Cooperative Interference Management in Wireless Networks

I-Hsiang Wang

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2011-103
http://www.eecs.berkeley.edu/Pubs/TechRpts/2011/EECS-2011-103.html

September 12, 2011

Cooperative Interference Management in Wireless Networks

by

I-Hsiang Wang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

and the Designated Emphasis

in

Communication, Computation and Statistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David N. C. Tse, Chair
Professor Suhas Diggavi
Professor Ilan Adler

Fall 2011

Cooperative Interference Management in Wireless Networks

Abstract

Cooperative Interference Management in Wireless Networks

by

I-Hsiang Wang

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

and the Designated Emphasis in Communication, Computation and Statistics

University of California, Berkeley

Professor David N. C. Tse, Chair

With the growing number of users along with ever-increasing demand for higher data rates and better quality of service in modern wireless networks, interference has become the major barrier against efficient utilization of limited resources. On the other hand, opportunities for cooperation among radios also increase with the growing number of users, which potentially lead to better interference management. In traditional wireless system design, however, such opportunities are usually neglected and only basic interference management schemes are employed, mainly due to lack of fundamental understanding of interference and cooperation.

In this dissertation, we study the fundamental aspects of cooperative interference management through the lens of network information theory. In the first and the second part, we characterize both qualitatively and quantitatively how limited cooperation between transmitting or receiving terminals helps mitigate interference in a canonical two-transmitter-two-receiver wireless system. We identify two regions regarding the gain from limited cooperation: linear and saturation regions. In the linear region cooperation is efficient and provides a *degrees-of-freedom* gain, which is either *one cooperation bit buys one bit over the air* or *two cooperation bits buy one bit over the air* until saturation. In the saturation region cooperation is inefficient and only provides a bounded *power* gain. The conclusions are drawn from the approximate characterization of the capacity regions.

In the third part, we investigate how intermediate relay nodes help resolve interference in delivering information from two sources to their respective destinations in multi-hop wireless networks. We focus on a linear deterministic approximate model for wireless networks, and when the minimum cut value between each source-destination pair is constrained to be 1, we completely characterize the capacity region. One of the interesting findings is that, at most four nodes need to take special coding operations so that interference can be canceled over-the-air or within-a-node, while other nodes can take oblivious operations. We also develop a systematic approach to identify these special nodes.

To my family.

Contents

List of Figures

List of Tables

Acknowledgments

I am greatly indebted to my adviser, collaborators, teachers, colleagues, friends, and family, for all the help and encouragement I have received during this wonderful journey.

First of all, I would like to express my deepest gratitude to my adviser, Professor David Tse, for his guidance and support throughout my study at Berkeley. Beyond his experience and profound knowledge in the field, he has taught me how to formulate and solve research problems by identifying the key issues and ideas behind complicated structures. His research philosophy of pursuing the simplest solution and the most intuitive interpretation helps me realize the essence of good scientific research. I am also greatly benefitted from his high standard on the quality of publications and presentations and learn how to give good talks from his valuable comments.

I would like to thank my colleague Sudeep Kamath for the collaboration on the problem of two unicast wireless networks. I would also like to appreciate Professor Borivoje Nikolic, my colleagues Vinayak Nagpal and Milos Jorgovanovic for the fruitful collaboration on the project of cooperative MIMO. I am grateful to Professor Amos Lapidoth for hosting me at ETH Zurich during the summer of 2010 and the collaboration on the problem of remote joint source channel coding. My gratitude further extends to Professor Suhas Diggavi and Professor Ilan Adler for serving on the dissertation committee and their insightful comments on the dissertation. Great thanks to Professor Jan Rabaey and Kannan Ramchandran for their considerable participation in the qualifying exam committee.

I would like to thank Professor Venkat Anantharam, Jean Walrand, Michael Gastpar, and Anant Sahai, for their excellent teaching during my education at Berkeley, which extends my horizons beyond my own expertise of research. I would also like to thank my fellow colleagues in the Wireless Foundations, specifically Guy Bresler, Changho Suh, Mohammad Ali Maddah-Ali, Bobak Nazer, and Pulkit Grover, and all my friends at Berkeley, especially Ting-Ta Yen and Ting-Ying Chung, for their help and friendship.

Finally, I would like to extend my deepest appreciation to my family and I-Hsuan, for all the love and support. Without them, I would not have gone so far.

Chapter 1

Introduction

In the past two decades, significant progress has been made in wireless communications. Theoretical advances led by information theory have given birth to various modern wireless communication techniques, including multiple-input-multiple-output (MIMO) techniques, code-division multiple access (CDMA), orthogonal frequency-division multiple access (OFDMA), etc. These ideas have been successfully implemented and deployed in real worlds, such as CDMA2000, WCDMA, WLAN, WiMAX, LTE, etc., and are now used in our daily lives. Wireless technologies have greatly changed the way we collect information and communicate with one another. As a result, the demand for higher data rates as well as better quality of service has been increasing rapidly.

With more and more users but fixed amount of resource, including spectrum, power, etc., *interference* has become the major factor that limits performance of wireless systems. Interference arises whenever multiple source-destination pairs are present in a network, and each destination is only interested in retrieving information from its own source(s). Due to the broadcast and superposition nature of wireless medium, one user's information-carrying signal causes interference to other users. When the strength of interference is comparable with the desired signal, it causes severe problems to the interfered source-desitination pair. Unfortunately, current wireless system designs either treat interference as noise, which degrades system performance when the interference becomes strong, or orthogonalize interferences from desired signals, which causes shortage of resource when the number of interferers grows. In short, only basic *interference management* schemes are employed in current wireless systems.

One of the main reasons why more advanced interference management schemes did not receive much attention is that, there was no such demand for them in the past as interference was not the bottleneck of system performance. More importantly, despite the advances in wireless communications, interference has not been well understood fundamentally. The simplest information theoretic model for studying interference is the two-user interference channel, introduced by Claude E. Shannon in 1961 [1]. Characterizing its capacity region has been open for 50 years, except for several special cases, including the strong interference regime [2], a class of deterministic interference channels [3], etc. The largest achievable rate region to date was reported by Han and Kobayashi in 1981 [4]. After 30 years, it is still the best we know even for this simplest two-user model.

From the past experience, to improve the design of wireless systems, information theoretic advances must be made first. Driven by the need of more advanced interference management schemes, in the past few years, significant progress has been made towards understanding the fundamental limit of delivering information in interference channels. In particular, for the two-user Gaussian interference channel, the capacity region was characterized to within 1 bit/s/Hz [5], regardless of channel parameters. The bounded-gap-to-optimality result consolidates the optimality of Han-Kobayashi scheme at high signal-to-noise ratio, and leads to an uniform approximation of the capacity region. Later, the outer-bounding technique in [5] was improved, and the sum capacity in a very weak interference regime and a mixed interference regime was characterized [6] [7] [8].

Theoretical advances in understanding the fundamental capacity limit of the two-user interference channel make a huge step towards better interference management techniques in wireless networks. However, in the above interference channel set-up, there are no intermediate radios between each source-destination pair, and the sources (destinations) are thought of as isolated radios, not allowed to communicate with one another. Hence, they have to combat interference on their own. Such assumptions are no longer appropriate for modern wireless networks. In various scenarios, sources/destinations are not isolated, and there may be more radios involved in the network. *Cooperation* among radios can thus be induced, and potentially they can better combat interference in a joint effort. For example, in cellular systems, base stations are able to exchange certain amount of information and cooperate through the infrastructure backhaul network, and in wireless ad-hoc networks, there are usually radios other than the sources and the destinations that can cooperatively help mitigate interferences for both users. Unfortunately, the fundamental limit of *cooperative interference management* has not been well understood even in the simplest two-user Gaussian interference channel with cooperation, not to mention more general wireless networks with additional radios other than sources and destinations.

Without better fundamental understanding, the potential of using cooperation to improve interference management cannot be further explored in practice. Therefore in this work, we focus on the fundamental limit of cooperative interference management in wireless networks. We first consider the two-user Gaussian interference channel (GIC) with receiver cooperation and the two-user GIC with transmitter cooperation, to understand how cooperation help mitigate interference in the canonical two-user system. We then study the problem of delivering information over a two-source tow-destination wireless network with arbitrary number of radio nodes, not limited to four (two sources and two destinations) as in the two-user interference channel.

1.1 Background

In this section, we briefly review some background knowledge for developing the dissertation. First we introduce a *linear deterministic model* for studying Gaussian networks [9], which will be used throughout the dissertation. Next, to better understand the two key elements of cooperative interference management, namely, *cooperation* and *interference*, we review some previous work on the single-source single-destination Gaussian relay network and the two-user Gaussian interference channel, and point out the key elements that will be helpful in developing our results.

1.1.1 Linear Deterministic Channel Model

The *linear deterministic model* [9] to be introduced below was originally proposed by Avestimehr *et al.*to study the approximate capacity limit of the single-source single-destination Gaussian relay network. It is analytically simpler than the Gaussian model, while it still captures the two key features of wireless medium, broadcast and superposition. The motivation to study such a model is that in contrast to fixed point-to-point channels where noise is the only source of uncertainty, in multiuser communication, the signal interactions are a critical source of uncertainty. Therefore, for a first level of understanding, our focus is on such signal interactions rather than the received noise. Later, the model is also shown to be useful in studying the two-user Gaussian interference channel [10]. For more intuitions about the model and its connection to the Gaussian model, please refer to [9].

Consider a general Gaussian network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ stand for the collection of nodes (radios) and directed edges (links) respectively, the the received signal at any node $\mathsf{v} \in \mathcal{V}$ is specified as follows:

$$y_{\mathsf{v}} = \sum_{(\mathsf{u},\mathsf{v})\in\mathcal{E}} h_{\mathsf{vu}} x_{\mathsf{u}} + z_{\mathsf{v}}, \tag{1.1}$$

where $h_{\mathsf{vu}} \in \mathbb{C}$ is the complex channel gain associated with the link $(\mathsf{u},\mathsf{v}) \in \mathcal{E}$, $x_{\mathsf{u}} \in \mathbb{C}$ denotes the transmit signal of the node u, and $z_{\mathsf{v}} \sim \mathcal{CN}(0,1)$ denotes the additive white Gaussian noise (circularly symmetric with unit variance) at the node v. We assume unit average transmit power constraint at each node. Hence for every link $(\mathsf{u},\mathsf{v}) \in \mathcal{E}$, the associated signal-to-noise ratio $\mathsf{SNR}_{\mathsf{uv}} = |h_{\mathsf{vu}}|^2$.

The corresponding linear deterministic network is defined over the same network topology $\mathcal{G}$, while the received signal at any node $\mathsf{v} \in \mathcal{V}$ becomes

$$\mathbf{y}_{\mathsf{v}} = \sum_{(\mathsf{u},\mathsf{v})\in\mathcal{E}} S^{(q-n_{\mathsf{vu}})} \mathbf{x}_{\mathsf{u}}, \tag{1.2}$$

where additions are modulo-two component-wise, $n_{\mathsf{vu}} \in \mathbb{N}$ is the channel "gain" associated with the link $(\mathsf{u},\mathsf{v}) \in \mathcal{E}$, $q = \max_{(\mathsf{u},\mathsf{v})\in\mathcal{E}}\{n_{\mathsf{vu}}\}$, $\mathbf{x}_{\mathsf{u}} \in \mathbb{F}_2^q$ denotes the transmit signal of the node

u, and $S \in \mathbb{F}_2^{q \times q}$ is the shift matrix

$$S = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \tag{1.3}$$

An interpretation of this model considers the *binary* expansion of signals. The effect of additive white Gaussian noise is modeled by *truncation* of the signal below the noise level. The effect of superposition with interference is modeled by the modulo-two component-wise addition of the bits in the binary expansion, where the carry-over in real addition is neglected for simplicity. Hence, we have the following correspondence:

$$n_{\mathsf{vu}} = \left(\lfloor \log |h_{\mathsf{vu}}|^2 \rfloor \right)^+ = \left(\lfloor \log \mathsf{SNR}_{\mathsf{uv}} \rfloor \right)^+ . \tag{1.4}$$

1.1.2 Single-Source Single-Destination Relay Network

The simplest information theoretic model for studying cooperation is the single relay channel, introduced by van der Meulen [11]. The capacity of the single relay channel remains open afterwards, even for the Gaussian relay channel, its capacity is unknown. The most general strategies for this simple network were developed by Cover and El Gamal [3], including decode-and-forward (DF) and compress-and-forward (CF). In the DF scheme, the relay decodes the message sent from the source, re-encodes it, and transmits the codeword to the destination so that the destination can better decode the message by incorporating relay's side information. In the CF scheme, on the other hand, the relay compresses its received signal, re-encodes the compression index, and sends it to the destination. Then the destination decodes the message with the help from the relay. It turns out that both DF and CF can achieve the capacity to within 1 bits/s/Hz, regardless of channel parameters [9]. Extension of the strategies in [3] to a more general relay network was made in [12]. The extension, however, seems not able to achieve the capacity to within a bounded gap for general Gaussian relay networks [9].

The recent work by Avestimehr *et al.*[9] connected Ford and Fulkerson's well-know max-flow min-cut theorem in single-unicast wired networks [13] to an *approximate*[1] max-flow min-cut theorem in Gaussian relay networks, by introducing the linear deterministic network as an intermediate model. The capacity of the linear deterministic network was characterized, and the result was a natural generalization of the max-flow min-cut theorem for wired networks. Based on the insights obtained from the deterministic analysis, a quantize-map-and-forward (QMF) scheme was proposed for Gaussian networks, where each relay quantizes the received signal at the noise level and maps it to a random Gaussian codeword for forwarding, and the

[1] The maximum information flow is within a bounded gap to the minimum cut-set value.

final destination decodes the source's message based on the received signal. In contrast to existing schemes, this scheme can achieve the cut-set upper bound to within a gap which is independent of the channel parameters. QMF scheme turns out to be useful in the two-user interference channel with receiver cooperation, developed in Chapter 2.

However, the extension of the result to the network with more than single source and more than single destination (that is, more than single unicast session in the network) is non-trivial, and QMF alone no longer suffices to achieve the capacity to within a bounded gap. Unlike single unicast networks where good understanding has been well established both for wired and wireless networks, even for two unicast wired networks, the capacity region remains open. In Chapter 6 we will present some results towards this direction for wireless networks.

1.1.3 Two-User Interference Channel

When there are more than single source and more than single destination in wireless networks, the simplest scenario is the two-user interference channel. As mentioned above, recent progress [5] has shown that the Han-Kobayashi scheme achieves the capacity of the two-user Gaussian IC to within 1 bit/s/Hz. The Han-Kobayashi scheme is nothing but a superposition coding scheme [4]. It turns out that a simple superposition of Gaussian random codes suffices to achieve the capacity to within 1 bit regardless of channel parameters. Such a scheme is the key building block of the two-user interference channel with receiver cooperation and the two-user interference channel with transmitter cooperation developed in Chapter 2 and Chapter 4 respectively. Therefore, below we first briefly discuss the simple Han-Kobayashi scheme used in the two-user Gaussian interference channel, adapted from [5].

The Han-Kobayashi strategy involves splitting the transmitted information of both users into two parts: *private* information to be decoded only at the intended receiver and *common* information that can be decoded at both receivers. By decoding the common information, part of the interference can be canceled off, while the remaining private information from the other user is treated as noise. The Han-Kobayashi strategy allows arbitrary splits of each users transmit power into the private and common information portions as well as time sharing between multiple such splits. Hence, the key is to find the most suitable power split. It turns out that the power of the private information of each user should be set such that it is received at the level of the Gaussian noise at the other receiver. In this way, the interference caused by the private information has a small effect on the other link as compared to the impairments already caused by the noise. At the same time, quite a lot of private information can be conveyed in the own link if the direct gain is appreciably larger than the cross gain. Moreover, the scheme involves only a single private-common split and no time-sharing is needed.

The linear deterministic model provides a clear and intuitive explanation about the power-split mentioned above. In the linear deterministic interference channel, at each transmitter, the bit levels of the transmitted signal can be divided naturally into two parts: private levels, which do not appear at the other receiver due to the shift matrix, and common levels, which appear at the other receiver. This gives a clear illustration why such power-split in the Gaussian model is reasonable. On the other hand, the capacity region of the linear deterministic interference channel is known, as it is a special case of the deterministic interference channel considered in [14]. Hence by mimicking the outer-bounding techniques in [14], tighter outer bounds can be derived in the Gaussian case [5].

When we deal with the case where the transmitters or the receiver are allowed to cooperate, the schemes and outer bounds will build upon the insights obtained in the case without cooperation, while more new elements will come into play due to cooperation. More details will be described in Chapter 2 and Chapter 4.

1.2 Organization of the Dissertation

The dissertation comprises of three parts. The contents of Part I and Part II are published in [15] and [16] respectively. The work of Part III is in collaboration with Sudeep U. Kamath and is published in [17]. Each part is divided into two chapters, where the first chapter describes the main contribution and the second chapter complements the first with detailed proofs.

In Part I and Part II we investigate how much interference can limited cooperation mitigate in the canonical two-transmitter-two-receiver system. In Chapter 2 and Chapter 4, we respectively investigate the two-user Gaussian interference channel with conferencing[2] receivers and the two-user Gaussian interference channel with conferencing transmitters. We characterize the capacity region to within a bounded gap in both scenarios by deriving tight outer bounds and good cooperative interference management strategies. Further, we show that there is an interesting reciprocity between the scenario with conferencing transmitters and the scenario with conferencing receivers, and their capacity regions are within a bounded gap to each other. Hence in the interference-limited regime, the behavior of the benefit brought by transmitter cooperation is the same as that by receiver cooperation. Qualitatively, we identify two regions regarding the gain from limited cooperation: linear and saturation regions. In the linear region cooperation is efficient and provides a *degrees-of-freedom* gain, which is either *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit* until saturation. In the saturation region cooperation is inefficient and provides only a *power* gain, which is bounded regardless of the rate at which the two terminals cooperate.

In Part III, we investigate the problem of two unicast wireless information flows by studying the two unicast linear deterministic networks. When the minimum cut value between each source-destination pair is constrained to be 1, we completely characterize the capacity region and conclude that there are exactly five possible capacity regions of this class of networks. Our achievability scheme is based on linear coding over an extension field with

[2] Cooperation is orthogonal to the signals in the interference channel.

at most four nodes performing special linear coding operations, namely interference neutralization and zero forcing, while all other nodes perform oblivious operations, that is, random linear coding.

Part I

Interference Mitigation through Receiver Cooperation

Chapter 2

Interference Channel with Receiver Cooperation

Interference is a major issue limiting the performance in wireless networks. Cooperation among receivers can help mitigate interference by forming distributed MIMO systems. The rate at which receivers cooperate, however, is limited in most scenarios. How much interference can one bit of receiver cooperation mitigate? In this chapter, we study the two-user Gaussian interference channel with conferencing receivers to answer this question in a simple setting. We identify two regions regarding the gain from receiver cooperation: linear and saturation regions. In the linear region, receiver cooperation is efficient and provides a *degrees-of-freedom* gain, which is either *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit*. In the saturation region receiver cooperation is inefficient and provides a *power* gain, which is bounded regardless of the rate at which receivers cooperate. The conclusion is drawn from the characterization of capacity region to within two bits/s/Hz, regardless of channel parameters. The proposed strategy consists of two parts: (1) the transmission scheme, where superposition encoding with a simple power split is employed, and (2) the cooperative protocol, where one receiver quantize-bin-and-forwards its received signal, and the other after receiving the side information decode-bin-and-forwards its received signal.

2.1 Introduction

In modern communication systems, interference is one of the fundamental factors that limit performance. The simplest information theoretic model for studying this issue is the *two-user interference channel*. Characterizing its capacity region is a long-standing open problem, except for several special cases (eg., the strong interference regime [2]). The largest achievable region to date is reported by Han and Kobayashi [4], and the core of the scheme is a superposition coding strategy. Recent progress has been made on both inner bounds and outer bounds: Etkin, Tse, and Wang characterize the capacity region of the two-user Gaussian interference channel to within one bit [5] by using a superposition coding scheme with a simple power-split configuration and by providing new upper bounds. The bounded-gap-to-optimality result [5] leads to an uniform approximation of the capacity region and provides a strong guarantee on the performance of the proposed scheme. Later, Motahari and Khandani [6], Shang, Kramer, and Chen [7], and Annapureddy and Veeravalli [8] independently improve the outer bounds and characterize the sum capacity in a very weak interference regime and a mixed interference regime.

In the above interference channel set-up, transmitters or receivers are not allowed to communicate with one another, and each user has to combat interference on its own. In various scenarios, however, nodes are not isolated, and transmitters/receivers can exchange certain amount of information. Cooperation among transmitters/receivers can help mitigate interference by forming distributed MIMO systems which provide two kinds of gains: *degrees-of-freedom* gain and *power* gain. The rate at which they cooperate, however, is limited, due to physical constraints. Therefore, one of the fundamental questions is, how much *interference* can limited *transmitter/receiver cooperation* mitigate? How much gain can it provide?

In this chapter, we consider a two-user Gaussian interference channel with *conferencing receivers* to answer this question regarding receiver cooperation. Conferencing among encoders/decoders has been studied in [18], [19], [20], [21], [22], and [23]. Our model is similar to those in [22] and [23] but in an interference channel set-up. The work in [22] characterizes the capacity region of the compund multiple access channel (MAC) with unidirectional conferencing between decoders. For general set-up (i.e., bidirectional conferencing), it provides achievable rates and finds the maximum achievable individual rate to within a bounded gap, but is not able to establish an uniform approximation result on the capacity region. The work in [23] considers one-sided Gaussian interference channels with unidirectional conferencing between decoders and characterizes the capacity region in strong interference regimes and the asymptotic sum capacity at high SNR. For general receiver cooperation, works including [24] and [25], investigate cooperation in interference channels with a set-up where the cooperative links are in the same band as the links in the interference channel. In particular, [25] characterizes the sum capacity of Gaussian interference channels with symmetric in-band receiver cooperation to within a bounded gap. Our work, on the other hand, is focused on the Gaussian interference channel with out-of-band (orthogonal) receiver cooperation, and studies its entire capacity region. Works on interference channels with additional relays [26] [27] [28] and two-hop interference-relay networks [29] are also related to our problem, since the receivers also serve as relays in our set-up.

We propose a strategy achieving the capacity region universally to within 2 bits/s/Hz per user, regardless of channel parameters. The two-bit gap is the worst-case gap which can be loose in some regimes, and it is vanishingly small at high SNR when compared to the capacity. The strategy consists of two parts: (1) the transmission scheme, describing how transmitters encode their messages, and (2) the cooperative protocol, describing how receivers exchange information and decode messages. For transmission, both transmitters use superposition coding [4] with the same common-private power split as in the case without cooperation [5]. For the cooperative protocol, it is appealing to apply the decode-forward or compress-forward schemes, originally proposed in [3] for the relay channel, like most works dealing with more complicated networks, including [21], [22], [23], [24], [12], etc. It turns out neither conventional compress-forward nor decode-forward achieves capacity to within a bounded gap for the problem at hand. On the other hand, [30], [31], [32], [33], and [9] observe that the conventional compress-forward scheme [3] may be improved by the destination directly decoding the sender's message instead of requiring to first decode the quantized signal of the relay. We use such an improved compress-forward scheme as part of our cooperative protocol. One of the receivers quantizes its received signal at an appropriate distortion, bins the quantization codeword and sends the bin index to the other receiver. The other receiver then decodes its own information based on its own received signal and the received bin index. After decoding, it bin-and-forwards the decoded common messages back to the former receiver and helps it decode. Note that although arbitrary number of rounds is allowed in the conferencing formulation, it turns out that two rounds are sufficient to achieve within 2 bits of the capacity.

We identify two regions regarding the gain from receiver cooperation: linear and saturation regions, as illustrated through a numerical example in Fig. 2.1. In the plot we fix the signal-to-noise ratios (SNR) and interference-to-noise ratios (INR) to be 20dB and 15dB respectively, and we plot the over-the-air user data rate versus the cooperation rate. In the linear region, receiver cooperation is *efficient*, in the sense that the growth of over-the-air user data rate is roughly linear with respect to the capacity of receiver-cooperative links. The gain in this region is the *degrees-of-freedom* gain that distributed MIMO systems provide. On the other hand, in the saturation region, receiver cooperation is *inefficient* in the sense that the growth of each user's over-the-air data rate becomes saturated as one increases the rate in receiver-cooperative links. The gain is the *power* gain that is bounded, regardless of the cooperation rate. We will focus on the system performance in the linear region, because not only that in most scenarios the rate at which receivers can cooperate is limited, but also that the gain from cooperation is more significant.

With the bounded-gap-to-optimality result, we find that the fundamental gain from cooperation in the linear region as follows: either *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit* until saturation, depending on channel parameters. In the symmetric set-up, at high SNR, when INR is below 50% of SNR in dB scale, one-bit cooperation per direction buys roughly one-bit gain per user until full receiver cooperation performance is reached, while when INR is between 67% and 200% of SNR in dB scale, one-bit cooperation per direction buys roughly half-bit gain per user. In the weak interference regime, for a given pair of $(\mathsf{SNR}, \mathsf{INR})$, when the receiver-cooperative link capacity $\mathsf{C}^{\mathsf{B}} > \log \mathsf{INR}$, cooperation between receivers can get a close-to-interference-free (that is, within a bounded gap) performance. In the strong interference regime, in contrast to that without cooperation, system performance can be boost *beyond* interference-free performance, by utilizing receiver-cooperative links not only for interference mitigation but also

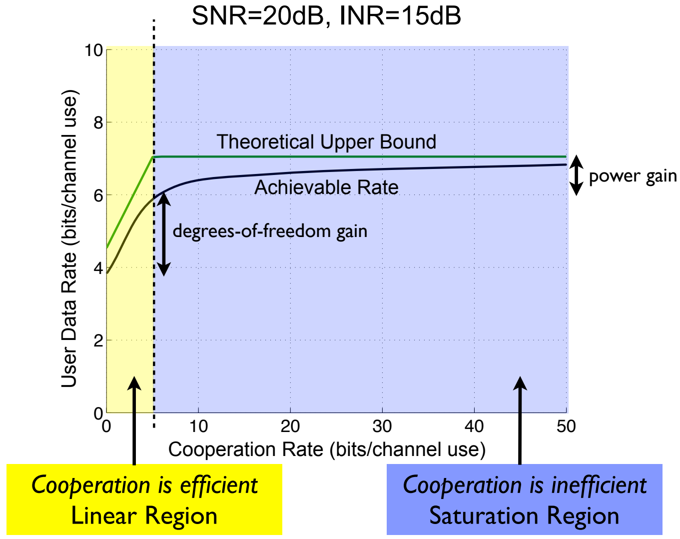


Figure 2.1: The Gain from Limited Receiver Cooperation

for forwarding desired information, since the cross link is stronger than the direct link.

The rest of this chapter is organized as follows. In Section 2.2, we introduce the channel model and formulate the problem. In Section 2.3, we provide intuitive discussions about achievability and motivate our two-round strategy. Then we give examples to illustrate why it is not a good idea to use cooperative protocols based on conventional compress-forward or decode-forward. In Section 2.4, we describe the strategy concretely and derive its achievable rates, and in Section 2.5 we show that the achievable rate region is within 2 bits per user to the outer bounds we provide. In addition, we characterize the capacity region of the compound MAC with conferencing receivers to within 1 bit, as a by-product. In Section 2.7, focusing on the symmetric set-up, we illustrate the fundamental gain from receiver cooperation by deriving the optimal number of *generalized degrees of freedom* (g.d.o.f.) and compare it with the achievable ones of suboptimal schemes.

2.2 Problem Formulation

2.2.1 Channel Model

The two-user Gaussian interference channel with conferencing receivers is depicted in Fig. 4.1.

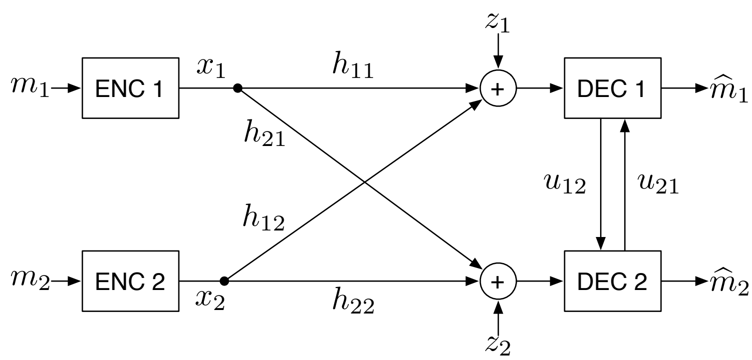


Figure 2.2: Channel Model

Transmitter-Receiver Links

The transmitter-receiver links are modeled as the *normalized* Gaussian interference channel:

$$y_1 = h_{11}x_1 + h_{12}x_2 + z_1$$
$$y_2 = h_{21}x_1 + h_{22}x_2 + z_2,$$

where the additive noise processes $\{z_i[n]\}$, $(i = 1, 2)$, are independent $\mathcal{CN}(0, 1)$, i.i.d. over time. In this chapter, we use [.] to denote time indices. Transmitter i intends to convey message m_i to receiver i by encoding it into a block codeword $\{x_i[n]\}_{n=1}^{N}$, with transmit power constraints

$$\frac{1}{N}\sum_{n=1}^{N}\left|x_i[n]\right|^2 \le 1,\ i = 1, 2,$$

for arbitrary block length N. Note that the outcome of each encoder depends solely on its own message. Messages m_1, m_2 are independent. Define channel parameters

$$\mathsf{SNR}_i := |h_{ii}|^2,\ \mathsf{INR}_i := |h_{ij}|^2,\ i, j = 1, 2,\ i \ne j.$$

Receiver-Cooperative Links

For $(i, j) = (1, 2), (2, 1)$, the receiver-cooperative links are noiseless with capacity $\mathsf{C}_{ij}^{\mathsf{B}}$ from receiver i to j. Encoding must satisfy causality constraints: for any time index $n = 1, 2, \ldots, N$, the cooperation signal from receiver 2 to 1, $u_{21}[n]$, is only a function of $\{y_2[1], \ldots, y_2[n-1], u_{12}[1], \ldots, u_{12}[n-1]\}$, and the cooperation signal from receiver 1 to 2, $u_{12}[n]$, is only a function of $\{y_1[1], \ldots, y_1[n-1], u_{21}[1], \ldots, u_{21}[n-1]\}$.

In the rest of this chapter, we use v^n to denote the sequence $\{v[1], \ldots, v[n]\}$.

2.2.2 Strategies, Rates, and Capacity Region

We give the basic definitions for the coding strategies, achievable rates of the strategy, and the capacity region of the channel.

Definition 2.1 (Strategy and Average Probability of Error). *An* (M_1, M_2, N)*-strategy consists of the following: for* $i, j = 1, 2,\ i \neq j$,

(1) message set $\mathcal{M}_i := \{1, 2, \ldots, M_i\}$ *for user* i*;*

(2) encoding function $e_i^{(N)} : \mathcal{M}_i \to \mathbb{C}^N,\ m_i \mapsto x_i^N$ *at transmitter* i*;*

(3) set of relay functions $\{r_i^{(n)}\}_{n=1}^N$ *such that* $u_{ij}[n] = r_i^{(n)}(y_i^{n-1}, u_{ji}^{n-1}) \in \{1, 2, \ldots, 2^{\mathsf{C}^{\mathsf{B}}_{ij}}\}$, $\forall n = 1, 2, \ldots, N$ *at receiver* i*;*

(4) decoding function $d_i^{(N)} : \mathbb{C}^N \times \{1, 2, \ldots, 2^{N\mathsf{C}^{\mathsf{B}}_{ji}}\} \to \mathcal{M}_i,\ (y_i^N, u_{ji}^N) \mapsto \widehat{m}_i$ *at receiver* i*.*

The average probability of error

$$P_e^{(N)} := \frac{1}{M_1 M_2} \sum_{\substack{m_1 \in \mathcal{M}_1 \\ m_2 \in \mathcal{M}_2}} \Pr \left\{ \begin{array}{l} d_1^{(N)}(y_1^N, u_{21}^N) \neq m_1 \text{ or} \\ d_2^{(N)}(y_2^N, u_{12}^N) \neq m_2 \end{array} \middle| \begin{array}{l} m_1, m_2 \\ \textit{are sent} \end{array} \right\}$$

Definition 2.2 (Achievable Rates and Capacity Region). *A rate tuple* (R_1, R_2) *is achievable if for any* $\epsilon > 0$ *and for all sufficiently large* N*, there exists an* (M_1, M_2, N) *strategy with* $M_i \geq 2^{NR_i}$*, for* $i = 1, 2$*, such that* $P_e^{(N)} < \epsilon$*. The capacity region* $\mathscr{C}$ *is the collections of all achievable* (R_1, R_2)*.*

2.2.3 Notations

We summarize below the notations used in the rest of this chapter.

- For a real number a, $(a)^+ := \max(a, 0)$ denotes its positive part.
- For sets $A, B \subseteq \mathbb{R}^k$ in an k-dimensional space, $A \oplus B := \{a + b : a \in A, b \in B\}$ denotes the direct sum of A and B. $\text{conv}\{A\}$ denotes the convex hull of the set A.
- With a little abuse of notations, for $x, y \in \mathbb{F}_q$, $x \oplus y$ denotes the modulo-q sum of x and y.
- Unless specified, all the logarithms $\log(.)$ are of base 2.

2.3 Motivation of Strategies

Before introducing our main result, we first provide intuitive discussions about achievability and motivate our two-round strategy (to be described in detail in Section 2.4) from a high-level perspective. Then we give examples to illustrate why cooperative protocols based on conventional compress-forward or decode-forward may not be good for cooperation between receivers to mitigate interference. Throughout the discussion in this section, we will make use of the *linear deterministic model* proposed in [34] [9].

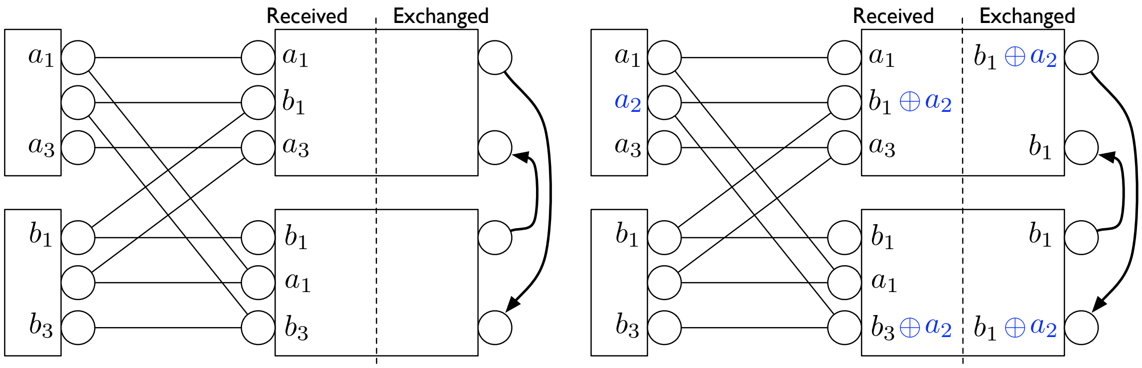
The linear deterministic model is a tool for studying Gaussian networks so that an uniform approximation of the capacity can be found. It is also used for the two-user interference channel [10]. The model captures the signal interaction in the original Gaussian scenario to some extent, and is useful for illustrating some subtle facts which are not easy to be uncovered in the Gaussian scenario. Throughout this chapter, all discussions involving the linear deterministic model are either aimed to elucidate a certain phenomenon that arises in the Gaussian scenario, or to provide an intuitive argument for a certain claim without rigorously proving it.

2.3.1 Optimal Strategy in the Linear Deterministic Channel

First, consider the following symmetric channel: $\mathsf{SNR}_1 = \mathsf{SNR}_2 = \mathsf{SNR}$, $\mathsf{INR}_1 = \mathsf{INR}_2 = \mathsf{INR}$, and $\mathsf{C}^{\mathsf{B}}_{12} = \mathsf{C}^{\mathsf{B}}_{21} = \mathsf{C}^{\mathsf{B}}$. Set INR to be 2/3 of SNR in dB scale, that is, $\log \mathsf{INR} = \frac{2}{3} \log \mathsf{SNR}$. Set $\mathsf{C}^{\mathsf{B}} = \frac{1}{3} \log \mathsf{SNR}$. The corresponding linear deterministic channel (LDC) is depicted in Fig. 2.3. The bits at the levels of transmitters/receivers can be thought of as chunks of binary expansions of the transmitted/received signals. Note that in this example, one bit in the LDC corresponds to $\frac{1}{3} \log \mathsf{SNR}$ bits in the Gaussian channel. Because $\mathsf{INR} < \mathsf{SNR}$, the least significant bit (LSB) of each transmitter appears below noise level at the other receiver and is invisible.

In the discussions below, bit $a_k \in \mathbb{F}_2$ denotes the bit sent at the k-th level from the most significant bit (MSB) at transmitter 1, and similarly $b_k \in \mathbb{F}_2$ denotes the bit sent at the k-th level at transmitter 2.

We begin with the baseline where two receivers are not allowed to cooperate. The transmitted signals are naturally broken down into two parts: (1) the common levels, which appear at both receivers, and (2) the private levels, which only appear at its own receiver. Each transmitter splits its message into common and private parts, which are linearly modulated onto the common and private levels of the signal respectively. Each receiver then decodes both user's common messages and its own private message by solving the linear equations it received. This is shown to be optimal in the two-user interference channel [10]. In this example (Fig. 2.3.(a)), bits a_1 and b_1 are common, while a_3 and b_3 are private. The sum capacity without cooperation is 4 bits. One cannot turn on the bit a_2 (or b_2) since the number of variables (bits) to be solved at the receiver 1, that is, $\{a_1, a_3, b_1\}$, has already met the maximum number of equations it has.

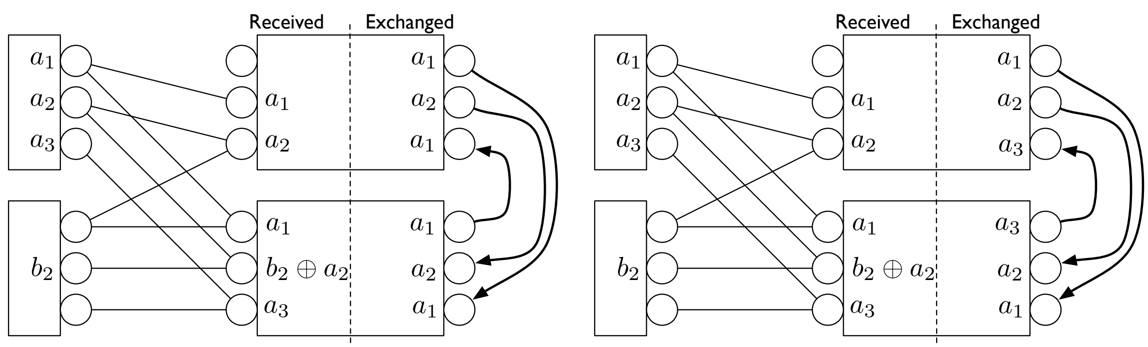


(a) Without Cooperation (b) With Cooperation

Figure 2.3: An Example Channel

With receiver cooperation, the natural split of transmitted signals does not change. This suggests that the encoding procedure and the aim of each decoder remain the same. Each receiver with the help from the other receiver, however, is able to decode more information because it has additional linear equations. Since each user's private message is not of interest to the other receiver, a natural scheme for receiver cooperation is to exchange linear combinations formed by the signals *above* the private signal level so that the undesired signal does not pollute the cooperative information. In this example, as illustrated in Fig. 2.3.(b), with one-bit cooperation in each direction in the LDC, the optimal sum rate is 5 bits, achieved by turning on one more bit a_2. This causes collisions at the second level at receiver 1 and at the third level at receiver 2, but they can be resolved with cooperation: receiver 1 sends $b_1 \oplus a_2$ to receiver 2, and receiver 2 sends b_1 to receiver 1. Now receiver 1 can solve (a_1, a_2, a_3, b_1), and receiver 2 can solve (b_1, b_3, a_1, a_2). In fact, the exchanged linear combinations are not unique. For example, receiver 1 can send $(b_1 \oplus a_2) \oplus a_1$ and receiver 2 can send $b_1 \oplus a_1$, and this again achieves the same rates. As long as receiver 1 does not send a linear combination containing the private bit a_3 and the sent linear combination is linearly independent of the signals at receiver 2 (and vice versa for the linear combination sent from receiver 2 to receiver 1), the scheme is optimal for this example channel. The above discussion regarding the scheme in the LDC naturally leads to an implementable one-round scheme in the Gaussian channel, where both receivers quantize and bin their received signals at their own private signal level.

In the above example, it is optimal that each receiver sends to each other linear combinations formed by its received signals above its private signal level. Is this optimal in general? The answer is no. Consider the following asymmetric example: $\mathsf{SNR}_2 = \mathsf{INR}_2$, SNR_1 is $2/3$ of SNR_2 in dB, and INR_1 is $1/3$ of SNR_2 in dB. $\mathsf{C}^{\mathsf{B}}_{12} = \frac{2}{3}\log \mathsf{SNR}_2$ and $\mathsf{C}^{\mathsf{B}}_{21} = \frac{1}{3}\log \mathsf{SNR}_2$. The corresponding LDC is depicted in Fig. 2.4, where one bit in the LDC corresponds to $\frac{1}{3}\log \mathsf{SNR}_2$ in the Gaussian channel. First consider the same scheme as that in the previous exmaple. Note that if receiver 2 just forwards signals above its private signal level, it can only forward a_1 to receiver 1 and achieves R_1 up to 2 bits. On the other hand, if receiver 2 forwards a_3 to receiver 1, which is below user 2's private signal level, it achieves $R_1 = 3$ bits. From this example, we see that once there is "useful" information (which should not be polluted by the receiver's own private bits) which lies *at or below* the private signal level (in this example, the bit a_3), the one-round scheme described in the previous example is suboptimal. To extract the useful information at or below the private signal level, one of the receivers (in this example, receiver 2) can first decode and then form linear combinations using (decoded) common messages *only*.



(a) Suboptimal Scheme (b) Optimal Scheme

Figure 2.4: An Asymmetric Example

It turns out that without loss of generality, the above situation (where there is useful information for the other receiver lies at or below the private signal level) only happens at most at one receiver. In other words, there exists a receiver where no useful information (for the other receiver) lies at or below the private signal level. The reason is the following:

1. It is straightforward to see that the capacity region is convex, and hence if a scheme can achieve $\max_{(R_1,R_2)\in\mathscr{C}} \{\mu_1 R_1 + \mu_2 R_2\}$ for all $\mu_1, \mu_2 \geq 0$, it is optimal.

2. If $\mu_1 \geq \mu_2$, we weigh user 1's rate more. Since the private bits are cheaper to support in the sense that they do not cause interference at receiver 2, user 1 should be transmitting at its full private rate, which is equal to the number of levels at or below the private signal level at receiver 1. Therefore, all levels at or below the private signal level are occupied by user 1's private bits and there is no useful information at receiver 1 for receiver 2.

3. Similarly if $\mu_1 \leq \mu_2$, there is no useful information at receiver 2 for receiver 1 at or below the private signal level.

Hence, the following two-round strategy turns out to be optimal in the LDC (the proof is omitted here): if $\mu_1 \geq \mu_2$, receiver 1 forms a certain number (no more than the cooperative link capacity) of linear combinations composed of the signals above its private signal level and sends them to receiver 2. After receiver 2 decodes, it forms a certain number of linear combinations composed of the decoded common bits and sends them to receiver 1. If $\mu_1 \le \mu_2$, the roles of receiver 1 and 2 are exchanged. Note that depending on the operating point in the capacity region, we use different configurations, implying that time-sharing is needed to achieve the full capacity region.

From the above discussion, a natural and implementable two-round strategy for Gaussian channels emerges. For the transmission, we use a superposition Gaussian random coding scheme with a simple power-split configuration, as described in [5]. For the cooperative protocol, one of the receivers quantize-and-bins its received signal at its private signal level and forwards the bin index; after the other receiver decodes with the helping side information, it bin-and-forwards the decoded common messages back to the first receiver and helps it decode. In Section 2.5, we shall prove that this strategy achieves the capacity region universally to within 2 bits per user.

2.3.2 Conventional Compress-Forward and Decode-Forward

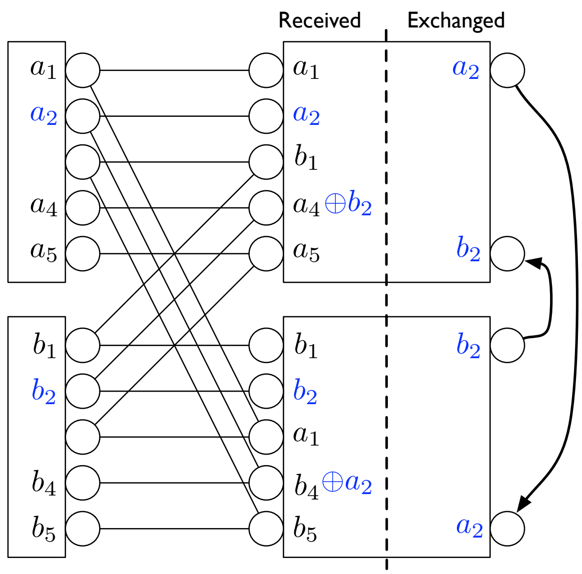
We have motivated the two-round strategy to be proposed formally in the next section from a high level perspective. Below we shall illustrate why conventional compress-forward (CF) and decode-forward (DF) are not good in certain regimes.

It is a standard approach to evaluate achievable rates of Gaussian relay networks using conventional compress-forward with Gaussian vector quantization (VQ) assuming joint Gaussianity of the received signals at relays and destination in the literature, including [21], [22], [23], [24], [12], etc. What if we replace the quantize-binning part in the two-round strategy proposed above by the conventional compress-forward with Gaussian VQ, as in [35], [21], and [22], where the two-round idea is also used?

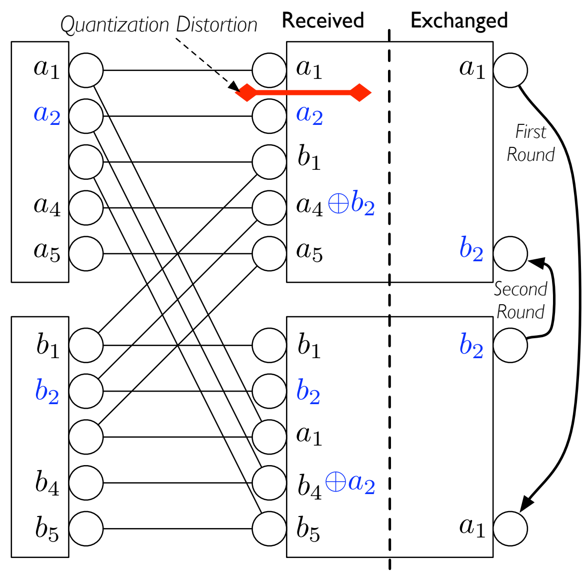
Let us consider another symmetric channel with $\log \mathsf{INR} = \frac{3}{5} \log \mathsf{SNR}$ and $\mathsf{C}^{\mathsf{B}} = \frac{1}{5} \log \mathsf{SNR}$. From its corresponding LDC in Fig. 2.5, one can see that the two received signals of the Gaussian channel, (y_1, y_2), are not jointly Gaussian. The reason is that, suppose they are jointly Gaussian, the conditional distribution of y_2 given y_1 should be marginally Gaussian. As Fig. 2.5 suggests, however, conditioning on receiver 1's signal results in a hole at the third level of receiver 2's signal, which was occupied by a_1. Therefore, transmitter 2's common codebook is not dense enough to make the conditional distribution of y_2 given y_1 marginally Gaussian. The incorrect assumption results in larger quantization distortions, as depicted in Fig. 2.5.(b)[1]. The information sent from receiver 1 to receiver 2, a_1, is *redundant*, and cannot help mitigate interference a_2. Hence, the achievable sum rate is 7 bits (4 bits for user 1 and 3 bit for user 2), while the one-round scheme in Fig. 2.5.(a) achieves 8 bits. Recall that 1 bit

[1]If we view the received signals as vectors of bits rather than binary expansions of Gaussian signals, we are not restricted to send the MSB a_1 to receiver 2 and a_2 can be sent instead. However, this kind of scheme cannot be implemented in the Gaussian scenario using conventional compress-forward with Gaussian VQ.

in the LDC corresponds to $\frac{1}{5}\log\mathsf{SNR}$ in the Gaussian channel, therefore the performance loss is unbounded as $\mathsf{SNR}\to\infty$. The main reason why conventional compress-forward does not work well is that the scheme does not well utilize the dependency between the two received signals.



(a) Optimal Scheme (One Round)



(b) Conventional Compress-Forward in First Round

Figure 2.5: Another Example Channel

Another standard approach is to use decode-forward for the two receivers to cooperate. Let us go back to the first example and consider the channel in Fig. 2.3. Note that there is no gain if we require both common messages to be decoded at one of the receivers at the first stage without cooperation. By symmetry we can assume that, without loss of generality, each receiver first decodes its own common message and then bin-and-forwards the decoded information to the other receiver. At the second stage, it then decodes the other user's common message with the help from cooperation and decodes its own private message. In the corresponding LDC, the common bit a_2 cannot be decoded at the first stage, and hence the total throughput using this strategy is at most 4 bits, which is again the same as that without cooperation. The reason why decode-forward is not good for the two receivers to cooperate is that, it is too costly to decode users' own common message at the first stage without the help from cooperation.

2.4 A Two-Round Strategy

In this section we describe the two-round strategy and derive its achievable rate region. The strategy consists of two parts: (1) the transmission scheme and (2) the cooperative protocol.

2.4.1 Transmission Scheme

We use a simple superposition coding scheme with Gaussian random codebooks. For each transmitter, it splits its own message into common and private (sub-)messages. Each common message is aimed at both receivers, while each private one is aimed at its own receiver. Each message is encoded into a codeword drawn from a Gaussian random codebook with a certain power. For transmitter i, the power for its private and common codes are Q_{ip} and $Q_{ic} = 1 - Q_{ip}$ respectively, for $i = 1, 2$. As [5] points out, since the private signal is undesired at the unintended receiver, a reasonable configuration is to make the private interference at or below the noise level so that it does not cause much damage and can still convey additional information in the direct link if it is stronger than the cross link. When the interference is stronger than the desired signal, simply set the whole message to be common. In a word, for $(i, j) = (1, 2)$ or $(2, 1)$, $Q_{ip} = \min\left\{\frac{1}{\mathsf{INR}_j}, 1\right\}$ if $\mathsf{SNR}_i > \mathsf{INR}_j$, and $Q_{ip} = 0$ otherwise.

2.4.2 Cooperative Protocol

The cooperative protocol is two-round. We briefly describe it as follows: for $(i, j) = (1, 2)$ or $(2, 1)$, at the first round, receiver j quantizes its received signal and sends out the bin index (the procedure is described in detail below). At the second round, receiver i receives this side information, decodes its desired messages (both users' common messages and its own private message) with the decoder described in detail below, randomly bins the decoded common messages, and sends the bin indices to receiver j. Finally receiver j decodes with the help from the receiver-cooperative link. We call this a two-round strategy $\mathrm{STG}_{j\to i\to j}$, meaning that the processing order is: receiver j quantize-and-bins, receiver i decode-and-bins, and receiver j decodes. Its achievable rate region is denoted by $\mathscr{R}_{j\to i\to j}$. By time-sharing, we can obtain achievable rate region $\mathscr{R} := \mathrm{conv}\left\{\mathscr{R}_{2\to 1\to 2} \cup \mathscr{R}_{1\to 2\to 1}\right\}$, convex hull of the union of two rate regions.

Remark 2.3 (Engineering Interpretation). *There is a simple way to understand the strategy from an engineering perspective. To achieve* $\max_{(R_1,R_2)\in\mathscr{R}}\left\{\mu_1 R_1 + \mu_2 R_2\right\}$ *for some non-negative* (μ_1, μ_2)*, the processing configuration can be easily determined: strategy* $\mathrm{STG}_{j\to i\to j}$ *should be used, where* $i = \arg\min_{l=1,2}\{\mu_l\}$ *and* $j = \arg\max_{l=1,2}\{\mu_l\}$*. In a word, the receiver which decodes last is the one we favor most. This is the high-level intuition we obtained from the discussion in the LDC in Section 2.3.1.*

In the following, we describe each component in detail, including quantize-binning, decode-binning, and their corresponding decoders. For simplicity, we consider strategy

$\text{STG}_{2\to1\to2}$.

Quantize-binning (receiver 2):

Upon receiving its signal from the transmitter-receiver link, receiver 2 does not decode messages immediately. Instead, serving as a relay, it first quantizes its signal by a pre-generated Gaussian quantization codebook with certain distortion, and then sends out a bin index determined by a pre-generated binning function. How should we set the distortion? As discussed in the previous section, note that both its own (user 2's) private signal and the noise it encounters are not of interest to receiver 1. Therefore, a natural configuration is to set the distortion level equal to *the aggregate power level of the noise and user 2's private signal*.

Decoder at receiver 1:

After retrieving the receiver-cooperative side information, that is, the bin index, receiver 1 decodes two common messages and its own private message, by searching in transmitters' codebooks for a codeword triple (indexed by user 1 and user 2's common messages and user 1's own private message) that is jointly typical with its received signal and some quantization point (codeword) in the given bin. If there is no such unique codeword triple, it declares an error.

Decode-binning (receiver 1):

After receiver 1 decodes, it uses two pre-generated binning functions to bin the two common messages and sends out these two bin indices to receiver 2.

Decoder at receiver 2:

After receiving these two bin indices, receiver 2 decodes two common messages and its own private message, by searching in the corresponding bins (containing common messages) and user 2's private codebook for a codeword triple that is jointly typical with its received signal.

Remark 2.4 (Difference from the Conventional CF)**.** *The action of receiver 2 as a relay is very similar to that of the relay in the conventional compress-forward with Gaussian vector quantization. Note that the main difference from the conventional compress-forward with Gaussian vector quantization lies in the* decoding *procedure (at receiver 1) and the chosen distortion. In the conventional Gaussian compress-forward, the decoder first searches in the bin for one quantization codeword that is jointly typical with its received signal from its own transmitter* only*, assuming that the two received signals are jointly Gaussian. This may not be true since a single user may not transmit at the capacity in its own link, which results in "holes" in signal space. As a consequence, this scheme may not utilize the dependency of two received signals well and cause larger distortions. Our scheme, on the other hand, utilizes the dependency in a better way by* jointly *deciding the quantization codeword and the message triple, consequently allows smaller distortions, and is able to reveal the* *beneficial side information to the other receiver. Quantize-binning and its corresponding decoding part of our scheme is very similar to* extended hash-and-forward *proposed in [31], in which it was pointed out that the scheme has no advantage over conventional compress-forward in a single-source single-relay setting. In the Gaussian single-relay channel (with orthogonal noise-free relay-destination link), the received signal at the relay and the destination are indeed jointly Gaussian when communicating at the quantize-map-and-forward achievable rate, and hence the performances of the two schemes are the same. Due to the above mentioned issues, however, we recognize in our problem where the channel consists of two source-destination pairs and two relays, the scheme has an unbounded advantage over the conventional compress-forward in certain regimes. Such phenomena are also observed in single-source single-destination Gaussian relay networks [9] [36] and interference-relay channels [28] [36].*

2.4.3 Achievable Rates

The following theorem establishes the achievable rates of strategy $\mathrm{STG}_{2\to1\to2}$. Let R_{ic} and R_{ip} denote the rates for user i's common message and private message respectively, for $i = 1, 2$.

Theorem 2.5 (Achievable Rate Region for $\mathrm{STG}_{2\to1\to2}$). *The rate tuple $(R_{1c}, R_{2c}, R_{1p}, R_{2p})$ satisfying the following constraints is achievable:*

<u>Constraints at receiver 1</u>*:*

$$R_{1p} \le \min\left\{ I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right) \right\} \tag{2.1}$$

$$R_{2c} \le \min\left\{ I\left(x_{2c}; y_1 | x_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_{2c}; y_1, \widehat{y}_2 | x_1\right) \right\} \tag{2.2}$$

$$\begin{aligned} R_{2c} + R_{1p} &\le \min\left\{ I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right) \right\} \\ R_{1c} + R_{1p} &\le \min\left\{ I\left(x_1; y_1 | x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{2c}\right) \right\} \\ R_{1c} + R_{2c} + R_{1p} &\le \min\left\{ I\left(x_1, x_{2c}; y_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) \right\} \end{aligned} \tag{2.3}$$

where

$$\xi_1 = I\left(\widehat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1\right).$$

For $i = 1, 2$, $x_{ic} \sim \mathcal{CN}\left(0, Q_{ic}\right)$ is the common codebook generating random variable. $x_1 = x_{1p} + x_{1c}$ is the superposition codebook generating variable, where $x_{1p} \sim \mathcal{CN}\left(0, Q_{1p}\right)$ is independent of x_{1c}. $\widehat{y}_2 \overset{d}{=} y_2 + \widehat{z}_2$ is the quantization codebook generating random variable, and $\widehat{z}_2 \sim \mathcal{CN}\left(0, \Delta_2\right)$, independent of everything else. Δ_2 is the quantization distortion at receiver 2.

Constraints at receiver 2:

$$\begin{aligned} R_{2p} &\leq I\left(x_2; y_2 | x_{2c}, x_{1c}\right) \\ R_{1c} + R_{2p} &\leq I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{2p} &\leq I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{1c} + R_{2p} &\leq I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}, \end{aligned}$$

where $x_2 = x_{2p} + x_{2c}$ is the superposition codebook generating variable, and $x_{2p} \sim \mathcal{CN}\left(0, Q_{2p}\right)$ is independent of x_{2c}.

Proof. For details, see Chapter 3.1. Here we give some high-level comments on these rate constraints. First, unlike interference channels without cooperation, here receiver 1 is required to decode m_{2c} correctly so that it can help receiver 2. This additional requirement gives the rate constraint (2.2) on R_{2c}.

Second, in the set of constraints at receiver 1, on the right-hand side they are all minimum of two terms. The second term corresponds to the case when the receiver-cooperative link is strong enough to convey the quantized $\widehat{y}_2^N$ correctly. The first term corresponds to the case when receiver 1 can only figure out a set of candidates of quantized $\widehat{y}_2^N$. Regarding the "rate loss" term ξ_1, in Section 2.3 we see that in the LDC as long as the quantization level is chosen such that no private signals pollute the cooperative information, there is no such penalty. In fact, $\xi_1 = I\left(\widehat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1\right)$ corresponds to the number of private bits polluting the cooperative linear combinations in the LDC if one chooses the quantization distortion to be too small. In the Gaussian channel, however, due to the carry-over of real additions, the private part will always "leak" into the levels above the quantization level and hence there is always at least a bounded rate loss even if we choose the quantization distortion properly.

Finally, in the set of constraints at receiver 2, since receiver 1 only helps receiver 2 decode m_{1c} and m_{2c}, there is no enhancement in R_{2p}. □

We shall use the following shorthand notations throughout the rest of the chapter: for $(i,j) = (1,2), (2,1)$,

$$\begin{aligned} \mathsf{SNR}_{ip} &:= |h_{ii}|^2 Q_{ip} = \mathsf{SNR}_i \cdot Q_{ip}, \\ \mathsf{INR}_{ip} &:= |h_{ij}|^2 Q_{jp} = \mathsf{INR}_i \cdot Q_{jp}. \end{aligned}$$

Next, we quantify the "rate loss" term ξ_1 in the set of rate constraints at receiver 1, in terms of distortions Δ_2:

$$\begin{aligned} \xi_1 &= I(\widehat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1) \\ &= h(\widehat{y}_2 | x_{1c}, x_1, x_{2c}, y_1) - h(\widehat{y}_2 | x_{1c}, x_1, x_{2c}, y_1, y_2) \\ &= h\left(h_{22} x_{2p} + z_2 + \widehat{z}_2 | h_{12} x_{2p} + z_1\right) - h\left(\widehat{z}_2\right) \\ &= \log\left(\frac{1+\Delta_2}{\Delta_2} + \frac{\mathsf{SNR}_{2p}}{(1+\mathsf{INR}_{1p})\Delta_2}\right) \end{aligned}$$

$$\leq \log\left(\frac{1+\Delta_2+\mathsf{SNR}_{2p}}{\Delta_2}\right), \tag{2.4}$$

Below we shall see why the intuition of quantizing at the private signal level works. By choosing $\Delta_2 = 1 + \mathsf{SNR}_{2p}$, the "rate loss" ξ_1 is upper bounded by 1. In particular, when $\mathsf{SNR}_2 \leq \mathsf{INR}_1$, we have $\mathsf{SNR}_{2p} = 0$ and hence $\xi_1 = 1$. On the other hand, note that for receiver 1 the unwanted signal power level in y_2 is exactly $1 + \mathsf{SNR}_{2p}$, and receiver 1 treats the unwanted signals as noise anyway. Hence, replacing $\widehat{y}_2$ by y_2 only increases the rate by a bounded gain.

Remark 2.6. *The above configuration of the distortion may not be optimal. The achievable rates can be further improved if we optimize over all possible distortions. For example, if the cooperative link capacity is large, one could lower the distortion level to yield a finer description of received signals. With the above simple configuration, however, we are able to show that it achieves the capacity region to within a bounded gap. Also note that in this chapter, we generate the quantization codebook in a slightly different way than that in conventional lossy source coding, where instead a "test channel"* $y_2 = \widehat{y}_2 + \widehat{z}_2$ *is used. With this choice the rate loss* ξ_1 *can be made smaller, while the calculations become more complicated.*

2.5 Characterization of the Capacity Region to within 2 Bits

The main result in this section is the characterization of the capacity region to within 2 bits per user universally, regardless of channel parameters. To prove it, first we provide outer bounds of the capacity region. Ideas about how to prove them are outlined, and details are left in Chapter 3. Then we make use of Theorem 2.5 to evaluate the achievable rate region, and show that it is within 2 bits per user to the proposed outer bounds.

2.5.1 Outer Bounds

To prove the outer bounds, the main idea is the following: first, upper bound the rates by mutual informations via Fano's inequality and data processing inequality; second, decompose them into two parts: (1) terms which are similar to those in Gaussian interference channels without cooperation, and (2) terms which correspond to the enhancement from cooperation. We use the genie-aided techniques in [5] to upper bound the first part and obtain namely the Z-channel bound (where the genie gives interfering symbols x_j^N to receiver i, $i \neq j$) and ETW-bound (where the genie gives the interference term caused by user i at receiver j, $s_i^N := h_{ji}x_i^N + z_j^N$ to receiver i). For the second part, we make use of the fact that u_{12}^N and u_{21}^N are both functions of (y_1^N, y_2^N), and other straightforward bounding techniques. The results are summarized in the following lemma.

Lemma 2.7. *$\mathscr{C} \subseteq \overline{\mathscr{C}}$, where $\overline{\mathscr{C}}$ consists of nonnegative rate tuples (R_1, R_2) satisfying the following inequalities*

$$R_1 \le \log(1+\mathsf{SNR}_1) + \min\left\{\mathsf{C}^{\mathsf{B}}_{21}, \log\left(1+\frac{\mathsf{INR}_2}{1+\mathsf{SNR}_1}\right)\right\} \tag{2.5}$$

$$R_2 \le \log(1+\mathsf{SNR}_2) + \min\left\{\mathsf{C}^{\mathsf{B}}_{12}, \log\left(1+\frac{\mathsf{INR}_1}{1+\mathsf{SNR}_2}\right)\right\} \tag{2.6}$$

$$R_1 + R_2 \le \log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \mathsf{C}^{\mathsf{B}}_{21} + \mathsf{C}^{\mathsf{B}}_{12} \tag{2.7}$$

$$R_1 + R_2 \le \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \mathsf{C}^{\mathsf{B}}_{12} \tag{2.8}$$

$$R_1 + R_2 \le \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \mathsf{C}^{\mathsf{B}}_{21} \tag{2.9}$$

$$R_1 + R_2 \le \log\left(1+\mathsf{SNR}_1+\mathsf{SNR}_2+\mathsf{INR}_1+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right) \tag{2.10}$$

$$\begin{aligned} 2R_1 + R_2 \le\ & \log\left(1+\mathsf{INR}_2+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) \\ & + \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \mathsf{C}^{\mathsf{B}}_{21} + \mathsf{C}^{\mathsf{B}}_{12} \end{aligned} \tag{2.11}$$

$$\begin{aligned} R_1 + 2R_2 \le\ & \log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\ & + \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} \end{aligned} \tag{2.12}$$

$$\begin{aligned} 2R_1 + R_2 \le\ & \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}+\mathsf{INR}_2+\mathsf{SNR}_1+\frac{\mathsf{INR}_1}{1+\mathsf{INR}_1}+\frac{|h_{11}h_{22}-h_{12}h_{21}|^2}{1+\mathsf{INR}_1}\right) \\ & + \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right) + \mathsf{C}^{\mathsf{B}}_{21} \end{aligned} \tag{2.13}$$

$$\begin{aligned} R_1 + 2R_2 \le\ & \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}+\mathsf{INR}_1+\mathsf{SNR}_2+\frac{\mathsf{INR}_2}{1+\mathsf{INR}_2}+\frac{|h_{11}h_{22}-h_{12}h_{21}|^2}{1+\mathsf{INR}_2}\right) \\ & + \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right) + \mathsf{C}^{\mathsf{B}}_{12} \end{aligned} \tag{2.14}$$

Proof. Details are left in Chapter 3.2. Below we give a short outline and intuitions. First of all, bounds (2.5), (2.6), and (2.10) are straightforward cut-set upper bounds of individual rates and sum rate respectively.

Bound (2.7) corresponds to the ETW-bound in Gaussian interference channels without cooperation. In the genie-aided channel, we upper bound the gain from receiver cooperation by $\mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21}$, that is, in both directions each bit is useful.

Bounds (2.8) and (2.9) correspond to the Z-channel bounds. In the genie-aided channel, since the genie gives interfering symbols x_j^N to receiver i, $i \neq j$, there is no interference at receiver i. Intuitively, the cooperation from receiver j to i is now providing only the power gain, and the genie can provide y_j^N to receiver i to upper bound this power gain. The gain from the cooperation from receiver i to j is upper bounded by $\mathsf{C}^{\mathsf{B}}_{ij}$.

Bounds (2.11) and (2.12) on R_i+2R_j are derived by giving side information s_i^N to receiver i and side information x_i^N and y_i^N to one of the receiver j's. In the genie-aided channel there is an underlying Z-channel structure, and hence the gain from one direction of the cooperation is absorbed into a power gain. The rest is upper bounded by $\mathsf{C}^{\mathsf{B}}_{12}+\mathsf{C}^{\mathsf{B}}_{21}$.

Bounds (2.13) and (2.14) on R_i+2R_j are derived by giving side information y_j^N and $\widetilde{s}_i^N := h_{ji}x_i^N + \widetilde{z}_j^N$, where $\widetilde{z}_j \sim \mathcal{CN}(0,1)$ and independent of everything else, to receiver i and side information y_i^N to one of the receiver j's. In the genie-aided channel, there is an underlying point-to-point MIMO channel, and hence the gain from both directions of cooperation is absorbed into the MIMO system. The rest is upper bounded by $\mathsf{C}^{\mathsf{B}}_{ij}$.

Note that the derivation of all bounds works for all INR's and SNR's. □

We make the following observations:

Remark 2.8 (Dependence on Phases)**.** *The sum-rate cut-set bound (2.10) not only depends on SNR's and INR's but also on the phases of channel coefficients, due to the term $|h_{11}h_{22}-h_{12}h_{21}|^2$. In particular, when the receiver-cooperative link capacities C^{B}'s are large, the two receivers become near-fully cooperated, and the system performance is constrained by that of the SIMO MAC; that is, it enters the saturation region. Therefore this bound becomes active and the outer bound depends on phases.*

Remark 2.9 (Strong Interference Regime)**.** *When $\mathsf{SNR}_1 \le \mathsf{INR}_2$ and $\mathsf{SNR}_2 \le \mathsf{INR}_1$, unlike the Gaussian interference channel of which the capacity region is equal to that of a compound MAC in the strong interference regime [2], here we cannot apply Sato's argument. Recall that when there is no cooperation, once user i's own message is decoded successfully at receiver i, it can produce $\widetilde{y}_j^N$ which has the same distribution as y_j^N. Since the error probability for decoding user j's message at receiver j only depends on the marginal distribution of y_j^N, it can be concluded that at receiver i one can achieve the same performance for decoding user j's message by using the same decoder as that in receiver j, and hence receiver i can decode user j's message successfully as well. When there is cooperation, however, the error probability for decoding user j's message at receiver j depends on the joint distribution of (y_j^N, u_{ij}^N). Note that the additive noise terms in $\widetilde{y}_j^N$ and y_j^N have different correlations with the noise term z_i^N, and u_{ij}^N can be highly correlated with z_i^N. As a consequence, the joint distributions of (y_j^N, u_{ij}^N) and $(\widetilde{y}_j^N, u_{ij}^N)$ are not guaranteed to be the same, and receiver i may not be able to achieve the same performance for decoding user j's message by using the same decoder as that in receiver j. Therefore, we cannot claim that the capacity region under strong interference condition is the same as that of compound MAC with conferencing receivers (CMAC-CR). Instead, we take the Z-channel bounds (2.8) and* (2.9)*, which are within 1 bit to the sum rate cut-set bound of CMAC-CR in strong interference regimes. This will be discussed in the last part of this section.*

2.5.2 Capacity Region to within 2 bits

Subsequently we investigate three qualitatively different cases, namely, weak interference, mixed interference, and strong interference[2], in the rest of this section. We summarize the main achievability result in the following theorem: (recall that $\overline{\mathscr{C}}$ is the outer bound region defined in Lemma 2.7)

Theorem 2.10 (Within Two-Bit Gap to Capacity Region)**.**

$$\mathscr{R} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R} \oplus \big([0,2] \times [0,2]\big),$$

Proof. Proved by Lemma 2.11, 2.14, and 2.17 in the rest of this section. □

2.5.3 Weak interference

In the case $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 > \mathsf{INR}_1$, the configuration of superposition coding is to split message m_i into m_{ic} and m_{ip}, for both users $i = 1, 2$. We first consider $\mathrm{STG}_{2\to1\to2}$: referring to Theorem 2.5, we obtain the set of achievable rates $(R_{1c}, R_{2c}, R_{1p}, R_{2p})$. The term $\xi_1 \leq 1$ bit, due to (2.4) in Section 2.4.3 and the chosen distortion $\Delta_2 = 1 + \mathsf{SNR}_{2p}$.

