國 立 交 通 大 學

電信工程研究所

碩 士 論 文

流體介質中的粒子通訊:相加性反高斯雜訊的通道容量界線

Molecular Communication in Fluid Media: Bounds on the Capacity of the Additive Inverse Gaussian Noise Channel

研 究 生：張惠婷

指導教授：莫詩台方　教授

中 華 民 國 一零一 年 三 月

流體介質中的粒子通訊:相加性反高斯雜訊的通道容量界線

Molecular Communication in Fluid Media: Bounds on the Capacity of the Additive Inverse Gaussian Noise Channel

研 究 生：張惠婷　　Student：Hui-Ting Chang

指導教授：莫詩台方 教授　　Advisor：Prof. 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 of Science

in

Communication Enginering

March 2013

Hsinchu, Taiwan, Republic of China

中華民國一零一年三月

Information Theory Laboratory
Dept. of Electrical and Computer Engineering
National Chiao Tung University

Master Thesis

Molecular Communication in Fluid Media: Bounds on the Capacity of the Additive Inverse Gaussian Noise Channel

Chang Hui-Ting

Advisor:	Prof. Dr. Stefan M. Moser National Chiao Tung University, Taiwan
Graduation Committee:	Prof. Dr. Yu Ted Su National Chiao Tung University, Taiwan Prof. Dr. Scott Ch. Huang National Tsing Hua University, Taiwan

流體介質中的粒子通訊：
相加性反高斯雜訊的通道容量界線

研究生：張惠婷　　　　　　　　　　指導教授：莫詩台方 教授

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，我們研究一個相當新且近代的通道模型，該通道利用常速流體當中的化學粒子交換來做為溝通的訊息。這些粒子由傳送端出發至接收端的路徑，我們將其視為一維空間來做模擬。很典型的通訊應用像是我們將奈米級的儀器置入血管中，以完成傳遞訊息的任務。在這個情況下，我們不再依賴電磁波傳遞訊息，而是將訊息放在釋放粒子的時間點上。一旦粒子被傳送端釋放進入流體中時，會在介質中行布朗運動，這將會對粒子到達接收端的時間產生不確定性，這樣的不確定性就是我們的雜訊。我們用反高斯分布來描述這樣的雜訊。此篇研究將重點放在相加性雜訊通道以描述基本的通道容量趨勢。

我們深入研究此模型，並分析出新的通道容量上界與下界。 這些界線是漸進緊的，也就是說，如果平均延遲的限制可放寬至無限大，或是介質流體流速趨近無限大，則相對應的漸進通道容量可被精確的推導出來。

Molecular Communication in Fluid Media: Bounds on the Capacity of the Additive Inverse Gaussian Noise Channel

Student: Chang Hui-Ting Advisor: Prof. Stefan M. Moser

Institute of Communication Engineering
National Chiao Tung University

Abstract

In this thesis a very recent and new channel model is investigated that describes communication based on the exchange of chemical molecules in a liquid medium with constant drift. They travel from the transmitter to the receiver at two ends of a one-dimensional axis. A typical application of such communication are nano-devices inside a blood vessel communicating with each other. In this case, we no longer transmit our signal via electromegnetic waves, but we put our information on the emission time of the molecules. Once a molecule is emitted in the fluid medium, it will be affected by Brownian motion, which causes uncertainty of the molecule's arrival time at the receiver. We characterize this noise with an inverse Gaussian distribution. Here we focus solely on an additive noise channel to describe the fundamental channel capacity behavior.

This new model is investigated and new analytical upper and lower bounds on the capacity are presented. The bounds are asymptotically tight, i.e., if the average-delay constraint is loosened to infinity or if the drift velocity of the liquid medium tends to infinity, the corresponding asymptotic capacities are derived precisely.

Acknowledgments

Never in my life did I realize that learning could be so enjoyable, until I started my first lesson with Prof. Stefan M. Moser. in 2008. I was so much encouraged to ask questions and discuss our ideas during his wonderful classes. Great thanks to his kind acceptance for me as a member of Information Theory Laboratory (IT Lab), NCTU. Under his patient guidance, I had my happiest research period. He always pays fully attention on my meeting presentation and gave very useful suggestions for improvement.

He always shows warm welcome whenever I knock his office door for questions or even advice of life. Moreover, I cherish the moments when we sat down together at his place with his wife Yin-Tia and son Matthias. There, we shared about life, educations, coffee, trips and all sorts of thing. These are the moments which lighten my thoughts deeply. I would never live my life as faithfully and clearly like now without the help from them. To me, he is my Prof. Albus Dumbledore. Thank you for giving me so many wonderful lessons.

Big thanks to Yuan-Chu. She is a strong support and a very nice accompany during my time in NCTU. Without her, this thesis wouldn't be complete.

Thanks to Gu-Rong Lin's advice during my master study. He is a very helpful former lab member who has a very kind heart. Thanks to Hsuan-Ying Lin, who is always patient for my questions and has helped me a lot.

At the end but not the least, thanks to the financial support of my parents. Without them, it would be more difficult to finish my master study.

Flight back to Taiwan, 18 March 2013

Chang Hui-Ting

Contents

List of Figures

Chapter 1

Introduction

1.1 General Molecular Communication Channel Model

Usually, we transmit our signal with electromagnetic waves in the air or in wires. Recently, people are more and more interested in communication within nanoscale networks. But when we want to transmit our signal via these tiny devices, we face some problems that the antenna size of them are restricted and the energy that could be stored in them is very little. Therefore, we solve these problems with providing a different type of communication instead. This thesis focuses on a channel which operates in a fluid medium with a constant drift velocity. The transmitter is a point source with many molecules to be emitted. The receiver waits on the other side for the molecules' arrival. The information is encoded in the emission time of the molecules, X, which takes value in a finite set. One application example is blood vessel, which has a blood drift. The nanoscale device could be any medical inspection device that is inserted in our body.

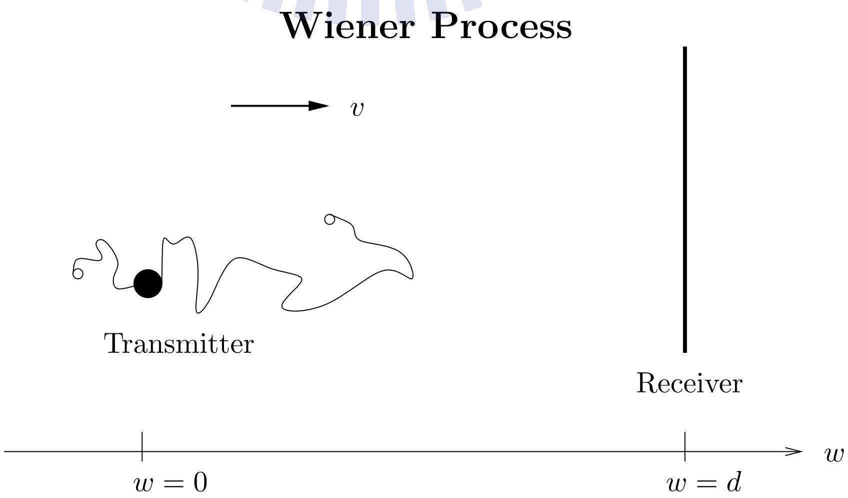


Figure 1.1: Wiener process of molecular communication channel.

Once the nanoscale molecules are emitted in the fluid medium, they are effected by Brownian motion which causes uncertainty of the arrival time at the receiver. We describe this type of channel noise with an inverse Gaussian distribution. Now, consider a channel as shown in Figure 1.1 where w is the position parameter, d is the receiver's position on w axis and v is the drift velocity, $v > 0$. The transmitter is placed at the origin of w axis. It emits a molecule into a fluid with positive drift velocity, v. The information is put on the releasing time. In order to know this information, the receiver ideally subtracts the average traveling time, $\frac{d}{v}$, from the arrival time. Note that once a molecule arrives at the receiver, it is absorbed and never returns to the fluid. Moreover, every molecule is independent of each other.

This molecular communication channel model was proposed by Srinivas, Adve and Eckford [1].

1.2 Mathematical Model

Let $W(x)$ be the position of a molecule at time x that travels via a Brownian motion medium. Let $0 \leq x_1 < x_2 < \cdots < x_k$ be a sequence of time indices ordered from small to large. Then, $W(x)$ is a Wiener process if the position increment $R_i = W(x_{i-1}) - W(x_i)$ are independent random variables with

$$R_i \sim \mathcal{N}\big(v(x_i - x_{i-1}), \sigma^2(x_i - x_{i-1})\big) \tag{1.1}$$

where $\sigma^2 = \frac{D}{2}$ with D being the diffusion coefficient, which depends on the temperature and the stickiness of the fluid and the size of the particles. Assuming the molecule is released at time $x = 0$ at position $W(0) = 0$, the position at time $\tilde{x}$ is $W(\tilde{x}) \sim \mathcal{N}\big(v\tilde{x}, \sigma^2\tilde{x}\big)$. The probability density function (PDF) of W is given by:

$$f_W(w; \tilde{x}) = \frac{1}{\sqrt{2\pi\sigma^2\tilde{x}}} \exp\left(-\frac{(w - v\tilde{x})^2}{2\sigma^2\tilde{x}}\right). \tag{1.2}$$

In our communication system, instead of looking at the position of the molecule at a certain time, we turn our focus on its arriving time at the receiver for a fixed distance d.

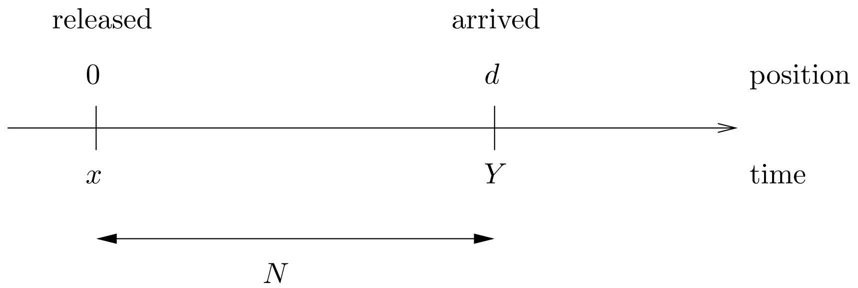


Figure 1.2: The relation between the molecule's time and position.

We release the molecule at time x from the origin, $W(x) = 0$ and $x \geq 0$. Assuming that after traveling for a random time N, the molecule arrives at the receiver for the first time at time Y,

$$Y = x + N. \tag{1.3}$$

Hence, our channel model is characterized by an additive noise in the form of the random propagation time N. This is the only uncertainty we have in the system. When we assume a positive drift velocity $v > 0$, the distribution of the traveling time N is well known to be an *inverse Gaussian* (IG) distribution. As a result, we call this channel the *additive inverse Gaussian noise (AIGN) channel*. Since the PDF of N is

$$f_N(n) = \begin{cases} \sqrt{\frac{\lambda}{2\pi n^3}} \exp\left(-\frac{\lambda(n-\mu)^2}{2\mu^2 n}\right) & n > 0, \\ 0 & n \leq 0, \end{cases} \tag{1.4}$$

we get the conditional probability density of output Y given the channel input $X = x$ as

$$f_{Y|X}(y|x) = \begin{cases} \sqrt{\frac{\lambda}{2\pi(y-x)^3}} \exp\left(-\frac{\lambda(y-x-\mu)^2}{2\mu^2(y-x)}\right) & y > x, \\ 0 & y \leq x. \end{cases} \tag{1.5}$$

There are two important parameters for the inverse Gaussian distribution: the average traveling time

$$\mu = \frac{d}{v} = \frac{\text{distance between transmitter and receiver}}{\text{drift velocity}}, \tag{1.6}$$

and a parameter

$$\lambda = \frac{d^2}{\sigma^2} \tag{1.7}$$

that describes the impact of the noise. Usually we write $N \sim \text{IG}(\mu, \lambda)$. By calculation, we get

$$\mathsf{E}[N] = \mu \quad = \frac{d}{v}, \tag{1.8}$$

$$\text{Var}(N) = \frac{\mu^3}{\lambda} = \frac{d\sigma^2}{v^3}. \tag{1.9}$$

If the drift velocity v increases, the variance decreases, in other words, the distribution is more centered. If the drift velocity is slowed down, we will have a more spread-out noise distribution. Without loss of generality, we normalize the propagation distance to $d = 1$.

For practical reasons, we constrain the transmitter to have an average delay m on the emission of a molecule, i.e., the input X is subject to the constraint:

$$\mathsf{E}[X] \leq m. \tag{1.10}$$

Note that a peak constraint would also be of large practical interest, but for simplicity we focus on average constraint at the moment.

1.3 Capacity

Since we introduced a new type of channel, the AIGN channel, we are interested in how much information it can transmit. In [2], Shannon showed that for memoryless channels with continuous input and output alphabets and an corresponding conditional PDF describing the channel, and under an input constraint $\mathsf{E}[X] \leq m$, the channel capacity is given by

$$\triangleq \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} I(X;Y) \tag{1.11}$$

where the supremum is taken over all input probability distributions $f(\cdot)$ on X that satisfy the mean constraint $\mathsf{E}[X] \leq m$. By $I(X;Y)$ we denote the mutual information between X and Y. For the AIGN channel, we have

$$\begin{aligned}
\sup_{f_X(x)\colon \mathsf{E}[X]\leq m} I(X;Y) &= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} \Big\{h(Y) - h(Y|X)\Big\} && (1.12)\\
&= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} \Big\{h(Y) - h(X+N|X)\Big\} && (1.13)\\
&= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} \Big\{h(Y) - h(N|X)\Big\} && (1.14)\\
&= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} h(Y) - h(N) && (1.15)\\
&= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} h(Y) - h_{\mathrm{IG}(\mu,\lambda)}, && (1.16)
\end{aligned}$$

where (1.15) holds because N and X are independent. The mean constraint (1.10) of the input signal translates to an average constraint for Y:

$$\begin{aligned}
\mathsf{E}[Y] &= \mathsf{E}[X+N] && (1.17)\\
&= \mathsf{E}[X] + \mathsf{E}[N] && (1.18)\\
&= \mathsf{E}[X] + \mu && (1.19)\\
&\leq m + \mu. && (1.20)
\end{aligned}$$

Chapter 2

Mathematical Preliminaries

In this chapter, we will introduce some mathematical properties of the inverse Gaussian random variable and other useful lemmas for future use in this thesis.

2.1 Properties of the Inverse Gaussian Distribution

In [1], the differential entropy of an inverse Gaussian random variable was given in a complicated form that is unwielding for analytical analysis. So we try to modify the original expression and derive a cleaner form for mathematical derivation.

Proposition 2.1 (Differential Entropy of the Inverse Gaussian Distribution)**.**

$$h_{\mathrm{IG}(\mu,\lambda)} = \log\left(2K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)\mu\right) + \frac{3}{2}\frac{\frac{\partial}{\partial\gamma}K_\gamma\left(\frac{\lambda}{\mu}\right)\Big|_{\gamma=-\frac{1}{2}}}{K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)} + \frac{\lambda}{2\mu}\frac{K_{\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)+K_{-\frac{3}{2}}\left(\frac{\lambda}{\mu}\right)}{K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)} \tag{2.1}$$

$$= \frac{1}{2}\log\frac{2\pi\mu^3}{\lambda} + \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2} \tag{2.2}$$

$$= \frac{1}{2}\log\frac{2\pi\sigma^2 d}{v^3} + \frac{3}{2}\exp\left(\frac{2dv}{\sigma^2}\right)\mathrm{Ei}\left(-\frac{2dv}{\sigma^2}\right) + \frac{1}{2} \tag{2.3}$$

where $K_\gamma(\cdot)$ is the order-γ modified Bessel function of the second kind, and $\mathrm{Ei}(\cdot)$ *is the exponential integral function defined as*

$$\mathrm{Ei}(-x) \triangleq -\int_x^\infty \frac{e^{-t}}{t}\,\mathrm{d}t = \int_{-\infty}^{-x}\frac{e^t}{t}\,\mathrm{d}t, \quad x>0. \tag{2.4}$$

In `MATLAB`, the exponential integral function is implement as `expint(x)`$= -\mathrm{Ei}(-x)$.

