Gaussian Interference Channels: Examining the Achievable Rate Region

by

Ali Haghi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Interference is assumed to be one of the main barriers to improving the throughput of communication systems. Consequently, interference management plays an integral role in wireless communications. Although the importance of interference has promoted numerous studies on the interference channel, the capacity region of this channel is still unknown.

The focus of this thesis is on Gaussian interference channels. The two-user Gaussian Interference Channel (GIC) represents the standard model of a wireless system in which two independent transmitter-receiver pairs share the bandwidth. Three important problems are investigated: the boundary of the best-known achievable rate region, the complexity of sum-rate optimal codes, and the role of causal cooperation in enlarging the achievable rate region.

The best-known achievable rate region for the two-user GIC is due to the Han-Kobayashi (HK) scheme. The HK achievable rate region includes the rate regions achieved by all other known schemes. However, mathematical expressions that characterize the HK rate region are complicated and involve a time sharing variable and two arbitrary power splitting variables. Accordingly, the boundary points of the HK rate region, and in particular the maximum HK sum-rate, are not known in general. The second chapter of this thesis studies the sum-rate of the HK scheme with Gaussian inputs, when time sharing is not used. Note that the optimal input distribution is unknown. However, for all cases where the sum-capacity is known, it is achieved by Gaussian inputs. In this thesis, we examine the HK scheme with Gaussian inputs. For the weak interference class, this study fully characterizes the maximum achievable sum-rate and shows that the weak interference class is partitioned into five parts. For each part, the optimal power splitting and the corresponding maximum achievable sum-rate are expressed in closed forms. In the third chapter, we show that the same approach can be adopted to characterize an arbitrary weighted sum-rate. Moreover, when time sharing is used, we expressed the entire boundary in terms of the upper concave envelope of a function. Consequently, the entire boundary of the HK rate region with Gaussian inputs is fully characterized.

The decoding complexity of a given coding scheme is of paramount importance in wireless communications. Most coding schemes proposed for the interference channel

take advantage of joint decoding to achieve a larger rate region. However, decoding complexity escalates considerably when joint decoding is used. The fourth chapter studies the achievable sum-rate of the two-user GIC when joint decoding is replaced by successive decoding. This achievable sum-rate is known when interference is mixed. However, when interference is strong or weak, it is not well understood. First, this study proves that when interference is strong and transmitters' powers satisfy certain conditions, the sum-capacity can be achieved by successive decoding. Second, when interference is weak, a novel rate-splitting scheme is proposed that does not use joint decoding. It is proved that the difference between the sum-rate of this scheme and that of the HK scheme is bounded. This study sheds light on the structure of sum-rate optimal codes.

Causal cooperation among nodes in a communication system is a promising approach to increasing overall system performance. To guarantee causality, delay is inevitable in cooperative communication systems. Traditionally, delay granularity has been limited to one symbol; however, channel delay is in fact governed by channel memory and can be shorter. For example, the delay requirement in Orthogonal Frequency-Division Multiplexing (OFDM), captured in the cyclic prefix, is typically much shorter than the OFDM symbol itself. This perspective is used in the fifth chapter to study the two-user GIC with full-duplex transmitters. Among other results, it is shown that under a mild condition, the maximum multiplexing gain of this channel is in fact two.

Acknowledgments

I am honored to be supervised by Professor Amir Keyvan Khandani. He is an exceptionally talented professor who can think beyond mathematical expressions. Amir is also the kindest and the most supportive professor I have ever met. From the first moment that I came to Canada, I have felt that there is someone who cares about me and is always available to help me. I hope that I can continue to learn from him.

I am grateful to my committee members, Professors Vincent Poor, Mary Thompson, Ravi Mazumdar, and Patric Mitran for taking the time to carefully read my thesis and to provide me with valuable feedback. I am also grateful to professors Penny Haxell, Levent Tuncel, and Henry Wolkowicz from whom I learned Combinatorics and Optimization. I learned a lot from Mary McPherson who help me to better articulate my ideas. She is exceptionally nice and always ready to help students as much as she can.

I would like to thank my friend at the University of Waterloo. In particular, the members of the Coding and Signal Transmission laboratory: Hossein Bagheri, Abolfazl Motahari, and Ali Hesammohseni. I am in debt to Mohsen Raeeszadeh who taught me how to use LaTex to write a thesis. Moreover, I like to thank my friends with whom I spent my enjoyable student lifetime: Masoud Ansari, Hamidreza Mohebi, Mehdi Jalalmaab, Emad Zilouchian, Ali Zibaeenejad, Mahdi Khodaian, Hadi Zibaeenejad, and Meysam Shahrbaf.

I had the chance of learning from some amazing teachers. I am always thankful for what they did to me. My beloved teachers: Bahman Eslahpazir, Mohammadreza Aref, Mahmoud Shahabadi, and Said Nader-Esfahani.

Finally, I am truly in debt to my wife who inspires me to work harder. I usually explain my results to her, and although she is not an engineer, she is a very good listener who provides me with feedback on parts that requires further clarification. Her reassuring words have put out the flames of frustration in my life. You are the best thing that happened to me.

To my parents,

and

to my beloved wife.

Table of Contents

List of Figures **xi**

List of Tables **xv**

List of Abbreviations **xvi**

Nomenclature **xvii**

1 Introduction **1**

1.1 Boundary of the HK Rate Region 2

1.2 Complexity of Sum-Rate Optimal Codes 3

1.3 Causal Cooperation among Transmitters 3

1.4 Outline of Thesis and Main Contributions 4

2 Maximum Han-Kobayashi Sum-Rate **7**

2.1 Introduction 7

2.2 Channel Model and Preliminaries 10

2.2.1 Classes of Interference and the Corresponding Sum-Capacity . . . 11

2.2.2 Han-Kobayashi Coding Scheme 14

2.2.3 Sum-Capacity versus Maximum HK Sum-Rate 15

2.3 Maximum HK Sum-Rate without Time Sharing 18

2.3.1 Main Results 18

2.3.2 The Optimization Problem Corresponding to the Maximum HK Sum-Rate . . . 22
2.3.3 The Proposed Optimization Technique for Maximizing the HK Sum-Rate . . . 28
2.3.4 Three Categories of Points Corresponding to Optimal Power Splitting 32
2.3.5 A Sufficient Condition for Optimal Power Splitting . . . 33
2.3.6 Maximum HK Sum-Rate over Stationary Points . . . 34
2.3.7 Maximum HK Sum-Rate over Boundary Points . . . 39
2.3.8 Maximum HK Sum-Rate over Non-Differentiable Points . . . 44
2.3.9 Solving the Optimization Problem Corresponding to the Maximum HK Sum-Rate . . . 58
2.4 Conclusion . . . 71

3 Boundary of the Han-Kobayashi Rate Region **72**

3.1 Introduction . . . 72
3.2 Preliminaries . . . 74
3.2.1 Time Sharing versus Time/Frequency Division . . . 75
3.3 Boundary of the HK Rate Region . . . 77
3.3.1 Main Results . . . 78
3.3.2 Properties of the HK Rate Region . . . 79
3.3.3 The Optimization Problem Corresponding to the Maximum Weighted HK Sum-rate . . . 84
3.3.4 Rederiving Existing Results . . . 91
3.4 Conclusion . . . 94

4 Rate Splitting and Successive Decoding for Gaussian Interference Channels **95**

4.1 Introduction . . . 95

4.2 Preliminaries 98

4.2.1 The Underlying Optimization Problem Corresponding to Maximum Sum-Rate 98

4.3 Strong Interference Class 101

4.3.1 Is Rate Splitting Required? 102

4.3.2 How Many Splits Are Required? 104

4.3.3 Maximum Sum-Rate Loss 113

4.4 Weak Interference Class 118

4.4.1 Is Rate Splitting Required? 119

4.4.2 How Many Splits Are Required? 125

4.4.3 Maximum Sum-Rate Loss 135

4.5 Conclusion 144

5 Delay in Cooperative Communications: Multiplexing Gain of Gaussian Interference Channels with Full-Duplex Transmitters 145

5.1 Introduction 145

5.2 Preliminaries 148

5.2.1 Channel Model 148

5.2.2 Causal Cooperation 154

5.3 Interference Cancellation with Full-Duplex Transmitters 156

5.3.1 The Two-User GIC with One Full-Duplex Transmitter 157

5.3.2 The Two-User GIC with Two Full-Duplex Transmitters 159

5.3.3 Optimal Power Allocation 171

5.4 Simulation Results 185

5.5 Conclusion 189

6 Conclusion and Future Research Directions **190**

6.1 Conclusion 190

6.2 Future Research Directions 192

Bibliography **194**

List of Figures

2.1 Classes of interference and the corresponding sum-capacity expressions. . 11

2.2 All sub-classes of interference for which the sum-capacity is known. . . . 13

2.3 For fixed values of P_1 and P_2, the weak interference class is partitioned into four sub-classes. These sub-classes and their corresponding maximum sum-rate expressions are demonstrated in the ab-plane. 17

2.4 For fixed values of a and b, the weak interference class is partitioned into four sub-classless. These sub-classless and the corresponding maximum sum-rate are demonstrated in the P_1P_2-plane. 18

2.5 The maximum achievable sum-rate of the HK scheme ($R^{\max}_{\text{sum-HK}}$) for the two-user GIC with weak interference. The weak interference class is partitioned into five sub-classes. For fixed (a, b), these sub-classes are demonstrated in the P_1P_2-plane, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is characterized. 20

2.6 The maximum achievable sum-rate of the HK scheme for the two-user GIC with weak interference. The weak interference class is partitioned into five sub-classes, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is characterized. 21

2.7 To find the maximum of $\min\{f_1(x), f_2(x)\}$ over $[0, 1]$, it is sufficient to check all stationary points like x_s and all boundary points like x_b and all non-differentiable points like x_{nd}. 29

2.8 The behavior of $h_1(\lambda_1, \lambda_2)$ over the boundary. 35

2.9 The behavior of $h_2(\lambda_1, \lambda_2)$ over the boundary. 37

2.10 The behavior of $h_3(\lambda_1, \lambda_2)$ over the boundary 38

2.11 Four sub-categories of the boundary points: the optimal point and the maximum sum-rate corresponding to each sub-category. 39

2.12 The achievable sum-rate of the HK scheme over the boundary of the feasible region, for the barely weak interference sub-class with $c \geq 0$. 42

2.13 The sum-rate of the HK scheme achieved by investigating only the boundary points: Quadrant I of the P_1P_2-plane, is partitioned into three regions. In each region, exactly one of the $C\left(\frac{P_1}{1+aP_2}\right) + C\left(\frac{P_2}{1+bP_1}\right), C(P_1 + aP_2), C(P_2 + bP_1)$ is the achievable sum-rate. 44

2.14 Three sub-categories of non-differentiable points in the $\lambda_1\lambda_2$-plane, when $c \geq 0$. 46

2.15 Three sub-categories of non-differentiable points in the $\lambda_1\lambda_2$-plane, when $c < 0$. 47

2.16 The non-zero power splitting II sub-class demonstrated in the P_1P_2-plane. 61

2.17 For the non-zero power splitting II sub-class, the achievable sum-rate corresponding to $\mathcal{ND}_1$ is greater than the achievable sum-rate corresponding to all other sub-categories. 62

2.18 The non-zero power splitting I sub-class, projected onto the P_1P_2-plane. For this sub-class, $C(P_1+aP_2)+g_1(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2})$, which corresponds to $\mathcal{ND}_2$, is greater than the sum-rate corresponding to all other sub-categories. 65

2.19 The barely weak interference sub-class is partitioned into four sub-classes, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is demonstrated. 67

2.20 The maximum achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing for all values of a and b. 70

3.1 The achievable rate region $\mathcal{G}_0$ and its extreme points. 83

3.2 Depending on the value of μ, $R_1 + \mu R_2$ is maximized at one of the extreme points. 85

3.3 Behavior of $R_\mu = R_1 + \mu R_2 = D_1 + \mu(D_4 - 2D_1)$ over the feasible region and the six optimal power splittings that maximize R_μ. 86

3.4 Behavior of D_3 over the feasible region. 88

3.5 Behavior of $R_\mu = R_1 + \mu R_2 = (1-\mu)D_4 + (2\mu - 1)D_3$ over the feasible region: the optimal power splittings that maximize R_μ are shown by solid black dots. 89

4.1 The sum-capacity of the strong interference class. 101

4.2 Comparison of $R_{\text{sum}}^{\text{NRS}}$ with the sum-capacity for the strong interference class. 104

4.3 Regions in the P_1P_2-plane for which SD can achieve the sum-capacity of the strong interference class. The label associated with each point shows the theorem and the value of N corresponding to the point. 110

4.4 Regions in the ab-plane for which SD can achieve the sum-capacity of the strong interference class. The label associated with each point shows the theorem and the value of N that corresponds to the point. 112

4.5 The feasible region of the optimization problem (4.51). 115

4.6 Comparison of the achievable sum-rate $R_{\text{sum-SD}}$ with the sum-capacity. . 116

4.7 Comparison of the sum-capacity and the sum-rate achieved using SD for the symmetric two-user GIC with strong interference. 118

4.8 The maximum achievable sum-rate when rate splitting is not used: Quadrant I of the P_1P_2-plane is partitioned into three regions. In each region, $R_{\text{sum}}^{\text{NRS}}$ is demonstrated. 121

4.9 The weak interference class is partitioned into five sub-classes. For each suc-class, $\Delta R_{\text{sum}} \doteq R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{NRS}}$ is demonstrated. 124

4.10 Quadrant I of the P_1P_2-plane is partitioned into rectangles. Each rectangle determines the decoding orders $(\mathbf{S}_1, \mathbf{S}_2)$ and the number of splits $(N+1)$. 125

4.11 The relation between rectangles $REC(m,n)$ and sub-classes A, B, C, D, and E. 127

4.12 The achievable sum-rate $R_{\text{sum}}^{\text{RS-SD}}$. 131

4.13 The sub-class E is partitioned by hyperplanes L_i. On the boundary of each part, $R_{\text{sum-HK}}^{\max} = R_{\text{sum}}^{\text{RS-SD}}$. Inside each part, the maximum of $R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}}$ occurs when $(P_1, P_2) = \big(P_{1,\text{W}}^{\text{opt}}(N), P_{2,\text{W}}^{\text{opt}}(N)\big)$ for $N > 1$. 137

4.14 The function $g_{\min}(P_1, P_2)$ over the sub-class E. 138

5.1 Two groups (A and B) of wireless transmitters sharing M sub-carriers of OFDMA. 149

5.2 M parallel GICs formed across M sub-carriers of OFDMA. 151

5.3 The equivalent GIC with full-duplex transmitters. 152

5.4 The interference, caused by $\boldsymbol{T}_B$, reaches $\boldsymbol{R}_A$ directly by I_{di} and indirectly by I_{in}. The filter $\boldsymbol{F}_1$ can guarantee that $I_{di} + I_{in} = 0$. 156

5.5 The feasible region of the optimization problem (5.108) and the optimal solution on the boundary. 179

5.6 The optimal power allocation of the optimization problem (5.108), when $n_1^i \leq m_1^i$ and $n_2^i \geq m_2^i$. 183

5.7 The average achievable sum-rate (per complex sub-carrier) of the symmetric two-user GIC for four different coding schemes, with $M = 512$ and $P_1 = P_2 = M \times 10^3$. 187

5.8 The power available for $\boldsymbol{T}_A$ to transmit its own message $\boldsymbol{S}_1$, when optimal power allocation is used. 188

List of Tables

2.1 Sub-classes in the weak interference class and the corresponding sum-capacity expressions and maximum sum-rate expressions. 16

2.2 The weak interference class is partitioned into five sub-classes. For each sub-class, the optimal power splitting $(\lambda_1^\star, \lambda_2^\star)$ and the corresponding optimal sum-rate $R_{\text{sum-HK}}^{\max}$ are given. 22

2.3 The achievable sum-rate corresponding to four corner points of the boundary. 43

2.4 Sub-categories, their corresponding optimal power splittings and achievable sum-rate expressions, for the barely weak interference sub-class. . . . 59

3.1 The optimal power splittings. 91

4.1 The achievable sum-rate of the strong interference class corresponding to four decoding orders. 103

4.2 The achievable sum-rate of the weak interference class corresponding to four decoding orders. 120

4.3 The weak interference class is partitioned into five sub-classes. For each sub-class, $R_{\text{sum-HK}}^{\max}$ is compared with $R_{\text{sum}}^{\text{NRS}}$. 123

List of Abbreviations

GIC	Gaussian Interference Channel
MAC	Multiple Access Channel
BC	Broadcast Channel
RS	Rate Splitting
SD	Successive Decoding
HK	Han-Kobayashi
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
TD	Time Division
TS	Time Sharing
CTS	Coded Time Sharing
FD	Frequency Division
SND	Simultaneous Non-Unique Decoding
SNR	Signal-To-Noise Ratio
TIN	Treat Interference as Noise

Nomenclature

$a \doteq b$	b is the definition of a
$\mathbb{R}$	The set of real numbers
$\mathbb{R}^n$	The n-dimensional Euclidean space
$C(x)$	$\frac{1}{2}\log(1+x)$
$[x]^+$	$\max\{x, 0\}$
$[x]_a^b$	For non-negative numbers a, b, and x such that $a \leq b$, $[x]_a^b \doteq \min\{\max\{x, a\}, b\}$
$\mathbb{1}(S)$	For a statement S, $\mathbb{1}(S) = 1$ if S is true, otherwise $\mathbb{1}(S) = 0$.
$\mathcal{C}[f](x)$	For a function $f(x)$, $\mathcal{C}[f](x)$ represents the upper concave envelope of $f(x)$.
$N(m, \sigma)$	Gaussian distribution with mean m and variance σ^2
$\mathbb{E}_Z$	The expectation with respect to the random variable Z
$\oplus$	Addition modulo 2
$\lvert Q \rvert$	Cardinality of the set Q
$[1:n]$	The set of integers from 1 to n
$S_1^{1:N}$	$\{S_1^1, S_1^2, ..., S_1^N\}$
$P(S_1)$	For a random variable S_1, $P(S_1)$ represents the power of S_1
$P(S_1^{1:N})$	$P(S_1^{1:N}) \doteq \sum_{i=1}^{N} P(S_1^i)$
$\mathrm{diag}(P_1, P_2, ..., P_M)$	An $M \times M$ matrix in which $(P_1, P_2, ..., P_M)$ is the main diagonal and all other entries are zero
$\nabla(R_1)$	The gradient of the function R_1
$\mathbf{C}[i]$	For a square matrix $\mathbf{C}$, $\mathbf{C}[i]$ is the number that represents the i^{th} element of the main diagonal
$\mathbf{S}[i]$	For a vector $\mathbf{S} = [S_1, S_2, ..., S_M]^T$, $\mathbf{S}[i] \doteq S_i$
$\Leftrightarrow$	If and only if

Chapter 1

Introduction

The accelerated improvement of wireless technology, in which numerous wireless devices employ the same frequency band, has made interference an intrinsic part of today's communication systems. The first study of a communication system that considered interference as an intrinsic element, was in Shannon's work on the two-way channel [1]. His work was followed by that of several other scholars, and nowadays, the interference channel is the accepted model of a communication system in which interference, signal, and noise interact with each other [2–7].

The importance of interference in wireless communication has promoted many studies on the interference channel. The two-user Gaussian Interference Channel (GIC) is of particular interest. This channel models a practical wireless network consisting of two independent receiver-transmitter pairs. Each transmitter tries to send its message to its corresponding receiver, but it inevitably causes interference for the unintended receiver. Both receivers suffer from Gaussian noise as well.

Although the capacity region of the Gaussian interference channel has been studied for more than 40 years, it is only known for some specific cases. For example, with strong interference, the whole capacity region is known to be achieved by decoding the inference [7–9]. On the other hand, with very weak interference, the sum-capacity is achieved by treating the interference as noise [10–12].

This thesis is intended to provide a better understanding of the capacity region of the two-user Gaussian interference channel. The main contribution is to address three important aspects of this channel: (1) the boundary of the best-known achievable rate region,

(2) the complexity of sum-rate optimal codes, and (3) the role of causal cooperation in enlarging the achievable rate region.

1.1 Boundary of the HK Rate Region

One challenging aspect of characterizing the capacity region is to find a tight inner bound corresponding to a particular coding scheme. A general coding scheme, based on the idea of rate splitting, was first proposed by Calieal [5]. This scheme was then improved by Han and Kobayashi [6], whose main contribution was joint decoding at the receivers. In fact, Carlieal used successive decoding instead of joint decoding but Han and Kobayashi proved that joint decoding at the receivers can increase the achievable rate region.

For the two-user GIC, the Han-Kobayashi (HK) scheme results in the best-known inner bound. By optimizing over a time-sharing variable and two power splitting variables, the HK scheme can include all known achievable results as its special cases. However, the optimization problem involving the underlying variables has yet to be clarified. In fact, [13] states

> “ Unfortunately, the optimization among such myriads of possibilities is not well-understood”.

This thesis aims to shed light on this issue by investigating the HK scheme and finding the optimal power-splitting policy that maximizes the weighted sum-rate. Consequently, the boundary of the HK rate region with Gaussian inputs is fully characterized. This important has been investigated for more than 30 years.

The other challenging aspect of characterizing the capacity region is to find tight outer bounds. For the two-user GIC, various outer bounds have been derived using different techniques [10, 13–17]. Unlike achievable schemes, where the HK scheme results in the best-known inner bound, no converse scheme results in the best-known outer bound. In fact, each outer bound can be tighter or looser than other outer bounds, depending on the channel parameters. The outer bound obtained in [13] is of particular interest. Using a genie that provides information about the intended message to each receiver, [13] proves that a sub-region of the HK scheme is within 1 bit of the capacity region. In this thesis, our focus is on the achievable schemes. We use the existing outer bounds to check the optimality of the achievable schemes under certain conditions.

1.2 Complexity of Sum-Rate Optimal Codes

Joint decoding is used in the HK scheme to enlarge the achievable rate region. Joint decoding is a powerful coding scheme; however, it considerably increases decoding complexity. The decoding complexity of the joint decoding of k messages of a random coding scheme is proportional to $2^{nR_{\text{sum}}}$, where n is the block length, and R_{sum} is the sum of the rates corresponding to the messages that are jointly decoded, i.e., $R_{\text{sum}} = \sum_{i=1}^{k} R_i$. However, the decoding complexity of the successive decoding of the same set of messages is proportional to $2^{nR_{\max}}$, where $R_{\max} = \max(R_1, R_2, ..., R_k)$ [18]. Therefore, practical coding schemes employ successive decoding in their decoder to decrease the complexity of decoding. Moreover, there exist numerous studies regarding the construction of high performance point-to-point codes [19–23], whereas there are fewer studies on multiuser codebooks, which are jointly decoded. Thus, this study compares the performance of successive decoding, which employs existing point-to-point codes, with that of joint decoding, which employs multiuser codebooks.

Rate Splitting (RS) and Successive Decoding (SD) can reduce decoding complexity and have been used to investigate the multiple access channel and the interference channel [24,25]. The capacity region of the two-user multiple access channel can be achieved by RS and SD [18, 26]. However, for the two-user Gaussian interference channel (GIC), RS and SD cannot achieve even the Simultaneous Non-unique Decoding (SND) rate region [27].

RS and SD have been used to investigate the maximum achievable sum-rate of the two-user GIC. For instance, when interference is mixed, it is known that the sum-capacity can be achieved with SD [10]. When interference is strong or weak, the performance of RS and SD has not been well-understood. This study characterizes the maximum achievable sum-rate when joint decoding is replaced by successive decoding, and shows that, under a set of mild conditions on transmitters' powers, RS and SD can achieve the sum-rate of the HK scheme.

1.3 Causal Cooperation among Transmitters

Cooperation among nodes in a communication system is a promising approach to increasing overall system performance. Full-duplex transmitters can not only double the rate of wireless communication systems, but also facilitate collaborative signaling and cooperative communication [28]. For the two-user interference channel, full-duplex transmitters can take advantage of the signal they receive from each other to mitigate interference at their receivers, and this simple cooperation among the transmitters can enlarge the achievable rate region. In the context of cognitive radio channels, the role of cooperation in enlarging the capacity region of the GIC has been studied, and rate-splitting along with Gelfand-Pinsker binning has been used to improve the achievable rate region [29], [30]. Moreover, the capacity region of the two-user Gaussian interference channel with conferencing encoders is established in [31] to within a constant gap. To investigate the effect of causal cooperation, the achievable rate region of two-user interference channels with cribbing encoders is studied in [32–34].

Furthermore, multiplexing gain has been used as a measure to investigate the role of partial non-causal cooperation in wireless networks in the high Signal-To-Noise Ratio (SNR) regime. It is proved that, for the K-user GIC, as the cooperation among transmitters increases from no cooperation to perfect cooperation, the multiplexing gain increases from $\frac{1}{2}K$ to K [35]. However, practical cooperation among different nodes requires the causal delay consideration as an essential constraint. The signal transmitted by a node will be received and processed by other nodes with some delay, and the minimum acceptable delay can significantly affect the potential gains of cooperative communication systems. For instance, in the two-user GIC, when only transmitters cooperate non-causally, the channel behaves like the broadcast channel, and the maximum multiplexing gain of two is achievable [36, 37]. Similarly, non-causal cooperation among the receivers achieves the multiple-access-channel multiplexing gain of two [38]. This study investigates the two-user GIC with full-duplex transmitters to show that causal cooperation among transmitters can increase the multiplexing gain.

1.4 Outline of Thesis and Main Contributions

The main objective of this thesis is to provide a better understanding of the capacity region of the two-user GIC. To this end, three important problems are investigated: the boundary of the HK rate region with Gaussian inputs, the complexity of sum-rate optimal codes, and the role of cooperation in enlarging the achievable rate region.

The first problem is addressed in Chapters 2 and 3. In Chapter 2, we investigate the maximum HK sum-rate. Note that the HK scheme results in the best-known achievable rate region. However, mathematical expressions that characterize the achievable sum-rate of the HK scheme are complicated, involving two power-splitting variables and one time-sharing variable. For simplicity, we first investigate the maximum HK sum-rate with Gaussian inputs when time sharing is not used. Then in Chapter 3, we return to time sharing and investigate its role in increasing the achievable sum-rate. Note that, when interference is strong or mixed, the maximum HK sum-rate is known. However, for the weak interference class, the maximum HK sum-rate has remained unknown. The main contribution of Chapter 2 is the characterization of the explicit power-splitting policy that maximizes the HK sum-rate when interference is weak. We first describe the optimization problem that corresponds to the maximum HK sum-rate and highlight the challenges in solving the optimization problem. In particular, the fact that the objective function is non-differentiable over the feasible region is discussed. Then we explain our idea for solving the problem. The idea is to partition the entire feasible region into several parts such that, inside each part, the objective function is differentiable. In other words, we partition the feasible region into several parts such that all non-differentiable points lie on the boundary of the parts. Relying on this idea, we solve the optimization problem and fully characterize the maximum HK sum-rate. Chapter 2 shows that, depending on the values of channel parameters, five different power-splitting policies maximize the HK sum-rate. In fact, we partition the weak interference class into five sub-classes, and for each sub-class, we fully characterize the optimal power-splitting policy and the corresponding maximum sum-rate.

In Chapter 3, we generalize the results of Chapter 2 and characterize the maximum weighted sum-rate of the HK scheme with Gaussian inputs. In other words, we fully characterize the optimal power-splitting policy that maximizes a linear combination of R_1 and R_2. Note that the time-sharing variable can increase the sum-rate of the HK scheme. We first highlight that time sharing and time division are not essentially the same. In fact, time division, which convexifies the achievable rate region, is a special case of time sharing. Then we show that the role of time sharing in increasing the achievable weighted sum-rate can be described in terms of the upper concave envelope of a function of transmitters' powers. This characterization can reveal several important properties.

For instance, it allows to identify the regions in which time sharing does not increase the achievable sum-rate. More importantly, we can identify regions in which simple time division with power control is as good as the general time sharing strategy.

In Chapter 4, we address the second problem. To achieve the sum-capacity of the two-user GIC, most coding schemes take advantage of joint decoding. However, decoding complexity increases when joint decoding is used. Rate splitting and successive decoding provide alternatives that can reduce this complexity. On the other hand, this complexity reduction is achieved at a price. Some points of the capacity region cannot be achieved by successive decoding. In this study, we investigate the achievable sum-rate, when joint decoding is replaced by successive decoding. First, we express the optimization problem that corresponds to the maximum achievable sum-rate. Chapter 4 highlights the challenges in solving the optimization problem. In particular, it is shown that the optimization problem is non-convex. Then a method is proposed for solving this optimization problem. We explicitly determine the number of required splits and the amount of power allocated to each split. We then show that the sum-rate loss, caused by replacing joint decoding with successive decoding, is bounded, even when transmitters' powers approach infinity.

Chapter 5 addresses the third problem, namely the role of cooperation in enlarging the achievable rate region. It is known that causal cooperation among transmitters of the two-user GIC does not increase the multiplexing gain [39]. This result is obtained with the traditional delay assumption. To guarantee causality, delay granularity has been assumed to be limited to one symbol; however, channel delay is in fact determined by channel memory and can be much shorter. Using this perspective, we investigate the two-user GIC with full-duplex transmitters, and reach the following conclusion: with a new constraint of causal delay, which is slightly different from the traditional one, the role of delay is captured more accurately. As a result, the maximum multiplexing gain is in fact two, rather than the limit of one, previously proved under the traditional constraint of causal delay [39]. Furthermore, we study the optimal power allocation that maximizes the achievable sum-rate and examine its effect through several numerical simulations.

Finally, Chapter 6 summarizes the main contributions of this thesis and discusses future research directions.

Chapter 2

Maximum Han-Kobayashi Sum-Rate

Chapters 2 and 3 investigate the best-known achievable rate region proposed for the two-user Gaussian interference channel, i.e., the Han-Kobayashi (HK) region. Chapter 2 characterizes the maximum sum-rate achieved by the HK scheme, when time sharing is not used. The main challenge in this characterization is to solve a non-differentiable optimization problem. A method is proposed for solving the optimization problem and is discussed in detail. Chapter 3 extends the method and characterizes the maximum weighted sum-rate achieved by the HK scheme.

2.1 Introduction

Shannon's work on the two-way channel [1] is one of the first studies of a communication system that considered interference as an essential element. In wireless communications, interference is assumed to be one of the main challenges that hinders overall system performance. The two-user Gaussian Interference Channel (GIC) models a practical wireless network consisting of two independent transmitter-receiver pairs. Each transmitter aims to send a message to its receiver, thereby inevitably causing interference for the other receiver.

Although the two-user interference channel has been studied for more than 40 years, its capacity region is known only for a few specific cases. Several coding schemes have been proposed for the two-user GIC, such as time division with power control, Treating Interference as Noise (TIN), Simultaneous Non-unique Decoding (SND), and Han-Kobayashi

(HK) [37, Chapter 6].

Rate splitting provides a general frame work to enlarge the achievable rate region of this channel. In fact, Carleial was the first to propose a scheme based on rate splitting and successive decoding [5]. This scheme was subsequently improved by HK [6], whose main contribution was the use of joint decoding at receivers. Their work proved that joint decoding at receivers can increase the achievable rate region. For the two-user GIC, the HK scheme has four main ingredients: (1) rate splitting, (2) power splitting, (3) joint decoding, and (4) time sharing. By modifying the power-splitting policies and using different time-sharing strategies, the HK scheme can include all known coding schemes as its special cases. However, the optimization among the power-splitting variables and time-sharing variables is complicated. In fact, [13] states that

> " Unfortunately, the optimization among such myriads of possibilities is not well-understood, ... it is not very clear how much improvement can be obtained and in which parameter regime would one get significant improvement".

This chapter is intended to provide a better understanding of this issue by investigating the HK scheme and finding the optimal power splitting that maximizes the sum-rate.

The sum-capacity of the interference channel is known for only a few special cases. When interference is strong, the sum-capacity is achieved by decoding both messages at both receivers [7, 8]. When interference is mixed, the sum-capacity is achieved if one transmitter sends only the private message and the other transmitter sends only the public message [10]. When interference is weak, the sum-capacity is not known in general. For a small part of the weak interference class, the sum-capacity is achieved by treating interference as noise [10–12]. In all cases where the sum-capacity is known, it is achieved by the HK scheme with Gaussian inputs and no time sharing. However, for the weak interference class, the maximum sum-rate of the HK scheme with Gaussian inputs and no time sharing is not known.

When interference is weak, the maximum achievable sum-rate of the HK scheme, even when inputs are Gaussian, is unknown. This problem has been studied in [40–42]. Reference [40] studies the two-user symmetric GIC when the HK scheme with Gaussian inputs and no time sharing is used. Among all possible power-splitting policies, reference

[40] investigates only two special cases: the symmetric power splitting and an asymmetric power splitting in which exactly one user allocates all its power to its public message. Moreover, reference [41] studies the achievable sum-rate of the two-user GIC when the HK scheme with Gaussian inputs and no time sharing is used. For some parts of the weak interference class, reference [41] finds a closed form expression for the optimal power splitting that maximizes the achievable sum-rate.

This chapter studies the achievable sum-rate of the two-user GIC, when the HK scheme uses Gaussian inputs. Note that the optimal distribution of the inputs is not known. However, in all cases where the sum-capacity is known, it is achieved by the HK scheme with Gaussian inputs. In this thesis, we always assume that inputs are Gaussian. First, the full characterization of the achievable sum-rate is found, when no time sharing is used [43]. It is shown that when interference is weak, the achievable sum-rate can have five distinct closed-form expressions. For each expression, the optimal power splitting that achieves the maximum sum-rate is found. Moreover, for given channel gains and given transmit powers, the optimal strategy that achieves the maximum sum-rate is derived. In doing so, we characterize an optimization problem that formulates the maximum HK sum-rate. We show that this optimization problem is challenging, as it involves a non-differentiable objective function. The main contribution of Chapter 3 is the characterization of the solution to the optimization problem. Since the proof is involved, we divide the proof into several steps and examine each step separately. Moreover, we use several figures to visually illustrate each step.

In Chapter4, we show that the approach used in Chapter 3 for finding the maximum sum-rate can be adopted to find the support function of the HK rate region, i.e, the maximum of any linear combination of the individual rates. Accordingly, we express the optimal power-splitting strategy that achieves any boundary point of the HK scheme with Gaussian inputs and no time sharing. More importantly, we examine the role of the time-sharing variable in enlarging the achievable rate region. We show that, for the weak interference class, the optimization problem over the time-sharing variable and the power-splitting variables can be decoupled. Relying on this idea, we can significantly decrease the complexity of the HK rate region for the weak interference class.

The rest of this chapter is organized as follows. Section 2.2 introduces the channel model and existing results and reviews different classes of interference. Then we examine the existing results on the capacity region of each class. In Section 2.3, the maximum sum-rate of the HK scheme is studied for the two-user GIC with weak interference. This section, which demonstrates how power is allocated among public and private messages, contains the main contributions of Chapter 3. Finally, Section 2.4 concludes the chapter.

2.2 Channel Model and Preliminaries

In this chapter, the following notations are used. Random variables are denoted by upper case letters. $N(m, \sigma)$ represents the Gaussian distribution with mean m and variance σ^2. The notation $[1 : n]$ represents the set of integers from 1 to n, and $a \doteq b$ means b is the definition of a. $C(x) \doteq \frac{1}{2}\log(1 + x)$ where $\log(x) \doteq \log_2(x)$. The notation $\mathbb{1}(x \geq y) = 1$ if $x \geq y$, otherwise $\mathbb{1}(x \geq y) = 0$. Moreover, $[x]^+ \doteq \max\{x, 0\}$. The expectation with respect to a random variable Z is expressed by $\mathbb{E}_Z$. For a set Q, $|Q|$ denotes the size of the set. Finally, $\oplus$ represents addition modulo 2 and $\mathbb{R}_+^2$ represents the set of all (R_1, R_2), such that both R_1 and R_2 are non-negative real numbers.

The two-user GIC is modeled by the following expressions:

$$\begin{aligned} &Y_1 = X_1 + \sqrt{a}X_2 + Z_1,\ a \in \mathbb{R}_+, \\ &Y_2 = X_2 + \sqrt{b}X_1 + Z_2,\ b \in \mathbb{R}_+, \\ &Z_i \sim N(0, 1),\ \mathbb{E}[(X_i)^2] \leq P_i,\ i \in \{1, 2\}, \end{aligned} \tag{2.1}$$

where X_i is transmitted by the i^{th} transmitter and Y_i is received by the i^{th} receiver. The i^{th} encoder assigns a codeword $X_i^n(m_i)$ to each message $m_i \in [1 : 2^{nR_i}]$, where n is the length of the codeword and R_i is the rate of the i^{th} transmitter. The gains of the cross-link channels, which affect the power of interference, are represented by $\sqrt{a}$ and $\sqrt{b}$. Additionally, the i^{th} transmitter has limited power P_i to transmit its message. The capacity region of the two-user GIC is the closure of all $(R_1, R_2) \in \mathbb{R}_+^2$, such that each receiver is able to decode its intended message with arbitrarily small probability of error as n approaches infinity.

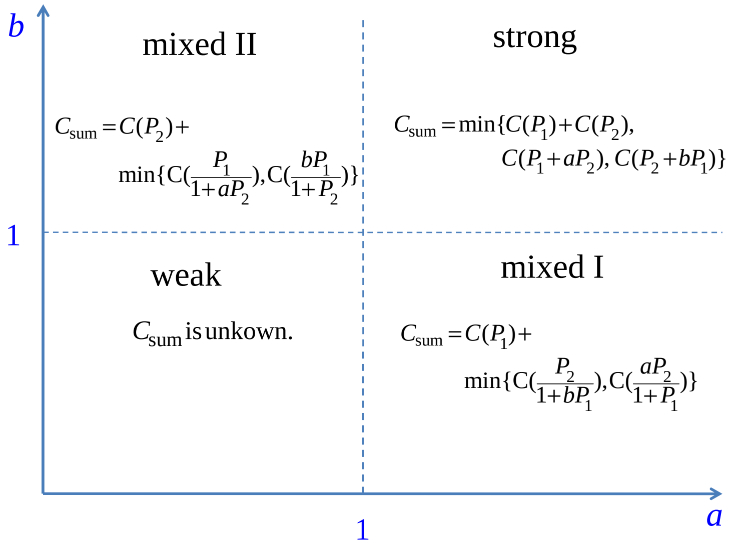

Figure 2.1: Classes of interference and the corresponding sum-capacity expressions.

2.2.1 Classes of Interference and the Corresponding Sum-Capacity

Based on the values of a, b, P_1, and P_2, the interference is categorized into several classes as shown in Figure 2.1. Note that each class is a region in $\mathbb{R}^4_+$. However, to demonstrate each class, we use one of the following ways: either for a given (P_1, P_2), the projection of the class onto the ab-plane is depicted or for a given (a, b), the projection of the class onto the P_1P_2-plane is depicted. Four main classes of interference are defined as follows: If $a \geq 1$ and $b \geq 1$, the interference is *strong*. If either $0 < a < 1$ and $b \geq 1$ or $0 < b < 1$ and $a \geq 1$, the interference is *mixed*. For more clarity, we refer to the class corresponding to $0 < b < 1$ and $a \geq 1$ as *mixed I*. Similarly, we refer to the class corresponding to $0 < a < 1$ and $b \geq 1$ as *mixed II*. Moreover, if $a < 1$ and $b < 1$, the interference is *weak*.

To investigate one class, we partition it into some sub-classes. For instance, in the strong interference class, $a \geq 1 + P_1$ and $b \geq 1 + P_2$ specify the *very strong* interference sub-class. The weak interference class is the focus of this chapter. Therefore, we focus on some sub-classes within the weak interference class, namely *very weak, somewhat weak, and barely weak* sub-classes. For $a < 1$ and $b < 1$, the *very weak* interference sub-class is specified by $P_1\sqrt{b} + P_2\sqrt{a} \leq \frac{1-\sqrt{a}-\sqrt{b}}{\sqrt{ab}}$ [10, 37]. As shown in Figure 2.4, we refer to $P_1 \leq \frac{1-a}{ab}$, $P_2 \leq \frac{1-b}{ab}$ as the *somewhat weak* interference sub-class and $P_1 > \frac{1-a}{ab}$, $P_2 > \frac{1-b}{ab}$ as the *barely weak* interference sub-class.

The sum-capacity of the two-user GIC is known, when the interference is strong [7], or when the interference is mixed [10]. However, when the interference is weak, the sum-capacity is not known in general. Define C_{sum} as the sum-capacity of the two-user GIC. Figure 2.1 shows the main classes and the corresponding sum-capacity expressions. For the weak interference class, the sum-capacity is known only for the very weak interference sub-class [10, 11]. This sub-class is shown in Figure 2.2.

Moreover, we can partition each class into some sub-classes, such that inside each sub-class, C_{sum} is given by a single expression. To this end, we define the following sub-classes. For the strong interference class, the entire capacity region is achieved by SND [7, 37], therefore, C_{sum} is given by

$$C_{\text{sum}} = \min \left\{ \begin{matrix} C(P_1 + aP_2), C(P_2 + bP_1), \\ C(P_1) + C(P_2) \end{matrix} \right\}. \tag{2.2}$$

Note that the strong interference class can be partitioned into three sub-classes, such that in each sub-class, C_{sum} is given by one of the terms inside the $\min\{\}$ in (2.2). In fact, when the interference is *very strong*, i.e., $a \geq 1 + P_1$ and $b \geq 1 + P_2$, (2.2) reduces to

$$C_{\text{sum}} = C(P_1) + C(P_2). \tag{2.3}$$

In addition, when $1 \leq b < 1 + P_2$ and $a \geq b$, (2.2) reduces to

$$C_{\text{sum}} = C(P_2 + bP_1). \tag{2.4}$$

We refer to this sub-class as the *mixedly strong I* sub-class.

Similarly, we refer to $1 \leq a < 1 + P_1$ and $b \geq a$ as the *mixedly strong II* sub-class. In this sub-class, (2.2) reduces to

$$C_{\text{sum}} = C(P_1 + aP_2). \tag{2.5}$$

Figure 2.2 shows how the entire strong interference class is partitioned into these three sub-classes.

Moreover, for the mixed I and mixed II classes, the sum-capacity is known [10]. In fact, [10] shows that for the mixed I class, C_{sum} is given by

$$C_{\text{sum}} = C(P_1) + \min\left\{C\left(\frac{P_2}{1 + bP_1}\right), C\left(\frac{aP_2}{1 + P_1}\right)\right\}, \tag{2.6}$$

and for the mixed II class, C_{sum} is given by

$$C_{\text{sum}} = C(P_2) + \min\left\{C\left(\frac{P_1}{1 + aP_2}\right), C\left(\frac{bP_1}{1 + P_2}\right)\right\}. \tag{2.7}$$

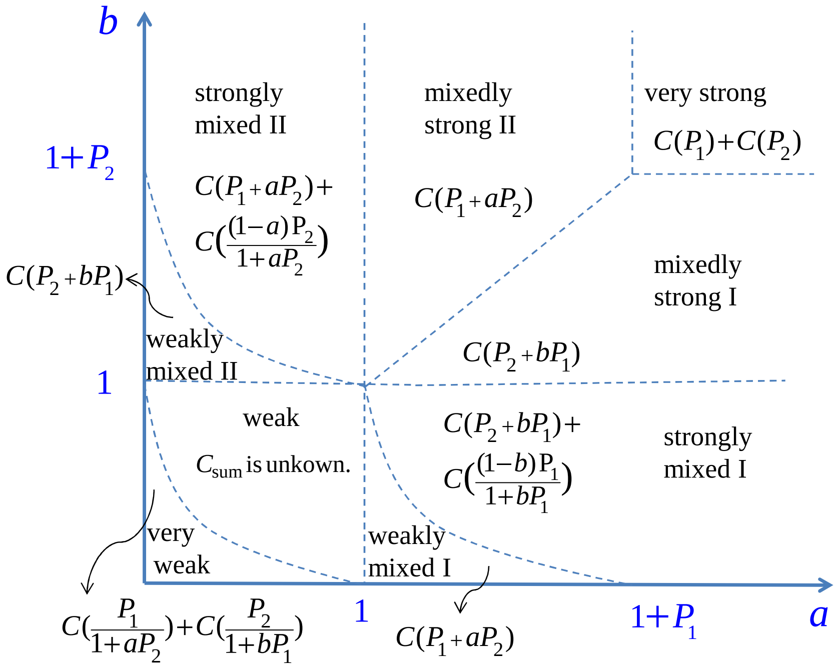

Figure 2.2: All sub-classes of interference for which the sum-capacity is known.

Note that we can partition the mixed I class into two sub-classes, such that for each sub-class the sum-capacity is given by one of the terms inside the min{} in equation (2.6). Accordingly, we define the following sub-classes inside the mixed I and mixed II classes: the *weakly mixed I* sub-class satsifies $0 \le b < 1$ and $1 \le a \le \frac{1+P_1}{1+bP_1}$. The *strongly mixed I* sub-class satisfies $0 \le b < 1$ and $a > \frac{1+P_1}{1+bP_1}$. Note that for the weakly mixed I sub-class, (2.6) reduces to

$$C_{\text{sum}} = C(P_1 + aP_2), \tag{2.8}$$

and for the strongly mixed I sub-class, (2.6) reduces to

$$C_{\text{sum}} = C(P_2 + bP_1) + C\Big(\frac{(1-b)P_1}{1+bP_1}\Big). \tag{2.9}$$

Similarly, the *weakly mixed II* sub-class satisfies $0 \le a < 1$, $1 \le b \le \frac{1+P_2}{1+aP_2}$, and the *strongly mixed II* sub-class satisfies $0 \le a < 1$, $b > \frac{1+P_2}{1+aP_2}$. For the weakly mixed II sub-class, (2.7) reduces to

$$C_{\text{sum}} = C(P_2 + bP_1), \tag{2.10}$$

and for the strongly mixed II sub-class, (2.7) reduces to

$$C_{\text{sum}} = C(P_1 + aP_2) + C\Big(\frac{(1-a)P_2}{1+aP_2}\Big). \tag{2.11}$$

Figure 2.2 shows how the mixed I and mixed II classes are partitioned into four subclasses.

Note that whenever the sum-capacity is known, it is achieved by the HK scheme with Gaussian inputs. Therefore, a step toward finding the sum-capacity can be the full characterization of the achievable sum-rate of the HK scheme. In the following, we briefly review the HK scheme. In particular, we review all cases for which the maximum achievable sum-rate of the HK scheme is known.

2.2.2 Han-Kobayashi Coding Scheme

The HK scheme divides each message M_i, $i \in \{1, 2\}$ into two parts: public and private. Following the notation of [37], M_{ii} represents the private message at rate R_{ii} and M_{i0} represents the public message at rate R_{i0}. Consequently, $R_i = R_{ii} + R_{i0}$. Each encoder uses superposition coding to encode its message: M_{i0} is encoded by the cloud center U_i and (M_{i0}, M_{ii}) is encoded by the satellite codeword X_i. Then, using two power-splitting variables, λ_1 and λ_2, each transmitter splits its available power between its public and private messages. In fact, since the total power of the i^{th} transmitter is limited to P_i, the total power is divided between the messages: $\lambda_i P_i$ is allocated to M_{ii} and $(1 - \lambda_i)P_i$ is allocated to M_{i0}, where $0 \le \lambda_i \le 1$.

The i^{th} decoder uses joint decoding and finds the unique $(\hat{m}_{i0}, \hat{m}_{ii})$ and some $\hat{m}_{(i\oplus1)0}$, such that $(u_i^n(\hat{m}_{i0}), u_{i\oplus1}^n(\hat{m}_{(i\oplus1)0}), x_i^n(\hat{m}_{i0}, \hat{m}_{ii}), y_i^n)$ are jointly typical. Note that we did not consider the time-sharing variable Q. Moreover, the optimal input distribution is still an open problem. In this study, we assume that both transmitters use Gaussian distribution. For fixed values of (λ_1, λ_2), the average probability of error at decoders approaches zero as the block length goes to infinity, if R_1 and R_2 satisfy the following inequalities:

$$
\begin{aligned}
R_1 <& C\Big(\frac{P_1}{1+a\lambda_2P_2}\Big),\\
R_2 <& C\Big(\frac{P_2}{1+b\lambda_1P_1}\Big),\\
R_1+R_2 <& C\Big(\frac{P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2}{1+b\lambda_1P_1}\Big),\\
R_1+R_2 <& C\Big(\frac{P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big)+C\Big(\frac{\lambda_1P_1}{1+a\lambda_2P_2}\Big),\\
R_1+R_2 <& C\Big(\frac{\lambda_1P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big),\\
2R_1+R_2 <& C\Big(\frac{P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_1P_1}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big),\\
R_1+2R_2 <& C\Big(\frac{P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big)+C\Big(\frac{\lambda_2P_2}{1+b\lambda_1P_1}\Big)+C\Big(\frac{\lambda_1P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big).
\end{aligned}
\tag{2.12}
$$

It is worth mentioning that (2.12) is in fact the simplified description of the HK constraints, presented in [44]. Define $\mathcal{G}_0$ as all the rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy all the constraints of (2.12). Clearly, $\mathcal{G}_0$ is a function of $(P_1, P_2, \lambda_1, \lambda_2)$.

Moreover, if R_1 and R_2 satisfy all inequalities of (2.12), then the maximum achievable sum-rate for a fixed power splitting (λ_1, λ_2) is denoted by $R_{\text{sum-HK}}$, as expressed in the following equation:

$$
R_{\text{sum-HK}}(\lambda_1,\lambda_2) \doteq \max_{R_1,R_2\in\mathcal{G}_0} R_1+R_2. \tag{2.13}
$$

In this chapter, we investigate the maximum achievable sum-rate of the HK scheme with Gaussian inputs. We first investigate the following optimization problem:

$$
R_{\text{sum-HK}}^{\max} \doteq \max_{\lambda_1,\lambda_2\in[0,1]} R_{\text{sum-HK}}(\lambda_1,\lambda_2). \tag{2.14}
$$

This optimization problem characterizes the optimal power allocation that maximizes the HK sum-rate without time sharing. Although time sharing can strictly increase the HK sum-rate, for all the cases that the sum-capacity is known, it is achieved by the HK scheme without time sharing. Therefore, in the following, we first review the existing results on the maximum HK sum-rate without time sharing.

2.2.3 Sum-Capacity versus Maximum HK Sum-Rate

It is important to note that for all cases in which the sum-capacity C_{sum} is known, we have $C_{\text{sum}} = R_{\text{sum-HK}}^{\max}$. This means, although the time sharing variable Q can increase the

Sub-class	C_{sum}	$R_{\text{sum-HK}}^{\max}$	**Optimal** $(\lambda_1^*,\lambda_2^*)$	**Ref.**
$P_1\sqrt{b}+P_2\sqrt{a}\le\frac{1-\sqrt{a}-\sqrt{b}}{\sqrt{ab}}$	$C(\frac{P_1}{1+aP_2})+C(\frac{P_2}{1+bP_1})$	$C(\frac{P_1}{1+aP_2})+C(\frac{P_2}{1+bP_1})$	(1,1)	[10-12]
$0\le P_1\le\frac{1-a}{ab}, 0\le P_2\le\frac{1-b}{ab}$	**unknown**	$C(\frac{P_1}{1+aP_2})+C(\frac{P_2}{1+bP_1})$	(1,1)	[41]
$0\le P_1\le\frac{1-a}{ab}, P_2>\frac{1-b}{ab}$	**unknown**	$C(P_2+bP_1)$	(0,1)	[41]
$P_1>\frac{1-a}{ab}, 0\le P_2\le\frac{1-b}{ab}$	**unknown**	$C(P_1+aP_2)$	(1,0)	[41]
$P_1>\frac{1-a}{ab}, P_2>\frac{1-b}{ab}$	**unknown**	**Theorem 1**	**Theorem 1**	

Table 2.1: Sub-classes in the weak interference class and the corresponding sum-capacity expressions and maximum sum-rate expressions.

achievable sum-rate, for all cases that the sum-capcity is knwon, it is achieved without any time sahring. Let $R_{\text{sum-HK}}^{\text{max-Q}}$ denote the achievable sum-rate when time sharing is used. Although $R_{\text{sum-HK}}^{\text{max-Q}} \geq R_{\text{sum-HK}}^{\max}$, for all cases that the sum-capacity is known, we have $C_{\text{sum}} = R_{\text{sum-HK}}^{\max} = R_{\text{sum-HK}}^{\text{max-Q}}$.

For instance, in the strong and the mixed interfere classes, $R_{\text{sum-HK}}^{\text{max-Q}}$ is known. For these classes, C_{sum} is known and $R_{\text{sum-HK}}^{\text{max-Q}} = R_{\text{sum-HK}}^{\max} = C_{\text{sum}}$, as shown in Figure 2.1. However, for the weak interference class, $R_{\text{sum-HK}}^{\max}$ is known only for a few sub-classes. A primary goal of this chapter is to find $R_{\text{sum-HK}}^{\max}$ for the entire weak interference class.

We first review all known results for the weak interference class. When interference is somewhat weak, reference [41] shows that treating interference as noise achieves $R_{\text{sum-HK}}^{\max}$. Therefore, in this sub-class, $R_{\text{sum-HK}}^{\max}$ is given by

$$R_{\text{sum-HK}}^{\max} = C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big). \tag{2.15}$$

Moreover, when $a \leq \frac{1}{1+P_1 b}$ and $b > \frac{1}{1+P_2 a}$, $R_{\text{sum-HK}}^{\max}$ is given by the same expression corresponding to the mixed II class, i.e., $b \geq 1$ and $a < 1$ [41]. Therefore,

$$R_{\text{sum-HK}}^{\max} = C(P_2) + \min\Big\{C\Big(\frac{P_1}{1+aP_2}\Big), C\Big(\frac{bP_1}{1+P_2}\Big)\Big\}. \tag{2.16}$$

Note that, for the weak interference class, $C\Big(\frac{P_1}{1+aP_2}\Big) \geq C\Big(\frac{bP_1}{1+P_2}\Big)$, and therefore, for $a \leq \frac{1}{1+P_1 b}$ and $b > \frac{1}{1+P_2 a}$, we have

$$R_{\text{sum-HK}}^{\max} = C(P_2 + bP_1). \tag{2.17}$$

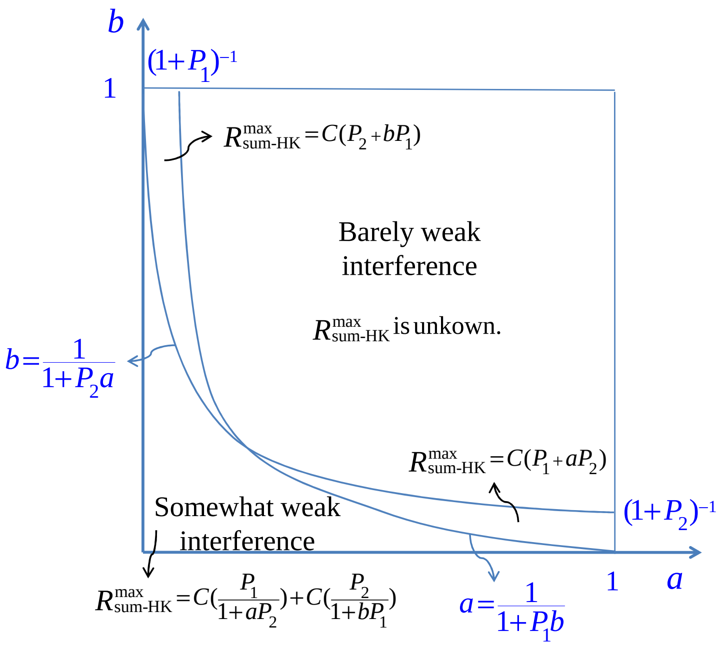

Figure 2.3: For fixed values of P_1 and P_2, the weak interference class is partitioned into four sub-classes. These sub-classes and their corresponding maximum sum-rate expressions are demonstrated in the ab-plane.

Similarly, when $a > \frac{1}{1+P_1 b}$ and $b \leq \frac{1}{1+P_2 a}$, $R^{\max}_{\text{sum-HK}}$ is given by

$$R^{\max}_{\text{sum-HK}} = C(P_1 + aP_2). \tag{2.18}$$

Table 2.1 summarizes all sub-classes for which $R^{\max}_{\text{sum-HK}}$ is known.

Note that, for the barely weak interference sub-class, $R^{\max}_{\text{sum-HK}}$ has been unknown. This chapter characterizes $R^{\max}_{\text{sum-HK}}$ for the barely weak interference sub-class. In Theorem 2.1, we partition the barely weak interference sub-class into four smaller sub-classes, and for each sub-class, we characterize $R^{\max}_{\text{sum-HK}}$.

Figure 2.3 demonstrates the sub-classes of Table 2.1 in the ab-plane. As mentioned earlier, it is traditional to use the ab-plane to investigate different interference classes. In fact, Figure 2.3 shows all sub-classes of Table 2.1 in the ab-plane, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is depicted. However, it turns out that it would be easier to investigate all these sub-classes in the P_1P_2-plane. Figure 2.4 shows all sub-classes of Table 2.1 in the P_1P_2-plane. In the following, as the main contribution of this chapter, we explicitly determine the optimal power splitting policy that maximizes the sum-rate.

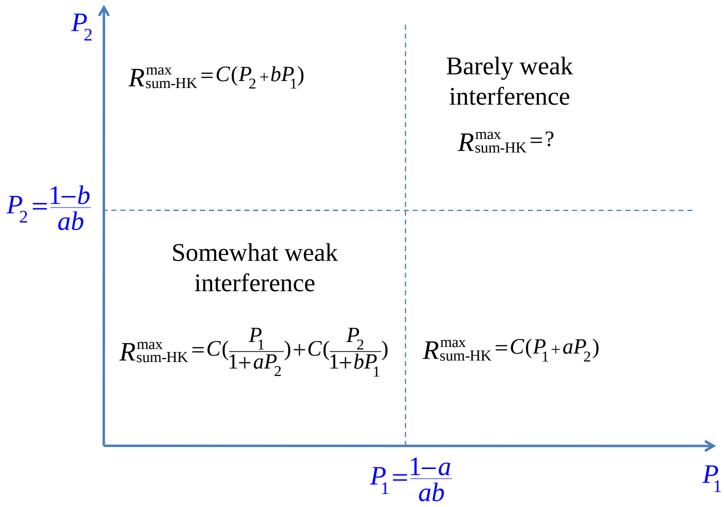

Figure 2.4: For fixed values of a and b, the weak interference class is partitioned into four sub-classless. These sub-classless and the corresponding maximum sum-rate are demonstrated in the P_1P_2-plane.

2.3 Maximum HK Sum-Rate without Time Sharing

In this section, the maximum achievable sum-rate of the two-user GIC is investigated when the HK scheme is used. The mathematical optimization problem that characterizes the maximum sum-rate of the HK scheme is presented. Our main result is the solution to this optimization problem. Note that Chapter 3 does not consider time sharing. Chapter 4 shows how time sharing increases the achievable sum-rate.

2.3.1 Main Results

Theorem 2.1 is the main result of this chapter. In this theorem, we characterize the achievable sum-rate of the two-user GIC when the HK scheme is used.

Theorem 2.1. *For the two-user Gaussian interference channel, when interference is weak, the maximum achievable sum-rate of the HK scheme with Gaussian inputs and no* *time sharing is given by*

$$\begin{aligned}
&R^{\max}_{\text{sum-HK}} = \\
\max\Big\{ &C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big), \\
&C(P_1 + aP_2), \\
&C(P_2 + bP_1), \\
&C(P_1 + aP_2) + g(\tilde{\lambda}_1, \tilde{\lambda}_2)\mathbb{1}(\tilde{\lambda}_1 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq 0)\mathbb{1}(\hat{\lambda}_2 \geq \tilde{\lambda}_2), \\
&C(P_1 + aP_2) + g(\hat{\lambda}_1, \hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0)\mathbb{1}(\hat{\lambda}_2 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq \hat{\lambda}_2)\Big\},
\end{aligned} \tag{2.19}$$

where $g(\lambda_1, \lambda_2) \doteq C(\frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1+a\lambda_2 P_2}) - C(b\lambda_1 P_1)$ *and* $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ *is given by*

$$\begin{aligned}
\tilde{\lambda}_1 &= ab - \frac{1-a}{P_1}, \\
\tilde{\lambda}_2 &= ab - \frac{1-b}{P_2}.
\end{aligned} \tag{2.20}$$

Moreover, $\hat{\lambda}_2$ *is the non-negative solution of the following equation:*

$$(\lambda_2^2) + 2\frac{(1+bP_1c)}{(bP_1m + P_2)}(\lambda_2) + \frac{(1+bP_1c)(abP_1c + a - 1)}{abP_1m(bP_1m + P_2)} = 0, \tag{2.21}$$

and $\hat{\lambda}_1$ *is given by*

$$\hat{\lambda}_1 = m\hat{\lambda}_2 + c, \tag{2.22}$$

where m *and* c *are given by*

$$m \doteq \frac{P_2\big((1-a) + P_1(1-ab)\big)}{P_1\big(1 - b + P_2(1-ab)\big)}, \tag{2.23}$$

$$c \doteq \frac{P_1(1-b) - P_2(1-a)}{P_1\big(1 - b + P_2(1-ab)\big)}. \tag{2.24}$$

Theorem 2.1 shows that the maximum sum-rate of the HK scheme can have five distinct mathematical expressions, depending on the values of a, b, P_1, and P_2. In fact, this theorem partitions the weak interference class into five sub-classes. For each sub-class, Theorem 2.1 computes $R^{\max}_{\text{sum-HK}}$. Note that each sub-class is a region in $\mathbb{R}^4_+$. We demonstrate these sub-classes in two different planes: the P_1P_2-plane and the ab-plane.

Figure 2.5 shows the P_1P_2-plane. This figure shows that quadrant I of the P_1P_2-plane is partitioned into five regions, such that in each region, one of the expressions given in

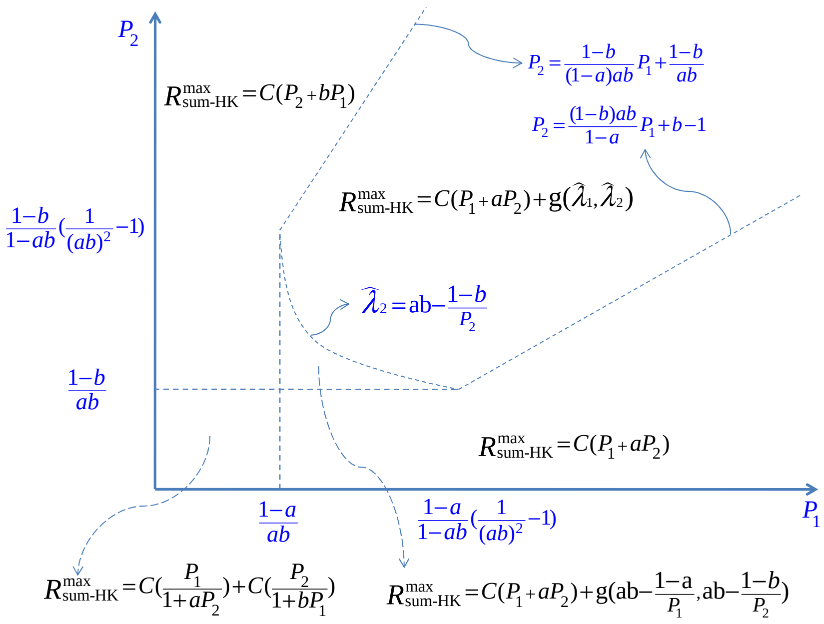

Figure 2.5: The maximum achievable sum-rate of the HK scheme ($R^{\max}_{\text{sum-HK}}$) for the two-user GIC with weak interference. The weak interference class is partitioned into five sub-classes. For fixed (a, b), these sub-classes are demonstrated in the P_1P_2-plane, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is characterized.

(2.19) is the maximum achievable sum-rate. Similarly, Figure 2.6 shows that quadrant I of the ab-plane is partitioned into five regions, such that in each region, one of the expressions given in (2.19) is the maximum achievable sum-rate.

Theorem 2.1 demonstrates the maximum achievable sum-rate expressions but does not show the optimal power splitting. Each sum-rate expression is achieved by a particular pair of power-splitting variables (λ_1, λ_2) as explained in the following theorem:

Theorem 2.2. *For the two-user Gaussian interference channel, when interference is weak, the maximum achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing is given by*

$$R^{\max}_{\text{sum-HK}} = R_{\text{sum-HK}}(\lambda_1^\star, \lambda_2^\star), \tag{2.25}$$

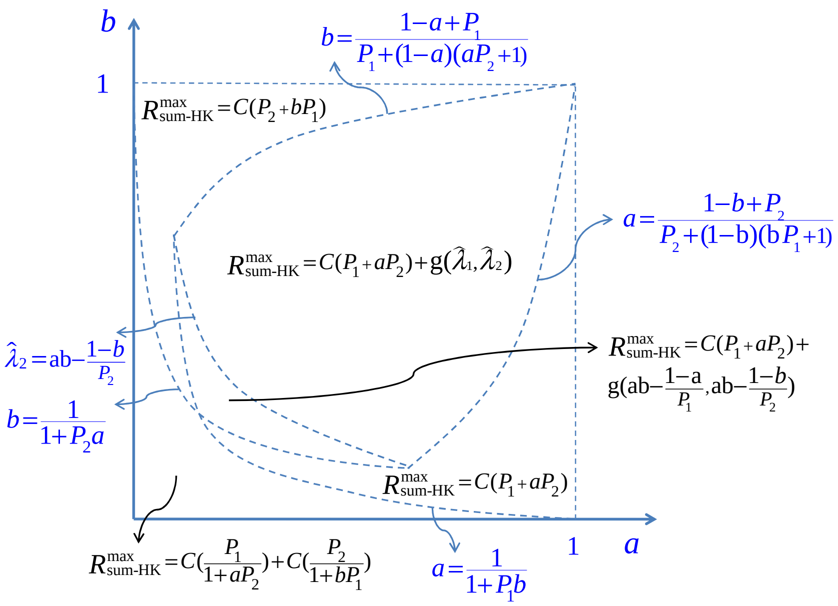

Figure 2.6: The maximum achievable sum-rate of the HK scheme for the two-user GIC with weak interference. The weak interference class is partitioned into five sub-classes, and for each sub-class, $R_{\text{sum-HK}}^{\max}$ is characterized.

where $R_{\text{sum-HK}}$ *is defined in (2.13) and the optimal power splitting* $(\lambda_1^\star, \lambda_2^\star)$ *is given by*

$$(\lambda_1^\star,\lambda_2^\star)=\begin{cases}(0,0) & \text{if } (a,b,P_1,P_2)\in \textit{somewhat weak sub-class},\\ (\lambda_1^\star\geq c,0) & \text{if } (a,b,P_1,P_2)\in \textit{weakly mixed I sub-class},\\ (0,\lambda_2^\star\geq c') & \text{if } (a,b,P_1,P_2)\in \textit{weakly mixed II sub-class},\\ (\tilde{\lambda}_1,\tilde{\lambda}_2) & \text{if } (a,b,P_1,P_2)\in \textit{power splitting I sub-class},\\ (\hat{\lambda}_1,\hat{\lambda}_2) & \text{if } (a,b,P_1,P_2)\in \textit{power splitting II sub-class},\end{cases} \tag{2.26}$$

where c *and* c' *are given by*

$$\begin{aligned} c &\doteq \frac{P_1(1-b)-P_2(1-a)}{P_1\big(1-b+P_2(1-ab)\big)},\\ c' &\doteq \frac{P_2(1-a)-P_1(1-b)}{P_2\big(1-a+P_1(1-ab)\big)},\end{aligned} \tag{2.27}$$

and the descriptions of all sub-classes are given in Table 2.2.

Table 2.2 shows how the entire weak interference class is partitioned into five sub-classes. For each sub-class, the maximum achievable sum-rate and $(\lambda_1^\star, \lambda_2^\star)$, the optimal values of λ_1 and λ_2 that result in the maximum achievable sum-rate, are specified. Note that the optimal power splitting is unique for three sub-classes, namely somewhat weak,

Sub-class Name	Sub-class Description	$R^{\max}_{\text{sum-HK}}$	Optimal $(\lambda_1^*,\lambda_2^*)$
Somewhat Weak Interference	$0 \le P_1 \le \frac{1-a}{ab}$, $0 \le p_2 \le \frac{1-b}{ab}$.	$C(\frac{P_1}{1+aP_2})+C(\frac{P_2}{1+bP_1})$	$(1,1)$
Weakly Mixed Interference I	$P_1 > \frac{1-a}{ab}$, $0 \le P_2 \le \max\{\frac{1-b}{ab}, \frac{(1-b)ab}{1-a}P_1+b-1\}$.	$C(P_1+aP_2)$	$(0,1)$
Weakly Mixed Interference II	$P_2 > \frac{1-b}{ab}$, $0 \le P_1 \le \max\{\frac{1-a}{ab}, \frac{(1-a)ab}{1-b}P_1+a-1\}$.	$C(P_2+bP_1)$	$(1,0)$
Non-zero Power Splitting I	$P_1 > \frac{1-a}{ab}$, $p_2 > \frac{1-b}{ab}$, $\hat{\lambda}_2 \ge ab-\frac{1-b}{P_2}$.	$C(P_1+aP_2)+$ $g(\lambda_1^*,\lambda_2^*)$	$\lambda_1^* = ab-\frac{1-a}{P_1}$, $\lambda_2^* = ab-\frac{1-b}{P_2}$.
Non-zero Power Splitting II	$P_1 > \frac{(1-a)ab}{1-b}P_1+a-1$, $P_2 > \frac{(1-b)ab}{1-a}P_1+b-1$, $\hat{\lambda}_2 < ab-\frac{1-b}{P_2}$.	$C(P_1+aP_2)+$ $g(\hat{\lambda}_1,\hat{\lambda}_2)$	$(\hat{\lambda}_1,\hat{\lambda}_2)$

Table 2.2: The weak interference class is partitioned into five sub-classes. For each sub-class, the optimal power splitting $(\lambda_1^\star, \lambda_2^\star)$ and the corresponding optimal sum-rate $R^{\max}_{\text{sum-HK}}$ are given.

power splitting I, and power splitting II. However, for the weakly mixed I sub-class any $(\lambda_1^\star, 0)$ that satisfies $c \le \lambda_1^\star \le 1$ is an optimal power splitting. Similarly, for the weakly mixed II sub-class, any $(0, \lambda_2^\star)$ that satisfies $c' \le \lambda_1^\star \le 1$ is an optimal power splitting.

To prove Theorem 2.1 and Theorem 2.2, we first need to derive an optimization problem that characterizes the maximum sum-rate of the HK scheme, as provided in the following.

