Genie-based Outer Bounds for Interference Channels

LIU, Sida

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Information Engineering

The Chinese University of Hong Kong
September 2016

Abstract of thesis entitled:

Genie-based Outer Bounds for Interference Channels

Submitted by LIU, Sida

for the degree of Doctor of Philosophy

at The Chinese University of Hong Kong in September 2016

In multi-user information theory, the interference channel is a classical model for communication between two or more transmitter-receiver pairs over a shared medium. Determining the capacity region of the interference channel remains a major open question in this field. Several inner and outer bounds have been proposed for the capacity region. Among those, Kramer (2004) developed a genie-based outer bound for the degraded Gaussian interference channel studied by Sato (1978). Later genie approaches are used to show the sum-capacity in a weak interference regime of the Gaussian interference channel. In this thesis, two outer bounds are developed, both for discrete and continuous settings, using the genie idea. The genie-based outer bound is shown to be sum-rate optimal for a specific class of discrete interference channels with low interference. In the Gaussian setting, one of the outer-bound developed in this thesis, *enhanced genie-based outer bound* turns out to be tight in all cases where sum-capacity has been previously established; thus unifying the converse arguments.

We study the optimality of Gaussian signaling using perturbations along Hermite polynomials, an idea introduced by Abbe and Zheng (2009). By generalizing the above approach we derive a larger regime under which Gaussian signaling is not optimal for the coding scheme of treating interference as noise.

This thesis also examines the weighted sum-rate for Gaussian Z interference channels. We present a conjecture that is equivalent to testing the optimality of Gaussian signaling with power control at the corner point of capacity region.

摘要

在多用户信息理論中，干擾信道是一個用來模擬兩對或更多傳輸接受端通過共享媒介進行通訊的古典模型。解決干擾信道的信道容量一直是這個領域中懸而未決的問題。幾種信道容量的內界和外界已經被提出和研究。在其中，Krammer（2004）推導出一個對於降級高斯干擾信道的基於精靈方法的外界。之後，精靈方法被用於證明高斯干擾信道的信道容量和。此篇論文中，兩種基於精靈方法的外界被推導出。這兩種精靈外界同時適用於離散和高斯的設定。在離散設定中，對於某一類離散干擾信道的信道容量和，第一種精靈外界是最優的。在高斯設定中，對於所有信道容量和已知的區間，另一種進階版精靈外界都是最優的，由此統一了信道容量和的反面證明。

我們用Hermite多項式微擾的方法研究了高斯信號的最優性。這個方法最初由Abbe和Zheng（2009）提出。對於干擾當作噪音的編碼方式，我們推廣了以上方法並且得到了一個更大的高斯信號非最優的區間。

此論文也研究了高斯Z型干擾信道的權重信道容量和。我們提出了一個猜想，這個猜想被證明是等價於在高斯Z型干擾信道中高斯信號與功率控制的最優性。

Acknowledgement

First and foremost I want to thank my advisor, Professor Chandra M. Nair. He has been a helpful mentor since I undertook a research project with him as an undergraduate. Throughout my Ph.D. pursuit, he has been supportive with his expertise in information theory, his insight in academic research and his patience in teaching me. I appreciate all his guidance in academic and beyond, his influence on my growth and the excellent model he provided for me to learn from.

I would also like to acknowledge Professor Max H. M. Costa, Professor Raymond Yeung, Professor Sidharth Jaggi and Professor Pascal Vontobel. It has been an honor to have these outstanding scholars on my thesis committee. Thanks to them for the time on the thesis review and comments to improve the thesis.

Many thanks goes to the former and current group members and fellow labmates in CUHK: Dr. Yanlin Geng, Dr. Fan Cheng, Dr. Chengwei Guo, Dr. Lingxiao Xia, Dr. Yang Liu, Mr. Chenglong Ma, Mr. Yuanming Yu, Mr. Mehdi Yazdanpanah, Mr. Yannan Wang and many other friends. Together with them, I had a lot of fun.

Finally, I would like to express my gratitude to my parents for their unconditional deep love.

This work is dedicated to my parents.

Contents

List of Figures

Chapter 1

Introduction

The ground-breaking work done by Claude E. Shannon in his paper "A Mathematical Theory of Communication" (1948) founded the discipline of information theory. Communication from one point to another was modeled as a three-stage process:

1. Encoding: There is a set of finitely many possible messages that may need to be sent. The encoding process maps each message to a codeword, a sequence of transmit symbols from a transmit alphabet.

2. Channel: This models the physical medium that corrupts the transmit symbol. The relationship between the received symbol and the transmitted symbol is often characterized by a probability transition matrix that yields the transition probabilities between output symbols and input symbols.

3. Decoding: This is the process of estimating the message from the sequence of received symbols.

Shannon's channel coding theorem has successfully quantified the maximum reliable rate of information flow through a channel, called channel capacity. The point to point communication model can be directly extended to a network setting. The first model in network information theory is the two-way channel studied by Shannon (1961).

During 1970s to 1980s, more channels were proposed and studied including the multiple access channel, the broadcast channel, and the interference channel. However determining the channel capacity region for most of these channels remains open. After decades when researchers had little interest in this field, network information theory was revived since 1990s thanks to development in wireless technology and advance of data processing ability.

This thesis focuses on the interference channel.

1.1 Discrete memoryless interference channel

The interference channel was first introduced by Ahlswede (1974). It is a classical model for communication consisting of two pairs of transmitters and receivers over a shared medium. Each receiver wants to send a private message to its intended receiver; however the sharing of the medium causes it to suffer interference from the other communication pair. The characterization of the capacity region is a classical and fundamental open problem in the area of multi-terminal information theory.

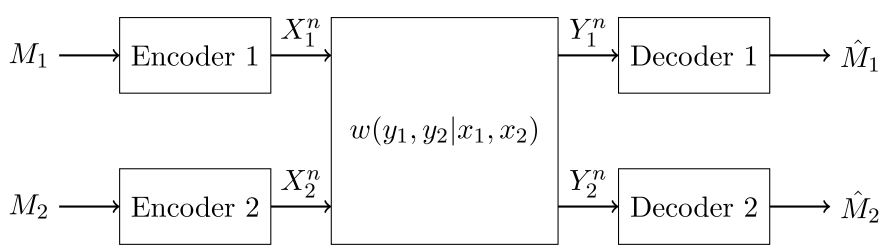

Figure 1.1: Discrete memoryless interference channel

Consider a discrete memoryless interference channel (DM-IC) depicted in Figure 1.1. The input and output alphabet are over two finite sets $\mathcal{X}$, $\mathcal{Y}$. Interference and noise are characterized by the transition probability $w(y_1, y_2|x_1, x_2)$. An (R_1, R_2, n) rate coding scheme for the discrete memoryless interference channel consists of

- Two message sets $\{1, 2, \ldots, \lfloor 2^{nR_i} \rfloor\}$, $i = 1, 2$. The messages are assumed to be independent of each other and uniformly distributed over their message sets.

- Two encoders: Each encoder maps a message $M_i \in \{1, 2, \ldots, \lfloor 2^{nR_i} \rfloor\}$ to X_i^n using an encoding function $\Psi_i : \{1, 2, \ldots, \lfloor 2^{nR_i} \rfloor\} \mapsto \mathcal{X}_i^n$, $i = 1, 2$.

- Two decoders: Each decoder maps received n-letter sequence Y_i^n to an estimate of the message $\hat{M}_i$ in $\{1, 2, \ldots, \lfloor 2^{nR_i} \rfloor\}$ using a decoding function, $\Phi_i : \mathcal{Y}_i^n \mapsto \{1, 2, \ldots, \lfloor 2^{nR_i} \rfloor\}$, $i = 1, 2$.

A rate pair (R_1, R_2) is said to be achievable if there is a sequence of (R_1, R_2, n) coding schemes such that error probability $P_e(n) := Pr\{(M_1, M_2) \neq (\hat{M}_1, \hat{M}_2)\} \to 0$ as $n \to \infty$. The *capacity region* C is the closure of the set of achievable rate pairs $(R_1, R_2) \in \mathbb{R}^2$. The sum-capacity is defined as $C_{sum} = \max_{(R_1, R_2) \in C} R_1 + R_2$. Note that receivers decode messages independently, which means that the capacity only depends on the marginals $w(y_1|x_1, x_2)$ and $w(y_2|x_1, x_2)$ rather than $w(y_1, y_2|x_1, x_2)$. It is also assumed that there is no feedback from the receivers to the transmitter or co-operation between the two transmitters.

A more detailed problem introduction and additional prior results on interference channels can be found in Chapter 6 [9].

1.1.1 Strong interference

Definition 1.1.1 ([3])**.** A DM-IC is said to have *very strong interference* if

$$I(X_1; Y_1|X_2) \leq I(X_1; Y_2) \tag{1.1}$$

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1) \tag{1.2}$$

for all $p_1(x_1)p_2(x_2)$.

This definition sheds some light on the intensity of interference: Consider equation (1.1). The left hand side is the rate that can be achieved for channel X_1 to Y_1 without interference from X_2. The right hand side is the rate that can be achieved for channel X_1 to Y_2 by treating the intended signal X_2 as noise. The inequality indicates that interference is so strong that decoding interference would be optimal (and can indeed shown to be the case [3]). This intuition will lead to definition of very weak interference later in the following chapter.

Definition 1.1.2 ([7])**.** A DM-IC is said to have *strong interference* if

$$I(X_1;Y_1|X_2) \le I(X_1;Y_2|X_2) \tag{1.3}$$

$$I(X_2;Y_2|X_1) \le I(X_2;Y_1|X_1) \tag{1.4}$$

for all $p(x_1)p(x_2)$.

It is clear that very strong interference channels also have strong interference since $I(X_1;Y_2|X_2) \ge I(X_1;Y_2)$ and $I(X_2;Y_1|X_1) \ge I(X_2;Y_1)$.

Theorem 1.1.1 (Sato (1978) [17], Costa, El Gamal (1987) [7])**.** The capacity region of the DM-IC with strong interference is the union of rate pairs (R_1, R_2) such that

$$R_1 \le I(X_1;Y_1|X_2,Q) \tag{1.5}$$

$$R_2 \le I(X_2;Y_2|X_1,Q) \tag{1.6}$$

$$R_1 + R_2 \le \min\{I(X_1,X_2;Y_1|Q), I(X_1,X_2;Y_2|Q)\} \tag{1.7}$$

for some $p(q)p(x_1|q)p(x_2|q)$ with $|\mathcal{Q}| \le 4$.

For each distribution, the above constraints give a pentagonal region in $\mathbb{R}^2$. The capacity region is given by the union of these pentagons. The auxiliary random variable Q is a time/frequency sharing random variable to mix different strategies. Since Q has a cardinality bound, this characterization of the capacity region is computable by searching over a finite dimensional space. The optimal achievable coding scheme is *simultaneous-nonunique-decoding* and the converse is given by traditional single-letter argument using strong interference condition [7].

From this theorem, the sum-capacity for strong interference channel is obtained as follows

$$C_{sum} = \max_{p(q)p(x_1|q)p(x_2|q)} \min\{I(X_1,X_2;Y_1|Q), I(X_1,X_2;Y_2|Q)\}$$

1.2 Gaussian interference channel

The Gaussian interference channel (GIC) model in Figure 1.2 is widely used in wireless communications. The Gaussian interference channel with outputs Y_1, Y_2 and inputs X_1, X_2, $i = 1, 2$ are given by

$$Y_1 = X_1 + bX_2 + Z_1$$

$$Y_2 = X_2 + aX_1 + Z_2$$

where Z_1 and Z_2, used to model channel noise, are normally distributed random variables with mean 0 and variance 1, denoted as $\mathcal{N}(0, 1)$. Note that one can assume Z_1, Z_2 to have arbitrary correlation since the capacity only depends on the marginal distribution.

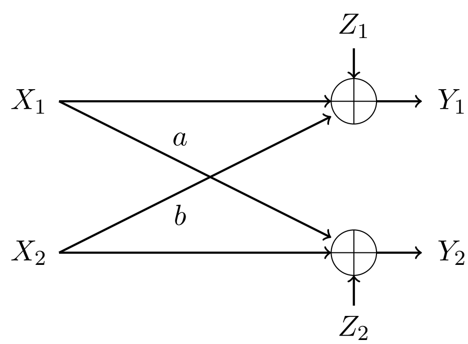

Figure 1.2: Gaussian interference channel

The input and output alphabets are assumed to be real numbers. The capacity of the Gaussian interference channel is often studied under the assumption that the input codewords satisfy an average power constraint, i.e.

$$\frac{1}{2^{nR_i}} \sum_{m=1}^{2^{nR_i}} \|\Psi_i(m)\|^2 \leq nP_i, \; i = 1, 2.$$

where m is the message to be send and Ψ_i is the encoding function.

For $a \geq 1$ and $b \geq 1$, the GIC satisfies strong interference condition (1.3) and (1.4). Hence, from Theorem 1.1.1, the capacity is simultaneous-nonunique-decoding region and it is not hard to show that the optimal input distribution is $X_i \sim \mathcal{N}(0, P_i)$, $i = 1, 2$ and $Q = \emptyset$.

For $a \geq 1$ and $b < 1$ (or $a < 1$ and $b \geq 1$), the sum-capacity of the GIC can be inferred from [17]

$$\min\left\{\frac{1}{2}\log(1 + a^2 P_1 + P_2), \frac{1}{2}\log(1 + \frac{P_1}{b^2 P_2 + 1}) + \frac{1}{2}\log(1 + P_2)\right\}$$

and the optimal input distribution is $X_i \sim \mathcal{N}(0, P_i)$, $i = 1, 2$ and $Q = \emptyset$.

For $a < 1$ and $b < 1$, and if in addition the following condition holds

$$a(1 + b^2 P_2) + b(1 + a^2 P_1) < 1, \tag{1.8}$$

then the sum-capacity of the GIC is given by

$$\frac{1}{2}\log(1 + \frac{P_1}{b^2 P_2 + 1}) + \frac{1}{2}\log(1 + \frac{P_2}{a^2 P_1 + 1}).$$

This is the rate obtained by the *treating-interference-as-noise* strategy with Gaussian inputs. This result was established independently by [18], [2], [13]; and uses a genie-based approach.

1.3 Han–Kobayashi inner bound

The best known inner bound is Han-Kobayashi inner bound.

Theorem 1.3.1 (Han–Kobayashi [11], [4])**.** A rate pair (R_1, R_2) is achievable for the DM-IC if

$$\begin{aligned}
R_1 &\leq I(X_1; Y_1 | U_2, Q) \\
R_2 &\leq I(X_2; Y_2 | U_1, Q) \\
R_1 + R_2 &\leq I(X_1, U_2; Y_1 | Q) + I(X_2; Y_2 | U_1, U_2, Q) \\
R_1 + R_2 &\leq I(X_2, U_1; Y_2 | Q) + I(X_1; Y_1 | U_1, U_2, Q) \\
R_1 + R_2 &\leq I(X_1, U_2; Y_1 | U_1, Q) + I(X_2, U_1; Y_2 | U_2, Q) \\
2R_1 + R_2 &\leq I(X_1, U_2; Y_1 | Q) + I(X_1; Y_1 | U_1, U_2, Q) + I(X_2, U_1; Y_2 | U_2, Q) \\
R_1 + 2R_2 &\leq I(X_2, U_1; Y_2 | Q) + I(X_2; Y_2 | U_1, U_2, Q) + I(X_1, U_2; Y_1 | U_1, Q)
\end{aligned}$$

for some $p(q)p(u_1, x_1 | q)p(u_2, x_2 | q)$.

In particular, by setting $U_1 = U_2 = \emptyset$, the Han–Kobayashi inner bound reduces to the *treating-interference-as-noise (TIN)* inner bound:

$$R_1 \leq I(X_1; Y_1|Q)$$
$$R_2 \leq I(X_2; Y_2|Q)$$

for some $p(q)p(x_1|q)p(x_2|q)$.

By setting $U_i = X_i,\ i = 1, 2$, the Han–Kobayashi inner bound reduces to the *simultaneous-nonunique-decoding* inner bound, which is tight for strong interference channels.

The Han–Kobayashi inner bound subsumes all known inner bounds and is optimal for some classes of channels such as the strong interference channel [17] and the injective deterministic interference channel [8]. However, a recent work [14] has shown that there are some DM-ICs for which the Han–Kobayashi inner bound is strictly sub-optimal.

1.4 Existing outer bounds

1.4.1 An outer bound using traditional techniques

Theorem 1.4.1 (Outer bound [12])**.** It can be shown that any achievable rate pair (R_1, R_2) must satisfy

$$\begin{aligned} R_1 &\leq \min\{I(U_2X_1; Y_1|Q), I(X_1; Y_1|X_2Q)\} \\ R_2 &\leq \min\{I(U_1X_2; Y_2|Q), I(X_2; Y_2|X_1Q)\} \\ R_1 + R_2 &\leq I(U_2X_1; Y_1|Q) + I(X_2; Y_2|U_2X_1Q) \\ R_1 + R_2 &\leq I(U_1X_2; Y_2|Q) + I(X_1; Y_1|U_1X_2Q), \end{aligned} \tag{1.9}$$

for some $p(q, u_1, u_2)p(x_1|u_1, u_2, q)p(x_2|u_1, u_2, q)$ such that the following statements hold:

1. X_1, X_2 are conditionally independent of Q,
2. For every $Q = q$, X_1 and X_2 are conditionally independent of U_1,
3. For every $Q = q$, X_1 and X_2 are conditionally independent of U_2,

4. $Q, U_1, U_2 \to (X_1, X_2) \to (Y_1, Y_2)$ forms a Markov chain.

This outer bound is tight for the sum-capacity of mixed Gaussian interference channels ($a > 1, b < 1$ or $a < 1, b > 1$). The proof [12] uses traditional techniques like Csiszar sum lemma and identification of auxiliaries $U_{1i} = (X_2^{n\setminus i}, Y_{11}^{i-1}, Y_{2i+1}^{n}), U_{2i} = (X_1^{n\setminus i}, Y_{11}^{i-1}, Y_{2i+1}^{n})$.

1.4.2 An outer bound for injective semi-deterministic interference channels

Figure 1.3 is a semi-deterministic interference channel. Fix $x_1 \in \mathcal{X}_1$, $y_1(x_1, t_2)$ is a one-to-one function of t_2. Similarly for $y_2(x_2, t_1)$.

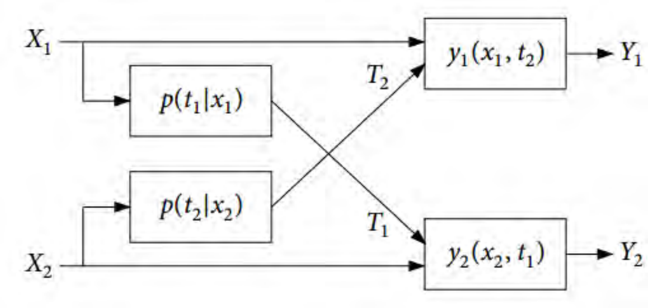

Figure 1.3: Injective semi-deterministic interference channel

Theorem 1.4.2 ([19])**.** Any achievable rate pair (R_1, R_2) for the injective semi-deterministic IC must satisfy the inequalities

$$
\begin{aligned}
R_1 &\leq H(Y_1|X_2, Q) - H(T_2|X_2), \\
R_2 &\leq H(Y_2|X_1, Q) - H(T_1|X_1), \\
R_1 + R_2 &\leq H(Y_1|Q) + H(Y_2|U_2, X_1, Q) - H(T_1|X_1) - H(T_2|X_2), \\
R_1 + R_2 &\leq H(Y_1|U_1, X_2, Q) + H(Y_2|Q) - H(T_1|X_1) - H(T_2|X_2), \\
R_1 + R_2 &\leq H(Y_1|U_1, Q) + H(Y_2|U_2, Q) - H(T_1|X_1) - H(T_2|X_2), \\
2R_1 + R_2 &\leq H(Y_1|Q) + H(Y_1|U_1, X_2, Q) + H(Y_2|U_2, Q) \\
&\quad - H(T_1|X_1) - 2H(T_2|X_2), \\
R_1 + 2R_2 &\leq H(Y_2|Q) + H(Y_2|U_2, X_1, Q) + H(Y_1|U_1, Q) \\
&\quad - 2H(T_1|X_1) - H(T_2|X_2)
\end{aligned}
$$

for some $p(q)p(x_1|q)p(x_2|q)p_{T_1|X_1}(u_1|x_1)p_{T_2|X_2}(u_2|x_2)$.

The GIC is a special class of this semi-deterministic IC with $T_1 = aX_1 + Z_2$ and $T_2 = bX_2 + Z_1$. For GICs, [19] showed that the gap between Han–Kobayashi inner bound and this outer bound is less than half a bit.

Chapter 2

Genie-based outer bounds

Genie-based arguments were first used to establish the capacity of injective deterministic ICs [8]. Recently, they has been employed to show a half-bit gap for the Han–Kobayashi region in [10] and also to establish the sum-capacity of the Gaussian interference channel in [18], [2], [13] for a subset of the weak interference regime. Motivated by these works, two outer bounds on weighted sum-capacity are derived in this chapter.

The capacity region is characterized using tangent lines which are given by the maximal weighted sum rate $\max R_1 + \lambda R_2$. Thus we consider outer bound on maximal weighted sum rate for $\lambda \geq 1$. (When $\lambda \leq 1$, the maximal weighted sum rate considered instead is $\max \frac{1}{\lambda} R_1 + R_2$. The outer bound on it can be obtained similarly.)

2.1 A genie-based outer bound

Capacity regions can be easily characterized in many multi-user settings as a limit of n-letter expressions using Fano's inequality. However these limits are infeasible to compute without knowing explicit convergence behaviour. On the other hand information-theorists seek computable characterizations of capacity regions.

