國 立 交 通 大 學

電信工程學系

碩 士 論 文

無線光通道容量分析

On the Capacity of Free-Space Optical Communication Channel

研 究 生： 林鼎傑

指導教授： 莫詩台方　助理教授

中華民國九十七年六月

無線光通道容量分析

On the Capacity of Free-Space Optical Communication Channel

研 究 生：林鼎傑　　　　　　　　　Student：Lin Ding-Jie

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

國 立 交 通 大 學
電 信 工 程 學 系 碩 士 班

碩 士 論 文

A Thesis

Submitted to Institute of Communication Engineering

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Electrical Engineering

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

無線光通道容量分析

研究生：林鼎傑　　　　指導教授：莫詩台方　博士

國立交通大學電信工程學系碩士班

中文摘要

本篇論文探討無線光通道的容量大小。此通道有下列幾項特性: 輸入端的是以光的強度當做記量的單位，所以其值為非負值；輸出端會受到雜訊的影響，在此我們把此雜訊假設成相加白色高斯雜訊，因為在接收端會受到許多不同獨立來源的雜訊，所以我們把它假設成白色高斯雜訊；此外，電池電力的限制和使用上眼睛安全的問題，我們必須有平均能量與峰值能量大小的限制。

我們的任務是改善之前所提出的上下限，讓上下限間的距離變得更靠近，我們就可夾擊出真正的通道容量。首先我們先利用數值分析的方式求出一個數值的通道容量下限，然後跟之前的結果比較。之後在根據先前的比較，用理論分析的方式去求得公式。

我們所得到的數值下限與之前論文所推導出的上限在低訊雜比(SNR) 的情況下相當接近，如果我們放鬆對平均能量的限制，我們可以得到更好結果。而我們理論分析出來的下限，在中雜訊比的地方有明顯的改善。

On the Capacity of Free-Space Optical Communication Channel

Student: Lin Ding-Jie Advisor: Prof. Stefan M. Moser

Institute of Communication Engineering
National Chiao Tung University

Abstract

In this thesis we study the free-space optical communication channeld. This channel is characterized by inputs that are nonnegative representing transmitted optical intensity, and by outputs that are corrupted by additive white Gaussian noise, because in free space the disturbances arise from many independent sources. Due to battery and safety reasons the inputs are simultaneously constrained on both their average and peak power.

Our task is trying to find upper and lower bounds on the capacity of this channel. We use a numerical approach to compute the channel capacity and compare it with previous bounds. Based on the result of this comparison, we try to derive improved analytical lower bounds on capacity.

Our numerical lower bounds are very close to previous upper bounds in lower SNR region (below 5 dB) with improving performance if we loosen the average-power constraint. Our analytical lower bounds are good in medium SNR region (between 5 dB and 15 dB).

Acknowledgments

I would like to thank my advisor, Prof. Stefan M. Moser, for providing me this project and offer. He provided me a nice place, the Information Theory Lab, NCTU and all facilities I need, including an iMac G5, which is the most lovely and the coolest machine I have ever used, and also printers, etc. With his valuable help and instructive suggestion, I have always been able to find some way to solve problems in math or in programming. I also would like to thank him for his patience when I sometimes had difficulties following his brilliant ideas so that he had to slow down, explain once again.

Next, I would like to thank my classmates, Lin Hsuan-Yin and Lin Gu-Rong who have helped me a lot in anything concerning this project and share a lot of cool and interest things.

Finally, I would like to thank my family for their support during my whole education career since my childhood.

Hsinchu, June 15, 2008

Ding-Jie Lin

Contents

List of Figures

Chapter 1

Introduction

In this thesis we focus on a channel model that describes a short optical communication link in free space like, *e.g.*, the link between a remote control and a TV. We assume that the line-of-sight component is dominant and ignore any effects due to multiple-path propagation like fading or inter-symbol interference. Since in ambient light conditions the received signal is disturbed by a high number of independent sources, we model the noise to be Gaussian distributed.

In optical communication not only battery-power is limited, but we have to restrict the maximum allowed peak power for eye safety reasons. We therefore assume simultaneously two constraints: an average-power constraint $\mathcal{E}$ and a limitation on the maximum allowed peak power A. Besides, the ratio of $\mathcal{E}$ to A will play an important role in our investigations.

In this work we study the channel capacity of such an optical communication channel. The **information channel capacity** of a continuous memoryless channel [1] with input $X \in \mathbb{R}$ and output $Y \in \mathbb{R}$ is defined as follows:

$$\mathcal{C} \triangleq \sup_f I(X,Y) \tag{1.1}$$

where the supremum is over all possible input distributions f. There is also an operational definition of channel capacity as the highest rate in nats[1] per channel use at which information can be sent with arbitrarily low probability of error. Shannon's second theorem establishes that the information channel capacity is equal to the operational channel capacity. Thus we drop the word information in most discussions of channel capacity.

In this work we study the channel capacity and try to find numerical lower bounds. After comparing numerical lower bounds with previous bounds in [2], we will find that the upper bounds in [2] are quite tight. Hence, we will concentrate our efforts on the lower bound.

Now we will introduce briefly the technics which are used in [2]. The upper bounds in [2] were derived as follows: using a dual expression of channel capacity based on relative

[1]The base of the logarithm is e, the entropy is measured in nats.

entropy, the maximum in (1.1) is changed into a minimization over output distributions. So the problem is reduced to find an output distribution that will lead to a good upper bound. The lower bounds in [2] are derived as follows: the supremum in (1.1) is lower-bounded by choosing a particular input distribution. Then the expression is further simplified by the application of the Entropy Power Inequality. Unfortunately, the Entropy Power Inequality seems to loosen the bound, so in this work we have to find new approaches to improve on the lower bound.

The remainder of this thesis is structured as follows. After some brief remarks about our notation, we will define the considered channel model more in detail in the subsequent chapter. Chapter 3 contains some mathematical preliminaries. In Chapter 4 we review those bounds of [2] that we need for comparison. In Chapter 5 we will then show some numerical simulations which will give us a general picture about which way we should improve our bounds, *i.e*, make the upper bound lower or try to pull up the lower bound. Next, we will present our new lower bounds and their detailed derivations in Chapter 6 and 7, respectively.

For random quantities we use uppercase letters and for their realizations lowercase letters. Scalars are typically denoted using Greek letters or lowercase Roman letters. A few exceptions are the following symbols: C stands for capacity, h denotes the differential entropy, $\mathcal{E}$ stands for average power, A denotes peak power, and $\delta(\cdot)$ stands for the Kronecker delta. Moreover R, f, denote probability density functions (PDF):

- $R(\cdot)$ denotes a distribution on the channel output,
- $f_{\text{bin}}(\cdot)$, $f_{\text{tri}}(\cdot)$ denotes a distribution on the input signal.

We shall denote the mean-η, variance-σ^2 real Gaussian random variable by $\mathcal{N}_{\mathbb{R}}(\eta, \sigma^2)$. All rates specified in this paper are in nats per channel use, and all logarithms are natural logarithms.

Chapter 2

Channel Model

In this chapter, we will first introduce physical properties of this channel, then try to use mathematical equation to describe this channel. This chapter is based on [2].

2.1 Physical Description

In free-space optical communication the input signal usually is transmitted by means of some light emitting diodes (LED) or laser diodes (LD). Conventional and most inexpensive diodes emit light of wavelength between 850 and 950 nanometers, *i.e.*, the light lies in the infrared spectrum. For such high frequencies, amplitude or even frequency modulation (as usually used in radio communication) is technically very demanding. Therefore practical systems often apply the much simpler intensity modulation where the signal is modulated onto the optical intensity of the emitted light. Since the optical intensity is proportional to the optical power, in such a system the instantaneous optical power is proportional to the electrical input current (and not to its square as is usually the case for radio communication).

The receiver then directly measures the incident optical intensity of the incoming signal, *i.e.*, it produces an electrical current at the output which is proportional to the detected intensity.

For our model we will neglect the impact of fading or inter-symbol interference due to multiple-path propagation and assume that the direct line-of-sight path is dominant. However, we do take into account that the signal is corrupted by additive noise. The dominant noise source is assumed to be strong ambient light since there is no protective medium like, *e.g.*, a fiber cable. Even if optical filters are applied to reduce the impact of this noise, it typically has much larger power than the actual signal and causes high-intensity shot noise in the electrical output signal. In a first approximation this shot noise can be assumed to be independent of the signal itself.

The maximum allowed optical peak power of the transmitted signal has to be constrained due to eye safety reasons and due to the danger of potential thermal skin damage. Moreover, as battery power always is limited, we also constrain the maximum allowed optical average power.

2.2 Mathematical Channel Model

We will now try to translate the physical channel description into a simplified time-discrete channel model. The channel output $\tilde{Y}_k$ at time k, modeling a sample of the electrical output signal, is given by

$$\tilde{Y}_k = x_k + \tilde{Z}_k, \tag{2.1}$$

where $\tilde{Y}_k \in \mathbb{R}$ is the time-k channel output; $x_k \in \mathbb{R}^+$ is the time-k channel input and represents a sample of the electrical input current that is proportional to the optical intensity. The random process $\{\tilde{Z}_k\}$ models the additive noise. From our description in Section 2.1 we know that this noise is mainly caused by strong ambient light and we neglect multiple-path propagation effect. We assume that the noise is a constant intensity term η and some intensity-fluctuations around η. Because these fluctuations are caused by many independent sources, *e.g.*, shot noise or thermal noise, it is reasonable to assume that they are independent and identically distributed (IID) zero-mean Gaussian with a given variance σ^2. *I.e.*,

$$\{\tilde{Z}_k\} \sim \text{IID } \mathcal{N}_{\mathbb{R}}(\eta, \sigma^2). \tag{2.2}$$

Since η is constant, we can shift the whole expression in (2.1) by $-\eta$ without loss of generality and define a new channel output random variable Y_k which represents the disturbance of the electrical output signal. Moreover we note that our channel model is memoryless and therefore we drop the time-index k.

The channel output Y becomes

$$Y = x + Z, \tag{2.3}$$

where x is the channel input that is proportional to the optical intensity and cannot be negative,

$$x \in \mathbb{R}_0^+, \tag{2.4}$$

and where the additive noise is

$$Z \sim \mathcal{N}_{\mathbb{R}}(0, \sigma^2). \tag{2.5}$$

It is important to note that in contrast to the input x, the output Y may very well be negative since the noise introduced at the receiver is not necessarily positive.

The restrictions on the optical peak and average intensity are translated into a peak-power and an average-power constraint on the input, respectively:

$$\Pr[X > \mathsf{A}] = 0, \tag{2.6}$$

$$\mathsf{E}[X] \leq \mathcal{E}. \tag{2.7}$$

Be aware that the average-power constraint is in this case a constraint on the mean of the channel input and not—as is usually the case for channels modeling electro-magnetic radio communication—on the second moment. We shall denote the ratio between average power and peak power by α

$$\alpha \triangleq \frac{\mathcal{E}}{\mathsf{A}}. \tag{2.8}$$

where $0 < \alpha \leq 1$. If $\alpha = 1$, this means that the average-power constraint is inactive in the sense that we can use the maximum power all the time. On the other hand, if $\alpha \ll 1$, it means that the total power is limited in the sense that we have to limit our input power.

Chapter 3

Mathematical Preliminaries

3.1 $\mathcal{Q}$-Function

In the following we will extensively use the so-called $\mathcal{Q}$-function. This is a well-behaved function $\mathbb{R} \to (0,1)$ which maps every $\xi \in \mathbb{R}$ to the probability that a standard Gaussian random variable $\mathcal{N}_{\mathbb{R}}(0,1)$ takes on a value that exceeds ξ.

Definition 3.1. *The $\mathcal{Q}$-function is defined by*

$$\mathcal{Q}(\xi) \triangleq \int_{\xi}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \, \mathrm{d}t, \qquad \forall \xi \in \mathbb{R}. \tag{3.1}$$

The following properties of $\mathcal{Q}(\cdot)$ are well-known.

Lemma 3.1. *The $\mathcal{Q}$-function satisfies*

$$\mathcal{Q}(-\xi) + \mathcal{Q}(\xi) = 1, \qquad \forall \xi \in \mathbb{R}, \tag{3.2}$$

and

$$\mathcal{Q}(0) = \frac{1}{2}. \tag{3.3}$$

Moreover the first and second derivative of $\mathcal{Q}(\cdot)$ are given by

$$\mathcal{Q}'(\xi) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}, \qquad \forall \xi \in \mathbb{R}, \tag{3.4}$$

and

$$\mathcal{Q}''(\xi) = \frac{\xi}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}, \qquad \forall \xi \in \mathbb{R}. \tag{3.5}$$

Thus $\mathcal{Q}$ is monotonically decreasing for all $\xi \in \mathbb{R}$, concave over $(-\infty, 0]$, and convex over $[0, \infty)$ Also the $\mathcal{Q}$-function can be bounded by

$$\frac{1}{\sqrt{2\pi}\xi} e^{-\frac{\xi^2}{2}} \left(1 - \frac{1}{\xi^2}\right) < \mathcal{Q}(\xi) < \frac{1}{\sqrt{2\pi}\xi} e^{-\frac{\xi^2}{2}}, \qquad \xi > 0, \tag{3.6}$$

and

$$\mathcal{Q}(\xi) \leq \frac{1}{2} e^{-\frac{\xi^2}{2}}, \qquad \xi \geq 0. \tag{3.7}$$

Proof. See *e.g.*, [2, Section 3].

