

which follows because $\xi \mapsto \log\xi - \mathrm{Ei}(-\xi)$ is monotonically increasing.

Therefore, we can rewrite (A.5) as follows:

$$
\begin{aligned}
I(\mathbf{X};Y) \le -1 + \mathsf{E}\left[\log\left(\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right) - \mathrm{Ei}\left(-\frac{|\mathbf{d}^\mathsf{T}\mathbf{X}|^2}{\|\mathbf{X}\|^2}\right)\right] + \alpha\left(\log\beta - \log\sigma^2 + \gamma\right)\\
+ \log\Gamma\left(\alpha,\frac{\nu}{\beta}\right) + \epsilon_\nu + \frac{1}{\beta}\left((1+\|\mathbf{d}\|^2)\mathcal{E}+\sigma^2\right) + \frac{\nu}{\beta} \qquad \text{(A.18)}
\end{aligned}
$$

and Lemma 4.2 is proved.

Appendix B

Derivation of Lemma 4.3

To derive (4.11) and (4.12), we first define the set $\mathcal{B}$ as

$$\mathcal{B} \triangleq \{x_1 : 0 \leq |x_1| \leq a|x_i|\} \tag{B.1}$$

for an arbitrary value $a > 1$ and

$$\mathcal{E}_1 \triangleq \mathsf{E}\left[|X|^2\right]; \tag{B.2}$$

assume that the first user escapes to infinity, *i.e.*, if $\mathcal{E} \uparrow \infty$ then $\mathcal{E}_1 \uparrow \infty$. Furthermore, the LHS in (4.11) can be upper-bound as follows:

$$\begin{aligned}
&\overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_1||X_1||d_i||X_i|}{|X_1|^2+\cdots+|X_m|^2}\right] \\
&\leq |d_1||d_i| \overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|X_1||X_i|}{|X_1|^2+|X_i|^2}\right] && \text{(B.3)} \\
&\leq |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \mathsf{E}\left[\frac{|X_1||X_i|}{|X_1|^2+|X_i|^2}\right] && \text{(B.4)} \\
&= |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int \frac{|x_1||x_i|}{|x_1|^2+|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) && \text{(B.5)} \\
&\leq |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}} \frac{|x_1||x_i|}{|x_1|^2+|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \\
&\quad + |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}^c} \frac{|x_1||x_i|}{|x_1|^2+|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i). && \text{(B.6)}
\end{aligned}$$

Here (B.3) follows because we drop some terms in the denominator; in (B.4), we define A_1 as the set of all input distributions of the first user that escape to infinity, and take the supremum over all Q_{x_i} which are independent on Q_{x_1} and without any constraint on the average power; and (B.6) follows from splitting the inner integration into two parts and the property that the supremum of a sum is always upper-bounded by the sum of the suprema.

We focus on the first term in (B.6) to continue:

$$|d_1||d_i| \sup_{Q_{x_i}} \varlimsup_{\mathcal{E}_1\uparrow\infty} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}} \underbrace{\frac{|x_1||x_i|}{|x_1|^2+|x_i|^2}}_{\leq\frac{1}{2}} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i)$$

$$\leq |d_1||d_i| \sup_{Q_{x_i}} \varlimsup_{\mathcal{E}_1\uparrow\infty} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}} \frac{1}{2}\mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.7}$$

$$\leq |d_1||d_i| \sup_{Q_{x_i}} \varlimsup_{\mathcal{E}_1\uparrow\infty} \int\left(\sup_{Q_{x_1}\in\mathcal{A}_1} \frac{1}{2}\int_{x_1\in\mathcal{B}} \mathsf{d}Q_{x_1}(x_1)\right)\mathsf{d}Q_{x_i}(x_i) \tag{B.8}$$

$$= |d_1||d_i| \sup_{Q_{x_i}} \int \varlimsup_{\mathcal{E}_1\uparrow\infty} \left(\sup_{Q_{x_1}\in\mathcal{A}_1} \frac{1}{2}\int_{x_1\in\mathcal{B}} \mathsf{d}Q_{x_1}(x_1)\right)\mathsf{d}Q_{x_i}(x_i) \tag{B.9}$$

$$= |d_1||d_i| \sup_{Q_{x_i}} \int\left(\varlimsup_{\mathcal{E}_1\uparrow\infty} \sup_{Q_{x_1}\in\mathcal{A}_1} \frac{1}{2}\Pr(|X_1|\leq a|x_i|)\right)\mathsf{d}Q_{x_i}(x_i) \tag{B.10}$$