To simplify calculations, note that the right-hand-side of (2.1), (2.2), and (2.3) are at most a bounded gap from their lower bounds $I\left(x_1; y_1|x_{1c}, x_{2c}\right)$, $I\left(x_{2c}; y_1|x_1\right)$, and $I\left(x_1; y_1|x_{2c}\right)$ respectively. Therefore, we replace these three constraints by

$$\begin{aligned} R_{1p} &\leq I\left(x_1; y_1|x_{1c}, x_{2c}\right), \\ R_{2c} &\leq I\left(x_{2c}; y_1|x_1\right), \\ R_{1c} + R_{1p} &\leq I\left(x_1; y_1|x_{2c}\right) \end{aligned}$$

in the following calculations. Next, rewriting $R_{ip} = R_i - R_{ic}$ for $i = 1, 2$, applying Fourier-Motzkin algorithm to eliminate R_{1c} and R_{2c}, and removing redundant terms (details omitted here), we obtain an achievable $\mathscr{R}_{2\to1\to2}$, which consists of nonnegative (R_1, R_2) satisfying:

$$R_1 \leq \min\left\{I\left(x_1; y_1|x_{2c}\right), I\left(x_1; y_1|x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2|x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\}$$
$$R_2 \leq \min\left\{I\left(x_2; y_2|x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}}, I\left(x_{2c}; y_1|x_1\right) + I\left(x_2; y_2|x_{1c}, x_{2c}\right)\right\}$$
$$R_1 + R_2 \leq I\left(x_1, x_{2c}; y_1\right) + I\left(x_2; y_2|x_{1c}, x_{2c}\right) + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \tag{2.15}$$
$$R_1 + R_2 \leq I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2|x_{1c}, x_{2c}\right) \tag{2.16}$$
$$R_1 + R_2 \leq I\left(x_1, x_{2c}; y_1|x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} + I\left(x_{1c}, x_2; y_2|x_{2c}\right) + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \tag{2.17}$$
$$R_1 + R_2 \leq I\left(x_1, x_{2c}; y_1, \widehat{y}_2|x_{1c}\right) + I\left(x_{1c}, x_2; y_2|x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \tag{2.18}$$

[2]The definitions of these cases are the following: (1) weak interference, where $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 > \mathsf{INR}_1$; (2) mixed interference, where $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 \leq \mathsf{INR}_1$; (3) strong interference, where $\mathsf{SNR}_1 \leq \mathsf{INR}_2$ and $\mathsf{SNR}_2 \leq \mathsf{INR}_1$.

$$R_1 + R_2 \le I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}^{\mathsf{B}}_{12} \tag{2.19}$$

$$R_1 + R_2 \le I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{2c}; y_1|x_1) + I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} \tag{2.20}$$

$$2R_1 + R_2 \le I(x_1, x_{2c}; y_1) + I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2|x_{2c}) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ + \mathsf{C}^{\mathsf{B}}_{12}$$

$$2R_1 + R_2 \le I(x_1, x_{2c}; y_1, \widehat{y}_2) + I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} \tag{2.21}$$

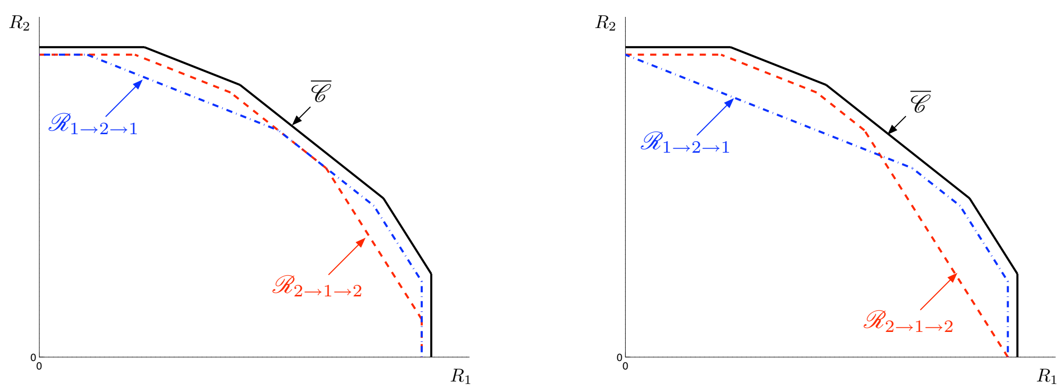
$$R_1 + 2R_2 \le I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}^{\mathsf{B}}_{12} + I(x_2; y_2|x_{1c}, x_{2c}) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+$$

$$R_1 + 2R_2 \le I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{2c}; y_1|x_1) + I(x_{1c}, x_2; y_2|x_{2c}) + I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+$$

$$R_1 + 2R_2 \le I(x_1, x_{2c}; y_1, \widehat{y}_2|x_{1c}) + I(x_{1c}, x_2; y_2) + I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12}$$

$$R_1 + 2R_2 \le I(x_1, x_{2c}; y_1, \widehat{y}_2|x_{1c}) + I(x_{2c}; y_1|x_1) + I(x_{1c}, x_2; y_2|x_{2c}) + I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}^{\mathsf{B}}_{12}$$

We will show that except (2.21), all bounds are within a bounded gap from the corresponding outer bounds in Lemma 2.7. By symmetry, however, one can write down $\mathscr{R}_{1\to 2\to 1}$ and see that the troublesome constraint (2.21) can be compensated by time-sharing with rate points in $\mathscr{R}_{1\to 2\to 1}$. Therefore the resulting $\mathscr{R} := \mathrm{conv}\left\{\mathscr{R}_{2\to 1\to 2} \cup \mathscr{R}_{1\to 2\to 1}\right\}$ is within a bounded gap from the outer bounds in Lemma 2.7. An illustration is provided in Fig. 2.6.



(a) Taking union is required, while time-sharing is not

(b) Time-sharing is required

Figure 2.6: Time-sharing to achieve approximate capacity region

We give the following lemma.

Lemma 2.11 (Rate Region in the Weak Interference Regime)**.**

$$\mathscr{R} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R} \oplus \left([0,2] \times [0,2]\right),$$

in the weak interference regime.

Proof. We need the following claims:

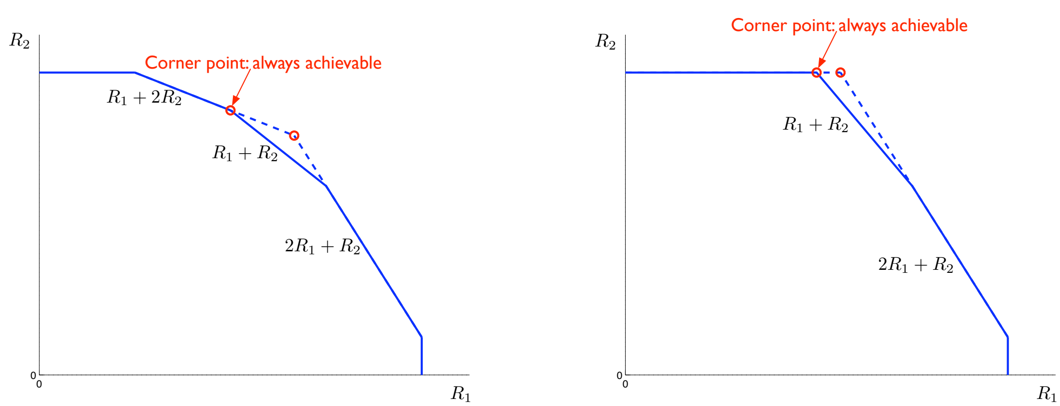
Claim 2.12. *In $\mathscr{R}_{2\to1\to2}$, whenever the $2R_1 + R_2$ bound (2.21) is active,*

(a) if $R_1 + 2R_2$ bounds are active, the corner point where $R_1 + R_2$ bound and $R_1 + 2R_2$ bound intersect can be achieved;

(b) if $R_1 + 2R_2$ bounds are not active, the corner point where $R_1 + R_2$ bound and R_2 bound intersect can be achieved.

Above two situations are illustrated in Fig. 2.7.

Proof. In both situations, we will argue that the value of $R_1 + R_2$ at the intersection of the dashed lines are always greater than the value of $R_1 + R_2$ at the desired corner point. Details are left in Chapter 3.3. □



(a) $R_1 + 2R_2$ bound is active (b) $R_1 + 2R_2$ bound is not active

Figure 2.7: Situations in $\mathscr{R}_{2\to1\to2}$

Therefore, the $2R_1 + R_2$ bound (2.21) and, by symmetry, its corresponding $R_1 + 2R_2$ bound in $\mathscr{R}_{1\to2\to1}$ do not show up in $\mathscr{R} = \text{conv}\{\mathscr{R}_{2\to1\to2} \cup \mathscr{R}_{1\to2\to1}\}$, and $\mathscr{R}$ is within 2 bits per user to the outer bounds in Lemma 2.7. To show this, we first look at the bounds in $\mathscr{R}_{2\to1\to2}$ except (2.21). We claim that

Claim 2.13. *The bounds in $\mathscr{R}_{2\to1\to2}$ except (2.21) satisfy:*

- *R_1 bound is within 2 bits to outer bounds;*
- *R_2 bound is within 2 bits to outer bounds;*
- *$R_1 + R_2$ bound is within 3 bits to outer bounds;*

- *$2R_1 + R_2$ bound is within 4 bits to outer bounds;*
- *$R_1 + 2R_2$ bound is within 5 bits to outer bounds.*

Proof. See Chapter 3.3. □

By symmetry, we obtain similar results for $\mathscr{R}_{1\to2\to1}$, and hence conclude that the bounds in $\mathscr{R}$ satisfies (1) both R_1 and R_2 bounds are within 2 bits; (2) R_1+R_2 bound is within 3 bits; (3) both $2R_1+R_2$ and R_1+2R_2 bound are within 5 bits to their corresponding outer bounds. This completes the proof. □

2.5.4 Mixed interference

In the case $\mathsf{SNR}_1 > \mathsf{INR}_2$ and $\mathsf{SNR}_2 \le \mathsf{INR}_1$, the configuration of superposition coding is to split message m_1 into m_{1c} and m_{1p}, while making the whole m_2 common. We first consider $\mathrm{STG}_{2\to1\to2}$: by Theorem 2.5, rates satisfying the following are achievable,

$$\begin{aligned}
R_{1p} &\le \min\left\{I\left(x_1;y_1|x_{1c},x_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_1;y_1,\widehat{y}_2|x_{1c},x_2\right)\right\} && (2.22)\\
R_2 &\le \min\left\{I\left(x_2;y_1|x_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_2;y_1,\widehat{y}_2|x_1\right)\right\} && (2.23)\\
R_2+R_{1p} &\le \min\left\{I\left(x_2,x_1;y_1|x_{1c}\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_2,x_1;y_1,\widehat{y}_2|x_{1c}\right)\right\}\\
R_{1c}+R_{1p} &\le \min\left\{I\left(x_1;y_1|x_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_1;y_1,\widehat{y}_2|x_2\right)\right\} && (2.24)\\
R_{1c}+R_2+R_{1p} &\le \min\left\{I\left(x_1,x_2;y_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, I\left(x_1,x_2;y_1,\widehat{y}_2\right)\right\}\\
R_{1c} &\le I\left(x_{1c};y_2|x_2\right)+\mathsf{C}^{\mathsf{B}}_{12}\\
R_2 &\le I\left(x_2;y_2|x_{1c}\right)+\mathsf{C}^{\mathsf{B}}_{12}\\
R_2+R_{1c} &\le I\left(x_2,x_{1c};y_2\right)+\mathsf{C}^{\mathsf{B}}_{12},
\end{aligned}$$

where $\xi_1 = 1$ since $\mathsf{SNR}_2 \le \mathsf{INR}_1$.

Again to simplify calculations, note that the right-hand-side of (2.22), (2.23), and (2.24) are at most a bounded constant number of bits greater than their lower bounds $I\left(x_1;y_1|x_{1c},x_2\right)$, $I\left(x_2;y_1|x_1\right)$, and $I\left(x_1;y_1|x_2\right)$ respectively. Therefore, we replace these three constraints by

$$\begin{aligned}
R_{1p} &\le I\left(x_1;y_1|x_{1c},x_2\right),\\
R_2 &\le I\left(x_2;y_1|x_1\right),\\
R_{1c}+R_{1p} &\le I\left(x_1;y_1|x_2\right)
\end{aligned}$$

in the following calculations. Next, rewriting $R_{1p} = R_1 - R_{1c}$, applying Fourier-Motzkin algorithm to eliminate R_{1c}, and removing redundant terms (details omitted here), we obtain an achievable $\mathscr{R}_{2\to1\to2}$, consists of nonnegative (R_1, R_2) satisfying:

$$R_1 \le \min\left\{I\left(x_1;y_1|x_2\right), I\left(x_1;y_1|x_{1c},x_2\right)+I\left(x_{1c};y_2|x_2\right)+\mathsf{C}^{\mathsf{B}}_{12}\right\}$$

$$
\begin{aligned}
R_2 &\le \min\left\{I\left(x_2;y_1|x_1\right), I\left(x_2;y_2|x_{1c}\right)+\mathsf{C}_{12}^{\mathsf{B}}\right\}\\
R_1+R_2 &\le I\left(x_1,x_2;y_1\right)+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\\
R_1+R_2 &\le I\left(x_1,x_2;y_1,\widehat{y}_2\right)\\
R_1+R_2 &\le I\left(x_1;y_1|x_{1c},x_2\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}_{12}^{\mathsf{B}}\\
R_1+R_2 &\le I\left(x_1,x_2;y_1|x_{1c}\right)+I\left(x_{1c};y_2|x_2\right)+\mathsf{C}_{12}^{\mathsf{B}}+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\\
R_1+R_2 &\le I\left(x_1,x_2;y_1,\widehat{y}_2|x_{1c}\right)+I\left(x_{1c};y_2|x_2\right)+\mathsf{C}_{12}^{\mathsf{B}}\\
R_1+2R_2 &\le I\left(x_1,x_2;y_1|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}_{12}^{\mathsf{B}}+\left(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1\right)^+\\
R_1+2R_2 &\le I\left(x_1,x_2;y_1,\widehat{y}_2|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}_{12}^{\mathsf{B}}
\end{aligned}
$$

Comparing $\mathscr{R}_{2\to1\to2}$ with the outer bounds in Lemma 2.7, one can easily conclude that

Lemma 2.14 (Rate Region in the Mixed Interference Regime)**.**

$$\mathscr{R}_{2\to1\to2} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R}_{2\to1\to2} \oplus \left([0,1.5]\times[0,1.5]\right),$$

in the mixed interference regime. Besides, $\mathscr{R}_{2\to1\to2} \subseteq \mathscr{R}$*.*

Proof. We investigate the bounds in $\mathscr{R}_{2\to1\to2}$ and claim that

Claim 2.15. *The bounds in* $\mathscr{R}_{2\to1\to2}$ *satisfy:*

- R_1 *bound is within* 1 *bit to outer bounds;*
- R_2 *bound is within* 1 *bit to outer bounds;*
- R_1+R_2 *bound is within* 3 *bits to outer bounds;*
- R_1+2R_2 *bound is within* 3 *bits to outer bounds.*

Proof. See Chapter 3.3 □

This completes the proof. □

2.5.5 Strong interference

In the case $\mathsf{SNR}_1 \le \mathsf{INR}_2$ and $\mathsf{SNR}_2 \le \mathsf{INR}_1$, it turns out that a one-round strategy $\mathrm{STG}_{\mathrm{OneRound}}$ described below suffices to achieve capacity to within a bounded gap. The transmission scheme is the same as that described in Section 2.4.1. The difference is that, both receivers quantize-and-bins their received signals and decode with the help from the side information, as described in Section 2.4.2. It is called one-round since both receivers decode after one-round exchange of informaion. Below is the coding theorem for this strategy:

Theorem 2.16. *The rate tuple* $(R_{1c}, R_{2c}, R_{1p}, R_{2p})$ *satisfying the following constraints are achievable for* $\text{STG}_{\text{OneRound}}$*:*

Constraints at receiver 1*:*

$$\begin{aligned}
R_{1p} &\le \min\left\{ I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right)\right\}\\
R_{2c} + R_{1p} &\le \min\left\{ I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right)\right\}\\
R_{1c} + R_{1p} &\le \min\left\{ I\left(x_1; y_1 | x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{2c}\right)\right\}\\
R_{1c} + R_{2c} + R_{1p} &\le \min\left\{ I\left(x_1, x_{2c}; y_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right)\right\}
\end{aligned}$$

Constraints at receiver 2*: Above constraints with index "1" and "2" exchanged.*

Proof. The proof follows the same line as the proof of Theorem 2.5. There is no rate constraint for R_{jc} at receiver i for $(i, j) = (1, 2)$ or $(2, 1)$, since decoding m_{jc} incorrectly at receiver i does not account for an error. □

Now, in the strong interference regime, the configuration of superposition coding is to make the whole message m_i common for both users $i = 1, 2$; in a word, there is no superposition eventually. One-round strategy $\text{STG}_{\text{OneRound}}$ yields achievable rate region $\mathscr{R}_{\text{OneRound}}$, which consists of nonnegative (R_1, R_2) satisfying

$$R_2 \le \min\left\{ I\left(x_2; y_1 | x_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_2; y_1, \widehat{y}_2 | x_1\right)\right\} \tag{2.25}$$
$$R_1 \le \min\left\{ I\left(x_1; y_1 | x_2\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_2\right)\right\} \tag{2.26}$$
$$R_1 + R_2 \le \min\left\{ I\left(x_1, x_2; y_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1, x_2; y_1, \widehat{y}_2\right)\right\} \tag{2.27}$$
$$R_1 \le \min\left\{ I\left(x_1; y_2 | x_2\right) + (\mathsf{C}^{\mathsf{B}}_{12} - \xi_2)^+, I\left(x_1; y_2, \widehat{y}_1 | x_2\right)\right\} \tag{2.28}$$
$$R_2 \le \min\left\{ I\left(x_2; y_2 | x_1\right) + (\mathsf{C}^{\mathsf{B}}_{12} - \xi_2)^+, I\left(x_2; y_2, \widehat{y}_1 | x_1\right)\right\} \tag{2.29}$$
$$R_2 + R_1 \le \min\left\{ I\left(x_2, x_1; y_2\right) + (\mathsf{C}^{\mathsf{B}}_{12} - \xi_2)^+, I\left(x_2, x_1; y_2, \widehat{y}_1\right)\right\}, \tag{2.30}$$

where $\xi_i = 1$, for both $i = 1, 2$.

Comparing $\mathscr{R}_{\text{OneRound}}$ with the outer bounds in Lemma 2.7, one can easily conclude that

Lemma 2.17 (Rate Region in the Strong Interference Regime)**.**

$$\mathscr{R}_{\text{OneRound}} \subseteq \mathscr{C} \subseteq \overline{\mathscr{C}} \subseteq \mathscr{R}_{\text{OneRound}} \oplus \left([0,1] \times [0,1]\right),$$

in the strong interference regime. Besides, $\mathscr{R}_{\text{OneRound}} \subseteq \mathscr{R}$*.*

Proof. We investigate the bounds in $\mathscr{R}_{\text{OneRound}}$ and claim that:

Claim 2.18. *The bounds in* $\mathscr{R}_{\text{OneRound}}$ *satisfy:*

- R_1 *bound is within* 1 *bit to outer bounds;*

- *R_2 bound is within 1 bit to outer bounds;*
- *$R_1 + R_2$ bound is within 2 bits to outer bounds.*

Proof. See Chapter 3.3. □

This completes the proof. □

2.5.6 Approximate Capacity of Compound MAC with Conferencing Receivers

As a side-product of this work, we characterize the capacity region of the *compound multiple access channel with conferencing receivers* (CMAC-CR) to within 1 bit. The channel is defined as follows.

Definition 2.19. *A compound multiple access channel with conferencing receivers (CMAC-CR), is a channel with the same set-up as depicted in Fig. 4.1., while both receivers aim to decode both m_1 and m_2.*

We give straightforward cut-set upper bounds as follows:

Lemma 2.20. *If (R_1, R_2) is achievable, it must satisfy the following constraints:*

$$
\begin{aligned}
R_1 &\le \min\left\{\log(1+\mathsf{SNR}_1)+\mathsf{C}^{\mathsf{B}}_{21}, \log(1+\mathsf{INR}_2)+\mathsf{C}^{\mathsf{B}}_{12}, \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)\right\}\\
R_2 &\le \min\left\{\log(1+\mathsf{SNR}_2)+\mathsf{C}^{\mathsf{B}}_{12}, \log(1+\mathsf{INR}_1)+\mathsf{C}^{\mathsf{B}}_{21}, \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_1\right)\right\}\\
R_1+R_2 &\le \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+\mathsf{C}^{\mathsf{B}}_{21}\\
R_1+R_2 &\le \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)+\mathsf{C}^{\mathsf{B}}_{12}\\
R_1+R_2 &\le \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right).
\end{aligned}
$$

Proof. These are straightforward cut-set bounds. We omit the details here. □

For achievability, we adapt the one-round scheme proposed above with no superposition coding at transmitters. Therefore, the rate region is exactly the same as (2.25)-(2.30). Hence, we conclude that

Theorem 2.21 (Within 1 bit to CMAC-CR Capacity Region)**.** *The scheme achieves the capacity of compound MAC with conferencing receivers to within 1 bit.*

Proof. Following the same line in the proof of Lemma 2.17, we can conclude that the bounds in $\mathscr{R}_{\text{OneRound}}$ satisfy:

- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;

- $R_1 + R_2$ bound is within 1 bit to outer bounds.

This completes the proof. □

This result implies that for the Gaussian compound MAC with conferencing receivers, a simple one-round strategy suffices to achieve the capacity region to within 1 bit universally, regardless of channel parameters.

2.6 One-Round Strategy versus Two-Round Strategy

In Section 2.5 we show that for the two-user Gaussian interference channel with conferencing receivers, the two-round strategy proposed in Section 2.4 along with time-sharing achieves the capacity region to within 2 bits universally. One of the drawbacks of the two-round strategy, however, is the decoding latency. The quantize-binning receiver cannot proceed to decoding until the other receiver decodes and forwards the bin indices back. The decoding latency is two times the block length, which can be large. To avoid such large delay, fortunately in some cases, the one-round strategy $\text{STG}_{\text{OneRound}}$ described in Section 2.5.5 suffices. One of such cases is the strong interference regime. This can be easily justified in the corresponding linear deterministic channel (LDC). At strong interference, all transmitted signals in the LDC are common. There is no useful information lies below the noise level since the signal is corrupted by the noise. Hence, quantize-binning at the noise level is sufficient to convey the useful information.

Another such cases is the symmetric set-up, where $\mathsf{SNR} = \mathsf{SNR}_1 = \mathsf{SNR}_2$, $\mathsf{INR} = \mathsf{INR}_1 = \mathsf{INR}_2$, and $\mathsf{C}^{\mathsf{B}} = \mathsf{C}^{\mathsf{B}}_{12} = \mathsf{C}^{\mathsf{B}}_{21}$.

For the symmetric set-up, a natural performance measure is the symmetric capacity, defined as follows:

Definition 2.22 (Symmetric Capacity)**.**

$$C_{\text{sym}} := \sup \{R : (R, R) \in \mathscr{C}\}.$$

It turns out that the one-round strategy suffices to achieve C_{sym} to within a bounded gap.

Theorem 2.23 (Bounded Gap to the Symmetric Capacity)**.**
The one-round strategy $\text{STG}_{\text{OneRound}}$ *can achieve the symmtric capacity to within 3 bits.*

Proof. See Chapter 3.4. □

The justification in the corresponding LDC is again simple. Since the performance measure in which we are interested is the symmetric capacity, we can without loss of generality assume that both transmitters are transmitting at full private rate, that is, the entropy of each user's private signals is equal to the number of levels below the private signal level. Therefore at each receiver, there is no useful information below the private signal level, and quantize-binning at the private signal level suffices to convey the useful information.

2.7 Generalized Degrees of Freedom Characterization

With the characterization of the capacity region to within a bounded gap, we attempt to answer the original fundamental question: how much interference can one bit of receiver cooperation mitigate? For simplicity, we consider the symmetric set-up.

By Lemma 2.7 and Theorem 2.10, we have the characterization of the symmetric capacity to within 2 bits:

Corollary 2.24 (Approximate Symmetric Capacity)**.** *Let* $\overline{C}_{\text{sym}}$ *be the minimum of the below four terms:*

$$\begin{aligned}
&\log(1+\mathsf{SNR})+\min\left\{\mathsf{C}^{\mathsf{B}},\log\left(1+\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right)\right\},\\
&\log\left(1+\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\mathsf{C}^{\mathsf{B}},\\
&\frac{1}{2}\log\left(1+\mathsf{SNR}+\mathsf{INR}\right)+\frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\frac{1}{2}\mathsf{C}^{\mathsf{B}},\\
&\frac{1}{2}\log\left(1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2\right).
\end{aligned}$$

Then, $\overline{C}_{\text{sym}}-2\le C_{\text{sym}}\le\overline{C}_{\text{sym}}$.

2.7.1 Generalized Degrees of Freedom

To study the behavior of the system performance in the linear region, we use the notion of *generalized degrees of freedom* (g.d.o.f.), which is originally proposed in [5]. A natural extension from the definition in [5] would be the following: let

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log\mathsf{INR}}{\log\mathsf{SNR}}=\alpha;\ \lim_{\mathsf{SNR}\to\infty}\frac{\mathsf{C}^{\mathsf{B}}}{\log\mathsf{SNR}}=\kappa,$$

and define the number of generalized degrees of freedom per user as

$$d:=\lim_{\substack{\text{fix }\alpha,\kappa\\ \mathsf{SNR}\to\infty}}\frac{C_{\text{sym}}}{\log\mathsf{SNR}},\tag{2.31}$$

if the limit exists. With fixed α and κ, however, there are certain channel realizations under which (2.31) has different values and hence the limit does not exist. This happens when $\alpha=1$, where the phases of the channel gains matter both in inner and outer bounds. In particular, its value can depend on whether the system MIMO matrix is well-conditioned or not.

From the above discussion we see that the limit does not exist, since for different channel phases and different INR settings the value of (2.31) may be different. The reason is that, the original notion proposed in [5] cannot capture the impact of *phases* in MIMO situations, while from Lemma 2.7 and Theorem 2.10, or Corollary 2.24, we see that our results depend on phases heavily, if the receiver-cooperative link capacity C^{B} is so large that MIMO sum-rate cut-set bound becomes active. Therefore, instead of claiming that the limit (2.31) exists for *all* channel realizations, we pose a reasonable distribution, namely, i.i.d. uniform distribution, on the phases, show that the limit exists *almost surely*, and define the limit to be the number of *generalized degrees of freedom* per user.

Lemma 2.25. *Let*

$$|h_{ij}| = g_{ij},\ \angle h_{ij} = \Theta_{ij},\ \forall i, j \in \{1, 2\},$$

where g_{ij}'s are deterministic and Θ_{ij}'s are i.i.d. uniformly distributed over $[0, 2\pi]$. Then the limit (2.31) exists almost surely, and is defined as the number of generalized degrees of freedom (per user) in the system.

Proof. We leave the proof in Chapter 3.5. □

Now that the number of g.d.o.f. is well-defined, we can give the following theorem:

Theorem 2.26 (Generalized Degrees of Freedom Per User)**.** *We have a direct consequence from Corollary 2.24:*

For $0 \le \alpha < 1$,

$$d = \min\left\{1,\ \max(\alpha, 1-\alpha) + \kappa,\ 1 - \frac{\alpha}{2} + \frac{\kappa}{2}\right\}.$$

For $\alpha \ge 1$,

$$d = \min\left\{\alpha,\ 1 + \kappa,\ \frac{\alpha}{2} + \frac{\kappa}{2}\right\}.$$

Numerical plots for g.d.o.f. are given in Fig. 2.8. We observe that at different values of α, the gain from cooperation varies. By investigating the g.d.o.f., we conclude that at high SNR, when INR is below 50% of SNR in dB scale, one-bit cooperation per direction buys roughly one-bit gain per user until full receiver cooperation performance is reached, while when INR is between 67% and 200% of SNR in dB scale, one-bit cooperation per direction buys roughly half-bit gain per user until saturation.

2.7.2 Gain from Limited Receiver Cooperation

The fundamental behavior of the gain from receiver cooperation is explained in the rest of this section, by looking at two particular points: $\alpha = \frac{1}{2}$ and $\alpha = \frac{2}{3}$. Furthermore, we use the linear deterministic channel (LDC) for illustration.

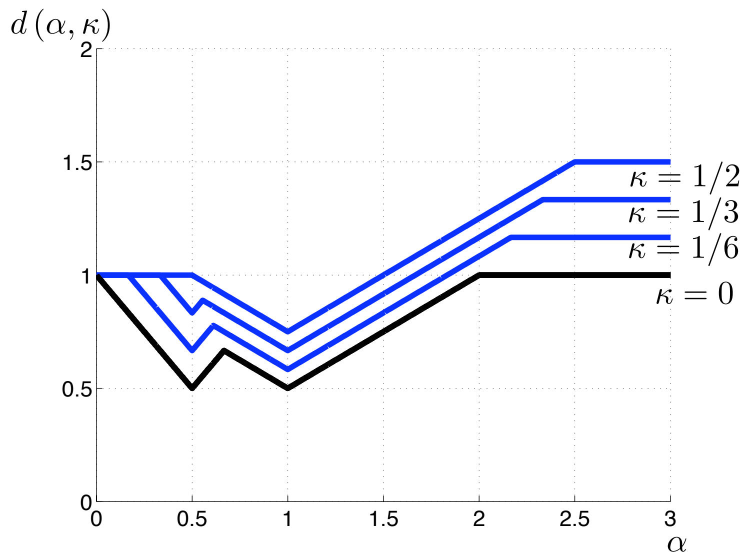


Figure 2.8: Generalized Degrees of Freedom

At $\alpha = \frac{1}{2}$, the plot of d versus κ is given in Fig. 2.9.(a). The slope is 1 until full receiver cooperation performance is reached, implying that one-bit cooperation buys one over-the-air bit per user. We look at a particular point $\kappa = \frac{1}{4}$ and use its corresponding LDC (Fig. 2.9.(b)) to provide insights. Note that 1 bit in the LDC corresponds to $\frac{1}{4}\log \mathsf{SNR}$ in the Gaussian channel, and since $\mathsf{C}^{\mathsf{B}} \approx \frac{1}{4}\log \mathsf{SNR}$, in the corresponding LDC each receiver is able to sent one-bit information to the other. Without cooperation, the optimal way is to turn on bits not causing interference, that is, the *private* bits a_3, a_4, b_3, b_4. We cannot turn on more bits without cooperation since it causes collisions, for example, at the fourth level of receiver 2 if we turn on a_2 bit. Now with receiver cooperation, we want to support two more bits a_2, b_2. Note that prior to turning on a_2, b_2, there are "holes" left in receiver signal spaces, and turning on each of these bits only causes one collision at one receiver. Therefore, we need 1 bit in each direction to resolve the collision at each receiver. We can achieve 3 bits per user in the corresponding LDC and $d = \frac{3}{4}$ in the Gaussian channel. We cannot turn on more bits in the LDC since it causes collisions while no cooperation capability is left.

At $\alpha = \frac{2}{3}$, the plot of d versus κ is given in Fig. 2.9.(c). The slope is $\frac{1}{2}$ until full receiver cooperation performance is reached, implying that two-bit cooperation buys one over-the-air bit per user. We look at a particular point $\kappa = \frac{1}{3}$ and use its corresponding LDC (Fig. 2.9.(d)) to provide insights. Note that now 1 bit in the LDC corresponds to $\frac{1}{3}\log \mathsf{SNR}$ in the Gaussian channel, and since $\mathsf{C}^{\mathsf{B}} \approx \frac{1}{3}\log \mathsf{SNR}$, in the corresponding LDC each receiver is able to sent one-bit information to the other. Without cooperation, the optimal way is to turn on bits a_1, a_3, b_1, b_3. We cannot turn on more bits without cooperation since it causes collisions, for example, at the second level of receiver 2 if we turn on a_2 bit. Now with

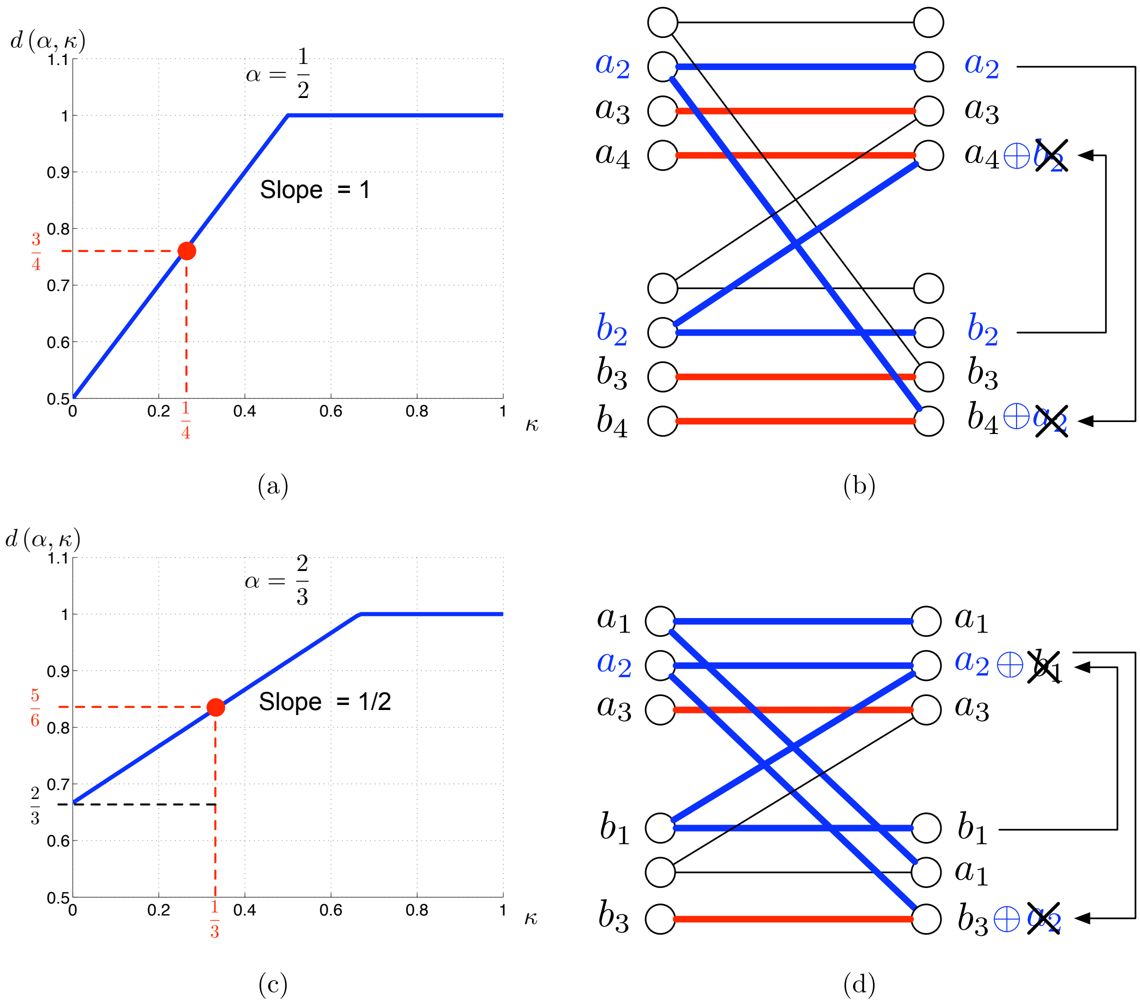


Figure 2.9: Gain from Cooperation

receiver cooperation, we want to support one more bit a_2. Note that prior to turning on a_2, there are no "holes" left in receiver signal spaces, and turning on a_2 causes collisions at *both* receivers. Therefore, we need 2 bits in total to resolve collisions at both receivers. We can achieve 5 bits in total in the corresponding LDC and $d = \frac{5}{6}$ in the Gaussian channel. We cannot turn on more bits in the LDC since it causes collision while no cooperation capability is left.

From above examples and illustrations, we see that whether *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit* depends on whether there are "holes" in receiver signal spaces before increasing data rates. The "holes" play a central role not only in why conventional compress-forward is suboptimal in certain regimes, as mentioned in Section 2.3.2, but also in the fundamental behavior of the gain from receiver cooperation. We notice that in [25], there is a similar behavior about the gain from in-band receiver cooperation as discussed in Section III-B of [25]. We conjecture that the behavior can be explained via the concept of "holes" as well.

2.7.3 Comparison with Suboptimal Strategies

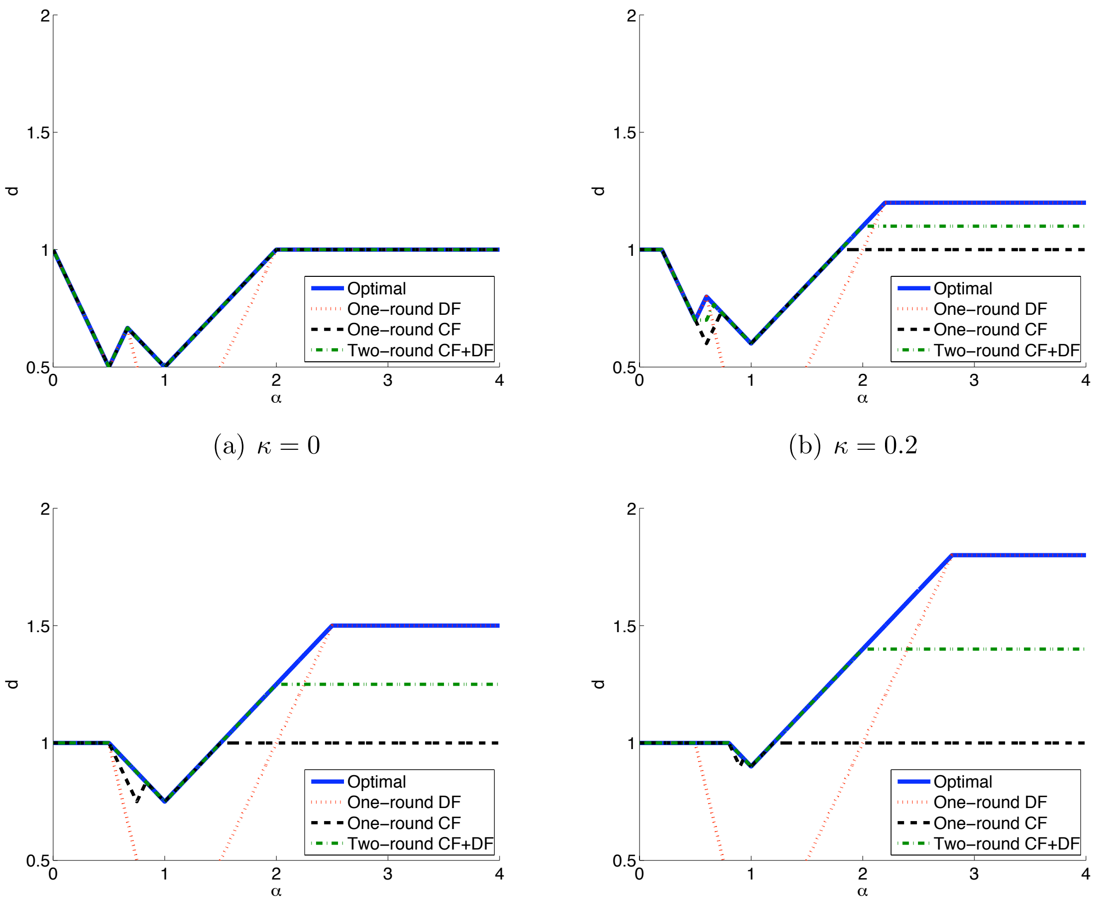
Pointed out by the motivating examples in Section 2.3.2, conventional compress-forward and decode-forward are not good for receiver cooperation to mitigate interference in certain regimes, which are used in [22] and [23]. These suboptimal schemes include:

(1) One-round compress-forward (CF) strategy: the conventional compress-forward is used for the two receivers to first exchange information and then decode.

(2) One-round decode-forward (DF) strategy: at the first stage both receivers decode one of the common messages with stronger signal strength without help from the receiver-cooperative links, by treating other signals as noise. Both then bin-and-forward the decoded information to each other. At the second stage, both receivers make use of the bin index send over receiver-cooperative links to decode and enhance the rate.

(3) Two-round CF+DF strategy: at the first stage one of the receivers, say, receiver 1, compresses its received signal and forwards it to the other receiver. At the second stage, receiver 2 decodes with the side information received at the first round, and then bin-and-forwards the decoded information to receiver 1. Then at the third stage receiver 1 decodes with the help from receiver-cooperative links.

Comparisons of these strategies in terms of the number of generalized degrees of freedom for different scaling exponents α of $\log \mathsf{INR}$ and κ of C^{B} are depicted in Fig. 2.10. None of them achieves the optimal g.d.o.f. universally. Note that although the two-round CF+DF strategy outperforms one-round CF/DF strategies, it cannot achieve the optimal number of g.d.o.f. for all α's and κ's. By Theorem 2.23, the one-round strategy based on our cooperative protocol, on the other hand, is sufficient to achieve the symmetric capacity to within 3 bits universally and hence achieves the optimal number of g.d.o.f. for all α's and κ's.



(a) $\kappa = 0$ (b) $\kappa = 0.2$

(c) $\kappa = 0.5$ (d) $\kappa = 0.8$

Figure 2.10: Number of Generalized Degrees of Freedom

Chapter 3

Proofs of Part I

In this chapter we fill in the details of various proofs mentioned in the previous chapter.

3.1 Proof of Theorem 2.5

We will first describe the strategy in detail and analyze the error probability rigorously.

3.1.1 Description of the Strategy

In the following, consider all $i,j \in \{1,2\}$ and $i \neq j$.

Codebook generation:

Transmitter i splits its message $m_i \to (m_{ic}, m_{ip})$. Consider block length-N encoding. First we generate $2^{NR_{ic}}$ common codewords $\{x_{ic}^N(m_{ic}),\ 1 \le m_{ic} \le 2^{NR_{ic}}\}$, according to distribution $p\left(x_{ic}^N\right) = \prod_{n=1}^N p\big(x_{ic}[n]\big)$ with $x_{ic}[n] \sim \mathcal{CN}(0, Q_{ic})$ for all n. Then for each common codeword $x_{ic}^N(m_{ic})$ serving as a *cloud center*, we generate $2^{NR_{ip}}$ codewords $\{x_i^N(m_{ic}, m_{ip}),\ 1 \le m_{ip} \le 2^{NR_{ip}}\}$, according to conditional distribution $p\left(x_i^N | x_{ic}^N\right) = \prod_{n=1}^N p\big(x_i[n] | x_{ic}[n]\big)$ such that for all n, $x_i[n] = x_{ic}[n] + x_{ip}[n]$, where $x_{ip}[n] \sim \mathcal{CN}(0, Q_{ip})$ and independent of everything else. The power split configuration is such that $Q_{ip} + Q_{ic} = 1$, $\mathsf{INR}_{jp} := Q_{ip}|h_{ji}|^2 \le 1$ if $\mathsf{SNR}_i > \mathsf{INR}_j$, and no such split if $\mathsf{SNR}_i \le \mathsf{INR}_j$. Hence, $Q_{ip} = \min\left\{1, \frac{1}{\mathsf{INR}_j}\right\}$ if $\mathsf{SNR}_i > \mathsf{INR}_j$, and $Q_{ip} = 0$ otherwise.

For receiver 2 serving as relay, it generates a quantization codebook $\widehat{\mathscr{Y}_2}$, of size $\left|\widehat{\mathscr{Y}_2}\right| = 2^{N\widehat{R}_2}$, randomly according to marginal distribution $p(\widehat{y}_2^N)$, marginalized over joint distribution $p\left(y_2^N, x_{1c}^N, x_1^N, x_{2c}^N\right) p(\widehat{y}_2^N | y_2^N, x_{1c}^N, x_1^N, x_{2c}^N)$, where

$$p(\widehat{y}_2^N | y_2^N, x_{1c}^N, x_1^N, x_{2c}^N) = \prod_{n=1}^N p\Big(\widehat{y}_2[n] \big| y_2[n], x_{1c}[n], x_1[n], x_{2c}[n]\Big).$$

The conditional distribution is such that for all n, $\widehat{y}_2[n] = y_2[n] + \widehat{z}_2[n]$, where $\widehat{z}_2[n] \sim \mathcal{CN}(0, \Delta_2)$, independent of everything else. Parameters $\widehat{R}_2$ and Δ_2 are to be specified later. For each element in codebook $\widehat{\mathscr{Y}_2}$, map it into $\{1, \ldots, 2^{N\mathsf{C}^{\mathsf{B}}_{21}}\}$ through a uniformly generated random mapping $b_2 : \widehat{\mathscr{Y}_2} \to \{1, \ldots, 2^{N\mathsf{C}^{\mathsf{B}}_{21}}\}, \widehat{y}_2^N \mapsto l_{21}$ (*binning*).

For receiver 1 serving as relay, it generates two binning functions $b_1^{(1c)}$ and $b_1^{(2c)}$ independently according to uniform distributions, such that the message set $\{1 \le m_{ic} \le 2^{NR_{ic}}\}$ is partitioned into $2^{\lambda_1^{(ic)}N\mathsf{C}^{\mathsf{B}}_{12}}$ bins, for $i = 1, 2$, where $0 \le \lambda_1^{(ic)} \le 1$, $\lambda_1^{(1c)} + \lambda_1^{(2c)} = 1$, and

$$
\begin{aligned}
b_1^{(ic)} : \{1, \ldots, 2^{NR_{ic}}\} &\to \{1, \ldots, 2^{\lambda_1^{(ic)}N\mathsf{C}^{\mathsf{B}}_{12}}\}, \\
m_{ic} &\mapsto l_{12}^{(ic)} \in \{1, \ldots, 2^{\lambda_1^{(ic)}N\mathsf{C}^{\mathsf{B}}_{12}}\}.
\end{aligned}
$$

The superscript notation "(ic)" denotes which message set is partitioned into bins, while the subscript "1" denotes the binning procedure is at receiver 1.

Encoding:

Transmitter i sends out signals according to its message and the codebook. Receiver 2, serving as relay, chooses the quantization codeword which is jointly typical with y_2^N (if there is more than one, it chooses the one with the smallest index), and then sends out the bin index l_{21} for the quantization codeword. After decoding (m_{1c}, m_{1p}, m_{2c}) (to be specified below), receiver 1 sends out bin indices $\left(l_{12}^{(1c)}, l_{12}^{(2c)}\right)$ according to binning functions $\left(b_1^{(1c)}, b_1^{(2c)}\right)$.

Decoding at receiver 1:

To draw comparison with the decoding procedure in the conventional compress-forward, the above decoding can be interpreted as a two-stage procedure as follows. It first constructs a *list* of message triples (both users' common messages and its own private message), each element of which indices a codeword triple that is jointly typical with its received signal from the transmitter-receiver link. Then, for each message triple in this list, it constructs an *ambiguity set* of quantization codewords, each element of which is jointly typical with the codeword triple and the received signal. Finally, it searches through all ambiguity sets and finds one that contains a quantization codeword with the same bin index it received. If there is no such unique ambiguity set, it declares an error. The two-stage interpretation is illustrated in Fig. 3.1.

To be specific, upon receiving signal y_1 and receiver-cooperative side information l_{21}, receiver i constructs a list of candidates $L_i(y_i^N)$, defined as

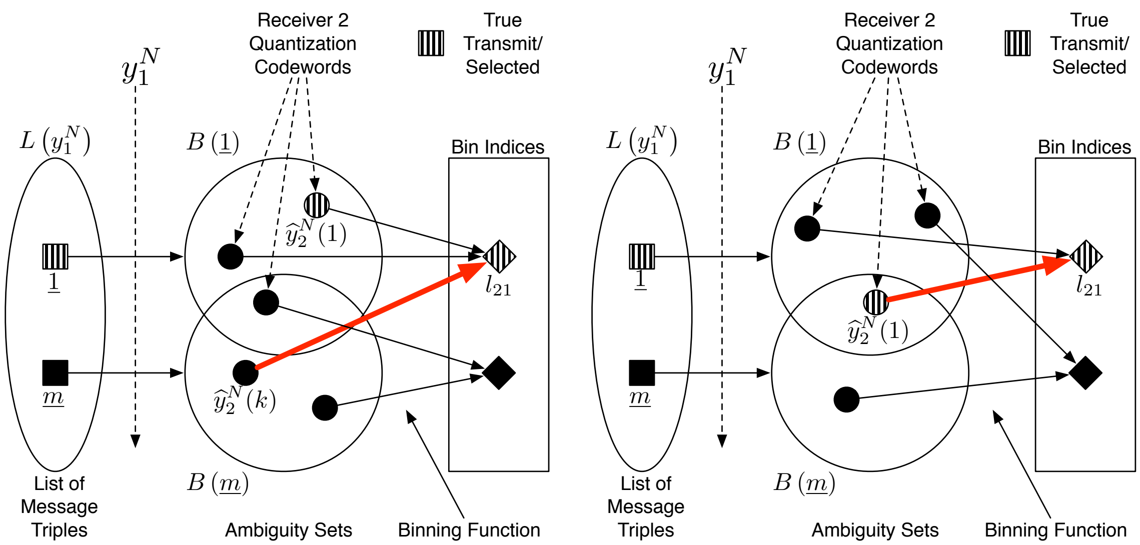
$$
L(y_1^N) := \left\{ \underline{m} := (m_{1c}, m_{1p}, m_{2c}) \middle| \left(x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c}), y_1^N\right) \in A_\epsilon^{(N)} \right\},
$$

where $A_\epsilon^{(N)}$ denotes the set of jointly ϵ-typical N-sequences, correspondingly [37].

For each element $\underline{m} \in L(y_1^N)$, construct an ambiguity set of quantization codewords $B(\underline{m})$, defined as

$$
B(\underline{m}) := \left\{ \widehat{y}_2^N \in \widehat{\mathscr{Y}_2} \,\middle|\, \left(\widehat{y}_2^N, x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c}), y_i^N\right) \in A_\epsilon^{(N)} \right\}.
$$

Declare the transmitted message is $\underline{\widehat{m}}$ if there exists an unique $\underline{\widehat{m}}$ such that $\exists\ \widehat{y}_2^N \in B(\underline{\widehat{m}})$ with $b_2(\widehat{y}_2^N) = l_{21}$. Otherwise, declare an error.



(a) Error Event (1) (b) Error Event (2)

Figure 3.1: Decoding at Receiver 1 and Error Events

Decoding at receiver 2:

After receiving bin indices $\left(l_{12}^{(1c)}, l_{12}^{(2c)}\right)$, receiver 2 searches for an unique message triple (m_{2c}, m_{2p}, m_{1c}) such that $\left(x_{2c}^N(m_{2c}), x_2^N(m_{2c}, m_{2p}), x_{1c}^N(m_{1c}), y_2^N\right) \in A_\epsilon^{(N)}$, and $b_1^{(ic)}(m_{ic}) = l_{12}^{(ic)}$, for $i = 1, 2$. If there is no such unique triple, it declares an error.

3.1.2 Analysis

Error probability analysis at receiver 1:

Without loss of generality, assume that all transmitted messages are 1's. For simplicity, we first focus on the case where receiver 1 aims to decode while receiver 2 serves as a relay to help it decode.

At receiver 1, due to law of large numbers, the probability that the truly transmitted $\underline{1} := (m_{1c} = 1, m_{2c} = 1, m_{1p} = 1) \notin L(y_1^N)$ goes to zero as $N \to \infty$. Besides, the probability that $B(\underline{1})$ does not contain the truly selected $\widehat{y}_2^N$ is also negligible when N is sufficiently large. Consider the following error events:

First, there is no quantization codeword jointly typical with received signals. This probability goes to zero as $N \to \infty$ if $\widehat{R}_2 \geq I(\widehat{y}_2; y_2)$, which is a known result in the source coding literature.

Second, there exists $\underline{m} \neq \underline{1}$ such that both of them are in the candidate list $L(y_1^N)$, and the ambiguity set $B(\underline{m})$ contains some quantization codeword $\widehat{y}_2^N$ with bin index $b_2(\widehat{y}_2^N) = l_{21}$. This event can further be distinguished into two cases: First, this $\widehat{y}_2^N \in B(\underline{m})$ is not the actual selected quantization codeword (illustrated in Fig. 3.1.(a)); second, this $\widehat{y}_2^N \in B(\underline{m})$ is indeed the selected quantization codeword (illustrated in Fig. 3.1.(b)). In the following we analyze the error probability of these two typical error events.

Again, refer to Fig. 3.1. for illustration. Define error events as follows: for any nonempty $S \subseteq \{1c, 1p, 2c\}$,

$E_S^{(1)}$:= the event that there exists some $\underline{m} \neq \underline{1}$, (where $m_s \neq 1, \forall s \in S$ and $m_s = 1, \forall s \notin S$), such that $\underline{m} \in L(y_1^N)$ and $B(\underline{m})$ contains some $\widehat{y}_2^N(k)$, $k \in \{1, 2, \ldots, 2^{N\widehat{R}_2}\}$ with $b_2(\widehat{y}_2^N(k)) = l_{21}$. Note: this $\widehat{y}_2^N(k)$ is not the truly selected quantization codeword $\widehat{y}_2^N(1)$.

$E_S^{(2)}$:= the event that there exists some $\underline{m} \neq \underline{1}$, (where $m_s \neq 1, \forall s \in S$ and $m_s = 1, \forall s \notin S$), such that $\underline{m} \in L(y_1^N)$ and $B(\underline{m})$ contains $\widehat{y}_2^N(1)$.

Probability of $E_S^{(1)}$

Consider the probability of the error event $E_S^{(1)}$: it can be upper bounded as follows:

$$\begin{aligned}
\Pr\left\{E_S^{(1)}\right\} &\leq \sum_{\substack{\underline{m}: m_s \neq 1,\\ \forall s \in S}} \sum_{k \neq 1} \Pr\left\{\underline{m} \in L(y_1^N), \widehat{y}_2^N(k) \in B(\underline{m}), b_2(\widehat{y}_2^N(k)) = l_{21}\right\} \\
&= \sum_{\substack{\underline{m}: m_s \neq 1,\\ \forall s \in S}} \sum_{k \neq 1} \Pr\left\{\left(\widehat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}, b_2(\widehat{y}_2^N(k)) = l_{21}\right\} \\
&\overset{\text{(a)}}{=} 2^{-N\mathsf{C}_{21}^{\mathsf{B}}} \sum_{\substack{\underline{m}: m_s \neq 1,\\ \forall s \in S}} \sum_{k \neq 1} \Pr\left\{\left(\widehat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\} \\
&\leq 2^{N\left(\sum_{s \in S} R_s\right)} 2^{-N\mathsf{C}_{21}^{\mathsf{B}}} \sum_{k \neq 1} \Pr\left\{\left(\widehat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\};
\end{aligned}$$

where (a) is due to the independent uniform binning.

For notational convenience we use $\underline{x}^N(\underline{m})$ to denote the vector of codewords corresponding to message $\underline{m}$, that is, $\left(x_{1c}^N(m_{1c}), x_1^N(m_{1c}, m_{1p}), x_{2c}^N(m_{2c})\right)$.

Note that for $k \neq 1$, $\widehat{y}_2^N(k)$ is independent of $\left(\underline{x}^N(\underline{m}), y_1^N\right)$. We then make use of Theorem 15.2.2 in [37], which upper bounds the volume of conditional joint ϵ-typical set $A_\epsilon^{(N)}\left(\widehat{y}_2 \middle| \underline{x}^N, y_1^N\right)$ given that $\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}$, to further upper bound

$$\sum_{k \neq 1} \Pr\left\{\left(\widehat{y}_2^N(k), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\} \leq 2^{N\widehat{R}_2} \int_{\left(\widehat{y}_2^N, \underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(\widehat{y}_2^N\right) p\left(\underline{x}^N, y_1^N\right) d\widehat{y}_2^N d\underline{x}^N dy_1^N$$

$$
\begin{aligned}
&\leq 2^{N\widehat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \int_{\widehat{y}_2^N \in A_\epsilon^{(N)}\left(\widehat{y}_2 \middle| \underline{x}^N, y_1^N\right)} p\left(\widehat{y}_2^N\right) dy_2^N \\
&\leq 2^{N\widehat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \int_{\widehat{y}_2^N \in A_\epsilon^{(N)}\left(\widehat{y}_2 \middle| \underline{x}^N, y_1^N\right)} 2^{-N(h(\widehat{y}_2)-\epsilon)} dy_2^N \\
&\overset{(b)}{\leq} 2^{N(h(\widehat{y}_2|x_{1c},x_1,x_{2c},y_1)+2\epsilon)} \cdot 2^{-N(h(\widehat{y}_2)-\epsilon)} \cdot 2^{N\widehat{R}_2} \int_{\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \\
&= 2^{N\widehat{R}_2} 2^{-N(I(\widehat{y}_2;x_{1c},x_1,x_{2c},y_1)-3\epsilon)} \int_{\left(\underline{x}^N, y_1^N\right) \in A_\epsilon^{(N)}} p\left(\underline{x}^N, y_1^N\right) d\underline{x}^N dy_1^N \\
&= \Pr\left\{\underline{m} \in L(y_1^N)\right\} \cdot 2^{N\widehat{R}_2} 2^{-N(I(\widehat{y}_2;x_{1c},x_1,x_{2c},y_1)-3\epsilon)},
\end{aligned}
$$

where (b) is due to Theorem 15.2.2 in [37]. Besides, according to the results in [38],

$$
\Pr\left\{\underline{m} \in L(y_1^N)\right\} \leq \begin{cases}
2^{-N(I(x_1;y_1|x_{1c},x_{2c})-\epsilon')} & S = \{1p\} \\
2^{-N(I(x_1;y_1|x_{2c})-\epsilon')} & S = \{1c\} \\
2^{-N(I(x_{2c};y_1|x_1)-\epsilon')} & S = \{2c\} \\
2^{-N(I(x_{2c},x_1;y_1|x_{1c})-\epsilon')} & S = \{1p, 2c\} \\
2^{-N(I(x_1;y_1|x_{2c})-\epsilon')} & S = \{1p, 1c\} \\
2^{-N(I(x_1,x_{2c};y_1)-\epsilon')} & S = \{2c, 1c\} \\
2^{-N(I(x_1,x_{2c};y_1)-\epsilon')} & S = \{1p, 2c, 1c\}
\end{cases},
$$

where $\epsilon' = 4\epsilon$. Note that unlike in the interference channel without cooperation as in [38], here we require receiver 1 to decode m_{2c} correctly. Hence, the event when $S = \{2c\}$ does cause an error. Therefore, the probability of the first kind of error event vanishes as $N \to \infty$ if

$$
\begin{aligned}
R_{1p} &\leq I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + \varphi \\
R_{2c} &\leq I\left(x_{2c}; y_1 | x_1\right) + \varphi \\
R_{2c} + R_{1p} &\leq I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + \varphi \\
R_{1c} + R_{1p} &\leq I\left(x_1; y_1 | x_{2c}\right) + \varphi \\
R_{1c} + R_{2c} + R_{1p} &\leq I\left(x_1, x_{2c}; y_1\right) + \varphi,
\end{aligned}
$$

where $\varphi = \mathsf{C}_{21}^{\mathsf{B}} - \widehat{R}_2 + I(\widehat{y}_2; x_{1c}, x_1, x_{2c}, y_1)$.

On the other hand, since we can alternatively upper bound $\Pr\left\{E_S^{(1)}\right\}$ as follows:

$$
\begin{aligned}
&\Pr\left\{E_S^{(1)}\right\} \leq \\
&\sum_{\substack{m: m_s \neq 1, \\ \forall s \in S}} \Pr\left\{\underline{m} \in L(y_1^N)\right\} \cdot \Pr\left\{\exists\, k \neq 1, \widehat{y}_2^N(k) \in B(\underline{m}), b_2(\widehat{y}_2^N(k)) = l_{21} \middle| \underline{m} \in L(y_1^N)\right\}
\end{aligned}
$$

$$\leq 2^{N\left(\sum_{s\in S} R_s\right)} \Pr\left\{\underline{m} \in L(y_1^N)\right\}.$$

the probability of the first kind of error event vanishes as $N \to \infty$ if

$$\begin{aligned}
R_{1p} &\leq I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + \varphi^+ \\
R_{2c} &\leq I\left(x_{2c}; y_1 | x_1\right) + \varphi^+ \\
R_{2c} + R_{1p} &\leq I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + \varphi^+ \\
R_{1c} + R_{1p} &\leq I\left(x_1; y_1 | x_{2c}\right) + \varphi^+ \\
R_{1c} + R_{2c} + R_{1p} &\leq I\left(x_1, x_{2c}; y_1\right) + \varphi^+.
\end{aligned}$$

Finally, plug in $\widehat{R}_2 = I(\widehat{y}_2; y_2)$ and by Markov relation: $(x_{1c}, x_1, x_{2c}, y_1) - y_2 - \widehat{y}_2$, we get the rate loss term

$$\begin{aligned}
\xi_1 &:= \widehat{R}_2 - I(\widehat{y}_2; x_{1c}, x_1, x_{2c}, y_1) = I(\widehat{y}_2; y_2) - I(\widehat{y}_2; x_{1c}, x_1, x_{2c}, y_1) \\
&= I(\widehat{y}_2; y_2 | x_{1c}, x_1, x_{2c}, y_1).
\end{aligned}$$

Probability of $E_S^{(2)}$

Consider the probability of the error event $E_S^{(2)}$:

$$\begin{aligned}
&\Pr\left\{E_S^{(2)}\right\} \\
&\leq \sum_{\underline{m}: m_s \neq 1, \forall s \in S} \Pr\left\{\widehat{y}_2^N(1) \in B(\underline{m}), \underline{m} \in L(y_1^N)\right\} \\
&= \sum_{\underline{m}: m_s \neq 1, \forall s \in S} \Pr\left\{\left(\widehat{y}_2^N(1), \underline{x}^N(\underline{m}), y_1^N\right) \in A_\epsilon^{(N)}\right\} \\
&\leq \begin{cases}
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}) - \epsilon')}, & S = \{1p\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_1; y_1, \widehat{y}_2 | x_{2c}) - \epsilon')}, & S = \{1c\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_{2c}; y_1, \widehat{y}_2 | x_1) - \epsilon')}, & S = \{2c\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}) - \epsilon')}, & S = \{1p, 2c\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_1; y_1, \widehat{y}_2 | x_{2c}) - \epsilon')}, & S = \{1p, 1c\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_1, x_{2c}; y_1, \widehat{y}_2) - \epsilon')}, & S = \{2c, 1c\} \\
2^{N\left(\sum\limits_{s\in S} R_s\right)} \cdot 2^{-N(I(x_1, x_{2c}; y_1, \widehat{y}_2) - \epsilon')}, & S = \{1p, 2c, 1c\}
\end{cases}
\end{aligned}$$

where $\epsilon' = 4\epsilon$. Note that the event when $S = \{2c\}$ does cause an error. Hence, the probability of the second kind of error event vanishes as $N \to \infty$ if

$$\begin{aligned} R_{1p} &\le I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right) \\ R_{2c} &\le I\left(x_{2c}; y_1, \widehat{y}_2 | x_1\right) \\ R_{2c} + R_{1p} &\le I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right) \\ R_{1c} + R_{1p} &\le I\left(x_1; y_1, \widehat{y}_2 | x_{2c}\right) \\ R_{1c} + R_{2c} + R_{1p} &\le I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right). \end{aligned}$$

Error probability analysis at receiver 2:

After receiving the two bin indices, receiver 2 can decode (m_{1c}, m_{2c}, m_{2p}), with effectively smaller candidate message sets, (namely, the bins,) for m_{1c} and m_{2c}. Following the same line as [38], it can be shown that (we omit the detailed analysis here), for all $0 \le \lambda_1^{(ic)} \le 1$ and $\lambda_1^{(1c)} + \lambda_1^{(2c)} = 1$, the following region is achievable:

$$\begin{aligned} R_{2p} &\le I\left(x_2; y_2 | x_{2c}, x_{1c}\right) \\ R_{1c} + R_{2p} &\le I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \lambda_1^{(1c)} \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{2p} &\le I\left(x_2; y_2 | x_{1c}\right) + \lambda_1^{(2c)} \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{1c} + R_{2p} &\le I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}. \end{aligned}$$

Note that the performance of decoding the private message m_{2p} does not gain from cooperation, since receiver 1 does not decode the private message m_{2p}.

Taking convex hull over all possible $\lambda_1^{(1c)} \in [0,1]$. Note that the bounds for R_{2p} and $R_{2c} + R_{1c} + R_{2p}$ remain unchanged. Project the three-dimensional rate region to a two-dimensional space for any fixed $R_{2p} = r_{2p}$, we see that the convexifying procedure results in the following region:

$$\begin{aligned} R_{1c} + r_{2p} &\le I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + r_{2p} &\le I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{1c} + r_{2p} &\le I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}. \end{aligned}$$

Hence the following rate region is achievable for receiver 2 to decode successfully:

$$\begin{aligned} R_{2p} &\le I\left(x_2; y_2 | x_{2c}, x_{1c}\right) \\ R_{1c} + R_{2p} &\le I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{2p} &\le I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2c} + R_{1c} + R_{2p} &\le I\left(x_2, x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}. \end{aligned}$$

3.2 Proof of Lemma 2.7

(1) *Bounds (2.5) on* R_1 *and* (2.6) *on* R_2

Proof. One can directly use cut-set bounds. As an alternative, we give the following proof in which the decomposition of mutual informations is made clear.

We have the following bounds by Fano's inequality, data-processing inequality, and chain rule: if R_1 is achievable,

$$\begin{aligned}
&N(R_1 - \epsilon_N) \\
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N) \overset{(b)}{\leq} I(x_1^N; y_1^N, u_{21}^N, x_2^N) \overset{(c)}{=} I(x_1^N; y_1^N, u_{21}^N | x_2^N) \\
&\overset{(d)}{=} I(x_1^N; y_1^N | x_2^N) + I(x_1^N; u_{21}^N | y_1^N, x_2^N) = h(h_{11}x_1^N + z_1^N) - h(z_1^N) + I(x_1^N; u_{21}^N | y_1^N, x_2^N) \\
&\overset{(e)}{\leq} N \log(1 + \mathsf{SNR}_1) + I(x_1^N; u_{21}^N | y_1^N, x_2^N),
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to Fano's inequality and data processing inequality. (b) is due to the genie giving side information x_1^N to receiver 1, ie., *conditioning reduces entropy*. (c) is due to the fact that x_1^N and x_2^N are independent. (d) is due to chain rule. (e) is due to the fact that i.i.d. Gaussian distribution maximizes differential entropy under covariance constraints.

To upper bound $I(x_1^N; u_{21}^N | y_1^N, x_2^N)$, which corresponds to the enhancement from cooperation, we make use of the fact that u_{21}^N is a function of (y_1^N, y_2^N):

$$\begin{aligned}
&I(x_1^N; u_{21}^N | y_1^N, x_2^N) = h(x_1^N | y_1^N, x_2^N) - h(x_1^N | u_{21}^N, y_1^N, x_2^N) \\
&\overset{(a)}{\leq} h(x_1^N | y_1^N, x_2^N) - h(x_1^N | u_{21}^N, y_1^N, x_2^N, y_2^N) \\
&\overset{(b)}{=} h(x_1^N | y_1^N, x_2^N) - h(x_1^N | y_1^N, x_2^N, y_2^N) \\
&= I(x_1^N; y_2^N | y_1^N, x_2^N) = h(y_2^N | y_1^N, x_2^N) - h(y_2^N | y_1^N, x_2^N, x_1^N) \\
&= h(h_{21}x_1^N + z_2^N | h_{11}x_1^N + z_1^N) - h(z_2^N) \\
&\leq N \log\left(1 + \frac{\mathsf{INR}_2}{1 + \mathsf{SNR}_1}\right).
\end{aligned}$$

(a) is due to the fact that conditioning reduces entropy. (b) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N).

Besides, it is trivial to see that $I(x_1^N; u_{21}^N | y_1^N, x_2^N) \leq H(u_{21}^N) \leq N\mathsf{C}_{21}^{\mathsf{B}}$. Hence, (and similarly for R_2), we have shown bounds (2.5) and (2.6). □

(2) *Bound (2.7) on* $R_1 + R_2$

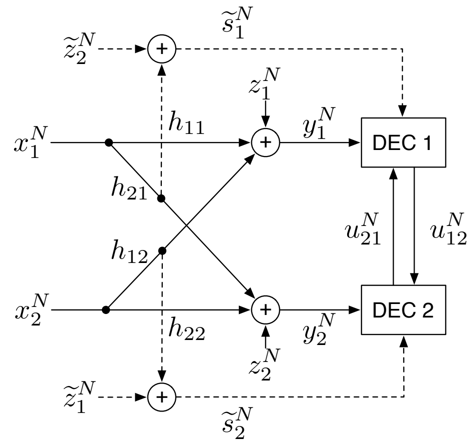


Figure 3.2: Side Information Structure for Bound (2.7)

Proof. Define

$$s_1 := h_{21}x_1 + z_2,\ s_2 := h_{12}x_2 + z_1,$$
$$\widetilde{s}_1 := h_{21}x_1 + \widetilde{z}_2,\ \widetilde{s}_2 := h_{12}x_2 + \widetilde{z}_1,$$

where $\widetilde{z}_1, \widetilde{z}_2$ are i.i.d. $\mathcal{CN}(0,1)$'s, independent of everything else. Note that s_i and $\widetilde{s}_i$ have the same marginal distribution, for $i = 1, 2$.