3.2 Properties of the Channel Model

First, we note that it has been shown in [3] that the capacity-achieving input distribution is unique.

Lemma 3.2. *Fix a peak power* A *and an average power* $\mathcal{E}$*. Let* $\mathcal{C}(\mathsf{A},\mathcal{E})$ *denote the capacity-cost function of the free-space optical intensity channel* (2.3) *given as*

$$\mathcal{C}(\mathsf{A},\mathcal{E}) = \sup_{f} I(f, W). \tag{3.8}$$

where W *denotes the channel law, and where the supremum is over all input distributions* $f \in \mathcal{P}\left(\mathbb{R}_0^+\right)$ *that satisfy the peak- and average-power constraints* (2.6) *and* (2.7)*. Then there exists a unique input distribution* $f^*_{\mathsf{A},\mathcal{E}}$ *that achieves the supremum.*

Proof. See [3].

Using Lemma 3.2 together with the symmetry of the channel law and the concavity of channel capacity in the input distribution, we can make the following observation.

Lemma 3.3. *Consider a peak-power constraint* A *and an average-power constraint* $\mathcal{E}$ *such that* $\alpha = \frac{\mathcal{E}}{\mathsf{A}} \geq \frac{1}{2}$*. Then the optimal input distribution in will have average power equal to half of the peak power*

$$\frac{\mathsf{E}_{f^*}[X]}{\mathsf{A}} = \frac{1}{2}, \tag{3.9}$$

irrespective of α*.* I.e., *the average-power constraint is inactive for all* $\alpha \in \left(\frac{1}{2}, 1\right]$*, and in particular*

$$\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) = \mathcal{C}\left(\mathsf{A}, \frac{\mathsf{A}}{2}\right), \quad \frac{1}{2} \leq \alpha \leq 1. \tag{3.10}$$

Proof. See [2, Appendix D].

3.3 Jensen's Inequality

Lemma 3.4. *(Jensen's inequality) Consider a convex function* $g(\cdot)$ *and a random variable* X*. Then*

$$\mathsf{E}[g(X)] \geq g(\mathsf{E}[X]). \tag{3.11}$$

Moreover, if $g(\cdot)$ *is strictly convex, then equality in* (3.11) *implies that* $X = \mathsf{E}[X]$ *with probability 1 (*i.e., X *is a constant).*

Proof. See *e.g.*, [1, Theorem 2.6.2].

Chapter 4

Previous Results

In this chapter, we review some previous results based on [2]. In Section 4.1 we review the bounds on channel capacity with both an average- and a peak-power constraint. In Section 4.2 we review the bounds on channel with a strong peak-power and inactive average-power constraint. In Section 4.3 we review the bounds on channel capacity just with average-power constraint.

4.1 Bounds on Channel Capacity with both an Average- and a Peak-Power Constraint

We use the channel model (2.3) defined in Section 2.2 and assume that the input distribution is constrained by (2.6) and (2.7). Then the channel capacity $\mathcal{C}(\mathsf{A},\mathcal{E})$ is bounded as follows:

$$\mathcal{C}(\mathsf{A},\alpha\mathsf{A}) \geq \frac{1}{2}\log\left(1+\mathsf{A}^2\frac{e^{2\alpha\mu^*}}{2\pi e\sigma^2}\left(\frac{1-e^{-\mu^*}}{\mu^*}\right)^2\right), \tag{4.1}$$

$$\mathcal{C}(\mathsf{A},\alpha\mathsf{A}) \leq \frac{1}{2}\log\left(1+\alpha(1-\alpha)\frac{\mathsf{A}^2}{\sigma^2}\right), \tag{4.2}$$

$$\begin{aligned}\mathcal{C}(\mathsf{A},\alpha\mathsf{A}) \leq & \left(1-\mathcal{Q}\left(\frac{\delta+\alpha\mathsf{A}}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+(1-\alpha)\mathsf{A}}{\sigma}\right)\right)\cdot \\ & \cdot\log\left(\frac{\mathsf{A}}{\sigma}\frac{e^{\frac{\mu\delta}{\mathsf{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathsf{A}}\right)}}{\sqrt{2\pi}\mu(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right))}\right)-\frac{1}{2}+\mathcal{Q}\left(\frac{\delta}{\sigma}\right) \\ & +\frac{\delta}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}}+\frac{\sigma}{\mathsf{A}}\frac{\mu}{\sqrt{2\pi}}\left(e^{-\frac{\delta^2}{2\sigma^2}}-e^{-\frac{(\mathsf{A}+\delta)^2}{2\sigma^2}}\right) \\ & +\mu\alpha\left(1-2\mathcal{Q}\left(\frac{\delta+\frac{\mathsf{A}}{2}}{\sigma}\right)\right).\end{aligned} \tag{4.3}$$

Here μ^* is the positive solution to

$$\alpha = \frac{1}{\mu^*}-\frac{e^{-u^*}}{1-e^{-\mu^*}} \tag{4.4}$$

the function $\mathcal{Q}(\cdot)$ is defined in (3.1); and $\mu \geq 0$ and $\delta \geq 0$ are free parameters.

In [2] the authors suggest the following (suboptimal) choice for these parameters:

$$\delta = \delta(\mathsf{A}) \triangleq \sigma \log\left(1 + \frac{\mathsf{A}}{\sigma}\right), \tag{4.5}$$

$$\mu = \mu(\mathsf{A}) \triangleq \mu^* \left(1 - e^{-\alpha \frac{\delta^2}{2\sigma^2}}\right), \tag{4.6}$$

where μ^* is the positive solution to (4.4).
For this choice of parameters the bounds (4.1), (4.2), and (4.3) are depicted in Figures 5.1, 5.2, and 5.3.

4.2 Bounds on Channel Capacity with a Strong Peak-Power and Inactive Average-Power Constraint

From [2] we know that for $\frac{1}{2} < \alpha \leq 1$ the average-power constraint is inactive and $\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) = \mathcal{C}(\mathsf{A}, \frac{1}{2}\mathsf{A})$. Thus we can obtain the results in this section by simply deriving bounds for the case $\alpha = \frac{1}{2}$. Indeed the lower bound is valid for $\alpha \in [\frac{1}{2}, 1]$ and the upper bounds are even valid for any $\alpha \in (0, 1]$.

$$\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) \geq \frac{1}{2} \log\left(1 + \frac{\mathsf{A}^2}{2\pi e \sigma^2}\right), \tag{4.7}$$

$$\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) \leq \frac{1}{2} \log\left(1 + \frac{\mathsf{A}^2}{4\sigma^2}\right), \tag{4.8}$$

$$\begin{aligned}\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) \leq{}& \left(1 - 2\mathcal{Q}\left(\frac{\delta + \frac{\mathsf{A}}{2}}{\sigma}\right)\right) \log \frac{\mathsf{A} + 2\delta}{\sigma\sqrt{2\pi}\left(1 - 2\mathcal{Q}(\frac{\delta}{\sigma})\right)} - \frac{1}{2} \\ &+ \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma} e^{-\frac{\delta^2}{2\sigma^2}}.\end{aligned} \tag{4.9}$$

Here the function $\mathcal{Q}(\cdot)$ is defined in (3.1), and $\delta > 0$ is a free parameter. In [2] the authors suggest the following (suboptimal) choice for δ:

$$\delta = \delta(\mathsf{A}) \triangleq \sigma \log\left(1 + \frac{\mathsf{A}}{\sigma}\right). \tag{4.10}$$

For this choice (4.7), (4.8), and (4.9) are depicted in Figure 5.4.

4.3 Bounds on Channel Capacity with an Average-Power Constraint

Finally, we consider the case with an average-power constraint only. Here we change our notation $\mathcal{C}(\mathsf{A}, \alpha\mathsf{A})$ to $\mathcal{C}(\mathcal{E})$. The channel capacity $\mathcal{C}(\mathcal{E})$ is bounded as follows:

$$\mathcal{C}(\mathcal{E}) \geq \frac{1}{2}\log\left(1+\frac{\mathcal{E}^2 e}{2\pi\sigma^2}\right), \tag{4.11}$$

$$\begin{aligned}\mathcal{C}(\mathcal{E}) \leq{}& \log\left(\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) - \log(\sqrt{2\pi}\sigma)\\ &- \frac{\delta\mathcal{E}}{2\sigma^2} + \frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta}{\sigma}\right) - \frac{\mathcal{E}}{\delta}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)\\ &+ \frac{1}{\beta}\left(\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}\right), \qquad \delta \leq -\frac{\sigma}{\sqrt{e}},\end{aligned} \tag{4.12}$$

$$\begin{aligned}\mathcal{C}(\mathcal{E}) \leq{}& \log\left(\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) + \frac{1}{2}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\\ &+ \frac{\delta}{2\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} + \frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta+\mathcal{E}}{\sigma}\right)\right)\\ &+ \frac{1}{\beta}\left(\delta+\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right) - \frac{1}{2}\log 2\pi e\sigma^2, \qquad \delta \geq 0.\end{aligned} \tag{4.13}$$

where $\beta > 0$ and δ are free parameters. Note that bound (4.12) only holds for $\delta \leq -\sigma e^{-\frac{1}{2}}$, while bound (4.13) only holds for $\delta \geq 0$. A suboptimal choice for the free parameters in bound (4.12) is

$$\delta = \delta(\mathcal{E}) \triangleq -2\sigma\sqrt{\log\frac{\sigma}{\mathcal{E}}}, \qquad \text{for } \frac{\mathcal{E}}{\sigma} \leq e^{-\frac{1}{4e}} \approx -0.4\text{ dB}, \tag{4.14}$$

$$\begin{aligned}\beta = \beta(\mathcal{E}) \triangleq{}& \frac{1}{2}\left(\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}\right)\\ &+ \frac{1}{2}\sqrt{\left(\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}\right)^2 + 4\left(\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}\right)\sqrt{2\pi}\sigma e^{\frac{\delta^2}{2\sigma^2}}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)},\end{aligned} \tag{4.15}$$

and for the free parameters in bound (4.13) is

$$\delta = \delta(\mathcal{E}) \triangleq \sigma\log\left(1+\frac{\mathcal{E}}{\sigma}\right), \tag{4.16}$$

$$\begin{aligned}\beta = \beta(\mathcal{E}) \triangleq{}& \frac{1}{2}\left(\delta+\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right)\\ &+ \frac{1}{2}\left(\left(\delta+\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right)^2\right.\\ &\left.+4\left(\delta+\mathcal{E}+\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right)\sqrt{2\pi}\sigma e^{\frac{\delta^2}{2\sigma^2}}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)^{\frac{1}{2}}.\end{aligned} \tag{4.17}$$

For this choice of parameters the bounds (4.11), (4.12), and (4.13) are depicted in Figure 5.5.

Chapter 5

Numerical Lower Bounds

In this chapter, we will use a numerical approach to compute the channel capacity and compare it with previous results, in particular, with the bounds shown in Chapter 4. In Section 5.1 we will give readers a rough picture about how we programed the simulation. In Section 5.2 we will introduce some input distributions and show that these distributions satisfy the peak- and average-power constraint. In Section 5.3.4 we present the numerical results and make some comments.

5.1 Introduction

The process of calculating the lower bound is based on the following ideas: considering the definition of capacity from Lemma 3.2

$$\mathcal{C}(\mathsf{A}, \mathcal{E}) = \sup_f I(f, W) \tag{5.1}$$

where the supremum is over all distributions f on the channel input that satisfy the power constraints, we drop the optimization and choose one particular f. This leads to a natural lower bound on capacity.

$$\begin{aligned}
\mathcal{C}(\mathsf{A}, \mathcal{E}) &= \sup_f I(f, W) &(5.2)\\
&\geq I(f, W)|_{\text{for a specific } f} &(5.3)\\
&= (h(Y) - h(Y|X))|_{\text{for a specific } f} &(5.4)\\
&= h(X + Z)|_{\text{for a specific } f} - h(Z). &(5.5)
\end{aligned}$$

Now we need to choose a distribution f that is reasonably close to the capacity-achieving input distribution in order to get a tight lower bound. However, we might still have the problem that the evaluation of $h(Y)$ is hard and complicated. The reason for this is that even for relatively "easy" distributions f, the corresponding distribution on the channel output may be difficult to compute, let alone $h(Y)$.

So we use a numerical approach to get an impression of the channel capacity. The idea is as follows: remember the definition of entropy

$$h(Y) = -\int_{-\infty}^{\infty} R(y) \cdot \log R(y)\, \mathrm{d}y \tag{5.6}$$

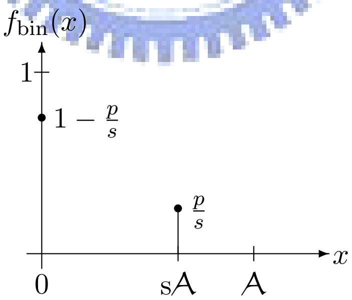
where $R(y)$ is the output distribution. We can choose a simple input distribution which we describe our choices in Section 5.2 and then calculate the corresponding output distribution. We then use MATLAB to numerically evaluate the integration in (5.6). Because the output distribution is a continuous function, we can use the command "quad" in MATLAB to compute the entropy.

Note that we often experience numerical problems: When we calculate some parts of the entropy, *e.g.*, the probability of the tail and the probability of a "valley" between two "mountains", it might happen that the calculations result is a Not-a-Number value. We know that they are very small values, so we can ignore them and set them to zero.