Here (2.1) is taken from [1]. The concise expression (2.2) and (2.3) are derived below.

Proof. We divide the right hand side of (2.1) into three parts. From [3, (8.469.3)] the first part can be simplified as follows:

$$\log\left(2K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)\mu\right) = \log\left(2\left(\left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)\right)\mu\right) \tag{2.5}$$

$$= \log\left(\left(\frac{2\pi\mu^3}{\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)\right) \tag{2.6}$$

$$= \frac{1}{2}\log\frac{2\pi\mu^3}{\lambda} - \frac{\lambda}{\mu} \tag{2.7}$$

From formula [3, (8.486(1).21)], we get:

$$\left.\frac{\partial K_\gamma\left(\frac{\lambda}{\mu}\right)}{\partial\gamma}\right|_{\gamma=-\frac{1}{2}} = \left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(\frac{\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) \tag{2.8}$$

such that the second term of (2.1) can be written as

$$\frac{3}{2}\frac{\left.\frac{\partial}{\partial\gamma}K_\gamma\left(\frac{\lambda}{\mu}\right)\right|_{\gamma=-\frac{1}{2}}}{K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)} = \frac{3}{2}\frac{\left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(\frac{\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right)}{\left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)} \tag{2.9}$$

$$= \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right). \tag{2.10}$$

From formula [3, (8.486.16)]

$$K_{-\nu}(z) = K_\nu(z) \tag{2.11}$$

and formula [3, (8.468)]

$$K_{n+\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2z}}e^{-z}\sum_{k=0}^{n}\frac{(n+k)!}{k!(n-k)!(2z)^k} \tag{2.12}$$

we get the result:

$$K_{-\frac{3}{2}}\left(\frac{\lambda}{\mu}\right) = K_{\frac{3}{2}}\left(\frac{\lambda}{\mu}\right) \tag{2.13}$$

$$= K_{1+\frac{1}{2}}\left(\frac{\lambda}{\mu}\right) \tag{2.14}$$

$$= \left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)\left(\frac{1!}{0!1!\left(\frac{2\lambda}{\mu}\right)^0} + \frac{2!}{1!0!\left(\frac{2\lambda}{\mu}\right)^1}\right) \tag{2.15}$$

$$= \left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)\left(1+\frac{\mu}{\lambda}\right). \tag{2.16}$$

Therefore, the third term of (2.1) will be

$$\frac{\lambda}{2\mu}\frac{K_{\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)+K_{-\frac{3}{2}}\left(\frac{\lambda}{\mu}\right)}{K_{-\frac{1}{2}}\left(\frac{\lambda}{\mu}\right)} = \frac{\lambda}{2\mu}\left(1+\frac{\left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)\left(1+\frac{\mu}{\lambda}\right)}{\left(\frac{\pi\mu}{2\lambda}\right)^{\frac{1}{2}}\exp\left(-\frac{\lambda}{\mu}\right)}\right) \tag{2.17}$$

$$= \frac{\lambda}{2\mu}\left(2+\frac{\mu}{\lambda}\right) \tag{2.18}$$

$$= \frac{\lambda}{\mu}+\frac{1}{2}. \tag{2.19}$$

As a result,

$$h_{\mathrm{IG}(\mu,\lambda)} = \frac{1}{2}\log\frac{2\pi\mu^3}{\lambda} - \frac{\lambda}{\mu} + \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{\lambda}{\mu} + \frac{1}{2} \tag{2.20}$$

$$= \frac{1}{2}\log\frac{2\pi\mu^3}{\lambda} + \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2} \tag{2.21}$$

□

Next, when we want to make an IG random variable add with another IG random variable and end up also in IG distributed, there is a specific way to reach it. Only certain type of IGs will add up to be IG distributed.

Proposition 2.2 (Additivity of the IG distribution)**.** *Let M be a linear combination of random variables M_i:*

$$M = \sum_{i=0}^{l} c_i M_i, \quad c_i > 0, \tag{2.22}$$

where

$$M_i \sim \mathrm{IG}(\mu_i, \lambda_i), \quad i = 1, \ldots, l. \tag{2.23}$$

Here we assume that M_i are not necessarily independent, but summed up under the constraint that

$$\frac{\lambda_i}{c_i\mu_i^2} = \kappa, \quad \text{for all } i. \tag{2.24}$$

Then

$$M \sim \mathrm{IG}\left(\sum_i c_i\mu_i, \kappa\left(\sum_i c_i\mu_i\right)^2\right) \tag{2.25}$$

Proof. The proof can be found in [4, Sec. 2.4, p. 13]. □

Remark 2.3. *If we simply add two inverse Gaussian random variable, as long as they are in the same fluid, which means they have the same v and σ^2, the result is still inverse Gaussian.*

Consider a Wiener process $X(t)$ beginning with $X(0) = x_0$ with positive drift v and variance σ^2. Choose a and b so that $x_0 < a < b$ and consider the first passage *time T_1 from x_0 to a and T_2 from a to b. Then T_1 and T_2 are independent inverse Gaussian variables with parameters*

$$\mu_1 = \frac{a - x_0}{v}, \quad \lambda_1 = \frac{(a - x_0)^2}{\sigma^2} \tag{2.26}$$

and

$$\mu_2 = \frac{b - a}{v}, \quad \lambda_2 = \frac{(b - a)^2}{\sigma^2}. \tag{2.27}$$

Now consider $T_3 = T_1 + T_2$, therefore, $c_1 = c_2 = 1$ and

$$\frac{\lambda_i}{\mu_i} = \frac{v^2}{\sigma^2} = \text{constant}, \tag{2.28}$$

T_3 is also an inverse Gaussian variable. That is

$$T_3 \sim \mathrm{IG}\left(\mu_1 + \mu_2, \frac{v^2(\mu_1 + \mu_2)^2}{\sigma^2}\right). \tag{2.29}$$

Since $\mu_1 + \mu_2 = \frac{b - x_0}{v}$,

$$T_3 \sim \mathrm{IG}\left(\frac{b - x_0}{v}, \frac{(b - x_0)^2}{\sigma^2}\right). \tag{2.30}$$

The last observation also follows directly from the realization that T_3 is the first passage time from x_0 to b [4].

Proposition 2.4 (Scaling)**.** *If $N \sim \mathrm{IG}(\mu, \lambda)$, then for any $k > 0$*

$$kN \sim \mathrm{IG}(k\mu, k\lambda). \tag{2.31}$$

Proof. The proof can be found in [4, Sec. 2.4, p. 13]. □

Proposition 2.5. *If N is a random variable distributed as $\mathrm{IG}(\mu,\, \lambda)$. Then*

$$\mathsf{E}[N] = \mu; \tag{2.32}$$

$$\mathsf{E}\left[\frac{1}{N}\right] = \frac{1}{\mu} + \frac{1}{\lambda}; \tag{2.33}$$

$$\mathsf{E}\left[N^2\right] = \mu^2 + \frac{\mu^3}{\lambda}; \tag{2.34}$$

$$\mathsf{E}\left[\frac{1}{N^2}\right] = \frac{1}{\mu^2} + \frac{3}{\lambda^2} + \frac{3}{\mu\lambda}; \tag{2.35}$$

$$\mathrm{Var}\,(N) = \frac{\mu^3}{\lambda}; \tag{2.36}$$

$$\mathrm{Var}\left(\frac{1}{N}\right) = \frac{1}{\mu\lambda} + \frac{2}{\lambda^2}; \tag{2.37}$$

$$\mathsf{E}[N^\nu] = \sqrt{\frac{2\lambda}{\pi}} e^{\frac{\lambda}{\mu}} \mu^{\nu - \frac{1}{2}} K_{\nu - \frac{1}{2}}\left(\frac{\lambda}{\mu}\right), \qquad \nu \in \mathbb{R}. \tag{2.38}$$

Remark 2.6. *From* (2.11)*, we can also write*

$$\mathsf{E}\left[N^{-\nu}\right] = \sqrt{\frac{2\lambda}{\pi}} e^{\frac{\lambda}{\mu}} \mu^{-\nu-\frac{1}{2}} K_{\nu+\frac{1}{2}}\left(\frac{\lambda}{\mu}\right). \tag{2.39}$$

Proof. The proofs are based on [4, (2.6)], [5, Proposition 2.15], [4, (8.36)] and [3, 3.471 9.]. □

Proposition 2.7. *If* $N \sim \mathrm{IG}(\mu, \lambda)$*, then*

$$\mathsf{E}[\log N] = e^{\frac{2\lambda}{\mu}}\, \mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \log \mu; \tag{2.40}$$

$$\mathsf{E}\left[\frac{N}{\mu} + \frac{\mu}{N}\right] = 2 + \frac{\mu}{\lambda}. \tag{2.41}$$

Proof. A proof is shown in [6]. □

Proposition 2.8. *If N_i are IID $\sim$ IG(μ, λ), then the sample mean from that distribution will be*

$$\frac{1}{n}\sum_{i=1}^{n} N_i = \bar{N} \sim \mathrm{IG}(\mu, n\lambda), \quad \text{for } i = 1, \ldots, n. \tag{2.42}$$

Proof. A proof can be found in [4, Sec. 5.1, p. 56]. □

Lemma 2.9. *Under the three constraints*

$$\mathsf{E}[\log X] = \alpha_1, \tag{2.43}$$

$$\mathsf{E}[X] = \alpha_2, \tag{2.44}$$

$$\mathsf{E}\left[X^{-1}\right] = \alpha_3, \tag{2.45}$$

where α_1, α_2 and α_3 are some fixed values, the maximum entropy distribution is the inverse Gaussian distribution.

Proof. From [7, Chap. 12] we know that if we have the three constraints above, the optimal distribution to maximize the entropy will have the form

$$f(x) = e^{\lambda_0 + \lambda_1 \log x + \lambda_2 x + \frac{\lambda_3}{x}} \tag{2.46}$$

$$= x^{\lambda_1} e^{\lambda_0 + \lambda_2 x + \frac{\lambda_3}{x}}, \tag{2.47}$$

which is exactly the form of the inverse Gaussian. □

2.2 Power Inverse Gaussian and Its Properties

The power inverse Gaussian (PIG) distribution parameterized by an arbitrarily fixed real number $\eta \neq 0$ has the PDF given by

$$R(y) = \sqrt{\frac{\alpha}{2\pi\beta^3}} \left(\frac{y}{\beta}\right)^{-\left(1+\frac{\eta}{2}\right)} \exp\left(-\frac{\alpha}{2\eta^2\beta}\left(\left(\frac{y}{\beta}\right)^{\frac{\eta}{2}} - \left(\frac{y}{\beta}\right)^{-\frac{\eta}{2}}\right)^2\right), \quad (2.48)$$

where

$$0 < y < \infty, \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty. \quad (2.49)$$

When $\eta = 1$, we will have $Y \sim \mathrm{IG}(\beta, \alpha)$. Therefore, we can take the inverse Gaussian distribution as a special case of the power inverse Gaussian [6].

Proposition 2.10. *If a power inverse Gaussian random variable Y is distributed as* (2.48)*, then*

$$\mathsf{E}[\log Y] = \frac{1}{\eta} e^{\frac{2\alpha}{\eta^2\beta}} \operatorname{Ei}\left(-\frac{2\alpha}{\eta^2\beta}\right) + \log\beta; \quad (2.50)$$

$$\mathsf{E}\left[\left(\frac{Y}{\beta}\right)^{\eta} + \left(\frac{Y}{\beta}\right)^{-\eta}\right] = 2 + \frac{\eta^2\beta}{\alpha}. \quad (2.51)$$

Proof. See [6]. □

Proposition 2.11 (Differential Entropy of the Power Inverse Gaussian)**.** *If a power inverse Gaussian random variable Y is distributed as* (2.48)*, its differential entropy will be*

$$h(Y_{PIG}) = -\log\sqrt{\frac{\alpha}{2\pi\beta^3}} + \left(\frac{1}{\eta} + \frac{1}{2}\right) e^{\frac{2\alpha}{\eta^2\beta}} \operatorname{Ei}\left(-\frac{2\alpha}{\eta^2\beta}\right) + \frac{1}{2}. \quad (2.52)$$

Proof. The claim can be derived simply by plugging in Proposition 2.10. □

Proposition 2.12. *If a power inverse Gaussian random variable Y is distributed as* (2.48)*,*

$$\mathsf{E}\left[Y^b\right] = \frac{\beta^{b-1/2}}{\eta}\sqrt{\frac{2\alpha}{\pi}} \exp\left(\frac{\alpha}{\eta^2\beta}\right) K_{\frac{b}{\eta}-\frac{1}{2}}\left(\frac{\alpha}{\eta^2\beta}\right). \quad (2.53)$$

Proof. See [8, Ch. 18]. □

Lemma 2.13. *If Y is a random variable that satisfied the following two constraints:*

$$\mathsf{E}[\log Y] = \frac{1}{\eta} e^{\frac{2\alpha}{\eta^2\beta}} \operatorname{Ei}\left(-\frac{2\alpha}{\eta^2\beta}\right) + \log\beta \quad (2.54)$$

and

$$\mathsf{E}\left[\left(\frac{Y}{\beta}\right)^{\eta} + \left(\frac{Y}{\beta}\right)^{-\eta}\right] = 2 + \frac{\eta^2\beta}{\alpha} \quad (2.55)$$

where η is an fixed real number, $\eta \neq 0$, $\beta > 0$ and $\alpha > 0$, the distribution of Y that maximizes the entropy is the power inverse Gaussian distribution.

2.3 Related Lemmas and Propositions

In this section, we will show the lemmas and properties which will be used in our proof of bounds.

The first one is the *data processing theorem for relative entropy*. It can also be called *relative entropy processing theorem*, or *monotonicity theorem of relative entropy.*

Lemma 2.14 (Data Processing Theorem for Relative Entropy)**.** *Let Q_{X_1} and Q_{X_2} be two input distributions of a communication channel, and Q_{Y_1} and Q_{Y_2} the two corresponding output distributions. Then*

$$D(Q_{X_1}\|Q_{X_2}) \geq D(Q_{Y_1}\|Q_{Y_2}), \tag{2.56}$$

where

$$D(Q_{X_1}\|Q_{X_2}) \triangleq \int Q_{X_1}(x) \log \frac{Q_{X_1}(x)}{Q_{X_2}(x)}\, \mathrm{d}x. \tag{2.57}$$

Proof. See, e.g., [9, (B.102)]. □

Lemma 2.4 says that due to the noise introduced in a channel, two output distributions are more difficult to distinguish from each other that the two corresponding inputs distribution.

Next, we will list some propositions related to the $\mathcal{Q}$-function.