2.3.2 The Optimization Problem Corresponding to the Maximum HK Sum-Rate

The HK scheme imposes several bounds on the achievable sum-rate. For a given interference class, some of these bounds may not be active. Consequently, one can simplify the mathematical expression that characterizes the maximum HK sum-rate. In the following theorem, we show that, for the weak interference class, exactly three bounds are active.

Theorem 2.3. *For the two-user Gaussian interference channel with weak interference, the maximum achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing is given by the following optimization problem:*

$$
\begin{aligned}
&R_{\text{sum-HK}}^{\max} = \\
&\max_{\lambda_1,\lambda_2\in[0,1]} \Bigg[C(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}) + C(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}) + \\
&\quad \min\Big\{ C(\frac{\bar{\lambda}_1 P_1 + a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}), C(\frac{\bar{\lambda}_2 P_2 + b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}), \\
&\qquad C(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}) + C(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}) \Big\} \Bigg]
\end{aligned} \tag{2.28}
$$

.

Proof. To be presented after Lemma 2.1. □

We first need to find a compact upper bound on the achievable sum-rate of the HK scheme. If we directly use (2.12) to find an upper bound on the maximum achievable sum-rate of the HK scheme, we would obtain the following optimization problem:

$$
\begin{aligned}
&R_{\text{sum-HK}}^{\max} = \\
&\max_{\lambda_1,\lambda_2\in[0,1]} \Bigg[\min\Big\{ C(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}) + C(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}), \\
&\quad C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}\Big), \\
&\quad C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big) + C\Big(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}\Big), \\
&\quad C\Big(\frac{\lambda_1 P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2 + \bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big), \\
&\quad \frac{1}{2}\Big(C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}\Big) + \\
&\qquad C\Big(\frac{\lambda_2 P_2 + \bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big) + C\Big(\frac{P_2}{1+b\lambda_1 P_1}\Big)\Big), \\
&\quad \frac{1}{2}\Big(C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big) + C\Big(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}\Big) + \\
&\qquad C\Big(\frac{\lambda_1 P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{P_1}{1+a\lambda_2 P_2}\Big)\Big) \Big\} \Bigg].
\end{aligned} \tag{2.29}
$$

Note that (2.29) involves minimization over six different functions, whereas (2.28) involves minimization over only three functions. We prove that, for the weak interference class, (2.29) and (2.28) are equivalent. To do so, we first look at the HK scheme in detail. The HK scheme jointly decodes the common messages and the intended private message at each receiver. However, we show that, if the common messages are decoded first while treating the private messages as noise, and then the private messages are decoded, the HK achievable sum-rate does not decrease. The following lemma describes this point.

Lemma 2.1. *For the two-user interference channel, if the HK scheme first decodes the common messages while treating the private messages as noise and then decodes the private messages, it achieves the sum-rate of the classical HK scheme, in which common messages and the intended private message are jointly decoded at each receiver.*

Proof. This idea has been mentioned in [13, 40] without formal proof. For the sake of completeness, we provide its proof. As explained in the previous section, in the HK scheme, decoder 1 finds the unique message pair $(\hat{m}_{10}, \hat{m}_{11})$ and some $\hat{m}_{20}$, such that $(u_1^n(\hat{m}_{10}), u_2^n(\hat{m}_{20}), x_1^n(\hat{m}_{10}, \hat{m}_{11}), y_1^n)$ are jointly typical. It can be shown [37] that the probability of error for the first receiver approaches zero as the block length n goes to infinity, if

$$\begin{aligned} R_{11} &< I(X_1; Y_1|U_1, U_2), \\ R_{11} + R_{10} &< I(X_1; Y_1|U_2), \\ R_{11} + R_{20} &< I(X_1, U_2; Y_1|U_1), \\ R_{11} + R_{10} + R_{20} &< I(X_1, U_2; Y_1). \end{aligned} \tag{2.30}$$

However, if successive decoding is used instead of joint decoding, i.e., decoder 1 first finds the unique pair $(\hat{m}_{10}, \hat{m}_{20})$, and then finds the unique $\hat{m}_{11}$, the probability of error approaches zero, if

$$\begin{aligned} R_{11} &< I(X_1; Y_1|U_1, U_2), \\ R_{10} &< I(U_1; Y_1|U_2), \\ R_{20} &< I(U_2; Y_1|U_1), \\ R_{10} + R_{20} &< I(U_1, U_2; Y_1). \end{aligned} \tag{2.31}$$

Note that we have

$$I(X_1; Y_1|U_2) = I(U_1; Y_1|U_2) + I(X_1; Y_1|U_1, U_2), \tag{2.32}$$

$$I(X_1, U_2; Y_1|U_1) = I(U_2; Y_1|U_1) + I(X_1; Y_1|U_1, U_2), \tag{2.33}$$

$$I(X_1, U_2; Y_1) = I(U_1, U_2; Y_1) + I(X_1; Y_1|U_1, U_2). \tag{2.34}$$

Therefore, the rate region characterized by (2.31) is a sub-set of the rate region characterized by (2.30). Similarly, one can prove the same argument for decoder 2. However, using Fourier-Motzkin elimination, one can show that both schemes impose the same set of constraints on $R_1 + R_2 = R_{11} + R_{22} + R_{10} + R_{20}$. In fact, both schemes impose the following four constraints on the sum-rate:

$$R_1 + R_2 < I(X_1; Y_1|U_2) + I(X_2; Y_2|U_1), \tag{2.35}$$

$$R_1 + R_2 < I(X_1; Y_1|U_1, U_2) + I(U_1, X_2; Y_2), \tag{2.36}$$

$$R_1 + R_2 < I(X_2; Y_2|U_1, U_2) + I(U_2, X_1; Y_1), \tag{2.37}$$

$$R_1 + R_2 < I(U_2, X_1; Y_1|U_1) + I(U_1, X_2; Y_2|U_2). \tag{2.38}$$

□

This lemma facilitates finding a compact upper bound on the achievable sum-rate of the HK scheme. We use this lemma to prove Theorem 2.3.

Proof. Here we provide the proof of Theorem 2.3. According to Lemma 2.1, for the two-user GIC, the rate of private messages should satisfy the following constraints:

$$\begin{aligned} R_{11} &\leq C\Big(\frac{\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big), \\ R_{22} &\leq C\Big(\frac{\lambda_2 P_2}{1 + b\lambda_1 P_1}\Big). \end{aligned} \tag{2.39}$$

Similarly, the rate of common messages should satisfy the following constraints:

$$\begin{aligned} R_{10} &\leq C\Big(\frac{\bar{\lambda}_1 P_1}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big), \\ R_{20} &\leq C\Big(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big), \\ R_{10}+R_{20} &\leq C\Big(\frac{\bar{\lambda}_1 P_1 + a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big), \\ R_{20} &\leq C\Big(\frac{\bar{\lambda}_2 P_2}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big), \\ R_{10} &\leq C\Big(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big), \\ R_{10}+R_{20} &\leq C\Big(\frac{\bar{\lambda}_2 P_2 + b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big). \end{aligned} \tag{2.40}$$

Note that the first three bounds in (2.40) are the MAC bounds at receiver Y_1 when common messages with the power of $\lambda_1 P_1 + a\lambda_2 P_2$ are treated as noise. Similarly, the last three bounds in (2.40) are the MAC bounds at receiver Y_2 when common messages with the power of $\lambda_2 P_2 + b\lambda_1 P_1$ are treated as noise.

From (2.39), it is clear that there is only one constraint on $R_{11} + R_{22}$ as follows:

$$R_{11} + R_{22} \leq C\Big(\frac{\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1 + b\lambda_1 P_1}\Big). \tag{2.41}$$

However, (2.40) imposes six constraints on $R_{12} + R_{21}$ as follows:

$$R_{10} + R_{20} \leq \min \Big\{$$

$$C\Big(\frac{\bar{\lambda}_1 P_1}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + b\lambda_2 P_2}\Big), \tag{2.42}$$

$$C\Big(\frac{\bar{\lambda}_1 P_1 + a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big), \tag{2.43}$$

$$C\Big(\frac{\bar{\lambda}_2 P_2}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big) + C\Big(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big), \tag{2.44}$$

$$C\Big(\frac{\bar{\lambda}_2 P_2 + b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big), \tag{2.45}$$

$$C\Big(\frac{\bar{\lambda}_1 P_1}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\bar{\lambda}_2 P_2}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big), \tag{2.46}$$

$$C\Big(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big)\Big\}. \tag{2.47}$$

Note that constraint (2.42) is always looser than (2.43). Similarly, constraint (2.44) is always looser than (2.45). Moreover, (2.46) is the sum of "direct" individual rates, whereas (2.47) is the sum of "cross" individual rates. In the following, we show that constraint (2.46) is always looser than (2.47).

$$\begin{aligned}
&C\Big(\frac{\bar{\lambda}_1 P_1}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\bar{\lambda}_2 P_2}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big)\\
&\geq C\Big(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big)\\
\Leftrightarrow&(1+P_1+a\lambda_2 P_2)(1+P_2+b\lambda_1 P_1)\\
&\geq (1+\lambda_1 P_1 + aP_2)(1+\lambda_2 P_2 + bP_1)\\
\Leftrightarrow&1+P_1(1+b\lambda_1)+P_2(1+a\lambda_2)+\\
&P_1^2 b\lambda_1 + P_2^2 a\lambda_2 + P_1 P_2(1+\lambda_1\lambda_2 ab)\\
&\overset{(a)}{\geq} 1+P_1(b+\lambda_1)+P_2(a+\lambda_2)+P_1^2 b\lambda_1 + P_2^2 a\lambda_2 + P_1P_2(\lambda_1\lambda_2+ab),
\end{aligned} \tag{2.48}$$

where (a) is valid because for $0 \leq a \leq 1, 0 \leq b \leq 1, 0 \leq \lambda_1 \leq 1$, and $0 \leq \lambda_2 \leq 1$, the following inequalities are valid:

$$\begin{aligned}
1+b\lambda_1 &\geq b+\lambda_1,\\
1+a\lambda_2 &\geq a+\lambda_2,\\
1+\lambda_1\lambda_2 ab &\geq \lambda_1\lambda_2+ab.
\end{aligned} \tag{2.49}$$

The above arguments show that $R_{12}+R_{21}$ is upper bounded by only (2.43), (2.45), and (2.47). Moreover, there is only one constraint (2.41) on $(R_{10}+R_{20})$. Since $R_{\text{sum-HK}}(\lambda_1,\lambda_2) = (R_{11}+R_{22})+(R_{10}+R_{20})$, the HK scheme imposes only three constraints on $R_{\text{sum-HK}}(\lambda_1,\lambda_2)$. Therefore, for the weak interference class, $R^{\max}_{\text{sum-HK}}$ is given by

$$\begin{aligned}
R^{\max}_{\text{sum-HK}} &= \max_{\lambda_1,\lambda_2\in[0,1]} R_{\text{sum-HK}}(\lambda_1,\lambda_2)\\
&= \max_{\lambda_1,\lambda_2\in[0,1]} \Bigg[C(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}) + C(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1})+\\
&\qquad \min\Big\{C(\frac{\bar{\lambda}_1 P_1 + a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}), C(\frac{\bar{\lambda}_2 P_2 + b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}),\\
&\qquad C(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}) + C(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1})\Big\}\Bigg],
\end{aligned} \tag{2.50}$$

and this completes the proof of Theorem 2.3. $\square$

We frequently use the following lemma which facilitates deriving compact expressions.

Lemma 2.2. *If P_1, P_2, and N are all positive real numbers, then we have*

$$C\Big(\frac{P_1}{N}\Big)+C\Big(\frac{P_2}{P_1+N}\Big)=C\Big(\frac{P_1+P_2}{N}\Big). \tag{2.51}$$

Proof.

$$\begin{aligned}C\Big(\frac{P_1}{N}\Big)+C\Big(\frac{P_2}{P_1+N}\Big) &=\frac{1}{2}\log\Big(\frac{P_1+N}{N}\Big)+\frac{1}{2}\log\Big(\frac{P_2+P_1+N}{p_1+N}\Big)\\ &=\frac{1}{2}\log\Big(\frac{P_2+P_1+N}{N}\Big)\\ &=C\Big(\frac{P_1+P_2}{N}\Big) \end{aligned}\tag{2.52}$$

□

2.3.3 The Proposed Optimization Technique for Maximizing the HK Sum-Rate

To prove Theorem 2.1, we first review an optimization technique to find the global maximum of an arbitrary function. Note that, according to Fermat's theorem (also known as Interior extremum theorem), the global maximum of a differentiable function f over a feasible region A is achieved at one of the following points: a stationary point or a boundary point [45]. In particular, assume that $f_1(x)$ and $f_2(x)$ are both functions from R^+ to R^+ which are differentiable over [0,1]. Now, consider the following optimization problem:

$$\max_{0\le x\le 1,}\min\Big\{f_1(x), f_2(x)\Big\}. \tag{2.53}$$

Define $f_{\min}(x)\doteq\min\Big\{f_1(x), f_2(x)\Big\}$. We can thus rewrite the optimization problem as

$$\max_{0\le x\le 1,} f_{\min}(x). \tag{2.54}$$

If $f_{\min}(x)$ were differentiable over $[0,1]$, then the optimal solution $x^\star$ would be either a stationary point ($\frac{d}{dx}g(x^\star)=0$) or a boundary point ($x^\star=0$ or $x^\star=1$). Since $f_{\min}(x)=\min\Big\{f_1(x), f_2(x)\Big\}$, $f_{\min}(x)$ may not be differentiable over $[0,1]$; however, since $f_1(x)$ and $f_2(x)$ are both differentiable, the only points at which $f_{\min}(x)$ may not be differentiable is when $f_1(x)=f_2(x)$. Therefore, if $x^\star$ is the optimal solution of (2.53), it belongs to

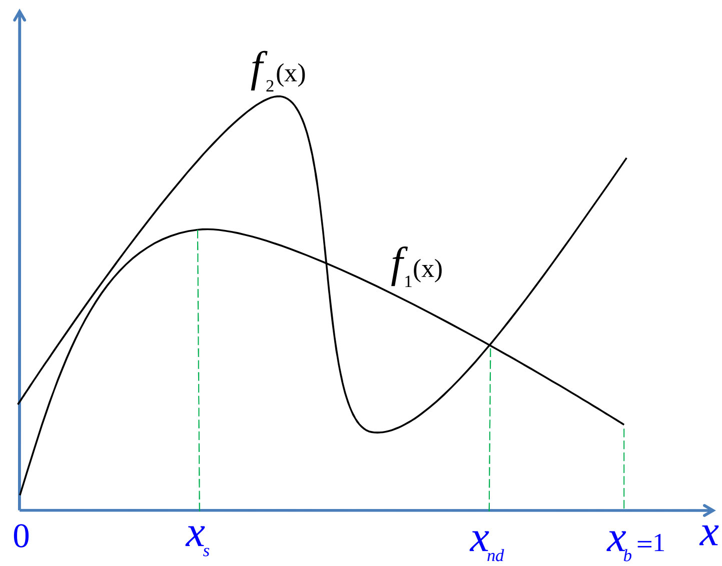

Figure 2.7: To find the maximum of $\min\{f_1(x), f_2(x)\}$ over $[0,1]$, it is sufficient to check all stationary points like x_s and all boundary points like x_b and all non-differentiable points like x_{nd}.

one of the following categories: I- stationary points, II- boundary points, and III- non-differentiable points. Consequently, the search for the optimal solution of (2.53), in the feasible region $[0,1]$, can be restricted to the three categories of points mentioned above.

Relying on this perspective, we can solve the optimization problem in Theorem 2.3. Define $h_0(\lambda_1,\lambda_2) \doteq C(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}) + C(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1})$. In fact, $h_0(\lambda_1,\lambda_2)$ represents the sum-rate of private messages. Moreover, define $h_1(\lambda_1,\lambda_2)$, $h_2(\lambda_1,\lambda_2)$, and $h_3(\lambda_1,\lambda_2)$ as follows:

$$\begin{aligned} h_1(\lambda_1,\lambda_2) \doteq & h_0(\lambda_1,\lambda_2) + C\Big(\frac{\bar{\lambda}_1 P_1 + a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) \\ \overset{(a)}{=} & C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}\Big), \end{aligned} \tag{2.55}$$

$$\begin{aligned} h_2(\lambda_1,\lambda_2) \doteq & h_0(\lambda_1,\lambda_2) + C\Big(\frac{\bar{\lambda}_2 P_2 + b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big) \\ \overset{(b)}{=} & C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big) + C\Big(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}\Big), \end{aligned} \tag{2.56}$$

$$\begin{aligned} h_3(\lambda_1,\lambda_2) \doteq & h_0(\lambda_1,\lambda_2) + C\Big(\frac{a\bar{\lambda}_2 P_2}{1+\lambda_1 P_1 + a\lambda_2 P_2}\Big) + \\ & C\Big(\frac{b\bar{\lambda}_1 P_1}{1+\lambda_2 P_2 + b\lambda_1 P_1}\Big) \\ \overset{(c)}{=} & C\Big(\frac{\lambda_1 P_1 + a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2 + b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big), \end{aligned} \tag{2.57}$$

where (a), (b), and (c) are valid by Lemma 2.2. Then the optimization problem of Theorem 2.3 is equivalent to

$$
\begin{aligned}
R_{\text{sum-HK}}^{\max} &= \max_{\lambda_1,\lambda_2\in[0,1]} R_{\text{sum-HK}}(\lambda_1,\lambda_2) \\
&= \max_{\lambda_1,\lambda_2\in[0,1]} \min\Big\{h_1(\lambda_1,\lambda_2), h_2(\lambda_1,\lambda_2), h_3(\lambda_1,\lambda_2)\Big\}. \qquad (2.58)
\end{aligned}
$$

Similar to (2.53), the search for the optimal solution of (2.58) can be restricted to three categories of points, namely stationary points, boundary points, and non-differentiable points. In the following, we describe each category.

In order to analyze this optimization problem, it is useful to know the condition under which one function inside the min is less than the other function. The following lemma describes this condition.

Lemma 2.3. *For $h_1(\lambda_1,\lambda_2), h_2(\lambda_1,\lambda_2)$ and $h_3(\lambda_1,\lambda_2)$ defined in (2.55-2.57), we have*

$$
\begin{aligned}
\text{A) } h_1(\lambda_1,\lambda_2) \le h_3(\lambda_1,\lambda_2) &\Leftrightarrow P_2 \ge \frac{1-b}{ab-\lambda_2} \text{ or } \lambda_1 = 1 \qquad (2.59)\\
&\Leftrightarrow \lambda_2 \le ab - \frac{1-b}{P_2} \text{ or } \lambda_1 = 1. \qquad (2.60)
\end{aligned}
$$

$$
\begin{aligned}
\text{B) } h_2(\lambda_1,\lambda_2) \le h_3(\lambda_1,\lambda_2) &\Leftrightarrow P_1 \ge \frac{1-a}{ab-\lambda_1} \text{ or } \lambda_2 = 1 \qquad (2.61)\\
&\Leftrightarrow \lambda_1 \le ab - \frac{1-a}{P_1} \text{ or } \lambda_2 = 1. \qquad (2.62)
\end{aligned}
$$

$$
\begin{aligned}
\text{C) } h_1(\lambda_1,\lambda_2) \le h_2(\lambda_1,\lambda_2) \Leftrightarrow\ & P_1((1-b)(1-\lambda_1)) \\
& + P_1P_2((1-ab)(\lambda_2-\lambda_1)) \\
& \le P_2((1-a)(1-\lambda_2)) \qquad (2.63)\\
& \Leftrightarrow \lambda_1 \ge m\lambda_2 + c, \qquad (2.64)
\end{aligned}
$$

where m and c are given in 2.23 and 2.24, respectively.

Proof. The proof is straightforward. In fact, for part A, we have

$$\begin{aligned}
h_1(\lambda_1,\lambda_2) \leq & h_3(\lambda_1,\lambda_2)\\
&\Leftrightarrow C\Big(\frac{P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2}{1+b\lambda_1P_1}\Big)\\
&\leq C\Big(\frac{\lambda_1P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big)\\
&\overset{(a)}{\Leftrightarrow} C\Big(\frac{\bar{\lambda}_1P_1}{1+a\lambda_2P_2+\lambda_1P_1++a\bar{\lambda}_2P_2}\Big)\\
&\leq C\Big(\frac{b\bar{\lambda}_1P_1}{1+b\lambda_1P_1+\lambda_2P_2}\Big)\\
&\Leftrightarrow \lambda_1=1 \text{ or } P_2\geq\frac{1-b}{ab-\lambda_2}\\
&\Leftrightarrow \lambda_1=1 \text{ or } \lambda_2\leq ab-\frac{1-b}{P_2},
\end{aligned} \tag{2.65}$$

where (a) is valid by Lemma 2.2.

The proof of part B follows similarly. For part C, we have

$$\begin{aligned}
h_1(\lambda_1,\lambda_2) \leq & h_2(\lambda_1,\lambda_2)\\
&\Leftrightarrow C\Big(\frac{P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2}{1+b\lambda_1P_1}\Big)\\
&\leq C\Big(\frac{\lambda_1P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2}\Big)+C\Big(\frac{\lambda_2P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\Big)\\
&\overset{(b)}{\Leftrightarrow} C\Big(\frac{\bar{\lambda}_1P_1+a\bar{\lambda}_2P_2}{1+a\lambda_2P_2+\lambda_1P_1}\Big)\\
&\leq C\Big(\frac{\bar{\lambda}_2P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1+\lambda_2P_2}\Big)\\
&\Leftrightarrow P_1(1-b)(1-\lambda_1)+P_1P_2(1-ab)(\lambda_2-\lambda_1)\\
&\leq P_2(1-a)(1-\lambda_2)\\
&\Leftrightarrow \lambda_1\geq m\lambda_2+c,
\end{aligned} \tag{2.66}$$

where (b) is valid by Lemma 2.2. This completes the proof.

□

In the following, we investigate three different categories of points in detail. The optimal power splitting belongs to one of these categories.

2.3.4 Three Categories of Points Corresponding to Optimal Power Splitting

To find the optimal solution of (2.58), we need to investigate the following three categories of points:

I- Stationary Points: If (λ_1, λ_2) is a stationary point of $\min\{h_1(), h_2(), h_3()\}$, then it is a stationary point of $h_1()$ or $h_2()$ or $h_3()$. Therefore, the category of stationary points represents (λ_1, λ_2), such that (λ_1, λ_2) is a stationary point of $h_1()$ or $h_2()$ or $h_3()$ inside the region $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$. Moreover, a stationary point (λ_1, λ_2) corresponding to $h_1()$ can be the optimal solution of (2.58), if we have $h_1(\lambda_1, \lambda_2) \leq \min\left\{h_2(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\right\}$. Similar arguments follow for $h_2()$ and $h_3()$. Since we have three functions, we have three sub-categories of stationary points, namely $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$. These sub-categories are described by the following sets:

$$\begin{aligned}
\mathcal{S}_1 \doteq \Big\{ &(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), \nabla h_1(\lambda_1, \lambda_2) = 0, \\
&h_1(\lambda_1, \lambda_2) \leq \min\left\{h_2(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\right\}\Big\}, \quad (2.67)\\
\mathcal{S}_2 \doteq \Big\{ &(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), \nabla h_2(\lambda_1, \lambda_2) = 0, \\
&h_2(\lambda_1, \lambda_2) \leq \min\left\{h_1(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\Big\}, \quad (2.68)\\
\mathcal{S}_3 \doteq \Big\{ &(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), \nabla h_3(\lambda_1, \lambda_2) = 0, \\
&h_3(\lambda_1, \lambda_2) \leq \min\left\{h_1(\lambda_1, \lambda_2), h_2(\lambda_1, \lambda_2)\Big\}. \quad (2.69)
\end{aligned}$$

II- Boundary Points: Since $0 \leq \lambda_1 \leq 1$ and $0 \leq \lambda_2 \leq 1$, the boundary of the feasible region consists of four line segments. Each line segment is a sub-category of boundary points, as described by the following sets:

$$\mathcal{B}_1 \doteq \{(\lambda_1, 0) : 0 \leq \lambda_1 \leq 1\}, \quad (2.70)$$

$$\mathcal{B}_2 \doteq \{(\lambda_1, 1) : 0 \leq \lambda_1 \leq 1\}, \quad (2.71)$$

$$\mathcal{B}_3 \doteq \{(0, \lambda_2) : 0 \leq \lambda_2 \leq 1\}, \quad (2.72)$$

$$\mathcal{B}_4 \doteq \{(1, \lambda_2) : 0 \leq \lambda_2 \leq 1\}. \quad (2.73)$$

III- Non-differentiable Points: This category includes all (λ_1, λ_2) where

$$\min\{h_1(\lambda_1, \lambda_2), h_2(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\}$$

can be non-differentiable. This category is the union of all (λ_1, λ_2) for which two of $h_i()$s are equal and are less than the third one. Since we have three functions, we have three sub-categories of non-differentiable points, as described by the following sets:

$$\mathcal{ND}_1 \doteq \Big\{(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), h_1(\lambda_1, \lambda_2) = h_2(\lambda_1, \lambda_2) \leq h_3(\lambda_1, \lambda_2)\Big\}, \tag{2.74}$$

$$\mathcal{ND}_2 \doteq \Big\{(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), h_2(\lambda_1, \lambda_2) = h_3(\lambda_1, \lambda_2) \leq h_1(\lambda_1, \lambda_2)\Big\}, \tag{2.75}$$

$$\mathcal{ND}_3 \doteq \Big\{(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), h_3(\lambda_1, \lambda_2) = h_1(\lambda_1, \lambda_2) \leq h_2(\lambda_1, \lambda_2)\Big\}. \tag{2.76}$$

Note that, if (λ_1, λ_2) belongs to one of the sub-categories of non-differentiable points, it is not necessarily a non-differentiable point of $\min\{h_1(), h_2(), h_3()\}$. However, if (λ_1, λ_2) is a non-differentiable point of $\min\{h_1(), h_2(), h_3()\}$, it necessarily belongs to one of the sub-categories of non-differentiable points.

2.3.5 A Sufficient Condition for Optimal Power Splitting

If $(\lambda_1^\star, \lambda_2^\star)$ is the optimal solution of (2.58), it must belong to one of the three categories of points, listed above. In the following, we investigate each category in detail and find all points of each category that can maximize the sum-rate. By comparing the achievable sum-rate of all these points, we find the optimal solution of (2.58). To demonstrate our proof more clearly, we investigate each category in a separate lemma. We first find a sufficient condition under which the point (λ_1, λ_2) is the optimal solution of (2.58).

Lemma 2.4. *Sufficient condition for optimality: Let $m \in \{1, 2, 3\}$ be a fixed integer. If $(\lambda_1^\star, \lambda_2^\star)$ is the optimal solution of*

$$\max_{\lambda_1, \lambda_2 \in [0,1]} h_m(\lambda_1, \lambda_2), \tag{2.77}$$

and if we have

$$h_m(\lambda_1^\star, \lambda_2^\star) \leq h_j(\lambda_1^\star, \lambda_2^\star), \tag{2.78}$$

for every $j \in \{1, 2, 3\}$, then $(\lambda_1^\star, \lambda_2^\star)$ is the optimal solution of

$$\max_{\lambda_1, \lambda_2 \in [0,1]} \min\Big\{h_1(\lambda_1, \lambda_2), h_2(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\Big\}. \tag{2.79}$$

Proof. Note that, for $(\lambda_1, \lambda_2) = (\lambda_1^\star, \lambda_2^\star)$, according to (2.78), we have

$$\min\left\{h_1(\lambda_1^\star, \lambda_2^\star), h_2(\lambda_1^\star, \lambda_2^\star), h_3(\lambda_1^\star, \lambda_2^\star)\right\} = h_m(\lambda_1^\star, \lambda_2^\star). \tag{2.80}$$

Let us denote the optimal solution of (2.79) by $(\lambda_1^*, \lambda_2^*)$. According to (2.80), we have

$$\min\left\{h_1(\lambda_1^*, \lambda_2^*), h_2(\lambda_1^*, \lambda_2^*), h_3(\lambda_1^*, \lambda_2^*)\right\} \geq h_m(\lambda_1^\star, \lambda_2^\star). \tag{2.81}$$

Note that (2.81) implies that

$$h_m(\lambda_1^*, \lambda_2^*) \geq h_m(\lambda_1^\star, \lambda_2^\star). \tag{2.82}$$

On the other hand, according to (2.77), $(\lambda_1^\star, \lambda_2^\star)$ is the optimal solution of $\max\limits_{\lambda_1,\lambda_2\in[0,1]} h_m(\lambda_1, \lambda_2)$. Therefore, we have

$$h_m(\lambda_1^*, \lambda_2^*) \leq h_m(\lambda_1^\star, \lambda_2^\star). \tag{2.83}$$

Comparing (2.82) with (2.83), we conclude that $h_m(\lambda_1^\star, \lambda_2^\star) = h_m(\lambda_1^*, \lambda_2^*)$. This completes the proof. □

In the following, we use this sufficient condition to characterize the maximum achievable sum-rate for some parts of the weak interference class.

2.3.6 Maximum HK Sum-Rate over Stationary Points

Next, we investigate the first category of points, i.e., stationary points. We show that over the feasible region $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$, the optimization problem (2.58) has no stationary points, as described in the following lemma:

Lemma 2.5. *Stationary points: Over $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$, no stationary points exist, that is, the equation $\nabla(h_i(\lambda_1, \lambda_2)) = 0$, $i \in \{1, 2, 3\}$ has no solutions. Therefore, the optimal solution of (2.58) is either over the boundary points or over the non-differentiable points.*

Proof. Let us start with $\mathcal{S}_1$ and $\mathcal{S}_2$. To find all solutions of $\nabla(h_1(\lambda_1, \lambda_2)) = 0$, we first calculate $\nabla(h_1(\lambda_1, \lambda_2))$ as follows:

$$\begin{aligned}\nabla(h_1(\lambda_1, \lambda_2)) =& \frac{\partial h_1(\lambda_1, \lambda_2)}{\partial \lambda_1}\hat{i} + \frac{\partial h_1(\lambda_1, \lambda_2)}{\partial \lambda_2}\hat{j} \\ =& \frac{-b\lambda_2 P_2 P_1 \hat{i}}{(1 + bP_1\lambda_1)(1 + bP_1\lambda_1 + \lambda_2 P_2)} + \frac{P_2(1 - a - ab\lambda_1 P_1)\hat{j}}{(1 + aP_2\lambda_2)(1 + bP_1\lambda_1 + \lambda_2 P_2)}.\end{aligned} \tag{2.84}$$

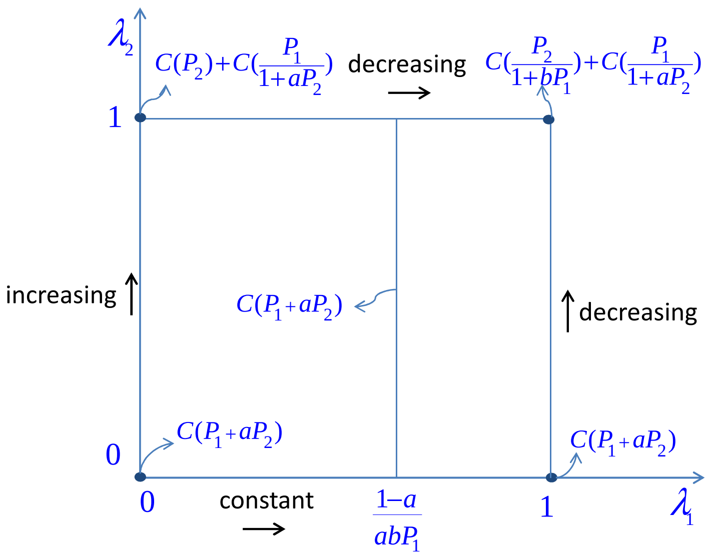

Figure 2.8: The behavior of $h_1(\lambda_1, \lambda_2)$ over the boundary.

Therefore, $\nabla(h_1(\lambda_1, \lambda_2)) = (0, 0)$ has no solutions over $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$.

Similarly, one can calculate $\nabla(h_2(\lambda_1, \lambda_2))$ as follows:

$$\begin{aligned}\nabla(h_2(\lambda_1, \lambda_2)) =& \frac{\partial h_2(\lambda_1, \lambda_2)}{\partial \lambda_1}\hat{i} + \frac{\partial h_2(\lambda_1, \lambda_2)}{\partial \lambda_2}\hat{j} \\ =& \frac{P_1(1-b-ab\lambda_2 P_2)\hat{i}}{(1+bP_1\lambda_1)(1+aP_2\lambda_2+\lambda_1 P_1)} + \frac{-a\lambda_1 P_2 P_1 \hat{j}}{(1+aP_2\lambda_2)(1+aP_2\lambda_2+\lambda_1 P_1)}.\end{aligned} \tag{2.85}$$

One can show that $\nabla(h_2(\lambda_1, \lambda_2)) = (0, 0)$ has no solutions over $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$.

Next, we consider $\mathcal{S}_3$. We first calculate $\nabla(h_3(\lambda_1, \lambda_2))$ as follows:

$$\begin{aligned}\nabla(h_3(\lambda_1, \lambda_2)) &= \frac{\partial h_3(\lambda_1, \lambda_2)}{\partial \lambda_1}\hat{i} + \frac{\partial h_3(\lambda_1, \lambda_2)}{\partial \lambda_2}\hat{j} \\ &= \frac{P_1(1-b-abP_2)\hat{i}}{(1+bP_1\lambda_1)(1+aP_2+P_1\lambda_1)} + \frac{P_2(1-a-abP_1)\hat{j}}{(1+aP_2\lambda_2)(1+bP_1+P_2\lambda_2)}.\end{aligned} \tag{2.86}$$

Clearly, $\nabla(h_3(\lambda_1, \lambda_2)) = 0$ has no solutions in $0 < \lambda_1 < 1$, $0 < \lambda_2 < 1$, and this completes the proof. □

An interesting observation about Lemma 2.5 is the behavior of $h_1(\lambda_1, \lambda_2)$, $h_2(\lambda_1, \lambda_2)$, and $h_3(\lambda_1, \lambda_2)$. According to Lemma 2.5, none of these functions has a stationary point inside the feasible region. Therefore, they all achieve their maximums over the boundary. Figure 2.8 demonstrates the behavior of $h_1(\lambda_1, \lambda_2)$ over the boundary. Note that, according to (2.84), as (λ_1, λ_2) moves from (0,0) to (0,1), the value of $h_1(\lambda_1, \lambda_2)$ increases from $C(P_1 + aP_2)$ to $C(P_2) + C(\frac{P_1}{1+aP_2})$. Moreover, as (λ_1, λ_2) moves from (1,0) to (1,1), the value of $h_1(\lambda_1, \lambda_2)$ decreases from $C(P_1 + aP_2)$ to $C(\frac{P_2}{1+bP_1}) + C(\frac{P_1}{1+aP_2})$. Therefore, $h_1(\lambda_1, \lambda_2)$ achieves its maximum value of $C(P_2) + C(\frac{P_1}{1+aP_2})$ at $(\lambda_1 = 0, \lambda_2 = 1)$ and its minimum value of $C(\frac{P_2}{1+bP_1}) + C(\frac{P_1}{1+aP_2})$ at $(\lambda_1 = 1, \lambda_2 = 1)$. Moreover, according to (2.84), $\nabla h_1(\lambda_1, \lambda_2)$ equals zero in the direction of $\hat{j}$, for $\lambda_1 = \frac{1-a}{abP_1}$. Therefore the function $h_1(\lambda_1, \lambda_2)$ remains constant over the line $\lambda_1 = \frac{1-a}{abP_1}$, as depicted in Figure 2.8.

Similarly, Figure 2.9 demonstrates the behavior of $h_2(\lambda_1, \lambda_2)$ over the boundary. Note that, according to (2.84), as (λ_1, λ_2) moves from (0,0) to (1,0), the value of $h_2(\lambda_1, \lambda_2)$ increases from $C(P_2 + bP_1)$ to $C(P_1) + C(\frac{P_2}{1+bP_1})$. Moreover, as (λ_1, λ_2) moves from (0,1) to (1,1), the value of $h_1(\lambda_1, \lambda_2)$ decreases from $C(P_2 + bP_1)$ to $C(\frac{P_2}{1+bP_1}) + C(\frac{P_1}{1+aP_2})$. Therefore, $h_2(\lambda_1, \lambda_2)$ achieves its maximum value of $C(P_1) + C(\frac{P_2}{1+bP_1})$ at $(\lambda_1 = 1, \lambda_2 = 0)$ and its minimum value of $C(\frac{P_2}{1+bP_1}) + C(\frac{P_1}{1+aP_2})$ at $(\lambda_1 = 1, \lambda_2 = 1)$.

The behavior of $h_3(\lambda_1, \lambda_2)$ can be used to find $R^{\max}_{\text{sum-HK}}$. The sign of $\nabla(h_3(\lambda_1, \lambda_2))$, corresponding to both directions $\hat{i}$ and $\hat{j}$, does not depend on λ_1 or λ_2 and depends only on (a, b, P_1, P_2). Therefore, for each direction, $h_3(\lambda_1, \lambda_2)$ is either strictly increasing or strictly decreasing, as shown in Figure 2.10. Consequently, $h_3(\lambda_1, \lambda_2)$ achieves its maximum at one of the four corner points of the feasible region, namely $(\lambda_1 = 0, \lambda_2 = 0)$, $(\lambda_1 = 0, \lambda_2 = 1)$, $(\lambda_1 = 1, \lambda_2 = 0)$, and $(\lambda_1 = 1, \lambda_2 = 1)$. This property can be used in conjunction with Lemma 2.4 to find $R^{\max}_{\text{sum-HK}}$, as explained in the following remark.

Remark 2.1. *In this remark, we partition the weak interference class into four sub-classes. Using Lemma 2.4, we characterize $R^{\max}_{\text{sum-HK}}$ for three sub-classes. For one sub-class, namely the barely weak interference sub-class, Lemma 2.4 cannot characterize $R^{\max}_{\text{sum-HK}}$.*

A) *If $P_1 \le \frac{1-a}{ab}$ and $P_2 \le \frac{1-b}{ab}$, then $\nabla(h_3(\lambda_1, \lambda_2))$ has positive values in both directions $\hat{i}$ and $\hat{j}$. Therefore, $h_3(\lambda_1, \lambda_2)$ achieves its maximum when $(\lambda_1 = 1, \lambda_2 = 1)$, that is, when the entire interference is treated as noise in both decoders. The maximum value of $h_3(\lambda_1, \lambda_2)$ is*

$$\begin{aligned}\max_{\lambda_1,\lambda_2\in[0,1]} h_3(\lambda_1, \lambda_2) =& h_3(1, 1)\\ =& C\Big(\frac{P_1}{1 + aP_2}\Big) + C\Big(\frac{P_2}{1 + bP_1}\Big),\end{aligned} \tag{2.87}$$

as shown in Figure 2.10A. One can check that $h_3(1, 1) = h_1(1, 1) = h_2(1, 1)$. By

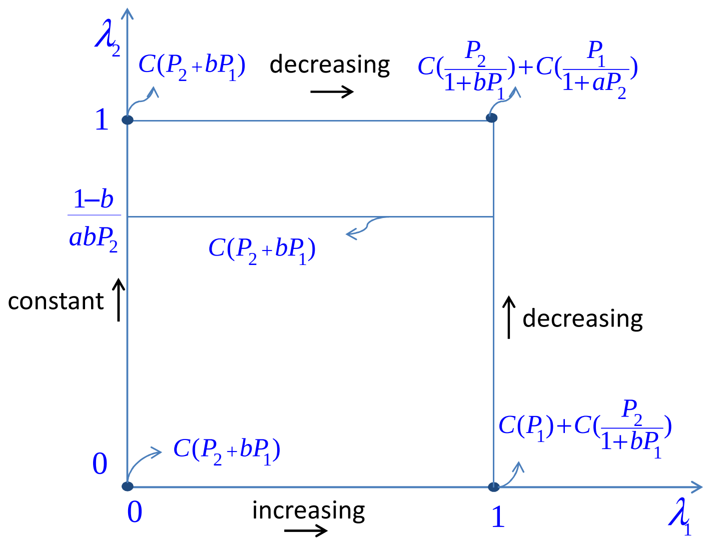

Figure 2.9: The behavior of $h_2(\lambda_1, \lambda_2)$ over the boundary.

Lemma 2.4, this means if $P_1 \leq \frac{1-a}{ab}$ and $P_2 \leq \frac{1-b}{ab}$, treating interference as noise maximizes the achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing. Therefore, we have

$$R^{\max}_{\text{sum-HK}} = C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big). \tag{2.88}$$

B) *If $P_1 \leq \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$, then $\nabla(h_3(\lambda_1, \lambda_2))$ has negative value in the direction of $\hat{i}$ and positive value in the direction of $\hat{j}$. Therefore, $h_3(\lambda_1, \lambda_2)$ achieves its maximum when $\lambda_1 = 0$ and $\lambda_2 = 1$, that is when the entire interference is teated as noise in the first decoder and the entire interference is decoded in the second decoder. As a result, the maximum value of $h_3(\lambda_1, \lambda_2)$ is given by*

$$\max_{\lambda_1,\lambda_2 \in [0,1]} h_3(\lambda_1, \lambda_2) = h_3(0,1) = C\Big(P_2 + bP_1\Big), \tag{2.89}$$

as shown in Figure 2.10B. One can check that $h_3(0,1) \leq h_1(0,1)$ and $h_3(0,1) \leq h_2(0,1)$. By Lemma 2.4, this means, if $P_1 \leq \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$, $(\lambda_1, \lambda_2) = (0,1)$ is the optimal solution of (2.58), and the maximum achievable sum-rate is given by:

$$R^{\max}_{\text{sum-HK}} = C\Big(P_2 + bP_1\Big). \tag{2.90}$$

C) *If $P_1 > \frac{1-a}{ab}$ and $P_2 \leq \frac{1-b}{ab}$, one can show that $(\lambda_1, \lambda_2) = (1,0)$ is the optimal solution of (2.58), and the maximum achievable sum-rate is given by*

$$R^{\max}_{\text{sum-HK}} = C\Big(P_1 + aP_2\Big). \tag{2.91}$$

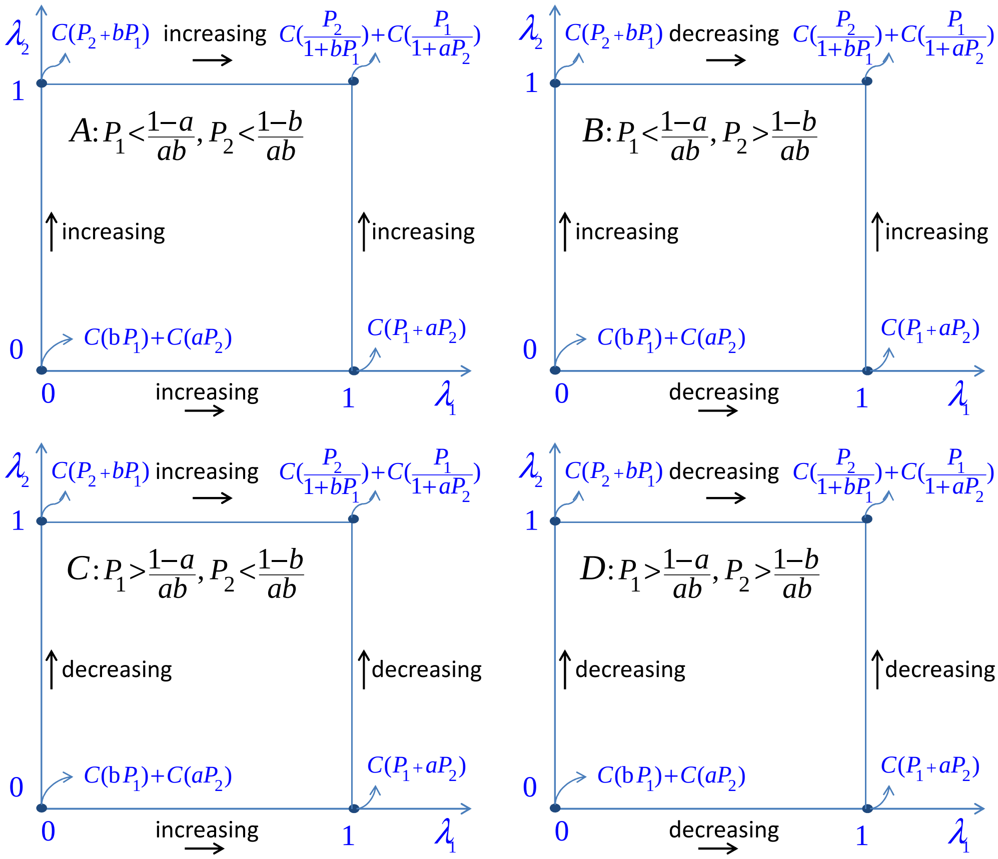

Figure 2.10: The behavior of $h_3(\lambda_1, \lambda_2)$ over the boundary

This is in agreement with the result of [41].

D) *If* $P_1 > \frac{1-a}{ab}$ *and* $P_2 > \frac{1-b}{ab}$*, i.e., for the barely weak interference sub-class,* $\nabla(h_3(\lambda_1, \lambda_2))$ *has negative values in both directions* $\hat{\imath}$ *and* $\hat{\jmath}$*. Therefore,* $h_3(\lambda_1, \lambda_2)$ *achieves its maximum when* $(\lambda_1 = 0, \lambda_2 = 0)$*, that is, when the entire interference is decoded at both decoders. The maximum value of* $h_3(\lambda_1, \lambda_2)$ *is*

$$\max_{\lambda_1,\lambda_2\in[0,1]} h_3(\lambda_1, \lambda_2) = h_3(0,0) = C(aP_2) + C(bP_1), \tag{2.92}$$

as shown in Figure 2.10D. However, we cannot use Lemma 2.4, because the following inequalities are not satisfied:

$$\begin{aligned} h_3(0,0) &\le h_1(0,0), \\ h_3(0,0) &\le h_2(0,0). \end{aligned} \tag{2.93}$$

For the barely weak interference sub-class, we have

$$\begin{aligned} h_3(0,0) &= C(aP_2) + C(bP_1) \ge h_1(0,0) = C(P_1 + aP_2), \\ h_3(0,0) &= C(aP_2) + C(bP_1) \ge h_2(0,0) = C(P_2 + bP_1), \end{aligned} \tag{2.94}$$

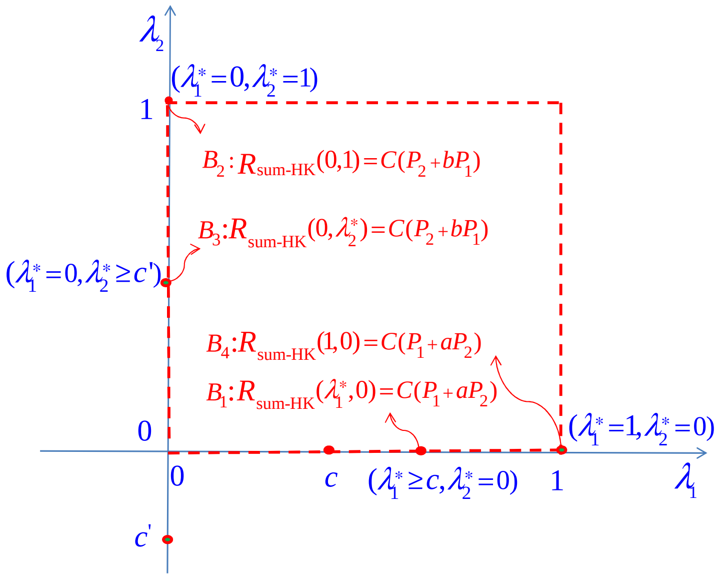

Figure 2.11: Four sub-categories of the boundary points: the optimal point and the maximum sum-rate corresponding to each sub-category.

and consequently, (2.92) is not the maximum achievable sum-rate. In fact, we will later show that for the weak interference class, $(\lambda_1, \lambda_2) = (0, 0)$ *is never the optimal solution of (2.58), i.e., SND does not achieve* $R^{\max}_{\text{sum-HK}}$*, as will be explained in Corollary 2.1. Note that, for the barely weak interference sub-class, the maximum achievable sum-rate has been unknown. In the rest of our analysis, we focus on the barely weak interference sub-class, that is, we assume that* $P_1 > \frac{1-a}{ab}$ *and* $P_2 > \frac{1-b}{ab}$.

2.3.7 Maximum HK Sum-Rate over Boundary Points

Now that we have investigated the behavior of $h_1(\lambda_1, \lambda_2)$, $h_2(\lambda_1, \lambda_2)$, and $h_3(\lambda_1, \lambda_2)$ over the boundary, we investigate the behavior of

$$\min\{h_1(\lambda_1\lambda_2), h_2(\lambda_1, \lambda_2), h_3(\lambda_1, \lambda_2)\}$$

and find all local maximum points over the boundary.

Lemma 2.6. *Boundary points: For the boundary points, when* $P_1 > \frac{1-a}{ab}$ *and* $P_2 > \frac{1-b}{ab}$*, define* $c \doteq \frac{P_1(1-b)-P_2(1-a)}{P_1\left(1-b+P_2(1-ab)\right)}$ *and* $c' \doteq \frac{P_2(1-a)-P_1(1-b)}{P_2\left(1-a+P_1(1-ab)\right)}$*, then we have*

2.6-A: For the sub-category of boundary points $\mathcal{B}_1$*, i.e.,* $\lambda_2 = 0$*, the optimal* λ_1 *is not unique. In fact, any* $\lambda_1^\star \in \left[[c]^+, 1\right]$ *is an optimal solution, and the corresponding maximum sum-rate is given by* $C(P_1 + aP_2)$*, as shown in Figure 2.11.*

2.6-B: For the sub-category of boundary points $\mathcal{B}_2$, i.e., $\lambda_2 = 1$, $\lambda_1^\star = 0$ is the unique optimal solution, and the corresponding maximum sum-rate is given by $C(P_2 + bP_1)$, as shown in Figure 2.11.

2.6-C: For the sub-category of boundary points $\mathcal{B}_3$, i.e., $\lambda_1 = 0$, the optimal λ_2 is not unique. In fact, any $\lambda_2^\star \in \Big[[c']^+, 1\Big]$ is an optimal solution, and the corresponding maximum sum-rate is given by $C(P_2 + bP_1)$, as shown in Figure 2.11.

2.6-D: For the sub-category of boundary points $\mathcal{B}_4$, i.e., $\lambda_1 = 1$, $\lambda_2^\star = 0$ is the unique optimal solution, and the corresponding maximum sum-rate is given by $C(P_1 + aP_2)$, as shown in Figure 2.11.

Proof. 2.6-A: When $\lambda_2 = 0$ and $0 \le \lambda_1 \le 1$, the optimization problem (2.58) reduces to

$$\begin{aligned}\max_{\lambda_1,\lambda_2\in[0,1]} R_{\text{sum-HK}}(\lambda_1,\lambda_2) &= \max_{0\le\lambda_1\le1} \min\Big\{h_1(\lambda_1,0), h_2(\lambda_1,0), h_3(\lambda_1,0)\Big\}\\ &\overset{(a)}{=} \max_{0\le\lambda_1\le1} \min\Big\{h_1(\lambda_1,0), h_2(\lambda_1,0)\Big\},\end{aligned} \tag{2.95}$$

where (a) is valid because, for $\lambda_2 = 0$ and $P_2 > \frac{1-b}{ab}$, according to Lemma 2.3, we have $h_1(\lambda_1, 0) < h_3(\lambda_1, 0)$. To solve the optimization problem (2.95), we first characterize $\min\Big\{h_1(\lambda_1,0), h_2(\lambda_1,0)\Big\}$ as follows:

Note that $h_1(\lambda_1, 0) = C(\lambda_1 P_1) + C\Big(\frac{\bar{\lambda}_1 P_1 + aP_2}{1+\lambda_1 P_1}\Big) = C(P_1 + aP_2)$. Therefore, $h_1(\lambda_1, 0)$ is a constant function for all values of λ_1. On the other hand, $h_2(\lambda_1, 0) = C(\lambda_1 P_1) + C\Big(\frac{b\bar{\lambda}_1 P_1 + P_2}{1+b\lambda_1 P_1}\Big)$. Therefore, we have $\frac{\partial h_1(\lambda_1,0)}{\partial \lambda_1} = \frac{P_1}{1+P_1\lambda_1} - \frac{P_1 b}{1+P_1 b\lambda_1} \ge 0$. This implies that $h_2(\lambda_1, 0)$ is an increasing function over $\lambda_1 \in [0, 1]$. Finally, according to Lemma 2.3, $h_1(\lambda_1, 0) \le h_2(\lambda_1, 0)$ if and only if $\lambda_1 \ge c = \frac{P_1(1-b)-P_2(1-a)}{P_1\big(1-b+P_2(1-ab)\big)}$. Consequently, we have

$$\min\Big\{h_1(\lambda_1,0), h_2(\lambda_1,0)\Big\} = \begin{cases} h_2(\lambda_1,0) & \text{if } \lambda_1 < c\\ h_1(\lambda_1,0) & \text{if } \lambda_1 \ge c.\end{cases} \tag{2.96}$$

Moreover, since $h_2(\lambda_1, 0)$ is an increasing function, we conclude that

$$\begin{aligned}\max_{0\le\lambda_1\le1} \min\Big\{h_1(\lambda_1,0), h_2(\lambda_1,0)\Big\} =& h_1(\lambda_1^\star, 0)\\ =& C(P_1 + aP_2),\end{aligned} \tag{2.97}$$

and any $\lambda_1^\star \ge \max\{c, 0\}$ is an optimal solution. This completes the proof of 2.6-A of Lemma 2.6. Figure 2.11 shows that any $\lambda_1^\star$ that is greater than c achieves the maximum sum-rate over the boundary sub-category $\mathcal{B}_1$.

2.6-B: When $\lambda_2 = 1$ and $0 \leq \lambda_1 \leq 1$, the optimization problem (2.58) reduces to

$$\begin{aligned}\max_{\lambda_1,\lambda_2\in[0,1]} R_{\text{sum-HK}}(\lambda_1,\lambda_2) &= \max_{0\leq\lambda_1\leq1} \min\Big\{h_1(\lambda_1,1), h_2(\lambda_1,1), h_3(\lambda_1,1)\Big\} \\ &\overset{(a)}{=} \max_{0\leq\lambda_1\leq1} h_2(\lambda_1,1), \qquad (2.98)\end{aligned}$$

where (a) is valid, because by Lemma 2.3, for $\lambda_2 = 1$, we have $h_2(\lambda_1,1) = h_3(\lambda_1,1) < h_1(\lambda_1,1)$. Moreover, according to (2.56), $h_2(\lambda_1,1) = C\left(\frac{P_2+b\bar{\lambda}_1P_1}{1+b\lambda_1P_1}\right)+C\left(\frac{\lambda_1P_1}{1+aP_2}\right)$. Therefore, $\frac{\partial h_2(\lambda_1,1)}{\partial\lambda_1} = \frac{P_1}{1+P_1\lambda_1+aP_2} - \frac{P_1b}{1+P_1b\lambda_1} = \frac{P_1(1-b-abP_2)}{(1+P_1\lambda_1+aP_2)(1+P_1b\lambda_1)}$. Since $P_2 > \frac{1-b}{ab}$, we see that $\frac{\partial h_2(\lambda_1,1)}{\partial\lambda_1}$ is strictly negative over $[0,1]$. Therefore, $h_2(\lambda_1,1)$ achieves its maximum when $\lambda_1 = 0$. The maximum of (2.98) is $C(P_2 + bP_1)$. This completes the proof of 2.6-B of Lemma 2.6. Figure 2.11 shows that $(\lambda_1,\lambda_2) = (0,1)$ achieves the maximum sum-rate, over the boundary sub-category $\mathcal{B}_2$.

Note that the proof of 2.6-C and 2.6-D follows by exchanging the indices 1, 2, as well as exchanging the cross-link gains a and b, in the proof of 2.6-A and 2.6-B, respectively. Figure 2.11 summarizes all parts of this lemma. It demonstrates the optimal point over each sub-category of the boundary points. This completes the proof of Lemma 2.6. □

Lemma 2.6 completely characterizes the sum-rate corresponding to the boundary of the feasible region. The constants c and c' determine the optimal points over the boundary. Note that if c is positive, then c' is negative, and therefore, c' does not restrict the optimal points over the boundary. Similarly, if c' is positive, then c is negative, and therefore, c does not restrict the optimal points over the boundary. Figure 2.12 shows the achievable sum-rate over the boundary, when c is positive. Note that for $(\lambda_1 = 0, \lambda_2 = 0)$, the achievable sum-rate is given by $\min\{C(P_1 + aP_2), C(P_2 + bP_1)\} = C(P_2 + bP_1)$. If λ_1 remains zero, but λ_2 starts to increase, the achievable sum-rate remains constant. However, if λ_2 remains zero, but λ_1 starts to increase, the achievable sum-rate increases, until $\lambda_1 = c$. At this point, the achievable sum-rate is given by $\min\{C(P_1+aP_2), C(P_2+bP_1)\} = C(P_1 + aP_2)$. If λ_1 increases further, the achievable sum-rate remains constant, until (λ_1,λ_2) reaches the point $(\lambda_1 = 1, \lambda_2 = 0)$. If (λ_1,λ_2) moves from $(0,1)$ to (1,1), then the achievable sum-rate decreases to $C(\frac{P_1}{1+aP_2})+C(\frac{P_2}{1+bP_1})$. Note that, for the barely weak interference sub-class, we have $C(\frac{P_1}{1+aP_2}) + C(\frac{P_2}{1+bP_1}) \leq \min\{C(P_1 + aP_2), C(P_2 + bP_1)\}$. This means, the sum-rate achieved by treating interference as noise is less than the sum-rate achieved by SND. Moreover, the sum-rate achieved by SND is less than the sum-rate achieved by $(\lambda_1 \geq c, \lambda_2 = 0)$.

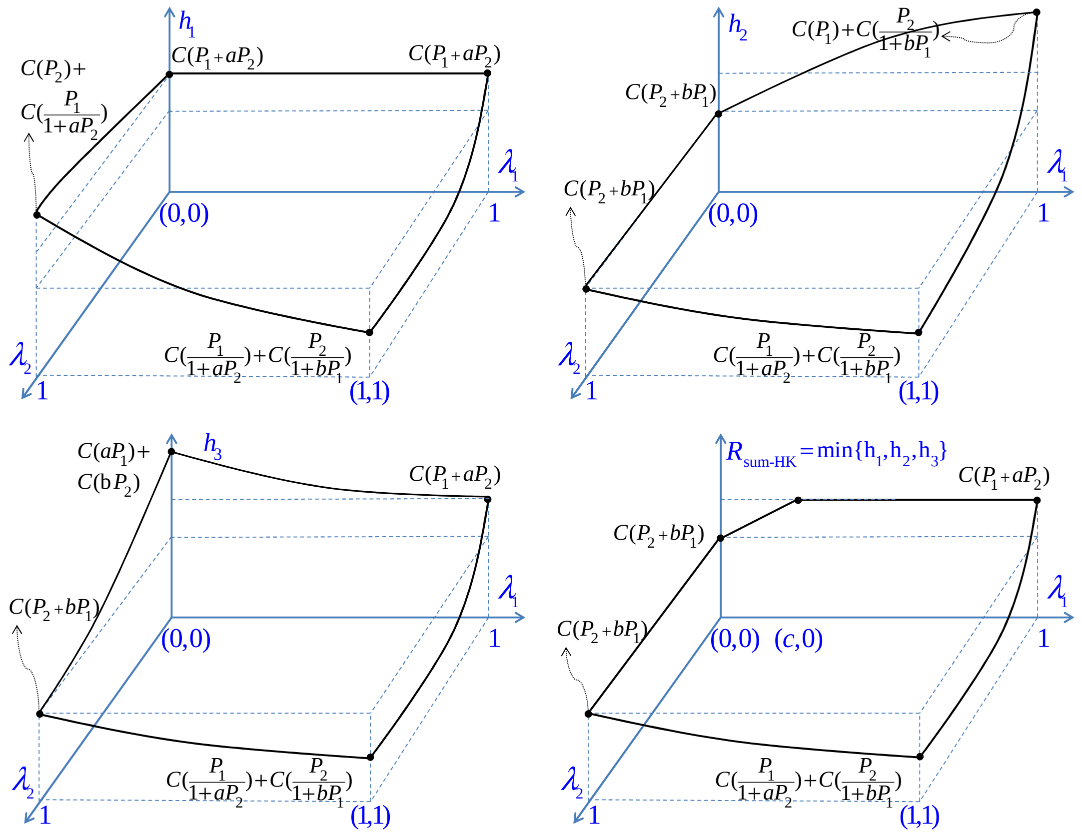

Figure 2.12: The achievable sum-rate of the HK scheme over the boundary of the feasible region, for the barely weak interference sub-class with $c \geq 0$.

Note that the sum-capacity of the two-user GIC is known for some sub-classes, as shown in Figure 2.2. For all such sub-classes, the sum-capacity is is equal to $R^{\max}_{\text{sum-HK}}$. Moreover, the optimal (λ_1, λ_2) belongs to one corner point of the feasible region. For the sarong interference class, $(\lambda_1 = 0, \lambda_2 = 0)$ leads to $R^{\max}_{\text{sum-HK}}$. For the mixed I interference class, $(\lambda_1 = 1, \lambda_2 = 0)$ leads to $R^{\max}_{\text{sum-HK}}$, and for the mixed II interference class, $(\lambda_1 = 0, \lambda_2 = 1)$ leads to $R^{\max}_{\text{sum-HK}}$. Finally, for the very weak interference sub-class, $(\lambda_1 = 1, \lambda_2 = 1)$ leads to $R^{\max}_{\text{sum-HK}}$. For the weak interference class, the following corollary compares the achievable sum-rates corresponding to the four corner points of the feasible region.

Corollary 2.1. *For the two-user GIC with weak interference, the HK scheme can achieve the following sum-rate:*

$$\begin{aligned} R^{\text{bnd}}_{\text{sum}} \doteq \max \Big\{ & C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big), \\ & C(P_1 + aP_2), C(P_2 + bP_1) \Big\}. \end{aligned} \tag{2.99}$$

Table 2.3 shows the achievable sum-rate corresponding to the four corner points of the

(λ_1,λ_2)	$h_1(\lambda_1,\lambda_2)$	$h_2(\lambda_1,\lambda_2)$	$h_3(\lambda_1,\lambda_2)$	$R_{\text{sum-HK}}(\lambda_1,\lambda_2) =$ $\min\{h_1(), h_2(), h_3()\}$
$(0,0)$	$C(P_1+aP_2)$	$C(P_2+bP_1)$	$C(aP_2)+$ $C(bP_1)$	$R_{\text{sum-SND}}$
$(0,1)$	$C(\frac{P_1}{1+aP_2})+$ $C(P_2)$	$C(P_2+bP_1)$	$C(P_2+bP_1)$	$C(P_2+bP_1)$
$(1,0)$	$C(P_1+aP_2)$	$C(\frac{P_2}{1+bP_1})+$ $C(P_1)$	$C(P_1+aP_2)$	$C(P_1+aP_2)$
$(1,1)$	$C(\frac{P_1}{1+aP_2})+$ $C(\frac{P_2}{1+bP_1})$	$C(\frac{P_1}{1+aP_2})+$ $C(\frac{P_2}{1+bP_1})$	$C(\frac{P_1}{1+aP_2})+$ $C(\frac{P_2}{1+bP_1})$	$C(\frac{P_1}{1+aP_2})+$ $C(\frac{P_2}{1+bP_1})$

Table 2.3: The achievable sum-rate corresponding to four corner points of the boundary.

feasible region. Note that the sum-rate corresponding to $(\lambda_1 = 0, \lambda_2 = 0)$ is the sum-rate achieved by SND, denoted by $R_{\text{sum-SND}}$. For the weak interference class, $R_{\text{sum-SND}}$ is given by

$$R_{\text{sum-SND}} = \min\Big\{C(P_1 + aP_2), C(P_2 + bP_1), C(aP_2) + C(bP_1)\Big\}. \tag{2.100}$$

For the weak interference class, this sum-rate is smaller than the sum-rate corresponding to $(\lambda_1 = 1, \lambda_2 = 0)$ or $(\lambda_1 = 0, \lambda_2 = 1)$, as shown in Table 2.3. Therefore, although SND achieves the sum-capacity for every (a, b, P_1, P_2) that belongs to the strong interference class, SND achieves the sum-capacity for no (a, b, P_1, P_2) that belongs to the weak interference class. Consequently, the sum-rate (2.99) is achieved by just considering the three corner points of the boundary of the feasible region, namely $(\lambda_1 = 0, \lambda_2 = 0)$, $(\lambda_1 = 0, \lambda_2 = 1)$, and $(\lambda_1 = 1, \lambda_2 = 0)$. In fact, when $P_1 \leq \frac{1-a}{ab}$ or $P_2 \leq \frac{1-b}{ab}$, Remark 2.1 shows that the maximum sum-rate of HK scheme is given by (2.99). However, for the barely weak interference sub-class, i.e., $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$, the maximum sum-rate of HK is not known.

Figure 2.13 shows quadrant I of the P_1P_2-plane. This quadrant is divided into three

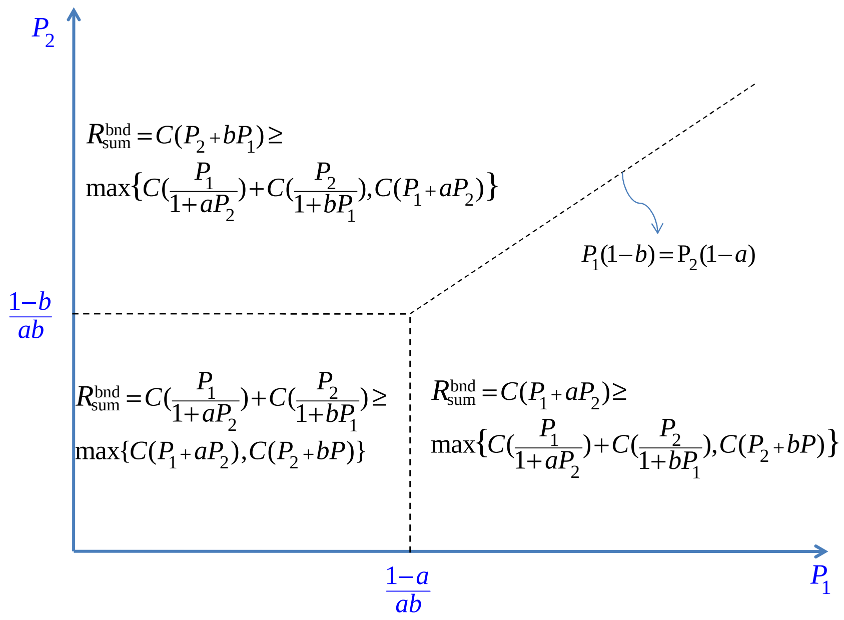

Figure 2.13: The sum-rate of the HK scheme achieved by investigating only the boundary points: Quadrant I of the P_1P_2-plane, is partitioned into three regions. In each region, exactly one of the $C\left(\frac{P_1}{1+aP_2}\right) + C\left(\frac{P_2}{1+bP_1}\right), C(P_1 + aP_2), C(P_2 + bP_1)$ is the achievable sum-rate.

regions. In each region, exactly one of the $C\left(\frac{P_1}{1+aP_2}\right)+C\left(\frac{P_2}{1+bP_1}\right), C(P_1+aP_2), C(P_2+bP_1)$ is greater than the other two, as shown in the figure. Note that the line $P_1(1-b) = P_2(1-a)$ separates two regions: the region in which $C(P_1 + aP_2)$ is the maximum of the three and the region in which $C(P_2 + bP_1)$ is the maximum of the three. Lemma 2.5 and 2.6 studied all stationary points and all boundary points, respectively. Figure 2.13 demonstrates the summary of these lemmas. To solve the optimization problem (2.58), all that is left is to investigate the last category of points, i.e., the non-differentiable points.

2.3.8 Maximum HK Sum-Rate over Non-Differentiable Points

As highlighted in (2.74-2.76), there exist three sub-categories of non-differentiable points, namely $\mathcal{ND}_1$, $\mathcal{ND}_2$, and $\mathcal{ND}_3$. We characterize each sub-category inside the $\lambda_1\lambda_2$-plane. For sub-category $\mathcal{ND}_1$, we have $h_1(\lambda_1, \lambda_2) = h_2(\lambda_1, \lambda_2) \leq h_3(\lambda_1, \lambda_2)$. According to

Lemma 2.3, for $\lambda_1, \lambda_2 \in (0,1)$, we have

$$h_1(\lambda_1,\lambda_2) = h_2(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 = m\lambda_2 + c, \tag{2.101}$$

$$h_1(\lambda_1,\lambda_2) \leq h_3(\lambda_1,\lambda_2) \Leftrightarrow \lambda_2 \leq ab - \frac{1-b}{P_2}, \tag{2.102}$$

$$h_2(\lambda_1,\lambda_2) \leq h_3(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 \leq ab - \frac{1-a}{P_1}. \tag{2.103}$$

Therefore, the subcategory $\mathcal{ND}_1$ can be expressed by

$$\mathcal{ND}_1 = \Big\{(\lambda_1,\lambda_2) : \lambda_1,\lambda_2 \in (0,1), \lambda_1 = m\lambda_2 + c, \\ 0 < \lambda_1 \leq ab - \frac{1-a}{P_1}, 0 < \lambda_2 \leq ab - \frac{1-b}{P_2}\Big\}. \tag{2.104}$$

All points that belong to the sub-category $\mathcal{ND}_1$ lie on the line $\lambda_1 = m\lambda_2 + c$, and are shown by the blue solid line in Figure 2.14. In fact, $\mathcal{ND}_1$ is a line segment that has two end points. One end point is given by $(\lambda_1 = ab - \frac{1-a}{P_1}, \lambda_2 = ab - \frac{1-b}{P_2})$, as shown in Figure 2.14. Depending on the value of c, the other end point can have two cases. If $c \geq 0$, the other endpoint is given by $(\lambda_1 = c, \lambda_2 = 0)$, as shown Figure 2.14. However, if $c < 0$, the other endpoint is given by $(\lambda_1 = 0, \lambda_2 = c')$, as shown Figure 2.15.