$$= |d_1||d_i| \sup_{Q_{x_i}} \int 0\mathsf{d}Q_{x_i}(x_i) = 0. \tag{B.11}$$

Here (B.7) follows the fact that

$$\frac{r_1 r_i}{r_1^2+r_i^2} \leq \frac{1}{2} \tag{B.12}$$

and that $r_1 \mapsto \frac{r_1 r_i}{r_1^2+r_i^2}$ is monotonically decreasing if $r_1 > r_i$; (B.8) follows by taking the supremum into the first integral which can only enlarge the expression; in (B.9), we exchange limit and integration which needs justification: define

$$g_{\mathcal{E}_1}(x_i) \triangleq \sup_{Q_{x_1}\in\mathcal{A}_1} \frac{1}{2}\int_{x_1\in\mathcal{B}} \mathsf{d}Q_{x_1}(x_1) \tag{B.13}$$

$$\leq \sup_{Q_{x_1}\in\mathcal{A}_1} \frac{1}{2}\int \mathsf{d}Q_{x_1}(x_1) \tag{B.14}$$

$$= \frac{1}{2} \triangleq g_{\text{upper}}(x_i) \tag{B.15}$$

and then note that

$$\int g_{\text{upper}}(x_i)\mathsf{d}Q_{x_i}(x_i) = \int \frac{1}{2}\mathsf{d}Q_{x_i}(x_i) = \frac{1}{2}, \tag{B.16}$$

i.e., $g_{\text{upper}}(\cdot)$ is independent of $\mathcal{E}_1$ and integrable, therefore, we are allowed to swap limit and integration by the Dominated Convergence Theorem in [10]; and (B.10) follows from Proposition 3.4 since Q_{x_1} escapes to infinity.

Next, we upper-bound the LHS in (4.11) as follows:

$$\varlimsup_{\mathcal{E}\uparrow\infty} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_1||X_1||d_i||X_i|}{|X_1|^2+\cdots+|X_m|^2}\right]$$

$$\leq |d_1||d_i| \sup_{Q_{x_i}} \varlimsup_{\mathcal{E}_1\uparrow\infty} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}^c} \frac{|x_1||x_i|}{|x_1|^2+|x_i|^2}\mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.17}$$

$$\leq |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}^c} \frac{(a|x_i|)|x_i|}{(a|x_i|)^2+|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.18}$$

$$= |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int_{x_1\in\mathcal{B}^c} \frac{a}{a^2+1} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.19}$$

$$\leq |d_1||d_i| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}_1} \int\int \frac{a}{a^2+1} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.20}$$

$$= |d_1||d_i| \sup_{Q_{x_i}} \int \frac{a}{a^2+1} \mathsf{d}Q_{x_i}(x_i) \tag{B.21}$$

$$= |d_1||d_i| \frac{a}{a^2+1} < \tau \tag{B.22}$$

for any $\tau > 0$ if we choose a large enough. Here (B.18) follows the fact that $r_1 \mapsto \frac{r_1 r_i}{r_1^2+r_i^2}$ is monotonically decreasing if $r_1 > r_i$. Since $a > 1$ is arbitrary, we get

$$\overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_1||X_1||d_i||X_i|}{|X_1|^2+\cdots+|X_m|^2}\right] = 0. \tag{B.23}$$

Moreover, we upper-bound the LHS in (4.12) with the same steps as (4.11):

$$\overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\right]$$

$$\leq |d_i||d_j| \overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|X_i||X_j|}{|X_1|^2+|X_i|^2+|X_j|^2}\right] \tag{B.24}$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}\cdot Q_{x_j}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \mathsf{E}\left[\frac{|X_i||X_j|}{|X_1|^2+|X_i|^2+|X_j|^2}\right] \tag{B.25}$$

$$= |d_i||d_j| \sup_{Q_{x_i}\cdot Q_{x_j}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int\int \frac{|x_i||x_j|}{|x_1|^2+|x_i|^2+|x_j|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i)\mathsf{d}Q_{x_j}(x_j) \tag{B.26}$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int \frac{|x_i|^2}{|x_1|^2+2|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.27}$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}} \frac{|x_i|^2}{|x_1|^2+2|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i)$$

$$+ |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}^c} \frac{|x_i|^2}{|x_1|^2+2|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i). \tag{B.28}$$

Here (B.24) follows because we drop some terms in the denominator; in (B.25), we define A_1 as the set of all input distributions of the first user that escape to infinity, and take the supremum over all $Q_{x_i}\cdots Q_{x_j}$ which are independent on Q_{x_1} and without any constraint on the average power; (B.27) follows the fact that

$$\frac{r_i r_j}{r_1^2+r_i^2+r_j^2} \leq \frac{r_i^2}{r_1^2+2r_i^2} \leq \frac{1}{2} \tag{B.29}$$

and that $r_1 \mapsto \frac{r_i^2}{r_1^2+2r_i^2}$ is monotonically decreasing if $r_1 > r_i$; and (B.28) follows from the same step as (B.6).