Outer bounds to a capacity region are computable regions that contain the capacity region. These outer bounds usually satisfy the *ten-*

sorization property, i.e. their multi-letter extensions coincide with the single-letter one. Usually, the outer bounds are obtained by upper bounding an n-letter region which tends to the capacity region, by a tensorizing functional whose single-letter region is computable.

In this chapter we develop outer bounds by giving additional information (usually said to be provided by *genies*) to the receivers prior to finding a tensorizing expression. With the help of genies, the n-letter expression of the capacity region can be upper bounded by a n-letter genie-based outer bound. Then this n-letter genie-based outer bound is further single-letterized to a 1-letter genie-based outer bound so that the express now becomes computable. We will show that in later chapter this genie-based outer bound can be tight.

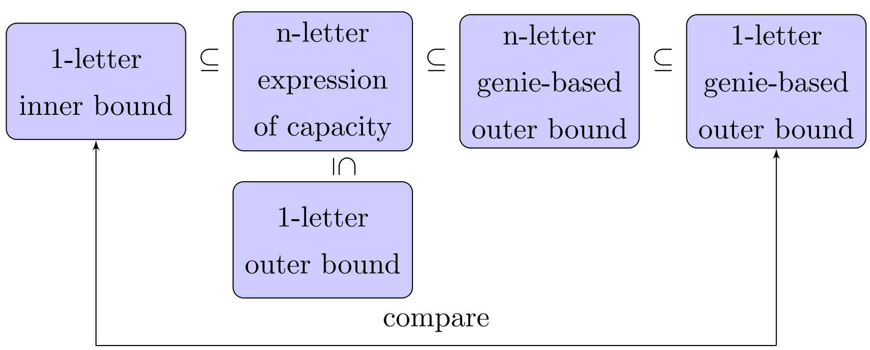

Figure 2.1: How a "genie" could help

Before we present the outer bound, we define the notion of upper concave envelope, which will be used to express the genie-based outer bound. The upper concave envelope of a function $f(x)$ over domain $\mathcal{D}$ is defined as

$$\mathcal{C}[f](x) := \inf\{g(x) : g(y) \text{ is concave in } \mathcal{D}, \text{ and } g(y) \geq f(y)\ \forall y \in \mathcal{D}.\}.$$

The following theorem provides an outer bound to the capacity region of the interference channel with genie random variables denoted by T_1, T_2 carrying information about X_1 and X_2 respectively. The structure of the genie-aided channel is depicted in Figure 2.2.

Theorem 2.1.1 (Genie based outer bound)**.** Consider a discrete mem-

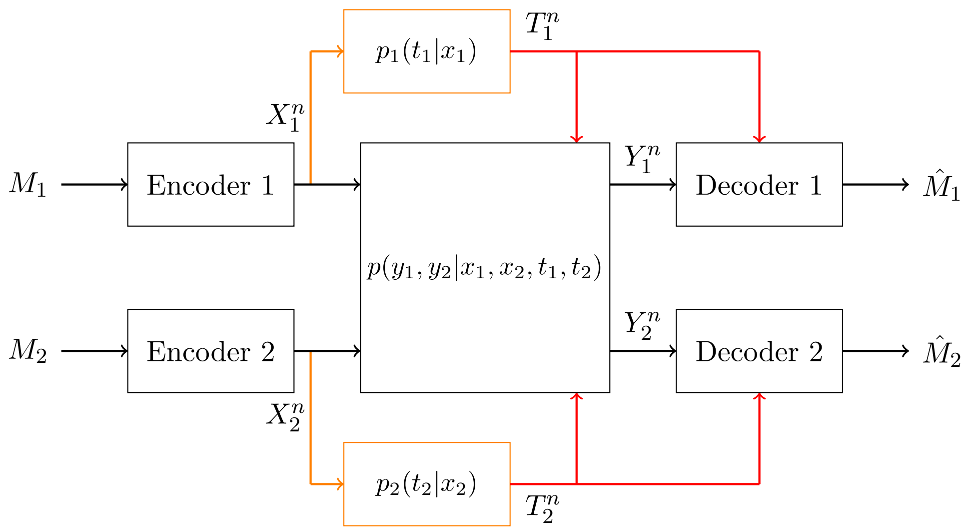

Figure 2.2: Discrete memoryless interference channel with genies

oryless interference channel characterized by $w(y_1, y_2|x_1, x_2)$. Let T_1, T_2 be any pair of random variables such that the joint distributions satisfy $p(y_1, t_1, y_2, t_2|x_1, x_2) = p(t_1|x_1)p(t_2|x_2)p(y_1, y_2|t_1, t_2, x_1, x_2)$, and their marginals distributions are consistent with the given channel transition probabilities, i.e. $p(y_1|x_1, x_2) = w(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2) = w(y_2|x_1, x_2)$. The achievable weighted sum-rate can be upper bounded as follows:

$$\begin{aligned} R_1 + \lambda R_2 \leq \max_{p_1(x_1)p_2(x_2)} & I(X_1; T_1 Y_1) + \lambda I(X_2; T_2 Y_2) \\ & + \mathcal{C}[I(X_1; T_1|X_2 T_2) - \lambda I(X_1; Y_2|T_2 X_2)] \\ & - I(X_1; T_1|X_2 T_2) + \lambda I(X_1; Y_2|T_2 X_2) \\ & + \mathcal{C}[I(X_2; T_2|X_1 T_1) - I(X_2; Y_1|T_1 X_1)] \\ & - I(X_2; T_2|X_1 T_1) + I(X_2; Y_1|T_1 X_1), \end{aligned} \tag{2.1}$$

where $\mathcal{C}[I(X_2; T_2|X_1 T_1) - I(X_2; Y_1|T_1 X_1)]$ denotes the upper concave envelope of the function $I(X_2; T_2|X_1 T_1) - I(X_2; Y_1|T_1 X_1)$ evaluated with respect to the space of product distributions $p_1(x_1)p_2(x_2)$. Similarly, $\mathcal{C}[I(X_1; T_1|X_2 T_2) - \lambda I(X_1; Y_2|T_2 X_2)]$ denotes the upper concave envelope of the function $I(X_1; T_1|X_2 T_2) - \lambda I(X_1; Y_2|T_2 X_2)$ evaluated with respect to the same space of product distributions $p_1(x_1)p_2(x_2)$.

Proof. See Appendix 2.A. □

2.2 Enhanced genie-based outer bound

In the previous section, one can observe that from the Markov structure, the genie random variable T_1 (T_2) carries information about X_1 (X_2) and helps receiver 1 (2) to decode its message. Indeed, one can use another genie random variable S_2 (S_1) carrying information of X_2 (X_1) to help receiver 1 (2) to decode its message. The pair of genies T_1 and S_2 (T_2 and S_1) helping receiver 1 (2) to decode message would be potentially better than the single genie in the previous subsection. The enhanced structure of genie-aided channel is depicted in Figure 2.3. We obtain a single-letter outer bound based on this scenario and it is presented below.

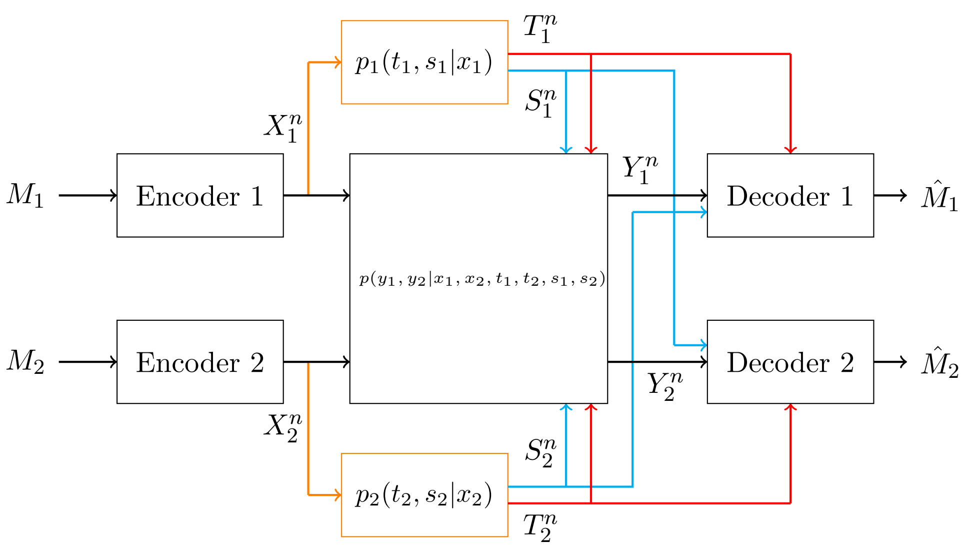

Figure 2.3: DM-IC with two genies per decoder

Theorem 2.2.1 (Enhanced genie-based outer bound)**.** Consider a discrete memoryless interference channel with transition probability marginals $w(y_1, y_2|x_1, x_2)$. Let T_1, S_1, T_2, S_2 be any random variables such that

$p(y_1, t_1, s_1, y_2, t_2, s_2|x_1, x_2)$ decomposes as

$$p(t_1, s_1|x_1)p(t_2, s_2|x_2)p(y_1, y_2|t_1, t_2, s_1, s_2, x_1, x_2).$$

Further we require that

- the marginals are consistent with the given channel transition probabilities, that is,
  $p(y_1|x_1, x_2) = w(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2) = w(y_2|x_1, x_2)$.
- for each $i = 1, 2$, T_i, S_i has degraded order, *i.e.* either $X_i \to T_i \to S_i$ or $X_i \to S_i \to T_i$ must form a Markov chain.

The weighted sum-capacity of this DMIC can be upper bounded as following:

$$\begin{aligned} R_1 + \lambda R_2 \leq \max_{p_1(x_1)p_2(x_2)} & I(X_1; T_1, Y_1|S_2) + \lambda I(X_2; T_2, Y_2|S_1) \\ & + \mathcal{C}[I(X_1; T_1|X_2, T_2, S_1) - \lambda I(X_1; Y_2|X_2, T_2, S_1)] \\ & - I(X_1; T_1|X_2, T_2, S_1) + \lambda I(X_1; Y_2|X_2, T_2, S_1) \\ & + \mathcal{C}[I(X_2; T_2|X_1, T_1, S_2) - I(X_2; Y_1|X_1, T_1, S_2)] \\ & - I(X_2; T_2|X_1, T_1, S_2) + I(X_2; Y_1|X_1, T_1, S_2) \end{aligned} \tag{2.2}$$

where $\mathcal{C}[\cdot]$ denotes as before the upper concave envelope of a function over the space of product distributions $p_1(x_1)p_2(x_2)$.

Proof. See Appendix 2.B. □

Genie random variables are different from traditional auxiliaries. For an outer bound involving traditional auxiliaries, the region containing the capacity region is usually obtained by taking the union over all possible distribution of the auxiliaries. (This is the reason that outer bounds can only be computable if there are cardinality bounds on auxiliaries) Whereas in this genie-based outer bound, any feasible genie produces a valid outer bound. Therefore the challenge is to find genies that lead to plausibly tight outer bounds.

Remark 2.2.1. The last degradation requirement in Theorem 2.2.1 that two genies together with channel input must form the Markov chain is an assumption to make the single-letterization go through. In the Gaussian interference channel, since we typically take the genies to be the signals with additive Gaussian noise, this degradation condition is automatically satisfied.

We will show in Sections 4.1, 4.2 and 4.4 that this outer bound turns out to be tight for Gaussian interference channels in regimes where the sum-capacity is known: strong interference regime ($a \geq 1$, $b \geq 1$), mixed interference regime ($a \geq 1$, $b < 1$ or $a < 1$, $b \geq 1$), weak interference sub-regime ($a < 1$, $b < 1$ and $a(1+b^2P_2)+b(1+a^2P_1) \leq 1$). Indeed this is the first outer bound that unifies all the results on sum-capacity.

Appendix

2.A Proof of Theorem 2.1.1

Proof of theorem 2.1.1. Consider a sequence of codebooks with growing block length n such that their decoding error probabilities tend to zero as n goes to infinity. The distribution on the n-tuples is given by

$$p(m_1, m_2, x_1^n, x_2^n, y_1^n, t_1^n, y_2^n, t_2^n)$$
$$= p(m_1, x_1^n)p(m_2, x_2^n)\prod_{i=1}^{n} p(t_{1i}|x_{1i})p(y_{1i}|x_{1i}, x_{2i}, t_{1i})p(t_{2i}|x_{2i})p(y_{2i}|x_{1i}, x_{2i}, t_{2i}).$$

Keep in mind that the channel capacity of an interference channel depends only on the marginals $q(y_1|x_1, x_2)$ and $q(y_2|x_1, x_2)$ and that the distribution above is consistent with the marginal distributions by assumption. For $\lambda \geq 1$,

$$\begin{aligned}
&n(R_1 + \lambda R_2) \\
&= H(M_1) + \lambda H(M_2) \\
&\leq I(M_1; Y_1^n) + \lambda I(M_2; Y_2^n) + n\epsilon \qquad \text{(by Fano's inequality)} \\
&\leq I(X_1^n; Y_1^n) + \lambda I(X_2^n; Y_2^n) + n\epsilon \\
&\leq I(X_1^n; Y_1^n T_1^n) + \lambda I(X_2^n; Y_2^n T_2^n) + n\epsilon \\
&= I(X_1^n; T_1^n) + I(X_1^n; Y_1^n|T_1^n) + \lambda I(X_2^n; T_2^n) + \lambda I(X_2^n; Y_2^n|T_2^n) + n\epsilon \\
&= H(T_1^n) - H(T_1^n|X_1^n) + H(Y_1^n|T_1^n) - H(Y_1^n|T_1^n X_1^n) \\
&\quad + \lambda H(T_2^n) - \lambda H(T_2^n|X_2^n) + \lambda H(Y_2^n|T_2^n) - \lambda H(Y_2^n|T_2^n X_2^n) + n\epsilon.
\end{aligned}$$

Then, for the term $H(T_1^n) - \lambda H(Y_2^n|X_2^n T_2^n)$, note that

$$\begin{aligned} & H(T_1^n) - \lambda H(Y_2^n|X_2^n T_2^n) \\ =& H(T_1^n|T_2^n X_2^n) - \lambda H(Y_2^n|X_2^n T_2^n) \\ & \qquad \big(\text{since } T_1^n \text{ is independent of } (T_2^n, X_2^n)\big) \\ =& \sum_i H(T_{1i}|T_1^{i-1} T_2^n X_2^n) - \lambda H(Y_{2i}|Y_{2,i+1}^n X_2^n T_2^n) \\ \leq& \sum_i H(T_{1i}|Y_{2,i+1}^n T_1^{i-1} T_2^n X_2^n) - \lambda H(Y_{2i}|T_1^{i-1} Y_{2,i+1}^n X_2^n T_2^n) \\ & \qquad \text{(Csiszar-sum lemma)} \\ =& \sum_i H(T_{1i}|U_i X_{2i} T_{2i}) - \lambda H(Y_{2i}|U_i X_{2i} T_{2i}). \\ & \qquad \big(U_i := (Y_{2,i+1}^n, T_1^{i-1}, T_2^{n\setminus i}, X_2^{n\setminus i})\big) \end{aligned}$$

Consider a Bayesian network representation in Figure 2.A.1 of the variables. Any path from X_{1i} to X_{2i} is d-separated by X_{2i+1}^n. Hence

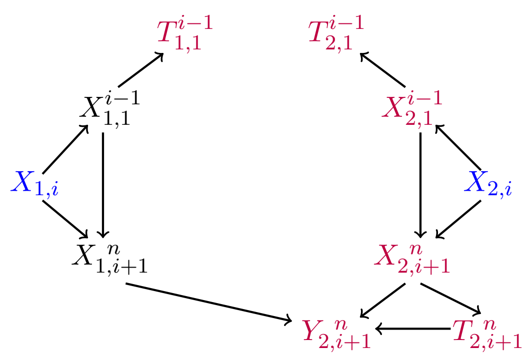

Figure 2.A.1: Bayesian network of dependence

we have Markov chain $X_{1i} \to U_i \to X_{2i}$.

Similarly

$$\begin{aligned} & H(T_2^n) - H(Y_1^n|X_1^n T_1^n) \\ & \quad = \sum_i H(T_{2i}|V_i X_{1i} T_{1i}) - H(Y_{1i}|V_i X_{1i} T_{1i}) \end{aligned}$$

where $V_i = (Y_{1,i+1}^n, T_2^{i-1}, T_1^{n\setminus i}, X_1^{n\setminus i})$ and $X_{1i} \to V_i \to X_{2i}$.

Secondly, from the n-tuple distribution we get that

$$H(T_1^n|X_1^n) = \sum_{i=1}^{n} H(T_{1i}|X_{1i}X_1^{n\backslash i}T_1^{i-1}) = \sum_{i=1}^{n} H(T_{1i}|X_{1i}),$$
$$H(T_2^n|X_2^n) = \sum_{i=1}^{n} H(T_{2i}|X_{2i}X_2^{n\backslash i}T_2^{i-1}) = \sum_{i=1}^{n} H(T_{2i}|X_{2i}).$$

Following chain rule and that conditioning reduces entropy,

$$H(Y_1^n|T_1^n) \le \sum_{i=1}^{n} H(Y_{1i}|T_{1i}),$$
$$H(Y_2^n|T_2^n) \le \sum_{i=1}^{n} H(Y_{2i}|T_{2i}).$$

Combining the above arguments, using routine manipulations, we obtain that

$$\begin{aligned}
&n(R_1 + \lambda R_2)\\
&\le H(T_1^n) - H(T_1^n|X_1^n) + H(Y_1^n|T_1^n) - H(Y_1^n|T_1^n X_1^n)\\
&+ H(T_2^n) - \lambda H(T_2^n|X_2^n) + \lambda H(Y_2^n|T_2^n) - \lambda H(Y_2^n|T_2^n X_2^n)\\
&\quad + (\lambda - 1)H(T_2^n) + n\epsilon\\
&\le \sum_i H(T_{2i}|V_i X_{1i} T_{1i}) - H(Y_{1i}|V_i X_{1i} T_{1i}) - H(T_{1i}|X_{1i}) + H(Y_{1i}|T_{1i})\\
&\quad + H(T_{1i}|U_i X_{2i} T_{2i}) - \lambda H(Y_{2i}|U_i X_{2i} T_{2i}) - \lambda H(T_{2i}|X_{2i}) + \lambda H(Y_{2i}|T_{2i})\\
&\quad + (\lambda - 1)H(T_{2i}) + n\epsilon\\
&= \sum_i I(X_{2i}; T_{2i}|V_i X_{1i} T_{1i}) + I(V_i X_{1i}; Y_{1i}|T_{1i}) - (\lambda - 1)H(T_{2i}|X_{2i})\\
&\quad + I(X_{1i}; T_{1i}|U_i X_{2i} T_{2i}) + \lambda I(U_i X_{2i}; Y_{2i}|T_{2i}) + (\lambda - 1)H(T_{2i}) + n\epsilon\\
&= \sum_i I(X_{2i}; T_{2i}|X_{1i} T_{1i}) - I(V_i; T_{2i}|X_{1i} T_{1i})\\
&\qquad\qquad\qquad (\text{since } I(V_i X_{2i}; T_{2i}|X_{1i} T_{1i}) = I(X_{2i}; T_{2i}|X_{1i} T_{1i})\ )\\
&\quad + I(X_{1i}; Y_{1i}|T_{1i}) + I(V_i; Y_{1i}|T_{1i} X_{1i})\\
&\quad + I(X_{1i}; T_{1i}|X_{2i} T_{2i}) - I(U_i; T_{1i}|X_{2i} T_{2i})\\
&\qquad\qquad\qquad (\text{since } I(U_i X_{1i}; T_{1i}|X_{2i} T_{2i}) = I(X_{1i}; T_{1i}|X_{2i} T_{2i})\ )\\
&\quad + \lambda I(X_{2i}; Y_{2i}|T_{2i}) + \lambda I(U_i; Y_{2i}|T_{2i} X_{2i}) + (\lambda - 1)I(X_{2i}; T_{2i}) + n\epsilon
\end{aligned}$$

$$= \sum_i \lambda I(X_{2i};T_{2i}) - I(V_i;T_{2i}|X_{1i}T_{1i}) + I(X_{1i};Y_{1i}|T_{1i}) + I(V_i;Y_{1i}|T_{1i}X_{1i})$$
$$+ I(X_{1i};T_{1i}) - I(U_i;T_{1i}|X_{2i}T_{2i}) + \lambda I(X_{2i};Y_{2i}|T_{2i})$$
$$+ \lambda I(U_i;Y_{2i}|T_{2i}X_{2i}) + n\epsilon$$
$$\text{(since } (X_1,T_1) \text{ and } (X_2,T_2) \text{ are independent)}$$
$$= \sum_i I(X_{1i};T_{1i}Y_{1i}) + \lambda I(X_{2i};T_{2i}Y_{2i})$$
$$- I(V_i;T_{2i}|X_{1i}T_{1i}) + I(V_i;Y_{1i}|T_{1i}X_{1i})$$
$$- I(U_i;T_{1i}|X_{2i}T_{2i}) + \lambda I(U_i;Y_{2i}|T_{2i}X_{2i}) + n\epsilon$$