5.2 Input Distributions

5.2.1 Input Distributions in the Situation of an Average- and a Peak-Power Constraint

In this section we will introduce two input distributions with peak-power constraint. First, we use the binary distribution which is the most basic distribution and easy to handle. The distribution is defined as follows:

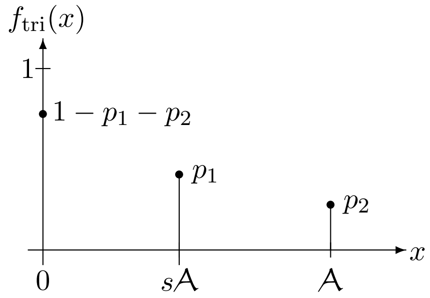


$$f_{\text{bin}}(X) = \left(1 - \frac{p}{s}\right) \cdot \delta(x) + \frac{p}{s} \cdot \delta(x - s\mathsf{A}), \qquad \alpha \le s \le 1\,; \tag{5.7}$$

$$\mathsf{E}[X] = \left(1 - \frac{p}{s}\right) \cdot 0 + \frac{p}{s} \cdot s\mathsf{A} \;=\; p\mathsf{A}. \tag{5.8}$$

From (5.7) we can see that we fix one symbol at 0 and change another symbol from $\alpha\mathsf{A}$ to A. The parameter s is limited by the peak-power constraint and by the basic property of probability that the summation of all events must be one. From (5.8) we check that the distribution satisfies the average-power constraint.

Next, we introduce a more elaborate distribution, the ternary input distribution. Note that our purpose here is not to find the maximum input distribution which achieves the channel capacity. Thus our choice is not necessarily optimal. We fix two of the three symbols at 0 and A, so we can only modify the position of the third symbol. The distribution is defined as follows:



$$f_{\text{tri}}(x) = (1 - p_1 - p_2) \cdot \delta(x) + p_1 \cdot \delta(x - s\mathsf{A}) + p_2 \cdot \delta(x - \mathsf{A})\,; \tag{5.9}$$

$$p_1 = \frac{\alpha - p_2}{s}, \qquad (\alpha - p_2) \le s \le 1\,; \tag{5.10}$$

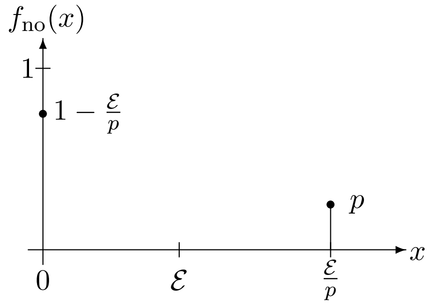
$$p_2 \le \alpha\,; \tag{5.11}$$

$$\begin{aligned} \mathsf{E}[X] &= (1 - p_1 - p_2) \cdot 0 + p_1 \cdot s\mathsf{A} + p_2 \cdot \mathsf{A} \\ &= \frac{\alpha - p_2}{s} \cdot s\mathsf{A} + p_2 \cdot \mathsf{A} = \alpha\mathsf{A}. \end{aligned} \tag{5.12}$$

The basic structure is the same as for the binary distribution, but now we have one more symbol. We use p_2 to control how much power we want to give the symbol which costs most power, then we use s to control how much of the remaining power we want to give the symbol which costs fewer power. Note that we need to satisfy (5.11) and (5.10) in order to guarantee that our distribution satisfies the average-power constraint and the basic property of probability that the summation of all events must be one. From (5.12) we check that the distribution satisfies the average-power constraint.

5.2.2 Input Distribution in the Situation of an Average-Power Constraint Only

In this section we will introduce a distribution just with average-power constraint. This is slightly different from Section 5.2.1. Without peak-power constraint, we can use higher transmit power to combat the noise in order to get more capacity. Note that the capacity here is higher than in the situation with both an average- and a peak-power constraint. We define the distribution as follows



$$f_{\text{no}}(X) = \left(1-\frac{\mathcal{E}}{p}\right)\cdot\delta(x) + \frac{\mathcal{E}}{p}\cdot\delta(x-\frac{\mathcal{E}}{p}), \qquad 0<p<1\,; \tag{5.13}$$

$$\mathsf{E}[X] = \left(1-\frac{\mathcal{E}}{p}\right)\cdot 0 + p\cdot\frac{\mathcal{E}}{p} \;=\; \mathcal{E}. \tag{5.14}$$

The basic structure is the same as for the binary distribution, but now we have no peak-power constraint. So we remove the notation A, then change to $\mathcal{E}$ and use p to satisfy the average-power constraint. From (5.14) we check that the distribution satisfies the average-power constraint.

5.3 Simulation Results

We will now use the input distributions in Sections 5.2.1 and 5.2.2 to compute numerical lower bounds. In the first case we have an average- and a peak-power constraint. We will use three different α values to show how much capacity we lose. Before we start to introduce our simulation results, we first explain how the α value effects our capacity. Recall the expectation of binary distribution in (5.8) and set s equal 1 as follows:

$$\begin{aligned}\mathsf{E}[X] &= \sum x\cdot f(x)\\ &= (1-p)\cdot 0 + p\cdot\mathsf{A} = p\mathsf{A},\end{aligned} \tag{5.15}$$

and recall the definition of α as follows:

$$\alpha = \frac{\mathcal{E}}{\mathsf{A}} \Longrightarrow \mathcal{E} = \alpha\mathsf{A}. \tag{5.16}$$

By (5.15) and (5.16) we can see that $p=\alpha$ in this distribution. If α is very small, we have fewer chance to transmit signal. Although in high-SNR region we can not use peak-power to transmit, that means we set s small than 1 in order to get higher chance to transmit. But in the low-SNR region, if we do the same way as the high-SNR region. The noise will destroy the signal, because the signal power is small than noise too much. Now we use $\alpha = 0.02$ to represent strong average-power constraint, $\alpha = 0.1$ to represent normal average-power constraint, and $\alpha = 0.4$ to represent loose average-power constraint.

In the second case we have a strong peak-power and an inactive average-power constraint. According to Lemma 3.3, we just use $\alpha = \frac{1}{2}$ to compute the channel capacity, because it can achieve the maximum lower bound when α equals $\frac{1}{2}$.

Finally we ignore the peak-power constraint and just focus on the average-power constraint.

5.3.1 Simulation Results in the Situation of an Average- and a Peak-Power Constraint

Now we start from first case: the input distribution satisfies both average- and peak-power constraint. From Figure 5.1 with $\alpha = 0.02$, we see that the numerical lower bound is almost the same as (4.2) in the region below 5 dB. When the peak power is larger than 5 dB, we can observe that the gap between numerical lower bound and upper bound becomes larger and larger.

From Figure 5.2 with $\alpha = 0.1$, we can see that when the peak power is lower than 5 dB, the upper bound (4.2) is still very close to numerical lower bound. Besides, the ternary input distribution starts to exceed the binary input distribution about 6 dB, in the sense that we can get a tighter bound due to the more complex input distribution.

From Figure 5.3 with $\alpha = 0.4$, the upper bound (4.2) is still a nice approximation of channel capacity below 5 dB and we observe that the ternary input distribution is not only close to the upper bound (4.2), but also to the upper bound (4.3) above 5 dB. This shows that when α is large, the upper bound (4.3) is also tight for medium SNR.

5.3.2 Simulation Results in the Situation of a Peak-Power Constraint Only

The second case is the input distribution with strong peak-power and inactive average-power constraints. From Figure 5.4 with $\alpha = 0.5$, we see that the numerical lower bound is almost the same as the upper bound (4.8) in the region which is below 5 dB. The upper bound (4.9) is also close to the ternary input distribution after 6 dB. The whole situation is similar to the previous case with $\alpha = 0.4$.

5.3.3 Simulation Results in the Situation of an Average Peak-Power Constraint Only

The last case is the input distribution just with a average-power constraint. From Figure 5.5, we can see that the numerical lower bound is not as tight as in the previous cases in the Sections 5.3.1 and 5.3.2, but the upper bound (4.12) becomes close only for values below -20 dB.

5.3.4 Summary

After we compare these numerical lower bounds with the previous results in [2], we learn that the lower bounds in [2] are too loose, while the upper bounds seem tight. Hence, we need to find an analytic way to improve on the lower bounds.

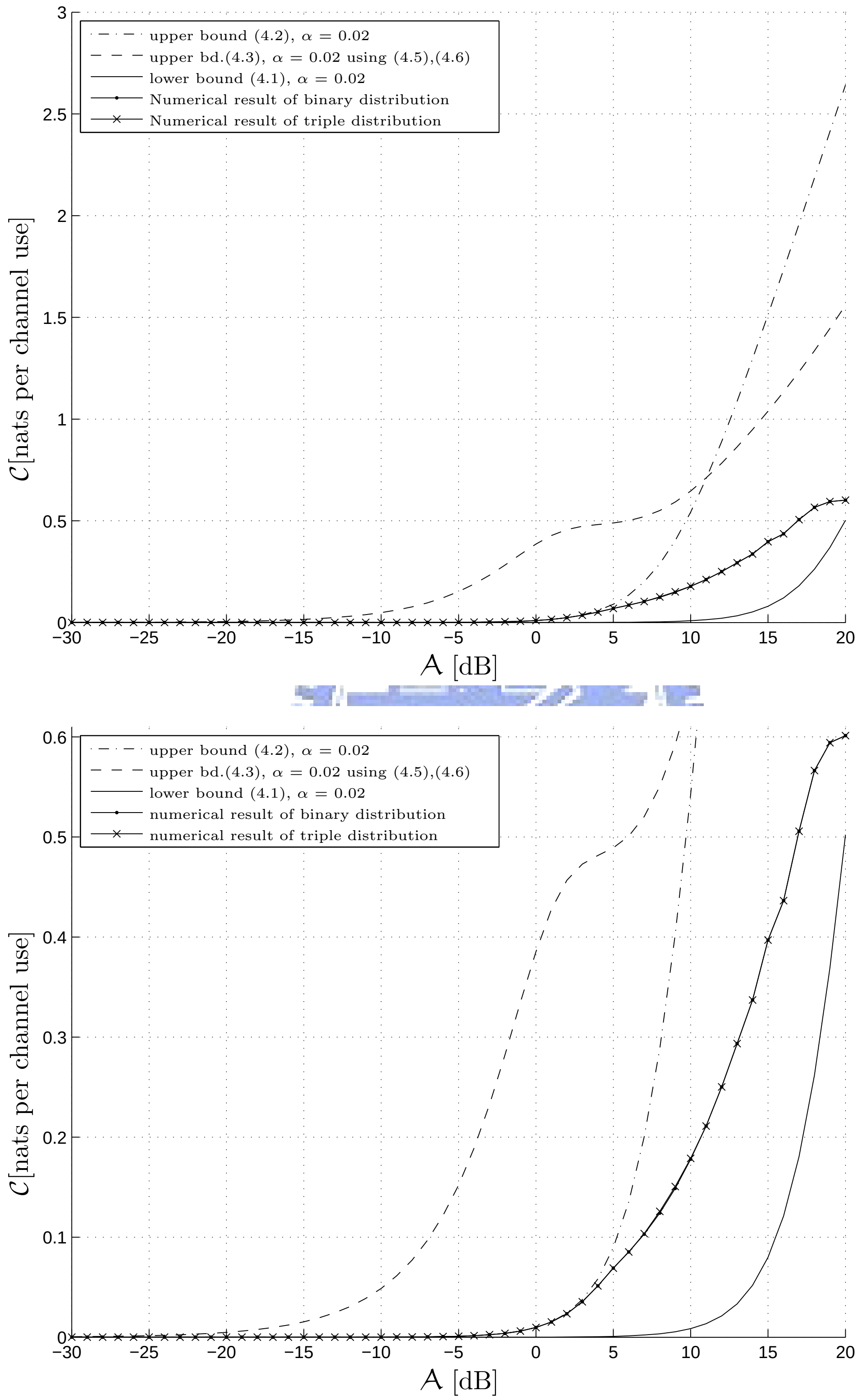


Figure 5.1: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.02$.