For the sub-category $\mathcal{ND}_2$, we have $h_2(\lambda_1,\lambda_2) = h_3(\lambda_1,\lambda_2) \leq h_1(\lambda_1,\lambda_2)$. According to Lemma 2.3, for $\lambda_1, \lambda_2 \in (0,1)$, we have

$$h_2(\lambda_1,\lambda_2) = h_3(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 = ab - \frac{1-a}{P_1}, \tag{2.105}$$

$$h_2(\lambda_1,\lambda_2) \leq h_1(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 \leq m\lambda_2 + c, \tag{2.106}$$

$$h_3(\lambda_1,\lambda_2) \leq h_1(\lambda_1,\lambda_2) \Leftrightarrow \lambda_2 \geq ab - \frac{1-b}{P_2}. \tag{2.107}$$

Therefore, the subcategory $\mathcal{ND}_2$ can be expressed by

$$\mathcal{ND}_2 = \Big\{(\lambda_1,\lambda_2) : \lambda_1,\lambda_2 \in (0,1), \lambda_1 = ab - \frac{1-a}{P_1}, \lambda_2 \geq ab - \frac{1-b}{P_2}\Big\}. \tag{2.108}$$

Consequently, all points that belong to the sub-category $\mathcal{ND}_2$ lie on the vertical line $\lambda_1 = ab - \frac{1-a}{P_1}$, as shown by the blue dashed line in Figure 2.14.

Finally, for the sub-category $\mathcal{ND}_3$, we have $h_3(\lambda_1,\lambda_2) = h_1(\lambda_1,\lambda_2) \leq h_2(\lambda_1,\lambda_2)$. According to Lemma 2.3, for $\lambda_1, \lambda_2 \in (0,1)$, we have

$$h_3(\lambda_1,\lambda_2) = h_1(\lambda_1,\lambda_2) \Leftrightarrow \lambda_2 = ab - \frac{1-b}{P_2}, \tag{2.109}$$

$$h_3(\lambda_1,\lambda_2) \leq h_2(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 \geq ab - \frac{1-a}{P_1}, \tag{2.110}$$

$$h_1(\lambda_1,\lambda_2) \leq h_2(\lambda_1,\lambda_2) \Leftrightarrow \lambda_1 \geq m\lambda_2 + c. \tag{2.111}$$

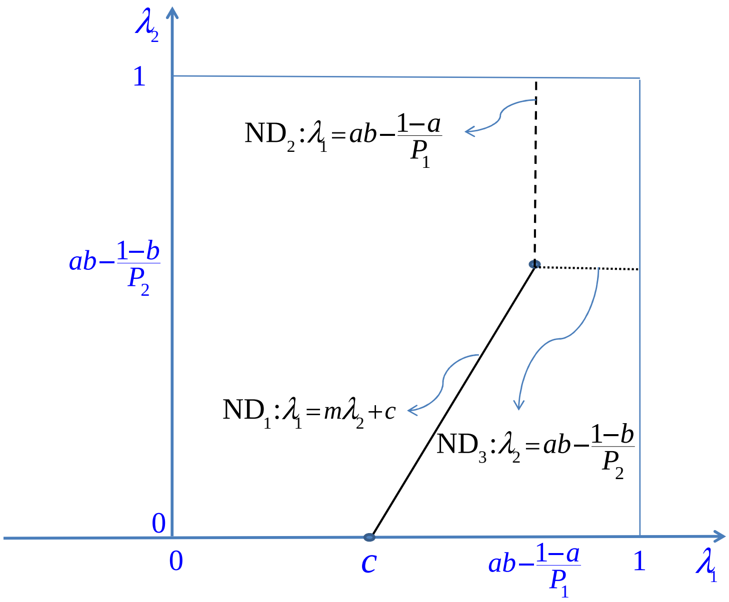

Figure 2.14: Three sub-categories of non-differentiable points in the $\lambda_1\lambda_2$-plane, when $c \geq 0$.

Therefore, the subcategory $\mathcal{ND}_3$ can be expressed by

$$\mathcal{ND}_3 = \Big\{(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \in (0,1), \lambda_2 = ab - \frac{1-b}{P_2}, \lambda_1 \geq ab - \frac{1-a}{P_1}\Big\}. \tag{2.112}$$

Consequently, all points of the sub-category $\mathcal{ND}_3$ lie on the horizontal line $\lambda_2 = ab - \frac{1-b}{P_2}$ and are shown by the blue dotted line in Figure 2.14.

Lemma 2.5 shows that there exists no stationary points. Corollary 2.1 shows that by investigating the boundary points, the maximum achievable sum-rate is given by

$$R_{\text{sum}}^{\text{bnd}} = \max\Big\{C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big), C(P_1 + aP_2), C(P_2 + bP_1)\Big\}.$$

We now investigate the three sub-categories of non-differentiable points to see, if we can achieve a sum-rate greater than the sum-rate corresponding to the boundary points. The following lemma describes the result.

Lemma 2.7. *Non-differentiable points: Over the non-differentiable points, when $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$, we have*

2.7-A: For the non-differentiable sub-category $\mathcal{ND}_1$, the optimal solution of (2.58), is given by

$$(\lambda_1^\star, \lambda_2^\star) \in \{(c,0), (0,c'), (\tilde{\lambda}_1, \tilde{\lambda}_2), (\hat{\lambda}_1, \hat{\lambda}_2)\}, \tag{2.113}$$

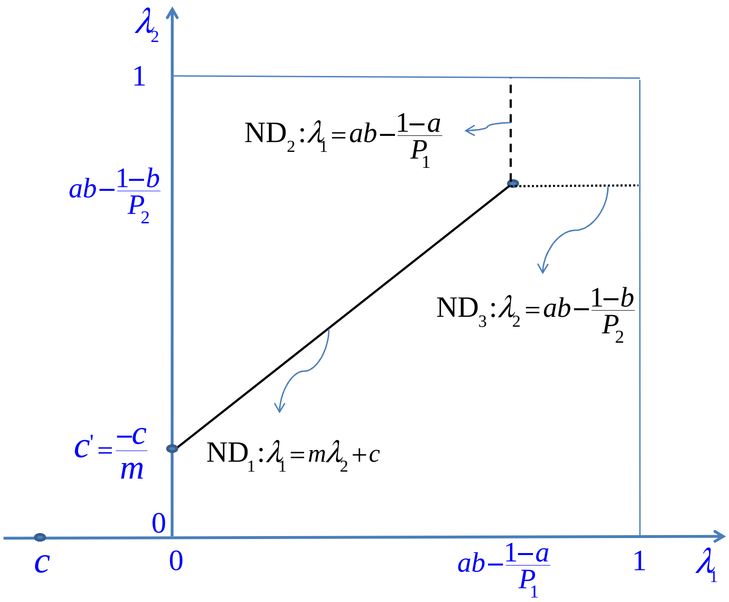

Figure 2.15: Three sub-categories of non-differentiable points in the $\lambda_1\lambda_2$-plane, when $c < 0$.

where $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ *is given by*

$$\tilde{\lambda}_1 \doteq ab - \frac{1-a}{P_1},$$
$$\tilde{\lambda}_2 \doteq ab - \frac{1-b}{P_2}. \tag{2.114}$$

Moreover, $\hat{\lambda}_1 \doteq m\hat{\lambda}_2 + c$*, where* $m = \frac{P_2\big((1-a)+P_1(1-ab)\big)}{P_1\big(1-b+P_2(1-ab)\big)}$, $c = \frac{P_1(1-b)-P_2(1-a)}{P_1\big(1-b+P_2(1-ab)\big)}$*, and* $\hat{\lambda}_2$ *is the non-negative solution of the following second order equation:*

$$(\lambda_2^2) + 2\frac{(1+bP_1c)}{(bP_1m+P_2)}(\lambda_2) + \frac{(1+bP_1c)(abP_1c+a-1)}{abP_1m(bP_1m+P_2)} = 0. \tag{2.115}$$

The maximum achievable sum-rate corresponding to the this sub-category is given by

$$\max\{h_1(c,0), h_1(0,c'), h_1(\tilde{\lambda}_1,\tilde{\lambda}_2), h_1(\hat{\lambda}_1,\hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0)(\hat{\lambda}_2 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq \hat{\lambda}_2)\} =$$
$$\max\Big\{C(P_1+aP_2), C(P_2+bP_1),$$
$$C(P_1+aP_2) + g_1(\tilde{\lambda}_1,\tilde{\lambda}_2),$$
$$C(P_1+aP_2) + g_1(\hat{\lambda}_1,\hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0)(\hat{\lambda}_2 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq \hat{\lambda}_2)\Big\}, \tag{2.116}$$

where the function $g_1(\lambda_1, \lambda_2)$ *is defined by*

$$g_1(\lambda_1,\lambda_2) \doteq C(\frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1+a\lambda_2 P_2}) - C(b\lambda_1 P_1). \tag{2.117}$$

Moreover, $(\hat{\lambda}_1, \hat{\lambda}_2)$ is an acceptable power splitting, i.e., $\hat{\lambda}_1, \hat{\lambda}_2 \in [0,1]$, that belongs to $\mathcal{ND}_1$ if and only if

$$\frac{(1-b)ab}{1-a}P_1 + b - 1 \leq P_2, \tag{2.118}$$

$$\frac{(1-a)ab}{1-b}P_2 + a - 1 \leq P_1, \tag{2.119}$$

$$\hat{\lambda}_2 \leq ab - \frac{1-b}{P_2}. \tag{2.120}$$

2.7-B: For the non-differentiable sub-category $\mathcal{ND}_3$, the optimal solution of (2.58) is given by

$$\begin{aligned}\lambda_1^\star &= \tilde{\lambda}_1 \doteq ab - \frac{1-a}{P_1},\\ \lambda_2^\star &= \tilde{\lambda}_2 \doteq ab - \frac{1-b}{P_2},\end{aligned} \tag{2.121}$$

and the corresponding achievable sum-rate is given by

$$R_{\text{sum-HK}}(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}) = h_1(ab\frac{1-a}{P_1}, ab - \frac{1-b}{P_2}). \tag{2.122}$$

Moreover, $(\tilde{\lambda}_1, \tilde{\lambda}_2) = (ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2})$ is an acceptable power splitting, i.e., $\tilde{\lambda}_1, \tilde{\lambda}_2 \in [0,1]$ if and only if

$$P_1 \geq \frac{1-a}{ab}, \tag{2.123}$$

$$P_2 \geq \frac{1-b}{ab}. \tag{2.124}$$

2.7-C: For the non-differentiable sub-category $\mathcal{ND}_2$, the optimal solution of (2.58) is the same as 2.7-B.

Proof. 2.7-A: When $h_1(\lambda_1, \lambda_2) = h_2(\lambda_1, \lambda_2) \leq h_3(\lambda_1, \lambda_2)$, the optimization problem (2.58) reduces to

$$\begin{aligned} &\max_{\lambda_1,\lambda_2\in[0,1]} && R_{\text{sum-HK}}(\lambda_1, \lambda_2) = \\ &\max_{\lambda_1,\lambda_2\in[0,1]} && h_1(\lambda_1, \lambda_2) \\ &\text{subject to} && h_1(\lambda_1, \lambda_2) \leq h_3(\lambda_1, \lambda_2). \end{aligned} \tag{2.125}$$

Since $h_1(\lambda_1, \lambda_2) = h_2(\lambda_1, \lambda_2)$, by Lemma 2.3, we know that the optimal λ_1 and λ_2 are linearly dependent, and we have

$$\lambda_1 = m\lambda_2 + c. \tag{2.126}$$

where

$$m \doteq \frac{P_2\big((1-a)+P_1(1-ab)\big)}{P_1\big(1-b+P_2(1-ab)\big)}, \tag{2.127}$$

$$c \doteq \frac{P_1(1-b)-P_2(1-a)}{P_1\big(1-b+P_2(1-ab)\big)}. \tag{2.128}$$

Therefore, the optimization problem (2.125) reduces to

$$\begin{aligned} \max_{0\le\lambda_2\le1} \quad & h_1(m\lambda_2+c,\lambda_2) \\ \text{subject to} \quad & h_1(m\lambda_2+c,\lambda_2) \le h_3(m\lambda_2+c,\lambda_2). \end{aligned} \tag{2.129}$$

To solve the optimization problem (2.129), note that the feasible region is a line segment, as shown in Figure 2.14. Therefore, the optimal point is either a stationary point on this line segment or one of the two ends of this line segment. One of the end points is $(\tilde{\lambda}_1, \tilde{\lambda}_2)$. This point achieves the sum-rate of $h_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$. The other end point can have two cases, depending on the value of c. If c is positive, the other end point is given by $(c, 0)$, as shown in Figure 2.14. This point achieves the sum-rate of $h_1(c, 0)$. Note that according to (2.128), we have

$$\begin{aligned} c =& \frac{P_1(1-b)-P_2(1-a)}{P_1\big(1-b+P_2(1-ab)\big)} \\ \le& \frac{P_1(1-b)}{P_1\big(1-b+P_2(1-ab)\big)} \\ \le& 1. \end{aligned} \tag{2.130}$$

However, if c is negative, the other end point is given by $(0, c' = \frac{-c}{m})$, as shown in Figure 2.15. This point achieves the sum-rate of $h_1(0, c')$. Note that According to (2.128) and (2.127), we have

$$\begin{aligned} c' = \frac{-c}{m} =& \frac{P_2(1-a)-P_1(1-b)}{P_2\big(1-a+P_1(1-ab)\big)} \\ \le& \frac{P_2(1-a)}{P_2\big(1-a+P_1(1-ab)\big)} \\ \le& 1. \end{aligned} \tag{2.131}$$

Let us denote the stationary point that belongs to $\mathcal{ND}_1$ by $(\hat{\lambda}_1, \hat{\lambda}_2)$. Therefore, if $c \ge 0$, the maximum achievable sum-rate corresponding to $\mathcal{ND}_1$ is given by

$$\max\{h_1(c,0), h_1(\tilde{\lambda}_1, \tilde{\lambda}_2), h_1(\hat{\lambda}_1, \hat{\lambda}_2)\}, \tag{2.132}$$

and if $c < 0$, it is given by

$$\max\{h_1(0,c'), h_1(\tilde{\lambda}_1, \tilde{\lambda}_2), h_1(\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{2.133}$$

Therefore, for $\mathcal{ND}_1$, the optimal solution of (2.58) is given by

$$(\lambda_1^\star, \lambda_2^\star) \in \{(c,0), (0,c'), (\tilde{\lambda}_1, \tilde{\lambda}_2), (\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{2.134}$$

Note that, since $h_1(c,0) = C(P_1 + aP_2)$ and $h_1(0,c') = C(P_2 + bP_2)$, we have

$$h_1(c,0) \geq h_1(0,c') \Leftrightarrow P_1(1-a) \geq P_2(1-b) \Leftrightarrow c \geq 0. \tag{2.135}$$

Consequently, the maximum achievable sum-rate corresponding to $\mathcal{ND}_1$ is given by

$$\max\{h_1(c,0), h_1(0,c'), h_1(\tilde{\lambda}_1, \tilde{\lambda}_2), h_1(\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{2.136}$$

Note that $(c,0)$, $(0,c')$, and $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ necessarily belong to $\mathcal{ND}_1$. However, $h_1(m\lambda_2 + c, \lambda_2)$ may not have any stationary points that belongs to $\mathcal{ND}_1$. In the following, we prove that $h_1(m\lambda_2 + c, \lambda_2)$ can have at most two stationary points, namely $\check{\lambda}_2$ and $\hat{\lambda}_2$, where $\check{\lambda}_2 \leq \hat{\lambda}_2$. Moreover, $\check{\lambda}_2$ is negative, and consequently, does not belong to $\mathcal{ND}_1$. However, $\hat{\lambda}_2$ belongs to $\mathcal{ND}_1$ if and only if

$$\frac{(1-b)ab}{1-a}P_1 + b - 1 \leq P_2, \tag{2.137}$$

$$\frac{(1-a)ab}{1-b}P_2 + a - 1 \leq P_1, \tag{2.138}$$

$$\hat{\lambda}_2 \leq ab - \frac{1-b}{P_2}. \tag{2.139}$$

To find the stationary points, we investigate $\frac{\partial h_1(m\lambda_2 + c, \lambda_2)}{\partial \lambda_2} = 0$. According to (2.55), for $\lambda_1 = m\lambda_2 + c$, we have

$$h_1(m\lambda_2 + c, \lambda_2) = C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1 + b(m\lambda_2 + c)P_1}\Big). \tag{2.140}$$

Consequently,

$$\begin{aligned}\frac{\partial h_1(m\lambda_2 + c, \lambda_2)}{\partial \lambda_2} = &- \frac{aP_2}{1 + a\lambda_2 P_2} \\ &+ \frac{P_2(1 + bcP_1)}{\big(1 + bP_1(m\lambda_2 + c)\big)\big(1 + bP_1(m\lambda_2 + c) + P_2\lambda_2\big)}.\end{aligned} \tag{2.141}$$

To solve $\frac{\partial h_1(m\lambda_2+c,\lambda_2)}{\partial\lambda_2}=0$, we need to solve

$$\frac{aP_2}{1+a\lambda_2 P_2}=\frac{P_2(1+bcP_1)}{\big(1+bP_1(m\lambda_2+c)\big)\big(1+bP_1(m\lambda_2+c)+P_2\lambda_2\big)},\tag{2.142}$$

which is equivalent to

$$\begin{aligned}abP_1m(bP_1m+P_2)(\lambda_2^2)+2abP_1m(1+bP_1c)(\lambda_2)&\\ +(1+bP_1c)(abP_1c+a-1)&=0\\ \Leftrightarrow (\lambda_2^2)+2\frac{(1+bP_1c)}{(bP_1m+P_2)}(\lambda_2)+\frac{(1+bP_1c)(abP_1c+a-1)}{abP_1m(bP_1m+P_2)}&=0.\end{aligned}\tag{2.143}$$

Let us denote the solutions of (2.143) by $\check{\lambda}_2$ and $\check{\lambda}_2$, such that $\mathrm{Re}\Big\{\tilde{\lambda}_2\Big\}\le\mathrm{Re}\Big\{\hat{\lambda}_2\Big\}$. In fact, we can express $\check{\lambda}_2$ and $\hat{\lambda}_2$ as follows:

$$\begin{aligned}\hat{\lambda}_2=&\frac{1+bP_1c}{bP_1m+P_2}\Big(-1+\sqrt{1-\frac{(bP_1m+P_2)(abP_1c+a-1)}{(1+bP_1c)(abP_1m)}}\Big),\\ \check{\lambda}_2=&\frac{1+bP_1c}{bP_1m+P_2}\Big(-1-\sqrt{1-\frac{(bP_1m+P_2)(abP_1c+a-1)}{(1+bP_1c)(abP_1m)}}\Big).\end{aligned}\tag{2.144}$$

Note that $\check{\lambda}_2$ and $\hat{\lambda}_2$ are functions of a,b,P_1, and P_2. We find the constraints on (a,b,P_2,P_2) under which the equation (2.143) has exactly one non-negative solution that belongs to $\mathcal{ND}_1$. Note that, we have

$$\check{\lambda}_2+\hat{\lambda}_2=-2\frac{1+bP_1c}{bP_1m+P_2},\tag{2.145}$$

$$\check{\lambda}_2\hat{\lambda}_2=\frac{(1+bP_1c)(abP_1c+a-1)}{abP_1m(bP_1m+P_2)}.\tag{2.146}$$

We claim that $\check{\lambda}_2+\hat{\lambda}_2<0$. Note that according to (2.127), $m\ge 0$, and consequent, $bP_1m+P_2>0$. Moreover, according to (2.128), we can simplify $1+bP_1c$ as follows:

$$1+bP_1c=\frac{(1-b)(1+bP_1+P_2)}{1-b+P_2(1-ab)}>0.\tag{2.147}$$

Therefore, we have

$$\check{\lambda}_2+\hat{\lambda}_2=-2\frac{(1+bP_1c)}{(bP_1m+P_2)}<0.\tag{2.148}$$

Note that by (2.148), we can conclude that $\check{\lambda}_2$ cannot be a non-negative real number. Therefore, equation (2.143) does not have two non-negative solutions. Moreover, equation (2.143) has exactly one non-negative solution if $\hat{\lambda}_2$ is a non-negative number. Note that

$\hat{\lambda}_2 > 0$ if and only if $\check{\lambda}_2\hat{\lambda}_2 < 0$, which is valid if and only if

$$\begin{aligned}(abP_1c + a - 1) &< 0\\ \Leftrightarrow P_1c &< \frac{1-a}{ab}.\end{aligned} \tag{2.149}$$

Note that (2.149) is valid if and only if

$$\begin{aligned}\frac{P_1(1-b) - P_2(1-a)}{\big(1-b+P_2(1-ab)\big)} &< \frac{1-a}{ab}\\ \Leftrightarrow \frac{(1-b)ab}{1-a}P_1 + b - 1 &\leq P_2.\end{aligned} \tag{2.150}$$

Therefore, $\hat{\lambda}_2$ is non-negative if and only if (2.150) is satisfied.

Note that $(\hat{\lambda}_1 = m\hat{\lambda}_2 + c, \hat{\lambda}_2)$ is an acceptable power splitting if both $\hat{\lambda}_1$ and $\hat{\lambda}_2$ belong to $[0,1]$. We already showed that $\hat{\lambda}_2$ is non-negative if and only if (2.150) is satisfied. Similarly, it follows that that $\hat{\lambda}_1$ is non-negative if and only if

$$\frac{(1-a)ab}{1-b}P_2 + a - 1 \leq P_1. \tag{2.151}$$

We now show that $\hat{\lambda}_2 \leq 1$. Note that $\hat{\lambda}_2$ is the nonnegative root of the equation (2.143). Since (2.143) has one negative root $\check{\lambda}_2$, we can conclude that $\hat{\lambda}_2 \leq 1$, if for $\lambda_2 = 1$, the value of equation (2.143) is nonnegative, that is

$$\begin{aligned}&f(P_1,P_2) \doteq\\ &\qquad abP_1m(bP_1m + P_2)(\lambda_2^2) + 2abP_1m(1+bP_1c)(\lambda_2)\\ &\qquad + (1+bP_1c)(abP_1c + a - 1)|_{(\lambda_2=1)} \geq 0.\end{aligned} \tag{2.152}$$

Note that we only need to prove (2.152), for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. To this end, we first show that $f(P_1,P_2) \geq 0$, when $P_1 = \frac{1-a}{ab}$ and $P_2 = \frac{1-b}{ab}$. Then we show that $f(P_1,P_2)$ is an increasing function of P_1 and P_2, for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$.

By inserting (2.127) and (2.128) into (2.152), we see that

$$\begin{aligned}&f(P_1,P_2) =\\ &\qquad \frac{1}{1-b+P_2(1-ab)}\Big(ab^2(1-ab)P_1^2P_2 + ab(1-ab)P_1P_2^2\\ &\qquad + 2ab(1-ab)P_1P_2 + ab^2(1-b)P_1^2 + ab(1-a)P_2^2\\ &\qquad + (1-a)(ab+b-1)P_2 + (1-b)(2ab-b)P_1\\ &\qquad - (1-a)(1-b)\Big).\end{aligned} \tag{2.153}$$

First, note that, for $P_1 = \frac{1-a}{ab}$ and $P_2 = \frac{1-b}{ab}$, we have

$$f(P_1, P_2) = \frac{1-a}{ab} > 0. \tag{2.154}$$

Moreover, since $\frac{1}{1-b+P_2(1-ab)} \geq 0$, to show that $f(P_1, P_2)$ remains positive for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$, it is sufficient to show that the numerator in (2.153) remains positive. Let us denote the numerator in (2.153) by

$$\begin{aligned} f_N(P_1, P_2) \doteq& \\ & ab^2(1-ab)P_1^2P_2 + ab(1-ab)P_1P_2^2 \\ & + 2ab(1-ab)P_1P_2 + ab^2(1-b)P_1^2 + ab(1-a)P_2^2 \\ & + (1-a)(ab+b-1)P_2 + (1-b)(2ab-b)P_1 \\ & + (1-a)(1-b). \end{aligned} \tag{2.155}$$

One can check that $\frac{\partial f_N(P_1,P_2)}{\partial P_1}$ is an increasing function of P_1, when $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. Moreover, we have

$$\frac{\partial f_N(P_1, P_2)}{\partial P_1}\Big|_{(P_1=\frac{1-a}{ab}, P_2=\frac{1-b}{ab})} = \frac{1-a}{ab} \geq 0, \tag{2.156}$$

which proves that $\frac{\partial f_N(P_1,P_2)}{\partial P_1}$ is positive, for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. Therefore, we have

$$\frac{\partial f_N(P_1, P_2)}{\partial P_1} \geq 0. \tag{2.157}$$

Similarly, one can check that $\frac{\partial f_N(P_1,P_2)}{\partial P_2}$ is an increasing function of P_2, when $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. Moreover, we have

$$\frac{\partial f_N(P_1, P_2)}{\partial P_2}\Big|_{(P_1=\frac{1-a}{ab}, P_2=\frac{1-b}{ab})} = \frac{1-a}{ab} \geq 0, \tag{2.158}$$

which proves that $\frac{\partial f_N(P_1,P_2)}{\partial P_2}$ is positive, for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. Therefore, we have

$$\frac{\partial f_N(P_1, P_2)}{\partial P_2} \geq 0. \tag{2.159}$$

Note that, (2.154), (2.159), and(2.159) prove that $f(P_1, P_2)$ is greater than zero, for $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. This proves that $\hat{\lambda}_2 \leq 1$, as intended.

Next, we prove that $\hat{\lambda}_1 \leq 1$. Note that $\hat{\lambda}_1 \doteq m\hat{\lambda}_2 + c$. According to (2.127), $m \geq 0$. Therefore, $\hat{\lambda}_1$ takes its maximum value when $\hat{\lambda}_2$ take its maximum value. Moreover, we have proved that $\hat{\lambda}_2 \leq 1$. Consequently, we have

$$\begin{aligned}\hat{\lambda}_1 =& m\hat{\lambda}_2 + c\\ \leq& m + c\\ =& 1,\end{aligned} \tag{2.160}$$

where the last equality is valid, according to the definitions of m and c, given in (2.127) and (2.128), respectively.

Constraints (2.150) and (2.151) are the necessary and sufficient conditions for $\hat{\lambda}_1, \hat{\lambda}_2 \in [0, 1]$. However, $\hat{\lambda}_1, \hat{\lambda}_2$ should belong to $\mathcal{ND}_1$. Therefore, we should have

$$\hat{\lambda}_1 \leq ab - \frac{1-a}{P_1}, \tag{2.161}$$

$$\hat{\lambda}_2 \leq ab - \frac{1-b}{P_2}, \tag{2.162}$$

as shown in Figure 2.14. Note that $(\lambda_1 = ab - \frac{1-a}{P_1}, \lambda_2 = ab - \frac{1-b}{P_2})$ is one of the end points of $\mathcal{ND}_1$. Therefore, we have

$$ab - \frac{1-a}{P_1} = m(ab - \frac{1-b}{P_2}) + c. \tag{2.163}$$

Consequently, (2.161) is satisfied if and only if (2.162) is satisfied.

Note that the three constraints (2.150), (2.151), (2.162) represent the necessary and sufficient conditions for $(\hat{\lambda}_1, \hat{\lambda}_2) \in \mathcal{ND}_1$. In fact, (2.150) guarantees that $\hat{\lambda}_2 \geq 0$, (2.151) guarantees that $\hat{\lambda}_1 \geq 0$, and (2.162) guarantees that $\hat{\lambda}_2 \leq \tilde{\lambda}_2$. If these three constraints are satisfied, the stationary point that belongs to $\mathcal{ND}_1$ is given by

$$(\lambda_1^\star, \lambda_2^\star) = (m\hat{\lambda}_2 + c, \hat{\lambda}_2), \tag{2.164}$$

and the corresponding achievable sum-rate is given by

$$R_{\text{sum-HK}}(m\hat{\lambda}_2 + c, \hat{\lambda}_2) = h_1(m\hat{\lambda}_2 + c, \hat{\lambda}_2). \tag{2.165}$$

Note that we can simplify the achievable sum-rate given by $h_1(\lambda_1, \lambda_2)$ as follows:

$$\begin{aligned}h_1(\lambda_1, \lambda_2) =& C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1 + b\lambda_1 P_1}\Big)\\ =& C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) + C(a\lambda_2 P_2)\\ &+ C(\frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1 + a\lambda_2 P_2}) - C(b\lambda_1 P_1)\\ =& C(P_1 + aP_2) + g_1(\lambda_1, \lambda_2),\end{aligned} \tag{2.166}$$

where the function $g_1(\lambda_1, \lambda_2)$ is defined by

$$g_1(\lambda_1, \lambda_2) \doteq C(\frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1 + a\lambda_2 P_2}) - C(b\lambda_1 P_1). \tag{2.167}$$

Consequently, the achievable sum-rate expressed in (2.165) is equal to

$$R_{\text{sum-HK}}(m\hat{\lambda}_2 + c, \hat{\lambda}_2) = C(P_1 + aP_2) + g_1(m\hat{\lambda}_2 + c, \hat{\lambda}_2). \tag{2.168}$$

Similarly, one can simplify $h_2(\lambda_1, \lambda_2)$ as follows:

$$\begin{aligned}
h_2(\lambda_1, \lambda_2) =& C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1}\Big) + C\Big(\frac{\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big) \\
=& C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1}\Big) + C(b\lambda_1 P_1) \\
& + C(\frac{(1-b)\lambda_1 P_1 + a\lambda_2 P_2}{1 + b\lambda_1 P_1}) - C(a\lambda_2 P_2) \\
=& C(P_2 + bP_1) + g_2(\lambda_1, \lambda_2),
\end{aligned} \tag{2.169}$$

where the function $g_2(\lambda_1, \lambda_2)$ is defined by

$$g_2(\lambda_1, \lambda_2) \doteq C(\frac{(1-b)\lambda_1 P_1 + a\lambda_2 P_2}{1 + b\lambda_1 P_1}) - C(a\lambda_2 P_2). \tag{2.170}$$

Since we have $h_1(m\hat{\lambda}_2 + c, \hat{\lambda}_2) = h_2(m\hat{\lambda}_2 + c, \hat{\lambda}_2)$, we can equivalently express the maximum achievable sum-rate by

$$\begin{aligned}
R_{\text{sum-HK}}(m\hat{\lambda}_2 + c, \hat{\lambda}_2) =& h_2(m\hat{\lambda}_2 + c, \hat{\lambda}_2) \\
=& C(P_2 + bP_1) + g_2(m\hat{\lambda}_2 + c, \hat{\lambda}_2).
\end{aligned} \tag{2.171}$$

If the three constraints (2.150), (2.151), (2.162) are satisfied, then $\mathcal{ND}_1$ includes exactly one stationary point $(\hat{\lambda}_1, \hat{\lambda}_2)$, and therefore, the maximum achievable sum-rate corresponding to $\mathcal{ND}_1$ is given by

$$\max\{h_1(c, 0), h_1(0, c'), h_1(\tilde{\lambda}_1, \tilde{\lambda}_2), h_1(\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{2.172}$$

However, if these three constraints are not satisfied, then $\mathcal{ND}_1$ does not include any stationary point, and therefore, the maximum achievable sum-rate corresponding to $\mathcal{ND}_1$ is given by

$$\max\{h_1(c, 0), h_1(0, c'), h_1(\tilde{\lambda}_1, \tilde{\lambda}_2)\}. \tag{2.173}$$

Therefore, we can use the function $\mathbb{1}()$ and express the maximum achievable sum-rate corresponding to $\mathcal{ND}_1$ by

$$\begin{aligned}&\max\{h_1(c,0), h_1(0,c'), h_1(\tilde{\lambda}_1,\tilde{\lambda}_2),\\ &\quad h_1(\hat{\lambda}_1,\hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0)(\hat{\lambda}_2 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq \hat{\lambda}_2)\}.\end{aligned} \tag{2.174}$$

Note that we have

$$h_1(c,0) = C(P_1 + aP_2), \tag{2.175}$$

$$h_1(0,c') = C(P_2 + bP_1), \tag{2.176}$$

$$h_1(\tilde{\lambda}_1,\tilde{\lambda}_2) = C(P_1 + aP_2) + g_1(\tilde{\lambda}_1,\tilde{\lambda}_2), \tag{2.177}$$

$$h_1(\hat{\lambda}_1,\hat{\lambda}_2) = C(P_1 + aP_2) + g_1(\hat{\lambda}_1,\hat{\lambda}_2), \tag{2.178}$$

where the last two equalities are valid by (2.166). Therefore, (2.174) is equivalent to

$$\begin{aligned}\max\Big\{&C(P_1 + aP_2), C(P_2 + bP_1),\\ &C(P_1 + aP_2) + g_1(\tilde{\lambda}_1,\tilde{\lambda}_2),\\ &C(P_1 + aP_2)+\\ &g_1(\hat{\lambda}_1,\hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0)(\hat{\lambda}_2 \geq 0)\mathbb{1}(\tilde{\lambda}_2 \geq \hat{\lambda}_2)\Big\}.\end{aligned} \tag{2.179}$$

This completes the proof of 2.7-A of Lemma 2.7.

2.7-B: When $h_1(\lambda_1,\lambda_2) = h_3(\lambda_1,\lambda_2) \leq h_2(\lambda_1,\lambda_2)$, the optimization problem (2.58) reduces to

$$\begin{aligned}&\max_{\lambda_1,\lambda_2\in[0,1]} && R_{\text{sum-HK}}(\lambda_1,\lambda_2) =\\ &\max_{\lambda_1,\lambda_2\in[0,1]} && h_1(\lambda_1,\lambda_2)\\ &\text{subject to} && h_1(\lambda_1,\lambda_2) \leq h_2(\lambda_1,\lambda_2).\end{aligned} \tag{2.180}$$

Since $h_1(\lambda_1,\lambda_2) = h_3(\lambda_1,\lambda_2)$, by Lemma 2.3, we have

$$\lambda_2 = ab - \frac{1-b}{P_2}. \tag{2.181}$$

Therefore, the optimization problem (2.180) reduces to

$$\begin{aligned}&\max_{0\leq\lambda_1\leq 1} && h_1(\lambda_1, ab - \frac{1-b}{P_2})\\ &\text{subject to} && h_3(\lambda_1, ab - \frac{1-b}{P_2}) \leq h_2(\lambda_1, ab - \frac{1-b}{P_2}).\end{aligned} \tag{2.182}$$

To solve the optimization problem (2.180), we investigate $\frac{\partial h_1\left(\lambda_1, ab-\frac{1-b}{P_2}\right)}{\partial \lambda_1} = 0$. According to (2.55), for $\lambda_2 = ab - \frac{1-b}{P_2}$, we have

$$h_1(\lambda_1, ab - \frac{1-b}{P_2}) = C\Big(\frac{P_1 + a(1-ab+\frac{1-b}{P_2})P_2}{1 + a(ab - \frac{1-b}{P_2})P_2}\Big) + C\Big(\frac{(ab - \frac{1-b}{P_2})P_2}{1+b\lambda_1 P_1}\Big). \quad (2.183)$$

Clearly, (2.183) is a decreasing function of λ_1. Therefore, the optimal λ_1 is the smallest λ_1 that satisfies $h_3(\lambda_1, \lambda_2) \le h_2(\lambda_1, \lambda_2)$. According to Lemma 2.3, $h_3(\lambda_1, \lambda_2) \le h_2(\lambda_1, \lambda_2)$ is equivalent to $\lambda_1 \ge ab - \frac{1-a}{P_1}$. Consequently, the optimal λ_1 that maximizes (2.180) is given by:

$$\lambda_1^\star = \tilde{\lambda}_1 \doteq ab - \frac{1-a}{P_1}. \quad (2.184)$$

This means the optimal solution of (2.180) is given by

$$\begin{aligned} \lambda_1^\star &= \tilde{\lambda}_1 = ab - \frac{1-a}{P_1}, \\ \lambda_2^\star &= \tilde{\lambda}_2 = ab - \frac{1-b}{P_2}, \end{aligned} \quad (2.185)$$

and the achievable sum-rate is given by

$$R_{\text{sum-HK}}(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}) = h_1(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}). \quad (2.186)$$

Note that, according to Lemma 2.3, for $(\lambda_1, \lambda_2) = (\tilde{\lambda}_1, \tilde{\lambda}_1)$, we have $h_1(\tilde{\lambda}_1, \tilde{\lambda}_1) = h_2(\tilde{\lambda}_1, \tilde{\lambda}_1) = h_3(\tilde{\lambda}_1, \tilde{\lambda}_1)$. Therefore, (2.186) can be expressed as

$$\begin{aligned} R_{\text{sum-HK}}(\tilde{\lambda}_1, \tilde{\lambda}_1) =& h_1(\tilde{\lambda}_1, \tilde{\lambda}_1) \\ =& h_2(\tilde{\lambda}_1, \tilde{\lambda}_1) \\ =& h_3(\tilde{\lambda}_1, \tilde{\lambda}_1). \end{aligned} \quad (2.187)$$

Similar to (2.166), we can simplify (2.187). In fact, the achievable sum-rate is equal to

$$\begin{aligned} R_{\text{sum-HK}}(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}) =& C(P_1 + aP_2) + g_1(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}) \\ =& C(P_2 + bP_2) + g_2(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}), \end{aligned} \quad (2.188)$$

where the functions $g_1()$ and $g_2()$ are defined in (2.167) and (2.170), respectively.

Note that $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ is an acceptable power splitting if both $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ belong to $[0, 1]$. Since $(\tilde{\lambda}_1, \tilde{\lambda}_2) = (ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2})$, we have

$$\tilde{\lambda}_1 \in [0, 1] \Leftrightarrow P_1 \geq \frac{1-a}{ab}, \tag{2.189}$$

$$\tilde{\lambda}_1 \in [0, 1] \Leftrightarrow P_2 \geq \frac{1-b}{ab}. \tag{2.190}$$

This completes the proof of 2.7-B of Lemma 2.7. Note that the proof of 2.7-C follows from the proof of 2.7-B, if we exchange indices 1 with 2 and cross-link gains a with b. Therefore, the proof of Lemma 2.7 is complete.

□

2.3.9 Solving the Optimization Problem Corresponding to the Maximum HK Sum-Rate

Now that we have investigated all the three categories of points, we can prove Theorems 2.1 and 2.2. In fact, it is sufficient to compare the achievable sum-rates corresponding to all sub-categories. In Lemma 2.5-2.7, we calculated the achievable sum-rate of all these sub-categories. By comparing these achievable sum-rates, we can now prove Theorem 2.1 and Theorem 2.2 as follows:

Proof. First, note that $R^{\max}_{\text{sum-HK}}$ is only unknown for the barely weak-sub-class, as depicted in Figure 2.4. In the following, we show that the barely weak-sub-class can be partitioned into four parts. For each part, we characterize the optimal power splitting and find the maximum achievable sum-rate $R^{\max}_{\text{sum-HK}}$. Note that the optimal power splitting belongs to one of the sub-categories investigated in Lemma 2.6 and 2.7. Table 2.4 summarizes the results of these Lemmas. Note that for any (a, b, P_1, P_2) in the barely weak sub-class, the optimal power splitting belongs to one of the sub-categories of Table 2.4. Therefore, we should find the constraints under which the achievable sum-rate of one sub-category is greater than that of all other sub-categories.

Note that in Table 2.4, the optimal power splitting corresponding to $\mathcal{ND}_1$ is $(\hat{\lambda}_1, \hat{\lambda}_2)$. In Lemma 2.7, we proved that the optimal power splitting of this sub-category can have four cases and is given by

$$(\lambda_1^\star, \lambda_2^\star) \in \{(c, 0), (0, c'), (\tilde{\lambda}_1, \tilde{\lambda}_2), (\hat{\lambda}_1, \hat{\lambda}_2)\}.$$

Sub-category	Optimal $(\lambda_1^\star, \lambda_2^\star)$	Achievable sum-rate $R_{\text{sum-HK}}(\lambda_1^\star, \lambda_2^\star)$
$\mathcal{B}_1$	$(\lambda_1^\star \geq c, 0)$	$C(P_1 + aP_2)$
$\mathcal{B}_2$	$(0, 1)$	$C(P_2 + bP_1)$
$\mathcal{B}_3$	$(0, \lambda_2^\star \geq c')$	$C(P_2 + bP_1)$
$\mathcal{B}_4$	$(1, 0)$	$C(P_1 + aP_2)$
$\mathcal{ND}_1$	$(\hat{\lambda}_1, \hat{\lambda}_2)$	$h_1(\hat{\lambda}_1, \hat{\lambda}_2)$
$\mathcal{ND}_2$	$(\tilde{\lambda}_1, \tilde{\lambda}_2)$	$h_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$
$\mathcal{ND}_3$	$(\tilde{\lambda}_1, \tilde{\lambda}_2)$	$h_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$

Table 2.4: Sub-categories, their corresponding optimal power splittings and achievable sum-rate expressions, for the barely weak interference sub-class.

Note that $(c, 0)$ and $(0, c')$ belong to the boundary. Moreover, $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ belongs to $\mathcal{ND}_2$. Therefore, if we do not consider $(c, 0)$, $(0, c')$, and $(\tilde{\lambda}_1, \tilde{\lambda}_2)$ in $\mathcal{ND}_1$, the maximum achievable sum-rate does not decrease.

First we characterize the constraints under which the $R_{\text{sum-HK}}^{\max}$ is given by $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$, i.e., the optimal power splitting belongs to the sub-category $\mathcal{ND}_1$. Note that, according to Lemma 2.7, the sum-rate $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is achievable if and only if

$$\frac{(1-b)ab}{1-a}P_1 + b - 1 \leq P_2, \tag{2.191}$$

$$\frac{(1-a)ab}{1-b}P_2 + a - 1 \leq P_1, \tag{2.192}$$

$$\hat{\lambda}_2 \leq ab - \frac{1-b}{P_2}. \tag{2.193}$$

These three constraints demonstrate a region in $\mathbb{R}_+^4$ which can be demonstrated in the P_1P_2-plane. Note that this region is a subset of the barely weak interference sub-class. We refer to this region as the *non-zero power splitting II* sub-class, as can be seen in Figure 2.16. For this sub-class, $(m\hat{\lambda}_2 + c, \hat{\lambda}_2)$ is an acceptable power splitting that belongs to the non-differentiable sub-category $\mathcal{ND}_1$ and results in the maximum achievable sum-rate given by $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$. We prove that for any (a, b, P_1, P_2) that belongs to this sub-class, we have $R^{\max}_{\text{sum-HK}} = h_1(\hat{\lambda}_1, \hat{\lambda}_2)$. To this end, we should show that if (a, b, P_1, P_2) belongs to the non-zero power splitting II sub-class, then $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is greater than all other sum-rates listed in Table 2.4.

Note that the non-zero power splitting II is inside the barely weak interference region. For the barely weak interference sub-class, the maximum sum-rate achieved by investigating the boundary points is given by

$$\max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\}, \tag{2.194}$$

as shown in Figure 2.13. Therefore, we need to prove that

$$h_1(\hat{\lambda}_1, \hat{\lambda}_2) \geq \max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\}. \tag{2.195}$$

We present the proof for the case

$$\max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\} = C(P_1 + aP_2). \tag{2.196}$$

Note that (2.196) is valid if and only if $P_1(1-b) \geq P_2(1-a)$. Due to the symmetry of the problem, the proof of (2.195) for $P_1(1-b) \leq P_2(1-a)$ follows by exchanging index 1 with 2 and channel gain a with b.

Figure 2.17 demonstrates the proof of $h_1(\hat{\lambda}_1, \hat{\lambda}_2) \geq C(P_1 + aP_2)$, for $P_1(1-b) \geq P_2(1-a)$. In fact, in the barely weak interference sub-class, when $P_1(1-b) \geq P_2(1-a)$, we have

$$\max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\} = C(P_1 + aP_2), \tag{2.197}$$

as shown in Figure 2.13. Note that $(\hat{\lambda}_1, \hat{\lambda}_2)$ is the optimal solution of (2.58) if we restrict our search to the points that lie on $\mathcal{ND}_1$. Note that since $P_1(1-b) \geq P_2(1-a)$, we know that $c \geq 0$. Since $(c, 0)$ lies on the line segment $\mathcal{ND}_1$, we have

$$R_{\text{sum-HK}}(\hat{\lambda}_1, \hat{\lambda}_2) \geq R_{\text{sum-HK}}(c, 0), \tag{2.198}$$

as shown in Figure 2.17. On the other hand, in Lemma 2.6, we proved that when $\lambda_2 = 0$, we have

$$R_{\text{sum-HK}}(\lambda_1, 0) \leq R_{\text{sum-HK}}(c, 0) = C(P_1 + aP_2). \tag{2.199}$$

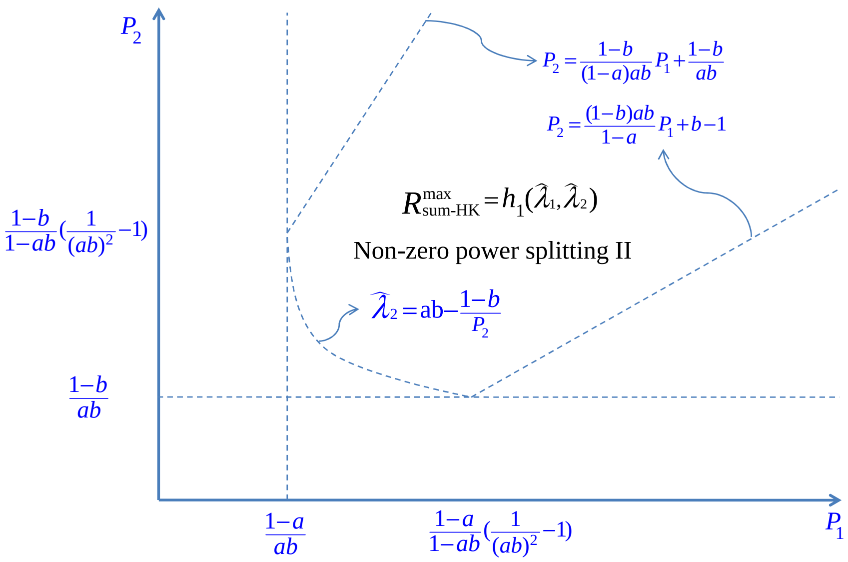

Figure 2.16: The non-zero power splitting II sub-class demonstrated in the P_1P_2-plane.

Comparing (2.198) and (2.199), we conclude that

$$R_{\text{sum-HK}}(\hat{\lambda}_1, \hat{\lambda}_2) \geq R_{\text{sum-HK}}(c, 0) = C(P_1 + aP_2). \tag{2.200}$$

Similarly, one can show that

$$R_{\text{sum-HK}}(\hat{\lambda}_1, \hat{\lambda}_2) \geq R_{\text{sum-HK}}(0, c') = C(P_2 + bP_1). \tag{2.201}$$

Therefore, we have

$$\begin{aligned} R_{\text{sum-HK}}(m\hat{\lambda}_2 + c, \hat{\lambda}_2) &= h_1(\hat{\lambda}_1, \hat{\lambda}_2) \\ &= h_2(\hat{\lambda}_1, \hat{\lambda}_2) \\ &\geq \max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\}. \end{aligned} \tag{2.202}$$

Therefore, we have shown that, for the non-zero power splitting II sub-class, $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is greater than the sum-rate achieved by the four sub-categories of the boundary, i.e., $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$, and $\mathcal{B}_4$. The proof will be complete if we show that it is also greater than the sum-rate of the $\mathcal{ND}_2$ and $\mathcal{ND}_3$ sub-categories. In the proof of Part 2.7-B, we show that the optimal power splitting over non-differentiable points expressed in $\mathcal{ND}_2$ and $\mathcal{ND}_3$ is given by $(\lambda_1, \lambda_2) = (ab - \frac{1-a}{P_1}, ab - \frac{1-a}{P_1})$. Therefore, we need to show that

$$R_{\text{sum-HK}}(\hat{\lambda}_1, \hat{\lambda}_2) \geq R_{\text{sum-HK}}(ab - \frac{1-a}{P_1}, ab - \frac{1-a}{P_1}). \tag{2.203}$$

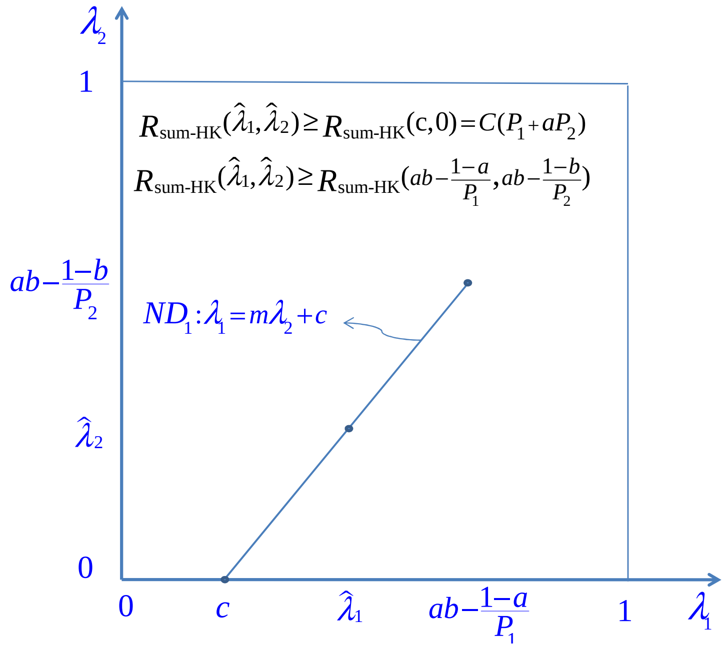

Figure 2.17: For the non-zero power splitting II sub-class, the achievable sum-rate corresponding to $\mathcal{ND}_1$ is greater than the achievable sum-rate corresponding to all other sub-categories.

However, $(\lambda_1, \lambda_2) = (ab - \frac{1-a}{P_1}, ab - \frac{1-a}{P_1})$ lies on the line $\lambda_1 = m\lambda_2 + c$, as shown in Figure 2.17. Over this line, $(\hat\lambda_1, \hat\lambda_2)$ is the optimal solution of (2.58). Therefore, (2.203) is valid, and this proves that over the non-zero power splitting II sub-class, we have

$$R^{\max}_{\text{sum-HK}} = h_1(\hat\lambda_1, \hat\lambda_2) = C(P_1 + aP_2) + g_1\hat\lambda_1, \hat\lambda_2). \tag{2.204}$$

Second, we characterize the constraints under which the $R^{\max}_{\text{sum-HK}}$ is given by $h_1(\tilde\lambda_1, \tilde\lambda_2)$, i.e., the optimal power splitting belongs to the sub-category $\mathcal{ND}_2$. Therefore, we need to compare the sum-rate corresponding to $\mathcal{ND}_2$ with the sum-rate corresponding to all other sub-categories.

Remember that we only investigate the barely weak interference sub-class, in which $P_1 > \frac{1-a}{ab}$ and $P_2 > \frac{1-b}{ab}$. According to Lemma 2.7, for the barely weak interference sub-class, both $\tilde\lambda_1$ and $\tilde\lambda_2$ are acceptable power splittings.

Moreover, note that

$$\begin{aligned} h_1(\tilde\lambda_1, \tilde\lambda_2) =& h_2(\tilde\lambda_1, \tilde\lambda_2) \\ =& C(P_1 + aP_2) + g_1(\tilde\lambda_1, \tilde\lambda_2) \\ =& C(P_2 + bP_2) + g_2(\tilde\lambda_1, \tilde\lambda_2). \end{aligned} \tag{2.205}$$

Therefore, we have

$$h_1(\tilde{\lambda}_1, \tilde{\lambda}_2) \geq \max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\}, \tag{2.206}$$

if and only if

$$g_1(\tilde{\lambda}_1, \tilde{\lambda}_2) \geq 0, \tag{2.207}$$

$$g_2(\tilde{\lambda}_1, \tilde{\lambda}_2) \geq 0. \tag{2.208}$$

According to (2.167), we have

$$\begin{aligned} g_1(\lambda_1, \lambda_2) &\geq 0 \\ \Leftrightarrow \frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1 + a\lambda_2 P_2} &\geq b\lambda_1 P_1 \\ \Leftrightarrow \lambda_1 &\leq \frac{1}{bP_1}(\frac{1}{a} - 1). \end{aligned} \tag{2.209}$$

Similarly, according to (2.170), we have

$$\begin{aligned} g_2(\lambda_1, \lambda_2) &\geq 0 \\ \Leftrightarrow \frac{(1-b)\lambda_1 P_1 + a\lambda_2 P_2}{1 + b\lambda_1 P_1} &\geq a\lambda_2 P_2 \\ \Leftrightarrow \lambda_2 &\leq \frac{1}{aP_2}(\frac{1}{b} - 1). \end{aligned} \tag{2.210}$$

Therefore, we have

$$h_1(\tilde{\lambda}_1, \tilde{\lambda}_2) \geq \max\Big\{C(P_1 + aP_2), C(P_2 + bP_1)\Big\}, \tag{2.211}$$

if and only if

$$\begin{aligned} \tilde{\lambda}_1 &\leq \frac{1}{bP_1}(\frac{1}{a} - 1), \\ \tilde{\lambda}_2 &\leq \frac{1}{aP_2}(\frac{1}{b} - 1), \end{aligned} \tag{2.212}$$

which can be re-written as

$$\begin{aligned} ab - \frac{1-a}{P_1} &\leq \frac{1}{bP_1}(\frac{1}{a} - 1), \\ ab - \frac{1-b}{P_2} &\leq \frac{1}{aP_2}(\frac{1}{b} - 1). \end{aligned} \tag{2.213}$$

Note that (2.213) is equivalent to

$$\begin{aligned} P_1 &\leq \frac{1-a}{1-ab}(\frac{1}{(ab)^2} - 1), \\ P_2 &\leq \frac{1-b}{1-ab}(\frac{1}{(ab)^2} - 1). \end{aligned} \tag{2.214}$$

Therefore, for the barely weak interference sub-class, $h_1(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2}) \geq \max\Big\{C(P_1+aP_2), C(P_2+bP_1)\Big\}$ if and only if

$$\begin{aligned}\frac{1-a}{ab} &< P_1 \leq \frac{1-a}{1-ab}(\frac{1}{(ab)^2}-1),\\ \frac{1-b}{ab} &< P_2 \leq \frac{1-b}{1-ab}(\frac{1}{(ab)^2}-1).\end{aligned} \tag{2.215}$$

This region is depicted in Figure 2.18. In this region, $h_1(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2})$ is greater than the sum-rate corresponding to all four sub-categories of the boundary. Finally, we compare $h_1(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2})$ with $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$, i.e., the sum-rate corresponding to $\mathcal{ND}_1$.

According to Lemma 2.7, $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is the sum-rate corresponding to $\mathcal{ND}_1$ if (a, b, P_1, P_2) belongs to the power splitting II sub-class. Moreover, inside this sub-class, $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is greater than the sum-rates corresponding to all other sub-categories. Therefore, we only need to consider the compliment of the Power splitting II sub-class. Consequently, the constraints under which $h_2(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2})$ is greater than all other sum-rates corresponding to other sub-categories is specified by

$$\begin{aligned}\frac{1-a}{ab} &<P_1 \leq \frac{1-a}{1-ab}(\frac{1}{(ab)^2}-1),\\ \frac{1-b}{ab} &<P_2 \leq \frac{1-b}{1-ab}(\frac{1}{(ab)^2}-1),\\ \hat{\lambda}_2 &>ab-\frac{1-b}{P_2}.\end{aligned} \tag{2.216}$$

Since $\hat{\lambda}_2 > ab-\frac{1-b}{P_2}$ implies that $P_1 \leq \frac{1-a}{1-ab}(\frac{1}{(ab)^2}-1)$ and $P_2 \leq \frac{1-b}{1-ab}(\frac{1}{(ab)^2}-1)$, (2.216) is equivalent to

$$\begin{aligned}\frac{1-a}{ab} &< P_1,\\ \frac{1-b}{ab} &< P_2,\\ \hat{\lambda}_2 &> ab-\frac{1-b}{P_2},\end{aligned} \tag{2.217}$$

as can be seen in Figure 2.18. We refer this region as the *non-zero power splitting I* sub-class. For this sub-class, we have

$$R^{\max}_{\text{sum-HK}} = h_1(\tilde{\lambda}_1, \tilde{\lambda}_2) = C(P_1+aP_2) + g_1(\tilde{\lambda}_1, \tilde{\lambda}_2). \tag{2.218}$$

Third, we characterize the constraints under which the $R^{\max}_{\text{sum-HK}}$ is given by $C(P_1 + aP_2)$, i.e., the optimal power splitting belongs to the sub-category $\mathcal{B}_1$ or $\mathcal{B}_4$. Since we have

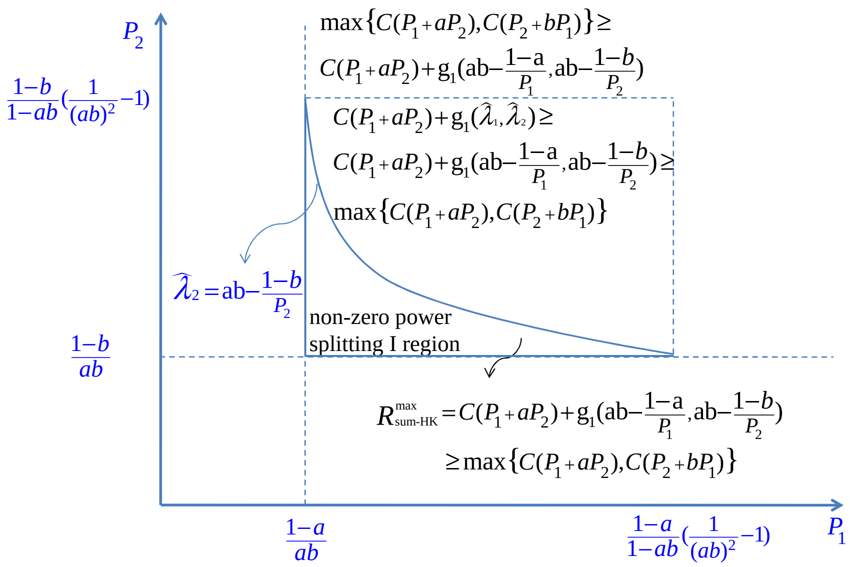

Figure 2.18: The non-zero power splitting I sub-class, projected onto the P_1P_2-plane. For this sub-class, $C(P_1+aP_2)+g_1(ab-\frac{1-a}{P_1}, ab-\frac{1-b}{P_2})$, which corresponds to $\mathcal{ND}_2$, is greater than the sum-rate corresponding to all other sub-categories.

characterized the Power splitting I and II sub-classes in which $h_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$ and $h_1(\hat{\lambda}_1, \hat{\lambda}_2)$ show $R^{\max}_{\text{sum-HK}}$, respectively, we only need to compare $C(P_1+aP_2)$ with $C(P_2+bP_1)$. Note that $C(P_1+aP_2) \geq C(P_2+bP_1)$ if and only if $P_1(1-b) \geq P_2(1-a)$. Therefore, $C(P_1+aP_2)$ is greater than other subcategories and equals $R^{\max}_{\text{sum-HK}}$ if and only if

$$\begin{aligned} \frac{(1-b)ab}{1-a}P_1 + b - 1 &\geq P_2, \\ \frac{1-b}{ab} &\leq P_2, \end{aligned} \tag{2.219}$$

as shown in Figure 2.19.

Similarly, $C(P_2+bP_1)$ is greater than other subcategories and equals $R^{\max}_{\text{sum-HK}}$ if and only if

$$\begin{aligned} \frac{(1-a)ab}{1-b}P_2 + a - 1 &\geq P_1, \\ \frac{1-a}{ab} &\leq P_1, \end{aligned} \tag{2.220}$$

as shown in Figure 2.19. In fact, Figure 2.19 shows that the entire barely weak interference sub-class is partitioned into four sub-classes. For each sub-class, the expression that shows $R^{\max}_{\text{sum-HK}}$ is demonstrated.

Note that Figure 2.4 demonstrates $R^{\max}_{\text{sum-HK}}$ for the entire weak interference class, except the barely weak interference sub-class. On the other hand, Figure 2.19 demonstrates

$R_{\text{sum-HK}}^{\max}$ only for the barely weak interference sub-class. By comparing these two figures, we see that $C(P_1 + aP_2)$ corresponds to $R_{\text{sum-HK}}^{\max}$ for two adjacent sub-classes. In fact, $R_{\text{sum-HK}}^{\max} = C(P_1 + aP_2)$ if

$$\begin{aligned} \frac{(1-b)ab}{1-a}P_1 + b - 1 &\geq P_2, \\ \frac{1-b}{ab} &\leq P_2, \end{aligned} \tag{2.221}$$

as shown in Figure 2.19. On the other hand, $R_{\text{sum-HK}}^{\max} = C(P_1 + aP_2)$ if

$$\begin{aligned} P_1 &> \frac{1-a}{ab}, \\ P_2 &\leq \frac{1-b}{ab}, \end{aligned} \tag{2.222}$$

as shown in Figure 2.4. Therefore, $R_{\text{sum-HK}}^{\max} = C(P_1 + aP_2)$ for the union of the regions expressed by (2.221) and (2.222). Therefore, for the weak interference class, we have $R_{\text{sum-HK}}^{\max} = C(P_1 + aP_2)$ if

$$\begin{aligned} &P_1 > \frac{1-a}{ab}, \\ &P_2 \leq \max\{\frac{1-b}{ab}, \frac{(1-b)ab}{1-a}P_1 + b - 1\}, \end{aligned} \tag{2.223}$$

as shown in Figure 2.5. We denote to this sub-class as *weakly mixed interference I* sub-class.

Similarly, for the weak interference class, we have $R_{\text{sum-HK}}^{\max} = C(P_2 + bP_1)$ if

$$\begin{aligned} &P_2 > \frac{1-b}{ab}, \\ &P_1 \leq \max\{\frac{1-a}{ab}, \frac{(1-a)ab}{1-b}P_2 + a - 1\}, \end{aligned} \tag{2.224}$$

as shown in Figure 2.5. We denote to this sub-class as *weakly mixed interference II* sub-class. As one can see in Figure 2.5, the entire weak interference class is partitioned into five sub-classes. For each sub-class, the optimal power splitting and the maximum sum-rate is shown in Table 2.2. This completes the proof. □

Theorem1 investigates the maximum achievable sum-rate of a general two-user GIC, when HK scheme with Gaussian inputs and no time sharing is used. Therefore, it can be used to characterize the maximum achievable sum-rate for some particular classes of the two-user GIC. For instance, define the class of semi-symmetric two-user GICs as

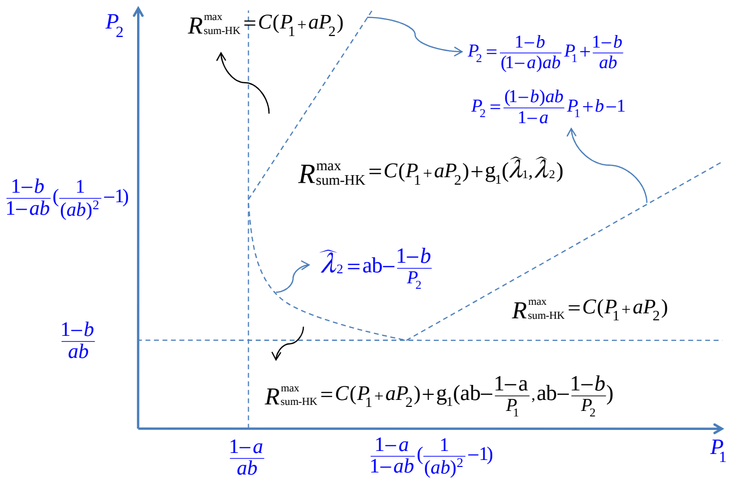

Figure 2.19: The barely weak interference sub-class is partitioned into four sub-classes, and for each sub-class, $R^{\max}_{\text{sum-HK}}$ is demonstrated.

all two-user GICs in which $P_1(1-b) = P_2(1-a)$. Note that the two-user symmetric GIC, in which $P_1 = P_2$ and $a = b$, is a special member of this class. Over the barely weak interference sub-class, when $P_1(1-b) = P_2(1-a)$, the optimal solution is always a non-differentiable point. In fact, for the class of semi-symmetric two-user GICs, the optimal power splitting $(\lambda_1^\star, \lambda_2^\star)$ is always symmetric, i.e., $\lambda_1^\star = \lambda_2^\star$. The following theorem investigates the achievable sum-rate of the semi-symmetric two-user GIC.

Theorem 2.4. *For a two-user semi-symmetric GIC, the maximum achievable sum-rate of the HK scheme with Gaussian inputs is given by*

$$R^{\max}_{\text{sum-HK}} = \begin{cases} C\left(\frac{P_1}{1+aP_2}\right) + C\left(\frac{P_2}{1+bP_1}\right) & \text{if } P_1 \leq \frac{1-a}{ab}, \\ C(P_1 + aP_2) + g(\lambda_s) & \text{if } \frac{1-a}{ab} < P_1 \leq \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2}, \\ C(P_1 + aP_2) + g(\hat{\lambda}) & \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2} < P_1, \end{cases} \tag{2.225}$$

where $g(\lambda) = C(\frac{P_1\lambda}{1+a\lambda P_2}) - C(b\lambda P_1)$, *and*

$$\lambda_s = ab - \frac{1-a}{P_1}, \tag{2.226}$$

$$\hat{\lambda} = \frac{1-a}{1-ab}\frac{\sqrt{ab}-ab}{abP_1}. \tag{2.227}$$

Moreover, the optimal power splitting is given by

$$(\lambda_1^\star, \lambda_2^\star) = \begin{cases} (1,1) & \text{if } P_1 \leq \frac{1-a}{ab}, \\ (\lambda_s, \lambda_s) & \text{if } \frac{1-a}{ab} < P_1 \leq \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2}, \\ (\hat{\lambda}, \hat{\lambda}) & \text{if } \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2} < P_1. \end{cases} \tag{2.228}$$

Proof. In a two-user semi-symmetric GIC, if $P_1 \leq \frac{1-a}{ab}$, then we have $P_2 \leq \frac{1-b}{ab}$. Therefore, the maximum sum-rate is achieved by treating interference as noise. When $P_1 \leq \frac{1-a}{ab}$, if $\hat{\lambda}_2 \geq ab - \frac{1-b}{P_2}$, then the maximum sum-rate is achieved by $(\lambda_1^\star, \lambda_2^\star) = (ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2})$. Note that since $P_1(1-b) = P_2(1-a)$, we have $ab - \frac{1-a}{P_1} = ab - \frac{1-b}{P_2}$. Finally, if $\hat{\lambda}_2 < ab - \frac{1-b}{P_2}$, then the maximum sum-rate is achieved by $(\lambda_1^\star, \lambda_2^\star) = (m\hat{\lambda}_2 + c, \hat{\lambda}_2)$, where m, c, and $\hat{\lambda}_2$ are given by (2.127), (2.128), and (2.144), respectively.

Note that, since $P_1(1-b) = P_2(1-a)$, we can easily check that $m = 1$ and $c = 0$. Therefore,

$$\begin{aligned} \hat{\lambda}_2 =& \frac{1+bP_1c}{bP_1m+P_2}\Big(-1+\sqrt{1-\frac{(bP_1m+P_2)(abP_1c+a-1)}{(1+bP_1c)(abP_1m)}}\Big) \\ =& \frac{1}{bP_1+P_2}\Big(-1+\sqrt{1-\frac{(bP_1+P_2)(a-1)}{abP_1}}\Big) \\ \overset{(a)}{=}& \frac{1-a}{(1-ab)P_1}\Big(-1+\sqrt{\frac{1}{ab}}\Big) \\ =& \frac{1-a}{(1-ab)P_1}\frac{\sqrt{ab}-ab}{ab}, \end{aligned} \tag{2.229}$$

where (a) is valid because $bP_1 + P_2 = P_1\frac{1-ab}{1-a}$. Moreover, $\hat{\lambda}_2 \geq ab - \frac{1-b}{P_2}$ is valid if and only if

$$\begin{aligned} \frac{1-a}{1-ab}\frac{\sqrt{ab}-ab}{abP_1} \geq& ab - \frac{1-b}{P_2} \\ \Leftrightarrow \frac{1-a}{1-ab}\frac{\sqrt{ab}-ab}{abP_1} \geq& ab - \frac{1-a}{P_1} \\ \Leftrightarrow P_1 \leq& \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2}. \end{aligned} \tag{2.230}$$

This completes the proof. $\square$

On interesting observation about Theorem 2.4 is the value of $g(\hat{\lambda}) = C(\frac{P_1\hat{\lambda}}{1+a\hat{\lambda}P_2}) -$

$C(b\hat{\lambda}P_1)$. Note that, according to (2.227), we have

$$P_1\hat{\lambda} = \frac{1-a}{1-ab}\frac{\sqrt{ab}-ab}{ab},$$
$$P_2\hat{\lambda} = P_1\frac{1-b}{1-a}\hat{\lambda} = \frac{1-b}{1-ab}\frac{\sqrt{ab}-ab}{ab}. \tag{2.231}$$

Therefore, $g(\hat{\lambda})$ does not depend on P_1 and P_2. In fact, we have

$$g(\hat{\lambda}) = \log\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}} = 2C\Big(\frac{(1-\sqrt{a})(1-\sqrt{b})}{\sqrt{a}+\sqrt{b}}\Big). \tag{2.232}$$

This implies that for fixed values of a and b and large values of P_1, i.e., $P_1 > \frac{(1-a)(\sqrt{ab}-(ab)^2)}{(1-ab)(ab)^2}$, the achievable sum-rate is given by $C(P_1 + aP_2)$ plus a constant term $2C\Big(\frac{(1-\sqrt{a})(1-\sqrt{b})}{\sqrt{a}+\sqrt{b}}\Big)$.

Corollary 2.2. *For a two-user symmetric GIC, in which $P_1 = P_2 = P$ and $a = b$, the maximum achievable sum-rate of the HK scheme with Gaussian inputs is given by*

$$R^{\max}_{\text{sum-HK}} = \begin{cases} 2C\Big(\frac{P}{1+aP}\Big) & \text{if } P \le \frac{1-a}{a^2}, \\ C\Big(P(a+1)\Big) + g(\lambda_s) & \text{if } \frac{1-a}{a^2} < P \le \frac{1-a^3}{(1+a)a^3}, \\ C\Big(P(a+1)\Big) + g(\hat{\lambda}) & \text{if } \frac{1-a^3}{(1+a)a^3} < P, \end{cases} \tag{2.233}$$

where $g(\lambda) = C(\frac{P\lambda}{1+a\lambda P}) - C(a\lambda P)$, and

$$\lambda_s = a^2 - \frac{1-a}{P}, \tag{2.234}$$
$$\hat{\lambda} = \frac{1-a}{a(1+a)P}. \tag{2.235}$$

Moreover, the optimal power splitting is given by

$$(\lambda_1^\star, \lambda_2^\star) = \begin{cases} (1,1) & \text{if } P_1 \le \frac{1-a}{a^2}, \\ (\lambda_s, \lambda_s) & \text{if } \frac{1-a}{a^2} < P \le \frac{1-a^3}{(1+a)a^3}, \\ (\hat{\lambda}, \hat{\lambda}) & \text{if } \frac{1-a^3}{(1+a)a^3} < P. \end{cases} \tag{2.236}$$

Note that [40] investigates the two-user symmetric GIC, and shows that if power is allocated symmetrically, (2.233) is the maximum achievable sum-rate of the HK scheme. However, Corollary 2.2 shows that (2.233) is indeed the maximum achievable sum-rate of the HK scheme and no non-symmetric power splitting can achieve a higher sum-rate.