Now since $V_i \to (X_{1i},T_{1i},X_{2i}) \to (Y_{1i},T_{2i})$ and $U_i \to (X_{1i},X_{2i},T_{2i}) \to (Y_{2i},T_{1i})$, one can rewrite the above as

$$\begin{aligned}
&n(R_1 + \lambda R_2)\\
&\le \sum_i I(X_{1i};T_{1i}Y_{1i}) + \lambda I(X_{2i};T_{2i}Y_{2i})\\
&\quad - I(X_{2i};T_{2i}|X_{1i}T_{1i}) + I(X_{2i};Y_{1i}|T_{1i}X_{1i})\\
&\quad + I(X_{2i};T_{2i}|V_i,X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|V_i,T_{1i}X_{1i})\\
&\quad - I(X_{1i};T_{1i}|X_{2i}T_{2i}) + \lambda I(X_{1i};Y_{2i}|T_{2i}X_{2i})\\
&\quad + I(X_{1i};T_{1i}|U_i,X_{2i}T_{2i}) - \lambda I(X_{1i};Y_{2i}|U_i,T_{2i}X_{2i}) + n\epsilon\\
&\le \sum_i I(X_{1i};T_{1i}Y_{1i}) + \lambda I(X_{2i};T_{2i}Y_{2i}) + n\epsilon\\
&\quad - I(X_{2i};T_{2i}|X_{1i}T_{1i}) + I(X_{2i};Y_{1i}|T_{1i}X_{1i})\\
&\quad + \mathcal{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]\\
&\quad - I(X_{1i};T_{1i}|X_{2i}T_{2i}) + \lambda I(X_{1i};Y_{2i}|T_{2i}X_{2i})\\
&\quad + \mathcal{C}[I(X_{1i};T_{1i}|X_{2i}T_{2i}) - \lambda I(X_{1i};Y_{2i}|T_{2i}X_{2i})],
\end{aligned}$$

where $\mathcal{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]$ is the upper concave envelope of the function $I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})$ defined on the space of distributions $p_1(x_1)p_2(x_2)$. It is easy to see from the definition of the upper concave envelope that

$$\begin{aligned}
&\mathcal{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]\\
&= \sup_{\substack{U: X_{1i} \to U \to X_{2i}\\ U \to (X_{1i},X_{2i}) \to (Y_{1i},T_{2i},T_{1i})}} I(X_{1i};T_{1i}|U,X_{2i}T_{2i})
\end{aligned}$$

$$- I(X_{1i}; Y_{2i}|U, T_{2i}X_{2i}).$$

Thus for any valid choice of genies T_1, T_2, we obtain an outer bound to the sum-rate given by

$$\begin{aligned}
& R_1 + \lambda R_2 \\
\leq & \max_{p_1(x_1)p_2(x_2)} I(X_1; T_1Y_1) + \lambda I(X_2; T_2Y_2) \\
& \quad + \mathcal{C}[I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)] \\
& \quad - I(X_2; T_2|X_1T_1) + I(X_2; Y_1|T_1X_1) \\
& \quad + \mathcal{C}[I(X_1; T_1|X_2T_2) - \lambda I(X_1; Y_2|T_2X_2)] \\
& \quad - I(X_1; T_1|X_2T_2) + \lambda I(X_1; Y_2|T_2X_2)
\end{aligned} \tag{2.3}$$

□

2.B Proof of Theorem 2.2.1

Proof of Theorem 2.2.1. The proof is basically following Csiszar sum lemma and manipulation of mutual information.

$$\begin{aligned}
& n(R_1 + \lambda R_2) - n\epsilon \\
\leq & H(M_1) + \lambda H(M_2) \\
\leq & I(X_1^n; Y_1^n) + \lambda I(X_1^n; Y_1^n) \\
\leq & I(X_1^n; Y_1^n T_1^n S_2^n) + \lambda I(X_2^n; Y_2^n T_2^n S_1^n) \\
= & I(X_1^n; T_1^n) + I(X_1^n; Y_1^n|T_1^n S_2^n) + \lambda I(X_2^n; T_2^n) + \lambda I(X_2^n; Y_2^n|T_2^n S_1^n) \\
= & \underline{H(T_1^n)} - H(T_1^n|X_1^n) + H(Y_1^n|T_1^n S_2^n) \underline{- H(Y_1^n|T_1^n S_2^n X_1^n)} \\
& \quad \underline{+ \lambda H(T_2^n)} - \lambda H(T_2^n|X_2^n) + \lambda H(Y_2^n|T_2^n S_1^n) \underline{- \lambda H(Y_2^n|T_2^n S_1^n X_2^n)}
\end{aligned}$$

Note that

$$\begin{aligned}
& H(T_1^n) - \lambda H(Y_2^n|T_2^n S_1^n X_2^n) \\
= & H(T_1^n|S_1^n) + I(T_1^n; S_1^n) - \lambda H(Y_2^n|T_2^n S_1^n X_2^n) \\
= & H(T_1^n|T_2^n S_1^n X_2^n) + I(T_1^n; S_1^n) - \lambda H(Y_2^n|T_2^n S_1^n X_2^n) \\
\leq & \sum_i H(T_{1i}|T_1^{i-1} Y_{2,i+1}^n T_2^n S_1^n X_2^n) - \lambda H(Y_{2i}|T_1^{i-1} Y_{2,i+1}^n T_2^n S_1^n X_2^n)
\end{aligned}$$

$$+ I(T_1^n; S_1^n)$$

The last inequality is due to Csiszar sum identity. We have

$$\begin{aligned}
& n(R_1 + \lambda R_2) - n\epsilon \\
\le & \sum_i H(T_{1i}|T_1^{i-1} Y_{2,i+1}^n T_2^n S_1^n X_2^n) - \lambda H(Y_{2i}|T_1^{i-1} Y_{2,i+1}^n T_2^n S_1^n X_2^n) \\
& + H(T_{2i}|T_2^{i-1} Y_{1,i+1}^n T_1^n S_2^n X_1^n) - H(Y_{1i}|T_2^{i-1} Y_{1,i+1}^n T_1^n S_2^n X_1^n) \\
& - H(T_{1i}|X_{1i}) + H(Y_{1i}|T_{1i} S_{2i}) - \lambda H(T_{2i}|X_{2i}) + \lambda H(Y_{2i}|T_{2i} S_{1i}) \\
& + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) + (\lambda - 1) H(T_2^n)
\end{aligned}$$

Use substitution $U_{1i} = T_1^{i-1} S_1^{n\backslash i}$, $V_{1i} = X_2^{n\backslash i} T_2^{n\backslash i} Y_{2,i+1}^n$,$U_{2i} = T_2^{i-1} S_2^{n\backslash i}$, $V_{2i} = X_1^{n\backslash i} T_1^{n\backslash i} Y_{1,i+1}^n$,

$$\begin{aligned}
& n(R_1 + \lambda R_2) - n\epsilon \\
= & \sum_i H(T_{1i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i}) - \lambda H(Y_{2i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i}) \\
& + H(T_{2i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i}) - H(Y_{1i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i}) \\
& - \underline{H(T_{1i}|X_{1i} S_{1i})} - I(T_{1i}; S_{1i}|X_{1i}) + H(Y_{1i}|T_{1i} S_{2i}) \\
& - \underline{\lambda H(T_{2i}|X_{2i} S_{2i})} - \lambda I(T_{2i}; S_{2i}|X_{2i}) + \lambda H(Y_{2i}|T_{2i} S_{1i}) \\
& + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) + (\lambda - 1) H(T_{2i}) \\
\le & \sum_i \underline{H(T_{1i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i})} - \lambda H(Y_{2i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i}) \\
& - \underline{H(T_{1i}|X_{1i} U_{1i} V_{1i} T_{2i} S_{1i} X_{2i})} + H(Y_{1i}|T_{1i} S_{2i}) \\
& + \underline{H(T_{2i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i})} - H(Y_{1i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i}) \\
& - \underline{\lambda H(T_{2i}|X_{2i} U_{2i} V_{2i} T_{1i} S_{2i} X_{1i})} + \lambda H(Y_{2i}|T_{2i} S_{1i}) \\
& - I(T_{1i}; S_{1i}|X_{1i}) - \lambda I(T_{2i}; S_{2i}|X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) \\
& + (\lambda - 1) H(T_{2i}) \\
= & \sum_i I(X_{1i}; T_{1i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i}) \\
& - \underline{\lambda H(Y_{2i}|U_{1i} V_{1i} T_{2i} S_{1i} X_{2i} X_{1i})} + H(Y_{1i}|T_{1i} S_{2i}) \\
& + I(X_{2i}; T_{2i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i}) \\
& - \underline{H(Y_{1i}|U_{2i} V_{2i} T_{1i} S_{2i} X_{1i} X_{2i})} + \lambda H(Y_{2i}|T_{2i} S_{1i})
\end{aligned}$$

$$- I(T_{1i}; S_{1i}|X_{1i}) - \lambda I(T_{2i}; S_{2i}|X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n)$$
$$+ (\lambda - 1) I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i}; T_{2i})$$

The inequality is due to the fact conditional entropy is less than original entropy. Use mutual information to rewrite above as follow,

$$\begin{aligned}
& n(R_1 + \lambda R_2) - \epsilon \\
= & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\
& - \underline{\lambda H(Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}X_{1i})} + H(Y_{1i}|T_{1i}S_{2i}) \\
& + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \\
& - \underline{H(Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}X_{2i})} + \lambda H(Y_{2i}|T_{2i}S_{1i}) \\
& - I(T_{1i}; S_{1i}|X_{1i}) - \lambda I(T_{2i}; S_{2i}|X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) \\
& + (\lambda - 1) I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i}; T_{2i}) \\
= & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\
& - \underline{\lambda H(Y_{2i}|T_{2i}S_{1i}X_{2i}X_{1i})} + \underline{H(Y_{1i}|X_{1i}T_{1i}S_{2i})} + I(X_{1i}; Y_{1i}|T_{1i}S_{2i}) \\
& + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \\
& - \underline{H(Y_{1i}|T_{1i}S_{2i}X_{1i}X_{2i})} + \underline{\lambda H(Y_{2i}|X_{2i}T_{2i}S_{1i})} + \lambda I(X_{2i}; Y_{2i}|T_{2i}S_{1i}) \\
& - I(T_{1i}; S_{1i}|X_{1i}) - \lambda I(T_{2i}; S_{2i}|X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) \\
& + (\lambda - 1) I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i}; T_{2i}) \\
= & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\
& + I(X_{2i}; Y_{1i}|X_{1i}T_{1i}S_{2i}) + I(X_{1i}; Y_{1i}|T_{1i}S_{2i}) \\
& + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \\
& + \lambda I(X_{1i}; Y_{2i}|X_{2i}T_{2i}S_{1i}) + \lambda I(X_{2i}; Y_{2i}|T_{2i}S_{1i}) \\
& - I(T_{1i}; S_{1i}|X_{1i}) - \lambda I(T_{2i}; S_{2i}|X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) \\
& + (\lambda - 1) I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i}; T_{2i})
\end{aligned}$$

Add and subtract the terms $I(X_{1i}; T_{1i}|T_{2i}S_{1i}X_{2i})$ $(= I(X_{1i}; T_{1i}|S_{1i}))$

and $I(X_{2i};T_{2i}|T_{1i}S_{2i}X_{1i})$ $(= I(X_{2i};T_{2i}|S_{2i}))$. We obtain

$$\begin{aligned}
&n(R_1+\lambda R_2)-n\epsilon\\
=&\sum_i I(X_{1i};T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})-\lambda I(X_{1i};Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})\\
&-I(X_{1i};T_{1i}|T_{2i}S_{1i}X_{2i})+\lambda I(X_{1i};Y_{2i}|X_{2i}T_{2i}S_{1i})\\
&+I(X_{2i};T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})-I(X_{2i};Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})\\
&-I(X_{2i};T_{2i}|T_{1i}S_{2i}X_{1i})+I(X_{2i};Y_{1i}|X_{1i}T_{1i}S_{2i})\\
&+\underline{I(X_{1i};T_{1i}|S_{1i})}+\underline{I(X_{2i};T_{2i}|S_{2i})}+\underline{I(X_{1i};Y_{1i}|T_{1i}S_{2i})}+\underline{\lambda I(X_{2i};Y_{2i}|T_{2i}S_{1i})}\\
&-I(T_{1i};S_{1i}|X_{1i})-\lambda I(T_{2i};S_{2i}|X_{2i})+I(T_1^n;S_1^n)+I(T_2^n;S_2^n)\\
&+(\lambda-1)I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i};T_{2i})\\
=&\sum_i I(X_{1i};T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})-\lambda I(X_{1i};Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})\\
&-I(X_{1i};T_{1i}|T_{2i}S_{1i}X_{2i})+\lambda I(X_{1i};Y_{2i}|X_{2i}T_{2i}S_{1i})\\
&+I(X_{2i};T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})-I(X_{2i};Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})\\
&-I(X_{2i};T_{2i}|T_{1i}S_{2i}X_{1i})+I(X_{2i};Y_{1i}|X_{1i}T_{1i}S_{2i})\\
&+I(X_{1i};T_{1i}|S_{1i})+I(X_{2i};T_{2i}|S_{2i})+I(X_{1i};Y_{1i}T_{1i}|S_{2i})-\underline{I(X_{1i};T_{1i}|S_{2i})}\\
&+\lambda I(X_{2i};Y_{2i}T_{2i}|S_{1i})-\underline{\lambda I(X_{2i};T_{2i}|S_{1i})}-\underline{I(T_{1i};S_{1i}|X_{1i})}-\underline{\lambda I(T_{2i};S_{2i}|X_{2i})}\\
&+I(T_1^n;S_1^n)+I(T_2^n;S_2^n)+(\lambda-1)I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i};T_{2i})
\end{aligned}$$

Note (X_{1i},T_{1i}) and S_{2i} are independent. So $I(X_{1i};T_{1i}|S_{2i})=I(X_{1i};T_{1i})$ and $I(X_{2i};T_{2i}|S_{1i})=I(X_{2i};T_{2i})$.

$$\begin{aligned}
&n(R_1+\lambda R_2)-n\epsilon\\
=&\sum_i I(X_{1i};T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})-\lambda I(X_{1i};Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i})\\
&-I(X_{1i};T_{1i}|T_{2i}S_{1i}X_{2i})+\lambda I(X_{1i};Y_{2i}|X_{2i}T_{2i}S_{1i})\\
&+I(X_{2i};T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})-I(X_{2i};Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i})\\
&-I(X_{2i};T_{2i}|T_{1i}S_{2i}X_{1i})+I(X_{2i};Y_{1i}|X_{1i}T_{1i}S_{2i})\\
&+I(X_{1i};T_{1i}|S_{1i})+I(X_{2i};T_{2i}|S_{2i})+I(X_{1i};Y_{1i}T_{1i}|S_{2i})+\lambda I(X_{2i};Y_{2i}T_{2i}|S_{1i})\\
&-I(T_{1i};S_{1i}X_{1i})-\lambda I(T_{2i};S_{2i}X_{2i})+I(T_1^n;S_1^n)+I(T_2^n;S_2^n)\\
&+(\lambda-1)I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i};T_{2i})
\end{aligned}$$

When genies has degraded order, say $X_1 \to T_1 \to S_1$, we have

$$\begin{aligned} I(T_1^n; S_1^n) &= H(S_1^n) - H(S_1^n|T_1^n) \\ &\le \sum_i H(S_{1i}) - H(S_{1i}|S_1^{i-1}T_1^n) \\ &= \sum_i H(S_{1i}) - H(S_{1i}|T_{1i}) \\ &= \sum_i I(T_{1i}; S_{1i}) \end{aligned}$$

since $S_1^{i-1}T_1^{n\backslash i} \to X_{1i} \to T_{1i} \to S_{1i}$.

$$\begin{aligned} & n(R_1 + \lambda R_2) - \epsilon \\ \le & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\ & - I(X_{1i}; T_{1i}|T_{2i}S_{1i}X_{2i}) + \lambda I(X_{1i}; Y_{2i}|X_{2i}T_{2i}S_{1i}) \\ & + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \\ & - I(X_{2i}; T_{2i}|T_{1i}S_{2i}X_{1i}) + I(X_{2i}; Y_{1i}|X_{1i}T_{1i}S_{2i}) \\ & + I(X_{1i}; T_{1i}|S_{1i}) + I(X_{2i}; T_{2i}|S_{2i}) + I(X_{1i}; Y_{1i}T_{1i}|S_{2i}) + \lambda I(X_{2i}; Y_{2i}T_{2i}|S_{1i}) \\ & - I(T_{1i}; S_{1i}X_{1i}) - \lambda I(T_{2i}; S_{2i}X_{2i}) + I(T_1^n; S_1^n) + I(T_2^n; S_2^n) \\ & + (\lambda - 1) I(X_{1i}X_{2i}U_{2i}V_{2i}T_{1i}S_{2i}; T_{2i}) \\ \le & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\ & - I(X_{1i}; T_{1i}|T_{2i}S_{1i}X_{2i}) + \lambda I(X_{1i}; Y_{2i}|X_{2i}T_{2i}S_{1i}) \\ & + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \\ & - I(X_{2i}; T_{2i}|T_{1i}S_{2i}X_{1i}) + I(X_{2i}; Y_{1i}|X_{1i}T_{1i}S_{2i}) \\ & + \underline{I(X_{1i}; T_{1i}|S_{1i})} + \underline{I(X_{2i}; T_{2i}|S_{2i})} + I(X_{1i}; Y_{1i}T_{1i}|S_{2i}) + \lambda I(X_{2i}; Y_{2i}T_{2i}|S_{1i}) \\ & - \underline{I(T_{1i}; S_{1i}X_{1i})} - \underline{\lambda I(T_{2i}; S_{2i}X_{2i})} + \underline{I(T_{1i}; S_{1i})} + \underline{I(T_{2i}; S_{2i})} \\ & + (\lambda - 1) I(X_{2i}S_{2i}; T_{2i}) \\ = & \sum_i I(X_{1i}; T_{1i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) - \lambda I(X_{1i}; Y_{2i}|U_{1i}V_{1i}T_{2i}S_{1i}X_{2i}) \\ & - I(X_{1i}; T_{1i}|T_{2i}S_{1i}X_{2i}) + \lambda I(X_{1i}; Y_{2i}|X_{2i}T_{2i}S_{1i}) \\ & + I(X_{2i}; T_{2i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) - I(X_{2i}; Y_{1i}|U_{2i}V_{2i}T_{1i}S_{2i}X_{1i}) \end{aligned}$$

$$- I(X_{2i}; T_{2i}|T_{1i}S_{2i}X_{1i}) + I(X_{2i}; Y_{1i}|X_{1i}T_{1i}S_{2i})$$
$$+ I(X_{1i}; Y_{1i}T_{1i}|S_{2i}) + \lambda I(X_{2i}; Y_{2i}T_{2i}|S_{1i}).$$

It is easy to verify that $X_{2i} \rightarrow (U_{1i}, V_{1i}) \rightarrow X_{1i}$, $X_{2i} \rightarrow (U_{2i}, V_{2i}) \rightarrow X_{1i}$. Hence by using concave envelope, we have

$$\begin{aligned} R_1 + \lambda R_2 \leq \max_{p_1(x_1)p_2(x_2)} & I(X_1; Y_1T_1|S_2) + \lambda I(X_2; Y_2T_2|S_1) \\ & + \mathcal{C}\left[I(X_1; T_1|T_2S_1X_2) - \lambda I(X_1; Y_2|T_2S_1X_2)\right] \\ & - I(X_1; T_1|T_2S_1X_2) + \lambda I(X_1; Y_2|X_2T_2S_1) \\ & + \mathcal{C}\left[I(X_2; T_2|T_1S_2X_1) - I(X_2; Y_1|T_1S_2X_1)\right] \\ & - I(X_2; T_2|T_1S_2X_1) + I(X_2; Y_1|X_1T_1S_2) \end{aligned}$$

□

Chapter 3

Very weak interference

As discussed in Chapter 1, for channels with *very strong interference*, the best coding strategy for receivers is to decode both messages from the intended signal as well as interference. In contrast, when interference is very weak, one would expect that receivers do not decode any information from interference. This chapter will study channels with *very weak interference*. We are going to focus on sum-capacity only.

3.1 Definition of very weak interference

Definition 3.1.1 (Very weak interference)**.** An interference channel is said to have *very weak interference* if

$$I(U_1; Y_1) \geq I(U_1; Y_2 | X_2) \tag{3.1}$$

$$I(U_2; Y_2) \geq I(U_2; Y_1 | X_1) \tag{3.2}$$

for all auxiliaries (U_1, U_2) with joint distribution $p(u_1, u_2, x_1, x_2) = p(u_1)p(x_1|u_1)p(u_2)p(x_2|u_2)$.

The definition can be interpreted in the following way: the left hand side of (3.1) represents the rate required for Y_1 to decode U_1, which can be considered as partial information about X_1 according to the Markov structure, without other help. The right hand side of

(3.1) represents the rate required for Y_2 to decode the same U_1 under the most favourable situation where Y_2 is fully aware of its intended message X_2. The inequality indicates that the interference is so weak that decoding any part of X_1 by Y_2 would decreases the sum rate, compared to decoding by Y_1. Thus, to maximize $R_1 + R_2$, *intuitively* the best strategy is to treat interference as noise and not decode any part of the interference.

An equivalent definition to the very weak interference is given by the following lemma. This condition is easier to check.

Lemma 3.1.1. The channel has very weak interference if and only if $I(X_1;Y_1) - I(X_1;Y_2|X_2)$ is concave in $p_1(x_1)$ (for a fixed $p_2(x_2)$), and $I(X_2;Y_2) - I(X_2;Y_1|X_1)$ is concave in $p_2(x_2)$ (for a fixed $p_1(x_1)$).

Proof. Since $U_1 \to X_1 \to (X_2, Y_1, Y_2)$ is Markov, observe that

$$\begin{aligned} &I(U_1;Y_1) \geq I(U_1;Y_2|X_2) \iff \\ &\quad I(X_1;Y_1) - I(X_1;Y_2|X_2) \geq I(X_1;Y_1|U_1) - I(X_1;Y_2|U_1,X_2), \end{aligned}$$

which is equivalent to concavity w.r.t. $p_1(x_1)$. Similar reasoning holds for the second equation w.r.t. $p_2(x_2)$. □

3.2 Han–Kobayashi sum-rate for very weak interference channels

Using Fourier-Motzkin elimination, the Han–Kobayashi sum-rate is given by the following lemma.