For the first term in (B.28), we have

$$|d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}} \underbrace{\frac{|x_i|^2}{|x_1|^2+2|x_i|^2}}_{<\frac{1}{2}} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i)$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}} \frac{1}{2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.30}$$

$$\leq 0. \tag{B.31}$$

Here (B.31) follows from the derivation of (B.8)–(B.11).

Combined with (B.31), the LHS in (4.12) becomes:

$$\overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\right]$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}^c} \frac{|x_i|^2}{|x_1|^2+2|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.32}$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}^c} \frac{|x_i|^2}{(a|x_i|)^2+2|x_i|^2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.33}$$

$$= |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int_{x_1\in\mathcal{B}^c} \frac{1}{a^2+2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.34}$$

$$\leq |d_i||d_j| \sup_{Q_{x_i}} \overline{\lim_{\mathcal{E}_1\uparrow\infty}} \sup_{Q_{x_1}\in\mathcal{A}} \int\int \frac{1}{a^2+2} \mathsf{d}Q_{x_1}(x_1)\mathsf{d}Q_{x_i}(x_i) \tag{B.35}$$

$$= |d_i||d_j| \sup_{Q_{x_i}} \int \frac{1}{a^2+2} \mathsf{d}Q_{x_i}(x_i) \tag{B.36}$$

$$= |d_i||d_j| \frac{1}{a^2+2} \leq \tau \tag{B.37}$$

for any $\tau > 0$ if we choose a large enough. Here (B.33) follows the fact $r_1 \mapsto \frac{r_i^2}{r_1^2+r_i^2}$ is monotonically decreasing if $r_1 > r_i$. Since $a > 1$ is arbitrary, we obtain

$$\overline{\lim_{\mathcal{E}\uparrow\infty}} \sup_{Q_{\mathcal{E}}\in\mathcal{A}} \mathsf{E}\left[\frac{|d_i||X_i||d_j||X_j|}{|X_1|^2+\cdots+|X_m|^2}\right] = 0. \tag{B.38}$$

Bibliography

[1] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.

[2] Amos Lapidoth and Stefan M. Moser, “The expected logarithm of a noncentral chi-square random variable,” website. [Online]. Available: http://moser.cm.nctu.edu.tw/explog.html

[3] Amos Lapidoth and Stefan M. Moser, “Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, October 2003.

[4] Gu-Rong Lin, “Capacity analysis of multiple-access Rician fading channel,” Master’s thesis, Information Theory Lab, Department of Communication Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan, June 2009, supervised by Prof. Dr. Stefan M. Moser. [Online]. Available: http://moser.cm.nctu.edu.tw/publications.html

[5] Gu-Rong Lin and Stefan M. Moser, “The fading number of a multiple-access Rician fading channel,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4983–4991, August 2011.

[6] Stefan M. Moser, *Duality-Based Bounds on Channel Capacity*, ser. ETH Series in Information Theory and its Applications. Konstanz: Hartung-Gorre Verlag, January 2005, vol. 1, ISBN 3–89649–956–4, edited by Amos Lapidoth. [Online]. Available: http://moser.cm.nctu.edu.tw/publications.html

[7] Stefan M. Moser, “Capacity analysis of multiple-access OFDM channels,” Department of Communication Engineering, National Chiao Tung University (NCTU), Zhudong, Taiwan, final report of project *4G Wireless Access Technology* G1-95003, from January 1, to December 31, 2006, funded by Industrial Technology Research Institute (ITRI).

[8] Stefan M. Moser, “The fading number of multiple-input multiple-output fading channels with memory,” *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2716–2755, June 2009.

[9] Mohsen Pourahmadi, *Foundations of Time Series Analysis and Prediction Theory.* New York: John Wiley & Sons, 2001.

[10] Hilary A. Priestley, *Introduction to Integration.* Oxford: Oxford University Press, 1997.

[11] Vignesh Sethuraman, Ligong Wang, Bruce Hajek, and Amos Lapidoth, "Low-snr capacity of noncoherent fading channels," *IEEE Transactions on Information Theory*, vol. 55, no. 4, pp. 1555–1574, April 2009.

[12] Claude E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[13] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Richard A. Silverman, Ed. Upper Saddle River, NJ: Prentice Hall, 1962.