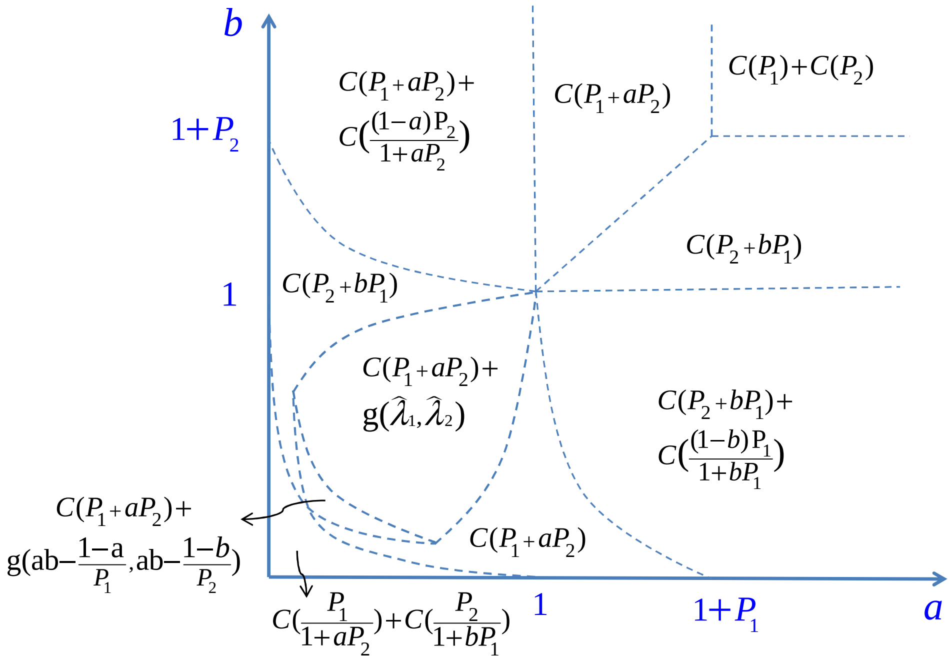

Figure 2.20: The maximum achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing for all values of a and b.

Next, we characterize the maximum achievable sum-rate of the HK scheme for all values of a and b. Note that, when interference is weak, Theorem 2.1 completely characterizes the maximum achievable sum-rate of the two-user GIC achieved by the HK scheme with Gaussian inputs and no time sharing, as shown in Figure 2.6. Moreover, the maximum achievable sum-rate expressions for the mixed and strong interference classes are already known, as shown in Figure 2.2. Comparing Figure 2.6 with Figure 2.2, we characterize the maximum achievable sum-rate of the HK scheme with Gaussian inputs and no time sharing, for all values of a and b, as shown in Figure 2.20.

One interesting observation about Figure 2.20 is the region that corresponds to $R^{\max}_{\text{sum-HK}} = C(P_1 + aP_2)$. Figure 2.6 shows that, for the weakly mixed I sub-class, we have $R^{\max}_{\text{sum-HK}} = C(P_1 + aP_2)$. On the other hand, Figure 2.2 shows that, for the mixed weak I sub-class, we also have $R^{\max}_{\text{sum-HK}} = C(P_1 + aP_2)$. Consequently, these two sub-classes can be merged together, as shown in Figure 2.20. Note that for the weakly mixed I sub-class, it is known that $C_{\text{sum}} = R^{\max}_{\text{sum-HK}} = C(P_1 + aP_2)$. However, for the mixed weak I sub-class, C_{sum} is unknown. Similar arguments follow for the region that corresponds to $R^{\max}_{\text{sum-HK}} = C(P_2 + bP_1)$. In the next chapter, we show that a similar approach can be used to find the maximum of any linear combination of R_1 and R_2.

2.4 Conclusion

This chapter studied the maximum achievable sum-rate of the HK scheme with Gaussian inputs for the class of weak interference. We fully characterized the maximum sum-rate without time sharing. We showed that when interference is weak, depending on the values of P_1 and P_2, five distinct power-splitting policies can maximize the achievable sum-rate. For each power splitting policy, the corresponding maximum sum-rate expression is explicitly determined. In the next chapter, we show that time sharing increases the maximum achievable sum-rate, and the corresponding increase can be expressed using the upper concave envelope of a function of P_1 and P_2.

Chapter 3

Boundary of the Han-Kobayashi Rate Region

In the previous chapter, we characterized the maximum HK sum-rate. In this chapter, we first generalize the results of the previous chapter and characterize the maximum of an arbitrary weighted sum-rate. Moreover, we show that the role of the time-sharing strategy in enlarging the achievable rate region can be described in terms of calculating the upper concave envelope of a function of P_1 and P_2.

3.1 Introduction

Recall that, for the two-user Gaussian Interference Channel (GIC), the Han-Kobayashi (HK) scheme has two arbitrary variables: power splitting and time sharing. In this scheme, each message is divided into public and private messages, and using two power-splitting variables, λ_1 and λ_2, the available power of each transmitter is shared between its public and private messages. Moreover, a time-sharing variable Q can exploit different strategies to enlarge the achievable rate region. However, the optimization problem involving all possible power splits and all time-sharing strategies that characterizes the boundary of this region is not well-understood. In particular, [13] states

> "even if we restrict ourselves to use only Gaussian codebooks, we need to consider all possible power splits and different time-sharing strategies among them. This is in general very complicated".

This chapter addresses this issue by investigating the HK scheme with Gaussian inputs and finding the optimal power splitting that results in boundary points of the achievable rate region.

The boundary of the HK rate region is known for only a few particular cases. When interference is strong, it is known that the rate HK region in which ($\lambda_1 = 0, \lambda_2 = 0$) characterizes the capacity region [6–8]. Moreover, this region is a polygon, and therefore, the entire boundary can be easily characterized.

There is a one-to-one correspondence between a closed set and its support function [46]. Let $\mathcal{G}$ denote the region achieved by the HK scheme with Gaussian inputs. For $\mathcal{G}$, the support function is a mapping from $\mathbb{R}_+^2$ to $\mathbb{R}_+^1$, defined by

$$h_{\mathcal{G}}(\mu) = \max\{R_1 + \mu R_2 | (R_1, R_2) \in \mathcal{G}\}. \tag{3.1}$$

Therefore, by characterizing the maximum of $R_1 + \mu R_2$, one can fully catheterize $\mathcal{G}$. However the maximum of $R_1 + \mu R_2$ is not known in general. For $\mu = 1$, the maximum sum-rate is known for only a few particular cases. For the few cases where the sum-capacity is known, it equals the maximum sum-rate of the HK scheme. Unfortunately, the sum-capacity is not known in general, but only for strong interference [7] and mixed interference [10]. For weak interference, the sum-capacity is an open problem and is known for only a small part of the weak interference class [10–12]. For weak interference, not only is the boundary of the HK rate region unknown, but its corresponding maximum sum-rate is also unknown [40–42]. This chapter fully characterizes the boundary of the HK scheme with Gaussian inputs, even when time sharing is used, a problem that has been unsolved for more than 30 years.

This chapter studies the HK scheme with "Gaussian" inputs. Note that the optimal distribution of the inputs is not known. In fact, for all cases where the capacity is known, it has been achieved using the HK scheme with "Gaussian" inputs. First, the full characterization of the achievable rate region is found, when no time sharing is used. It is shown that, when interference is weak, the optimal power splitting that achieves a boundary point is not unique and belongs to a set with a finite size that can be explicitly characterized. Moreover, we examine the role of the time-sharing variable Q and the Frequency Division (FD) technique in enlarging the achievable rate region.

The rest of this chapter is organized as follows. In Section 3.2, the existing results are reviewed. In particular, the difference between time sharing and time division is highlighted. In Section 3.3, the boundary of the HK rate region is studied for the two-user GIC with weak interference. This section, which demonstrates how optimization over power splitting and time sharing is performed, contains the main contributions of this chapter. Moreover, in this section, using upper concave envelope, we show how time sharing increases the achievable rate region. Finally, Section 3.4 concludes the chapter.

3.2 Preliminaries

In this chapter, the following notations are used. The notation $m \doteq n$ means n is the definition of m, and $C(x) \doteq \frac{1}{2}\log(1+x)$. Moreover, for non-negative numbers a, b, x such that $a \leq b$, $[x]_a^b \doteq \min\{\max\{x, a\}, b\}$. For a set Λ, $|\Lambda|$ shows the size of Λ. For a function $f : \mathbb{R}_+^2 \to \mathbb{R}_+^1$, $\mathcal{C}[f]$ represents the upper concave envelope of f, i.e. the smallest concave function that is bigger than f. Note that, by Caratheodory's theorem,

$$\mathcal{C}[f](P_1, P_2) = \sup_{\theta_i, \alpha_i \beta_i \in [0,1]} \sum_{i=1}^{3} \theta_i f\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}\Big), \tag{3.2}$$

subject to $\sum_{i=1}^3 \theta_i = \sum_{i=1}^3 \alpha_i = \sum_{i=1}^3 \beta_i = 1$.

In this chapter, we investigate the weak interference class, i.e., when $a < 1$ and $b < 1$. Recall that, for the two-user GIC, the HK scheme results in the best-known achievable rate region. As stated in the previous chapter, this region is described by [6, 44, 47]

$$\begin{aligned}
R_1 <& D_1 \doteq C\Big(\frac{P_1}{1+a\lambda_2 P_2}\Big),\\
R_2 <& D_2 \doteq C\Big(\frac{P_2}{1+b\lambda_1 P_1}\Big),\\
R_1+R_2 <& D_3^1 \doteq C\Big(\frac{P_1+a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big)+C\Big(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}\Big),\\
R_1+R_2 <& D_3^2 \doteq C\Big(\frac{P_2+b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big)+C\Big(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}\Big),\\
R_1+R_2 <& D_3^3 \doteq C\Big(\frac{\lambda_1 P_1+a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big)+C\Big(\frac{\lambda_2 P_2+b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big),\\
2R_1+R_2 <& D_4 \doteq C\Big(\frac{P_1+a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big)+C\Big(\frac{\lambda_1 P_1}{1+a\lambda_2 P_2}\Big)+C\Big(\frac{\lambda_2 P_2+b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big),\\
R_1+2R_2 <& D_5 \doteq C\Big(\frac{P_2+b\bar{\lambda}_1 P_1}{1+b\lambda_1 P_1}\Big)+C\Big(\frac{\lambda_2 P_2}{1+b\lambda_1 P_1}\Big)+C\Big(\frac{\lambda_1 P_1+a\bar{\lambda}_2 P_2}{1+a\lambda_2 P_2}\Big).
\end{aligned} \tag{3.3}$$

3.2.1 Time Sharing versus Time/Frequency Division

One of the contributions of Han and Kobayashi is the introduction of the time-sharing variable Q which can enlarge the achievable rate region. It is important to highlight that the role of the time-sharing variable Q is not necessarily equivalent to the convex hull operation of the FD technique [10, 12, 37].

Following [10], we define

$$D_3 \doteq \min\{D_3^1, D_3^2, D_3^2\}. \tag{3.4}$$

Let the vector

$$\mathbf{D}(P_1, P_2, \lambda_1, \lambda_2) \doteq (D_1, D_2, D_3, D_4, D_5)^t, \tag{3.5}$$

where D_is are defined in (3.3). The rate region $\mathcal{G}_0$ is defined as follows:

$$\mathcal{G}_0 = \{\mathbf{R} \in \mathbb{R}_+^2 | \mathbf{AR} \leq \mathbf{D}\}, \tag{3.6}$$

where $\mathbf{R} \doteq (R_1, R_2)^t$, and $\mathbf{A}$ is defined as

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix}^t.$$

$\mathcal{G}_0$ is a polytope which has at most 7 extreme points. In fact, $\mathcal{G}_0$ represents the region achieved by a fix power splitting (λ_1, λ_2). Observe that $(0,0)$, $(C(P_1), 0)$, and $(0, C(P_2))$ are three extreme points of $\mathcal{G}_0$. For this region, the maximum of $R_1 + \mu R_2$ is denoted by $R_{\mu\text{-HK}}$ and is expressed by

$$R_{\mu\text{-HK}}(P_1, P_2, \lambda_1, \lambda_2) \doteq \max_{R_1, R_2 \in \mathcal{G}_0} R_1 + \mu R_2. \tag{3.7}$$

We can enlarge the achievable rate region $\mathcal{G}_0$ using different techniques. For instance, define $\mathcal{G}_1$ as the union of the $\mathcal{G}_0(P_1, P_2, \lambda_1, \lambda_2)$, where the union is taken over all $\lambda_1, \lambda_2 \in [0,1]$, as explained in the following:

$$\mathcal{G}_1 \doteq \bigcup_{\lambda_1, \lambda_2 \in [0,1]} \mathcal{G}_0(P_1, P_2, \lambda_1, \lambda_2). \tag{3.8}$$

For this region, the maximum of $R_1 + \mu R_2$ is denoted by $R^{\max}_{\mu\text{-HK}}$, as given by the following expression:

$$R^{\max}_{\mu\text{-HK}} \doteq \max_{R_1, R_2 \in \mathcal{G}_1} R_1 + \mu R_2. \tag{3.9}$$

Note that we have

$$R^{\max}_{\mu\text{-HK}}(P_1, P_2) = \max_{\lambda_1, \lambda_2 \in [0,1]} R_{\mu\text{-HK}}(P_1, P_2, \lambda_1, \lambda_2). \tag{3.10}$$

One can enlarge $\mathcal{G}_0$ using the time-sharing variable Q. Define $\mathcal{G}_Q$ as

$$\mathcal{G}_Q = \{\mathbf{R} \in \mathbb{R}^2_+ | \mathbf{A}\mathbf{R} \leq \mathbf{D}_Q\}, \tag{3.11}$$

where $\mathbf{D}_Q \doteq \sum_{i=1}^{5} q_i \mathbf{D}\Big(\frac{\alpha_i P_1}{q_i}, \frac{\beta_i P_2}{q_i}, \lambda_{1i}, \lambda_{2i}\Big)$, and we have $\lambda_{1i}, \lambda_{2i}, \alpha_i, \beta_i, q_i \in [0,1]$, such that $\sum_{i=1}^{5} q_i = \sum_{i=1}^{5} \alpha_i = \sum_{i=1}^{5} \beta_i = 1$. It is proved that, using more than 5 $q_i s$ does not enlarge $\mathcal{G}_Q$ [48]. This scheme is called *Coded Time Sharing* (CTS) [37]. We denote the maximum of $R_1 + \mu R_2$ of the HK scheme with Gaussian inputs and with CTS by $R^{\text{max-Q}}_{\mu\text{-HK}}$, as expressed in the following:

$$R^{\text{max-Q}}_{\mu\text{-HK}} \doteq \max_{R_1, R_2 \in \mathcal{G}_Q} R_1 + \mu R_2. \tag{3.12}$$

Moreover, we can enlarge $\mathcal{G}_0$ by using the Time Division (TD) or FD technique. Define $\mathcal{G}_{FD}$ as

$$\mathcal{G}_{FD} = \Big\{\mathbf{R} | \mathbf{R} = \sum_{i=1}^{3} \theta_i \mathbf{R}_i, A\mathbf{R}_i \leq \mathbf{D}\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}, \lambda_{1i}, \lambda_{2i}\Big)\Big\}, \tag{3.13}$$

for $\mathbf{R}_i \in \mathbb{R}_+^2$ and $\lambda_{1i}, \lambda_{2i}, \alpha_i, \beta_i, \theta_i \in [0,1]$, such that $\sum_{i=1}^3 \theta_i = \sum_{i=1}^3 \alpha_i = \sum_{i=1}^3 \beta_i = 1$. Intuitively, in the FD scheme, the entire bandwidth is divided into 3 sub-bands, where the i^{th} sub-band has θ_i percentage of the bandwidth. The first transmitter allocates α_i percentage of its power to the i^{th} sub-band and the second transmitter allocates β_i percentage of its power to the i^{th} sub-band. Finally, $(\lambda_{1i}, \lambda_{2i})$ represents the power splitting used in the i^{th} sub-band. It is known that $\mathcal{G}_{FD}$ is a closed and convex region and increasing the number of sub-bands to more than 3 does not enlarge $\mathcal{G}_{FD}$ [10, 48]. We denote the maximum weighted sum-rate of the HK scheme with Gaussian inputs and with FD by $R_{\mu\text{-HK}}^{\text{max-FD}}$, as expressed in the following:

$$R_{\mu\text{-HK}}^{\text{max-FD}} \doteq \max_{R_1, R_2 \in \mathcal{G}_{\text{FD}}} R_1 + \mu R_2. \tag{3.14}$$

One can see that $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \mathcal{G}_{FD} \subseteq \mathcal{G}_Q$, and therefore, $R_{\mu\text{-HK}}(\lambda_1, \lambda_2) \leq R_{\mu\text{-HK}}^{\text{max}} \leq R_{\mu\text{-HK}}^{\text{max-FD}} \leq R_{\mu\text{-HK}}^{\text{max-Q}}$. However, for the weak interference class, [10] proves that CTS and FD result in the same achievable rate region, i.e., $\mathcal{G}_{FD} = \mathcal{G}_Q$. Therefore, we can conclude the following corollary:

Corollary 3.1. *For the two-user GIC with weak interference,*

$$R_{\mu\text{-HK}}^{\text{max-FD}} = R_{\mu\text{-HK}}^{\text{max-Q}}.$$

This corollary is used to find $R_{\mu\text{-HK}}^{\text{max-Q}}$. Solving the optimization problem (3.12) is complicated. However, in the next section, we solve (3.14) in two steps. In the first step, we optimize over $\lambda_{1j}, \lambda_{2j}$, for a fixed j. In the second step, we show that the optimization over θ_j, α_j, and β_j is equivalent to calculating the upper concave envelope with respect to (P_1, P_2).

3.3 Boundary of the HK Rate Region

This section characterizes the entire boundary of the HK rate region. The main results are given in the following two theorems. The first theorem shows the set of optimal power splittings. The second theorem discusses the role of time sharing in enlarging the achievable rate region.

3.3.1 Main Results

Theorem 3.1. *For the two-user GIC, when interference is weak, the maximum of $R_1 + \mu R_2$ achieved by the HK scheme with Gaussian inputs and without CTS is given by*

$$R^{\max}_{\mu\text{-HK}}(P_1, P_2) = \max_{\lambda_1,\lambda_2\in\Lambda_\mu} R_{\mu\text{-HK}}(\lambda_1, \lambda_2), \tag{3.15}$$

where Λ_μ is a finite set representing the optimal power splittings that maximize $R_1+\mu R_2$. More importantly, for a fixed μ, one can explicitly find all elements of Λ_μ.

Theorem 3.1 demonstrates that the optimal power splitting, and consequently, the maximum of $R_1 + \mu R_2$ can have up to $|\Lambda_\mu|$ distinct mathematical expressions, depending on the values of P_1 and P_2. In fact, this theorem partitions the weak interference class into $|\Lambda_\mu|$ sub-classes. For each sub-class, Theorem 3.1 demonstrates $R^{\max}_{\mu\text{-HK}}$ and the corresponding optimal power-splitting variables.

Note that according to (3.10), $R^{\max}_{\mu\text{-HK}}(P_1, P_2)$ is obtained by maximizing $R_{\mu\text{-HK}}(\lambda_1, \lambda_2)$ over all (λ_1, λ_2). Theorem 3.1 claims that one can restrict the search for optimal power-splitting variables to the finite set Λ_μ. We show that the set of optimal power splitting points can be partitioned into three categories of points: points that correspond to stationary points inside the feasible region, points that lie on the boundary of the feasible region, and points at which the function $R_{\mu\text{-HK}}(\lambda_1, \lambda_2)$ is non-differentiable. Before proving this theorem, we state our second result. The next theorem shows how CTS increases $R_1 + \mu R_2$.

Theorem 3.2. *For the two-user GIC, when interference is weak, the maximum of $R_1 + \mu R_2$ achieved by the HK scheme with Gaussian inputs and with CTS is given by*

$$R^{\max\text{-Q}}_{\mu\text{-HK}}(P_1, P_2) = \mathcal{C}[R^{\max}_{\mu\text{-HK}}](P_1, P_2). \tag{3.16}$$

Proof. When interference is weak, we have

$$
\begin{aligned}
R_{\mu\text{-HK}}^{\text{max-Q}} &\overset{(a)}{=} R_{\mu\text{-HK}}^{\text{max-FD}} \\
&= \max_{R_1, R_2 \in \mathcal{G}_{\text{FD}}} R_1 + \mu R_2 \\
&= \max_{\theta_i, \alpha_i \beta_i, \lambda_1^i, \lambda_2^i \in [0,1]} \sum_{i=1}^{3} \theta_i R_{\mu\text{-HK}}\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}, \lambda_1^i, \lambda_2^i\Big) \\
&= \max_{\theta_i, \alpha_i \beta_i \in [0,1]} \sum_{i=1}^{3} \theta_i \max_{\lambda_1^i, \lambda_2^i \in [0,1]} R_{\mu\text{-HK}}\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}, \lambda_1^i, \lambda_2^i\Big) \\
&= \max_{\theta_i, \alpha_i \beta_i \in [0,1]} \sum_{i=1}^{3} \theta_i R_{\mu\text{-HK}}^{\max}\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}\Big) \\
&\overset{(b)}{=} \mathcal{C}[R_{\mu\text{-HK}}^{\max}](P_1, P_2),
\end{aligned}
\tag{3.17}
$$

where (a) is valid by Corollary 3.1 and (b) is valid by (3.2). □

Theorem 3.2 shows that when CTS is used, the maximum of $R_1 + \mu R_2$ increases from $R_{\mu\text{-HK}}^{\max}(P_1, P_2)$ to $\mathcal{C}[R_{\mu\text{-HK}}^{\max}](P_1, P_2)$. Note that, by the definition of the upper concave envelop, we have $R_{\mu\text{-HK}}^{\max}(P_1, P_2) \leq \mathcal{C}[R_{\mu\text{-HK}}^{\max}](P_1, P_2)$. Moreover, this theorem clarifies the role of time sharing in increasing the achievable rate region. For instance, if $R_{\mu\text{-HK}}^{\max}(P_1, P_2)$ is concave, then time sharing does not increase it. In fact, for mixed and strong interference, the achievable sum-rate of the HK scheme without time sharing is a concave function of (P_1, P_2), and therefore, time sharing does not increase it. However, when interference is weak, $R_{\mu\text{-HK}}^{\max}(P_1, P_2)$ is not concave and time sharing can be useful.

In the following, we discus two interesting properties of the HK achievable rate region. We show that similar to the achievable rate region of the multiple access channel, which corresponds to a pentagon, the achievable rate region of the HK scheme with no time sharing is a polygon with seven extreme points. These properties are used to prove Theorem 3.1.

3.3.2 Properties of the HK Rate Region

To prove Theorem 3.1, we first explore some properties of the HK rate region, as stated in the following lemmas.

Lemma 3.1. *For the HK rate region defined in (3.3), we have*

$$D_4 + D_5 = D_3^1 + D_3^2 + D_3^3, \tag{3.18}$$

$$D_4 + D_5 \geq 3D_3, \tag{3.19}$$

$$D_1 + D_3^3 \geq D_4, \tag{3.20}$$

$$D_2 + D_3^3 \geq D_5, \tag{3.21}$$

$$D_1 + D_5 \geq 2D_3, \tag{3.22}$$

$$D_2 + D_4 \geq 2D_3, \tag{3.23}$$

$$D_3^2 + D_1 \geq D_4 \qquad \text{if } a < 1, ab < 1, \tag{3.24}$$

$$D_3^2 + D_2 \geq D_5 \qquad \text{if } a < 1, ab < 1, \tag{3.25}$$

$$D_3^1 + D_1 \geq D_4 \qquad \text{if } b < 1, ab < 1, \tag{3.26}$$

$$D_3^1 + D_2 \geq D_5 \qquad \text{if } b < 1, ab < 1, \tag{3.27}$$

$$D_3 + D_1 \geq D_4 \qquad \text{if } b < 1, a < 1, \tag{3.28}$$

$$D_3 + D_2 \geq D_5 \qquad \text{if } b < 1, a < 1, \tag{3.29}$$

$$D_2 + D_1 \geq D_3 \qquad \text{if } a < 1, b < 1, \tag{3.30}$$

$$D_3^1 = D_3^2 \Leftrightarrow \lambda_1 = (1-c)\lambda_2 + c, \tag{3.31}$$

$$D_3^2 = D_3^3 \Leftrightarrow \lambda_1 = \tilde{\lambda}_1 \text{ or } \lambda_2 = 1, \tag{3.32}$$

$$D_3^3 = D_3^1 \Leftrightarrow \lambda_2 = \tilde{\lambda}_2 \text{ or } \lambda_1 = 1, \tag{3.33}$$

where $c \doteq \frac{P_1(1-b)-P_2(1-a)}{P_1(1-b+P_2(1-ab))}$ *and* $(\tilde{\lambda}_1, \tilde{\lambda}_2) \doteq (ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2})$.

Proof. The proof is straightforward. In fact, (3.18) is validated by direct calculation, and (3.19) is the direct consequence of (3.18). Note that

$$\begin{aligned} 3D_3 =& 3\min\{D_3^1, D_3^2, D_3^3\} \\ \leq& D_3^1 + D_3^2 + D_3^3 \\ =& D_4 + D_5. \end{aligned} \tag{3.34}$$

To prove (3.20), we calculate $D_4 - D_1 - D_3^3$.

$$
\begin{aligned}
D_4 - D_1 - D_3^3 \overset{(a)}{=} & C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big) \\
& - C\Big(\frac{\lambda_1 P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) - C\Big(\frac{P_1}{1 + a\lambda_2 P_2}\Big) \\
\overset{(b)}{=} & C\Big(\frac{\bar{\lambda}_1 P_1}{1 + \lambda_1 P_1 + aP_2}\Big) - C\Big(\frac{\bar{\lambda}_1 P_1}{1 + \lambda_1 P_1 + a\lambda_2 P_2}\Big) \\
\overset{(c)}{\leq} & 0, \qquad (3.35)
\end{aligned}
$$

where (a) is valid by (3.3), (b) is valid by Lemma 2.2 of the previous chapter, and (c) is valid because $\lambda_2 \leq 1$. (3.21) can be proved similarly.

To prove (3.22), note that

$$
\begin{aligned}
D_1 + D_5 \overset{(a)}{=} & D_1 + D_3^1 + D_3^2 + D_3^3 - D_4 \\
\overset{(b)}{\geq} & D_3^1 + D_3^2 \\
\overset{(c)}{\geq} & 2D_3, \qquad (3.36)
\end{aligned}
$$

where (a), (b), and (c) are valid by (3.18), (3.20), and (3.4), respectively. (3.23) can be proved similar to (3.22).

To prove (3.24), we directly calculate $D_3^2 + D_1 - D_4$, as follows:

$$
\begin{aligned}
D_3^2 + D_1 - D_4 \overset{(a)}{=} & C\Big(\frac{P_1}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{P_2 + b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1}\Big) \\
& - C\Big(\frac{P_1 + a\bar{\lambda}_2 P_2}{1 + a\lambda_2 P_2}\Big) - C\Big(\frac{\lambda_2 P_2 + b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1}\Big) \\
\overset{(b)}{=} & - C\Big(\frac{a\bar{\lambda}_2 P_2}{1 + P_1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\bar{\lambda}_2 P_2}{1 + bP_1 + \lambda_2 P_2}\Big) \\
= & - C\Big(\frac{\bar{\lambda}_2 P_2}{\frac{1}{a} + \frac{P_1}{a} + \lambda_2 P_2}\Big) + C\Big(\frac{\bar{\lambda}_2 P_2}{1 + bP_1 + \lambda_2 P_2}\Big) \\
\overset{(c)}{\geq} & 0, \qquad (3.37)
\end{aligned}
$$

where (a) is valid by (3.3), (b) is valid by Lemma 2.2 of the previous chapter, and (c) is valid if $a + abP_1 \leq 1 + P_1$, which is satisfied because we have assumed that $a < 1$ and $ab < 1$. (3.25) can be proved similarly.

To prove (3.26), we calculate $D_3^1 + D_1 - D_4$, as follows:

$$
\begin{aligned}
D_3^1 + D_1 - D_4 &\overset{(a)}{=} C\Big(\frac{P_1}{1 + a\lambda_2 P_2}\Big) + C\Big(\frac{\lambda_2 P_2}{1 + b\lambda_1 P_1}\Big) \\
&\quad - C\Big(\frac{\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big) - C\Big(\frac{\lambda_2 P_2 + b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1}\Big) \\
&\overset{(b)}{=} + C\Big(\frac{\bar{\lambda}_1 P_1}{1 + \lambda_1 P_1 + a\lambda_2 P_2}\Big) - C\Big(\frac{b\bar{\lambda}_1 P_1}{1 + b\lambda_1 P_1 + \lambda_2 P_2}\Big) \\
&= + C\Big(\frac{\bar{\lambda}_1 P_1}{1 + \lambda_1 P_1 + a\lambda_2 P_2}\Big) - C\Big(\frac{\bar{\lambda}_1 P_1}{\frac{1}{b} + \lambda_1 P_1 + \lambda_2 \frac{P_2}{b}}\Big) \\
&\overset{(c)}{\geq} 0,
\end{aligned} \tag{3.38}
$$

where (a) is valid by (3.3), (b) is valid by Lemma 2.2 of the previous chapter, and (c) is valid if $b + abP_2 \leq 1 + P_2$, which is satisfied because we have assumed that $b < 1$ and $ab < 1$. (3.27) can be proved similarly.

To prove (3.28), note that when $a < 1$ and $b < 1$, we have

$$
\begin{aligned}
D_3^1 + D_1 &\overset{(a)}{\geq} D_4, \\
D_3^2 + D_1 &\overset{(b)}{\geq} D_4, \\
D_3^3 + D_1 &\overset{(c)}{\geq} D_4,
\end{aligned}
$$

where (a), (b), and (c) are valid by (3.26), (3.24), and (3.20), respectively. Therefore, (3.28) is valid. (3.29) can be proved similarly.

To prove (3.30), we can write

$$
\begin{aligned}
D_1 + D_2 &\overset{(a)}{\geq} D_4 + D_5 - 2D_3 \\
&\overset{(b)}{=} D_3^1 + D_3^2 + D_3^3 - 2D_3 \\
&\geq D_3,
\end{aligned} \tag{3.39}
$$

where (a) is valid by (3.28) and (3.29) and (b) is valid by (3.18).

Observe that (3.31), (3.32), and (3.33) are valid by Lemma 2.3, proved in the previous chapter. This completes the proof. □

Lemma 3.2. *For the two-user GIC with weak interference, the HK rate region $\mathcal{G}_0$, characterized in (3.3), is a polygon with exactly seven extreme points if $(\lambda_1, \lambda_2) \notin \{(1,1), (\tilde{\lambda}_1, \tilde{\lambda}_2)\}$. Moreover, if $(\lambda_1, \lambda_2) = (1,1)$, $\mathcal{G}_0$ has four extreme points, and if $(\lambda_1, \lambda_2) = (\tilde{\lambda}_1, \tilde{\lambda}_2)$, $\mathcal{G}_0$ has six extreme points.*

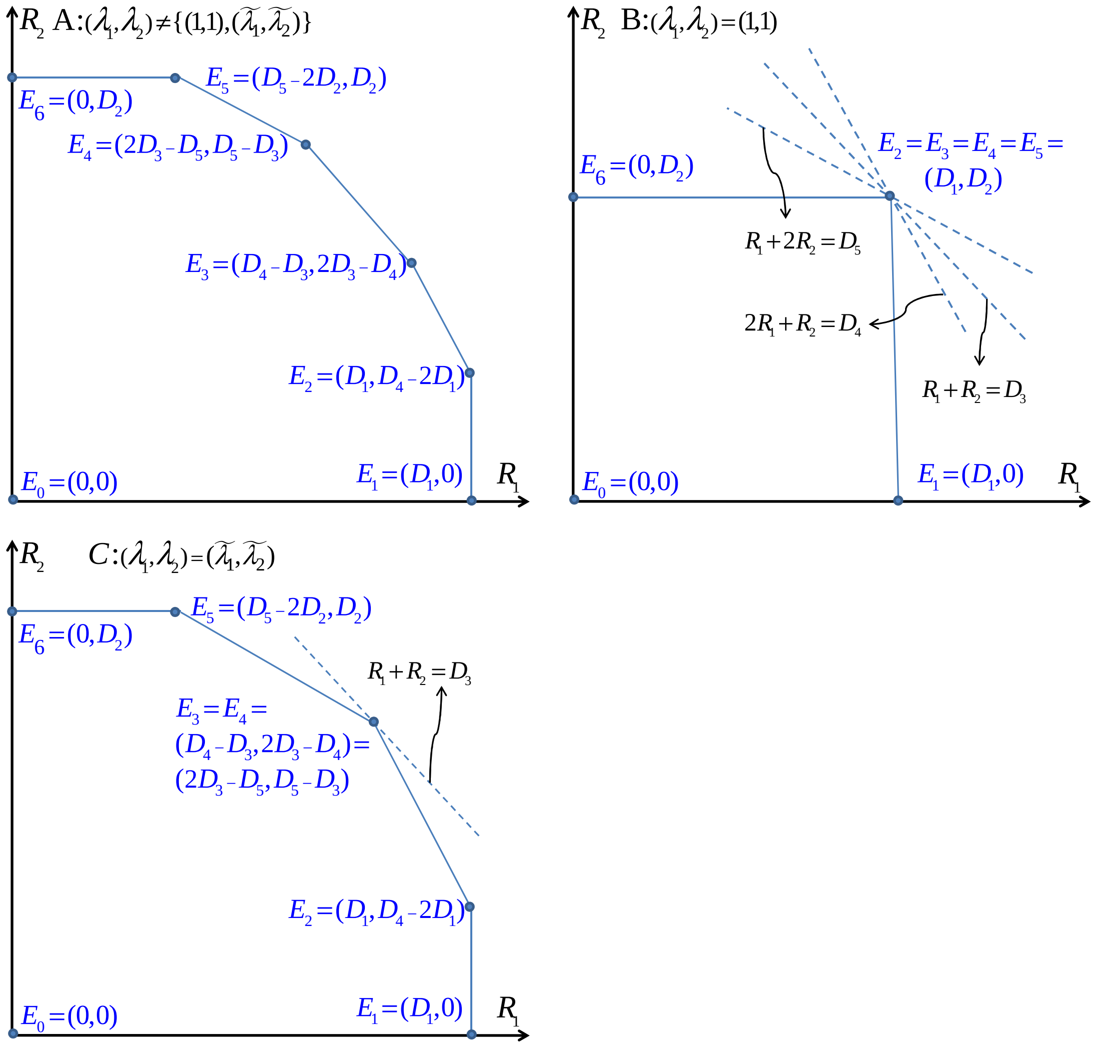

Figure 3.1: The achievable rate region $\mathcal{G}_0$ and its extreme points.

Proof. The proof can be established using Lemma 3.1. For instance, $\mathcal{G}_0$ can have six extreme points if and only if $D_4 + D_5 \leq 3D_3$. However, according to (3.19), $D_4 + D_5 \geq 3D_3$. Therefore, $\mathcal{G}_0$ can have six extreme points if and only if

$$\begin{aligned} D_4 + D_5 =& D_3^1 + D_3^2 + D_3^3 \\ =& 3D_3. \end{aligned} \tag{3.40}$$

On the other hand, $D_3^1 + D_3^2 + D_3^3 = 3D_3$ if and only if

$$D_3^1 = D_3^2 = D_3^3. \tag{3.41}$$

Note that, according to (3.31), (3.32), and (3.33), one can satisfy (3.41) if and only if

$$(\lambda_1, \lambda_2) \in \{(1,1), (\tilde{\lambda}_1, \tilde{\lambda}_2)\}. \tag{3.42}$$

Therefore, if $(\lambda_1, \lambda_2) \notin \{(1,1), (\tilde{\lambda}_1, \tilde{\lambda}_2)\}$, then $\mathcal{G}_0$ cannot have six extreme points. One can check that if $(\lambda_1, \lambda_2) = (1,1)$, then $\mathcal{G}_0$ has four extreme points, as shown in Figure 3.1. Similarly, if $(\lambda_1, \lambda_2) = (\tilde{\lambda}_1, \tilde{\lambda}_2)$, then $\mathcal{G}_0$ has six extreme points, as shown in Figure 3.1.

Following a similar line of arguments, one can see that $\mathcal{G}_0$ cannot have five extreme points. Moreover, $\mathcal{G}_0$ can have four extreme points if and only if $(\lambda_1, \lambda_2) = (1,1)$. This completes the proof. □

The properties of the HK rate region can be used to describe the optimization problem that corresponds to the maximum of an arbitrary weighted sum-rate. In the next section, we use linear programming tools to describe that optimization problem.

3.3.3 The Optimization Problem Corresponding to the Maximum Weighted HK Sum-rate

To prove Theorem 3.1, we express an optimization problem that characterizes the maximum of $R_1 + \mu R_2$, as explained in the following theorem.

Theorem 3.3. *For the two-user GIC with weak interference, $R^{\max}_{\mu\text{-HK}}$ is given by the following optimization problem:*

$$R^{\max}_{\mu\text{-HK}} = \begin{cases} \max\limits_{\lambda_1,\lambda_2\in[0,1]} D_1 + \mu(D_4 - 2D_1) & \text{if } 0 < \mu \le \frac{1}{2} \\ \max\limits_{\lambda_1,\lambda_2\in[0,1]} D_4 - D_3 + \mu(2D_3 - D_4) & \text{if } \frac{1}{2} < \mu \le 1 \\ \max\limits_{\lambda_1,\lambda_2\in[0,1]} 2D_3 - D_5 + \mu(D_5 - D_3) & \text{if } 1 < \mu \le 2 \\ \max\limits_{\lambda_1,\lambda_2\in[0,1]} D_5 - 2D_2 + \mu D_2 & \text{if } 2 < \mu, \end{cases}$$

where D_1, D_2, D_3, D_4, and D_5 are defined in (3.3).

Proof. Assume that $(\lambda_1, \lambda_2) \notin \{(1,1), (\tilde{\lambda}_1, \tilde{\lambda}_2)\}$. By Lemma 3.2, we know that the feasible region $\mathcal{G}_0$ has seven extreme points, as shown in Figure 3.2. Since the objective function $R_1 + \mu R_2$ is a linear function, it achieves its maximum at one of the extreme points of the feasible region. In fact, $R^{\max}_{\mu\text{-HK}}$ is the solution of the optimization problem (3.9), and $R_1 + \mu R_2$ obtains its maximum at E_2, E_3, E_4, and E_5, if $\mu \le 0.5$, $0.5 < \mu \le 1$, $1 < \mu \le 2$,

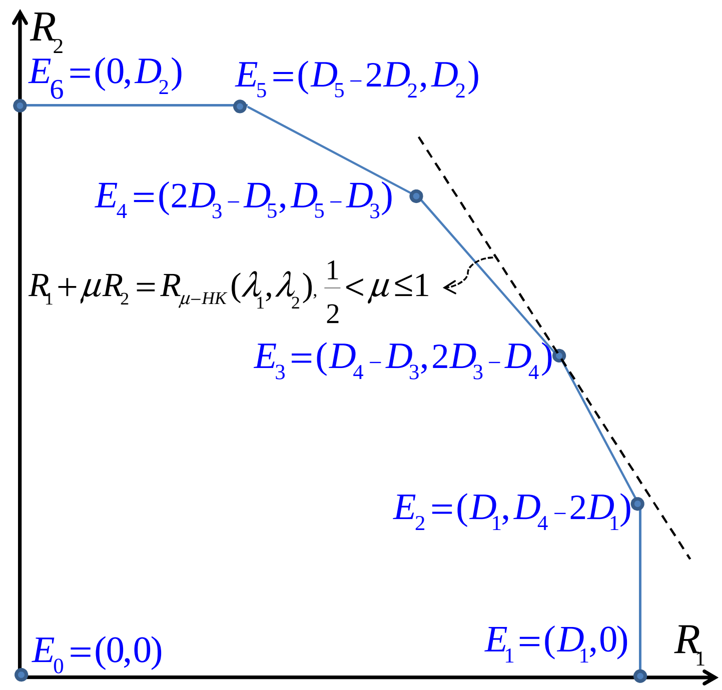

Figure 3.2: Depending on the value of μ, $R_1 + \mu R_2$ is maximized at one of the extreme points.

and $2 < \mu$, respectively, as stated in Theorem 3.3. Moreover, if $(\lambda_1, \lambda_2) = (1, 1)$, one can show that $E_2 = E_3 = E_4 = E_5$. Similarly, if $(\lambda_1, \lambda_2) = (ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2})$, then $E_3 = E_4$. Consequently, for these two cases, optimization of Theorem 3.3 holds. This completes the proof. □

To prove Theorem 3.1, we need to solve four optimization problems, the problems given in Theorem 3.3 for different values of μ. According to the interior extremum theorem, the global maximum of a function f over a feasible region $\mathcal{A}$ is achieved at one of the following points: a stationary point or a boundary point or a point at which the function f is non-differentiable [45, 46]. Note that the feasible region of the optimization problems of Theorem 3.3 is $\lambda_1, \lambda_2 \in [0, 1]$. Therefore, the boundary of the feasible region, denoted by $\mathcal{B}$, is the boundary of a unit square which can be represented as $\mathcal{B} = \mathcal{B}_1 \cup$

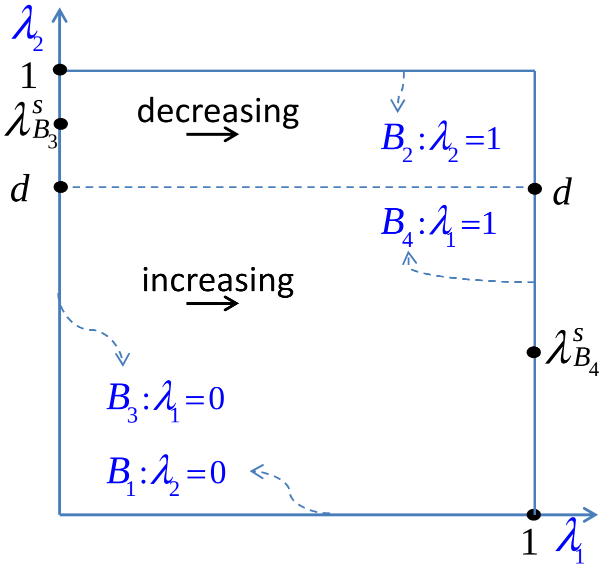

Figure 3.3: Behavior of $R_\mu = R_1 + \mu R_2 = D_1 + \mu(D_4 - 2D_1)$ over the feasible region and the six optimal power splittings that maximize R_μ.

$\mathcal{B}_2 \cup \mathcal{B}_2 \cup \mathcal{B}_4$, where

$$\mathcal{B}_1 \doteq \{(\lambda_1, 0) : 0 \le \lambda_1 \le 1\},$$

$$\mathcal{B}_2 \doteq \{(\lambda_1, 1) : 0 \le \lambda_1 \le 1\},$$

$$\mathcal{B}_3 \doteq \{(0, \lambda_2) : 0 \le \lambda_2 \le 1\},$$

$$\mathcal{B}_4 \doteq \{(1, \lambda_2) : 0 \le \lambda_2 \le 1\}.$$

Relying on this idea, we solve the optimization problems corresponding to $\mu \le 0.5$, and $0.5 < \mu \le 1$ in two separate lemmas. The other two optimization problems of Theorem 3.3, corresponding to $1 < \mu \le 2$ and $2 < \mu$, can be solved similarly. Lemma 3.3 investigates the case in which $\mu \le \frac{1}{2}$. This lemma proves Theorem 3.1, for $\mu \le \frac{1}{2}$. It shows that for this range of μ, $|\Lambda_\mu| \le 6$.

Lemma 3.3. *If $R^\star_\mu$ is the optimal solution of the optimization problem*

$$R^\star_\mu = \max_{\lambda_1, \lambda_2 \in [0,1]} D_1 + \mu(D_4 - 2D_1), \tag{3.43}$$

then $R^\star_\mu = \max_{\lambda_1, \lambda_2 \in \Lambda_\mu} R_{\mu\text{-HK}}(\lambda_1, \lambda_2)$, where Λ_μ is given by

$$\Lambda_\mu = \{(0, d), (0, 1), (1, 0), (1, d), (0, [\lambda^s_{\mathcal{B}_3}]^1_d), (1, [\lambda^s_{\mathcal{B}_4}]^d_0)\}, \tag{3.44}$$

where $d \doteq [\frac{1-b}{abP_2}]^1_0$. Moreover, $\lambda^s_{\mathcal{B}_3}$ and $\lambda^s_{\mathcal{B}_4}$, which are stationary points corresponding to local maximums over $\mathcal{B}_3$ and $\mathcal{B}_4$, respectively, can be obtained by solving the following *equations:*

$$\frac{\partial\Big(D_1(0,\lambda_2)+\mu(D_4(0,\lambda_2)-2D_1(0,\lambda_2))\Big)}{\partial\lambda_2}=0, \tag{3.45}$$

$$\frac{\partial\Big(D_1(1,\lambda_2)+\mu(D_4(1,\lambda_2)-2D_1(1,\lambda_2))\Big)}{\partial\lambda_2}=0. \tag{3.46}$$

Proof. Note that we have

$$\begin{aligned}\frac{\partial\big(D_1+\mu(D_4-2D_1)\big)}{\partial\lambda_1}&=\mu\frac{\partial D_4}{\partial\lambda_1}\\&=\frac{-\mu P_1(ab\lambda_2P_2+b-1)}{(1+b\lambda_1P_1)(1+\lambda_1P_1+a\lambda_2P_2)}.\end{aligned}$$

This shows that, with respect to λ_1, $D_1+\mu(D_4-2D_1)$ is increasing if $\lambda_2\le d\doteq[\frac{1-b}{abP_2}]_0^1$, and is decreasing if $\lambda_2>d$ (see Figure 3.3). Therefore, the optimal $\lambda_1^\star$ belongs to $\{0,1\}$. If $\lambda_1^\star=0$, then by taking derivative with respect to λ_2, one can show that $\lambda_2^\star\in\{d,1,[\lambda_{\mathcal{B}_3}^s]_d^1\}$. Similarly, if $\lambda_1^\star=1$, then $\lambda_2^\star\in\{0,d,[\lambda_{\mathcal{B}_4}^s]_d^1\}$, as shown in Figure 3.3. One can check that equations (3.45) and (3.46) can have at most one solution in $[0,1]$ that corresponds to a local maximum. This completes the proof. $\square$

Solving the optimization problem of Theorem 3.3 corresponding to $\frac{1}{2}<\mu\le1$ is more challenging because the function $D_3=\min\{D_3^1,D_3^2,D_3^3\}$ is not differentiable over the feasible region. However, we use properties (3.31-3.33) and partition the feasible region into up to three parts, namely $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$, where

$$\mathcal{I}_j\doteq\{(\lambda_1,\lambda_2):D_3=D_3^j\}. \tag{3.47}$$

Note that D_3 is differentiable within each partition. Figure 3.4 shows how this partitioning is performed, depending on the values of P_1 and P_2. D_3 can be non-differentiable only at the boundary between two adjacent partitions. As shown in Figure 3.4, adjacent partitions are separated by black solid line segments. These three lines are expressed by

$$\mathcal{N}_1\doteq\{(\lambda_1,\lambda_2):\lambda_1=(1-c)\lambda_2+c\}, \tag{3.48}$$

$$\mathcal{N}_2\doteq\{(\lambda_1,\lambda_2):\lambda_1=\tilde{\lambda}_1\}, \tag{3.49}$$

$$\mathcal{N}_3\doteq\{(\lambda_1,\lambda_2):\lambda_2=\tilde{\lambda}_2\}. \tag{3.50}$$

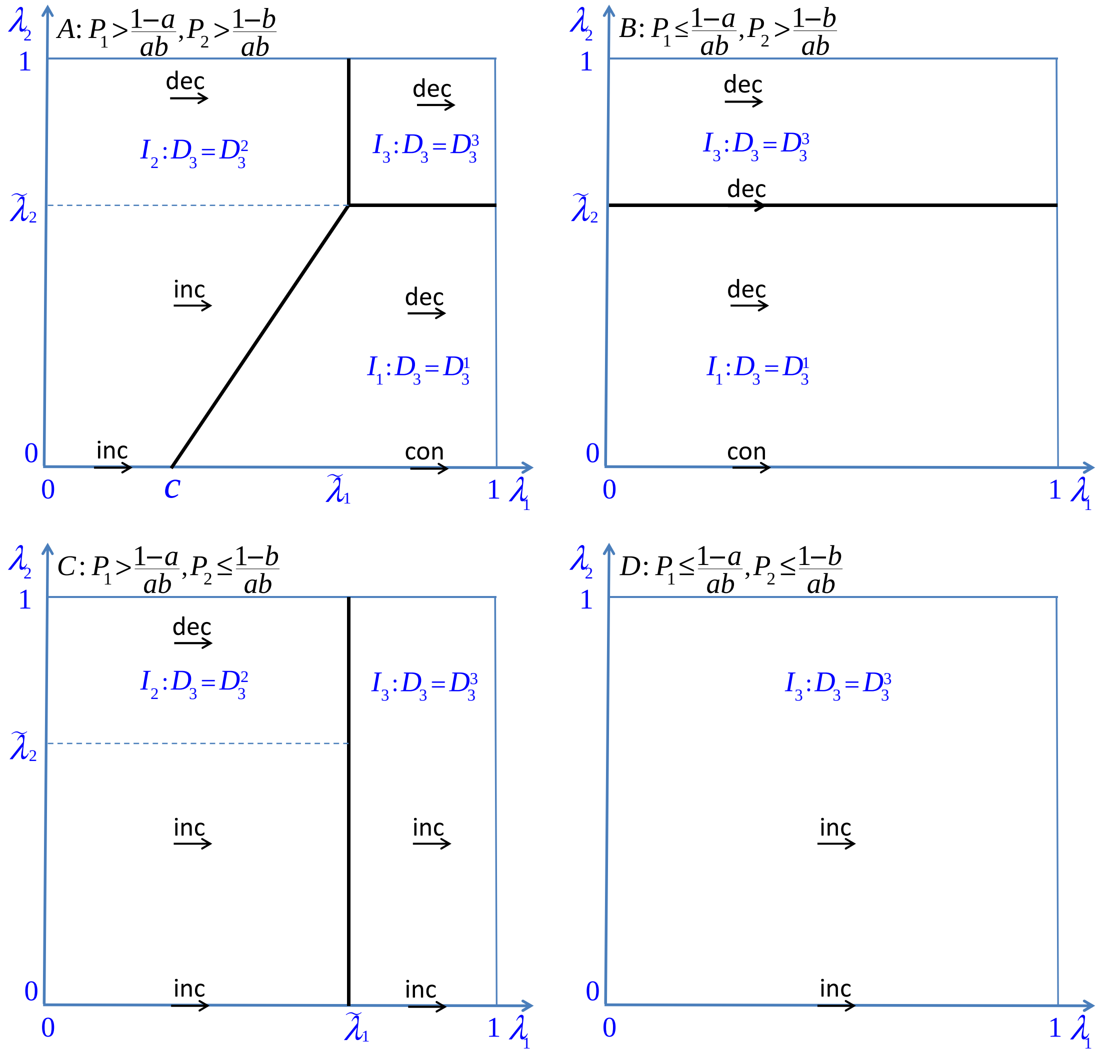

Figure 3.4: Behavior of D_3 over the feasible region.

Moreover, we explore the behavior of D_3 with respect to λ_1. Note that we have

$$\frac{\partial D_3^1(\lambda_1, \lambda_2)}{\partial \lambda_1} = \frac{-b\lambda_2 P_2 P_1}{(1 + bP_1\lambda_1)(1 + bP_1\lambda_1 + \lambda_2 P_2)}, \tag{3.51}$$

$$\frac{\partial D_3^2(\lambda_1, \lambda_2)}{\partial \lambda_1} = \frac{P_1(1 - b - ab\lambda_2 P_2)}{(1 + bP_1\lambda_1)(1 + aP_2\lambda_2 + \lambda_1 P_1)}, \tag{3.52}$$

$$\frac{\partial D_3^3(\lambda_1, \lambda_2)}{\partial \lambda_1} = \frac{P_1(1 - b - abP_2)}{(1 + bP_1\lambda_1)(1 + aP_2 + P_1\lambda_1)}. \tag{3.53}$$

Therefore, for each partition, we can check if D_3 is increasing (inc), decreasing (dec), or constant (con) with respect to λ_1, as shown in Figure 3.4. Relying on this perspective, we prove Theorem 3.1 for $0.5 < \mu \le 1$, in the following lemma:

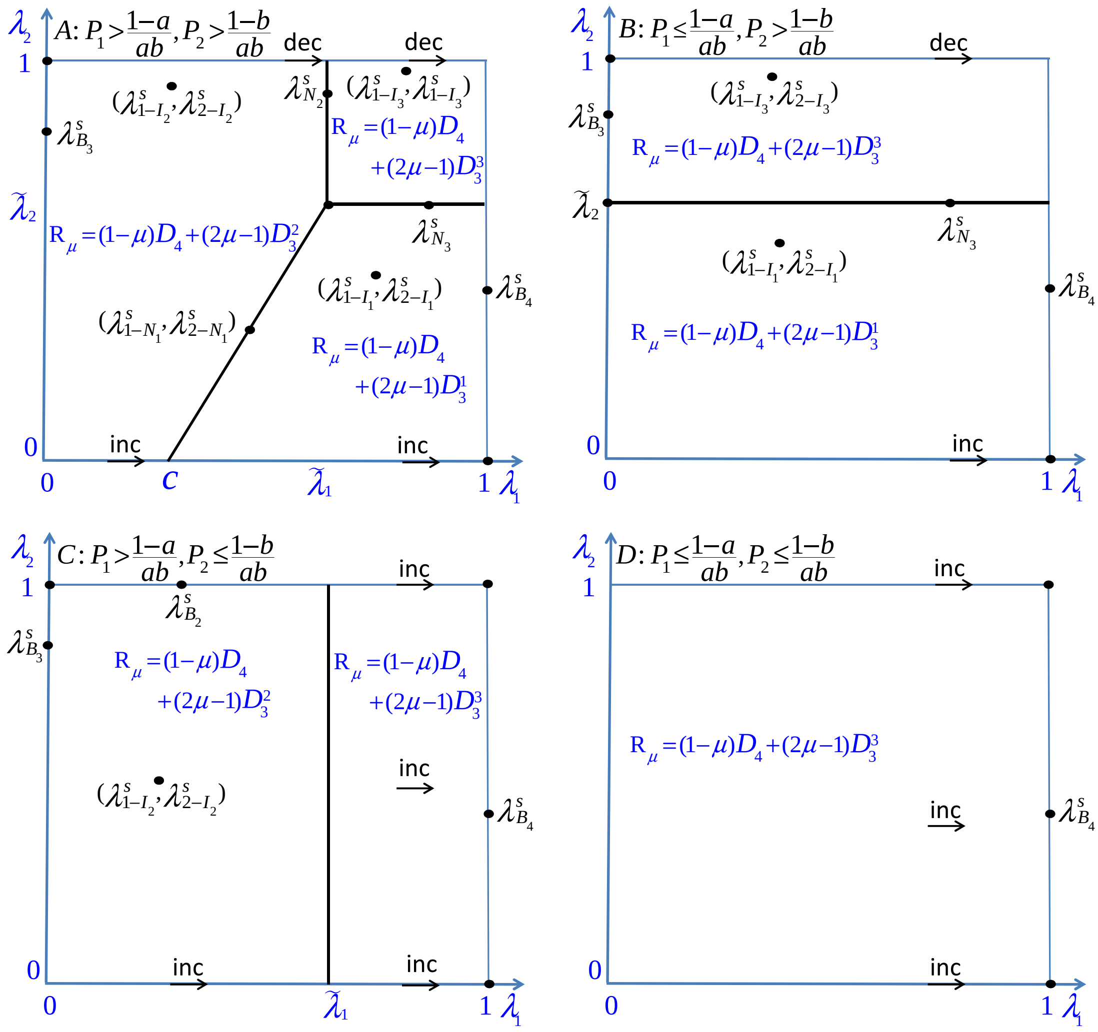

Figure 3.5: Behavior of $R_\mu = R_1 + \mu R_2 = (1-\mu)D_4 + (2\mu-1)D_3$ over the feasible region: the optimal power splittings that maximize R_μ are shown by solid black dots.

Lemma 3.4. *If $R^\star_\mu$ is the optimal value of the optimization problem*

$$\begin{aligned} R^\star_\mu &= \max_{\lambda_1,\lambda_2\in[0,1]} D_4 - D_3 + \mu(2D_3 - D_4) \\ &= \max_{\lambda_1,\lambda_2\in[0,1]} (1-\mu)D_4 + (2\mu-1)D_3, \end{aligned} \tag{3.54}$$

then $R^\star_\mu = \max_{\lambda_1,\lambda_2\in\Lambda_\mu} R_{\mu\text{-HK}}(\lambda_1,\lambda_2)$, where Λ_μ is given by

$$\begin{aligned} \Lambda_\mu = \Big\{&(0,1),(0,[\lambda^s_{\mathcal{B}_3}]_0^1),(1,0),(1,[\lambda^s_{\mathcal{B}_4}]_0^1),(1,1),([\lambda^s_{\mathcal{B}_2}]_0^1,1) \\ &([\tilde\lambda_1]_0^1,[\lambda^s_{\mathcal{N}_2}]_0^1),([\tilde\lambda_1]_0^1,[\tilde\lambda_2]_0^1),([\lambda^s_{\mathcal{N}_3}]_0^1,[\tilde\lambda_2]_0^1), \\ &([\lambda^s_{1-\mathcal{N}_1}]_0^1,[\lambda^s_{2-\mathcal{N}_1}]_0^1),([\lambda_{1-\mathcal{I}_1}]_0^1,[\lambda_{2-\mathcal{I}_1}]_0^1), \\ &([\lambda_{1-\mathcal{I}_2}]_0^1,[\lambda_{2-\mathcal{I}_2}]_0^1),([\lambda_{1-\mathcal{I}_3}]_0^1,[\lambda_{2-\mathcal{I}_3}]_0^1)\Big\}. \end{aligned} \tag{3.55}$$

Figure 3.5 demonstrates all optimal power splittings of Λ_μ, and Table 3.1 provides their corresponding definitions.

Proof. Once the feasible region is partitioned, we need to solve an optimization problem over each partition. For instance, if $P_1 > \frac{1-a}{ab}$ and $P_2\frac{1-b}{ab}$, we solve three optimization problems, corresponding to three feasible regions, namely $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$, as shown in Figure 3.5.

The optimal power splitting is either a stationary point inside one of the partitions or a point on the boundary of the partitions. To find the stationary point inside $\mathcal{I}_j$, we should solve the equation

$$\begin{aligned} \nabla\big(R^\star_\mu\big) &= (0,0) \\ \Leftrightarrow \nabla\big((1-\mu)D_4 + (2\mu-1)D_3\big) &= (0,0) \\ \Leftrightarrow \frac{\partial(1-\mu)D_4(\lambda_1,\lambda_2) + (2\mu-1)D_3^j(\lambda_1,\lambda_2)}{\partial\lambda_1} &= 0 \text{ and} \\ \frac{\partial(1-\mu)D_4(\lambda_1,\lambda_2) + (2\mu-1)D_3^j(\lambda_1,\lambda_2)}{\partial\lambda_2} &= 0. \end{aligned} \tag{3.56}$$

In fact, $(\lambda_{1-\mathcal{I}_j},\lambda_{2-\mathcal{I}_j})$ is the solution of (3.56) that corresponds to a local maximum inside $\mathcal{I}_j$. Other elements of Λ_μ belong to the boundary of partitions. Note that the boundary of all partitions are line segments. Therefore, an optimal power splitting on the boundary of partitions is either a vertex of the boundary or a local maximum over the boundary which has a derivative of zero along the direction of the boundary. Table 3.1 shows all

Power splitting	Given by the solution of
$\lambda^s_{\mathcal{B}_2}$	$\frac{\partial\big((1-\mu)D_4(\lambda_1,1)+(2\mu-1)D_3(\lambda_1,1)\big)}{\partial\lambda_1}=0$
$\lambda^s_{\mathcal{B}_3}$	$\frac{\partial\big((1-\mu)D_4(0,\lambda_2)+(2\mu-1)D_3(0,\lambda_2)\big)}{\partial\lambda_2}=0$
$\lambda^s_{\mathcal{B}_4}$	$\frac{\partial\big((1-\mu)D_4(1,\lambda_2)+(2\mu-1)D_3(1,\lambda_2)\big)}{\partial\lambda_2}=0$
$\lambda^s_{\mathcal{N}_2}$	$\frac{\partial\big((1-\mu)D_4(\tilde{\lambda}_1,\lambda_2)+(2\mu-1)D_3(\tilde{\lambda}_1,\lambda_2)\big)}{\partial\lambda_2}=0$
$\lambda^s_{\mathcal{N}_3}$	$\frac{\partial\big((1-\mu)D_4(\lambda_1,\tilde{\lambda}_2)+(2\mu-1)D_3(\lambda_1,\tilde{\lambda}_2)\big)}{\partial\lambda_1}=0$
$\lambda^s_{2-\mathcal{N}_1}$	$\frac{\partial\big((1-\mu)D_4((1-c)\lambda_2+c,\lambda_2)+(2\mu-1)D_3((1-c)\lambda_2+c,\lambda_2)\big)}{\partial\lambda_2}=0$
$\lambda^s_{1-\mathcal{N}_1}$	$\lambda^s_{1-\mathcal{N}_1}=(1-c)\lambda^s_{2-\mathcal{N}_1}+c$
$\lambda_{1-\mathcal{I}_j}$	$\frac{\partial(1-\mu)D_4(\lambda_1,\lambda_2)+(2\mu-1)D_3^j(\lambda_1,\lambda_2)}{\partial\lambda_1}=0, j\in\{1,2,3\}$
$\lambda_{2-\mathcal{I}_j}$	$\frac{\partial(1-\mu)D_4(\lambda_1,\lambda_2)+(2\mu-1)D_3^j(\lambda_1,\lambda_2)}{\partial\lambda_2}=0, j\in\{1,2,3\}$

Table 3.1: The optimal power splittings.

the optimal power splittings. For instance, $(\lambda^s_{\mathcal{B}_2}, 1)$ is a point on the boundary section $\mathcal{B}_2$. The point $(\tilde{\lambda}_1, \lambda^s_{\mathcal{N}_2})$ lies on non-differentiable points $\mathcal{N}_2$. Finally, the point $(\lambda_{1-\mathcal{I}_j}, \lambda_{2-\mathcal{I}_j})$ is a stationary point inside $\mathcal{I}_j$. $\square$

In the next section, we show that Theorems 3.1 and 3.2 can be used to rederive several known results about the HK achievable rate region.

3.3.4 Rederiving Existing Results

Using Theorems 3.1 and 3.2, we prove some known results. First, note that the set Λ_μ, given in (3.55), leads to a full characterization of $R_1+\mu R_2$ for $0.5<\mu\leq 1$. For instance, for $\mu = 1$, Chapter 1 shows that the set Λ_μ reduces to

$$\Lambda_1 = \{(0,1),(1,0),([\tilde{\lambda}_1]_0^1,[\tilde{\lambda}_2]_0^1),(1,1),([\lambda^s_{1-\mathcal{N}_1}]_0^1,[\lambda^s_{2-\mathcal{N}_1}]_0^1)\}.$$

Consequently, the maximum sum-rate ($\max\limits_{\lambda_1,\lambda_2\in\Lambda_1} R_{1\text{-HK}}(\lambda_1,\lambda_2)$) equals the maximum of five distinct functions of (P_1, P_2), as proved in the previous chapter.

Remark 3.1. *Sason proposes a coding scheme that achieves a maximum sum-rate of $C(P_1 + P_2)$ for all values of a and b [49]. For the symmetric channel, we compare this achievable sum-rate with $R^{\max}_{\text{sum-HK}}$, given in (2.233). When P is small, i.e., $P \le \frac{1-2a}{a^2}$, one can see that $R^{\max}_{\text{sum-HK}} \ge C(2P)$. For this range, the HK scheme with no time sharing outperforms the Sason's scheme. On the other hand, when P goes to infinity, $R^{\max}_{\text{sum-HK}}$ approaches $\frac{1}{2}\big(\log(P) + \log(a+1) + g_s(\hat{\lambda})\big) = \frac{1}{2}\big(\log(P) + \log(\frac{(1+a)^3}{4a})\big)$, whereas $C(2P)$ approaches $\frac{1}{2}\big(\log(P) + \log(2)\big)$. One can see that, if $a \le \sqrt{5}-2$, then $R^{\max}_{\text{sum-HK}} \ge C(2P)$, and if $a > \sqrt{5}-2$, then $R^{\max}_{\text{sum-HK}} < C(2P)$. This implies that for large values of a and P, Sason's scheme outperforms the HK scheme with no time sharing.*

The observation that Sason's scheme can sometimes achieve a higher sum-rate is a special case of the following argument: if the FD technique is used, the achievable sum-rate increases from $R^{\max}_{\text{sum-HK}}(P_1, P_2)$ to $\mathcal{C}[R^{\max}_{\text{sum-HK}}](P_1, P_2)$, where

$$\mathcal{C}[R^{\max}_{\text{sum-HK}}](P_1,P_2) = \max_{\theta_i,\alpha_i\beta_i\in[0,1]} \sum_{i=1}^{3} \theta_i R^{\max}_{\text{sum-HK}}(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}),$$
$$\text{subject to} \sum \theta_i = \sum \alpha_i = \sum \beta_i = 1.$$

In this scheme, the entire bandwidth is divided into 3 sub-bands, and in each sub-band the HK scheme is used. In the i^{th} sub-band, which has θ_i percentage of the bandwidth, the first transmitter uses α_i percentage of its total power and the second transmitter uses β_i percentage of its total power. According to Caratheodory's theorem, the rate region achieved by th FD technique will not enlarge if more than three sub-bands are used [10, 48]. Sason's achievable sum-rate of $C(P_1 + P_2)$ can be directly achieved by this scheme. Note that $f_1(P_1, P_1) \doteq C(P_1 + aP_2)$ and $f_2(P_1, P_1) \doteq C(P_2 + bP_1)$ are both achievable sum-rates. Therefore, using the FD technique, $\mathcal{C}[\max\{f_1, f_2\}](P_1, P_2)$ is also achievable. In the following, we show that

$$\mathcal{C}[\max\{f_1, f_2\}](P_1,P_2) = C(P_1+P_2). \tag{3.57}$$

Let $f \doteq \max\{f_1, f_2\}$. By Caratheodory's theorem,

$$\mathcal{C}[f](P_1, P_2) = \sup_{\theta_i, \alpha_i \beta_i \in [0,1]} \sum_{i=1}^{3} \theta_i f\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}\Big), \tag{3.58}$$

subject to $\sum_{i=1}^{3} \theta_i = \sum_{i=1}^{3} \alpha_i = \sum_{i=1}^{3} \beta_i = 1$. Therefore, we have

$$\begin{aligned}
\mathcal{C}[f](P_1, P_2) = & \sup_{\theta_i, \alpha_i \beta_i \in [0,1]} \sum_{i=1}^{3} \theta_i f\Big(\frac{\alpha_i P_1}{\theta_i}, \frac{\beta_i P_2}{\theta_i}\Big) \\
\geq & \hat{\theta}_1 f\Big(\frac{\hat{\alpha}_1 P_1}{\hat{\theta}_1}, \frac{\hat{\beta}_1 P_2}{\hat{\theta}_1}\Big) + \hat{\theta}_2 f\Big(\frac{\hat{\alpha}_2 P_1}{\hat{\theta}_2}, \frac{\hat{\beta}_2 P_2}{\hat{\theta}_2}\Big) \\
\geq & \hat{\theta}_1 f_1\Big(\frac{\hat{\alpha}_1 P_1}{\hat{\theta}_1}, \frac{\hat{\beta}_1 P_2}{\hat{\theta}_1}\Big) + \hat{\theta}_2 f_2\Big(\frac{\hat{\alpha}_2 P_1}{\hat{\theta}_2}, \frac{\hat{\beta}_2 P_2}{\hat{\theta}_2}\Big) \\
= & \hat{\theta}_1 C(P_1 + P_2) + \hat{\theta}_2 C(P_1 + P_2) \\
= & C(P_1 + P_2),
\end{aligned} \tag{3.59}$$

where

$$\hat{\theta}_1 = \frac{P_1}{P_1 + P_2}, \tag{3.60}$$

$$\hat{\theta}_2 = \frac{P_2}{P_1 + P_2}, \tag{3.61}$$

$$\hat{\alpha}_1 = 1, \tag{3.62}$$

$$\hat{\alpha}_2 = 0, \tag{3.63}$$

$$\hat{\beta}_1 = 0, \tag{3.64}$$

$$\hat{\beta}_2 = 1. \tag{3.65}$$

On the other hand, $\mathcal{C}[f](P_1, P_2)$ is the smallest concave function which lies above $C(P_1 + aP_2)$ and $C(P_2 + bP_1)$. Since $C(P_1 + P_2)$ is concave and is larger than $C(P_1 + aP_2)$ and $C(P_2 + bP_1)$, we have

$$\mathcal{C}[f](P_1, P_2) \leq C(P_1 + P_2). \tag{3.66}$$

Comparing (3.66) with (3.59), we conclude that

$$\mathcal{C}[f](P_1, P_2) = C(P_1 + P_2). \tag{3.67}$$

Therefore, for all values of $a < 1$ and $b < 1$, one can achieve $C(P_1 + P_2)$. On the other hand, if the first receiver decides to decode the entire interference, then a MAC bound on the sum-rate implies that $R_1 + R_2 < C(P_1 + aP_2)$. Note that since $a < 1$,

$C(P_1 + aP_2) < C(P_1 + P_2)$. This shows that, for weak interference, if the HK scheme requires one of the receivers to decode the entire interference, then the achievable sum-rate will not be optimal. This is in contrast to the strong and mixed interference classes, in which to achieve the sum-capacity, at least one of the receivers must decode the entire interference.