Lemma 3.2.1 (Han–Kobayashi sum-rate inner bound)**.** The sum-capacity $R_1 + R_2$ of interference channel is achievable if it satisfies

$$R_1 + R_2 \leq I(X_1;Y_1|U_2,Q) + I(X_2;Y_2|U_1,Q), \tag{3.3}$$

$$R_1 + R_2 \leq I(U_2,X_1;Y_1|Q) + I(X_2;Y_2|U_2,U_1,Q), \tag{3.4}$$

$$R_1 + R_2 \leq I(U_1,X_2;Y_2|Q) + I(X_1;Y_1|U_2,U_1,Q), \tag{3.5}$$

$$R_1 + R_2 \leq I(U_2,X_1;Y_1|U_1,Q) + I(U_1X_2;Y_2|U_2,Q), \tag{3.6}$$

for some $p(q)p(u_1,x_1|q)p(u_2,x_2|q)$.

For very weak interference channels, the Han–Kobayashi sum-rate can be significantly simplified.

Theorem 3.2.1. The maximum achievable sum-rate of the Han-Kobayashi inner bound (3.3)-(3.6), denoted as S_{HK} for a DMIC, reduces to

$$S_{HK} = \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1) + I(X_2;Y_2)$$

under the very weak interference condition.

Proof. Treating interference as noise, or in particular, setting $Q = U_1 = U_2 = \emptyset$ (*i.e.* the trivial random variable) gives that $\max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1) + I(X_2;Y_2)$ is achievable. This indicates that

$$S_{HK} \geq \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1) + I(X_2;Y_2). \tag{3.7}$$

Next, note that equation (3.6) satisfies:

$$\begin{aligned}
&I(U_2, X_1; Y_1|U_1, Q) + I(U_1, X_2; Y_2|U_2, Q) \\
&\overset{(a)}{=} I(U_2, X_1; Y_1|Q) - I(U_1; Y_1|Q) + I(U_1, X_2; Y_2|Q) - I(U_2; Y_2|Q) \\
&= I(X_1; Y_1|Q) + I(U_2; Y_1|X_1, Q) - I(U_2; Y_2|Q) \\
&\quad + I(X_2; Y_2|Q) + I(U_1; Y_2|X_2, Q) - I(U_1; Y_1|Q) \\
&\overset{(b)}{\leq} I(X_1; Y_1|Q) + I(X_2; Y_2|Q),
\end{aligned}$$

where (a) holds because $U_1 \to X_1 \to (U_2, X_2, Y_1, Y_2)$, $U_2 \to X_2 \to (U_1, X_1, Y_1, Y_2)$ form Markov chains conditioned on $Q = q$ and (b) holds as an immediate consequence of the definition of very weak interference. Since S_{HK} has to be smaller than the maximum of any of the four expressions, and that the average over Q is dominated by the maximum value, we have $S_{HK} \leq \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1) + I(X_2;Y_2)$. Combining this with (3.7), the proposition is established. □

Remark 3.2.1. In Gaussian settings, X_1, X_2 need to satisfy power constraints and in general,

$$\max_{E[X_i^2] \leq P_i} I(X_1;Y_1) + I(X_2;Y_2) \neq \max_{E[X_i^2] \leq P_i} I(X_1;Y_1|Q) + I(X_2;Y_2|Q)$$

Thus Han–Kobayashi sum-rate reduces to $\max_{E[X_i^2]\leq P_i} I(X_1;Y_1|Q) + I(X_2;Y_2|Q)$.

Remark 3.2.2. To characterize the entire HK region, one needs to maximize $\lambda R_1 + R_2$. Treating-interference-as-noise weighted sum-rate, $\max_{p_1(x_1)p_2(x_2)} \lambda I(X_1;Y_1) + I(X_2;Y_2)$, might not be equivalent to HK weighted sum-rate under very weak interference. Therefore the definition of very weak interference is tailored for sum-rate (*i.e.* $\lambda = 1$).

3.3 Examples

3.3.1 Binary skewed-Z interference channel

The first example is a DMIC with binary input and output.

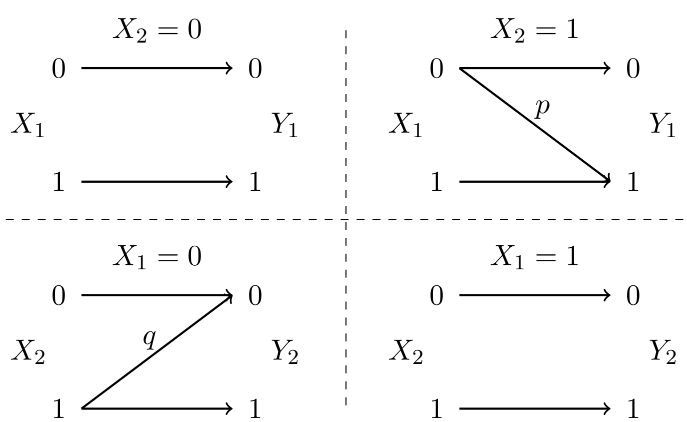

Figure 3.1: Binary skewed-Z interference channel (BSZIC)

Consider a DMIC with transition probabilities as depicted in Figure 3.1 with parameters $p, q \in [0, 1]$. We call such a channel a *Binary Skewed-Z Interference Channel (BSZIC)*.

Lemma 3.3.1. A BSZIC with parameter (p, q) has very weak interference if and only if

$$0 \leq p + q \leq 1.$$

Proof. From Lemma 3.1.1, it suffices to determine the conditions under which $I(X_1;Y_1) - I(X_1;Y_2|X_2)$ is concave in $p_1(x_1)$ for all fixed $p_2(x_2)$

and $I(X_2;Y_2) - I(X_2;Y_1|X_1)$ is concave in $p_2(x_2)$ for all fixed $p_1(x_1)$.

Let $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ denote the binary entropy function. Let $P(X_2 = 0) = a$ and $P(X_1 = 0) = x$. We need to determine the set of values of $p, q \in [0,1]$ such that $I(X_1;Y_1) - I(X_1;Y_2|X_2)$ is concave in x for all $a \in [0,1]$.

$$\begin{aligned} &I(X_1;Y_1) - I(X_1;Y_2|X_2) \\ &\quad = H(x(1-\bar{a}p)) - xH(1-\bar{a}p) - \bar{a}H(xq) + \bar{a}xH(q), \end{aligned}$$

where $\bar{a} = 1 - a$. Note that the second and the last terms are linear in x. After taking the second derivative, one could see that the concavity of the above expression with respect to x is equivalent to showing that

$$\frac{1-\bar{a}p}{1-x(1-\bar{a}p)} \geq \frac{\bar{a}q}{1-xq},$$
$$i.e. \quad (1-\bar{a}p)(1-xq) \geq \bar{a}q(1-x(1-\bar{a}p)).$$

The above condition must hold for all $x \in [0,1]$. Since both sides of the inequality are linear in x, it suffices to verify the inequality only at $x = 0$ and $x = 1$. Substituting them in, we obtain the following two conditions, respectively.

$$\begin{cases} 1-\bar{a}p \geq \bar{a}q, \\ (1-\bar{a}p)(1-q) \geq pq\bar{a}^2. \end{cases}$$

Both conditions have to be satisfied for all $a \in [0,1]$. Keeping in mind that $p, q \in [0,1]$, it is easy to check that this is equivalent to $0 \leq p + q \leq 1$. The same condition can be derived from the concavity of $I(X_2;Y_2) - I(X_2;Y_1|X_1)$. □

The sum-capacity of BSZIC under a sub-regime of very weak interference region can be obtained using genie-based outer bound (2.1).

Theorem 3.3.1. Treating interference as noise is sum-rate optimal for BSZIC when channel parameters (p,q) satisfy

$$p + q + 3pq \leq 1.$$

The regime of parameters is shown in Figure 3.2.

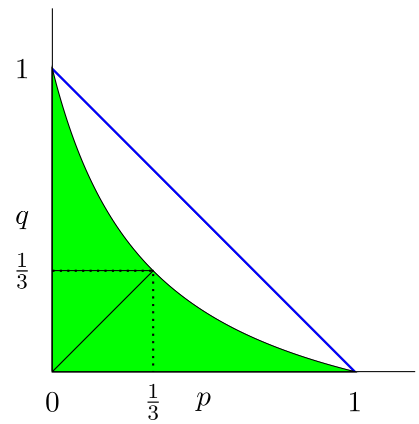

Figure 3.2: Regime of parameters where the sum-capacity is established for the Skewed-Z interference channel

The following lemma aids in our proof of the theorem.

Lemma 3.3.2. Let $\mathcal{C}[f](x, y)$ denote the upper concave envelope of $f(x, y)$ over the space of product distributions where $Pr(X_1 = 0) = x$ and $Pr(X_2 = 1) = y$. Suppose $f(x, y)$ is linear in x. Let $g_0(y) = f(0, y)$ and $g_1(y) = f(1, y)$, then $f(x, y) = (1 - x)g_0(y) + xg_1(y)$ and

$$\mathcal{C}[f](x, y) = (1 - x)\mathcal{C}[g_0](y) + x\mathcal{C}[g_1](y),$$

where $\mathcal{C}[g_0](y)$, $\mathcal{C}[g_1](y)$ denotes the upper concave envelope of $g_0(y)$, $g_1(y)$, w.r.t. $y \in [0, 1]$.

Proof. For a generic variable $x \in [0, 1]$, let $\bar{x} = 1 - x$. Now consider a maximizing convex combination at $(x\bar{y}, xy, \bar{x}\bar{y}, \bar{x}y)$, *i.e.* a weight vector $\{\alpha_i\}$ and product distributions $(x_i\bar{y}_i, x_iy_i, \bar{x}_i\bar{y}_i, \bar{x}_iy_i)$ such that $\sum_i \alpha_i(x_i\bar{y}_i, x_iy_i, \bar{x}_i\bar{y}_i, \bar{x}_iy_i) = (x\bar{y}, xy, \bar{x}\bar{y}, \bar{x}y)$ and that $\sum_i \alpha_i f(x_i, y_i) = \mathcal{C}[f](x, y)$. Note $\sum_i \alpha_i\bar{x}_i = \sum_i \alpha_i\bar{x}_i(\bar{y}_i+y_i) = \bar{x}\bar{y}+\bar{x}y = \bar{x}$, $\sum_i \alpha_ix_i = x$, $\sum_i \alpha_i\bar{x}_iy_i = \bar{x}y$ and $\sum_i \alpha_ix_iy_i = xy$. Therefore,

$$\begin{aligned}\mathcal{C}[f](x, y) &= \sum_i \alpha_i f(x_i, y_i) \\ &= \sum_i (\alpha_i\bar{x}_i f(0, y_i) + \alpha_ix_i f(1, y_i))\end{aligned}$$

$$
\begin{aligned}
&= \bar{x}\left(\sum_i \frac{\alpha_i \bar{x}_i}{\bar{x}} f(0, y_i)\right) + x\left(\sum_i \frac{\alpha_i x_i}{x} f(1, y_i)\right) \\
&\leq \bar{x}\mathcal{C}[g_0](\sum_i \frac{\alpha_i \bar{x}_i}{\bar{x}} y_i) + x\mathcal{C}[g_1](\sum_i \frac{\alpha_i x_i}{x} y_i) \\
&= \bar{x}\mathcal{C}[g_0](y) + x\mathcal{C}[g_1](y).
\end{aligned}
$$

The other direction is immediate as one can always mix the convex combination that achieves $\mathcal{C}[g_0](y)$ and the convex combination that achieves $\mathcal{C}[g_1](y)$ to obtain $(1-x)\mathcal{C}[g_0](y) + x\mathcal{C}[g_1](y)$. □

Proof of Theorem 3.3.1. Let $p_1^*(x_1)p_2^*(x_2)$ be the maximizing input for equation (2.1) when $\lambda = 1$ and $Pr(X_1 = 0) = x^*$, $Pr(X_2 = 1) = y^*$ at $p_1^*(x_1)p_2^*(x_2)$. We will show the existence of a valid pair of genies (T_1, T_2) corresponds to any point of the green region of Figure 3.2 such that the following two conditions hold:

1. $X_i \to Y_i \to T_i$, at $p_1^*(x_1)p_2^*(x_2)$, $i = 1, 2$.
2. $I(X_2; T_2|X_1, T_1) - I(X_2; Y_1|T_1, X_1)$ and $I(X_1; T_1|X_2, T_2) - I(X_1; Y_2|T_2, X_2)$ are concave w.r.t. product distributions $p_1(x_1)p_2(x_2)$.

The above conditions immediately imply that (2.1) reduces to

$$
\begin{aligned}
R_1 + R_2 &\leq I(X_1^*; Y_1) + I(X_2^*; Y_2) \\
&\leq \max_{p_1(x_1)p_2(x_2)} I(X_1; Y_1) + I(X_2; Y_2),
\end{aligned}
$$

which is achievable by treating interference as noise. This establishes the sum-rate capacity.

Remark 3.3.1. One should note that the above conditions, though sufficient, are not necessary for the genie-based sum-rate outer bound to match HK sum-rate inner bound. The second condition could be relaxed to requiring that the functions match their corresponding concave envelopes at $p_1^*(x_1)p_2^*(x_2)$. Requiring the functions to be concave everywhere vanishes the gap terms in (2.1).

For the first condition to hold, given that the valid genies should also satisfy $T_2 \to X_2 \to X_1 \to T_1$ and the channel transition probabilities $q(y_1|x_1, x_2)$, $q(y_2|x_1, x_2)$, one could verify that distributions $p_1(x_1, x_2, y_1, t_1)$ and $p_2(x_1, x_2, y_2, t_2)$ must be of the form given in Table 3.1, where $\{a_i\}, \{b_i\}$ are two generic probability vectors of size $|T_1|$ and $\{c_i\}, \{d_i\}$ are two generic probability vectors of size $|T_2|$. $Pr(X_1 = 0) = x$, $Pr(X_2 = 1) = y$.

X_1	X_2	Y_1	T_1	Probability
0	0	0	i	$x(1-y)((1-p)a_i + pb_i))$
1	0	1	i	$(1-x)(1-y)b_i$
0	1	0	i	$xy(1-p)a_i$
0	1	1	i	$xypb_i$
1	1	1	i	$(1-x)yb_i$,
X_1	X_2	Y_2	T_2	Probability
1	1	1	i	$(1-x)y((1-q)c_i + qd_i))$
0	1	0	i	$xyqd_i$
1	0	0	i	$(1-x)(1-y)d_i$
0	1	1	i	$xy(1-q)c_i$
0	0	0	i	$x(1-y)d_i$,

Table 3.1: Generic probability distribution for genies that satisfy the Markov conditions

Remark 3.3.2. Suppose the Markov chains hold for $Pr(X_1 = 0) = x_*$, $Pr(X_2 = 1) = y_*$, note that our final joint distributions are independent of (x_*, y_*). This is because if the Markov chains hold for some (x_*, y_*), they continue to hold for any other product distribution. This is a chance observation (peculiar to the Binary skewed-Z interference channel) that greatly simplified our analysis.

Next, we will discuss the concavity condition for genies. Define

$f(x,y)$, $\tilde{f}(x,y)$ as

$$f(x,y) := (I(X_2;T_2|X_1,T_1) - I(X_2;Y_1|X_1,T_1))|_{Pr(X_1=0)=x, Pr(X_2=1)=y} ,$$
$$\tilde{f}(x,y) := (I(X_1;T_1|X_2,T_2) - I(X_1;Y_2|X_2,T_2))|_{Pr(X_1=0)=x, Pr(X_2=1)=y} .$$

For a generic variable $x \in [0,1]$, let $\bar{x} = 1-x$ and $L(x) = -x\log_2 x$. Then

$$\begin{aligned} f(x,y) = \sum_i & L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) - yL(\bar{q}c_i + qd_i) \\ & -(xpb_i + x\bar{p}a_i)L\left(\frac{ypb_i}{pb_i+\bar{p}a_i}\right) - x(pb_i+\bar{p}a_i)L\left(\frac{\bar{y}pb_i+\bar{p}a_i}{pb_i+\bar{p}a_i}\right) \\ & +xy(pb_i+\bar{p}a_i)L\left(\frac{pb_i}{pb_i+\bar{p}a_i}\right) + xy(pb_i+\bar{p}a_i)L\left(\frac{\bar{p}a_i}{pb_i+\bar{p}a_i}\right) \end{aligned}$$

Note that $f(x,y)$ is linear in x. Therefore we could write it as the linear combination of two functions $g_0(y) = f(0,y)$ and $g_1(y) = f(1,y)$ as in Proposition 3.3.2.

$$\begin{aligned} g_0(y) &:= \sum_i L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) - yL(\bar{q}c_i + qd_i) , \\ g_1(y) &:= \sum_i L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) - yL(\bar{q}c_i + qd_i) \\ & \quad - (pb_i+\bar{p}a_i)L\left(\frac{ypb_i}{pb_i+\bar{p}a_i}\right) - (pb_i+\bar{p}a_i)L\left(\frac{\bar{y}pb_i+\bar{p}a_i}{pb_i+\bar{p}a_i}\right) \\ & \quad + y(pb_i+\bar{p}a_i)L\left(\frac{pb_i}{pb_i+\bar{p}a_i}\right) + y(pb_i+\bar{p}a_i)L\left(\frac{\bar{p}a_i}{pb_i+\bar{p}a_i}\right) . \end{aligned}$$

Similar for $\tilde{f}(x,y)$, define $\tilde{g}_0(x)$, $\tilde{g}_1(x)$ such that $\tilde{f}(x,y) = (1-y)\tilde{g}_0(x) + y\tilde{g}_1(x)$.

It is sufficient to consider only binary genies. *i.e.* $T_1, T_2 \in \{0,1\}$.

Then, by Proposition 3.3.2, the concavity condition is equivalent to that $g_0(y)$, $g_1(y)$ are concave for all $y \in (0,1)$ and $\tilde{g}_0(x)$, $\tilde{g}_1(x)$ are concave for all $x \in (0,1)$. Since $g_0(y)$, $\tilde{g}_0(x)$ are already concave w.r.t. y, x, respectively. The condition is further reduced to that $g_1(y)$ and

$\tilde{g}_1(x)$ are concave. *i.e.* their second derivatives be non-positive:

$$\sum_{i=0}^{1} -\frac{\bar{q}^2(c_i-d_i)^2}{\bar{y}d_i+y(\bar{q}c_i+pd_i)}+\frac{pb_i}{y}+\frac{p^2b_i^2}{\bar{y}pb_i+\bar{p}a_i}\leq 0 \tag{3.8}$$

$$\sum_{i=0}^{1} -\frac{\bar{p}^2(a_i-b_i)^2}{\bar{x}b_i+y(\bar{p}a_i+pb_i)}+\frac{qd_i}{x}+\frac{q^2d_i^2}{\bar{x}qd_i+\bar{q}c_i}\leq 0 \tag{3.9}$$

Note that in (3.8), either d_0 or d_1 has to be 0 in order to cancel $\frac{pb_i}{y}$ while $y \to 0^+$. Similarly, either b_0 or b_1 has to be zero because of (3.9). Without loss of generosity, we assume that $d_0 = 0$ and $b_0 = 0$. Setting $a_0 = a$, $a_1 = \bar{a}$, $c_0 = c$ and $c_1 = \bar{c}$, (3.8) becomes equivalent to, for all $y \in (0,1)$,

$$\begin{aligned}
&-\frac{\bar{q}c}{y}+\frac{p}{y}-\frac{\bar{q}^2(\bar{c}-1)^2}{\bar{y}+y(\bar{q}\bar{c}+q)}+\frac{p^2}{\bar{y}p+\bar{p}\bar{a}}\leq 0\\
\Leftrightarrow&\frac{p-\bar{p}c}{y}-\frac{\bar{p}^2c^2}{1-y\bar{p}c}+\frac{p^2}{\bar{y}p+\bar{p}\bar{a}}\leq 0\\
\Leftrightarrow&\frac{p}{y}+\frac{p^2}{\bar{y}p+\bar{p}\bar{a}}\leq\frac{\bar{p}c}{y}+\frac{\bar{p}^2c^2}{1-y\bar{p}c}\\
\Leftrightarrow&\frac{p^2+p\bar{p}\bar{a}}{\bar{y}p+\bar{p}\bar{a}}\leq\frac{\bar{p}c}{1-y\bar{p}c}\\
\Leftrightarrow&(p^2+p\bar{p}\bar{a})(1-y\bar{p}c)\leq(\bar{p}c)(\bar{y}p+\bar{p}\bar{a}),\forall y\in(0,1)
\end{aligned} \tag{3.10}$$

As the expression is linear in y on both sides, it suffices to check the validity of (3.10) for when $y = 0$ and $y = 1$, *i.e.* (3.10) is equivalent to

$$\begin{cases} p \leq \bar{q}c, \\ p+\dfrac{p^2}{\bar{p}\bar{a}}\leq\dfrac{\bar{q}c}{1-\bar{q}c}. \end{cases}$$

Rearranging the first inequality we get

$$\begin{cases} \dfrac{p}{\bar{p}} \leq \dfrac{\bar{q}c}{1-\bar{q}c}, \\ p+\dfrac{p^2}{\bar{p}\bar{a}}\leq\dfrac{\bar{q}c}{1-\bar{q}c}. \end{cases}$$

Note that $p+\frac{p^2}{\bar{p}\bar{a}} = p(1+\frac{p/\bar{a}}{\bar{p}}) \geq p(1+\frac{p}{\bar{p}}) = \frac{p}{\bar{p}}$. Therefore, the first inequality is redundant and we are left with a single constraint

$$p+\frac{p^2}{\bar{p}\bar{a}}\leq\frac{\bar{q}c}{1-\bar{q}c}.$$

Similarly, inequality (3.9) is equivalent to the following,

$$q + \frac{q^2}{\bar{q}\bar{c}} \leq \frac{\bar{p}a}{1 - \bar{p}a}.$$

Further, without loss of generality, we assume $p \leq q$. Putting all the conditions together, we get

$$0 \leq a \leq 1 \tag{3.11}$$

$$0 \leq c \leq 1 \tag{3.12}$$

$$0 \leq p \leq q \leq 1 \tag{3.13}$$

$$0 \leq p + q \leq 1 \tag{3.14}$$

$$p + \frac{p^2}{\bar{p}\bar{a}} \leq \frac{\bar{q}c}{1 - \bar{q}c} \tag{3.15}$$

$$q + \frac{q^2}{\bar{q}\bar{c}} \leq \frac{\bar{p}a}{1 - \bar{p}a} \tag{3.16}$$

Rearranging (3.15), we have

$$\bar{p}a \leq \frac{\bar{p}\bar{q}c - p\bar{p}}{\bar{q}c - p^2\bar{q}c - p\bar{p}}$$

$$\frac{\bar{p}a}{1 - \bar{p}a} \leq \frac{\bar{q}c - p}{p\bar{q}c}$$

Note

$$\frac{\bar{q}c - p}{p\bar{q}c} = \frac{1 - p/\bar{q}c}{p} \leq \frac{\bar{p}}{1 - \bar{p}}$$

This means (3.11) is redundant.