Definition 2.15. *The $\mathcal{Q}$-function is defined by*

$$\mathcal{Q}(\alpha) \triangleq \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{t^2/2}\, \mathrm{d}t. \tag{2.58}$$

Note that $\mathcal{Q}(\alpha)$ is the probability that a standard Gaussian random variable will exceed the value α and is therefore monotonically decreasing with an increasing argument.

Proposition 2.16 (Bounds for the $\mathcal{Q}$-function)**.**

$$\frac{1}{\sqrt{2\pi}\alpha} e^{-\frac{\alpha^2}{2}} \left(1 - \frac{1}{\alpha^2}\right) < \mathcal{Q}(\alpha) < \frac{1}{\sqrt{2\pi}\alpha} e^{-\frac{\alpha^2}{2}}, \qquad \alpha > 0; \tag{2.59}$$

$$\mathcal{Q}(\alpha) \leq \frac{1}{2} e^{-\frac{\alpha^2}{2}}, \qquad \alpha \geq 0. \tag{2.60}$$

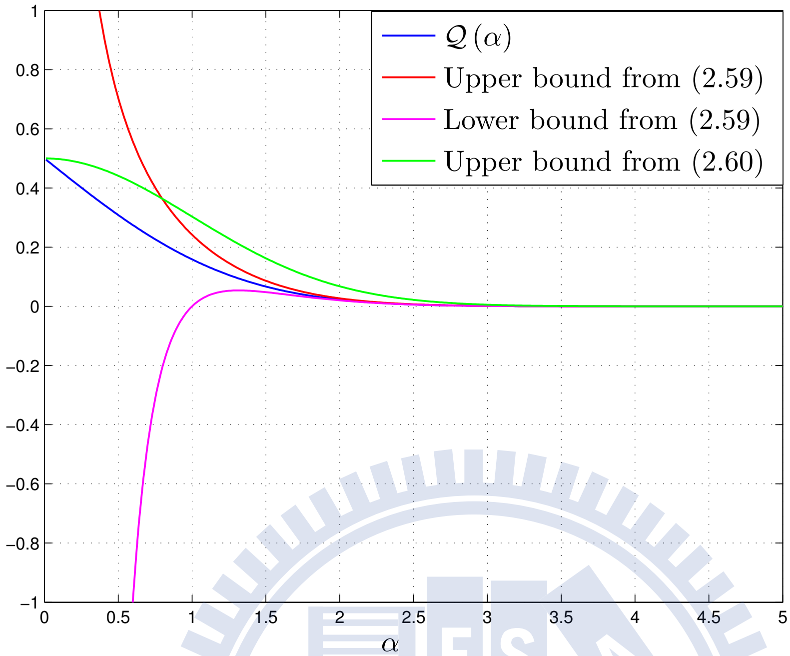
Proposition 2.17. *Let $\Phi(\cdot)$ denote the cumulative distribution function (CDF) of the standard normal distribution:*

$$\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-t^2/2}\, \mathrm{d}t. \tag{2.61}$$

Then

$$\mathcal{Q}(\alpha) = \Phi(-\alpha), \tag{2.62}$$

$$\mathcal{Q}(\alpha) + \mathcal{Q}(-\alpha) = 1. \tag{2.63}$$



In `MATLAB`, we can use `y=qfunc(x)` to get the value of the $\mathcal{Q}$-function.

Proposition 2.18 (Upper and Lower Bound for Exponential Integral Function)**.** *We have*

$$\frac{1}{2}e^{-x}\ln\left(1+\frac{2}{x}\right) < \mathrm{E}_1(x) = -\,\mathrm{Ei}(-x) < e^{-x}\ln\left(1+\frac{1}{x}\right), \quad x>0, \tag{2.64}$$

or

$$-e^{-x}\ln\left(1+\frac{1}{x}\right) < -\mathrm{E}_1(x) = \mathrm{Ei}(-x) < -\frac{1}{2}e^{-x}\ln\left(1+\frac{2}{x}\right), \quad x>0. \tag{2.65}$$

Chapter 3

Known Bounds to the Capacity of the AIGN Channel

The entropy maximizing distribution $f^*(y)$ with a mean constraint $\mathsf{E}[Y] \leq m+\mu$ is the exponential distribution with parameter $\frac{1}{m+\mu}$ [7, (12.21)]:

$$f^*(y) = \frac{1}{m+\mu} e^{-\frac{y}{m+\mu}}, \quad y \geq 0. \tag{3.1}$$

The entropy of such a distribution is

$$h^*(Y) = 1 + \ln(m+\mu). \tag{3.2}$$

This can be used to derive a upper bound on the capacity of the AIGN channel:

$$\begin{aligned} &\triangleq \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} I(X;Y) && (3.3)\\ &= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} \left\{ h(Y) - h_{\text{IG}(\mu,\lambda)} \right\} && (3.4)\\ &= \sup_{f_X(x)\colon \mathsf{E}[X]\leq m} h(Y) - h_{\text{IG}(\mu,\lambda)} && (3.5)\\ &= 1 + \ln(m+\mu) - h_{\text{IG}(\mu,\lambda)}. && (3.6) \end{aligned}$$

In [1], to derive a lower bound, we drop the maximization and the additivity property of the IG distribution is used. We choose an input signal X to be IG in such a way, according to Lemma 2.2, that the output Y will be also inverse Gaussian.

Since the noise distribution is $N \sim \text{IG}(\mu,\lambda)$, i.e., $\kappa = \frac{\lambda}{\mu^2}$. Setting $X \sim \text{IG}(m,\lambda_x)$, we must satisfy

$$\kappa = \frac{\lambda}{\mu^2} = \frac{\lambda_x}{m^2}. \tag{3.7}$$

Hence we need

$$\lambda_x = \lambda \frac{m^2}{\mu^2}. \tag{3.8}$$

As a result, the input X is chosen as:

$$X \sim \mathrm{IG}\left(m, \lambda \frac{m^2}{\mu^2}\right). \tag{3.9}$$

The corresponding output Y is then also IG distributed:

$$Y \sim \mathrm{IG}\left(m+\mu, \frac{\lambda}{\mu^2}(\mu+m)^2\right). \tag{3.10}$$

The distribution of Y is not necessarily an entropy maximizing distribution for a given mean, $m+\mu$.

Combing the upper and lower bound on capacity, we have the following:

$$h_{\mathrm{IG}\left(m+\mu, \frac{\lambda}{\mu^2}(m+\mu)^2\right)} - h_{\mathrm{IG}(\mu,\lambda)} \leq \quad \leq 1 + \log(\mu+m) - h_{\mathrm{IG}(\mu,\lambda)}. \tag{3.11}$$

Using Proposition 2.1, we have

$$\begin{aligned} h_{\mathrm{IG}\left(m+\mu, \frac{\lambda}{\mu^2}(m+\mu)^2\right)} &= \frac{1}{2}\log\frac{2\pi\mu^2(m+\mu)}{\lambda} \\ &\quad + \frac{3}{2}\exp\left(\frac{2\lambda(m+\mu)}{\mu^2}\right)\cdot \mathrm{Ei}\left(-\frac{2\lambda(m+\mu)}{\mu^2}\right) + \frac{1}{2}, \end{aligned} \tag{3.12}$$

which then results in the following bounds:

$$\begin{aligned} &\geq \frac{1}{2}\log\frac{m+\mu}{\mu} \\ &\quad + \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\left(\exp\left(\frac{2\lambda m}{\mu^2}\right)\mathrm{Ei}\left(-\frac{2\lambda(m+\mu)}{\mu^2}\right) - \mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right)\right) && (3.13) \\ &= \frac{1}{2}\log\frac{mv+d}{d} \\ &\quad + \frac{3}{2}\exp\left(\frac{2dv}{\sigma^2}\right)\left(\exp\left(\frac{2mv^2}{\sigma^2}\right)\mathrm{Ei}\left(-\frac{2v(mv+d)}{\sigma^2}\right) - \mathrm{Ei}\left(-\frac{2dv}{\sigma^2}\right)\right); && (3.14) \\ &\leq \frac{1}{2}\log\frac{\lambda(m+\mu)^2}{2\pi\mu^3} - \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2} && (3.15) \\ &= \frac{1}{2}\log\frac{v(mv+d)^2}{2\pi d\sigma^2} - \frac{3}{2}\exp\left(\frac{2dv}{\sigma^2}\right)\mathrm{Ei}\left(-\frac{2dv}{\sigma^2}\right) + \frac{1}{2}. && (3.16) \end{aligned}$$

In (3.13) and (3.15), we express the capacity as a function of m, μ and λ, while in (3.14) and (3.16) we show the same expression as a function of m, v, σ^2 and d. These bounds are depicted in Fig. 3.3 and Fig. 3.4

We see in Fig. 3.3 that the known upper bound performs not good at high velocities and at very low velocities.

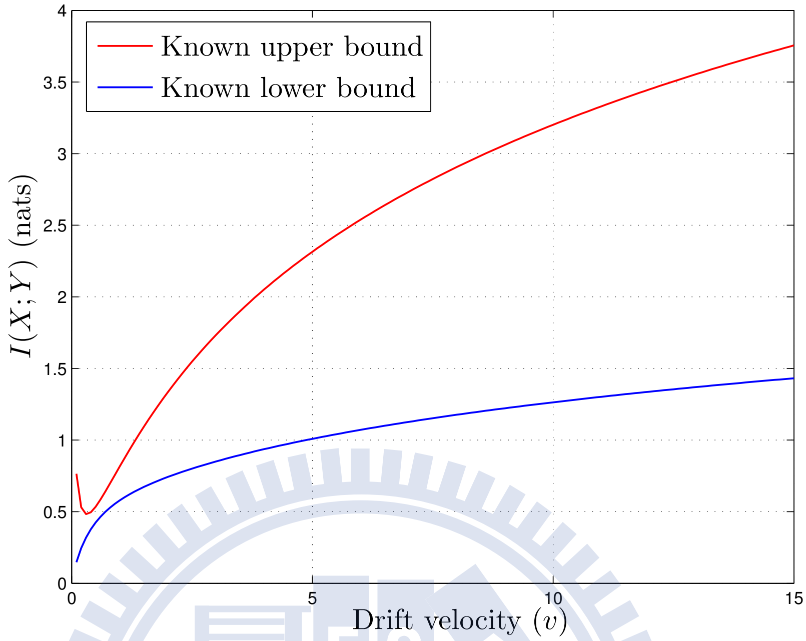


Figure 3.3: Upper and lower bound (3.14) and (3.16) of the AIGN channel for the choice: $m = 1$, $d = 1$, and $\sigma^2 = 1$.

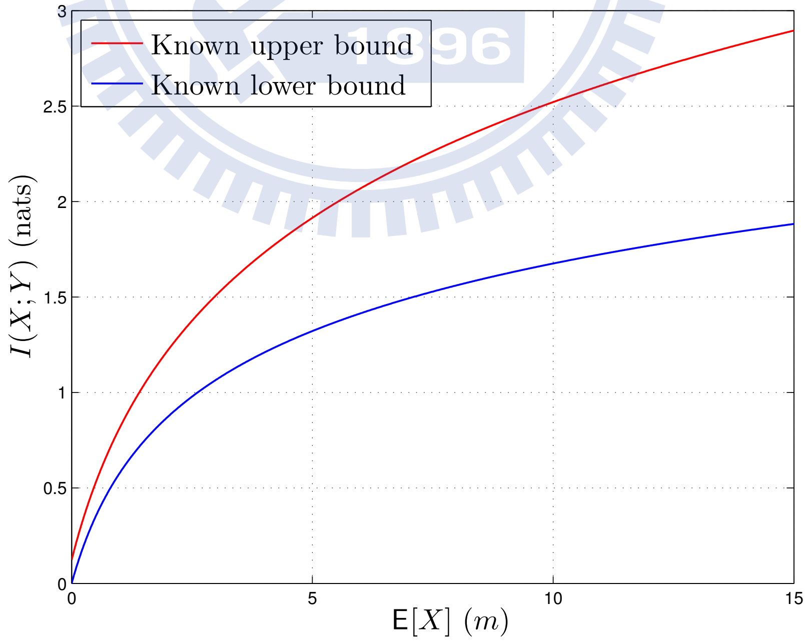


Figure 3.4: Upper and lower bound in (3.14) and (3.16) of the AIGN channel for the choice: $v = 1$, $d = 1$, and $\sigma^2 = 1$.

Chapter 4

Our Different Trials of Lower Bounds

4.1 Lower Bounds of $h(Y)$ Based on $h(X)$

Theorem 4.1. *A lower bound on the output entropy of the inverse Gaussian channel is as follow:*

$$h(Y) \geq h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \log\frac{m}{m+\mu} + \frac{\lambda m}{2\mu^2} \tag{4.1}$$

$$= h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{(mv)^2}{2\sigma^2}\mathsf{E}\left[\frac{1}{X}\right] + \log\frac{mv}{mv+1} + \frac{mv^2}{2\sigma^2}. \tag{4.2}$$

Proof. Our approach of lower bounding $h(Y)$ is based on the data process inequality of relative entropy as given in Lemma 2.14

$$D(Q_X \| Q_{X_{\text{IG}}}) \geq D(Q_Y \| Q_{Y_{\text{IG}}}). \tag{4.3}$$

We pick Q_{X_2} as

$$X_{\text{IG}} \sim \text{IG}\left(m, \lambda\frac{m^2}{\mu^2}\right), \tag{4.4}$$

such that Q_{Y_2} will be

$$Y_{\text{IG}} \sim \text{IG}\left(m+\mu, \frac{\lambda}{\mu^2}(\mu+m)^2\right) \tag{4.5}$$

and keep Q_{X_1} and Q_{Y_1} arbitrary distributed as Q_X and Q_Y.