The time-sharing variable Q can also enlarge the achievable rate region. Furthermore, this region includes the rate region achieved by the FD technique [6]. However, under some constraints, these two regions are in fact equal. For instance, when interference is weak, FD and Q result in the same achievable rate region, as stated in Corollary 3.1. Therefore, one can characterize the maximum achievable sum-rate, even when time sharing is used, as explained in the following corollary:

Corollary 3.2. *When interference is weak, the maximum achievable sum-rate of the HK scheme with Gaussian inputs (and with time sharing) is given by* $\mathcal{C}[R^{\max}_{\text{sum-HK}}](P_1, P_2)$*, where the function* $R^{\max}_{\text{sum-HK}}(P_1, P_2)$ *is given in (2.19).*

Remark 3.2. *Explicit calculation of the upper concave envelope of a function is in general very complicated. However, under some constraints, one can use supporting hyperplanes and explicitly characterize the upper concave envelope. Using this idea, Costa and Nair [42] characterized the maximum achievable sum-rate of the symmetric channel, for some ranges of channel parameters. Following a similar approach, one can explicitly characterize* $\mathcal{C}[R^{\max}_{\text{sum-HK}}](P_1, P_2)$*, for some ranges of channel parameters. Moreover, it is known that representing the achievable rate region in terms of upper concave envelope can help characterize the capacity region [50–53].*

3.4 Conclusion

This chapter examined the boundary of the HK rate region relying on Gaussian inputs. When no time sharing is used, we characterized the boundary for the class of weak interference. When time sharing is used, we expressed the entire boundary in terms of the upper concave envelope of a function of (P_1, P_2). Therefore, we fully characterized the entire boundary of the HK region.

Chapter 4

Rate Splitting and Successive Decoding for Gaussian Interference Channels

This chapter investigates the structure of sum-rate optimal codes proposed for the two-user Gaussian Interference Channel (GIC). It describes an optimization problem that corresponds to the maximum achievable sum-rate through rate splitting and successive decoding. First, the complexity of the optimization problem, and in particular the non-convexity of the problem, is highlighted. Then an optimization method is proposed to solve the problem under a set of mild conditions. The main result of this chapter is the closed form expression for the optimal power allocation that achieves the sum-capacity.

4.1 Introduction

Most coding schemes proposed for the two-user GIC employ joint decoding to enlarge the achievable rate region. For instance, the Simultaneous Non-unique Decoding (SND) scheme [37] and the well-known Han-Kobayashi (HK) scheme [6] employ joint decoding; however, joint decoding increases decoding complexity.

To decrease decoding complexity, practical coding schemes employ Successive Decoding (SD). Moreover, there exists a considerable amount of literature regarding the construction of high performance point-to-point codes [19–23], whereas there is much less research on multiuser codebooks, which can be jointly decoded. Thus, it is important to have a comprehensive understanding of the performance of SD, which employs existing point-to-point codes, in comparison to joint decoding, which employs multiuser codes.

Rate Splitting (RS) and successive decoding can reduce decoding complexity and have been used to investigate the multiple access channel and the interference channel [24, 25]. RS and SD have been used in a wide range of problems in information theory [54–57]. The capacity region of the two-user Gaussian multiple access channel can be achieved using RS and SD. In fact, if each message is split into two parts and decoding is done in the proper order, the boundary of the capacity region can be achieved using SD. Moreover, even the boundary of the capacity region of the K-user Gaussian multiple access channel can be achieved using RS and SD [18, 26, 58].

For the interference channel, a misconception exists that RS and SD can achieve the entire SND rate region or even the HK rate region. Reference [59] explains this misconception and highlights that, when several receivers have to decode a rate-splitting codebook, the entire capacity region may not be achieved. In particular, it is proved that, for the two-user GIC, RS and SD cannot achieve even the SND rate region [27]. Moreover, [27] proposes a sliding window decoding scheme that achieves the performance of the simultaneous non-unique decoding inner bound.

The problem of sum-rate maximization has been studied in the literature [60–63]. In particular, RS and SD have been used to investigate the maximum achievable sum-rate of the two-user GIC. For instance, [64] proposes an algorithm based on RS and SD which is derived by first investigating the deterministic interference channel [65, 66]. For the symmetric two-user GIC, [64] provides numerical evaluations to show that the sum-rate of the SD algorithm is above that of the single-split schemes and below that of the HK scheme. In addition, [10] shows that, when interference is mixed, the sum-capacity can be achieved using SD. However, when interference is strong or weak, the performance of RS and SD has not been well-understood. This study shows that, under a mild condition on transmitters' powers, RS and SD can achieve the sum-rate of the HK scheme [67, 68].

This study examines the achievable sum-rate of the two-user GIC when SD is used instead of joint decoding. Although it is known that a corner point of the SND rate region cannot be achieved using SD [27], this chapter shows that SD can achieve the maximum sum-rate of the HK scheme. First, this chapter investigates the strong interference class and shows that, if transmitters' powers satisfy certain conditions, RS and SD achieve the sum-capacity of the channel. The order of decoding at the receivers, the number of the required splits, and the amount of power allocated to each split are described as closed-form expressions. Moreover, when SD is strictly inferior to joint decoding, this study calculates the maximum sum-rate loss when joint decoding is replaced by SD. It is shown that the maximum sum-rate loss does not depend on transmitters' powers and remains constant as powers approach infinity. Second, this chapter investigates the weak interference class. Similar to the strong interference class, it is shown that, if transmitters' powers satisfy certain conditions, the maximum sum-rate of the HK scheme can be achieved using SD. It is shown that for a wide range of channel gains and transmitters' powers, a single-split scheme can achieve the sum-rate of the HK scheme. For a small region, the single-split scheme actually achieves the sum-capacity. Moreover, we propose a coding scheme based on RS and SD in which both transmitters divide their messages into $N+1$ parts, where N can be any positive integer. We show that this scheme can achieve the sum-rate of the HK scheme. Once again, the order of decoding at the receivers, the number of required splits, and the amount of power allocated to each split are described as closed-form expressions. When SD is strictly inferior to the HK scheme, this study calculates the maximum sum-rate difference. It is shown that the maximum sum-rate difference does not depend on transmitters' power and remains constant as transmitters' powers approach infinity.

The HK scheme results in the best-known achievable rate region. Unfortunately, the mathematical expressions that characterize the HK rate region are complicated and involve some arbitrary power splitting variables. In contrast, our SD scheme does not have arbitrary variables and results in simple characterization of the achievable sum-rate. Consequently, our scheme provides insight into structures of sum-rate optimal codes.

Joint decoding is also used in parallel channels. An important question about parallel channels is separability: is it necessary to jointly encode and decode across all subchannels to achieve the capacity region? Can separate encoding and decoding achieve the entire capacity region? In fact, it is known that parallel Gaussian point-to-point channels, parallel Gaussian multiple access channels, and parallel Gaussian broadcast channels are separable [37], [69] and there is no need for joint coding. However, parallel

Gaussian interference channels are not separable and separate decoding can considerably decrease the achievable rate region [70]. Specific cases of parallel Gaussian interference channels are studied by [71, 72] and the optimality of separate coding is investigated for each case. Note that this chapter does not investigate parallel channels. Rather, in this chapter, joint decoding is performed over one GIC, and the decoders jointly decode some messages that are transmitted over a single channel.

The structure of this chapter is as follows. In Section 4.2, the channel model and preliminaries are introduced. This section expresses the optimization problem that corresponds to maximizing the achievable sum-rate. Although it is shown that the optimization problem is non-convex and involves a discrete optimization, we provide closed-form expressions for the optimal solution. In Section 4.3 and 4.4 the achievable sum-rate is studied for the strong and weak interference classes, respectively. These sections, which demonstrate how many splits are required and how much power should be allocated to each split, highlight the main contributions of Chapter 4. This chapter concludes in Section 4.5.

4.2 Preliminaries

The following notations are used in this chapter. $S_1^{1:N}$ represents $\{S_1^1, S_1^2, ..., S_1^N\}$. For a random variable S_1, $P(S_1)$ represents the power of S_1 and for a set $S_1^{1:N}$, $P(S_1^{1:N}) \doteq \sum_{i=1}^{N} P(S_1^i)$. For a statement Q, $\mathbb{1}(Q) = 1$ if Q is true, otherwise $\mathbb{1}(Q) = 0$.

The two-user GIC is defined in Chapter 2. Based on the values of a and b, the interference is divided into some classes, namely weak, strong, and mixed, as defined in Chapter 2. In this chapter, we investigate the achievable sum-rate of each class separately.

4.2.1 The Underlying Optimization Problem Corresponding to Maximum Sum-Rate

We formulate the achievable sum-rate of the two-user GIC, when RS and SD are used. The i^{th} transmitter, $i \in \{1, 2\}$, splits its message M_i into N_i parts, namely $M_i^1, M_i^2, ..., M_i^{N_i}$. Then, M_i^j is encoded by X_i^j according to $N(0, P_i^j)$ where P_i^j is the power allocated to M_i^j.

Moreover, R_i^j represents the rate of M_i^j and $R_i = \sum_{j=1}^{N_i} R_i^j$. Finally, all N_i codewords are superimposed and $X_i = \sum_{j=1}^{N_i} X_i^j$ is transmitted. Transmitters' powers are bounded by P_1 and P_2, i.e.,

$$\begin{aligned}\sum_{j=1}^{N_1} P_1^j &\le P_1,\\ \sum_{j=1}^{N_2} P_2^j &\le P_2.\end{aligned} \tag{4.1}$$

The order of decoding at the receivers can affect the sum-rate achieved using SD. The first receiver successively decodes all parts of M_1 using a specific order $\mathbf{S_1}$ where $\mathbf{S_1} \doteq (S_1^1, S_1^2, ..., S_1^{N_1+N_2})$. In fact, each S_1^j represents exactly one of the X_i^j, $i \in [1:2]$ and $j \in [1:N_i]$, such that $\mathbf{S_1}$ is a permutation of $\{X_1^1, X_1^2, ..., X_1^{N_1}\} \cup \{\sqrt{a}X_2^1, \sqrt{a}X_2^2, ..., \sqrt{a}X_2^{N_2}\}$. First, $S_1^{N_1+N_2}$ is decoded by the first receiver, while considering all other splits as noise. After decoding $S_1^{N_1+N_2}$, $S_1^{N_1+N_2-1}$ is decoded, while considering $S_1^{1:N_1+N_2-2}$ as noise. The first receiver follows $\mathbf{S_1}$ until all parts of M_1 are decoded. Note that some parts of M_2 may not be decoded. For instance, if $S_1^1 = \sqrt{a}X_2^1$, then the first receiver does not decode X_2^1. Similarly, the second receiver successively decodes all parts of M_2 using a specific order $\mathbf{S_2} \doteq (S_2^1, S_2^2, ..., S_2^{N_1+N_2})$ where $\mathbf{S_2}$ is a permutation of $\{X_2^1, X_2^2, ..., X_2^{N_2}\} \cup \{\sqrt{b}X_1^1, \sqrt{b}X_1^2, ..., \sqrt{b}X_1^{N_1}\}$.

The first receiver must decode X_1^j, but the second receiver only decodes X_1^j if according to the order $\mathbf{S_2}$, X_1^j is required to be decoded. Therefore, the first receiver imposes a constraint on R_1^j, but the second receiver only imposes a constraint on R_1^j if it decodes X_1^j. Mathematically,

$$R_1^j \le c_1^j \doteq C\Big(\frac{P(S_1^{K_1^j})}{1 + P(S_1^{1:K_1^j-1})}\Big), \tag{4.2}$$

so that X_1^j can be reliably decoded at the first receiver where $S_1^{K_1^j} = X_1^j$ in the decoding order $\mathbf{S_1}$. Similarly, if X_1^j is decoded at the second receiver, then

$$R_1^j \le d_1^j \doteq C\Big(\frac{P(S_2^{L_1^j})}{1 + P(S_2^{1:L_1^j-1})}\Big), \tag{4.3}$$

so that X_1^j can be reliably decoded at the second receiver where $S_2^{L_1^j} = \sqrt{b}X_1^j$ in the decoding order $\mathbf{S_2}$. Therefore,

$$R_1^j \le \begin{cases} \min\{c_1^j, d_1^j\} & \text{if the second receiver decodes } X_1^j, \\ c_1^j & \text{otherwise.} \end{cases} \tag{4.4}$$

Similarly, X_2^j should be decoded by the second receiver, but the first receiver only decodes it if the decoding order $\mathbf{S_1}$ requires decoding of X_2^j. Therefore,

$$R_2^j \le \begin{cases} \min\{c_2^j, d_2^j\} & \text{if the first receiver decodes } X_2^j, \\ c_2^j & \text{otherwise.} \end{cases} \tag{4.5}$$

where $c_2^j \doteq C\Big(\frac{P(S_2^{L_2^j})}{1+P(S_2^{1:L_2^j-1})}\Big)$, $d_2^j \doteq C\Big(\frac{P(S_1^{K_2^j})}{1+P(S_1^{1:K_2^j-1})}\Big)$, $S_1^{K_2^j} = \sqrt{a}X_2^j$, and $S_2^{L_2^j} = X_2^j$.

To find the maximum sum-rate achieved using SD, the following optimization problem is investigated.

$$R_{\text{sum-SD}}^{\text{opt}} \doteq \max_{N_1,N_2,P_1^j,P_2^j,\mathbf{S_1},\mathbf{S_2}} \Big(\sum_{j=1}^{N_1} R_1^j + \sum_{j=1}^{N_2} R_2^j\Big),$$
$$\text{subject to (4.1), (4.4), (4.5).} \tag{4.6}$$

This optimization problem is not convex, and finding the general solution can be difficult. However, in this chapter, we characterize the optimal solution of this problem, for a wide range of a, b, P_1, and P_2.

Note that when interference is mixed, the optimal solution of (4.6) can be easily found. In fact, [10] shows that, for the mixed class in which $a \ge 1$ and $0 < b < 1$, the sum-capacity is given by

$$C_{\text{sum}} = C(P_1) + \min\Big\{C\Big(\frac{P_2}{1+bP_1}\Big), C\Big(\frac{aP_2}{1+P_1}\Big)\Big\}. \tag{4.7}$$

On the other hand, consider the following solution to the optimization problem (4.6).

$$\begin{aligned} (N_1, N_2) &= (1,1), \\ \mathbf{S_1} &= (X_1, \sqrt{a}X_2), \\ \mathbf{S_2} &= (\sqrt{b}X_1, X_2). \end{aligned} \tag{4.8}$$

This solution leads to the following achievable rates:

$$\begin{aligned} R_1 &= c_1^1 = C(P_1), \\ R_2 &= \min\{c_2^1, d_2^1\} = \min\Big\{C\Big(\frac{P_2}{1+bP_1}\Big), C\Big(\frac{aP_2}{1+P_1}\Big)\Big\}. \end{aligned} \tag{4.9}$$

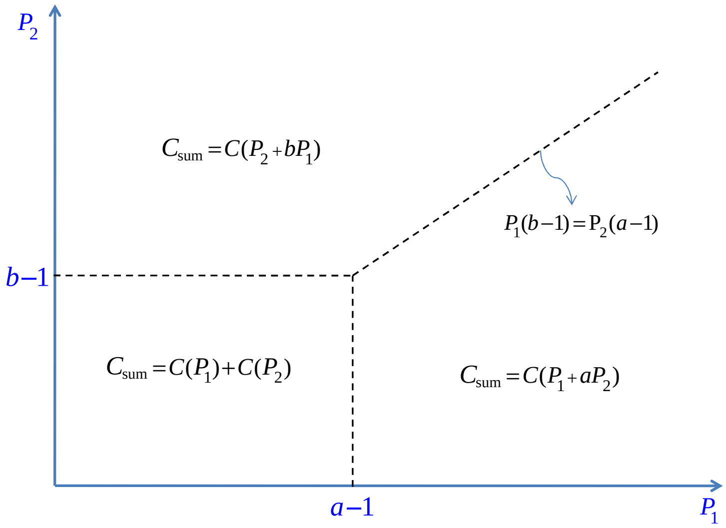

Figure 4.1: The sum-capacity of the strong interference class.

Comparing (4.9) with (4.7), we conclude that the solution (4.8) achieves the sum-capacity, and therefore, is the optimal solution of (4.6). Similarly, one can show that for the mixed class in which $b \geq 1$ and $0 < a < 1$,

$$\begin{aligned} (N_1, N_2) =& (1, 1), \\ \mathbf{S_1} =& (\sqrt{a}X_2, X_1), \\ \mathbf{S_2} =& (X_2, \sqrt{b}X_1), \end{aligned} \tag{4.10}$$

shows the optimal solution of (4.6) that achieves the sum-capacity.

In the following, two distinct cases are studied, namely the strong interference class and the weak interference class. We calculate closed-form expressions for the number of splits, the optimal power allocated to each split, and the achievable rate of each split.

4.3 Strong Interference Class

The strong interference class is the case defined by $a \geq 1$ and $b \geq 1$. The sum-capacity of this class is known. In fact, for the strong interference class, the entire capacity region is achieved using SND [7, 37], and the sum-capacity is given by

$$C_{\text{sum}} = \min \left\{ \begin{matrix} C(P_1 + aP_2), C(P_2 + bP_1), \\ \\ C(P_1) + C(P_2) \end{matrix} \right\}. \tag{4.11}$$

Consequently,

$$C_{\text{sum}} = \begin{cases} C(P_1) + C(P_2) & \text{if } P_1 \le a-1, P_2 \le b-1, \\ C(P_1 + aP_2) & \text{if } P_1 \ge \max\{a-1, P_2 \frac{(a-1)}{(b-1)}\}, \\ C(P_2 + bP_1) & \text{if } P_2 \ge \max\{b-1, P_1 \frac{(b-1)}{(a-1)}\}, \end{cases} \tag{4.12}$$

as shown in Figure 4.1.

The main goal of this section is to show that, the sum-rate achieved using RS and SD equals the sum-capacity for a wide range of (a, b, P_1, P_2). In other words, the optimal solution of (4.6) equals C_{sum} for a wide range of (a, b, P_1, P_2). In doing so, we first show that without using rate splitting, one can achieve C_{sum} for some values of (a, b, P_1, P_2). Then we show that by using rate splitting, but without any joint decoding, one can achieve C_{sum} for a wide range of (a, b, P_1, P_2).

4.3.1 Is Rate Splitting Required?

We calculate the achievable sum-rate when no RS is used. Our main goal is to show that, for some values of (a, b, P_1, P_2), RS is not required. In doing so, we first solve the optimization problem (4.6) for $N_1 = N_2 = 1$. Then we compare the results with the sum-capacity expression given in (4.12). The following theorem characterizes the maximum achievable sum-rate when no rate splitting is used.

Theorem 4.1. *For the two-user GIC with strong interference, the maximum sum-rate achieved with no rate splitting is given by*

$$\begin{aligned} R_{\text{sum}}^{\text{NRS}} = \min \Big\{ & C(\frac{bP_1}{1+P_2}) + C(\frac{aP_2}{1+P_1}), \\ & C(P_1 + aP_2), \\ & C(P_2 + bP_1), \\ & C(P_1) + C(P_2) \Big\}. \end{aligned} \tag{4.13}$$

Proof. When no RS is used, we have $N_1 = N_2 = 1$. Therefor, the optimization (4.6) reduces to

$$R_{\text{sum}}^{\text{NRS}} \doteq \max_{\mathbf{S_1}, \mathbf{S_2}} R_1 + R_2. \tag{4.14}$$

Decoding order $\mathbf{S_1}$	Decoding order $\mathbf{S_2}$	$R_1 + R_2$
$(\sqrt{a}X_2, X_1)$	$(\sqrt{b}X_1, X_2)$	$C(\frac{P_1}{1+aP_2}) + C(\frac{P_2}{1+bP_1})$
$(X_1, \sqrt{a}X_2)$	$(\sqrt{b}X_1, X_2)$	$C(P_1) + \min\{C(\frac{P_2}{1+bP_1}), C(\frac{aP_2}{1+P_1})\}$ $= C(P_1) + C(\frac{P_2}{1+bP_1})$
$(\sqrt{a}X_2, X_1)$	$(X_2, \sqrt{b}X_1)$	$\min\{C(\frac{P_1}{1+aP_2}), C(\frac{bP_1}{1+P_2})\}+$ $C(P_2)$ $= C(\frac{P_1}{1+aP_2}) + C(P_2)$
$(X_1, \sqrt{a}X_2)$	$(X_2, \sqrt{b}X_1)$	$\min\{C(P_1), C(\frac{bP_1}{1+P_2})\}+$ $\min\{C(P_2), C(\frac{aP_2}{1+P_1})\}$ $= \min\{C(\frac{bP_1}{1+P_2}) + C(\frac{aP_2}{1+P_1}),$ $C(P_1 + aP_2), C(P_2 + bP_1),$ $C(P_1) + C(P_2)\}$

Table 4.1: The achievable sum-rate of the strong interference class corresponding to four decoding orders.

There exist four possibilities for $\mathbf{S_1}$ and $\mathbf{S_2}$, as shown in Table 4.1. This table shows the achievable sum-rate corresponding to the four possible decoding orders. In the first case, both receivers treat the interference as noise. In the second case, the first receiver decodes the interference whereas the second receiver treats the interference as noise. In the third case, the second receiver decodes the interference whereas the first receiver treats the interference as noise. Finally, in the fourth case, both receiver decode the interference. Note that the sum-rate corresponding to the fourth decoding orders, i.e., $\mathbf{S_1} = (X_1, \sqrt{a}X_2), \mathbf{S_2} = (X_2, \sqrt{b}X_1)$, is greater than the sum-rate achieved by other decoding orders. Therefore, for all values of (a, b, P_1, P_2), the maximum achievable sum-rate is given by the rate expression corresponding to fourth decoding orders, as stated in (4.13). This completes the proof $\square$

For fixed values of a and b, Figure 4.2 demonstrates $R_{\text{sum}}^{\text{NRS}}$ in the P_1P_2-plane. By

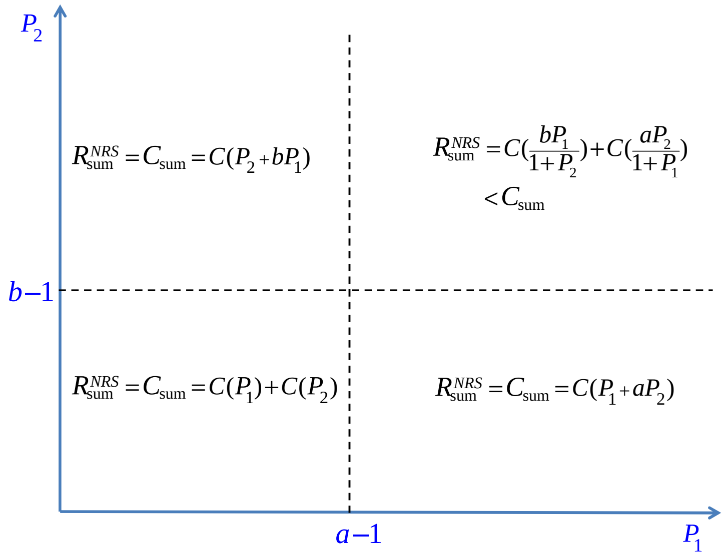

Figure 4.2: Comparison of $R_{\text{sum}}^{\text{NRS}}$ with the sum-capacity for the strong interference class.

comparing this figure with Figure 4.1, we can compare $R_{\text{sum}}^{\text{NRS}}$ with C_{sum}. Note that, $C_{\text{sum}} > R_{\text{sum}}^{\text{NRS}}$ if and only if

$$\begin{aligned} &P_1 > a-1, \\ &P_2 > b-1. \end{aligned} \tag{4.15}$$

Moreover, when interference is very strong, i.e., $1+P_1 \leq a$ and $1+P_2 \leq b$, Figure 4.2 shows that without any rate splitting, the sum-capacity is achieved. We highlight this observation in the following corollary.

Corollary 4.1. *For the two-user GIC, when interference is very strong, the sum-capacity can be achieved using SD.*

In the next sub-section, we propose a novel coding scheme based on RS and SD that achieves a sum-rate better than $R_{\text{sum}}^{\text{NRS}}$. We show that our scheme achieves the sum-capacity for a wide range of (a, b, P_1, P_2).

4.3.2 How Many Splits Are Required?

In this sub-section, we propose a coding scheme that divides both messages into $N+1$ parts. We show that, to achieve the sum-capacity, N should be properly chosen according to the value of (P_1, P_2).

To find the optimal solution of (4.6), the following decoding orders are proposed. For $a \neq 1$ and $b \neq 1$, let

$$(S_1^1, S_1^2, S_1^3, S_1^4, ...) = (X_1^1, \sqrt{a}X_2^1, X_1^2, \sqrt{a}X_2^2, ...),$$
$$(S_2^1, S_2^2, S_2^3, S_2^4, ...) = (X_2^1, \sqrt{b}X_1^1, X_2^2, \sqrt{b}X_1^2, ...). \tag{4.16}$$

Since the optimization problem (4.6) is non-convex, it may be difficult to find the optimal power allocations. The main idea is to use proper power allocations, such that

$$c_1^j = d_1^j \quad \text{if } X_1^j \text{ is decoded by the second receiver,}$$
$$c_2^j = d_2^j \quad \text{if } X_2^j \text{ is decoded by the first receiver.} \tag{4.17}$$

Intuitively, these extra constraints prevent power loss. If $c_1^j > d_1^j$, then we have allocated some power to enhance the channel between the first transmitter and the first receiver. However, since $d_1^j < c_1^j$, the capacity of the channel between the first transmitter and the second receiver restricts the achievable rate of the channel between the first transmitter and the first receiver.

Relying on (4.17), we characterize a feasible solution to the optimization problem (4.6). As highlighted earlier, due to non-convexity of (4.6), characterizing the optimal solution can be difficult. However, we show that, for a wide range of (a, b, P_1, P_2), the feasible solution that satisfies (4.17) is in fact the optimal solution of (4.6). The idea to prove the optimality of our solution is to use some form of duality certificate. Instead of proving the optimality directly, we show that our solution achieves the sum-capacity. In the following, we first propose our feasible solution. Then we show that for a wide range of (a, b, P_1, P_2) our solution achieves the sum-capacity.

According to (4.16), we have $c_1^1 = C(P_1^1)$ and $d_1^1 = C(\frac{bP_1^1}{1+P_2^1})$. To satisfy (4.17), we have $c_1^1 = d_1^1$, and consequently, P_2^1 is found as follows:

$$c_1^1 = d_1^1$$
$$\Rightarrow P_2^1 = b - 1. \tag{4.18}$$

Similarly, we have $c_2^1 = C(P_2^1)$ and $d_2^1 = C(\frac{aP_2^1}{1+P_1^1})$. By letting $c_2^1 = d_2^1$, P_1^1 is found.

$$c_2^1 = d_2^1$$
$$\Rightarrow P_1^1 = a - 1. \tag{4.19}$$

Generally, for $k \geq 2$, c_1^k and d_1^k are given by

$$c_1^k \overset{(a)}{=} C\Big(\frac{P_1^k}{1+\sum_{j=1}^{k-1} P_1^j + a(\sum_{j=1}^{k-1} P_2^j)}\Big),$$
$$d_1^k \overset{(b)}{=} C\Big(\frac{bP_1^k}{1+\sum_{j=1}^{k} P_2^j + b(\sum_{j=1}^{k-1} P_1^j)}\Big), \tag{4.20}$$

where (a) and (b) is calculated based on the decoding orders given in (4.16). Next, by letting $c_1^k = d_1^k$, P_2^k is found, as follows.

$$\begin{aligned} c_1^k =& d_1^k \\ \Rightarrow \Big(1+\sum_{j=1}^{k} P_2^j + b(\sum_{j=1}^{k-1} P_1^j)\Big) =& b\Big(1+\sum_{j=1}^{k-1} P_1^j + a(\sum_{j=1}^{k-1} P_2^j)\Big) \\ \Rightarrow P_2^k =& (b-1) + (ab-1)\sum_{j=1}^{k-1} P_2^j \\ \overset{(a)}{\Rightarrow} P_2^k =& (b-1)(ab)^{k-1}, k \in \{1,2,3,...\}, \end{aligned} \tag{4.21}$$

where (a) is justified by induction on k. Similarly, c_2^k and d_2^k are given by

$$c_2^k = C\Big(\frac{aP_2^k}{1+\sum_{j=1}^{k} P_1^j + a(\sum_{j=1}^{k-1} P_2^j)}\Big),$$
$$d_2^k = C\Big(\frac{P_2^k}{1+\sum_{j=1}^{k-1} P_2^j + b(\sum_{j=1}^{k-1} P_1^j)}\Big), \tag{4.22}$$

and by letting $c_2^k = d_2^k$, P_1^k is found.

$$P_1^k = (a-1)(ab)^{k-1}, k \in \{1,2,3,...\}. \tag{4.23}$$

Moreover, by inserting (4.21) and (4.23) into (4.20), c_1^k and d_1^k simplify to

$$c_1^k = d_1^k = C(a-1). \tag{4.24}$$

Similarly, by inserting (4.21) and (4.23) into (4.22), c_2^k and d_2^k simplify to

$$c_2^k = d_2^k = C(b-1). \tag{4.25}$$

Note that the values of c_1^k and c_2^k do not depend on k.

With this power allocation, the constraints (4.4) on R_1^k and (4.5) on R_2^k simplify to

$$R_1^k \leq \min\{c_1^k, d_1^k\} \overset{(a)}{=} C(a-1),$$
$$R_2^k \leq \min\{c_2^k, d_2^k\} \overset{(b)}{=} C(b-1), \tag{4.26}$$

where (a) is valid by (4.24), and (b) is valid by (4.25).

For the strong interference class, define $P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$ as

$$
\begin{aligned}
P_{1,\mathrm{S}}^{\mathrm{opt}}(N) &\doteq \sum_{j=1}^{N} P_1^j, \\
P_{2,\mathrm{S}}^{\mathrm{opt}}(N) &\doteq \sum_{j=1}^{N} P_2^j,
\end{aligned} \tag{4.27}
$$

where P_1^j and P_2^j are given by (4.23) and (4.21), and N is a positive integer. Therefore,

$$
\begin{aligned}
P_{1,\mathrm{S}}^{\mathrm{opt}}(N) &= \frac{a-1}{ab-1}\left((ab)^N - 1\right), \\
P_{2,\mathrm{S}}^{\mathrm{opt}}(N) &= \frac{b-1}{ab-1}\left((ab)^N - 1\right).
\end{aligned} \tag{4.28}
$$

To simplify the notations, we define $P_{1,\mathrm{S}}^{\mathrm{opt}}(0) = 0$ and $P_{2,\mathrm{S}}^{\mathrm{opt}}(0) = 0$. In fact, if $P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$, for some positive integer N, then each transmitter can split its message into exactly N parts and can allocate a proper amount of power to each of these N parts such that (4.17) is satisfied. This power allocation has the property that all splits of M_1 can achieve the same rate, i.e., $C(a-1)$, and all splits of M_2 can achieve the same rate, i.e., $C(b-1)$. Therefore, based on the proposed decoding orders (4.16) and power allocations (4.21,4.23), SD results in the following achievable sum-rate.

$$
R_1 + R_2 = NC(a-1) + NC(b-1). \tag{4.29}
$$

The following theorem shows that if $P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$ for some positive integer N, then SD can achieve the sum-capacity of the strong interference class.

Theorem 4.2. *For the two-user GIC with strong interference, if $P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$ for some positive integer N, then splitting of M_1 and M_2 into N parts and allocating power according to (4.21,4.23) and decoding according to (4.16) is sum-rate optimal.*

Proof. For $P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$, since both $P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$ are strictly increasing functions of N, $P_1 \geq P_{1,\mathrm{S}}^{\mathrm{opt}}(1) = a-1$ and $P_2 \geq P_{2,\mathrm{S}}^{\mathrm{opt}}(1) = b-1$. Therefore, interference is strong but not very strong, and C_{sum} is given by

$$
C_{\mathrm{sum}} \overset{(a)}{=} \min\{C(P_1 + aP_2), C(P_2 + bP_1)\}, \tag{4.30}
$$

where (a) is valid by (4.12). For such values of P_1 and P_2,

$$\begin{aligned}
C(P_1+aP_2) =& \frac{1}{2}\log(1+P_1+aP_2) \\
\overset{(a)}{=}& \frac{1}{2}\log\big(1+\frac{a-1}{ab-1}\left((ab)^N-1\right)+a(\frac{b-1}{ab-1})\left((ab)^N-1\right)\big) \\
=& \frac{1}{2}\log((ab)^N) \\
=& N\frac{1}{2}\log(a)+N\frac{1}{2}\log(b) \\
=& NC(a-1)+NC(b-1),
\end{aligned} \tag{4.31}$$

where (a) is valid by (4.28). Similarly,

$$C(P_2+bP_1)=NC(a-1)+NC(b-1). \tag{4.32}$$

Since $C(P_1+aP_2)$=$C(P_2+bP_1)$=$NC(a-1)+NC(b-1)$, the sum-capacity is given by

$$\begin{aligned}
C_{\text{sum}} =& \min\{C(P_1+aP_2),C(P_2+bP_1)\} \\
=& NC(a-1)+NC(b-1),
\end{aligned} \tag{4.33}$$

but this sum-rate is achieved using the proposed SD, as explained in (4.29). This completes the proof. □

Theorem 4.1 and Theorem 4.2 show that if P_1 and P_2 satisfy certain conditions, then SD achieves the sum-capacity of the channel. In the next theorem, we propose a novel RS scheme that divides both messages into $N+1$ parts. We show that N should be properly chosen according to (P_1,P_2). The next theorem, uses Theorem 4.2 to find even more values of P_1 and P_2 for which SD is sum-rate optimal.

In the rest of this chapter, we deal with many calculations that involve the function $C(x)$. We frequently use the following property of this function: if x and y are non-negative real numbers, we have

$$C(x+y)=C(x)+C(\frac{y}{1+x}). \tag{4.34}$$

Theorem 4.3. *For the two-user GIC with strong interference, if one of the following conditions holds for some non-negative integer N, then allocating power according to (4.21,4.23) and decoding according to (4.16) is sum-rate optimal.*

Condition A:

$$P_{2,\mathrm{S}}^{\mathrm{opt}}(N) \leq P_2 < P_{2,\mathrm{S}}^{\mathrm{opt}}(N+1),\ P_{1,\mathrm{S}}^{\mathrm{opt}}(N+1) \leq P_1. \tag{4.35}$$

Condition B:

$$P_{1,\mathrm{S}}^{\mathrm{opt}}(N) \leq P_1 < P_{1,\mathrm{S}}^{\mathrm{opt}}(N+1),\ P_{2,\mathrm{S}}^{\mathrm{opt}}(N+1) \leq P_2. \tag{4.36}$$

Proof. We prove this theorem when condition A holds. The proof, corresponding to condition B, can be obtained by changing indices 1 and 2. The main idea is to use a portion of P_1 and a portion of P_2 for the first splits of M_1 and M_2 such that the remaining powers satisfy conditions of Theorem 4.2. Therefore, we express P_1 and P_2 as follows:

$$\begin{aligned} P_1 &= P_{1,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_1, \\ P_2 &= P_{2,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_2, \end{aligned} \tag{4.37}$$

and since condition A holds, we have

$$\begin{aligned} \Delta P_1 &\geq P_{1,\mathrm{S}}^{\mathrm{opt}}(N+1) - P_{1,\mathrm{S}}^{\mathrm{opt}}(N) \stackrel{(a)}{=} (a-1)(ab)^N, \\ \Delta P_2 &< P_{2,\mathrm{S}}^{\mathrm{opt}}(N+1) - P_{2,\mathrm{S}}^{\mathrm{opt}}(N) \stackrel{(b)}{=} (b-1)(ab)^N, \end{aligned} \tag{4.38}$$

where (a) and (b) are valid by (4.28). In fact, for each value of N, (4.37) describes a power region in the P_1P_2-plane. For this region of powers, the first transmitter uses ΔP_1 to transmit X_1^{N+1}. Similarly, the second transmitter uses ΔP_2 to transmit X_2^{N+1}. Then each receiver successively decodes both X_1^{N+1} and X_2^{N+1}. After this step, the remaining power of each transmitter satisfies Theorem 4.2, i.e., $P_1 - \Delta P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N)$ and $P_2 - \Delta P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N)$. In fact, according to the decoding order (4.16), we have

$$\begin{aligned} R_1 =& NC(a-1) + \min\{c_1^{N+1}, d_1^{N+1}\} \\ =& NC(a-1) + \min\Big\{C\Big(\frac{\Delta P_1}{1 + P_{1,\mathrm{S}}^{\mathrm{opt}}(N) + aP_{2,\mathrm{S}}^{\mathrm{opt}}(N)}\Big), \\ & \qquad\qquad C\Big(\frac{b\Delta P_1}{1 + P_{2,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_2 + bP_{1,\mathrm{S}}^{\mathrm{opt}}(N)}\Big)\Big\} \\ \stackrel{(a)}{=}& NC(a-1) + \min\Big\{C\Big(\frac{\Delta P_1}{(ab)^N}\Big), C\Big(\frac{b\Delta P_1}{(ab)^N + \Delta P_2}\Big)\Big\} \\ \stackrel{(b)}{=}& NC(a-1) + C\Big(\frac{\Delta P_1}{(ab)^N}\Big), \end{aligned} \tag{4.39}$$

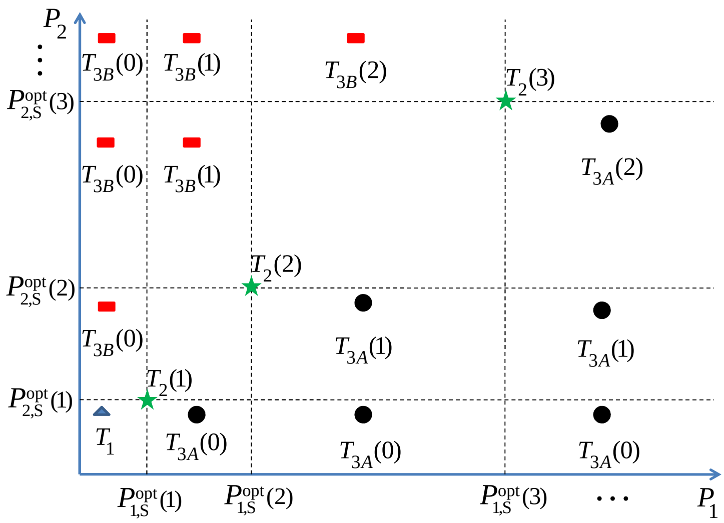

Figure 4.3: Regions in the P_1P_2-plane for which SD can achieve the sum-capacity of the strong interference class. The label associated with each point shows the theorem and the value of N corresponding to the point.

where (a) is valid since

$$\begin{aligned}(ab)^N =& 1 + P_{1,\mathrm{S}}^{\mathrm{opt}}(N) + aP_{2,\mathrm{S}}^{\mathrm{opt}}(N)\\ =& 1 + P_{2,\mathrm{S}}^{\mathrm{opt}}(N) + bP_{1,\mathrm{S}}^{\mathrm{opt}}(N),\end{aligned} \tag{4.40}$$

and (b) is valid since $\Delta P_2 < (b-1)(ab)^N$. Similarly,

$$\begin{aligned}R_2 =& NC(b-1) + \min\Bigg\{C\Big(\frac{\Delta P_2}{1 + P_{2,\mathrm{S}}^{\mathrm{opt}}(N) + bP_{1,\mathrm{S}}^{\mathrm{opt}}(N)}\Big),\\ & C\Big(\frac{a\Delta P_2}{1 + P_{1,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_1 + aP_{2,\mathrm{S}}^{\mathrm{opt}}(N)}\Big)\Bigg\}\\ \stackrel{(a)}{=}& NC(b-1) + \min\left\{C\Big(\frac{\Delta P_2}{(ab)^N}\Big), C\Big(\frac{a\Delta P_2}{(ab)^N + \Delta P_1}\Big)\right\}\\ \stackrel{(b)}{=}& NC(b-1) + C\Big(\frac{a\Delta P_2}{(ab)^N + \Delta P_1}\Big),\end{aligned} \tag{4.41}$$

where (a) is valid by (4.40), and (b) is valid because $\Delta P_1 \geq (a-1)(ab)^N$. Therefore, the following sum-rate is achievable

$$\begin{aligned}R_1 + R_2 =& NC(a-1) + NC(b-1)\\ & + C\left(\frac{\Delta P_1}{(ab)^N}\right) + C\Big(\frac{a\Delta P_2}{(ab)^N + \Delta P_1}\Big).\end{aligned} \tag{4.42}$$

Moreover, we know that SND achieves the sum-capacity of the strong interference channel. Therefore, for the values of P_1 and P_2 satisfying condition A, the sum-rate is upper-bounded by

$$\begin{aligned}
R_{\text{sum-SND}} =&\min\{C(P_2+bP_1), C(P_1+aP_2)\}\\
=&C(P_1+aP_2)\\
=&C(P_{1,\mathrm{S}}^{\mathrm{opt}}(N)+aP_{2,\mathrm{S}}^{\mathrm{opt}}(N)+\Delta P_1+a\Delta P_2)\\
=&C(P_{1,\mathrm{S}}^{\mathrm{opt}}(N)+aP_{2,\mathrm{S}}^{\mathrm{opt}}(N))+C\Big(\frac{\Delta P_1+a\Delta P_2}{1+P_{1,\mathrm{S}}^{\mathrm{opt}}(N)+aP_{2,\mathrm{S}}^{\mathrm{opt}}(N)}\Big)\\
=&NC(a-1)+NC(b-1)+C\Big(\frac{\Delta P_1+a\Delta P_2}{(ab)^N}\Big).
\end{aligned} \tag{4.43}$$

One can use (4.34) and check that (4.42) and (4.43) are equal, and this completes the proof. □

Results of Theorems 4.1, 4.2, and 4.3 describe conditions under which SD achieves the sum-capacity. These conditions can be interpreted in two ways. For fixed a and b, Figure 4.3 visualizes regions in the P_1P_2-plane for which SD achieves the sum-capacity. On the other hand, for fixed P_1 and P_2, Figure 4.4 shows regions in the ab-plane for which SD achieves the sum-capacity. For each value of N, Theorem 4.2 demonstrates a point in the P_1P_2-plane or in the ab-plane. These points are shown by stars in Figure 4.3 and Figure 4.4. For instance, the star $T_1(1)$ satisfies the condition of Theorem 4.1 for $N=1$. Theorem 4.2 describes the very strong interference region. This region is filled with a triangle, labeled T_2, in both Figure 4.3 and Figure 4.4. For each value of N, Theorem 4.3, under condition A, also describes a region. For instance, for $N=0$, Theorem 4.3 describes the region $P_1>a-1$ and $0<P_2<b-1$. For fixed values of a and b, this region is filled with three circles in Figure 4.3. These circles are labeled $T_{3A}(0)$. On the other hand, for fixed values of P_1 and P_2, this region is expressed by $a<P_1+1$ and $b>P_2+1$ and is filled with one circle labeled $T_{3A}(0)$ in Figure 4.4. The circle labeled $T_{3A}(i)$ represents a point that satisfies condition A of Theorem 4.3 for $N=i$. Figure 4.3 and Figure 4.4 show the regions characterized by Theorem 4.3A only for $N\in\{0,1,2\}$. Similarly, the regions characterized by Theorem 4.3B for $N\in\{0,1,2\}$ are demonstrated in Figure 4.3 and Figure 4.4 and are filled with rectangles. The rectangle labeled $T_{3B}(i)$ represents a point that satisfies condition B of Theorem 4.3 for $N=i$.

Next, we summarize the results of Theorem 4.3. In Theorem 4.3, we proposed a novel coding scheme, and we showed that, for a wide range of (a,b,P_1,P_2), our scheme achieves the sum-capacity. Let $R_{\text{sum-SD}}$ represent the achievable sum-rate of this scheme.

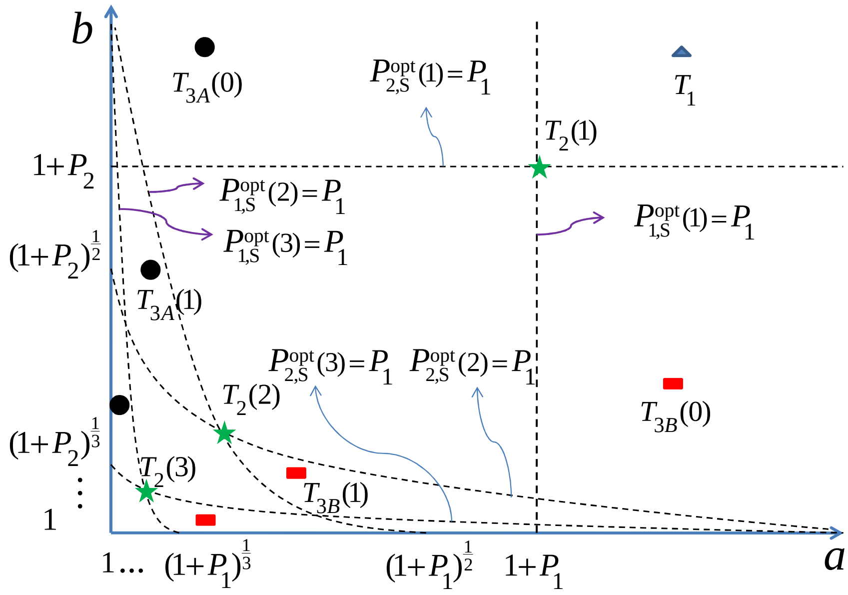

Figure 4.4: Regions in the ab-plane for which SD can achieve the sum-capacity of the strong interference class. The label associated with each point shows the theorem and the value of N that corresponds to the point.

In the following, we explicitly characterizer $R_{\text{sum-SD}}$. Consider a pair of power allocation (P_1, P_2). We can uniquely determine (P_1, P_2) as follows:

$$\begin{aligned} P_1 &= P_{1,\text{S}}^{\text{opt}}(N) + \Delta P_1, \\ P_2 &= P_{2,\text{S}}^{\text{opt}}(N) + \Delta P_2, \end{aligned} \tag{4.44}$$

where N is the greatest non-negative integer such that $\Delta P_1 \geq 0$ and $\Delta P_2 \geq 0$. Note that N, ΔP_1, and ΔP_2 are unique. Then, by dividing each message into $N+1$ parts, the following sum-rate is achievable by the scheme proposed in Theorem 4.3.

$$R_{\text{sum-SD}} = NC(a-1) + NC(b-1) + R_{\text{sum}}^{N+1}, \tag{4.45}$$

where $R_{\text{sum}}^{N+1} \doteq R_1^{N+1} + R_2^{N+1}$ is given by

$$R_{\text{sum}}^{N+1} =$$

$$\begin{cases} C\left(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\right) & \text{if } \Delta P_1 \geq (a-1)(ab)^N, \Delta P_2 \leq (b-1)(ab)^N, \\ C\left(\frac{\Delta P_2 + b\Delta P_1}{(ab)^N}\right) & \text{if } \Delta P_1 \leq (a-1)(ab)^N, \Delta P_2 \geq (b-1)(ab)^N, \\ C\left(\frac{\Delta P_1}{(ab)^N}\right) + C\left(\frac{\Delta P_2}{(ab)^N}\right) & \text{if } \Delta P_1 \leq (a-1)(ab)^N, \Delta P_2 \leq (b-1)(ab)^N. \end{cases} \tag{4.46}$$

The first line of (4.46) is exactly equivalent to the condition A of Theorem 4.3. Similarly, the second line of (4.46) is equivalent to the condition B of Theorem 4.3. The third line shows the case, in which SD does not achieve the sum-capacity.

The proof of achievability of the third line follows similar to (4.39). In fact, one can see that SD achieves the following rates:

$$\begin{aligned} R_1 =& NC(a-1) + C\Big(\frac{\Delta P_1}{(ab)^N}\Big), \\ R_2 =& NC(b-1) + C\Big(\frac{\Delta P_2}{(ab)^N}\Big). \end{aligned} \tag{4.47}$$

4.3.3 Maximum Sum-Rate Loss

According to Figure 4.3, the only regions in the P_1P_2-plane for which sum-capacity is not achieved using SD are as follows:

$$\begin{aligned} &P_1 = P_{1,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_1, 0 < \Delta P_1 < (a-1)(ab)^N, \\ &P_2 = P_{2,\mathrm{S}}^{\mathrm{opt}}(N) + \Delta P_2, 0 < \Delta P_2 < (b-1)(ab)^N, \\ &N \geq 1. \end{aligned} \tag{4.48}$$

A natural question is the maximum difference between the optimal sum-rate and the sum-rate achieved using SD. Interestingly, the next theorem shows that the maximum sum-rate difference is only a function of channel gains, i.e., a and b, and does not depend on the number of splits $N+1$.

Theorem 4.4. *For the two-user GIC with strong interference, if joint decoding is replaced by SD, the maximum sum-rate loss is given by* $\Delta R_{\mathrm{sum}}^{\max} = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$.

Proof. First, note that $\Delta R_{\mathrm{sum}}^{\max}$ represents the maximum difference between C_{sum} and $R_{\text{sum-SD}}$. Since, for the strong interference class $C_{\mathrm{sum}} = R_{\text{sum-SND}}$, $\Delta R_{\mathrm{sum}}^{\max}$ is given by

$$\Delta R_{\mathrm{sum}}^{\max} \doteq \max_{P_1>0, P_2>0} \Big(R_{\text{sum-SND}} - R_{\text{sum-SD}}\Big). \tag{4.49}$$

Second, if P_1 and P_2 are not in the region described by (4.48), then Theorems 4.1, 4.2, and 4.3 show that SD is sum-rate optimal and there is no sum-rate loss. If P_1 and

P_2 belong to the region described by (4.48), the sum-rate of our proposed SD, $R_{\text{sum-SD}}$, the sum-rate of SND, $R_{\text{sum-SND}}$, and the sum-rate difference, ΔR_{sum}, are as follows:

$$\begin{aligned}
R_{\text{sum-SD}} =& NC(a-1) + NC(b-1) + C\Big(\frac{\Delta P_1}{(ab)^N}\Big) + C\Big(\frac{\Delta P_2}{(ab)^N}\Big),\\
R_{\text{sum-SND}} =& NC(a-1) + NC(b-1)\\
&+ \min\Big\{C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\Big), C\Big(\frac{\Delta P_2 + b\Delta P_1}{(ab)^N}\Big)\Big\},\\
\Delta R_{\text{sum}}^N \doteq& R_{\text{sum-SND}} - R_{\text{sum-SD}}\\
=&\min\Big\{C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\Big), C\Big(\frac{\Delta P_2 + b\Delta P_1}{(ab)^N}\Big)\Big\}\\
&- C\Big(\frac{\Delta P_1}{(ab)^N}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big).
\end{aligned} \tag{4.50}$$

Therefore, to find the maximum sum-rate loss, the following optimization problem is solved.

$$\begin{aligned}
\Delta R_{\text{sum}}^{\max} &= \max_{\Delta P_1, \Delta P_2} \Delta R_{\text{sum}}^N,\\
\text{subject to } & 0 \leq \Delta P_1 \leq (a-1)(ab)^N, N \geq 1,\\
& 0 \leq \Delta P_2 \leq (b-1)(ab)^N, N \geq 1.
\end{aligned} \tag{4.51}$$

Let us review an optimization technique. According to interior extremum theorem, the global maximum of a differentiable function f over a feasible region $\mathcal{A}$ is achieved at one of the following points: an stationary point or a boundary point [45, 46]. In particular, consider the function $\Delta R_{\text{sum}}^N(\Delta P_1, \Delta P_2)$, defined in (4.50). First note that this function is not necessarily differentiable. The function $\min\{\}$ can make $\Delta R_{\text{sum}}^N(\Delta P_1, \Delta P_2)$ non-differentiable. However, $\Delta R_{\text{sum}}^N(\Delta P_1, \Delta P_2)$ can be non-differentiable only if

$$\begin{aligned}
C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\Big) =& C\Big(\frac{\Delta P_2 + b\Delta P_1}{(ab)^N}\Big)\\
\Rightarrow (a-1)\Delta P_2 =& (b-1)\Delta P_1.
\end{aligned} \tag{4.52}$$

Consequently, all non-differentiable points of the function $\Delta R_{\text{sum}}^N(\Delta P_1, \Delta P_2)$ lie on $(a-1)\Delta P_2 = (b-1)\Delta P_1$.

The feasible region of the optimization problem (4.51) is a rectangle, as shown in Figure 4.5. Observe that $(a-1)\Delta P_2 = (b-1)\Delta P_1$ is a line inside the feasible region that

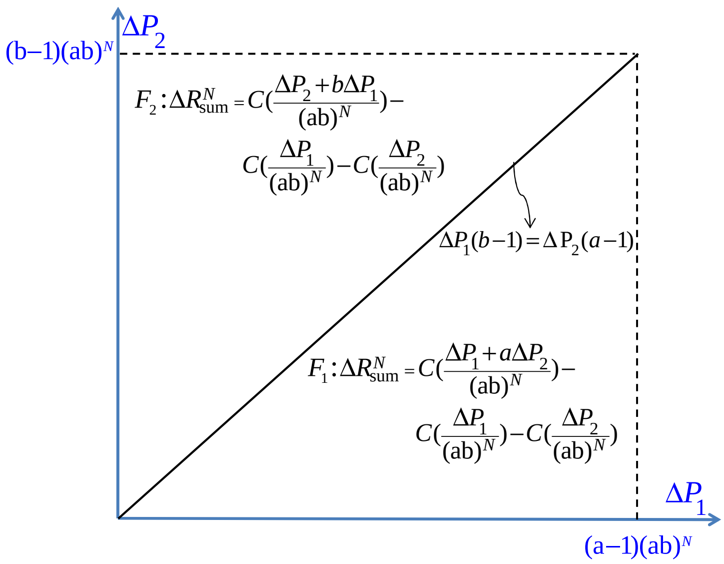

Figure 4.5: The feasible region of the optimization problem (4.51).

divides the feasible region into two parts, namely F_1 and F_2, where

$$\begin{aligned} F_1 = \{(\Delta P_1, \Delta P_2) :& 0 \leq \Delta P_1 \leq (a-1)(ab)^N, \\ & 0 \leq \Delta P_2 \leq (b-1)(ab)^N, \\ & (a-1)\Delta P_2 \leq (b-1)\Delta P_1\}, \end{aligned} \tag{4.53}$$

$$\begin{aligned} F_2 = \{(\Delta P_1, \Delta P_2) :& 0 \leq \Delta P_1 \leq (a-1)(ab)^N, \\ & 0 \leq \Delta P_2 \leq (b-1)(ab)^N, \\ & (a-1)\Delta P_2 \geq (b-1)\Delta P_1\}. \end{aligned} \tag{4.54}$$

We solve the optimization problem (4.51) in three steps. First, we find the optimal solution over F_1. Second, we find the optimal solution over F_2. Finally, we compare the results together. To do so, we first solve the following problem

$$\begin{aligned} \Delta R_{\text{sum}}^{\max} &= \max_{\Delta P_1, \Delta P_2} \Delta R_{\text{sum}}^N, \\ &\text{subject to } (\Delta P_1, \Delta P_2) \in F_1. \end{aligned} \tag{4.55}$$

Inside F_1, ΔR_{sum}^N is a differentiable function. According to interior extremum theorem, the optimal solution of (4.51) is either an stationary point, or a point over the boundary.

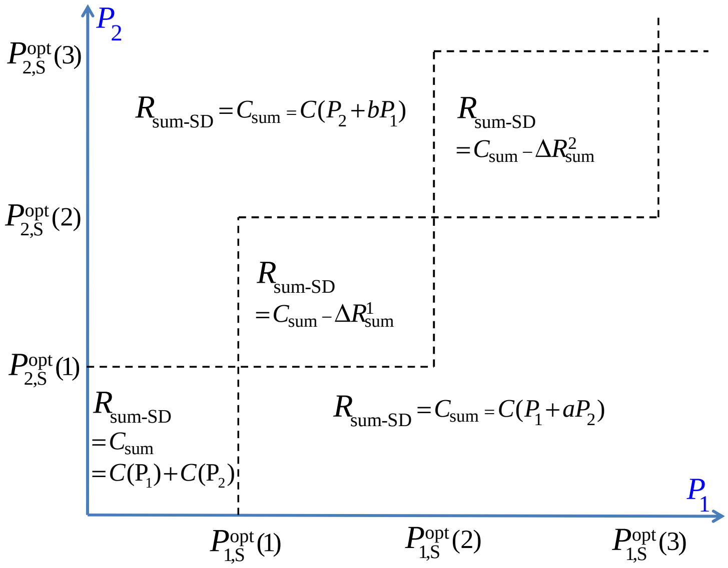

Figure 4.6: Comparison of the achievable sum-rate $R_{\text{sum-SD}}$ with the sum-capacity.

We can see that the function ΔR_{sum}^N has no stationary points.

$$\begin{aligned}\Delta R_{\text{sum}}^N =& \min\left\{C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\Big), C\Big(\frac{\Delta P_2 + b\Delta P_1}{(ab)^N}\Big)\right\} \\ & - C\Big(\frac{\Delta P_1}{(ab)^N}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big) \\ =& C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^N}\Big) - C\Big(\frac{\Delta P_1}{(ab)^N}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big) \\ =& C\Big(\frac{a\Delta P_2}{(ab)^N + \Delta P_1}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big). \end{aligned} \tag{4.56}$$

(4.56) shows that $\Delta R_{\text{sum}}^N(\Delta P_1, \Delta P_2)$ is a decreasing function of ΔP_1. Therefore, ΔR_{sum}^N has no stationary points.

To investigate the boundary, first note that F_1 is a right triangle. Over the two legs of the right angle, we have

$$\Delta R_{\text{sum}}^N = R_{\text{sum-SND}} - R_{\text{sum-SD}} \overset{(a)}{=} 0, \tag{4.57}$$

where (a) is valid by Theorem 4.3. Consequently, ΔR_{sum}^N achieves its maximum over the line

$$(b-1)\Delta P_1 = (a-1)\Delta P_2. \tag{4.58}$$

In fact, by letting the derivatives equal zero, we find the following point that maximizes the sum-rate loss over $(b-1)\Delta P_1 = (a-1)\Delta P_2$:

$$\begin{aligned}\frac{\partial}{\partial \Delta P_2}\left(C\Big(\frac{a\Delta P_2}{(ab)^N + (\Delta P_2)\frac{a-1}{b-1}}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big)\right) =&0\\ \Rightarrow \Delta P_2^{\text{opt}} =& \frac{(\sqrt{ab}-1)(b-1)(ab)^N}{(ab-1)},\end{aligned} \tag{4.59}$$

Moreover, since $(b-1)\Delta P_1 = (a-1)\Delta P_2$, we have

$$\Delta P_1^{\text{opt}} = \frac{(\sqrt{ab}-1)(a-1)(ab)^N}{(ab-1)}. \tag{4.60}$$

Inserting (4.60) and (4.59), into (4.56), we see that

$$\begin{aligned}\Delta R_{\text{sum}}^N &\doteq R_{\text{sum-SND}} - R_{\text{sum-SD}}\\ &= C\Big(\frac{a\Delta P_2}{(ab)^N + \Delta P_1}\Big) - C\Big(\frac{\Delta P_2}{(ab)^N}\Big)\\ &\overset{(a)}{=} C\Big(\frac{b-1}{1+\sqrt{\frac{b}{a}}}\Big) - C\Big(\frac{b-1}{1+\sqrt{ab}}\Big)\\ &= \frac{1}{2}\log\Big(1+\frac{b-1}{1+\sqrt{\frac{b}{a}}}\Big) - \frac{1}{2}\log\Big(1+\frac{b-1}{1+\sqrt{ab}}\Big)\\ &= \frac{1}{2}\log\Big(\frac{(1+\sqrt{ab})^2}{(\sqrt{a}+\sqrt{b})^2}\Big)\\ &= \log\Big(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\Big),\end{aligned} \tag{4.61}$$

where (a) is valid by (4.60) and (4.59).

Similarly, one can show that over F_2, the optimal solution that maximizes ΔR_{sum}^N is given by (4.60). Therefore, (4.60) represents the optimal solution of the original problem (4.51), and (4.61) represents the maximum sum-rate loss, as claimed in Theorem 4.4.

Note that (4.60) and (4.59) show the optimal solution $(\Delta P_1^{\text{opt}}, \Delta P_2^{\text{opt}})$ that maximizes the optimization problem (4.51) and the value of the maximum sum-rate loss is given by (4.61). Moreover, ΔP_1^{opt} and ΔP_2^{opt} are functions of N, whereas the maximum sum-rate loss is not. This means, for each $N \geq 1$, there is exactly one pair of $(\Delta P_1^{\text{opt}}, \Delta P_2^{\text{opt}})$, and for all $N \geq 1$, these pairs result in the same maximum sum-rate loss. $\square$

Theorems 4.2-4.3 show that for a wide range of (a, b, P_1, P_2), $C_{\text{sum}} - R_{\text{sum-SD}} = 0$. Theorem 4.4 shows that, for values of (a, b, P_1, P_2) that $C_{\text{sum}} - R_{\text{sum-SD}} > 0$, we know that $C_{\text{sum}} - R_{\text{sum-SD}}$ is bounded. Figures 4.6 compares C_{sum} with $R_{\text{sum-SD}}$.

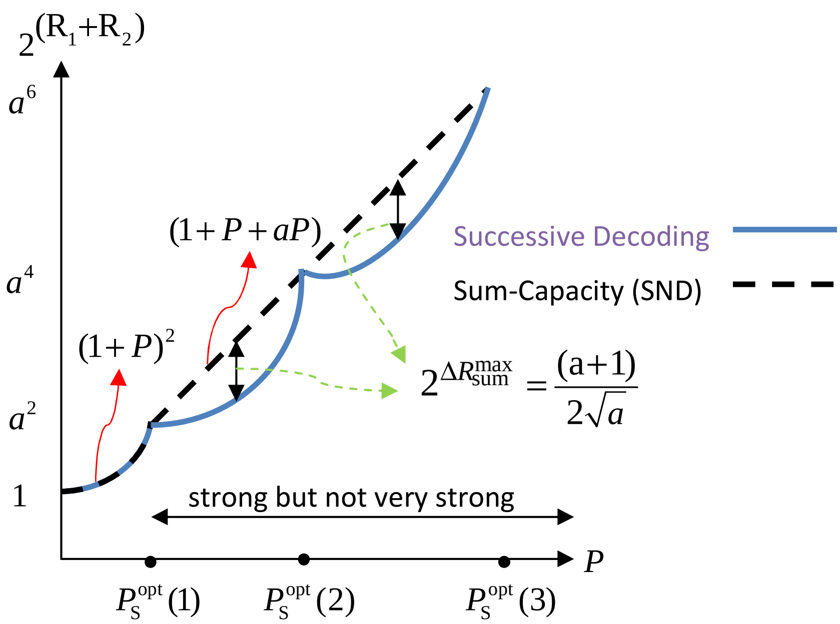

Figure 4.7: Comparison of the sum-capacity and the sum-rate achieved using SD for the symmetric two-user GIC with strong interference.

Moreover, we demonstrate the results of previous theorems by considering the symmetric Gaussian interference channel in which $P_1 = P_2 = P$ and $a = b$. Figure 4.7 investigates the strong interference class and compares the sum-rate achieved using our proposed SD and the sum-capacity achieved using SND. It shows that when interference is very strong, i.e., $P \leq a-1$, SD achieves the sum-capacity. When interference is strong but not very strong, if $P = P_{\text{S}}^{\text{opt}}(N) \doteq \frac{a-1}{a^2-1}(a^{2N}-1)$, SD still achieves the sum-capacity. Moreover, Figure 4.7 depicts the sum-rate loss when the proposed SD scheme is used. In fact, according to Theorem 4.4, the maximum sum-rate loss equals $\log\left(\frac{a+1}{2\sqrt{a}}\right)$ and does not depend on P. Figure 4.7 shows that this maximum loss is seen exactly once in every interval $\left(P_{\text{S}}^{\text{opt}}(N), P_{\text{S}}^{\text{opt}}(N+1)\right)$.

4.4 Weak Interference Class

In this section, we investigate the weak interference class. The weak interference class is more challenging than the strong interference class. The sum-capacity of the weak interference class is unknown. For the strong interference class, the maximum HK sum-rate is achieved by decoding the entire interference at both receivers. For the weak interference class, [43] shows that to achieve the maximum HK sum-rate, a specific portion of the interference should be decoded by each receiver. This portion varies as (a, b, P_1, P_2) varies inside the weak interference class. For the strong interference class, a fixed decoding order, given in (4.16), achieves the sum-capacity for a wide range of transmitters' powers. For the weak interference, we show that different decoding orders should be used, depending on the value of (a, b, P_1, P_2).

The structure of this section is as follows. We first show that, without any RS and joint decoding, the maximum sum-rate of the HK scheme is achievable for a wide range of (a, b, P_1, P_2). Second, to achieve the maximum sum-rate of the HK scheme for a wider range of (a, b, P_1, P_2), we propose a novel scheme in which both transmitters divide their messages into some parts.

4.4.1 Is Rate Splitting Required?

We calculate the achievable sum-rate when no RS is used. Our main goal is to show that, for a wide range of (a, b, P_1, P_2), RS is not required. In doing so, we first solve the optimization problem (4.6) for $N_1 = N_2 = 1$. Then we compare the result with the maximum achievable sum-rate of the HK scheme.

Theorem 4.5. *For the two-user GIC with weak interference, the maximum sum-rate achieved with no rate splitting is given by*

$$R_{\text{sum}}^{\text{NRS}} = \max\Big\{C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big), \\ C(P_1 + aP_2), \quad C(P_2 + bP_1)\Big\}. \tag{4.62}$$

Proof. When no RS is used, we have $N_1 = N_2 = 1$. Therefor, the optimization (4.6) reduces to

$$R_{\text{sum}}^{\text{NRS}} \doteq \max_{\mathbf{S_1},\mathbf{S_2}} R_1 + R_2. \tag{4.63}$$

There exists four possibilities for $\mathbf{S_1}$ and $\mathbf{S_2}$, as shown in Table 4.2. This table shows the achievable sum-rate corresponding to the four possible decoding orders. In the first case, both receivers treat the interference as noise. In the second case, the first receiver decodes the interference, while the second receiver treats the interference as noise. In

Decoding order $\mathbf{S_1}$	Decoding order $\mathbf{S_2}$	$R_1 + R_2$
$(\sqrt{a}X_2, X_1)$	$(\sqrt{b}X_1, X_2)$	$C(\frac{P_1}{1+aP_2}) + C(\frac{P_2}{1+bP_1})$
$(X_1, \sqrt{a}X_2)$	$(\sqrt{b}X_1, X_2)$	$C(P_1) + \min\{C(\frac{P_2}{1+bP_1}), C(\frac{aP_2}{1+P_1})\}$ $= C(P_1 + aP_2)$
$(\sqrt{a}X_2, X_1)$	$(X_2, \sqrt{b}X_1)$	$\min\{C(\frac{P_1}{1+aP_2}), C(\frac{bP_1}{1+P_2})\}+$ $C(P_2) \qquad = C(P_2 + bP_1)$
$(X_1, \sqrt{a}X_2)$	$(X_2, \sqrt{b}X_1)$	$\min\{C(P_1), C(\frac{bP_1}{1+P_2})\} +$ $\min\{C(P_2), C(\frac{aP_2}{1+P_1})\}$ $= C(\frac{bP_1}{1+P_2}) + C(\frac{aP_2}{1+P_1})$

Table 4.2: The achievable sum-rate of the weak interference class corresponding to four decoding orders.

other words, since $\mathbf{S_2} = (\sqrt{b}X_1, X_2)$, the second receiver does not decode X_1. Consequently, R_1 is "not" required to be smaller than $C(bP_1)$. In fact, $R_1 = C(P_1)$ and $R_2 = \min\{C(\frac{P_2}{1+bP_1}), C(\frac{aP_2}{1+P_1})\} = C(\frac{aP_2}{1+P_1})$, and therefore, $R_1 + R_2 = C(P_1 + aP_2)$. In the third case, the second receiver decodes the interference, while the first receiver treats the interference as noise. Therefore, we have $R_1 + R_2 = C(P_2 + bP_1)$. In the fourth case, both receivers decode the interference. Note that the sum-rate corresponding to this order is smaller than the sum-rate achieved by other decoding orders. Therefore, the maximum achievable sum-rate is the maximum of the three rate expressions corresponding to the first three decoding orders, as stated in (4.62).

This completes the proof. $\square$

Remark 4.1. *The sum-rate achieved by $R_{\text{sum}}^{\text{NRS}}$ is greater than the sum-rate achieved using SND: For the weak interference class, $R_{\textit{sum-SND}}$ is given by*

$$R_{\textit{sum-SND}} = \min\Big\{C(P_1 + aP_2), C(P_2 + bP_1), C(aP_2) + C(bP_1)\Big\}. \tag{4.64}$$

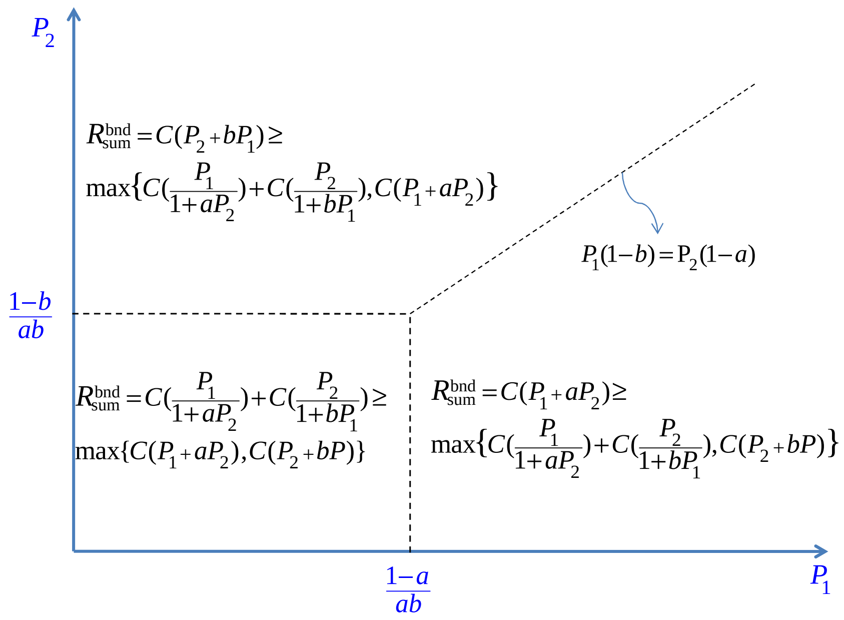

Figure 4.8: The maximum achievable sum-rate when rate splitting is not used: Quadrant I of the P_1P_2-plane is partitioned into three regions. In each region, $R_{\text{sum}}^{\text{NRS}}$ is demonstrated.