Combining with (3.16) we have the condition

$$\frac{q\bar{q}\bar{c} + q^2}{\bar{c}} \leq \frac{\bar{q}c - p}{pc}$$

$$(1 - pq)\bar{q}c^2 - (1 + p)\bar{q}c + p \leq 0 \tag{3.17}$$

This inequality must holds for some $c \in [0, 1]$.

When $c = \frac{1+p}{2(1-pq)}$. $0 \leq c \leq 1$ is given by the following

$$0 \leq \frac{1 + p}{2(1 - pq)} = \frac{1 + p}{1 + (1 - 2pq)} \leq \frac{1 + p}{1 + (1 - q)} \leq \frac{1 + p}{1 + (1 - \bar{p})} = 1$$

where first inequality is due to $p \leq \frac{1}{2}$ and the second one is due to $q \leq \bar{p}$. So we can let $c = \frac{1+p}{2(1-pq)}$.

Then inequality (3.17) gives

$$p - \frac{(1+p)^2\bar{q}}{4(1-pq)} \leq 0$$

$$q \leq \frac{1-p}{1+3p}$$

To satisfy (3.13), we need $p \leq \frac{1-p}{1+3p}$. That is $0 \leq p \leq \frac{1}{3}$.

Same analysis can be applied to the case $q \leq p$.

Hence we derive the conditions for the existence of smart and useful genie,

$$\begin{aligned} 0 \leq p \leq \frac{1}{3}, \\ p \leq q \leq \frac{1-p}{1+3p}, \end{aligned} \quad \text{or} \quad \begin{aligned} 0 \leq q \leq \frac{1}{3}, \\ q \leq p \leq \frac{1-q}{1+3q}. \end{aligned}$$

It is easy to verify that this region is equivalent to requiring $p+q+3pq \leq 1$ and $p, q \geq 0$. □

The above theorem provides sum-rate capacity for a certain range of (p, q) for BSZIC by showing genie-based outer bound (2.1) matches treating-interference-as-noise inner bound. However, it is not true for any (p, q) that satisfies very weak interference conditions. The following lemma indicates that either genie-based outer bound (2.1) is not always tight or treating interference as noise is not always sum-rate optimal for BSZIC with very weak interference.

Lemma 3.3.3. For the binary skewed-Z interference channel when $p = q = \frac{1}{2}$, the genie based outer bound is strictly greater than the treating interference as noise inner bound.

Proof. Define $f(x, y)$, $g_0(y)$ and $g_1(y)$ in the same way as before. The joint distribution is the same as defined in Table 3.1.

Setting $p = q = \frac{1}{2}$ and taking second derivative of $g_1(y)$, we get

$$\begin{aligned}
\frac{d^2 g_1(y)}{dy^2} &= \sum_i \left(-\frac{(c_i - d_i)^2}{2y(c_i - d_i) + 4d_i} + \frac{b_i}{2y} + \frac{b_i^2}{2\bar{y}b_i + 2a_i} \right) \\
&= -\sum_i \frac{(c_i - d_i)^2}{2y(c_i - d_i) + 4d_i} + \sum_i \frac{b_i}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\
&\geq -\sum_i \frac{c_i^2 + d_i^2}{2y(c_i - d_i) + 4d_i} + \sum_i \frac{b_i}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\
&= -\sum_i \frac{c_i^2}{2yc_i - 2yd_i + 4d_i} - \sum_i \frac{d_i^2}{2yc_i - 2yd_i + 4d_i} \\
&\quad + \frac{1}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\
&\geq -\sum_i \frac{c_i^2}{2yc_i} - \sum_i \frac{d_i^2}{-2yd_i + 4d_i} + \frac{1}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\
&= -\frac{1}{2y} - \frac{1}{-2y + 4} + \frac{1}{2y} + \frac{\bar{y} + 1}{2} \left(\sum_i \frac{\bar{y}b_i + a_i}{\bar{y} + 1} \frac{b_i^2}{(\bar{y}b_i + a_i)^2} \right) \\
&\overset{(a)}{\geq} -\frac{1}{-2y + 4} + \frac{\bar{y} + 1}{2} \left(\sum_i \frac{\bar{y}b_i + a_i}{\bar{y} + 1} \frac{b_i}{\bar{y}b_i + a_i} \right)^2 \\
&= -\frac{1}{-2y + 4} + \frac{1}{2(\bar{y} + 1)} \\
&= 0,
\end{aligned}$$

where (a) holds because $\mathrm{E}(X^2) \geq \mathrm{E}(X)^2$. Thus $g_1(y)$ is convex in general. The only hope for the outer bound to work would be that $g_1(y)$ was a straight line. Next we analyze if this is possible.

Note $\frac{d^2 g_1(y)}{dy^2} = 0$ would imply that $c_i d_i = 0$ (for the first inequality to be equality) and $a_i = b_i$ (for the inequality labeled (a) to be an equality).

For the symmetric condition to hold, define $\tilde{f}(x, y)$ as

$$I(X_1; T_1 | X_2 T_2) - I(X_1; Y_2 | T_2 X_2)|_{Pr(X_1=0)=x, Pr(X_2=1)=y}$$

Split $\tilde{f}(x, y)$ in same way as for $f(x, y)$,

$$\tilde{f}(x, y) = (1 - y)\tilde{g}_0(x) + y\tilde{g}_1(x)$$

Computing derivative of $\tilde{g}_1(x)$, we have

$$\frac{d^2\tilde{g}_1(x)}{dx^2} \geq 0$$

with equality holding only *iff* $a_i b_i = 0$ and $c_i = d_i$.

Clearly, both equalities cannot hold at the same time. At least one of g_1 and $\tilde{g}_1$ is strictly convex. Therefore, for any $(x, y) \in (0,1)^2$,

$$\begin{aligned}
&\mathfrak{C}[f](x,y) + \mathfrak{C}[\tilde{f}](x,y) \\
=& x\mathfrak{C}[g_0](y) + (1-x)\mathfrak{C}[g_1](y) + y\mathfrak{C}[\tilde{g}_0](x) + (1-y)\mathfrak{C}[\tilde{g}_1](x) \\
>& xg_0(y) + (1-x)g_1(y) + y\tilde{g}_0(x) + (1-y)\tilde{g}_1(x) \\
=& f(x,y) + f_{c_i,d_i,a_i,b_i}(y,x)
\end{aligned}$$

Hence outer bound (2.1) is strictly larger than $I(X_1;Y_1)+I(X_2;Y_2)$ for any independent distribution of X_1, X_2. $\square$

3.3.2 Gaussian Z interference channel

The second example of very weak interference channel is Gaussian Z interference channel.

Set $b = 0$ in Gaussian interference channel model.

$$\begin{aligned}
Y_1 &= X_1 + Z_1 \\
Y_2 &= X_2 + aX_1 + Z_2
\end{aligned}$$

where X_1, X_2 are independent continues random variables with $E[X_1^2] \leq P_1$ and $E[X_2^2] \leq P_2$. Z_1, Z_2 are independent Gaussian noise $\mathcal{N}(0,1)$ and $0 \leq a \leq 1$. See Figure 3.3.

Lemma 3.3.4. A Gaussian Z interference channel as described in Figure 3.3 has very weak interference.

Proof. Let $U_1 \to X_1 \to (Y_1, Y_2)$, $U_2 \to X_2 \to (Y_1, Y_2)$. Then

$$I(U_2; Y_1|X_1) = I(U_2; X_1 + Z_1|X_1) = I(U_2; Z_1) = 0$$

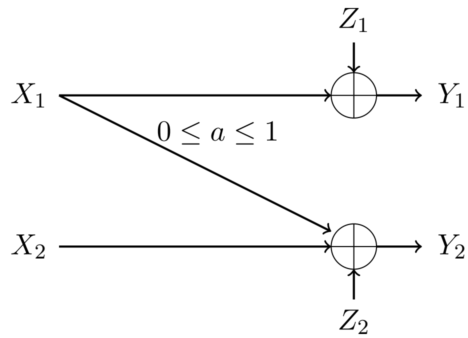

Figure 3.3: Gaussian Z interference channel

Hence $I(U_2; Y_1|X_1) \le I(U_2; Y_2)$. The second inequality is established as follow,

$$\begin{aligned} I(U_1; Y_2|X_2) &= I(U_1; X_2 + aX_1 + Z_2|X_2) \\ &= I(U_1; aX_1 + Z_2) \\ &= I(U_1; X_1 + \frac{1}{a} Z_2) \\ &\le I(U_1; X_1 + Z_1) \\ &= I(U_1; Y_1) \end{aligned}$$

where the inequality holds because $U_1 \to X_1 + Z_1 \to X_1 + \frac{1}{a} Z_2$ is stochastically degraded. $\square$

Sum-capacity is known since it satisfies (1.8). In fact, the corner point $(R_1, R_2) = \left(\frac{1}{2}\log(1 + P_1), \frac{1}{2}\log\left(1 + \frac{P_2}{1+a^2 P_1}\right)\right)$ of capacity region attains the sum-capacity.

3.4 Mixed interference

For the sake of completeness, we discuss DMICs of which one sender produces strong interference while the other produces very weak interference. Sum-capacity of mixed Gaussian interference channel is known.

Lemma 3.4.1. Consider a DMIC satisfying

$$I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2) \ \forall p(x_1)p(x_2), \tag{3.18}$$
$$I(U_2;Y_2) \geq I(U_2;Y_1|X_1) \ \forall p(x_1)p(u_2)p(x_2|u_2) \tag{3.19}$$

The sum-capacity of this channel is

$$\max_{p(x_1)p(x_2)} \min\{I(X_1;Y_1) + I(X_2;Y_2|X_1), I(X_1X_2;Y_2)\}$$

Proof. To show the sum-capacity, we use the traditional outer bound provided in (1.9).

$$R_1 + R_2 \leq I(X_1;Y_1|X_2Q) + I(X_2;Y_2|X_1Q) \tag{3.20a}$$
$$R_1 + R_2 \leq I(U_2X_1;Y_1|Q) + I(U_1X_2;Y_2|Q) \tag{3.20b}$$
$$R_1 + R_2 \leq I(U_2X_1;Y_1|Q) + I(X_2;Y_2|U_2X_1Q) \tag{3.20c}$$
$$R_1 + R_2 \leq I(U_1X_2;Y_2|Q) + I(X_1;Y_1|U_1X_2Q) \tag{3.20d}$$

for some $p(q)p(u_1,x_1|q)p(u_2,x_2|q)$.

By the data processing inequality and the strong interference condition (3.18),

$$\begin{aligned} I(U_1X_2;Y_2|Q) &\leq I(X_1X_2;Y_2|Q) \\ I(U_1X_2;Y_2|Q) + I(X_1;Y_1|U_1X_2Q) &\leq I(U_1X_2;Y_2|Q) + I(X_1;Y_2|U_1X_2Q) \\ &= I(X_1X_2;Y_2|Q). \end{aligned}$$

This indicates setting $U_1 = X_1$ is optimal for the sum-rate outer bound since both (3.20b) and (3.20b) are maximized. Then the outer bound reduces to

$$R_1 + R_2 \leq I(X_1;Y_1|X_2Q) + I(X_2;Y_2|X_1Q) \tag{3.21a}$$
$$R_1 + R_2 \leq I(U_2X_1;Y_1|Q) + I(X_2;Y_2|U_2X_1Q) \tag{3.21b}$$
$$R_1 + R_2 \leq I(X_1X_2;Y_2|Q) \tag{3.21c}$$

Now observe that from the weak interference condition (3.19).

$$\begin{aligned} I(U_2;Y_1|X_1Q) &\leq I(U_2;Y_2|Q) \\ &\leq I(U_2;X_1Y_2|Q) = I(U_2;Y_2|X_1Q) \end{aligned}$$

It follows

$$\begin{aligned}
&I(U_2X_1;Y_1|Q) + I(X_2;Y_2|U_2X_1Q)\\
&\quad = I(X_1;Y_1|Q) + I(U_2;Y_1|X_1Q) + I(X_2;Y_2|U_2X_1Q)\\
&\quad \leq I(X_1;Y_1|Q) + I(U_2;Y_2|X_1Q) + I(X_2;Y_2|U_2X_1Q)\\
&\quad = I(X_1;Y_1|Q) + I(X_2;Y_2|X_1Q)
\end{aligned}$$

and which indicates setting $U_2 = \emptyset$ is optimal. Therefore (3.21a) is redundant and the outer bound reduces to

$$\begin{aligned}
R_1 + R_2 &\leq \min\{I(X_1;Y_1|Q) + I(X_2;Y_2|X_1Q), I(X_1X_2;Y_2|Q)\}\\
&\leq \max_{p(x_1)p(x_2)} \min\{I(X_1;Y_1) + I(X_2;Y_2|X_1), I(X_1X_2;Y_2)\}
\end{aligned}$$

This rate is achievable by Han–Kobayashi sum-rate inner bound (3.3)-(3.6) with auxiliaries chosen as $U_1 = X_1$, $U_2 = \emptyset$ and $Q = \emptyset$. □

3.5 Open questions about very weak interference conditions

There are two questions of interest. The first one is whether our definition on very weak interference can extend to n-letter.

Question: Consider two very weak interference channels with transitional probability $w_1(y_{11}, y_{21}|x_{11}, x_{21})$ and $w_2(y_{12}, y_{22}|x_{12}, x_{22})$. Does the product channel $w_1 \otimes w_2$ also has very weak interference?

The product channel $w_1 \otimes w_2$ has two input (x_{11}, x_{12}) and (x_{21}, x_{22}), two output (y_{11}, y_{12}) and (y_{21}, y_{22}), and transitional probability

$$\begin{aligned}
&w(y_{11}, y_{12}, y_{21}, y_{22}|x_{11}, x_{12}, x_{21}, x_{22})\\
&\qquad = w_1(y_{11}, y_{21}|x_{11}, x_{21})w_2(y_{12}, y_{22}|x_{12}, x_{22}).
\end{aligned}$$

The Han–Kobayashi (HK) sum-rate is equivalent to the treating-interference-as-noise (TIN) sum-rate for very weak interference channels. If the answer to this question is yes, then the n-letter HK sum-rate reduces to the n-letter TIN sum-rate for very weak interference channels. Then it can be concluded that both HK and TIN n-letter expressions converge to sum-capacity at same rate of convergence.

Another question is whether TIN tensorizes, i.e. n-letter TIN is equivalent one-letter TIN.

Question: Consider two very weak interference channels with transitional probability w_1 and w_2 and their product channel $w_1 \otimes w_2$. Is the following true?

$$\begin{aligned} &\max_{p(q,x_{11},x_{12},x_{21},x_{22})} I(X_{11}X_{12};Y_{11}Y_{12}|Q) + I(X_{21}X_{22};Y_{21}Y_{22}|Q) \\ &\quad \le \max_{p(q_1,x_{11},x_{21})} I(X_{11};Y_{11}|Q_1) + I(X_{21};Y_{21}|Q_1) \\ &\qquad + \max_{p(q_2,x_{12},x_{22})} I(X_{12};Y_{12}|Q_2) + I(X_{22};Y_{22}|Q_2) \end{aligned}$$

If the answer is yes, then the one-letter TIN sum-rate is the sum-capacity of very weak interference channel.

Chapter 4

Gaussian interference channels

This chapter examines Gaussian interference channels (GICs). First, we employ the enhanced genie-based outer bound on sum-capacity ($\lambda = 1$) of GICs. We will see that, like the Han–Kobayashi inner bound, the enhanced genie-based outer bound is tight for all the regimes where capacity has been established. In the second part, we study the optimality of Gaussian signalling for the TIN sum rate using Hermite polynomials. At last, we consider Gaussian Z interference channels. A hypothesis related to optimality of Han–Kobayashi region with Gaussian signalling is proposed.

4.1 Optimality for Gaussian Interference Channel with Strong Interference

As introduced in Chapter 1, when $a \geq 1$, $b \geq 1$, Gaussian interference channels have strong interference. The sum-capacity is therefore given by

$$\min\{\frac{1}{2}\log(1+P_1+b^2P_2), \frac{1}{2}\log(1+P_2+a^2P_1)\}.$$

In the enhanced genie-based outer bound (2.2), set $S_1 = \emptyset$, $T_2 = \emptyset$ and $S_2 = X_2$. Since $a \geq 1$, we could find independent $\dot{Z}_1, \ddot{Z}_1 \sim$

$\mathcal{N}(0,1)$ such that $Z_1 = \frac{1}{a}\dot{Z}_1 + \frac{\sqrt{a^2-1}}{a}\ddot{Z}_1$. Thus $Y_1 = X_1 + bX_2 + \frac{1}{a}\dot{Z}_1 + \frac{\sqrt{a^2-1}}{a}\ddot{Z}_1$. Let $T_1 = X_1 + \frac{1}{a}\dot{Z}_1$. Then the genie-based sum-rate outer bound becomes

$$\begin{aligned}
R_1 + R_2 \le & \max_{p_1(x_1)p_2(x_2)} I(X_1;T_1,Y_1|S_2) + I(X_2;T_2,Y_2|S_1) \\
& + \mathcal{C}[I(X_1;T_1|X_2,T_2,S_1) - I(X_1;Y_2|X_2,T_2,S_1)] \\
& - I(X_1;T_1|X_2,T_2,S_1) + I(X_1;Y_2|X_2,T_2,S_1) \\
& + \mathcal{C}[I(X_2;T_2|X_1,T_1,S_2) - I(X_2;Y_1|X_1,T_1,S_2)] \\
& - I(X_2;T_2|X_1,T_1,S_2) + I(X_2;Y_1|X_1,T_1,S_2) \\
= & \max_{p_1(x_1)p_2(x_2)} I(X_1;X_1 + \frac{1}{a}\dot{Z}_1, X_1 + \frac{1}{a}\dot{Z}_1 + \frac{\sqrt{a^2-1}}{a}\ddot{Z}_1) \\
& + I(X_2;Y_2) + \mathcal{C}[I(X_1;X_1 + \frac{1}{a}\dot{Z}_1) - I(X_1;aX_1 + Z_2)] \\
& - I(X_1;X_1 + \frac{1}{a}\dot{Z}_1) + I(X_1;Y_2|X_2) \\
\overset{*}{=} & \max_{p_1(x_1)p_2(x_2)} I(X_2;Y_2) + I(X_1;Y_2|X_2) \\
= & \max_{p_1(x_1)p_2(x_2)} I(X_1X_2;Y_2|X_2) \\
\overset{**}{=} & \frac{1}{2}\log(1 + P_2 + a^2 P_1)
\end{aligned}$$

where (*) holds because $X_1 \to X_1 + \frac{1}{a}\dot{Z}_1 \to X_1 + \frac{1}{a}\dot{Z}_1 + \frac{\sqrt{a^2-1}}{a}\ddot{Z}_1$ is degraded and (**) holds because the Gaussian inputs maximize the expression. Symmetrically, since $b \ge 1$ we can get

$$R_1 + R_2 \le \frac{1}{2}\log(1 + P_1 + b^2 P_2).$$

Hence the enhanced genie-based outer bound is tight for Gaussian interference channels with strong interference.

4.2 Optimality for Gaussian Interference Channel with Mixed Interference

When $a \ge 1$ and $b < 1$ or $a < 1$ and $b \ge 1$, the Gaussian interference channel has strong interference from only one sender-receiver pair. For

$a \geq 1$ and $b < 1$, the sum-rate capacity is

$$\min\{\frac{1}{2}\log(1+P_2+a^2P_1), \frac{1}{2}\log\left(1+\frac{P_1}{b^2P_2+1}\right)+\frac{1}{2}\log\left(1+P_2\right)\}.$$

Since $a \geq 1$, the argument in previous section shows enhanced genie-based outer bound is no greater than $\frac{1}{2}\log(1+P_2+a^2P_1)$.

Since $b \leq 1$ and capacity does not depend on correlation between Z_1 and Z_2, we could find independent $\ddot{Z}_1 \sim \mathcal{N}(0,1)$ such that $Z_1 = bZ_2+\sqrt{1-b^2}\ddot{Z}_1$. Setting $T_1=\emptyset$, $S_1=X_1$, $T_2=X_2+\frac{1}{b}Z_1=X_2+Z_2+\frac{\sqrt{1-b^2}}{b}\ddot{Z}_1$ and $S_2=\emptyset$, the enhanced genie-based sum-rate outer bound becomes

$$\begin{aligned}
R_1+R_2 \leq & \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1)+I(X_2;Y_2,X_2+Z_2+\frac{\sqrt{1-b^2}}{b}\ddot{Z}_1|X_1) \\
& +\mathcal{C}\left[I(X_2;X_2+\frac{1}{b}Z_1|X_1)-I(X_2;bX_2+Z_1|X_1)\right] \\
& -I(X_2;X_2+\frac{1}{b}Z_1|X_1)+I(X_2;bX_2+Z_1|X_1) \\
= & \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1)+I(X_2;X_2+Z_2,X_2+Z_2+\frac{\sqrt{1-b^2}}{b}\ddot{Z}_1|X_1) \\
\overset{*}{=} & \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1)+I(X_2;X_2+Z_2|X_1) \\
= & \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1)+I(X_2;X_2+aX_1+Z_2|X_1) \\
= & \max_{p_1(x_1)p_2(x_2)} I(X_1;Y_1)+I(X_2;Y_2|X_1) \\
\overset{**}{=} & \frac{1}{2}\log\left(1+\frac{P_1}{b^2P_2+1}\right)+\frac{1}{2}\log\left(1+P_2\right),
\end{aligned}$$

where (*) holds because $X_2 \to X_2+Z_2 \to X_2+Z_2+\frac{\sqrt{1-b^2}}{b}\ddot{Z}_1$ is degraded and (**) holds because the Gaussian inputs maximize the expression.