The left hand side of (4.3) can be evaluated as follows:

$$D(Q_X \| Q_{X_{\text{IG}}}) = -h(X) - \mathsf{E}_{Q_X}\left[\log\left(\left(\frac{\lambda\frac{m^2}{\mu^2}}{2\pi X^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda\frac{m^2}{\mu^2}(X-m)^2}{2m^2 X}\right)\right)\right] \tag{4.6}$$

$$= -h(X) - \frac{1}{2}\log\frac{\lambda m^2}{2\pi\mu^2} + \frac{3}{2}\mathsf{E}_{Q_X}[\log X] + \mathsf{E}_{Q_X}\left[\frac{\lambda(X-m)^2}{2\mu^2 X}\right] \tag{4.7}$$

$$= -h(X) - \frac{1}{2}\log\frac{\lambda m^2}{2\pi\mu^2} + \frac{3}{2}\mathsf{E}_{Q_X}[\log X] + \mathsf{E}_{Q_X}\left[\frac{\lambda(X^2-2mX+m^2)}{2\mu^2 X}\right] \tag{4.8}$$

$$= -h(X) + \frac{3}{2}\mathsf{E}_{Q_X}[\log X] + \frac{\lambda}{2\mu^2}\mathsf{E}_{Q_X}[X] + \frac{\lambda m^2}{2\mu^2}\mathsf{E}_{Q_X}\left[\frac{1}{X}\right] - \frac{1}{2}\log\frac{\lambda m^2}{2\pi\mu^2} - \frac{\lambda m}{\mu^2} \tag{4.9}$$

The right hand side of (4.3) will be:

$$D(Q_Y \| Q_{Y_{\mathrm{IG}}})$$

$$= -h(Y) - \mathsf{E}_{Q_Y}\left[\log\left(\left(\frac{\frac{\lambda}{\mu^2}(\mu+m)^2}{2\pi Y^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\frac{\lambda}{\mu^2}(\mu+m)^2(Y-\mu-m)^2}{2(\mu+m)^2 Y}\right)\right)\right] \tag{4.10}$$

$$= -h(Y) - \frac{1}{2}\log\frac{\lambda(m+\mu)^2}{2\pi\mu^2} + \frac{3}{2}\mathsf{E}_{Q_Y}[\log Y] + \mathsf{E}_{Q_Y}\left[\frac{\lambda(Y-\mu-m)^2}{2\mu^2 Y}\right] \tag{4.11}$$

$$= -h(Y) - \frac{1}{2}\log\frac{\lambda(m+\mu)^2}{2\pi\mu^2} + \frac{3}{2}\mathsf{E}_{Q_Y}[\log Y] + \mathsf{E}_{Q_Y}\left[\frac{\lambda(Y^2+\mu^2+m^2-2\mu Y-2mY+2m\mu)}{2\mu^2 Y}\right] \tag{4.12}$$

$$= -h(Y) + \frac{3}{2}\mathsf{E}_{Q_Y}[\log Y] + \frac{\lambda}{2\mu^2}\mathsf{E}_{Q_Y}[Y] + \frac{\lambda(m+\mu)^2}{2\mu^2}\mathsf{E}_{Q_Y}\left[\frac{1}{Y}\right] - \frac{1}{2}\log\frac{\lambda(m+\mu)^2}{2\pi\mu^2} - \frac{\lambda(m+\mu)}{\mu^2}. \tag{4.13}$$

After rearranging both side of the data processing inequality, we get:

$$\begin{aligned} h(Y) \geq{} & h(X) + \log\frac{m}{m+\mu} - \frac{\lambda}{\mu} \\ & - \frac{3}{2}\mathsf{E}_{Q_X}[\log X] - \frac{\lambda}{2\mu^2}\mathsf{E}_{Q_X}[X] - \frac{\lambda m^2}{2\mu^2}\mathsf{E}_{Q_X}\left[\frac{1}{X}\right] \\ & + \frac{3}{2}\mathsf{E}_{Q_Y}[\log Y] + \frac{\lambda}{2\mu^2}\mathsf{E}_{Q_Y}[Y] + \frac{\lambda(m+\mu)^2}{2\mu^2}\mathsf{E}_{Q_Y}\left[\frac{1}{Y}\right] \end{aligned} \tag{4.14}$$

$$= h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{\lambda(m+\mu)^2}{2\mu^2}\mathsf{E}\left[\frac{1}{Y}\right] + \frac{\lambda}{2\mu^2}\mathsf{E}[Y-X] + \log\frac{m}{m+\mu} - \frac{\lambda}{\mu} \tag{4.15}$$

$$= h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{\lambda(m+\mu)^2}{2\mu^2}\mathsf{E}\left[\frac{1}{Y}\right] + \log\frac{m}{m+\mu} - \frac{\lambda}{2\mu}. \tag{4.16}$$

To get rid of $\mathsf{E}\left[\frac{1}{Y}\right]$, we use Jensen's inequality:

$$\mathsf{E}\left[\frac{1}{Y}\right] \geq \frac{1}{\mathsf{E}[Y]} = \frac{1}{\mathsf{E}[X]+\mathsf{E}[N]} \geq \frac{1}{m+\mu}, \tag{4.17}$$

which yields

$$h(Y) \geq h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{\lambda(m+\mu)}{2\mu^2} + \log\frac{m}{m+\mu} - \frac{\lambda}{2\mu} \tag{4.18}$$

$$= h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \log\frac{m}{m+\mu} + \frac{\lambda m}{2\mu^2}. \tag{4.19}$$

The second bound (4.2) then follows simply by substituting $\mu = \frac{d}{v}$ and $\lambda = \frac{d^2}{\sigma^2}$ into (4.1). □

4.2 Capacity Lower Bounds Based on Theorem 4.1

In this section, it is our goal to continue trying to make the lower bound only a function of X, independent of Y. I.e., we need to replace $\mathsf{E}[\log Y]$ by further lower-bounding it. We use two different approaches.

4.2.1 Lower Bound 1: Taylor Expansion

We use a Taylor expansion of the logarithm:

$$\log(1+u) = -\sum_{n=0}^{\infty}(-1)^{n+1}\frac{u^n}{n} \qquad \text{for } -1 < u \leq 1, \tag{4.20}$$

$$= u - \frac{u^2}{2} + \frac{u^3}{3} - \frac{u^4}{4} + \frac{u^5}{5} - \frac{u^6}{6} + \cdots \tag{4.21}$$

$$\log(1+u) \geq u - \frac{u^2}{2}, \qquad \forall u \geq 0, \tag{4.22}$$

$$\log\left(1+\frac{N}{x}\right) \geq \frac{N}{x} - \frac{N^2}{2x^2}, \qquad \frac{N}{x} \geq 0. \tag{4.23}$$

We use this bound in the following way:

$$\mathsf{E}[\log Y] - \mathsf{E}[\log X] = \mathsf{E}\left[\log\frac{N+X}{X}\right] \tag{4.24}$$

$$= \mathsf{E}_X\left[\mathsf{E}_N\left[\log\left(1+\frac{N}{x}\right)\middle| X = x\right]\right] \tag{4.25}$$

$$\geq \mathsf{E}_X\left[\mathsf{E}_N\left[\frac{N}{x} - \frac{N^2}{2x^2}\middle| X = x\right]\right] \tag{4.26}$$

$$= \mathsf{E}_X\left[\frac{\mathsf{E}[N]}{X} - \frac{\mathsf{E}\left[N^2\right]}{2X^2}\right] \tag{4.27}$$

$$= \mathsf{E}_X\left[\frac{\mu}{X} - \frac{\mathsf{E}\left[N^2\right]}{2X^2}\right] \tag{4.28}$$

$$= \mathsf{E}_X\left[\frac{\mu}{X} - \frac{\mu^2}{2X^2}\left(1+\frac{\mu}{\lambda}\right)\right] \tag{4.29}$$

$$= \mu \mathsf{E}\left[\frac{1}{X}\right] - \frac{\mu^2}{2}\left(1+\frac{\mu}{\lambda}\right)\mathsf{E}\left[\frac{1}{X^2}\right], \tag{4.30}$$

where in (4.29) we use (2.34). This bound can now be plugged into Theorem 4.1: plugging (4.30) into (4.1), we get a new lower bound as a function only of X:

$$\begin{aligned} h(Y) &\geq h(X) + \frac{3\mu}{2}\mathsf{E}\left[\frac{1}{X}\right] - \frac{3\mu^2}{4}\left(1+\frac{\mu}{\lambda}\right)\mathsf{E}\left[\frac{1}{X^2}\right] - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] \\ &\quad + \log\frac{m}{m+\mu} + \frac{\lambda m}{2\mu^2} \qquad (4.31)\\ &= h(X) + \frac{1}{2}\left(3\mu - \frac{\lambda m^2}{\mu^2}\right)\mathsf{E}\left[\frac{1}{X}\right] - \frac{3\mu^2}{4}\left(1+\frac{\mu}{\lambda}\right)\mathsf{E}\left[\frac{1}{X^2}\right] \\ &\quad + \log\frac{m}{m+\mu} + \frac{\lambda m}{2\mu^2} \qquad (4.32)\\ &= h(X) + \frac{1}{2}\left(\frac{3d}{v} - \left(\frac{mv}{\sigma}\right)^2\right)\mathsf{E}\left[\frac{1}{X}\right] - \frac{3d}{4v^3}(dv+\sigma^2)\mathsf{E}\left[\frac{1}{X^2}\right] \\ &\quad + \log\frac{mv}{mv+1} + \frac{mv^2}{2\sigma^2}. \qquad (4.33) \end{aligned}$$

Theorem 4.2 (A lower bound on capacity of the inverse Gaussian channel)**.**

$$\begin{aligned} &\geq h(X) + \frac{1}{2}\left(3\mu - \frac{\lambda m^2}{\mu^2}\right)\mathsf{E}\left[\frac{1}{X}\right] - \frac{3\mu^2}{4}\left(1+\frac{\mu}{\lambda}\right)\mathsf{E}\left[\frac{1}{X^2}\right] \\ &\quad + \log\left(\frac{m}{m+\mu}\left(\frac{\lambda}{2\pi\mu^3}\right)^{\frac{1}{2}}\right) - \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{\lambda m}{2\mu^2} - \frac{1}{2} \qquad (4.34)\\ &= h(X) + \frac{1}{2}\left(\frac{3d}{v} - \left(\frac{mv}{\sigma}\right)^2\right)\mathsf{E}\left[\frac{1}{X}\right] - \frac{3d}{4v^3}(dv+\sigma^2)\mathsf{E}\left[\frac{1}{X^2}\right] \\ &\quad + \log\frac{mv^{\frac{5}{2}}}{\sqrt{2\pi}\sigma(mv+1)} - \frac{3}{2}\exp\left(\frac{2v}{\sigma^2}\right)\mathrm{Ei}\left(-\frac{2v}{\sigma^2}\right) + \frac{mv^2}{2\sigma^2} - \frac{1}{2}. \qquad (4.35) \end{aligned}$$

Proof.

$$\begin{aligned} \geq I(X;Y) &= h(Y) - h(Y|X) \qquad (4.36)\\ &= h(Y) - h(X+N|X) \qquad (4.37)\\ &= h(Y) - h_{\mathrm{IG}(\mu,\lambda)}. \qquad (4.38) \end{aligned}$$

The proof is complete if we plug equation (4.32) and (4.33) separately and Proposition 2.1 in (4.38). □

We plug $X \sim$IG(m,β) into our lower bound of Theorem 4.2. Unfortunately, this lower bound is not tighter than the known lower bound and is even decreasing as v gets larger.

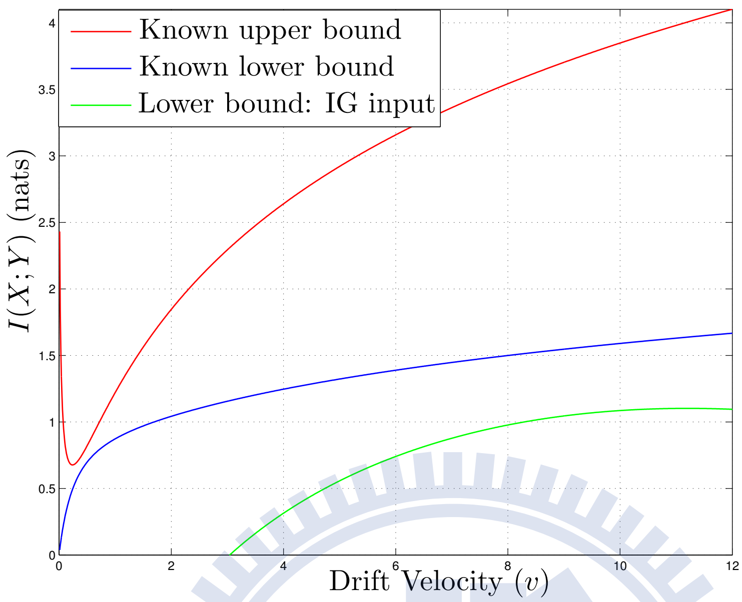


Figure 4.5: $m = 2$, $\sigma^2 = 1$, $d = 1$

4.2.2 Lower Bound 2

Another approach of lower bounding $\mathsf{E}[\log Y] - \mathsf{E}[\log X]$ is as follows:

$$\mathsf{E}[\log Y] - \mathsf{E}[\log X] = \mathsf{E}\left[\log \frac{X+N}{X}\right] \tag{4.39}$$

$$= \mathsf{E}[\log N] + \mathsf{E}\left[\log\left(\frac{1}{N} + \frac{1}{X}\right)\right] \tag{4.40}$$

$$\geq e^{\frac{2\lambda}{\mu}} \operatorname{Ei}\left(-\frac{2\lambda}{\mu}\right) + \log \mu + \mathsf{E}\left[\log\left(\frac{1}{\mu} + \frac{1}{X}\right)\right] \tag{4.41}$$

$$\geq e^{\frac{2\lambda}{\mu}} \operatorname{Ei}\left(-\frac{2\lambda}{\mu}\right) + \log \mu + \log\left(\frac{1}{\mu} + \frac{1}{m}\right). \tag{4.42}$$

Therefore, we derive the capacity lower bound as follow:

$$\geq h(Y) - h(N) \tag{4.43}$$

$$\geq h(X) + \frac{3}{2}\left(\mathsf{E}[\log Y] - \mathsf{E}[\log X]\right) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \log\frac{m}{m+\mu} + \frac{\lambda m}{2\mu^2} - h(N) \tag{4.44}$$

$$\geq h(X) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{\lambda m}{2\mu^2} - \frac{1}{2} + \frac{1}{2}\log\frac{(m+\mu)\lambda}{2\pi m \mu^3}, \tag{4.45}$$

where equation (4.44) comes from (4.19) and the equation (4.45) comes from (4.42) and (2.2).

Corollary 4.3.

$$\geq h(X) - \frac{\lambda m^2}{2\mu^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{\lambda m}{2\mu^2} + \frac{1}{2}\log\frac{(m+\mu)\lambda}{2\pi e m \mu^3} \tag{4.46}$$

$$= h(X) - \frac{m^2 v^2}{2\sigma^2}\mathsf{E}\left[\frac{1}{X}\right] + \frac{m v^2}{2\sigma^2} + \frac{1}{2}\log\frac{(mv+d)v^2}{2\pi e m \sigma^2 d}. \tag{4.47}$$

As a choice for the input X, we choose a power inverse Gaussian distribution described in Section 2.2:

$$f_X(x) = \sqrt{\frac{\alpha}{2\pi\beta^3}}\left(\frac{x}{\beta}\right)^{-(1+\frac{\eta}{2})}\exp\left(-\frac{\alpha}{2\eta^2\beta}\left(\left(\frac{x}{\beta}\right)^{\frac{\eta}{2}} - \left(\frac{x}{\beta}\right)^{-\frac{\eta}{2}}\right)^2\right), \tag{4.48}$$

where

$$0 < x < \infty, \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty. \tag{4.49}$$

From Proposition 2.12 and the delay constraint of X we have:

$$\mathsf{E}[X] = \beta\sqrt{\frac{2\tilde{\alpha}}{\pi}}e^{\tilde{\alpha}}K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha}) \overset{!}{=} m. \tag{4.50}$$

Defining $\tilde{\alpha} \triangleq \frac{\alpha}{\eta^2\beta}$ now leads to

$$\beta = \frac{m}{\sqrt{\frac{2\tilde{\alpha}}{\pi}}e^{\tilde{\alpha}}K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})}. \tag{4.51}$$

From Proposition 2.11, we have

$$h(X) = -\frac{1}{2}\log\frac{\tilde{\alpha}\eta^2}{2\pi\beta^2} + \left(\frac{1}{\eta}+\frac{1}{2}\right)e^{2\tilde{\alpha}}\,\mathrm{Ei}(-2\tilde{\alpha}) + \frac{1}{2}, \tag{4.52}$$

and

$$\mathsf{E}\left[\frac{1}{X}\right] = \sqrt{\frac{2\tilde{\alpha}}{\pi\beta^2}}e^{\tilde{\alpha}}K_{-\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha}) \tag{4.53}$$

$$= \frac{m}{\beta^2}\frac{K_{-\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})}{K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})} \tag{4.54}$$

$$= \frac{2\tilde{\alpha}}{m\pi}e^{2\tilde{\alpha}}K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})\,K_{-\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha}). \tag{4.55}$$