For the weak interference class, this sum-rate is smaller than $R_{\text{sum}}^{\text{NRS}}$ given in (4.62). Therefore, although SND achieves the sum-capacity for the strong interference class, SND fails to achieve $R_{\text{sum}}^{\text{NRS}}$ for the weak interference class.

Figure 4.8 shows quadrant I of the P_1P_2-plane. This quadrant is divided into three regions. In each region, exactly one of $C\left(\frac{P_1}{1+aP_2}\right)+C\left(\frac{P_2}{1+bP_1}\right)$, $C(P_1+aP_2)$, and $C(P_2+bP_1)$ is greater than the others, as shown in the figure. Note that the region in which $R_{\text{sum}}^{\text{NRS}}$ equals $C(P_1 + aP_2)$ and the region in which $R_{\text{sum}}^{\text{NRS}}$ equals $C(P_2 + bP_1)$ are separated by the line $P_1(1 - b) = P_2(1 - a)$.

The main goal of this section is to find out when RS is required. To this end, we need to compare $R_{\text{sum}}^{\text{NRS}}$ with the maximum sum-rate of the HK scheme with Gaussian inputs, denoted by $R_{\text{sum-HK}}^{\max}$. We have

$$R_{\text{sum}}^{\text{NRS}} \leq R_{\text{sum-SD}}^{\text{opt}} \leq R_{\text{sum-HK}}^{\max}. \tag{4.65}$$

Therefore, wherever we have $R_{\text{sum}}^{\text{NRS}} = R_{\text{sum-HK}}^{\max}$, we have found an optimal solution of the optimization problem (4.6).

The maximum sum-rate of the HK scheme with Gaussian inputs, $R_{\text{sum-HK}}^{\max}$, was characterized in Chapter 2. In the following theorem, we review this characterization. To make comparison simpler, we use a slightly different notation here.

Theorem 4.6. *For the two-user Gaussian interference channel, when interference is weak, let $R^{\max}_{\text{sum-HK}}$ denote the maximum achievable sum-rate of the HK scheme with Gaussian inputs, without time sharing. Then $R^{\max}_{\text{sum-HK}}$ is given by*

$$\begin{aligned} &R^{\max}_{\text{sum-HK}}(P_1, P_2) = \\ &\max\Big\{C\Big(\frac{P_1}{1+aP_2}\Big) + C\Big(\frac{P_2}{1+bP_1}\Big), \\ &\qquad C(P_1 + aP_2), \\ &\qquad C(P_2 + bP_1), \\ &\qquad C(P_1 + aP_2) + g_1(\tilde{\lambda}_1, \tilde{\lambda}_2)\mathbb{1}(\tilde{\lambda}_1 \geq 0, \tilde{\lambda}_2 \geq 0), \\ &\qquad C(P_1 + aP_2) + g_1(\hat{\lambda}_1, \hat{\lambda}_2)\mathbb{1}(\hat{\lambda}_1 \geq 0, \hat{\lambda}_2 \geq 0, \tilde{\lambda}_2 \geq \hat{\lambda}_2)\Big\}, \end{aligned} \tag{4.66}$$

where

$$g_1(\lambda_1, \lambda_2) \doteq C\Big(\frac{(1-a)\lambda_2 P_2 + b\lambda_1 P_1}{1 + a\lambda_2 P_2}\Big) - C(b\lambda_1 P_1), \tag{4.67}$$

$$(\tilde{\lambda}_1, \tilde{\lambda}_2) \doteq \Big(ab - \frac{1-a}{P_1}, ab - \frac{1-b}{P_2}\Big), \tag{4.68}$$

$$\hat{\lambda}_2 \doteq \frac{1 + bP_1 c}{bP_1\alpha + P_2}\Big(-1 + \sqrt{1 + \frac{(bP_1\alpha + P_2)(1 - abP_1 c - a)}{(1 + bP_1 c)(abP_1\alpha)}}\Big), \tag{4.69}$$

$$\hat{\lambda}_1 \doteq \alpha\hat{\lambda}_2 + c, \tag{4.70}$$

$$c \doteq \frac{P_1(1-b) - P_2(1-a)}{P_1\big(1 - b + P_2(1 - ab)\big)}, \tag{4.71}$$

$$\alpha \doteq 1 - c. \tag{4.72}$$

As expected, (4.66) shows that $R^{\text{NRS}}_{\text{sum}} \leq R^{\max}_{\text{sum-HK}}$. More importantly, Theorem 4.6 shows that the maximum HK sum-rate has five distinct mathematical expressions, depending on the value of (a, b, P_1, P_2). Table 4.3 partitions the entire weak interference class into five sub-classes, namely A, B, C, D, and E. For each sub-class, the maximum H-K sum-rate is demonstrated. By comparing $R^{\text{NRS}}_{\text{sum}}$ with $R^{\max}_{\text{sum-HK}}$, we characterize three sub-classes inside the weak interfere class, for which $R^{\text{NRS}}_{\text{sum}} = R^{\max}_{\text{sum-HK}}$, as explained in the following theorem.

Theorem 4.7. *For the two-user GIC with weak interference, if (a, b, P_1, P_2) belongs to the union of sub-classes A, B, and C, then $R^{\text{NRS}}_{\text{sum}} = R^{\max}_{\text{sum-HK}}$.*

Proof. Theorem 4.5 and Theorem 4.6 characterize $R^{\text{NRS}}_{\text{sum}}$ and $R^{\max}_{\text{sum-HK}}$, respectively. Table 4.3 partitions the weak interference class into five sub-classes. For each sub-class, we

Sub-class	Description	$R^{\max}_{\text{sum-HK}}$	$R^{\text{NRS}}_{\text{sum}}$
A	$0 \le P_1 \le \frac{1-a}{ab}$, $0 \le P_2 \le \frac{1-b}{ab}$.	$C(\frac{P_1}{1+aP_2})+ C(\frac{P_2}{1+bP_1})$	$C(\frac{P_1}{1+aP_2})+ C(\frac{P_2}{1+bP_1})$
B	$P_1 > \frac{1-a}{ab}$, $0 \le P_2 \le \max\{\frac{1-b}{ab}, \frac{(1-b)ab}{1-a}P_1 + b - 1\}$	$C(P_1 + aP_2)$	$C(P_1 + aP_2)$
C	$P_2 > \frac{1-b}{ab}$, $0 \le P_1 \le \max\{\frac{1-a}{ab}, \frac{(1-a)ab}{1-b}P_2 + a - 1\}$	$C(P_2 + bP_1)$	$C(P_2 + bP_1)$
D	$P_1 > \frac{1-a}{ab}$, $P_2 > \frac{1-b}{ab}$, $\hat{\lambda}_2 > ab - \frac{1-b}{P_2}$	$C(P_1 + aP_2)$ $+g_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$	$\max\{$ $C(P_1+aP_2)$, $C(P_2+bP_1)\}$
E	$P_1 > \frac{(1-a)ab}{1-b}P_2+a-1$, $P_2 > \frac{(1-b)ab}{1-b}P_1+b-1$, $\hat{\lambda}_2 \le ab - \frac{1-b}{P_2}$	$C(P_1 + aP_2)$ $+g_1(\hat{\lambda}_1, \hat{\lambda}_2)$	$\max\{$ $C(P_1+aP_2)$, $C(P_2+bP_1)\}$

Table 4.3: The weak interference class is partitioned into five sub-classes. For each sub-class, $R^{\max}_{\text{sum-HK}}$ is compared with $R^{\text{NRS}}_{\text{sum}}$.

can compare $R^{\text{NRS}}_{\text{sum}}$ with $R^{\max}_{\text{sum-HK}}$, as shown in Table 4.3. For the first three sub-classes, we have $R^{\text{NRS}}_{\text{sum}} = R^{\max}_{\text{sum-HK}}$. This completes the proof. □

Note that the sum-capacity of the weak interference channel is not known in general. For the small sub-class of the very weak interference, characterized by $P_1\sqrt{b} + P_2\sqrt{a} \le \frac{1-\sqrt{a}-\sqrt{b}}{\sqrt{ab}}$, treating interference as noise is sum-rate optimal [10, 11]. This sub-class is strictly inside sub-class A. Therefore, for the very weak interference sub-class, $R^{\text{NRS}}_{\text{sum}}$ achieves the sum-capacity, as stated in the following corollary.

Corollary 4.2. *For the two-user GIC with very weak interference, $R^{\text{NRS}}_{\text{sum}}$ achieves the*

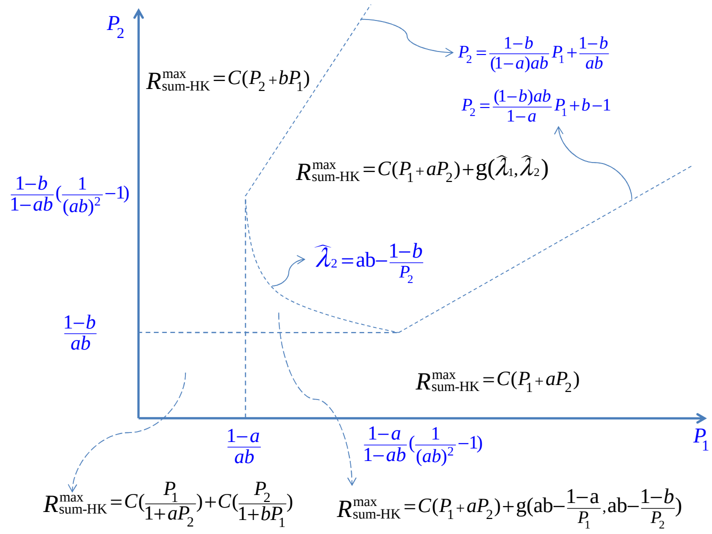

Figure 4.9: The weak interference class is partitioned into five sub-classes. For each suc-class, $\Delta R_{\text{sum}} \doteq R^{\max}_{\text{sum-HK}} - R^{\text{NRS}}_{\text{sum}}$ is demonstrated.

sum-capacity.

Moreover, Figure 4.9 shows that the entire weak interference class is partitioned into five sub-classes. For each sub-class, $\Delta R_{\text{sum}} \doteq R^{\max}_{\text{sum-HK}} - R^{\text{NRS}}_{\text{sum}}$ is shown. For two sub-classes, namely D and E, we have $\Delta R_{\text{sum}} \geq 0$. To characterize ΔR_{sum}, let us define

$$g_2(\lambda_1, \lambda_2) \doteq C(\frac{(1-b)\lambda_1 P_1 + a\lambda_2 P_2}{1 + b\lambda_1 P_1}) - C(a\lambda_2 P_2). \tag{4.73}$$

Using direct calculation, one can show that

$$C(P_1 + aP_2) + g_1(\lambda_1, \lambda_2) = C(P_2 + bP_1) + g_2(\lambda_1, \lambda_2). \tag{4.74}$$

Consequently, for sub-class D, ΔR_{sum} is given by

$$\begin{aligned}\Delta R_{\text{sum}} =& C(P_1 + aP_2) + g_1(\tilde{\lambda}_1, \tilde{\lambda}_2) - \max\{C(P_1 + aP_2), C(P_2 + bP_1)\}\\ =& C(P_2 + bP_1) + g_2(\tilde{\lambda}_1, \tilde{\lambda}_2) - \max\{C(P_1 + aP_2), C(P_2 + bP_1)\}\\ =& \min\{g_1(\tilde{\lambda}_1, \tilde{\lambda}_2), g_2(\tilde{\lambda}_1, \tilde{\lambda}_2)\}.\end{aligned} \tag{4.75}$$

Similarly, for sub-class E, ΔR_{sum} is given by

$$\Delta R_{\text{sum}} = \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{4.76}$$

In the next sub-section, we propose a novel coding scheme based on RS and SD that achieves a sum-rate better than $R^{\text{NRS}}_{\text{sum}}$. We show that the proposed scheme achieves $R^{\max}_{\text{sum-HK}}$, for sub-classes A, B, C, and D.

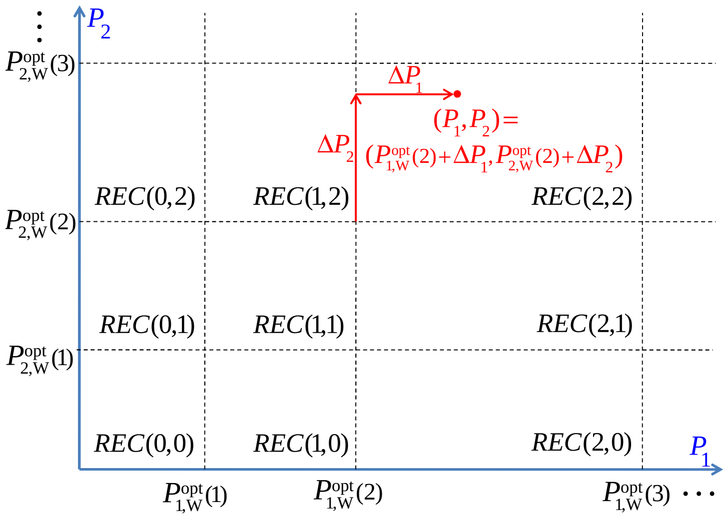

Figure 4.10: Quadrant I of the P_1P_2-plane is partitioned into rectangles. Each rectangle determines the decoding orders $(\mathbf{S}_1, \mathbf{S}_2)$ and the number of splits $(N + 1)$.

4.4.2 How Many Splits Are Required?

In this sub-section, we propose a novel coding scheme that divides both messages into $N + 1$ parts. We show that to achieve the HK sum-rate, N should be properly chosen. In fact, the number of splits depends on the value of (P_1, P_2). Note that (P_1, P_2) is a point in the first quadrant of $\mathbb{R}^2_+$. We partition the entire $\mathbb{R}^2_+$ into rectangles, as shown in Figure 4.10. The point (P_1, P_2) lies in one of these rectangles, denoted by $REC(m, n)$. As demonstrated in Figure 4.10, the partitioning is created by vertical lines $P_1 = P^{\text{opt}}_{1,\text{W}}(N)$ and horizontal lines $P_2 = P^{\text{opt}}_{2,\text{W}}(N)$, where

$$P^{\text{opt}}_{1,\text{W}}(N) \doteq \frac{1-a}{1-ab}\left(\frac{1}{(ab)^N} - 1\right), \tag{4.77}$$

$$P^{\text{opt}}_{2,\text{W}}(N) \doteq \frac{1-b}{1-ab}\left(\frac{1}{(ab)^N} - 1\right). \tag{4.78}$$

If the point (P_1, P_2) lies on $REC(m, n)$, we let

$$N = \min\{m, n\}, \tag{4.79}$$

and divide each message into $N+1$ parts. According to (4.79), N is a function of (P_1, P_2). In other words, (P_1, P_2) determines the number of splits. There is a close relation between the partitions introduced in Table 4.3 and the rectangles $REC(m, n)$. For instance, sub-class A is exactly $REC(0, 0)$, and sub-class D is a part of $REC(1, 1)$. This relation is demonstrated in Figure 4.11.

We represent (P_1, P_2) as follows:

$$(P_1, P_2) = (P_{1,\mathrm{W}}^{\mathrm{opt}}(N) + \Delta P_1, P_{2,\mathrm{W}}^{\mathrm{opt}}(N) + \Delta P_2), \tag{4.80}$$

as shown in Figure 4.10. Note that according to (4.80), each (P_1, P_2) has a unique representation. Furthermore, according to the value of m and n, we have the following constraints on ΔP_1 and ΔP_2:

If $m = n$, then we have

$$\Delta P_1 < P_{1,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{1,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-a}{(ab)^{N+1}}, \tag{4.81}$$

$$\Delta P_2 < P_{2,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{2,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-b}{(ab)^{N+1}}. \tag{4.82}$$

If $m > n$, then we have

$$\Delta P_1 > P_{1,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{1,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-a}{(ab)^{N+1}}, \tag{4.83}$$

$$\Delta P_2 < P_{2,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{2,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-b}{(ab)^{N+1}}. \tag{4.84}$$

If $m < n$, then we have

$$\Delta P_1 < P_{1,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{1,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-a}{(ab)^{N+1}}, \tag{4.85}$$

$$\Delta P_2 > P_{2,\mathrm{W}}^{\mathrm{opt}}(N+1) - P_{2,\mathrm{W}}^{\mathrm{opt}}(N) = \frac{1-b}{(ab)^{N+1}}. \tag{4.86}$$

We already noticed in Theorem 4.5, that the optimal decoding orders $\mathbf{S}_1$ and $\mathbf{S}_2$ depend on the value of (P_1, P_2). Figure 4.8 shows that depending on the value of (P_1, P_2), three different decoding orders can be optimal. Relying on this idea, we use the following three decoding orders, based on the value of (P_1, P_2).

If (P_1, P_2) lies on $REC(m, n)$, then $\mathbf{S}_1 = (S_1^1, S_1^2, S_1^3, S_1^4, ...)$ and $\mathbf{S}_2 = (S_2^1, S_2^2, S_2^3, S_2^4, ...)$ are given by the following:

If $m = n$, then we let

$$\begin{aligned}\mathbf{S}_1 &= (\sqrt{a}X_2^1, X_1^1, \sqrt{a}X_2^2, X_1^2, ..., \sqrt{a}X_2^{N+1}, X_1^{N+1}),\\ \mathbf{S}_2 &= (\sqrt{b}X_1^1, X_2^1, \sqrt{b}X_1^2, X_2^2, ..., \sqrt{b}X_1^{N+1}, X_2^{N+1}).\end{aligned} \tag{4.87}$$

If $m > n$, then we let

$$\begin{aligned}\mathbf{S}_1 &= (X_1^1, \sqrt{a}X_2^1, \sqrt{a}X_2^2, X_1^2, ..., \sqrt{a}X_2^{N+1}, X_1^{N+1}),\\ \mathbf{S}_2 &= (\sqrt{b}X_1^1, X_2^1, \sqrt{b}X_1^2, X_2^2, ..., \sqrt{b}X_1^{N+1}, X_2^{N+1}).\end{aligned} \tag{4.88}$$

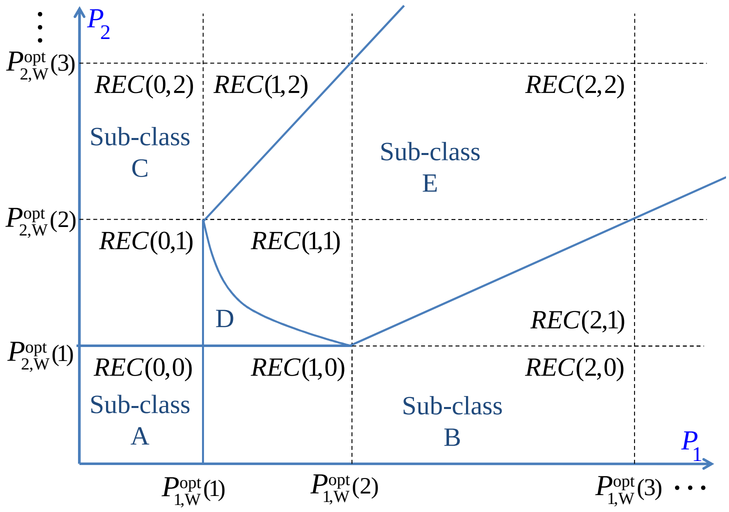

Figure 4.11: The relation between rectangles $REC(m,n)$ and sub-classes A, B, C, D, and E.

If $m < n$, then we let

$$\begin{aligned}
\mathbf{S}_1 &= (\sqrt{a}X_2^1, X_1^1, \sqrt{a}X_2^2, X_1^2, ..., \sqrt{a}X_2^{N+1}, X_1^{N+1}),\\
\mathbf{S}_2 &= (X_2^1, \sqrt{b}X_1^1, \sqrt{b}X_1^2, X_2^2\ , ..., \sqrt{b}X_1^{N+1}, X_2^{N+1}).
\end{aligned} \tag{4.89}$$

Observe that only the first two elements of $\mathbf{S}_1$ and $\mathbf{S}_2$ have changed in (4.87-4.89).

We have determined the number of splits and the decoding orders, based on the value of (P_1, P_2) in (4.79) and (4.87-4.89), respectively. We also need to determine the value of the optimal power allocations, i.e., P_1^j and P_2^j. Using (4.87-4.89) and (4.17), we characterize a feasible solution to the optimization problem (4.6).

Similar to the strong interference class, we use (4.17) to characterize a feasible solution to the optimization problem (4.6). According to (4.2) and for the decoding orders given in (4.87-4.89), we have

$$\begin{aligned}
c_1^{N+1} &= C\Big(\frac{P_1^{N+1}}{1+\sum_{k=1}^{N} P_1^j + a(\sum_{k=1}^{N+1} P_2^j)}\Big),\\
d_1^{N+1} &= C\Big(\frac{bP_1^{N+1}}{1+\sum_{k=1}^{N}\ P_2^j + b(\sum_{k=1}^{N} P_1^j)}\Big).
\end{aligned} \tag{4.90}$$

By letting $c_1^{N+1} = d_1^{N+1}$, we calculate P_2^{N+1}, as follows:

$$\begin{aligned} c_1^{N+1} =& d_1^{N+1} \\ \Rightarrow b(1+\sum_{k=1}^{N} P_1^j + a(\sum_{k=1}^{N+1} P_2^j)) =& 1+\sum_{k=1}^{N} P_2^j + b(\sum_{k=1}^{N} P_1^j) \\ \Rightarrow P_2^{N+1} =& 1-b+(1-ab)P_2. \end{aligned} \tag{4.91}$$

Similarly, by letting $c_2^{N+1} = d_2^{N+1}$, we have

$$P_1^{N+1} = 1-a+(1-ab)P_1. \tag{4.92}$$

Inserting (4.91) and (4.92) into (4.90), c_1^{N+1} and d_1^{N+1} simplify to

$$c_1^{N+1} = d_1^{N+1} = C\Big(\frac{(1-ab)P_1+1-a}{a(P_2+bP_1+1)}\Big). \tag{4.93}$$

Similarly, one can show that

$$c_2^{N+1} = d_2^{N+1} = C\Big(\frac{(1-ab)P_2+1-b}{b(P_1+aP_2+1)}\Big). \tag{4.94}$$

Following this scheme, let $c_1^j = d_1^j$ and $c_2^j = d_2^j$ for all $j \geq 2$. Consequently, we calculate P_1^j and P_2^j for $2 \leq j \leq N$, as follows:

$$P_1^j = 1-a+(1-ab)(P_1 - \sum_{k=j+1}^{N+1} P_1^j), \tag{4.95}$$

$$P_2^j = 1-b+(1-ab)(P_2 - \sum_{k=j+1}^{N+1} P_2^j). \tag{4.96}$$

With this choice of values for P_1^j and P_2^j, the values of $c_1^j = d_1^j$ and $c_2^j = d_2^j$ simplify to

$$\begin{aligned} c_1^j = d_1^j = C\Big(\frac{(1-ab)P_1+1-a}{a(P_2+bP_1+1)}\Big), \\ c_2^j = d_2^j = C\Big(\frac{(1-ab)P_2+1-b}{b(P_1+aP_2+1)}\Big). \end{aligned} \tag{4.97}$$

Note that the values of c_1^j and c_2^j are independent of j. This is a direct consequence of (4.95) and (4.96). Moreover, according to (4.1), we have $\sum_{k=1}^{N+1} P_1^k = P_1$ and $\sum_{k=1}^{N+1} P_2^k = P_2$. Therefore, by choosing the values of P_1^k and P_2^k according to (4.95) and (4.96), respectively, P_1^1 and P_2^1 are determined by

$$P_1^1 = P_1 - \sum_{k=2}^{N+1} P_1^k, P_2^1 = P_2 - \sum_{k=2}^{N+1} P_2^k. \tag{4.98}$$

Using (4.95-4.96), one can show that P_1^1 and P_2^1 are given by

$$\begin{aligned} P_1^1 &= \Delta P_1 (ab)^N, \\ P_2^1 &= \Delta P_2 (ab)^N. \end{aligned} \tag{4.99}$$

For a given (P_1, P_2), (4.79) determines the number of splits, (4.87-4.89) determine the decoding order, and (4.95-4.96) determine the power allocation. Consequently, one can calculate the achievable sum-rate of this scheme, as stated in the next theorem.

Theorem 4.8. *For the two-user GIC with weak interference, if (P_1, P_2) lies on $REC(m, n)$, then rate splitting and power allocation according to (4.95,4.96) and successive decoding according to (4.87-4.89) achieves the following sum-rate:*

$$R_{\text{sum}}^{\text{RS-SD}} = NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + R_{\text{sum}}^1. \tag{4.100}$$

where $N = \min\{m, n\}$ and $R_{\text{sum}}^1 \doteq R_1^1 + R_2^1$ is given by

$$R_{\text{sum}}^1 = \begin{cases} C\big(\frac{\Delta P_1 (ab)^N}{1+a\Delta P_2 (ab)^N}\big) + C\big(\frac{\Delta P_2 (ab)^N}{1+b\Delta P_1 (ab)^N}\big) & \text{if } m = n, \\ C\big(\Delta P_1 (ab)^N + a\Delta P_2 (ab)^N\big) & \text{if } m > n, \\ C\big(\Delta P_2 (ab)^N + b\Delta P_1 (ab)^N\big) & \text{if } m < n. \end{cases} \tag{4.101}$$

Proof. For $2 \le j \le N+1$, we have

$$\begin{aligned} R_1^j + R_2^j &= \min\{c_1^j, d_1^j\} + \min\{c_2^j, d_2^j\} \\ &\overset{(a)}{=} c_1^j + c_2^j \\ &\overset{(b)}{=} C\Big(\frac{(1-ab)P_1 + 1 - a}{a(P_2 + bP_1 + 1)}\Big) + C\Big(\frac{(1-ab)P_2 + 1 - b}{b(P_1 + aP_2 + 1)}\Big) \\ &= \frac{1}{2}\log\Big(\big(1 + \frac{(1-ab)P_1 + 1 - a}{a(P_2 + bP_1 + 1)}\big)\big(1 + \frac{(1-ab)P_2 + 1 - b}{b(P_1 + aP_2 + 1)}\big)\Big) \\ &= \frac{1}{2}\log\Big(\frac{1}{ab}\Big) \\ &= C\big(\frac{1-a}{a}\big) + C\big(\frac{1-b}{b}\big). \end{aligned} \tag{4.102}$$

where (a) is valid by (4.17), and (b) is valid by (4.97). Therefore,

$$\begin{aligned} R_{\text{sum}}^{\text{RS-SD}} &= \sum_{j=1}^{N+1} \big(R_1^j + R_2^j\big) = \sum_{j=2}^{N+1} \big(R_1^j + R_2^j\big) + R_1^1 + R_2^1 \\ &\overset{(a)}{=} NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + R_{\text{sum}}^1, \end{aligned} \tag{4.103}$$

where (a) is valid by (4.102).

Moreover, for $m = n$,

$$\begin{aligned} R_{\text{sum}}^1 &= R_1^1 + R_2^1 = c_1^1 + c_2^1 \\ &\stackrel{(a)}{=} C\Big(\frac{\Delta P_1 (ab)^N}{1 + a\Delta P_2 (ab)^N}\Big) + C\Big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\Big). \end{aligned} \tag{4.104}$$

where (a) is valid because, according to (4.87), c_1^1 and c_2^1 are given by

$$\begin{aligned} c_1^1 &= C\Big(\frac{P_1^1}{1 + aP_2^1}\Big) = C\Big(\frac{\Delta P_1 (ab)^N}{1 + a\Delta P_2 (ab)^N}\Big), \\ c_2^1 &= C\Big(\frac{P_2^1}{1 + bP_1^1}\Big) = C\Big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\Big). \end{aligned} \tag{4.105}$$

This completes the proof for $m = n$.

For $m > n$,

$$\begin{aligned} R_{\text{sum}}^1 &= R_1^1 + R_2^1 = c_1^1 + \min\{c_2^1, d_2^1\} \\ &\stackrel{(a)}{=} c_1^1 + d_2^1 \\ &\stackrel{(b)}{=} C\big(\Delta P_1 (ab)^N\big) + C\Big(\frac{a\Delta P_2 (ab)^N}{1 + \Delta P_1 (ab)^N}\Big) \\ &= C\big(\Delta P_1 (ab)^N + a\Delta P_2 (ab)^N\big) \end{aligned} \tag{4.106}$$

where (a) and (b) are valid because, according to (4.88), c_1^1, c_2^1, and d_2^1 are given by

$$\begin{aligned} c_1^1 &= C\big(P_1^1\big) = C\big(\Delta P_1 (ab)^N\big), \\ c_2^1 &= C\Big(\frac{P_2^1}{1 + bP_1^1}\Big) = C\Big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\Big), \\ d_2^1 &= C\Big(\frac{aP_2^1}{1 + P_1}\Big) = C\Big(\frac{a\Delta P_2 (ab)^N}{1 + \Delta P_1 (ab)^N}\Big). \end{aligned} \tag{4.107}$$

This completes the proof for $m > n$. The proof for $m < n$ follows similarly. $\square$

The previous theorem characterizes $R_{\text{sum}}^{\text{RS-SD}}$. In order to compare the performance of $R_{\text{sum}}^{\text{RS-SD}}$ with $R_{\text{sum-Hk}}^{\max}$, we need to simplify the expressions given in Theorem 4.8, as stated in the following theorem.

Theorem 4.9. *For the two-user GIC with weak interference, if (P_1, P_2) lies on $REC(m, n)$, then rate splitting and power allocation according to (4.95,4.96) and successive decoding*

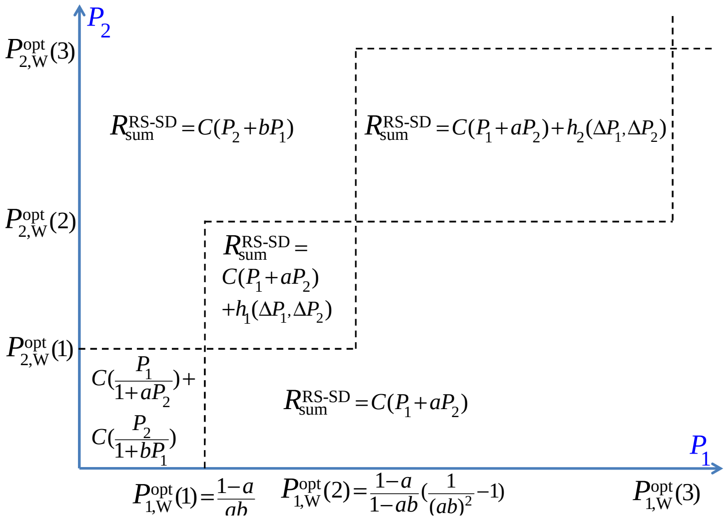

Figure 4.12: The achievable sum-rate $R_{\text{sum}}^{\text{RS-SD}}$.

according to (4.87-4.89) achieve the following sum-rate:

$$R_{\text{sum}}^{\text{RS-SD}} = \begin{cases} C\left(\frac{P_1}{1+aP_2}\right) + C\left(\frac{P_2}{1+bP_1}\right) & \textit{if } m = n = 0, \\ C(P_1 + aP_2) + h_1^N(\Delta P_1, \Delta P_2) & \textit{if } m = n \geq 1, \\ C(P_1 + aP_2) & \textit{if } m > n, \\ C(P_2 + bP_1) & \textit{if } m < n, \end{cases} \tag{4.108}$$

where $h_1^N(\Delta P_1, \Delta P_2) \doteq C\left(\frac{\Delta P_2 (ab)^N}{1+b\Delta P_1 (ab)^N}\right) - C\left(a\Delta P_2 (ab)^N\right)$ *is a non-negative function.*

Proof. To prove this theorem, we simplify the expression of $R_{\text{sum}}^{\text{RS-SD}}$ given in (4.100). Note that, according to (4.77) and (4.78), we have

$$\begin{aligned} (ab)^{-N} &= 1 + P_{1,\text{W}}^{\text{opt}}(N) + aP_{2,\text{W}}^{\text{opt}}(N) \\ &= 1 + P_{2,\text{W}}^{\text{opt}}(N) + bP_{1,\text{W}}^{\text{opt}}(N). \end{aligned} \tag{4.109}$$

Moreover,

$$\begin{aligned} C(P_1 + aP_2) =& C(P_{1,\text{W}}^{\text{opt}}(N) + aP_{2,\text{W}}^{\text{opt}}(N) + \Delta P_1 + a\Delta P_2) \\ =& C\left(P_{1,\text{W}}^{\text{opt}}(N) + aP_{2,\text{W}}^{\text{opt}}(N)\right) + C\left(\frac{\Delta P_1 + a\Delta P_2}{1 + P_{1,\text{W}}^{\text{opt}}(N) + aP_{2,\text{W}}^{\text{opt}}(N)}\right) \\ \overset{(a)}{=}& NC\left(\frac{1-a}{a}\right) + NC\left(\frac{1-b}{b}\right) + C\left(\frac{\Delta P_1 + a\Delta P_2}{(ab)^{-N}}\right), \end{aligned} \tag{4.110}$$

where (a) is valid by (4.109). Consequently, for $m > n$, we have

$$R_{\text{sum}}^{\text{RS-SD}} \stackrel{(a)}{=} NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + \big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^{-N}}\big)$$
$$\stackrel{(b)}{=} C(P_1 + aP_2), \tag{4.111}$$

where (a) is valid by (4.101), and (b) is valid by (4.110). Similarly, for $m < n$, one can see that

$$R_{\text{sum}}^{\text{RS-SD}} = C(P_2 + bP_2). \tag{4.112}$$

For $m = n = 0$, according to (4.101)

$$\begin{aligned} R_{\text{sum}}^{\text{RS-SD}} &= C\big(\frac{\Delta P_1}{1 + a\Delta P_2}\big) + C\big(\frac{\Delta P_2}{1 + b\Delta P_1}\big) \\ &= C\big(\frac{P_1}{1 + aP_2}\big) + C\big(\frac{P_2}{1 + bP_1}\big). \end{aligned} \tag{4.113}$$

Finally, for $m = n \geq 1$,

$$\begin{aligned} R_{\text{sum}}^{\text{RS-SD}} \stackrel{(a)}{=} & NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + C\big(\frac{\Delta P_1 (ab)^N}{1 + a\Delta P_2 (ab)^N}\big) + C\big(a\Delta P_2 (ab)^N\big) \\ & + C\big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\big) - C\big(a\Delta P_2 (ab)^N\big) \\ = & NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + C\Big(\frac{\Delta P_1 + a\Delta P_2}{(ab)^{-N}}\Big) \\ & + C\big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\big) - C\big(a\Delta P_2 (ab)^N\big) \\ \stackrel{(b)}{=} & C(P_1 + aP_2) + h_1^N(\Delta P_1, \Delta P_2), \end{aligned} \tag{4.114}$$

where (a) is valid by (4.101) and (b) is valid by (4.110). Moreover,

$$\begin{aligned} h_1^N(\Delta P_1, \Delta P_2) \doteq & C\big(\frac{\Delta P_2 (ab)^N}{1 + b\Delta P_1 (ab)^N}\big) - C\big(a\Delta P_2 (ab)^N\big) \\ \stackrel{(a)}{\geq} & 0. \end{aligned} \tag{4.115}$$

where (a) is valid by (4.81). This completes the proof. □

The previous theorem characterizes simplified expressions for the achievable sum-rate. Note that for $m = n \geq 1$, $R_{\text{sum}}^{\text{RS-SD}}$ is given in terms of $C(P_1 + aP_2)$ plus a nonnegative function. Similarly, one can show that, $R_{\text{sum}}^{\text{RS-SD}}$ can be given in terms of $C(P_2 + bP_1)$ plus a nonnegative function too. In fact, for $m = n \geq 1$, we have

$$\begin{aligned}
R_{\text{sum}}^{\text{RS-SD}} \overset{(a)}{=} & NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + C\big(\frac{\Delta P_2(ab)^N}{1+b\Delta P_1(ab)^N}\big) + C\big(b\Delta P_1(ab)^N\big) \\
& + C\big(\frac{\Delta P_1(ab)^N}{1+a\Delta P_2(ab)^N}\big) - C\big(b\Delta P_1(ab)^N\big) \\
= & NC\big(\frac{1-a}{a}\big) + NC\big(\frac{1-b}{b}\big) + C\Big(\frac{\Delta P_2 + b\Delta P_1}{(ab)^{-N}}\Big) \\
& + C\big(\frac{\Delta P_1(ab)^N}{1+a\Delta P_2(ab)^N}\big) - C\big(b\Delta P_1(ab)^N\big) \\
\overset{(b)}{=} & C(P_2 + bP_1) + h_2^N(\Delta P_1, \Delta P_2),
\end{aligned} \tag{4.116}$$

where (a) is valid by (4.101) and (b) is valid by (4.112). Moreover,

$$\begin{aligned}
h_2^N(\Delta P_1, \Delta P_2) \doteq & C\big(\frac{\Delta P_1(ab)^N}{1+a\Delta P_2(ab)^N}\big) - C\big(b\Delta P_1(ab)^N\big) \\
\overset{(a)}{\geq} & 0.
\end{aligned} \tag{4.117}$$

where (a) is valid by (4.82). Therefore, for $m = n \geq 1$, we have

$$\begin{aligned}
R_{\text{sum}}^{\text{RS-SD}} = & C(P_1 + aP_2) + h_1^N(\Delta P_1, \Delta P_2) \\
= & C(P_2 + bP_1) + h_2^N(\Delta P_1, \Delta P_2) \\
= & \max\{C(P_1 + aP_2), C(P_2 + bP_1)\} \\
& + \min\{h_1^N(\Delta P_1, \Delta P_2), h_2^N(\Delta P_1, \Delta P_2)\}.
\end{aligned} \tag{4.118}$$

Relying on this observation, we compare this simplified sum-rate with $R_{\text{sum}}^{\text{NRS}}$ and show that $R_{\text{sum}}^{\text{RS-SD}} \geq R_{\text{sum}}^{\text{NRS}}$, as explained in the following remark.

Remark 4.2. $R_{\text{sum}}^{\text{RS-SD}} \geq R_{\text{sum}}^{\text{NRS}}$*: To compare* $R_{\text{sum}}^{\text{RS-SD}}$ *with* $R_{\text{sum}}^{\text{NRS}}$*, we can compare Figure 4.12 with Figure 4.8. For* $m \neq n$*, we have* $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum}}^{\text{NRS}}$*. Moreover, for* $m = n = 0$*, we have* $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum}}^{\text{NRS}}$*. However, for* $m = n \geq 1$

$$\begin{aligned}
R_{\text{sum}}^{\text{RS-SD}} - R_{\text{sum}}^{\text{NRS}} \overset{(a)}{=} & \max\{C(P_1 + aP_2), C(P_2 + bP_1)\} \\
& + \min\{h_1^N(\Delta P_1, \Delta P_2), h_2^N(\Delta P_1, \Delta P_2)\} \\
& - \max\{C(P_1 + aP_2), C(P_2 + bP_1)\} \\
= & \min\{h_1^N(\Delta P_1, \Delta P_2), h_2^N(\Delta P_1, \Delta P_2)\} \\
\geq & 0,
\end{aligned} \tag{4.119}$$

where (a) *is valid by* (4.118). *Since both* $h_1^N(\Delta P_1, \Delta P_2)$ *and* $h_2^N(\Delta P_1, \Delta P_2)$ *are nonnegative functions,* (4.119) *shows that for* $m = n \geq 1$*, we have* $R_{\text{sum}}^{\text{RS-SD}} \geq R_{\text{sum}}^{\text{NRS}}$.

Remark 4.2 shows that our proposed scheme can achieve a higher sum-rate compared to $R_{\text{sum}}^{\text{NRS}}$. Next, we compare $R_{\text{sum}}^{\text{RS-SD}}$ with $R_{\text{sum-HK}}^{\max}$ to show that for a wide range of (a, b, P_1, P_2), $R_{\text{sum}}^{\text{RS-SD}}$ achieves the maximum HK sum-rate. In fact, Tables 4.3 shows that the HK scheme partitions the weak interference class into five sub-classes. Next theorem proves that for the first four sub-classes, we have $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$.

Theorem 4.10. *For the two-user GIC with weak interference, if* (a, b, P_1, P_2) *belongs to the union of sub-classes* A*,* B*,* C*, and* D*, then* $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$.

Proof. Figure 4.12 demonstrates $R_{\text{sum}}^{\text{RS-SD}}$ inside all rectangles $REC(m, n)$. By comparing Figure 4.12 with Figure 4.9, we see that for sub-classes A, B, and C, we have $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$. For sub-class D, according to (4.114), $R_{\text{sum}}^{\text{RS-SD}} = C(P_1 + aP_2) + h_1(\Delta P_1, \Delta P_2)$. On the other hand, according to Theorem 4.6, $R_{\text{sum-HK}}^{\max} = C(P_1 + aP_2) + g_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$. One can verify that $h_1^1(\Delta P_1, \Delta P_2) = g_1(\tilde{\lambda}_1, \tilde{\lambda}_2)$ and conclude that $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$. In fact, according to (4.68), we have

$$\tilde{\lambda}_1 P_1 = abP_1 - (1 - a), \tag{4.120}$$

$$\tilde{\lambda}_2 P_2 = abP_2 - (1 - b). \tag{4.121}$$

On the other hand, since sub-class D is inside $REC(1, 1,)$

$$\begin{aligned}
\Delta P_1 ab =& (P_1 - P_{1,\text{W}}^{\text{opt}}(1))ab \\
=& (P_1 - \frac{1-a}{ab})ab \\
=& abP_1 - (1 - a) \\
\overset{(a)}{=}& \tilde{\lambda}_1 P_1,
\end{aligned} \tag{4.122}$$

where (a) is valid by (4.120). Similarly,

$$\Delta P_2 ab = \tilde{\lambda}_2 P_2. \tag{4.123}$$

Inserting (4.122) and (4.123) into (4.67), we have

$$\begin{aligned}
g_1(\tilde{\lambda}_1, \tilde{\lambda}_2) =& C(\frac{(1-a)\tilde{\lambda}_2 P_2 + b\tilde{\lambda}_1 P_1}{1 + a\tilde{\lambda}_2 P_2}) - C(b\tilde{\lambda}_1 P_1) \\
=& C(\frac{(1-a)\Delta P_2 ab + b\Delta P_1 ab}{1 + a\Delta P_2 ab}) - C(b\Delta P_1 ab).
\end{aligned} \tag{4.124}$$

On the other hand,

$$h_1^1(\Delta P_1, \Delta P_2) \doteq C\big(\frac{\Delta P_2(ab)}{1+b\Delta P_1(ab)}\big) - C\big(a\Delta P_2(ab)\big). \tag{4.125}$$

Observe that (4.125) and (4.124) are equal because

$$\begin{aligned} h_1^1(\Delta P_1, \Delta P_2) =& g_1(\tilde{\lambda}_1, \tilde{\lambda}_2) \\ &\Leftrightarrow C\big(\frac{\Delta P_2(ab)}{1+b\Delta P_1(ab)}\big) - C\big(a\Delta P_2(ab)\big) \\ &= C\big(\frac{(1-a)\Delta P_2 ab + b\Delta P_1 ab}{1+a\Delta P_2 ab}\big) - C(b\Delta P_1 ab) \\ &\Leftrightarrow C\big(\frac{\Delta P_2(ab)}{1+b\Delta P_1(ab)}\big) + C(b\Delta P_1 ab) \\ &= C\big(\frac{(1-a)\Delta P_2 ab + b\Delta P_1 ab}{1+a\Delta P_2 ab}\big) + C\big(a\Delta P_2(ab)\big) \\ &\Leftrightarrow C(\Delta P_2 ab + b\Delta P_1 ab) \\ &= C(\Delta P_2 ab + b\Delta P_1 ab). \end{aligned} \tag{4.126}$$

This shows that, for sub-class D, $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$. The proof is complete.

□

4.4.3 Maximum Sum-Rate Loss

The previous theorem shows that inside sub-class E, $R_{\text{sum}}^{\text{RS-SD}} \leq R_{\text{sum-HK}}^{\max}$. However, we can show that even in this sub-class, $R_{\text{sum}}^{\text{RS-SD}}$ is close to $R_{\text{sum-HK}}^{\max}$. First, we show that there exist hyperplanes inside sub-class E, for which we have $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$. Second, we show that inside sub-class E, $R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}}$ is bounded.

Theorem 4.11. *For the two-user GIC with weak interference, if (a, b, P_1, P_2) belongs to $REC(N, N)$ and also belongs to the hyperplane L_N characterized by*

$$\begin{aligned} L_N \doteq& \{(a, b, P_1, P_2) \in \mathbb{R}_+^4 : \\ &\hat{\lambda}_2 = (ab)^N - \frac{\frac{1-b}{1-ab}\big(1-(ab)^N\big)}{P_2}\}, \end{aligned} \tag{4.127}$$

for some positive integer N, then $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$.

Proof. First, remember that sub-class E represents all $(a, b, P_1, P_2) \in \mathbb{R}_+^4$ that satisfy the description given in Table 4.3. Observe that $\hat{\lambda}_2$, given in (4.69), is a function of

(a, b, P_1, P_2). Therefore, for a fixed N, the equation $\hat{\lambda}_2 = (ab)^N - \frac{\frac{1-b}{1-ab}\left(1-(ab)^N\right)}{P_2}$ represents a hyperplane in $\mathbb{R}_+^4$.

Second, note that if N=1, (4.127) simplifies to

$$\begin{aligned}\hat{\lambda}_2 =& ab - \frac{\frac{1-b}{1-ab}\left(1-ab\right)}{P_2}\\ =& ab - \frac{1-b}{P_2}\\ =& \tilde{\lambda}_2 \end{aligned} \tag{4.128}$$

Moreover,

$$\begin{aligned}\hat{\lambda}_1 =& \alpha\hat{\lambda}_2 + c\\ =& \alpha\tilde{\lambda}_2 + c\\ =& \tilde{\lambda}_1 \end{aligned} \tag{4.129}$$

Consequently, for $N = 1$, Theorem 4.11 reduces to equality (4.126).

For $N > 1$, we have

$$\begin{aligned} R_{\text{sum}}^{\text{RS-SD}} &= C(P_1 + aP_2) + h_1^N(\Delta P_1, \Delta P_2).\\ R_{\text{sum-HK}}^{\max} &= C(P_1 + aP_2) + g_1(\hat{\lambda}_1, \hat{\lambda}_2). \end{aligned}$$

We claim that if (4.127) is satisfied, then $h_1^N(\Delta P_1, \Delta P_2) = g_1(\hat{\lambda}_1, \hat{\lambda}_2)$, and consequently, we have $R_{\text{sum}}^{\text{RS-SD}} = R_{\text{sum-HK}}^{\max}$. To prove this claim, on one hand, if $\hat{\lambda}_2 = (ab)^N - \frac{\frac{1-b}{1-ab}\left(1-(ab)^N\right)}{P_2}$, then $\hat{\lambda}_1$ is given by

$$\begin{aligned}\hat{\lambda}_1 =& \alpha\hat{\lambda}_2 + c\\ =& (1-c)\left((ab)^N - \frac{\frac{1-b}{1-ab}\left(1-(ab)^N\right)}{P_2}\right) + c\\ \overset{(a)}{=}& (ab)^N - \frac{\frac{1-a}{1-ab}\left(1-(ab)^N\right)}{P_1}, \end{aligned} \tag{4.130}$$

where (a) is valid by (4.72). On the other hand, for $m = n = N$, we have

$$\begin{aligned}\Delta P_1 (ab)^N =& (P_1 - P_{1,\text{W}}^{\text{opt}}(N))(ab)^N\\ =& \left(P_1 - \frac{1-a}{1-ab}\left(\frac{1}{(ab)^N} - 1\right)\right)(ab)^N\\ =& (ab)^N P_1 - \frac{1-a}{1-ab}\left(1-(ab)^N\right)\\ \overset{(a)}{=}& \hat{\lambda}_1 P_1, \end{aligned} \tag{4.131}$$

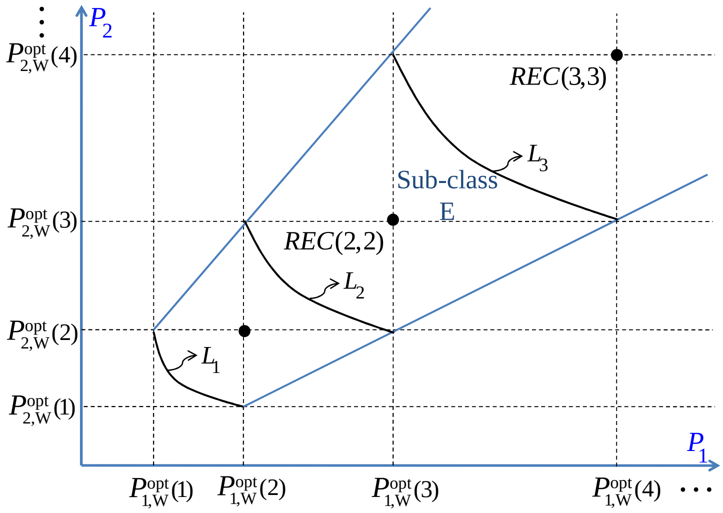

Figure 4.13: The sub-class E is partitioned by hyperplanes L_i. On the boundary of each part, $R^{\rm max}_{\text{sum-HK}} = R^{\text{RS-SD}}_{\rm sum}$. Inside each part, the maximum of $R^{\rm max}_{\text{sum-HK}} - R^{\text{RS-SD}}_{\rm sum}$ occurs when $(P_1, P_2) = \big(P^{\rm opt}_{1,\rm W}(N), P^{\rm opt}_{2,\rm W}(N)\big)$ for $N > 1$.

where (a) is valid by (4.130). Similarly, one can show that

$$\Delta P_2(ab)^N = \hat{\lambda}_2 P_2. \tag{4.132}$$

Therefore, $g_1(\hat{\lambda}_1, \hat{\lambda}_2)$ is given by

$$\begin{aligned} g_1(\hat{\lambda}_1, \hat{\lambda}_2) =& C(\frac{(1-a)\hat{\lambda}_2 P_2 + b\hat{\lambda}_1 P_1}{1 + a\hat{\lambda}_2 P_2}) - C(b\hat{\lambda}_1 P_1) \\ =& C(\frac{(1-a)\Delta P_2(ab)^N + b\Delta P_1(ab)^N}{1 + a\Delta P_2(ab)^N}) - \\ & C(b\Delta P_1(ab)^N). \end{aligned} \tag{4.133}$$

On the other hand,

$$h_1^N(\Delta P_1, \Delta P_2) \doteq C\big(\frac{\Delta P_2(ab)^N}{1 + b\Delta P_1(ab)^N}\big) - C\big(a\Delta P_2(ab)^N\big). \tag{4.134}$$

Similar to (4.126), one can see that (4.133) and (4.134) are equal. This completes the proof. □

The hyperplane L_N is demonstrated in Figure 4.13, for $N = 1$, $N = 2$, and $N = 3$. Theorem 4.11 shows that the hyperplanes L_i, partition the sub-class E into many parts. Over the boundary of each part, we have $R^{\text{RS-SD}}_{\rm sum} = R^{\rm max}_{\text{sum-HK}}$. In the next theorem, we show that inside each part, the maximum difference between $R^{\text{RS-SD}}_{\rm sum}$ and $R^{\rm max}_{\text{sum-HK}}$ is limited to

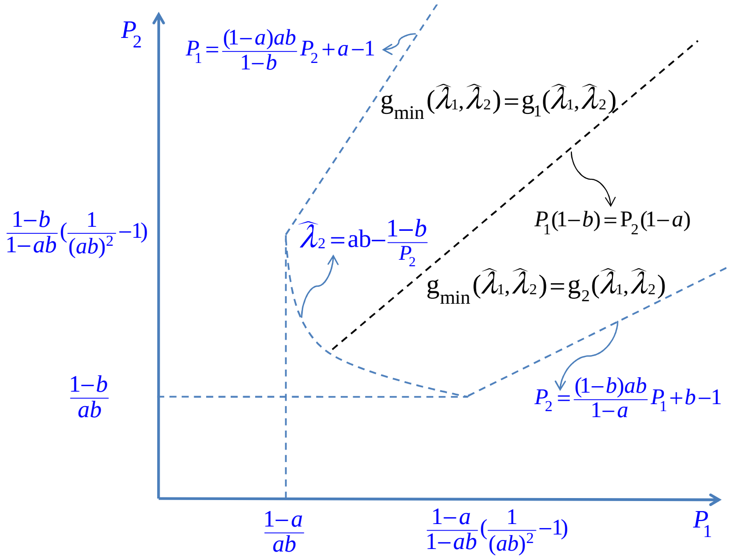

Figure 4.14: The function $g_{\min}(P_1, P_2)$ over the sub-class E.

$\log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$. Interestingly, there exists exactly one (P_1, P_2) inside each part that leads to this maximum difference. Note that the maximum difference is the same for all parts.

Theorem 4.12. *For the two-user GIC with weak interference, if joint decoding is replaced by SD, the maximum sum-rate loss is given by* $\Delta R_{\text{sum}}^{\max} = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$.

Proof. Our goal is to show that

$$\max_{P_1,P_2}\left(R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}}\right) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right). \tag{4.135}$$

Note that $R_{\text{sum}}^{\text{RS-SD}} < R_{\text{sum-HK}}^{\max}$ only in the sub-class E. Therefore, we can restrict (P_1, P_2) to the sub-class E. Let $\mathcal{E}$ represents the sub-class E, for fixed values of a and b. We have

$$\begin{aligned}\mathcal{E} = \{(P_1, P_2) :& P_1 > \frac{(1-a)ab}{1-b}P_2 + a - 1,\\ & P_2 > \frac{(1-b)ab}{1-b}P_1 + b - 1,\\ & \hat{\lambda}_2 \le ab - \frac{1-b}{P_2}\}.\end{aligned} \tag{4.136}$$

Then the optimization problem (4.135) is equivalent to

$$\begin{aligned}&\max_{P_1,P_2}\left(R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}}\right)\\ &\text{subject to } (P_1, P_2) \in \mathcal{E}.\end{aligned} \tag{4.137}$$

We can partition $\mathcal{E}$ into parts, namely $E_1, E_2, ...$, as shown in Figure 4.13. Our idea to solve this optimization problem is as follows. Instead of looking for the optimal solution over $\mathcal{E}$, we look for the optimal solution over each E_i. Let ΔR_{E_i} be defined as

$$\begin{aligned}\Delta R_{E_i} &= \max_{P_1,P_2}\big(R^{\max}_{\text{sum-HK}} - R^{\text{RS-SD}}_{\text{sum}}\big)\\ &\text{subject to } (P_1, P_2) \in E_i.\end{aligned} \tag{4.138}$$

Since E_is form a partitioning of $\mathcal{E}$, we conclude that (4.137) is equivalent to

$$\max_i\Big(\Delta R_{E_i}\Big). \tag{4.139}$$

In the following, we show that we have

$$\Delta R_{E_i} = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right), \tag{4.140}$$

and therefore, (4.137) is equivalent to

$$\max_i\Big(\Delta R_{E_i}\Big) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right). \tag{4.141}$$

To solve (4.138) and characterize E_i, we first note that over the boundary of each part, we have $R^{\text{RS-SD}}_{\text{sum}} = R^{\max}_{\text{sum-HK}}$.

Moreover, according to Remark 4.2, $R^{\text{RS-SD}}_{\text{sum}} \geq R^{\text{NRS}}_{\text{sum}}$. Therefore, we have

$$\begin{aligned}R^{\max}_{\text{sum-HK}} - R^{\text{RS-SD}}_{\text{sum}} &\leq R^{\max}_{\text{sum-HK}} - R^{\text{NRS}}_{\text{sum}}\\ &= \Delta R_{\text{sum}}\\ &\overset{(a)}{=} \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\},\end{aligned} \tag{4.142}$$

where (a) is valid by (4.76). Define

$$g_{\min}(P_1, P_2) \doteq \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\}. \tag{4.143}$$

According to (4.69) and (4.70), $\hat{\lambda}_2$ and $\hat{\lambda}_1$ are functions of (P_1, P_2), and therefor, $g_{\min}()$ is a function of (P_1, P_2). Figure 4.14 demonstrates the function $g_{\min}()$ over the sub-class E. Observe that, we have

$$\max_{P_1,P_2}\big(R^{\max}_{\text{sum-HK}} - R^{\text{RS-SD}}_{\text{sum}}\big) \leq \max_{P_1,P_2} g_{\min}(\hat{\lambda}_1, \hat{\lambda}_2), \tag{4.144}$$

subject to (P_1, P_2) belongs to the sub-class E. Instead of solving (4.135) directly, we solve

$$\max_{P_1,P_2} g_{\min}(P_1, P_2) \tag{4.145}$$

In the following, we first prove that

$$\max_{P_1,P_2} g_{\min}(P_1, P_2) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right), \tag{4.146}$$

then, we show that only if $(P_1, P_2) = \left(P_{1,\mathrm{W}}^{\mathrm{opt}}(N), P_{2,\mathrm{W}}^{\mathrm{opt}}(N)\right)$, we have

$$R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}} = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right). \tag{4.147}$$

To show that (4.146) is valid, we note an optimization technique. According to interior extremum theorem, the global maximum of a differentiable function f over a feasible region $\mathcal{A}$ is achieved at one of the following points: an stationary point or a boundary point [45, 46]. In particular, $g_1(\hat{\lambda}_1, \hat{\lambda}_2)$ and $g_2(\hat{\lambda}_1, \hat{\lambda}_2)$ are both differentiable functions of (P_1, P_2). However, $g_{\min}(P_1, P_2)$ can be non-differentiable. In fact, over the hyperplane $P_1(1-b) = P_2(1-a)$, we have $g_1(\hat{\lambda}_1, \hat{\lambda}_2) = g_2(\hat{\lambda}_1, \hat{\lambda}_2)$. Consequently, if $g_{\min}(P_1, P_2)$ is not differentiable at (P_1, P_2), then (P_1, P_2) belongs to the hyperplane $P_1(1-b) = P_2(1-a)$. Therefore, if $(P_1^\star, P_2^\star)$ maximizes the optimization problem (4.146), then $(P_1^\star, P_2^\star)$ is either an stationary point, or a point on the boundary, or a non-differentiable point on the hyperplane $P_1(1-b) = P_2(1-a)$.

We solve the optimization problem (4.146) in three steps. First, we note that $g_{\min}(P_1, P_2)$ has no stationary points inside the sub-class E. Then we show that over the hyperplane $P_1(1-b) = P_2(1-a)$, which include all non-differentiable points, we have $g_{\min}(P_1, P_2) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$. Finally, we show that over the boundary of the sub-class E, we have $g_{\min}(P_1, P_2) \leq \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$.

To show that $g_{\min}(P_1, P_2)$ has no stationary point, we should investigate $\nabla_{(P_1,P_2)} g_1(\hat{\lambda}_1, \hat{\lambda}_2)$. Direct calculation shows that $\nabla_{(P_1,P_2)} g_1(\hat{\lambda}_1, \hat{\lambda}_2) = (0, 0)$ has no solution in the sub-class E. Similarly, $\nabla_{(P_1,P_2)} g_2(\hat{\lambda}_1, \hat{\lambda}_2) = (0, 0)$ has no solution in the sub-class E. Consequently, $g_{\min}(P_1, P_2)$ has no stationary point in the sub-class E.

Next, we investigate the non-differentiable points of $g_{\min}(P_1, P_2)$. If (P_1, P_2) belongs to the hyperplane $P_1(1-b) = P_2(1-a)$, according to (4.72), $c = 0$ and $\alpha = 1$. Consequently, by (4.69),

$$\begin{aligned}
\hat{\lambda}_2 =& \frac{1+bP_1c}{bP_1\alpha+P_2}\Big(-1+\sqrt{1-\frac{(bP_1\alpha+P_2)(abP_1c+a-1)}{(1+bP_1c)(abP_1\alpha)}}\Big)\\
=& \frac{1}{bP_1+P_2}\Big(-1+\sqrt{1-\frac{(bP_1+P_2)(a-1)}{abP_1}}\Big)\\
\overset{(a)}{=}& \frac{1-a}{(1-ab)P_1}\Big(-1+\sqrt{\frac{1}{ab}}\Big)\\
=& \frac{1-a}{(1-ab)P_1}\frac{\sqrt{ab}-ab}{ab}.
\end{aligned} \tag{4.148}$$

where (a) is valid because $bP_1+P_2 = P_1\frac{1-ab}{1-a}$. Moreover, by (4.70)

$$\begin{aligned}
\hat{\lambda}_1 =& \alpha\hat{\lambda}_2 + c\\
=& \hat{\lambda}_2\\
=& \frac{1-a}{(1-ab)P_1}\frac{\sqrt{ab}-ab}{ab}.
\end{aligned} \tag{4.149}$$

Inserting this value of $(\hat{\lambda}_1, \hat{\lambda}_2)$ into (4.67) and (4.73), we see that

$$g_1(\hat{\lambda}_1, \hat{\lambda}_2) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right), \tag{4.150}$$

$$g_2(\hat{\lambda}_1, \hat{\lambda}_2) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right). \tag{4.151}$$

Consequently,

$$\begin{aligned}
R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{NRS}} =& g_{\min}(P_1, P_2)\\
=& \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\}\\
=& \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right).
\end{aligned} \tag{4.152}$$

Therefore, when $P_1(1-b) = P_2(1-a)$, the value of $g_{\min}(P_1, P_2)$ is independent of (P_1, P_2).

Finally, we investigate $g_{\min}(P_1, P_2)$ over the boundary. The boundary of sub-class E is characterized by the following three hyperplanes:

$$P_1 = \frac{(1-a)ab}{1-b}P_2 + a - 1, \tag{4.153}$$

$$P_2 = \frac{(1-b)ab}{1-b}P_1 + b - 1, \tag{4.154}$$

$$\hat{\lambda}_2 = ab - \frac{1-b}{P_2}. \tag{4.155}$$

For fixed values of a and b, these hyperplanes are lines in the P_1P_2-plane, as shown in Figure 4.14.

If (P_1, P_2) belongs to the hyperplane (4.153), then (4.69) shows that $\hat{\lambda}_2 = 0$, and consequently,

$$\begin{aligned} g_1(\hat{\lambda}_1, \hat{\lambda}_2) =& C(\frac{(1-a)\hat{\lambda}_2 P_2 + b\hat{\lambda}_1 P_1}{1 + a\hat{\lambda}_2 P_2}) - C(b\hat{\lambda}_1 P_1) \\ =& C(\frac{0 + b\hat{\lambda}_1 P_1}{1+0}) - C(b\hat{\lambda}_1 P_1) \\ =& 0. \end{aligned} \tag{4.156}$$

Therefore, when (4.153) is satisfied,

$$\begin{aligned} R^{\max}_{\text{sum-HK}} - R^{\text{NRS}}_{\text{sum}} =& g_{\min}(P_1, P_2) \\ =& \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\} \\ =& 0. \end{aligned} \tag{4.157}$$

Similarly, if (P_1, P_2) belongs to the hyperplane (4.154), then $\hat{\lambda}_1 = 0$, $g_2(\hat{\lambda}_1, \hat{\lambda}_2) = 0$. Consequently, when (4.154) is satisfied,

$$\begin{aligned} R^{\max}_{\text{sum-HK}} - R^{\text{NRS}}_{\text{sum}} =& g_{\min}(P_1, P_2) \\ =& \min\{g_1(\hat{\lambda}_1, \hat{\lambda}_2), g_2(\hat{\lambda}_1, \hat{\lambda}_2)\} \\ =& 0. \end{aligned} \tag{4.158}$$

If (P_1, P_2) belongs to the hyperplane (4.155), then $(\hat{\lambda}_1, \hat{\lambda}_2) = (\tilde{\lambda}_1, \tilde{\lambda}_2)$. Note that, for fixed values of a and b, this hyperplane is demonstrated by L_1, shown in Figure 4.13. One can see that as (P_1, P_2) moves over L_1 and goes from $(P^{\text{opt}}_{1,\text{W}}(2), P^{\text{opt}}_{2,\text{W}}(1))$ to $(P^{\text{opt}}_{1,\text{W}}(1), P^{\text{opt}}_{2,\text{W}}(2))$, the value of $g_1()$ continuously increases from 0 to $C(\frac{(1-a)(1-b)}{a})$. Similarly, as (P_1, P_2) moves over L_1 and goes from $(P^{\text{opt}}_{1,\text{W}}(1), P^{\text{opt}}_{2,\text{W}}(2))$ to $(P^{\text{opt}}_{1,\text{W}}(2), P^{\text{opt}}_{2,\text{W}}(1))$, the value of $g_2()$ continuously increases from 0 to $C(\frac{(1-a)(1-b)}{b})$. Consequently, $g_{\min}() = \min\{g_1(), g_2()\}$, achieves its maximum when $g_1() = g_2()$. Direct calculation shows that $g_1() = g_2()$ occurs when

$$P_1(1-b) = P_2(1-a). \tag{4.159}$$

Note that according to (4.152), when (4.159) is satisfied, we have $g_{\min}(P_1, P_2) = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$. Therefore, over L_1 we have

$$g_{\min}(P_1, P_2) \leq \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right) \tag{4.160}$$

Examining (4.157), (4.158), and (4.160), we conclude that over the boundary of sub-class E, we have

$$g_{\min}(P_1, P_2) \leq \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right) \tag{4.161}$$

and equality occurs if (4.159) is satisfied.

Since $g_{\min}(P_1, P_2)$ has no stationary points, it achieves its maximum value over the boundary or at a non-differentiable points. (4.152) and (4.161) show that this maximum value is attained over the hyperplane $P_1(1-b) = P_2(1-a)$, and therefore, (4.146) is valid.

Next, we prove that (4.147) is valid. Note that by (4.142) and (4.146), $R^{\max}_{\text{sum-HK}} - R^{\text{RS-SD}}_{\text{sum}} = \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right)$ only if we have

$$R^{\text{RS-SD}}_{\text{sum}} = R^{\text{NRS}}_{\text{sum}}. \tag{4.162}$$

On the other hand, Remark 4.2 shows that $R^{\text{RS-SD}}_{\text{sum}} \geq R^{\text{NRS}}_{\text{sum}}$. According to (4.119) and for sub-class E, $R^{\text{RS-SD}}_{\text{sum}} = R^{\text{NRS}}_{\text{sum}}$ only if

$$\begin{aligned} h_1^N(\Delta P_1, \Delta P_2) &= 0, \\ h_2^N(\Delta P_1, \Delta P_2) &= 0. \end{aligned} \tag{4.163}$$

According to (4.115), $h_1^N(\Delta P_1, \Delta P_2) = 0$ if and only if $\Delta P_2 = 0$. Similarly, according to (4.117), $h_2^N(\Delta P_1, \Delta P_2) = 0$ if and only if $\Delta P_1 = 0$. Note that when $\Delta P_1 = 0$ and $\Delta P_2 = 0$, we have

$$\begin{aligned} P_1 &= P^{\text{opt}}_{1,\text{W}}(N), \\ P_2 &= P^{\text{opt}}_{2,\text{W}}(N). \end{aligned} \tag{4.164}$$

Consequently, for sub-class E, we have $R^{\text{RS-SD}}_{\text{sum}} = R^{\text{NRS}}_{\text{sum}}$ if and only if (4.164) is satisfied. Observe that in the sub-class E, (4.164) can be satisfied for $N \geq 2$, as shown in

Figure 4.13. It is straightforward to see that if $(P_1, P_2) = \left(P_{1,\mathrm{W}}^{\mathrm{opt}}(N), P_{2,\mathrm{W}}^{\mathrm{opt}}(N)\right)$, we have $P_1(1-b) = P_2(1-a)$, and Consequently, by (4.152),

$$\begin{aligned} R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{RS-SD}} =& R_{\text{sum-HK}}^{\max} - R_{\text{sum}}^{\text{NRS}} \\ =& \log\left(\frac{1+\sqrt{ab}}{\sqrt{a}+\sqrt{b}}\right). \end{aligned} \tag{4.165}$$

This completes the proof.

□

4.5 Conclusion

This chapter studied the role of RS and SD in the two-user GIC when interference is strong or weak. It was proved that, for a wide range of (a, b, P_1, P_2), the sum-rate of the HK scheme can be achieved using RS and SD. When SD is strictly inferior to the HK scheme, the maximum sum-rate loss was calculated and was shown to remain constant as P_1 and P_2 approach infinity. This study revealed some interesting structures of sum-rate optimal codes. The extension of the results of this chapter to more than two-user channels can be an interesting future work.

Chapter 5

Delay in Cooperative Communications: Multiplexing Gain of Gaussian Interference Channels with Full-Duplex Transmitters

This chapter investigates the role of cooperation among transmitters of the two-user Gaussian interference channel in enlarging the achievable rate region. In particular, we focus on causal cooperation among transmitters, in which a delay constraint guarantees causality. We review the existing results and highlight the importance of the delay constraint. The main contribution of this chapter is a more accurate analysis of delay in cooperative communications. We introduce a new constraint of causal delay. This new constraint allows the coding scheme to achieve a higher multiplexing gain.

5.1 Introduction

The importance of interference in wireless communications has generated major interest in the interference channel. Different coding schemes have been proposed for the two-user GIC, that maximize the achievable sum-rate under certain conditions. For example, under strong interference, in which each cross-link channel gain is greater than the corresponding direct-link channel gain, the optimal coding scheme is to decode the interference as well as the desired signal [7–9]. In contrast, when cross-link channel gains are much smaller than direct-link channel gains, the sum-capacity is achieved by simply treating the interference as noise [10–12].

Cooperation among nodes in a communication system is a promising approach to increasing the overall system performance. Full-duplex transmitters can not only double the rate of wireless communication systems, but also facilitate collaborative signaling and cooperative communication [73–76]. In the two-user interference channel, full-duplex transmitters can take advantage of the signal they receive from each other to mitigate the interference at their receivers, and this simple cooperation among transmitters can enlarge the achievable rate region. In the context of cognitive radio channels, the role of cooperation in enlarging the capacity region of the GIC has been studied and rate splitting along with Gelfand-Pinsker binning has been used to improve the achievable rate region [29], [30]. Moreover, the capacity region of the two-user GIC with conferencing encoders is established in [31] to within a constant gap. To investigate causal cooperation, the achievable rate region of the two-user interference channel with cribbing encoders is studied in [32–34, 77, 78].