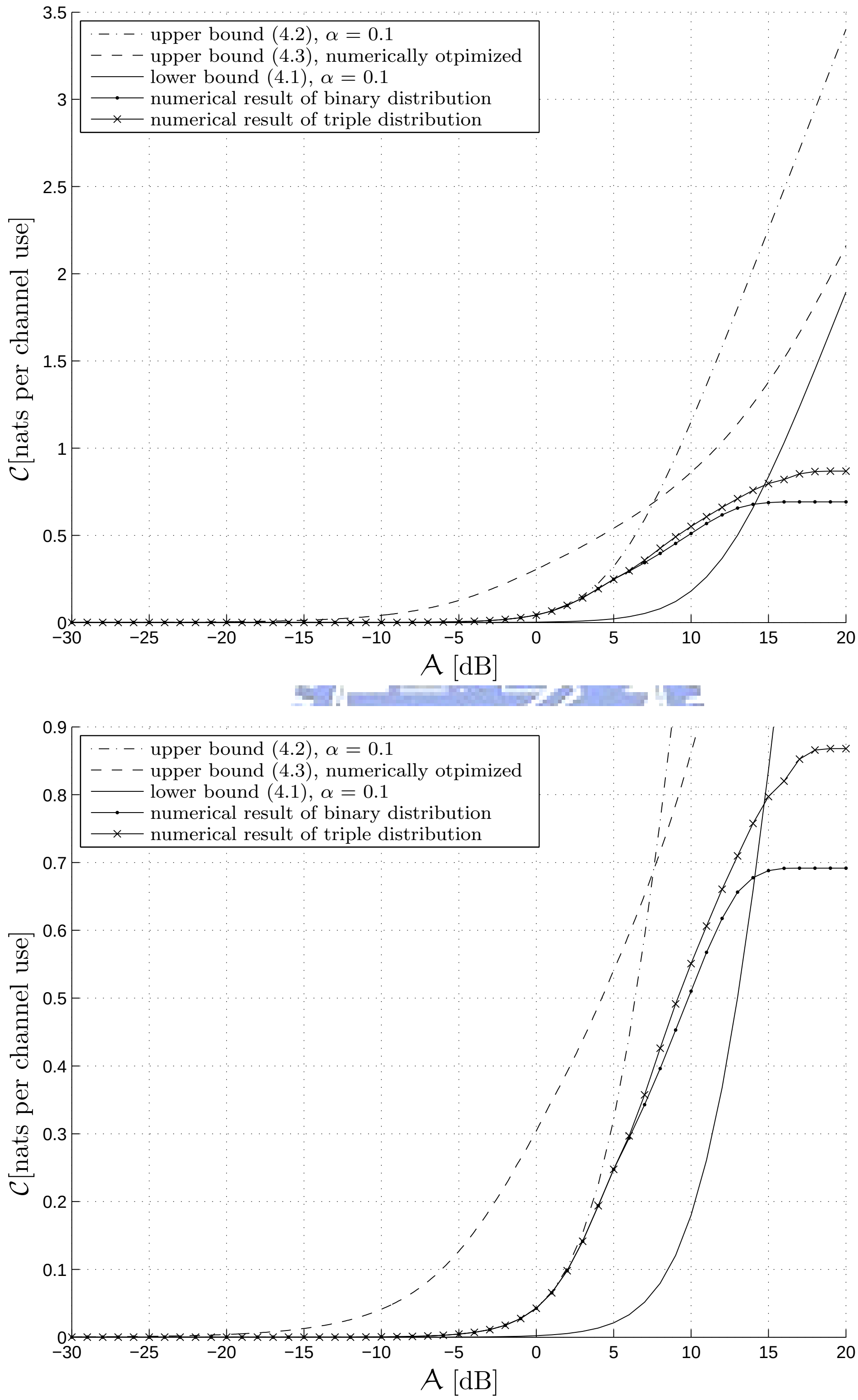


Figure 5.2: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.1$.

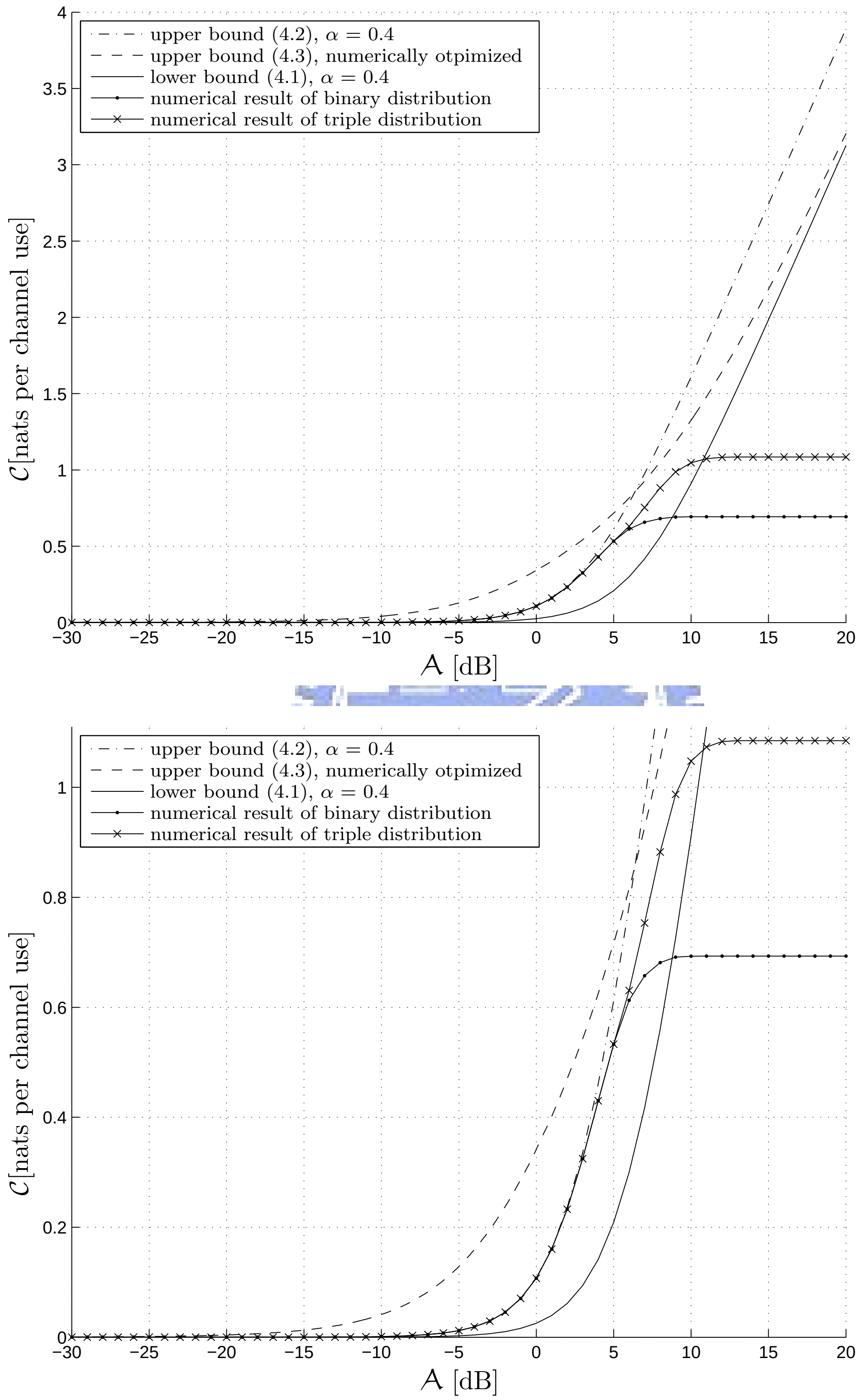


Figure 5.3: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.4$.

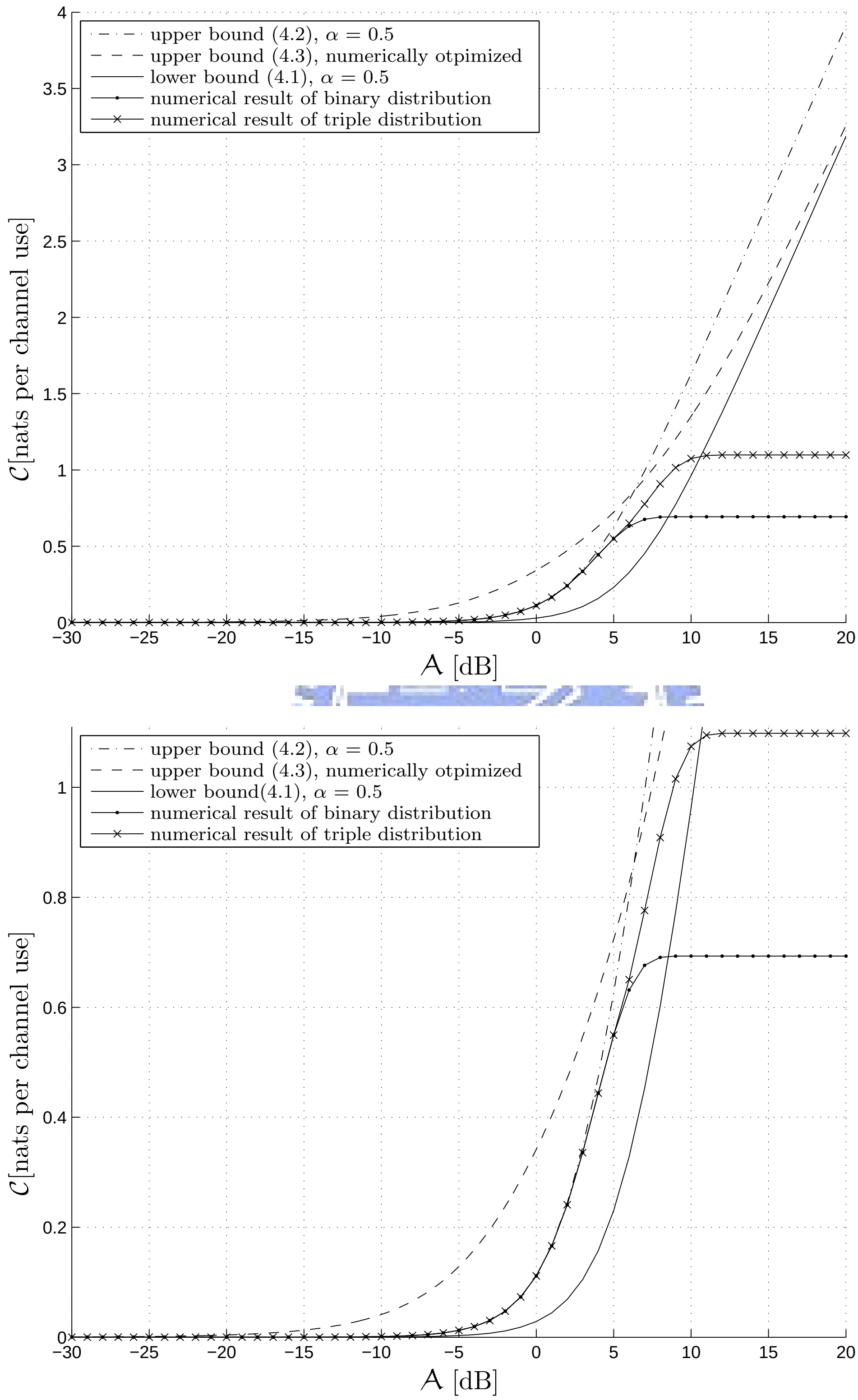


Figure 5.4: Bounds on the capacity of the free-space optical intensity channel with strong peak-power and inactive average-power constraint.

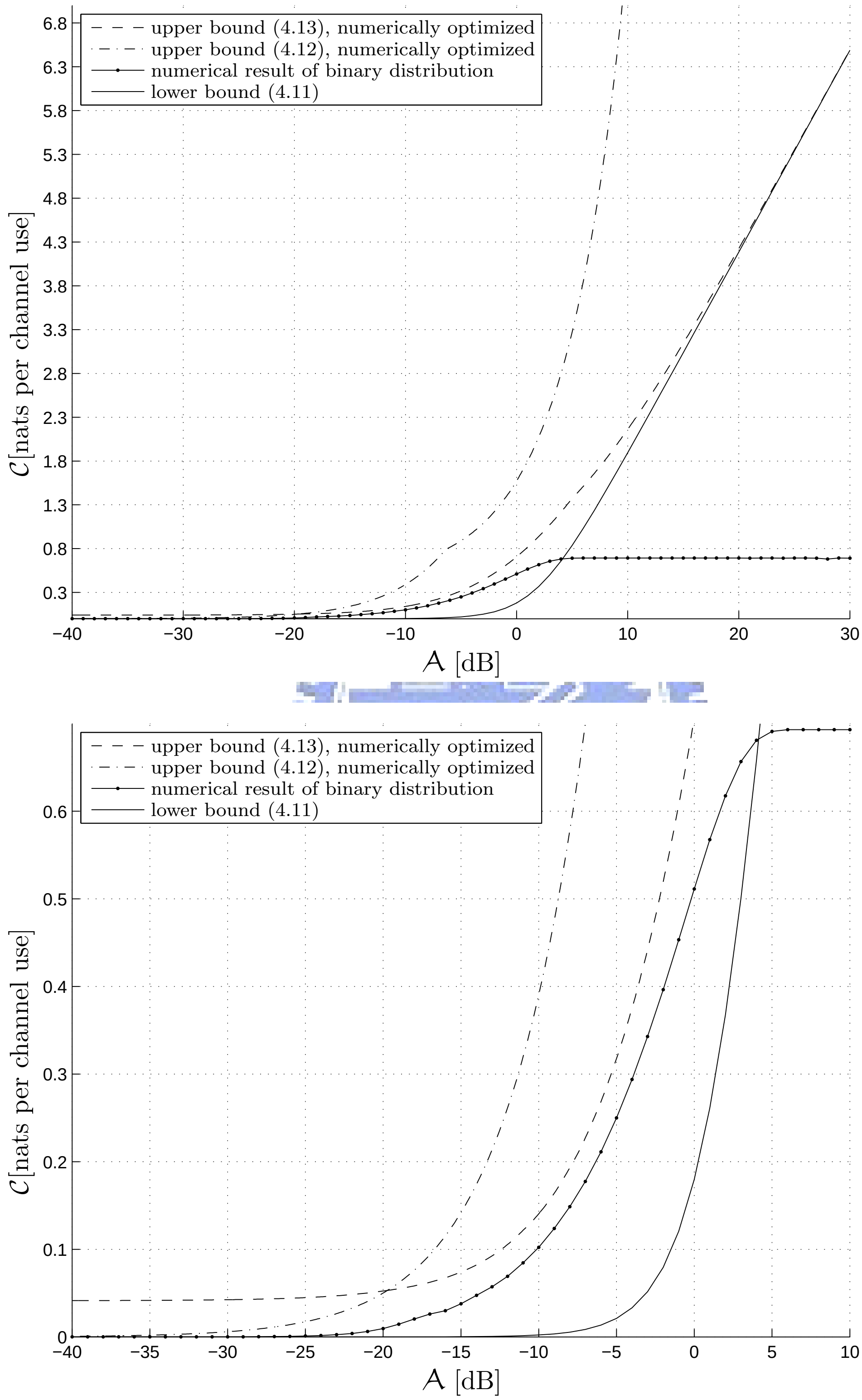


Figure 5.5: Bounds on the capacity of the free-space optical intensity channel just with average-power constraint.