Plugging (4.51), (4.52) and (4.55) into (4.46), we derive the lower bound as follows:

$$\geq -\frac{1}{2}\log\frac{\tilde{\alpha}\eta^2}{2\pi e\beta^2} + \left(\frac{1}{\eta}+\frac{1}{2}\right)e^{2\tilde{\alpha}}\,\mathrm{Ei}(-2\tilde{\alpha}) - \frac{m\lambda\tilde{\alpha}}{\pi\mu^2}e^{2\tilde{\alpha}}K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})\,K_{-\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})$$
$$+ \frac{\lambda m}{2\mu^2} + \frac{1}{2}\log\frac{(m+\mu)\lambda}{2\pi e m\mu^3} \tag{4.56}$$

$$= \frac{1}{2}\log\frac{\pi m(m+\mu)\lambda}{2\mu^3} - \log\tilde{\alpha} - \log|\eta| - \tilde{\alpha} - \log K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha}) + \frac{\lambda m}{2\mu^2}$$
$$+ \left(\frac{1}{\eta}+\frac{1}{2}\right)e^{2\tilde{\alpha}}\,\mathrm{Ei}(-2\tilde{\alpha}) - \frac{m\lambda\tilde{\alpha}}{\pi\mu^2}e^{2\tilde{\alpha}}K_{\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha})\,K_{-\frac{1}{\eta}-\frac{1}{2}}(\tilde{\alpha}). \tag{4.57}$$

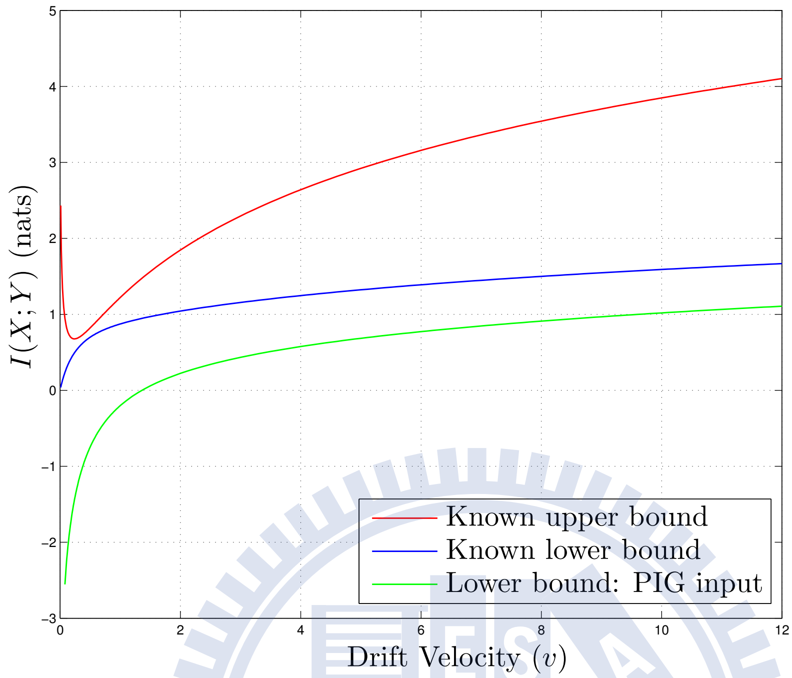


Figure 4.6: $m = 2$, $\sigma^2 = 1$, $d = 1$

The parameters α, β and η are freely choosable. We optimize their value numerically and get a lower bound shown in Fig. 4.6.

Unfortunately, the lower bound with power inverse Gaussian is not tighter than the known lower bound.

4.3 Lower Bound Based on the Convolution of Exponential and Inverse Gaussian Distribution

We start with a new approach based on the choice of the input $X \sim \text{Exp}\left(\frac{1}{m}\right)$. Because it is an additive channel, the resulting output will have a PDF that is the convolution of the exponential PDF with the inverse Gaussian PDF. This can actually be computed explicitly [10, Eq.(18)]:

$$f_Y(y) = \frac{1}{m}e^{-\frac{Y}{m}+\frac{\lambda}{\mu}}\left(e^{-kd\lambda}\Phi\left(\frac{kY-d}{\sqrt{Y/\lambda}}\right)+e^{kd\lambda}\Phi\left(-\frac{kY+d}{\sqrt{Y/\lambda}}\right)\right) \tag{4.58}$$

$$= \frac{1}{m}e^{-\frac{y}{m}+\frac{\lambda}{\mu}}\left(e^{-k\lambda}\left(1-\mathcal{Q}\left(\frac{ky-1}{\sqrt{y/\lambda}}\right)\right)+e^{k\lambda}\mathcal{Q}\left(\frac{ky+1}{\sqrt{y/\lambda}}\right)\right) \tag{4.59}$$

$$= \frac{1}{m}e^{-\frac{y}{m}+\frac{\lambda}{\mu}}\left(e^{-k\lambda}\mathcal{Q}\left(-\sqrt{k\lambda}\left(\sqrt{ky}-\frac{1}{\sqrt{ky}}\right)\right)\right)$$

$$+ e^{k\lambda}\mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{ky}+\frac{1}{\sqrt{ky}}\right)\right), \tag{4.60}$$

where

$$k = \sqrt{v^2 - \frac{2\sigma^2}{m}} = d\sqrt{\frac{v^2}{d^2} - \frac{2\sigma^2}{d^2 m}} = \sqrt{\frac{1}{\mu^2} - \frac{2}{\lambda m}}. \tag{4.61}$$

We remind that without loss of generality, we assume to have a unit length $d = 1$ between transmitter and receiver. Note that to make sure that the bound is real, it must constrain the channel parameters to satisfy

$$m \geq \frac{2\mu^2}{\lambda}. \tag{4.62}$$

This now yields the following lower bound on capacity:

$$\triangleq \max_{f_X(x)} I(X;Y) \tag{4.63}$$

$$\geq I(X;Y)\big|_{X\sim\exp(\frac{1}{m})} \tag{4.64}$$

$$= (h(Y) - h(Y|X))\big|_{X\sim\exp(\frac{1}{m})} \tag{4.65}$$

$$= h(Y)\big|_{X\sim\exp(\frac{1}{m})} - h(N) \tag{4.66}$$

$$= -\mathsf{E}_Y[\log f_Y(Y)] - h(N) \tag{4.67}$$

$$\begin{aligned} = -\mathsf{E}_Y\Bigg[\log\Bigg(1 - \mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{kY}-\sqrt{\frac{1}{kY}}\right)\right) \\ + e^{2k\lambda}\mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{kY}+\sqrt{\frac{1}{kY}}\right)\right)\Bigg)\Bigg] \\ + \log m + \frac{m+\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N) \end{aligned} \tag{4.68}$$

$$\begin{aligned} \geq -\mathsf{E}_Y\left[\log\left(1 + e^{2k\lambda}\mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{kY}+\sqrt{\frac{1}{kY}}\right)\right)\right)\right] + \log m + \frac{m+\mu}{m} \\ - \frac{\lambda}{\mu} + k\lambda - h(N) \end{aligned} \tag{4.69}$$

$$\geq -\mathsf{E}_Y\left[\log\left(1 + e^{2k\lambda}\frac{1}{2}e^{-\frac{1}{2}k\lambda\left(kY+2+\frac{1}{kY}\right)}\right)\right] + \log m + \frac{m+\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N) \tag{4.70}$$

$$= -\mathsf{E}_Y\left[\log\left(1 + \frac{1}{2}e^{-\frac{1}{2}k\lambda\left(kY-2+\frac{1}{kY}\right)}\right)\right] + \log m + \frac{m+\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N) \tag{4.71}$$

$$= -\mathsf{E}_Y\left[\log\left(1 + \frac{1}{2}e^{-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2}\right)\right] + \log m + \frac{m+\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N). \tag{4.72}$$

Here in (4.69) we lower-bound the first $\mathcal{Q}$-function by 0 because $\mathcal{Q}(\cdot)$ is nonnegative. Then based on that $\log(\cdot)$ is a monotonic function, in (4.70) we apply the upper bound for the $\mathcal{Q}$-function (2.60).

To further bound this expectation over Y, we provides two methods. The first method is simply using

$$\mathsf{E}_Y\left[\log\left(1+\frac{1}{2}e^{-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2}\right)\right] \le \log\frac{3}{2}, \tag{4.73}$$

since

$$-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2 \le 0. \tag{4.74}$$

Therefore, the first method gives us the result

$$\ge \log m - \log\frac{3}{2} + \frac{\mu}{m} - \frac{\lambda}{\mu} + k\lambda + \frac{1}{2}\log\frac{\lambda e}{2\pi} - \frac{3}{2}\log\mu - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\operatorname{Ei}\left(-\frac{2\lambda}{\mu}\right). \tag{4.75}$$

Another option is applying Jensen's inequality:

$$\ge -\mathsf{E}_Y\left[\log\left(1+\frac{1}{2}e^{-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2}\right)\right] + \log m + 1 + \frac{\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N) \tag{4.76}$$

$$\ge -\log\left(\mathsf{E}_Y\left[1+\frac{1}{2}e^{-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2}\right]\right) + \log m + 1 + \frac{\mu}{m} - \frac{\lambda}{\mu} + k\lambda - h(N). \tag{4.77}$$

Making a variable transformation $t = ky$ we now evaluate the expectation as follows:

$$\begin{aligned}
&\mathsf{E}_Y\left[e^{-\frac{1}{2}k\lambda\left(\sqrt{kY}-\frac{1}{\sqrt{kY}}\right)^2}\right] \\
&= \int_0^\infty \frac{1}{km}e^{\frac{\lambda}{\mu}-\frac{t}{km}-k\lambda}e^{-\frac{1}{2}k\lambda\left(\sqrt{t}-\frac{1}{\sqrt{t}}\right)^2} \\
&\quad\cdot\left(1-\mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{t}-\frac{1}{\sqrt{t}}\right)\right)+e^{2k\lambda}\mathcal{Q}\left(\sqrt{k\lambda}\left(\sqrt{t}+\frac{1}{\sqrt{t}}\right)\right)\right)\mathrm{d}t && (4.78)\\
&\le \int_0^\infty \frac{1}{km}e^{\frac{\lambda}{\mu}-\frac{t}{km}}\left(1+\frac{1}{2}e^{-\frac{1}{2}k\lambda\left(t-2+\frac{1}{t}\right)}\right)e^{-\frac{1}{2}k\lambda\left(t+\frac{1}{t}\right)}\,\mathrm{d}t && (4.79)\\
&= \frac{1}{km}e^{\frac{\lambda}{\mu}}\left(\int_0^\infty e^{-\frac{t}{km}-\frac{1}{2}k\lambda t-\frac{1}{2}k\lambda\frac{1}{t}}\,\mathrm{d}t+\frac{1}{2}e^{k\lambda}\int_0^\infty e^{-\frac{t}{km}-k\lambda t-k\lambda\frac{1}{t}}\,\mathrm{d}t\right) && (4.80)\\
&= \frac{2}{m}e^{\frac{\lambda}{\mu}}\sqrt{\frac{\lambda m}{2+k^2\lambda m}}\;{}_{1}\left(\sqrt{\frac{2\lambda}{m}+k^2\lambda^2}\right) \\
&\quad+\frac{1}{m}e^{\frac{\lambda}{\mu}+k\lambda}\sqrt{\frac{\lambda m}{1+k^2\lambda m}}\;{}_{1}\left(2\sqrt{\frac{\lambda}{m}+k^2\lambda^2}\right), && (4.81)
\end{aligned}$$

and therefore

$$\begin{aligned}
&\ge \log\frac{m}{\lambda}+\frac{\mu}{m}-\frac{\lambda}{\mu}+k\lambda+\frac{3}{2}\log\frac{\lambda}{\mu}+\frac{1}{2}\log\frac{e}{2\pi}-\frac{3}{2}e^{\frac{2\lambda}{\mu}}\operatorname{Ei}\left(-\frac{2\lambda}{\mu}\right) \\
&\quad-\log\left(1+\frac{1}{m}e^{\frac{\lambda}{\mu}}\sqrt{\frac{\lambda m}{2+k^2\lambda m}}\;{}_{1}\left(\sqrt{\frac{2\lambda}{m}+k^2\lambda^2}\right)\right. \\
&\qquad\left.+\frac{1}{2m}e^{\frac{\lambda}{\mu}+k\lambda}\sqrt{\frac{\lambda m}{1+k^2\lambda m}}\;{}_{1}\left(2\sqrt{\frac{\lambda}{m}+k^2\lambda^2}\right)\right). && (4.82)
\end{aligned}$$

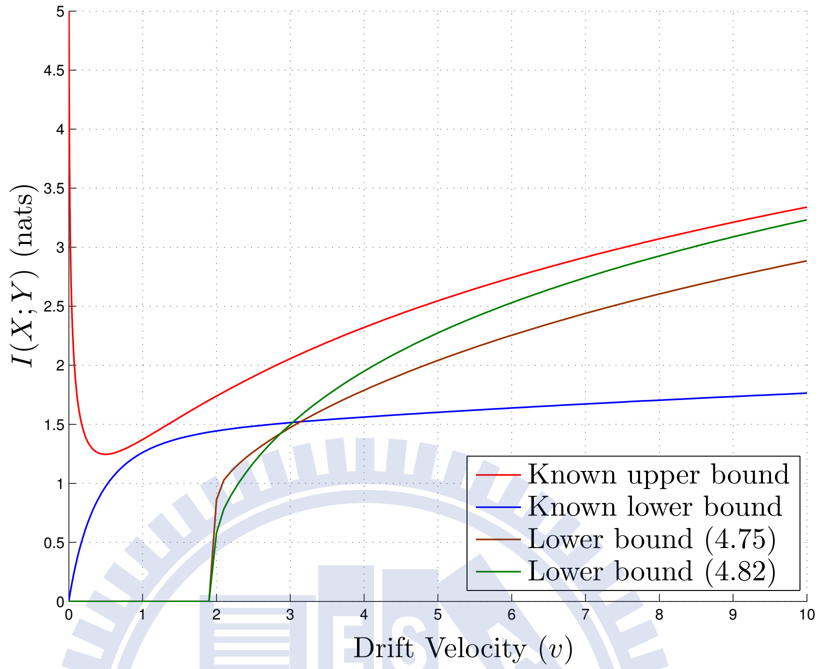


Figure 4.7: $m = 2$, $\sigma^2 = 1$, $d = 1$

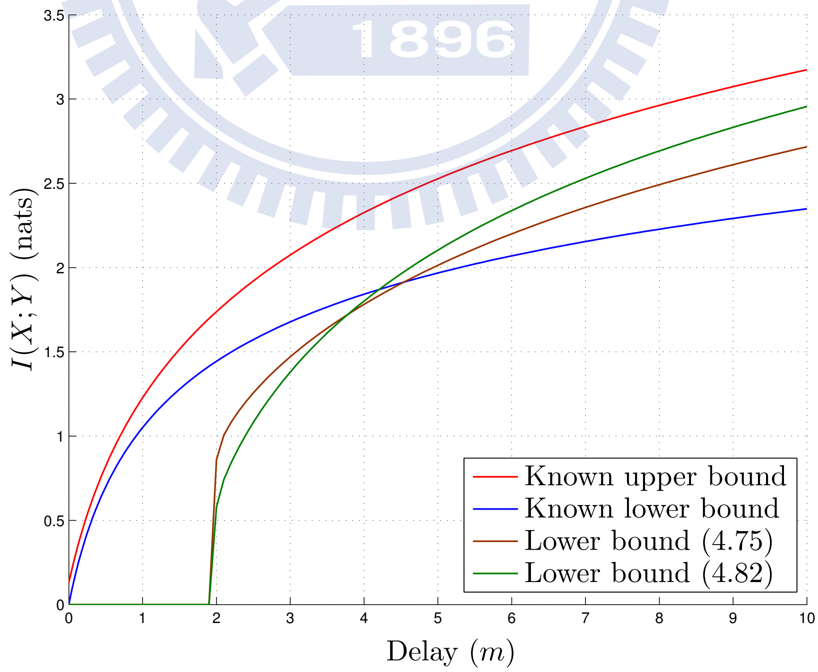


Figure 4.8: $v = 2$, $\sigma^2 = 1$, $d = 1$

This bound is shown in Figs. 4.7 and 4.8.