A genie gives side information $\widetilde{s}_i^N$ to receiver i (refer to Fig. 3.2.) Making use of Fano's inequality, data processing inequality, and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$\begin{aligned}
&N(R_1 + R_2 - \epsilon_N)\\
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N)\\
&\overset{(b)}{=} I(x_1^N; y_1^N) + I(x_2^N; y_2^N) + I(x_1^N; u_{21}^N|y_1^N) + I(x_2^N; u_{12}^N|y_2^N)\\
&\overset{(c)}{\leq} I(x_1^N; y_1^N, \widetilde{s}_1^N) + I(x_2^N; y_2^N, \widetilde{s}_2^N) + H(u_{21}^N) + H(u_{12}^N)\\
&\overset{(d)}{\leq} h(y_1^N, \widetilde{s}_1^N) - h(s_2^N, \widetilde{z}_2^N) + h(y_2^N, \widetilde{s}_2^N) - h(s_1^N, \widetilde{z}_1^N) + N\mathsf{C}_{21}^{\mathsf{B}} + N\mathsf{C}_{12}^{\mathsf{B}}\\
&\overset{(e)}{=} h(y_1^N|\widetilde{s}_1^N) + h(\widetilde{s}_1^N) - h(s_2^N) - h(\widetilde{z}_2^N) + h(y_2^N|\widetilde{s}_2^N)\\
&\quad + h(\widetilde{s}_1^N) - h(s_1^N) - h(\widetilde{z}_1^N) + N\mathsf{C}_{21}^{\mathsf{B}} + N\mathsf{C}_{12}^{\mathsf{B}}\\
&= h(y_1^N|\widetilde{s}_1^N) - h(\widetilde{z}_2^N) + h(y_2^N|\widetilde{s}_2^N) - h(\widetilde{z}_1^N) + N(\mathsf{C}_{21}^{\mathsf{B}} + \mathsf{C}_{12}^{\mathsf{B}})\\
&\overset{(f)}{\leq} N\big\{\text{RHS of (2.7)}\big\},
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) follows from Fano's inequality and data processing inequality. (b) is due to chain rule. (c) is due to the genie giving side information $\widetilde{s}_i^N$ to receiver i, $i = 1, 2$, and $I(x_i^N; u_{ji}^N | y_i^N) \le H(u_{ji}^N)$. (d) is due to the fact that $H(u_{ji}^N) \le N\mathsf{C}_{ji}^{\mathsf{B}}$. (e) is due to chain rule. (f) is due to the fact that i.i.d. Gaussian distribution maximizes conditional entropy subject to conditional variance constraints. Note that alternatively the genie can give side informations s_i^N to receiver i, as in [5].

Hence, we have shown bound (2.7). □

(3) *Bounds (2.8) and* (2.9) *on* $R_1 + R_2$

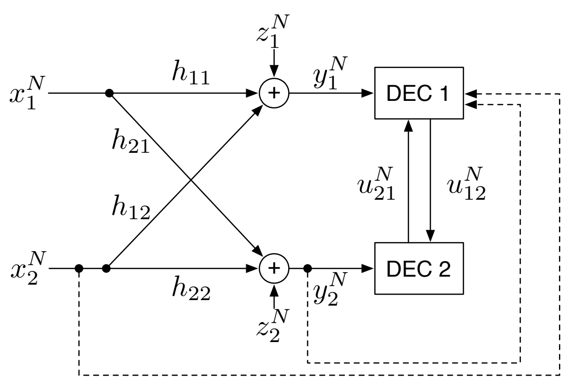


Figure 3.3: Side Information Structure for Bound (2.8)

Proof. A genie gives side information x_2^N and y_2^N to receiver 1 (refer to Fig. 3.3.) Making use of Fano's inequality, data processing inequality, the fact that u_{21}^N is a function of (y_1^N, y_2^N), and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$
\begin{aligned}
&N(R_1 + R_2 - \epsilon_N) \\
&\le I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) \\
&\overset{(a)}{\le} I(x_1^N; y_1^N, u_{21}^N, y_2^N, x_2^N) + I(x_2^N; y_2^N) + I(x_2^N; u_{12}^N | y_2^N) \\
&\overset{(b)}{\le} I\left(x_1^N; y_1^N, u_{21}^N, y_2^N | x_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) + H\left(u_{12}^N\right) \\
&\overset{(c)}{=} I\left(x_1^N; y_1^N, y_2^N | x_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N, s_1^N\right) - h\left(z_1^N, z_2^N\right) + h\left(y_2^N\right) - h\left(s_1^N\right) + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N | s_1^N\right) - h\left(z_1^N, z_2^N\right) + h\left(y_2^N\right) + H\left(u_{12}^N\right) \\
&\le N\big\{\text{RHS of (2.8)}\big\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to chain rule and the genie giving side information x_2^N and y_2^N to receiver 1. (b) is due to the fact that x_1^N and x_2^N are independent, and $I(x_2^N; u_{12}^N | y_2^N) \leq H(u_{12}^N)$. (c) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N).

Hence, (and similarly if we gives side information x_1^N to receiver 2), we have shown bounds (2.8) and (2.9). □

(4) *Bound (2.10) on* $R_1 + R_2$

Proof. This is straightforward cut-set upper bound: if (R_1, R_2) is achievable,

$$
\begin{aligned}
& N(R_1 + R_2 - \epsilon_N) \\
& \leq I\left(x_1^N, x_2^N; y_1^N, y_2^N\right) = h\left(y_1^N, y_2^N\right) - h\left(z_1^N, z_2^N\right) \\
& \leq N\big\{\text{RHS of (2.10)}\big\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$.

Hence, we have shown bound (2.10). □

(5) *Bounds (2.11) on* $2R_1 + R_2$ *and* (2.12) *on* $R_1 + 2R_2$

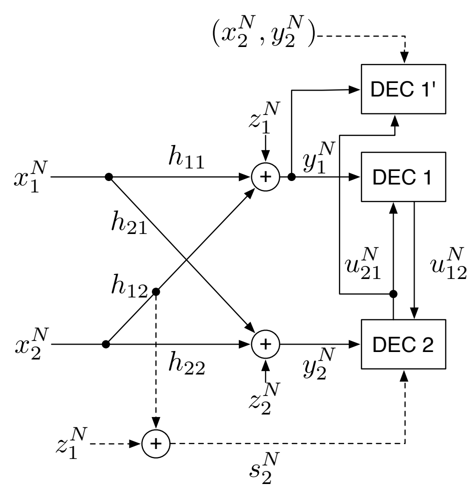


Figure 3.4: Side Information Structure for Bound (2.11)

Proof. A genie gives side information x_2^N and y_2^N to one of the two receiver 1's, and side information s_2^N to receiver 2 (refer to Fig. 3.4.) Making use of Fano's inequality, data processing inequality, the fact that u_{21}^N is a function of (y_1^N, y_2^N), and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$N(2R_1 + R_2 - \epsilon_N)$$

$$
\begin{aligned}
&\overset{(a)}{\leq} I(x_1^N; y_1^N, u_{21}^N) + I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) \\
&\overset{(b)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N, x_2^N) + I(x_1^N; y_1^N) + I(x_2^N; y_2^N, s_2^N) + I(x_1^N; u_{21}^N|y_1^N) + I(x_2^N; u_{12}^N|y_2^N) \\
&\overset{(c)}{\leq} I(x_1^N; y_1^N, u_{21}^N, y_2^N|x_2^N) + I(x_1^N; y_1^N) + I(x_2^N; y_2^N, s_2^N) + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\overset{(d)}{=} I(x_1^N; y_1^N, y_2^N|x_2^N) + I(x_1^N; y_1^N) + I(x_2^N; y_2^N, s_2^N) + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N, s_1^N\right) - h\left(z_1^N, z_2^N\right) + h(y_1^N) - h(s_2^N) + h(y_2^N, s_2^N) - h(s_1^N, z_1^N) \\
&\quad + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&= h\left(h_{11}x_1^N + z_1^N|s_1^N\right) - h\left(z_1^N, z_2^N\right) + h(y_1^N) + h(y_2^N|s_2^N) - h(z_1^N) + H\left(u_{21}^N\right) + H\left(u_{12}^N\right) \\
&\leq N\big\{\text{RHS of (2.11)}\big\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) follows from Fano's inequality and data processing inequality. (b) is due to chain rule and the genie giving side information x_2^N and y_2^N to one of the receiver 1's and side information s_2^N to receiver 2. (c) is due to the fact that x_1^N, x_2^N are independent and $I(x_i^N; u_{ji}^N|y_i^N) \leq H(u_{ji}^N)$. (d) is due to the fact that u_{21}^N is a function of (y_1^N, y_2^N). Hence, (and similarly for $R_1 + 2R_2$), we have shown bounds (2.11) and (2.12). □

(6) *Bounds (2.13) on* $2R_1 + R_2$ *and* (2.14) *on* $R_1 + 2R_2$

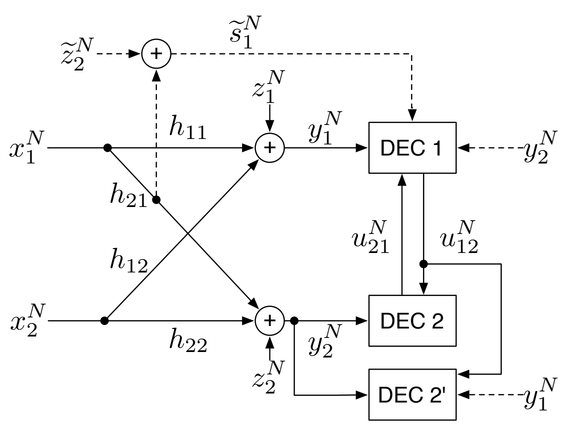


Figure 3.5: Side Information Structure for Bound (2.14)

Proof. A genie gives side information $\tilde{s}_1^N, y_2^N$ to receiver 1, and side information y_1^N to one of the receiver 2's (refer to Fig. 3.5.) Making use of Fano's inequality, data processing inequality, the fact that u_{12}^N, u_{21}^N are functions of (y_1^N, y_2^N), and the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints, we have: if (R_1, R_2) is achievable,

$$N(R_1 + 2R_2 - \epsilon_N)$$

$$\le I(x_1^N; y_1^N, u_{21}^N) + I(x_2^N; y_2^N, u_{12}^N) + I(x_2^N; y_2^N, u_{12}^N)$$
$$\overset{(a)}{\le} I(x_1^N; y_1^N, u_{21}^N, y_2^N, \widetilde{s}_1^N) + I(x_2^N; y_2^N, u_{12}^N, y_1^N) + I(x_2^N; y_2^N) + I(x_2^N; u_{12}^N|y_2^N)$$
$$\overset{(b)}{\le} I(x_1^N; y_1^N, u_{21}^N, y_2^N|\widetilde{s}_1^N) + I(x_1^n; \widetilde{s}_1^N) + I(x_2^N; y_2^N, u_{12}^N, y_1^N) + I(x_2^N; y_2^N) + H(u_{12}^N)$$
$$\overset{(c)}{\le} I(x_1^N; y_1^N, y_2^N|\widetilde{s}_1^N) + I(x_2^N; y_1^N, y_2^N) + h(\widetilde{s}_1^N) - h(z_2^N) + h(y_2^N) - h(s_1^N) + N\mathsf{C}_{12}^{\mathsf{B}}$$
$$\overset{(d)}{\le} I(x_1^N; y_1^N, y_2^N|\widetilde{s}_1^N) + I(x_2^N; y_1^N, y_2^N|x_1^N, \widetilde{s}_1^N) + h(y_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}}$$
$$= I(x_1^N, x_2^N; y_1^N, y_2^N|\widetilde{s}_1^N) + h(y_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}}$$
$$= h(y_1^N, y_2^N|\widetilde{s}_1^N) + h(y_2^N) - h(z_1^N, z_2^N) - h(z_2^N) + N\mathsf{C}_{12}^{\mathsf{B}}$$
$$\overset{(e)}{\le} N\big\{\text{RHS of (2.14)}\big\},$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to the genie giving side information $\widetilde{s}_1^N, y_2^N$ to receiver 1, and side information y_1^N to one of the receiver 2's. (b) is due to chain rule and the fact that $I(x_2^N; u_{12}^N|y_2^N) \le H(u_{12}^N)$. (c) is due to the fact that u_{21}^N and u_{12}^N are both functions of (y_1^N, y_2^N), and that $H(u_{12}^N) \le N\mathsf{C}_{12}^{\mathsf{B}}$. (d) is due to the fact that conditioning reduces entropy and that x_2^N and $(x_1^N, \widetilde{s}_1^N)$ are independent. (e) is due to the fact that Gaussian distribution maximizes conditional entropy subject to conditional variance constraints.

Hence, (and similarly for $2R_1 + R_2$), we have shown bounds (2.14) and (2.13). □

3.3 Proof of Claim 2.12, Claim 2.13, Claim 2.15, and Claim 2.18

3.3.1 Proof of Claim 2.12

Proof. To show (a), since we have four possible $R_1 + 2R_2$ bounds, we distinguish into 4 cases:

(1) If the bound

$$R_1 + 2R_2 \le I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{1c}, x_2; y_2) + I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$\begin{aligned}&3R_1^* + 3R_2^*\\ &= \left\{I(x_1, x_{2c}; y_1, \widehat{y}_2) + I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2|x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}}\right\}\\ &\quad + \left\{I(x_1, x_{2c}; y_1|x_{1c}) + I(x_{1c}, x_2; y_2) + I(x_2; y_2|x_{1c}, x_{2c}) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+\right\}\\ &= \{I(x_1, x_{2c}; y_1, \widehat{y}_2) + I(x_2; y_2|x_{1c}, x_{2c})\} + \left\{I(x_1; y_1|x_{1c}, x_{2c}) + I(x_{1c}, x_2; y_2) + \mathsf{C}_{12}^{\mathsf{B}}\right\}\end{aligned}$$

$$
\begin{aligned}
&+ \left\{ I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \right\} \\
&= (2.16) + (2.19) + (2.17),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

(2) If the bound

$$
\begin{aligned}
R_1 + 2R_2 \leq\ & I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \\
& + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+
\end{aligned}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$
\begin{aligned}
&3R_1^* + 3R_2^* \\
&= \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&\quad + \left\{ \begin{array}{l} I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) \\ + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \end{array} \right\} \\
&= \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \right\} \\
&\quad + \left\{ I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&\quad + \left\{ I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \right\} \\
&= (2.16) + (2.20) + (2.17),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

(3) If the bound

$$
R_1 + 2R_2 \leq I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12}
$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$
\begin{aligned}
&3R_1^* + 3R_2^* \\
&= \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&\quad + \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&= \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \right\} + \left\{ I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&\quad + \left\{ I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \right\} \\
&= (2.16) + (2.19) + (2.18),
\end{aligned}
$$

which is greater than three times the active sum rate bound.

(4) If the bound

$$R_1 + 2R_2 \leq I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}$$

is active, note that the point (R_1^*, R_2^*) where the $R_1 + 2R_2$ bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$\begin{aligned}
&3R_1^* + 3R_2^* \\
&= \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&\quad + \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&= \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right)\right\} \\
&\quad + \left\{I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&\quad + \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&= (2.16) + (2.20) + (2.18),
\end{aligned}$$

which is greater than three times the active sum rate bound.

Hence, we conclude that in case (a), the corner point where $R_1 + R_2$ bound and $R_1 + 2R_2$ bound intersect can be achieved.

To show (b), since we have two possible R_2 bounds, we distinguish into 2 cases:

(1) If the bound

$$R_2 \leq I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}}$$

is active, note that the point (R_1^*, R_2^*) where the R_2 bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$\begin{aligned}
&2R_1^* + 2R_2^* \\
&= \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&\quad + \left\{I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&= I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \\
&\quad + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&\quad + I\left(x_2; y_2 | x_{1c}\right) + I\left(x_{1c}; y_2 | x_{2c}\right) - I\left(x_{1c}, x_2; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= (2.16) + (2.19) + \left[I\left(x_{1c}; y_2 | x_{2c}\right) - I\left(x_{1c}; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}}\right] \\
&\overset{(**)}{\geq} (2.16) + (2.19),
\end{aligned}$$

which is greater than two times the active sum rate bound. $(**)$ is due to

$$\begin{aligned}
I\left(x_{1c}; y_2 | x_{2c}\right) &= I\left(x_{1c}; y_2, x_{2c}\right) - I\left(x_{1c}; x_{2c}\right) \\
&= I\left(x_{1c}; y_2, x_{2c}\right) \geq I\left(x_{1c}; y_2\right),
\end{aligned}$$

since x_{1c} and x_{2c} are independent.

(2) If the bound

$$R_2 \leq I\left(x_{2c}; y_1 | x_1\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right)$$

is active, note that the point (R_1^*, R_2^*) where the R_2 bound and the $2R_1 + R_2$ bound (2.21) intersect, satisfies

$$\begin{aligned}
&2R_1^* + 2R_2^* \\
&= \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&\quad + \left\{I\left(x_{2c}; y_1 | x_1\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right)\right\} \\
&= \left\{I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right)\right\} \\
&\quad + \left\{I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
&= (2.16) + (2.20),
\end{aligned}$$

which is greater than two times the active sum rate bound.

Hence, we conclude that in case (b), the corner point where $R_1 + R_2$ bound and R_2 bound intersect can be achieved. □

3.3.2 Proof of Claim 2.13

Proof. (Keep in mind $\Delta_2 = 1 + \mathsf{SNR}_{2p}$ and $\mathsf{INR}_{ip} \leq 1$, $i = 1, 2$)

(1) R_1 bound: We have two bounds. First, $I\left(x_1; y_1 | x_{2c}\right) = \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_{1p}}\right)$, which is within 2 bits to the upper bound $\log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right)$. Second,

$$\begin{aligned}
&I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1 + \mathsf{INR}_{1p}}\right) + \log\left(\frac{1 + \mathsf{SNR}_{2p} + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&\geq \log\left(\frac{1 + \mathsf{SNR}_1 + \mathsf{INR}_2}{1 + \mathsf{INR}_{1p}}\right) - 1.
\end{aligned}$$

Hence, if the second bound is active, it is within 2 bits to the upper bound $\log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right)$.

(2) R_2 bound: We have two bounds. First, $I\left(x_2; y_2 | x_{1c}\right) + \mathsf{C}_{12}^{\mathsf{B}} = \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}}$. If the first bound is active, it is within 1 bit to the upper bound $\log\left(1 + \mathsf{SNR}_2\right) + \mathsf{C}_{12}^{\mathsf{B}}$. Second,

$$\begin{aligned}
&I\left(x_{2c}; y_1 | x_1\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \\
&= \log\left(\frac{1 + \mathsf{INR}_1}{1 + \mathsf{INR}_{1p}}\right) + \log\left(\frac{1 + \mathsf{SNR}_{2p} + \mathsf{INR}_{2p}}{1 + \mathsf{INR}_{2p}}\right) \geq \log\left(\frac{1 + \mathsf{SNR}_2 + \mathsf{INR}_1}{1 + \mathsf{INR}_{2p}}\right) - 1.
\end{aligned}$$

Hence, the second bound is within 2 bits to the upper bound $\log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_1\right)$.

(3) $R_1 + R_2$ bound: We have six bounds for $R_1 + R_2$, investigated as follows:

- First,

$$
\begin{aligned}
&I\left(x_1, x_{2c}; y_1\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \\
&= \log\left(\frac{1 + \mathsf{SNR}_1 + \mathsf{INR}_1}{1 + \mathsf{INR}_{1p}}\right) + \log\left(1 + \frac{\mathsf{SNR}_{2p}}{1 + \mathsf{INR}_{2p}}\right) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+,
\end{aligned}
$$

which is within $2 + 1 = 3$ bits to the upper bound (2.9).

- Second,

$$
\begin{aligned}
&I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) \\
&= \log\left(\frac{(1 + \Delta_2)(1 + \mathsf{SNR}_1 + \mathsf{INR}_1) + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2}{(1 + \Delta_2)(1 + \mathsf{INR}_{1p}) + \mathsf{SNR}_{2p}}\right) \\
&\quad + \log\left(1 + \frac{\mathsf{SNR}_{2p}}{1 + \mathsf{INR}_{2p}}\right) \\
&\overset{(a)}{\geq} \log\left(\frac{1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2}{4\Delta_2}\right) + \log\left(1 + \mathsf{SNR}_{2p}\right) - 1 \\
&= \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2\right) + \log\left(\frac{1 + \mathsf{SNR}_{2p}}{\Delta_2}\right) - 3 \\
&\overset{(b)}{\geq} \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2\right) - 3,
\end{aligned}
$$

where (a) is due to $(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p} \leq (1+\Delta_2)2+\mathsf{SNR}_{2p} = 4+3\mathsf{SNR}_{2p} \leq 4\Delta_2$ since $\Delta_2 = 1 + \mathsf{SNR}_{2p}$ and $\mathsf{INR}_{1p} \leq 1$. (b) is due to $\Delta_2 = 1 + \mathsf{SNR}_{2p}$.

This lower bound is within 3 bits to the upper bound (2.10).

- Third,

$$
\begin{aligned}
&I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ \\
&= \log\left(\frac{1 + \mathsf{SNR}_{1p} + \mathsf{INR}_1}{1 + \mathsf{INR}_{1p}}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \log\left(\frac{1 + \mathsf{SNR}_{2p} + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+,
\end{aligned}
$$

which is within $2 + 1 = 3$ bits to the upper bound (2.7).

- Fourth,

$$
\begin{aligned}
&I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\
&= I\left(x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12} \\
&\geq I\left(x_{2c}; \widehat{y}_2 | x_{1c}\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}^{\mathsf{B}}_{12}
\end{aligned}
$$

$$\stackrel{(a)}{\geq} I\left(x_{2c}; y_2 | x_{1c}\right) - 1 + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}}$$
$$\stackrel{(b)}{\geq} I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}} - 1$$
$$= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1 + \mathsf{INR}_{1p}}\right) + \log\left(\frac{1 + \mathsf{SNR}_2 + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 1,$$

which is within 3 bits to the upper bound (2.8). Note that (a) is due to

$$\begin{aligned}
I\left(x_{2c}; \widehat{y}_2 | x_{1c}\right) &= \log\left(\frac{1 + \Delta_2 + \mathsf{INR}_{2p} + \mathsf{SNR}_2}{1 + \Delta_2 + \mathsf{INR}_{2p} + \mathsf{SNR}_{2p}}\right) \\
&\geq \log\left(\frac{1 + \mathsf{INR}_{2p} + \mathsf{SNR}_2}{1 + (1 + \mathsf{SNR}_{2p}) + \mathsf{INR}_{2p} + \mathsf{SNR}_{2p}}\right) \\
&\geq \log\left(\frac{1 + \mathsf{INR}_{2p} + \mathsf{SNR}_2}{1 + \mathsf{INR}_{2p} + \mathsf{SNR}_{2p}}\right) - 1 \\
&= I\left(x_{2c}; y_2 | x_{1c}\right) - 1.
\end{aligned}$$

(b) is due to

$$\begin{aligned}
& I\left(x_{2c}; y_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) \\
&= I\left(x_{2c}; y_2, x_{1c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) \\
&\geq I\left(x_{2c}; y_2\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) \\
&= I\left(x_{1c}, x_2, x_{2c}; y_2\right) = I\left(x_{1c}, x_2; y_2\right).
\end{aligned}$$

- Fifth,

$$\begin{aligned}
& I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1 + \mathsf{INR}_{1p}}\right) + \log\left(\frac{1 + \mathsf{SNR}_2 + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}},
\end{aligned}$$

which is within 2 bits to the upper bound (2.8).

- Sixth,

$$\begin{aligned}
& I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
= \;& \log\left(1 + \tfrac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
\geq \;& \log\left(1 + \tfrac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) + \mathsf{C}_{12}^{\mathsf{B}}\;,
\end{aligned}$$

which is within 3 bits to the upper bound (2.8).

(4) $2R_1 + R_2$ bound: The bound

$$\begin{aligned} & I\left(x_1, x_{2c}; y_1\right) + I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \\ = \;& \log\left(\tfrac{1+\mathsf{SNR}_1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(1 + \tfrac{\mathsf{SNR}_{1p}}{1+\mathsf{INR}_{1p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \log\left(\tfrac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \;, \end{aligned}$$

which is within $3 + 1 = 4$ bits to the upper bound (2.11).

(5) $R_1 + 2R_2$ bound: We have four bounds for $R_1 + 2R_2$, investigated as follows:

- First,

$$\begin{aligned} & I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \\ = \;& \log\left(\tfrac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1 + \tfrac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \;, \end{aligned}$$

  which is within $3 + 1 = 4$ bits to the upper bound (2.12).

- Second,

$$\begin{aligned} & I\left(x_1, x_{2c}; y_1 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) \\ & + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \\ & = \left\{ \begin{array}{l} \log\left(\tfrac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) \\ + \log\left(\tfrac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1 + \tfrac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\} \\ & \geq \left\{ \begin{array}{l} \log\left(\tfrac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) \\ + \log\left(1 + \tfrac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} + \left(\mathsf{C}_{21}^{\mathsf{B}} - \xi_1\right)^+ \end{array} \right\}, \end{aligned}$$

  which is within $4 + 1 = 5$ bits to the upper bound (2.12).

- Third,

$$\begin{aligned} & I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{1c}, x_2; y_2\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\ & = \left\{ \begin{array}{l} \log\left(\tfrac{(1+\Delta_2)(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p}}\right) \\ + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1 + \tfrac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\ & \geq \left\{ \begin{array}{l} \log\left(\tfrac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{4\Delta_2}\right) \\ + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1 + \mathsf{SNR}_{2p}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 1 \end{array} \right\} \\ & \geq \left\{ \begin{array}{l} \log\left(1 + \mathsf{SNR}_{1p} + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_{2p} + |h_{11}h_{22} - h_{12}h_{21}|^2 Q_{1p}\right) \\ + \log\left(\tfrac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 3 \end{array} \right\}, \end{aligned}$$

  which is within $1 + 3 = 4$ bits to the upper bound (2.14).

- Fourth,

$$
\begin{aligned}
&I\left(x_1, x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_{2c}; y_1 | x_1\right) + I\left(x_{1c}, x_2; y_2 | x_{2c}\right) + I\left(x_2; y_2 | x_{1c}, x_{2c}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= \left\{ \begin{array}{l} \log\left(\frac{(1+\Delta_2)(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{(1+\Delta_2)(1+\mathsf{INR}_{1p})+\mathsf{SNR}_{2p}}\right) \\ +\log\left(\frac{1+\mathsf{INR}_1}{1+\mathsf{INR}_{1p}}\right) + \log\left(\frac{1+\mathsf{SNR}_{2p}+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \log\left(1+\frac{\mathsf{SNR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \end{array} \right\} \\
&\ge \left\{ \begin{array}{l} \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{4\Delta_2}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) + \log\left(1+\mathsf{SNR}_{2p}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 1 \end{array} \right\} \\
&\ge \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{(1+\mathsf{INR}_{1p})(1+\mathsf{INR}_{2p})}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 3 \end{array} \right\},
\end{aligned}
$$

  which is within $2+3=5$ bits to the upper bound (2.14).

Therefore, we see that the bounds in $\mathscr{R}_{2\to1\to2}$ except (2.21) satisfies:

- R_1 bound is within 2 bits to outer bounds;
- R_2 bound is within 2 bits to outer bounds;
- $R_1 + R_2$ bound is within 3 bits to outer bounds;
- $2R_1 + R_2$ bound is within 4 bits to outer bounds;
- $R_1 + 2R_2$ bound is within 5 bits to outer bounds.

□

3.3.3 Proof of Claim 2.15

Proof. (Keep in mind $\Delta_2 = 1$ and $\mathsf{INR}_{2p} \le 1$)

(1) R_1 bound: We have two bounds. First, $I\left(x_1; y_1 | x_2\right) = \log\left(1+\mathsf{SNR}_1\right)$, which is within 1 bit to the upper bound $R_1 \le \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)$. Second,

$$
\begin{aligned}
&I\left(x_1; y_1 | x_{1c}, x_2\right) + I\left(x_{1c}; y_2 | x_2\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&= \log\left(1+\mathsf{SNR}_{1p}\right) + \log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
&\ge \log\left(\frac{1+\mathsf{SNR}_1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}_{12}^{\mathsf{B}}.
\end{aligned}
$$

Hence, if the second bound is active, it is within 1 bit to the upper bound $\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)$.

(2) R_2 bound: We have two bounds. First, $I(x_1; y_1|x_1) = \log(1 + \mathsf{INR}_1)$, which is within 1 bit to the upper bound $R_2 \le \log(1 + \mathsf{SNR}_2 + \mathsf{INR}_1)$. Second, $I(x_2; y_2|x_{1c}) + \mathsf{C}^{\mathsf{B}}_{12} = \log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_{2p}}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}^{\mathsf{B}}_{12}$, which is within 1 bit to the upper bound $R_2 \le \log(1 + \mathsf{SNR}_2) + \mathsf{C}^{\mathsf{B}}_{12}$.

(3) $R_1 + R_2$ bound: We have five bounds, investigated as follows:

- First,

$$I(x_1, x_2; y_1) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+ = \log(1 + \mathsf{SNR}_1 + \mathsf{INR}_1) + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+,$$

  which is within $1 + \xi_1 = 2$ bits to the upper bound (2.9).

- Second,

$$I(x_1, x_2; y_1, \widehat{y}_2) = \log\left(\frac{2(1 + \mathsf{SNR}_1 + \mathsf{INR}_1) + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2}{2}\right)$$

  which is within 1 bit to the upper bound (2.10).

- Third,

$$\begin{aligned}&I(x_1; y_1|x_{1c}, x_2) + I(x_{1c}, x_2; y_2) + \mathsf{C}^{\mathsf{B}}_{12}\\&= \log(1 + \mathsf{SNR}_{1p}) + \log\left(\frac{1 + \mathsf{SNR}_2 + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}^{\mathsf{B}}_{12}\end{aligned}$$

  which is within 1 bit to the upper bound (2.8).

- Fourth,

$$\begin{aligned}&I(x_1, x_2; y_1|x_{1c}) + I(x_{1c}; y_2|x_2) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+\\&= \log(1 + \mathsf{SNR}_{1p} + \mathsf{INR}_1) + \log\left(\frac{1 + \mathsf{INR}_2}{1 + \mathsf{INR}_{2p}}\right) + \mathsf{C}^{\mathsf{B}}_{12} + \left(\mathsf{C}^{\mathsf{B}}_{21} - \xi_1\right)^+,\end{aligned}$$

  which is within $2 + \xi_1 = 3$ bits to the upper bound (2.7).

- Fifth,

$$\begin{aligned}&I(x_1, x_2; y_1, \widehat{y}_2|x_{1c}) + I(x_{1c}; y_2|x_2) + \mathsf{C}^{\mathsf{B}}_{12}\\&= I(x_2; y_1, \widehat{y}_2|x_{1c}) + I(x_1; y_1, \widehat{y}_2|x_{1c}, x_2) + I(x_{1c}; y_2|x_2) + \mathsf{C}^{\mathsf{B}}_{12}\\&\ge I(x_2; y_1, \widehat{y}_2|x_{1c}) + I(x_1; y_1|x_{1c}, x_2) + I(x_{1c}; y_2|x_2) + \mathsf{C}^{\mathsf{B}}_{12}\\&= \left\{\begin{array}{l}\log\left(\frac{2(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{2(1+\mathsf{SNR}_{1p})+\mathsf{INR}_{2p}}\right)\\ + \log(1 + \mathsf{SNR}_{1p}) + \log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right) + \mathsf{C}^{\mathsf{B}}_{12}\end{array}\right\}\end{aligned}$$

$$\geq \left\{ \begin{array}{l} \log\left(\frac{1+\mathsf{SNR}_{1p}+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{3(1+\mathsf{SNR}_{1p})}\right) \\ +\log\left(1+\mathsf{SNR}_{1p}\right)+\log\left(\frac{1+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}$$

$$\geq \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right) \\ +\log\left(\frac{1}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12}-\log 3 \end{array} \right\}.$$

Hence, if this bound is active, it is within $1+\log 3=\log 6$ bits to the upper bound (2.10).

(4) R_1+2R_2 bound: We have two bounds. First,

$$\begin{aligned} &I\left(x_1,x_2;y_1|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+ \\ &=\log\left(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1\right)+\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12}+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+, \end{aligned}$$

which is within $2+\xi_1=3$ bits to the upper bound (2.12)

Second,

$$\begin{aligned} &I\left(x_1,x_2;y_1,\widehat{y}_2|x_{1c}\right)+I\left(x_{1c},x_2;y_2\right)+\mathsf{C}^{\mathsf{B}}_{12} \\ &=\left\{ \begin{array}{l} \log\left(\frac{2(1+\mathsf{SNR}_{1p}+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_{2p}+|h_{11}h_{22}-h_{12}h_{21}|^2 Q_{1p}}{2}\right) \\ +\log\left(\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2}{1+\mathsf{INR}_{2p}}\right)+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\}, \end{aligned}$$

which is within 2 bits to the upper bound (2.14).

Therefore, we see that the bounds in $\mathscr{R}_{2\to1\to2}$ satisfies:

- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- R_1+R_2 bound is within 3 bits to outer bounds;
- R_1+2R_2 bound is within 3 bits to outer bounds.

□

3.3.4 Proof of Claim 2.18

Proof. (Keep in mind that $\Delta_1=\Delta_2=1$)

(1) R_1 bound: We have four bounds. First,

$$I\left(x_1;y_1|x_2\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+=\log\left(1+\mathsf{SNR}_1\right)+\left(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1\right)^+$$

which is within $\xi_1 = 1$ bit to the upper bound $\log(1+\mathsf{SNR}_1)+\mathsf{C}^{\mathsf{B}}_{21}$. Second,

$$I(x_1;y_2|x_2)+(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2)^+ = \log(1+\mathsf{INR}_2)+(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2)^+$$
$$\geq \log(1+\mathsf{SNR}_1+\mathsf{INR}_2)-1.$$

Hence if this bound is active, it is within 1 bit to the upper bound $\log(1+\mathsf{SNR}_1+\mathsf{INR}_2)$. Finally,

$$I(x_1;y_1,\widehat{y}_2|x_2) = \log\left(\frac{2+2\mathsf{SNR}_1+\mathsf{INR}_2}{2}\right)$$
$$I(x_1;y_2,\widehat{y}_1|x_2) = \log\left(\frac{2+\mathsf{SNR}_1+2\mathsf{INR}_2}{2}\right),$$

which are both within 1 bit to the upper bound $\log(1+\mathsf{SNR}_1+\mathsf{INR}_2)$.

(2) R_2 bound: By symmetry we have the same gap result as (1).

(3) R_1+R_2 bound: We have four bounds. First,

$$I(x_1,x_2;y_1)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+ = \log(1+\mathsf{SNR}_1+\mathsf{INR}_1)+(\mathsf{C}^{\mathsf{B}}_{21}-\xi_1)^+,$$

which is within $1+\xi_1 = 2$ bits to the upper bound (2.9). Second,

$$I(x_2,x_1;y_2)+(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2)^+ = \log(1+\mathsf{SNR}_2+\mathsf{INR}_2)+(\mathsf{C}^{\mathsf{B}}_{12}-\xi_2)^+,$$

which is within $1+\xi_2 = 2$ bits to the upper bound (2.8). Finally,

$$I(x_1,x_2;y_1,\widehat{y}_2) = \log\left(\frac{2(1+\mathsf{SNR}_1+\mathsf{INR}_1)+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2}{2}\right)$$
$$I(x_2,x_1;y_2,\widehat{y}_1) = \log\left(\frac{2(1+\mathsf{SNR}_2+\mathsf{INR}_2)+\mathsf{SNR}_1+\mathsf{INR}_1+|h_{11}h_{22}-h_{12}h_{21}|^2}{2}\right),$$

which are both within 1 bit to the upper bound (2.10).

Therefore, we see that the bounds in $\mathscr{R}_{\text{OneRound}}$ satisfies:

- R_1 bound is within 1 bit to outer bounds;
- R_2 bound is within 1 bit to outer bounds;
- R_1+R_2 bound is within 2 bits to outer bounds.

□

3.4 Proof of Theorem 2.23

From Section 2.5.5, we have shown that when $\mathsf{SNR} \le \mathsf{INR}$,

$$R_{\text{sym,OneRound}} \le C_{\text{sym}} \le \overline{C}_{\text{sym}} \le R_{\text{sym,OneRound}} + 1.$$

Hence we focus on the case $\mathsf{SNR} > \mathsf{INR}$ in the rest of the proof.

By symmetry and by Theorem 2.16, if $R_{\text{sym,OneRound}} \ge 0$ satisfies the following, it is achievable:

$$\begin{aligned}
R_{\text{sym,OneRound}} &\le \min\left\{I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right)\right\},\\
R_{\text{sym,OneRound}} &\le \min\left\{I\left(x_1; y_1 | x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{2c}\right)\right\},\\
2R_{\text{sym,OneRound}} &\le \min\left\{I\left(x_1, x_{2c}; y_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right)\right\}\\
&\quad + \min\left\{I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right)\right\}.
\end{aligned}$$

Note that since

$$\begin{aligned}
I\left(x_1; y_1 | x_{1c}, x_{2c}\right) \le I\left(x_1; y_1, \widehat{y}_2 | x_{1c}, x_{2c}\right) \le I\left(x_1; y_1 | x_{1c}, x_{2c}\right) + \text{constant},\\
I\left(x_1; y_1 | x_{2c}\right) \le I\left(x_1; y_1, \widehat{y}_2 | x_{2c}\right) \le I\left(x_1; y_1 | x_{2c}\right) + \text{constant},
\end{aligned}$$

a sufficient condition for achievable $R_{\text{sym,OneRound}}$ is

$$\begin{aligned}
R_{\text{sym,OneRound}} &\le \min\left\{I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right), I\left(x_1; y_1 | x_{2c}\right)\right\}\\
R_{\text{sym,OneRound}} &\le \frac{1}{2}\min\left\{I\left(x_1, x_{2c}; y_1\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+, I\left(x_1, x_{2c}; y_1, \widehat{y}_2\right)\right\} + \frac{1}{2} I\left(x_1; y_1 | x_{1c}, x_{2c}\right)
\end{aligned}$$

(1) $I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+$:

$$I\left(x_{2c}, x_1; y_1 | x_{1c}\right) + (\mathsf{C}^{\mathsf{B}}_{21} - \xi_1)^+ = \log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}}{1 + \mathsf{INR}_p}\right) + (\mathsf{C}^{\mathsf{B}} - \xi)^+,$$

and its gap to the outer bound $\log\left(1 + \mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \mathsf{C}^{\mathsf{B}}$:

$$\text{gap} \le \log\left(1 + \mathsf{INR}_p\right) + \xi \le 1 + 1 = 2.$$

(2) $I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right)$:

$$\begin{aligned}
&I\left(x_{2c}, x_1; y_1, \widehat{y}_2 | x_{1c}\right)\\
&= I\left(x_{2c}; y_1, \widehat{y}_2 | x_{1c}\right) + I\left(x_1; y_1, \widehat{y}_2 | x_{2c}, x_{1c}\right)\\
&\ge I\left(x_{2c}; \widehat{y}_2 | x_{1c}\right) + I\left(x_1; y_1 | x_{2c}, x_{1c}\right)\\
&= \log\left(\frac{1 + \Delta + \mathsf{SNR} + \mathsf{INR}_p}{1 + \Delta + \mathsf{SNR}_p + \mathsf{INR}_p}\right) + \log\left(\frac{1 + \mathsf{SNR}_p + \mathsf{INR}_p}{1 + \mathsf{INR}_p}\right)
\end{aligned}$$

$$\overset{\text{(a)}}{\geq} \log\left(\frac{1+\mathsf{SNR}}{2+2\mathsf{SNR}_p+2\mathsf{INR}_p}\right)+\log\left(\frac{1+\mathsf{SNR}_p+\mathsf{INR}_p}{1+\mathsf{INR}_p}\right)$$
$$=\log\left(\frac{1+\mathsf{SNR}}{1+\mathsf{INR}_p}\right)-1,$$

where (a) is due to $\Delta = 1+\mathsf{SNR}_p$.

Therefore, the gap to the outer bound $\log(1+\mathsf{SNR}+\mathsf{INR})$:

$$\begin{aligned}\text{gap} &\leq 1+\log\left(\frac{1+\mathsf{SNR}+\mathsf{INR}}{1+\mathsf{SNR}}\right)+\log(1+\mathsf{INR}_p)\\ &\leq 1+\log\left(\frac{2+2\mathsf{SNR}}{1+\mathsf{SNR}}\right)+\log(1+1)=3,\end{aligned}$$

since $\mathsf{SNR} > \mathsf{INR}$ and $\mathsf{INR}_p \leq 1$.

(3) $I(x_1;y_1|x_{2c})$:

$$I(x_1;y_1|x_{2c}) = \log(1+\mathsf{SNR}+\mathsf{INR}_p)-\log(1+\mathsf{INR}_p),$$

and its gap to the outer bound $\log(1+\mathsf{SNR}+\mathsf{INR})$:

$$\begin{aligned}\text{gap} &\leq \log\left(\frac{1+\mathsf{SNR}+\mathsf{INR}}{1+\mathsf{SNR}+\mathsf{INR}_p}\right)+\log(1+\mathsf{INR}_p)\\ &\leq \log\left(\frac{2+2\mathsf{SNR}}{1+\mathsf{SNR}}\right)+\log(1+1)=2.\end{aligned}$$

(4) $\frac{1}{2}I(x_1,x_{2c};y_1)+\frac{1}{2}(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+ + \frac{1}{2}I(x_1;y_1|x_{1c},x_{2c})$:

$$\begin{aligned}&\frac{1}{2}I(x_1,x_{2c};y_1)+\frac{1}{2}(\mathsf{C}_{21}^{\mathsf{B}}-\xi_1)^+ + \frac{1}{2}I(x_1;y_1|x_{1c},x_{2c})\\ &=\frac{1}{2}\log(1+\mathsf{SNR}+\mathsf{INR})+\frac{1}{2}(\mathsf{C}^{\mathsf{B}}-\xi)^+ + \frac{1}{2}\log(1+\mathsf{SNR}_p+\mathsf{INR}_p)-\log(1+\mathsf{INR}_p),\end{aligned}$$

and its gap to the outer bound $\frac{1}{2}\log(1+\mathsf{SNR}+\mathsf{INR})+\frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\frac{1}{2}\mathsf{C}^{\mathsf{B}}$:

$$\text{gap} \leq \frac{1}{2}\xi + \log(1+\mathsf{INR}_{\mathrm{p}}) \leq 1.5.$$

(5) $\frac{1}{2}I(x_1,x_{2c};y_1,\widehat{y}_2)+\frac{1}{2}I(x_1;y_1|x_{1c},x_{2c})$:

$$\begin{aligned}&\frac{1}{2}I(x_1,x_{2c};y_1,\widehat{y}_2)+\frac{1}{2}I(x_1;y_1|x_{1c},x_{2c})\\ &=\frac{1}{2}\log\left(\frac{\Delta(1+\mathsf{SNR}+\mathsf{INR})+1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2}{\Delta(1+\mathsf{INR}_p)+1+\mathsf{SNR}_p+\mathsf{INR}_p}\right)\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\log\left(\frac{1+\mathsf{SNR}_p+\mathsf{INR}_p}{1+\mathsf{INR}_p}\right)\\
&\geq \frac{1}{2}\log\left(\frac{1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2}{4\Delta}\right)+\frac{1}{2}\log\left(\frac{\Delta}{1+\mathsf{INR}_p}\right)\\
&=\frac{1}{2}\log\left(1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2\right)-\frac{1}{2}\log\left(1+\mathsf{INR}_p\right)-1.
\end{aligned}
$$

Therefore, the gap to the outer bound $\frac{1}{2}\log\left(1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2\right)$:

$$\text{gap} \leq \frac{1}{2}\log\left(1+\mathsf{INR}_{\rm p}\right)+1 \leq 1.5.$$

From (1) - (5), we conclude that when $\mathsf{SNR} > \mathsf{INR}$,

$$R_{\rm sym,OneRound} \leq C_{\rm sym} \leq \overline{C}_{\rm sym} \leq R_{\rm sym,OneRound}+3.$$

This completes the proof.

3.5 Proof of Lemma 2.25

Proof. From Corollary 2.24 we see that except the term

$$V := \frac{1}{2}\log\left(1+2\mathsf{SNR}+2\mathsf{INR}+|h_{11}h_{22}-h_{12}h_{21}|^2\right),$$

all terms scaled by $\log\mathsf{SNR}$ converges *everywhere* as $\mathsf{SNR}\to\infty$ with α,κ fixed. Note that

$$
\begin{aligned}
&|h_{11}h_{22}-h_{12}h_{21}|^2\\
&=|g_{11}e^{j\Theta_{11}}g_{22}e^{j\Theta_{22}}-g_{12}e^{j\Theta_{12}}g_{21}e^{j\Theta_{21}}|^2\\
&=\Big[g_{11}g_{22}\cos\left(\Theta_{11}+\Theta_{22}\right)-g_{12}g_{21}\cos\left(\Theta_{12}+\Theta_{21}\right)\Big]^2\\
&\quad+\Big[g_{11}g_{22}\sin\left(\Theta_{11}+\Theta_{22}\right)-g_{12}g_{21}\sin\left(\Theta_{12}+\Theta_{21}\right)\Big]^2\\
&=g_{11}^2g_{22}^2+g_{12}^2g_{21}^2-2g_{11}g_{22}g_{12}g_{21}\cos\left(\Theta_{11}+\Theta_{22}-\Theta_{12}-\Theta_{21}\right)\\
&=\mathsf{SNR}^2+\mathsf{INR}^2-2(\cos\Theta)\mathsf{SNR}\mathsf{INR},
\end{aligned}
$$

where $\Theta=\Theta_{11}+\Theta_{22}-\Theta_{12}-\Theta_{21} \mod 2\pi$. Obviously Θ is uniformly distributed over $[0,2\pi]$. Now, consider the limit

$$L(\alpha,\kappa) := \lim_{\substack{\text{fix }\alpha,\kappa\\ \mathsf{SNR}\to\infty}} \frac{V}{\log\mathsf{SNR}}.$$

We have the following upper and lower bounds for V due to the fact that $\big||h_{11}||h_{22}| - |h_{12}||h_{21}|\big| \leq |h_{11}h_{22} - h_{12}h_{21}| \leq |h_{11}||h_{22}| + |h_{12}||h_{21}|$:

$$V \geq \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + (\mathsf{SNR} - \mathsf{INR})^2\right);$$
$$V \leq \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + (\mathsf{SNR} + \mathsf{INR})^2\right).$$

Hence, when $\alpha < 1$, taking limits at both sides yields $1 \leq L(\alpha,\kappa) \leq 1$ and implies $L(\alpha,\kappa) = 1$. Similarly, when $\alpha > 1$, taking limits at both sides yields $\alpha \leq L(\alpha,\kappa) \leq \alpha$ and implies $L(\alpha,\kappa) = \alpha$. When $\alpha = 1$, note that

$$\begin{aligned} V &= \frac{1}{2}\log\left(1 + 2\mathsf{SNR} + 2\mathsf{INR} + \mathsf{SNR}^2 + \mathsf{INR}^2 - 2(\cos\Theta)\mathsf{SNR}\mathsf{INR}\right) \\ &= \frac{1}{2}\log\left((1 + \mathsf{SNR} + \mathsf{INR})^2 - 4\cos^2\frac{\Theta}{2}\mathsf{SNR}\mathsf{INR}\right), \end{aligned}$$

and therefore $L(\alpha,\kappa) = 1$ if $\Theta \neq 0, 2\pi$. Since the event $\{\Theta = 0, 2\pi\}$ is of zero measure, the limit $L(\alpha,\kappa)$ exists almost surely. □

Part II

Interference Mitigation through Transmitter Cooperation

Chapter 4

Interference Channel with Transmitter Cooperation

In Part I, we show how limited receiver cooperation helps mitigate interference. The scenario with transmitter cooperation, however, is more difficult to tackle. In this chapter we study the two-user Gaussian interference channel with conferencing transmitters to make progress towards this direction. We characterize the capacity region to within 6.5 bits/s/Hz, regardless of channel parameters. Based on the bounded-gap-to-optimality result, we show that there is an interesting reciprocity between the scenario with conferencing transmitters and the scenario with conferencing receivers, and their capacity regions are within a bounded gap to each other. Hence in the interference-limited regime, the behavior of the benefit brought by transmitter cooperation is the same as that by receiver cooperation.

4.1 Introduction

In modern wireless communication systems, interference has become the major factor that limits the performance. Interference arises whenever multiple transmitter-receiver pairs are present, and each receiver is only interested in retrieving information from its own transmitter. Due to the broadcast and superposition nature of wireless channels, one user's information-carrying signal causes interference to other users. The *interference channel* is the simplest information theoretic model for studying this issue, where each transmitter (receiver) is assumed to be isolated from other transmitters (receivers). In various practical scenarios, however, they are not isolated, and *cooperation* among transmitters or receivers can be induced. For example, in downlink cellular systems, base stations are connected via infrastructure backhaul networks.

In Part I, we have studied the two-user Gaussian interference channel with conferencing receivers to understand how limited receiver cooperation helps mitigate interference. We proposed good coding strategies, proved tight outer bounds, and characterized the capacity region to within 2 bits/s/Hz. Based upon the bounded-gap-to-optimality result, we identify two regions regarding the gain from receiver cooperation: linear and saturation regions. In the linear region, receiver cooperation is *efficient*, in the sense that the growth of each user's "over-the-air" data rate is roughly linear with respect to the capacity of receiver-cooperative links. The gain in this region is the *degrees-of-freedom* gain that distributed MIMO systems provide. In the saturation region, receiver cooperation is *inefficient* in the sense that the growth of each user's over-the-air data rate becomes saturated as one increases the rate in receiver-cooperative links. The gain is the *power* gain which is bounded, independent of the channel strength. Furthermore, until saturation the degree-of-freedom gain is either *one cooperation bit buys one over-the-air bit* or *two cooperation bits buy one over-the-air bit*.

In this chapter, we study its reciprocal problem, the two-user Gaussian interference channel with conferencing transmitters, to investigate how limited transmitter cooperation helps mitigate interference. A natural cooperative strategy between transmitters is that, prior to each block of transmission, two transmitters hold a conference to tell each other part of their messages. Hence the messages are classified into two kinds: (1) *cooperative* messages, which are those known to both transmitters due to the conference, and (2) *noncooperative* ones, which are those unknown to the other transmitter since the cooperative link capacities are finite. On the other hand, messages can also be classified based on their target receivers: (1) *common* messages, which are those aimed at both receivers, and (2) *private* ones, which are those aimed at their own receiver. Hence in total there are four kinds of messages for each user, and seven codes for the whole system[1]. Now the question is, how do we encode these messages?

Generally speaking, Gaussian interference channels with transmitter cooperation are more difficult to tackle than Gaussian interference channels with receiver cooperation. Take the following extreme case. When transmitters can cooperate in an unlimited fashion, the scenario reduces to the MIMO Gaussian broadcast channel. When receivers can cooperate in an unlimited fashion, the scenario reduces to the MIMO Gaussian multiple access channel. The capacity region of the latter is fully characterized in the 70's [39] [40], while that of the former has not been solved until recently [41]. This is due to difficulties both in achievability and outer bounds.

Similar phenomenon arises between Gaussian interference channels with conferencing transmitters and Gaussian interference channels with conferencing receivers. Compared with the scenario with conferencing receivers [15] where each user just has two kinds of messages (common and private), in the scenario with conferencing transmitters not only does the message structure in the strategy become more complicated due to the collaboration among transmitters, but it is also more difficult to prove the outer bounds since the transmitters are potentially correlated. In order to overcome the difficulties, we first study an auxiliary problem in the *linear deterministic* setting [34] [9]. We first characterize the capacity region of the linear deterministic interference channel with conferencing transmitters, and then

[1] There is only one cooperative common code carrying both cooperative common messages.

make use of the intuition there to design good coding strategies and to prove outer bounds in the Gaussian scenario. Eventually the proposed strategy in the Gaussian setting is a simple superposition of a pair of *noncooperative* common and private codewords and a pair of *cooperative* common and private codewords. For the noncooperative part, Han-Kobayashi scheme [4] is employed, and the common-private split is such that the private interference is at or below the noise level at the unintended receiver [5]. For the cooperative part, we use a simple linear beamforming strategy for encoding the private messages, superimposed upon the common codewords. By choosing the power split and beamforming vectors cleverly, the strategy achieves the capacity region universally to within 6.5 bits, regardless of channel parameters. The 6.5-bit gap is the worst-case gap which can be loose in some regimes, and it is vanishingly small at high SNR when compared to the capacity.

With the bounded-gap-to-optimality result, we observe an interesting uplink-downlink reciprocity between the scenario with conferencing receivers and the scenario with conferencing transmitters: for the original and reciprocal channels, the capacity regions are within a bounded gap to each other. Hence the fundamental gain from transmitter cooperation at high SNR is the same as that from receiver cooperation [15].

Related Works

Conferencing among transmitters is first studied by Willems [18] in the context of *multiple access channels*, where the capacity region is characterized. The capacity of the Gaussian MAC with conferencing transmitters, however, has not been characterized explicitly in a computable form until recently by Bross *et al.*[19], where the authors show that the optimization on auxiliary random variables can be reduced to finding the optimal Gaussian input distribution. On the other hand, the extension to the compound MAC has been done by Marić *et al.*[20].

Works on Gaussian *interference channel* with transmitter cooperation can be roughly divided into two categories. One set of works investigate cooperation in interference channels with a set-up where the cooperative links share the same band as the links in the interference channel. Høst-Madsen [24] proposes cooperative strategies based on decode-forward, compress-forward, and dirty paper coding, and derives the achievable rates. The recent work by Prabhakaran *et al.*[42] characterizes the sum capacity of Gaussian interference channels with reciprocal in-band transmitter cooperation to within a bounded gap. The other set of works focus on conferencing transmitters, that is, cooperative links are orthogonal to each other as well as the links in the interference channel. Some works are dedicated to achievable rates. Cao *et al.*[43] derive an achievable rate region based on superposition coding and dirty paper coding. Some works consider special cases of the channel. One such special case attracting particularly broad interest is the *cognitive interference channel*, where one of the transmitters (the cognitive user) is assumed to have full knowledge about the other's transmission (the primary user). It is equivalent to the case where transmitter cooperation is unidirectional and unlimited. As for the cognitive interference channel, Marić *et al.*[20] characterize the capacity region in the strong interference regime. Wu *et al.*[44] and Jovičić *et al.*[45] independently characterize the capacity region when the interference at the primary receiver is weak. Very recently, Rini *et al.*[46] characterize the capacity region to within a bounded gap, regardless of channel parameters. On the other hand, works on the case with limited cooperative capacities are not rich in the literature. Bagheri *et al.*[47] investigate symmetric Gaussian interference channel with unidirectional limited transmitter cooperation, and characterize the sum capacity to within a bounded gap.

Our main contribution in this chapter is characterizing the capacity *region* of the two-user Gaussian interference channel with conferencing transmitters to within a bounded gap for *arbitrary* channel strength and cooperative link capacities. The rest of the chapter is organized as follows. After we formulate the problem in Section 4.2, we investigate the auxiliary linear deterministic channel in Section 4.3. Then we carry the intuitions and techniques to solve the original problem in Section 4.4 and characterize the capacity region to within a bounded gap. In Section 4.5 we discuss the interesting uplink-downlink reciprocity.

4.2 Problem Formulation

4.2.1 Channel Model

The Gaussian interference channel with conferencing transmitters is depicted in Fig. 4.1.

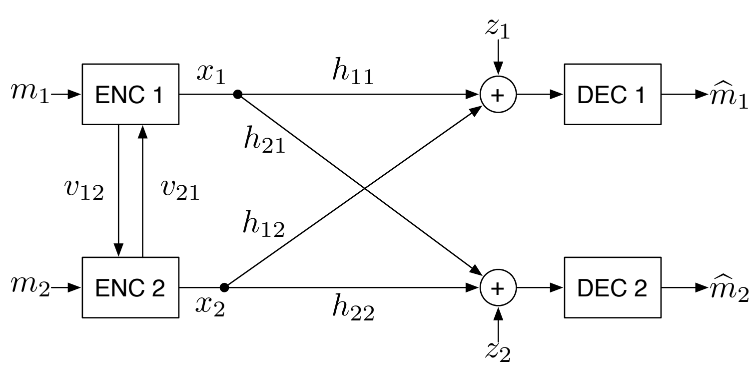


Figure 4.1: Channel Model

The links among transmitters and receivers are modeled as the *normalized* Gaussian interference channel:

$$y_1 = h_{11}x_1 + h_{12}x_2 + z_1,\ y_2 = h_{21}x_1 + h_{22}x_2 + z_2,$$

where the additive noise processes $\{z_i[k]\}$, $(i = 1, 2)$, are independent $\mathcal{CN}(0, 1)$, i.i.d. over time. In this chapter, we use [.] to denote time indices. Transmitter i intends to convey message m_i to receiver i by encoding it into a block codeword $\{x_i[k]\}_{k=1}^N$, with transmit power constraints

$$\frac{1}{N}\sum_{k=1}^{N}\left|x_i[k]\right|^2 \leq 1,\ i=1,2,$$

for arbitrary block length N. Note that outcome of the encoder depends on both messages. Messages m_1, m_2 are independent. Define channel parameters

$$\mathsf{SNR}_i := |h_{ii}|^2,\ \mathsf{INR}_i := |h_{ij}|^2,\ i,j=1,2,\ i\neq j.$$

The cooperative links between transmitters are noiseless with finite capacity $\mathsf{C}^{\mathsf{B}}_{ij}$ from transmitter i to j. Encoding must satisfy causality constraints: for any time index $k = 1,2,\ldots,N$, $v_{ij}[k]$ is only a function of $\{m_i, v_{ji}[1],\ldots,v_{ji}[k-1]\}$.

4.2.2 Notations

We summarize below the notations used in the rest of this chapter.

- For a real number a, $(a)^+ := \max(a,0)$ denotes its positive part.
- For a real number a, $\lfloor a \rfloor$ denotes the closest integer that is not greater than a.
- For sets $A,B\subseteq\mathbb{R}^k$ in k-dimensional space, $A\oplus B := \{a+b : a\in A, b\in B\}$ denotes the direct sum of A and B.
- With a little abuse of notations, for $x,y\in\mathbb{F}_q$, $x\oplus y$ denotes the modulo-q sum of x and y.
- Unless specified, all the logarithms $\log(.)$ are of base 2.

4.3 Linear Deterministic Interference Channel with Conferencing Transmitters

As discussed in Section 4.1, we shall first study an auxiliary problem, *linear deterministic* interference channel with conferencing transmitters, to overcome the complications both in achievability and outer bounds.

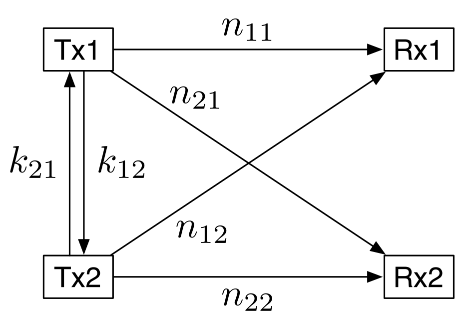
The corresponding linear deterministic channel (LDC) is parametrized by nonnegative integers n_{11}, n_{21}, n_{22}, n_{12}, k_{12}, and k_{21}, where

$$n_{ij} := \left(\lfloor \log|h_{ij}|^2 \rfloor\right)^+,\ i,j\in\{1,2\}$$

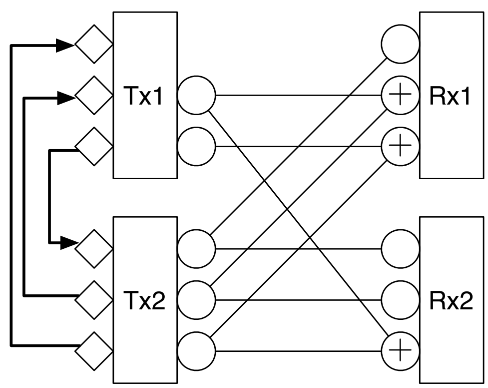
correspond to the channel gains in logarithmic-two scale, and

$$k_{12} := \lfloor \mathsf{C}^{\mathsf{B}}_{12} \rfloor, \; k_{21} := \lfloor \mathsf{C}^{\mathsf{B}}_{21} \rfloor$$

correspond to the cooperative link capacities. An illustration is depicted in Fig. 4.2(a) along with an example in Fig. 4.2(b). Each circle or diamond represents a bit. The bit emitting from a single circle at transmitters will broadcast noiselessly through the edges to the circles at receivers. Multiple incoming bits at a circle are summed up using modulo-two addition and produce a single received bit. The diamonds represent the bits exchanged between transmitters. In Fig. 4.2(b), Tx1 can send one bit to Tx2, and Tx2 can send two bits to Tx1. For more details about this model, we point the readers to reference [34] [9] [10].



(a) Channel Model



(b) Example Channel

Figure 4.2: Linear Deterministic Interference Channel with Conferencing Transmitters.

The following theorem characterizes the capacity region of this channel.

Theorem 4.1. *Nonnegative* (R_1, R_2) *is achievable if and only if it satisfies the following:*

$$R_1 \le \min\left\{ \max(n_{11}, n_{12}), \; n_{11} + k_{12} \right\}$$

$$R_2 \le \min\left\{ \max(n_{22}, n_{21}), \; n_{22} + k_{21} \right\}$$

$$R_1 + R_2 \le (n_{11} - n_{21})^+ + \max(n_{22}, n_{21}) + k_{12} \tag{4.1}$$

$$R_1 + R_2 \le (n_{22} - n_{12})^+ + \max(n_{11}, n_{12}) + k_{21} \tag{4.2}$$

$$R_1 + R_2 \le \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} + k_{12} + k_{21} \tag{4.3}$$

$$R_1 + R_2 \le \begin{cases} \max\{n_{11} + n_{22}, n_{12} + n_{21}\}, & \text{if } n_{11} + n_{22} \ne n_{12} + n_{21} \\ \max\{n_{11}, n_{12}, n_{21}, n_{22}\}, & \text{if } n_{11} + n_{22} = n_{12} + n_{21} \end{cases} \tag{4.4}$$

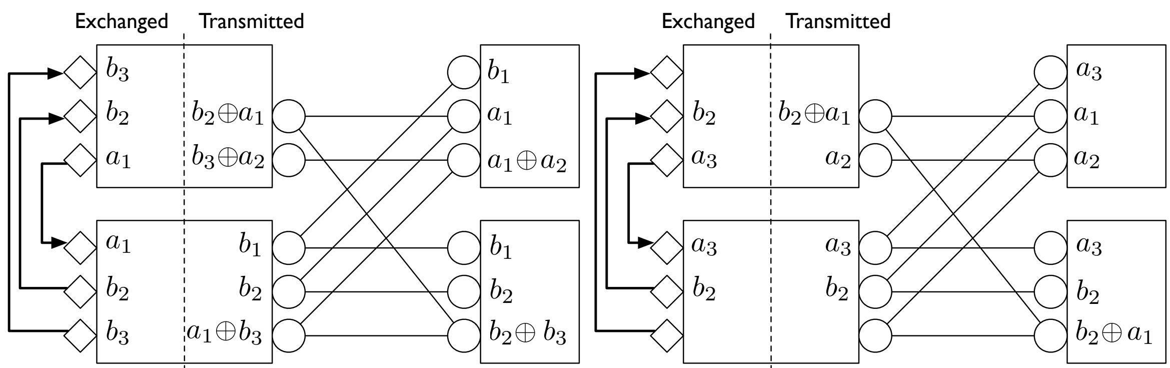
$$2R_1 + R_2 \le \max(n_{11}, n_{12}) + \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} + (n_{11} - n_{21})^+ + k_{12} + k_{21} \tag{4.5}$$

$$R_1 + 2R_2 \le \max(n_{22}, n_{21}) + \max\left\{n_{12}, (n_{11} - n_{21})^+\right\} + (n_{22} - n_{12})^+ + k_{21} + k_{12}$$

$$2R_1 + R_2 \leq n_{21} + \max\left\{n_{11} + (n_{22} - n_{21})^+, n_{12}\right\} + (n_{11} - n_{21})^+ + k_{12} \tag{4.6}$$
$$R_1 + 2R_2 \leq n_{12} + \max\left\{n_{22} + (n_{11} - n_{12})^+, n_{21}\right\} + (n_{22} - n_{12})^+ + k_{21}$$

4.3.1 Motivating Examples

Before going into technical details of proving the achievability and outer bounds, we first give several examples to motivate the scheme as well as the outer bounds. In the discussions below, bit $a_i \in \mathbb{F}_2$ denotes an information bit for user 1, and similarly $b_i \in \mathbb{F}_2$ denotes an information bit for user 2. The index i denotes the i-th level from the most significant bit (MSB) at the user's transmitter. If i becomes larger than the total number of levels available at the user's transmitter, the corresponding bit has to be relayed to the final destination via the other transmitter, as we will see in the sequel.



(a) Achieving $R_1 = 2, R_2 = 3$ (b) Achieving $R_1 = 3, R_2 = 1$

Figure 4.3: Coding Strategies for Example Channel in Fig. 4.2

Achievability

The first example channel is depicted in Fig. 4.2(b), where $n_{11} = 2, n_{12} = 3, n_{21} = 1, n_{22} = 3, k_{12} = 1, k_{21} = 2$. We shall use this example to argue intuitively the need of *cooperative common* messages, shed some light on how cooperative messages should be encoded, and illustrate the two-fold usage of transmitter cooperation - nulling out interferences and relaying additional information.

To achieve the rate point $(R_1, R_2) = (2, 3)$, one simple strategy is depicted in Fig. 4.3(a). In this coding scheme, we identify the message structure in Table 4.1. Note that transmitter 2 sends b_2 and b_3 to transmitter 1 so that it can carry out proper precoding to null out interference b_2 and b_3 at receiver 1. Similarly transmitter 1 sends a_1 to transmitter 2 so that it can null out interference a_1 at receiver 2.

Table 4.1: Message Structure in Fig. 4.3(a)

Cooperative common	Cooperative private	Noncooperative common	Noncooperative private
None	a_1 and (b_2, b_3)	b_1	a_2

On the other hand, to achieve the rate point $(R_1, R_2) = (3, 1)$, one simple strategy is depicted in Fig. 4.3(b). In this coding scheme, we identify the message structure in Table 4.2. Note that to support a third bit a_3 for user 1, it has to occupy the topmost circle level

Table 4.2: Message Structure in Fig. 4.3(b)

Cooperative common	Cooperative private	Noncooperative common	Noncooperative private
a_3	b_2	a_1	a_2

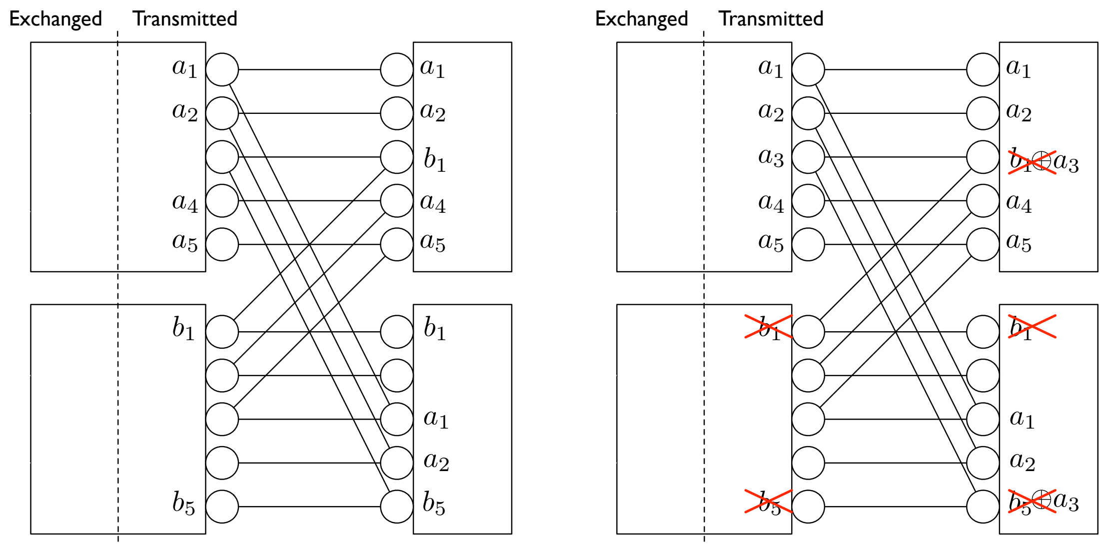
at transmitter 2 and both receivers, since the direct link from transmitter 1 to receiver 1 has only two levels. Hence, receiver 2 inevitably will decode bit a_3, which is then classified as cooperative common. From this example we see that *cooperative common* messages are needed, and their signal should occupy the levels that appear at both receivers cleanly. For the cooperative private parts, the example suggests that one should design precoding cleverly such that interference is nulled out at the unintended receiver. Based upon these intuitions, we propose an explicit scheme in Section 4.3.4.

In the above example, we can see that the usage of the cooperative links is two-fold: (1) null out interference, as in Fig. 4.3(a), and (2) relay additional bits, as the link from transmitter 1 to 2 in Fig. 4.3(b). This observation is also useful in motivating outer bounds.