Chapter 6

Analytical Lower Bounds

According to Chapter 5, we know that the upper bounds in [2] are pretty tight in the low-SNR region. It is hard to improve the upper bounds, but the gap between the lower bounds and the numerical results is still large. We will now state new lower bounds on the capacity of the channel defined in Section 2.2.

6.1 Lower Bounds with both an Average- and a Peak-Power Constraint

Theorem 6.1. *Assume the channel model* (2.3) *defined in Section 2.2 and assume that the input is constrained by* (2.6) *and* (2.7) *where the average-to-peak-power ration* $\alpha = \frac{\mathcal{E}}{\mathsf{A}}$ *lies in* $(0, 1]$. *Then the channel capacity* $\mathcal{C}(\mathsf{A}, \mathcal{E})$ *is bounded as follows:*

$$\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) \geq -\log\left(\frac{\left(1-\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\left(\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\frac{\alpha}{s}\left(1-\frac{\alpha}{s}\right)}{\sigma\sqrt{\pi}} e^{-\frac{s^2\mathsf{A}^2}{4\sigma^2}}\right) - \frac{1}{2}\log\left(2\pi e\sigma^2\right), \tag{6.1}$$

$$\begin{aligned}\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) \geq -\log\Bigg(&\frac{(1-p_1-p_2)^2}{2\sigma\sqrt{\pi}} + \frac{{p_1}^2}{2\sigma\sqrt{\pi}} + \frac{{p_2}^2}{2\sigma\sqrt{\pi}} \\ &+ \frac{p_1(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A})^2}{4\sigma^2}} + \frac{p_2(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{\mathsf{A}^2}{4\sigma^2}} \\ &+ \frac{p_1 p_2}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A}+\mathsf{A})^2}{4\sigma^2}}\Bigg) - \frac{1}{2}\log\left(2\pi e\sigma^2\right).\end{aligned} \tag{6.2}$$

where $s \in (\alpha, 1)$ *in* (6.1) *and* $p_2 \in (0, \alpha)$, $s \in (\alpha - p_2, 1)$ *in* (6.2) *are free parameter.*

The new lower bounds (6.1) and (6.2) are depicted in Figure 6.1, 6.2, 6.3, and 6.4. In these plots we have chosen the free parameters by numerical optimization.

Now we compare our new lower bounds with the previous results in [2] and see how much we have improved.

We start from the first case: input distribution satisfying average- and peak-power constraint. From Figure 6.1 with $\alpha = 0.02$, we can see that we can get a little gain between 12 dB and 20 dB, but the bound is useless for SNR value below 12 dB.

From Figure 6.2 with $\alpha = 0.1$, we can get more gain starting from 5 dB and the negative value region is reduced. Besides, the ternary input distribution start to exceed the binary input distribution about 12 dB.

From Figure 6.3 with $\alpha = 0.4$, we can get even more gain starting from 2 dB and the analytic lower bounds are getting close to the numerical lower bounds. Unfortunately, there are still negative values below 2 dB, but in the region between 2 dB and 8 dB, we almost reduce half of the distance between upper bound (4.2) and lower bound (4.1).

As the next case we describe input distributions which satisfy a strong peak-power and an inactive average-power constraint. From Figure 6.4 with $\alpha = 0.5$, we know that we can reduce half of the distance between upper bound (4.8) and lower bound (4.7) in the region between 2 dB and 9 dB, but we still have negative values below 1 dB.

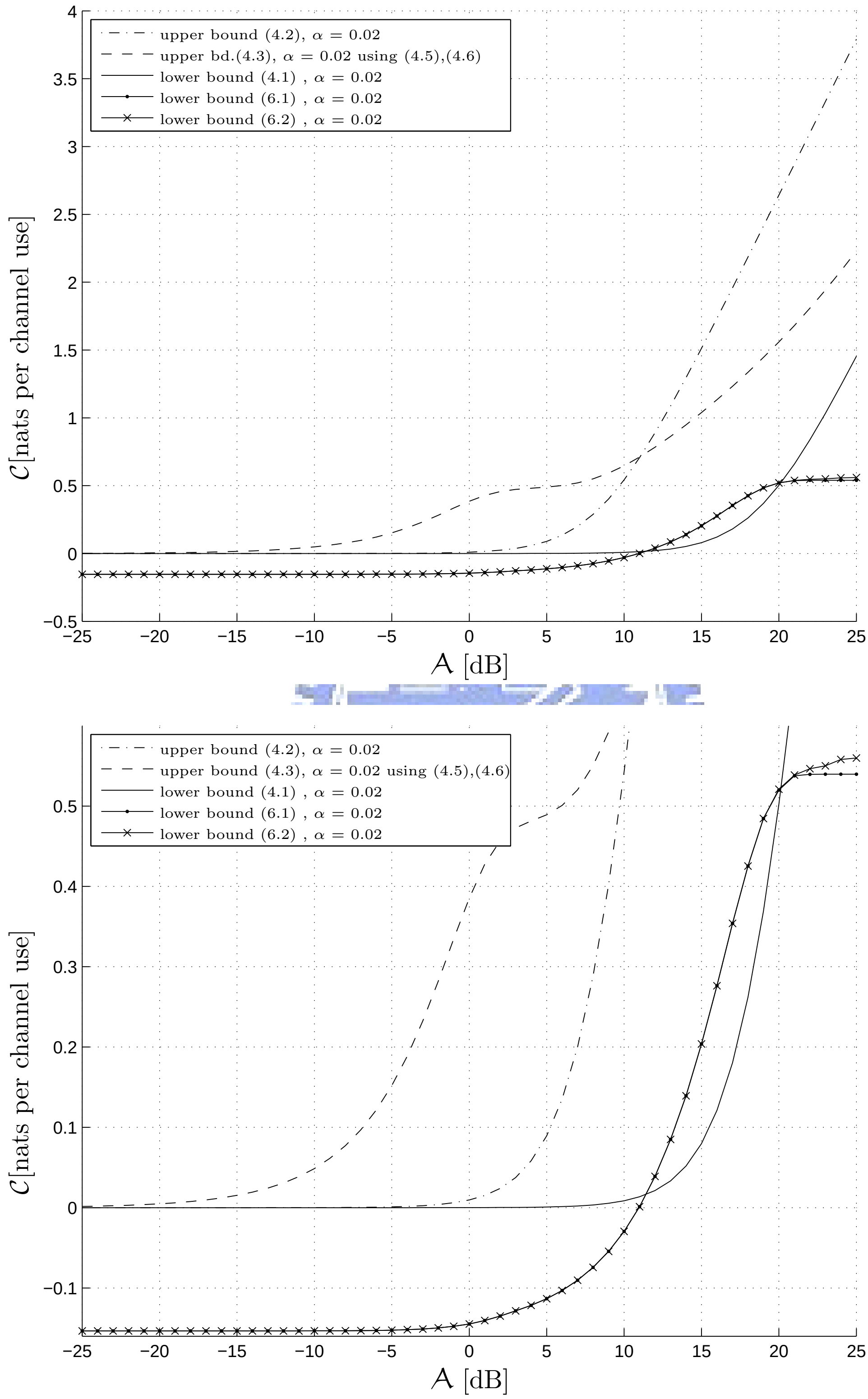


Figure 6.1: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.02$.

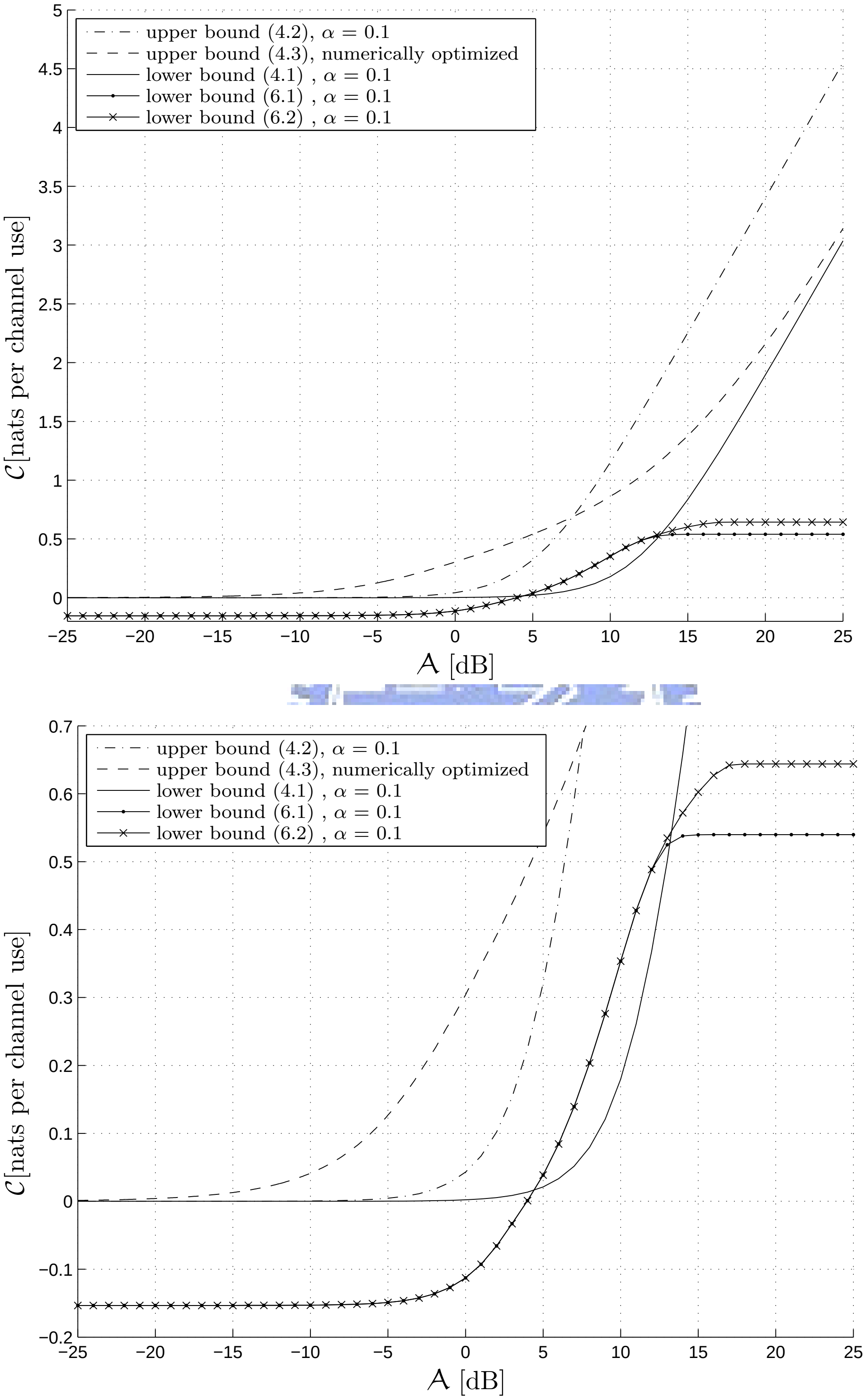


Figure 6.2: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.1$.