Chapter 5

Our Different Trials of Upper Bounds

In Chapter 6, we will show that the known upper bound proposed in [1] is quite tight in high drift velocity v and m. Therefore, the main goal in this chapter will focus on low v. Since we had a rather separated upper and lower bound in low v, we attempt to derive a better upper bound that behaves closer to the capacity. The method we use here is duality-based bound:

$$\leq \mathsf{E}_{Q^*}\left[D\big(W(\cdot|X)\big\|R(\cdot)\big)\right], \tag{5.1}$$

where $D(\cdot\|\cdot)$ is defined in (2.57), $W(\cdot|X)$ is the channel law, $R(\cdot)$ is the output distribution and Q^* is the capacity-achieving input distribution [11, Ch. 7]. The only thing which is known to us is the channel law $W(\cdot|X)$.

Note that the $R(\cdot)$ here can be any distribution on the output. It doesn't need to be an output distribution corresponding to a certain input distribution. This gives us quite a lot of freedom, but we still have three main tasks here to find a good or reasonable upper bound: First is making a clever choice of $R(\cdot)$ that helps us to get a better upper bound. Second is being able to analytically evaluate $D\big(W(\cdot|X)\big\|R(\cdot)\big)$, which is usually difficult. Third is evaluating or further upper-bounding the expectation over the capacity-achieving input on the right hand side of (5.1) without knowing Q^* generally. The third task could be solved by the properties of Q^*, e.g., the input moment constraint.

This chapter shows some of the trials we did by choosing $R(\cdot)$ to be exponential, inverse Gaussian, power inverse Gaussian and shifted Gamma distribution.

5.1 Exponential Distribution as Output Distribution

Based on (5.1), when we plug an exponential distribution $\text{Exp}(\beta)$ as output distribution, we get the following bound:

$$\leq \mathsf{E}_{Q^*}\left[D(f_{Y|X}(\cdot|X)\|f_Y(\cdot))\right] \tag{5.2}$$

$$= \mathsf{E}_{Q^*}\left[-h(Y|X=x) - \mathsf{E}_{Y|X}[\log f_Y(\cdot)]\right] \tag{5.3}$$
$$= -h(N) + \log\beta + \frac{\mathsf{E}[Y]}{\beta} \tag{5.4}$$
$$= -h(N) + \log\beta + \frac{\mathsf{E}[X+N]}{\beta} \tag{5.5}$$
$$\leq -h(N) + \log\beta + \frac{m+\mu}{\beta} \tag{5.6}$$
$$\triangleq q(\mu,\lambda,\beta) \tag{5.7}$$

We optimize over β

$$\frac{\partial q(\mu,\lambda,\beta)}{\partial\beta} = \frac{1}{\beta} - \frac{m+\mu}{\beta^2} \tag{5.8}$$
$$\overset{!}{=} 0 \tag{5.9}$$

and get the optimal result:

$$\beta^* = m+\mu, \tag{5.10}$$

which is exactly the upper bound in [1].

5.2 Inverse Gaussian Distribution as Output Distribution

In the communication environment described in Chapter 1, we have a mean constraint on the input delay, $\mathsf{E}[X] \leq m$. Therefore, another capacity upper bound is derived by choosing an output distribution as $\text{IG}(m+\mu,\beta)$ into the duality-based upper bound (5.1).

Lemma 5.1 (Upper bound with $\text{IG}(m+\mu,\beta)$ as the output distribution)**.**

$$\leq \frac{1}{2}\log\frac{2\pi}{\beta} + \frac{3}{2}\mathsf{E}[\log(X+N)] - \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{X+N}\right] - h(N). \tag{5.11}$$

Proof.

$$\leq \mathsf{E}_{Q^*}\left[D(f_{Y|X}(\cdot|X)\|f_Y(\cdot))\right] \tag{5.12}$$
$$= \mathsf{E}_{Q^*}\left[-h(Y|X=x) - \mathsf{E}_{Y|X}[\log f_Y(\cdot)]\right] \tag{5.13}$$
$$= -h(N) - \mathsf{E}_{Q^*}\left[\mathsf{E}_{Y|X}\left[\log\left(\sqrt{\frac{\beta}{2\pi Y^3}}\exp\left(-\frac{\beta(Y-m-\mu)^2}{2(m+\mu)^2 Y}\right)\right)\right]\right] \tag{5.14}$$
$$= -h(N) + \frac{1}{2}\log\frac{2\pi}{\beta} + \frac{3}{2}\mathsf{E}[\log Y] + \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{Y}\right] - \frac{\beta}{m+\mu} \tag{5.15}$$
$$= -h(N) + \frac{1}{2}\log\frac{2\pi}{\beta} + \frac{3}{2}\mathsf{E}[\log(X+N)] - \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{X+N}\right]. \tag{5.16}$$

□

There are two methods how we can further bound Lemma 5.1. First method, apply Jensen's inequality on $\mathsf{E}[\log(X+N)]$:

$$\leq -h(N) + \frac{1}{2}\log\frac{2\pi}{\beta} + \frac{3}{2}\log(m+\mu) - \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{X+N}\right] \tag{5.17}$$

$$\triangleq f(m,\mu,\lambda,\beta). \tag{5.18}$$

To optimize over β, we perform partial differential for $f(\cdot)$ over β:

$$\frac{\partial f(m,\mu,\lambda,\beta)}{\partial\beta} = -\frac{1}{2\beta} - \frac{1}{2(m+\mu)} + \frac{1}{2}\mathsf{E}\left[\frac{1}{X+N}\right] \tag{5.19}$$

$$\overset{!}{=} 0. \tag{5.20}$$

Solving equation (5.20), we get

$$\beta^* = \frac{m+\mu}{(m+\mu)\mathsf{E}\left[\frac{1}{X+N}\right] - 1}. \tag{5.21}$$

Hence, we plug the optimal β back in (5.17) and get:

$$\leq -h(N) + \frac{1}{2}\log\left(\frac{2\pi}{m+\mu}\left((m+\mu)\mathsf{E}\left[\frac{1}{X+N}\right] - 1\right)\right) + \frac{3}{2}\log(m+\mu) + \frac{1}{2} \tag{5.22}$$

$$= \frac{3}{2}\log\frac{m+\mu}{\mu} + \frac{1}{2}\log\lambda - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2}\log\left(\mathsf{E}\left[\frac{1}{X+N}\right] - \frac{1}{m+\mu}\right) \tag{5.23}$$

$$\leq \frac{3}{2}\log\frac{m+\mu}{\mu} + \frac{1}{2}\log\lambda - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2}\log\left(\mathsf{E}\left[\frac{1}{N}\right] - \frac{1}{m+\mu}\right) \tag{5.24}$$

$$= \frac{3}{2}\log\frac{m+\mu}{\mu} - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2}\log\left(1 + \frac{m\lambda}{\mu(m+\mu)}\right). \tag{5.25}$$

Here, (5.23) is derived by plugging in Proposition 2.1. Since X is nonnegative, dropping X results in (5.24). And in (5.25), we use (2.33).

In a second method, from Lemma 5.1, we do not apply Jensen's inequality on $\mathsf{E}[\log(X+N)]$ yet. Instead, we upper-bound it as follow:

$$\mathsf{E}[\log(N+X)] = \mathsf{E}[\log N] + \mathsf{E}_X\left[\mathsf{E}_{Y|X}\left[\log\left(1+\frac{x}{N}\right)\right]\middle|X=x\right] \tag{5.26}$$

$$\leq \mathsf{E}[\log N] + \mathsf{E}_X\left[\log\left(\mathsf{E}_{Y|X}\left[1+\frac{x}{N}\right]\right)\middle|X=x\right] \tag{5.27}$$

$$= \mathsf{E}[\log N] + \mathsf{E}_X\left[\log\left(1+X\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right)\middle|X=x\right] \tag{5.28}$$

$$\leq \mathsf{E}[\log N] + \log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) \tag{5.29}$$

$$= e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \log\mu + \log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right), \tag{5.30}$$

where (5.27) follows by Jensen's inequality, (5.28) applies (2.33) in Proposition 2.5, (5.29) follows from Jensen's inequality again together with $\mathsf{E}[X] \leq m$, and (5.30) simply plug in Proposition 2.7. This leads us to:

$$\leq -\frac{1}{2}\log\frac{2\pi\mu^3}{\lambda} - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) - \frac{1}{2} + \frac{1}{2}\log\frac{2\pi}{\beta} + \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right)$$

$$+\frac{3}{2}\log\mu + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) - \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{X+N}\right] \tag{5.31}$$

$$= \frac{1}{2}\log\frac{\lambda}{e\beta} + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) - \frac{\beta}{2(m+\mu)} + \frac{\beta}{2}\mathsf{E}\left[\frac{1}{X+N}\right] \tag{5.32}$$

$$\triangleq g(m,\mu,\lambda,\beta). \tag{5.33}$$

Here, the equation (5.31) follows from the entropy of IG in (2.2) and (5.30). To optimize over β, we perform partial differential for $g(\cdot)$ over β:

$$\frac{\partial g(m,\mu,\lambda,\beta)}{\partial\beta} = -\frac{1}{2\beta} - \frac{1}{2(m+\mu)} + \frac{1}{2}\mathsf{E}\left[\frac{1}{X+N}\right] \tag{5.34}$$

$$\stackrel{!}{=} 0. \tag{5.35}$$

Solving equation (5.35), we get

$$\beta'^{*} = \frac{m+\mu}{(m+\mu)\mathsf{E}\left[\frac{1}{X+N}\right] - 1}, \tag{5.36}$$

which is the same as (5.21)! We plug the optimal β' back in (5.32) and get:

$$\leq \frac{1}{2}\log\lambda + \frac{1}{2}\log\left(\mathsf{E}\left[\frac{1}{X+N}\right] - \frac{1}{m+\mu}\right) + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) \tag{5.37}$$

$$\leq \frac{1}{2}\log\lambda + \frac{1}{2}\log\left(\mathsf{E}\left[\frac{1}{N}\right] - \frac{1}{m+\mu}\right) + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) \tag{5.38}$$

$$= \frac{1}{2}\log\lambda + \frac{1}{2}\log\left(\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{m+\mu}\right) + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right) \tag{5.39}$$

$$= \frac{1}{2}\log\left(1+\frac{m\lambda}{\mu(m+\mu)}\right) + \frac{3}{2}\log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right). \tag{5.40}$$

We plot the lower bound according to drift velocity v and average-delay constraint m. As mentioned at the beginning of this chapter, the known upper bound is quite tight at the high m and v. Therefore, the figures below focus on the improvement in low v and m. From Figure 5.9, the known upper bound has a rising peak as v decreases, while our upper bound with IG output distribution has a decreasing tendency and will cross the known one. From Figure 5.10, our upper bound also crosses the known upper bound at low m.

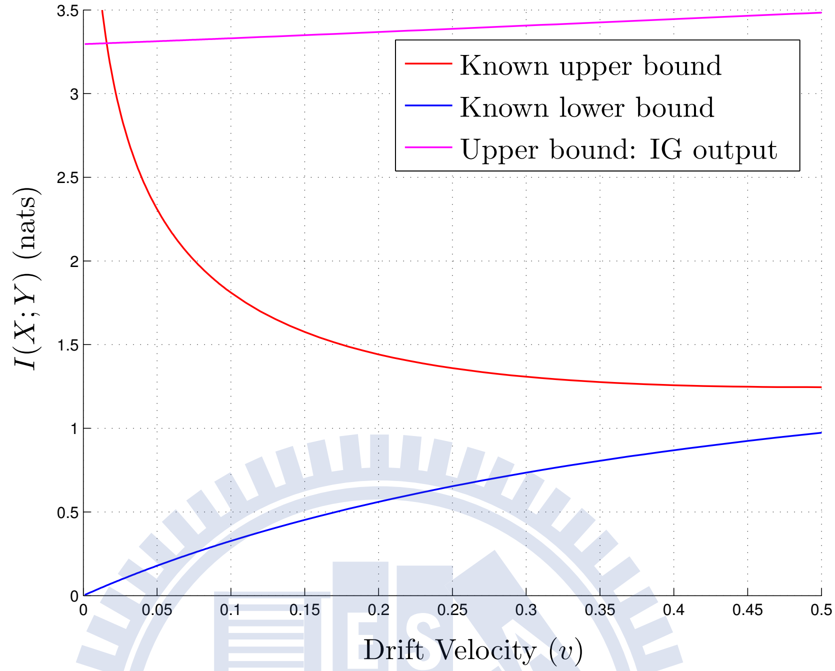


Figure 5.9: $m = 2$, $\sigma^2 = 1$, $d = 1$

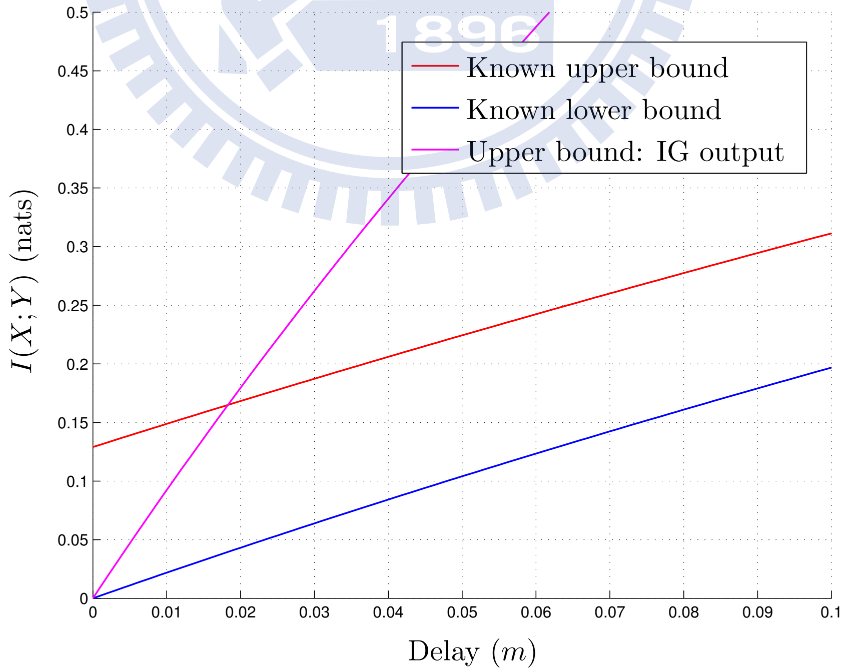


Figure 5.10: $v = 2$, $\sigma^2 = 1$, $d = 1$

5.3 Power Inverse Gaussian Distribution as Output Distribution

The PDF of the power inverse Gaussian distribution (PIG) is given by

$$f_Y(y) = \sqrt{\frac{\alpha}{2\pi\beta^3}} \left(\frac{\beta}{y}\right)^{\left(1+\frac{\eta}{2}\right)} \exp\left(-\frac{\alpha}{2\eta^2\beta}\left(\left(\frac{y}{\beta}\right)^{\frac{\eta}{2}} - \left(\frac{\beta}{y}\right)^{\frac{\eta}{2}}\right)^2\right), \tag{5.41}$$

where $y > 0$. The free parameters here are $\alpha, \beta > 0$, and $\eta \in \mathbb{R} \setminus \{0\}$. Using (5.1), we derive the capacity upper bound with output distribution PIG.