Fundamental Tradeoff on $2R_1 + R_2$ and $R_1 + 2R_2$

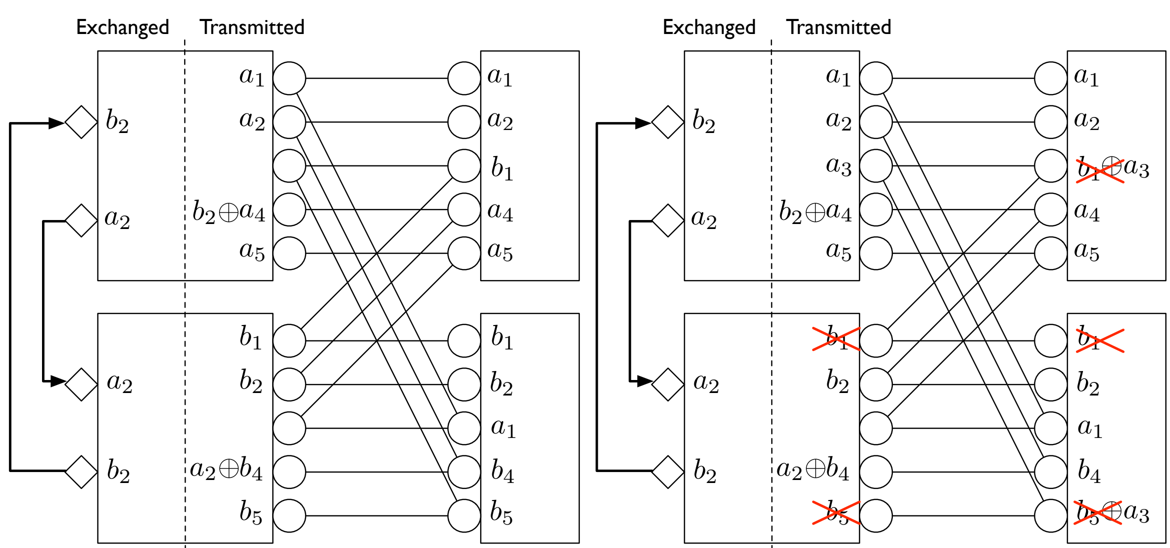
For outer bounds, the main difference from the interference channel *without* cooperation [14] [10] is that there are two different types of bounds on $2R_1 + R_2$ (and $R_1 + 2R_2$ correspondingly). Below we demonstrate the two different types of fundamental tradeoff on $2R_1 + R_2$ through two examples.

The first type of tradeoff does not involve the information that flows in the cooperative links. Consider the example channel with $n_{11} = n_{22} = 5$, $n_{12} = n_{21} = 3$, $k_{12} = k_{21} = 1$. We first consider the case *without* cooperation. Two corner points of the capacity region are $(R_1, R_2) = (4, 2)$ and $(5, 0)$, and the optimal strategies are depicted in Fig. 4.4(a) and (b) respectively. To enhance user 1's rate from 4 to 5 bits, the bit a_3 has to be turned on and causes collisions at the third level at receiver 1 and the fifth level at receiver 2. Transmitter 2 then has to turn off bit b_1 to avoid destroying bit a_3, and b_5 cannot be decoded since it is



(a) Achieving $R_1 = 4, R_2 = 2$ (b) Achieving $R_1 = 5, R_2 = 0$

Figure 4.4: Example Channel without Transmitter Coopertaion: Tradoff from $(R_1, R_2) = (4, 2)$ to $(5, 0)$



(a) Achieving $R_1 = 4, R_2 = 4$ (b) Achieving $R_1 = 5, R_2 = 2$

Figure 4.5: Example Channel with Transmitter Coopertaion: Tradoff from $(R_1, R_2) = (4, 4)$ to $(5, 2)$

corrupted by a_3. Now consider the case *with* cooperation. Two corner points of the capacity region are $(R_1, R_2) = (4, 4)$ and $(5, 2)$, and the optimal strategies are depicted in Fig. 4.5(a) and (b) respectively. Note that to enhance user 1's rate from 4 to 5 bits, again the bit a_3 has to be turned on and again causes collisions at the same places as in the case without cooperation. Note that the bits exchanged in the cooperative links remain the same, and hence the information that flows in the cooperative links is not involved in this tradeoff. Furthermore, the tradeoff is qualitatively the same as that in the case *without* cooperation. Later we will see that this type of outer bound on $2R_1 + R_2$ can be generalized from the $2R_1 + R_2$ bound in deterministic interference channel *without* cooperation [14], and the proof technique is quite similar.

The second type of tradeoff is a new phenomenon in interference channel with cooperation, and it involves the information that flows in the cooperative links. Consider the example channel in Fig. 4.2(b). The two rate points $(R_1, R_2) = (2, 3)$ and $(R_1, R_2) = (3, 1)$ are on the boundary of the capacity region. To enhance user 1's rate from 2 to 3 bits, since the number of levels from transmitter 1 to receiver 1 is only 2, the third bit a_3 has to be relayed from transmitter 2 to receiver 1. Hence, the topmost level at transmitter 2 has to be occupied by information exclusively for user 1, that is, a_3, and at receiver 2 the topmost level is no longer available for user 2. On the other hand, since the cooperative link from transmitter 1 to transmitter 2 is now occupied by a_3, the opportunity of nulling out the interference at the third level at receiver 2 is eliminated. As a consequence, the only available level for user 2 at receiver 2 is the second level, and user 2 has to back off its rate from 3 to 1. Note that the key difference from the first type of tradeoff is that, at the rate point $(R_1, R_2) = (2, 3)$ the cooperative link from transmitter 1 to 2 is used for *nulling out interference*, while at $(R_1, R_2) = (3, 1)$ it is used for *relaying additional bits*. Hence, the information that flows in the cooperative links *is* involved in this tradeoff, and the tradeoff is qualitatively different from that in the case *without* cooperation. As we will show later, to prove this type of outer bound on $2R_1 + R_2$, we need to develop a new technique for giving side information to the receivers.

4.3.2 Outer Bounds

To prove the converse part of Theorem 4.1, instead of giving full details of the proof[2], here we describe the techniques used in the proof. These techniques will be reused for proving outer bounds in the Gaussian problem.

1) Bounds on R_1 and R_2: These bounds are straightforward cut-set bounds.

2) Bounds on $R_1 + R_2$: Bound (4.4) is a standard cut-set bound. Its value is the rank of the *system transfer matrix*, with both transmit signals as input and both receive signals as output assuming full cooperation. It is quite straightforward to see that $n_{11} + n_{22} \neq n_{12} + n_{21}$

[2]We will provide full details when we deal with the Gaussian problem.

if and only if the system matrix is full rank, and the value of its rank is the right-hand side of (4.4).

Bound (4.1) is obtained by providing side information $\left(m_2, v_{12}^N\right)$ to receiver 1 so that receiver 1 is not interfered by transmitter 2 at all. This leads to the part $(n_{11} - n_{21})^+ + \max(n_{22}, n_{21})$, which is identical to the Z-channel bound in interference channel without cooperation. Giving the side information enhances the sum rate by at most $H\left(v_{12}^N|m_2\right) \leq Nk_{12}$ bits. Similar arguments works for bound (4.2).

Bound (4.3) is obtained by providing side information $\left(v_{12}^N, v_{21}^N, s_1^N\right)$ and $\left(v_{12}^N, v_{21}^N, s_2^N\right)$ to receiver 1 and 2 respectively, where s_1^N denotes the interference caused by transmitter 1 at receiver 2 (and vice versa for s_2^N). Giving side information $\left(v_{12}^N, v_{21}^N\right)$ to both receivers enhances the sum rate by at most $H\left(v_{12}^N, v_{21}^N\right) \leq N(k_{12} + k_{21})$ bits. Finally, we are able to prove the bound by making use of the Markov relations observed first in [18] for the MAC with conferencing transmitters, which states that given the conferencing signals, the transmit signals and messages at two transmitters are independent: $\left(m_i, x_i^N\right) - \left(v_{12}^N, v_{21}^N\right) - \left(m_j, x_j^N\right)$, for $(i,j) = (1,2)$ or $(2,1)$.

3) Bounds on $2R_1+R_2$ and R_1+2R_2: By symmetry we shall focus on the bounds on $2R_1+R_2$.

For linear deterministic interference channel *without* cooperation, the outer bound on $2R_1 + R_2$ is proved by first creating a copy of receiver 1 and then giving proper side information to these three receivers [14]. The side information structure is the following: give side information $\left(x_2^N, s_1^N\right)$ to one of the two receiver 1's and side information s_2^N to receiver 2.

As discussed in the previous section, there are two types of tradeoff on $2R_1 + R_2$. They correspond to bound (4.5) and bound (4.6) respectively. Bound (4.5) can be obtained via a similar technique as that in [14]. The side information structure is the following: give side information $\left(m_2, s_1^N, v_{12}^N, v_{21}^N\right)$ to one of the two receiver 1's, and $\left(s_2^N, v_{12}^N, v_{21}^N\right)$ to receiver 2. The role of $\left(m_2, v_{12}^N\right)$ is the same as x_2^N, and the additional side information $\left(v_{12}^N, v_{21}^N\right)$ is to make the transmitters conditionally independent. We then make use of the above Markov property to complete the proof.

Bound (4.6), which corresponds to the second type of tradeoff discussed earlier, is obtained by splitting receiver 2's signal into two parts: $y_2^N = \left(y_{2\alpha}^N, y_{2\beta}^N\right)$, where $y_{2\alpha}^N$ is the part of transmitter 2's signal that is not corrupted by s_1^N, the interference from transmitter 1. Then we apply a cut-set bound argument on one of the two receiver 1's and $y_{2\alpha}^N$, and provide side information $\left(m_2, v_{12}^N, s_1^N\right)$ to the other receiver 1. Fig. 4.6 provides an illustration. Since this kind of side information structure has not been reported in literature, we detail the proof below:

Proof. If (R_1, R_2) is achievable, by Fano's inequality,

$$
\begin{aligned}
&N\left(2R_1 + R_2 - \epsilon_N\right) \leq 2I\left(m_1; y_1^N\right) + I\left(m_2; y_2^N\right) \\
&\leq I\left(m_1; y_1^N, s_1^N|m_2, v_{12}^N\right) + I\left(m_1; v_{12}^N|m_2\right) + I\left(m_1; y_1^N\right) + I\left(m_2; y_{2\alpha}^N, y_{2\beta}^N\right) \\
&= H\left(x_1^N, s_1^N|m_2, v_{12}^N\right) + I\left(m_1; y_1^N\right) + I\left(m_2; y_{2\alpha}^N\right) + I\left(m_2; y_{2\beta}^N|y_{2\alpha}^N\right) + I\left(m_1; v_{12}^N|m_2\right)
\end{aligned}
$$

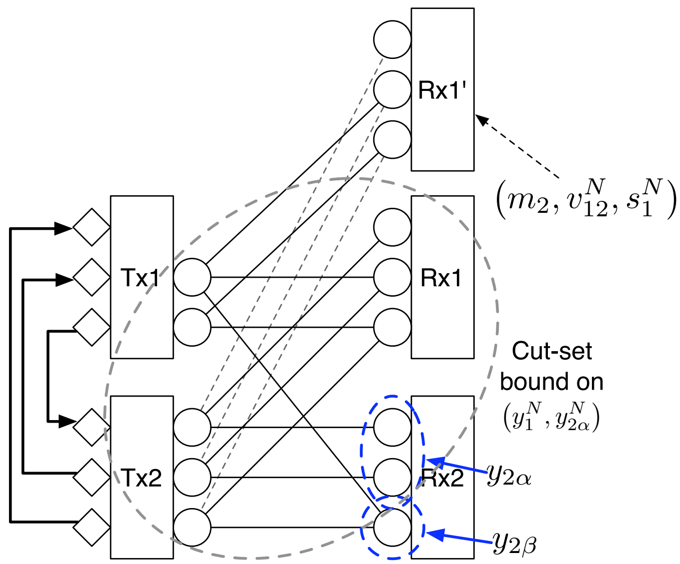


Figure 4.6: Side Information Structure for Outer Bound (4.6)

$$
\begin{aligned}
&\overset{(a)}{\leq} H\left(x_1^N, s_1^N|m_2, v_{12}^N\right) + I\left(m_1, m_2; y_1^N, y_{2\alpha}^N\right) + I\left(m_2, v_{12}^N; y_{2\beta}^N|y_{2\alpha}^N\right) + H\left(v_{12}^N|m_2\right)\\
&\overset{(b)}{\leq} H\left(x_1^N, s_1^N|m_2, v_{12}^N\right) + H\left(y_1^N, y_{2\alpha}^N\right) + H\left(y_{2\beta}^N\right) - H\left(y_{2\beta}^N|y_{2\alpha}^N, m_2, v_{12}^N\right) + H\left(v_{12}^N\right)\\
&\overset{(c)}{=} H\left(x_1^N, s_1^N|m_2, v_{12}^N\right) + H\left(y_1^N, y_{2\alpha}^N\right) + H\left(y_{2\beta}^N\right) - H\left(s_1^N|m_2, v_{12}^N\right) + H\left(v_{12}^N\right)\\
&= H\left(x_1^N|s_1^N, m_2, v_{12}^N\right) + H\left(y_1^N, y_{2\alpha}^N\right) + H\left(y_{2\beta}^N\right) + H\left(v_{12}^N\right)\\
&\leq N\left\{(n_{11} - n_{21})^+ + \max\left\{n_{11} + (n_{22} - n_{21})^+, n_{12}\right\} + n_{21} + k_{12}\right\}
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to a simple fact that $I\left(m_1; y_1^N\right) + I\left(m_2; y_{2\alpha}^N\right) \leq I\left(m_1, m_2; y_1^N, y_{2\alpha}^N\right)$. (b) is due to that conditioning reduces entropy. (c) holds since $y_{2\alpha}^N$ is a function of $\left(m_2, v_{12}^N\right)$. □

Let us revisit the example in Fig. 4.2(b) and demonstrate that bound (4.6) is active. Plugging the channel parameters into Theorem 4.1, we see that *without* bound (4.6), the region is

$$R_1 \leq 3,\ R_2 \leq 3,\ R_1 + R_2 \leq 5,$$

and the rate point $(3, 1)$ is not on its boundary. In this example, $y_{2\alpha}$ spans the topmost two levels at receiver 2. Hence, $H\left(y_1^N, y_{2\alpha}^N\right) \leq 4N$, and $2R_1 + R_2 \leq 1 + 4 + 1 + 1 = 7$ which is active in the capacity region:

$$R_1 \leq 3,\ R_2 \leq 3,\ R_1 + R_2 \leq 5,\ 2R_1 + R_2 \leq 7.$$

4.3.3 Achievability via Linear Reciprocity

Unlike the linear deterministic interference channel with conferencing receivers, it is not straightforward to directly show that linear strategies achieves the capacity in the case with conferencing transmitters. We can overcome this by using *linear reciprocity* of linear deterministic networks [48] and prove the achievability part of Theorem 4.1. We sketch the idea of the proof as follows.

First it is not hard to show that linear strategies are optimal for the reciprocal channel, that is, the linear deterministic interference channel with conferencing receivers. In such linear strategies, each user modulates its information bits (message) onto the transmit signal vector via a linear transformation. Each receiver, serving as a relay, linearly transforms its received signal and sends it to the other receiver through the finite-capacity link. Since the channel is linear and deterministic, the exchanged signals between receivers are again linear transformations of the transmit information bits. Finally, each receiver solves all its received linear equations of the transmit information bits (one set from the other receiver and the other from the transmitters) and recovers its desired message. Note that the decoding process is again a linear transformation. By choosing these linear transformations (encoding, relaying, and decoding) properly, the scheme achieves the capacity.

Next by linear reciprocity, we immediately show that the capacity region of the reciprocal channel (the linear deterministic interference channel with conferencing receivers) is an achievable region of the original channel. The strategy is again linear. Each transmitter sends a linear transformation of its information bits to the other transmitter through the finite-capacity link. Then it sends out a linear transformation of the received bits from the other transmitter and its own information bits to the receivers. Finally, each receiver solves the linear equations it receives to recover its desired message. It remains to show that this region coincides with that given in Theorem 4.1, which is a straightforward calculation.

Note that in such linear strategies, there is no need to split the messages at the transmitters, and the decoding process at the receivers can be viewed as *treating interference as noise*. This is first observed in Lecture Notes 6 in [49] for linear deterministic interference channels without cooperation. This implies that the complicated message structure described in Section 4.1 is not necessary for linear deterministic interference channel with conferencing receivers or transmitters.

To this end, there are two paths towards constructing good coding strategies in the Gaussian scenario. The first approach is deriving structured lattice strategies based on the capacity-achieving linear strategies of the corresponding linear deterministic channels. This approach, however, requires an explicit description of linear transformations in the capacity-achieving linear strategies for the LDC. The second approach is deriving Gaussian random coding strategies, which is the conventional approach for additive white Gaussian noise networks. In this chapter, we will take the second approach. For this purpose, however, the proof of achievability via linear reciprocity does not give much insight. Below we give an alternative proof of achievability, which provides guidelines for designing good Gaussian random coding strategies in the Gaussian interference channel with conferencing transmitters.

4.3.4 Alternative Proof of Achievability

To get a better handle to deal with the the design of good Gaussian random coding schemes, we propose a general coding strategy that applies to both the linear deterministic channel (LDC) and the Gaussian channel. The strategy is based on Marton's coding scheme for general broadcast channels [50] and superposition coding. It is described as follows: (Notations: subscript o stands for *cooperative common*, subscript h stands for *cooperative private*, subscript c stands for *noncooperative common*, and subscript p stands for *noncooperative private*.)

1) First, generate the cooperative common vector codeword $\underline{x}_o^N(m_{1o}, m_{2o})$ according to $p\left(\underline{x}_o^N\right) = \prod_{k=1}^N p\left(\underline{x}_o[k]\right)$. Denote $m_o := (m_{1o}, m_{2o})$.

2) Second, for each cooperative common m_o, generate the cooperative vector codeword $\underline{x}_{oh}^N(m_{1h}, m_{2h}, m_o)$ based on Marton's coding scheme according to conditional distribution $p\left(\underline{x}_{oh}^N, u_1^N, u_2^N | \underline{x}_o^N(m_o)\right) = \prod_{k=1}^N p\left(\underline{x}_{oh}[k], u_1[k], u_2[k] | \underline{x}_o(m_o)[k]\right)$, where the auxiliary codewords are $u_1^N(\widetilde{m}_{1h}, m_o)$ and $u_2^N(\widetilde{m}_{2h}, m_o)$.

3) Third, at transmitter i, for $i = 1, 2$, generate the noncooperative common codeword $x_{ic}^N(m_{ic})$ according to distribution $p\left(x_{ic}^N\right) = \prod_{k=1}^N p\left(x_{ic}[k]\right)$.

4) Fourth, at transmitter i, for each m_{ic} generate the noncooperative codeword $x_{icp}^N(m_{ip}, m_{ic})$ according to $p\left(x_{icp}^N | x_{ic}^N(m_{ic})\right) = \prod_{k=1}^N p\left(x_{icp}[k] | x_{ic}(m_{ic})[k]\right)$.

5) Finally, superimpose these two codewords to form the transmit codewords:

$$x_1^N(m_o, m_{1h}, m_{2h}, m_{1c}, m_{1p}) = \underline{x}_{oh}^N[1] + x_{1cp}^N$$
$$x_2^N(m_o, m_{1h}, m_{2h}, m_{2c}, m_{2p}) = \underline{x}_{oh}^N[2] + x_{2cp}^N.$$

Remark 4.2. *Note that in Step 4), say at transmitter 1, we can use Gelfand-Pinsker coding (dirty paper coding) to generate the noncooperative private codeword so that it can be protected against known interference at transmitter 1, which is caused by the cooperative private auxiliary codeword of the other user, that is, u_2. Throughout the chapter, however, we will choose (u_1, u_2) cleverly such that the effect of u_2 is zero-forced* exactly *in LDC and* approximately *in the Gaussian setting, and hence Gelfand-Pinsker coding does not provide significant improvement.*

For decoding, receiver 1 looks for a unique message tuple $(m_o, \widetilde{m}_{1h}, m_{1c}, m_{1p})$ such that

$$\left(y_1^N, \underline{x}_o^N(m_o), u_1^N(\widetilde{m}_{1h}, m_o), x_{1c}^N(m_{1c}), x_{1cp}^N(m_{1p}, m_{1c}), x_{2c}^N(\widehat{m}_{2c})\right)$$

is jointly typical, for some $\widehat{m}_{2c}$. Receiver 2 uses the same decoding rule with index 1 and 2 exchanged.

Based on the above strategy, we have the following coding theorem:

Theorem 4.3 (Achievable Rates)**.** *A nonnegative rate tuple* $(R_{1o}, R_{1h}, R_{1c}, R_{1p}, R_{2o}, R_{2h}, R_{2c}, R_{2p})$ *is achievable if it satisfies the following for some nonnegative* $(\widetilde{R}_{1h}, \widetilde{R}_{2h})$*: (denote* $R_o := R_{1o} + R_{2o}$ *)*

Constraints at Receiver 1:

$$\begin{aligned}
R_{1p} &\le I\left(x_{1cp}; y_1 | x_{1c}, x_{2c}, u_1, \underline{x}_o\right)\\
\widetilde{R}_{1h} &\le I\left(u_1; y_1 | x_{1cp}, x_{1c}, x_{2c}, \underline{x}_o\right)\\
\widetilde{R}_{1h} + R_{1p} &\le I\left(u_1, x_{1cp}; y_1 | x_{1c}, x_{2c}, \underline{x}_o\right)\\
R_{2c} + R_{1p} &\le I\left(x_{2c}, x_{1cp}; y_1 | x_{1c}, u_1, \underline{x}_o\right)\\
R_{1c} + R_{1p} &\le I\left(x_{1c}, x_{1cp}; y_1 | x_{2c}, u_1, \underline{x}_o\right)\\
R_{2c} + \widetilde{R}_{1h} &\le I\left(x_{2c}, u_1; y_1 | x_{1cp}, x_{1c}, \underline{x}_o\right)\\
R_{2c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(x_{2c}, u_1, x_{1cp}; y_1 | x_{1c}, \underline{x}_o\right)\\
R_{1c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(x_{1c}, x_{1cp}, u_1; y_1 | x_{2c}, \underline{x}_o\right)\\
R_{1c} + R_{2c} + R_{1p} &\le I\left(x_{1c}, x_{1cp}, x_{2c}; y_1 | u_1, \underline{x}_o\right)\\
R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(x_{1c}, x_{1cp}, x_{2c}, u_1; y_1 | \underline{x}_o\right)\\
R_o + \widetilde{R}_{1h} &\le I\left(\underline{x}_o, u_1; y_1 | x_{1cp}, x_{1c}, x_{2c}\right)\\
R_o + \widetilde{R}_{1h} + R_{1p} &\le I\left(\underline{x}_o, u_1, x_{1cp}; y_1 | x_{1c}, x_{2c}\right)\\
R_o + R_{2c} + \widetilde{R}_{1h} &\le I\left(\underline{x}_o, x_{2c}, u_1; y_1 | x_{1cp}, x_{1c}\right)\\
R_o + R_{2c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(\underline{x}_o, x_{2c}, u_1, x_{1cp}; y_1 | x_{1c}\right)\\
R_o + R_{1c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(\underline{x}_o, x_{1c}, x_{1cp}, u_1; y_1 | x_{2c}\right)\\
R_o + R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} &\le I\left(\underline{x}_o, x_{1c}, x_{1cp}, x_{2c}, u_1; y_1\right)
\end{aligned}$$

Constraints at Receiver 2: Above with index 1 and 2 exchanged.

Constraints at Transmitters:

$$\begin{aligned}
R_{1h} &\le \widetilde{R}_{1h}\\
R_{2h} &\le \widetilde{R}_{2h}\\
R_{1o} + R_{1h} &\le \mathsf{C}^{\mathsf{B}}_{12}\\
R_{2o} + R_{2h} &\le \mathsf{C}^{\mathsf{B}}_{21}\\
\widetilde{R}_{1h} + \widetilde{R}_{2h} - R_{1h} - R_{2h} &\ge I\left(u_1; u_2 | \underline{x}_o\right),
\end{aligned}$$

for some $(u_1, u_2) - \underline{x}_{oh} - (y_1, y_2)$*,* $x_1 = x_{oh}[1] + x_{1cp}$*,* $x_2 = x_{oh}[2] + x_{2cp}$*, and*

$$\begin{aligned} &p\left(\underline{x}_{oh}, \underline{x}_o, u_1, u_2, x_{1c}, x_{1cp}, x_{2c}, x_{2cp}\right) \\ &= p\left(\underline{x}_o\right) p\left(\underline{x}_{oh}, u_1, u_2 | \underline{x}_o\right) p\left(x_{1c}\right) p\left(x_{1cp} | x_{1c}\right) p\left(x_{2c}\right) p\left(x_{2cp} | x_{2c}\right). \end{aligned}$$

Proof. The proof is quite straightforward. It involves standard error probability analysis of superposition coding and Marton's coding scheme, and hence is omitted here. Note we have in total 5 independent messages to be decoded at each receiver, and hence in general there should be $2^5 - 1 = 31$ inequalities. However, say at receiver 1, decoding m_{2c} incorrectly is not accounted as an error. Furthermore due to the superposition coding of $\widetilde{m}_{1h}$ upon $\underline{x}_o^N$ and the superposition coding of m_{1p} upon x_{1c}^N, we remove the inequality on R_{2c} and the $2^3 + 2^3 - 2 = 14$ inequalities involving R_o but not $\widetilde{R}_{1h}$ or involving R_{1c} but not R_{1p}. Hence in total we have $31 - 1 - 14 = 16$ inequalities at each receiver. □

Below we show that with proper choices of distribution $p\left(\underline{x}_{oh}, \underline{x}_o, u_1, u_2, x_{1c}, x_{1cp}, x_{2c}, x_{2cp}\right)$, the above coding strategy can achieve the capacity region of LDC. We shall distinguish into two cases: (1) system transfer matrix is full-rank, and (2) system transfer matrix is not full-rank.

System matrix is full-rank: $n_{11} + n_{22} \neq n_{12} + n_{21}$

In this case, for the cooperative part, we shall set $\underline{x}_o$ to be running over all transmit levels, and choose $(u_1, u_2) | \underline{x}_o \overset{\text{d}}{=} (y_{1h}, y_{2h})$ occupying the following numbers of least significant bits (LSB) at receiver 1 and 2 respectively:

$$g_1 = \max\left\{n_{11} - (n_{21} - n_{22})^+, n_{12} - (n_{22} - n_{21})^+\right\} \tag{4.7}$$

$$g_2 = \max\left\{n_{22} - (n_{12} - n_{11})^+, n_{21} - (n_{11} - n_{12})^+\right\} \tag{4.8}$$

Then we choose $\underline{x}_h$ occupying the levels at transmitters so that it results in (y_{1h}, y_{2h}) at receivers. The cooperative codeword is generated according to the distribution of $\underline{x}_{oh} \overset{\text{d}}{=} \underline{x}_o + \underline{x}_h$. The addition here is bit-wise modulo-two. We observe the following:

Claim 4.4. *Under i.i.d. Bernoulli half inputs, mutual information* $I\left(u_1; u_2 | \underline{x}_o\right) = 0$ *with the above choice if* $n_{11} + n_{22} \neq n_{12} + n_{21}$*, and hence* u_1 *and* u_2 *are independent conditioned on* $\underline{x}_o$*.*

Remark 4.5 (Comments on Claim 4.4)**.** *Intuitively speaking, choosing* $p\left(u_1, u_2, \underline{x}_o\right)$ *as above has an effect that the interference caused by the other user's cooperative private signal is nulled out at the target receiver. The reason is that, the received signal contributed by the cooperative private signals at a receiver, say, receiver 1, is* $y_{1h} = u_1$*, and the conditional independence of* u_1 *and* u_2 *given cooperative common signal* $\underline{x}_o$ *implies that within* y_{1h} *there is no dependency on* u_2*, and hence, the interference caused by* u_2 *is nulled out.*

Proof. For the case $\{n_{11} \geq n_{12}, n_{22} \geq n_{21}\}$, (g_1, g_2) becomes

$$g_1 = \max\left\{n_{11} - (n_{21} - n_{22})^+, n_{12} - (n_{22} - n_{21})^+\right\} = n_{11}$$
$$g_2 = \max\left\{n_{22} - (n_{12} - n_{11})^+, n_{21} - (n_{11} - n_{12})^+\right\} = n_{22}$$

and the rank of the full system transfer matrix is $n_{11} + n_{22} = g_1 + g_2$. Hence, under i.i.d. Bernoulli half inputs,

$$\begin{aligned} I(u_1; u_2|\underline{x}_o) &= H(u_1|\underline{x}_o) + H(u_2|\underline{x}_o) - H(u_1, u_2|\underline{x}_o) \\ &= n_{11} + n_{22} - (n_{11} + n_{22}) = 0. \end{aligned}$$

Similar argument works for the case $n_{11} \leq n_{12}$, $n_{22} \leq n_{21}$.

For the case $\{n_{11} \leq n_{12}, n_{22} \geq n_{21}\}$, (g_1, g_2) becomes

$$g_1 = \max(n_{11}, n_{12} + n_{21} - n_{22})$$
$$g_2 = \max(n_{11} + n_{22} - n_{12}, n_{21})$$

and the rank of the transfer matrix is $g_1 + g_2$ again, since the subsystem (lies in the original system with (y_{1h}, y_{2h}) as output and the corresponding levels at transmitters as input) has channel parameters

$$n'_{11} = n_{11}, n'_{21} = n_{21}, n'_{12} = g_1, n'_{22} = g_2.$$

If $n_{11} + n_{22} > n_{12} + n_{21}$, then $g_1 = n_{11}$ and $g_2 = n_{11} + n_{22} - n_{12}$, and hence $n'_{11} = n'_{12}$, $n'_{22} > n'_{21}$.

Similarly, if $n_{11} + n_{22} < n_{12} + n_{21}$, then $g_1 = n_{12} + n_{21} - n_{22}$ and $g_2 = n_{21}$, and hence $n'_{11} < n'_{12}$, $n'_{22} = n'_{21}$.

Similar argument works for the case $n_{11} \geq n_{12}$, $n_{22} \leq n_{21}$.

Therefore, under i.i.d. Bernoulli half inputs, u_1 and u_2 are independent conditioned on $\underline{x}_o$ from the analysis of the previous cases. □

For the noncooperative part, we set $x_{1cp} \overset{\mathrm{d}}{=} x_{1c} + x_{1p}$ such that x_{1c} and x_{1p} are independent. x_{1c} is allowed to occupy all levels at transmitter 1, while x_{1p} is allowed to occupy only the $(n_{11} - n_{21})^+$ LSB's. The addition here is bit-wise modulo-two. The same design is applied to user 2. As commented in Remark 4.5, with the above choice of y_{1h} and y_{2h} in (4.7) and (4.8), there are no redundancy between y_{1h} and y_{2h} and hence no interference from u_2 at receiver 1. Similar situation happens at receiver 2. Now take all inputs to be i.i.d. Bernoulli half across levels, we obtain a set of achievable rates from Theorem 4.3. After Fourier-Motzkin elimination, we show that the achievable region coincides with the region given in Theorem 4.1.

Lemma 4.6. *The above strategy achieves the region given in Theorem 4.1 when* $n_{11} + n_{22} \neq n_{12} + n_{21}$.

Proof. The details are left in Chapter 5.1. □

System matrix is not full-rank: $n_{11} + n_{22} = n_{12} + n_{21}$

In this case, for the cooperative part, we shall again set $\underline{x}_o$ to be running over all transmit levels. The difference lies in the *cooperative private* part. Here we also choose $(u_1, u_2)\,|\underline{x}_o \overset{\mathrm{d}}{=} (y_{1h}, y_{2h})$, but occupying the following numbers of LSB's at receiver 1 and 2 respectively:

$$g_1 = (n_{11} - n_{21})^+ \,,\ g_2 = (n_{22} - n_{12})^+ \,.$$

Claim 4.7. *Under i.i.d. Bernoulli half inputs, mutual information $I(u_1; u_2|\underline{x}_o) = 0$ with the above choice if $n_{11} + n_{22} = n_{12} + n_{21}$.*

Proof. Since y_{ih} only occupies levels that appear at receiver i but do not appear at the other receiver, for $i = 1, 2$, hence they are conditionally independent given $\underline{x}_o$ under Bernoulli half i.i.d. inputs. □

For the noncooperative part, we use the same scheme as the previous case. Now take all inputs to be i.i.d. Bernoulli half across levels, we obtain achievable rates from Theorem 4.3. After Fourier-Motzkin elimination, we have the following lemma:

Lemma 4.8. *The above strategy achieves the region given in Theorem 4.1 when $n_{11} + n_{22} = n_{12} + n_{21}$.*

Proof. The details are left in Chapter 5.1. □

We conclude this section with the following remark.

Remark 4.9 (Implications on the Gaussian Problem)**.** *The two numbers g_1 and g_2 provide clues in determining the power allocated to the cooperative private codewords and the design of beamforming vectors in the Gaussian scenario. Take user 1 as an example. When $n_{11}+n_{22} \neq n_{12} + n_{21}$,*

$$\begin{aligned} g_1 &= \max\left\{n_{11} - (n_{21} - n_{22})^+ , n_{12} - (n_{22} - n_{21})^+\right\} \\ &= \max\left\{n_{11} + n_{22}, n_{12} + n_{21}\right\} - \max\left\{n_{22}, n_{21}\right\}, \end{aligned}$$

which corresponds to $\frac{|h_{11}h_{22} - h_{12}h_{21}|^2}{1+\mathsf{SNR}_2+\mathsf{INR}_2}$. On the other hand, when $n_{11} + n_{22} = n_{12} + n_{21}$,

$$g_1 = (n_{11} - n_{21})^+ = (n_{12} - n_{22})^+ \,,$$

which corresponds to $\frac{\mathsf{SNR}_1}{\mathsf{INR}_2} = \frac{\mathsf{INR}_1}{\mathsf{SNR}_2} = \frac{\mathsf{SNR}_1+\mathsf{INR}_1}{\mathsf{SNR}_2+\mathsf{INR}_2}$. This implies that the power of u_1 conditioned on $\underline{x}_o$ should be proportional to

$$\frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + \mathsf{SNR}_1 + \mathsf{INR}_1}{1 + \mathsf{SNR}_2 + \mathsf{INR}_2} = \frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + |h_{11}|^2 + |h_{22}|^2}{1 + \mathsf{SNR}_2 + \mathsf{INR}_2},$$

and that the beamforming vector should be a combination of zero-forcing and matched-filter vectors.

In the next section, we will incorporate the insights obtained in analyzing the linear deterministic interference channel with conferencing transmitters to derive capacity results for the Gaussian interference channel with conferencing transmitters, and show that the inner and outer bounds are within a bounded gap.

4.4 Gaussian Interference Channel with Conferencing Transmitters

With the full understanding in the linear deterministic interference channel with conferencing transmitters, now we have enough clues to crack the original Gaussian problem. As for the outer bounds, we shall mimic the genie-aided techniques and the structure of side informations in Section 4.3.2 to develop the proofs. As for the achievability, we shall mimic the choice of auxiliary random variables and level allocation in Section 4.3.4 to construct good schemes in the Gaussian scenario. Moreover, the achievable rate regions obtained prior to Fourier-Motzkin elimination can be made equivalent symbolically, and hence the proof of achieving approximate capacity in the Gaussian channel follows closely to the proof of achieving exact capacity in the linear deterministic channel. Although the Gaussian interference channel with conferencing transmitters and its corresponding linear deterministic channel are strongly related in coding strategies, proof of achievability, and outer bounds, unlike the two-user Gaussian interference channel [10], their capacity regions are *not* within a bounded gap. Similar situations happen in MIMO channel, Gaussian relay networks [9], and the Gaussian interference channel with conferencing receivers [15], where explicit/implicit MIMO structures lie in the channel model.

Our main result is summarized in the following lemma and theorem.

Lemma 4.10 (Outer Bounds). *If (R_1, R_2) is achievable, it satisfies the following constraints:*

$$R_1 \le \min\left\{\log\left(1+\mathsf{SNR}_1\right)+\mathsf{C}^{\mathsf{B}}_{12}, \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)\right\} \tag{4.9}$$

$$R_2 \le \min\left\{\log\left(1+\mathsf{SNR}_2\right)+\mathsf{C}^{\mathsf{B}}_{21}, \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}\right)\right\} \tag{4.10}$$

$$R_1+R_2 \le \log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}\right)+\mathsf{C}^{\mathsf{B}}_{12} \tag{4.11}$$

$$R_1+R_2 \le \log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)+\mathsf{C}^{\mathsf{B}}_{21} \tag{4.12}$$

$$R_1+R_2 \le \left\{\begin{array}{l} \log\left(1+\frac{\mathsf{SNR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}}{1+\mathsf{INR}_2}+\mathsf{INR}_1\right)+\mathsf{C}^{\mathsf{B}}_{12} \\ +\log\left(1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)+\mathsf{C}^{\mathsf{B}}_{21} \end{array}\right\} \tag{4.13}$$

$$R_1+R_2 \le \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2 \\ +2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}+|h_{11}h_{22}-h_{12}h_{21}|^2 \end{array}\right) \tag{4.14}$$

$$2R_1 + R_2 \leq \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) \\ +\log\left(1+\frac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)+\mathsf{C}^{\mathsf{B}}_{12}+\mathsf{C}^{\mathsf{B}}_{21} \end{array} \right\} \tag{4.15}$$

$$R_1 + 2R_2 \leq \left\{ \begin{array}{l} \log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) \\ +\log\left(1+\frac{\mathsf{SNR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}}{1+\mathsf{INR}_2}+\mathsf{INR}_1\right)+\mathsf{C}^{\mathsf{B}}_{21}+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\} \tag{4.16}$$

$$2R_1 + R_2 \leq \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{SNR}_2 \\ +\mathsf{INR}_1\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{INR}_2+2\left(1+\mathsf{INR}_2\right)\sqrt{\mathsf{SNR}_1\mathsf{INR}_1} \end{array} \right) \\ +\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+1+\mathsf{C}^{\mathsf{B}}_{12} \end{array} \right\} \tag{4.17}$$

$$R_1 + 2R_2 \leq \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+\mathsf{SNR}_+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{SNR}_2 \\ +\mathsf{INR}_1\mathsf{INR}_2+\mathsf{SNR}_2\mathsf{INR}_1+2\left(1+\mathsf{INR}_1\right)\sqrt{\mathsf{SNR}_2\mathsf{INR}_2} \end{array} \right) \\ +\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+1+\mathsf{C}^{\mathsf{B}}_{21} \end{array} \right\} \tag{4.18}$$

Theorem 4.11 (Bounded Gap to Capacity)**.** *Outer bounds in Lemma 4.10 is within* $\log 90 \approx 6.5$ *bits per user to the capacity region.*

4.4.1 Outer Bounds

Details of the proof of Lemma 4.10 are left in Chapter 5.4. It follows closely to the techniques we develop in the proofs of the LDC outer bounds. The only twist is how to mimic the proof of bound (4.6), which is a new type of outer bound that does not appear in the case without cooperation. It corresponds to bound (4.17) here. Recall that in the proof there, we split receiver 2's signal into two parts: $y_2^N = \left(y_{2\alpha}^N, y_{2\beta}^N\right)$, where $y_{2\alpha}^N$ is the part of transmitter 2's signal that is not corrupted by s_1^N, the interference from transmitter 1. Such split is not possible in the Gaussian channel due to additive noise and carry-over in real addition. As shown in Chapter 5.4, we will overcome this by providing the following side information to receiver 2:

$$\widetilde{y}_2^N := h_{22} x_2^N + \widetilde{z}_2^N,$$

where $\widetilde{z}_2 \sim \mathcal{CN}\left(0, 1+\mathsf{INR}_2\right)$, i.i.d. over time and is independent of everything else. This mimics the signal $y_{2\alpha}^N$ in LDC, and helps us prove the $2R_1 + R_2$ outer bound.

4.4.2 Coding Strategy and Achievable Rates

We shall employ the coding strategy proposed in Section 4.3.4. The analysis in the linear deterministic setting suggests that, for the cooperative private messages, in the Gaussian setting one may choose its bearing auxiliary random variables u_1 and u_2 to be conditionally independent given $\underline{x}_o$. This implies that a simple linear beamforming strategy is sufficient.

On the other hand, the interference should be zero-forced approximately. Based on this observation, we implement the following strategy.

For the cooperative common signal, recall that in the LDC we allow $\underline{x}_o$ to run over all transmit levels. To mimic it, in the Gaussian setting we choose $\underline{x}_o$ to be Gaussian with zero mean and a covariance matrix which has diagonal entries (values of transmit power) that are comparable with the total transmit power. For simplicity, we choose the covariance matrix to be diagonal:

$$K_{\underline{x}_o} = \mathrm{diag}\left(Q_{1o}, Q_{2o}\right), \; Q_{io} = 1/4, \; i = 1, 2.$$

Here the value 1/4 is just a heuristic choice such that the transmit power constraints will be satisfied.

For the cooperative private signal, from the discussion in Remark 4.9, we shall make it a superposition of *zero-forcing* vectors

$$\underline{v}_{1z} = \begin{bmatrix} h_{22} \\ -h_{21} \end{bmatrix}, \; \underline{v}_{2z} = \begin{bmatrix} -h_{12} \\ h_{11} \end{bmatrix}$$

and *matched-filter* vectors

$$\underline{v}_{1m} = \begin{bmatrix} h_{11}^* \\ h_{12}^* \end{bmatrix}, \; \underline{v}_{2m} = \begin{bmatrix} h_{21}^* \\ h_{22}^* \end{bmatrix}.$$

For the auxiliary random variables u_1 and u_2, we make them distributed as identical copies of user 1 and user 2's desired cooperative signals received at receiver 1 and 2 respectively. For example, u_1 would be the sum of the transmit cooperative common signal and user 1's cooperative private signal projected onto the channel vector $[h_{11} \; h_{12}]$.

Hence we choose $(\underline{x}_{oh}, \underline{x}_o, u_1, u_2)$ be jointly Gaussian such that

$$\begin{aligned} \underline{x}_{oh} &\overset{\mathrm{d}}{=} \underline{x}_o + \underbrace{w_{1z}\underline{v}_{1z} + w_{2z}\underline{v}_{2z} + w_{1m}\underline{v}_{1m} + w_{2m}\underline{v}_{2m}}_{\underline{x}_h} \\ u_1 &\overset{\mathrm{d}}{=} [h_{11} \; h_{12}] \left(\underline{x}_o + \underline{v}_{1z}w_{1z} + \underline{v}_{1m}w_{1m}\right) \\ u_2 &\overset{\mathrm{d}}{=} [h_{21} \; h_{22}] \left(\underline{x}_o + \underline{v}_{2z}w_{2z} + \underline{v}_{2m}w_{2m}\right), \end{aligned}$$

where w_{1z}, w_{2z}, w_{1m}, and w_{2m} are independent Gaussians and independent of everything else, with variances θ_{1z}, θ_{2z}, θ_{1m}, and θ_{2m} respectively. Their values are chosen such that the total transmit power constraint will be met and the conditional variances of u_1 and u_2 conditioned on $\underline{x}_o$ behave as we predicted in Remark 4.9. With this guideline, we choose

$$\theta_{1z} = \frac{1/4}{(1 + \mathsf{SNR}_2 + \mathsf{INR}_2)}$$

$$\theta_{1m} = \frac{1/4}{(\mathsf{SNR}_1 + \mathsf{INR}_1)(1 + \mathsf{SNR}_2 + \mathsf{INR}_2)}$$
$$\theta_{2z} = \frac{1/4}{(1 + \mathsf{SNR}_1 + \mathsf{INR}_1)}$$
$$\theta_{2m} = \frac{1/4}{(\mathsf{SNR}_2 + \mathsf{INR}_2)(1 + \mathsf{SNR}_1 + \mathsf{INR}_1)}.$$

Again, the factor $1/4$ is just a heuristic choice such that the transmit power constraints will be satisfied.

For the noncooperative part, we set $x_{ic} \sim \mathcal{CN}(0, Q_{ic})$, where $Q_{ic} = 1/4 - Q_{ip}$, for $i = 1, 2$. $x_{icp} \overset{\mathrm{d}}{=} x_{ic} + x_{ip}$, where $x_{ip} \sim \mathcal{CN}(0, Q_{ip})$ is independent of x_{ic} and $Q_{ip} = \min(1/4, 1/\mathsf{INR}_j)$, for $(i, j) = (1, 2)$ or $(2, 1)$. The choice of Q_{ip} is such that the interference caused by the other user's *noncooperative* private signal is at or below the noise level at the receiver.

At this stage, we shall check that the total transmit power constraint is met with the above heuristic choices of factors. We only need to show that the power for $\underline{x}_h$ at each transmitter is at most $1/2$, which is pretty straightforward[3].

Note that the variances of u_1 and u_2 conditioned on $\underline{x}_o$ are

$$K_{u_1|\underline{x}_o} = \frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + \mathsf{SNR}_1 + \mathsf{INR}_1}{4(1 + \mathsf{SNR}_2 + \mathsf{INR}_2)},$$
$$K_{u_2|\underline{x}_o} = \frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + \mathsf{SNR}_2 + \mathsf{INR}_2}{4(1 + \mathsf{SNR}_1 + \mathsf{INR}_1)},$$

matching our prediction in Remark 4.9.

With this encoding, the interference caused by the other user's *cooperative* private signal should be nulled out approximately, that is, its variance is at or below the noise level. To see this, the received signals are jointly distributed with $(\underline{x}_o, u_1, u_2, x_{1c}, x_{1cp}, x_{2c}, x_{2cp})$ such that

$$y_1 \overset{\mathrm{d}}{=} u_1 + \widehat{z}_1 + h_{11}x_{1cp} + h_{12}x_{2cp} + z_1$$
$$y_1 \overset{\mathrm{d}}{=} u_2 + \widehat{z}_2 + h_{21}x_{1cp} + h_{22}x_{2cp} + z_2,$$

where the interferences caused by undesired cooperative private signals are

$$\widehat{z}_1 = (h_{11}h_{21}^* + h_{12}h_{22}^*)\, w_{2m},\ \ \widehat{z}_2 = (h_{21}h_{11}^* + h_{22}h_{12}^*)\, w_{1m},$$

at receiver 1 and 2 respectively. Note that the variance of these terms are upper bounded by a constant, since

$$|h_{11}h_{21}^* + h_{12}h_{22}^*|^2 = |h_{21}h_{11}^* + h_{22}h_{12}^*|^2$$

[3]We have $Q_{1h} = \frac{\mathsf{SNR}_2 + \frac{\mathsf{SNR}_1}{\mathsf{SNR}_1 + \mathsf{INR}_1}}{4(1+\mathsf{SNR}_2+\mathsf{INR}_2)} + \frac{\mathsf{INR}_1 + \frac{\mathsf{INR}_2}{\mathsf{SNR}_2 + \mathsf{INR}_2}}{4(1+\mathsf{SNR}_1+\mathsf{INR}_1)} \leq \frac{1}{2}$, and vice versa for Q_{2h}.

$$= (\mathsf{SNR}_1 + \mathsf{INR}_1)(\mathsf{SNR}_2 + \mathsf{INR}_2) - |h_{11}h_{22} - h_{12}h_{21}|^2$$
$$\leq (\mathsf{SNR}_1 + \mathsf{INR}_1)(\mathsf{SNR}_2 + \mathsf{INR}_2).$$

Hence,

$$\sigma_1^2 := \mathrm{Var}(\widehat{z}_1) = \frac{|h_{11}h_{21}^* + h_{12}h_{22}^*|^2}{4(\mathsf{SNR}_1 + \mathsf{INR}_1)(1 + \mathsf{SNR}_2 + \mathsf{INR}_2)} \leq \frac{1}{4}$$
$$\sigma_2^2 := \mathrm{Var}(\widehat{z}_2) = \frac{|h_{21}h_{11}^* + h_{22}h_{12}^*|^2}{4(1 + \mathsf{SNR}_1 + \mathsf{INR}_1)(\mathsf{SNR}_2 + \mathsf{INR}_2)} \leq \frac{1}{4}$$

and in effect the interference is nulled out approximately.

Remark 4.12. *When the cooperative link capacities are sufficiently large and the channel becomes a two-user Gaussian MIMO broadcast channel with two transmit antennas and single receive antenna at each receiver, the proposed scheme in Theorem 4.3 is capacity-achieving. Dirty paper coding among cooperative private messages is needed to achieve the capacity of Gaussian MIMO broadcast channel* exactly *[41], that is, u_1 and u_2 is not independent conditioned on $\underline{x}_o$ and $\underline{x}_o$ is made zero. As shown in Chapter 5.2 and 5.3, however, linear beamforming strategies along with superposition coding suffice to achieve the capacity* approximately. *We conjecture that dirty paper coding among cooperative private messages will lead to a better rate region and smaller gap to the outer bounds, while the procedure of computing the achievable region becomes complicated.*

We have designed a coding strategy and its configuration which met the observation and intuition from the analysis of LDC, and it turns out that it achieves the capacity to within a bounded gap. This completes the proof of Theorem 4.11. The proof is broken into two parts: (1) the computation of the achievable rate region, and (2) the evaluation of the gap among inner and outer bounds. Details are left in Chapter 5.2 and Chapter 5.3 respectively.

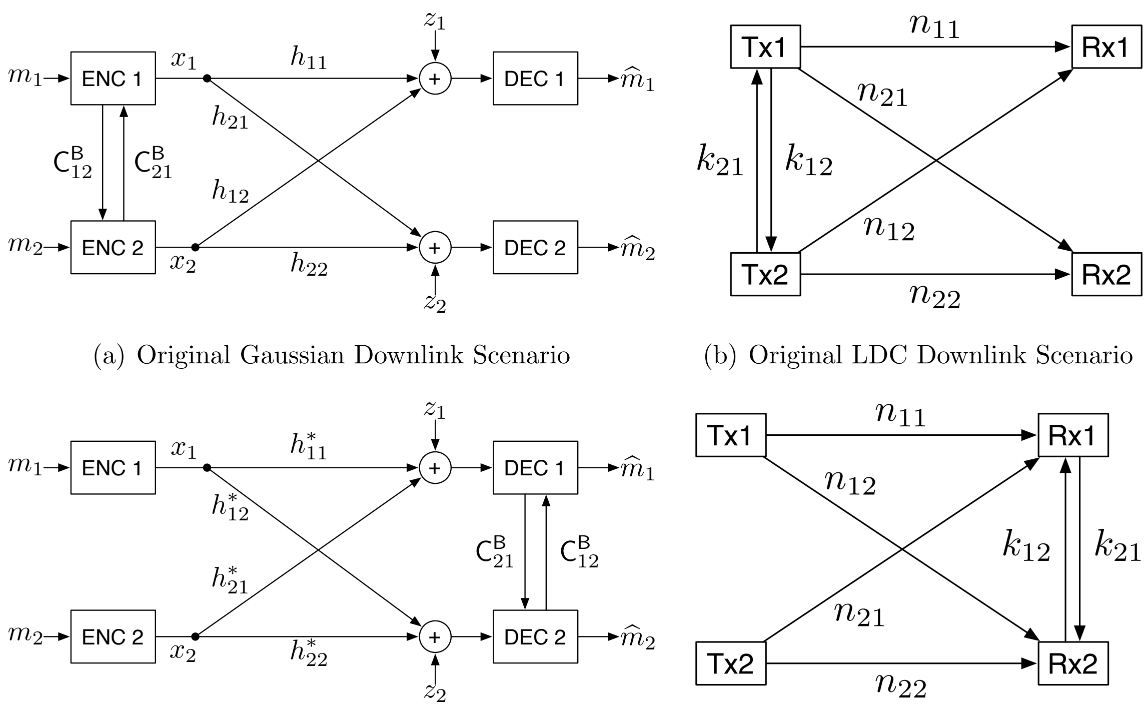
4.5 Uplink-Downlink Reciprocity

Recall that in Section 4.3.3 we have demonstrated the reciprocity between linear deterministic interference channel with conferencing receivers and linear deterministic interference channel with conferencing transmitters. In this section, we show that a similar reciprocity holds in the Gaussian case.

For the channel described in Section 4.2, we define its *reciprocal* channel as the Gaussian interference channel with conferencing receivers [15] with the 2-by-2 channel matrix

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}^H = \begin{bmatrix} h_{11}^* & h_{21}^* \\ h_{12}^* & h_{22}^* \end{bmatrix}$$

and cooperative link capacities $\mathsf{C}^{\mathsf{B}}_{21}$ *from receiver 1 to 2* and $\mathsf{C}^{\mathsf{B}}_{12}$ *from receiver 2 to 1.* Note that for the reciprocal channel, the channel matrix is the Hermitian of the original one and the cooperative link capacities are swapped. Motivated by backhaul cooperation in cellular networks where cooperation is among base stations, we term the interference channel with conferencing receivers the *uplink* scenario, and the interference channel with conferencing transmitters the *downlink* scenario. The original downlink and the reciprocal uplink scenarios are depicted in Fig. 4.7.



(a) Original Gaussian Downlink Scenario (b) Original LDC Downlink Scenario

(c) Reciprocal Gaussian Uplink Scenario (d) Reciprocal LDC Uplink Scenario

Figure 4.7: Uplink-Downlink Reciprocity

Theorem 4.13. *The capacity regions of the original and the reciprocal channels are within a bounded gap, regardless of channel parameters.*

Proof. Details are left in Chapter 5.5. □

The reciprocity implies immediately that the gain from transmitter cooperation shares the same characteristics as that from receiver cooperation, that is, the degree-of-freedom gain is either one bit or half a bit per cooperation bit until saturation, and the power gain is bounded no matter how large the cooperative link capacities are after saturation.

Remark 4.14. *As mentioned in Section 4.3.3, there is an* exact *reciprocity between the linear deterministic downlink scenario and the uplink scenario. Not only are the capacity regions of the original and the reciprocal channel the same, but the capacity-achieving linear schemes are also reciprocal. On the other hand, for the Gaussian downlink scenario and the uplink scenario, combining the results in this chapter and Chapter 2, it seems such reciprocity in the proposed strategies does not exist, since the message structures are different. Although the strategies proposed in this chapter and Chapter 2 are not reciprocal, we conjecture that such reciprocity may be obtained via structured lattice strategies derived from capacity-achieving linear schemes of the corresponding linear deterministic channels. Such conversion has been applied successfully in [51] to construct lattice coding strategies for many-to-one and one-to-many Gaussian interference channels.*

Chapter 5

Proofs of Part II

In this chapter we fill in the details of various proofs mentioned in the previous chapter.

5.1 Proof of Achievability in Theorem 4.1

5.1.1 Proof of Lemma 4.6

Plugging in the configuration, we have the following achievable rates from Theorem 4.3: for some nonnegative $(\widetilde{R}_{1h}, \widetilde{R}_{2h})$, (notations are listed in Table 5.1)

Table 5.1: Notations

p_1	t_1	m_1	l_1	s_1
$(n_{11}-n_{21})^+$	$\max(n_{12}, p_1)$	$\max(n_{11}, n_{12})$	$\max(n_{11}, g_1)$	$\max(n_{12}, g_1)$
p_2	t_2	m_2	l_2	s_2
$(n_{22}-n_{12})^+$	$\max(n_{21}, p_2)$	$\max(n_{22}, n_{21})$	$\max(n_{22}, g_2)$	$\max(n_{21}, g_2)$

Constraints at Transmitters:

$$
\begin{aligned}
R_{1h} &\le \widetilde{R}_{1h} \\
R_{2h} &\le \widetilde{R}_{2h} \\
R_{1o} + R_{2o} &= R_o \\
R_{1o} + R_{1h} &\le k_{12} \\
R_{2o} + R_{2h} &\le k_{21} \\
\widetilde{R}_{1h} + \widetilde{R}_{2h} - R_{1h} - R_{2h} &\ge 0
\end{aligned}
$$

Constraints at Receiver 1:

$$
\begin{array}{rrr}
R_{1p} \le p_1 & \widetilde{R}_{1h} \le g_1 & R_o + \widetilde{R}_{1h} \le m_1 \\
R_{2c} + R_{1p} \le t_1 & \widetilde{R}_{1h} + R_{1p} \le g_1 & R_o + \widetilde{R}_{1h} + R_{1p} \le m_1 \\
R_{1c} + R_{1p} \le n_{11} & R_{2c} + \widetilde{R}_{1h} \le s_1 & R_o + R_{2c} + \widetilde{R}_{1h} \le m_1 \\
R_{1c} + R_{2c} + R_{1p} \le m_1 & R_{2c} + \widetilde{R}_{1h} + R_{1p} \le s_1 & R_o + R_{2c} + \widetilde{R}_{1h} + R_{1p} \le m_1 \\
 & R_{1c} + \widetilde{R}_{1h} + R_{1p} \le l_1 & R_o + R_{1c} + \widetilde{R}_{1h} + R_{1p} \le m_1 \\
 & R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} \le m_1 & R_o + R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} \le m_1
\end{array}
$$

Constraints at Receiver 2: Above with index 1 and 2 exchanged.

After Fourier-Motzkin elimination, we have the following achievable rates and identify all redundant terms. The claims used below to show the redundancy are proved in the end of this section. We use symbol "Red" to denote "redundant".

(1) R_1 and R_2:

$$
\begin{array}{llll}
R_1 \le n_{11} + k_{12} & & R_2 \le n_{22} + k_{21} & \\
R_1 \le m_1 & & R_2 \le m_2 & \\
R_1 \le p_1 + t_2 + k_{12} & (\mathsf{Red}) & R_2 \le p_2 + t_1 + k_{21} & (\mathsf{Red}) \\
R_1 \le p_1 + s_2 + k_{12} & (\mathsf{Red}) & R_2 \le p_2 + s_1 + k_{21} & (\mathsf{Red}) \\
R_1 \le g_1 + t_2 + k_{12} & (\mathsf{Red}) & R_2 \le g_2 + t_1 + k_{21} & (\mathsf{Red}) \\
R_1 \le g_1 + s_2 + k_{12} & (\mathsf{Red}) & R_2 \le g_2 + s_1 + k_{21} & (\mathsf{Red}).
\end{array}
$$

To show the redundancy, we need to prove the following claim

Claim 5.1.

- $p_1 + t_2 \ge n_{11},\ p_2 + t_1 \ge n_{22}$
- $g_1 \ge p_1,\ g_2 \ge p_2$
- $s_1 \ge t_1,\ s_2 \ge t_2$

(2) $R_1 + R_2$:

$$
\begin{aligned}
&R_1 + R_2 \le t_1 + t_2 + k_{12} + k_{21} \\
&R_1 + R_2 \le p_1 + m_2 + k_{12} \\
&R_1 + R_2 \le p_2 + m_1 + k_{21} \\
&R_1 + R_2 \le g_1 + m_2 \\
&R_1 + R_2 \le g_2 + m_1
\end{aligned}
$$

$$R_1 + R_2 \leq s_1 + t_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$R_1 + R_2 \leq s_2 + t_1 + k_{12} + k_{21} \quad \text{(Red)}$$
$$R_1 + R_2 \leq s_1 + s_2 + k_{12} + k_{21} \quad \text{(Red)}$$

(3) $2R_1 + R_2$ and $R_1 + 2R_2$:

$$2R_1 + R_2 \leq p_1 + m_1 + t_2 + k_{12} + k_{21}$$
$$2R_1 + R_2 \leq g_1 + m_1 + t_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + R_2 \leq p_1 + m_1 + s_2 + k_{12}$$
$$2R_1 + R_2 \leq g_1 + m_1 + s_2 + k_{12} \quad \text{(Red)}$$
$$2R_1 + R_2 \leq p_1 + l_1 + m_2 + k_{12}$$
$$R_1 + 2R_2 \leq p_2 + m_2 + t_1 + k_{12} + k_{21}$$
$$R_1 + 2R_2 \leq g_2 + m_2 + t_1 + k_{12} + k_{21} \quad \text{(Red)}$$
$$R_1 + 2R_2 \leq p_2 + m_2 + s_1 + k_{21}$$
$$R_1 + 2R_2 \leq g_2 + m_2 + s_1 + k_{21} \quad \text{(Red)}$$
$$R_1 + 2R_2 \leq p_2 + l_2 + m_1 + k_{21}.$$

(4) $2R_1 + 2R_2$:

$$2R_1 + 2R_2 \leq p_1 + s_1 + t_2 + m_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq l_1 + t_1 + p_2 + m_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq p_1 + s_1 + s_2 + m_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq l_1 + t_1 + g_2 + m_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq t_1 + m_1 + p_2 + s_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq s_1 + m_1 + p_2 + s_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq p_1 + m_1 + l_2 + t_2 + k_{12} + k_{21} \quad \text{(Red)}$$
$$2R_1 + 2R_2 \leq g_1 + m_1 + l_2 + t_2 + k_{12} + k_{21} \quad \text{(Red)}$$

To show the redundancy, we need to prove the following claim:

Claim 5.2.

- $s_1 + t_2 \geq p_2 + m_1$; $s_2 + t_1 \geq p_1 + m_2$
- $l_1 + t_1 \geq p_1 + m_1$; $l_2 + t_2 \geq p_2 + m_2$

After removing the redundant terms, we have the following achievable region for LDC when $n_{11} + n_{22} \neq n_{12} + n_{21}$:

$$R_1 \leq \min \{n_{11} + k_{12}, m_1\}$$

$$
\begin{aligned}
R_2 &\leq \min\{n_{22}+k_{21}, m_2\} \\
R_1+R_2 &\leq \min\{g_1+m_2, g_2+m_1\} \\
R_1+R_2 &\leq t_1+t_2+k_{12}+k_{21} \\
R_1+R_2 &\leq \min\{p_1+m_2+k_{12}, p_2+m_1+k_{21}\} \\
2R_1+R_2 &\leq \min\{p_1+l_1+m_2+k_{12}, p_1+s_2+m_1+k_{12}\} \\
2R_1+R_2 &\leq p_1+m_1+t_2+k_{12}+k_{21} \\
R_1+2R_2 &\leq \min\{p_2+l_2+m_1+k_{21}, p_2+s_1+m_2+k_{21}\} \\
R_1+2R_2 &\leq p_2+m_2+t_1+k_{21}+k_{12}.
\end{aligned}
$$

To show that the above achievable region coincides with the rate region given in Theorem 4.1, the following facts are crucial:

Claim 5.3.

- $g_1+m_2=g_2+m_1=\max(n_{11}+n_{22}, n_{12}+n_{21})$.
- $s_2+m_1=l_1+m_2$; $s_1+m_2=l_2+m_1$

With these facts, referring to Table 5.1, and checking with the outer bounds, we complete the proof.

5.1.2 Proof of Lemma 4.8

We have the following achievable rates: for some nonnegative $(\widetilde{R}_{1h}, \widetilde{R}_{2h})$,

Constraints at Transmitters:

$$
\begin{aligned}
R_{1h} &\leq \widetilde{R}_{1h} \\
R_{2h} &\leq \widetilde{R}_{2h} \\
R_{1o}+R_{2o} &= R_o \\
R_{1o}+R_{1h} &\leq k_{12} \\
R_{2o}+R_{2h} &\leq k_{21} \\
\widetilde{R}_{1h}+\widetilde{R}_{2h}-R_{1h}-R_{2h} &\geq 0
\end{aligned}
$$

Constraints at Receiver 1:

$$
\begin{aligned}
R_{1p} &\leq p_1 & \widetilde{R}_{1h} &\leq p_1 & R_o+\widetilde{R}_{1h} &\leq m_1 \\
R_{2c}+R_{1p} &\leq t_1 & \widetilde{R}_{1h}+R_{1p} &\leq p_1 & R_o+\widetilde{R}_{1h}+R_{1p} &\leq m_1 \\
R_{1c}+R_{1p} &\leq n_{11} & R_{2c}+\widetilde{R}_{1h} &\leq t_1 & R_o+R_{2c}+\widetilde{R}_{1h} &\leq m_1
\end{aligned}
$$

$$R_{1c}+R_{2c}+R_{1p}\le m_1 \qquad R_{2c}+\widetilde{R}_{1h}+R_{1p}\le t_1 \qquad R_o+R_{2c}+\widetilde{R}_{1h}+R_{1p}\le m_1$$
$$R_{1c}+\widetilde{R}_{1h}+R_{1p}\le n_{11} \qquad R_o+R_{1c}+\widetilde{R}_{1h}+R_{1p}\le m_1$$
$$R_{1c}+R_{2c}+\widetilde{R}_{1h}+R_{1p}\le m_1 \quad R_o+R_{1c}+R_{2c}+\widetilde{R}_{1h}+R_{1p}\le m_1$$

Constraints at Receiver 2: Above with index 1 and 2 exchanged.

After Fourier-Motzkin elimination and removing redundant terms based on facts derived in the previous analysis, we have the following achievable region:

$$\begin{aligned}
R_1 &\le \min\{m_1, n_{11}+k_{12}\}\\
R_2 &\le \min\{m_2, n_{22}+k_{21}\}\\
R_1+R_2 &\le \min\{p_1+m_2, p_2+m_1, t_1+t_2+k_{12}+k_{21}\}\\
2R_1+R_2 &\le p_1+t_2+m_1+k_{12} \quad \text{(Red)}\\
R_1+2R_2 &\le p_2+t_1+m_2+k_{21} \quad \text{(Red)},
\end{aligned}$$

which coincides with the outer bounds. To prove this, we need the following facts:

Claim 5.4.

- $p_1+m_2=p_2+m_1=\max(n_{11},n_{22},n_{12},n_{21})$
- $p_1+t_2=p_2+n_{11}$; $p_2+t_1=p_1+n_{22}$

With the first fact $p_1+m_2=p_2+m_1=\max(n_{11},n_{22},n_{12},n_{21})$, we show that the sum rate inner bound coincides the outer bound. With the second fact $p_1+t_2=p_2+n_{11}$, we show that the $2R_1+R_2$ inner bound is redundant. Similarly the R_1+2R_2 inner bound is also redundant. This completes the proof.

5.1.3 Proof of the Claims

Proof of Claim 5.1

- $p_1+t_2\ge n_{11}$, $p_2+t_1\ge n_{22}$

  Proof.

  $$p_1+t_2\ge (n_{11}-n_{21})^+ + n_{21}\ge n_{11}, \quad p_2+t_1\ge (n_{22}-n_{12})^+ + n_{12}\ge n_{22}.$$

  □

- $g_1\ge p_1$, $g_2\ge p_2$

Proof.

$g_1 = \max\left\{n_{11} - (n_{21} - n_{22})^+, n_{12} - (n_{22} - n_{21})^+\right\} \geq n_{11} - (n_{21} - n_{22})^+ \geq n_{11} - n_{21}.$

On the other hand, $g_1 \geq 0$. Hence, $g_1 \geq (n_{11} - n_{21})^+ = p_1$. Similarly $g_2 \geq p_2$. □

- $s_1 \geq t_1$, $s_2 \geq t_2$

  Proof.

  $$s_1 = \max(n_{12}, g_1) \geq \max(n_{12}, p_1) = t_1,$$

  since $g_1 \geq p_1$. Similarly $s_2 \geq t_2$. □

Proof of Claim 5.2

- $s_1 + t_2 \geq p_2 + m_1$; $s_2 + t_1 \geq p_1 + m_2$

  Proof. If $n_{21} \leq n_{22}$,

  $$s_1 + t_2 = \max\left\{n_{12}, n_{11} - (n_{21} - n_{22})^+\right\} + t_2 = m_1 + t_2 \geq m_1 + p_2.$$

  If $n_{21} > n_{22}$ and $n_{22} \leq n_{12}$,

  $$\begin{aligned} s_1 + t_2 &= \max\{n_{12}, n_{11} + n_{22} - n_{21}\} + n_{21} = \max\{n_{12} + n_{21}, n_{11} + n_{22}\} \\ &\geq 0 + \max(n_{11}, n_{12}) = p_2 + m_1. \end{aligned}$$

  If $n_{21} > n_{22}$ and $n_{22} > n_{12}$,

  $$\begin{aligned} s_1 + t_2 &= \max\{n_{12}, n_{11} + n_{22} - n_{21}\} + n_{21} = \max\{n_{12} + n_{21}, n_{11} + n_{22}\} \\ &\geq \max\{n_{11} + n_{22} - n_{12}, n_{21}\} \geq \max\{n_{11} + n_{22} - n_{12}, n_{22}\} \\ &\geq n_{22} - n_{12} + \max(n_{11}, n_{12}) = p_2 + m_1. \end{aligned}$$

  Hence, $s_1 + t_2 \geq p_2 + m_1$. Similarly, $s_2 + t_1 \geq p_1 + m_2$. □

- $l_1 + t_1 \geq p_1 + m_1$; $l_2 + t_2 \geq p_2 + m_2$

  Proof. If $n_{21} \geq n_{22}$,

  $$l_1 + t_1 = \max\left\{n_{11}, n_{12} - (n_{22} - n_{21})^+\right\} + t_1 = m_1 + t_1 \geq m_1 + p_1.$$

  If $n_{21} < n_{22}$ and $n_{11} \geq n_{12}$,

  $$p_1 + m_1 = p_1 + n_{11} \leq t_1 + l_1.$$

  If $n_{21} < n_{22}$ and $n_{11} < n_{12}$, then $m_1 = t_1 = n_{12}$, hence

  $$p_1 + m_1 = p_1 + t_1 \leq n_{11} + t_1 \leq l_1 + t_1.$$

  Hence, $l_1 + t_1 \geq p_1 + m_1$. Similarly, $l_2 + t_2 \geq p_2 + m_2$ □

Proof of Claim 5.3

- $g_1 + m_2 = g_2 + m_1 = \max(n_{11} + n_{22}, n_{12} + n_{21})$.