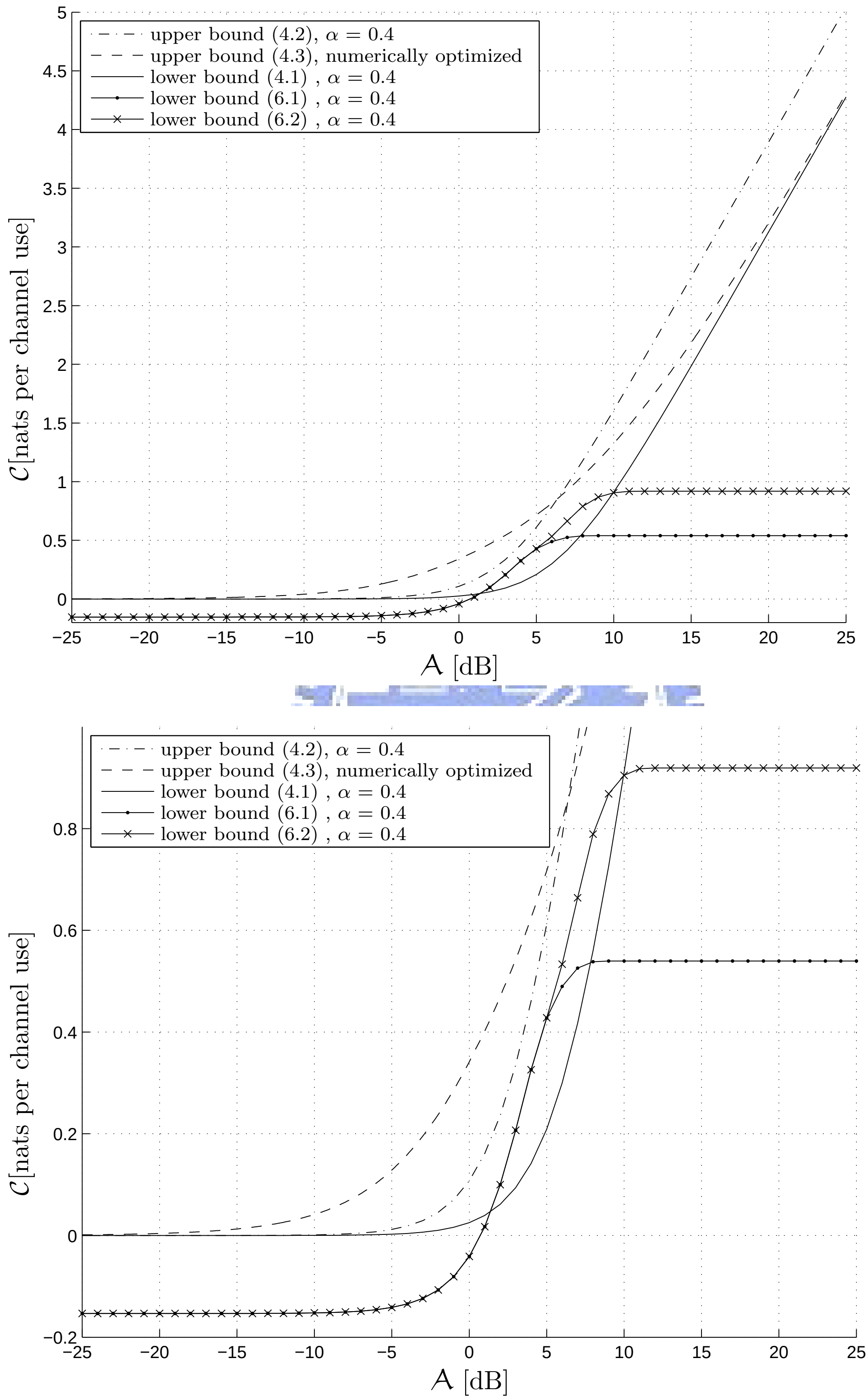
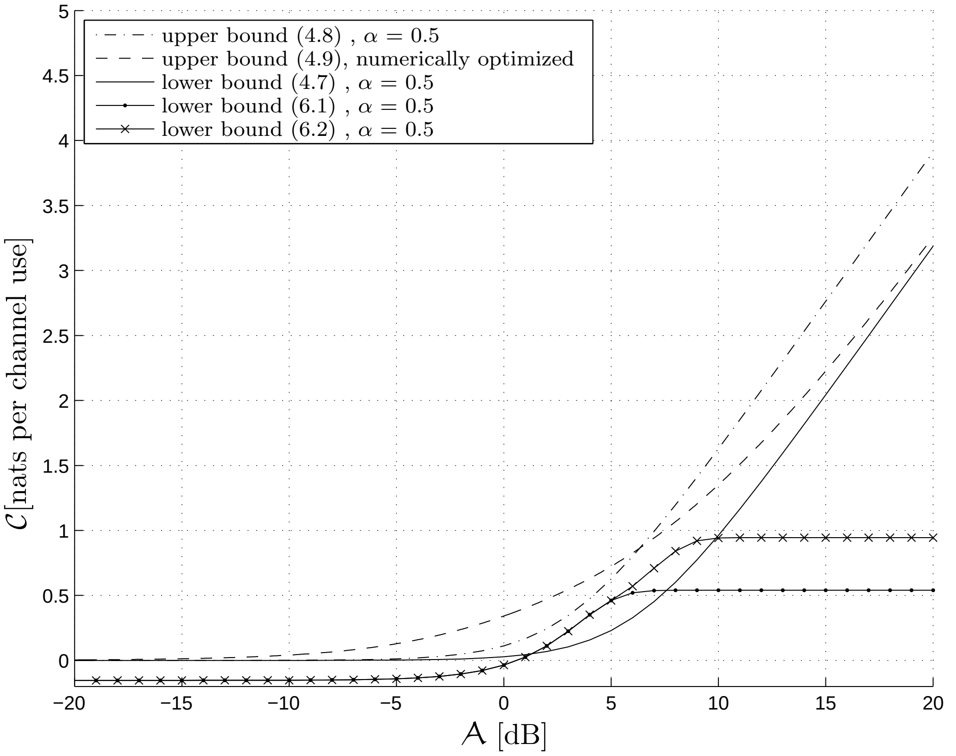


Figure 6.3: Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha = 0.4$.



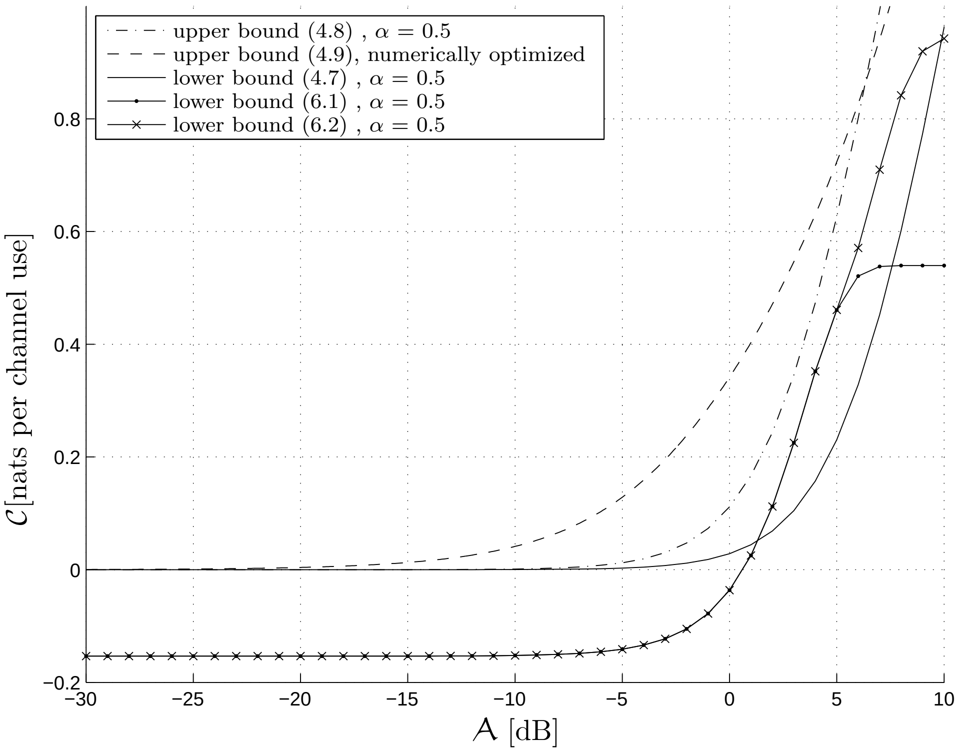


Figure 6.4: Bounds on the capacity of the free-space optical intensity channel with strong peak-power and inactive average-power constraint.

6.2 Lower Bounds with an Average-Power Constraint Only

Theorem 6.2. *Assume the channel model* (2.3) *defined in Section 2.2 and assume that the input is constrained by an average-power constraint* (2.7) *(but no constraint on the peak power). Then the channel capacity* $\mathcal{C}(\mathcal{E})$ *is bounded as follows:*

$$\mathcal{C}(\mathcal{E}) \geq -\log\left(\frac{(1-p)^2}{2\sigma\sqrt{\pi}} + \frac{p^2}{2\sigma\sqrt{\pi}} + \frac{p(1-p)}{\sigma\sqrt{\pi}} e^{-\frac{\mathcal{E}^2}{4p^2\sigma^2}}\right) - \frac{1}{2}\log\left(2\pi e\sigma^2\right), \tag{6.3}$$

$$\begin{aligned}\mathcal{C}(\mathcal{E}) \geq &-\Bigg[1 - p - \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) + 2p\mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) + \frac{2p\sigma}{\mathcal{E}\sqrt{2\pi}} e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} \\ &- \frac{2p(1-p)\sigma}{\mathcal{E}\sqrt{2\pi}} - \frac{2p^2\sigma}{\mathcal{E}\sqrt{2\pi}} e^{-\frac{\mathcal{E}^2}{2p^2\sigma^2}}\Bigg] \cdot \log\left(1 - p + pe^{-\frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right) \\ &+ \frac{\mathcal{E}^2}{2\sigma^2 p^2}\mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) - \frac{\mathcal{E}}{\sqrt{2\pi}\sigma p} e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}}.\end{aligned} \tag{6.4}$$

where $p \in [0,1]$ *in the* (6.3) *and* (6.4) *are free parameters.*

The new lower bounds (6.3) and (6.4) are depicted in Figure 6.5, where the free parameter have been chosen by numerical optimization.

Next we discuss the input distribution which satisfies average-power constraint only. We try two ways to find a new lower bound. In order to derive (6.3), we use Jensen's inequality in Lemma 3.4 to simplify the computation of entropy. From Figure 6.5 we improve the lower bound between -5 dB and 3 dB, but we have negative values below -5 dB.

In order to derive (6.4), we try to use bounds to simplify the computation of $h(Y)$. We divide the integration into several parts and find a close mathematical expressions for all parts separately, then integrate them in order to get a approximate entropy. Although we just get a little gain between -15 dB and -1 dB, we indeed improve our lower bound in the region where we used to have negative values.

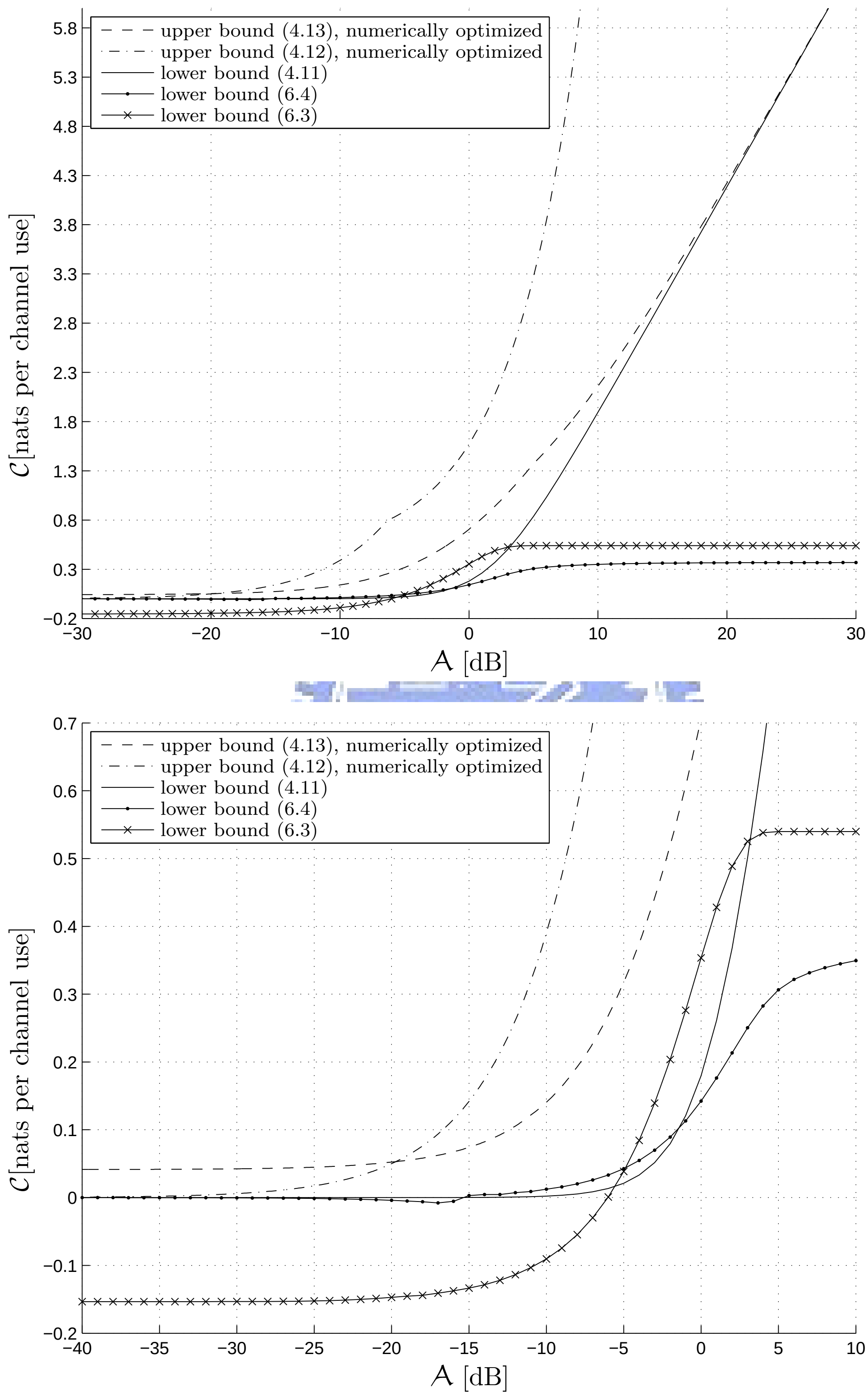


Figure 6.5: Bounds on the capacity of the free-space optical intensity channel just with peak-power constraint.

Chapter 7

Derivation of the Lower Bounds

The derivation of the lower bounds presented in Section 6 rely strongly on the Jensen's inequality in Lemma 3.4.