Hence the enhanced genie-based sum-rate outer bound is tight for Gaussian interference channels with mixed interference.

4.3 A closed form of the enhanced genie-based outer bound

Before going to the discussion in the weak interference regime, we provide a closed form of the outer bound under following settings of genie random variable.

We could find independent $\dot{Z}_1, \ddot{Z}_1, \tilde{Z}_1 \sim \mathcal{N}(0,1)$ such that $Z_1 = \rho_{11}\dot{Z}_1 + \rho_{12}\ddot{Z}_1 + \rho_{13}\tilde{Z}_1$ and independent $\dot{Z}_2, \ddot{Z}_2, \tilde{Z}_2 \sim \mathcal{N}(0,1)$ such that $Z_2 = \rho_{21}\dot{Z}_2 + \rho_{22}\ddot{Z}_2 + \rho_{23}\tilde{Z}_2$, where $\rho_{11}^2 + \rho_{12}^2 + \rho_{13}^2 = 1$ and $\rho_{21}^2 + \rho_{22}^2 + \rho_{23}^2 = 1$. Let $T_1 = X_1 + \eta_1\dot{Z}_1$, $S_2 = X_2 + \mu_2\ddot{Z}_1$, and $T_2 = X_2 + \eta_2\dot{Z}_2$, $S_1 = X_1 + \mu_1\ddot{Z}_2$.

Again taking advantage of the fact that capacity do not depend on correlation between noise Z_1 and Z_2, we may assume $\eta_1\dot{Z}_1$ and $\mu_1\ddot{Z}_2$ are correlated so that $X_1 \to T_1 \to S_1$ if $\mu_1 \geq \eta_1$ or $X_1 \to S_1 \to T_1$ if $\mu_1 \leq \eta_1$. So do $\eta_2\dot{Z}_2$ and $\mu_2\ddot{Z}_1$. Thus genie random variables T_i, S_i, $i = 1, 2$ are valid choices. It is ready to evaluate the enhanced genie-based outer bound under this setting.

Lemma 4.3.1. Assume $X_i \sim \mathcal{N}(0, P_i)$ is the optimal distribution for (2.2) under aforementioned setting, then sum-capacity must satisfy,

$$\begin{aligned}\min_{\rho,\eta,\mu} \quad & \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right) \\ & + \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right) \\ & + \left[\frac{1}{2}\log\left(1 + P_1\frac{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}{\mu_1^2\rho_{23}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_1}{\eta_1^2 \wedge \mu_1^2}\right)\right]_+ \\ & + \left[\frac{1}{2}\log\left(1 + P_2\frac{(\rho_{12} - b\mu_2)^2 + \rho_{13}^2}{\mu_2^2\rho_{13}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_2}{\eta_2^2 \wedge \mu_2^2}\right)\right]_+,\end{aligned} \tag{4.1}$$

where $a \wedge b = \min\{a, b\}$ and $[a]_+ = \max\{a, 0\}$, or equivalently,

$$\min_{\rho,\eta,\mu} \max\{B_1, B_2, B_3, B_4\},$$

where

$$B_1 = \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right);$$

$$B_2 = \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_1\frac{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}{\mu_1^2\rho_{23}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_1}{\eta_1^2 \wedge \mu_1^2}\right);$$

$$B_3 = \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\frac{(\rho_{12} - b\mu_2)^2 + \rho_{13}^2}{\mu_2^2\rho_{13}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_2}{\eta_2^2 \wedge \mu_2^2}\right);$$

$$B_4 = \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right)$$
$$+ \frac{1}{2}\log\left(1 + P_1\frac{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}{\mu_1^2\rho_{23}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_1}{\eta_1^2 \wedge \mu_1^2}\right)$$
$$+ \frac{1}{2}\log\left(1 + P_2\frac{(\rho_{12} - b\mu_2)^2 + \rho_{13}^2}{\mu_2^2\rho_{13}^2}\right) - \frac{1}{2}\log\left(1 + \frac{P_2}{\eta_2^2 \wedge \mu_2^2}\right);$$

Proof. See Appendix 4.A for detail. □

4.4 Optimality for Gaussian Interference Channel with weak Interference

The optimization problem in Lemma 4.3.1 is hard to solve analytically. We consider the case

$$\max\{B_1, B_2, B_3, B_4\} = B_1.$$

In this case, we are going to find the condition for which outer bound (4.1) is tight.

Let η_1^*, η_2^* be defined as

$$\frac{1}{\eta_1^*} = \frac{\rho_{11}}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{11}^2 + \rho_{13}^2}$$

$$\frac{1}{\eta_2^*} = \frac{\rho_{21}}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{21}^2 + \rho_{23}^2}$$

In order to satisfy $\max\{B_1, B_2, B_3, B_4\} = B_1$ when $\eta_1 = \eta_1^*$, $\eta_2 = \eta_2^*$, it requires

$$\frac{\rho_{23}^2 + (\rho_{22} - a\mu_1)^2}{\mu_1^2\rho_{23}^2} \le \frac{1}{\eta_1^{*2}} \tag{4.2}$$

$$\frac{\rho_{13}^2 + (\rho_{12} - b\mu_2)^2}{\mu_2^2\rho_{13}^2} \le \frac{1}{\eta_2^{*2}} \tag{4.3}$$

Under these conditions, it can be concluded that B_1 attains minimum by setting $\eta_1 = \eta_1^*$, $\eta_2 = \eta_2^*$. Outer bound (4.1) reduces to

$$\frac{1}{2}\log(1 + \frac{P_1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{11}^2 + \rho_{13}^2}) + \frac{1}{2}\log(1 + \frac{P_2}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{21}^2 + \rho_{23}^2})$$

Furthermore, if $\mu_2 \to \infty$, $\rho_{12} = 0$ and $\mu_1 \to \infty$, $\rho_{22} = 0$, the outer bound reduces to the treating interference as noise inner bound. To make conditions (4.2) and (4.3) continue to hold, it requires

$$\frac{a^2}{\rho_{23}^2} \le \frac{\rho_{11}^2}{(b^2P_2 + 1)^2}$$
$$\frac{b^2}{\rho_{13}^2} \le \frac{\rho_{21}^2}{(a^2P_1 + 1)^2}$$

or equivalently,

$$a(b^2P_2 + 1) + b(a^2P_1 + 1) \leq 1.$$

The following Figure 4.1 shows the enhanced genie-based outer bound for symmetric GICs with $P_1 = P_2 = 100$.

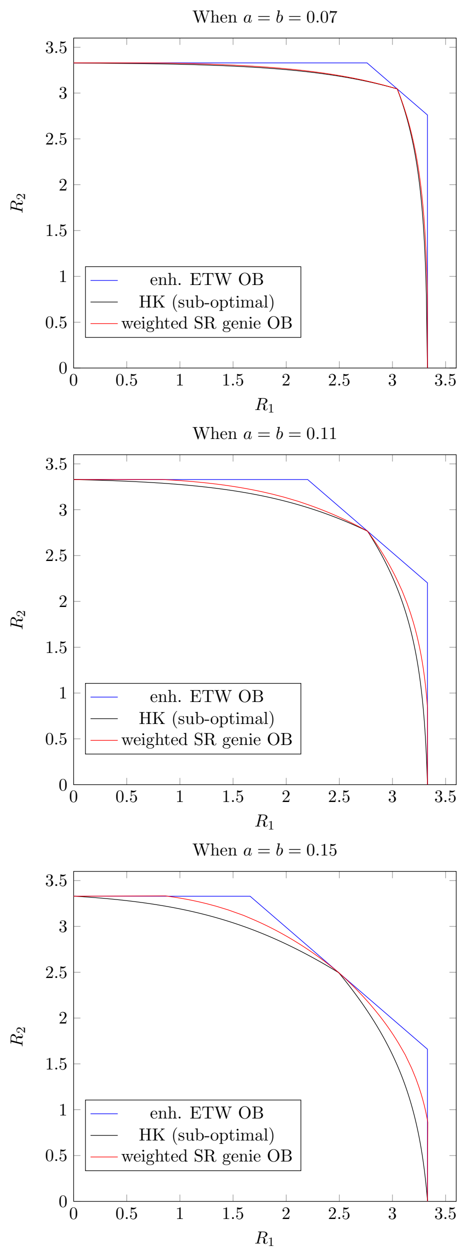

Figure 4.1: Inner and outer bounds for symmetric GICs

4.5 Hermite perturbation on Gaussian distribution for TIN

From the previous section, the treating-interference-as-noise (TIN) sum-rate inter bound is optimal in a sub-regime of Gaussian interference channel with weak interference. Moreover, Gaussian distribution is the optimal input distribution that maximizes TIN sum-rate in the sub-regime. However, Gaussian input is not always optimal. This section examines the sub-optimality of Gaussian input for TIN using Hermite polynomials to perturb Gaussian distribution, as proposed in [1]. For simplicity, we consider symmetric Gaussian interference channel, where $a = b$, $P_1 = P_2 = p$.

Denote normal probability density function $\mathcal{N}(0, p)$ as

$$g^p(x) = \frac{1}{\sqrt{2\pi p}} e^{-\frac{x^2}{2p}}.$$

Hermite polynomial with degree k and power p is defined as

$$H_k^p(x) = \frac{(-1)^k \sqrt{p}^k}{\sqrt{k!}} e^{x^2/2p} \frac{d^k}{dx^k} e^{-\frac{x}{2p}}, \quad k = 1, 2 \ldots,$$

$$H_0^p(x) = 1.$$

In particular,

$$H_1^p(x) = \frac{x}{\sqrt{p}},$$

$$H_2^p(x) = \frac{1}{\sqrt{2}}(\frac{x^2}{p} - 1),$$

$$H_4^p(x) = \frac{x^4}{p^2} - 6\frac{x^2}{p} + 3.$$

We may omit variable x without confusions and use g^p, H_k^p in shorthand.

One merit of using Hermite polynomial is its invariance under convolution.

Lemma 4.5.1. We have

$$g^p H_k^p * g^q H_l^q = c_{k,l}^{p,q} g^{p+q} H_{k+l}^{p+q}$$

where $c_{k,l}^{p,q} = \frac{\sqrt{(k+l)!p^k q^l}}{\sqrt{k!l!(p+q)^{k+l}}}$ is a constant.

Proof. By linear property of convolution for differentiation,

$$\begin{aligned}
g^p H_k^p * g^q H_l^q &= \frac{(-1)^k \sqrt{p}^k}{\sqrt{k!}} \frac{d^k}{dx^k} \frac{e^{-\frac{x}{2p}}}{\sqrt{2\pi p}} * \frac{(-1)^l \sqrt{q}^l}{\sqrt{l!}} \frac{d^l}{dx^l} \frac{e^{-\frac{x}{2q}}}{\sqrt{2\pi q}} \\
&= \frac{(-1)^k \sqrt{p}^k}{\sqrt{k!}} \frac{(-1)^l \sqrt{q}^l}{\sqrt{l!}} * \frac{d^{k+l}}{dx^{k+l}} \frac{e^{-\frac{x}{2(p+q)}}}{\sqrt{2\pi(p+q)}} \\
&= \frac{\sqrt{(k+l)!p^k q^l}}{\sqrt{k!l!(p+q)^{k+l}}} g^{p+q} H_{k+l}^{p+q}.
\end{aligned}$$

□

When $l = 0$, we have

$$g^p H_k^p * g^q = \frac{\sqrt{p^k}}{\sqrt{(p+q)^k}} g^{p+q} H_k^{p+q}.$$

The following two lemmas are used to compute difference of differential entropy after perturbation on Gaussian distribution.

Lemma 4.5.2. Let $g_\epsilon(x)$ be a probability density function and $h(g_\epsilon)$ denotes differential entropy of some random variable with distribution g_ϵ. Then

$$h(g_\epsilon) - h(g^p) = -D(g_\epsilon || g^p) + \frac{1}{2} \int \frac{x^2}{p} (g_\epsilon(x) - g^p(x)) \mathrm{d}x$$

Proof. By definition of KL divergence,

$$\begin{aligned}
D(g_\epsilon || g^p) &= \int g_\epsilon(x) \log g_\epsilon(x) - g_\epsilon(x) \log g^p(x) \\
&= -h(g_\epsilon) - \int g_\epsilon(x)(-\frac{1}{2} \log 2\pi p - \frac{x^2}{2p}) \\
&= -h(g_\epsilon) + \frac{1}{2} \log 2\pi p + \int g_\epsilon(x)(\frac{x^2}{2p}) \\
&= -h(g_\epsilon) + h(g^p) - 1/2 + \int g_\epsilon(x) \frac{x^2}{2p} \\
&= -h(g_\epsilon) + h(g^p) + \int (g_\epsilon(x) - g^p(x)) \frac{x^2}{2p}.
\end{aligned}$$

The lemma follows after rearrange terms. □

Lemma 4.5.3. Let $g_\epsilon(x) = g^P(x) + \epsilon \sum_{k,l} a_{k,l} H_l^{p_k}(x) g^{p_k}(x)$ be a perturbed probability density function. Then we can approximate KL divergence by

$$D(g_\epsilon || g) = \frac{\epsilon^2}{2} \int \frac{(\sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k})^2}{g^P} + o(\epsilon^2).$$

Proof. Write the KL divergence,

$$\begin{aligned} & D(g_\epsilon || g^P) \\ & = \int (g^P + \epsilon \sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k}) \log \left(1 + \epsilon \sum_{k,l} \frac{a_{k,l} H_l^{p_k} g^{p_k}}{g^P} \right). \end{aligned}$$

Expand $\log(1+x)$ in integrand,

$$\begin{aligned} & \log \left(1 + \epsilon \sum_{k,l} \frac{a_{k,l} H_l^{p_k} g^{p_k}}{g^P} \right) \\ & \quad = \epsilon \sum_{k,l} \frac{a_{k,l} H_l^{p_k} g^{p_k}}{g^P} - \frac{\epsilon^2}{2} \frac{(\sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k})^2}{(g^P)^2} + o(\epsilon^2), \end{aligned}$$

Then

$$\begin{aligned} & D(g_\epsilon || g) \\ & = \int \left(g^P + \epsilon \sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k} \right) \\ & \quad \left(\epsilon \sum_{k,l} \frac{a_{k,l} H_l^{p_k} g^{p_k}}{g^P} - \epsilon^2/2 \frac{(\sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k})^2}{(g^P)^2} + o(\epsilon^2) \right) \\ & = \frac{\epsilon^2}{2} \int \frac{(\sum_{k,l} a_{k,l} H_l^{p_k} g^{p_k})^2}{g^P} + o(\epsilon^2) \end{aligned}$$

□

Now we are ready to compute the change after certain special perturbation on Gaussian distribution for TIN. The following theorem states the result of this perturbation method.

Theorem 4.5.1. For the symmetric Gaussian interference channel with cross channel gain a and power constrain p, Gaussian signalling $X_i \sim \mathcal{N}(0,p)$ $i=1,2$ do **not** maximize TIN without power control

$$I(X_1;Y_1)+I(X_2;Y_2)$$

if there exists some $y \in [0,(1-r)^2]$ such that

$$\frac{a^4 r(2r^2+a^4 y)}{(r^2-a^4 y)^{\frac{5}{2}}}+\frac{2a^2(2-a^2 y)}{(1+a^2 y)^{\frac{5}{2}}}-\left(\frac{a^4(2+a^4 y)}{(1-a^4 y)^{\frac{5}{2}}}+\frac{(2+y)}{(1-y)^{\frac{5}{2}}}\right)>0$$

where $r=\frac{1+a^2 p}{1+p+a^2 p}$.

Proof. See Appendix 4.B. □

In theorem 4.5.1, numerical evidence suggests that setting $y=0$ is optimal. Then we have Gaussian is not optimal when $\frac{a^2}{1-a^2}>r$, or $2a^2(1+a^2p)>1$. This is the same regime as the moderate interference in [6].

4.6 Z-interference channel corner point

Consider the Gaussian Z-interference channel

$$\begin{aligned} Y_1 &= X_1 + Z_1 \\ Y_2 &= aX_1 + X_2 + Z_2, \end{aligned}$$

where $0<a<1$, $Z_i \sim \mathcal{N}(0,1)$ and the power constraints

$$\mathrm{E}[X_1^2] \le P_1, \quad \mathbb{E}[X_2^2] \le P_2.$$

Since Y_1 are independent of U_2 and $U_2=\emptyset$ maximize all terms involving U_2, after having redundant conditions removed, the Han–Kobayashi region reduces to

$$\begin{aligned} R_1 &\le I(X_1;Y_1|Q) \\ R_2 &\le I(X_2;Y_2|U_1,Q) \\ R_1+R_2 &\le I(X_1;Y_1|U_1,Q)+I(X_2,U_1;Y_2|Q) \end{aligned}$$

for some $p(q)p(u_1, x_1|q)p(u_2, x_2|q)$.

It is clear that the maximal achievable rate for each communication pair (X_i, Y_i) is $\frac{1}{2}\log(1+P_i)$, $i = 1, 2$. It can be easily shown that $\left(\frac{1}{2}\log(1+P_1), \frac{1}{2}\log\left(1+\frac{P_2}{1+a^2P_1}\right)\right)$ is a corner point of the GZIC capacity region and it attains the sum-capacity. The other corner point of the capacity region is

$$\left(\frac{1}{2}\log\left(1+\frac{a^2P_1}{1+P_2}\right), \frac{1}{2}\log(1+P_2)\right) \tag{4.4}$$

which is established in [6] and recently completed in [15]. [5] provides the slope at this corner point of Han–Kobayashi region with Gaussian signalling and power control. More precisely, [5] shows for all (R_1, R_2) in Han–Kobayashi region with Gaussian signalling and power control, $\max R_1 + \lambda R_2$ passes through the corner point (4.4) (i.e. this corner point is the maximizer of $\max R_1 + \lambda R_2$) whenever

$$\lambda \geq \lambda_{cr} := \max\left\{\frac{-\log a^2 - \frac{1-a^2}{(1+a^2P_1+P_2)}}{\log(1+P_2) - \frac{P_2}{1+P_2}}, \frac{(1-a^2)(1+P_2)}{a^2P_2}\right\} + 1. \tag{4.5}$$

Regarding to the optimality of Han–Kobayashi region with Gaussian signalling and power control, we come up with the following hypothesis.

Hypothesis 1. For some choice of P_1, P_2 there exists independent random vectors $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^n$ for some n, satisfying power constraints $\mathrm{E}(\|\mathbf{X}_1\|^2) \leq nP_1, \mathrm{E}(\|\mathbf{X}_2\|^2) \leq nP_2$, such that for some $\lambda \geq \lambda_{cr}$ (given by (4.5))

$$\begin{aligned} n\frac{\lambda-1}{2}\log(1+P_2) &< (\lambda-1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) \\ &\quad - \lambda h(a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}). \end{aligned} \tag{4.6}$$

Verification of the Hypothesis 1 can determine the optimality of Han–Kobayashi region with Gaussian signalling and power control.

Lemma 4.6.1. If Hypothesis 1 holds then (single-letter) Han–Kobayashi with Gaussian signaling and power control is not optimal.

Proof. Suppose there exists some $a, P_1, P_2, n, \lambda \geq \lambda_{cr}$, and independent random vectors $\mathbf{X}_1, \mathbf{X}_2$ satisfying power constraints such that

$$n\frac{\lambda-1}{2}\log(1+P_2) + n\delta = (\lambda-1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) \\ - \lambda h(a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}),$$

for some $\delta > 0$.

Let $\hat{P}_1 = P_1 + Q_1$ be the true power constraint on the transmitters. Take the transmitted sequence to be $\hat{\mathbf{X}}_1 = \mathbf{X}_1 + \mathbf{U}_1$ where $\mathbf{U} \sim \mathcal{N}(0, Q_1 I)$ independent of $\mathbf{X}_1$. Notice that the λ_{cr} for the parameters$(a, \hat{P}_1, P_2)$ is smaller than that of (a, P_1, P_2); therefore the inequality $\lambda \geq \lambda_{cr}$ continues to hold for the new parameter set.

By using multi-letter Han–Kobayashi scheme one can achieve the weighted sum rate

$$\begin{aligned} & n(R_1 + \lambda R_2) \\ &= I(\hat{\mathbf{X}}_1, \mathbf{X}_2; \mathbf{Y}_2) + (\lambda-1)I(\mathbf{X}_2; \mathbf{Y}_2|\mathbf{U}_1) - I(\hat{\mathbf{X}}_1; \mathbf{Y}_2|\mathbf{U}_1, \mathbf{X}_2) \\ &\quad + I(\hat{\mathbf{X}}_1; \mathbf{Y}_1|\mathbf{U}_1) \\ &= h(\mathbf{X}_2 + a\mathbf{U}_1 + a\mathbf{X}_1 + \mathbf{Z}) - h(\mathbf{Z}) + (\lambda-1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) \\ &\quad - \lambda h(a\mathbf{X}_1 + Z) + h(\mathbf{X}_1 + \mathbf{Z}) \\ &= h(\mathbf{X}_2 + a\mathbf{U}_1 + a\mathbf{X}_1 + \mathbf{Z}) - h(\mathbf{Z}) + n\frac{\lambda-1}{2}\log(1+P_2) + n\delta. \end{aligned}$$

Since $\lambda \geq \lambda_{cr}$, the corner point (4.4) attains maximal weighted sum-rate of single-letter Han–Kobayashi with Gaussian signalling and power control

$$\frac{\lambda-1}{2}\log(1+P_2) + \frac{1}{2}\log(1 + a^2(Q_1+P_1) + P_2).$$

Therefore to show the sub-optimality of the above expression it suffices to show that

$$\frac{n}{2}\log 2\pi e(1 + a^2(Q_1+P_1) + P_2) - h(\mathbf{X}_2 + a\mathbf{U}_1 + a\mathbf{X}_1 + \mathbf{Z}) \to 0$$

as $Q_1 \to \infty$.