  Proof. Note that

  $$\max(n_{22}, n_{21}) - (n_{21} - n_{22})^+ = n_{22},\ \max(n_{22}, n_{21}) - (n_{22} - n_{21})^+ = n_{21}.$$

  Hence,

  $$\begin{aligned} g_1 + m_2 &= \max\left\{n_{11} - (n_{21} - n_{22})^+, n_{12} - (n_{22} - n_{21})^+\right\} + \max(n_{22}, n_{21}) \\ &= \max\{n_{11} + n_{22}, n_{12} + n_{21}\}. \end{aligned}$$

  By symmetry, $g_2 + m_1 = \max(n_{11} + n_{22}, n_{12} + n_{21})$. □

- $s_2 + m_1 = l_1 + m_2$; $s_1 + m_2 = l_2 + m_1$

  Proof.

  $$\begin{aligned} s_2 + m_1 &= \max\left\{n_{21}, n_{22} - (n_{12} - n_{11})^+\right\} + \max(n_{11}, n_{12}) \\ &= \max\{n_{21} + \max(n_{11}, n_{12}), n_{22} + n_{11}\} \\ &= \max\{n_{21} + n_{11}, n_{21} + n_{12}, n_{22} + n_{11}\}; \\ l_1 + m_2 &= \max\left\{n_{11}, n_{12} - (n_{22} - n_{21})^+\right\} + \max(n_{22}, n_{21}) \\ &= \max\{n_{11} + \max(n_{22}, n_{21}), n_{12} + n_{21}\} \\ &= \max\{n_{11} + n_{22}, n_{11} + n_{21}, n_{12} + n_{21}\}. \end{aligned}$$

  Hence $s_2 + m_1 = l_1 + m_2$. By symmetry $s_1 + m_2 = l_2 + m_1$. □

Proof of Claim 5.4

- $p_1 + m_2 = p_2 + m_1 = \max(n_{11}, n_{22}, n_{12}, n_{21})$

  Proof. If $n_{11} \geq n_{21} \geq n_{22}$, then $n_{11} \geq n_{12}$ (otherwise contradicts the assumption $n_{11} + n_{22} = n_{12} + n_{21}$) and

  $$p_1 + m_2 = n_{11} - n_{21} + n_{21} = n_{11} = \max_{i,j \in \{1,2\}} \{n_{ij}\}.$$

  If $n_{21} \geq n_{22}$ and $n_{21} \geq n_{11}$, then $n_{21} \geq n_{12}$ (contradiction otherwise) and

  $$p_1 + m_2 = 0 + n_{21} = n_{21} = \max_{i,j \in \{1,2\}} \{n_{ij}\}.$$

If $n_{21} \le n_{22}$ and $n_{21} \le n_{11}$, then $n_{12} \ge n_{11}$ and $n_{12} \ge n_{22}$ (contradiction otherwise) and

$$p_1 + m_2 = n_{11} - n_{21} + n_{22} = n_{12} = \max_{i,j \in \{1,2\}} \{n_{ij}\}.$$

If $n_{11} \le n_{21} \le n_{22}$, then $n_{22} \ge n_{12}$ (contradiction otherwise) and

$$p_1 + m_2 = 0 + n_{22} = n_{22} = \max_{i,j \in \{1,2\}} \{n_{ij}\}.$$

Hence, $p_1 + m_2 = \max_{i,j \in \{1,2\}} \{n_{ij}\}$. Similarly, $p_2 + m_1 = \max_{i,j \in \{1,2\}} \{n_{ij}\}$. □

- $p_1 + t_2 = p_2 + n_{11}$; $p_2 + t_1 = p_1 + n_{22}$

  Proof. Note that $t_2 = \max\left\{n_{21}, (n_{22} - n_{12})^+\right\} = n_{21}$, since $n_{21} \ge 0$ and

$$n_{21} = n_{11} + n_{22} - n_{12} \ge n_{22} - n_{12}.$$

Hence,

$$p_1 + t_2 = (n_{11} - n_{21})^+ + n_{21} = \max(n_{11}, n_{21})$$

On the other hand,

$$p_2 + n_{11} = (n_{22} - n_{12})^+ + n_{11} = (n_{21} - n_{11})^+ + n_{11} = \max(n_{11}, n_{21})$$

Hence, $p_1 + t_2 = p_2 + n_{11}$. Similarly, $p_2 + t_1 = p_1 + n_{22}$. □

5.2 Proof of Theorem 4.11: Achievable Rate Region

Plug in Theorem 4.3 and evaluate, we obtain the following achievable rates:

Constraints at Transmitters:

$$\begin{aligned} R_{1h} &\le \widetilde{R}_{1h} \\ R_{2h} &\le \widetilde{R}_{2h} \\ R_{1o} + R_{2o} &= R_o \\ R_{1o} + R_{1h} &\le \mathsf{C}_{12}^{\mathsf{B}} \\ R_{2o} + R_{2h} &\le \mathsf{C}_{21}^{\mathsf{B}} \\ \widetilde{R}_{1h} + \widetilde{R}_{2h} - R_{1h} - R_{2h} &\ge 0, \end{aligned}$$

for some nonnegative $(\widetilde{R}_{1h}, \widetilde{R}_{2h})$.

Constraints at Receiver 1:

$$R_{1p} \leq \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$\widetilde{R}_{1h} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$\widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{2c} + R_{1p} \leq \log\left(1 + \frac{\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{1c} + R_{1p} \leq \log\left(1 + \frac{\mathsf{SNR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{2c} + \widetilde{R}_{1h} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}+\mathsf{INR}_{1c}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{2c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{1c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}+\mathsf{SNR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{1c} + R_{2c} + R_{1p} \leq \log\left(1 + \frac{\mathsf{SNR}_{1c}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1|\underline{x}_o}+\mathsf{SNR}_{1c}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + \widetilde{R}_{1h} \leq \log\left(1 + \frac{K_{u_1}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + R_{2c} + \widetilde{R}_{1h} \leq \log\left(1 + \frac{K_{u_1}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + R_{2c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1}+\mathsf{SNR}_{1c}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + R_{1c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1}+\mathsf{SNR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$R_o + R_{1c} + R_{2c} + \widetilde{R}_{1h} + R_{1p} \leq \log\left(1 + \frac{K_{u_1}+\mathsf{SNR}_{1c}+\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$

Constraints at Receiver 2: Above with index 1 and 2 exchanged.

Notice that $\mathsf{SNR}_{ic}+\mathsf{SNR}_{ip} = \mathsf{SNR}_i/4$, $\mathsf{INR}_{ic}+\mathsf{INR}_{ip} = \mathsf{INR}_i/4$, and $K_{u_i} \geq \mathsf{SNR}_i/4+\mathsf{INR}_i/4$ for $i = 1, 2$. For simplicity, we consider the subset of the above region:

Constraints at Transmitters: The same as above.

Constraints at Receiver 1: Exactly the same expressions as those in Section 5.1.1. but the notations are defined as

$$
\begin{aligned}
p_1 &:= \log\left(1+\frac{\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right), \; g_1 := \log\left(1+\frac{K_{u_1|\underline{x}_o}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)\\
t_1 &:= \log\left(1+\frac{\mathsf{INR}_{1c}+\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right), \; n_{11} := \log\left(1+\frac{\mathsf{SNR}_1/4}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)\\
s_1 &:= \log\left(1+\frac{K_{u_1|\underline{x}_o}+\mathsf{INR}_{1c}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right), \; l_1 := \log\left(1+\frac{K_{u_1|\underline{x}_o}+\mathsf{SNR}_1/4}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)\\
m_1 &:= \log\left(1+\frac{\mathsf{SNR}_1/4+\mathsf{INR}_{1c}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)
\end{aligned}
$$

instead of those in Table 5.1.

Constraints at Receiver 2: Above with index 1 and 2 exchanged.

Notice now the rate region is symbolically identical to that in LDC when the system matrix is full rank. Hence, after the Fourier-Motzkin procedure, we have the following achievable rates, which are also symbolically identical to those in LDC. The only difference is that, "redundancy" is replaced by "approximate redundancy". Proof of the claims to show the approximate redundancy will be given later in this section.

(1) R_1 and R_2:

R_1 constraints:

$$
\begin{aligned}
&R_1 \le n_{11} + \mathsf{C}_{12}^{\mathsf{B}} \\
&R_1 \le m_1 \\
&R_1 \le p_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} && \text{(approx. } \mathsf{Red}\text{)}\\
&R_1 \le p_1 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} && \text{(approx. } \mathsf{Red}\text{)}\\
&R_1 \le g_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} && \text{(approx. } \mathsf{Red}\text{)}\\
&R_1 \le g_1 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} && \text{(approx. } \mathsf{Red}\text{)}
\end{aligned}
$$

R_2 constraints: Above with index 1 and 2 exchanged.

To show the approximate redundancy, we need to prove the following claim:

Claim 5.5.

- $p_1 + t_2 \ge n_{11} - \log(9/4)$, $p_2 + t_1 \ge n_{22} - \log(9/4)$
- $p_1 + s_2 \ge n_{11} - \log(9/4)$, $p_2 + s_1 \ge n_{22} - \log(9/4)$

- $g_1 + t_2 \geq n_{11} - \log 9,\ g_2 + t_1 \geq n_{22} - \log 9$
- $g_1 + s_2 \geq n_{11} - \log 9,\ g_2 + s_1 \geq n_{22} - \log 9$

(2) $R_1 + R_2$:

$$\begin{aligned}
R_1 + R_2 &\leq t_1 + t_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} \\
R_1 + R_2 &\leq p_1 + m_2 + \mathsf{C}^{\mathsf{B}}_{12} \\
R_1 + R_2 &\leq p_2 + m_1 + \mathsf{C}^{\mathsf{B}}_{21} \\
R_1 + R_2 &\leq g_1 + m_2 \\
R_1 + R_2 &\leq g_2 + m_1 \\
R_1 + R_2 &\leq s_1 + t_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red}) \\
R_1 + R_2 &\leq s_2 + t_1 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red}) \\
R_1 + R_2 &\leq s_1 + s_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red})
\end{aligned}$$

To show the approximate redundancy, we need to prove the following claim:

Claim 5.6. $s_1 \geq t_1 - \log 5,\ s_2 \geq t_2 - \log 5$

(3) $2R_1 + R_2$ and $R_1 + 2R_2$:

$$\begin{aligned}
2R_1 + R_2 &\leq p_1 + m_1 + t_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} \\
2R_1 + R_2 &\leq g_1 + m_1 + t_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red}) \\
2R_1 + R_2 &\leq p_1 + m_1 + s_2 + \mathsf{C}^{\mathsf{B}}_{12} \\
2R_1 + R_2 &\leq g_1 + m_1 + s_2 + \mathsf{C}^{\mathsf{B}}_{12} && \text{(approx. } \mathsf{Red}) \\
2R_1 + R_2 &\leq p_1 + l_1 + m_2 + \mathsf{C}^{\mathsf{B}}_{12} \\
R_1 + 2R_2 &\leq p_2 + m_2 + t_1 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} \\
R_1 + 2R_2 &\leq g_2 + m_2 + t_1 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red}) \\
R_1 + 2R_2 &\leq p_2 + m_2 + s_1 + \mathsf{C}^{\mathsf{B}}_{21} \\
R_1 + 2R_2 &\leq g_2 + m_2 + s_1 + \mathsf{C}^{\mathsf{B}}_{21} && \text{(approx. } \mathsf{Red}) \\
R_1 + 2R_2 &\leq p_2 + l_2 + m_1 + \mathsf{C}^{\mathsf{B}}_{21}.
\end{aligned}$$

To prove the approximate redundancy, we need to show the following claim:

Claim 5.7. $g_1 \geq p_1 - \log 5,\ g_2 \geq p_2 - \log 5$

(4) $2R_1 + 2R_2$:

$$2R_1 + 2R_2 \leq p_1 + s_1 + t_2 + m_2 + \mathsf{C}^{\mathsf{B}}_{12} + \mathsf{C}^{\mathsf{B}}_{21}$$

$$
\begin{aligned}
2R_1 + 2R_2 &\leq l_1 + t_1 + p_2 + m_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq p_1 + s_1 + s_2 + m_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq l_1 + t_1 + g_2 + m_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq t_1 + m_1 + p_2 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq s_1 + m_1 + p_2 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq p_1 + m_1 + l_2 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} \\
2R_1 + 2R_2 &\leq g_1 + m_1 + l_2 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}}
\end{aligned}
$$

All the above are approximately redundant.

To show the approximate redundancy, we need to prove the following claim:

Claim 5.8.

- $s_1 + t_2 \geq p_2 + m_1 - \log 18,\ s_2 + t_1 \geq p_1 + m_2 - \log 18$
- $l_1 + t_1 \geq p_1 + m_1 - \log 12,\ l_2 + t_2 \geq p_2 + m_2 - \log 12$

We summarize in the lemma below an achievable rate region:

Lemma 5.9. *If* (R_1, R_2) *satisfies the following, it is achievable.*

$$
\begin{aligned}
R_1 &\leq \min\left\{m_1, n_{11} + \mathsf{C}_{12}^{\mathsf{B}} - 2\log 3\right\} \\
R_2 &\leq \min\left\{m_2, n_{22} + \mathsf{C}_{21}^{\mathsf{B}} - 2\log 3\right\} \\
R_1 + R_2 &\leq t_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} - 2\log 5 \\
R_1 + R_2 &\leq \min\left\{p_1 + m_2 + \mathsf{C}_{12}^{\mathsf{B}}, p_2 + m_1 + \mathsf{C}_{21}^{\mathsf{B}}\right\} - (\log 90)/2 \\
R_1 + R_2 &\leq \min\left\{g_1 + m_2, g_2 + m_1\right\} \\
2R_1 + R_2 &\leq p_1 + m_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} - \log 5 \\
2R_1 + R_2 &\leq \min\left\{p_1 + m_1 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} - \log 5, p_1 + l_1 + m_2 + \mathsf{C}_{12}^{\mathsf{B}}\right\} \\
R_1 + 2R_2 &\leq p_2 + m_2 + t_1 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} - \log 5 \\
R_1 + 2R_2 &\leq \min\left\{p_2 + m_2 + s_1 + \mathsf{C}_{21}^{\mathsf{B}} - \log 5, p_2 + l_2 + m_1 + \mathsf{C}_{21}^{\mathsf{B}}\right\}.
\end{aligned}
$$

In Section 5.3, we will show that the above achievable rate region is within a bounded gap from the outer bounds in Lemma 4.10. We close this section by the proof of the above mentioned claims.

Proof of the Claims

Prior to the proof of the above claims, we give a bunch of useful lemmas.

Lemma 5.10.

$$\begin{aligned}&\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right)\\&\geq\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right).\end{aligned}$$

Proof. Consider the Gaussian interference channel without cooperation. We take independent Gaussian input signals. Note that

$$\begin{aligned}&\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)\\&=I\left(x_1;y_1,y_2|x_2\right)+I\left(x_2;y_2\right)\\&\leq I\left(x_1;y_1,y_2|x_2\right)+I\left(x_2;y_2,y_1\right)=I\left(x_1,x_2;y_1,y_2\right)\\&=\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right).\end{aligned}$$

□

Corollary 5.11.

$$K_{u_1|\underline{x}_o}\geq\frac{\mathsf{SNR}_1}{4\left(1+\mathsf{INR}_2\right)},\ K_{u_2|\underline{x}_o}\geq\frac{\mathsf{SNR}_2}{4\left(1+\mathsf{INR}_1\right)}$$

Proof.

$$\begin{aligned}&1+K_{u_1|\underline{x}_o}\\&=\frac{3}{4}+\frac{1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2+\mathsf{SNR}_1+\mathsf{INR}_1}{4\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)}\\&\overset{(a)}{\geq}\frac{3}{4}+\frac{1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}}{4}=1+\frac{\mathsf{SNR}_1}{4\left(1+\mathsf{INR}_2\right)},\end{aligned}$$

where (a) is due to Lemma 5.10. Hence $K_{u_1|\underline{x}_o}\geq\frac{\mathsf{SNR}_1}{4(1+\mathsf{INR}_2)}$. Similarly $K_{u_2|\underline{x}_o}\geq\frac{\mathsf{SNR}_2}{4(1+\mathsf{INR}_1)}$. □

Lemma 5.12.

$$\begin{aligned}&2|h_{11}h_{22}-h_{12}h_{21}|^2+4\mathsf{SNR}_1\mathsf{SNR}_2\geq\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2,\\&2|h_{11}h_{22}-h_{12}h_{21}|^2+4\mathsf{INR}_1\mathsf{INR}_2\geq\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2.\end{aligned}$$

Proof.

$$\begin{aligned}|h_{11}h_{22}-h_{12}h_{21}|^2&\geq\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2-2\sqrt{\mathsf{SNR}_1\mathsf{SNR}_2\mathsf{INR}_1\mathsf{INR}_2}\\&:=x+y-2\sqrt{xy},\end{aligned}$$

where $x=\mathsf{SNR}_1\mathsf{SNR}_2$ and $y=\mathsf{INR}_1\mathsf{INR}_2$. Hence,

$$\begin{aligned}&2|h_{11}h_{22}-h_{12}h_{21}|^2+4\mathsf{SNR}_1\mathsf{SNR}_2-\left(\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2\right)\\&\geq 2x+2y-4\sqrt{xy}+3x-y=y\left(5u^2-4u+1\right)\geq 0,\end{aligned}$$

where $u:=\sqrt{x/y}$.

Similarly, we can prove that $2|h_{11}h_{22}-h_{12}h_{21}|^2+4\mathsf{INR}_1\mathsf{INR}_2\geq\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2$. □

Proof of Claim 5.5

- $p_1 + t_2 \geq n_{11} - \log(9/4)$, $p_2 + t_1 \geq n_{22} - \log(9/4)$

  Proof.

$$\begin{aligned} p_1 + t_2 &= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) + \log\left(1 + \frac{\mathsf{INR}_{2c}+\mathsf{SNR}_{2p}}{1+\sigma_2^2+\mathsf{INR}_{2p}}\right) \\ &= \log\left(\frac{(1+\sigma_1^2+\mathsf{SNR}_{1p}+\mathsf{INR}_{1p})(1+\sigma_2^2+\mathsf{INR}_2/4+\mathsf{SNR}_{2p})}{(1+\sigma_1^2+\mathsf{INR}_{1p})(1+\sigma_2^2+\mathsf{INR}_{2p})}\right) \\ &\geq \log\left(\frac{1+\sigma_1^2+\mathsf{SNR}_1/4+\mathsf{INR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) - \log\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right) \\ &\geq n_{11} - \log(9/4). \end{aligned}$$

  Similarly, $p_2 + t_1 \geq n_{22} - \log(9/4)$. □

- $p_1 + s_2 \geq n_{11} - \log(9/4)$, $p_2 + s_1 \geq n_{22} - \log(9/4)$

  Proof.

$$\begin{aligned} p_1 + s_2 &= \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) + \log\left(1 + \frac{K_{u_2|\underline{x}_o}+\mathsf{INR}_{2c}}{1+\sigma_2^2+\mathsf{INR}_{2p}}\right) \\ &= \log\left(\frac{(1+\sigma_1^2+\mathsf{SNR}_{1p}+\mathsf{INR}_{1p})\left(1+\sigma_2^2+\mathsf{INR}_2/4+K_{u_2|\underline{x}_o}\right)}{(1+\sigma_1^2+\mathsf{INR}_{1p})(1+\sigma_2^2+\mathsf{INR}_{2p})}\right) \\ &\geq n_{11} - \log(9/4). \end{aligned}$$

  Similarly, $p_2 + s_1 \geq n_{22} - \log(9/4)$. □

- $g_1 + t_2 \geq n_{11} - \log 9$, $g_2 + t_1 \geq n_{22} - \log 9$

  Proof.

$$\begin{aligned} g_1 + t_2 &= \log\left(\frac{\left(1+\sigma_1^2+\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}\right)(1+\sigma_2^2+\mathsf{INR}_2/4+\mathsf{SNR}_{2p})}{(1+\sigma_1^2+\mathsf{INR}_{1p})(1+\sigma_2^2+\mathsf{INR}_{2p})}\right) \\ &\geq \log\left(\frac{\sigma_1^2+\mathsf{INR}_{1p}+\left(1+K_{u_1|\underline{x}_o}\right)(1+\mathsf{INR}_2/4)}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \\ &\quad - \log\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right) \\ &\overset{(a)}{\geq} \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_{1p}+\mathsf{SNR}_1/16}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) - \log\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right) \\ &\geq n_{11} - \log 9, \end{aligned}$$

  where (a) is due to Corollary 5.11. Similarly, $g_2 + t_1 \geq n_{22} - \log 9$. □

- $g_1 + s_2 \geq n_{11} - \log 9$

  Proof.

$$\begin{aligned} g_1 + s_2 &= \log\left(\frac{\left(1+\sigma_1^2+\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}\right)\left(1+\sigma_2^2+\mathsf{INR}_2/4+K_{u_2|\underline{x}_o}\right)}{\left(1+\sigma_1^2+\mathsf{INR}_{1p}\right)\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right)}\right) \\ &\geq \log\left(\frac{\sigma_1^2+\mathsf{INR}_{1p}+\left(1+K_{u_1|\underline{x}_o}\right)\left(1+\mathsf{INR}_2/4\right)}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \\ &\quad -\log\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right) \\ &\geq \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_{1p}+\mathsf{SNR}_1/16}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) - \log\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right) \\ &\geq n_{11} - \log 9. \end{aligned}$$

  Similarly, $g_2 + s_1 \geq n_{22} - \log 9$. □

Remark 5.13. *If we want to follow the proofs in LDC closely, we can also prove the approximate redundancy by making use of the fact (to be proved later)*

$$s_2 \geq t_2 - \log 5,\ \ s_1 \geq t_1 - \log 5,\ \ g_1 \geq p_1 - \log 5,\ \ g_2 \geq p_2 - \log 5,$$

which results in a looser upper bound on the gap to outer bounds.

Proof of Claim 5.6

Proof.

$$\begin{aligned} s_1 &= \log\left(1+\frac{K_{u_1|\underline{x}_o}+\mathsf{INR}_{1c}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) = \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_1/4+K_{u_1|\underline{x}_o}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \\ &\overset{(a)}{\geq} \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_1/4+\frac{\mathsf{SNR}_1}{4(1+\mathsf{INR}_2)}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \overset{(b)}{\geq} \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_1/4+\frac{\mathsf{SNR}_{1p}}{5}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \\ &\geq t_1 - \log 5, \end{aligned}$$

where (a) is due to Corollary 5.11, and (b) is due to the fact that $\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2} \geq \frac{4\mathsf{SNR}_{1p}}{5}$.
Similarly $s_2 \geq t_2 - \log 5$. □

Proof of Claim 5.7

Proof.

$$g_1 = \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \overset{(a)}{\geq} \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_{1p}+\frac{\mathsf{SNR}_1}{4(1+\mathsf{INR}_2)}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$\overset{(b)}{\geq} \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_{1p}+\frac{\mathsf{SNR}_{1p}}{5}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right) \geq p_1 - \log 5,$$

where (a) is due to Corollary 5.11, and (b) is due to the fact that $\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2} \geq \frac{4\mathsf{SNR}_{1p}}{5}$.
Similarly $g_2 \geq p_2 - \log 5$. □

Proof of Claim 5.8

- $s_1 + t_2 \geq p_2 + m_1 - \log 18$, $s_2 + t_1 \geq p_1 + m_2 - \log 18$

Proof.

$$s_1 = \log\left(\frac{1+\sigma_1^2+\mathsf{INR}_1/4+K_{u_1|\underline{x}_o}}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$
$$t_2 = \log\left(\frac{1+\sigma_2^2+\mathsf{INR}_2/4+\mathsf{SNR}_{2p}}{1+\sigma_2^2+\mathsf{INR}_{2p}}\right)$$
$$p_2 = \log\left(\frac{1+\sigma_2^2+\mathsf{INR}_{2p}+\mathsf{SNR}_{2p}}{1+\sigma_2^2+\mathsf{INR}_{2p}}\right)$$
$$m_1 = \log\left(\frac{1+\sigma_1^2+\mathsf{SNR}_1/4+\mathsf{INR}_1/4}{1+\sigma_1^2+\mathsf{INR}_{1p}}\right)$$

Hence, it suffices to compare $L = \left(1+\sigma_1^2+\frac{\mathsf{INR}_1}{4}+K_{u_1|\underline{x}_o}\right)\left(1+\sigma_2^2+\frac{\mathsf{INR}_2}{4}+\mathsf{SNR}_{2p}\right)$ and $R = \left(1+\sigma_1^2+\frac{\mathsf{SNR}_1}{4}+\frac{\mathsf{INR}_1}{4}\right)(1+\sigma_2^2+\mathsf{INR}_{2p}+\mathsf{SNR}_{2p})$.
Note that from Lemma 5.12, if $\mathsf{SNR}_2 \geq \mathsf{INR}_2$,

$$\frac{\mathsf{INR}_1}{4}+K_{u_1|\underline{x}_o} = \frac{|h_{11}h_{22}-h_{12}h_{21}|^2+\mathsf{INR}_1\mathsf{INR}_2+\mathsf{SNR}_2\mathsf{INR}_1+\mathsf{SNR}_1+2\mathsf{INR}_1}{4\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)}$$
$$\geq \frac{\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2+4\mathsf{SNR}_2\mathsf{INR}_1+4\mathsf{SNR}_1+8\mathsf{INR}_1}{16\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)}$$
$$\geq \frac{\mathsf{SNR}_1\max\left(\mathsf{SNR}_2,1\right)+4\mathsf{INR}_1\max\left(\mathsf{SNR}_2,1\right)}{48\max\left(\mathsf{SNR}_2,1\right)}$$
$$\geq \frac{\mathsf{SNR}_1+\mathsf{INR}_1}{48}$$

Also, $\mathsf{INR}_2/4 \geq \mathsf{INR}_{2p}$. Hence, $s_1 + t_2 \geq p_2 + m_1 - \log 12$.

If $\mathsf{SNR}_2 < \mathsf{INR}_2$,

$$\begin{aligned}
R &= \left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right)\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) + (\mathsf{SNR}_1/4)\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) \\
&\quad + \mathsf{INR}_{2p}\left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right) + (\mathsf{SNR}_1/4)\mathsf{INR}_{2p} \\
&\leq 2\left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right)\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) + (\mathsf{SNR}_1/4)\,(5/4 + \mathsf{SNR}_{2p}) + \mathsf{SNR}_1/4 \\
&= 2\left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right)\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) + \frac{9}{16}\mathsf{SNR}_1 + \frac{\mathsf{SNR}_1\mathsf{SNR}_{2p}}{4},
\end{aligned}$$

and

$$\begin{aligned}
L &= \left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right)\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) + K_{u_1|\underline{x}_o}\left(1 + \sigma_2^2 + \mathsf{SNR}_{2p}\right) \\
&\quad + \left(1 + \sigma_1^2 + \mathsf{INR}_1/4\right)(\mathsf{INR}_2/4) + K_{u_1|\underline{x}_o}(\mathsf{INR}_2/4) \\
&\geq \frac{(1 + \sigma_1^2 + \mathsf{INR}_1/4)\,(1 + \sigma_2^2 + \mathsf{SNR}_{2p})}{2} + \frac{\mathsf{INR}_1}{8} \\
&\quad + \frac{K_{u_1|\underline{x}_o}\max(\mathsf{INR}_2, 1)}{4} + \frac{\mathsf{INR}_1\mathsf{INR}_2}{16} + \frac{K_{u_1|\underline{x}_o}}{2} \\
&\geq \frac{(1 + \sigma_1^2 + \mathsf{INR}_1/4)\,(1 + \sigma_2^2 + \mathsf{SNR}_{2p})}{2} + \frac{\mathsf{SNR}_1}{96} \\
&\quad + \frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + 2\mathsf{INR}_1\mathsf{INR}_2 + \mathsf{SNR}_1 + \mathsf{INR}_1}{48} \\
&\geq \frac{(1 + \sigma_1^2 + \mathsf{INR}_1/4)\,(1 + \sigma_2^2 + \mathsf{SNR}_{2p})}{2} + \frac{\mathsf{SNR}_1\mathsf{SNR}_2}{96} + \frac{\mathsf{SNR}_1}{32}.
\end{aligned}$$

Hence, $s_1 + t_2 \geq p_2 + m_1 - \log 18$.

In summary, $s_1 + t_2 \geq p_2 + m_1 - \log 18$, and similarly, $s_2 + t_1 \geq p_1 + m_2 - \log 18$. □

- $l_1 + t_1 \geq p_1 + m_1 - \log 12$, $l_2 + t_2 \geq p_2 + m_2 - \log 12$

Proof.

$$\begin{aligned}
l_1 &= \log\left(\frac{1 + \sigma_1^2 + \mathsf{INR}_{1p} + \mathsf{SNR}_1/4 + K_{u_1|\underline{x}_o}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \\
t_1 &= \log\left(\frac{1 + \sigma_1^2 + \mathsf{INR}_1/4 + \mathsf{SNR}_{1p}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \\
p_1 &= \log\left(\frac{1 + \sigma_1^2 + \mathsf{INR}_{1p} + \mathsf{SNR}_{1p}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \\
m_1 &= \log\left(\frac{1 + \sigma_1^2 + \mathsf{SNR}_1/4 + \mathsf{INR}_1/4}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right)
\end{aligned}$$

Hence, it suffices to compare

$$L = \left(1+\sigma_1^2+\mathsf{INR}_{1p}+\mathsf{SNR}_1/4+K_{u_1|\underline{x}_o}\right)\left(1+\sigma_1^2+\mathsf{INR}_1/4+\mathsf{SNR}_{1p}\right)$$

and

$$R = \left(1+\sigma_1^2+\mathsf{SNR}_1/4+\mathsf{INR}_1/4\right)\left(1+\sigma_1^2+\mathsf{INR}_{1p}+\mathsf{SNR}_{1p}\right).$$

Note that from Lemma 5.12, if $\mathsf{SNR}_2 \le \mathsf{INR}_2$,

$$\begin{aligned}
\frac{\mathsf{SNR}_1}{4}+K_{u_1|\underline{x}_o} &= \frac{|h_{11}h_{22}-h_{12}h_{21}|^2+\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{SNR}_1\mathsf{INR}_2+2\mathsf{SNR}_1+1\mathsf{INR}_1}{4\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)} \\
&\ge \frac{\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2+4\mathsf{SNR}_1\mathsf{INR}_2+8\mathsf{SNR}_1+4\mathsf{INR}_1}{16\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)} \\
&\ge \frac{\mathsf{INR}_1\max\left(\mathsf{INR}_2,1\right)+4\mathsf{SNR}_1\max\left(\mathsf{INR}_2,1\right)}{48\max\left(\mathsf{INR}_2,1\right)} \\
&\ge \frac{\mathsf{INR}_1+\mathsf{SNR}_1}{48}
\end{aligned}$$

Also, $\mathsf{INR}_1/4 \ge \mathsf{INR}_{1p}$. Hence, $l_1+t_1 \ge p_1+m_1-\log 12$.

If $\mathsf{SNR}_2 > \mathsf{INR}_2$,

$$\begin{aligned}
R &= \left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\left(\mathsf{INR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right) \\
&\quad +\mathsf{INR}_{1p}\left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)+\left(\mathsf{INR}_1/4\right)\mathsf{INR}_{1p} \\
&\le 2\left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\left(\mathsf{INR}_1/4\right)\left(5/4+\mathsf{SNR}_{1p}\right)+\mathsf{INR}_1/4 \\
&= 2\left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\frac{9}{16}\mathsf{INR}_1+\frac{\mathsf{INR}_1\mathsf{SNR}_{1p}}{4},
\end{aligned}$$

and

$$\begin{aligned}
L &= \left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\left(\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}\right)\left(\mathsf{INR}_1/4\right) \\
&\quad +\left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(\mathsf{INR}_1/4\right)+\left(\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right) \\
&\ge \left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\frac{\mathsf{INR}_1}{4}\left(1+K_{u_1|\underline{x}_o}+\frac{\mathsf{SNR}_1}{4}\right) \\
&\ge \left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\frac{\mathsf{INR}_1}{4}\left(1+\frac{\mathsf{SNR}_1}{48}\right) \\
&\ge \left(1+\sigma_1^2+\mathsf{SNR}_1/4\right)\left(1+\sigma_1^2+\mathsf{SNR}_{1p}\right)+\frac{\mathsf{INR}_1}{4}\left(1+\frac{\mathsf{SNR}_{1p}}{12}\right)
\end{aligned}$$

Hence, $l_1+t_1 \ge p_1+m_1-\log 12$.

In summary, $l_1+t_1 \ge p_1+m_1-\log 12$, and similarly, $l_2+t_2 \ge p_2+m_2-\log 12$. □

5.3 Proof of Theorem 4.11: Bounded Gap to Outer Bounds

(1) Bounds on R_1:

- Consider the outer bound

$$R_1 \leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\sqrt{\mathsf{SNR}_1 \mathsf{INR}_1}\right)$$

and the inner bound

$$R_1 \leq m_1 = \log\left(1 + \frac{\mathsf{SNR}_1/4 + \mathsf{INR}_{1c}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) = \log\left(\frac{1 + \sigma_1^2 + \frac{\mathsf{SNR}_1 + \mathsf{INR}_1}{4}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right)$$

Note that

$$\log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\sqrt{\mathsf{SNR}_1 \mathsf{INR}_1}\right) \leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + 1,$$
$$\log\left(\frac{1 + \sigma_1^2 + \frac{\mathsf{SNR}_1 + \mathsf{INR}_1}{4}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \geq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) - \log 9.$$

Hence the gap is at most $\log 9 + 1 = 2\log 3 + 1$.

- Consider the outer bound

$$R_1 \leq \log\left(1 + \mathsf{SNR}_1\right) + \mathsf{C}_{12}^{\mathsf{B}}$$

and the inner bound

$$\begin{aligned} R_1 &\leq n_{11} + \mathsf{C}_{12}^{\mathsf{B}} - 2\log 3 \\ &= \log\left(1 + \frac{\mathsf{SNR}_1/4}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) + \mathsf{C}_{12}^{\mathsf{B}} - 2\log 3 \end{aligned}$$

Note that

$$\log\left(1 + \frac{\mathsf{SNR}_1/4}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \geq \log\left(1 + \mathsf{SNR}_1\right) - \log 9.$$

Hence, the gap is at most $2\log 3 + \log 9 = 4\log 3 \approx 6.34$

In summary, the gap is at most $4\log 3 \approx 6.34$.

(2) R_2: similar to R_1, the gap is at most $4\log 3 \approx 6.34$.

(3) $R_1 + R_2$:

- Consider the outer bound (4.11) and the inner bound

$$R_1 + R_2 \le p_1 + m_2 + \mathsf{C}_{12}^{\mathsf{B}} - (\log 90)/2.$$

Note that

$$\log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + 2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}\right) \le \log(1 + \mathsf{SNR}_2 + \mathsf{INR}_2) + 1,$$
$$m_2 \ge \log(1 + \mathsf{SNR}_2 + \mathsf{INR}_2) - \log 9,$$
$$p_1 = \log\left(1 + \frac{\mathsf{SNR}_{1p}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \ge \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2}\right) - \log(9/4).$$

Hence the gap is at most $5\log 3 - 1 + (\log 10)/2 \approx 8.586$.

- Consider the outer bound (4.13) and the inner bound

$$R_1 + R_2 \le t_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} - (2\log 5).$$

Note that

$$\log\left(1 + \frac{\mathsf{SNR}_1 + 2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}}{1 + \mathsf{INR}_2} + \mathsf{INR}_1\right)$$
$$\le \log\left(1 + \frac{2\mathsf{SNR}_1 + \mathsf{INR}_1}{1 + \mathsf{INR}_2} + \mathsf{INR}_1\right) \le \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2} + \mathsf{INR}_1\right) + 1,$$

and

$$t_1 = \log\left(\frac{1 + \sigma_1^2 + \mathsf{INR}_1/4 + \mathsf{SNR}_{1p}}{1 + \sigma_1^2 + \mathsf{INR}_{1p}}\right) \ge \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_2} + \mathsf{INR}_1\right) - \log 9.$$

Hence, the gap is at most $4\log 3 + 2 + 2\log 5 \approx 12.984$.

- Consider the outer bound (4.14) and the inner bound

$$R_1 + R_2 \le \min\{g_1 + m_2, g_2 + m_1\}.$$

Note that

$$\log\left(\begin{array}{l} 1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + 2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1} \\ +2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2} + |h_{11}h_{22} - h_{12}h_{21}|^2 \end{array}\right)$$
$$\le \log\left(1 + 2\mathsf{SNR}_1 + 2\mathsf{INR}_1 + 2\mathsf{SNR}_2 + 2\mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2\right)$$
$$\le \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2\right) + 1,$$

and

$$g_1 + m_2 = \log\left(\frac{\left(1+\sigma_1^2+\mathsf{INR}_{1p}+K_{u_1|\underline{x}_o}\right)\left(1+\sigma_2^2+\frac{\mathsf{SNR}_2+\mathsf{INR}_2}{4}\right)}{\left(1+\sigma_1^2+\mathsf{INR}_{1p}\right)\left(1+\sigma_2^2+\mathsf{INR}_{2p}\right)}\right)$$
$$\geq \log\left(\left(1+K_{u_1|\underline{x}_o}\right)\left(1+\mathsf{SNR}_2+\mathsf{INR}_2\right)\right) - \log(81/4)$$
$$= \log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right) - \log 81.$$

Hence, the gap is at most $4\log 3 + 1 \approx 7.34$.

In summary, the gap is at most $4\log 3 + 2 + 2\log 5 \approx 12.984$.

(4) $2R_1 + R_2$:

- Consider the outer bound (4.15) and the inner bound

$$2R_1 + R_2 \leq p_1 + m_1 + t_2 + \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}} - (\log 5).$$

From previous arguments, one can directly see that the gap is at most

$$\log(9/4) + (2\log 3 + 1) + (2\log 3 + 1) + \log 5$$
$$= 6\log 3 + \log 5 \approx 11.832.$$

- Consider the outer bound (4.17) and the inner bounds

$$2R_1 + R_2 \leq p_1 + m_1 + s_2 + \mathsf{C}_{12}^{\mathsf{B}} - (\log 5)$$
$$2R_1 + R_2 \leq p_1 + l_1 + m_2 + \mathsf{C}_{12}^{\mathsf{B}}.$$

Note that

$$\log\begin{pmatrix} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{SNR}_2 \\ +\mathsf{INR}_1\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{INR}_2+2\left(1+\mathsf{INR}_2\right)\sqrt{\mathsf{SNR}_1\mathsf{INR}_1} \end{pmatrix}$$
$$\leq \log\begin{pmatrix} 1+2\mathsf{SNR}_1+2\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{SNR}_2 \\ +2\mathsf{INR}_1\mathsf{INR}_2+2\mathsf{SNR}_1\mathsf{INR}_2 \end{pmatrix}.$$

For the inner bounds,

$$m_1 + s_2 \geq \log\left(\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)\left(1+\mathsf{INR}_2/4+K_{u_2|\underline{x}_o}\right)\right) - (4\log 3 - 2)$$
$$= \log\begin{pmatrix} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\frac{\mathsf{SNR}_2}{4}+\frac{\mathsf{INR}_2}{2} \\ +\frac{\mathsf{SNR}_1\mathsf{INR}_2}{4}+\frac{|h_{11}h_{22}-h_{12}h_{21}|^2+\mathsf{INR}_1\mathsf{INR}_2}{4} \end{pmatrix} - (4\log 3 - 2)$$
$$\geq \log\begin{pmatrix} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\frac{\mathsf{SNR}_2}{4}+\frac{\mathsf{INR}_2}{2} \\ +\frac{\mathsf{SNR}_1\mathsf{INR}_2}{4}+\frac{\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2}{16} \end{pmatrix} - (4\log 3 - 2),$$

and

$$\begin{aligned} l_1 + m_2 &\geq \log\left((1 + \mathsf{SNR}_2 + \mathsf{INR}_2)\left(1 + \mathsf{SNR}_1/4 + K_{u_1|\underline{x}_o}\right)\right) - (4\log 3 - 2) \\ &= \log\left(\begin{array}{c} 1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_1}{2} + \frac{\mathsf{INR}_1}{4} \\ + \frac{\mathsf{SNR}_1 \mathsf{INR}_2}{4} + \frac{|h_{11}h_{22} - h_{12}h_{21}|^2 + \mathsf{SNR}_1\mathsf{SNR}_2}{4} \end{array}\right) - (4\log 3 - 2) \\ &\geq \log\left(\begin{array}{c} 1 + \mathsf{SNR}_2 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_1}{2} + \frac{\mathsf{INR}_1}{4} \\ + \frac{\mathsf{SNR}_1 \mathsf{INR}_2}{4} + \frac{\mathsf{SNR}_1\mathsf{SNR}_2 + \mathsf{INR}_1\mathsf{INR}_2}{16} \end{array}\right) - (4\log 3 - 2). \end{aligned}$$

Hence, the gap is at most $(4\log 3 + 3) + \log(9/4) + \log 5 = 6\log 3 + \log 5 + 1 \approx 12.832$.

In summary, the gap is at most $6\log 3 + \log 5 + 1 \approx 12.832$.

(5) $R_1 + 2R_2$: similar to $2R_1 + R_2$, the gap is at most $6\log 3 + \log 5 + 1 \approx 12.832$.

Combining the results, we characterize the capacity region to within a bounded gap, which is at most

$$\begin{aligned} &\max\left\{4\log 3, \frac{4\log 3 + 2 + 2\log 5}{2}, \frac{6\log 3 + \log 5 + 1}{3}\right\} \\ &= 2\log 3 + 1 + \log 5 = \log 90 \approx 6.5. \end{aligned}$$

5.4 Proof of Lemma 4.10

We first state a useful fact [18]:

Fact 5.14 (Conditional Independence among Messages). *The following Markov relations hold:*

$$m_1 \leftrightarrow \left(v_{12}^N, v_{21}^N\right) \leftrightarrow m_2;\ m_1 \leftrightarrow \left(v_{12}^N, v_{21}^N\right) \leftrightarrow x_2^N;\ m_2 \leftrightarrow \left(v_{12}^N, v_{21}^N\right) \leftrightarrow x_1^N.$$

The proof can be found in [18].

Below we start the proof of the outer bounds stated in Lemma 4.10.

Proof. (1) R_1 bound (4.9):

If R_1 is achievable, by Fano's inequality,

$$\begin{aligned} &N\left(R_1 - \epsilon_N\right) \\ &\leq I\left(m_1; y_1^N\right) \leq I\left(m_1; y_1^N | m_2\right) \\ &\leq I\left(m_1; y_1^N | m_2, v_{12}^N\right) + I\left(m_1; v_{12}^N | m_2\right) \\ &= h\left(y_1^N | m_2, v_{12}^N\right) - h\left(y_1^N | m_2, v_{12}^N, m_1\right) + H\left(v_{12}^N | m_2\right) - H\left(v_{12}^N | m_2, m_1\right) \\ &\overset{(a)}{=} h\left(h_{11}x_1^N + z_1^N | m_2, v_{12}^N\right) - h\left(z_1^N | m_2, v_{12}^N, m_1\right) + H\left(v_{12}^N | m_2\right) \end{aligned}$$

$$\stackrel{\text{(b)}}{\leq} h\left(h_{11}x_1^N + z_1^N\right) - h\left(z_1^N\right) + H\left(v_{12}^N\right)$$
$$\leq N \log\left(1 + |h_{11}|^2\right) + N\mathsf{C}_{12}^{\mathsf{B}} = N\left\{\log\left(1 + \mathsf{SNR}_1\right) + \mathsf{C}_{12}^{\mathsf{B}}\right\}$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to the fact that x_2^N is a function of $\left(m_2, v_{12}^N\right)$, and $\left(x_1^N, x_2^N, v_{12}^N\right)$ are all functions of (m_1, m_2). (b) is due to conditioning reduces entropy and the fact that z_1^N is independent of everything else.

On the other hand, if R_1 is achievable

$$\begin{aligned}
&N\left(R_1 - \epsilon_N\right)\\
&\leq I\left(m_1; y_1^N | m_2\right) = h\left(y_1^N | m_2\right) - h\left(y_1^N | m_2, m_1\right)\\
&\leq h\left(y_1^N\right) - h\left(z_1^N | m_2, m_1\right) = h\left(y_1^N\right) - h\left(z_1^N\right)\\
&\leq \max_{\|\rho| \leq 1}\left\{N \log\left(1 + |h_{11}|^2 + |h_{12}|^2 + 2\Re\left\{h_{11}h_{12}^*\rho\right\}\right)\right\}\\
&= N \log\left(1 + |h_{11}|^2 + |h_{12}|^2 + 2|h_{11}||h_{12}|\right)\\
&= N \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right),
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$.

(2) R_2 bound (4.10): They follow the same line as the R_1 bounds.

(3) $R_1 + R_2$ bound (4.11) and (4.12):

Let $s_1 := h_{21}x_1 + z_2$, and $s_2 := h_{12}x_2 + z_1$. If (R_1, R_2) is achievable, by Fano's inequality,

$$\begin{aligned}
&N\left(R_1 + R_2 - \epsilon_N\right)\\
&\leq I\left(m_1; y_1^N\right) + I\left(m_2; y_2^N\right)\\
&\leq I\left(m_1; y_1^N, s_1^N, v_{12}^N | m_2\right) + I\left(m_2, v_{12}^N; y_2^N\right)\\
&= I\left(m_1; y_1^N, s_1^N | v_{12}^N, m_2\right) + I\left(m_1; v_{12}^N | m_2\right) + h\left(y_2^N\right) - h\left(y_2^N | m_2, v_{12}^N\right)\\
&= h\left(y_1^N, s_1^N | v_{12}^N, m_2\right) - h\left(z_1^N, z_2^N\right) + H\left(v_{12}^N | m_2\right) + h\left(y_2^N\right) - h\left(s_1^N | m_2, v_{12}^N\right)\\
&= h\left(y_1^N | s_1^N, v_{12}^N, m_2\right) + h\left(y_2^N\right) - h\left(z_1^N, z_2^N\right) + H\left(v_{12}^N | m_2\right)\\
&\leq h\left(h_{11}x_1^N + z_1^N | h_{21}x_1^N + z_2^N\right) + h\left(y_2^N\right) - h\left(z_1^N, z_2^N\right) + H\left(v_{12}^N\right)\\
&\leq N \log\left(1 + \frac{|h_{11}|^2}{1 + |h_{21}|^2}\right) + N \log\left(1 + |h_{21}|^2 + |h_{22}|^2 + 2|h_{21}||h_{22}|\right) + N\mathsf{C}_{12}^{\mathsf{B}}\\
&= N\left\{\text{RHS of (4.11)}\right\},
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$.

Similarly, we prove the outer bound (4.12).

(4) $R_1 + R_2$ bound (4.13):

If (R_1, R_2) is achievable, by Fano's inequality,

$$
\begin{aligned}
& N\left(R_1+R_2-\epsilon_N\right) \\
& \leq I\left(m_1; y_1^N\right)+I\left(m_2; y_2^N\right) \\
& \leq I\left(m_1; y_1^N | v_{12}^N, v_{21}^N\right)+I\left(m_2; y_2^N | v_{12}^N, v_{21}^N\right)+I\left(m_1; v_{12}^N, v_{21}^N\right)+I\left(m_2; v_{12}^N, v_{21}^N\right) \\
& \overset{\text{(a)}}{\leq} I\left(m_1; y_1^N | v_{12}^N, v_{21}^N\right)+I\left(m_2; y_2^N | v_{12}^N, v_{21}^N\right)+I\left(m_1, m_2; v_{12}^N, v_{21}^N\right) \\
& \leq I\left(m_1; y_1^N, s_1^N | v_{12}^N, v_{21}^N\right)+I\left(m_2; y_2^N, s_2^N | v_{12}^N, v_{21}^N\right)+I\left(m_1, m_2; v_{12}^N, v_{21}^N\right) \\
& = h\left(y_1^N, s_1^N | v_{12}^N, v_{21}^N\right)-h\left(y_1^N, s_1^N | v_{12}^N, v_{21}^N, m_1\right) \\
& \quad + h\left(y_2^N, s_2^N | v_{12}^N, v_{21}^N\right)-h\left(y_2^N, s_2^N | v_{12}^N, v_{21}^N, m_2\right)+H\left(v_{12}^N, v_{21}^N\right) \\
& = h\left(y_1^N, s_1^N | v_{12}^N, v_{21}^N\right)-h\left(s_2^N, z_2^N | v_{12}^N, v_{21}^N, m_1\right) \\
& \quad + h\left(y_2^N, s_2^N | v_{12}^N, v_{21}^N\right)-h\left(s_1^N, z_1^N | v_{12}^N, v_{21}^N, m_2\right)+H\left(v_{12}^N, v_{21}^N\right) \\
& \overset{\text{(b)}}{=} h\left(y_1^N, s_1^N | v_{12}^N, v_{21}^N\right)-h\left(s_2^N | v_{12}^N, v_{21}^N, m_1\right)-h\left(z_2^N\right) \\
& \quad + h\left(y_2^N, s_2^N | v_{12}^N, v_{21}^N\right)-h\left(s_1^N | v_{12}^N, v_{21}^N, m_2\right)-h\left(z_1^N\right)+H\left(v_{12}^N, v_{21}^N\right) \\
& \overset{\text{(c)}}{=} h\left(y_1^N, s_1^N | v_{12}^N, v_{21}^N\right)-h\left(s_2^N | v_{12}^N, v_{21}^N\right)-h\left(z_2^N\right) \\
& \quad + h\left(y_2^N, s_2^N | v_{12}^N, v_{21}^N\right)-h\left(s_1^N | v_{12}^N, v_{21}^N\right)-h\left(z_1^N\right)+H\left(v_{12}^N, v_{21}^N\right) \\
& = h\left(y_1^N | s_1^N, v_{12}^N, v_{21}^N\right)+h\left(y_2^N | s_2^N, v_{12}^N, v_{21}^N\right)-h\left(z_1^N\right)-h\left(z_2^N\right)+H\left(v_{12}^N, v_{21}^N\right) \\
& \leq h\left(y_1^N | s_1^N\right)+h\left(y_2^N | s_2^N\right)-h\left(z_1^N\right)-h\left(z_2^N\right)+H\left(v_{12}^N\right)+H\left(v_{21}^N\right) \\
& \leq N\left\{\text{RHS of (4.13)}\right\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to the fact that m_1 and m_2 are independent. (b) is due to the fact that z_1^N and z_2^N are independent to everything else, respectively. (c) is due to the following fact regarding the Markov relations:

(5) $R_1 + R_2$ bound (4.14):
If (R_1, R_2) is achievable, by Fano's inequality,

$$
\begin{aligned}
& N\left(R_1+R_2-\epsilon_N\right) \\
& \leq I\left(m_1; y_1^N\right)+I\left(m_2; y_2^N\right) \leq I\left(m_1; y_1^N, y_2^N\right)+I\left(m_2; y_1^N, y_2^N\right) \\
& \leq I\left(m_1; y_1^N, y_2^N | m_2\right)+I\left(m_2; y_1^N, y_2^N\right) \\
& = I\left(m_1, m_2; y_1^N, y_2^N\right) = h\left(y_1^N, y_2^N\right)-h\left(z_1^N, z_2^N\right) \\
& \leq N\left\{\text{RHS of (4.14)}\right\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$.

(6) $2R_1 + R_2$ bound (4.15) and $R_1 + 2R_2$ bound (4.16):

If (R_1, R_2) is achievable, by Fano's inequality,

$$\begin{aligned}
&N\left(2R_1+R_2-\epsilon_N\right)\\
&\leq 2I\left(m_1;y_1^N\right)+I\left(m_2;y_2^N\right)\\
&\leq I\left(m_1;y_1^N\right)+I\left(m_1;y_1^N,s_1^N,v_{12}^N,v_{21}^N|m_2\right)+I\left(m_2;y_2^N,s_2^N,v_{12}^N,v_{21}^N\right)\\
&\leq I\left(m_1,v_{12}^N,v_{21}^N;y_1^N\right)+I\left(m_1;y_1^N,s_1^N|v_{12}^N,v_{21}^N,m_2\right)\\
&\quad+I\left(m_2;y_2^N,s_2^N|v_{12}^N,v_{21}^N\right)+I\left(m_1;v_{12}^N,v_{21}^N|m_2\right)+I\left(m_2;v_{12}^N,v_{21}^N\right)\\
&=h\left(y_1^N\right)-h\left(s_2^N|m_1,v_{12}^N,v_{21}^N\right)+I\left(m_1,m_2;v_{12}^N,v_{21}^N\right)\\
&\quad+h\left(h_{11}x_1^N+z_1^N,s_1^N|m_2,v_{12}^N,v_{21}^N\right)-h\left(z_1^N,z_2^N\right)\\
&\quad+h\left(y_2^N,s_2^N|v_{12}^N,v_{21}^N\right)-h\left(s_1^N,z_1^N|m_2,v_{12}^N,v_{21}^N\right)\\
&\overset{(a)}{=}h\left(y_1^N\right)-h\left(s_2^N|v_{12}^N,v_{21}^N\right)+H\left(v_{12}^N,v_{21}^N\right)\\
&\quad+h\left(h_{11}x_1^N+z_1^N|s_1^N,m_2,v_{12}^N,v_{21}^N\right)+h\left(y_2^N,s_2^N|v_{12}^N,v_{21}^N\right)-2h\left(z_1^N\right)-h\left(z_2^N\right)\\
&=h\left(y_1^N\right)+h\left(h_{11}x_1^N+z_1^N|s_1^N,m_2,v_{12}^N,v_{21}^N\right)\\
&\quad+H\left(v_{12}^N,v_{21}^N\right)+h\left(y_2^N|s_2^N,v_{12}^N,v_{21}^N\right)-2h\left(z_1^N\right)-h\left(z_2^N\right)\\
&\leq N\left\{\text{RHS of (4.15)}\right\},
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to the Markovity in Fact 5.14.

Similar arguments work for $R_1 + 2R_2$ bound (4.16).

(7) $2R_1 + R_2$ bound (4.17) and $R_1 + 2R_2$ bound (4.18):

Using the intuition from the study of linear deterministic channel, we give the following side information to receiver 2:

$$\widetilde{y}_2^N := h_{22}x_2^N + \widetilde{z}_2^N,$$

where $\widetilde{z}_2 \sim \mathcal{CN}\left(0, 1+\mathsf{INR}_2\right)$, i.i.d. over time and is independent of everything else.

Now, if (R_1, R_2) is achievable, by Fano's inequality,

$$\begin{aligned}
&N\left(2R_1+R_2-\epsilon_N\right)\\
&\leq 2I\left(m_1;y_1^N\right)+I\left(m_2;y_2^N\right)\\
&\leq I\left(m_1;y_1^N,s_1^N|m_2,v_{12}^N\right)+I\left(m_1;v_{12}^N|m_2\right)+I\left(m_1;y_1^N\right)+I\left(m_2;\widetilde{y}_2^N,y_2^N\right)\\
&=h\left(h_{11}x_1^N+z_1^N,s_1^N|m_2,v_{12}^N\right)-h\left(z_1^N,z_2^N\right)+I\left(m_1;y_1^N\right)\\
&\quad+I\left(m_2;\widetilde{y}_2^N\right)+I\left(m_2;y_2^N|\widetilde{y}_2^N\right)+I\left(m_1;v_{12}^N|m_2\right)\\
&\overset{(a)}{\leq}h\left(h_{11}x_1^N+z_1^N,s_1^N|m_2,v_{12}^N\right)-h\left(z_1^N,z_2^N\right)\\
&\quad+I\left(m_1,m_2;y_1^N,\widetilde{y}_2^N\right)+I\left(m_2,v_{12}^N;y_2^N|\widetilde{y}_2^N\right)+I\left(m_1;v_{12}^N|m_2\right)\\
&=h\left(h_{11}x_1^N+z_1^N,s_1^N|m_2,v_{12}^N\right)-h\left(z_1^N,z_2^N\right)+h\left(y_1^N,\widetilde{y}_2^N\right)
\end{aligned}$$

$$
\begin{aligned}
&\quad - h\left(z_1^N, \widetilde{z}_2^N\right) + h\left(y_2^N | \widetilde{y}_2^N\right) - h\left(y_2^N | \widetilde{y}_2^N, m_2, v_{12}^N\right) + H\left(v_{12}^N | m_2\right) \\
&= h\left(h_{11} x_1^N + z_1^N, s_1^N | m_2, v_{12}^N\right) + h\left(y_1^N, \widetilde{y}_2^N\right) + h\left(y_2^N | \widetilde{y}_2^N\right) \\
&\quad - h\left(s_1^N | \widetilde{z}_2^N, m_2, v_{12}^N\right) + H\left(v_{12}^N | m_2\right) - h\left(z_1^N, z_2^N\right) - h\left(z_1^N, \widetilde{z}_2^N\right) \\
&\overset{\text{(b)}}{=} h\left(h_{11} x_1^N + z_1^N, s_1^N | m_2, v_{12}^N\right) + h\left(y_1^N, \widetilde{y}_2^N\right) + h\left(y_2^N | \widetilde{y}_2^N\right) \\
&\quad - h\left(s_1^N | m_2, v_{12}^N\right) + H\left(v_{12}^N | m_2\right) - h\left(z_1^N, z_2^N\right) - h\left(z_1^N, \widetilde{z}_2^N\right) \\
&= h\left(h_{11} x_1^N + z_1^N | s_1^N, m_2, v_{12}^N\right) + h\left(y_1^N, \widetilde{y}_2^N\right) + h\left(y_2^N | \widetilde{y}_2^N\right) \\
&\quad + H\left(v_{12}^N | m_2\right) - h\left(z_1^N, z_2^N\right) - h\left(z_1^N, \widetilde{z}_2^N\right) \\
&\le N \left\{\text{RHS of } (4.17)\right\},
\end{aligned}
$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to a simple fact that $I\left(m_1; y_1^N\right) + I\left(m_2; \widetilde{y}_2^N\right) \le I\left(m_1, m_2; y_1^N, \widetilde{y}_2^N\right)$ and that conditioning reduces entropy. (b) holds since $\widetilde{z}_2^N$ is independent of $\left(m_2, v_{12}^N\right)$ and s_1^N.

Similar arguments work for $R_1 + 2R_2$ bound (4.18). □

5.5 Proof of Theorem 4.13

In Part I, we characterized the capacity region of the Gaussian interference channel with conferencing receivers to within 2 bits per user. Hence by Theorem 4.11, we only need to compare the outer bounds. Note that for the reciprocal channel, its channel parameters are

$$
\mathsf{SNR}_1' = \mathsf{SNR}_1,\ \mathsf{SNR}_2' = \mathsf{SNR}_2;\ \mathsf{INR}_1' = \mathsf{INR}_2,\ \mathsf{INR}_2' = \mathsf{INR}_1;\ \mathsf{C}_{12}^{\mathsf{B}\,\prime} = \mathsf{C}_{21}^{\mathsf{B}},\ \mathsf{C}_{21}^{\mathsf{B}\,\prime} = \mathsf{C}_{12}^{\mathsf{B}}.
$$

Plugging these into the outer bounds in Chapter 2, we obtain the outer bounds for the reciprocal channel as follows:

$$
\begin{aligned}
R_1 &\le \min\left\{\log\left(1 + \mathsf{SNR}_1\right) + \mathsf{C}_{12}^{\mathsf{B}}, \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right)\right\} \\
R_2 &\le \min\left\{\log\left(1 + \mathsf{SNR}_2\right) + \mathsf{C}_{21}^{\mathsf{B}}, \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2\right)\right\} \\
R_1 + R_2 &\le \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_1}\right) + \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_2}\right) + \mathsf{C}_{21}^{\mathsf{B}} + \mathsf{C}_{12}^{\mathsf{B}} \\
R_1 + R_2 &\le \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_1\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_1}\right) + \mathsf{C}_{21}^{\mathsf{B}} \\
R_1 + R_2 &\le \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right) + \log\left(1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_2}\right) + \mathsf{C}_{12}^{\mathsf{B}} \\
R_1 + R_2 &\le \log\left(1 + \mathsf{SNR}_1 + \mathsf{SNR}_2 + \mathsf{INR}_1 + \mathsf{INR}_2 + |h_{11}h_{22} - h_{12}h_{21}|^2\right) \\
2R_1 + R_2 &\le \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_2}{1 + \mathsf{INR}_2}\right) + \log\left(1 + \frac{\mathsf{SNR}_1}{1 + \mathsf{INR}_1}\right) + \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right) \\
&\quad + \mathsf{C}_{21}^{\mathsf{B}} + \mathsf{C}_{12}^{\mathsf{B}}
\end{aligned}
$$

$$R_1 + 2R_2 \leq \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + \log\left(1 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_2\right)$$
$$+ \mathsf{C}_{12}^{\mathsf{B}} + \mathsf{C}_{21}^{\mathsf{B}}$$
$$2R_1 + R_2 \leq \log\left(1 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2} + \mathsf{INR}_1 + \mathsf{SNR}_1 + \frac{\mathsf{INR}_2}{1+\mathsf{INR}_2} + \frac{|h_{11}h_{22} - h_{12}h_{21}|^2}{1+\mathsf{INR}_2}\right)$$
$$+ \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_2\right) + \mathsf{C}_{12}^{\mathsf{B}}$$
$$R_1 + 2R_2 \leq \log\left(1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1} + \mathsf{INR}_2 + \mathsf{SNR}_2 + \frac{\mathsf{INR}_1}{1+\mathsf{INR}_1} + \frac{|h_{11}h_{22} - h_{12}h_{21}|^2}{1+\mathsf{INR}_1}\right)$$
$$+ \log\left(1 + \mathsf{SNR}_2 + \mathsf{INR}_1\right) + \mathsf{C}_{21}^{\mathsf{B}}$$

(1) Bounds on R_1 and R_2:
Note that

$$\log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) \leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1 + 2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)$$
$$\leq \log\left(1 + \mathsf{SNR}_1 + \mathsf{INR}_1\right) + 1.$$

Hence the gap is at most 1 bit.

(2) Bounds on $R_1 + R_2$:
Note that:

(a)

$$\log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right)$$
$$= \log\left(\frac{(1+\mathsf{INR}_1)(1+\mathsf{INR}_2) + \mathsf{SNR}_1}{1+\mathsf{INR}_1}\right) + \log\left(\frac{(1+\mathsf{INR}_2)(1+\mathsf{INR}_1) + \mathsf{SNR}_2}{1+\mathsf{INR}_2}\right)$$
$$= \log\left(\frac{(1+\mathsf{INR}_1)(1+\mathsf{INR}_2) + \mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(\frac{(1+\mathsf{INR}_2)(1+\mathsf{INR}_1) + \mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)$$
$$= \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)$$
$$\leq \log\left(1 + \frac{\mathsf{SNR}_1 + 2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}}{1+\mathsf{INR}_2} + \mathsf{INR}_1\right) + \log\left(1 + \frac{\mathsf{SNR}_2 + 2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}}{1+\mathsf{INR}_1} + \mathsf{INR}_2\right)$$
$$\leq \log\left(1 + \mathsf{INR}_1 + \frac{2\mathsf{SNR}_1 + \mathsf{INR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{INR}_2 + \frac{2\mathsf{SNR}_2 + \mathsf{INR}_2}{1+\mathsf{INR}_1}\right)$$
$$\leq \log\left(1 + \mathsf{INR}_1 + \frac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) + \log\left(1 + \mathsf{INR}_2 + \frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right) + 2$$

(b)

$$\begin{aligned}
&\log\left(1+\mathsf{SNR}_2+\mathsf{INR}_1\right)+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right)\\
&=\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)\\
&\le\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)\\
&\le\log\left(1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+1
\end{aligned}$$

(c)

$$\begin{aligned}
&\log\left(1+\mathsf{SNR}_1+\mathsf{SNR}_2+\mathsf{INR}_1+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right)\\
&\le\log\left(\begin{array}{l}1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2\\ +2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}+|h_{11}h_{22}-h_{12}h_{21}|^2\end{array}\right)\\
&\le\log\left(1+2\mathsf{SNR}_1+2\mathsf{INR}_1+2\mathsf{SNR}_2+2\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right)\\
&\le\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2\right)+1
\end{aligned}$$

Hence the gap is at most 2 bits.

(3) Bounds on $2R_1+R_2$ and R_1+2R_2:
Note that:

(a)

$$\begin{aligned}
&\log\left(1+\mathsf{INR}_1+\frac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}\right)+\log\left(1+\frac{\mathsf{SNR}_1}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)\\
&=\left\{\log\left(1+\mathsf{INR}_2+\tfrac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)\right\}\\
&\le\left\{\begin{array}{l}\log\left(1+\tfrac{\mathsf{SNR}_2+2\sqrt{\mathsf{SNR}_2\mathsf{INR}_2}}{1+\mathsf{INR}_1}+\mathsf{INR}_2\right)\\ +\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1+2\sqrt{\mathsf{SNR}_1\mathsf{INR}_1}\right)\end{array}\right\}\\
&\le\left\{\begin{array}{l}\log\left(1+\mathsf{INR}_2+\tfrac{\mathsf{SNR}_2}{1+\mathsf{INR}_1}\right)+1+\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)\\ +\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_1\right)+1\end{array}\right\}
\end{aligned}$$

(b)

$$\left\{\begin{array}{l}\log\left(1+\tfrac{\mathsf{SNR}_2}{1+\mathsf{INR}_2}+\mathsf{INR}_1+\mathsf{SNR}_1+\tfrac{\mathsf{INR}_2}{1+\mathsf{INR}_2}+\tfrac{|h_{11}h_{22}-h_{12}h_{21}|^2}{1+\mathsf{INR}_2}\right)\\ +\log\left(1+\mathsf{SNR}_1+\mathsf{INR}_2\right)\end{array}\right\}$$

$$
\begin{aligned}
&= \left\{ \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+2\mathsf{INR}_2 \\ +\mathsf{SNR}_1\mathsf{INR}_2+\mathsf{INR}_1\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2 \end{array} \right) + \log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) \right\} \\
&\le \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+2\mathsf{INR}_2 \\ +\mathsf{SNR}_1\mathsf{INR}_2+\mathsf{INR}_1\mathsf{INR}_2+2\left(\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2\right) \end{array} \right) \\ +\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right) \end{array} \right\} \\
&\le \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2+\mathsf{SNR}_1\mathsf{SNR}_2+\mathsf{INR}_1\mathsf{INR}_2 \\ +\mathsf{SNR}_1\mathsf{INR}_2+2\left(1+\mathsf{INR}_2\right)\sqrt{\mathsf{SNR}_1\mathsf{INR}_1} \end{array} \right) \\ +\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+1 \end{array} \right\} \\
&\le \left\{ \log\left(\begin{array}{l} 1+2\mathsf{SNR}_1+2\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2 \\ +\mathsf{SNR}_1\mathsf{SNR}_2+2\mathsf{INR}_1\mathsf{INR}_2+2\mathsf{SNR}_1\mathsf{INR}_2 \end{array} \right) + \log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+1 \right\} \\
&\le \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+2\mathsf{SNR}_1+2\mathsf{INR}_1+\mathsf{SNR}_2+\mathsf{INR}_2 \\ +2\mathsf{SNR}_1\mathsf{INR}_2+4|h_{11}h_{22}-h_{12}h_{21}|^2+8\mathsf{INR}_1\mathsf{INR}_2 \end{array} \right) \\ +\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+1 \end{array} \right\} \\
&\le \left\{ \begin{array}{l} \log\left(\begin{array}{l} 1+\mathsf{SNR}_1+\mathsf{INR}_1+\mathsf{SNR}_2+2\mathsf{INR}_2 \\ +\mathsf{SNR}_1\mathsf{INR}_2+\mathsf{INR}_1\mathsf{INR}_2+|h_{11}h_{22}-h_{12}h_{21}|^2 \end{array} \right) \\ +\log\left(1+\tfrac{\mathsf{SNR}_1}{1+\mathsf{INR}_2}\right)+4 \end{array} \right\}
\end{aligned}
$$

Hence the gap is at most 4 bits.

In summary, we have

$$
\overline{\mathcal{C}}_{\mathrm{Rx}} \subset \overline{\mathcal{C}}_{\mathrm{Tx}} \subset \overline{\mathcal{C}}_{\mathrm{Rx}} \oplus [0,\tau]\times[0,\tau], \tag{5.1}
$$

where $\tau = 4/3$ bits.

Part III

Two Unicast Wireless Information Flows

Chapter 6

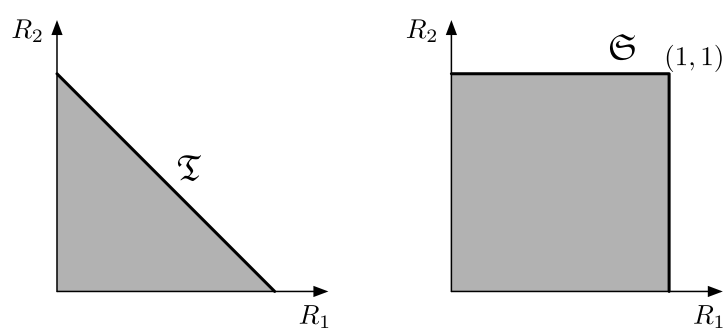
Two Unicast Linear Deterministic Network

We investigate the two unicast flow problem over layered linear deterministic networks with arbitrary number of nodes. When the minimum cut value between each source-destination pair is constrained to be 1, it is obvious that the triangular rate region $\{(R_1, R_2) : R_1, R_2 \geq 0, R_1 + R_2 \leq 1\}$ can be achieved, and that one cannot achieve beyond the square rate region $\{(R_1, R_2) : R_1, R_2 \geq 0, R_1 \leq 1, R_2 \leq 1\}$. Analogous to the work by Wang and Shroff for *wired* networks [52], we provide the necessary and sufficient conditions for the capacity region to be the triangular region and the necessary and sufficient conditions for it to be the square region. Moreover, we completely characterize the capacity region and conclude that there are exactly three more possible capacity regions of this class of networks, in contrast to the result in wired networks where only the triangular and square rate regions are possible. Our achievability scheme is based on linear coding over an extension field with at most four nodes performing special linear coding operations, namely interference neutralization and zero forcing, while all other nodes perform random linear coding.