Multiplexing gain has been used as a measure to investigate the role of partial non-causal cooperation (or cognitive message sharing) in wireless networks in the high Signal-To-Noise Ratio (SNR) regime. The multiplexing gain of multiple-input multiple-output (MIMO) Gaussian channels depends on the minimum number of transmits and receive antennas [79,80]. Furthermore, in the K-user GIC, as the cooperation among transmitters increases from no cooperation to perfect cooperation, the multiplexing gain increases from $\frac{1}{2}K$ to K [35]. However, practical cooperation among different nodes requires causal delay to be considered as an essential constraint. The signal transmitted by a node will be received and processed by other nodes with some delay, and the minimum acceptable delay can significantly affect the potential gains of cooperative communication systems. For instance, in the two-user GIC, when only transmitters cooperate non-causally (no delay constraint), i.e., each transmitter non-causally knows the other transmitter's message, the channel behaves similar to the broadcast channel, and the maximum multiplexing gain of two is achievable [36, 37]. Similarly, non-causal cooperation among the receivers achieves the Multiple-Access-Channel multiplexing gain of two [38].

In contrast, when cooperation is causal, Host-Madsen and Nosratina [39] have proved that the maximum achievable multiplexing gain is limited to one. Interestingly, this multiplexing gain can be achieved by half-duplex transmitters without any cooperation. Furthermore, even when all nodes are synchronized and operate in the full-duplex mode, as long as they satisfy the traditional constraint of causal delay, the maximum multiplexing gain remains limited to one [39,81]. Therefore, [39] states that "the multiplexing gains promised by the MIMO systems are critically dependent on a tight coordination among the transmit antennas on the one side, or among the receive antennas on the other side; a level of coordination that seemingly cannot be achieved by the wireless connections available to cooperative communication". Similarly, causal cooperation is known to increase the capacity region of the MIMO GIC, but not its multiplexing gain [82].

This study investigates the two-user GIC with full-duplex transmitters to reach the following conclusion: with a new constraint of causal delay, which is slightly different from the traditional one and captures the role of delay more accurately, the maximum multiplexing gain is in fact two [83]. We introduce this new constraint of causal delay and compare it with that of [39]. The causal delay constraint is traditionally applied to each symbol, whereas in this study, we apply this constraint to a block of M symbols that constitute one OFDM symbol. This new constraint plays an integral role in this study as it allows the coding scheme to achieve a higher multiplexing gain. Moreover, it is known that the channel delay does not affect the capacity of the point-to-point memoryless channel, the memoryless broadcast channel, and the memoryless multiple access channel [84]. However, we show that a small change in the delay of the channels between full-duplex transmitters of the two-user GIC can significantly change the sum-capacity.

To illustrate our results, we first consider a case in which only one of the transmitters operates in the full-duplex mode. Then, we consider the general case in which both transmitters are full-duplex and cooperative. We highlight the potentials (higher multiplexing gain) and limitations (caused by the closed loop between transmitters) that emerge when both transmitters are full-duplex. Furthermore, we study the optimal power allocation that maximizes the achievable sum-rate and examine its effect through some simulations. Interestingly, the simulation results reveal that, when full-duplex transmitters are designed to cancel interference, the achievable sum-rate of the symmetric GIC does not significantly degrade. In fact, we show that when the interference power increases, as the cross-link channel gains increase, the achievable sum-rate of full-duplex transmitters is almost flat and close to that of non-interfering transmitters.

The rest of this chapter is organized as follows: in Section 5.2, our notation and the channel model are introduced. Section 5.3, which contains our main contribution, investigates the achievable multiplexing gain of the two-user GIC with full-duplex transmitters. The case when only one of the transmitters is full-duplex is studied separately. Furthermore, the closed form expression of the optimal power allocation is computed. In Section 5.4, simulation results are presented to highlight the corresponding improvement in the sum-rate. Moreover, the role of optimal power allocation in increasing the achievable sum-rate is depicted in simulation results. Finally, Section 5.5 concludes the chapter.

5.2 Preliminaries

In this chapter, matrices including vectors, are denoted by boldface uppercase letters. $a \doteq b$ means that b is the definition of a. $\text{diag}(P_{1,1}, P_{1,2}, ...P_{1,M})$ demonstrates an $M \times M$ matrix in which $(P_{1,1}, P_{1,2}, ...P_{1,M})$ is the main diagonal and all other entries are zero. For a square complex matrix $\boldsymbol{C}_1$, $\boldsymbol{C}_1[i]$ is the complex number that represents the i^{th} element of the main diagonal. For a complex vector $\boldsymbol{S}_1 = [S_{1,1}, S_{1,2}, ..., S_{1,M}]^T$, $\boldsymbol{S}_1[i]$ is the complex number that represents the i^{th} element of the vector, i.e., $\boldsymbol{S}_1[i] \doteq S_{1,i}$. The notation $\nabla(R_1)$ represents the gradient of the function R_1. $[x]^+ = x$ if $x > 0$, otherwise $[x]^+ = 0$, and $\log(x) = \log_2(x)$. The notation $\mathbb{1}(x \leq y) = 1$ if $x \leq y$, otherwise $\mathbb{1}(x \leq y) = 0$. Finally, for a complex number z, $|z|$ represents the magnitude of z.

5.2.1 Channel Model

In the system model studied in this chapter, a bandwidth of B is divided into M orthogonal sub-carriers and is shared between $2M$ links (a link is composed of a transmitter and its corresponding receiver). Therefore, we assume that M orthogonal sub-carriers are shared by two groups of transmitter-receiver pairs where each group has exactly M links, as depicted in Figure 5.1. In other words, the two groups share the entire

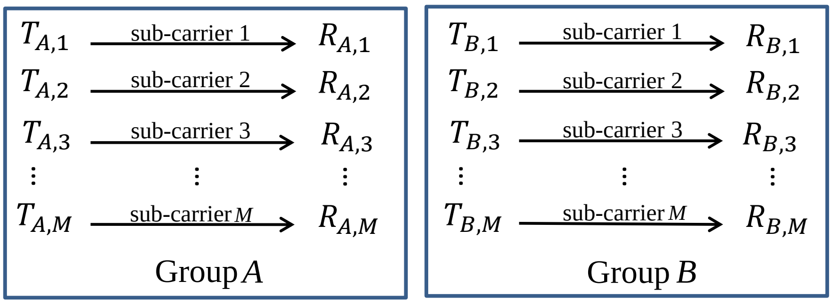

Figure 5.1: Two groups (A and B) of wireless transmitters sharing M sub-carriers of OFDMA.

bandwidth, based on Orthogonal Frequency-Division Multiple Access (OFDMA) with M sub-carriers. Each sub-carrier is used by both groups, and therefore, M parallel two-user GICs are formed, as shown in Figure 5.2. Note that OFDMA is used in many wireless standards. For instance, in the Long-Term Evolution (LTE) standard, used by many telecommunications providers, OFDMA is used in the down-link [85].

One of the main advantages of OFDM systems is their immunity to multi-path fading. When a signal $x(t)$ is transmitted through a channel with impulse response $c(t)$, the recieved signal $y(t)$ is expressed by a linear convolution as follows:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)c(t-\tau)d\tau + z(t), \tag{5.1}$$

where $z(t)$ represents the noise at the receiver. The multi-path delay can be modeled as

$$c(t) = \sum_{i=1}^{N_{path}} \zeta_i \delta(t-\tau_i), \tag{5.2}$$

where ζ_i and τ_i represent the gain and the delay of the i^{th} path, respectively. N_{path} is the number of paths, and $\delta(t)$ is the Dirac delta function. For this channel, the delay spread t_d is given by $t_d = \max\{\tau_i\} - \min\{\tau_i\}$. The receiver retrieves $x(t)$ from $y(t)$ using an equalizer; however, the complexity of implementing such an equalizer increases as N_{path} increases. The basic idea of OFDM is to transmit the message through narrow-band orthogonal sub-carriers, so that each sub-carrier experiences a complex gain, and consequently, the equalizer structure is simplified.

To realize an OFDM symbol of size M, a symbol of incoming data $\boldsymbol{S}(n) = [S_1(n), S_2(n), ..., S_i(n), ..., S$ is multiplied by an inverse Fast Fourier Transform (FFT) matrix to create the time-domain symbol $\boldsymbol{D}(n) = [D_1(n), D_2(n), ..., D_i(n), ..., D_M(n)]^T$, where n, i, and M represent the time index, the sub-carrier index and the symbol size, respectively. Note that one OFDM symbol conveys M messages. Furthermore, a cyclic prefix of size L_{cp} is added at the beginning of the OFDM symbol $\boldsymbol{D}(n)$. Then, a parallel to serial converter and a digital to analogue converter are used to generate the analogue signal $d(t)$, in which the OFDM symbol has duration t_0 and the cyclic prefix has duration t_{cp}. The cyclic prefix is used to avoid interference among sub-carriers. In fact, if the delay spread t_d is shorter than t_{cp} (or equivalently, if the channel impulse response is shorter than L_{cp}), the cyclic prefix turn the linear convolution into the cyclic convolution. Since circular convolution can be diagonalized in the Fourier basis [86], it can be verified that multi-path delay in the time domain is transformed into complex gains over sub-carriers in the frequency domain [87–89]. Therefore, by adding a redundancy of size L_{cp} to an OFDM symbol of size M, OFDM systems can effectively handle the multi-path fading. In OFDM systems, instead of dealing with the delay of each $D_i(n)$, the delay of $\boldsymbol{D}(n) = [D_1(n), D_2(n), ..., D_i(n), ..., D_M(n)]^T$ is managed by adding a cyclic prefix. This in turn results in the message embedded in each OFDM sub-carrier to be multiplied by a complex channel gain value, without any interaction with the rest of the messages embedded in other OFDM sub-carriers. This is the key idea that let us relax the traditional delay constraint, as will be further explained in Remark 5.1.

In this study, the i^{th} sub-carrier is used by both groups simultaneously; in group A, it is used by the i^{th} transmitter-receiver pair. Similarly, in group B, it is used by the i^{th} transmitter-receiver pair. From the receivers' points of view, the entire channel is similar to M parallel two-user GICs; therefore, in Figure 5.3, all transmitters of group A are gathered together and labeled $\boldsymbol{T}_A$ and all receivers of group A are labeled $\boldsymbol{R}_A$. In our notation, $\boldsymbol{T}_{A,i}$ and $\boldsymbol{R}_{A,i}$ represent the i^{th} transmitter and the i^{th} receiver of group A, respectively. Similarly, all transmitters of group B are labeled $\boldsymbol{T}_B$ and all receivers of group B are labeled $\boldsymbol{R}_B$.

Moreover, we assume that transmitters have full-duplex capability. When a signal is broadcasted from the i^{th} transmitter of $\boldsymbol{T}_A$, it is received by the other three nodes operating over the same sub-carrier, i.e., the i^{th} transmitter of $\boldsymbol{T}_B$, the i^{th} receiver of $\boldsymbol{R}_A$, and the i^{th} receiver of $\boldsymbol{R}_B$, passing through the corresponding channels with gains $\boldsymbol{C}_{12}[i]$, $\boldsymbol{G}_{11}[i]$, and $\boldsymbol{G}_{12}[i]$, respectively. Similarly, the broadcasted signal from $\boldsymbol{T}_{B,i}$ is received by $\boldsymbol{T}_{A,i}$, $\boldsymbol{R}_{B,i}$, and $\boldsymbol{R}_{A,i}$, affected by channel gains denoted by $\boldsymbol{C}_{21}[i]$, $\boldsymbol{G}_{22}[i]$, and $\boldsymbol{G}_{21}[i]$, respectively. Note that the self interference, i.e., the leakage from a transmitter's

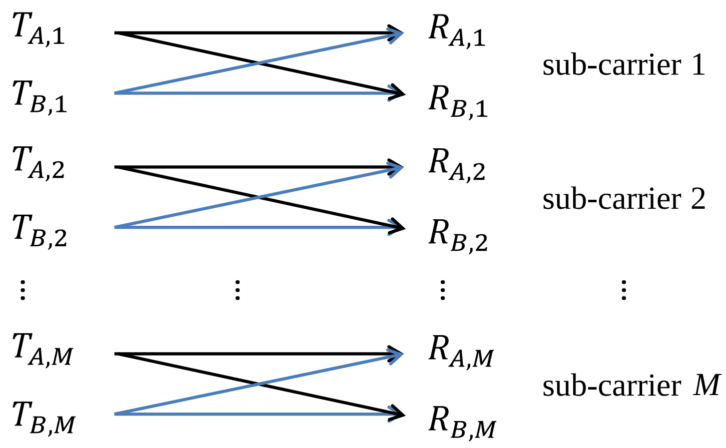

Figure 5.2: M parallel GICs formed across M sub-carriers of OFDMA.

antenna to the receiver of the same transmitter is assumed to be fully compensated. All considered channels between the nodes are illustrated with blue boxes in Figure 5.3. It is assumed that all channel gains are constant during the transmission of one OFDM symbol and are fully known by all transmitters and all receivers.

The goal of this study is to mitigate the interference through cooperation among transmitters when receivers simply treat the interference as noise. This interference mitigation is performed by a scheme that we call *superimposed interfere cancellation*. In this scheme, $\boldsymbol{T}_A$ superimposes a filtered version of the signal it receives from $\boldsymbol{T}_B$ on the original signal of $\boldsymbol{T}_A$. The filter is chosen such that the interference is canceled at $\boldsymbol{R}_A$. Similarly, $\boldsymbol{T}_B$ superimposes the signal it receives from $\boldsymbol{T}_A$ on its own signal to cancel the interference at $\boldsymbol{R}_B$. The filter at $\boldsymbol{T}_A$, which is used to cancel the interference at $\boldsymbol{R}_A$, is denoted by $\boldsymbol{F}_1$ and the filter at $\boldsymbol{T}_B$, which is used to cancel the interference at $\boldsymbol{R}_B$, is denoted by $\boldsymbol{F}_2$ (see Figure 5.3). Note that $\boldsymbol{F}_1$ represents M filters; each used by the corresponding transmitter of $\boldsymbol{T}_A$. The i^{th} transmitter of $\boldsymbol{T}_A$ is designed to cancel the interference only over the i^{th} sub-carrier at the i^{th} receiver of $\boldsymbol{R}_A$. As explained in the rest of this chapter, under some assumed conditions, the filter used by the i^{th} transmitter of $\boldsymbol{T}_A$, i.e., $\boldsymbol{F}_1[i]$, is a simple one tap filter that has a constant magnitude and a constant phase.

Note that $\boldsymbol{T}_A$ represents M distinct transmitters of group A, installed at different locations. Each transmitter communicates with a corresponding receiver over a narrow frequency band. Such a narrow frequency band is formed over a single OFDM sub-carrier, or a group of adjacent OFDM sub-carriers with equal gains. The requirement

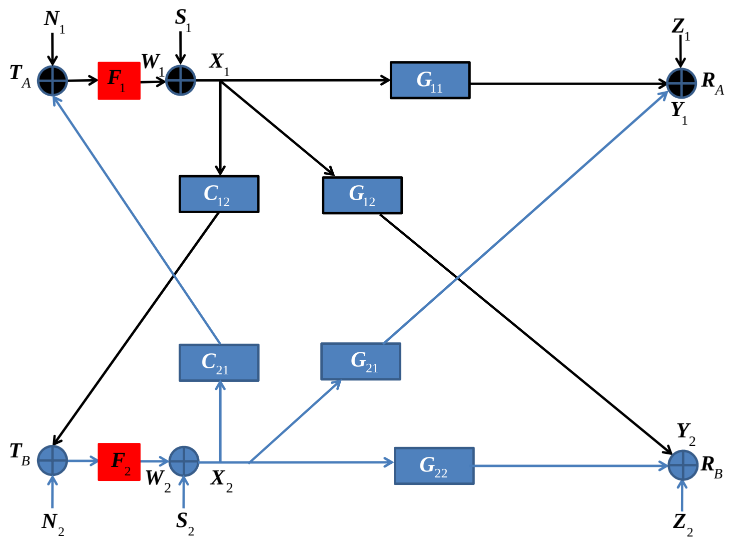

Figure 5.3: The equivalent GIC with full-duplex transmitters.

is that the channel formed over each of such narrow bands is frequency flat. Relying on this assumption, filters $\boldsymbol{F}_1[i]$ and $\boldsymbol{F}_2[i]$, for the i^{th} transmitter/receiver pair, will operate over a frequency flat channel. Consequently, each such filtering operation will be equivalent to multiplication by a complex number (phase and magnitude adjustment). Under this condition, each of these M pairs of filters can be implemented in the time domain, without introducing any additional delay, and without the need to filter/separate OFDM sub-carriers. Each such filter will introduce a phase and magnitude adjustment over the entire band. This will effectively provide the phase and magnitude adjustment required to cancel the (narrow band) interference term over the corresponding two-users GIC. Due to orthogonality of sub-carriers , each pair of filters, although operating over the entire band, will affect only its corresponding two-user GIC. In other words, such a filtering operation will be transparent to other transmitter/receiver pairs.

This model only requires that transmitters are physically separated, each operating over a narrow-band (flat) sub-channel. However, receivers can be either grouped together in one physical location, or be in separate locations. If the receivers are physically together, the model will correspond to the uplink in an OFDMA system. Note that in the uplink, coverage is typically governed by the limitation on the amount of power mobile nodes can transmit. Using a narrow band channel allows mobile units to concentrate their available power in a smaller band and satisfy the required link budget.

The case in which receivers are in separate locations corresponds to M physically separate transmitter receiver pairs (M two-users GIC) sharing the spectrum. Use of small cells in emerging wireless standards, such as LTE, would be an example for the application of such a set of separate transmitter/receiver pairs. Each link connects a micro/pico base-station to a client within a small cell. The model would capture $2M$ such small cells sharing the spectrum. In this case, frequency planning would ideally assign a pair of two-user CIG sharing a sub-carrier to small cells further away from each other, while neighboring small cells would be separated by assigning them to orthogonal sub-channels. This is aligned with our assumption in Theorem 5.1 that the product of the cross-link channel gains is smaller than the product of the direct-link channel gains.

It should be added that such filters can be implemented directly as part of the Radio Frequency (RF) front end as a simple tunable phase/magnitude adjustment of the incoming signal prior to combining it (in the RF domain) with the outgoing radio signal. Finally, note that although $\boldsymbol{F}_1[i]$ and $\boldsymbol{F}_2[i]$ are implemented in the time domain, in our notations, these are equivalently represented as complex multiplications in the frequency domain.

The signal received by $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$ are expressed as

$$\begin{aligned} \boldsymbol{Y}_1 &= \boldsymbol{G}_{11}\boldsymbol{X}_1 + \boldsymbol{G}_{21}\boldsymbol{X}_2 + \boldsymbol{Z}_1, \\ \boldsymbol{Y}_2 &= \boldsymbol{G}_{12}\boldsymbol{X}_1 + \boldsymbol{G}_{22}\boldsymbol{X}_2 + \boldsymbol{Z}_2, \end{aligned} \tag{5.3}$$

where $\boldsymbol{G}_{11}, \boldsymbol{G}_{21}, \boldsymbol{G}_{12}$, and $\boldsymbol{G}_{22}$ are diagonal $M \times M$ complex matrices, representing channel gains. $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are complex $M \times 1$ vectors, representing the outputs of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$, respectively. Furthermore, $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are $M \times 1$ vectors, representing the zero-mean unit-variance complex Gaussian noise of $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$, respectively. $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ have independent equal variance real and imaginary parts. As depicted in Figure 5.3, since $\boldsymbol{T}_A$ has full-duplex capability, its output, $\boldsymbol{X}_1$, is the sum of the $\boldsymbol{S}_1$, i.e., an $M \times 1$ Gaussian vector that represents the original message of $\boldsymbol{T}_A$, and $\boldsymbol{W}_1 = \boldsymbol{F}_1(\boldsymbol{C}_{21}\boldsymbol{X}_2 + \boldsymbol{N}_1)$. $\boldsymbol{W}_1$ is an $M \times 1$ vector, and it represents the filtered signal that $\boldsymbol{T}_A$ receives from $\boldsymbol{T}_B$. Similarly, $\boldsymbol{X}_2$ is the sum of $\boldsymbol{S}_2$ and $\boldsymbol{W}_2 = \boldsymbol{F}_2(\boldsymbol{C}_{12}\boldsymbol{X}_1 + \boldsymbol{N}_2)$, therefore,

$$\begin{aligned} \boldsymbol{X}_1 &= \boldsymbol{S}_1 + \boldsymbol{F}_1(\boldsymbol{C}_{21}\boldsymbol{X}_2 + \boldsymbol{N}_1), \\ \boldsymbol{X}_2 &= \boldsymbol{S}_2 + \boldsymbol{F}_2(\boldsymbol{C}_{12}\boldsymbol{X}_1 + \boldsymbol{N}_2), \end{aligned} \tag{5.4}$$

where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ represent zero-mean unit-variance complex Gaussian noise of the receivers of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$, respectively. Moreover, all noises are independent of each other and are independent of the inputs $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. Define $\boldsymbol{L}$ as the gain of the loop between transmitters. Therefore,

$$\boldsymbol{L} = \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12}. \tag{5.5}$$

Note that $\boldsymbol{F}_1$, $\boldsymbol{C}_{21}$, $\boldsymbol{F}_2$, and $\boldsymbol{C}_{12}$ are all diagonal matrices, and therefore, are commuting matrices. From (5.4), $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are expressed as functions of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ as follows:

$$\begin{aligned}\boldsymbol{X}_1 &= (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2))(\boldsymbol{I} - \boldsymbol{L})^{-1},\\ \boldsymbol{X}_2 &= (\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2 + \boldsymbol{F}_2\boldsymbol{C}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1))(\boldsymbol{I} - \boldsymbol{L})^{-1}.\end{aligned} \tag{5.6}$$

Note that $\boldsymbol{S}_j(n) = [S_{j,1}(n), S_{j,2}(n), ..., S_{j,M}(n)]^T$, $j \in \{1, 2\}$, represents an OFDM symbol of size M, transmitted through M orthogonal sub-carriers, and n represents the time index. In this chapter, whenever the time index n is clear from the context, it will be omitted.

Although the M transmitters of $\boldsymbol{T}_A$ can have different powers, we impose a power constraint on the total power transmitted by all of them. Consequently, the transmission power at $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$ are bounded by P_1 and P_2, respectively. The justification for this power constraint is further discussed, when we investigate the optimal power allocation. Since in OFDMA, the sub-carriers are orthogonal, the transmission power constraint is applied over the M orthogonal sub-carriers. Consequently, the total power of $\boldsymbol{X}_1$, which is the sum of powers of $\boldsymbol{X}_1[i]s, i \in \{1, 2, ..., M\}$, is restricted to P_1, and the total power of $\boldsymbol{X}_2$, which is the sum of powers of $\boldsymbol{X}_2[i]s, i \in \{1, 2, ..., M\}$, is restricted to P_2.

5.2.2 Causal Cooperation

Next, we examine the constraint of causal cooperation, used in this study. In a causal cooperation among transmitters, $\boldsymbol{X}_1(n) = [X_{1,1}(n), X_{1,2}(n), ..., X_{1,M}(n)]^T$ can be a function of its received signal, i.e., $\boldsymbol{W}_1(n-1), \boldsymbol{W}_1(n-2), ..., \boldsymbol{W}_1(1)$. Moreover, $\boldsymbol{T}_A$ can superimpose $\boldsymbol{W}_1(n)$ on $\boldsymbol{S}_1(n)$. Similarly, $\boldsymbol{X}_2(n)$ can only depend on $\boldsymbol{W}_2(n-1), ..., \boldsymbol{W}_2(1)$, and $\boldsymbol{T}_B$ can superimpose $\boldsymbol{W}_2(n)$ on $\boldsymbol{S}_2(n)$. Note that to achieve a multiplexing gain of two, transmitters do not need to use the past received signals. In this study, we show that by just superimposing $\boldsymbol{W}_j(n)$ on $\boldsymbol{S}_j(n)$, a multiplexing gain of two is achievable.

Remark 5.1. *Causal delay: Since $\boldsymbol{T}_A$ is not aware of the message of $\boldsymbol{T}_B$, to consider a causal scenario, it is traditionally assumed that $\boldsymbol{X}_1(n)$ can only depend on $\boldsymbol{W}_1(n-1), \boldsymbol{W}_1(n-2), ..., \boldsymbol{W}_1(1)$. With this traditional constraint of causal delay, [39] proves that cooperation among transmitters does not increase the multiplexing gain of the two-user GIC. However, in this study, $\boldsymbol{T}_A$ can filter the signal it receives from $\boldsymbol{T}_B$ and superimpose it on its own signal. Note that $\boldsymbol{T}_A$ does not decode $\boldsymbol{X}_2(n) = [X_{2,1}(n), X_{2,2}(n), ..., X_{2,M}(n)]^T$ and does not use it to encode $\boldsymbol{X}_1(n)$. The justification behind our assumption is the fact that the actual delay is determined by the channel memory length and not by one symbol length. In OFDM systems, as far as the maximum delay spread is less than the t_{cp}, the effect of multi-path delay is just M complex gains over the M parallel sub-channels in the frequency domain. Note that the message embedded in each OFDM sub-carrier will be detectable only after the entire OFDM symbol is received, however this extension in time is the same for all paths, and it is consistent with the OFDM structure. Due to the cyclic prefix, this results in a simple linear combination of the desired signal and the interference over each OFDM sub-carrier.*

In this setup, the role of the relaying of the interfering signal is equivalent to creating some additional paths in the propagation of the OFDM symbol, and consequently, as long as all the paths corresponding to any given OFDM symbol are received by the destination within a delay spread satisfying the cyclic prefix condition, the superposition principal over each sub-carrier will be valid. With this idea, we can capture the role of delay inside the OFDM symbol. A longer delay requires a longer cyclic prefix. For a fixed M, as the size of the cyclic prefix L_{cp} increases, the effective rate of the OFDM symbol decreases. In the next section, we investigate this issue.

In the scenario investigated in this chapter, the signal transmitted by $\boldsymbol{T}_B$, reaches $\boldsymbol{R}_A$ through two distinct paths: a direct path from $\boldsymbol{T}_B$ to $\boldsymbol{R}_A$ and an indirect path from $\boldsymbol{T}_B$ to $\boldsymbol{T}_A$ and then from $\boldsymbol{T}_A$ to $\boldsymbol{R}_A$. Therefore, as far as the total delay spread, including the delay from $\boldsymbol{T}_B$ to $\boldsymbol{T}_A$ and the processing delay in $\boldsymbol{T}_A$ and the delay from $\boldsymbol{T}_A$ to $\boldsymbol{R}_A$, is less than the cyclic prefix duration, the i^{th} transmitter of $\boldsymbol{T}_A$ can deploy a proper filter, i.e., $\boldsymbol{F}_1[i]$ to apply the required gain/phase shift in the indirect path. Note that since each transmitter operates over a sub-carrier of OFDMA, such filtering operation can be performed in time by operating over successive time samples of each OFDM symbol as these are received and relayed. With this filtering, the i^{th} receiver of $\boldsymbol{R}_A$ experiences an

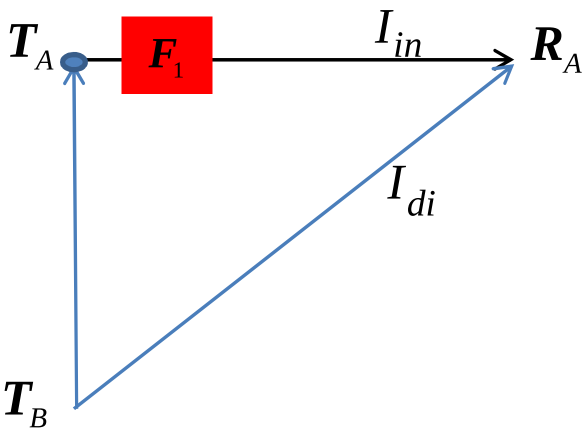

Figure 5.4: The interference, caused by $\boldsymbol{T}_B$, reaches $\boldsymbol{R}_A$ directly by I_{di} and indirectly by I_{in}. The filter $\boldsymbol{F}_1$ can guarantee that $I_{di} + I_{in} = 0$.

interference-free sub-channel. This is depicted in Figure 5.4 for the signals transmitted by $\boldsymbol{T}_B$ that reaches $\boldsymbol{R}_A$ trough two distinct paths. In Section 5.3, we characterize the direct interference $\boldsymbol{I}_{di}$ and the indirect interference $\boldsymbol{I}_{in}$. Then, we compute the filter $\boldsymbol{F}_1$ that satisfies $\boldsymbol{I}_{di} + \boldsymbol{I}_{in} = 0$.

In this chapter, multiplexing gain is used to investigate the role of cooperation in the achievable rate region of the two-user GIC. Intuitively, multiplexing gain is the factor in front of log(SNR) in the expression of the achievable sum-rate. Mathematically, the following is used as the definition of multiplexing gain [37]:

Definition 5.1. For the two-user GIC, a multiplexing gain of l is said to be achievable, if for $P_1 = P_2 = P$, there exists a coding scheme that achieves the rate tuple $(R_1(P), R_2(P))$, such that

$$\limsup_{P\to\infty} \frac{R_1(P) + R_2(P)}{\log(P)} = l. \tag{5.7}$$

In the next section, we show that a simple causal cooperation among full-duplex transmitters of the two-user GIC, can achieve a multiplexing gain of two.

5.3 Interference Cancellation with Full-Duplex Transmitters

In this section, we first study the case in which transmitters of only one group are full-duplex. Then, we investigate the general case in which transmitters of both groups are full-duplex.

5.3.1 The Two-User GIC with One Full-Duplex Transmitter

We first study the case in which only $\boldsymbol{T}_A$ is full-duplex. This would be equivalent to the general case depicted in Figure 5.3 if we let $\boldsymbol{F}_2 = \boldsymbol{N}_2 = 0$. In fact, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are given by

$$\boldsymbol{X}_1 = \boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2, \tag{5.8}$$

$$\boldsymbol{X}_2 = \boldsymbol{S}_2. \tag{5.9}$$

The signal received by receivers, i.e., $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are expressed as

$$\begin{aligned}\boldsymbol{Y}_1 &= \boldsymbol{G}_{11}\boldsymbol{X}_1 + \boldsymbol{G}_{21}\boldsymbol{X}_2 + \boldsymbol{Z}_1\\ &= \boldsymbol{G}_{11}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2) + \boldsymbol{G}_{21}\boldsymbol{S}_2 + \boldsymbol{Z}_1\\ &= \boldsymbol{G}_{11}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1) + (\boldsymbol{G}_{11}\boldsymbol{F}_1\boldsymbol{C}_{21} + \boldsymbol{G}_{21})\boldsymbol{S}_2 + \boldsymbol{Z}_1, \end{aligned}\tag{5.10}$$

$$\begin{aligned}\boldsymbol{Y}_2 &= \boldsymbol{G}_{12}\boldsymbol{X}_1 + \boldsymbol{G}_{22}\boldsymbol{X}_2 + \boldsymbol{Z}_2\\ &= \boldsymbol{G}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2) + \boldsymbol{G}_{22}\boldsymbol{S}_2 + \boldsymbol{Z}_2\\ &= \boldsymbol{G}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1) + (\boldsymbol{G}_{12}\boldsymbol{F}_1\boldsymbol{C}_{21} + \boldsymbol{G}_{22})\boldsymbol{S}_2 + \boldsymbol{Z}_2. \end{aligned}\tag{5.11}$$

As expressed in (5.10), the interference caused by $\boldsymbol{S}_2$, reaches $\boldsymbol{R}_A$ though two distinct paths. Directly, through $\boldsymbol{G}_{21}$, $\boldsymbol{S}_2$ causes the interference $\boldsymbol{I}_{di}$, which is expressed as

$$\boldsymbol{I}_{di} = \boldsymbol{G}_{21}\boldsymbol{S}_2, \tag{5.12}$$

and indirectly, through $\boldsymbol{G}_{11}\boldsymbol{F}_1\boldsymbol{C}_{21}$, $\boldsymbol{S}_2$ causes the interference $\boldsymbol{I}_{in}$, which is expressed as

$$\boldsymbol{I}_{in} = (\boldsymbol{G}_{11}\boldsymbol{F}_1\boldsymbol{C}_{21})\boldsymbol{S}_2. \tag{5.13}$$

To cancel the interference, we choose $\boldsymbol{F}_1$ such that

$$\boldsymbol{I}_{di} + \boldsymbol{I}_{in} = 0. \tag{5.14}$$

Therefore, $\boldsymbol{F}_1$ is given by

$$\boldsymbol{F}_1 = -\boldsymbol{G}_{21}(\boldsymbol{G}_{11}\boldsymbol{C}_{21})^{-1}. \tag{5.15}$$

Note that $\boldsymbol{R}_B$ simply treats the existing interference as noise. Consequently, the achievable rate of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$ are given by

$$R_1 = \sum_{i=1}^{M} \log(1 + \frac{A_1^i P_1^i}{J_1^i}),$$
$$R_2 = \sum_{i=1}^{M} \log(1 + \frac{A_2^i P_2^i}{B_2^i P_1^i + J_2^i}), \tag{5.16}$$

where P_1^i and P_2^i represent the power of $S_{1,i}$ and $S_{2,i}$, respectively. A_1^i and A_2^i represent the effective channel gains at the i^{th} receiver of $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$, respectively. $B_2^i P_1^i$ represents the power of the interference that the i^{th} receiver of $\boldsymbol{R}_B$ experiences. Finally, J_1^i and J_2^i determine the power of the effective noise at the i^{th} receiver of $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$, respectively. Moreover, according to (5.10) and (5.11), we have

$$A_1^i = |\boldsymbol{G}_{11}[i]|^2,$$
$$A_2^i = |(\boldsymbol{G}_{12}\boldsymbol{F}_1\boldsymbol{C}_{21} + \boldsymbol{G}_{22})[i]|^2,$$
$$B_2^i = |\boldsymbol{G}_{12}[i]|^2,$$
$$J_1^i = |(\boldsymbol{G}_{11}\boldsymbol{F}_1)[i]|^2 + 1,$$
$$J_2^i = |(\boldsymbol{G}_{12}\boldsymbol{F}_1)[i]|^2 + 1. \tag{5.17}$$

Moreover, since $\boldsymbol{X}_2 = \boldsymbol{S}_2$, the power constraint of $\boldsymbol{T}_B$ is simply $\sum_{i=1}^{M} P_2^i \leq P_2$. However, since $\boldsymbol{X}_1 = \boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2$, the power constraint of $\boldsymbol{T}_A$ is a constraint involving P_1^i and P_2^i. In fact, to show that the total power of $\boldsymbol{X}_1$ is restricted to P_1, we have

$$\sum_{i=1}^{M} P_1^i + |\boldsymbol{F}_1[i]|^2 + |(\boldsymbol{F}_1\boldsymbol{C}_{21})[i]|^2 P_2^i \leq P_1. \tag{5.18}$$

This constraint can be rewritten as follows:

$$\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \leq E_1, \tag{5.19}$$

where according to (5.18), C_1^i, D_1^i, and E_1 are expressed as

$$C_1^i = 1,$$
$$D_1^i = |(\boldsymbol{F}_1\boldsymbol{C}_{21})[i]|^2,$$
$$E_1 = P_1 - \sum_{i=1}^{M} |\boldsymbol{F}_1[i]|^2. \tag{5.20}$$

By canceling the interference using filter $\boldsymbol{F}_1$, given in (5.15), R_1 increases proportionally to $\log(P_1)$. The power of the noise of the i^{th} sub-carrier of $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$ has increased from 1 to J_1^i and J_2^i, respectively. Moreover, because of power constraint (5.19), $\sum_{i=1}^{M} P_1^i$ is strictly less than P_1. However, as can be seen from (5.16) and (5.20), the full-duplex transmitter can cancel interference at its receiver, and in the high SNR regime, R_1 is proportional to $\log(P_1)$ while R_2 is proportional to $\log(\frac{P_2}{P_1})$. In the next subsection, we show that if both transmitters are full-duplex, then both R_1 and R_2 can increase proportionally to $\log(P_1)$ and $\log(P_2)$, respectively.

5.3.2 The Two-User GIC with Two Full-Duplex Transmitters

In this subsection, we study the general case, where transmitters of both groups are full-duplex. Transmitters superimpose the signal they receive on their own messages such that their intended receiver will see no interference. The signal received at $\boldsymbol{R}_A$ is

$$
\begin{aligned}
\boldsymbol{Y}_1 =& \boldsymbol{G}_{11}\boldsymbol{X}_1 + \boldsymbol{G}_{21}\boldsymbol{X}_2 + \boldsymbol{Z}_1 \\
\overset{(a)}{=}& \boldsymbol{G}_{11}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2))(\boldsymbol{I}-\boldsymbol{L})^{-1} \\
&+ \boldsymbol{G}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2 + \boldsymbol{F}_2\boldsymbol{C}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1))(\boldsymbol{I}-\boldsymbol{L})^{-1} + \boldsymbol{Z}_1 \\
=& (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)(\boldsymbol{G}_{11} + \boldsymbol{G}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12})(\boldsymbol{I}-\boldsymbol{L})^{-1} \\
&+ (\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)(\boldsymbol{G}_{21} + \boldsymbol{G}_{11}\boldsymbol{F}_1\boldsymbol{C}_{21})(\boldsymbol{I}-\boldsymbol{L})^{-1} + \boldsymbol{Z}_1,
\end{aligned} \tag{5.21}
$$

where (a) is valid by (5.6). Similarly, $\boldsymbol{Y}_2$ can be expressed in terms of $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, and $\boldsymbol{Z}_2$, as follows:

$$
\begin{aligned}
\boldsymbol{Y}_2 =& \boldsymbol{G}_{12}\boldsymbol{X}_1 + \boldsymbol{G}_{22}\boldsymbol{X}_2 + \boldsymbol{Z}_2 \\
=& (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)(\boldsymbol{G}_{12} + \boldsymbol{G}_{22}\boldsymbol{F}_2\boldsymbol{C}_{12})(\boldsymbol{I}-\boldsymbol{L})^{-1} \\
&+ (\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)(\boldsymbol{G}_{22} + \boldsymbol{G}_{12}\boldsymbol{F}_1\boldsymbol{C}_{21})(\boldsymbol{I}-\boldsymbol{L})^{-1} + \boldsymbol{Z}_2.
\end{aligned} \tag{5.22}
$$

In addition, the power constraints for $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are expressed as

$$
\begin{aligned}
\sum_{i=1}^{M} \mathbb{E}[|X_{1,i}|^2] \le P_1, \\
\sum_{i=1}^{M} \mathbb{E}[|X_{2,i}|^2] \le P_2.
\end{aligned} \tag{5.23}
$$

Since $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are given by (5.6), we obtain

$$X_{1,i} = \frac{\big(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\big)[i]}{1 - \boldsymbol{L}[i]},$$
$$X_{2,i} = \frac{\big(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2 + \boldsymbol{F}_2\boldsymbol{C}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\big)[i]}{1 - \boldsymbol{L}[i]}. \tag{5.24}$$

Therefore, we can rewrite (5.23) as follows:

$$\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \le E_1,$$
$$\sum_{i=1}^{M} C_2^i P_1^i + D_2^i P_2^i \le E_2, \tag{5.25}$$

where P_1^i and P_2^i represent the power of $S_{1,i}$ and $S_{2,i}$, respectively, as expressed by the following equations:

$$\mathbb{E}[|S_{1,i}|^2] = P_1^i,$$
$$\mathbb{E}[|S_{2,i}|^2] = P_2^i. \tag{5.26}$$

According to (5.24), for $j \in \{1,2\}$, C_j^i, D_j^i, and E_j are given by

$$C_1^i = \left|\frac{1}{1 - \boldsymbol{L}[i]}\right|^2, \quad C_2^i = \left|\frac{(\boldsymbol{F}_2\boldsymbol{C}_{12})[i]}{1 - \boldsymbol{L}[i]}\right|^2,$$
$$D_1^i = \left|\frac{(\boldsymbol{F}_1\boldsymbol{C}_{21})[i]}{1 - \boldsymbol{L}[i]}\right|^2, \; D_2^i = \left|\frac{1}{1 - \boldsymbol{L}[i]}\right|^2,$$
$$E_j = P_j - \sum_{i=1}^{M} C_j^i |\boldsymbol{F}_1[i]|^2 - \sum_{i=1}^{M} D_j^i |\boldsymbol{F}_2[i]|^2. \tag{5.27}$$

A simple power allocation scheme is the uniform power allocation. When the entire power is allocated uniformly across all sub-carriers, $P_1^i = P_1^k = \breve{P}_1$ and $P_2^i = P_2^k = \breve{P}_2$ for $i,k \in \{1,2,...,M\}$. Therefore, according to (5.25), we obtain

$$\breve{P}_1 \sum_{i=1}^{M} C_j^i + \breve{P}_2 \sum_{i=1}^{M} D_j^i = E_j, j \in \{1,2\}, \tag{5.28}$$

which results in the following expressions for the uniform power allocation:

$$\breve{P}_1 = \frac{E_1 \sum_{i=1}^{M} D_2^i - E_2 \sum_{i=1}^{M} D_1^i}{\sum_{i=1}^{M} C_1^i \sum_{i=1}^{M} D_2^i - \sum_{i=1}^{M} C_2^i \sum_{i=1}^{M} D_1^i},$$
$$\breve{P}_2 = \frac{E_1 \sum_{i=1}^{M} C_2^i - E_2 \sum_{i=1}^{M} C_1^i}{\sum_{i=1}^{M} D_1^i \sum_{i=1}^{M} C_2^i - \sum_{i=1}^{M} D_2^i \sum_{i=1}^{M} C_1^i}. \tag{5.29}$$

Note that in the high SNR regime, in which $P_1 = P_2 = P$ approach infinity, we have

$$\lim_{P\to\infty} \frac{\breve{P}_1}{P} = c_1,$$
$$\lim_{P\to\infty} \frac{\breve{P}_2}{P} = c_2, \tag{5.30}$$

where c_1 and c_2 are positive constants. This means that in the high SNR regime, a fixed portion of the entire power is allocated to each sub-carrier. The following theorem shows that if both transmitters are full-duplex, using the uniform power allocation, a multiplexing gain of two is achievable.

Theorem 5.1. *When the magnitude of the product of cross-link channel gains is smaller than the magnitude of the product of direct-link channel gains, the maximum multiplexing gain of the two-user GIC with full-duplex transmitters is equal to two.*

Proof. The proof has two main parts. First, we show that full-duplex transmitters can use filters $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ to simultaneously cancel the interference at their receivers. To do so, transmitters use OFDM symbols of size M with cyclic prefix of size L_{cp}. Second, we show that in the high SNR regime, M and L_{cp} can be chosen such that the use of cyclic prefix does not reduce the multiplexing gain. As it will be shown later, we need to assume that

$$|\boldsymbol{G}_{12}[i]\boldsymbol{G}_{21}[i]| \leq |\boldsymbol{G}_{11}[i]\boldsymbol{G}_{22}[i]|, \tag{5.31}$$

for all $i \in \{1, 2, ..., M\}$. This means the magnitude of the product of cross-link channel gains is smaller than the magnitude of the product of direct-link channel gains.

According to (5.21), $\boldsymbol{T}_A$ can cancel the interference at $\boldsymbol{R}_A$, if $\boldsymbol{G}_{21} + \boldsymbol{G}_{11}\boldsymbol{F}_1\boldsymbol{C}_{21} = 0$. Consequently, $\boldsymbol{T}_A$ uses the following filter:

$$\boldsymbol{F}_1 = -\boldsymbol{G}_{21}(\boldsymbol{G}_{11}\boldsymbol{C}_{21})^{-1}. \tag{5.32}$$

Similarly, $\boldsymbol{T}_B$ can cancel the interference at $\boldsymbol{R}_B$, if the following filter is used by $\boldsymbol{T}_B$:

$$\boldsymbol{F}_2 = -\boldsymbol{G}_{12}(\boldsymbol{G}_{22}\boldsymbol{C}_{12})^{-1}. \tag{5.33}$$

When the interference is canceled, $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are given by

$$\boldsymbol{Y}_1 = (\boldsymbol{S}_1 + \boldsymbol{F_1N_1})(\boldsymbol{G}_{11} + \boldsymbol{G}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12})(\boldsymbol{I} - \boldsymbol{L})^{-1} + \boldsymbol{Z}_1,$$
$$\boldsymbol{Y}_2 = (\boldsymbol{S}_2 + \boldsymbol{F_2N_2})(\boldsymbol{G}_{22} + \boldsymbol{G}_{12}\boldsymbol{F}_1\boldsymbol{C}_{21})(\boldsymbol{I} - \boldsymbol{L})^{-1} + \boldsymbol{Z}_2. \tag{5.34}$$

Let R_1 and R_2 denote the achievable rate of $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$, respectively. Then, R_1 and R_2 are given by

$$R_j = \sum_{i=1}^{M} \log\Big(1 + \frac{A_j^i P_j^i}{J_j^i}\Big), \tag{5.35}$$

where A_j^i and J_j^i represent the effective gain and the power of the effective noise at the i^{th} sub-carrier of $\boldsymbol{Y}_j$, $i \in \{1, ..., M\}$ and $j \in \{1, 2\}$, respectively. According to (5.34), these quantities are described in terms of $\boldsymbol{L} = \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12}$ as follows:

$$\begin{aligned}
A_1^i &= \left|\frac{(\boldsymbol{G}_{11} + \boldsymbol{G}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12})[i]}{1 - \boldsymbol{L}[i]}\right|^2, \\
A_2^i &= \left|\frac{(\boldsymbol{G}_{22} + \boldsymbol{G}_{12}\boldsymbol{F}_1\boldsymbol{C}_{21})[i]}{1 - \boldsymbol{L}[i]}\right|^2, \\
J_1^i &= |\boldsymbol{F}_1[i]|^2 A_1^i + 1, \\
J_2^i &= |\boldsymbol{F}_2[i]|^2 A_2^i + 1.
\end{aligned} \tag{5.36}$$

Although the achievable sum-rate depends on the value of A_j^i and J_j^i, the achievable multiplexing gain does not. The achievable multiplexing gain can be computed by letting $P_1 = P_2 = P \to \infty$. Note that according to (5.30), for the uniform power allocation, $\lim_{P\to\infty} \frac{\log(P_1^i)}{\log(P)} = \lim_{P\to\infty} \frac{\log(P_2^i)}{\log(P)} = 1$ for all $i \in \{1, 2, ..., M\}$. Therefore, for OFDM symbols of size M, we have

$$\begin{aligned}
&\limsup_{P\to\infty} \frac{R_1(P) + R_2(P)}{M\log(P)} \\
=&\limsup_{P\to\infty} \frac{\sum_{i=1}^{M} \log(1 + \frac{A_1^i P_1^i}{J_1^i}) + \sum_{i=1}^{M} \log(1 + \frac{A_2^i P_2^i}{J_2^i})}{M\log(P)} \\
=&\limsup_{P\to\infty} \frac{\sum_{i=1}^{M} \log(P_1^i) + \sum_{i=1}^{M} \log(P_2^i)}{M\log(P)} \\
=&2,
\end{aligned} \tag{5.37}$$

which shows the achievability of a multiplexing gain of two.

For the above argument to be valid, we should show that the cyclic prefix does not decrease the achievable multiplexing gain. The addition of the cyclic prefix at the beginning of the OFDM symbol decreases the spectral efficiency. When a cyclic prefix of size L_{cp} is added to an OFDM symbol of size M, the effective rate would decrease by an efficiency factor of $\frac{M}{M+L_{cp}}$. By choosing a large M, or equivalently a large duration t_0, this efficiency factor tends to one.

In addition, the time duration, required to reach the steady state signals expressed in (5.6), should be smaller than the cyclic prefix duration t_{cp}. More precisely, let t_{lp} be the loop duration, i.e., the time required for the signal to go from $\boldsymbol{T}_A$ to $\boldsymbol{T}_B$ and to come back from $\boldsymbol{T}_B$ to $\boldsymbol{T}_A$. In the following, we show that if the cyclic prefix duration is greater than a certain multiple of t_{lp}, a multiplexing gain of two is achievable.

Assume that a time duration of $k \times t_{lp}$ has passed, where k is an integer. As can be seen in Figure 5.3, we have

$$\boldsymbol{X}_1 = \sum_{l=0}^{k} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l + \sum_{l=0}^{k} \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l. \tag{5.38}$$

When k goes to infinity, (5.38) will be equivalent to (5.6), as explained by the following expression:

$$\begin{aligned}&\sum_{l=0}^{\infty} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l + \sum_{l=0}^{\infty} \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l \\ &= (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2))(\boldsymbol{I} - \boldsymbol{L})^{-1}.\end{aligned} \tag{5.39}$$

However, for a finite k, we have

$$\begin{aligned}\boldsymbol{X}_1 =& \sum_{l=0}^{k} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l + \sum_{l=0}^{k} \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l \\ =& \sum_{l=0}^{\infty} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l + \sum_{l=0}^{\infty} \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l \\ &- \sum_{l=k+1}^{\infty} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l - \sum_{l=k+1}^{\infty} \boldsymbol{F_1}\boldsymbol{C_{21}}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l \\ =& (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2))(\boldsymbol{I} - \boldsymbol{L})^{-1} - \boldsymbol{N}_A,\end{aligned} \tag{5.40}$$

where the last equality is valid by (5.39). Therefore, after a time duration of $k \times t_{lp}$, $\boldsymbol{X}_1$ is equal to the steady state expression given in (5.6) and a noise term $\boldsymbol{N}_A$, where $\boldsymbol{N}_A = [N_{A,1}, N_{A,2}, ...N_{A,M}]^T$ is an $M \times 1$ vector, representing the noise experienced by receivers of $\boldsymbol{T}_A$, caused by approximating (5.38) by (5.6). According to (5.40), $\boldsymbol{N}_A$ is expressed as

$$\begin{aligned}\boldsymbol{N}_A =& \sum_{l=k+1}^{\infty} (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l + \sum_{l=k+1}^{\infty} \boldsymbol{F_1}\boldsymbol{C_{21}}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l \\ =& (\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1 + \boldsymbol{F}_1\boldsymbol{C}_{21}(\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2))\boldsymbol{L}^{k+1}(\boldsymbol{I} - \boldsymbol{L})^{-1}.\end{aligned} \tag{5.41}$$

Similarly, after a time duration of $k \times t_{lp}$, $\boldsymbol{T}_B$ will experience an additional noise term $\boldsymbol{N}_B$, which is given by

$$\begin{aligned} \boldsymbol{N}_B = & \sum_{l=k+1}^{\infty} (\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2)\boldsymbol{L}^l + \sum_{l=k+1}^{\infty} \boldsymbol{F_2 C_{12}}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1)\boldsymbol{L}^l \\ = & (\boldsymbol{S}_2 + \boldsymbol{F}_2\boldsymbol{N}_2 + \boldsymbol{F}_2\boldsymbol{C}_{12}(\boldsymbol{S}_1 + \boldsymbol{F}_1\boldsymbol{N}_1))\boldsymbol{L}^{k+1}(\boldsymbol{I} - \boldsymbol{L})^{-1}. \end{aligned} \tag{5.42}$$

In (5.40), (5.41), and (5.42), we have used the following geometric series:

$$\sum_{l=0}^{\infty} \boldsymbol{L}^l = (\boldsymbol{I} - \boldsymbol{L})^{-1} = (\boldsymbol{I} - \boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{F}_2\boldsymbol{C}_{12})^{-1}, \tag{5.43}$$

$$\sum_{l=k+1}^{\infty} \boldsymbol{L}^l = \boldsymbol{L}^{k+1}(\boldsymbol{I} - \boldsymbol{L})^{-1}, \tag{5.44}$$

where these equalities are valid if $|\boldsymbol{L}[i]| < 1$ for all $i \in \{1, 2, ..., M\}$. Note that

$$|\boldsymbol{L}[i]| = |(\boldsymbol{C}_{21}\boldsymbol{F}_1 C_{12}\boldsymbol{F}_2)[i]| \overset{(a)}{=} |(\boldsymbol{G}_{12}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{22}^{-1})[i]|, \tag{5.45}$$

where (a) is valid by (5.32) and (5.33) and the fact that $\boldsymbol{C}_{21}$, $\boldsymbol{C}_{12}$, $\boldsymbol{G}_{11}$, $\boldsymbol{G}_{22}$, $\boldsymbol{G}_{12}$, and $\boldsymbol{G}_{21}$ are all commuting matrices. Therefore, $|\boldsymbol{L}[i]| < 1$ is equivalent to

$$|(\boldsymbol{G}_{12}\boldsymbol{G}_{21}\boldsymbol{G}_{11}^{-1}\boldsymbol{G}_{22}^{-1})[i]| \leq 1. \tag{5.46}$$

Note that (5.46) means that over all sub-carriers, the magnitude of the product of cross-link channel gains should be smaller than that of the product of direct-link channel gains, i.e.,

$$|\boldsymbol{G}_{12}[i]\boldsymbol{G}_{21}[i]| \leq |\boldsymbol{G}_{11}[i]\boldsymbol{G}_{22}[i]|, \tag{5.47}$$

which was assumed in Theorem 1. (5.47) is reminiscent of the weak interference condition, in which each cross-link channel gain is smaller that the corresponding direct-link channel gain. The interference channel formed across the i^{th} sub-carrier is weak if we have

$$\begin{aligned} |\boldsymbol{G}_{12}[i]| \leq |\boldsymbol{G}_{22}[i]|, \\ |\boldsymbol{G}_{21}[i]| \leq |\boldsymbol{G}_{11}[i]|. \end{aligned} \tag{5.48}$$

Clearly, if the interference channels formed across all sub-carriers are weak, then the constraint $|\boldsymbol{L}[i]| < 1$ is satisfied for all $i \in \{1, 2, .., M\}$.

Receivers of $\boldsymbol{T}_A$ can consider $\boldsymbol{N}_A$ as an extra noise, in addition to $\boldsymbol{N}_1$. To achieve a multiplexing gain of two, it would be enough to keep the power of $N_{A,i}$ at the same level as the power of $N_{1,i}$. More precisely, as P_1 and P_2 go to infinity, we should make sure that $\max_i\{\mathbb{E}[|N_{A,i}|^2]\}$ does not approach infinity, so that in the high SNR regime, the effect of the power of $\boldsymbol{N}_A$ on the achievable sum-rate is negligible. Assume that power is allocated uniformly according to (5.29). Note that in the high SNR regime, both P_1^i and P_2^i are proportional to P as highlighted in (5.30). Consequently, for $P_1 = P_2 = P$, a multiplexing gain of two is achievable if we can find two positive constants b_1 and b_2 such that the following inequalities are satisfied:

$$\lim_{P\to\infty} \max_i \Big\{\mathbb{E}[|N_{1,i} + N_{A,i}|^2]\Big\} \le b_1, \tag{5.49}$$

$$\lim_{P\to\infty} \max_i \Big\{\mathbb{E}[|N_{2,i} + N_{B,i}|^2]\Big\} \le b_2. \tag{5.50}$$

The i^{th} receiver of $\boldsymbol{T}_A$ treats $N_{A,i} + N_{1,i}$ as the total noise that it observes. Therefore, if (5.49) is satisfied, then the power of the total noise experienced by the i^{th} receiver of $\boldsymbol{T}_A$ is upper bounded by b_1. Similarly, if (5.50) is satisfied, then the power of the total noise experienced by the i^{th} receiver of $\boldsymbol{T}_B$ is upper bounded by b_2. This means that the effect of $N_{A,i}$ and $N_{B,i}$ is equivalent to a bounded increase in the power level of $N_{1,i}$ and $N_{2,i}$, respectively, and therefore, does not decrease the multiplexing gain.

In the following, we show that if $|\boldsymbol{L}[i]| < 1$ for all $i \in \{1, 2, ..M\}$, we can always keep $\max_i\{\mathbb{E}[|N_{A,i} + N_{1,i}|^2]\}$ and $\max_i\{\mathbb{E}[|N_{B,i} + N_{2,i}|^2]\}$ to be small enough such that (5.49) and (5.50) are satisfied. In doing so, note that

$$\begin{aligned}
\mathbb{E}[|N_{A,i} + N_{1,i}|^2] &\le \mathbb{E}[|N_{A,i}|^2 + |N_{1,i}|^2 + 2|N_{1,i}N_{A,i}|] \\
&\overset{(a)}{=} \mathbb{E}[|N_{A,i}|^2] + 1 + \\
&\quad 2\mathbb{E}[|N_{1,i}(\boldsymbol{F}_1\boldsymbol{N}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|] \\
&= \mathbb{E}[|N_{A,i}|^2] + 1 + \\
&\quad 2\mathbb{E}[|(\boldsymbol{F}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]||N_{1,i}|^2] \\
&= \mathbb{E}[|N_{A,i}|^2] + 1 + 2|(\boldsymbol{F}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2.
\end{aligned} \tag{5.51}$$

where (a) is valid because $N_{1,i}$ is unit-variance noise and is independent of $\boldsymbol{S}_1[i]$, $\boldsymbol{S}_2[i]$, and $N_{2,i}$. Therefore, to upper bound $\mathbb{E}[|N_{A,i} + N_{1,i}|^2]$, we find upper bounds on $\mathbb{E}[|N_{A,i}|^2]$

and $2|(\boldsymbol{F}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2$. According to (5.51), we obtain

$$\lim_{P\to\infty}\max_i\Big\{\mathbb{E}[|N_{1,i}+N_{A,i}|^2]\Big\}\leq \\ 1+\lim_{P\to\infty}\max_i\Big\{\mathbb{E}[|N_{A,i}|^2]\Big\}+\lim_{P\to\infty}\max_i\Big\{2|(\boldsymbol{F}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2\Big\}. \tag{5.52}$$

First, we find an upper bound on the power of $N_{A,i}$. Note that the power of $N_{A,i}$ is proportional to $|\boldsymbol{L}[i]|^2$. According to (5.41), since $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, $\boldsymbol{N}_1$, and $\boldsymbol{N}_2$ are independent random variables, we have

$$\mathbb{E}[|N_{A,i}|^2]\leq \\ \Big(P_1^i+|\boldsymbol{F}_1[i]|^2+|\boldsymbol{F}_1[i]\boldsymbol{C}_{21}[i]|^2(P_2^i+|\boldsymbol{F}_2[i]|^2)\Big)\Big(|\boldsymbol{L}[i]|^{2(k+1)}|1-\boldsymbol{L}[i]|^{-2}\Big). \tag{5.53}$$

Note that according to (5.27), $E_1\leq P_1$ and $E_2\leq P_2$. Therefore, in (5.25), we can replace E_1 with P_1 and E_2 with P_2, and we have

$$\sum_{i=1}^{M}C_1^iP_1^i+D_1^iP_2^i\leq P_1, \\ \sum_{i=1}^{M}C_2^iP_1^i+D_2^iP_2^i\leq P_2. \tag{5.54}$$

Moreover, according to (5.27), C_1^i, C_2^i, D_1^i, and D_2^i are all non-negative values, therefore, it follows that

$$P_1^i\leq\frac{P_1}{C_1^i}, P_2^i\leq\frac{P_1}{D_1^i}, \tag{5.55}$$

$$P_1^i\leq\frac{P_2}{C_2^i}, P_2^i\leq\frac{P_2}{D_2^i}. \tag{5.56}$$

Inserting (5.55) into (5.53),

$$\mathbb{E}[|N_{A,i}|^2]\leq \\ \Big(\frac{P_1}{C_1^i}+|\boldsymbol{F}_1[i]|^2+|\boldsymbol{F}_1[i]\boldsymbol{C}_{21}[i]|^2(\frac{P_1}{D_1^i}+|\boldsymbol{F}_2[i]|^2)\Big)\Big(|\boldsymbol{L}[i]|^{2(k+1)}|1-\boldsymbol{L}[i]|^{-2}\Big). \tag{5.57}$$

Moreover, define λ as the maximum magnitude of the loop gains, given by

$$\lambda\doteq\max_{i\in\{1,2,..M\}}|\boldsymbol{L}[i]|. \tag{5.58}$$

Since we have assumed that over all sub-carriers, the product of cross-link channel gains is smaller than that of direct-link channel gains, by (5.45), we deduce that $|\boldsymbol{L}[i]|\leq 1$.

Therefore, we can conclude that $0 < \lambda < 1$. For $P_1 = P_2 = P$, let $n(P)$ be the smallest positive integer such that

$$\lambda^{2(n(P)+1)} \leq \frac{1}{P}. \tag{5.59}$$

Equivalent, $n(P)$ can be defined as

$$n(P) = \left\lfloor \frac{-\ln(P)}{2\ln(\lambda)} \right\rfloor, \tag{5.60}$$

where $\lfloor . \rfloor$ represents the floor function. Note that since $0 < \lambda < 1$, we can always choose a large $n(P)$ that satisfies (5.59). Consequently, for all $i \in \{1, 2, ..M\}$, we have

$$\begin{aligned} |(\boldsymbol{L}[i])|^{2(n(P)+1)}|(1 - \boldsymbol{L}[i])|^{-2} &\overset{(a)}{\leq} \lambda^{2(n(P)+1)}|1 - \boldsymbol{L}[i]|^{-2} \\ &\overset{(b)}{\leq} \frac{|1 - \boldsymbol{L}[i]|^{-2}}{P}, \end{aligned} \tag{5.61}$$

where (a) is valid because of (5.58), and (b) is valid because of (5.59). Inserting (5.61) into (5.57),

$$\begin{aligned} &\mathbb{E}[(N_{A,i})^2] \leq \\ &\Big(\frac{P_1}{C_1^i} + |\boldsymbol{F}_1[i]|^2 + |\boldsymbol{F}_1[i]\boldsymbol{C}_{21}[i]|^2(\frac{P_1}{D_1^i} + |\boldsymbol{F}_2[i]|^2)\Big)\Big(\frac{|1 - \boldsymbol{L}[i]|^{-2}}{P}\Big). \end{aligned} \tag{5.62}$$

Therefore, for $P_1 = P_2 = P$, we can bound the power of the noise $N_{A,i}$, as follows:

$$\begin{aligned} &\lim_{P\to\infty} \max_i \Big\{\mathbb{E}[|N_{A,i}|^2]\Big\} \\ &\overset{(a)}{\leq} \lim_{P\to\infty} \frac{\max_i \Big\{\frac{P}{C_1^i} + |\boldsymbol{F}_1[i]|^2 + |\boldsymbol{F}_1[i]\boldsymbol{C}_{21}[i]|^2(\frac{P}{D_1^i} + |\boldsymbol{F}_2[i]|^2)\Big\}}{\frac{P}{|1-\boldsymbol{L}[i]|^{-2}}} \\ &= \max_i \Big\{\frac{|1 - \boldsymbol{L}[i]|^{-2}}{C_1^i} + |\boldsymbol{F}_1[i]\boldsymbol{C}_{21}[i]|^2(\frac{|1 - \boldsymbol{L}[i]|^{-2}}{D_1^i})\Big\} \\ &\overset{(b)}{=} \max_i \Big\{|1 - \boldsymbol{L}[i]|^{-2}|1 - \boldsymbol{L}[i]|^2 + |1 - \boldsymbol{L}[i]|^{-2}|1 - \boldsymbol{L}[i]|^2\Big\} \\ &= 2, \end{aligned} \tag{5.63}$$

where (a) is valid by (5.62), and (b) is valid by (5.27). Therefore, we have

$$\lim_{P\to\infty} \max_i \Big\{\mathbb{E}[|N_{A,i}|^2]\Big\} \leq 2. \tag{5.64}$$

Similarly, after a time duration of $n(P) \times t_{lp}$, we have

$$\lim_{P\to\infty} \max_i \{\mathbb{E}[|N_{B,i}|^2]\} \leq 2. \tag{5.65}$$

Note that the bounds used in (5.55) and (5.56) are not sharp. As a result, two is not a sharp upper bound on the maximum power of the noise, as expressed in (5.64) and (5.65). However, this upper bound is enough to prove that a multiplexing gain of two is achievable.

Second, we find an upper bound on

$$|(\boldsymbol{F}_1\boldsymbol{L}^{n(P)+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2. \tag{5.66}$$

In particular, we show that

$$\lim_{P\to\infty}\max_i\{|(\boldsymbol{F}_1\boldsymbol{L}^{n(P)+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2\}=0. \tag{5.67}$$

Note that

$$\begin{aligned}
&\lim_{P\to\infty}\max_i\{|\boldsymbol{F}_1[i]|^2|\boldsymbol{L}[i]|^{2(n(P)+1)}|1-\boldsymbol{L}[i]|^{-2}\}\\
&\overset{(a)}{\le}\lim_{P\to\infty}\max_i\{|\boldsymbol{F}_1[i]|^2\frac{|(1-\boldsymbol{L}[i])|^{-2}}{P}\}\\
&=\lim_{P\to\infty}\frac{1}{P}\max_i\{|\boldsymbol{F}_1[i]|^2|1-\boldsymbol{L}[i]|^{-2}\}\\
&=0,
\end{aligned} \tag{5.68}$$

where (a) is valid by (5.61). Inserting (5.68) and (5.63) into (5.52),

$$\begin{aligned}
&\lim_{P\to\infty}\max_i\Big\{\mathbb{E}[|N_{1,i}+N_{A,i}|^2]\Big\}\\
&\quad\le 1+\lim_{P\to\infty}\max_i\Big\{\mathbb{E}[|N_{A,i}|^2]\Big\}+\lim_{P\to\infty}\max_i\Big\{2|(\boldsymbol{F}_1\boldsymbol{L}^{k+1}(\boldsymbol{I}-\boldsymbol{L})^{-1})[i]|^2\Big\}\\
&\quad\le 3.
\end{aligned} \tag{5.69}$$

This shows that (5.49) is satisfied with $b_1=3$. Similarly, one can show that

$$\lim_{P\to\infty}\max_i\Big\{\mathbb{E}[|N_{2,i}+N_{B,i}|^2]\Big\}\le 3, \tag{5.70}$$

which shows that (5.50) is also satisfied with $b_2=3$.

Consequently, a sufficient condition, under which a multiplexing gain of two is still achievable, is to make sure that $n(P)\times t_{lp}$ is smaller than t_{cp}, that is,

$$t_{cp}\ge n(P)\times t_{lp}, \tag{5.71}$$

or equivalently,

$$t_{cp}\ge\left\lfloor\frac{-\ln(P)}{2\ln(\lambda)}\right\rfloor\times t_{lp}. \tag{5.72}$$

Under this condition, the power of $N_{A,i}$ and the power of $N_{B,i}$ become negligible and a multiplexing gain of two is still achievable. This means, as P increases, $n(P)$ will increase, and consequently, t_{cp} should also increase such that (5.72) is satisfied. Therefore, L_{cp} should increase and this increase can reduce the achievable rate. However, as highlighted earlier, one can increase the size of the OFDM symbol M such that the ratio $\frac{M}{M+L_{cp}}$ tends to one. Note that

$$\lim_{P\to\infty} \frac{M}{M+L_{cp}} = \lim_{P\to\infty} \frac{1}{1+\frac{L_{cp}}{M}}. \tag{5.73}$$

Consequently, a multiplexing gain of two is achievable if M grows as P goes to infinity such that

$$\lim_{P\to\infty} \frac{L_{cp}}{M} = 0. \tag{5.74}$$

Note that L_{cp} is proportional to t_{cp}, and according to (5.72), t_{cp} is proportional to $\ln(P)$. Therefore, one can see that for,

$$M = \ln(\ln(P))\ln(P), \tag{5.75}$$

(5.74) is satisfied and a multiplexing gain of two is achievable.

For the converse part of the proof, note that even if transmitters are non-interfering and each transmitter non-causally knows all the messages of the other transmitters, the maximum multiplexing gain of the channel is limited to two, and this completes the proof. □

Remark 5.2. *Note that in the proof of Theorem 5.1, we forced the power of the noise at* $\boldsymbol{T}_A$ *and* $\boldsymbol{T}_B$ *to be bounded, as expressed in (5.49) and (5.50). It is worth mentioning that even if the power of the noise is proportional to* $\ln(P)$*, still the achievable multiplexing gain is two. In fact, one can replace (5.49) and (5.50) with*

$$\lim_{P\to\infty} \frac{\underset{i}{max}\Big\{\mathbb{E}[|N_{1,i}+N_{A,i}|^2]\Big\}}{\ln(P)} \leq b_1, \tag{5.76}$$

$$\lim_{P\to\infty} \frac{\underset{i}{max}\Big\{\mathbb{E}[|N_{2,i}+N_{B,i}|^2]\Big\}}{\ln(P)} \leq b_2. \tag{5.77}$$

This relaxation can lead to a smaller lower bound on the size of the cyclic prefix Lcp.

Remark 5.3. *Comparison with the relay channel: It is interesting that just one time slot delay, assumed by [39], decreases the achievable multiplexing gain from two to one. On the other hand, if transmitters of each group are non-causally provided with all of the other transmitters' messages, the achievable multiplexing gain will not be greater than two. This is reminiscent of the results of [90] in which the lookahead relay is investigated. [90] defines C_0 as the capacity of the relay channel when relay has access to the present signal transmitted by the main transmitter in addition to its past received signal. It is shown than having access to the present transmitted signal, allows C_0 to pass the cut-set bound of the classical relay channel in which relay does not have access to the present signal transmitted by the main transmitter. Note that in the definition of the capacity, the block length, and consequently, the delay go to infinity. Therefore, it might seem that the channel delay has no effect on the capacity of the channel. In fact, the capacity region of the memoryless multiple access channel does not depend on the channel delay [84]. However, [90] shows that for the relay channel, the channel delay can significantly change the capacity. Moreover, [90] defines C^* as the capacity of the relay channel, when relay has non-causal access to future signals of the main transmitter. Clearly, $C_0 \leq C^*$; however, [90] shows that under a condition on channel gains, $C_0 = C^*$, and a simple amplify and forward strategy achieves the capacity.*

Furthermore, we can compare the achievable sum-rate of full-duplex transmitters with that of non-interfering transmitters. If the $\boldsymbol{T}_{A,i}$ and $\boldsymbol{T}_{B,i}$ were non-interfering, then the SNR at the i^{th} sub-carrier of $\boldsymbol{R}_A$ would be $P_1^i|G_{11}|^2$, and the power constraint for $\boldsymbol{T}_A$ is given by $\sum_{i=1}^{M} P_1^i \leq P_1$. This case is investigated as an upper bound on the achievable sum-rate of the full-duplex interfering transmitters. On the other hand, when interfering full-duplex transmitters cancel the interference at their receivers, the SNR at $\boldsymbol{R}_A$ is calculated by (5.36). The power of the signal of the i^{th} sub-carrier that is intended for $\boldsymbol{R}_A$ is amplified by $(A_1^i)^2$, while the power of the noise is amplified by $\boldsymbol{F}_1[i]^2 A_1^i + 1$. Moreover, the power constraint $\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \leq E_1$, implies that a portion of the power of $\boldsymbol{T}_A$ is used to cancel interference, and only a portion of its total power is available to transmit its original messages across sub-carriers. Therefore, although interference is canceled, the SNR and the achievable sum-rate will decrease in comparison with the SNR and the achievable sum-rate of two non-interfering transmitters. However, as $P = P_1 = P_2$ goes to infinity, the effect of this decrease of the available power on the achievable sum-rate becomes insignificant. Consequently, both interfering and non-interfering transmitters achieve the same multiplexing gain of two.