Clearly since

$$\begin{aligned} h(\mathbf{X}_2 + a\mathbf{U}_1 + a\mathbf{X}_1 + \mathbf{Z}) &\geq h(a\mathbf{U}_1 + \mathbf{Z}) \\ &= \frac{n}{2}\log 2\pi e(1 + a^2 Q_1), \end{aligned}$$

we are done. □

Lemma 4.6.2. If Hypothesis 1 is not true then the title=When $a = b = 0.07$,(single-letter) Han–Kobayashi with Gaussian signaling (and power control) is optimal.

Proof. Clearly by Fano's inequality we obtain that any achievable rates R_1, R_2 must satisfy

$$\begin{aligned} & R_1 + \lambda R_2 \\ & \leq \lim_n \frac{1}{n} \sup_{\mathbf{X}_1, \mathbf{X}_2} I(\mathbf{X}_1; \mathbf{Y}_1) + \lambda I(\mathbf{X}_2; \mathbf{Y}_2) \\ & \leq \lim_n \frac{1}{n} \big(\sup h(\mathbf{Y}_2) - h(\mathbf{Y}_1 | \mathbf{X}_1) \\ & \qquad + \sup \big((\lambda - 1) h(\mathbf{Y}_2) - \lambda h(\mathbf{Y}_2 | \mathbf{X}_2) + h(\mathbf{Y}_1) \big) \big) \\ & \overset{(a)}{\leq} \frac{1}{2}\log(1 + P_2 + a^2 P_1) + \frac{\lambda - 1}{2}\log(1 + P_2), \end{aligned}$$

and the last expression matches the sum-rate of the (single-letter) Han–Kobayashi with Gaussian signaling and power control. Inequality (a) follows since the hypothesis is false. □

Remark 4.6.1. The inequality in Hypothesis 1 does not hold if either $\mathbf{X}_2 \sim \mathcal{N}(0, P_2 I)$ or if $\mathbf{X}_1 \sim \mathcal{N}(0, P_1 I)$. It is immediate that when $\mathbf{X}_1 \sim \mathcal{N}(0, P_1 I)$, the maximizing choice of $\mathbf{X}_2$ is $\mathbf{X}_2 \sim \mathcal{N}(0, P_2 I)$; and then one can verify that the inequality does not hold when $\mathbf{X}_2 \sim \mathcal{N}(0, P_2 I)$.

From the concavity of $h(\sqrt{t}\mathbf{X}_1 + \mathbf{Z})$ in t [16],

$$(\lambda - 1) h\Big(\sqrt{\frac{a^2}{1 + P_2}}\mathbf{X}_1 + \mathbf{Z}\Big) + h(\mathbf{X}_1 + \mathbf{Z}) \leq \lambda h\Big(\sqrt{\frac{\lambda - 1}{\lambda}\frac{a^2}{1 + P_2} + \frac{1}{\lambda}}\mathbf{X}_1 + \mathbf{Z}\Big).$$

Since $\lambda \geq \frac{1 - a^2 + P_2}{a^2 P_2}$,

$$\frac{\lambda - 1}{\lambda}\frac{a^2}{1 + P_2} + \frac{1}{\lambda} \leq a^2.$$

As $h(\sqrt{t}\mathbf{X}_1 + \mathbf{Z})$ is increasing in t, it follows

$$\begin{aligned}
&(\lambda-1)h(a\mathbf{X}_1 + \sqrt{1+P_2}\mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}) \\
&\leq \frac{n}{2}\log(1+P_2) + \lambda h(\sqrt{\frac{\lambda-1}{\lambda}\frac{a^2}{1+P_2} + \frac{1}{\lambda}}\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z}) \\
&\leq \frac{n}{2}\log(1+P_2).
\end{aligned}$$

Remark 4.6.2. An attempt to prove the converse of Hypothesis 1 is a path argument. For any $(\mathbf{X}_1, \mathbf{X}_2)$ with second moment (P_1, P_2), consider new input random variables $(\mathbf{X}_{1t}, \mathbf{X}_{2t}) = (\sqrt{1-t}\mathbf{X}_1, \sqrt{1-t}\mathbf{X}_2 + \sqrt{tP_2}\mathbf{Z})$. The converse of Hypothesis 1 follows if right hand side of (4.6) is increasing in t. However this is not true for certain $(\mathbf{X}_1, \mathbf{X}_2)$.

4.7 Discussion on the weighted sum-rate

Maximizers of weighted sum-rate can be used to characterize the boundary of Han–Kobayashi region. We are interested in the question that whether Gaussian signalling maximizes the weighted sum-rate for any $\lambda > 1$. (When $\lambda \leq 1$, the maximizing boundary point is the corner point $\left(\frac{1}{2}\log(1+P_1), \frac{1}{2}\log\left(1 + \frac{P_2}{1+a^2P_1}\right)\right)$, which is obtained with Gaussian signalling.)

Consider the weighted sum-rate $R_1 + \lambda R_2,\ \lambda > 1$. Using Fourier-Motzkin elimination as did in finding the Han-Kobayashi sum-rate, we have

$$\begin{aligned}
R_1 + \lambda R_2 &\leq I(X_1;Y_1|Q) + \lambda I(X_2;Y_2|U_1,Q) \\
R_1 + \lambda R_2 &\leq I(X_1;Y_1|U_1,Q) + I(X_2,U_1;Y_2|Q) + (\lambda-1)I(X_2;Y_2|U_1,Q).
\end{aligned}$$

As $I(U_1;Y_1) \geq I(U_1,Y_2)$, the first inequality is redundant. Thus the weighted sum-rate is

$$\begin{aligned}
&R_1 + \lambda R_2 \\
&\leq I(X_1;Y_1|U_1,Q) + I(X_2,U_1;Y_2|Q) + (\lambda-1)I(X_2;Y_2|U_1,Q) \\
&= h(Y_2|Q) - h(Z_1|Q) + h(X_1+Z_1|U_1,Q) - \lambda h(aX_1+Z_2|U_1,Q)
\end{aligned}$$

$$
\begin{aligned}
&+ (\lambda - 1)h(Y_2|U_1, Q) \\
&= \mathcal{C}\left[h(Y_2) - h(Z_1) + \mathcal{C}_{X_1}\left[(\lambda - 1)h(Y_2) - \lambda h(aX_1 + Z_2) + h(X_1 + Z_1)\right]\right].
\end{aligned}
\tag{4.7}
$$

The last equality makes use of nested concave envelops notation to remove auxiliaries. Keep in mind the inner concave envelop is taken over X_1.

Use Gaussian signalling as input, (4.7) becomes

$$
\begin{aligned}
\mathcal{C}_{P_1,P_2}\Bigg[&\frac{1}{2}\log(1 + a^2P_1 + P_2) \\
&+ \max_{x \le P_1}\left[\frac{\lambda - 1}{2}\log\frac{1 + a^2x + P_2}{1 + a^2x} + \frac{1}{2}\log\frac{1 + x}{1 + a^2x}\right]\Bigg].
\end{aligned}
\tag{4.8}
$$

By differentiating the function in inner bracket of (4.8) with respect to x, we can find that the behaviour of the function depends on λ. When $\lambda \ge \frac{1-a^2+P_2}{a^2P_2}$, it is decreasing in $(0, +\infty)$. When $\frac{1-a^2+P_2}{P_2} < \lambda < \frac{1-a^2+P_2}{a^2P_2}$, it is increasing in $\left(0, \frac{\frac{1-a^2+P_2}{a^2P_2} - \lambda}{\lambda - \frac{1-a^2+P_2}{P_2}}\right)$ and then decrease. When $\lambda \le \frac{1-a^2+P_2}{P_2}$, it is increasing in $(0, +\infty)$. We therefore can evaluate the maximum and compare it with non Gaussian signalling.

An observation is that the function inside inner concave envelope in (4.7) is not maximized by Gaussian signalling. A counterexample is $\lambda = 3.1641$, $a = 0.6759$, $P_1 = 4.6547$, $P_2 = 0.3417$. The inputs are mixed Gaussian distributions $X_1 \sim 0.5 * \mathcal{N}(1.0374, 2.2836) + 0.5 * \mathcal{N}(-1.0374, 4.8735)$, $X_2 \sim 0.5 * \mathcal{N}(0.3505, 0.4376) - 0.1752$.

Figure 4.1 plots the weighted sum-rate (4.8) without concave envelope for $\lambda = 3.1641$, $a = 0.6759$. When $P_1 \ge 4.6547$, $P_2 = 0.3417$, it can be observed that the function is not concave in (P_1, P_2). Thus power control is needed. Although the counterexample outperforms Gaussian signalling without power control, simulation shows that it is still less than Gaussian signalling with power control. This indicates that to prove Gaussian signalling is optimal, one should not try to maximize the leading term $h(Y_2) - h(Z_1)$ (this term is maximized by Gaussian) and the rest concave envelop function separately.

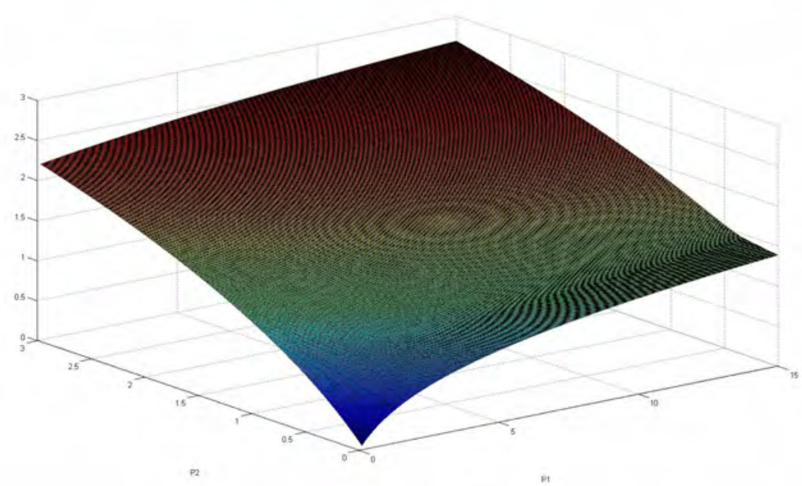

Figure 4.1: Gaussian signalling without power control

Appendix

4.A Proof of Lemma 4.3.1

Proof of Lemma 4.3.1. First, consider the expression inside concave envelop.

$$
\begin{aligned}
&I(X_1;T_1|T_2S_1X_2) - I(X_1;Y_2|T_2S_1X_2)\\
=&I(X_1;T_1|S_1) - I(X_1;Y_2|T_2S_1X_2)\\
=&I(X_1;T_1S_1) - I(X_1;Y_2S_1|T_2X_2) \quad \text{(since } I(X_1;S_1) = I(X_1;S_1|T_2X_2))\\
=&I(X_1;X_1+\eta_1\dot{Z}_1, X_1+\mu_1\ddot{Z}_2)\\
&\quad - I(X_1;X_1+\mu_1\ddot{Z}_2, X_2+aX_1+\rho_{21}\dot{Z}_2+\rho_{22}\ddot{Z}_2+\rho_{23}\tilde{Z}_2|\dot{Z}_2,X_2)\\
=&I(X_1;X_1+\eta_1\dot{Z}_1, X_1+\mu_1\ddot{Z}_2) - I(X_1;X_1+\mu_1\ddot{Z}_2, aX_1+\rho_{22}\ddot{Z}_2+\rho_{23}\tilde{Z}_2)\\
=&I(X_1;X_1+\eta_1\dot{Z}_1, X_1+\mu_1\ddot{Z}_2) - I(X_1;X_1+\mu_1\ddot{Z}_2, (\rho_{22}-a\mu_1)\ddot{Z}_2+\rho_{23}\tilde{Z}_2)\\
=&I(X_1;X_1+\eta_1\dot{Z}_1, X_1+\mu_1\ddot{Z}_2) - I(X_1;X_1+\mu_1\ddot{Z}_2|(\rho_{22}-a\mu_1)\ddot{Z}_2+\rho_{23}\tilde{Z}_2).
\end{aligned}
$$

As $X_1 \to T_1 \to S_1$ or $X_1 \to S_1 \to T_1$, the first term, when using Gaussian input, is

$$I(X_1;X_1+\eta_1\dot{Z}_1, X_1+\mu_1\ddot{Z}_2) = \frac{1}{2}\log\left(1+\frac{P_1}{\eta_1^2\wedge\mu_1^2}\right).$$

Now to simplify the second term, consider α such that

$$\mathrm{Cov}[X_1+\mu_1\ddot{Z}_2-\alpha((\rho_{22}-a\mu_1)\ddot{Z}_2+\rho_{23}\tilde{Z}_2), (\rho_{22}-a\mu_1)\ddot{Z}_2+\rho_{23}\tilde{Z}_2] = 0 \tag{4.9}$$

or

$$\alpha = \frac{\mu_1(\rho_{22}-a\mu_1)}{(\rho_{22}-a\mu_1)^2+\rho_{23}^2}.$$

Uncorrelated Gaussian random variables are independent. Thus

$$\begin{aligned}
&I(X_1; X_1 + \mu_1 \ddot{Z}_2 | (\rho_{22} - a\mu_1)\ddot{Z}_2 + \rho_{23}\tilde{Z}_2) \\
=&I(X_1; X_1 + \mu_1 \ddot{Z}_2 - \alpha((\rho_{22} - a\mu_1)\ddot{Z}_2 + \rho_{23}\tilde{Z}_2) | (\rho_{22} - a\mu_1)\ddot{Z}_2 + \rho_{23}\tilde{Z}_2) \\
=&I\left(X_1; X_1 + \frac{\mu_1 \rho_{23}^2}{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}\ddot{Z}_2 - \frac{\mu_1\rho_{23}(\rho_{22} - a\mu_1)}{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}\tilde{Z}_2\right) \\
=&\frac{1}{2}\log\left(1 + P_1 \frac{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}{\mu_1^2\rho_{23}^2}\right).
\end{aligned}$$

Note that when Gaussian signalling is used, the expression is either concave (when positive) or convex (when negative) as a result of stochastic degradation of Gaussian noise. Thus

$$\begin{aligned}
&\mathcal{C}\left[I(X_1; T_1 | T_2 S_1 X_2) - I(X_1; Y_2 | T_2 S_1 X_2)\right] \\
&\quad = \left[\frac{1}{2}\log\left(1 + \frac{P_1}{\eta_1^2 \wedge \mu_1^2}\right) - \frac{1}{2}\log\left(1 + P_1 \frac{(\rho_{22} - a\mu_1)^2 + \rho_{23}^2}{\mu_1^2\rho_{23}^2}\right)\right]_+.
\end{aligned}$$

Similar for the other concave envelop,

$$\begin{aligned}
&\mathcal{C}\left[I(X_2; T_2 | T_1 S_2 X_1) - I(X_2; Y_1 | T_1 S_2 X_1)\right] \\
&\quad = \left[\frac{1}{2}\log\left(1 + \frac{P_2}{\eta_2^2 \wedge \mu_2^2}\right) - \frac{1}{2}\log\left(1 + P_2 \frac{(\rho_{12} - b\mu_2)^2 + \rho_{13}^2}{\mu_2^2\rho_{13}^2}\right)\right]_+.
\end{aligned}$$

Now evaluate the leading terms. By the same technique in (4.9) and differential entropy of Gaussian random variables, it follows

$$\begin{aligned}
&I(X_1; Y_1 T_1 | S_2) \\
=&I(X_1; X_1 + bX_2 + \rho_{11}\dot{Z}_1 + \rho_{12}\ddot{Z}_1 + \rho_{13}\tilde{Z}_1, X_1 + \eta_1\dot{Z}_1 | X_2 + \mu_2\ddot{Z}_1) \\
=&I\bigg(X_1; X_1 + bX_2 + \rho_{11}\dot{Z}_1 + \rho_{12}\ddot{Z}_1 + \rho_{13}\tilde{Z}_1 - \frac{bP_2 + \rho_{12}\mu_2}{P_2 + \mu_2^2}\left(X_2 + \mu_2\ddot{Z}_1\right), \\
&\qquad\qquad X_1 + \eta_1\dot{Z}_1\bigg) \\
=&I\left(X_1; X_1 + \frac{b\mu_2^2 - \rho_{12}\mu_2}{P_2 + \mu_2^2}X_2 + \rho_{11}\dot{Z}_1 + \frac{P_2\rho_{12} - bP_2\mu_2}{P_2 + \mu_2^2}\ddot{Z}_1 + \rho_{13}\tilde{Z}_1, X_1 + \eta_1\dot{Z}_1\right) \\
=&\frac{1}{2}\log\frac{\det(K_1)}{\det(N_1)},
\end{aligned}$$

where

$$K_1 = \begin{pmatrix} P_1 + \frac{(b\mu_2 - \rho_{12})^2}{P_2 + \mu_2^2}P_2 + \rho_{11}^2 + \rho_{13}^2 & P_1 + \eta_1\rho_{11} \\ P_1 + \eta_1\rho_{11} & P_1 + \eta_1^2 \end{pmatrix},$$

$$N_1 = \begin{pmatrix} \frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{11}^2 + \rho_{13}^2 & \eta_1\rho_{11} \\ \eta_1\rho_{11} & \eta_1^2 \end{pmatrix}.$$

So

$$\begin{aligned} & I(X_1; Y_1T_1|S_2) \\ =& \frac{1}{2}\log \frac{\frac{(b\mu_2-\rho_{12})^2(P_1+\eta_1^2)}{P_2+\mu_2^2}P_2 + P_1 - \rho_{12}^2P_1 + P_1\eta_1^2 - 2\eta_1\rho_{11}P_1 + \eta_1^2\rho_{13}^2}{\frac{(b\mu_2-\rho_{12})^2\eta_1^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2\eta_1^2} \\ =& \frac{1}{2}\log\left(1 + P_1\left(\frac{1}{\eta_1^2} + \frac{1}{\frac{(b\mu_2-\rho_{12})^2}{P_2+\mu_2^2}P_2 + \rho_{13}^2}\left(\frac{\rho_{11}}{\eta_1} - 1\right)^2\right)\right). \end{aligned}$$

Similarly,

$$\begin{aligned} & I(X_2; Y_2T_2|S_1) \\ =& \frac{1}{2}\log\left(1 + P_2\left(\frac{1}{\eta_2^2} + \frac{1}{\frac{(a\mu_1-\rho_{22})^2}{P_1+\mu_1^2}P_1 + \rho_{23}^2}\left(\frac{\rho_{21}}{\eta_2} - 1\right)^2\right)\right). \end{aligned}$$

Hence we have the outer bound (4.1). □

4.B Proof of Theorem 4.5.1

Proof of Theorem 4.5.1. Consider the following perturbed distribution on X_1 and X_2,

$$X_1 \sim g^p(x) + \epsilon c H_1^p(x)g^p(x) + \epsilon\alpha H_2^{p_1}(x)g^{p_1}(x) + \epsilon\delta H_4^p(x)g^p(x),$$

$$X_2 \sim g^p(x) - \epsilon b H_1^p(x)g^p(x) + \epsilon\beta H_2^{p_2}(x)g^{p_2}(x) + \epsilon\delta H_4^p(x)g^p(x).$$

The above probability density function (pdf) is valid for $\delta > 0$ and ϵ small enough so that value of pdf at any x is positive. Also, orthogonal property of Hermite polynomial guarantees the integration of the pdf is 1. We may assume δ is negligible compared to α, β, b, c and thus omit δ terms.