6.1 Introduction

Characterizing the fundamental limit of delivering information from multiple sources to multiple destinations over networks is the holy grail in network information theory. The ultimate goal is to characterize the capacity region of multi-source-multi-destination information flows over arbitrary networks. Exploring *wired* network models yields fruitful understanding in this problem, and the capacity of single unicast [13] and multicast [53] are fully characterized. In wired networks, however, all links are orthogonal to one another, and such a model cannot fully capture the *broadcast* and *superposition* nature of wireless networks. In [9], a deterministic approach is proposed as a bridge for using results in wired networks to help understand wireless network information flow. The proposed *linear deterministic network* model turns out to be very useful for studying wireless networks as it preserves the broadcast and superposition aspects. Capacity of several traffic patterns are characterized completely in linear deterministic networks and approximately in Gaussian networks, including single unicast and multicast [9].

In the above mentioned problems where good understanding has been established, there is only one user's information flow in the network and no *interference* from other users. However, as for how multiple information flows interact as they interfere with one another, very little is known. To the best of our knowledge, even for the two unicast problem, there is no capacity results for general wired networks, let alone the general multi-source-multi-destination information flow problem. Instead of attempting directly to characterize the capacity region, a different route is taken in [52] to make progress in this problem, by studying the capacity region in the low rate regime. In [52], Wang and Shroff provide the necessary and sufficient condition for achievability of the rate pair $(1, 1)$ for two unicast flows over arbitrary wired networks with integer link capacities. They show that a simple outer bound called the Network Sharing Bound [54] turns out to be tight for the $(1, 1)$ point, i.e. if the Network Sharing Bound allows the achievability of $(1, 1)$, then in fact, $(1, 1)$ can be achieved. The result in [52] can be interpreted as follows. Let us consider a class of wired networks with integer link capacities where the minimum cut value between each source-destination pair is constrained to be 1, and hence rate pairs outside the square rate region $\mathfrak{S} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_1 \leq 1, R_2 \leq 1\}$ cannot be achieved. The main result of [52] provides the necessary and sufficient condition for the capacity region to be $\mathfrak{S}$. Moreover, since the triangular rate region $\mathfrak{T} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_1 + R_2 \leq 1\}$ can be achieved simply by time-division-access and routing, [52] completely characterizes the capacity region for this class of networks. There are only two possible capacity regions for this class of networks, the triangular region $\mathfrak{T}$ and the square region $\mathfrak{S}$. See Fig. 6.1 for an illustration of these rate regions.



(a) Network Sharing Bound = 1 (b) Network Sharing Bound ≥ 2

Figure 6.1: Capacity Regions for Wired Network

In this chapter, we take an initial step towards understanding the two unicast flow problem over linear deterministic networks [9] with arbitrary number of nodes. Our main result is an analog of [52] over linear deterministic networks. We assume that all channel strengths are zero or unity, that the network is layered and that each source can reach its own destination, and hence the minimum cut value between each source-destination pair is constrained to be 1. Similar to wired networks, rate pairs outside the square rate region $\mathfrak{S}$ cannot be achieved, and rate pairs inside the triangular rate region $\mathfrak{T}$ can be achieved by time-sharing between two users' single unicast flows. For this class of networks, we completely characterize the capacity region. We show that the capacity region of such a network must be one of the five regions depicted in Fig. 6.2, and provide the necessary and sufficient conditions for the capacity region to be each of them.

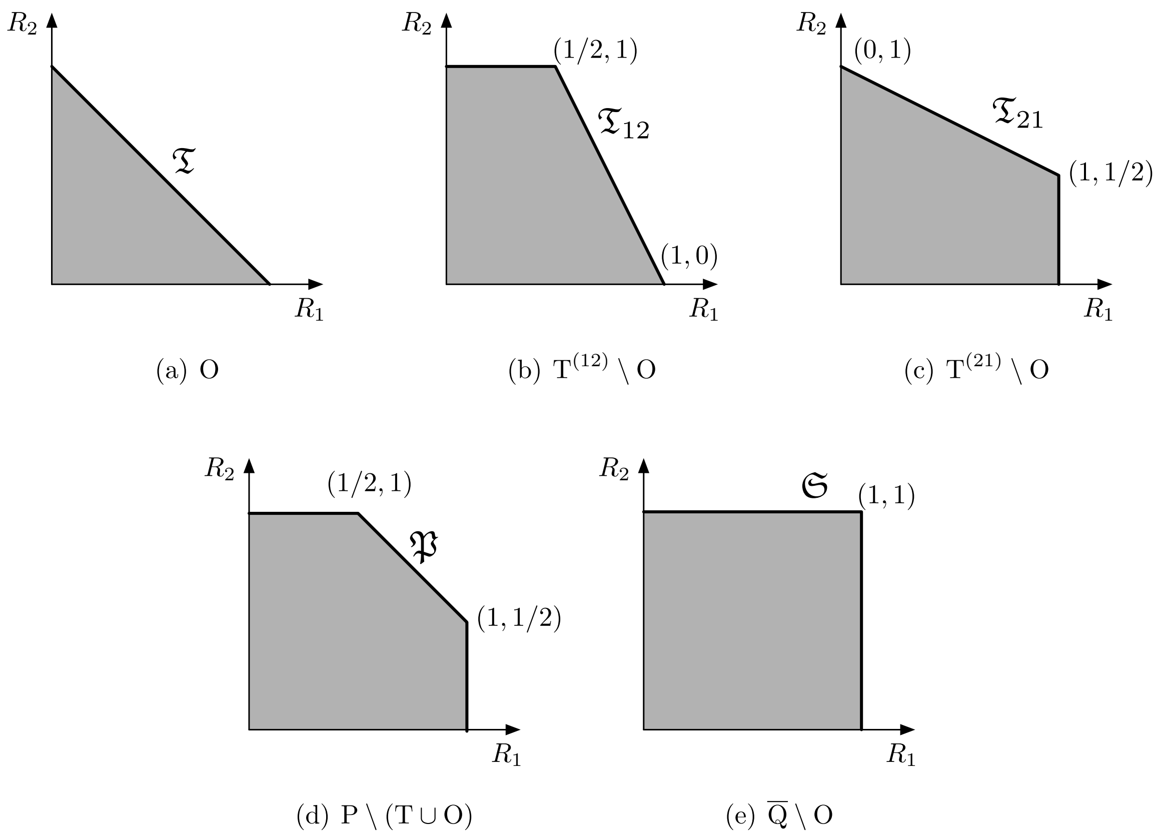


Figure 6.2: Capacity Regions for Linear Deterministic Network

Regarding when one can achieve beyond the trivially achievable $\mathfrak{T}$, we provide a novel sum rate outer bound on two unicast flows over linear deterministic networks, analogous to the Network Sharing Bound. This outer bound is intimately related to the Generalized

Network Sharing outer bound [55] for wired networks. We show that if this bound does not constrain the sum rate to be upper bounded by 1, then indeed, one can achieve *beyond* the triangular rate region $\mathfrak{T}$, and hence establish the necessary and sufficient condition for the capacity region being $\mathfrak{T}$. In contrast however, to achievability of the $(1, 1)$ point in [52], we find that we cannot always achieve $(1, 1)$. Instead, we show that once one can achieve beyond $\mathfrak{T}$, one can achieve either one of the two trapezoid rate regions: $\mathfrak{T}_{12} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_2 \leq 1, 2R_1 + R_2 \leq 2\}$ and $\mathfrak{T}_{21} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_1 \leq 1, R_1 + 2R_2 \leq 2\}$, and there are networks whose capacity regions are $\mathfrak{T}_{12}$ or $\mathfrak{T}_{21}$.

Regarding when one can achieve the full square $\mathfrak{S}$, we investigate the achievability of the $(1, 1)$ point, and find the necessary and sufficient conditions for it. For single source unicast and multicast problems, random linear coding over a large finite field at all nodes suffices to achieve capacity in wired as well as linear deterministic networks [53], [9]. This is no longer the case for the two-unicast problem since each destination is interested only in the message of its own source. Indeed, we can identify two nodes - one for each destination that must be able to decode the messages of their respective destinations. We call these two nodes *critical nodes* and their receptions are required to be completely free of *interference* from the other user. For this purpose, at certain nodes interference from the other user has to be cancelled "over-the-air", which is called *interference neutralization* in the literature [29] [56]. Other than the nodes performing interference neutralization, all other nodes may perform random linear coding. The parents of each critical node are the natural candidates to perform interference neutralization, although they are not the only ones. We introduce a systematic approach to capture the effect on the rest of the network caused by interference neutralization, and provide the graph-theoretic necessary and sufficient conditions for $(1, 1)$-achievability. Moreover, we show that if $(1, 1)$ cannot be achieved, then the capacity region is contained in the pentagon region $\mathfrak{P} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_1, R_2 \leq 1, R_1 + R_2 \leq 3/2\}$. Moreover, there are networks whose capacity regions are $\mathfrak{P}$.

Continuing further, we characterize the necessary and sufficient conditions for the capacity region to be $\mathfrak{T}_{12}$, $\mathfrak{T}_{21}$, and $\mathfrak{P}$ respectively. The outer bounds on $2R_1 + R_2, R_1 + 2R_2$ for the trapezoids $\mathfrak{T}_{12}, \mathfrak{T}_{21}$ respectively and that on $R_1 + R_2$ for the pentagon $\mathfrak{P}$ are inspired from the interference channel outer bounds [14]. The scheme we propose is linear over the *extension field* $\mathbb{F}_{2^r}$ for r sufficiently large. Note that unlike single multicast where a *random* (vector) linear scheme over the *base field* $\mathbb{F}_2$ suffices to achieve the capacity [9], in the two-unicast problem not only does the linear scheme operate on a larger field but also some nodes need to perform special linear coding (in contrast to random linear coding), including *interference neutralization* (over-the-air) and *zero forcing* (within-the-node). Later we will show by an example that both operating on a larger field and special coding at certain nodes are necessary for achieving capacity. It turns out that, fortunately, the number of nodes which are required to take special coding operation is bounded above by 4 and can be found explicitly. More specifically, they are usually parents of the two critical nodes and hence lie in two layers at most. Other than these special nodes, others can perform *random linear coding* (RLC) over the extension field.

Related Works

In the literature, the study of two unicast information flows over wireless networks using the deterministic approach begins with the investigation of the two-user interference channel [4] [14] [5] and its variants, including interference channels with cooperation [42] [25] [15] [16] and two-hop interference networks [29] [56]. Focusing on small networks (four nodes in total), researchers are able to characterize the capacity region exactly in the linear deterministic case [14] [15] [16] and to within a bounded gap in the Gaussian case [5] [15] [16], but the extension to larger networks seems non-trivial [29]. It is pointed out [29] that more advanced schemes, including interference neutralization, plays a key role in achieving capacity.

Another approach is directly looking at the Gaussian model but focusing on a cruder metric, degrees of freedom, instead of bounded gap to capacity. In [56], a systematic approach for interference neutralization called "aligned interference neutralization" is proposed for the 2x2x2 interference network, and it is shown that full degrees of freedom (one for each user) can be achieved *almost surely*. Later, in a recent independent work [57] such a scheme is employed and the authors characterize the degrees-of-freedom region of two unicast Gaussian networks almost surely. Interestingly, it is shown that [57] there are five possible degrees-of-freedom regions *almost surely* and they are the same as the five regions reported in this chapter. The connection between the two results is yet to be understood and explored. These degrees-of-freedom results, however, rely heavily either on the assumption that there is infinite channel diversity, or on the rationality/irrationality of the channel gains for the scheme to work.

The rest of the chapter is organized as follows. In Section 6.2, we formulate the problem and give several useful definitions. In Section 6.3, we state our main results, and in Section 6.4 we furnish examples that motivate linear scheme based on field extension and illustrate several important elements in achievability and outer bounds. Then we devote to details of achievability proof as well as outer bounds in Section 6.5 and 6.6, respectively. Finally, we conclude the chapter by discussing possible extensions to more general linear deterministic networks in Section 6.7.

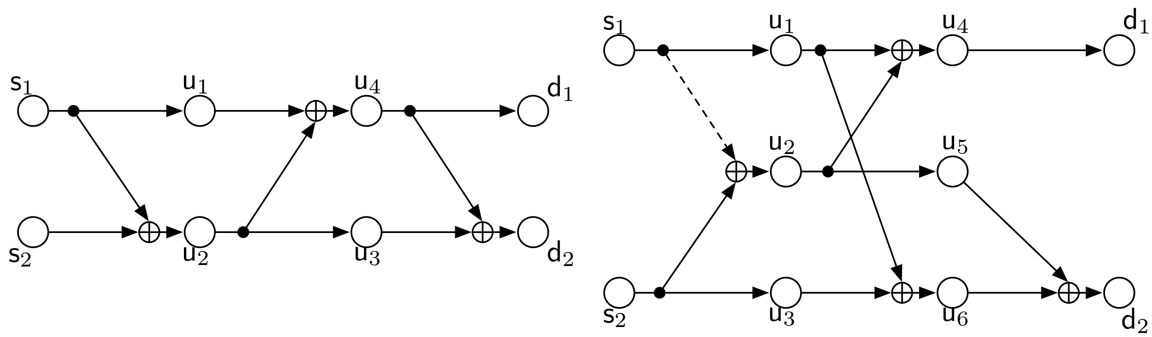
6.2 Problem Formulation

A two-source-two-destination layered network is a *directed, acyclic, layered* graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, i.e. where the collection of nodes $\mathcal{V}$ can be partitioned into $L+2$ *layers* ($L \geq 0$):

$$\mathcal{V} = \bigcup_{k=0}^{L+1} \mathcal{L}_k,\ \mathcal{L}_k \cap \mathcal{L}_j \neq \emptyset,\ \forall k \neq j,$$

such that for any edge $(\mathsf{u}, \mathsf{v}) \in \mathcal{E}$, $\exists\ k,\ 0 \leq k \leq L$ s.t. $\mathsf{u} \in \mathcal{L}_k, \mathsf{v} \in \mathcal{L}_{k+1}$. The first layer $\mathcal{L}_0 = \{\mathsf{s}_1, \mathsf{s}_2\}$ consists of the two source nodes, and the last layer $\mathcal{L}_{L+1} = \{\mathsf{d}_1, \mathsf{d}_2\}$ consists of the two destination nodes. Without loss of generality we assume each node in the network can be reached by at least one of the source nodes and can reach at least one of the destination nodes.

For each node $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$, we define nodes that can reach v as its *predecessors*. Let $\mathcal{P}(\mathsf{v})$ denote the set of predecessors that can reach v in one step. We will call the nodes in $\mathcal{P}(\mathsf{v})$ as the *parents* of v. Let $X_\mathsf{u}, Y_\mathsf{u} \in \mathbb{F}_2$ denote the transmission and reception of node u respectively. The reception of a node is the binary XOR of the transmission of its parents: $Y_\mathsf{v} = \bigoplus_{\mathsf{u} \in \mathcal{P}(\mathsf{v})} X_\mathsf{u}$. For example, in Fig. 6.3(a), the reception at node u_4 will be given by $Y_{\mathsf{u}_4} = X_{\mathsf{u}_1} \oplus X_{\mathsf{u}_2}$.



(a) Zigzag Network (b) An Asymmetric Network

Figure 6.3: Examples

The channel model we have used is a special case of the linear deterministic network from [9]. The simplification is that if there is a link from one node to another, then the channel strength is unity. We note that the essential nature of the linear deterministic network, namely broadcast and superposition, is preserved. As an example, in the network in Fig. 6.3(a), the transmission of u_2 is broadcasted to u_3 and u_4, and hence the two edges $(\mathsf{u}_2, \mathsf{u}_3)$ and $(\mathsf{u}_2, \mathsf{u}_4)$ carry the same signal. The reception of u_4, as mentioned above, is the binary XOR of the transmission of u_1 and u_2.

6.3 Main Result

If, for each $i = 1, 2$, s_i can reach d_i, then it is trivial to see that the *triangular* rate region $\mathfrak{T}$ can be achieved, and that one cannot achieve beyond the square rate region $\mathfrak{S}$. However, it is not clear under what conditions the triangular region or the square region is the capacity region. Our main result gives a complete answer to this question (and beyond). To state the result, we will need a few definitions.

A node is s_i-*reachable* if it can be reached by s_i. It is s_i-*only-reachable* if it can be reached by s_i but not s_j, $j \neq i$. It is $\mathsf{s}_1\mathsf{s}_2$-*reachable* if it can be reached by both s_1 and s_2.

For each node $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$,

- let $\mathcal{P}(\mathsf{v})$ denote the set of parents of v that are reachable from at least one of $\mathsf{s}_1, \mathsf{s}_2$,
- let $\mathcal{P}^{\mathsf{s}_i}(\mathsf{v}) \subseteq \mathcal{P}(\mathsf{v})$ denote the set of parents of v *reachable* by source s_i, $i = 1, 2$,
- let $\mathcal{K}(\mathsf{v}) := \{\mathsf{u} : \mathcal{P}(\mathsf{u}) = \mathcal{P}(\mathsf{v})\}$ denote the *clones* of v, the set of nodes that receive the same signal as v,
- let $\mathcal{K}^{\mathsf{s}_i}(\mathsf{v}) := \{\mathsf{u} : \mathcal{P}^{\mathsf{s}_i}(\mathsf{u}) = \mathcal{P}^{\mathsf{s}_i}(\mathsf{v})\}, i = 1, 2$, the set of nodes that have the same s_i-reachable parents as v. We called these nodes the s_i-clones of v.

The following table illustrates these sets of nodes for the node u_4 in the two example networks in Fig. 6.3. For the network in (b), we assume for now that there is no edge from s_1 to u_2.

	Fig. 6.3(a)	Fig. 6.3(b)
$\mathcal{P}(\mathsf{u}_4)$	$\{\mathsf{u}_1, \mathsf{u}_2\}$	$\{\mathsf{u}_1, \mathsf{u}_2\}$
$\mathcal{P}^{\mathsf{s}_1}(\mathsf{u}_4)$	$\{\mathsf{u}_1, \mathsf{u}_2\}$	$\{\mathsf{u}_1\}$
$\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_4)$	$\{\mathsf{u}_2\}$	$\{\mathsf{u}_2\}$
$\mathcal{K}(\mathsf{u}_4)$	$\{\mathsf{u}_4\}$	$\{\mathsf{u}_4\}$
$\mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_4)$	$\{\mathsf{u}_4\}$	$\{\mathsf{u}_4, \mathsf{u}_6\}$
$\mathcal{K}^{\mathsf{s}_2}(\mathsf{u}_4)$	$\{\mathsf{u}_3, \mathsf{u}_4\}$	$\{\mathsf{u}_4, \mathsf{u}_5\}$

For two sets of nodes $\mathcal{U}_1$ and $\mathcal{U}_2$, we say a collection of nodes $\mathcal{T}$ is a $(\mathcal{U}_1; \mathcal{U}_2)$-vertex-cut if in the graph obtained from the deletion of $\mathcal{T}$, there are no paths from any node in $\mathcal{U}_1 \setminus \mathcal{T}$ to any node in $\mathcal{U}_2 \setminus \mathcal{T}$. Note that this definition allows $\mathcal{T}$ to have nodes from $\mathcal{U}_1$ or $\mathcal{U}_2$.

We say a node $\mathsf{v} \in \mathcal{V}$ is *omniscient* if it satisfies either of (A) or (B) below:

(A) $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ is a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut.

(B) $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_2)$-vertex-cut and $\mathcal{K}^{\mathsf{s}_1}(\mathsf{v})$ is a $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut.

Theorem 6.1 (Characterization of $\mathfrak{T}$)**.** *Assume that* s_i *can reach* d_i *for* $i = 1, 2$.

(a) *If there exists an omniscient node in the network, then the capacity region is the triangular region* $\mathfrak{T}$.

(b) *Conversely, if no node in the network is omniscient, then the capacity region is strictly larger than* $\mathfrak{T}$. *Further, the capacity region contains at least one of the trapezoid regions* $\mathfrak{T}_{12}$ *and* $\mathfrak{T}_{21}$. *In particular,* $(2/3, 2/3)$ *is achievable and at least one of* $(1/2, 1)$ *and* $(1, 1/2)$ *is achievable.*

It turns out that we are able to give the necessary and sufficient condition for the capacity region to be either $\mathfrak{T}_{12}$ or $\mathfrak{T}_{21}$. Before describing the theorem, we need some extra definitions.

Definition 6.2 (Critical Nodes)**.** *For each* $i = 1, 2$, *we define the* critical node v_i^* *as any node with the smallest possible layer index such that* $\mathcal{K}(\mathsf{v}_i^*)$ *is a* $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_i)$*-vertex-cut.*

- Existence*:* $\{\mathsf{d}_i\} \subseteq \mathcal{K}(\mathsf{d}_i)$ *is a* $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_i)$*-vertex-cut.*
- Uniqueness up to clones*: if* u, w *are nodes in the same layer with* $\mathcal{K}(\mathsf{u})$ *and* $\mathcal{K}(\mathsf{w})$ *both being* $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_i)$*-vertex-cuts, then* $\mathcal{K}(\mathsf{u}) = \mathcal{K}(\mathsf{w})$*, i.e.* u *and* w *are clones.*

We use $\mathcal{L}_{k_i^*}$ *to denote the layer where critical nodes* v_i^* *lies, for* $i = 1, 2$*.*

For example in Fig. 6.3, $\mathsf{v}_1^* = \mathsf{u}_4, k_1^* = 2$ and $\mathsf{v}_2^* = \mathsf{d}_2, k_2^* = 3$ for both networks.

The critical nodes defined here are directly analogous to the edges performing the "reset" operation in the add-up-and-reset construction of Wang and Shroff [52].

Below we describe one scenario where we get a result similar to the one in [52]. This lemma strengthens part (b) of Theorem 6.1 in this special scenario.

Lemma 6.3. *Suppose in a network* s_2 *cannot reach* d_1*, i.e.* $k_1^* = 0$*. Then, the capacity region of this network is the triangle* $\mathfrak{T}$ *or the square* $\mathfrak{S}$ *depending on whether there is an omniscient node in the network or not, i.e. depending on whether* v_2^* *is omniscient or not (using Lemma 6.6). If* $k_2^* = 0$*, then* $(1, 1)$ *can be achieved by all nodes performing random linear coding. If* $k_2^* > 0$ *and there is no omniscient node, then* $(1, 1)$ *is achieved with high probability when all nodes except nodes in* $\mathcal{P}(\mathsf{v}_2^*)$ *performing random linear coding over a sufficiently large field.*

Next we define cut values and min-cut on the network.

Definition 6.4 (Cut Value and Min-Cut)**.** *Fix a set of nodes in layer* k*,* $\mathcal{U} \subseteq \mathcal{L}_k$*. Consider a partition of* $\mathcal{V}$ *into* $(\mathcal{T}, \mathcal{T}^c)$ *with* $\mathsf{s}_1, \mathsf{s}_2 \in \mathcal{T}$ *and* $\mathcal{U} \subseteq \mathcal{T}^c$*. Construct the transfer matrix* G *with rows indexed by elements of* $\mathcal{T}$ *and columns indexed by elements of* $\mathcal{T}^c$ *where the* (u, w) *entry of* G *is 1 if there is a directed edge from* u *to* w *and 0 otherwise. The* rank-mincut *[9] from* $\{\mathsf{s}_1, \mathsf{s}_2\}$ *to* $\mathcal{U}$ *is defined as the minimum rank of the transfer matrix* G *over all such partitions* $(\mathcal{T}, \mathcal{T}^c)$*, and is denoted by* $\mathrm{C}(\mathsf{s}_1.\mathsf{s}_2; \mathcal{U})$*.*

The following two lemmas provide some important properties of critical nodes. Their proofs are left in Chapter 7.

Lemma 6.5. *For* $i = 1, 2$*,* $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_i^*)) = 2$ *if* $k_i^* \geq 2$*.*

Lemma 6.6. *A network has an omniscient node if and only if one of the critical nodes* v_1^* *or* v_2^* *is omniscient.*

Once we define the cut value, we can define *primary min-cut nodes* for any set of nodes $\mathcal{U}$ with $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}) = 1$, due to the following lemma. What these primary min-cut nodes receive determines what $\mathcal{U}$ receive.

Lemma 6.7 (Primary Min-Cut)**.** *By* $\mathcal{U}_l, 0 \leq l < k$*, denote the set of nodes in layer* $\mathcal{L}_l$ *that can reach some node in* $\mathcal{U}$*. Let* l^* *be the minimum index such that* $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{l^*}) = 1$*. Then,* $\mathcal{U}_{l^*} \subseteq \mathcal{K}(\mathsf{u})$ *for any* $\mathsf{u} \in \mathcal{U}_{l^*}$*, i.e. nodes in* $\mathcal{U}_{l^*}$ *are all clones of each other.*

We then define any of the nodes in $\mathcal{K}(\mathsf{u})$ *as the* primary min-cut node *of* $\mathcal{U}$*, denoted by* $\mathsf{Pmc}(\mathcal{U})$*. It is unique up to clones.*

Comment: Note that the reception of any node in $\mathcal{U}$ is a function of the reception of $\mathsf{Pmc}(\mathcal{U})$.

For example, in Fig. 6.3(b) when there is an edge from s_1 to u_2, $\mathsf{Pmc}(\mathsf{u}_5) = \mathsf{u}_2$; when there is no edge from s_1 to u_2, $\mathsf{Pmc}(\mathsf{u}_5) = \mathsf{s}_2$. We also see that the critical node $\mathsf{v}_i^* = \mathsf{Pmc}(\mathsf{d}_i), i = 1, 2$.

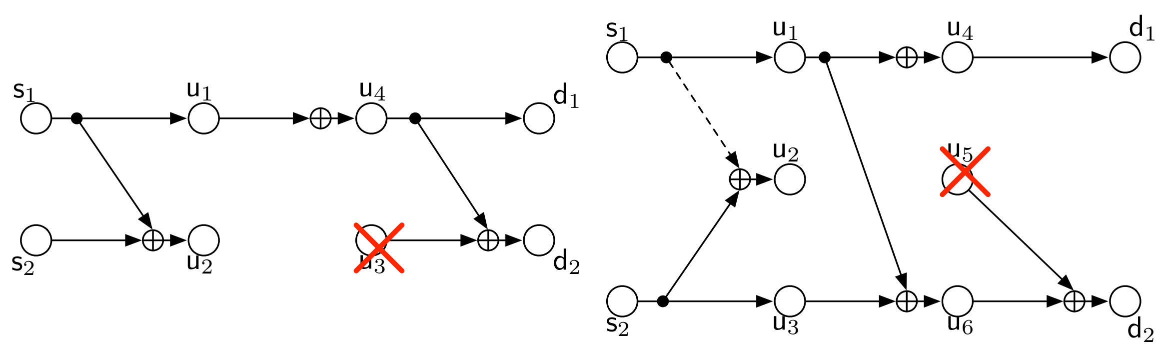
Next, we define induced graph $\mathcal{G}_{12}(\mathsf{w})$ for a node $\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ as follows. The purpose of these induced graph is two-fold: 1) to capture the effect on the rest of the network caused by interference neutralization for $(1,1)$-achievability, and 2) to capture the Markov relations that are useful in the derivation of outer bounds.

Definition 6.8 (Induced Graph $\mathcal{G}_{12}$).
If $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)) = 2$ *then* $\mathcal{G}_{12}(\mathsf{w}) := \mathcal{G}$. *If* $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)) = 1$, *then we define* $\mathcal{G}_{12}(\mathsf{w})$ *as the graph obtained by modifying only the parents of nodes in* $\mathcal{L}_{k_1^*}$ *as follows. For* $\mathsf{u} \in \mathcal{L}_{k_1^*}$,

$$\mathcal{P}_{\mathcal{G}_{12}(\mathsf{w})}(\mathsf{u}) = \begin{cases} \mathcal{P}(\mathsf{u}) & \text{if } \mathsf{w} \notin \mathcal{P}(\mathsf{u}) \\ \mathcal{P}(\mathsf{u}) \Delta \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) & \text{if } \mathsf{w} \in \mathcal{P}(\mathsf{u}), \end{cases}$$

where Δ *denotes symmetric set difference:* $A \Delta B := (A \setminus B) \cup (B \setminus A)$. *We then drop nodes in* $\mathcal{G}_{12}(\mathsf{w})$ *that cannot be reached by either of the two sources. In the rest of this chapter, a graph theoretic object with a graph (say,* $\mathcal{G}_{12}$*) in its subscript, like* $\mathcal{P}_{\mathcal{G}_{12}(\mathsf{w})}(\mathsf{u})$ *above, denote the graph theoretic object in the induced graph* $\mathcal{G}_{12}$. *Define* $\mathcal{R}(\mathsf{w})$ *as the set of nodes in* $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ *that can reach one of the two destinations in* $\mathcal{G}_{12}(\mathsf{w})$.

Similarly we can define $\mathcal{G}_{21}(\mathsf{w})$ with indices 1 and 2 swapped in the above definition.



(a) Zigzag Network (b) An Asymmetric Network

Figure 6.4: Induced Graph $\mathcal{G}_{12}$ for Example Networks in Fig. 6.3.

For example, induced graphs for the networks in Fig. 6.3 are depicted in Fig. 6.4. For $\mathcal{G}_{12}$ in (a), s_2 can no longer reach d_2, as u_4 is omniscient in the original network $\mathcal{G}$. In (b), node u_6 becomes omniscient in $\mathcal{G}_{12}$ while there is no omniscient node in the original network $\mathcal{G}$.

We will use $\mathcal{G}_{12}(\mathsf{w})$ when $k_1^* \leq k_2^*$ and $\mathcal{G}_{21}(\mathsf{w})$ when $k_2^* \leq k_1^*$. We will only use these graphs in relation to whether or not there is an omniscient node in $\mathcal{G}_{12}(\mathsf{w})$. Lemma 6.9 below allows us to drop the w and refer to any of the $\mathcal{G}_{12}(\mathsf{w})$ as $\mathcal{G}_{12}$ and talk about whether or not there is an omniscient node in $\mathcal{G}_{12}$.

Lemma 6.9. *Suppose, in a network with no omniscient node, and with $k_1^* \leq k_2^*$, there exists a node $\mathsf{w}_0 \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ such that there is an omniscient node in $\mathcal{G}_{12}(\mathsf{w}_0)$. Then for any node $\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, there is an omniscient node in $\mathcal{G}_{12}(\mathsf{w})$.*

Theorem 6.10 (Characterization of $\mathfrak{T}_{12}$ and $\mathfrak{T}_{21}$)**.** *Consider a network $\mathcal{G}$ in which no node is omniscient.*

(a) *If the network $\mathcal{G}$ satisfies the following conditions, then the capacity region is the trapezoid region $\mathfrak{T}_{12}$:*

- $\mathrm{T}_1^{(12)}$: $0 < k_1^* \leq k_2^*$.
- $\mathrm{T}_2^{(12)}$: $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)) = 1$. Let w_{12} denote $\mathsf{Pmc}(\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*))$.
- $\mathrm{T}_3^{(12)}$: Let $\mathsf{u}_{21} := \mathsf{Pmc}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$. u_{21} is omniscient in $\mathcal{G}_{12}$.
- $\mathrm{T}_4^{(12)}$: $\mathsf{w}_{12} = \mathsf{s}_2$, i.e., $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ cannot be reached by s_1.

We call the conjunction of the above conditions $T^{(12)}$. Symmetrically, if $\mathcal{G}$ satisfies the above condition with indices 1 *and* 2 *(in the superscript) exchanged, then the capacity region is the trapezoid region $\mathfrak{T}_{21}$.*

(b) *Conversely, if neither condition $T^{(12)}$ nor $T^{(21)}$ is satisfied, then the two trapezoid regions are strictly contained in the capacity region. Moreover, both* $(1/2, 1)$ *and* $(1, 1/2)$ *are achievable and hence the pentagon $\mathfrak{P}$.*

<u>Remark</u>: Based on Lemma 6.3, if $k_1^* = 0$, then the capacity region of this network is the triangle $\mathfrak{T}$ or the square $\mathfrak{S}$ depending on whether there is an omniscient node in the network. This is why in $\mathrm{T}_1^{(12)}$ we need to constrain $k_1^* > 0$.

Next we give the necessary and sufficient condition for the capacity region being the pentagon region $\mathfrak{P} := \{(R_1, R_2) : R_1, R_2 \geq 0, R_1 \leq 1, R_2 \leq 1, R_1 + R_2 \leq 3/2\}$.

Theorem 6.11 (Characterization of $\mathfrak{P}$ and $\mathfrak{S}$)**.** *Consider a network $\mathcal{G}$ in which no node is omniscient and neither $T^{(12)}$ nor $T^{(21)}$ is satisfied.*

(a) *Denote the conjunction of the below conditions by $P^{(12)}$:*

- $\mathrm{P}_1^{(12)} \equiv \mathrm{T}_1^{(12)}$, $\mathrm{P}_2^{(12)} \equiv \mathrm{T}_2^{(12)}$, $\mathrm{P}_3^{(12)} \equiv \mathrm{T}_3^{(12)}$

- $\mathrm{P}_4^{(12)}$: $\mathsf{w}_{12} \neq \mathsf{s}_2$ and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ forms an $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}$.

Similarly we define condition $P^{(21)}$ with indices 1 and 2 (in the superscript) exchanged. If the network $\mathcal{G}$ satisfies condition $P^{(12)}$ or $P^{(21)}$, then the capacity region is $\mathfrak{P}$.

(b) *Conversely, if neither condition $P^{(12)}$ nor $P^{(21)}$ is satisfied, then the pentagon region is strictly contained in the capacity region. Moreover, $(1,1)$ is achievable and hence the square $\mathfrak{S}$.*

We can easily see that $\mathrm{T}_4^{(12)} \vee \mathrm{P}_4^{(12)} = \{\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ forms an $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}.\}$ and hence $\mathrm{Q}^{(12)} := \mathrm{T}^{(12)} \vee \mathrm{P}^{(12)}$ is the conjunction of the following:

- $\mathrm{Q}_1^{(12)} \equiv \mathrm{T}_1^{(12)}$, $\mathrm{Q}_2^{(12)} \equiv \mathrm{T}_2^{(12)}$, $\mathrm{Q}_3^{(12)} \equiv \mathrm{T}_3^{(12)}$
- $\mathrm{Q}_4^{(12)}$: $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ forms an $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}$.

Corollary 6.12 (Complete Characterization of Capacity)**.** *As a corollary of Theorem 6.1, 6.10 and 6.11, we completely characterize all possible capacity regions of two unicast flows over the linear deterministic networks as formulated in Section 6.2, as follows: (Shorthand notations: $O := \{\exists$ an omniscient node$\}$, $T := T^{(12)} \vee T^{(21)}$, $P := P^{(12)} \vee P^{(21)}$, and $Q := Q^{(12)} \vee Q^{(21)} = T \vee P$. Also, in the context that no confusion will be caused, we use the same notation to denote the set of networks that satisfy the condition.)*

$$
\begin{array}{lcl}
O & \Longleftrightarrow & \mathfrak{T} \\
T^{(12)} \setminus O & \Longleftrightarrow & \mathfrak{T}_{12} \\
T^{(21)} \setminus O & \Longleftrightarrow & \mathfrak{T}_{21} \\
P \setminus (T \cup O) & \Longleftrightarrow & \mathfrak{P} \\
\overline{Q} \setminus O & \Longleftrightarrow & \mathfrak{S}
\end{array}
$$

Fig. 6.2 give an illustration of all these regions.

6.4 Motivating Examples

Before going into proofs of our main result, let us visit some examples to illustrate several important elements in our scheme as well as outer bounds.

6.4.1 Why Random Linear Coding Fails

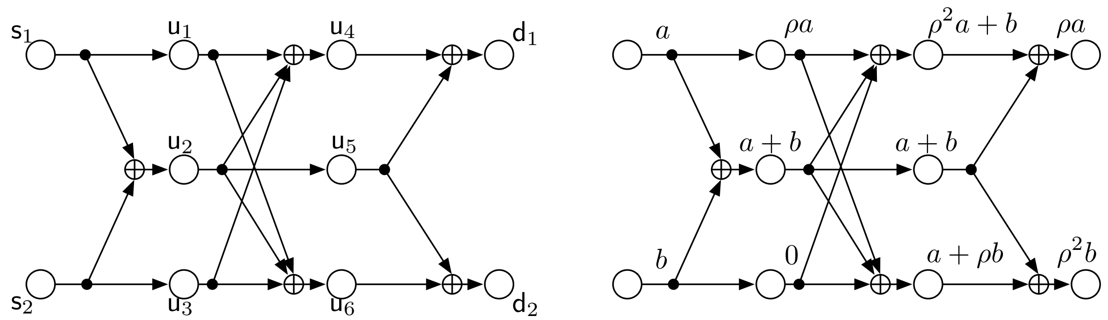
We first demonstrate, through a simple example, why random linear coding with its success in achieving the capacity of single multicast over wired and linear deterministic networks [53] [34], cannot achieve capacity for multiple unicast. Also, by the example we will show that most of the nodes in the network can perform random linear coding and only up to four nodes are needed to do special linear coding.

The example is depicted in Fig. 6.5(a). Random linear coding for achieving the $(1,1)$ point, in the context of this example, means that each node sends out a symbol in a large field of characteristic 2 and each intermediate node scales its reception by a randomly uniformly chosen coefficient from the field, independent of others, and transmits it. How and why we lift the symbols from the base field $\mathbb{F}_2$ to a larger field will be explained later. Random linear coding achieves the capacity of single multicast with high probability.

However, for two unicast if we perform random linear coding, in the network in Fig. 6.5(a), destinations d_1 and d_2 will receive linear combinations of the two symbols from sources, say a from source 1 and b from source 2, and their coefficients are non-zero with high probability. This is because both d_1 and d_2 can be reached by s_1 and s_2.

On the other hand, if nodes $\mathsf{u}_4, \mathsf{u}_5, \mathsf{u}_6$ choose their scaling coefficients more carefully, both d_1 and d_2 are able to receive a clean copy of their desired symbols. This is due to the fact that the reception of u_4 (which is the same as that of u_6) and the reception of u_5 are linearly independent with high probability under random linear coding at all other nodes in previous layers, since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_4, \mathsf{u}_5\right) = 2$. The scaling coefficients chosen at u_4 is such that the b-component in the transmission is cancelled over-the-air. Because the reception of u_4 and u_5 are linearly independent, the a-coefficient remains non-zero. Similarly, u_6 can choose its scaling coefficient so that d_2 receives a non-zero scaled-copy of symbol b.

We observe that in this example only nodes u_4 and u_6 need to perform linear coding carefully. It turns out that for arbitrary layered networks, at most 4 nodes need to perform special linear coding.



(a) Network (b) Linear Scheme over $\mathbb{F}_4$ achieving $(1,1)$

Figure 6.5: Examples

6.4.2 Why Field Extension is Necessary

We give an example to illustrate the limitation if we do not use field extension and stick to vector linear scheme in $\mathbb{F}_2$. The network is depicted in Fig. 6.5(a). Let the total number of channel uses be T, and source s_i would like to deliver B_i bits to its own destination d_i, $i = 1, 2$. We consider achieving beyond the triangular region $\mathfrak{T}$, and hence assume $B_1 + B_2 > T$. Therefore, at u_2, at least $B_1 + B_2 - T$ bits from each source get corrupted, while $B_1 - (B_1 + B_2 - T) = T - B_2$ bits from s_1 and $B_2 - (B_1 + B_2 - T) = T - B_1$ are clean. u_5's reception is just a function of what u_2 receives, and hence it cannot obtain more information than what u_2 possesses. In particular, u_5 cannot obtain the two length-$(B_1 + B_2 - T)$ chunks of bits of user 1 and user 2 that get corrupted at u_2. If u_5 does not transmit this corrupted chunk, u_4 needs to supply the clean chunk for user 1 to d_1 and u_6 needs to supply the clean chunk for user 2 to d_2, respectively. But the reception of u_4 and u_6 is identical, and therefore both of them should be able to decode these two chunks. As their reception has at most T bits, we have $2(B_1 + B_2 - T) \le T \implies 2B_1 + 2B_2 \le 3T$. If u_5 transmits this corrupted chunk, still u_4 needs to have the clean chunk for user 2 and u_6 needs to have the clean chunk for user 1. This is due to the property of $\mathbb{F}_2$. Hence we can again conclude $2B_1 + 2B_2 \le 3T$. Therefore, we see that this linear scheme over vector space of $\mathbb{F}_2$ cannot achieve beyond the pentagon $\mathfrak{P}$.

On the other hand, if instead we use a linear scheme over finite field $\mathbb{F}_4$, we are able to achieve $(1, 1)$. Recall that from the standard construction of extension field, the field $\mathbb{F}_4$ comprises $\{0, 1, \rho, \rho^2\}$ with the following addition and multiplication and one-to-one correspondence with $(\mathbb{F}_2)^2$:

$+$	0	1	ρ	ρ^2
0	0	1	ρ	ρ^2
1	1	0	ρ^2	ρ
ρ	ρ	ρ^2	0	1
ρ^2	ρ^2	ρ	1	0

$\times$	0	1	ρ	ρ^2
0	0	0	0	0
1	0	1	ρ	ρ^2
ρ	0	ρ	ρ^2	1
ρ^2	0	ρ^2	1	ρ

$\mathbb{F}_2 \times \mathbb{F}_2$	$\mathbb{F}_4$
$(0, 0)$	0
$(0, 1)$	1
$(1, 0)$	ρ
$(1, 1)$	ρ^2

Therefore, we can use two time slots to translate the following scalar coding scheme over $\mathbb{F}_4$ (depicted in Fig. 6.5(b)) back to a *nonlinear* coding scheme over $(\mathbb{F}_2)^2$: $a, b \in \mathbb{F}_4$,

	s_1	s_2	u_1	u_2	u_3	u_4	u_5	u_6	d_1	d_2
Transmits	a	b	ρa	$a + b$	0	$\rho^2 a + b$	$a + b$	$a + \rho b$		
Receives			a	$a + b$	b	$\rho^2 a + b$	$a + b$	$\rho^2 a + b$	ρa	$\rho^2 b$

Note that since the network is layered, one can without loss of generality assume that there is no processing delay within a node.

From the above example we see the benefit of working in the extension field is that, at each node there are more choices of scaling coefficients. In vector space $(\mathbb{F}_2)^r$, $r \ge 2$, the encoding matrix at each node has entries that are either 0 or 1, which limits the achievable rates of such a scheme.

6.4.3 Example of Networks with Different Capacity Regions

We provide examples of networks for each of the five possible capacity regions and use them to illustrate the important elements in our proposed scheme (including interference neutralization and zero forcing) as well as outer bounds.

1) Network with capacity region $\mathfrak{T}$

The example is depicted in Fig. 6.3(a). For achievability we know that $\mathfrak{T}$ can be achieved via time-sharing between rate pairs $(1,0)$ and $(0,1)$. For the outer bound, we notice that u_4 is omniscient, and the reception of the destination d_1 is a function of the reception of u_4. This means u_4 can decode the message of s_1. The reception of each node in $\mathcal{K}^{\mathsf{s}_2}(\mathsf{u}_4) = \{\mathsf{u}_4.\mathsf{u}_3\}$ is some function of the reception of node u_4 and the transmission of s_1. Since u_4 can now recover the transmission of s_1, and since $\mathcal{K}^{\mathsf{s}_2}(\mathsf{u}_4)$ forms a $(\mathsf{s}_2;\mathsf{d}_2)$-vertex-cut, u_4 can recover the reception of d_2, and thus, also the message of s_2. Therefore, the sum rate cannot be greater than the maximum entropy of the reception of u_4, which is 1.

2) Network with capacity region $\mathfrak{T}_{12}$

The example is depicted in Fig. 6.3(b) (without the dashed edge). We shall use this example to illustrate $(1/2,1)$-achievability as well as the outer bound. For achieving the rate pair $(1/2,1)$, we use the following scheme:

	s_1	s_2	u_1	u_2	u_3	u_4	u_5	u_6	d_1	d_2
Time 1 Transmits	a	b_1	a	b_1	b_1	$a+b_1$	b_1	0		
Time 1 Receives			a	b_1	b_1	$a+b_1$	b_1	$a+b_1$	$a+b_1$	b_1
Time 2 Transmits	a	b_2	a	0	b_2	a	0	$\boxed{b_2-b_1}$		
Time 2 Receives			a	b_2	b_2	a	0	$a+b_2$	a	$\boxed{b_2-b_1}$

Note that in the first time slot, all nodes transmit what they receive except for u_6. This is because the reception of u_6 contains a and hence it transmits 0 instead so that d_2 receives b_1. In the second time slot, u_2 has to keep silent so that u_4, the critical node for d_1, is able to decode a. u_5 hence receives 0, and b_2 needs to be provided by u_6. Still, it is necessary for u_6 to transmit a linear combination that does not contain a. Therefore, it makes use of the two linear combinations it receives over the two time slots, $a+b_1$ and $a+b_2$, to *zero-force* interference a and sends out b_2-b_1.

To see that the capacity region is $\mathfrak{T}_{12}$, we shall verify that the network satisfies $\mathrm{T}^{(12)}$. Obviously $\mathrm{T}_1^{(12)}$, $\mathrm{T}_2^{(12)}$, and $\mathrm{T}_4^{(12)}$ hold, as $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) = \{\mathsf{u}_2\}$. Induced graph $\mathcal{G}_{12}$ is $\mathcal{G}$ with edges $(\mathsf{u}_2,\mathsf{u}_4)$ and $(\mathsf{u}_2,\mathsf{u}_5)$ deleted. It can be seen that u_6 becomes omniscient in $\mathcal{G}_{12}$. Therefore $\mathrm{T}_3^{(12)}$ also holds.

3) Network with capacity region $\mathfrak{P}$

The example is the one depicted in Fig. 6.3(b) with an additional (dashed) edge $(\mathsf{s}_1,\mathsf{u}_2)$. To see that the capacity region is $\mathfrak{P}$, we shall verify that both $(1/2,1)$ and $(1,1/2)$ are achievable and the network satisfies $\mathrm{P}^{(12)}$. To achieve $(1/2,1)$, we use a similar scheme as above except that in the first time slot, u_5 and u_6 have to carry out interference neutralization to cancel a over the air. To achieve $(1,1/2)$, we use the following scheme:

	s_1	s_2	u_1	u_2	u_3	u_4	u_5	u_6	d_1	d_2
Time 1 Transmits	a_1	b	a_1	0	b	a_1	0	a_1+b		
Time 1 Receives			a_1	a_1+b	b	a_1	0	a_1+b	a_1	a_1+b
Time 2 Transmits	a_2	b	0	$\boxed{a_2-a_1}$	b	a_2-a_1	0	b		
Time 2 Receives			a_2	a_2+b	b	$\boxed{a_2-a_1}$	$\boxed{a_2-a_1}$	b	a_2-a_1	b

Note that in the first time slot, all nodes transmit what they receive except for u_2. This is because the reception of u_6 contains b and hence it transmits 0 instead so that u_4, the critical node for d_1, receives a_1. In the second time slot, u_2 makes use of the two linear combinations it receives over the two time slots, a_1+b and a_2+b, to zero-force interference b and sends out a_2-a_1. Meanwhile, u_1 keeps silent so that u_6 is able to decode b.

For the outer bound, obviously $\mathrm{P}_1^{(12)}$, $\mathrm{P}_2^{(12)}$, and $\mathrm{P}_4^{(12)}$ hold, as $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) = \{\mathsf{u}_2\}$ and $\mathsf{w}_{12} = \mathsf{u}_2$. Induced graph $\mathcal{G}_{12}$ is $\mathcal{G}$ with edges $(\mathsf{u}_2, \mathsf{u}_4)$ and $(\mathsf{u}_2, \mathsf{u}_5)$ deleted. It can be seen that u_6 becomes omniscient in $\mathcal{G}_{12}$. Therefore $\mathrm{P}_3^{(12)}$ also holds.

4) Network with capacity region $\mathfrak{S}$

The example is depicted in Fig. 6.5(a). Interference neutralization happens right at d_1 and d_2, which is carried out by $\mathsf{u}_4, \mathsf{u}_5, \mathsf{u}_6$. As explained in the previous subsection, such interference neutralization is not possible without coding in the extension field.

6.5 Proof of Achievability

In this section, we shall establish various achievability results *beyond* the trivially-achievable triangular rate region $\mathfrak{T}$. We assume that in the network no nodes are omniscient and describe a coding scheme that achieves $(1,1)$, $(1,1/2)$, or $(1/2,1)$. We will use a linear scheme over the finite field $\mathbb{F}_{2^r}$, for some $r > 0$. We map the r-length binary sequences in $(\mathbb{F}_2)^r$ to symbols in $\mathbb{F}_{2^r}$ such that the bitwise modulo-two addition in $(\mathbb{F}_2)^r$ translates to the addition operation in $\mathbb{F}_{2^r}$. Such a mapping is always possible by the standard construction of the extension field $\mathbb{F}_{2^r}$. Under such a mapping, we are able to abstract r usages in the original network to a single channel use in a network with the same topology, but with inputs and outputs in the extension field $\mathbb{F}_{2^r}$. A node is said to perform *Random Linear Coding* (RLC) over $\mathbb{F}_{2^r}$ if the coefficient(s) chosen by the node in the linear transformation is chosen uniformly at random in $\mathbb{F}_{2^r}$ and independently of the coefficients chosen by all its predecessors.

We will focus on schemes achieving rate pairs $(1,1)$ and $(1/2,1)$ respectively. To achieve $(1,1)$, each source aims to convey a symbol in $\mathbb{F}_{2^r}$ to its own destination over one symbol-time slot. The block length used by each node would be r. To achieve $(1/2,1)$, s_1 aims to deliver one symbol while s_2 aims to deliver two symbols to their respective destinations. The block length here would be $2r$. Note that the functions transforming an incoming r-block of bits (respectively $2r$-block) to an outgoing r-block of bits (respectively $2r$) is not a linear transformation over the vector space $(\mathbb{F}_2)^r$ (respectively $(\mathbb{F}_2)^{2r}$) and must necessarily be understood as operations over the extension field $\mathbb{F}_{2^r}$ for our proofs to work.

A scalar linear coding scheme over $\mathbb{F}_{2^r}$ is specified by the following collection of *linear coding coefficients*: $\{\alpha_{\mathsf{v}} \in \mathbb{F}_{2^r} : \mathsf{v} \in \mathcal{V} \setminus \{\mathsf{d}_1, \mathsf{d}_2\}\}$. Define for each $\mathsf{v} \in \mathcal{V}$ the *global coefficients* $\beta_{\mathsf{v},\mathsf{s}_1}, \beta_{\mathsf{v},\mathsf{s}_2} \in \mathbb{F}_{2^r}$ as follows.

- Initialize: $\beta_{\mathsf{s}_1,\mathsf{s}_1} := 1,\ \beta_{\mathsf{s}_1,\mathsf{s}_2} := 0,\ \beta_{\mathsf{s}_2,\mathsf{s}_1} := 0,\ \beta_{\mathsf{s}_2,\mathsf{s}_2} := 1.$
- For $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$, we define

$$\beta_{\mathsf{v},\mathsf{s}_1} := \sum_{\mathsf{u} \in \mathcal{P}(\mathsf{v})} \alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_1}, \quad \beta_{\mathsf{v},\mathsf{s}_2} := \sum_{\mathsf{u} \in \mathcal{P}(\mathsf{v})} \alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_2}.$$

If the messages of source s_1 and s_2 are a and b respectively, then the reception of node $\mathsf{v} \in \mathcal{V} \setminus \{\mathsf{s}_1, \mathsf{s}_2\}$ is given by $\beta_{\mathsf{v},\mathsf{s}_1} \cdot a + \beta_{\mathsf{v},\mathsf{s}_2} \cdot b$.

Recall Lemma 6.5 and 6.6. These two lemmas explain why we define critical nodes. Lemma 6.5 shows that the rank-influence from the sources to destination d_i drops precisely at the critical node v_i^* and hence, the nodes in $\mathcal{P}(\mathsf{v}_i^*)$ are natural candidates for special coding so as to cancel interference and arrange user i's symbol(s) to be received at v_i^* even while other nodes may perform random linear coding. Note that this kind of special coding is a linear operation over the finite field $\mathbb{F}_{2^r}$ making use of the superposition feature of the channel. Lemma 6.6 shows that the critical nodes suffice to capture the property of existence of an omniscient node in the network.

The reception of destination d_i is just a function of that of the critical node v_i^*. Hence we define the cloud $\mathcal{C}_i$, for $i = 1, 2$, to be the set of nodes that can be reached by some node in $\mathcal{K}(\mathsf{v}_i^*)$ and that can reach d_i. All nodes in the cloud receive functions of the reception of the critical node. Our scheme will ensure that v_i^* can decode what d_i aims to decode, $i = 1, 2$.

Below we provide several useful lemmas. Proofs of these lemmas can be found in Chapter 7.1.

Lemma 6.13. *If* u *is* s_i*-reachable, and all its predecessors do RLC with one symbol from each source, then* s_i*'s symbol has a non-zero coefficient in the reception of* u *with high probability.*

Lemma 6.14. *Consider* $\mathcal{U} \subseteq \mathcal{L}_k$ *with* $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}) = 2$*. Suppose each source transmits one symbol and all nodes in the network up to and including layer* $\mathcal{L}_{k-1}$ *perform RLC.*

(a) Then the nodes in $\mathcal{U}$ *can collectively decode both of the transmitted symbols with high probability.*

(b) If a node v *has* $\mathcal{U} \subseteq \mathcal{P}(\mathsf{v})$*, then with all nodes except nodes in* $\mathcal{U}$ *performing arbitray linear coding, nodes in* $\mathcal{U}$ *can arrange their transmission so that* v *receives any desired linear combination of the source symbols with high probability.*

(c) Let $\mathsf{u} \in \mathcal{U} \subseteq \mathcal{P}(\mathsf{v})$. *If nodes in* $\mathcal{P}(\mathsf{v}) \setminus \mathcal{U}$ *stay silent and nodes in* $\mathcal{U} \setminus \{\mathsf{u}\}$ *do RLC, then* u *is able to arrange its transmission so that* v *receives any linear combination linearly independent of the reception of* u *with high probability.*

(d) As a corollary of (c), if the node u *is* $\mathsf{s}_1\mathsf{s}_2$*-reachable, then* u *can adjust its transmission so that* v *can decode either* s_1 *or* s_2*'s symbol with high probability.*

Lemma 6.15. *If* $\mathcal{U} \subseteq \mathcal{L}_k$ *satisfies* $\mathrm{C}(\mathsf{s}_1.\mathsf{s}_2;\mathcal{U}) = 2$*, then for any* $\mathsf{u} \in \mathcal{U}$*, we can find some* $\mathsf{w} \in \mathcal{U}$ *such that* $\mathrm{C}(\mathsf{s}_1.\mathsf{s}_2;\mathsf{u},\mathsf{w}) = 2$.

Next, we shall prove achievability in different cases. Formal proofs of lemmas and claims are left in Chapter 7.1. Without loss of generality, we assume that $k_1^* \le k_2^*$. If $k_1^* = 0$, based on Lemma 6.3 we know that if there is no omniscient node, then $(1,1)$ is achievable. If $k_1^* = 1$, then by the definition of critical node v_1^*, both s_1 and s_2 are v_1^*'s parents and hence it is omniscient. Therefore we focus on $2 \le k_1^* \le k_2^*$ below. We shall distinguish into two cases: $k_1^* = k_2^*$ and $k_1^* < k_2^*$.

6.5.1 $k_1^* = k_2^* = k^*$

Special Patterns Implied by the Conditions

When the critical nodes are in the same layer, it turns out that if the network $\mathcal{G}$ satisfies the conditions given in Theorem 6.10 or Theorem 6.11, it has a special pattern. The fact is summarized in the following lemma. Let $\mathcal{P}_1 := \mathcal{P}(\mathsf{v}_1^*) \setminus \mathcal{P}(\mathsf{v}_2^*)$, $\mathcal{P}_2 := \mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}(\mathsf{v}_1^*)$, $\mathcal{P}_{12} := \mathcal{P}(\mathsf{v}_1^*) \cap \mathcal{P}(\mathsf{v}_2^*)$.

Lemma 6.16. *When* $k_1^* = k_2^* = k^*$ *and there is no omniscient node, we have the following equivalence relations.*

$$T^{(12)} \iff \begin{cases} \mathcal{P}_1 \text{ is } \mathsf{s}_1\text{-only-reachable} \\ \mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_2) = 1, \mathsf{u}_{21} := \mathsf{Pmc}(\mathcal{P}_2) \neq \mathsf{s}_i, i = 1,2 \\ \mathcal{P}_{12} \text{ is } \mathsf{s}_2\text{-only-reachable} \\ \mathcal{K}^{\mathsf{s}_1}(\mathsf{u}_{21}) \text{ forms } (\mathsf{s}_1;\mathsf{d}_1)\text{-vertex-cut.} \end{cases}$$

$$P^{(12)} \setminus T^{(21)} \iff \begin{cases} \mathcal{P}_1 \text{ is } \mathsf{s}_1\text{-only-reachable} \\ \mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_2) = 1, \mathsf{u}_{21} := \mathsf{Pmc}(\mathcal{P}_2) \neq \mathsf{s}_i, i = 1,2 \\ \mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_{12}) = 1, \mathsf{w}_{12} := \mathsf{Pmc}(\mathcal{P}_{12}) \neq \mathsf{s}_i, i = 1,2 \\ \mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12}) \text{ forms } (\mathsf{s}_2;\mathsf{d}_2)\text{-vertex-cut.} \\ \mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{u}_{21}) \text{ forms } (\mathsf{s}_1;\mathsf{d}_1)\text{-vertex-cut in } \mathcal{G}_{12}. \end{cases}$$

and the equivalence relation for $T^{(21)}$ *(*$P^{(21)} \setminus T^{(12)}$*) is the one for* $T^{(12)}$ *(*$P^{(12)} \setminus T^{(21)}$*) with indices "1" and "2" exchanged.*

Proof. Proof is detailed in Chapter 7. □

One direct consequence of the above lemma is that, $\mathrm{T}^{(12)} \cap \mathrm{T}^{(21)} = \mathrm{P}^{(12)} \cap \mathrm{P}^{(21)} = \emptyset$.

Proof of Achievability

In this case, it is sufficient to show that $\mathsf{v}_1^*, \mathsf{v}_2^*$ can decode the symbols desired by destinations $\mathsf{d}_1, \mathsf{d}_2$ respectively. This is because the network past layer $\mathcal{L}_{k^*}$ has no interference to d_i from any node in $\mathcal{K}(\mathsf{v}_j^*)$ for $(i,j) = (1,2)$ or $(2,1)$.

By definition, we can see that $\mathcal{K}(\mathsf{v}_1^*) \cup \mathcal{K}(\mathsf{v}_2^*) = \mathcal{L}_{k^*}$. For suppose, there exists node $\mathsf{u} \in \mathcal{L}_{k^*} \setminus (\mathcal{K}(\mathsf{v}_1^*) \cup \mathcal{K}(\mathsf{v}_2^*))$. As each node can reach at least one of the destinations, u can reach either d_1 or d_2 thus violating the definition of a critical node. Now, suppose $\mathcal{K}(\mathsf{v}_1^*) \cap \mathcal{K}(\mathsf{v}_2^*) \neq \emptyset$, then $\mathcal{K}(\mathsf{v}_1^*) = \mathcal{K}(\mathsf{v}_2^*)$ and so, both v_1^* and v_2^* are omniscient, violating the assumption. Hence $\mathcal{K}(\mathsf{v}_1^*)$ and $\mathcal{K}(\mathsf{v}_2^*)$ form a partition of $\mathcal{L}_{k^*}$.

Since neither v_1^* nor v_2^* is omniscient, $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \neq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ and $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \neq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. It can be stated equivalently as

$$\begin{aligned}&\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset \text{ or } \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \text{ and}\\ &\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset \text{ or } \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\end{aligned}$$

For notational convenience, let us define $\mathcal{P}_1^{\mathsf{s}_1} := \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ and $\mathcal{P}_2^{\mathsf{s}_2} := \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$.

Below we first show that $(1,1)$ is achievable when the network $\mathcal{G}$ does not fall into any of the above four patterns described in Lemma 6.16. Next we show that in the patterns corresponding to $\mathrm{P}^{(12)} \setminus \mathrm{T}^{(21)}$ and $\mathrm{P}^{(21)} \setminus \mathrm{T}^{(12)}$, both $(1,1/2)$ and $(1/2,1)$ can be achieved. Finally we show that in the pattern corresponding to $\mathrm{T}^{(12)}$, $(1/2,1)$ can be achieved, and in the pattern corresponding to $\mathrm{T}^{(21)}$, $(1,1/2)$ can be achieved.

As a first step, we show the following claim.

Claim 6.17. $(1,1)$ *is achievable if* $\mathcal{P}_1^{\mathsf{s}_1} = \emptyset$ *or* $\mathcal{P}_2^{\mathsf{s}_2} = \emptyset$ *or* $\mathcal{P}_{12} = \emptyset$*, under the assumption that there is no omniscient node.*

Proof. See Chapter 7. ☐

In the following we focus on the case where $\mathcal{P}_1^{\mathsf{s}_1} \neq \emptyset$, $\mathcal{P}_2^{\mathsf{s}_2} \neq \emptyset$, and $\mathcal{P}_{12} \neq \emptyset$. We then show the following claim.

Claim 6.18. *Consider the conditions*

A1 $\forall\ \mathsf{u}_1 \in \mathcal{P}_1^{\mathsf{s}_1}$, u_1 *is* s_1*-only-reachable.*

A2 $\forall\ \mathsf{u}_2 \in \mathcal{P}_2^{\mathsf{s}_2}$, u_2 *is* $\mathsf{s}_1\mathsf{s}_2$*-reachable.*

B1 $\forall\ \mathsf{u}_1 \in \mathcal{P}_1^{\mathsf{s}_1}$, u_1 *is* $\mathsf{s}_1\mathsf{s}_2$*-reachable.*

B2 $\forall\ \mathsf{u}_2 \in \mathcal{P}_2^{\mathsf{s}_2}$, u_2 *is* s_2*-only-reachable.*

Let $A = A1 \wedge A2$ *and* $B = B1 \wedge B2$*. Then the negation of* $A \vee B$ *implies that* $(1,1)$ *is achievable.*

Remark: Note that $A \vee B$ is implied by the disjunction of $\mathrm{T}^{(12)}$, $\mathrm{T}^{(21)}$, $\mathrm{P}^{(12)}$, and $\mathrm{P}^{(21)}$. Therefore this claim proves $(1,1)$-achievability for some cases.

Proof. See Chapter 7. □

So far we have demonstrated $(1,1)$-achievability when condition $A \vee B$ is not satisfied. Since A and B are disjoint, we can separate into two different cases. Besides, discussion on one case will lead to similar arguments for the other case by symmetry.

Case A: $\forall\ \mathsf{u}_1 \in \mathcal{P}_1^{\mathsf{s}_1}$, u_1 is s_1-only-reachable, and $\forall\ \mathsf{u}_2 \in \mathcal{P}_2^{\mathsf{s}_2}$, u_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable.

For this case, if $\mathcal{P}_1 \setminus \mathcal{P}_1^{\mathsf{s}_1} \neq \emptyset$, that is, there exists a node in $\mathcal{P}_1$ and it is s_2-only-reachable, then $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_1\right) = 2$. We can achieve $(1,1)$, by first arranging the transmission of $\mathcal{P}(\mathsf{v}_2^*)$ so that v_2^* can decode b and then arranging the transmission of $\mathcal{P}_1$ to form any linear combination of a and b; in particular, the one that combined with the transmission from $\mathcal{P}_{12}$ forms a at v_1^*. If $\mathcal{P}_2 \setminus \mathcal{P}_2^{\mathsf{s}_2} \neq \emptyset$, that is, there exists a node in $\mathcal{P}_2$ and it is s_1-only-reachable, then $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_2\right) = 2$. $(1,1)$ is then achievable by a similar argument as above.

We now narrow down to the case $\forall\ \mathsf{u}_1 \in \mathcal{P}_1$, u_1 is s_1-only-reachable, and $\forall\ \mathsf{u}_2 \in \mathcal{P}_2$, u_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable. If $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_2\right) = 2$, obviously $(1,1)$ is achievable, as v_1^* can always get a from $\mathcal{P}_1$ (whose transmission does not affect v_2^*) and one can arrange $\mathcal{P}_2$'s transmission (which does not affect v_1^*) to ensure v_2^* decode b. If $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{12}\right) = 2$, we can achieve $(1,1)$ by arranging the transmission of $\mathcal{P}_{12}$ so that their aggregate is a. Hence v_1^* can decode a. Then nodes in $\mathcal{P}_2$ just scale their received linear combinations so that a gets neutralized at v_2^* and only b is left. If $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{12}\right) = 1$, we identify $\mathsf{w}_{12} = \mathsf{Pmc}\left(\mathcal{P}_{12}\right)$. If $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ does not form a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut, we can arrange its parents' transmission so that w_{12} can decode a, and at the same time $\mathcal{P}_2$ can receive a linear combination with a non-zero b-coefficient. Hence nodes in $\mathcal{P}_{12}$ can send out a scaled version of a to neutralize a at v_2^* if necessary, and v_1^* can always obtain a from $\mathcal{P}_1$.

So far we have shown that in Case A, if one of the following is violated, then $(1,1)$ is achievable:

- $\mathcal{P}_1$ is s_1-only-reachable
- $\mathcal{P}_2$ is $\mathsf{s}_1\mathsf{s}_2$-reachable, and $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_2\right) = 1$
- $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{12}\right) = 1$, $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ forms a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut

To complete the proof of $(1,1)$-achievability, we need to show that if $\mathsf{u}_{21} := \mathsf{Pmc}\left(\mathcal{P}_2\right)$ is not ommiscient in $\mathcal{G}_{12}$, then $(1,1)$ can be achieved. We can simply arrange the transmission of $\mathcal{P}_{12}$ so that their aggregate become 0 at v_1^*. Effectively we are in $\mathcal{G}_{12}$ with this special linear coding operation. Since in $\mathcal{G}_{12}$, d_1 is s_1-only-reachable and u_{21} is the new critical node of d_2, by Lemma 6.3 we know that $(1,1)$ can be achieved in $\mathcal{G}_{12}$. We then translate the linear coding scheme in $\mathcal{G}_{12}$ back to a linear coding scheme in $\mathcal{G}$.

The next thing to show for Case A: if a network is in $\mathrm{P}^{(12)} \setminus \mathrm{T}^{(21)}$, then $(1, 1/2)$ can be achieved. To show it, we employ a two-time-slot coding scheme. We aim to deliver two symbols a_1 and a_2 for user 1 and one symbol b for user 2 over two symbol time slots. Symbols are drawn from the extension field $\mathbb{F}_{2^r}$. In the first time slot, we do RLC with s_1 transmitting a_1 and s_2 transmitting b, up to layer $\mathcal{L}_{k^*-1}$. Pick one node $\mathsf{w} \in \mathcal{P}_{12}$ and one node $\mathsf{u} \in \mathcal{P}_2$. Keep other nodes in $\mathcal{P}_{12}$ and $\mathcal{P}_2$ silent, while nodes in $\mathcal{P}_1$ do RLC. We turn off the transmission of w. v_1^* can then decode a_1. In the second time slot again we do RLC with s_1 transmitting a_2 and s_2 transmitting b, up to layer $\mathcal{L}_{k^*-1}$. We use the two linear combinations w receives over the two time slots to zero-force (ZF) b and produce a linear combination of a_1 and a_2: (the superscripts of β's denote the time indices)

$$\begin{aligned} &\beta^{(1)}_{\mathsf{w},\mathsf{s}_1} \cdot a_1 + \beta^{(1)}_{\mathsf{w},\mathsf{s}_2} \cdot b;\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} \cdot a_2 + \beta^{(2)}_{\mathsf{w},\mathsf{s}_2} \cdot b \\ &\overset{\text{ZF}}{\Longrightarrow}\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} \cdot a_2 + \frac{\beta^{(1)}_{\mathsf{w},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w},\mathsf{s}_2}} \cdot \beta^{(2)}_{\mathsf{w},\mathsf{s}_2} \cdot a_1. \end{aligned} \tag{6.1}$$

w then scales this ZF output and sends it out. Hence v_1^* can use a_1 as side information to decode a_2.

As for user 2, in the first time slot v_2^* receives a linear equation from u: $\beta^{(1)}_{\mathsf{u},\mathsf{s}_1} \cdot a_1 + \beta^{(1)}_{\mathsf{u},\mathsf{s}_2} \cdot b$. In the second time slot u receives $\beta^{(2)}_{\mathsf{u},\mathsf{s}_1} \cdot a_2 + \beta^{(2)}_{\mathsf{u},\mathsf{s}_2} \cdot b$. u makes use of of the two linear combinations to zero-force b and generate a linear combination of a_1 and a_2:

$$\begin{aligned} &\beta^{(1)}_{\mathsf{u},\mathsf{s}_1} \cdot a_1 + \beta^{(1)}_{\mathsf{u},\mathsf{s}_2} \cdot b;\ \beta^{(2)}_{\mathsf{u},\mathsf{s}_1} \cdot a_2 + \beta^{(2)}_{\mathsf{u},\mathsf{s}_2} \cdot b \\ &\overset{\text{ZF}}{\Longrightarrow}\ \beta^{(2)}_{\mathsf{u},\mathsf{s}_1} \cdot a_2 + \frac{\beta^{(1)}_{\mathsf{u},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{u},\mathsf{s}_2}} \cdot \beta^{(2)}_{\mathsf{u},\mathsf{s}_2} \cdot a_1. \end{aligned} \tag{6.2}$$

As long as the two linear combinations in (6.1) and (6.2) are not aligned, u can scale (6.2) properly to form a_1 at v_2^* in the second time slot. Then with reception of the first time slot, v_2^* can decode b.