$$\begin{aligned}
&\leq -h(N) - \mathsf{E}_{Q^*}\left[\mathsf{E}_{Y|X}[\log(f_Y(Y))]\right] && (5.42)\\
&= -h(N) - \left(1+\frac{\eta}{2}\right)\log\beta + \left(1+\frac{\eta}{2}\right)\mathsf{E}_{Q^*}[\log(X+N)] - \frac{1}{2}\log\alpha + \frac{1}{2}\log 2\pi \\
&\quad + \frac{3}{2}\log\beta + \frac{\alpha}{2\eta^2\beta^{1+\eta}}\mathsf{E}_{Q^*}[(X+N)^\eta] + \frac{\alpha}{2\eta^2\beta^{1-\eta}}\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] - \frac{\alpha}{\eta^2\beta} && (5.43)\\
&\triangleq g(\alpha,\beta,\eta). && (5.44)
\end{aligned}$$

We first optimize over α before doing any further bounding on the expectation over Q^*.

$$\begin{aligned}
\frac{\partial g(\alpha,\beta,\eta)}{\partial\alpha} &= -\frac{1}{2\alpha} + \frac{1}{2\eta^2\beta}\left(\beta^{-\eta}\mathsf{E}_{Q^*}[(X+N)^\eta] + \beta^\eta\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] - 2\right) && (5.45)\\
&\stackrel{!}{=} 0. && (5.46)
\end{aligned}$$

We solve the optimal α,

$$\alpha^* = \frac{\eta^2\beta}{\beta^{-\eta}\mathsf{E}_{Q^*}[(X+N)^\eta] + \beta^\eta\mathsf{E}_{Q^*}[(X+N)^{-\eta}] - 2} \tag{5.47}$$

and plug it back to (5.43):

$$\begin{aligned}
&\leq -h(N) - \left(1+\frac{\eta}{2}\right)\log\beta + \left(1+\frac{\eta}{2}\right)\mathsf{E}_{Q^*}[\log(X+N)] + \frac{1}{2}\log 2\pi + \frac{1}{2} + \log\beta \\
&\quad - \log|\eta| + \frac{1}{2}\log\left(\beta^{-\eta}\mathsf{E}_{Q^*}[(X+N)^\eta] + \beta^\eta\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] - 2\right) && (5.48)\\
&\triangleq b(\alpha,\beta,\eta). && (5.49)
\end{aligned}$$

Then, we continue with optimizing over β:

$$\begin{aligned}
\frac{\partial b(\beta,\eta)}{\partial\beta} &= -\frac{\eta}{2\beta} + \frac{-\eta\beta^{-\eta-1}\mathsf{E}[(X+N)^\eta] + \eta\beta^{\eta-1}\mathsf{E}[(X+N)^{-\eta}]}{2\left(\beta^{-\eta}\mathsf{E}_{Q^*}[(X+N)^\eta] + \beta^\eta\mathsf{E}_{Q^*}[(X+N)^{-\eta}] - 2\right)} && (5.50)\\
&\stackrel{!}{=} 0. && (5.51)
\end{aligned}$$

We solve the optimal β,

$$\beta^* = \mathsf{E}[(X+N)^\eta]^{\frac{1}{\eta}} \tag{5.52}$$

and plug it back to (5.48):

$$\leq -h(N) - \frac{1}{2}\log \mathsf{E}_{Q^*}[(X+N)^\eta] + \left(1+\frac{\eta}{2}\right)\mathsf{E}_{Q^*}[\log(X+N)] + \frac{1}{2}\log 2\pi + \frac{1}{2}$$
$$-\log|\eta| + \frac{1}{2}\log\left(\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\mathsf{E}_{Q^*}[(X+N)^\eta]\right] - 1\right) \tag{5.53}$$
$$= -h(N) + \left(1+\frac{\eta}{2}\right)\mathsf{E}_{Q^*}[\log(X+N)] + \frac{1}{2}\log 2\pi + \frac{1}{2} - \log|\eta|$$
$$+\frac{1}{2}\log\left(\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] - \frac{1}{\mathsf{E}_{Q^*}[(X+N)^\eta]}\right). \tag{5.54}$$

Firstly, we upper-bound

$$\left(1+\frac{\eta}{2}\right)\mathsf{E}_{Q^*}[\log(X+N)]$$
$$= \left(1+\frac{\eta}{2}\right)\left(\mathsf{E}[\log N] + \mathsf{E}_{Q^*}\left[\log\left(1+\frac{X}{N}\right)\right]\right) \tag{5.55}$$
$$\leq \left(1+\frac{\eta}{2}\right)\left(\mathsf{E}[\log N] + \log\left(1+\mathsf{E}_{Q^*}\left[\frac{X}{N}\right]\right)\right) \tag{5.56}$$
$$\leq \left(1+\frac{\eta}{2}\right)\left(\mathsf{E}[\log N] + \log\left(1+m\mathsf{E}\left[\frac{1}{N}\right]\right)\right) \tag{5.57}$$
$$= \left(1+\frac{\eta}{2}\right)\left(e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \log\mu + \log\left(1+m\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)\right)\right) \tag{5.58}$$

for $\eta \geq -2$. Here we apply Jensen's inequality in both (5.55) and (5.56). The mean constraint on the input $\mathsf{E}[X] \leq m$ is applied to get (5.57). With the help of Proposition 2.5, we have (5.58). Then we derive

$$\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] \leq \mathsf{E}\left[N^{-\eta}\right] \tag{5.59}$$
$$= \sqrt{\frac{2\lambda}{\pi}}e^{\frac{\lambda}{\mu}}\mu^{-\eta-\frac{1}{2}}K_{\eta+\frac{1}{2}}\left(\frac{\lambda}{\mu}\right) \tag{5.60}$$

for $0 \leq \eta$, and

$$\mathsf{E}_{Q^*}[(X+N)^\eta] \leq (\mathsf{E}_{Q^*}[X+N])^\eta \tag{5.61}$$
$$\leq (m+\mu)^\eta \tag{5.62}$$

for $0 \leq \eta \leq 1$. Here, (5.59) follows from X is non-negative, (5.60) follows from Proposition 2.5, (5.61) follows from Jensen's inequality and (5.62) follows from the mean constraint on the input $\mathsf{E}[X] \leq m$. As a result, we derive the capacity upper bound as follow:

$$\leq \frac{1}{2}\log\lambda + \frac{\eta-1}{2}\log\mu + \frac{\eta-1}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \left(1+\frac{\eta}{2}\right)\log\left(1+m\left(\frac{1}{\lambda}+\frac{1}{\mu}\right)\right)$$
$$-\log|\eta| + \frac{1}{2}\log\left(\mathsf{E}_{Q^*}\left[(X+N)^{-\eta}\right] - \frac{1}{\mathsf{E}_{Q^*}[(X+N)^\eta]}\right) \tag{5.63}$$
$$\leq \frac{1}{2}\log\lambda - \log\mu - e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{1}{2}\log\left(1+m\left(\frac{1}{\lambda}+\frac{1}{\mu}\right)\right)$$
$$+\frac{1}{2}\log\left(m+\mu-\frac{\mu\lambda}{\mu+\lambda}\right). \tag{5.64}$$

This upper bound is depicted in Figure 5.11 and Figure 5.12.

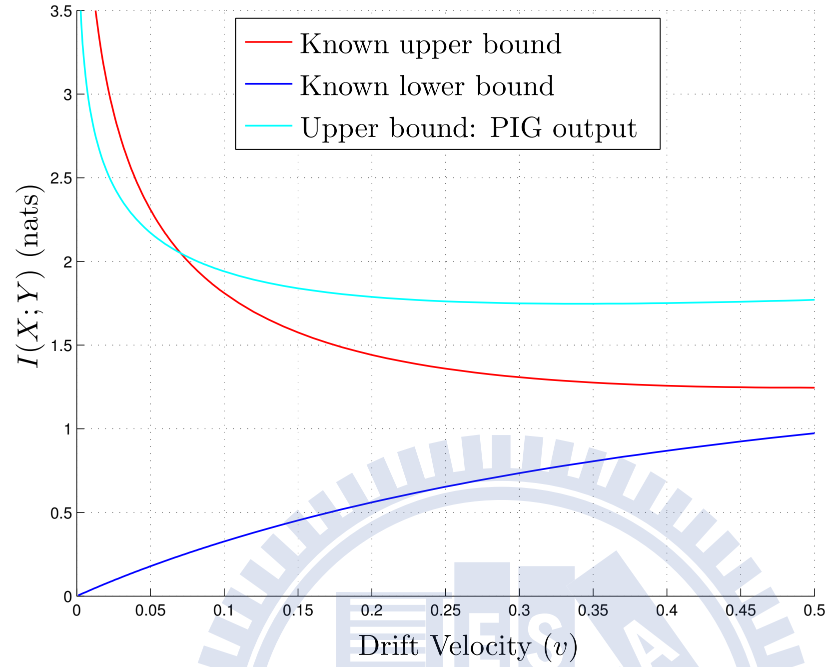


Figure 5.11: $m = 2$, $\sigma^2 = 1$, $d = 1$

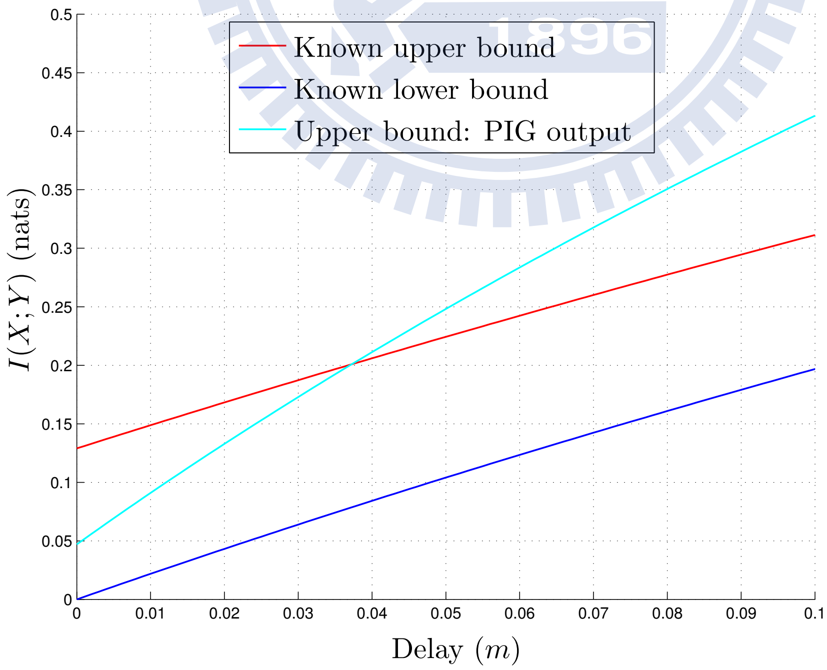


Figure 5.12: $v = 2$, $\sigma^2 = 1$, $d = 1$

5.4 Shifted Gamma Distribution as Output Distribution

The shifted Gamma distribution is actually a generalization of Gamma distribution and exponential distribution. We shift the gamma function left by δ and get the shifted gamma function as follow:

$$R_{\mathrm{SG}}(y) = \frac{(y+\delta)^{\alpha-1} e^{-\frac{y+\delta}{\beta}}}{\beta^{\alpha} \cdot \Gamma(\alpha, \frac{\delta}{\beta})}, \tag{5.65}$$

where $y \geq 0$, $\alpha > 0$, $\beta > 0$, $\delta \geq 0$. And gamma function is defined below:

$$\Gamma(\eta, \xi) \triangleq \int_{\xi}^{\infty} t^{\eta-1} e^{-t} \, \mathrm{d}t \tag{5.66}$$

where $\eta > 0$ and $\xi \geq 0$. When there is no shift, that means $\delta = 0$, we have an Gamma distribution:

$$R_{\mathrm{G}}(y) = \frac{y^{\alpha-1} e^{-\frac{y}{\beta}}}{\beta^{\alpha} \cdot \Gamma(\alpha, 0)}. \tag{5.67}$$

Moreover, when $\alpha = 1$, we have an exponential distribution with parameter $1/\beta$:

$$R_{\exp}(y) = \frac{1}{\beta} e^{-\frac{1}{\beta} y}. \tag{5.68}$$

Here we simply plug (5.65) as output distribution and the channel law

$$W(y|x) = \sqrt{\frac{\lambda}{2\pi(y-x)^3}} \exp\left(-\frac{\lambda(y-x-\mu)^2}{2\mu^2(y-x)}\right), \quad y > x, \tag{5.69}$$

in capacity upper bound (5.1). We get the following:

$$\leq -h(N) - \mathsf{E}_{X^*}\left[\mathsf{E}_{Y|X}\left[\log\left(\frac{(y+\delta)^{\alpha-1} e^{-\frac{y+\delta}{\beta}}}{\beta^{\alpha} \cdot \Gamma(\alpha, \frac{\delta}{\beta})}\right)\right]\right] \tag{5.70}$$

$$= -h(N) + (1-\alpha)\mathsf{E}[\log(N+X+\delta)] + \frac{\mathsf{E}[N+X+\delta]}{\beta} + \alpha \log \beta + \log \Gamma\left(\alpha, \frac{\delta}{\beta}\right) \tag{5.71}$$

$$\leq -h(N) + (1-\alpha)\mathsf{E}[\log(N+X+\delta)] + \frac{\mu+m+\delta}{\beta} + \alpha \log \beta + \log \Gamma\left(\alpha, \frac{\delta}{\beta}\right) \tag{5.72}$$

- For $\alpha \geq 1$: $\mathsf{E}[\log(N+X+\delta)] \geq \mathsf{E}[\log N]$

$$\leq -h(N) + (1-\alpha)\log \mu + (1-\alpha) e^{\frac{2\lambda}{\mu}} \operatorname{Ei}\left(-\frac{2\lambda}{\mu}\right) + \frac{\mu+m+\delta}{\beta} + \alpha \log \beta + \log \Gamma\left(\alpha, \frac{\delta}{\beta}\right). \tag{5.73}$$

- For $\alpha \geq 1$: $\mathsf{E}[\log(N + X + \delta)] \geq \log \delta$

$$\leq -h(N) + (1-\alpha)\log\delta + \frac{\mu + m + \delta}{\beta} + \alpha\log\beta + \log\Gamma\left(\alpha, \frac{\delta}{\beta}\right). \quad (5.74)$$

- For $0 < \alpha \leq 1$: applying Jensen's inequality $\mathsf{E}[\log(N + X + \delta)] \leq \log(\mu + m + \delta)$

$$\leq -h(N) + (1-\alpha)\log(\mu + m + \delta) + \frac{\mu + m + \delta}{\beta} + \alpha\log\beta + \log\Gamma\left(\alpha, \frac{\delta}{\beta}\right). \quad (5.75)$$

We optimize over α, β and δ for these three different boundings and get Figure 5.13. There we can see that the optimized (5.73), (5.74) and (5.75) are the same as known upper bound (3.16). This is because the shifted Gamma distribution contains exponential distribution as a special case.