More specifically, the derivation of the lower bounds is based on the following ideas: recall the definition of entropy

$$h(X) = -\int f(x) \cdot \log f(x)\, \mathrm{d}x. \tag{7.1}$$

It is hard to integrate the $\log(\cdot)$ term, especially when $f(x)$ is a summation of several distributions. However, we observe that $-\log(\cdot)$ is a convex function. So we can use Jensen's inequality to simplify the problem.

7.1 Lower Bound of Theorem 6.1

7.1.1 Lower Bound (6.1)

First we use the same input distribution as in (5.7) to calculate the output distribution. Using that the input distribution is independent of the noise, we can get the output distribution from (2.3) as follows:

$$\begin{aligned} R_{\text{bin}}(y) &= \int_{-\infty}^{\infty} f_x(y-z) \cdot f_z(z)\ \mathrm{d}z \qquad (7.2)\\ &= \frac{1-\frac{\alpha}{s}}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{\frac{\alpha}{s}}{\sqrt{2\pi}\sigma} e^{-\frac{(y-s\mathrm{A})^2}{2\sigma^2}}. \qquad (7.3)\end{aligned}$$

We calculate the differential output entropy using Jensen's inequality:

$$\begin{aligned} h(y) &= -\int_{-\infty}^{\infty} R_{\text{bin}}(y) \cdot \log R_{\text{bin}}(y)\ \mathrm{d}y \\ &\geq -\log \underbrace{\left(\int_{-\infty}^{\infty} R_{\text{bin}}(y) \cdot R_{\text{bin}}(y)\ \mathrm{d}y \right)}_{l_1}. \qquad (7.4)\end{aligned}$$

We next expand l_1 as follows:

$$
\begin{aligned}
l_1 &= \int_{-\infty}^{\infty} R_{\text{bin}}(y) \cdot R_{\text{bin}}(y)\ \mathrm{d}y \\
&= \int_{-\infty}^{\infty} \left(\frac{\left(1-\frac{\alpha}{s}\right)^2}{2\pi\sigma^2} e^{-\frac{y^2}{\sigma^2}} + \frac{\left(\frac{\alpha}{s}\right)^2}{2\pi\sigma^2} e^{-\frac{(y-s\mathsf{A})^2}{\sigma^2}} \right. \\
&\qquad \left. + \frac{\frac{\alpha}{s}\left(1-\frac{\alpha}{s}\right)}{\pi\sigma^2} e^{-\frac{y^2+(y-s\mathsf{A})^2}{2\sigma^2}} \right) \mathrm{d}y \\
&= \frac{\left(1-\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\left(\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\frac{\alpha}{s}\left(1-\frac{\alpha}{s}\right)}{\sigma\sqrt{\pi}} e^{-\frac{s^2\mathsf{A}^2}{4\sigma}}.
\end{aligned} \tag{7.5}
$$

Finally we get

$$
\begin{aligned}
\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) &= h(Y) - h(Y|X) \\
&= h(Y) - h(Z) \\
&\geq \frac{\left(1-\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\left(\frac{\alpha}{s}\right)^2}{2\sigma\sqrt{\pi}} + \frac{\frac{\alpha}{s}\left(1-\frac{\alpha}{s}\right)}{\sigma\sqrt{\pi}} e^{-\frac{s^2\mathsf{A}^2}{4\sigma}} \\
&\quad - \frac{1}{2}\log\left(2\pi e\sigma^2\right)
\end{aligned} \tag{7.6}
$$

where s is chosen by computer simulation.

7.1.2 Lower Bound (6.2)

Next, we will calculate the ternary input distribution. We use the same way as (7.2) to get the output distribution.

$$
R_{\text{tri}}(y) = \frac{(1-p_1-p_2)}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{p_1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-s\mathsf{A})^2}{2\sigma^2}} + \frac{p_2}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mathsf{A})^2}{2\sigma^2}}. \tag{7.7}
$$

We calculate the output entropy using Jensen's inequality:

$$
\begin{aligned}
h(y) &= -\int_{-\infty}^{\infty} R_{\text{tri}}(y) \cdot \log R_{\text{tri}}(y)\, \mathrm{d}y \\
&\geq -\log \underbrace{\left(\int_{-\infty}^{\infty} R_{\text{tri}}(y) \cdot R_{\text{tri}}(y)\, \mathrm{d}y \right)}_{l_2}.
\end{aligned} \tag{7.8}
$$

We next expand l_2 as follows:

$$
\begin{aligned}
l_2 &= \int_{-\infty}^{\infty} R_{\text{tri}}(y) \cdot R_{\text{tri}}(y)\, \mathrm{d}y \\
&= \int_{-\infty}^{\infty} \left(\frac{(1-p_1-p_2)^2}{2\pi\sigma^2} e^{-\frac{y^2}{\sigma^2}} + \frac{{p_1}^2}{2\pi\sigma^2} e^{-\frac{(y-s\mathsf{A})^2}{\sigma^2}} + \frac{{p_2}^2}{2\pi\sigma^2} e^{-\frac{(y-\mathsf{A})^2}{\sigma^2}} \right. \\
&\quad + \frac{(1-p_1-p_2)p_1}{\pi\sigma^2} e^{-\frac{y^2+(y-s\mathsf{A})}{2\sigma^2}} + \frac{(1-p_1-p_2)p_2}{\pi\sigma^2} e^{-\frac{y^2+(y-\mathsf{A})}{2\sigma^2}}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \frac{p_1 p_2}{\pi\sigma^2} e^{-\frac{(y-s\mathsf{A})^2+(y-\mathsf{A})}{2\sigma^2}} \Bigg) \,\mathrm{d}y \\
&= \frac{(1-p_1-p_2)^2}{2\sigma\sqrt{\pi}} + \frac{{p_1}^2}{2\sigma\sqrt{\pi}} + \frac{{p_2}^2}{2\sigma\sqrt{\pi}} + \frac{p_1(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A})^2}{4\sigma^2}} \\
&\quad + \frac{p_2(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{\mathsf{A}^2}{4\sigma^2}} + \frac{p_1 p_2}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A}-\mathsf{A})^2}{4\sigma^2}},
\end{aligned} \tag{7.9}
$$

then we get

$$
\begin{aligned}
\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) &= h(Y) - h(Y|X) \\
&= h(Y) - h(Z) \\
&\geq -\log\Bigg(\frac{(1-p_1-p_2)^2}{2\sigma\sqrt{\pi}} + \frac{{p_1}^2}{2\sigma\sqrt{\pi}} + \frac{{p_2}^2}{2\sigma\sqrt{\pi}} \\
&\qquad + \frac{p_1(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A})^2}{4\sigma^2}} + \frac{p_2(1-p_1-p_2)}{\sigma\sqrt{\pi}} e^{-\frac{\mathsf{A}^2}{4\sigma^2}} \\
&\qquad + \frac{p_1 p_2}{\sigma\sqrt{\pi}} e^{-\frac{(s\mathsf{A}+\mathsf{A})^2}{4\sigma^2}} \Bigg) \\
&\quad - \frac{1}{2}\log\left(2\pi e\sigma^2\right).
\end{aligned} \tag{7.10}
$$

where s and p_2 are chosen by computer simulation.

7.2 Lower Bound of Theorem 6.2

7.2.1 Lower Bound (6.3)

First we use the input distribution (5.13) to calculate the output distribution. By (2.3) and we know that the input distribution is independent of noise, we can get the output distribution:

$$
R_{\mathrm{no}}(y) = \frac{1-p}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{p}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\frac{\mathcal{E}}{p})^2}{2\sigma^2}} \tag{7.11}
$$

We calculate the output entropy using Jensen's inequality:

$$
\begin{aligned}
h(y) &= -\int_{-\infty}^{\infty} R_{\mathrm{no}}(y)\cdot\log R_{\mathrm{no}}(y)\,\mathrm{d}y \\
&\geq -\log \underbrace{\left(\int_{-\infty}^{\infty} R_{\mathrm{no}}(y)\cdot R_{\mathrm{no}}(y)\,\mathrm{d}y \right)}_{l_3}.
\end{aligned} \tag{7.12}
$$

We expand l_3

$$
\begin{aligned}
l_3 &= \int_{-\infty}^{\infty} R_{\mathrm{no}}(y)\cdot R_{\mathrm{no}}(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} \Bigg(\frac{(1-p)^2}{2\pi\sigma^2} e^{-\frac{y^2}{\sigma^2}} + \frac{p^2}{2\pi\sigma^2} e^{-\frac{(y-\frac{\mathcal{E}}{p})^2}{\sigma^2}}
\end{aligned}
$$

$$
\left. + \frac{p(1-p)}{\pi\sigma^2} e^{-\frac{y^2+(y-\frac{\mathcal{E}}{p})^2}{2\sigma^2}} \right) \mathrm{d}y
$$

$$
= \frac{(1-p)^2}{2\sigma\sqrt{\pi}} + \frac{p^2}{2\sigma\sqrt{\pi}} + \frac{p(1-p)}{\sigma\sqrt{\pi}} e^{-\frac{\mathcal{E}^2}{4\sigma p^2}}. \tag{7.13}
$$

Then we get

$$
\begin{aligned}
\mathcal{C}(\mathcal{E}) &= h(Y) - h(Y|X) \\
&= h(Y) - h(Z) &(7.14)\\
&\geq -\log\left(\frac{(1-p)^2}{2\sigma\sqrt{\pi}} + \frac{p^2}{2\sigma\sqrt{\pi}} + \frac{p(1-p)}{\sigma\sqrt{\pi}} e^{-\frac{\mathcal{E}^2}{4\sigma p^2}}\right) - \frac{1}{2}\log\left(2\pi e\sigma^2\right). &(7.15)
\end{aligned}
$$

where p is chosen by computer simulation.

7.2.2 Lower Bound (6.4)

We use the output distribution (7.11) and calculate the differential entropy.

$$
\begin{aligned}
h(Y) = &-\int_{-\infty}^{\infty} \left(\frac{1-p}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{p}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\frac{\mathcal{E}}{p})^2}{2\sigma^2}}\right) \cdot \\
&\qquad \cdot \log\left(\frac{1-p}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{p}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\frac{\mathcal{E}}{p})^2}{2\sigma^2}}\right) \mathrm{d}y \\
= &\frac{1}{2}\log 2\pi\sigma^2 + \frac{1}{2\sigma^2}\mathsf{E}\left[Y^2\right] \\
&- \int_{-\infty}^{\infty} R_{\mathrm{no}}(y) \cdot \log\left(1 - p + p e^{\frac{\mathcal{E}y}{p\sigma^2} - \frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right) \mathrm{d}y
\end{aligned}
$$

and the second moment of y is as follows

$$
\mathsf{E}\left[Y^2\right] = (1-p)\sigma^2 + p\left(\sigma^2 + \left(\frac{\mathcal{E}}{p}\right)\right) = \sigma^2 + \frac{\mathcal{E}}{p}
$$

and hence

$$
\begin{aligned}
h(Y) = &\frac{1}{2}\log 2\pi\sigma^2 + \frac{1}{2} + \frac{\mathcal{E}}{2p\sigma^2} \\
&- \int_{-\infty}^{\infty} R_{\mathrm{no}}(y) \cdot \underbrace{\log\left(1 - p\left(1 - e^{\frac{\mathcal{E}y}{p\sigma^2} - \frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right)\right)}_{g(y)} \mathrm{d}y, &(7.16)
\end{aligned}
$$

We expand $g(y)$ and try to find some bounds in different region of y.

$$
\begin{aligned}
g(y) &\triangleq \log\left(1 - p + p e^{\frac{\mathcal{E}y}{p\sigma^2} - \frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right) \\
&= \underbrace{\log\left(p + (1-p) \cdot \exp\underbrace{\left(-\frac{\mathcal{E}y}{p\sigma^2} + \frac{\mathcal{E}^2}{2\sigma^2 p^2}\right)}_{k_1}\right)}_{k_2} + \frac{\mathcal{E}y}{p\sigma^2} - \frac{\mathcal{E}^2}{2\sigma^2 p^2}. &(7.17)
\end{aligned}
$$

In the first region we focus on k_1: when $k_1 > 0$, it will make the exponential term larger than one and then k_2 will be positive. When $k_1 < 0$, it will make the exponential term smaller than one and then k_2 will be negative. So we now derive which y will make k_1 smaller than zero.

$$\text{if } -\frac{\mathcal{E}y}{p\sigma^2}+\frac{\mathcal{E}^2}{2\sigma^2p^2}\leq 0 \implies y\geq\frac{p\sigma^2}{\mathcal{E}}\cdot\frac{\mathcal{E}^2}{2\sigma^2p^2}=\frac{\mathcal{E}}{2p}, \tag{7.18}$$

We approximate $g(y)$ by linear functions as follows:

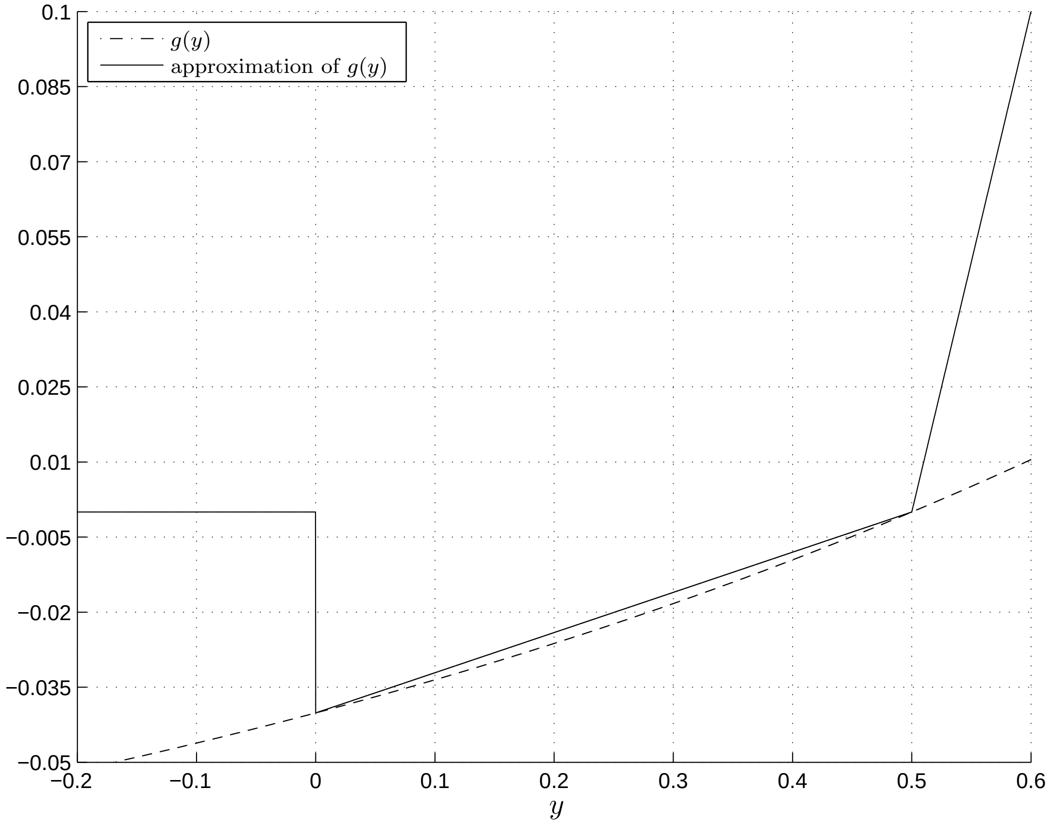


Figure 7.1: Approximation of $g(y)$ with $p = 0.1$ and $\mathcal{E} = 0.1$ dB

$$g(y)\leq\begin{cases}\log\left(1-p+pe^{-\frac{\mathcal{E}^2}{2\sigma^2p^2}}\right), & y<0,\\ -\log\left(1-p+pe^{-\frac{\mathcal{E}^2}{2\sigma^2p^2}}\right)\left(\frac{2py}{\mathcal{E}}-1\right), & 0\leq y\leq\frac{\mathcal{E}}{2p},\\ \frac{\mathcal{E}y}{p\sigma^2}-\frac{\mathcal{E}^2}{2\sigma^2p^2}, & y\geq\frac{\mathcal{E}}{2p},\end{cases}$$

We define the shorthand $\kappa\triangleq\log\left(1-p+pe^{-\frac{\mathcal{E}^2}{2\sigma^2p^2}}\right)$. Then we get for the region $\left[\frac{\mathcal{E}}{2p},\infty\right)$

$$\int_{\frac{\mathcal{E}}{2p}}^{\infty}R_{\text{no}}(y)\cdot\left(\frac{\mathcal{E}y}{p\sigma^2}-\frac{\mathcal{E}^2}{2\sigma^2p^2}\right)\,\mathrm{d}y$$

$$
\begin{aligned}
&= \int_{\frac{\mathcal{E}}{2p}}^{\infty} \left(\frac{1-p}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} + \frac{p}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(y-\frac{\mathcal{E}}{p}\right)^2}{2\sigma^2}} \right) \cdot \left(\frac{\mathcal{E}y}{p\sigma^2} - \frac{\mathcal{E}^2}{2\sigma^2 p^2} \right) \mathrm{d}y \\
&= \frac{(1-p)\mathcal{E}}{\sqrt{2\pi\sigma^2}p\sigma^2} \left(-\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \Big|_{\frac{\mathcal{E}}{2p}}^{\infty} \right) - \frac{\mathcal{E}^2(1-p)}{2\sigma^2 p^2} \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) \\
&\quad + \int_{-\frac{\mathcal{E}}{2p}}^{\infty} \frac{p}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} \left(\frac{\mathcal{E}t}{p\sigma^2} + \frac{\mathcal{E}^2}{p^2\sigma^2} - \frac{\mathcal{E}^2}{2p^2\sigma^2} \right) \mathrm{d}t \\
&= \frac{(1-p)\mathcal{E}}{\sqrt{2\pi}\sigma p} e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} - \frac{(1-p)\mathcal{E}^2}{2\sigma^2 p^2} \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) + \frac{p\mathcal{E}}{\sqrt{2\pi}\sigma p\sigma^2} \left(-\sigma^2 e^{-\frac{t^2}{2\sigma^2}} \Big|_{-\frac{\mathcal{E}}{2p}}^{\infty} \right) \\
&\quad + \frac{\mathcal{E}^2 p}{2\sigma^2 p^2} \left(1 - \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) \right) \\
&= \frac{\mathcal{E}}{\sqrt{2\pi}\sigma p} e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} + \frac{\mathcal{E}^2}{2\sigma^2 p^2} \left(p - \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) \right).
\end{aligned} \tag{7.19}
$$

For $y \in \left[0, \frac{\mathcal{E}}{2p}\right]$, we get

$$
\begin{aligned}
&\int_0^{\frac{\mathcal{E}}{2p}} R_{\text{no}}(y) \cdot \left(\log\left(1 - p + pe^{-\frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right) - \frac{2p\kappa}{\mathcal{E}} y \right) \mathrm{d}y \\
&= \kappa \cdot \int_0^{\frac{\mathcal{E}}{2p}} R_{\text{no}}(y)\,\mathrm{d}y - \frac{2p\kappa}{\mathcal{E}} \int_0^{\frac{\mathcal{E}}{2p}} y \cdot R_{\text{no}}(y)\,\mathrm{d}y \\
&= \kappa(1-p)\left(\frac{1}{2} - \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) \right) + \kappa p \left(\mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) - \mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) \right) \\
&\quad - \frac{2p\kappa}{\mathcal{E}} \frac{1-p}{\sqrt{2\pi}\sigma} \left(-\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \Big|_0^{\frac{\mathcal{E}}{2p}} \right) - \frac{2p\kappa}{\mathcal{E}} \int_{\frac{\mathcal{E}}{p}}^{-\frac{\mathcal{E}}{2p}} \left(t + \frac{\mathcal{E}}{p} \right) e^{-\frac{t^2}{2\sigma^2}} \frac{p}{\sqrt{2\pi}\sigma} \mathrm{d}t \\
&= \frac{\kappa(1-p)}{2} - \kappa \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) + 2\kappa p \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) - \kappa p \mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) \\
&\quad + \frac{2p\kappa(1-p)\sigma}{\mathcal{E}\sqrt{2\pi}} \left(e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} - 1 \right) - \frac{2p^2\kappa}{\mathcal{E}\sqrt{2\pi}\sigma} \left(-\sigma^2 e^{-\frac{t^2}{2\sigma^2}} \Big|_{-\frac{\mathcal{E}}{p}}^{-\frac{\mathcal{E}}{2p}} \right) \\
&\quad - \frac{2p^2\kappa}{\mathcal{E}} \frac{\mathcal{E}}{p} \left(\mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) - \mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) \right) \\
&= \frac{\kappa(1-p)}{2} - \kappa \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) + \kappa p \mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) + \frac{2p\kappa\sigma}{\mathcal{E}\sqrt{2\pi}} e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} \\
&\quad - \frac{2p\kappa\sigma(1-p)}{\mathcal{E}\sqrt{2\pi}} - \frac{2p^2\kappa\sigma}{\sqrt{2\pi}\mathcal{E}} e^{-\frac{\mathcal{E}^2}{2p^2\sigma^2}}.
\end{aligned} \tag{7.20}
$$

For $y \in \left(-\infty, \frac{\mathcal{E}}{2p}\right]$, we get

$$
\begin{aligned}
&\int_{-\infty}^{0} \kappa \cdot R_{\text{no}}(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{0} \frac{\kappa(1-p)}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} + \frac{\kappa p}{\sqrt{2\pi}\sigma} e^{-\frac{\left(y-\frac{\mathcal{E}}{p}\right)^2}{2\sigma^2}} \mathrm{d}y
\end{aligned}
$$

$$= \frac{\kappa(1-p)}{2} + \kappa p \mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) \tag{7.21}$$

Finally we plug (7.19), (7.20),and (7.21) into (7.16) and combine this with (7.14) we get the following bound:

$$\begin{aligned}\mathcal{C}(\mathcal{E}) \geq -\Bigg[1 - p - \mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) + 2p\mathcal{Q}\left(\frac{\mathcal{E}}{p\sigma}\right) + \frac{2p\sigma}{\mathcal{E}\sqrt{2\pi}}e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}} \\ - \frac{2p(1-p)\sigma}{\mathcal{E}\sqrt{2\pi}} - \frac{2p^2\sigma}{\mathcal{E}\sqrt{2\pi}}e^{-\frac{\mathcal{E}^2}{2p^2\sigma^2}}\Bigg] \cdot \log\left(1 - p + pe^{-\frac{\mathcal{E}^2}{2\sigma^2 p^2}}\right) \\ + \frac{\mathcal{E}^2}{2\sigma^2 p^2}\mathcal{Q}\left(\frac{\mathcal{E}}{2p\sigma}\right) - \frac{\mathcal{E}}{\sqrt{2\pi}\sigma p}e^{-\frac{\mathcal{E}^2}{8p^2\sigma^2}}.\end{aligned} \tag{7.22}$$

where p is chosen by computer simulation.

Chapter 8

Conclusion

In this thesis, we try two ways to obtain bounds on the channel capacity of the free-space optical intensity channel. We first numerically evaluate a lower bound to capacity for small to medium SNR. Comparing this bound with existing analytical upper and lower bounds we realize that these known analytical upper bounds are relatively tight, while the analytical lower bounds seem looser. Hence, we try to improve on the analytical lower bounds.

To that goal we try various approaches to simplify the task of computing the differential entropy of the channel output under the assumption of a binary input distribution. One approach is based on Jensen's inequality. This turns out to be less powerful, particularly, because we are interested in the low to medium SNR regime. Another way is to find linear functions that closely approximate from above some difficult logarithmic term in the entropy expression. This approach proved successful for a certain SNR regime.

A possible extension for future work is to use an input distribution having a slightly larger alphabet size. It also would be interesting to find new approaches that could replace the weak Jensen's inequality.

Appendix A

Jensen-Shannon Divergence

In this appendix, we will introduce an idea about improving our lower bound. From Chapter 6 and 7, we know that Jensen's inequality can help us to pull up medium SNR region's lower bound, but it also make the lower bound becoming negative in low SNR region. Here we provide an idea which is potential to get a new lower bound.

After we doing numerical simulation in Chapter 5, we know that the discrete input distribution performs well in low SNR region. So we try to analytical calculate this entropy, but combinative distribution is not easy to calculate. From [4] and [5], the authors provide a new way to calculate entropy of mixture of probability distributions. The way is Jensen-Shannon divergence and the definition as follows:

$$\sum_v \alpha_v \mathrm{D}(P_v \| Q) - \mathrm{D}\left(\sum_v \alpha_v P_v \| Q\right) \tag{A.1}$$

where $\mathrm{D}(\cdot\|\cdot)$ is Kullback-Leibler divergence. From (A.1) we can modify some term and then get the following result:

$$\sum_v \alpha_v \mathrm{D}(P_v \| \overline{P}) = \sum_v \alpha_v \mathrm{D}(P_v \| Q) - \mathrm{D}\left(\sum_v \alpha_v P_v \| Q\right) \tag{A.2}$$

$$= H\left(\sum_v \alpha_v P_v\right) - \sum_v \alpha_v H(P_v) \tag{A.3}$$

where $\overline{P} = \sum_v \alpha_v P_v$. In our situation, $\overline{P}$ represents the output distribution, $\alpha_v P_v$ represents the distribution of particular input signal which was disturbed by noise and Q is a free parameter. So we can combine the definition of channel capacity and rewrite these equation as follows:

$$\begin{aligned} \mathcal{C} &\geq I(X;Y) \\ &= H(Y) - H(Y|X) \\ &= H\left(\sum_v \alpha_v P_v\right) - \sum_v \alpha_v H(P_v) \end{aligned} \tag{A.4}$$

So we can combine (A.1) and (A.4), and we can get the capacity.

Unfortunately, the way from [4] and [5] only works for discrete system, but our noise is continue Gaussian noise. It can not help us in this situation, but if we can modify this method, we may get a nice lower bound.

Bibliography

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[2] A. Lapidoth, S. M. Moser, and M. A. Wigger, "On the capacity of free-space optical intensity channels," 2008, submitted.

[3] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[4] B. Fuglede and F. Topsøe, "Jensen-Shannon Divergence and Hilbert space embedding," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Chicago, IL, USA, June 27 – July 2, 2004, p. 31.

[5] F. Topsøe, "Jensen-Shannon Divergence and norm-based mesures of Discriminaion and Variation," inpreparation. [Online]. Available: http://www.math.ku.dk/ topsoe/ISIT2004JSD.pdf