The two linear combinations in (6.1) and (6.2) are aligned if and only if the determinant

$$\begin{vmatrix} \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} & \frac{\beta^{(1)}_{\mathsf{w},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w},\mathsf{s}_2}} \cdot \beta^{(2)}_{\mathsf{w},\mathsf{s}_2} \\ \beta^{(2)}_{\mathsf{u},\mathsf{s}_1} & \frac{\beta^{(1)}_{\mathsf{u},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{u},\mathsf{s}_2}} \cdot \beta^{(2)}_{\mathsf{u},\mathsf{s}_2} \end{vmatrix} = 0 \iff$$
$$\beta^{(1)}_{\mathsf{u},\mathsf{s}_1}\beta^{(1)}_{\mathsf{w},\mathsf{s}_2}\beta^{(2)}_{\mathsf{w},\mathsf{s}_1}\beta^{(2)}_{\mathsf{u},\mathsf{s}_2} = \beta^{(1)}_{\mathsf{u},\mathsf{s}_2}\beta^{(1)}_{\mathsf{w},\mathsf{s}_1}\beta^{(2)}_{\mathsf{w},\mathsf{s}_2}\beta^{(2)}_{\mathsf{u},\mathsf{s}_1},$$

which is of very low probability due to the same reason in Chapter 7.2.

Therefore, we show that in Case A, if $\mathrm{T}^{(12)} \vee \mathrm{P}^{(12)}$ is violated, then $(1, 1)$ can be achieved; if $\mathrm{T}^{(12)}$ is violated, then $(1, 1/2)$ can be achieved. It remains to show that if there is no omniscient node, then in Case A, $(1/2, 1)$ is always achievable.

We aim to deliver one symbol a for user 1 and two symbols b_1, b_2 for user 2 over two time slots. Pick nodes $\mathsf{u}_1 \in \mathcal{P}_1, \mathsf{u}_2 \in \mathcal{P}_2, \mathsf{w}_2 \in \mathcal{P}_{12}$. Both u_2 and w_2 zero-force user 1's symbol a

and form a linear combination of b_1, b_2. These two linear equations are linearly independent with high probability, as shown in Chapter 7.2. w_2 transmits in the first time slot, while u_1 and u_2 transmit in the second time slot. Therefore v_1^* can obtain a, v_2^* can solve b_1 and b_2, and $(1/2, 1)$ is achievable.

Case B: $\forall\ \mathsf{u}_2 \in \mathcal{P}_2^{\mathsf{s}_2}$, u_2 is s_2-only-reachable, and $\forall\ \mathsf{u}_1 \in \mathcal{P}_1^{\mathsf{s}_1}$, u_1 is $\mathsf{s}_1\mathsf{s}_2$-reachable.

Similar to Case A, we show that in Case B, if $\mathrm{T}^{(21)} \vee \mathrm{P}^{(21)}$ is violated, then $(1, 1)$ can be achieved; if $\mathrm{T}^{(21)}$ is violated, then $(1/2, 1)$ can be achieved. Besides, if there is no omniscient node, then in Case B, $(1, 1/2)$ is always achievable.

6.5.2 $k_1^* < k_2^*$

Since v_1^* is not omniscient, $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ does not form a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut, which is equivalent to

$$\exists\ \mathsf{u}_2 \in \mathcal{L}_{k_1^*} \setminus \mathcal{K}(\mathsf{v}_1^*) \text{ so that } \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \emptyset, \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*).$$

The following lemma makes sure that u_2 can still receive a linear combination where user 2's symbol has a non-zero coefficient with high probability even if $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ do some special linear coding.

Lemma 6.19. *Consider all nodes doing RLC for each source sending one symbol up to* $\mathcal{L}_{k_1^*-1}$ *including* $\mathcal{L}_{k_1^*-1}$*, except* $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$*. If* $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$*, then it is possible with high probability that* $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ *can arrange their transmission so that* v_1^* *receives a linear combination solely of user 1's symbol and* u_2 *receives a linear combination of at least user 2's symbol, that is, the coefficient of user 2's symbol is non-zero.*

$(1, 1)$-Achievability

For the $(1, 1)$-achievability we need to prove the following claim

Claim 6.20. $\neg Q^{(12)} \implies (1, 1)$ *is achievable.*

Proof. Since $\mathrm{Q}_1^{(12)}$ is satisfied, in a network that does not satisfy $\mathrm{Q}^{(12)}$, at least one of the following holds:

- $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 2$
- $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$ and $\mathcal{K}_{\mathcal{G}_{12}}^{\mathsf{s}_1}(\mathsf{u}_{21})$ does not form an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$
- $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$ and $\mathcal{K}_{\mathcal{G}_{12}}^{\mathsf{s}_1}(\mathsf{u}_{21})$ forms an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$ and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ does not form a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}$

Case 1: $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 2$

In this case, the idea is to arrange the transmission of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ so that their aggregate contains user 1's symbol a only. Mathematically, we aim to have

$$\sum_{\mathsf{w}\in\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} \alpha_{\mathsf{w}}\beta_{\mathsf{w},\mathsf{s}_1} \neq 0, \quad \sum_{\mathsf{w}\in\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} \alpha_{\mathsf{w}}\beta_{\mathsf{w},\mathsf{s}_2} = 0, \tag{6.3}$$

This is doable since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 2$. Other nodes in the same layer simply do RLC. Therefore v_1^* is able to decode a. Nodes in $\mathcal{C}_1$ do RLC and d_1 can decode a.

As for user 2, we look at v_2^*. If v_2^* has no parents in the cloud $\mathcal{C}_1$, we only need to guarantee that the parents of v_2^* can collectively decode b. If v_2^* has some parent(s) in the cloud $\mathcal{C}_1$, the parent(s) will inject user 1's symbol a to the reception of v_1^*. As v_2^* is not omniscient, there must exist $\mathsf{u}_1 \in \mathcal{L}_{k_2^*} \cap \mathcal{C}_1$ such that $\mathcal{P}(\mathsf{u}_1) \neq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$. If $\mathcal{P}(\mathsf{u}_1) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$, we can arrange the nodes in $\mathcal{P}(\mathsf{u}_1)$ to make sure that u_1 can decode user 1's symbol, and then arrange the nodes in $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}(\mathsf{u}_1)$ to neutralize user 1's symbol at v_2^*, given that these s_1-reachable nodes can still receive a linear combination with non-zero a-coefficient under the special coding carried out by $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. Then since some node in $\mathcal{P}(\mathsf{v}_2^*)$ will receive a linear combination with non-zero b-coefficient (eg., a successor of u_2 in Lemma 6.19), one can always ensure v_2^* to decode b. If $\mathcal{P}(\mathsf{u}_1) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset$, we can arrange the nodes in $\mathcal{P}(\mathsf{v}_2^*)$ to form a linear combination that only contains b at v_2^* given that the reception of $\mathcal{P}(\mathsf{v}_2^*)$ can collectively decode b. Then use nodes in $\mathcal{P}(\mathsf{u}_1) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ to place user 1's symbol at u_1 if necessary.

In summary, we want to guarantee that under the special linear coding carried out by $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ so that the neutralization criterion (6.3) is met, $\mathcal{P}(\mathsf{v}_2^*)$ can still collectively decode b and every node in $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ receives a linear combination with non-zero a-coefficient. The latter is quite obvious, as nodes affected by the special linear coding still receive linear combinations with non-zero a-coefficients. As for the former, note that if every node up to layer $k_2^* - 2$ does RLC, it holds with high probability since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_2^*)\right) = 2$. With the special linear coding carried out by $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ described above, however, we cannot claim it with the existing random linear network coding argument.

We shall use the following two lemmas to overcome the difficulty, by breaking the network into two stages: one from the source layer to the layer $\mathcal{L}_{k_1^*}$, and the other from layer $\mathcal{L}_{k_1^*}$ to layer $\mathcal{L}_{k_2^*-1}$. The first lemma claims that, under the special operation at $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ so that neutralization criterion (6.3) is satisfied, with high probability all nodes in layer $\mathcal{L}_{k_1^*}$ that can reach d_2 (call this set $\mathcal{U}$) receive a non-zero linear combination of a and b, and the subspace spanned by their reception has dimension two when all other nodes perform RLC. The second lemma claims that, once $\mathcal{U}$'s reception satisfies the above property, then $\mathcal{P}(\mathsf{v}_2^*)$ can collectively decode both a and b with high probability, when all nodes between $\mathcal{L}_{k_1^*}$ and $\mathcal{L}_{k_2^*-1}$ perform RLC. The lemmas are made concrete below.

Lemma 6.21 (Reception of $\mathcal{U}$)**.** *Let us recall that* $\mathcal{U} := \left\{\mathsf{u} \in \mathcal{L}_{k_1^*} : \mathsf{u} \text{ can reach } \mathsf{d}_1\right\}$. *Consider RLC with* s_1 *transmitting* a *and* s_2 *transmitting* b. *All nodes perform RLC up to layer* $\mathcal{L}_{k_1^*-1}$. *In* $\mathcal{L}_{k_1^*-1}$, *nodes except* $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ *also perform RLC. Then under special coding operation of*

$\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ such that neutralization criterion (6.3) is satisfied, with high probability all nodes in $\mathcal{U}$ receive a non-zero linear combination of a and b, and the subspace spanned by their reception has dimension two. Further, if node $\mathsf{u} \in \mathcal{U}$ is s_1-reachable, then its reception has a non-zero coefficient of s_1's symbol a with high probability.

Lemma 6.22 (Rank Conservation)**.** *Suppose $\mathcal{U}$ and $\mathcal{V}$ are the first and the last layers of a linear deterministic network and each node $\mathsf{u} \in \mathcal{U}$ possesses a linear combination of the symbols a, b given by $\lambda_{\mathsf{u}} \cdot a + \mu_{\mathsf{u}} \cdot b$. Suppose*

- *each node in $\mathcal{U}$ can reach some node in $\mathcal{V}$,*
- $\mathrm{C}\left(\mathcal{U}; \mathcal{V}\right) \geq 2$,
- *for each $\mathsf{u} \in \mathcal{U}$, we have $\lambda_{\mathsf{u}}, \mu_{\mathsf{u}}$ not both 0,*
- *the $|\mathcal{U}| \times 2$ matrix with rows given by $\begin{bmatrix} \lambda_{\mathsf{u}} & \mu_{\mathsf{u}} \end{bmatrix}$ for each $\mathsf{u} \in \mathcal{U}$ has full rank (i.e. rank 2).*

If all nodes in the network perform RLC, then nodes in $\mathcal{V}$ can collectively decode both the symbols a and b with high probability.

Case 2: $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$ and $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{u}_{21})$ does not form an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$

In this case, since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$, effectively they receive only one linear equation of a and b. Since $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ can be reached by s_2, the coefficient of b in this linear equation is non-zero in general. Hence we need to arrange their transmission so that $\sum_{\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} X_{\mathsf{w}} = 0$, that is,

$$\sum_{\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} \alpha_{\mathsf{w}} \beta_{\mathsf{w}, \mathsf{s}_1} = \sum_{\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} \alpha_{\mathsf{w}} \beta_{\mathsf{w}, \mathsf{s}_2} = 0, \tag{6.4}$$

Since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)\right) = 2$, v_1^* must have some s_1-only-reachable parents. Therefore v_1^* can decode a.

With such special operation in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, effectively we are in the induced graph $\mathcal{G}_{12}$. In other words, any linear coding scheme in the induced graph $\mathcal{G}_{12}$ can be translated to a linear coding scheme in $\mathcal{G}$ satisfying the neutralization criterion (6.4), in the sense that the reception of d_i remains the same in both schemes, for $i = 1, 2$. In $\mathcal{G}_{12}$, note that d_1 can only be reached by s_1 but not s_2. Hence by Lemma 6.3, as long as the critical node for destination d_2 in $\mathcal{G}_{12}$, u_{21}, is not omniscient, $(1, 1)$ is achievable. u_{21} is not omniscient in $\mathcal{G}_{12}$ by the assumption of this case.

Case 3: $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$ and $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{u}_{21})$ forms an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$ and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ does not form an $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}$

In this case the idea is to enable w_{12} to decode user 1's symbol a while keeping user 2's flow to v_2^*, making use of the fact that $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ does not form a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut in $\mathcal{G}$.

Effectively we impose the neutralization criterion on w_{12} instead of v_1^*, and carry out the special coding operation at $\mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})$ instead of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$.

As for user 1, obviously $\mathsf{w}_{12} \neq \mathsf{s}_2$, and hence it can be reached by s_1 due to the definition of critical nodes. Since $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{w}_{12})\right) = 2$, we can enable w_{12} to decode a by satisfying the following neutralization condition:

$$\sum_{\mathsf{w}\in\mathcal{P}(\mathsf{w}_{12})} \alpha_{\mathsf{w}}\beta_{\mathsf{w},\mathsf{s}_1} = \beta_{\mathcal{P}(\mathsf{w}_{12}),\mathsf{s}_1}, \quad \sum_{\mathsf{w}\in\mathcal{P}(\mathsf{w}_{12})} \alpha_{\mathsf{w}}\beta_{\mathsf{w},\mathsf{s}_2} = 0. \tag{6.5}$$

Once w_{12} decodes a, it simply sends out a scaled copy of a. With all other nodes performing RLC up to layer $\mathcal{L}_{k_1^*-1}$ (including $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$), critical node v_1^* can decode a.

As for user 2, note that depending on the value of $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})\right)$ being 2 or 1, $\beta_{\mathcal{P}(\mathsf{w}_{12}),\mathsf{s}_1}$ is either non-zero or zero. If $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})\right) = 2$, we use the two lemmas, Lemma 6.21 and 6.22, in the first case to show that the parents of v_2^* can recover both user's symbols with high probability under the special operation at $\mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})$. If $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})\right) = 1$, we construct *another* induced graph $\mathcal{G}_{12}'$ to capture the constraints that such special coding lays on the reception of other nodes in the same layer as w_{12}, which is similar to $\mathcal{G}_{12}$ in the second case. In $\mathcal{G}_{12}'$, the critical node for user 2 may no longer be v_2^*, as the min-cut value from the sources $\{\mathsf{s}_1, \mathsf{s}_2\}$ to the parents of v_1^* may drop to 1. Note that as in $\mathcal{G}_{12}$, destination d_1 is now s_1-only-reachable in $\mathcal{G}_{12}'$. Hence we only need the new critical node for destination d_2 is not omniscient in $\mathcal{G}_{12}'$.

The following lemma guarantees it in this case.

Lemma 6.23. *In this case (Case 3) when* $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})\right) = 1$*, for all possible* $\mathcal{G}_{12}'$*, the* s_1*-clones of* $\mathsf{Pmc}_{\mathcal{G}_{12}'}(\mathsf{d}_2)$ *do not form a* $(\mathsf{s}_1; \mathsf{d}_1)$*-vertex-cut.*

Combining the above three cases, we complete the proof for the claim and the $(1, 1)$-achievability. $\square$

$(1, 1/2)$-Achievability

For the $(1, 1/2)$-achievability we need to prove the following claim

Claim 6.24. $P^{(12)} \implies (1, 1/2)$ *is achievable.*

Proof. Consider two cases, distinguishing whether v_2^* has parents from the cloud or not.

1) $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 \neq \emptyset$: Under the condition that $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 \neq \emptyset$, we know that in $\mathcal{G}_{12}$ the critical node for d_2 is still v_2^*, as $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)\right) = 2$. This is because nodes in the cloud $\mathcal{C}_1$ become s_1-only-reachable in $\mathcal{G}_{12}$ while some nodes in $\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ are s_2-reachable in $\mathcal{G}_{12}$. $\mathrm{P}^{(12)}$ implies that the s_1-clones of v_2^* in $\mathcal{G}_{12}$ becomes a $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut. Therefore, some nodes in $\mathcal{P}(\mathsf{v}_2^*)$ must be dropped in generating $\mathcal{G}_{12}$ (as they cannot be reached by either one of the sources), and $\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*) \neq \mathcal{P}(\mathsf{v}_2^*)$.

We aim to deliver two symbols a_1, a_2 for user 1 and one symbol b for user 2 over two symbol time slots. Symbols are drawn from the extension field $\mathbb{F}_{2^r}$. In the first time slot we do RLC with s_1 transmitting a_1 and s_2 transmitting b, up to layer $\mathcal{L}_{k_1^*-1}$. Then we arrange the transmission of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ so that their aggregate becomes zero, as in Case 2. v_1^* can hence decode a_1, and transmit a scaled version of it. The rest of the nodes keep performing RLC. It is as if the communication is over the induced graph $\mathcal{G}_{12}$, and effectively nodes in $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ will receive nothing. As the s_1-clones of v_2^* form a $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$, v_2^* will receive a linear equation of a_1 and b, and both symbols have non-zero coefficients. Therefore, d_1 can decode a_1 in the first time slot.

In the second time slot, we do RLC with s_1 transmitting a_2 and s_2 transmitting b, up to the layer right before w_{12}. For those nodes in $\mathcal{K}(\mathsf{w}_{12})$ that can reach $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, instead of scaling their reception and transmitting it, they replace their reception by a linear combination of a_1 and a_2. This linear combination is obtained by zero-forcing b using the reception of the first and the second time slot:

$$\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1} \cdot a_1 + \beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2} \cdot b;\ \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_1} \cdot a_2 + \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_2} \cdot b$$
$$\overset{\text{ZF}}{\Longrightarrow}\ \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_1} \cdot a_2 + \frac{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}} \cdot \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_2} \cdot a_1.$$

The rest of the nodes perform RLC up to layer $\mathcal{L}_{k_1^*}$. Since v_1^* already obtains a_1 in the first time slot and it receives a linear combination of a_1, a_2 with non-zero a_2-coefficient in the second time slot, it can decode a_2. Onwards it transmits a scaled copy of a_2, while other nodes perform RLC. The nodes in $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$, unlike in the first time slot, will receive a linear combination of a_1, a_2, which is a scaled version of that transmitted by w_{12}. Hence, we can arrnage the transmission of $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ and $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1$ so that v_2^* can decode a_1. Therefore, using the reception from the first time slot, v_2^* can decode b.

2) $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 = \emptyset$: We aim to deliver two symbols a_1 and a_2 for user 1 and one symbol b for user 2 over two symbol time slots. Again the symbols are drawn from the extension field $\mathbb{F}_{2^r}$. In the first time slot, we do RLC with s_1 transmitting a_1 and s_2 transmitting b, up to layer $\mathcal{L}_{k_1^*-1}$. Then we arrange the transmission of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ so that their aggregate becomes zero, as in Case 2. It is as if the communication is over the induced graph $\mathcal{G}_{12}$. Since u_{21} is the critical node for the parents of v_2^* in $\mathcal{G}_{12}$, in the first time slot the they effectively receive only one equation

$$\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1} \cdot a_1 + \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2} \cdot b,$$

where $\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1}$ is non-zero with high probability since $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{u}_{21})$ forms a $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$ and hence u_{21} must be reachable by s_1. $\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2}$ is non-zero with high probability since v_1^* is not omniscient and hence s_2 must be able to reach v_2^* in $\mathcal{G}_{12}$.

In the second time slot, we do RLC with s_1 transmitting a_2 and s_2 transmitting b, up to the layer right before w_{12}. For those nodes in $\mathcal{K}(\mathsf{w}_{12})$ that can reach $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, instead of scaling their reception and transmitting it, they replace their reception by a linear combination of a_1 and a_2. This linear combination is obtained by zero-forcing b using the reception of the first and the second time slot:

$$\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}\cdot a_1+\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}\cdot b;\ \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_1}\cdot a_2+\beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_2}\cdot b$$
$$\overset{\text{ZF}}{\Longrightarrow}\ \beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_1}\cdot a_2+\frac{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}}\cdot\beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_2}\cdot a_1.$$

The rest of the nodes remain doing RLC, up to the layer right before v_2^*. In the second time slot, v_2^*'s parents receive at least two linear equations in $\{a_1,a_2,b\}$. Pick two nodes $\mathsf{u},\mathsf{w}\in\mathcal{P}(\mathsf{v}_2^*)$ such that $\mathrm{C}\left(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u},\mathsf{w}\right)=2$. Let their reception be

$$\beta_{\mathsf{u},a_1}\cdot a_1+\beta_{\mathsf{u},a_2}\cdot a_2+\beta_{\mathsf{u},b}\cdot b,$$
$$\beta_{\mathsf{w},a_1}\cdot a_1+\beta_{\mathsf{w},a_2}\cdot a_2+\beta_{\mathsf{w},b}\cdot b,$$

respectively. We shall show that the following determinant

$$\begin{vmatrix}\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1} & 0 & \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2}\\ \beta_{\mathsf{u},a_1} & \beta_{\mathsf{u},a_2} & \beta_{\mathsf{u},b}\\ \beta_{\mathsf{w},a_1} & \beta_{\mathsf{w},a_2} & \beta_{\mathsf{w},b}\end{vmatrix} = \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2}\begin{vmatrix}\beta_{\mathsf{u},a_1} & \beta_{\mathsf{u},a_2}\\ \beta_{\mathsf{w},a_1} & \beta_{\mathsf{w},a_2}\end{vmatrix}+\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1}\begin{vmatrix}\beta_{\mathsf{u},a_2} & \beta_{\mathsf{u},b}\\ \beta_{\mathsf{w},a_2} & \beta_{\mathsf{w},b}\end{vmatrix} \tag{6.6}$$

is non-zero with high probability.

Note that in the second time slot, we choose the scaling coefficients α's for all nodes up to the layer right before v_2^* in the same way as RLC. The only difference from RLC is that at the nodes in $\mathcal{K}(\mathsf{w}_{12})$ that can reach $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, the term scaled and transmitted is replaced by the zero-forced output $\beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_1}\cdot a_2+\frac{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}}\cdot\beta^{(2)}_{\mathsf{w}_{12},\mathsf{s}_2}\cdot a_1$. Suppose we do RLC, then u and w will receive

$$\beta^{(2)}_{\mathsf{u},\mathsf{s}_1}\cdot a_2+\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\cdot b,\ \text{and}\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1}\cdot a_2+\beta^{(2)}_{\mathsf{w},\mathsf{s}_2}\cdot b$$

respectively, where $D^{(2)}:=\begin{vmatrix}\beta^{(2)}_{\mathsf{u},\mathsf{s}_1} & \beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} & \beta^{(2)}_{\mathsf{w},\mathsf{s}_2}\end{vmatrix}\neq 0$ with high probability since $\mathrm{C}\left(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u},\mathsf{w}\right)=2$. As pointed out above, from the connection of the scheme to RLC, we see

$$\begin{vmatrix}\beta_{\mathsf{u},a_1} & \beta_{\mathsf{u},a_2}\\ \beta_{\mathsf{w},a_1} & \beta_{\mathsf{w},a_2}\end{vmatrix}=\frac{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}}\cdot D^{(2)}_Z,\quad \begin{vmatrix}\beta_{\mathsf{u},a_2} & \beta_{\mathsf{u},b}\\ \beta_{\mathsf{w},a_2} & \beta_{\mathsf{w},b}\end{vmatrix}=D^{(2)}_R,$$

where

$$D^{(2)}_Z:=\begin{vmatrix}\beta^{(2)}_{\mathsf{u},\mathsf{s}_1} & \left[\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\right]_Z\\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} & \left[\beta^{(2)}_{\mathsf{w},\mathsf{s}_2}\right]_Z\end{vmatrix},$$

and

$$D_R^{(2)} := \begin{vmatrix} \beta^{(2)}_{\mathsf{u},\mathsf{s}_1} & \left[\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\right]_R \\ \beta^{(2)}_{\mathsf{w},\mathsf{s}_1} & \left[\beta^{(2)}_{\mathsf{w},\mathsf{s}_2}\right]_R \end{vmatrix}.$$

Here $\left[\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\right]_R$ denotes the coefficient of b that u receives under a *virtual* RLC with the same coding operation as regular RLC except that nodes in $\mathcal{K}(\mathsf{w}_{12})$ that can reach $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ (call this set $\mathcal{Z}$) are transmitting scaled copies of a (with the same scaling coefficient as in regular RLC) instead of a linear combination of a, b. $\left[\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\right]_Z$ denotes the coefficient of b that u receives under a *virtual* RLC with the same coding operation as regular RLC except that the s_2-reachable predecessors of u in the same layer as w_{12} other than $\mathcal{Z}$ are transmitting scaled copies of the a-components in their reception (with the same scaling coefficient). Note that if there is no s_2-reachable predecessor of u in $\mathcal{Z}$, then $\left[\beta^{(2)}_{\mathsf{u},\mathsf{s}_2}\right]_Z = 0$. We have $D_Z^{(2)} + D_R^{(2)} = D^{(2)}$. The determinant in (6.6) equals to zero if and only if

$$\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2}\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1} D_Z^{(2)} + \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1}\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2} D_R^{(2)} = 0.$$

Suppose $\begin{vmatrix} \beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1} & \beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2} \\ \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1} & \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2} \end{vmatrix}$ is a zero polynomial, then we are done since $D^{(2)} \neq 0$ with high probability.

Note that $D_Z^{(2)}$ and $D_R^{(2)}$ cannot simultaneously be zero with high probability, as their sum is non-zero with high probability. First assume that $D_Z^{(2)} \neq 0$. The determinant in (6.6) equals to zero if and only if

$$\frac{D_R^{(2)}}{D_Z^{(2)}} = \frac{\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2}\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1}\beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2}}.$$

RHS and LHS are independent. We only need to consider the case $\begin{vmatrix} \beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_1} & \beta^{(1)}_{\mathsf{w}_{12},\mathsf{s}_2} \\ \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_1} & \beta^{(1)}_{\mathsf{u}_{21},\mathsf{s}_2} \end{vmatrix}$ is not a zero polynomial. The probability distribution of RHS is going to spread out over the spectrum (see Lemma 7.1 in Chapter 7.2). Hence the above equality happens with vanishing probability.

Similar conclusion can be drawn in the case $D_R^{(2)} \neq 0$. □

$(1/2, 1)$-Achievability

For the $(1/2, 1)$-achievability, we argue that if there is no omniscient node, then $(1/2, 1)$ is achievable. We shall use nodes reachable from u_2 in $\mathcal{P}(\mathsf{v}_2^*)$ to provide user 2's symbols. Define the collection of these nodes by $\mathcal{S}_{k_2^*-1}(\mathsf{u}_2)$. Consider the following two cases.

1) $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 \neq \emptyset$: If $\mathrm{Q}^{(12)}$ is violated, $(1,1)$ can be achieved and so can $(1/2,1)$. Hence we focus on the case in which $\mathrm{Q}^{(12)}$ is satisfied. Under the condition that $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 \neq \emptyset$, from the analysis of the previous case we know that some nodes in $\mathcal{P}(\mathsf{v}_2^*)$ must be dropped in generating $\mathcal{G}_{12}$, and $\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*) \neq \mathcal{P}(\mathsf{v}_2^*)$.

We aim to deliver one symbol a for user 1 and two symbols b_1, b_2 for user 2 over two time slots. In the first time slot, all nodes up to layer $k_2^* - 1$ perform RLC with s_1 transmitting a scaled copy of a, and s_2 transmitting a scaled copy of b_1, except that nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ perform special linear coding to make sure their aggregate transmission is zero. Hence effectively we are in $\mathcal{G}_{12}$, and the nodes in $\mathcal{P}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$, which form a $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$ and therefore lie in the cloud $\mathcal{C}_1$, can decode a. Since $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 2$, we can arrange the transmission of $\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ so that v_2^* can decode b_1 and so can d_2. But d_1 will receive nothing, as $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ is a $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$. In the second time slot, all nodes up to layer $k_2^* - 1$ perform RLC with s_1 transmitting a scaled copy of a, and s_2 transmitting a scaled copy of b_2. This time the nodes in $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ will receive a non-triavial linear combination of a and b_2 with a non-zero a-coefficient. Then we let nodes in $\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ transmit a scaled copy of a by choosing their scaling coefficients uniformly and independently, while using nodes in $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)$ to neutralize the symbol a in the reception of v_2^* and obtain a clean copy of b_2. Hence d_1 can decode a, and d_2 can decode b_2 in the second time slot.

2) $\mathcal{P}(\mathsf{v}_2^*) \cap \mathcal{C}_1 = \emptyset$: We aim to prove that $(1/2,1)$ is achievable in this case. User 1 has one symbol a and user 2 has two symbols b_1, b_2 to be delivered over two time slots.

In the first time slot, all nodes up to layer $k_2^* - 1$ perform RLC with s_1 transmitting a scaled copy of a, and s_2 transmitting a scaled copy of b_1. Note that because $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}(\mathsf{v}_2^*)) = 2$, we have that $\mathcal{P}(\mathsf{v}_2^*)$ can collectively decode both a and b_1 due to Lemma 6.14(a). In the second time slot, all nodes up to layer $k_1^* - 2$ perform RLC with s_1 sending a and s_2 sending b_2. Due to Lemma 6.14(a), we can arrange the transmission of $\mathcal{P}(\mathsf{v}_1^*)$ so that v_1^* receives only a in the second time slot, since $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}(\mathsf{v}_1^*)) = 2$. As $\emptyset \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, u_2 receives a linear combination with a non-zero coefficient of user 2's symbol due to Lemma 6.19. Further, all nodes perform RLC up to layer $k_2^* - 1$. As u_2 has a path to $\mathcal{P}(\mathsf{v}_2^*)$, some node in $\mathcal{P}(\mathsf{v}_2^*)$ receives a linear combination of the three symbols with a non-zero coefficient for b_2. Thus, $\mathcal{P}(\mathsf{v}_2^*)$ can collectively decode all three symbols a, b_1, b_2. Since this decoding is a linear operation, these nodes can arrange their transmissions so as to form b_1 and b_2 at v_2^*'s reception in first and second time slots respectively. All nodes in layer k_2^* onwards perform RLC with no mixing across time slots. Thus, d_2 can recover both b_1 and b_2. As nodes in $\mathcal{L}_{k_2^*-1} \cap \mathcal{C}_1$ perform RLC with no mixing across time slots, destination d_1 can recover both the symbols a and b_1.

6.6 Proof of Outer Bounds

6.6.1 Outer Bound on $R_1 + R_2$: the Omniscient Bound

We show that if a node v is omniscient, then it can decode both user's messages and hence, the achievable sum rate is upper bounded by 1. This explains the motivation for the name. Let v be omniscient and satisfy condition (A) in the definition of omniscient nodes: $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ is a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut. Since $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut, the reception of the destination d_1 is a function of the reception of v. This means v can decode the message of s_1. The reception of each node in $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ is some function of the reception of node v and the transmission of s_1. Since v can now recover the transmission of s_1, and since $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ forms a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut, v can recover the reception of d_2, and thus, also the message of s_2. We leave the formal proof of this outer bound in Chapter 7.

6.6.2 Outer Bounds on $2R_1 + R_2$ and $R_1 + 2R_2$

We want to show that if the condition $\mathrm{T}^{(12)}$ is satisfied, then $2R_1 + R_2 \leq 2$ for any achievable (R_1, R_2). We first show the following claim.

Claim 6.25. *If there exists random variables $\{Z_1, Z_{21}, Z_{22}\}$ in the network satisfying*

1) $H(Z_1) \leq 1, H(Z_{21}) \leq 1, H(Z_{22}) \leq 1.$
2) $X_{\mathsf{s}_1}^N \leftrightarrow Z_1^N \leftrightarrow Y_{\mathsf{d}_1}^N$ and $X_{\mathsf{s}_2}^N \leftrightarrow (Z_{21}^N, Z_{22}^N) \leftrightarrow Y_{\mathsf{d}_2}^N$
3) $X_{\mathsf{s}_1}^N \leftrightarrow (Z_{21}^N, X_{\mathsf{s}_2}^N) \leftrightarrow Y_{\mathsf{d}_1}^N$
4) $H\left(Z_1^N | X_{\mathsf{s}_1}^N\right) \geq H\left(Z_{22}^N\right)$
5) Z_{22}^N is a function of $X_{\mathsf{s}_2}^N$

then $2R_1 + R_2 \leq 2$ for any achievable (R_1, R_2).

Proof is detailed in Chapter 7.

We shall use the above claim to complete the proof of the outer bound $2R_1 + R_2 \leq 2$. We set $Z_1 := Y_{\mathsf{v}_1^*}$, $Z_{21} := Y_{\mathsf{u}_{21}}$, $Z_{22} := \sum_{\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}\left(\mathsf{v}_1^*\right)} X_{\mathsf{w}}$.

- Hence by the definition of the channels, condition 1) of the claim is satisfied.
- By the definition of v_1^*, we see that the first Markov chain in condition 2) is satisfied. By condition $\mathrm{T}_3^{(12)}$ and the definition of the induced graph $\mathcal{G}_{12}$ we see that the second Markov chain is also satisfied. Hence, condition 2) is satisfied.
- By conditions $\mathrm{T}_3^{(12)}$ and $\mathrm{T}_4^{(12)}$, we see that the Markov chain in condition 3) is satisfied.

- Condition 4) is satisfied with equality due to the definition of Z_{22} and condition $\mathrm{T}_4^{(12)}$.
- Condition 5) is satisfied due to the definition of Z_{22} and conditions $\mathrm{T}_2^{(12)}$ and $\mathrm{T}_4^{(12)}$.

Similarly, if the condition $\mathrm{T}^{(21)}$ is satisfied, then $R_1 + 2R_2 \leq 2$ by symmetry.

6.6.3 Outer Bound on $2R_1 + 2R_2$

We want to show that if the condition $\mathrm{P}^{(12)}$ is satisfied, then $2R_1+2R_2 \leq 3$ for any achievable (R_1, R_2). We first show the following claim.

Claim 6.26. *If there exists random variables $\{Z_{11}, Z_{12}, Z_{21}, Z_{22}\}$ in the network satisfying*

1) $H(Z_{11}) \leq 1,\ H(Z_{12}) \leq 1,\ H(Z_{21}) \leq 1,\ H(Z_{22}) \leq 1$
2) $X_{\mathsf{s}_1}^N \leftrightarrow Z_{11}^N \leftrightarrow Y_{\mathsf{d}_1}^N$ *and* $X_{\mathsf{s}_2}^N \leftrightarrow (Z_{21}^N, Z_{22}^N) \leftrightarrow Y_{\mathsf{d}_2}^N$
3) $X_{\mathsf{s}_1}^N \leftrightarrow (Z_{21}^N, Z_{22}^N, X_{\mathsf{s}_2}^N) \leftrightarrow Y_{\mathsf{d}_1}^N$ *and*
$X_{\mathsf{s}_2}^N \leftrightarrow (Z_{12}^N, X_{\mathsf{s}_1}^N) \leftrightarrow Y_{\mathsf{d}_2}^N$
4) $H\left(Z_{11}^N | X_{\mathsf{s}_1}^N\right) \geq H\left(Z_{22}^N | X_{\mathsf{s}_1}^N\right)$
5) Z_{22}^N *is a function of* Z_{12}^N

then $2R_1 + 2R_2 \leq 3$ for any achievable (R_1, R_2).

Proof is detailed in Chapter 7.

We shall use the above claim to complete the proof of the outer bound $2R_1 + 2R_2 \leq 3$. We set $Z_{11} := Y_{\mathsf{v}_1}$, $Z_{12} := Y_{\mathsf{w}_{12}}$, $Z_{21} := Y_{\mathsf{u}_{21}}$, $Z_{22} := \sum_{\mathsf{w}\in\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)} X_{\mathsf{w}}$.

- By the definition of the channels, condition 1) of the claim is satisfied.
- By the definition of v_1^*, we see that the first Markov chain in condition 2) is satisfied. By condition $\mathrm{P}_3^{(12)}$ and the definition of the induced graph $\mathcal{G}_{12}$ we see that the second Markov chain is also satisfied. Hence, condition 2) is satisfied.
- The first Markov chain in condition 3) is due to condition $\mathrm{P}_3^{(12)}$ and the definition of the induced graph $\mathcal{G}_{12}$. The second Markov chain is due to condition $\mathrm{P}_4^{(12)}$. Hence condition 3) is satisfied.
- Condition 4) is satisfied with equality due to the definition of Z_{22}.
- Condition 5) is satisfied due to the definition of Z_{12}, Z_{22} and conditions $\mathrm{P}_2^{(12)}$ and $\mathrm{P}_4^{(12)}$.

Similarly, if the condition $\mathrm{P}^{(21)}$ is satisfied, then $2R_1 + 2R_2 \leq 3$ by symmetry.

6.7 Concluding Remarks

In this chapter, we completely characterize the capacity region of two unicast information flows over a layered linear deterministic network with base field $\mathbb{F}_2$ under the unit-channel strength assumption. It turns out that when each source can reach its own destination, the capacity region is one of the five: the triangle $\mathfrak{T}$, the trapezoids $\mathfrak{T}_{12}, \mathfrak{T}_{21}$, the pentagon $\mathfrak{P}$, and the square $\mathfrak{S}$. The necessary and sufficient condition for the capacity region to be one of them elucidates when and how the connectivity of the network limits the amount of information deliverable to the destination under the presence of the other interfering information flow.

Our result extends to a more general linear deterministic channel setting where a general matrix in $\mathbb{F}_2$ (not necessarily a shift matrix) is associated to each edge in the network. Such generalization is made possible by looking at entries of the receive/transmit vectors, called "bubbles", and redefining omniscience, clone sets, parents, cuts, etc., for bubbles. This result will be detailed in a later version of [17].

Chapter 7

Proofs of Part III

7.1 Proof of Lemmas and Claims

By the phrase "with high probability", we mean a probability that goes to 1 as the size of the field $\mathbb{F}_{2^r}$ goes to infinity.

7.1.1 Proof of Lemma 6.3

If $k_1^* = 0$ and $k_2^* = 1$, then both s_1 and s_2 are v_2^*'s parents, and obviously v_2^* is omniscient, violating the assumption. Hence, $k_2^* = 0$ or $k_2^* \geq 2$. If $k_2^* = 0$, then there is no interference at destination d_1 from source s_2 and vice versa. In this case, clearly $(1, 1)$ is achievable.

If $k_2^* \geq 2$, we shall show that $(1, 1)$ can be achieved provided that there is no omniscient node. Nodes do RLC with s_1 transmitting a scaled copy of symbol $a \in \mathbb{F}_{2^r}$ and s_2 transmitting a scaled copy of symbol $b \in \mathbb{F}_{2^r}$, until layer $\mathcal{L}_{k_2^*-1}$. By definitions of v_2^* and $\mathcal{C}_1$, layer $\mathcal{L}_{k_2^*}$ is partitioned by $\mathcal{K}(\mathsf{v}_2^*)$ and $\mathcal{C}_1 \cap \mathcal{L}_{k_2^*}$. Since v_2^* is not omniscient,

$$\exists\ \mathsf{u}_1 \in \mathcal{C}_1 \cap \mathcal{L}_{k_2^*} \text{ such that } \mathcal{P}^{\mathsf{s}_1}(\mathsf{u}_1)(= \mathcal{P}(\mathsf{u}_1)) \neq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*).$$

Note that since $\mathsf{u}_1 \in \mathcal{C}_1$, $\mathcal{P}^{\mathsf{s}_1}(\mathsf{u}_1) = \mathcal{P}(\mathsf{u}_1)$. Also note that all nodes in $\mathcal{C}_1$ are s_1-only-reachable. Consider the following two cases:

1. $\mathcal{P}(\mathsf{u}_1) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset$. In this case we arrange nodes in $\mathcal{P}(\mathsf{v}_2^*)$ so that user 2's symbol b can be decoded at v_2^*. Then use nodes in $\mathcal{P}(\mathsf{u}_1) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ to provide user 1's symbol a at u_1 if necessary. u_1 and v_2^* and their successors do RLC.

2. $\mathcal{P}(\mathsf{u}_1) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$. In this case we first let nodes in $\mathcal{P}(\mathsf{u}_1)$ do RLC and place user 1's symbol a at u_1. Then, use nodes in $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}(\mathsf{u}_1)$ and nodes in $\mathcal{P}(\mathsf{v}_2^*)$ to neutralize user 1's symbol a and place user 2's symbol b at v_2^*. u_1 and v_2^* and their successors do RLC.

Hence, $(1, 1)$ is achievable when $k_1^* = 0$.

7.1.2 Proof of Lemma 6.5

Suppose $i = 1$. Suppose $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) \neq 2$. It cannot be larger than 2 by definition and it cannot be 0 because $\{\mathsf{s}_1, \mathsf{s}_2\}$ has paths to $\mathcal{P}(\mathsf{v}_1^*)$. So, suppose $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) = 1$.

Let $\mathcal{A} \subseteq \mathcal{V}$ be the set of nodes in the graph that can be reached by $\{\mathsf{s}_1, \mathsf{s}_2\}$ and can reach $\mathcal{P}(\mathsf{v}_1^*)$. Let $\mathcal{G}'$ be the graph induced by nodes in $\mathcal{A}$ and for $\mathcal{U} \subseteq \mathcal{A}$, let $\mathrm{C}_{\mathcal{G}'}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U})$ denote the mincut from $\{\mathsf{s}_1, \mathsf{s}_2\}$ to $\mathcal{U}$ in the graph $\mathcal{G}'$. Then, obviously $\mathrm{C}_{\mathcal{G}'}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) = 1$.

Note that for any partition of the vertices of $\mathcal{A}$ into $(B, \mathcal{A} \setminus B)$ with $\{\mathsf{s}_1, \mathsf{s}_2\} \subseteq B, \mathcal{P}(\mathsf{v}_1^*) \subseteq \mathcal{A} \setminus B$, if there exist nodes in the same layer $\mathsf{u}_1, \mathsf{u}_2 \in \mathcal{A}$ such that $\mathsf{u}_1 \in B$ and $\mathsf{u}_2 \in \mathcal{A} \setminus B$, then the rank of the transfer matrix across the cut $(B, \mathcal{A} \setminus B)$ is at least 2. Thus, if there exists a cut $(B, \mathcal{A} \setminus B)$ of value 1, then the cut must be of the form $B = (\cup_{l=0}^{t} \mathcal{L}_l) \cap \mathcal{A}$, for some $t \geq 0$. This tells us that if $\mathsf{u} \in \mathcal{A} \cap \mathcal{L}_{t+1}$, then $\mathcal{K}(\mathsf{u}) \supseteq \mathcal{A} \cap \mathcal{L}_{t+1}$, so that $\mathcal{K}(\mathsf{u})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut violating the definition of critical node v_1^*. Hence we complete the proof by contradiction.

7.1.3 Proof of Lemma 6.6

Suppose node v is omniscient, say $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ is a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut. Suppose v_1^* is not omniscient. As $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut, we have that either $\mathsf{v} \in \mathcal{K}(\mathsf{v}_1^*)$ or that v lies in a layer $\mathcal{L}_k$ with $k > k_1^*$. This follows from the definition of the critical node v_1^*. In the first case, we automatically have that v_1^* is omniscient. So, suppose v lies in layer $\mathcal{L}_k$ with $k > k_1^*$. Then, since v_1^* is not omniscient, there exists a path from s_2 to d_2 with a node $\mathsf{u}_{k_1^*}$ in layer $\mathcal{L}_{k_1^*}$ and a node u_k in layer $\mathcal{L}_k$ such that $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \mathcal{P}^{\mathsf{s}_2}(u_{k_1^*})$. Since node v is omniscient, we must have that $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}) = \mathcal{P}^{\mathsf{s}_2}(u_k)$. But this is impossible since u_k has an s_2-reachable parent from the path that does not lie in the cloud $\mathcal{C}_1$ which contains all the parents of v. This contradiction establishes that v_1^* must have been omniscient.

7.1.4 Proof of Lemma 6.7

First note that if we restrict attention to the induced subgraph $\mathcal{G}'$ obtained by deleting all nodes which can either not reach the set of nodes $\mathcal{U}$ or cannot be reached by at least one of s_1 and s_2, then the mincut value $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U})$ is preserved. Since each node can be reached by at least one of s_1 or s_2, we only have to delete nodes that cannot reach some node in $\mathcal{U}$.

Now, we are looking at a graph where the set of nodes in layer l is $\mathcal{U}_l$ for $0 \leq l < k$ and $\mathcal{U}$ for layer k.

Consider, for this graph, the set of all vertex bipartitions between $\{\mathsf{s}_1, \mathsf{s}_2\}$ and $\mathcal{U}$ which yield a transfer matrix of rank 1. All such bipartition cuts must be 'vertical', i.e. they are partitions of the form $(\mathcal{A}, \mathcal{A}^c)$ where $\mathcal{A} = \cup_{l=0}^{r} \mathcal{U}_l$ for some $l, 0 \leq l < k$. This is because any non-'vertical' cut yields a transfer matrix of rank at least 2. This establishes that the parents sets of all nodes in $\mathcal{U}_{l^*}$ are identical in this graph and so, also in the original graph because every node in the new graph $\mathcal{G}'$ has the same parent set as in the original graph. This concludes the proof of the lemma.

7.1.5 Proof of Lemma 6.9

We are in the scenario where $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)) = 1$. Consider $\mathcal{G}_{12}(\mathsf{w})$ for some node $\mathsf{w} \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ and suppose that there is no omniscient node in $\mathcal{G}_{12}(\mathsf{w})$. As $\mathcal{G}_{12}(\mathsf{w})$ has no paths from s_2 to d_1, by Lemma 6.3, we can achieve $(1,1)$ in $\mathcal{G}_{12}(\mathsf{w})$, with high probability, by all nodes except nodes in $\mathcal{P}(\mathsf{v}_2^*)$ performing RLC. Then, with high probability all nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ receive a non-trivial linear combination and each has a non-zero coefficient of symbol b sent by s_2. Take any such $(1,1)$ achieving scheme and any other node $\mathsf{w}' \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. Consider $\mathcal{G}_{12}(\mathsf{w}')$. Note w has no outgoing edges in $\mathcal{G}_{12}(\mathsf{w})$ and w' has no outgoing edges in $\mathcal{G}_{12}(\mathsf{w}')$. Let the reception of node w and w' in $\mathcal{G}_{12}(\mathsf{w})$ be $\beta_{\mathsf{w},\mathsf{s}_1} \cdot a + \beta_{\mathsf{w},\mathsf{s}_2} \cdot b$ and $\beta_{\mathsf{w}',\mathsf{s}_1} \cdot a + \beta_{\mathsf{w}',\mathsf{s}_2} \cdot b$ respectively, where $\beta_{\mathsf{w},\mathsf{s}_2}, \beta_{\mathsf{w}',\mathsf{s}_2} \neq 0$. Just make all nodes choose the same coefficients in $\mathcal{G}_{12}(\mathsf{w}')$ as in $\mathcal{G}_{12}(\mathsf{w})$ except for node w' which chooses $\alpha_{\mathsf{w}'}|_{\mathcal{G}_{12}(\mathsf{w}')} = \alpha_{\mathsf{w}} \cdot \frac{\beta_{\mathsf{w},\mathsf{s}_2}}{\beta_{\mathsf{w}',\mathsf{s}_2}}$. Then, the receptions of all nodes will be identical to those in $\mathcal{G}_{12}(\mathsf{w})$. This achieves $(1,1)$ in $\mathcal{G}_{12}(\mathsf{w}')$ and hence, there cannot be any omniscient node in this network either by the Omniscient node outer bound.

7.1.6 Proof of Lemma 6.13

Without loss of generality let $i = 1$. We shall prove this by induction on the layer index where u lies. Say $\mathsf{u} \in \mathcal{L}_k$. The node u receives $\beta_{\mathsf{u},\mathsf{s}_1} \cdot a + \beta_{\mathsf{u},\mathsf{s}_2} \cdot b$.

For $k = 1$, $\beta_{\mathsf{u},\mathsf{s}_1} = \alpha_{\mathsf{s}_1}\beta_{\mathsf{s}_1,\mathsf{s}_1} = \alpha_{\mathsf{s}_1}$. Since all predecessors of u are doing RLC, so does s_1 and hence α_{s_1} is chosen uniformly and randomly over $\mathbb{F}_{2^r}$. Therefore, $\Pr\{\beta_{\mathsf{u},\mathsf{s}_1} = 0\} = \Pr\{\alpha_{\mathsf{s}_1} = 0\} \to 0$ as $r \to \infty$.

Suppose for all nodes in $\mathcal{L}_l$, $l \geq 1$, that are reachable from s_1 the coefficient of user 1's symbol a is non-zero with high probability. Consider an s_1-reachable node in $\mathcal{L}_{l+1}$. We have

$$\beta_{\mathsf{u},\mathsf{s}_1} = \sum_{\mathsf{v} \in \mathcal{P}(\mathsf{u})} \alpha_{\mathsf{v}}\beta_{\mathsf{v},\mathsf{s}_1} = \sum_{\mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})} \alpha_{\mathsf{v}}\beta_{\mathsf{v},\mathsf{s}_1},$$

since for nodes that cannot be reached by s_1 the coefficient of a is always 0. Conditioned on a realization of $\{\beta_{\mathsf{v},\mathsf{s}_1} : \mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})\}$ where they are not all zero, $\beta_{\mathsf{u},\mathsf{s}_1}$ is uniformly distributed over $\mathbb{F}_{2^r}$ since $\{\alpha_{\mathsf{v}} | \mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})\}$ are chosen independently of one another and $\{\beta_{\mathsf{v},\mathsf{s}_1} : \mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})\}$, and uniformly over $\mathbb{F}_{2^r}$. Consequently,

$$\Pr\{\beta_{\mathsf{u},\mathsf{s}_1} = 0 | \{\beta_{\mathsf{v},\mathsf{s}_1} : \mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})\}\} \to 0$$

as $r \to \infty$, if $\{\beta_{\mathsf{v},\mathsf{s}_1} : \mathsf{v} \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{u})\}$ are not all zeros. By the induction assumption, the probability that they are all zeros also goes to zero as $r \to \infty$, and so we have $\Pr\{\beta_{\mathsf{u},\mathsf{s}_1} = 0\} \to 0$ as $r \to \infty$. This completes the proof by induction.

7.1.7 Proof of Lemma 6.14

Proof of Part (a)

Consider a super-sink d' with full access to the reception of all nodes in $\mathcal{U}$. Since $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}) = 2$, we have $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}') = 2$. Moreover, we can easily argue that both s_1 and s_2 can reach d' by contradiction, and hence $\mathrm{C}(\mathsf{s}_i; \mathsf{d}') = 1$, for $i = 1, 2$. Consider a multiple access flow problem with two sources $\mathsf{s}_1, \mathsf{s}_2$ and a single destination d'. The capacity region is the square region $\{(R_1, R_2) : R_1, R_2 \geq 0, R_1 \leq \mathrm{C}(\mathsf{s}_1; \mathsf{d}'), R_2 \leq \mathrm{C}(\mathsf{s}_2; \mathsf{d}'), R_1 + R_2 \leq \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}')\}$ and can be achieved via scalar random linear coding if the extension field size 2^r is sufficiently large [58]. Hence $(1, 1)$ can be achieved, and d' can decode both user's symbols and so can $\mathcal{U}$.

Proof of Part (b)

Fix the transmission from $\mathcal{P}(\mathsf{v}) \setminus \mathcal{U}$. We write the reception of v as

$$\underbrace{\sum_{\mathsf{u} \in \mathcal{U}} (\alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_1} \cdot a + \alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_2} \cdot b)}_{\text{To be determined}} + \underbrace{\sum_{\mathsf{u} \in \mathcal{P}(\mathsf{v}) \setminus \mathcal{U}} (\alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_1} \cdot a + \alpha_{\mathsf{u}} \beta_{\mathsf{u},\mathsf{s}_2} \cdot b)}_{\text{Given}}$$

From part (a) we know that $\mathcal{U}$ can collectively solve a and b with high probability, and hence it can construct any linear combination of a and b. Therefore, they can arrange their transmission by choosing the scaling coefficients α's carefully so that combined with the given part in v, the aggregate reception at v is the desired linear combination.

Proof of Part (c)

From part (a) we know the the subspace spanned by the received linear combinations of $\mathcal{U}$ has dimension 2 with high probability. The received linear combination of u spans an one-dimensional space with high probability. Note that $\mathcal{U} \setminus \{\mathsf{u}\} \neq \emptyset$.

Consider the subspace spanned by the received linear combination(s) of $\mathcal{U} \setminus \{\mathsf{u}\}$. This subspace is either has dimension 2 or has dimension 1 but not aligned with the reception of u. In the first case, after the nodes in $\mathcal{U} \setminus \{\mathsf{u}\}$ chose the scaling coefficients randomly, uniformly, and independently over $\mathbb{F}_{2^r}$, the resulting effective linear combination at v contributed by this part is uniformly distributed over the whole two-dimensional space. Hence it is not aligned with the reception of u with high probability. u can then choose its scaling coefficient properly so that any desired linear combination except those aligned with the reception of u can be formed at v. In the second case, it can be guaranteed that the resulting effective linear combination at v contributed by $\mathcal{U} \setminus \{\mathsf{u}\}$ is not aligned with the reception of u. Hence we arrive at the same conclusion as above.

Proof of Part (d)

This is a simple corollary of part (c). Since u is $\mathsf{s}_1\mathsf{s}_2$-reachable, with all its predecessors doing RLC it will receive a linear combination of a and b with non-zero coefficients for both symbols with high probability. Hence linear combinations consisting of purely a or b with high probability at v can be formed at v due to the conclusion in part (c).

7.1.8 Proof of Lemma 6.15

$\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}) = 2$. Note that all nodes can be reached by at least one of the source nodes. Fix node $\mathsf{u} \in \mathcal{U}$.

For sufficiently large block length N, if all nodes perform RLC with one symbol from each source, then by Lemma 6.13 and Lemma 6.14(a), we have the following with high probability:

- the subspace spanned by the received linear combination at u has dimension 1, and
- the subspace spanned by the received linear combinations at $\mathcal{U}$ has dimension 2.

Fix any choice of the coefficients so that the above hold. Pick any other node $\mathsf{w} \in \mathcal{U}$ such that the subspace spanned by the received linear combinations at u and w has dimension 2. Then, we must have $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}, \mathsf{w}) = 2$, or else u, w could not have received linearly independent linear combinations.

(Note: Lemma 6.15 is a purely graph-theoretic lemma. It is easier to prove it however using the random coding arguments in Lemma 6.13 and Lemma 6.14(a).)

7.1.9 Proof of Lemma 6.16

For all four cases, the direction "$\Leftarrow$" is quite obvious. Also note that when $k_1^* = k_2^* = k^*$ and there is no omniscient node in, the two clone-sets, $\mathcal{K}(\mathsf{v}_1^*)$ and $\mathcal{K}(\mathsf{v}_2^*)$, partition the whole layer $\mathcal{L}_{k^*}$. Therefore, it is sufficient to look at v_1^* and v_2^* only.

For the other direction "$\Rightarrow$", we shall prove the first and the third case, in which the superscript of the conditions is "(12)". To satisfy $\mathrm{T}_2^{(12)}$ and $\mathrm{T}_3^{(12)}$ (equivalently $\mathrm{P}_2^{(12)}$ and $\mathrm{P}_3^{(12)}$), we require $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$ as well as $\mathcal{K}_{\mathcal{G}_{12}}^{\mathsf{s}_1}(\mathsf{u}_{21})$ forms an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$. In generating $\mathcal{G}_{12}$, since there is only one node v_2^* (up to clones) in the same layer as v_1^*, the reorganization step will not involve any change in edges, as $M = 1$. There are two possible cases where $\mathcal{K}_{\mathcal{G}_{12}}^{\mathsf{s}_1}(\mathsf{u}_{21})$ forms an $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$: $\mathsf{u}_{21} \neq \mathsf{v}_2^*$, or $\mathsf{u}_{21} = \mathsf{v}_2^*$.

In the first case where $\mathsf{u}_{21} \neq \mathsf{v}_2^*$, we have $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)\right) = 1$. Hence all nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ are parents of v_2^*, and $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)\right) = 1$. Due to the fact that v_1^* is not omniscient, s_2 must be able to reach $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. On the other hand, nodes in $\mathcal{P}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ are all s_1-only-reachable and hence cannot belong to $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. Similarly nodes in $\mathcal{P}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ cannot be in $\mathcal{P}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. Therefore, we conclude that $\mathcal{U}_1$ is s_1-only-reachable, $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{W}\right) = \mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_2\right) = 1$. Then condition 3 and 4 in $\mathrm{T}^{(12)}$ ($\mathrm{P}^{(12)}$)

imply the rest of the conditions in the right-hand-side of the case $\mathrm{T}^{(12)}$ ($\mathrm{P}^{(12)}$). It is quite easy to see that in this case $\mathrm{P}^{(12)} \cap \mathrm{T}^{(21)} = \emptyset$.

In the second case where $\mathsf{u}_{21} = \mathsf{v}_2^*$, $\mathcal{P}(\mathsf{v}_1^*) \backslash \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ should be equal to the set of s_1-reachable parents of v_2^* in $\mathcal{G}_{12}$. If $\mathrm{T}_3^{(12)}$ is satisfied, then v_1^* and v_2^* will share the same s_1-reachable parents, contradicting the assumption that there is no omniscient node. If $\mathrm{P}_3^{(12)}$ is satisfied, then it must be the case that v_2^* has no parents in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. Therefore, $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) = \mathcal{U}_1$, all nodes in $\mathcal{W}$ are s_1-only-reachable, and all nodes in $\mathcal{U}_1$ are s_2-only-reachable. Then condition $\mathrm{P}_4^{(12)}$ implies that $\mathcal{K}^{\mathsf{s}_2}(\mathsf{w}_{12})$ forms an $(\mathsf{s}_2;\mathsf{d}_2)$-vertex-cut. Then it is easy to verify that $\mathrm{T}^{(21)}$ is satisfied. So considering $\mathrm{P}^{(12)} \setminus \mathrm{T}^{(21)}$, this pattern will not be included. Proof complete.

7.1.10 Proof of Claim 6.17

It is quite obvious that $(1,1)$ is achievable when $\mathcal{W} = \emptyset$. The assumption that there is no omniscient node combined with $\mathcal{U}_1^{\mathsf{s}_1} = \emptyset$ or $\mathcal{U}_2^{\mathsf{s}_2} = \emptyset$, implies the following three cases:

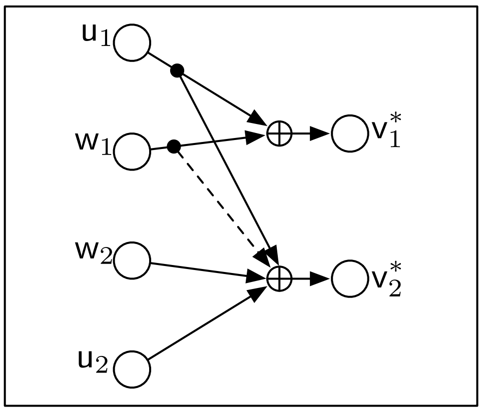
1) $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ and $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$: Pick $\mathsf{u}_1 \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*)$ and then find a node $\mathsf{w}_1 \in \mathcal{P}(\mathsf{v}_1^*)$ such that $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u}_1,\mathsf{w}_1) = 2$. Such a node exists by Lemma 6.15. Pick nodes $\mathsf{u}_2 \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ and $\mathsf{w}_2 \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*)$. Note that $\mathsf{u}_2, \mathsf{w}_2$ may be the same node.

(i) Suppose there exist u_2 and w_2 described as above such that $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u}_2,\mathsf{w}_2) = 2$: See Fig. 7.1(a) for an illustration. We first arrange the transmission of u_1 and w_1 so that only user 1's symbol appears at v_1^*. This can be done due to Lemma 6.14(a). Next we arrange the transmission of u_2 and w_2 so that the effect of user 1's symbol in the transmission of u_1 (and possibly w_2) at v_2^* can be neutralized, and user 2's symbol can appear cleanly. This can be done due to Lemma 6.14(b).

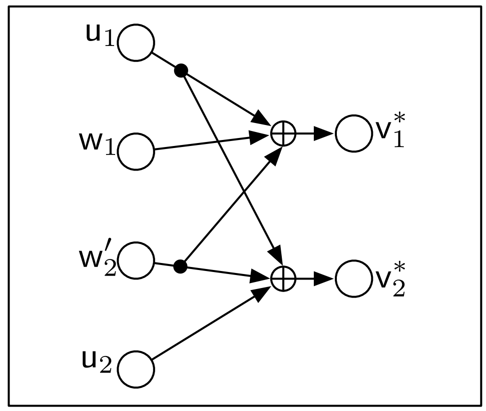
(ii) Suppose $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u}_2,\mathsf{w}_2) = 1$ for all u_2 and w_2 described as above: Then, we must have $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) = \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*)$, for if not, we can always find nodes $\mathsf{u}_2 \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ and $\mathsf{w}_2 \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*)$ such that $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u}_2,\mathsf{w}_2) = 2$. Thus, there must be a node $\mathsf{w}_2' \in \mathcal{P}(\mathsf{v}_1^*) \cap \mathcal{P}(\mathsf{v}_2^*)$ such that $\mathrm{C}(\mathsf{s}_1,\mathsf{s}_2;\mathsf{u}_2,\mathsf{w}_2') = 2$, by the definition of v_2^*. Note that w_2' may be the same node as u_1, w_1, or a clone of either one. Also note that now u_2 must be $\mathsf{s}_1\mathsf{s}_2$-reachable. See Fig. 7.1(b)(c) for an illustration. We further distinguish into two cases based on whether w_1 is a parent of v_2^* or not:

(1) If w_1 is not a parent of v_2^*, then it is s_2-only-reachable since $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$. We let u_1 and w_2' do RLC. Since u_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable, by Lemma 6.14(d), it can arrange its transmission so that v_2^* can decode s_2's symbol. We can then use w_1 to neutralize user 2's symbol in v_1^*'s reception if necessary. Since u_1 is s_1-reachable, v_1^* can obtain user 1's symbol cleanly after neutralization.

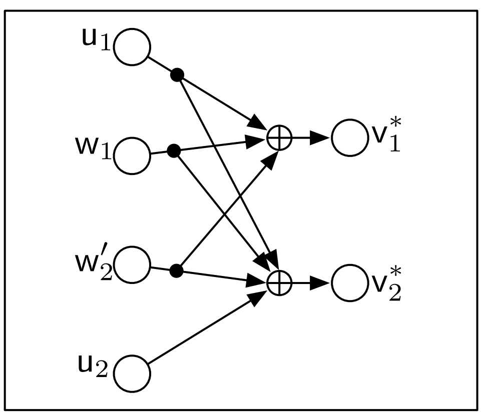
(2) If w_1 is a parent of v_2^*, then we first arrange the transmission of $\{\mathsf{u}_1,\mathsf{w}_1,\mathsf{w}_2'\}$ so that v_1^* can decode user 1's symbol. This can be done due to Lemma 6.14(a). Next, since



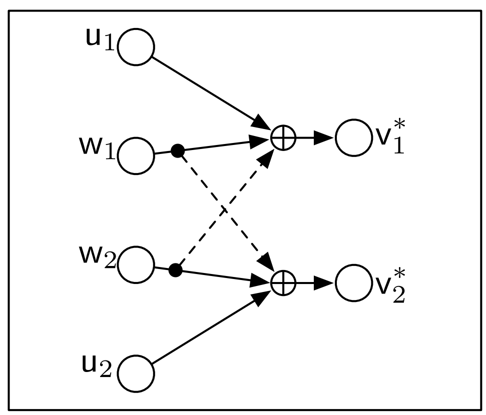
(a) Illustration of Case 1)(i). u_1 and w_2 are s_1-reachable and u_2 is s_2-reachable.



(b) Illustration of Case 1)(ii)(1). u_1 is s_1-reachable, u_2 is s_1s_2-reachable, and w_1 is s_1-only-reachable.



(c) Illustration of Case 1)(ii)(2). u_1 is s_1-reachable and u_2 is s_1s_2-reachable.



(d) Illustration of Case $\overline{A1} \cap \overline{B2}$. u_1 is s_1-reachable and u_2 is s_2-reachable.

Figure 7.1: Critical Nodes in the Same Layer

their aggregate at v_2^* has only user 1's symbol and u_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable, we can arrange the transmission of u_2 so that user 1's symbol is neutralized and only user 2's symbol is left.

2) $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ and $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset$: Similar to the previous case.

3) $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \subsetneq \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*)$ and $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$: Pick a node $\mathsf{u}_1 \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*)$ and a node $\mathsf{u}_2 \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*)$. By the definition of v_1^* and v_2^*, we shall be able to find $\mathsf{w}_1 \in \mathcal{P}(\mathsf{v}_1^*)$ and $\mathsf{w}_2 \in \mathcal{P}(\mathsf{v}_2^*)$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_1, \mathsf{w}_1) = \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_2, \mathsf{w}_2) = 2$. Note that $\mathsf{w}_1, \mathsf{w}_2$ may be the same node but $\mathsf{u}_1, \mathsf{u}_2$ are different nodes, although they may be clones. We shall show that $(1, 1)$ is achievable. First let $\{\mathsf{w}_1, \mathsf{w}_2\}$ do RLC. Then we can arrange the transmission of u_1 and u_2 such that v_1^* and v_2^* can obtain their desired symbols due to Lemma 6.14(c).

7.1.11 Proof of Claim 6.18

Note that

$$\begin{aligned}\neg(A \vee B) &= \neg(A1 \wedge A2) \wedge \neg(B1 \wedge B2) = (\neg A1 \vee \neg A2) \wedge (\neg B1 \vee \neg B2) \\ &= (\neg A1 \wedge \neg B1) \vee (\neg A1 \wedge \neg B2) \vee (\neg A2 \wedge \neg B1) \vee (\neg A2 \wedge \neg B2)\,.\end{aligned}$$

We distinguish into 4 cases.

1. $\neg A1 \wedge \neg B1$:

   In this case, there is a node in $\mathcal{P}_1^{\mathsf{s}_1}$ that is $\mathsf{s}_1\mathsf{s}_2$-reachable and there is another node in $\mathcal{P}_1^{\mathsf{s}_1}$ that is s_1-only-reachable. Hence $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_1^{\mathsf{s}_1}) = 2$. We can first arrange the transmission of $\mathcal{P}(\mathsf{v}_2^*)$ so that v_2^* can decode b. Since $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_1^{\mathsf{s}_1}) = 2$, we can arrange their transmission to form any linear combination of a and b; in particular, the one that combined with the transmission from $\mathcal{W}$ forms a at v_1^*. Hence $(1, 1)$ is achievable.

2. $\neg A1 \wedge \neg B2$:

   In this case there is a node $\mathsf{u}_1 \in \mathcal{P}_1^{\mathsf{s}_1}$ that is $\mathsf{s}_1\mathsf{s}_2$-reachable and there is a node $\mathsf{u}_2 \in \mathcal{P}_2^{\mathsf{s}_2}$ that is $\mathsf{s}_1\mathsf{s}_2$-reachable. Locate nodes $\mathsf{w}_1 \in \mathcal{P}(\mathsf{v}_1^*)$ and $\mathsf{w}_2 \in \mathcal{P}(\mathsf{v}_2^*)$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_1, \mathsf{w}_1) = \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_2, \mathsf{w}_2) = 2$. Note that $\mathsf{w}_1, \mathsf{w}_2$ may be the same node but $\mathsf{u}_1, \mathsf{u}_2$ will be different nodes although they may be clones. Then, let $\mathsf{w}_1, \mathsf{w}_2$ perform RLC while $\mathsf{u}_1, \mathsf{u}_2$ arrange their transmissions so that $\mathsf{v}_1^*, \mathsf{v}_2^*$ can decode their desired symbols. This can be done with high probability due to Lemma 6.14(d). See Fig. 7.1(d) for an illustration.

3. $\neg A2 \wedge \neg B1$:

   In this case there is a node in $\mathcal{P}_1^{\mathsf{s}_1}$ that is s_1-only-reachable and there is a node in $\mathcal{P}_2^{\mathsf{s}_2}$ that is s_2-only-reachable. Obviously $(1, 1)$ is achievable.

4. $\neg A2 \wedge \neg B2$:

   In this case, there is a node in $\mathcal{P}_2^{\mathsf{s}_2}$ that is s_2-only-reachable and there is another node in $\mathcal{P}_2^{\mathsf{s}_2}$ that is $\mathsf{s}_1\mathsf{s}_2$-reachable. Similar to the first case, $(1,1)$ is achievable.

Proof complete.

7.1.12 Proof of Lemma 6.19

We shall distinguish the condition $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ into two cases, (1) $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$, and (2) $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$.

$\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$

In this case, if $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \cap \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) = \emptyset$, then the special linear coding operation in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2)$ will not affect the coefficient of user 2's symbol b in the reception of u. Therefore the goal in the claim of this lemma can be met from $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) = 2$ and Lemma 6.14(b). Below we consider the case where $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \cap \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$.

If $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \neq \emptyset$, then we shall let the nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \cap \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ do RLC. Hence the parents of u_2 are all doing RLC. Since s_2 can reach u_2, the coefficient of b in the reception of u_2 is non-zero with high probability. Now we turn to v_1^*. As $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) = 2$ and all nodes other than $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2)$ are doing RLC, by Lemma 6.14(c) they can arrange their transmission so that v_1^* receives a linear combination consisting of a solely.

$\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2) \subsetneq \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$

In this case we let the nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2)$ do RLC. Hence the coefficient of b in the reception of u_2 is non-zero with high probability since all its predecessor are doing RLC. Then as $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathsf{v}_1^*)) = 2$ and all nodes other than $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{u}_2)$ are doing RLC, by Lemma 6.14(c) they can arrange their transmission so that v_1^* receives a linear combination consisting of a solely.