It is worth noting that the proof of Theorem 5.1 shows that the uniform power allocation achieves the maximum multiplexing gain. However, the uniform power allocation does not achieve the maximum sum-rate of the channel. The following sub-section investigates the optimal power allocation that achieves the maximum sum-rate and shows that the optimal power allocation is given by a generalization of the well-known water filling.

5.3.3 Optimal Power Allocation

The model used in the derivation of the power allocation relies on one main assumption: limiting the total power distributed among different sub-carriers, rather than limiting the power allocated to each sub-carrier. This assumption is justified noting that in reusing the spectrum in neighboring areas, the amount of interference is governed by the total amount of transmitted power. Note that such a power allocation strategy does not require a tight coordination among different links, and does not contradict the assumption that links operate autonomously. The reason is that issues such as power allocation, or structure of filters used in interference removal, depend only on factors that vary slowly with time. As a result, it is possible to use some form of central coordination to adjust the relevant system parameters according to a particular realization of such factors. On the other hand, instead of imposing a constraint on total power, one can impose a constraint on the power of each transmitter of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$. In the following, we study both cases.

First, consider the case in which the total power distributed among different sub-carriers is limited. After interference is canceled at all receivers, channels behave similar to two distinct parallel Gaussian point-to-point channels; however, the power constraint at each transmitter depends on the power allocated to other transmitters. Mathematically, to maximize the achievable sum-rate, the following optimization problem is solved:

$$\begin{aligned}
&\max_{P_1^i,P_2^i}\{R_1+R_2\} = \\
&\max_{P_1^i,P_2^i}\left\{\sum_{i=1}^{M}\log(1+\frac{A_1^iP_1^i}{J_1^i})+\sum_{i=1}^{M}\log(1+\frac{A_2^iP_2^i}{J_2^i})\right\} \\
&\text{subject to} \sum_{i=1}^{M} C_j^iP_1^i + D_j^iP_2^i - E_j \le 0, j\in\{1,2\}, \\
&\qquad\qquad -P_j^i \le 0, j\in\{1,2\}, i\in\{1,2,...,M\}.
\end{aligned} \tag{5.78}$$

where A_j^i, C_j^i, and D_j^i, and E_j are constants known by all transmitters and receivers, given in (5.27). As can be seen in (5.78), the achievable rate of $\boldsymbol{T}_A$, i.e., R_1, only depends on how P_1 is distributed over sub-carries; however, the power constraint for $\boldsymbol{T}_A$, i.e., $\sum_{i=1}^{M} C_1^iP_1^i + D_1^iP_2^i \le E_1$, shows that the power allocation used across sub-carries of $\boldsymbol{T}_B$, affects the power allocation over the sub-carries of $\boldsymbol{T}_A$. We denote the optimal power allocation, which maximizes (5.78), by $(\hat{P}_1^i, \hat{P}_2^i)$. Therefore, to find $(\hat{P}_1^i, \hat{P}_2^i)$, the Karush-Kuhn-Tucker (KKT) conditions are written and the result is explained in the following theorem:

Theorem 5.2. *The optimal power allocation $(\hat{P}_1^i, \hat{P}_2^i)$ that maximizes the achievable sum-rate of the parallel two-user GICs with full-duplex transmitters, when transmitters cooperate to cancel interference at their receivers, is given by*

$$\begin{aligned}
\hat{P}_1^i &= \left[\frac{1}{\mu_1C_1^i+\mu_2C_2^i} - \frac{J_1^i}{A_1^i}\right]^+, \\
\hat{P}_2^i &= \left[\frac{1}{\mu_1D_1^i+\mu_2D_2^i} - \frac{J_2^i}{A_2^i}\right]^+,
\end{aligned} \tag{5.79}$$

where μ_1 and μ_2 are KKT multipliers that are determined by

$$\begin{aligned}
\sum_{i=1}^{M} C_1^i\hat{P}_1^i + D_1^i\hat{P}_2^i \le E_1, \\
\sum_{i=1}^{M} C_2^i\hat{P}_1^i + D_2^i\hat{P}_2^i \le E_2.
\end{aligned} \tag{5.80}$$

Moreover, for the symmetric two-user GIC, where $\boldsymbol{C}_{12} = \boldsymbol{C}_{21}$, $\boldsymbol{G}_{12} = \boldsymbol{G}_{21} = \alpha\boldsymbol{I}$, and $P_1 = P_2$, when the available power at transmitters is high enough, i.e.,

$$\frac{E_1}{M} > \max_i\{\frac{J_1^i(C_1^i+D_1^i)}{A_1^i}\} - \sum_{i=1}^{M}\frac{J_1^i(C_1^i+D_1^i)}{MA_1^i}, \tag{5.81}$$

the closed-form expressions for $\hat{P}_1^i = \hat{P}_2^i$ and $\mu_1 = \mu_2$ are given by

$$\begin{aligned}\hat{P}_1^i = \hat{P}_2^i &= \frac{1}{\mu_1(C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i},\\ \mu_1 = \mu_2 &= \frac{M}{E_1 + \sum_{i=1}^{M} \frac{J_1^i}{A_1^i}(C_1^i + D_1^i)}.\end{aligned} \tag{5.82}$$

Proof. The KKT conditions for the optimization problem (5.78) are given by

Stationarity:

$$\begin{aligned}&\nabla(R_1 + R_2) = \mu_1 \nabla\Big(\sum_{i=1}^{M}(C_1^i P_1^i + D_1^i P_2^i)\Big) + \mu_2 \nabla\Big(\sum_{i=1}^{M}(C_2^i P_1^i + D_2^i P_2^i)\Big)\\ &+ \sum_{i=1}^{M} \lambda_1^i \nabla(-P_1^i) + \sum_{i=1}^{M} \lambda_2^i \nabla(-P_2^i).\end{aligned} \tag{5.83}$$

Primal feasibility:

$$\begin{aligned}&\sum_{i=1}^{M}(C_j^i P_1^i + D_j^i P_2^i) - E_j \leq 0, j \in \{1, 2\}.\\ &-P_j^i \leq 0, j \in \{1, 2\}, i \in \{1, 2, ..., M\}.\end{aligned} \tag{5.84}$$

Dual feasibility:

$$\mu_j \geq 0 \text{ and } \lambda_j^i \geq 0, j \in \{1, 2\}, i \in \{1, ..., M\}. \tag{5.85}$$

Complementary slackness:

$$\begin{aligned}&\mu_j\Big(\sum_{i=1}^{M}(C_j^i P_1^i + D_j^i P_2^i) - E_j\Big) = 0, j \in \{1, 2\},\\ &\lambda_j^i P_j^i = 0, j \in \{1, 2\}, i \in \{1, ..., M\}.\end{aligned} \tag{5.86}$$

First, note that KKT conditions are generally necessary conditions for optimality. However, for a maximization problem, if the feasible region is convex and the objective function is concave, then the KKT conditions are sufficient for optimality [45, 46]. The feasible region of the optimization problem (5.78) is a convex region. Moreover, the Hessian matrix of the objective function is given by

$$\begin{aligned}\nabla^2(R_1 + R_2) = \log(e)\mathrm{diag}\Big(&\frac{-(A_1^1)^2}{(J_1^1 + A_1^1 P_1^1)^2}, ..., \frac{-(A_1^M)^2}{(J_1^M + A_1^M P_1^M)^2},\\ &\frac{-(A_2^1)^2}{(J_2^1 + A_2^1 P_2^1)^2}, ..., \frac{-(A_2^M)^2}{(J_2^M + A_2^M P_2^M)^2}\Big).\end{aligned} \tag{5.87}$$

Note that the Hessian matrix is a $2M$ by $2M$ negative semidefinite matrix. This means the objective function is concave. Consequently, the KKT conditions are sufficient conditions.

Simplifying the stationarity condition leads to

$$\begin{aligned}
&\nabla\Big(\sum_{i=1}^{M}\log(1+\frac{A_1^iP_1^i}{J_1^i})\Big)+\nabla\Big(\sum_{i=1}^{M}\log(1+\frac{A_2^iP_2^i}{J_2^i})\Big)\\
&=\mu_1\nabla\Big(\sum_{i=1}^{M}(C_1^iP_1^i+D_1^iP_2^i)\Big)+\mu_2\nabla\Big(\sum_{i=1}^{M}(C_2^iP_1^i+D_2^iP_2^i)\Big)\\
&\quad+\sum_{i=1}^{M}\lambda_1^i\nabla(-P_1^i)+\sum_{i=1}^{M}\lambda_2^i\nabla(-P_2^i).
\end{aligned}\tag{5.88}$$

Calculating the gradient with respect to P_1^i and P_2^i, we have

$$\begin{aligned}
P_1^i+\frac{J_1^i}{A_1^i}&=\frac{1}{\mu_1C_1^i+\mu_2C_2^i-\lambda_1^i}.\\
P_2^i+\frac{J_2^i}{A_2^i}&=\frac{1}{\mu_1D_1^i+\mu_2D_2^i-\lambda_2^i}.
\end{aligned}\tag{5.89}$$

If $\frac{1}{\mu_1C_1^i+\mu_2C_2^i}\geq\frac{J_1^i}{A_1^i}$, let $\lambda_1^i=0$ and $P_1^i=\frac{1}{\mu_1C_1^i+\mu_2C_2^i}-\frac{J_1^i}{A_1^i}$. On the other hand, if $\frac{1}{\mu_1C_1^i+\mu_2C_2^i}<\frac{J_1^i}{A_1^i}$, let $\lambda_1^i=\mu_1C_1^i+\mu_2C_2^i-\frac{A_1^i}{J_1^i}$ and $P_1^i=0$. This choices of P_1^i and λ_1^i is equivalent to

$$P_1^i=\left[\frac{1}{\mu_1C_1^i+\mu_2C_2^i}-\frac{J_1^i}{A_1^i}\right]^+,\tag{5.90}$$

$$\lambda_1^i=\left[\mu_1C_1^i+\mu_2C_2^i-\frac{A_1^i}{J_1^i}\right]^+.\tag{5.91}$$

Similarly, let

$$P_2^i=\left[\frac{1}{\mu_1D_1^i+\mu_2D_2^i}-\frac{J_2^i}{A_2^i}\right]^+,\tag{5.92}$$

$$\lambda_2^i=\left[\mu_1D_1^i+\mu_2D_2^i-\frac{A_2^i}{J_2^i}\right]^+.\tag{5.93}$$

Note that

$$\begin{aligned}
\nabla(R_1+R_2)&=\nabla\Big(\sum_{i=1}^{M}\log(1+\frac{A_1^iP_1^i}{J_1^i})\Big)+\nabla\Big(\sum_{i=1}^{M}\log(1+\frac{A_2^iP_2^i}{J_2^i})\Big)\\
&=\sum_{i=1}^{M}\log(e)\frac{A_1^i}{J_1^i+A_1^iP_1^i}\hat{j}_1^i+\sum_{i=1}^{M}\log(e)\frac{A_2^i}{J_2^i+A_2^iP_2^i}\hat{j}_2^i,
\end{aligned}\tag{5.94}$$

where, $\hat{j}_1^i$ and $\hat{j}_2^i$ represent $2M$ orthonormal vectors. Therefor, $\nabla(R_1+R_2)=0$ has no solution for $P_1^i\geq 0$ and $P_2^i\geq 0$. This means that the optimal solution of (5.78) is achieved over the boundary of the feasible region, in which at least one of the inequalities of (5.80) is satisfied with equality.

If $\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i = E_1$ and $\sum_{i=1}^{M} C_2^i P_1^i + D_2^i P_2^i < E_2$, then by complementary slackness, $\mu_2 = 0$. In addition, $\mu_1 \geq 0$ is determined by

$$\begin{aligned} E_1 &= \sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \\ &= \sum_{i=1}^{M} C_1^i \left[\frac{1}{\mu_1 C_1^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_1^i \left[\frac{1}{\mu_1 D_1^i} - \frac{J_2^i}{A_2^i}\right]^+ \\ &= \sum_{i=1}^{M} \left[\frac{1}{\mu_1} - \frac{C_1^i J_1^i}{A_1^i}\right]^+ + \left[\frac{1}{\mu_1} - \frac{D_1^i J_2^i}{A_2^i}\right]^+ . \end{aligned} \tag{5.95}$$

Similarly, if $\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i < E_1$ and $\sum_{i=1}^{M} C_2^i P_1^i + D_2^i P_2^i = E_2$, then by complementary slackness, $\mu_1 = 0$. In addition, $\mu_2 \geq 0$ is determined by

$$\begin{aligned} E_2 &= \sum_{i=1}^{M} C_2^i P_1^i + D_2^i P_2^i \\ &= \sum_{i=1}^{M} C_2^i \left[\frac{1}{\mu_2 C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_2^i \left[\frac{1}{\mu_2 D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ \\ &= \sum_{i=1}^{M} \left[\frac{1}{\mu_1} - \frac{C_2^i J_1^i}{A_1^i}\right]^+ + \left[\frac{1}{\mu_1} - \frac{D_2^i J_2^i}{A_2^i}\right]^+ . \end{aligned} \tag{5.96}$$

Finally, if $\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i = E_1$ and $\sum_{i=1}^{M} C_2^i P_1^i + D_2^i P_2^i = E_2$, then μ_1 and μ_2 are the non-negative solutions of the following set of equations:

$$\begin{aligned} \sum_{i=1}^{M} C_1^i \left[\frac{1}{\mu_1 C_1^i + \mu_2 C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + \sum_{i=1}^{M} D_1^i \left[\frac{1}{\mu_1 D_1^i + \mu_2 D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ &= E_1, \\ \sum_{i=1}^{M} C_2^i \left[\frac{1}{\mu_1 C_1^i + \mu_2 C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + \sum_{i=1}^{M} D_2^i \left[\frac{1}{\mu_1 D_1^i + \mu_2 D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ &= E_2. \end{aligned} \tag{5.97}$$

One can easily verify that this solution satisfies all the KKT conditions.

For a symmetric two-user GIC, where $\boldsymbol{C}_{12} = \boldsymbol{C}_{21}$, $\boldsymbol{G}_{12} = \boldsymbol{G}_{21} = \alpha \boldsymbol{I}$, and $P_1 = P_2$, it can be verified that $A_1^i = A_2^i$, $C_1^i = D_2^i$, $C_2^i = D_1^i$, $J_1^i = J_2^i$, and $E_1 = E_2$. For a symmetric two-user GIC, (5.90), (5.92), and (5.97) are all symmetric expressions, which imply that

$$\begin{aligned} P_1^i &= P_2^i, \\ \mu_1 &= \mu_2. \end{aligned} \tag{5.98}$$

Therefore, the optimal power allocation is given by

$$\hat{P}_1^i = \hat{P}_2^i = \left[\frac{1}{\mu_1 (C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i}\right]^+ = \left[\frac{1}{\mu_2 (C_2^i + D_2^i)} - \frac{J_2^i}{A_2^i}\right]^+, \tag{5.99}$$

where $\mu_1 = \mu_2$ can be computed from

$$
\begin{aligned}
E_1 =& \sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \\
=& \sum_{i=1}^{M} C_1^i \Big[\frac{1}{\mu_1(C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i}\Big]^+ + D_1^i \Big[\frac{1}{\mu_1(C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i}\Big]^+ \\
=& \sum_{i=1}^{M} \Big[\frac{C_1^i}{\mu_1(C_1^i + D_1^i)} - \frac{C_1^i J_1^i}{A_1^i}\Big]^+ + \Big[\frac{D_1^i}{\mu_1(C_1^i + D_1^i)} - \frac{D_1^i J_1^i}{A_1^i}\Big]^+ \\
=& \sum_{i=1}^{M} \Big[\frac{1}{\mu_1} - \frac{(C_1^i + D_1^i) J_1^i}{A_1^i}\Big]^+ .
\end{aligned}
\tag{5.100}
$$

Note that the last equality is a standard water filling problem in which E_1 is the total amount of water and $\frac{1}{\mu_1}$ represents the level of the water. For this equation, if we have

$$
\frac{E_1}{M} + \sum_{i=1}^{M} \frac{J_1^i (C_1^i + D_1^i)}{M A_1^i} > \max_i \{\frac{J_1^i (C_1^i + D_1^i)}{A_1^i}\}, \tag{5.101}
$$

then $\frac{1}{\mu_1(C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i} \geq 0$ and $\frac{1}{\mu_2(C_2^i + D_2^i)} - \frac{J_2^i}{A_2^i} \geq 0$ for all $i \in \{1, 2, ..., M\}$. Therefore, the optimal power allocation is given by

$$
\hat{P}_1^i = \hat{P}_2^i = \frac{1}{\mu_1(C_1^i + D_1^i)} - \frac{J_1^i}{A_1^i} = \frac{1}{\mu_2(C_2^i + D_2^i)} - \frac{J_2^i}{A_2^i}, \tag{5.102}
$$

where $\mu_1 = \mu_2$ can be computed from (5.100) as follows:

$$
\begin{aligned}
\sum_{i=1}^{M} \Big(\frac{1}{\mu_1} - \frac{(C_1^i + D_1^i) J_1^i}{A_1^i}\Big) =& E_1 \\
\Rightarrow \frac{M}{\mu_1} =& E_1 + \sum_{i=1}^{M} (C_1^i + D_1^i) \frac{J_1^i}{A_1^i} \\
\Rightarrow \mu_1 =& \frac{M}{E_1 + \sum_{i=1}^{M} \frac{J_1^i}{A_1^i} (C_1^i + D_1^i)} .
\end{aligned}
\tag{5.103}
$$

One can see that this solution satisfies all KKT conditions, and due to the sufficiency of KKT conditions, the proof is complete. □

As mentioned earlier, to improve coverage in the uplink, mobile nodes may increase their transmit power without accounting for the total interference caused to the larger network. A second power allocation scheme, discussed next, accounts for such scenarios. In this power allocation, we investigate the case in which a power constraint is imposed on every transmitter of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$. This means, instead of the two power constraints given in (5.23), we impose $2M$ power constraints as follows:

$$\begin{aligned}\mathbb{E}[|X_{1,i}|^2] &\le Q_1^i,\\ \mathbb{E}[|X_{2,i}|^2] &\le Q_2^i,\end{aligned} \tag{5.104}$$

for $i \in \{1, 2, ..., M\}$, where Q_1^i and Q_2^i represent the power constraints on the i^{th} transmitter of $\boldsymbol{T}_A$ and $\boldsymbol{T}_B$, respectively. Similar to (5.25), we can rewrite (5.104) as follows:

$$\begin{aligned}C_1^i P_1^i + D_1^i P_2^i &\le E_1^i,\\ C_2^i P_1^i + D_2^i P_2^i &\le E_2^i,\end{aligned} \tag{5.105}$$

where C_j^i and D_j^i are the same quantities given in (5.27). The only new quantity is E_j^i which is defined by

$$E_j^i = Q_j^i - C_j^i |\boldsymbol{F}_1[i]|^2 - D_j^i |\boldsymbol{F}_2[i]|^2. \tag{5.106}$$

With these new power constraints, the optimization problem dealing with the maximum sum-rate is given by

$$\begin{aligned}&\max_{P_1^i, P_2^i}\{R_1 + R_2\} =\\ &\max_{P_1^i, P_2^i}\left\{\sum_{i=1}^{M} \log(1 + \frac{A_1^i P_1^i}{J_1^i}) + \sum_{i=1}^{M} \log(1 + \frac{A_2^i P_2^i}{J_2^i})\right\}\\ &\text{subject to } C_j^i P_1^i + D_j^i P_2^i - E_j^i \le 0,\\ &\qquad - P_j^i \le 0, j \in \{1, 2\}, i \in \{1, 2, ..., M\}.\end{aligned} \tag{5.107}$$

where A_j^i, J_j^i are given in (5.27).

Note that in the previous optimization problem given in (5.78), the achievable sum-rate of different sub-carriers depend on each other through the two power constraints given in (5.78). However, with the $2M$ power constraints of the optimization problem (5.107), the achievable sum-rate of different sub-carriers become independent of each other. Therefore, the optimization problem (5.107) is equivalent to

$$\begin{aligned}
&\max_{P_1^i, P_2^i}\{R_1 + R_2\} = \\
&\sum_{i=1}^{M}\Big(\max_{P_1^i, P_2^i}\Big\{\log(1+\frac{A_1^i P_1^i}{J_1^i}) + \log(1+\frac{A_2^i P_2^i}{J_2^i})\Big\}\Big) \\
&\text{subject to } C_j^i P_1^i + D_j^i P_2^i - E_j^i \leq 0, \\
&\qquad -P_j^i \leq 0, j \in \{1,2\}, i \in \{1,2,...,M\}.
\end{aligned} \tag{5.108}$$

Theorem 5.3. *The optimal solution of the optimization problem (5.108) is given by*

$$P_1^i = \left[\frac{1}{\mu_1^i C_1^i + \mu_2^i C_2^i} - \frac{J_1^i}{A_1^i}\right]^+, \tag{5.109}$$

$$P_2^i = \left[\frac{1}{\mu_1^i D_1^i + \mu_2^i D_2^i} - \frac{J_2^i}{A_2^i}\right]^+, \tag{5.110}$$

where μ_1^i and μ_2^i are the KKT multipliers determined by the power constraints (5.105).

Proof. Note that according to (5.108), we have a separate optimization problem for each i. For each P_1^i and P_2^i, the feasible region of this optimization problem is a convex region. The feasible region is a polygon with at most four edges as depicted in Figure 5.5. Moreover, the objective function is concave. Therefore, the KKT conditions are sufficient. Define $R_1^i + R_2^i$ as

$$R_1^i + R_2^i \doteq \log\Big(1+\frac{A_1^i P_1^i}{J_1^i}\Big) + \log\Big(1+\frac{A_2^i P_2^i}{J_2^i}\Big). \tag{5.111}$$

Note that we have

$$\begin{aligned}
\nabla(R_1^i + R_2^i) =& \nabla\Big(\log\Big(1+\frac{A_1^i P_1^i}{J_1^i}\Big) + \log\Big(1+\frac{A_2^i P_2^i}{J_2^i}\Big)\Big) \\
=& \log(e)\frac{A_1^i}{J_1^i + A_1^i P_1^i}\hat{j}_1^i + \log(e)\frac{A_2^i}{J_2^i + A_2^i P_2^i}\hat{j}_2^i,
\end{aligned} \tag{5.112}$$

where $\hat{j}_1^i$ and $\hat{j}_2^i$ are two orthonormal vectors corresponding to P_1^i and P_2^i, respectively.

One can see that the equation $\nabla(R_1^i + R_2^i) = (0,0)$ has no solution for $P_1^i \geq 0$ and $P_2^i \geq 0$. Consequently, the optimal solution of the optimization problem (5.108) is attained over the boundary of the feasible region, and therefore, satisfies at least one of the following equalities:

$$\begin{aligned}
C_1^i P_1^i + D_1^i P_2^i = E_1^i, \\
C_2^i P_1^i + D_2^i P_2^i = E_2^i,
\end{aligned} \tag{5.113}$$

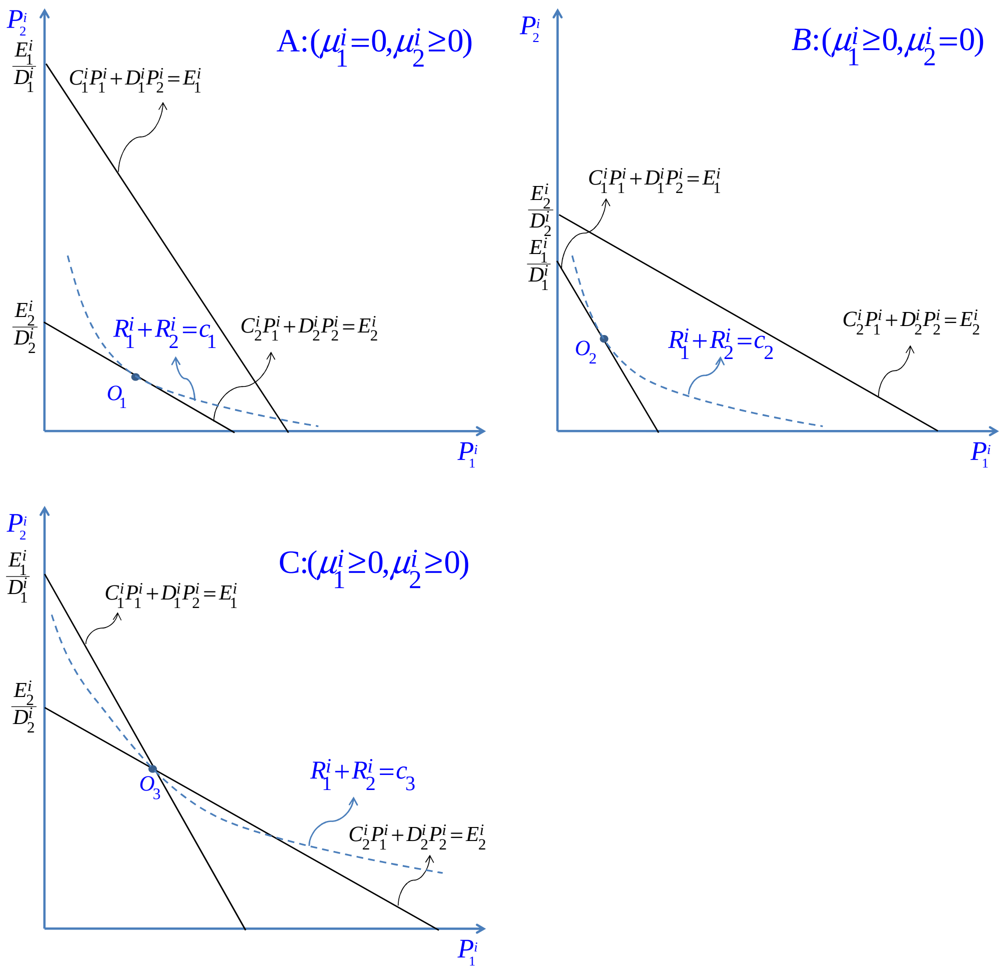

Figure 5.5: The feasible region of the optimization problem (5.108) and the optimal solution on the boundary.

for all $i \in \{1, 2, ..., M\}$. One can write the KKT conditions for this new problem. Note that since we have $2M$ power constraints, we have $2M$ corresponding KKT multipliers given by μ_j^i, where $j \in \{1, 2\}$ and $i \in \{1, 2, ..., M\}$. Similar to the previous optimization problem, one can write the stationarity condition and show that the optimal power allocation satisfies

$$\begin{aligned} P_1^i &= \left[\frac{1}{\mu_1^i C_1^i + \mu_2^i C_2^i} - \frac{J_1^i}{A_1^i}\right]^+, \\ P_2^i &= \left[\frac{1}{\mu_1^i D_1^i + \mu_2^i D_2^i} - \frac{J_2^i}{A_2^i}\right]^+. \end{aligned} \tag{5.114}$$

If $C_1^i P_1^i + D_1^i P_2^i = E_1^i$ and $C_2^i P_1^i + D_2^i P_2^i < E_2^i$, by complementary slackness, $\mu_2^i = 0$. Moreover, $\mu_1^i > 0$ is determined by

$$\begin{aligned} E_1^i =& C_1^i P_1^i + D_1^i P_2^i \\ =& C_1^i \left[\frac{1}{\mu_1^i C_1^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_1^i \left[\frac{1}{\mu_1^i D_1^i} - \frac{J_2^i}{A_2^i}\right]^+ \\ =& \left[\frac{1}{\mu_1^i} - \frac{C_1^i J_1^i}{A_1^i}\right]^+ + \left[\frac{1}{\mu_1^i} - \frac{D_1^i J_2^i}{A_2^i}\right]^+. \end{aligned} \tag{5.115}$$

Similarly, if $C_1^i P_1^i + D_1^i P_2^i < E_1^i$ and $C_2^i P_1^i + D_2^i P_2^i = E_2^i$, by complementary slackness, $\mu_1^i = 0$. Moreover, $\mu_2^i \geq 0$ is determined by

$$\begin{aligned} E_2^i =& C_2^i P_1^i + D_2^i P_2^i \\ =& C_2^i \left[\frac{1}{\mu_2^i C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_2^i \left[\frac{1}{\mu_2^i D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ \\ =& \left[\frac{1}{\mu_2^i} - \frac{C_2^i J_1^i}{A_1^i}\right]^+ + \left[\frac{1}{\mu_2^i} - \frac{D_2^i J_2^i}{A_2^i}\right]^+. \end{aligned} \tag{5.116}$$

Finally, if $C_1^i P_1^i + D_1^i P_2^i = E_1^i$ and $C_2^i P_1^i + D_2^i P_2^i = E_2^i$, then μ_1 and μ_2 are the non-negative solutions of the following set of equations:

$$\begin{aligned} C_1^i \left[\frac{1}{\mu_1^i C_1^i + \mu_2^i C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_1^i \left[\frac{1}{\mu_1^i D_1^i + \mu_2^i D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ = E_1^i, \\ C_2^i \left[\frac{1}{\mu_1^i C_1^i + \mu_2^i C_2^i} - \frac{J_1^i}{A_1^i}\right]^+ + D_2^i \left[\frac{1}{\mu_1^i D_1^i + \mu_2^i D_2^i} - \frac{J_2^i}{A_2^i}\right]^+ = E_2^i. \end{aligned} \tag{5.117}$$

In Figure 5.5, these three cases are shown. Figure 5.5A shows the case in which $\mu_1^i = 0$ and $\mu_2^i \geq 0$. The optimal power allocation is demonstrated by the point O_1, in which the contour curve $R_1^i + R_2^i = c_1$ tangentially touches the line $C_1^i P_1^i + D_1^i P_2^i = E_1^i$.

Figure 5.5B shows the second case in which $\mu_1^i \geq 0$ and $\mu_2^i = 0$. The optimal power allocation is demonstrated by the point O_2, in which the contour curve $R_1^i + R_2^i = c_2$ tangentially touches the line $C_2^i P_1^i + D_2^i P_2^i = E_2^i$. Finally, Figure 5.5$C$ shows the third case in which $\mu_1^i \geq 0$ and $\mu_2^i \geq 0$. The optimal power allocation is demonstrated by the point O_3, in which the contour curve $R_1^i + R_2^i = c_3$ passes through the intersection of $C_1^i P_1^i + D_1^i P_2^i = E_1^i$ and $C_2^i P_1^i + D_2^i P_2^i = E_2^i$. This competes the proof $\square$

We can further investigate the solution of the optimization problem (5.108), and find the optimal power allocation explicitly such that the KKT multipliers are eliminated. This solution can reveal the conditions under which exactly one of the possible three cases depicted in Figure 5.5 determines the optimal solution. In doing so, we solve equations (5.115), (5.116), and (5.117).

To solve equation (5.115), define $m_1^i \doteq \frac{C_1^i J_1^i}{A_1^i}$ and $n_1^i \doteq \frac{D_1^i J_2^i}{A_2^i}$. Then (5.115) is equivalent to

$$\left[\frac{1}{\mu_1^i} - m_1^i\right]^+ + \left[\frac{1}{\mu_1^i} - n_1^i\right]^+ = E_1^i. \tag{5.118}$$

Note that (5.118) is a standard water filling equation, and therefore, μ_1^i is given by

$$\mu_1^i = \begin{cases} \frac{1}{E_1^i + \min\{m_1^i, n_1^i\}} & \text{if } E_1^i \leq |m_1^i - n_1^i|, \\ \frac{2}{E_1^i + m_1^i + n_1^i} & \text{otherwise.} \end{cases} \tag{5.119}$$

Therefore, inserting (5.119) and $\mu_2^i = 0$ into (5.114), the optimal power allocation $(\hat{P}_1^i, \hat{P}_2^i)$ is given by

$$\hat{P}_1^i = \begin{cases} \frac{E_1^i}{C_1^i} \mathbb{1}(m_1^i \leq n_1^i) & \text{if } E_1^i \leq |m_1^i - n_1^i|, \\ \frac{E_1^i - m_1^i + n_1^i}{2C_1^i} & \text{otherwise.} \end{cases} \tag{5.120}$$

$$\hat{P}_2^i = \begin{cases} \frac{E_1^i}{D_1^i} \mathbb{1}(n_1^i \leq m_1^i) & \text{if } E_1^i \leq |m_1^i - n_1^i| \\ \frac{E_1^i + m_1^i - n_1^i}{2D_1^i} & \text{otherwise.} \end{cases} \tag{5.121}$$

Note that (5.120) and (5.121) represent the optimal solution of the optimization problem (5.108), if and only if $C_2^i P_1^i + D_2^i P_2^i < E_2^i$, that is, for $E_1^i \leq |m_1^i - n_1^i|$, we should have

$$C_2^i \frac{E_1^i}{C_1^i} \mathbb{1}(m_1^i \leq n_1^i) + D_2^i \frac{E_1^i}{D_1^i} \mathbb{1}(n_1^i \leq m_1^i) < E_2^i, \tag{5.122}$$

and for $E_1^i > |m_1^i - n_1^i|$, we should have

$$\begin{aligned} E_2^i >& C_2^i \frac{E_1^i - m_1^i + n_1^i}{2C_1^i} + D_2^i \frac{E_1^i + m_1^i - n_1^i}{2D_1^i} \\ =& \Big(\frac{C_2^i}{2C_1^i} + \frac{D_2^i}{2D_1^i}\Big) E_1^i + \frac{(n_1^i - m_1^i)(D_1^i - C_1^i)}{2C_1^i D_1^i}. \end{aligned} \tag{5.123}$$

In the first quadrant of the $E_1^i E_2^i$-plane, (5.122) and (5.123) specify a region. Label the region characterized by (5.122) and (5.123) as $\mathcal{R}_1$. Note that $\mathcal{R}_1$ represents the region inside which the optimal power allocation is given by (5.120) and (5.121).

Similarly, one can solve (5.116) and find the optimal power allocation. Define $m_2^i \doteq \frac{C_2^i J_1^i}{A_1^i}$ and $n_2^i \doteq \frac{D_2^i J_2^i}{A_2^i}$. By solving (5.116) and inserting into (5.114), the optimal power allocation $(\hat{P}_1^i, \hat{P}_2^i)$ is given by

$$\hat{P}_1^i = \begin{cases} \frac{E_2^i}{C_2^i} \mathbb{1}(m_2^i \leq n_2^i) & \text{if } E_2^i \leq |m_2^i - n_2^i|, \\ \frac{E_2^i - m_2^i + n_2^i}{2C_2^i} & \text{otherwise.} \end{cases} \tag{5.124}$$

$$\hat{P}_2^i = \begin{cases} \frac{E_2^i}{D_2^i} \mathbb{1}(n_2^i \leq m_2^i) & \text{if } E_2^i \leq |m_2^i - n_2^i|, \\ \frac{E_2^i + m_2^i - n_2^i}{2D_2^i} & \text{otherwise.} \end{cases} \tag{5.125}$$

Note that (5.120) and (5.121) demonstrate the optimal solution of the optimization problem (5.108), if and only if $C_1^i P_1^i + D_1^i P_2^i < E_1^i$, that is, for $E_2^i \leq |m_2^i - n_2^i|$, we should have

$$C_1^i \frac{E_2^i}{C_2^i} \mathbb{1}(m_2^i \leq n_2^i) + D_1^i \frac{E_2^i}{D_2^i} \mathbb{1}(n_2^i \leq m_2^i) < E_1^i, \tag{5.126}$$

and for $E_2^i > |m_2^i - n_2^i|$, we should have

$$\begin{aligned} E_1^i >& C_1^i \frac{E_2^i - m_2^i + n_2^i}{2C_2^i} + D_1^i \frac{E_2^i + m_2^i - n_2^i}{2D_2^i} \\ =& \Big(\frac{C_1^i}{2C_2^i} + \frac{D_1^i}{2D_2^i}\Big) E_2^i + \frac{(n_2^i - m_2^i)(D_2^i - C_2^i)}{2C_2^i D_2^i}. \end{aligned} \tag{5.127}$$

Let us label the region characterized by (5.126) and (5.127) as $\mathcal{R}_2$. In fact, $\mathcal{R}_2$ represents the region inside which the optimal power allocation is given by (5.124) and (5.124).

Finally, we solve equation (5.117). Note that (5.117) is not a standard water filling equation and finding μ_1^i and μ_2^i from (5.117) can be complicated. However, as depicted in Figure 5.5, there exist exactly three cases for (μ_1^i, μ_2^i). We have already shown that if (5.122) and (5.123) are satisfied, then $(\mu_1^i = 0, \mu_2^i \geq 0)$. Similarly, if (5.126) and (5.127)

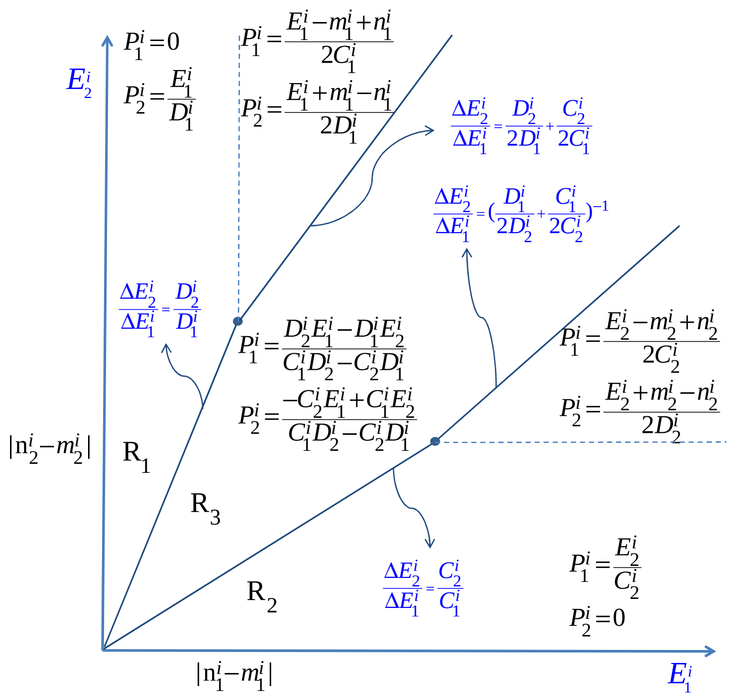

Figure 5.6: The optimal power allocation of the optimization problem (5.108), when $n_1^i \le m_1^i$ and $n_2^i \ge m_2^i$.

are satisfied, then $(\mu_1^i \le 0, \mu_2^i = 0)$. Therefore, for all other cases, the optimal power allocation satisfies $(\mu_1^i \ge 0, \mu_2^i \ge 0)$. In other words, the first quadrant of $E_1^i E_2^i$-plane can be partitioned into three regions such that inside each region, the closed form expression of the optimal power allocation is explicitly given, as shown in Figure 5.6.

We have already investigated two regions inside the first quadrant of $E_1^i E_2^i$-plane, namely $\mathcal{R}_1$ and $\mathcal{R}_2$. For the remaining region, that we label as $\mathcal{R}_3$, the optimal power allocation is the solution of (5.117). In fact, $\mathcal{R}_3$ is characterized by the following expressions: for $E_2^i \le |m_2^i - n_2^i|$,

$$\begin{aligned} C_2^i \frac{E_1^i}{C_1^i} \mathbb{1}(m_1^i \le n_1^i) + D_2^i \frac{E_1^i}{D_1^i} \mathbb{1}(n_1^i \le m_1^i) \ge E_2^i, \\ C_1^i \frac{E_2^i}{C_2^i} \mathbb{1}(m_2^i \le n_2^i) + D_1^i \frac{E_2^i}{D_2^i} \mathbb{1}(n_2^i \le m_2^i) \ge E_1^i, \end{aligned} \tag{5.128}$$

and for $E_1^i > |m_1^i - n_1^i|$,

$$\begin{aligned} \left(\frac{C_2^i}{2C_1^i} + \frac{D_2^i}{2D_1^i}\right) E_1^i + \frac{(n_1^i - m_1^i)(D_1^i - C_1^i)}{2C_1^i D_1^i} \ge E_2^i, \\ \left(\frac{C_1^i}{2C_2^i} + \frac{D_1^i}{2D_2^i}\right) E_2^i + \frac{(n_2^i - m_2^i)(D_2^i - C_2^i)}{2C_2^i D_2^i} \ge E_1^i. \end{aligned} \tag{5.129}$$

To find the solution of (5.117), instead of finding μ_1^i and μ_2^i, we directly find P_1^i and

P_2^i. Note that (5.117) is equivalent to

$$C_1^i P_1^i + D_1^i P_2^i = E_1^i,$$
$$C_2^i P_1^i + D_2^i P_2^i = E_2^i. \tag{5.130}$$

Therefore, we can directly find the optimal power allocation $(\hat{P}_1^i, \hat{P}_2^i)$ as follows:

$$\hat{P}_1^i = \frac{D_2^i E_1^i - D_1^i E_2^i}{C_1^i D_2^i - C_2^i D_1^i},$$
$$\hat{P}_2^i = \frac{-C_2^i E_1^i + C_1^i E_2^i}{C_1^i D_2^i - C_2^i D_1^i}. \tag{5.131}$$

Note that (5.131) represents the optimal power allocation, if and only if $(E_1^i, E_2^i) \in \mathcal{R}_3$, as depicted in Figure 5.6. Moreover, one can easily check that inside the region $\mathcal{R}_3$, expressions of (5.131) assign positive values to $\hat{P}_1^i$ and $\hat{P}_2^i$.

In order to demonstrate how $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ partition the first quadrant of $E_1^i E_2^i$-plane, we need to know whether m_1^i is smaller than n_1^i or not. Similarly, we need to know whether m_1^i is smaller than n_1^i or not. The case in which $n_1^i \leq m_1^i$ and $n_2^i > m_2^i$ is depicted in Figure 5.6. In this figure, the first quadrant of the $E_1^i E_2^i$-plane is partitioned into three regions, namely $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$. For each region, the optimal power allocation is explicitly given. In $\mathcal{R}_1$, the optimal power allocation is only a function of E_1^i and is independent of E_2^i. This can be justified by noting that in $\mathcal{R}_1$, the value of E_2^i is large enough such that the power constraint $C_2^i P_1^i + D_2^i P_2^i \leq E_2^i$ is inactive, as shown in Figure 5.5A. Similarly, in $\mathcal{R}_3$, the value of E_1^i is large enough such that the power constraint $C_1^i P_1^i + D_1^i P_2^i \leq E_1^i$ is inactive, and therefore, the optimal power allocation is independent of E_1^i. In $\mathcal{R}_3$, both power constraints are satisfied with equality, as shown in Figure 5.5C, and the optimal power allocation is a function of both E_1^i and E_2^i.

One interesting observation about Figure 5.6 is to note that the regions $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ demonstrate a valid partitioning of the first quadrant of $E_1^i E_2^i$-plane. To do so, we should make sure that for large values of E_1^i and E_2^i, the lines that determine the boundaries of $\mathcal{R}_1$ and $\mathcal{R}_2$ do not intersect. As can be seen in Figure 5.6, the boundary of $\mathcal{R}_1$ is a line that has a slope given by

$$\frac{\Delta E_2^i}{\Delta E_1^i} = \frac{D_2^i}{2D_1^i} + \frac{C_2^i}{2C_1^i}. \tag{5.132}$$

On the other hand, the boundary of $\mathcal{R}_2$ is a line that has a slope given by

$$\frac{\Delta E_2^i}{\Delta E_1^i} = \left(\frac{D_1^i}{2D_2^i} + \frac{C_1^i}{2C_2^i}\right)^{-1}. \tag{5.133}$$

To make sure that $\mathcal{R}_1$ and $\mathcal{R}_2$ have no intersection, we should make sure that (5.132) is greater than or equal to (5.133). Note that (5.132) represents the arithmetic mean of $\frac{D_2^i}{D_1^i}$ and $\frac{C_2^i}{C_1^i}$. However, (5.133) represents the harmonic mean of $\frac{D_2^i}{D_1^i}$ and $\frac{C_2^i}{C_1^i}$, and by the power mean inequality, we know that the arithmetic mean is always grater than or equal to the harmonic mean.

Next, to conclude this section, we compare the optimal power allocation (5.79) with the uniform power allocation (5.29). For the symmetric two-user GIC, in which $E_1 = E_2 = E$, (5.82) shows that the optimal power allocation is given by

$$P_1^i = P_2^i = \frac{E}{M(C_1^i + D_1^i)} + \frac{1}{M}\sum_{k=1}^{M} \frac{J_1^k}{A_1^k}\frac{C_1^k + D_1^k}{C_1^i + D_1^i} - \frac{J_1^i}{A_1^i}. \tag{5.134}$$

Moreover, (5.29) shows that for the symmetric two-user GIC, the uniform power allocation is given by

$$P_1^i = P_2^i = \breve{P}_1 = \frac{E}{\sum_{k=1}^{M}(C_1^k + D_1^k)}. \tag{5.135}$$

Comparing (5.134) and (5.135), we see that the optimal power allocated to the i^{th} channel will increase, if $C_1^i + D_1^i$ decreases. On the other hand, if all parallel GICs are identical such that $C_1^i + D_1^i$ and $\frac{J_1^i}{A_1^i}$ are independent of i, the uniform power allocation and the optimal power allocation will be the same. However, when parallel GICs are different, the optimal power allocation achieves a higher sum-rate. In the next section, we demonstrate some simulation results to compare the achievable sum-rate of the optimal power allocation and that of the uniform power allocation.

5.4 Simulation Results

The considered system model is simulated based on an OFDM system with $M = 512$ sub-carriers and a cyclic prefix of size $L_{cp} = 16$. Figure 5.7 considers parallel symmetric two-user GICs, in which $\boldsymbol{C}_{12} = \boldsymbol{C}_{21}$, $\boldsymbol{G}_{12} = \boldsymbol{G}_{21}$, $\boldsymbol{G}_{11} = \boldsymbol{G}_{22}$. Moreover, $P_1 = P_2 = M \times 10^3$. Since noise power is normalized to one, this power value corresponds to an average power of 30db per sub-carrier, which is typical in wireless systems supporting modulations with high spectral efficiency. For $1 \leq i \leq M$, $|\boldsymbol{G}_{11}[i]|$ and $|\boldsymbol{G}_{12}[i]|$ are distributed according to a Rayleigh distribution with means of $\frac{\sqrt{\pi}}{2}$ and $\alpha\frac{\sqrt{\pi}}{2}$, respectively, where $0 \leq \alpha \leq 1$

represents the cross-link channel gain. Furthermore, it is assumed that all channel gains are fully known at all transmitters and all receivers.

In the simulation, 10000 symbols of OFDM are generated, and for the GIC formed over each symbol, independent channel gains are realized according to the Rayleigh distribution. Note that the independence assumption is justified by the fact that different transmitters operate in different physical locations, resulting in physically separate links for different transmitter/receiver pairs. Then, the average achievable sum-rate per complex sub-carrier is calculated for four different scenarios. We have considered two different power allocations: the uniform power allocation and the optimal power allocation given in (5.79). Both power allocations satisfy the power constraint (5.25). OFDM symbols are transmitted through the channel as depicted in Figure 5.3. Furthermore, the additive white Gaussian noise with zero mean and unit variance is added to the received signals of every receiver in the system.

Figure 5.7 compares the achievable sum-rate of four different cases when the cross-link channel gain α goes from zero to one. Note that to satisfy $|\boldsymbol{L}[i]| < 1$ for all $i \in \{1, 2, ..M\}$, α should be smaller than one. The red line shows the case where transmitters are not full-duplex and interference is treated as noise. As can be seen, when the power of the interference increases as α goes to 1, the achievable sum-rate decreases significantly. Both the black line and the blue line show the case in which full-duplex transmitters are used to cancel the interference at their corresponding receivers. The black line shows the sum-rate when power is allocated optimally according to (5.79), whereas in the blue line, power is allocated uniformly. When power is allocated uniformly, i.e., $P_1^i = P_2^i = \frac{E_1}{\sum_{i=1}^{M}(C_1^i + D_1^i)}$, the achievable sum-rate is less than that of the optimal power allocation but still considerably more than the case in which transmitters are not full-duplex and interference is not canceled.

The green line, which shows the case in which transmitters do not interfere with each other, is considered as an upper bound. In this case, and with the specified values of P_1 and P_2, each complex sub-carrier can achieve around 9 bits per transmission. This achievable rate is motivated by the new trend for using higher order modulation such as 512-QAM and above. The black line represents the sum-rate of full-duplex transmitters with optimal power allocation as described in (5.79). As seen in Figure 5.7, the achievable sum-rate of full-duplex transmitters is strictly less than that of non-interfering transmit-

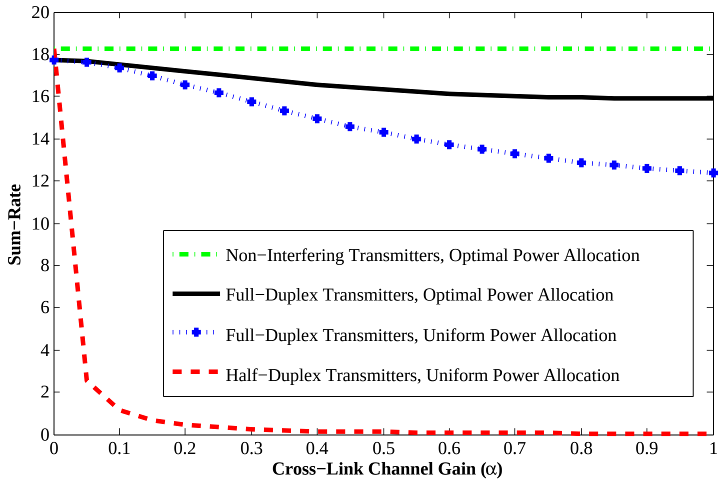

Figure 5.7: The average achievable sum-rate (per complex sub-carrier) of the symmetric two-user GIC for four different coding schemes, with $M = 512$ and $P_1 = P_2 = M \times 10^3$.

ters. In fact, although cooperative transmitters can completely cancel the interference, this cancellation is achieved at a price. To combat multi-path fading, a cyclic prefix of size $L_{cp} = 16$ is used. Consequently, the effective achievable rate is reduced by a factor of $\frac{512}{512+16}$. Moreover, a portion of the power of each transmitter is used to cancel the interference and for each transmitter less power is available to transmit its original message. Therefore, although full-duplex transmitters can completely cancel the interference, this cancellation reduces the available power to transmit the original message.

To clarify this power loss, Figure 5.8 shows the percentage of the power that is used to transmit the messages of each group. According to (5.6), $\boldsymbol{S}_1(\boldsymbol{I}-\boldsymbol{L})^{-1}$ is the signal that conveys the original messages of the transmitters of $\boldsymbol{T}_A$ and $\boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2(\boldsymbol{I}-\boldsymbol{L})^{-1}$ is the signal used to cancel the interference at $\boldsymbol{R}_A$. Define P_{S_1} as the power of $\boldsymbol{S}_1(\boldsymbol{I}-\boldsymbol{L})^{-1}$, i.e., $P_{S_1} = \sum_{i=1}^{M} C_1^i P_1^i$, where C_1^i is defined in (5.27). The ratio $\frac{P_{S_1}}{P_1}$ represents the percentage of P_1 that is used to transmit the original M messages of $\boldsymbol{T}_A$, conveyed by $\boldsymbol{S}_1 = [S_{1,1}, S_{1,2}, ...S_{1,M}]^T$. Figure 5.8 shows that, as the cross-link channel gain α goes to one, more power is required to cancel the interference. In fact, when $\alpha = 1$, the power of the interference is maximized, and consequently, more power is required to transmit $\boldsymbol{F}_1\boldsymbol{C}_{21}\boldsymbol{S}_2(\boldsymbol{I}-\boldsymbol{L})^{-1}$ such that the interference is canceled at $\boldsymbol{R}_A$. Figure 5.8 shows that, for most values of α, at least half of the total power is used by $\boldsymbol{T}_A$ to transmit the original messages. Therefore, the maximum sum-rate loss for most values of α, due to the power loss, is limited to $2\log_2(2) = 2$. This is clearly seen in Figure 5.7, as the

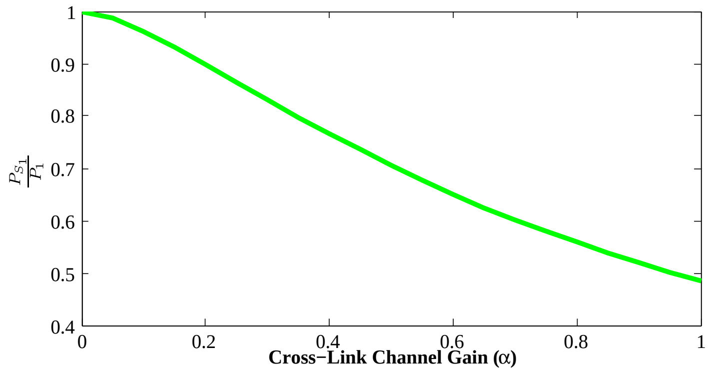

Figure 5.8: The power available for $\boldsymbol{T}_A$ to transmit its own message $\boldsymbol{S}_1$, when optimal power allocation is used.

achievable sum-rate of the full-duplex transmitters with optimal power allocation drops from approximately 17.7 to 15.8 bits per transmission.

Remark 5.4. *The achievable sum-rate does not significantly depend on the cross-link channel gain α: As can be seen in Figure 5.7, the achievable rate of the two-user GIC with half-duplex transmitters decreases significantly as the cross-link channel gain increases. This rate loss is expected, since as α increases, the power of the interference increases, and consequently, the* SNR *at the receivers decreases. However, the achievable sum-rate of full-duplex transmitters does not change significantly. In fact, as α increases, the power of the interference received by $\boldsymbol{R}_A$ increases. Therefore, more power is required to cancel the interference and less power remains available at each transmitter to transmit its own messages. For instance, consider the uniform power allocation, i.e., $P_1^i = P_2^i = \frac{E_1}{\sum_{i=1}^{M}(C_1^i+D_1^i)}$. If $\alpha \to 1$, both C_1^i and D_1^i will increase according to (5.27), and therefore, P_1^i and P_2^i will decrease. Similarly, when the power is allocated optimally according to (5.79), as α increases, P_1^i and P_2^i decrease. This is clearly depicted in Figure 5.8. The power constraint $\sum_{i=1}^{M} C_1^i P_1^i + D_1^i P_2^i \leq E_1$ implies that $\sum_{i=1}^{M} D_1^i P_2^i$ is a portion of P_1 that is used to cancel the interference at $\boldsymbol{R}_A$, and $P_{S_1} = \sum_{i=1}^{M} C_1^i P_1^i$ is a portion of P_1 that is used to transmit the original message $\boldsymbol{S}_1$. Figure 5.8 shows that when optimal power allocation is used, as the cross-link channel gain increases, less power remains available for the transmission of $\boldsymbol{S}_1$.*

Interestingly, the overall SNR*, i.e., $\frac{P_1^i A_1^i}{J_1^i}$, does not vary significantly, and since* $\log($SNR$)$ *determines the achievable sum-rate, a small change in* SNR *does not lead to a major change in the achievable sum-rate. Thus, full-duplex transmitters can guarantee an almost* *constant rate for different fading gains. Moreover, the achievable sum-rate of full-duplex transmitters is shown to be close to that of non-interfering transmitters.*

5.5 Conclusion

In this chapter, a new perspective was introduced that captures the role of the delay in cooperative communications more accurately. Relying on this perspective, the role of cooperation in increasing the achievable sum-rate of the two-user GIC was investigated. We showed that, in the context of OFDM systems, the traditional constraint of causal delay can be slightly modified. Then, we showed that when full-duplex transmitters causally cooperate with each other to cancel the interference, a multiplexing gain of two is achievable. Moreover, we computed the optimal power allocation that maximizes the achievable sum-rate when interference has been canceled. Simulation results were included to highlight the role of interference cancellation in improving the achievable sum-rate and the impact of interference cancellation on the optimal power allocation. The new perspective introduced in this study can shed light on the role of delay in a wide range of scenarios related to cooperative communications or multi-hop networks.

Chapter 6

Conclusion and Future Research Directions

6.1 Conclusion

In our attempt to offer a better understanding of the capacity region of the two-user GIC, we investigated three important aspects of this channel, as briefly explained in what follows.

In Chapters 2 and 3, we characterized the boundary of the HK region. In doing so, we first derived an optimization problem that corresponds to the maximum sum-rate achieved by the HK scheme with Gaussian input and no time sharing. The general optimization problem that demonstrates the maximum HK sum-rate is complicated. Our first contribution is a simpler characterization of this optimization problem for the weak interference class. However, even the simplified optimization problem is still difficult to solve and involves a non-differentiable objective function. We have thus used an optimization technique to overcome this difficulty. In fact, by partitioning the feasible region, we were able to solve the optimization problem. Consequently, we explicitly derived the optimal power allocation that maximizes the HK sum-rate. For the weak interference class, we showed that, depending on transmitters' powers, different power allocation policies maximize the HK sum-rate. This situation is in contrast to the strong and mixed classes, where a unique power allocation policy maximizes the sum-rate.

Chapter 3 extended the results of Chapter 2 and characterized the optimal power allocation policy that maximizes an arbitrary weighted HK sum-rate. Moreover, we described the role of time sharing in increasing the HK sum-rate. For strong and mixed classes, the time sharing variable Q does not increase the maximum HK sum-rate. However, for the weak interference class, we showed that time sharing can strictly increase the achievable sum-rate. We proved that the role of time sharing in increasing the sum-rate can be expressed in terms of calculating the upper concave envelope of a function of P_1 and P_2.

In Chapter 4, we discussed the complexity of sum-rate optimal codes. Most coding schemes proposed for the two-user GIC employ joint decoding to increase the achievable sum-rate. However, joint decoding significantly increases decoding complexity. In Chapter 4, we showed that joint decoding can be replaced by rate splitting and successive decoding. In doing so, we first characterized an optimization problem that corresponds to the maximum sum-rate achieved by rate splitting and successive decoding. We highlighted that the optimization problem is complicated and involves a non-convex optimization. We thus used an optimization technique to find a feasible solution for the optimization problem. Then an optimality certificate was used to investigate the optimality of the solution. Our main contribution is the closed-form expressions for the optimal power allocation, optimal number of splits, and optimal decoding order. We showed that the sum-rate loss, caused by replacing joint decoding with successive decoding, is bounded and remains constant as transmitters' powers approach infinity.

In Chapter 5, we discussed the role of causal cooperation among transmitters in enlarging the achievable region. In cooperative communications, a delay constraint is used to guarantee causality. Traditionally, delay granularity has been limited to one symbol; however, channel delay is in fact governed by channel memory and can be shorter. With this perspective, we introduced a new constraint to guarantee that cooperation is causal. In chapter 5, our main contribution is a more-accurate analysis of delay in cooperative communications. We showed that the new constraint allows the coding scheme proposed for the two-user GIC to increase the multiplexing gain.

6.2 Future Research Directions

This dissertation gives rise to several interesting research questions, as will be briefly discussed below.

In Chapter 2, we focused on the two-user GIC. One possible research direction is to characterize the boundary of the HK scheme for the K-user GIC. Note that general understanding of the achievable region of the K-user GIC is limited. Most of the results on the K-user GIC correspond only to interference alignment and the achievable multiplexing gain. Therefore, obtaining solid understanding of the achievable region is of paramount importance.

Another research direction is to develop optimization techniques that can address the maximum HK sum-rate. In this thesis, we used the partitioning idea to solve the optimization problem. Another useful optimization technique is the min-max idea. One can replace the non-differentiable objective function with a new function that involves minimization over new variables. Then, by replacing the order of maximization and minimization, one might be able to solve the optimization problem. This idea has been used to investigate the boundary of the Marton's rate region [53]. It would be interesting to see whether a similar approach can characterize the boundary of the HK rate region.

In Chapter 4, we focused on the maximum sum-rate through rate splitting and successive decoding. It would be worthwhile to use this idea and characterize the entire boundary of the achievable region. In fact, characterizing the maximum of an arbitrary weighted sum-rate involves an optimization problem, which is slightly more complicated. In Chapter 4, the symmetry of the sum-rate results in simplified closed-form expressions. However, by characterizing the maximum of an arbitrary weighted sum-rate, one can demonstrate how power should be allocated to achieve a boundary point of the achievable region. Another interesting direction would be to generalize the results of Chapter 4 to the K-user GIC. This generalization could shed light on the characterization of the HK scheme for the K-user GIC.

In Chapter 5, we introduced a new delay constraint that guarantees causality. We showed that this new constraint allows the coding scheme proposed for the two-user GIC to achieve a higher multiplexing gain. This new constraint can be used to analyze the multiplexing gain of multi-hop networks. Therefore, the application of this idea to other multi-hop networks would be a useful future research area.

Bibliography

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob*, vol. 1, 1961, pp. 611–644.

[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *The Annals of Probability*, pp. 805–814, 1974.

[3] H. Sato, "Two-user communication channels," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 295–304, May 1977.

[4] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 228–231, 1979.

[5] A. Carleial, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 60–70, 1978.

[6] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, Jan 1981.

[7] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, 1981.

[8] A. Carleial, "A case where interference does not reduce capacity," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, Sep 1975.

[9] M. H. Costa and A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference." *IEEE Transactions on Information Theory*, vol. 33, no. 5, pp. 710–711, 1987.

[10] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 620–643, 2009.

[11] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3032–3050, 2009.

[12] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum–rate capacity for Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 689–699, 2009.

[13] R. H. Etkin, D. N. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.

[14] H. Sato, "On degraded Gaussian two-user channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 637–640, 1978.

[15] M. H. Costa, "On the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 607–615, 1985.

[16] R. S. Cheng and S. Verdu, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE transactions on information theory*, vol. 39, no. 2, pp. 609–612, 1993.

[17] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 581–586, 2004.

[18] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 364–375, 1996.

[19] S. Shamai and S. Verdú, "Capacity of channels with uncoded side information," *European Transactions on Telecommunications*, vol. 6, no. 5, pp. 587–600, 1995.

[20] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, 1982.

[21] J. H. Conway and N. Sloane, "Fast quantizing and decoding and algorithms for lattice quantizers and codes," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 227–232, 1982.

[22] G. D. Forney Jr, "Geometrically uniform codes," *IEEE Transactions on Information Theory*, vol. 37, no. 5, pp. 1241–1260, 1991.

[23] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the awgn channel," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 273–278, 1998.

[24] J. Lee, H. Kwon, and I. Kang, "Interference mitigation in mimo interference channel via successive single-user soft decoding," in *Information Theory and Applications Workshop*, 2012, pp. 180–185.

[25] A. Yedla, P. S. Nguyen, H. D. Pfister, and K. R. Narayanan, "Universal codes for the Gaussian mac via spatial coupling," in *Allerton Conference on Communication, Control, and Computing*, 2011, pp. 1801–1808.

[26] B. Rimoldi, "Generalized time sharing: a low-complexity capacity-achieving multiple-access technique," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2432–2442, Sep 2001.

[27] L. Wang, E. Sasoglu, and Y.-H. Kim, "Sliding-window superposition coding for interference networks," in *International Symposium on Information Theory*, 2014, pp. 2749–2753.

[28] A. Khandani, "Two-way (true full-duplex) wireless," in *Canadian Workshop on Information Theory*, June 2013, pp. 33–38.

[29] I. Maric, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3536–3548, 2007.

[30] I. Marić, A. Goldsmith, G. Kramer *et al.*, "On the capacity of interference channels with one cooperating transmitter," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 405–420, 2008.

[31] I.-H. Wang and D. N. Tse, "Interference mitigation through limited transmitter cooperation," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2941–2965, 2011.

[32] V. M. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 156–186, 2011.

[33] S. Yang and D. Tuninetti, "Interference channel with generalized feedback (aka with source cooperation): Part i: Achievable region," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2686–2710, 2011.

[34] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "On the capacity of the two-user Gaussian causal cognitive interference channel," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2512–2541, May 2014.

[35] V. S. Annapureddy, A. El Gamal, and V. V. Veeravalli, "Degrees of freedom of the k-user interference channel with transmitter cooperation," in *International Symposium on Information Theory Proceedings*. IEEE, 2010, pp. 385–389.

[36] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[37] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge University Press, 2011.

[38] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[39] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *International Symposium on Information Theory Proceedings*. IEEE, 2005, pp. 2065–2069.

[40] O. Mehanna, J. Marcos, and N. Jindal, "On achievable rates of the two-user symmetric Gaussian interference channel," in *Allerton Conference on Communication, Control, and Computing*, Sept 2010, pp. 1273–1279.

[41] D. Tuninetti and Y. Weng, "On the Han-Kobayashi achievable region for Gaussian interference channels," in *International Symposium on Information Theory*, 2008, pp. 240–244.

[42] M. H. Costa and C. Nair, "On the achievable rate sum for symmetric Gaussian interference channels," in *Information Theory and Applications Workshop, San-Diego, California, USA*, 2012.

[43] A. Haghi and A. Khandani, "The maximum han-kobayashi sum-rate for Gaussian interference channel," *International Symposium on Information Theory*, June 2016.

[44] H.-F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the han–kobayashi region for theinterference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 7, pp. 3188–3195, 2008.

[45] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming.* Springer Science & Business Media, 2008, vol. 116.

[46] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[47] H.-F. Chong, M. Motani, and H. K. Garg, "A comparison of two achievable rate regions for the interference channel," in *Proceedings of USCD-ITA Workshop*, 2006.

[48] X. Shang and B. Chen, "A new computable achievable rate region for the Gaussian interference channel," in *International Symposium on Information Theory*, 2007, pp. 2191–2195.

[49] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1345–1356, 2004.

[50] C. Nair, "Upper concave envelopes and auxiliary random variables," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 1, pp. 12–20, 2013.

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

[52] S. Liu, C. Nair, and L. Xia, "Interference channels with very weak interference," in *IEEE International Symposium on Information Theory.* IEEE, 2014, pp. 1031–1035.

[53] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2087–2104, April 2014.

[54] P. P. Bergmans and T. M. Cover, "Cooperative broadcasting," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 317–324, 1974.

[55] A. D. Wyner, "Recent results in the shannon theory," *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 2–10, 1974.

[56] R. G. GALLAGER, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory*, vol. 31, no. 2, 1985.

[57] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems.* Cambridge University Press, 2011.

[58] A. Grant, B. Rimoldi, R. Urbanke, and P. Whiting, "Single user coding for the discrete memoryless multiple access channel," in *IEEE International Symposium on Information Theory.* IEEE, 1995, p. 448.

[59] O. Fawzi and I. Savov, "Rate-splitting in the presence of multiple receivers," *arXiv preprint arXiv:1207.0543*, 2012.

[60] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "Qos and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," in *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2. IEEE, 2002, pp. 477–486.

[61] H. Boche and S. Stańczak, "Optimal qos tradeoff and power control in cdma systems," in *Iwenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 2. IEEE, 2004, pp. 1078–1088.

[62] S.-J. Oh and A. C. Soong, "Qos-constrained information-theoretic sum capacity of reverse link cdma. systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 3–7, 2006.

[63] M. Ebrahimi, M. A. Maddah-Ali, and A. K. Khandani, "Power allocation and asymptotic achievable sum-rates in single-hop wireless networks," in *40th Annual Conference on Information Sciences and Systems*. IEEE, 2006, pp. 498–503.

[64] Y. Zhao, C. W. Tan, A. S. Avestimehr, S. N. Diggavi, and G. J. Pottie, "On the maximum achievable sum-rate with successive decoding in interference channels," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3798–3820, 2012.

[65] A. S. Avestimehr, S. N. Diggavi, and D. N. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, 2011.

[66] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," *European transactions on telecommunications*, vol. 19, no. 4, pp. 333–354, 2008.

[67] A. Haghi and A. K. Khandani, "Achievable sum-rate of the two-user Gaussian interference channel through rate-splitting and successive decoding," in *53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2015, pp. 1438–1445.

[68] ——, "Rate-splitting and successive decoding for Gaussian interference channels," in *Iran Workshop on Communication and Information Theory (IWCIT)*. IEEE, 2016.

[69] D. N. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *IEEE International Symposium on Information Theory (ISIT)*, Jun 1997.

[70] V. R. Cadambe and S. A. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3983–3990, 2009.

[71] L. Sankar, X. Shang, E. Erkip, and H. V. Poor, "Ergodic fading interference channels: Sum-capacity and separability," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2605–2626, 2011.

[72] A. Haghi and A. K. Khandani, "The separability and ergodic sum-rate of parallel Gaussian interference channels," in *IEEE International Symposium on Information Theory*, June 2014, pp. 851–855.

[73] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[74] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[75] O. Simeone, J. Gambini, Y. Bar-Ness, and U. Spagnolini, "Cooperation and cognitive radio," in *International Conference on Communications*, June 2007, pp. 6511–6515.

[76] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[77] S. I. Bross, Y. Steinberg, and S. Tinguely, "The causal cognitive interference channel," in *International Zurich Seminar on Communications*, 2010, p. 53.

[78] M. Mirmohseni, B. Akhbari, and M. R. Aref, "On the capacity of interference channel with causal and noncausal generalized feedback at the cognitive transmitter," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2813–2837, 2012.

[79] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[80] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of mimo communications-a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, 2004.

[81] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Transactions on Information theory*, vol. 52, no. 4, pp. 1522–1544, 2006.

[82] C. Huang and S. A. Jafar, "Degrees of freedom of the mimo interference channel with cooperation and cognition," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4211–4220, 2009.

[83] A. Haghi, N. Mohammadizadeh, and A. K. Khandani, "Delay in cooperative communications: Higher multiplexing gain in Gaussian interference channels with full-duplex transmitters," in *53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2015, pp. 1486–1493.

[84] T. M. Cover, R. J. McEliece, and E. C. Posner, "Asynchronous multiple-access channel capacity," *IEEE Transactions on Information Theory*, vol. 27, no. 4, pp. 409–413, 1981.

[85] J. Zyren and W. McCoy, "Overview of the 3gpp long term evolution physical layer," *Freescale Semiconductor, Inc., white paper*, 2007.

[86] G. H. Golub and C. F. Van Loan, *Matrix computations.* JHU Press, 2012, vol. 3.

[87] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, no. 5, pp. 628–634, October 1971.

[88] Z. Wang and G. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[89] R. v. Nee and R. Prasad, *OFDM for wireless multimedia communications.* Artech House, Inc., 2000.

[90] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3413–3431, Oct 2007.