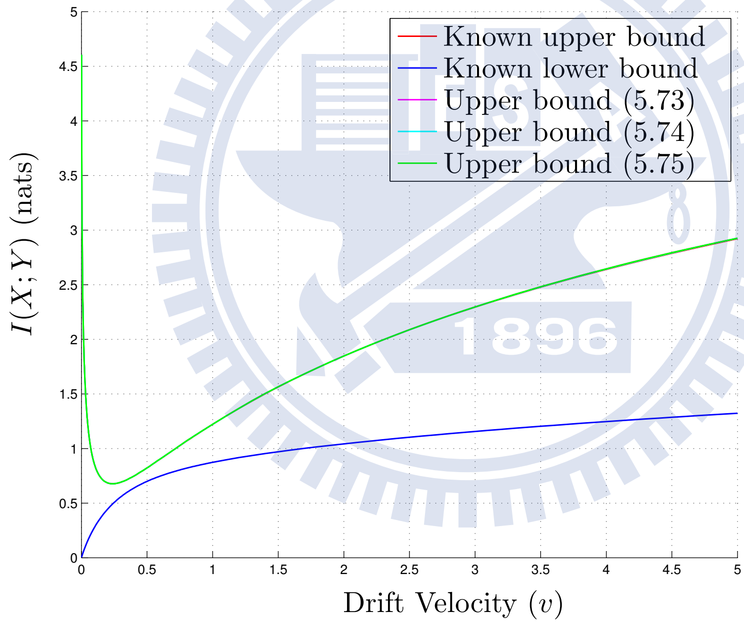


Figure 5.13: $m = 2$, $\sigma^2 = 1$, $d = 1$

Chapter 6

Asymptotic Capacity of AIGN Channel

In this chapter, we try to figure out how the capacity behaves when the drift velocity v and the average-delay constraint m tend to infinity.

6.1 When v Large

We first pick the known upper bound introduced in Chapter 3 since it is asymptotically tight in v:

$$
\begin{aligned}
(v) &\leq 1 + \log(m+\mu) - h_{\mathrm{IG}(\mu,\lambda)} && (6.1)\\
&= 1 + \log(m+\mu) - \frac{1}{2}\log\frac{2\pi\sigma^2 d}{v^3} - \frac{3}{2}\exp\left(\frac{2dv}{\sigma^2}\right)\mathrm{Ei}\left(-\frac{2dv}{\sigma^2}\right) - \frac{1}{2} && (6.2)\\
&< \frac{1}{2} + \log\left(m + \frac{d}{v}\right) - \frac{1}{2}\log 2\pi\sigma^2 d + \frac{3}{2}\log v + \frac{3}{2}\log\left(1 + \frac{\sigma^2}{2dv}\right) && (6.3)\\
&= \log\left(m + \frac{d}{v}\right) + \frac{1}{2}\log\frac{e}{2\pi\sigma^2 d} + \frac{3}{2}\log v - \frac{3}{2}\log\left(1 + \frac{\sigma^2}{2dv}\right), && (6.4)
\end{aligned}
$$

where (6.3) is simply plugging in the upper bound of $\mathrm{Ei}(\cdot)$ from Proposition 2.18. Its asymptotic upper bound is:

$$
(v) \leq \frac{3}{2}\log v + \frac{1}{2}\log\frac{\lambda m^2 e}{2\pi} + o(1). \quad (6.5)
$$

On the other hand, we pick the lower bound in Section 4.3:

$$
\begin{aligned}
(v) \geq{}& \log\frac{m}{\lambda} + \frac{1}{mv} - \lambda v + k\lambda + \frac{3}{2}\log\lambda v + \frac{1}{2}\log\frac{e}{2\pi} - \frac{3}{2}e^{2\lambda v}\,\mathrm{Ei}(-2\lambda v)\\
&- \log\left(1 + \frac{1}{m}e^{\lambda v}\sqrt{\frac{\lambda m}{2 + k^2\lambda m}}\;{}_1\left(\sqrt{\frac{2\lambda}{m} + k^2\lambda^2}\right)\right.\\
&\left.\qquad + \frac{1}{2m}e^{\lambda v + k\lambda}\sqrt{\frac{\lambda m}{1 + k^2\lambda m}}\;{}_1\left(2\sqrt{\frac{\lambda}{m} + k^2\lambda^2}\right)\right), \quad (6.6)
\end{aligned}
$$

where $k = \sqrt{v^2 - \frac{2\sigma^2}{m}}$. When v goes to infinity, the result is as follow:

$$(v) \geq \frac{3}{2}\log v + \frac{1}{2}\log\frac{\lambda m^2 e}{2\pi} + o(1). \tag{6.7}$$

From (6.5) and (6.7), we can observe that the upper and the lower bound coincide, which proves the following result:

Theorem 6.1 (Asymptotic Capacity of AIGN Channel with v Large)**.** *The capacity of the AIGN channel defined in Section 1.1 and 1.2 is asymptotically, when the drift velocity v of the fluid medium tends to infinity while all other parameters are kept constant, as follow:*

$$\lim_{v\uparrow\infty}\left\{ (v) - \frac{3}{2}\log v\right\} = \frac{1}{2}\log\frac{\lambda m^2 e}{2\pi}. \tag{6.8}$$

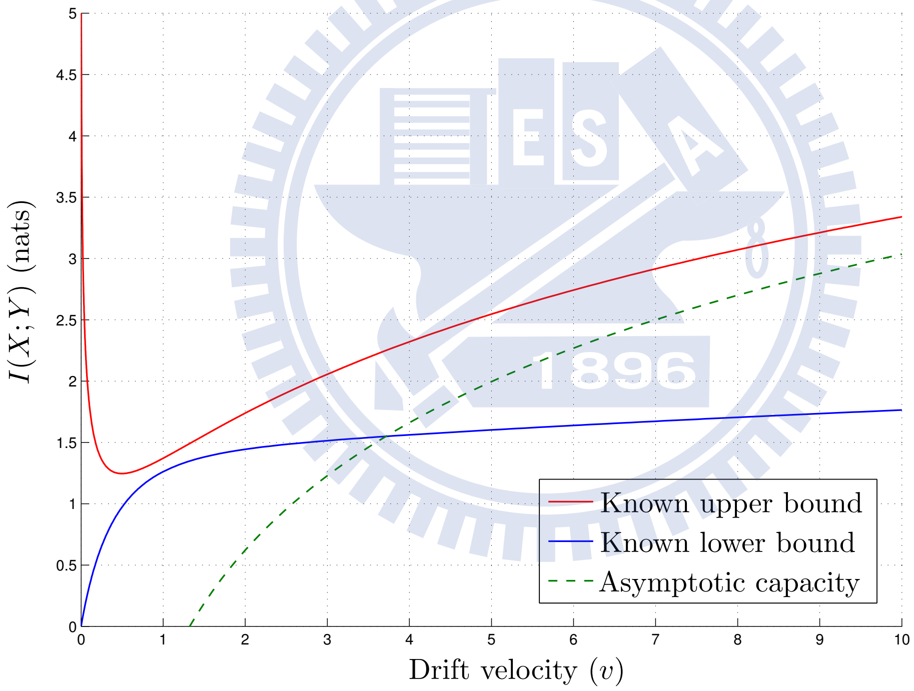


Figure 6.14: Setting parameter: $m = 2$, $d = 1$ and $\sigma^2 = 1$.

6.2 When m Large

First, we use the known upper bound introduced in Chapter 3 since it is also asymptotically tight in m:

$$(m) \leq 1 + \log(m + \mu) - h_{IG(\mu,\lambda)} \tag{6.9}$$

$$= 1 + \log m + \log\left(1 + \frac{\mu}{m}\right) - h_{IG(\mu,\lambda)}, \tag{6.10}$$

where $h(N)$ is independent of m. When m goes to infinity, the upper bound becomes:

$$(m) \leq \log m + \frac{1}{2}\log\frac{e\lambda}{2\pi\mu^3} - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + o(1) \tag{6.11}$$

On the other hand, we also pick the lower bound in Section 4.3.

$$\begin{aligned}(m) \geq{}& \log\frac{m}{\lambda} + \frac{\mu}{m} - \frac{\lambda}{\mu} + k\lambda + \frac{3}{2}\log\frac{\lambda}{\mu} + \frac{1}{2}\log\frac{e}{2\pi} - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right)\\ &- \log\left(1 + \frac{1}{m}e^{\frac{\lambda}{\mu}}\sqrt{\frac{\lambda m}{2+k^2\lambda m}}\;{}_1\left(\sqrt{\frac{2\lambda}{m}+k^2\lambda^2}\right)\right.\\ &\left.+ \frac{1}{2m}e^{\frac{\lambda}{\mu}+k\lambda}\sqrt{\frac{\lambda m}{1+k^2\lambda m}}\;{}_1\left(2\sqrt{\frac{\lambda}{m}+k^2\lambda^2}\right)\right) &(6.12)\\ ={}& \log m + \frac{1}{2}\log\frac{e\lambda}{2\pi\mu^3} - \frac{3}{2}e^{\frac{2\lambda}{\mu}}\,\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right) + o(1) &(6.13)\end{aligned}$$

where $k = \sqrt{\frac{1}{\mu^2} - \frac{2}{\lambda m}}$. Here, (6.13) is because the last term of (6.12) goes to 0 once we let m goes to infinity. From (6.11) and (6.13), we can observe that the upper and the lower bound coincide, which proves the following result.

Theorem 6.2 (Asymptotic Capacity of AIGN Channel with m Large)**.** *The capacity of the AIGN channel defined in Section 1.1 and 1.2 is asymptotically, when the average-delay constraint m is loosened to infinity while all other parameters are kept constant, as follow:*

$$\begin{aligned}\lim_{m\uparrow\infty}\{\ (m) - \log m\} &= 1 - h_{IG(\mu,\lambda)} &(6.14)\\ &= \frac{1}{2}\log\frac{\lambda e}{2\pi\mu^3} - \frac{3}{2}\exp\left(\frac{2\lambda}{\mu}\right)\mathrm{Ei}\left(-\frac{2\lambda}{\mu}\right). &(6.15)\end{aligned}$$

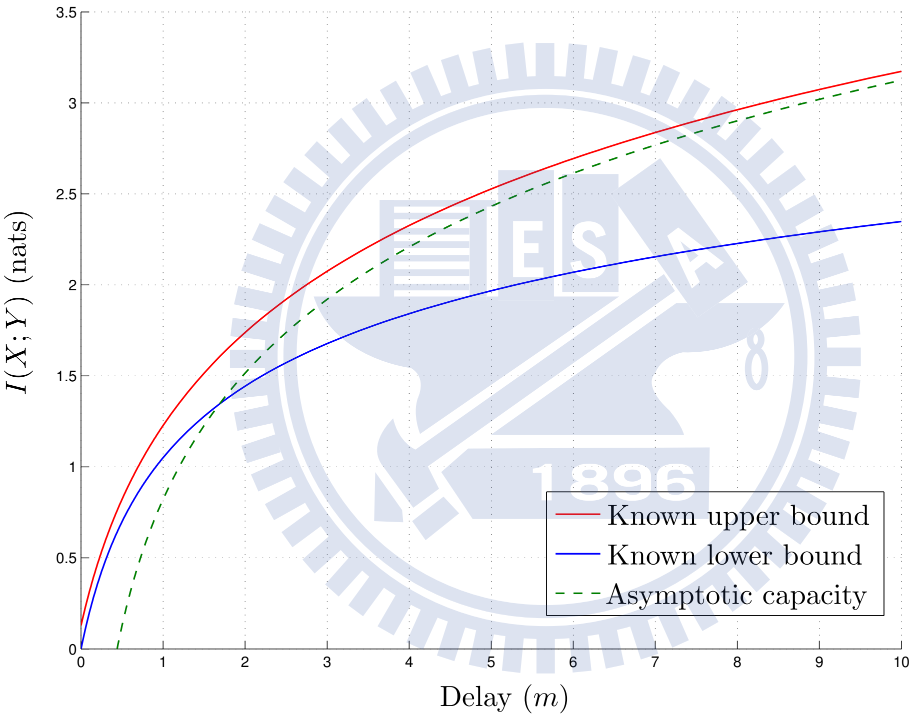


Figure 6.15: Setting parameter: $m = 2$, $d = 1$ and $\sigma^2 = 1$.

Chapter 7

Discussion and Conclusion

In this thesis, a new type of channel, the additive inverse Gaussian noise channel, has been investigated. We introduced its interesting properties and related lemmas. Several methods of upper-bounding and lower-bounding its channel capacity are provided.

We have found out that the upper bounds (3.15) and (3.16) from literature [1] are very tight by providing analytical lower bound that is tight in the asymptotic regime. Therefore, we focused on the low drift velocity v and low average-delay m regime. Note that (3.16) has a strange increasing behavior when $v \to 0$. Here we provided an upper bound (5.40) that is better than (3.16) in low v. Moreover, (3.15) does not tend to 0 as $m \to 0$, while we can show an improved upper bound (5.40) that does tends to 0.

The lower bounds (3.15) and (3.14) in [1], are not tight enough in both high v and high m. With the help of [10], we were able to compute the exact output distribution of an exponential input. This lower bound (4.82) was much tighter than the known bound with respect to both v and m. It turned out that together with the known upper bound, this lower bound allowed us to derive the asymptotic capacity at high v (6.8) and m (6.15).

For future research, we propose the following problems related to the additive inverse Gaussian noise channel:

- Derivation of the exact slope of the asymptotic capacity when $m \to 0$.
- Derivation of the channel capacity behavior for $v \to 0$.
- Inclusion of a peak-delay constraint to the system.
- Extension to nonadditive channels.

Bibliography

[1] K. V. Srinivas, Raviraj S. Adve, and Andrew W. Eckford, "Molecular communication in fluid media: The additive inverse Gaussian noise channel," December 2010, arXiv:1012.0081v2 [cs.IT]. [Online]. Available: http://arxiv.org/abs/1012.0081v2

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

[3] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed., Alan Jeffrey, Ed. San Diego: Academic Press, 2000.

[4] Raj S. Chhikara and J. Leroy Folks, *The Inverse Gaussian Distribution — Theory, Methodology, and Applications.* New York: Marcel Dekker, Inc., 1989.

[5] V. Seshadri, *The Inverse Gaussian Distribution — A Case Study in Exponential Families.* Oxford: Clarendon Press, 1993.

[6] Toshihiko Kawamura and Kosei Iwase, "Characterizations of the distributions of power inverse Gaussian and others based on the entropy maximization principle," *Journal of the Japan Statistical Society*, vol. 33, no. 1, pp. 95–104, January 2003.

[7] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, 2nd ed. New York: John Wiley & Sons, 2006.

[8] Norman L. Johnson and N. Balakrishnan, *Advances in the Theory and Practice of Statistics.* New York: John Wiley & Sons, 1997.

[9] 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

[10] Wolf Schwarz, "On the convolution of inverse Gaussian and exponential random variables," *Communications in Statistics — Theory and Methods*, vol. 31, no. 12, pp. 2113–2121, 2002.

[11] Amos Lapidoth, Stefan M. Moser, and Michèle Wigger, *Advanced Optical Wireless Communication Systems.* Cambridge: Cambridge University Press, May 2012, ch. Channel capacity.