7.1.13 Proof of Lemma 6.21

The special coding operation performed by nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ is as follows: Nodes choose their coefficients independently and uniformly over the set of coefficients satisfying $\sum_{\mathsf{u}\in\mathcal{U}} \alpha_{\mathsf{u}}\beta_{\mathsf{u},\mathsf{s}_2} = 0$. Under this special coding, it is easy to show the first part of the assertion, namely, that each node receives a non-trivial linear combination. Because the reception of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ has full rank, the linear constraint leaves the sum $\sum_{\mathsf{u}\in\mathcal{U}} \alpha_{\mathsf{u}}\beta_{\mathsf{u},\mathsf{s}_1}$ non-zero with high probability. This allows us to argue that any s_1-reachable node receives a non-zero coefficient for the symbol a transmitted by source s_1 inspite of the special coding.

Find node $\mathsf{u} \in \mathcal{U}$ such that $\mathsf{u} \notin \mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*)$. and u is s_2-reachable. Such a node exists because v_1^* is not omniscient. Then, find node $\mathsf{w} \in \mathcal{U}$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}, \mathsf{w}) = 2$. If all nodes in

$\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ performed random linear coding, then u, w jointly can decode both symbols a and b with high probability.

Let $\mathcal{P}_1 := \mathcal{P}(\mathsf{u}) \setminus \mathcal{P}(\mathsf{w}), \mathcal{P}_{12} := \mathcal{P}(\mathsf{u}) \cap \mathcal{P}(\mathsf{w}), \mathcal{P}_2 := \mathcal{P}(\mathsf{w}) \setminus \mathcal{P}(\mathsf{u})$.

- As $\mathcal{P}(\mathsf{u}) \neq \emptyset$, we have $\mathcal{P}_1 \cup \mathcal{P}_{12} \neq \emptyset$.
- As $\mathcal{P}(\mathsf{w}) \neq \emptyset$, we have $\mathcal{P}_{12} \cup \mathcal{P}_2 \neq \emptyset$.
- As $\mathcal{P}(\mathsf{u}) \neq \mathcal{P}(\mathsf{w})$ (since $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}, \mathsf{w}) = 2$), we have $\mathcal{P}_1 \cup \mathcal{P}_2 \neq \emptyset$.

Note that the conditions on the sets $\mathcal{P}_1, \mathcal{P}_{12}, \mathcal{P}_2$ are symmetric.

Reception of node u:

$\left(\sum_{\mathsf{x}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_1}\right) \cdot a + \left(\sum_{\mathsf{x}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2}\right) \cdot b$

Reception of node w:

$\left(\sum_{\mathsf{x}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_1}\right) \cdot a + \left(\sum_{\mathsf{x}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2}\right) \cdot b$

Let $D := \begin{vmatrix} \sum_{\mathsf{x}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_1} & \sum_{\mathsf{x}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2} \\ \sum_{\mathsf{x}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_1} & \sum_{\mathsf{x}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2} \end{vmatrix}$. D is non-zero if and only if u, w can jointly decode both symbols a and b.

For two nodes x, y, denote the determinant $\begin{vmatrix} \beta_{\mathsf{x},\mathsf{s}_1} & \beta_{\mathsf{x},\mathsf{s}_2} \\ \beta_{\mathsf{y},\mathsf{s}_1} & \beta_{\mathsf{y},\mathsf{s}_2} \end{vmatrix}$ by $\beta(\mathsf{x}, \mathsf{y})$. Some algebra allows the determinant D to be expressed as:

$$D = \sum_{\mathsf{x}\in\mathcal{P}_1} \sum_{\mathsf{y}\in\mathcal{P}_{12}} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x}, \mathsf{y}) + \sum_{\mathsf{x}\in\mathcal{P}_{12}} \sum_{\mathsf{y}\in\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x}, \mathsf{y}) + \sum_{\mathsf{x}\in\mathcal{P}_2} \sum_{\mathsf{y}\in\mathcal{P}_1} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x}, \mathsf{y}). \tag{7.1}$$

Note that D is also symmetric in the sets $\mathcal{P}_1, \mathcal{P}_{12}, \mathcal{P}_2$. Let the special coding set $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ be denoted by $\mathcal{P}$. We are given that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}) = 2$. The constraint placed on the coding coefficients of nodes in $\mathcal{P}$ is $\sum_{\mathsf{x}\in\mathcal{P}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2} = 0$.

- Suppose $\mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2) \neq \emptyset$.

  Because we have with high probability, $\beta_{\mathsf{x},\mathsf{s}_2} \neq 0 \forall x \in \mathcal{P}$, we can view the special coding as all nodes in $\mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2$ performing random linear coding while nodes in $\mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2)$ performing restricted coding. In this case, parent nodes of u and w perform RLC and so, the claim is obviously true.

- Suppose $\mathcal{P} \subseteq \mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2$ and suppose there are two non-empty sets among $\mathcal{P} \cap \mathcal{P}_1, \mathcal{P} \cap \mathcal{P}_{12}, \mathcal{P} \cap \mathcal{P}_2$.

  Without loss of generality, assume $\mathcal{P} \cap \mathcal{P}_1 \neq \emptyset$. Fix $\mathsf{x}_0 \in \mathcal{P} \cap \mathcal{P}_1$. Find $\mathsf{x}_1 \in \mathcal{P}$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_0, \mathsf{x}_1) = 2$. If $\mathsf{x}_1 \in \mathcal{P}_{12}$ or $\mathsf{x}_1 \in \mathcal{P}_2$, then we have $x_0 \in \mathcal{P} \cap \mathcal{P}_1$, and $x_1 \in \mathcal{P} \cap \mathcal{P}_{12}$ or $x_1 \in \mathcal{P} \cap \mathcal{P}_2$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_0, \mathsf{x}_1) = 2$.

If $\mathsf{x}_1 \in \mathcal{P}_1$, then pick any node x_2 in the non-empty set $\mathcal{P} \cap (\mathcal{P}_{12} \cup \mathcal{P}_2)$. By submodularity, we have $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_0, \mathsf{x}_1, \mathsf{x}_2) + \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_2) \leq \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_0, \mathsf{x}_2) + \mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2)$. Since the two terms on the left are 2 and 1 respectively, at least one term on the right must be greater than 1 and thus, 2.

Thus, we can always find nodes $\mathsf{x}_0 \in \mathcal{P} \cap E, \mathsf{x}_1 \in \mathcal{P} \cap F$, where $(E, F) = (\mathcal{P}_1, \mathcal{P}_{12}), (\mathcal{P}_{12}, \mathcal{P}_2)$ or $(\mathcal{P}_2, \mathcal{P}_1)$. such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_0, \mathsf{x}_1) = 2$.

Suppose, without loss of generality, we have $\mathsf{x}_1 \in \mathcal{P} \cap \mathcal{P}_1, \mathsf{x}_2 \in \mathcal{P} \cap \mathcal{P}_{12}$ so that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2) = 2$. We set $\alpha_{\mathsf{x}_1} = \beta_{\mathsf{x}_1,\mathsf{s}_2}^{-1} \left(\sum_{\mathsf{x} \in \mathcal{P} \setminus \{\mathsf{x}_1\}} \alpha_{\mathsf{x}} \beta_{\mathsf{x},\mathsf{s}_2} \right)$.

Then, evaluating Equation (7.1) with this substitution for α_{x_1} gives us a polynomial in $(\alpha_{\mathsf{x}} : \mathsf{x} \in \mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2 \setminus \{\mathsf{x}_1\})$ with coefficients being rational functions in $(\beta_{\mathsf{x},\mathsf{s}_1}, \beta_{\mathsf{x},\mathsf{s}_2} : \mathsf{x} \in \mathcal{P}_1 \cup \mathcal{P}_{12} \cup \mathcal{P}_2 \setminus \{\mathsf{x}_1\})$ which are themselves polynomials in the coding coefficients from the past stages. This polynomial has a coefficient for $\alpha_{\mathsf{x}_2}^2$ only in the sum

$$\begin{aligned}
&\sum_{\mathsf{x} \in \mathcal{P}} \sum_{\mathsf{y} \in \mathcal{P}_{12} \cup \mathcal{P}_2} \alpha_{\mathsf{x}} \alpha_{\mathsf{y}} \beta(\mathsf{x}, \mathsf{y}) \\
&= \sum_{\mathsf{x} \in \mathcal{P} \setminus \{\mathsf{x}_1\}} \sum_{\mathsf{y} \in \mathcal{P}_{12} \cup \mathcal{P}_2} \alpha_{\mathsf{y}} \left[\alpha_{\mathsf{x}} \beta(\mathsf{x}, \mathsf{y}) + \alpha_{\mathsf{x}} \beta_{\mathsf{x}_1,\mathsf{s}_2}^{-1} \beta_{\mathsf{x},\mathsf{s}_2} \beta(\mathsf{x}_1, \mathsf{y}) \right] \\
&= \sum_{\mathsf{x} \in \mathcal{P} \setminus \{\mathsf{x}_1\}} \sum_{\mathsf{y} \in \mathcal{P}_{12} \cup \mathcal{P}_2} \alpha_{\mathsf{x}} \alpha_{\mathsf{y}} \frac{\beta_{\mathsf{y},\mathsf{s}_2}}{\beta_{\mathsf{x}_1,\mathsf{s}_2}} \beta(\mathsf{x}_1, \mathsf{x})
\end{aligned}$$

where the last equality follows from the identity $\beta_{\mathsf{x},\mathsf{s}_2} \beta(\mathsf{x}_1, \mathsf{y}) + \beta_{\mathsf{x}_1,\mathsf{s}_2} \beta(\mathsf{x}, \mathsf{y}) + \beta_{\mathsf{y},\mathsf{s}_2} \beta(\mathsf{x}_1, \mathsf{x}) = 0$.

Putting $\mathsf{x} = \mathsf{y} = \mathsf{x}_2$ gives the coefficient of $\alpha_{\mathsf{x}_2}^2$ to be $\frac{\beta_{\mathsf{x}_2,\mathsf{s}_2}}{\beta_{\mathsf{x}_1,\mathsf{s}_2}} \beta(\mathsf{x}_1, \mathsf{x}_2)$ which is not identically zero since each of $\beta_{\mathsf{x}_2,\mathsf{s}_2}, \beta_{\mathsf{x}_1,\mathsf{s}_2}, \beta(\mathsf{x}_1, \mathsf{x}_2)$ are not identically zero, the first two because $\mathsf{x}_1, \mathsf{x}_2$ lie in $\mathcal{P}$ and so are s_2-reachable and the third because $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2) = 2$.

Thus, D is not identically zero and hence, evaluates to a non-zero value with high probability.

- Finally, suppose $\mathcal{P} \subseteq \mathcal{P}_1$ or $\mathcal{P} \subseteq \mathcal{P}_{12}$ or $\mathcal{P} \subseteq \mathcal{P}_2$.
  - First, suppose $\mathcal{P} \subseteq \mathcal{P}_1$. Fix $\mathsf{x}_1 \in \mathcal{P}$. There exists $\mathsf{x}_2 \in \mathcal{P}$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2) = 2$. Force $\alpha_{\mathsf{x}_1} = \beta_{\mathsf{x}_1,\mathsf{s}_2}^{-1} \left(\sum_{\mathsf{x} \in \mathcal{P} \setminus \{\mathsf{x}_1\}} \alpha_{\mathsf{x}} \beta_{\mathsf{x},\mathsf{s}_2} \right)$.

$$\begin{aligned}
&D \\
&= \sum_{\mathsf{x} \in \mathcal{P}_1} \sum_{\mathsf{y} \in \mathcal{P}_{12} \cup \mathcal{P}_2} \alpha_{\mathsf{x}} \alpha_{\mathsf{y}} \beta(\mathsf{x}, \mathsf{y}) + \sum_{\mathsf{x} \in \mathcal{P}_{12}} \sum_{\mathsf{y} \in \mathcal{P}_2} \alpha_{\mathsf{x}} \alpha_{\mathsf{y}} \beta(\mathsf{x}, \mathsf{y})
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\mathsf{x}\in\mathcal{P}} \sum_{\mathsf{y}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) \\
&\quad + \sum_{\mathsf{x}\in\mathcal{P}_1\backslash\mathcal{P}} \sum_{\mathsf{y}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) + \sum_{\mathsf{x}\in\mathcal{P}_{12}} \sum_{\mathsf{y}\in\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) \\
&= \sum_{\mathsf{x}\in\mathcal{P}\backslash\{\mathsf{x}_1\}} \sum_{\mathsf{y}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\frac{\beta_{\mathsf{y},\mathsf{s}_2}}{\beta_{\mathsf{x}_1,\mathsf{s}_2}}\beta(\mathsf{x}_1,\mathsf{x}) \\
&\quad + \sum_{\mathsf{x}\in\mathcal{P}_1\backslash\mathcal{P}} \sum_{\mathsf{y}\in\mathcal{P}_{12}\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) + \sum_{\mathsf{x}\in\mathcal{P}_{12}} \sum_{\mathsf{y}\in\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y})
\end{aligned}
$$

As $\mathsf{u} \notin \mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, we have that some node $\mathsf{y}_0 \in \mathcal{P}_{12}$ is s_2-reachable. Then, $\beta_{\mathsf{y}_0,\mathsf{s}_2}$ is not identically zero and the coefficient of $\alpha_{\mathsf{x}_2}\alpha_{\mathsf{y}_0}$ is not identically zero.

– Suppose $\mathcal{P} \subseteq \mathcal{P}_{12}$. Then, fix $\mathsf{x}_1 \in \mathcal{P}$. There exists $\mathsf{x}_2 \in \mathcal{P}$ such that $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2\right) = 2$. Force $\alpha_{\mathsf{x}_1} = \beta_{\mathsf{x}_1,\mathsf{s}_2}^{-1}\left(\sum_{\mathsf{x}\in\mathcal{P}\backslash\{\mathsf{x}_1\}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2}\right)$.

$$
\begin{aligned}
D = &\sum_{\mathsf{x}\in\mathcal{P}\backslash\{\mathsf{x}_1\}} \sum_{\mathsf{y}\in\mathcal{P}_1\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\frac{\beta_{\mathsf{y},\mathsf{s}_2}}{\beta_{\mathsf{x}_1,\mathsf{s}_2}}\beta(\mathsf{x}_1,\mathsf{x}) + \sum_{\mathsf{x}\in\mathcal{P}_{12}\backslash\mathcal{P}} \sum_{\mathsf{y}\in\mathcal{P}_1\cup\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) \\
&+ \sum_{\mathsf{x}\in\mathcal{P}_1} \sum_{\mathsf{y}\in\mathcal{P}_2} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}).
\end{aligned}
$$

Again, since $\mathsf{u} \notin \mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, we have that some node $\mathsf{y}_0 \in \mathcal{P}_1$ is s_2-reachable. Then, $\beta_{\mathsf{y}_0,\mathsf{s}_2}$ is not identically zero and the coefficient of $\alpha_{\mathsf{x}_2}\alpha_{\mathsf{y}_0}$ is not identically zero.

– Now, suppose $\mathcal{P} \subseteq \mathcal{P}_2$. Then, fix $\mathsf{x}_1 \in \mathcal{P}$. There exists $\mathsf{x}_2 \in \mathcal{P}$ such that $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathsf{x}_1, \mathsf{x}_2\right) = 2$. Force $\alpha_{\mathsf{x}_1} = \beta_{\mathsf{x}_1,\mathsf{s}_2}^{-1}\left(\sum_{\mathsf{x}\in\mathcal{P}\backslash\{\mathsf{x}_1\}} \alpha_{\mathsf{x}}\beta_{\mathsf{x},\mathsf{s}_2}\right)$.

$$
\begin{aligned}
D = &\sum_{\mathsf{x}\in\mathcal{P}\backslash\{\mathsf{x}_1\}} \sum_{\mathsf{y}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\frac{\beta_{\mathsf{y},\mathsf{s}_2}}{\beta_{\mathsf{x}_1,\mathsf{s}_2}}\beta(\mathsf{x}_1,\mathsf{x}) + \sum_{\mathsf{x}\in\mathcal{P}_2\backslash\mathcal{P}} \sum_{\mathsf{y}\in\mathcal{P}_1\cup\mathcal{P}_{12}} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}) \\
&+ \sum_{\mathsf{x}\in\mathcal{P}_1} \sum_{\mathsf{y}\in\mathcal{P}_{12}} \alpha_{\mathsf{x}}\alpha_{\mathsf{y}}\beta(\mathsf{x},\mathsf{y}).
\end{aligned}
$$

Again, as u is s_2-reachable, we have that some node $\mathsf{y}_0 \in \mathcal{P}_1 \cup \mathcal{P}_{12}$ is s_2-reachable. Then, $\beta_{\mathsf{y}_0,\mathsf{s}_2}$ is not identically zero and the coefficient of $\alpha_{\mathsf{x}_2}\alpha_{\mathsf{y}_0}$ is not identically zero.

7.1.14 Proof of Lemma 6.22

For $A \subseteq \mathcal{U}$, define $f(A)$ as the rank of the $|A| \times 2$ matrix with rows given by $\begin{bmatrix}\lambda_{\mathsf{u}} & \mu_{\mathsf{u}}\end{bmatrix}$ for $\mathsf{u} \in A$, and define $g(A) = \mathrm{C}\left(A; \mathcal{V}\right)$. Then, $f(\cdot), g(\cdot)$ are rank functions of two matroids on the same ground set $\mathcal{U}$. The given conditions tell us that both these matroids have rank at least two and every singleton subset has rank 1 in both matroids. We will first show that there exist a two-element subset of $\mathcal{U}$ that has rank 2 in both matroids.

Find two elements $\mathsf{x}, \mathsf{y} \in \mathcal{U}$, such that $f(\{\mathsf{x},\mathsf{y}\}) = 2$. If $g(\{\mathsf{x},\mathsf{y}\}) = 2$, we have found the desired two-element subset. Else, we must have $g(\{\mathsf{x},\mathsf{y}\}) = 1$. Then, there exists an element $\mathsf{z} \in \mathcal{U}$ such that $g(\{\mathsf{x},\mathsf{z}\}) = 2$. If $f(\{\mathsf{x},\mathsf{z}\}) = 2$, we have the required 2-element subset. Else if we have $f(\{\mathsf{x},\mathsf{z}\}) = 1$, then by submodularity, we must have

$f(\{\mathsf{z}\}) + f(\{\mathsf{x},\mathsf{y},\mathsf{z}\}) \leq f(\{\mathsf{x},\mathsf{z}\}) + f(\{\mathsf{y},\mathsf{z}\})$

$g(\{\mathsf{y}\}) + g(\{\mathsf{x},\mathsf{y},\mathsf{z}\}) \leq g(\{\mathsf{x},\mathsf{y}\}) + g(\{\mathsf{y},\mathsf{z}\})$

These give $f(\{\mathsf{y},\mathsf{z}\}), g(\{\mathsf{y},\mathsf{z}\}) \geq 2$, and thus, $\{\mathsf{y},\mathsf{z}\}$ is the required subset of $\mathcal{U}$ that has rank 2 in both matroids.

Thus, we have two nodes $\mathsf{x}, \mathsf{y} \in \mathcal{U}$ such that $\begin{vmatrix} \lambda_\mathsf{x} & \mu_\mathsf{x} \\ \lambda_\mathsf{y} & \mu_\mathsf{y} \end{vmatrix} \neq 0$ and $\mathrm{C}(\mathsf{x},\mathsf{y};\mathcal{V}) = 2$.

Again, for $A \subseteq \mathcal{V}$, the function defined by $h(A) = \mathrm{C}(\mathsf{x},\mathsf{y};A)$ is the rank function of a matroid over ground set $\mathcal{V}$ that has rank two. Thus, there exist $\mathsf{u},\mathsf{w} \in \mathcal{V}$ such that $\mathrm{C}(\mathsf{x},\mathsf{y};\mathsf{u},\mathsf{w}) = 2$.

Thus, when all nodes perform RLC except nodes in $\mathcal{U} \setminus \{\mathsf{x},\mathsf{y}\}$ remain silent, we have that u,v can jointly recover both symbols a and b.

Now, if all nodes perform RLC, the a and b coefficients of the receptions of nodes u,v would be polynomials in the random coding coefficients with a determinant that is a polynomial that is not identically zero. QED.

7.1.15 Proof of Lemma 6.23

If v_2^* has a parent from user 1's cloud $\mathcal{C}_1$, then in either $\mathcal{G}_{12}$ or $\mathcal{G}'_{12}$ since this node in the cloud becomes s_1-only-reachable while v_2^* can be reached by s_2, v_2^* remains to be the critical node for user 2, ie., $\mathsf{Pmc}_{\mathcal{G}_{12}}(\mathsf{d}_2) = \mathsf{Pmc}_{\mathcal{G}'_{12}}(\mathsf{d}_2) = \mathsf{v}_2^*$. Since the s_1-clones of v_2^* form an $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$ but not in $\mathcal{G}$, the only possibility is that some parents of v_2^* are not in the cloud $\mathcal{C}_1$ and are dropped in $\mathcal{G}_{12}$. These nodes are descendants of $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$, which becomes s_1-only-reachable in $\mathcal{G}'_{12}$. Therefore in $\mathcal{G}'_{12}$, v_2^* has some s_1-reachable parents that is not in the cloud $\mathcal{C}_1$, and hence the s_1-clones of v_2^* do not form an $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}'_{12}$.

In the rest of the proof we deal with the case where v_2^* has no parents from user 1's cloud $\mathcal{C}_1$. Hence "s_1-clones of v_2^* form an $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}_{12}$" implies that $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$. We shall show that, for all possible $\mathcal{G}'_{12}$, either $\mathrm{C}_{\mathcal{G}'_{12}}(\mathsf{s}_1,\mathsf{s}_2;\mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)) = 2$, which implies that $\mathsf{Pmc}_{\mathcal{G}'_{12}}(\mathsf{d}_2) = \mathsf{v}_2^*$ and s_1-clones of v_2^* do not form an $(\mathsf{s}_1;\mathsf{d}_1)$-vertex-cut in $\mathcal{G}'_{12}$, or directly prove the statement.

Below a few notations are given before we proceed. $\mathcal{U} := \left\{\mathsf{u} \in \mathcal{L}_{k_1^*} : \mathsf{u} \text{ can reach } \mathsf{d}_2\right\}$. $\mathcal{U}|_{\mathcal{G}'_{12}}$ and $\mathcal{U}|_{\mathcal{G}_{12}}$ denote the nodes in the same layer as v_1^* that can reach d_2 in $\mathcal{G}'_{12}$ and $\mathcal{G}_{12}$ respectively. Recall that $\mathcal{R}$ is the set of nodes in $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*)$ that can reach one of the two destinations in $\mathcal{G}_{12}$. Define the following subsets of $\mathcal{U}$: (use short-hand notations $\mathcal{P}$ for $\mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*)$ and $\mathcal{S} := \mathcal{P} \setminus \mathcal{R}$)

$$\mathcal{U}_\mathcal{P} := \{\mathsf{u} : \mathcal{P}(\mathsf{u}) \supseteq \mathcal{P}\}, \; \mathcal{U}_\mathcal{Q} := \{\mathsf{u} : \mathcal{P}(\mathsf{u}) \cap \mathcal{P} = \emptyset\}$$

$$\mathcal{U}_{\mathcal{R}} := \{\mathsf{u} : \mathcal{P}(\mathsf{u}) \cap \mathcal{P} \neq \emptyset, \mathcal{P}(\mathsf{u}) \cap \mathcal{P} \subseteq \mathcal{R}\}$$
$$\mathcal{U}_{\mathcal{S}} := \{\mathsf{u} : \mathcal{S} \subseteq \mathcal{P}(\mathsf{u}) \cap \mathcal{P} \subsetneq \mathcal{P}\}$$

Note that these four sets form a partition of $\mathcal{U}$, and $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*) \cap \mathcal{U} \subseteq \mathcal{U}_{\mathcal{P}}$.

Let us consider the following two cases: 1) $\mathcal{R} \neq \emptyset$, and 2) $\mathcal{R} = \emptyset$. Note that when generating induced graphs $\mathcal{G}_{12}$ and $\mathcal{G}'_{12}$, some nodes may be dropped as they are no longer reachable from the sources. Consequently $\mathcal{U}|_{\mathcal{G}_{12}}$ and $\mathcal{U}|_{\mathcal{G}'_{12}}$ may be strictly contained in $\mathcal{U}$. In the following discussion, we shall further distinguish into these cases.

1) $\mathcal{R} \neq \emptyset$:

We shall show that in this case, $\mathrm{C}_{\mathcal{G}'_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)) = 2$.

(A) $\mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}$: Since no nodes are dropped in $\mathcal{U}$ when generating $\mathcal{G}'_{12}$, no nodes will be dropped in the later layers and $\mathrm{C}\left(\mathcal{U}; \mathcal{P}(\mathsf{v}_2^*)\right)$ remains the same in $\mathcal{G}$ and $\mathcal{G}'_{12}$. As $\mathrm{C}\left(\mathcal{U}; \mathcal{P}(\mathsf{v}_2^*)\right) \geq 2$ and all non-vertical cuts have cut-values at least 2, we only need to show that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}\right) = 2$.

(i) $\mathcal{U}|_{\mathcal{G}_{12}} = \mathcal{U}$:

In this case, since $\mathcal{U}|_{\mathcal{G}_{12}} = \mathcal{U}$, we have $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)\right) = \mathrm{C}\left(\mathcal{U}; \mathcal{P}(\mathsf{v}_2^*)\right) \geq 2$. Therefore, $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$ implies that $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$.

Suppose that $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}}\right) = 2$. Since nodes in $\mathcal{U}_{\mathcal{R}}$ will not be affected in generating $\mathcal{G}_{12}$, $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}}\right) = 2$. Hence $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}_{12}}\right) = 2$, contradicting the above fact. Besides, $\mathcal{U}_{\mathcal{R}} \neq \emptyset$. Therefore, $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}}\right) = 1$.

Find a node $\mathsf{u} \in \mathcal{U}$ such that $\mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right) = 2$. Below we show that this node $\mathsf{u} \in \mathcal{U}_{\mathcal{S}}$ by contradiction. Suppose $\mathsf{u} \in \mathcal{U}_{\mathcal{Q}}$. As the nodes in $\mathcal{U}_{\mathcal{Q}}$ will not be affected in generating $\mathcal{G}_{12}$, we have $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right) = \mathrm{C}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right) = 2$, contradicting the above fact that $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$. Next, suppose $\mathsf{u} \in \mathcal{U}_{\mathcal{P}}$. Let us first consider the min-cut value from $\{\mathsf{s}_1, \mathsf{s}_2\}$ to the collection of parents of $\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}$, denoted by $\mathcal{P}\left(\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right)$. It is 2 in $\mathcal{G}$. In $\mathcal{G}_{12}$, nodes in $\mathcal{S}$ are dropped, but nodes in $\mathcal{R}$ and nodes in $\mathcal{P}\left(\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right) \setminus \mathcal{P}$ are not. Therefore the min-cut value is again 2 since nodes in $\mathcal{R}$ receive the same linear combination as those in $\mathcal{P}$ under any RLC scheme. Second, it is clear that in $\mathcal{G}_{12}$, $\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}$ are not clones as u has no parents in $\mathcal{R}$. Hence, $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right) = 2$, again contradicting the above fact that $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$.

Hence, $\mathsf{u} \in \mathcal{U}_{\mathcal{S}}$ for all such u. We use the same argument as above to show that the min-cut value from $\{\mathsf{s}_1, \mathsf{s}_2\}$ to $\mathcal{P}\left(\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}\right)$ is again 2 in $\mathcal{G}_{12}$. Then $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$ implies that $\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}$ become clones in $\mathcal{G}_{12}$. Next, we turn to look at $\mathcal{G}'_{12}$. First, obviously $\mathcal{U}_{\mathcal{R}} \cup \{\mathsf{u}\}$ are not clones in $\mathcal{G}'_{12}$, as u has some parents in $\mathcal{S}$ which are not dropped in $\mathcal{G}'_{12}$. Second, $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathcal{U})\right) = 2$ as $\mathcal{R}$ becomes s_1-only-reachable in $\mathcal{G}'_{12}$ while s_2 can reach some other node in $\mathcal{P}_{\mathcal{G}'_{12}}(\mathcal{U})$. Combining the above two, we have shown that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}\right) = 2$.

(ii) $\mathcal{U}|_{\mathcal{G}_{12}} \neq \mathcal{U}$:

Some nodes in $\mathcal{U}$ are dropped in generating $\mathcal{G}_{12}$ and hence $\mathcal{U} \cap \mathcal{K}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$. The nodes in this intersection will be come s_1-only-reachable in $\mathcal{G}'_{12}$. Since some nodes in $\mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}$ can be reached by s_2 in $\mathcal{G}'_{12}$, we conclude that $\mathrm{C}_{\mathcal{G}'_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}) = 2$.

(B) $\mathcal{U}|_{\mathcal{G}'_{12}} \neq \mathcal{U}$: Some nodes in $\mathcal{U}$ are dropped in generating $\mathcal{G}'_{12}$, and the collection of these nodes is $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$. In the same layer as $\mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})$, consider the collection of predecessors of nodes in $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$. It must be equal to $\mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})$, otherwise nodes in $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$ would not be dropped in generating $\mathcal{G}'_{12}$. Hence, $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$ cannot be reached by $\mathcal{K}(\mathsf{w}_{12})$, and has no parents in $\mathcal{P}$. Therefore $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}} \subseteq \mathcal{U}_{\mathcal{Q}}$. Nodes in $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$ hence will not be dropped in $\mathcal{G}_{12}$ and $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}_{\mathcal{R}} \cup \mathcal{U}_{\mathcal{Q}}) = 2$, as nodes in $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$ can only be reached by $\mathcal{P}^{\mathsf{s}_2}(\mathsf{w}_{12})$ while nodes in $\mathcal{U}_{\mathcal{R}}$ can be reached by w_{12} and its s_1-only-reachable parents. Hence, the only possibility such that $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$ is that $\mathcal{U}|_{\mathcal{G}_{12}} \neq \mathcal{U}$ and $\mathrm{C}_{\mathcal{G}_{12}}(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$.

Note that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathcal{U}|_{\mathcal{G}'_{12}}; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = \mathrm{C}\left(\mathcal{U}|_{\mathcal{G}'_{12}}; \mathcal{P}(\mathsf{v}_2^*)\right)$ and $\mathrm{C}_{\mathcal{G}_{12}}\left(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)\right) = \mathrm{C}\left(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}(\mathsf{v}_2^*)\right)$. Also note that $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}_{12}}$ will not be dropped in $\mathcal{G}'_{12}$, and $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$ will not be dropped in $\mathcal{G}_{12}$. Hence $\mathcal{U}$ is partitioned by $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}}$, $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}_{12}}$, and $\mathcal{U}|_{\mathcal{G}_{12}} \cap \mathcal{U}|_{\mathcal{G}'_{12}}$. Furthermore, $\mathcal{U}_{\mathcal{R}} \subseteq \mathcal{U}|_{\mathcal{G}_{12}} \cap \mathcal{U}|_{\mathcal{G}'_{12}}$.

We first show that $\mathrm{C}\left(\mathcal{U}|_{\mathcal{G}'_{12}}; \mathcal{P}(\mathsf{v}_2^*)\right) \geq 2$. Define a function of the subsets of $\mathcal{U}$ by

$$f(\mathcal{A}) := \mathrm{C}\left(\mathcal{A}; \mathcal{P}(\mathsf{v}_2^*)\right), \ \mathcal{A} \subseteq \mathcal{U}.$$

Since f is submodular, we have

$$\begin{aligned} 3 &\overset{\text{(a)}}{\leq} f\left(\mathcal{U}|_{\mathcal{G}_{12}} \cap \mathcal{U}|_{\mathcal{G}'_{12}}\right) + f\left(\mathcal{U}\right) \leq f\left(\mathcal{U}|_{\mathcal{G}_{12}}\right) + f\left(\mathcal{U}|_{\mathcal{G}'_{12}}\right) \\ &\overset{\text{(b)}}{=} 1 + f\left(\mathcal{U}|_{\mathcal{G}'_{12}}\right) \implies f\left(\mathcal{U}|_{\mathcal{G}'_{12}}\right) \geq 2. \end{aligned}$$

(a) is due to $f(\mathcal{U}) \geq 2$ and $f\left(\mathcal{U}|_{\mathcal{G}_{12}} \cap \mathcal{U}|_{\mathcal{G}'_{12}}\right) \geq 1$ since $\mathcal{U}_{\mathcal{R}} \subseteq \mathcal{U}|_{\mathcal{G}_{12}} \cap \mathcal{U}|_{\mathcal{G}'_{12}}$. (b) is due to $f\left(\mathcal{U}|_{\mathcal{G}_{12}}\right) = 1$.

Next we show that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{U}|_{\mathcal{G}'_{12}}\right) = 2$. This is easy to see, since $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}_{12}}$ will become s_1-only-reachable in $\mathcal{G}'_{12}$ and some other nodes in $\mathcal{U}|_{\mathcal{G}'_{12}}$ can be reached by s_2.

Combining the above arguments, we conclude that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 2$.

2) $\mathcal{R} = \emptyset$:

In this case, $\mathcal{U} = \mathcal{U}_{\mathcal{P}} \cap \mathcal{U}_{\mathcal{Q}}$. For notational convenience, we denote $\mathcal{P}(\mathcal{U}) \setminus \mathcal{P}$ by $\mathcal{Q}$. Since in $\mathcal{G}_{12}$ the nodes in $\mathcal{P}$ no longer connects to $\mathcal{U}_{\mathcal{P}}$, $\mathcal{P}(\mathcal{U}|_{\mathcal{G}_{12}}) = \mathcal{Q}$. Note that if $\mathcal{U}_{\mathcal{P}} \neq \emptyset$, then $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathcal{U}|_{\mathcal{G}'_{12}})\right) = 2$ since $\mathcal{P} \subseteq \mathcal{P}(\mathcal{U})$, and nodes in $\mathcal{P}$ become s_1-only-reachable in $\mathcal{G}'_{12}$ while some other nodes in $\mathcal{P}(\mathcal{U}|_{\mathcal{G}'_{12}})$ can be reached by s_2.

$\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$ implies that: (A) $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{Q}) = 1$, (B) $\mathrm{C}_{\mathcal{G}_{12}}(\mathcal{Q}; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$, or (C) $\mathrm{C}_{\mathcal{G}_{12}}(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$. Below we discuss the three cases respectively.

(A) $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{Q}) = 1$: Suppose $\mathcal{U}_\mathcal{P} = \emptyset$, then $\mathcal{P}(\mathcal{U}) = \mathcal{Q}$, and $\mathrm{C}_{\mathcal{G}_{12}}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{Q}) = 1$ implies $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathcal{Q}) = 1$ contradicting the definition of v_2^*. Hence $\mathcal{U}_\mathcal{P} \neq \emptyset$, implying that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathcal{U}|_{\mathcal{G}'_{12}})\right) = 2$.

If $\mathcal{U}|_{\mathcal{G}'_{12}}$ are not clones in $\mathcal{G}'_{12}$ and $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathcal{U}|_{\mathcal{G}'_{12}}; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) \geq 2$, then $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 2$.

If $\mathcal{U}|_{\mathcal{G}'_{12}}$ become clones in $\mathcal{G}'_{12}$, then $\mathsf{u}'_{21} := \mathsf{Pmc}_{\mathcal{G}'_{12}}(\mathsf{d}_1)$ must belong to this new clone set. Its parent set is $\mathcal{P} \cup \mathcal{Q}|_{\mathcal{G}'_{12}}$, as some nodes in $\mathcal{Q}$ may be dropped in $\mathcal{G}'_{12}$. $\mathcal{P}$ becomes s_1-only-reachable in $\mathcal{G}'_{12}$, while v_1^* has some s_1-only-reachable parents not in $\mathcal{P}$. Hence, $\mathcal{K}_{\mathcal{G}'_{12}}\mathsf{v}_1^* \cap \mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}'_{12}}(\mathsf{u}'_{21}) = \emptyset$, and $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}'_{12}}(\mathsf{u}'_{21})$ does not form a $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}'_{12}$.

If $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathcal{U}|_{\mathcal{G}'_{12}}; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 1$, then $\mathcal{U}|_{\mathcal{G}'_{12}} \neq \mathcal{U}$, that is, some nodes in $\mathcal{U}_\mathcal{Q}$ are dropped in $\mathcal{G}'_{12}$. But no nodes in $\mathcal{U}_\mathcal{P}$ will be dropped. A node in $\mathcal{U}_\mathcal{P}$ is s_1-reachable in $\mathcal{G}'_{12}$, and is an predecessor of u'_{21}. This node cannot lie in $\mathcal{K}(\mathsf{v}_1^*)$, otherwise $\mathcal{Q}$ contains some s_1-only-reachable nodes implying that all nodes in $\mathcal{Q}$ are s_1-only-reachable, contradicting the fact that in $\mathcal{G}'_{12}$ some nodes in $\mathcal{Q}|_{\mathcal{G}'_{12}}$ can be reached by s_2. Hence this node is not an predecessor of any node in the cloud $\mathcal{C}_1$. In $\mathcal{G}'_{12}$, u'_{21} has a s_1-reachable parent whose predecessors include this node in $\mathcal{U}_\mathcal{P}$, and this parent is not in the cloud $\mathcal{C}_1$. Therefore, $\mathcal{K}^{\mathsf{s}_1}_{\mathcal{G}'_{12}}(\mathsf{u}'_{21})$ does not form a $(\mathsf{s}_1; \mathsf{d}_1)$-vertex-cut in $\mathcal{G}'_{12}$.

(B) $\mathrm{C}_{\mathcal{G}_{12}}(\mathcal{Q}; \mathcal{U}|_{\mathcal{G}_{12}}) = 1$: In this case, $\mathcal{U}|_{\mathcal{G}_{12}}$ become clones in $\mathcal{G}_{12}$. Suppose $\mathcal{U}_\mathcal{P} = \emptyset$. Then $\mathcal{U}|_{\mathcal{G}_{12}} = \mathcal{U}$, and $\mathcal{U}$ are clones in $\mathcal{G}$, contradicting the definition of v_2^*. Hence $\mathcal{U}_\mathcal{P} \neq \emptyset$, implying that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}(\mathcal{U}|_{\mathcal{G}'_{12}})\right) = 2$. Moreover, we see that $\mathcal{U}_\mathcal{Q}$ are clones in $\mathcal{G}$.

Suppose $\mathcal{U}|_{\mathcal{G}'_{12}} \neq \mathcal{U}$. We know that $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}} \subseteq \mathcal{U}_\mathcal{Q}$. Since $\mathcal{U}_\mathcal{Q}$ are clones in $\mathcal{G}$, we conclude that $\mathcal{U} \setminus \mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}_\mathcal{Q}$, implying that all nodes in $\mathcal{U}_\mathcal{Q}$ and $\mathcal{Q}$ will be dropped in $\mathcal{G}'_{12}$. This contradicts the fact that some nodes in $\mathcal{U}|_{\mathcal{G}'_{12}}$ can be reached by s_2. Therefore, $\mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}$.

In $\mathcal{G}'_{12}$, nodes in $\mathcal{U}_\mathcal{P}$ have parents in $\mathcal{P}$. Therefore obviously $\mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}$ are not clones in $\mathcal{G}'_{12}$. Combining the above discussions, we conclude that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 2$.

(C) $\mathrm{C}_{\mathcal{G}_{12}}(\mathcal{U}|_{\mathcal{G}_{12}}; \mathcal{P}_{\mathcal{G}_{12}}(\mathsf{v}_2^*)) = 1$: In this case, we must have $\mathcal{U}_{\mathcal{G}_{12}} \neq \mathcal{U}$. If $\mathcal{U}|_{\mathcal{G}'_{12}} = \mathcal{U}$, we use the same argument in Case 1)(A)(ii) to show that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 2$. If If $\mathcal{U}|_{\mathcal{G}'_{12}} \neq \mathcal{U}$, we use the same argument in Case 1)(B)(ii) to show that $\mathrm{C}_{\mathcal{G}'_{12}}\left(\mathsf{s}_1, \mathsf{s}_2; \mathcal{P}_{\mathcal{G}'_{12}}(\mathsf{v}_2^*)\right) = 2$.

Proof of the claim is now complete.

7.2 $(1/2, 1)$-Achievability in Case A when $k_1^* = k_2^* = k^*$

We first state a useful lemma.

Lemma 7.1. *Let* $p(\alpha_1, \alpha_2, \ldots, \alpha_n), q(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}_2[\alpha_1, \alpha_2, \ldots, \alpha_n]$ *such that* p, q *are not identically equal to zero or to each other. If* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *are chosen independently and uniformly over* $\mathbb{F}_{2^k}$*, then*

- $q(\alpha_1, \alpha_2, \ldots, \alpha_n) \neq 0$ *with probability at least* $1 - O(\frac{1}{2^k})$*, so the rational function* $\frac{p}{q}$ *is well-defined with high probability,*
- *and* $P(\frac{p}{q} = \gamma) = O(\frac{1}{2^k})$ *for all* $\gamma \in \mathbb{F}_{2^k}$.

Proof. We use a standard result from finite fields which states that if a multivariate polynomial g in n variables over finite field $\mathbb{F}$ with degree in each variable at most d, is evaluated at an argument chosen uniformly over the set of possible arguments, then it yields zero with probability at most $\frac{nd}{|\mathbb{F}|}$, provided of course that the polynomial is not identically zero.

This proves the first item in the lemma with $g = q$ and the second item in the lemma for the case $\gamma = 0$ using $g = p$.

For $\gamma = 1$, we use the fact that $p - q$ is not identically zero to get $P(\frac{p}{q} = 1) = O(\frac{1}{2^k})$.

For any other $\gamma \in \mathbb{F}_{2^k}$, we notice that $p - \gamma q$ cannot possibly be identically zero unless both p and q are identically zero. This is because p, q have coefficients from $\mathbb{F}_2$ while $\gamma \neq 0, 1$. This establishes that $p - \gamma q$ evaluates to zero with probability atmost $O(\frac{1}{2^k})$. □

We start the proof of $(1/2, 1)$-achievability below.

Here, we have

- $\mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset, \mathcal{P}^{\mathsf{s}_2}(v_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$,
- $\mathsf{u}_1 \in \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_1^*) \setminus \mathcal{P}^{\mathsf{s}_1}(\mathsf{v}_2^*) \neq \emptyset$ and $\mathsf{u}_2 \in \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_2^*) \setminus \mathcal{P}^{\mathsf{s}_2}(\mathsf{v}_1^*) \neq \emptyset$,
- u_1 is s_1-only-reachable and u_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable,
- $\mathsf{w}_2 \in \mathcal{P}(\mathsf{v}_2^*)$ such that $\mathrm{C}(\mathsf{s}_1, \mathsf{s}_2; \mathsf{u}_2, \mathsf{w}_2) = 2$, and w_2 is a parent of v_1^*, and w_2 is s_2-reachable.

We will use RLC for the transmission of all nodes in layers 0 through $k^* - 2$. The RLC is performed without mixing across the time steps. In the first time step, s_1 transmits the symbol a while s_2 transmits the symbol b_1. In the second time step, s_1 transmits symbol a while s_2 transmits the symbol b_2.

Suppose now that w_2 is $\mathsf{s}_1\mathsf{s}_2$-reachable.

Consider the scheme where w_2 and u_2 both zero-force user 1's symbol a. u_1 and u_2 transmit in the first time slot, thus causing no interference at v_1^* and v_2^*. w_2 transmits in the second time slot.

We have $\beta^{(1)}_{\mathsf{u}_1,s_2} = \beta^{(1)}_{\mathsf{u}_1,s_2} = 0$, while $\beta^{(1)}_{\mathsf{u}_1,\mathsf{s}_1} \neq 0$ with high probability from Lemma 6.13. Thus u_1 can decode s_1's symbol a with high probability.

Now, the receptions of u_2 in the two time slots are $\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1}\cdot a+\beta^{(1)}_{\mathsf{u}_2,s_2}\cdot b_1$ and $\beta^{(2)}_{\mathsf{u}_2,\mathsf{s}_1}\cdot a+\beta^{(2)}_{\mathsf{u}_2,s_2}\cdot b_2$ respectively. Similarly, the receptions of w_2 are $\beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1}\cdot a+\beta^{(1)}_{\mathsf{w}_2,s_2}\cdot b_1$ and $\beta^{(2)}_{\mathsf{w}_2,\mathsf{s}_1}\cdot a+\beta^{(2)}_{\mathsf{w}_2,s_2}\cdot b_2$. Note that the coefficients of these symbols are all non-zero with high probability from Lemma 6.13.

The zero-forcing yields:

- Transmission of u_2 : $\beta^{(1)}_{\mathsf{u}_2,s_2}\beta^{(2)}_{\mathsf{u}_2,\mathsf{s}_1}\cdot b_1 - \beta^{(2)}_{\mathsf{u}_2,s_2}\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1}\cdot b_2$
- Transmission of w_2 : $\beta^{(1)}_{\mathsf{w}_2,s_2}\beta^{(2)}_{\mathsf{w}_2,\mathsf{s}_1}\cdot b_1 - \beta^{(2)}_{\mathsf{w}_2,s_2}\beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1}\cdot b_2$

To show that v_2^* can decode, we only need to show that the determinant $\begin{vmatrix}\beta^{(1)}_{\mathsf{u}_2,s_2}\beta^{(2)}_{\mathsf{u}_2,\mathsf{s}_1} & -\beta^{(2)}_{\mathsf{u}_2,s_2}\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1}\\ \beta^{(1)}_{\mathsf{w}_2,s_2}\beta^{(2)}_{\mathsf{w}_2,\mathsf{s}_1} & -\beta^{(2)}_{\mathsf{w}_2,s_2}\beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1}\end{vmatrix}$ is non-zero, ie that $\beta^{(1)}_{\mathsf{u}_2,s_2}\beta^{(2)}_{\mathsf{u}_2,\mathsf{s}_1}\beta^{(2)}_{\mathsf{w}_2,s_2}\beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1} \neq \beta^{(2)}_{\mathsf{u}_2,s_2}\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1}\beta^{(1)}_{\mathsf{w}_2,s_2}\beta^{(2)}_{\mathsf{w}_2,\mathsf{s}_1}$ or

$$\frac{\beta^{(1)}_{\mathsf{u}_2,s_2}\beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1}}{\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1}\beta^{(1)}_{\mathsf{w}_2,s_2}} \neq \frac{\beta^{(2)}_{\mathsf{u}_2,s_2}\beta^{(2)}_{\mathsf{w}_2,\mathsf{s}_1}}{\beta^{(2)}_{\mathsf{u}_2,\mathsf{s}_1}\beta^{(2)}_{\mathsf{w}_2,s_2}}$$

Note that the coefficients with 1 superscript are independent of the coefficients with 2 superscript. So, LHS and RHS are two independent and identically distributed random variables taking values in $\mathbb{F}_{2^r}$.

By Lemma 6.14, we have that the determinant $\begin{vmatrix}\beta^{(1)}_{\mathsf{u}_2,\mathsf{s}_1} & \beta^{(1)}_{\mathsf{u}_2,s_2}\\ \beta^{(1)}_{\mathsf{w}_2,\mathsf{s}_1} & \beta^{(1)}_{\mathsf{w}_2,s_2}\end{vmatrix} \neq 0$ with high probability. So, the above random variable is not equal to 1 with high probability.

Now, we note that the random variable is a ratio of two polynomials with coefficients from $\mathbb{F}_2$, a ratio that is not identically 1. The equality stating that the ratio equals $\gamma \in \mathbb{F}_{2^r}, \gamma \neq 0, 1$ is an equality stating that a polynomial not identically zero evaluates to 0. If all coefficients are chosen indpendently and uniformly at random, this polynomial evaluates to 0 with probability $O\left(\frac{1}{|\mathbb{F}_{2^r}|}\right)$. Thus, the random variable does not concentrate on any given value $\gamma \in \mathbb{F}_{2^r}$ and so, two independent and identically distributed copies of the random variable are unequal with high probability.

Suppose that w_2 is s_2-only-reachable. Then, $\mathsf{u}_1, \mathsf{u}_2$ transmit in the first time slot with u_2 zero-forcing user 1's symbol a. In the second time slot, w_2 which can recover both b_1 and b_2 with high probability, provides a linearly independent signal to u_2's transmission.

7.3 Formal Proofs of Outer Bounds

7.3.1 Proof of the Omniscient Bound

Since $\mathcal{K}(\mathsf{v})$ is a $(\mathsf{s}_1, \mathsf{s}_2; \mathsf{d}_1)$-vertex-cut, the received signal at d_1, Y_{d_1} is a function of Y_{v}. On the other hand, since $\mathcal{K}^{\mathsf{s}_2}(\mathsf{v})$ is a $(\mathsf{s}_2; \mathsf{d}_2)$-vertex-cut, we have that $Y^N_{\mathsf{d}_2}$ is a function of $X^N_{\mathsf{s}_1}$ and Y^N_{v}. Hence we have the Markov chains

$$X^N_{\mathsf{s}_1} \leftrightarrow Y^N_{\mathsf{v}} \leftrightarrow Y^N_{\mathsf{d}_1} \tag{7.2}$$

$$X_{\mathsf{s}_2}^N \leftrightarrow \left(Y_{\mathsf{v}}^N, X_{\mathsf{s}_1}^N\right) \leftrightarrow Y_{\mathsf{d}_1}^N \tag{7.3}$$

By Fano's inequality and the data processing inequality, we have for any scheme of block length N,

$$\begin{aligned}
&N\left(R_1+R_2-\epsilon_N\right)\\
&\le I\left(X_{\mathsf{s}_1}^N;Y_{\mathsf{d}_1}^N\right)+I\left(X_{\mathsf{s}_2}^N;Y_{\mathsf{d}_2}^N\right)\\
&\le I\left(X_{\mathsf{s}_1}^N;Y_{\mathsf{v}}^N\right)+I\left(X_{\mathsf{s}_2}^N;Y_{\mathsf{v}}^N,X_{\mathsf{s}_1}^N\right)\quad\text{(from (7.2) and (7.3))}\\
&\le I\left(X_{\mathsf{s}_1}^N;Y_{\mathsf{v}}^N\right)+I\left(X_{\mathsf{s}_2}^N;Y_{\mathsf{v}}^N|X_{\mathsf{s}_1}^N\right)\\
&=H\left(Y_{\mathsf{v}}^N\right)-H\left(Y_{\mathsf{v}}^N|X_{\mathsf{s}_1}^N\right)+H\left(Y_{\mathsf{v}}^N|X_{\mathsf{s}_1}^N\right)\\
&=H\left(Y_{\mathsf{v}}^N\right)\le N,
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. Hence $R_1 + R_2 \le 1$.

7.3.2 Proof of Claim 6.25

Proof. If (R_1, R_2) is achievable, from data processing inequality and Fano's inequality, we have

$$\begin{aligned}
&N\left(2R_1+R_2-\epsilon_N\right)\\
&\le I\left(X_{\mathsf{s}_1}^N;Y_{\mathsf{d}_1}^N\right)+I\left(X_{\mathsf{s}_1}^N;Y_{\mathsf{d}_1}^N\right)+I\left(X_{\mathsf{s}_2}^N;Y_{\mathsf{d}_2}^N\right)\\
&\overset{(a)}{\le} I\left(X_{\mathsf{s}_1}^N;Z_{21}^N,X_{\mathsf{s}_2}^N\right)+I\left(X_{\mathsf{s}_1}^N;Z_1^N\right)+I\left(X_{\mathsf{s}_2}^N;Z_{21}^N,Z_{22}^N\right)\\
&\overset{(b)}{=} I\left(X_{\mathsf{s}_1}^N;Z_{21}^N|X_{\mathsf{s}_2}^N\right)+I\left(X_{\mathsf{s}_1}^N;Z_1^N\right)+I\left(X_{\mathsf{s}_2}^N;Z_{21}^N,Z_{22}^N\right)\\
&=H\left(Z_{21}^N|X_{\mathsf{s}_2}^N\right)+H\left(Z_1^N\right)-H\left(Z_1^N|X_{\mathsf{s}_1}^N\right)+H\left(Z_{21}^N,Z_{22}^N\right)\\
&\quad-H\left(Z_{21}^N,Z_{22}^N|X_{\mathsf{s}_2}^N\right)\\
&\overset{(c)}{=} H\left(Z_1^N\right)+H\left(Z_{21}^N|X_{\mathsf{s}_2}^N\right)-H\left(Z_{21}^N|X_{\mathsf{s}_2}^N\right)+H\left(Z_{21}^N,Z_{22}^N\right)\\
&\quad-H\left(Z_1^N|X_{\mathsf{s}_1}^N\right)\\
&\overset{(d)}{\le} H\left(Z_1^N\right)+H\left(Z_{21}^N,Z_{22}^N\right)-H\left(Z_{22}^N\right)\\
&=H\left(Z_1^N\right)+H\left(Z_{21}^N|Z_{22}^N\right)\overset{(e)}{\le} 2N
\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. (a) is due to condition 2) and 3). (b) is due to the fact that $X_{\mathsf{s}_1}^N$ and $X_{\mathsf{s}_2}^N$ are independent. (c) is due to condition 5) and rearranging terms. (d) is due to condition 4). (e) is due to condition 1). □

7.3.3 Proof of Claim 6.26

Proof. If (R_1, R_2) is achievable, from data processing inequality and Fano's inequality, we have

$$
\begin{aligned}
&N\left(2R_1 + R_2 - \epsilon_{1,N}\right) \\
&\le I\left(X_{\mathsf{s}_1}^N; Y_{\mathsf{d}_1}^N\right) + I\left(X_{\mathsf{s}_1}^N; Y_{\mathsf{d}_1}^N\right) + I\left(X_{\mathsf{s}_2}^N; Y_{\mathsf{d}_2}^N\right) \\
&\overset{(a)}{\le} I\left(X_{\mathsf{s}_1}^N; Z_{21}^N, Z_{22}^N, X_{\mathsf{s}_2}^N\right) + I\left(X_{\mathsf{s}_1}^N; Z_{11}^N\right) \\
&\quad + I\left(X_{\mathsf{s}_2}^N; Z_{21}^N, Z_{22}^N\right) \\
&\overset{(b)}{=} I\left(X_{\mathsf{s}_1}^N; Z_{21}^N, Z_{22}^N | X_{\mathsf{s}_2}^N\right) + I\left(X_{\mathsf{s}_1}^N; Z_{11}^N\right) \\
&\quad + I\left(X_{\mathsf{s}_2}^N; Z_{21}^N, Z_{22}^N\right) \\
&= H\left(Z_{21}^N, Z_{22}^N | X_{\mathsf{s}_2}^N\right) + H\left(Z_{11}^N\right) - H\left(Z_{11}^N | X_{\mathsf{s}_1}^N\right) \\
&\quad + H\left(Z_{21}^N, Z_{22}^N\right) - H\left(Z_{21}^N, Z_{22}^N | X_{\mathsf{s}_2}^N\right) \\
&\overset{(c)}{\le} H\left(Z_{11}^N\right) - H\left(Z_{22}^N | X_{\mathsf{s}_1}^N\right) + H\left(Z_{22}^N\right) + H\left(Z_{21}^N | Z_{22}^N\right) \\
&\overset{(d)}{=} H\left(Z_{11}^N\right) + H\left(Z_{21}^N | Z_{22}^N\right) + I\left(X_{\mathsf{s}_1}^N; Z_{22}^N\right) \\
&\overset{(e)}{\le} 2N + I\left(X_{\mathsf{s}_1}^N; Z_{12}^N\right),
\end{aligned}
$$

where $\epsilon_{1,N} \to 0$ as $N \to \infty$. (a) is due to condition 2) and 3). (b) is due to the fact that $X_{\mathsf{s}_1}^N$ and $X_{\mathsf{s}_2}^N$ are independent. (c) is due to cancellation of terms and condition 4). (d) is due to $I\left(X_{\mathsf{s}_1}^N; Z_{22}^N\right) = H\left(Z_{22}^N\right) - H\left(Z_{22}^N | X_{\mathsf{s}_1}^N\right)$. (e) is due to condition 1) and 5).

We see that we cannot upper bound $2R_1 + R_2$ by 2 in this case. On the other hand,

$$
\begin{aligned}
&N\left(R_2 - \epsilon_{2,N}\right) \le I\left(X_{\mathsf{s}_2}^N; Y_{\mathsf{d}_2}^N\right) \\
&\overset{(a)}{\le} I\left(X_{\mathsf{s}_2}^N; Z_{12}^N, X_{\mathsf{s}_1}^N\right) \overset{(b)}{=} I\left(X_{\mathsf{s}_2}^N; Z_{12}^N | X_{\mathsf{s}_1}^N\right) = H\left(Z_{12}^N | X_{\mathsf{s}_1}^N\right).
\end{aligned}
$$

where $\epsilon_{2,N} \to 0$ as $N \to \infty$. (a) is due to condition (3). (b) is due to the fact that $X_{\mathsf{s}_1}^N$ and $X_{\mathsf{s}_2}^N$ are independent.

Combining the above two, we have

$$
\begin{aligned}
&N\left(2R_1 + 2R_2 - \epsilon_N\right) \\
&\le 2N + I\left(X_{\mathsf{s}_1}^N; Z_{12}^N\right) + H\left(Z_{12}^N | X_{\mathsf{s}_1}^N\right) = 2N + H\left(Z_{12}^N\right) \\
&\overset{(a)}{\le} 3N,
\end{aligned}
$$

where $\epsilon_N = \epsilon_{1,N} + \epsilon_{2,N} \to 0$ as $N \to \infty$. (a) is due to condition 1). Proof complete. □

Chapter 8

Conclusions and Future Work

In Part I and Part II, we characterized both qualitatively and quantitatively how limited cooperation between transmitting terminals or receiving terminals helps mitigate interference in the canonical two-user Gaussian interference channel. From the fundamental perspective, the study sheds light on how limited cooperation can be used for better interference management in wireless networks. It also points out the potential impact on wireless system design. One important insight gained from the study is that, in different regimes of interference, the gain from limited cooperation can be quite different. This leads to a broader optimization framework taking the resource used for cooperation into account. Although the study is on the additive white Gaussian noise channel model with orthogonal cooperation, which is a fairly reasonable model for cellular systems with backhaul cooperation, we believe that the same principles can be readily applied to fading wireless channels as well as various cooperation scenarios, such as in-band cooperation in wireless ad-hoc networks.

Nevertheless, from the practical point of view, it is still quite an open question regarding how much cooperation gain one can extract in a deployed wireless system. Various efforts have been pushed towards settling this question and a comprehensive survey can be found in [59]. Along this direction, one of the ongoing research topics is to pursue practical coding and system design for cooperative wireless systems. Some partial results can be found in our work [60] and the references therein.

In Part III, we investigated how intermediate relay nodes help resolve interference in delivering information from two sources to their respective destinations in multi-hop wireless networks. This belongs to a broader class of problem, namely, the two unicast infomation flow over wireless networks. Given that even the understanding in the two-unicast wired networks is limited, we focused on a special class of layered linear deterministic networks without any other restrictions on the model except that the min-cut between each source-destination pair is constrained to be 1. We completely characterized the capacity region of this class of two-unicast networks, which provides an analogous result to that of the wired networks [52]. However, for general two-unicast linear deterministic networks, the characterization of the capacity region remains open, as it is also open for two-unicast wired networks. Hence, the connection of the result in Part III to the Gaussian two-unicast networks is yet to be explored. To make progress, we believe that a better understanding of two-unicast wired networks is necessary. On the other hand, extensions of the result to networks with more than two source-destination pairs also seem non-trivial, and more work has to be done towards that direction.

Bibliography

[1] Claude E. Shannon. Two-way communication channels. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1:611–644, 1961.

[2] H. Sato. The capacity of the gaussian interference channel under strong interference. *IEEE Transactions on Information Theory*, 27(6):786–788, November 1981.

[3] Thomas M. Cover and Abbas A. El Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5):572–584, September 1979.

[4] Te Sun Han and Kingo Kobayashi. A new achievable rate region for the interference channel. *IEEE Transactions on Information Theory*, 27(1):49–60, January 1981.

[5] Raul Etkin, David N. C. Tse, and Hua Wang. Gaussian interference channel capacity to within one bit. *IEEE Transactions on Information Theory*, 54(12):5534–5562, December 2008.

[6] Abolfazl Seyed Motahari and Amir Keyvan Khandani. Capacity bounds for the gaussian interference channel. *IEEE Transactions on Information Theory*, 55(2):620 – 643, February 2009.

[7] Xiaohu Shang, Gerhard Kramer, and Biao Chen. A new outer bound and the noisy-interference sum–rate capacity for gaussian interference channels. *IEEE Transactions on Information Theory*, 55(2):689 – 699, February 2009.

[8] V. Sreekanth Annapureddy and Venugopal 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*, 55(7):3032–3050, July 2009.

[9] Amir Salman Avestimehr, Suhas N. Diggavi, and David N. C. Tse. Wireless network information flow: A deterministic approach. *IEEE Transactions on Information Theory*, 57(4):1872–1905, April 2011.

[10] Guy Bresler and David N. C. Tse. The two-user gaussian interference channel: a deterministic view. *European Transactions on Telecommunications*, 19(4):333 – 354, June 2008.

[11] Edward C. van der Meulen. Three-terminal communication channels. *Advances in Applied Probability*, 3(1):120–154, September 1971.

[12] Gerhard Kramer, Michael Gastpar, and Piyush Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, September 2005.

[13] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404, 1956.

[14] Abbas A. El Gamal and Max Henrique M. Costa. The capacity region of a class of determinisitic interference channels. *IEEE Transactions on Information Theory*, 28(2):343–346, March 1982.

[15] I-Hsiang Wang and David N. C. Tse. Interference mitigation through limited receiver cooperation. *IEEE Transactions on Information Theory*, 57(5):2913–2940, May 2011.

[16] I-Hsiang Wang and David N. C. Tse. Interference mitigation through limited transmitter cooperation. *IEEE Transactions on Information Theory*, 57(5):2941–2965, May 2011.

[17] I-Hsiang Wang, Sudeep U. Kamath, and David N. C. Tse. Two unicast information flows over linear deterministic networks. *Proceedings of IEEE International Symposium on Information Theory*, July 2011. Extended Preprint `http://arxiv.org/abs/1105.6326`.

[18] Frans M. J. Willems. The discrete memoryless mulitple access channel with partially cooperating encoders. *IEEE Transactions on Information Theory*, 29(3):441–445, May 1983.

[19] Shraga I. Bross, Amos Lapidoth, and Michéle A. Wigger. The gaussian mac with conferencing encoders. *Proceedings of IEEE International Symposium on Information Theory*, July 2008.

[20] Ivana Marić, Roy D. Yates, and Gerhard Kramer. Capacity of interference channels with partial transmitter cooperation. *IEEE Transactions on Information Theory*, 53(10):3536–3548, October 2007.

[21] Ron Dabora and Sergio D. Servetto. Broadcast channels with cooperating decoders. *IEEE Transactions on Information Theory*, 52(12):5438–5454, December 2006.

[22] Osvaldo Simeone, Deniz Gündüz, H. Vincent Poor, Andrea J. Goldsmith, and Shlomo Shamai. Compound multiple access channels with partial cooperation. *IEEE Transactions on Information Theory*, 55(6):2425–2441, June 2009.

[23] Wei Yu and Lei Zhou. Gaussian z-interference channel with a relay link: Achievability region and asymptotic sum capacity. *Arxiv Preprint*, 2010, `http://arxiv.org/abs/1006.5087`.

[24] Anders Høst-Madsen. Capacity bounds for cooperative diversity. *IEEE Transactions on Information Theory*, 52(4):1522–1544, April 2006.

[25] Vinod Prabhakaran and Pramod Viswanath. Interference channels with destination cooperation. *IEEE Transactions on Information Theory*, 57(1):187–209, January 2011.

[26] Ivana Marić, Ron Dabora, and Andrea J. Goldsmith. On the capacity of the interference channel with a relay. *Proceedings of IEEE International Symposium on Information Theory*, July 2008.

[27] Onur Sahin, Elza Erkip, and Osvaldo Simeone. Interference channel with a relay: Models, relaying strategies, bounds. *Proceedings of Information Theory and Applications Workshop*, February 2009.

[28] Peyman Razaghi and Wei Yu. Universal relaying for the interference channel. *Proceedings of Information Theory and Applications Workshop*, 2010.

[29] Soheil Mohajer, Suhas N. Diggavi, Christina Fragouli, and David N. C. Tse. Approximate capacity of a class of gaussian interference-relay networks. *IEEE Transactions on Information Theory*, 57(5):2837–2864, May 2011.

[30] Thomas M. Cover and Young-Han Kim. Capacity of a class of deterministic relay channels. *Proceedings of IEEE International Symposium on Information Theory*, June 2007.

[31] Young-Han Kim. Coding techniques for primitive relay channels. *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 2007.

[32] Amichai Sanderovich, Shlomo Shamai, Yossef Steinberg, and Gerhard Kramer. Communication via decentralized processing. *IEEE Transactions on Information Theory*, 54(7):3008–3023, July 2008.

[33] Clemens Schnurr, Sławomir Stańczak, and Tobias J. Oechtering. Coding theorems for the restricted half-duplex two-way relay channel with joint decoding. *Proceedings of IEEE International Symposium on Information Theory*, pages 2688–2692, July 2008.

[34] Amir Salman Avestimehr, Suhas N. Diggavi, and David N. C. Tse. Wireless network information flow. *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 2007.

[35] Stark C. Draper, Brendan J. Frey, and Frank R. Kschischang. Interactive decoding of a broadcast message. *Proceedings of Allerton Conference on Communication, Control, and Computing*, October 2003.

[36] Sung Hoon Lim, Young-Han Kim, Abbas A. El Gamal, and Sae-Young Chung. Noisy network coding. *IEEE Transactions on Information Theory*, 57(5):3132–3152, May 2011.

[37] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2nd edition, 2006.

[38] Hon-Fah Chong, Mehul Motani, Hari Krishna Garg, and Hesham El Gamal. On the han-kobayashi region for the interference channel. *IEEE Transactions on Information Theory*, 54(7):3188 – 3195, July 2008.

[39] Rudolf Ahlswede. Multi-way communication channels. *Proceedings of 2nd International Symposium on Information Theory*, pages 23–52, September 1971.

[40] H. H. J. Liao. *Multiple Access Channels*. PhD thesis, University of Hawaii, Honolulu, September 1972.

[41] Hanan Weingarten, Yossef Steinberg, and Shlomo Shamai. The capacity region of the multiple-input-multiple-output broadcast channel. *IEEE Transactions on Information Theory*, 52(9):3936–3964, September 2006.

[42] Vinod Prabhakaran and Pramod Viswanath. Interference channels with source cooperation. *IEEE Transactions on Information Theory*, 57(1):156–186, January 2011.

[43] Yi Cao and Biao Chen. An achievable rate region for interference channels with conferencing. *Proceedings of IEEE International Symposium on Information Theory*, July 2007, `http://arxiv.org/abs/0704.1873`.

[44] Wei Wu, Sriram Vishwanath, and Ari Arapostathis. Capacity of a class of cognitive radio channels: Interference channels with degraded message sets. *IEEE Transactions on Information Theory*, 53(11):4391–4399, November 2007.

[45] Aleksandar Jovičić and Pramod Viswanath. Cognitive radio: An information-theoretic perspective. *IEEE Transactions on Information Theory*, 55(9):3945–3958, September 2009.

[46] Stefano Rini, Daniela Tuninetti, and Natasha Devroye. State of the cognitive interference channel: a new unified inner bound, and capacity to within 1.87 bits. *Proceedings of International Zurich Seminar on Communications*, 2010, `http://arxiv.org/abs/0910.3028`.

[47] Hossein Bagheri, Abolfazl Seyed Motahari, and Amir Keyvan Khandani. On the symmetric gaussian interference channel with partial unidirectional cooperation. *Preprint*, September 2009. `http://arxiv.org/abs/0909.2777`.

[48] Adnan Raja, Vinod Prabhakaran, and Pramod Viswanath. Reciprocity in linear deterministic networks under linear coding. July 2009, `http://arxiv.org/abs/0907.1432`.

[49] Abbas A. El Gamal and Young-Han Kim. Lecture notes on network information theory. January 2010, `http://arxiv.org/abs/1001.3404`.

[50] Katalin Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Transactions on Information Theory*, 25(3):306–311, May 1979.

[51] Guy Bresler, Abhay Parekh, and David N. C. Tse. The approximate capacity of the many-to-one and one-to-many gaussian interference channels. *IEEE Transactions on Information Theory*, 56(9):4566–4592, September 2010.

[52] Chih-Chun Wang and Ness B. Shroff. Pairwise intersession network coding on directed networks. *IEEE Transactions on Information Theory*, 56(8):3879–3900, August 2010.

[53] Rudolf Ahlswede, Ning Cai, Shuo-Yen R. Li, and Raymond W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[54] Xijin Yan, Jun Yang, and Zhen Zhang. An outer bound for multisource multisink network coding with minimum cost consideration. *IEEE Transactions on Information Theory*, 52(6):2373–2385, June 2006.

[55] Sudeep U. Kamath, David N. C. Tse, and Venkat Anantharam. Generalized network sharing outer bound and the two-unicast problem. *to appear in International Symposium on Network Coding 2011*, July 2011.

[56] Tiangao Gou, Syed A. Jafar, Sang-Woon Jeon, and Sae-Young Chung. Aligned interference neutralization and the degrees of freedom of the 2x2x2 interference channel. *Arxiv Preprint*, December 2010, `http://arxiv.org/abs/1012.2350`.

[57] Ilan Shomorony and Amir Salman Avestimehr. Two-unicast wireless networks: Characterizing the degrees-of-freedom. *Submitted to IEEE Transactions on Information Theory*, March 2011, `http://arxiv.org/abs/1102.2498`.

[58] Minji Kim and Muriel Médard. Algebraic network coding approach to deterministic wireless relay networks. *Proceedings of Allerton Conference on Communication, Control, and Computing*, September 2010.

[59] Patrick Marsch and Gerhard P. Fettweis, editors. *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Cambridge University Press, July 2011.

[60] Vinayak Nagpal, I-Hsiang Wang, Milos Jorgovanovic, David N. N. Tse, and Borivoje Nikolić. Quantize-map-and-forward relaying: Coding and system design. *Proceedings of Allerton Conference on Communication, Control, and Computing*, pages 443–450, September 2010.