First, use Lemma 4.5.1 to compute pdf of sums of random variables.

$$
\begin{aligned}
&aX_2 + Z \sim \\
&\quad \left(g^{a^2p} - \epsilon b H_1^{a^2p} g^{a^2p} + \epsilon\beta H_2^{a^2p_2} g^{a^2p_2} + \epsilon\delta H_4^{a^2p} g^{a^2p}\right) * g^1 \\
&= g^{a^2p+1} - \epsilon b \frac{\sqrt{a^2p}}{\sqrt{a^2p+1}} H_1^{a^2p+1} g^{a^2p+1} + \epsilon\beta \frac{a^2p_2}{a^2p_2+1} H_2^{a^2p_2+1} g^{a^2p_2+1}.
\end{aligned}
$$

$$
\begin{aligned}
&aX_1 + Z \sim \\
&\quad (g^{a^2p} + \epsilon c H_1^{a^2p} g^{a^2p} + \epsilon\alpha H_2^{a^2p_1} g^{a^2p_1} + \epsilon\delta H_4^{a^2p} g^{a^2p}) * g^1 \\
&= g^{a^2p+1} + \epsilon c \frac{\sqrt{a^2p}}{\sqrt{a^2p+1}} H_1^{a^2p+1} g^{a^2p+1} + \epsilon\alpha \frac{a^2p_1}{a^2p_1+1} H_2^{a^2p_1+1} g^{a^2p_1+1}.
\end{aligned}
$$

$$
\begin{aligned}
&X_1 + aX_2 + Z \sim \\
&(g^p(x) + \epsilon c H_1^p(x) g^p(x) + \epsilon\alpha H_2^{p_1}(x) g^{p_1}(x)) \\
&* \left(g^{a^2p+1} - \epsilon b \frac{\sqrt{a^2p}}{\sqrt{a^2p+1}} H_1^{a^2p+1} g^{a^2p+1} \right. \\
&\qquad\qquad \left. + \epsilon\beta \frac{a^2p_2}{a^2p_2+1} H_2^{a^2p_2+1} g^{a^2p_2+1}\right) \\
&= g^{p+a^2p+1} + \epsilon \left[c\sqrt{\frac{p}{p+a^2p+1}} H_1^{p+a^2p+1} g^{p+a^2p+1} \right. \\
&\qquad\qquad + \alpha \frac{p_1}{p_1+a^2p+1} H_2^{p_1+a^2p+1} g^{p_1+a^2p+1} \\
&\qquad\qquad - b\sqrt{\frac{a^2p}{p+a^2p+1}} H_1^{p+a^2p+1} g^{p+a^2p+1} \\
&\qquad\qquad \left. + \beta \frac{a^2p_2}{p+a^2p_2+1} H_2^{p+a^2p_2+1} g^{p+a^2p_2+1} \right] \\
&\quad + \epsilon^2 \left[-bc \frac{\sqrt{a^2p}}{\sqrt{p+a^2p+1}} \sqrt{\frac{2p}{p+a^2p+1}} H_2^{p+a^p+1} g^{p+a^p+1} \right. \\
&\qquad\qquad + \beta c \frac{a^2p_2\sqrt{3p}}{(p+a^2p_2+1)^{3/2}} H_3^{p+a^2p_2+1} g^{p+a^2p_2+1} \\
&\qquad\qquad - \alpha b \frac{\sqrt{a^2p}\sqrt{3}p_1}{(1+a^2p+p_1)^{3/2}} g^{p_1+a^p+1} H_3^{p_1+a^p+1} \\
&\qquad\qquad \left. + \alpha\beta \frac{a^2p_2p_1\sqrt{6}}{(p_1+a^2p_2+1)^2} g^{p_1+a^2p_2+1} H_4^{p_1+a^2p_2+1} \right].
\end{aligned}
$$

$$
\begin{aligned}
&X_2 + aX_1 + Z \sim \\
&(g^p(x) - \epsilon b H_1^p(x) g^p(x) + \epsilon\beta H_2^{p_2}(x) g^{p_2}(x)) \\
&\quad * \left(g^{a^2p+1} + \epsilon c \frac{\sqrt{a^2p}}{\sqrt{a^2p+1}} H_1^{a^2p+1} g^{a^2p+1} \right. \\
&\qquad\qquad \left. + \epsilon\alpha \frac{a^2p_1}{a^2p_1+1} H_2^{a^2p_1+1} g^{a^2p_1+1} \right) \\
&= g^{p+a^2p+1} + \epsilon \left[-b\sqrt{\frac{p}{p+a^2p+1}} H_1^{p+a^2p+1} g^{p+a^2p+1} \right. \\
&\qquad + \beta \frac{p_2}{p_2+a^2p+1} H_2^{p_2+a^2p+1} g^{p_2+a^2p+1} \\
&\qquad + c\sqrt{\frac{a^2p}{p+a^2p+1}} H_1^{p+a^2p+1} g^{p+a^2p+1} \\
&\qquad \left. + \alpha \frac{a^2p_1}{p+a^2p_1+1} H_2^{p+a^2p_1+1} g^{p+a^2p_1+1} \right] \\
&\quad + \epsilon^2 \left[-bc \frac{\sqrt{a^2p}}{\sqrt{p+a^2p+1}} \sqrt{\frac{2p}{p+a^2p+1}} H_2^{p+a^p+1} g^{p+a^p+1} \right. \\
&\qquad - \alpha b \frac{a^2p_1\sqrt{3p}}{(p+a^2p_1+1)^{3/2}} H_3^{p+a^2p_1+1} g^{p+a^2p_1+1} \\
&\qquad + \beta c \frac{\sqrt{a^2p}\sqrt{3}p_2}{(1+a^2p+p_2)^{3/2}} g^{p_2+a^p+1} H_3^{p_2+a^p+1} \\
&\qquad \left. + \alpha\beta \frac{a^2p_2p_1\sqrt{6}}{(p_1+a^2p_2+1)^2} g^{p_1+a^2p_2+1} H_4^{p_1+a^2p_2+1} \right].
\end{aligned}
$$

Then, we compute change in differential entropy

$$
\begin{aligned}
&h(X_1 + aX_2 + Z_1) - h(g^{p+a^2p+1}) \\
&= -D(g_\epsilon || g) + \frac{1}{2} \int \frac{x}{p+a^2p+1} (g_\epsilon(x) - g(x)) \mathrm{d}x \\
&= -\frac{\epsilon^2}{2} \int \frac{1}{g^{p+a^2p+1}} \left[\sqrt{\frac{p}{p+a^2p+1}} (c-ab) g^{p+a^2p+1} H_1^{p+a^2p+1} \right. \\
&\quad \left. + \frac{\beta a^2 p_2}{p+a^2p_2+1} g^{p+a^2p_2+1} H_2^{p+a^2p+1} + \frac{\alpha p_1}{p_1+a^2p_1} g^{p_1+a^2p+1} H_2^{p_1+a^2p+1} \right]^2 \\
&\quad + \frac{\epsilon}{2} \left[\frac{\sqrt{2}\alpha p_1}{p+a^2p+1} + \frac{\sqrt{2}\beta a^2 p_2}{p+a^2p+1} \right] - \frac{\epsilon^2}{2} \left(\frac{2apbc}{p+a^2p+1} \right).
\end{aligned}
$$

$$
\begin{aligned}
& h(aX_2+Z_1)-h(g^{a^2p+1}) \\
=& -D(g_\epsilon||g)+1/2\int\frac{x}{p+a^2p+1}(g_\epsilon-g) \\
=& -\frac{\epsilon^2}{2}\int\frac{1}{g^{1+a^2p}}\left[-b\frac{\sqrt{a^2p}}{\sqrt{a^2p+1}}H_1^{a^2p+1}g^{a^2p+1}+\frac{\beta a^2p_2}{a^2p_2+1}H_2^{a^2p_2+1}g^{a^2p_2+1}\right]^2 \\
& \quad +\frac{\epsilon}{2}\frac{\sqrt{2}\beta a^2p_2}{1+a^2p}.
\end{aligned}
$$

Hence TIN total increment after perturbation is

$$
\begin{aligned}
& \Delta \\
=& -\frac{\epsilon^2}{2}\int\frac{1}{g^{p+a^2p+1}}\left(\left[\sqrt{\frac{p}{p+a^2p+1}}(c-ab)g^{p+a^2p+1}H_1^{p+a^2p+1}\right.\right. \\
& +\left.\frac{\beta a^2p_2}{p+a^2p_2+1}g^{p+a^2p_2+1}H_2^{p+a^2p_2+1}+\frac{\alpha p_1}{p_1+a^2p+1}g^{p_1+a^2p+1}H_2^{p_1+a^2p+1}\right]^2 \\
& +\left[-\sqrt{\frac{p}{p+a^2p+1}}(-b+ac)g^{p+a^2p+1}H_1^{p+a^2p+1}\right. \\
& +\left.\left.\frac{\alpha a^2p_1}{p+a^2p_1+1}g^{p+a^2p_1+1}H_2^{p+a^2p_1+1}+\frac{\beta p_2}{p_2+a^2p+1}g^{p_2+a^2p+1}H_2^{p_2+a^2p+1}\right]^2\right) \\
& +\frac{\epsilon^2}{2}\int\frac{1}{g^{1+a^2p}}\left(\left[-b\frac{\sqrt{a^2p}}{\sqrt{a^2p+1}}H_1^{a^2p+1}g^{a^2p+1}+\frac{\beta a^2p_2}{a^2p_2+1}H_2^{a^2p_2+1}g^{a^2p_2+1}\right]^2\right. \\
& +\left.\left[c\frac{\sqrt{a^2p}}{\sqrt{a^2p+1}}H_1^{a^2p+1}(x)g^{a^2p+1}(x)+\frac{\alpha a^2p_1}{a^2p_1+1}H_2^{a^2p_1+1}(x)g^{a^2p_1+1}(x)\right]^2\right) \\
& +\frac{\epsilon}{\sqrt{2}}\left[\frac{(\alpha p_1+\beta p_2)(1+a^2)}{p+a^2p+1}-\frac{a^2(\alpha p_1+\beta p_2)}{1+a^2p}\right]-\epsilon^2\left(\frac{2apbc}{p+a^2p+1}\right)
\end{aligned}
$$

Note $\int g^P(x)f_{odd}(x)=0$ for odd function $f_{odd}(x)$. Also, it is easy to verify

$$
\frac{g^{P_1}g^{P_2}}{g^P}=\frac{P}{\sqrt{P_1P+P_2P-P_1P_2}}g^{\frac{P_1P_2P}{P_1P+P_2P-P_1P_2}}.
$$

We can simplify the increment.

$$
\begin{aligned}
& \Delta \\
=& -\frac{\epsilon^2}{2}\left(\frac{\beta^2a^4p_2^2}{(p+a^2p_2+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)a^4(p-p_2)^2}{2\sqrt{(1+a^2p_2+p)(1+p+2a^2p-a^2p_2)^5}}\right.
\end{aligned}
$$

$$+\frac{\alpha^2 p_1^2}{(p_1+a^2p+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)(p-p_1)^2}{2\sqrt{(1+a^2p+p_1)(1+a^2p+2p-p_1)^5}}$$
$$+\frac{2\beta a^2p_2\alpha p_1(1+a^2p+p)[2(1+a^2p+p)^2+(p_1-p)a^2(p_2-p)]}{2\sqrt{[(1+a^2p+p)^2-a^2(p_2-p)(p_1-p)]^5}}$$
$$+\frac{\alpha^2 a^4 p_1^2}{(p+a^2p_1+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)a^4(p-p_1)^2}{2\sqrt{(1+a^2p_1+p)(1+p+2a^2p-a^2p_1)^5}}$$
$$+\frac{\beta^2 p_2^2}{(p_2+a^2p+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)(p-p_2)^2}{2\sqrt{(1+a^2p+p_2)(1+a^2p+2p-p_2)^5}}$$
$$\left.+\frac{2\beta a^2p_2\alpha p_1(1+a^2p+p)[2(1+a^2p+p)^2+(p_1-p)a^2(p_2-p)]}{2\sqrt{[(1+a^2p+p)^2-a^2(p_2-p)(p_1-p)]^5}}\right)$$
$$+\frac{\epsilon^2}{2}\left(\frac{\beta^2a^4p_2^2(1+a^2p)(2(1+a^2p)^2+a^4(p-p_2)^2)}{2\sqrt{(1+a^2p_2)^5(1+2a^2p-a^2p_2)^5}}\right.$$
$$\left.+\frac{\alpha^2a^4p_1^2(1+a^2p)(2(1+a^2p)^2+a^4(p-p_1)^2)}{2\sqrt{(1+a^2p_1)^5(1+2a^2p-a^2p_1)^5}}\right)-\epsilon^2/2\frac{p(b^2+c^2)}{(a^2p+1)(p+a^2p+1)}.$$

Note X_1, X_2 need to satisfy power constrain.

$$\int x^2(g^p(x)+\epsilon H_1^p(x)g^p(x)+\epsilon\alpha H_2^{p_1}(x)g^{p_1}(x)+\epsilon\delta H_4^p(x)g^p(x))$$
$$+\int x^2(g^p(x)-\epsilon H_1^p(x)g^p(x)+\epsilon\beta H_2^{p_2}(x)g^{p_2}(x)+\epsilon\delta H_4^p(x)g^p(x))$$
$$=2p+\epsilon\sqrt{2}(\alpha p_1+\beta p_2)\le 2p$$

So $\alpha p_1+\beta p_2\le 0$. To make the increment positive, we need $\alpha p_1+\beta p_2=0$. Then increment is

$$\Delta$$
$$=-\frac{\epsilon^2\alpha^2p_1^2}{2}\left(\frac{a^4}{(p+a^2p_2+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)a^4(p-p_2)^2}{2\sqrt{(1+a^2p_2+p)(1+p+2a^2p-a^2p_2)^5}}\right.$$
$$+\frac{1}{(p_1+a^2p+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)(p-p_1)^2}{2\sqrt{(1+a^2p+p_1)(1+a^2p+2p-p_1)^5}}$$
$$-\frac{2a^2(1+a^2p+p)[2(1+a^2p+p)^2+(p_1-p)a^2(p_2-p)]}{\sqrt{[(1+a^2p+p)^2-a^2(p_2-p)(p_1-p)]^5}}$$
$$+(\frac{a^4}{(p+a^2p_1+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)a^4(p-p_1)^2}{2\sqrt{(1+a^2p_1+p)(1+p+2a^2p-a^2p_1)^5}}$$

$$+\frac{1}{(p_2+a^2p+1)^2}\frac{2(1+a^2p+p)^3+(1+a^2p+p)(p-p_2)^2}{2\sqrt{(1+a^2p+p_2)(1+a^2p+2p-p_2)^5}})\Bigg)$$
$$+\frac{\epsilon^2\alpha^2p_1^2}{2}\Bigg(\frac{a^4(1+a^2p)(2(1+a^2p)^2+a^4(p-p_2)^2)}{2\sqrt{(1+a^2p_2)^5(1+2a^2p-a^2p_2)^5}}$$
$$+\frac{a^4(1+a^2p)(2(1+a^2p)^2+a^4(p-p_1)^2)}{2\sqrt{(1+a^2p_1)^5(1+2a^2p-a^2p_1)^5}}\Bigg)-\frac{\epsilon^2}{2}\frac{p(b^2+c^2)}{(a^2p+1)(p+a^2p+1)}.$$

Since α is arbitrary, $\Delta>0$ is equivalent to

$$\begin{aligned}0<&\frac{a^4(1+a^2p)(2(1+a^2p)^2+a^4(p-p_2)^2)}{2\sqrt{(1+a^2p_2)^5(1+2a^2p-a^2p_2)^5}}\\&+\frac{a^4(1+a^2p)(2(1+a^2p)^2+a^4(p-p_1)^2)}{2\sqrt{(1+a^2p_1)^5(1+2a^2p-a^2p_1)^5}}\\&+\frac{2a^2(1+a^2p+p)[2(1+a^2p+p)^2+(p_1-p)a^2(p_2-p)]}{\sqrt{[(1+a^2p+p)^2-a^2(p_2-p)(p_1-p)]^5}}\\&-\Bigg(\frac{a^4(1+a^2p+p)(2(1+a^2p+p)^2+a^4(p-p_2)^2)}{2\sqrt{(1+a^2p_2+p)^5(1+p+2a^2p-a^2p_2)^5}}\\&+\frac{(1+a^2p+p)(2(1+a^2p+p)^2+(p-p_1)^2)}{2\sqrt{(1+a^2p+p_1)^5(1+a^2p+2p-p_1)^5}}\\&+\frac{a^4(1+a^2p+p)(2(1+a^2p+p)^2+a^4(p-p_1)^2)}{2\sqrt{(1+a^2p_1+p)^5(1+p+2a^2p-a^2p_1)^5}}\\&+\frac{(1+a^2p+p)(2(1+a^2p+p)^3+(p-p_2)^2)}{2\sqrt{(1+a^2p+p_2)^5(1+a^2p+2p-p_2)^5}}\Bigg).\end{aligned}$$

In particular, if $p_1+p_2=2p$ and denote $x=p_1-p\in[-p,p]$, then the condition is

$$\begin{aligned}0<\max_{x\in[-p,p]}&\frac{a^4(1+a^2p)(2(1+a^2p)^2+a^4x^2)}{((1+a^2p)^2-a^4x^2)^{\frac{5}{2}}}+\frac{2a^2(1+a^2p+p)(2(1+a^2p+p)^2-a^2x^2)}{((1+a^2p+p)^2+a^2x^2)^{\frac{5}{2}}}\\&-\Bigg(\frac{a^4(1+a^2p+p)(2(1+a^2p+p)^2+a^4x^2)}{((1+a^2p+p)^2-a^4x^2)^{\frac{5}{2}}}+\frac{(1+a^2p+p)(2(1+a^2p+p)^2+x^2)}{((1+a^2p+p)^2-x^2)^{\frac{5}{2}}}\Bigg).\end{aligned}$$

Denote $y=\frac{x}{1+p+a^2p}$, $r=\frac{1+a^2p}{1+p+a^2p}\in[\frac{a^2}{1+a^2},1]$, then the condition is equivalent to

$$0<\max_{y\in[0,(1-r)^2]}\frac{a^4r(2r^2+a^4y)}{(r^2-a^4y)^{\frac{5}{2}}}+\frac{2a^2(2-a^2y)}{(1+a^2y)^{\frac{5}{2}}}-\Bigg(\frac{a^4(2+a^4y)}{(1-a^4y)^{\frac{5}{2}}}+\frac{(2+y)}{(1-y)^{\frac{5}{2}}}\Bigg)$$

□

Chapter 5

Conclusion

Characterizing capacity of interference channel has been a fundamental open problem in information theory. This thesis was set out to provide outer bounds on capacity using genie-based techniques and has examined the tightness of these outer bound in comparison with Han–Kobayashi inner bound. One of the major difficulties in analysis of capacity region of interference channel is the computation of Han–Kobayashi inner bound which involves optimizing over space of all probability distributions with certain Markov structure. To overcome this challenge, the thesis focuses on sum-capacity in discrete settings where interference is weak and in Gaussian settings.

Two genie-based outer bounds are developed in Chapter 2. The first outer bound is obtained by providing to each decoder additional information about its intended sender, then single-letterizing the n-letter expression, identifying auxiliary random variables, and at last using concave envelop to suppress these auxiliaries. The second outer bound is an enhanced version of the first one in the sense that information about interference is also provided to each decoder.

In chapter 3, a class of interference channels, called very weak interference channels, are defined and studied. Han–Kobayashi sum-rate for a very weak interference channel reduces to treating-interference-as-noise sum rate. Discrete and continuous examples of this class of channels are also provided. In discrete case, it is shown that for this particular example, the genie-based outer bound matches treating-interference-as-noise inner bound in a sub-regime of very weak interference regime.

Chapter 4 discusses Gaussian interference channels. The enhanced genie-based outer bound is applied to Gaussian interference channels and it turns out the outer bound is tight for sum rate in all regimes where the sum-capacity has been established, including regimes where treating-interference-as-noise is optimal. Then the optimality of Gaussian signalling for both treating-interference-as-noise sum rate of the symmetric Gaussian interference channels and Han–Kobayashi weighted sum-rate of Gaussian Z interference channels are also discussed. For the symmetric Gaussian interference channels, we use perturbation method by Hermite polynomials to discover a condition where Gaussian signalling is sub-optimal. For Gaussian Z interference channels, we propose a hypothesis about certain information inequality and this hypothesis is equivalent to the optimality of Han–Kobayashi with Gaussian signalling around the corner point of capacity region.

There is still need for a lot of effort in order to completely understand capacity regions of interference channels. The analysis of tightness of the genie-based outer bounds in general settings are still challenging due to the large search space of genies and behaviours of concave envelops. In Gaussian settings, optimality of Gaussian signalling is still an interesting topic to study which may reveal more properties about differential entropy of Gaussian random variables.

Bibliography

[1] E. Abbe and L. Zheng. Coding along hermite polynomials for gaussian noise channels. In *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pages 1644–1648, June 2009.

[2] V. Annapureddy and V. Veeravalli. Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region. *Information Theory, IEEE Transactions on*, 55(7):3032 –3050, july 2009.

[3] A. Carleial. A case where interference does not reduce capacity (corresp.). *IEEE Transactions on Information Theory*, 21(5):569–570, Sep 1975.

[4] H.-F. Chong, M. Motani, H. K. Garg, and H. E. Gamal. On the han-kobayashi region for the interference channel. *IEEE Transactions on Information Theory*, 54(7):3188–3194, 2008.

[5] M. COSTA and C. NAIR. Gaussian z-interference channel: Around the corner.

[6] M. H. Costa. On the gaussian interference channel. *Information Theory, IEEE Transactions on*, 31(5):607–615, 1985.

[7] M. H. Costa and A. El Gamal. The capacity region of the discrete memoryless interference channel with strong interference. *IEEE Transactions on Information Theory*, 33(5):710–711, 1987.

[8] A. El Gamal and M. Costa. The capacity region of a class of deterministic interference channels (corresp.). *Information Theory, IEEE Transactions on*, 28(2):343–346, 1982.

[9] A. El Gamal and Y.-H. Kim. *Network Information Theory*. Cambridge University Press, 2012.

[10] R. Etkin, D. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. *Information Theory, IEEE Transactions on*, 54(12):5534 –5562, dec. 2008.

[11] T. Han and K. Kobayashi. A new achievable rate region for the interference channel. *Information Theory, IEEE Transactions on*, 27(1):49 – 60, jan 1981.

[12] S. Liu, C. Nair, and L. Xia. Very weak interference channels, 2014. available at http://chandra.ie.cuhk.edu.hk/pub/papers/manuscripts/VWIC.pdf.

[13] A. Motahari and A. Khandani. Capacity bounds for the gaussian interference channel. *Information Theory, IEEE Transactions on*, 55(2):620 –643, feb. 2009.

[14] C. Nair, L. Xia, and M. Yazdanpanah. Sub-optimality of han-kobayashi achievable region for interference channels. In *Information Theory (ISIT), 2015 IEEE International Symposium on*, pages 2416–2420, June 2015.

[15] Y. Polyanskiy and Y. Wu. Wasserstein continuity of entropy and outer bounds for interference channels. *CoRR*, abs/1504.04419, 2015.

[16] O. Rioul. Information theoretic proofs of entropy power inequalities. *Information Theory, IEEE Transactions on*, 57(1):33–55, 2011.

[17] H. Sato. On the capacity region of a discrete two-user channel for strong interference (corresp.). *Information Theory, IEEE Transactions on*, 24(3):377 – 379, may 1978.

[18] X. Shang, G. Kramer, and B. Chen. A new outer bound and the noisy-interference sum-rate capacity for gaussian interference channels. *Information Theory, IEEE Transactions on*, 55(2):689 –699, feb. 2009.

[19] E. Telatar and D. Tse. Bounds on the capacity region of a class of interference channels. In *2007 IEEE International Symposium on Information Theory*, pages 2871–2874, June 2007.