

On the Benefit of Network Coding in Wireless Relay Networks

FONG, Lik Hang Silas

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

The Chinese University of Hong Kong
January 2011

UMI Number: 3492007

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3492007

Copyright 2011 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Abstract of thesis entitled:

On the Benefit of Network Coding in Wireless Relay Networks
Submitted by FONG, Lik Hang Silas
for the degree of Doctor of Philosophy
at The Chinese University of Hong Kong in January 2011

Our investigation of wireless relay networks begins by studying the *two-way relay channel* (TRC), in which a user and a base station exchange their messages with the help of a middle relay. We model the TRC as a three-node point-to-point relay network and propose practical symbol-level network coding schemes for the three-node network. We obtain several rate regions achievable by the network coding schemes and show that the use of symbol-level network coding rather than routing alone always enlarges the achievable rate region. In particular, the use of symbol-level network coding always increases the maximum equal-rate throughput. Furthermore, we model a cellular relay network consisting of multiple users, multiple relays and multiple base stations as a collection of two-node point-to-point systems and three-node point-to-point relay networks where each point-to-point channel is modeled as a bandlimited Gaussian channel. We propose several practical symbol-level network coding schemes on the network and investigate the benefit of symbol-level network coding by simulation. Our simulation results show that the use of symbol-level network coding rather than routing alone increases the average maximum equal-rate throughput over all users.

Next, we investigate several models of TRC including the discrete memoryless TRC, the Gaussian TRC and the bandlimited Gaussian TRC, and prove an outer bound on the capacity region of each of the TRC models. In particular, the outer bound on the capacity region of the bandlimited Gaussian TRC is a theoretical outer bound on the capacity region achievable by physical-layer network coding (PNC). Furthermore, we model a cellular relay network consisting of multiple users, multiple relays and multiple base stations as a collection of two-node point-to-point systems and three-node networks, where each two-node point-to-point system consists of two bandlimited Gaussian channels and each three-node network consists of a bandlimited Gaussian TRC. We obtain performance bounds of PNC on the cellular relay network by simulation and our simulation results show that the average maximum equal-rate throughput over all users under every PNC strategy investigated is generally worse than the average equal-rate throughput over all users under some routing strategy. This is possibly due to larger interference among the nodes under the PNC strategies compared with the routing strategy.

Abstract of thesis in Chinese entitled:

On the Benefit of Network Coding in Wireless Relay Networks

Submitted by FONG, Lik Hang Silas

for the degree of Doctor of Philosophy

at The Chinese University of Hong Kong in January 2011

為了研究無線中繼網絡(wireless relay networks)，我們首先考慮由三個節點(nodes)組成的雙向中繼信道(two-way relay channel)，以下簡稱為 TRC。TRC 是由一個用戶(user)、一個發射站(base station)和一個中繼(relay)所組成，中繼的角色是協助用戶與發射站的雙向信息傳遞。我們用一個三節點中繼網絡(three-node network)去模仿 TRC，然後提出切實可行的符號級網絡編碼(symbol-level network coding)傳送方式。我們計算了那些符號級網絡編碼傳送方式的速度可達到領域(achievable rate region)，及證明了符號級網絡編碼的速度可達到領域相對於純粹路由(routing alone)的速度可達到領域更大，尤其是使用網絡編碼可達至一個更高的等速率傳送速度(equal-rate throughput)。此外，我們用多個二節點系統(two-node system)和多個三節點中繼網絡(three-node relay network)去模仿一個多個用戶、多個發射站和多個中繼的無線中繼網絡。在那個無線中繼網絡中，每一個點到點信道都是一個帶限高斯通道(bandlimited Gaussian channel)。我們在那個無線中繼網絡裡提出切實可行的符號級網絡編碼(symbol-level network coding)，再用模擬軟件去研究在無線中繼網絡中使用網絡編碼究竟有沒有好處。我們的模擬結果顯示在無線中繼網絡裡使用網絡編碼可以提高用戶的最大等速率傳送速度的平均值(average maximum equal-rate throughput over all users)。

接著，我們研究幾種不用的 TRC，包括 discrete memoryless TRC， Gaussian TRC 和 bandlimited Gaussian TRC，然後找出了每個 TRC 的速度可達到領域上限，其中 bandlimited Gaussian TRC 的等速率速度上限是使用物理級網絡編碼(physical-layer network coding，簡稱 PNC)的等速率速度上限。此外，我們用多個二節點系統和多個三節點網絡去模仿一個多個用戶、多個發射站和多個中繼的無線中繼網絡，而在那無線中繼網絡中，每一個在二節點系統裡的點到點信道都是一個帶限高斯通道(bandlimited Gaussian channel)，而每一個三節點網絡都包含一個 bandlimited Gaussian TRC。然後，我們用模擬軟件去研究在無線中繼網絡裡使用 PNC 究竟有沒有好處。我們的模擬結果顯示在無線中繼網絡裡使用 PNC 的最大等速率傳送速度的平均值是低於使用某種純路由的等速率傳送速度的平均值，這很可能是由於在無線中繼網絡裡使用 PNC 相對於純路由會增加了節點與節點之間的互相干擾。

Acknowledgement

I wish to express my gratitude to my supervisor, Prof. Raymond Yeung, for his precious advice on my thesis and his guidance during my Doctor of Philosophy degree in Information Engineering.

This research project is partially supported by a grant from Qualcomm and was initiated in September, 2008 at Qualcomm China. I wish to express my gratitude to Dr. Mingxi Fan for his guidance during my intern in Qualcomm China.

I would like to express my gratitude to my parents for their support for my education throughout many years. I would also like to thank my colleagues for answering me plenty of questions during the degree.

Last but not least, I would like to thank the almighty God, the source of all intelligence, for His guidance throughout my life.

Contents

Abstract	i
Acknowledgement	iii
Acknowledgement	iv
1 Introduction	1
1.1 Symbol-Level Network Coding	4
1.2 Physical-Layer Network Coding	6
1.3 Contributions	8
I Symbol-Level Network Coding	12
2 Three-Node Relay Network	13
2.1 Four-Session Network Coding Scheme	16
2.1.1 Achievable Rate Region \mathcal{T}_4	17
2.1.2 Evaluation of \mathcal{T}_4	18
2.1.3 Advantage of Network Coding	24
2.2 Five-Session Network Coding Scheme	27
2.2.1 Achievable Rate Region \mathcal{T}_5	28
2.2.2 Evaluation of \mathcal{T}_5 when $\frac{A}{C} + \frac{B}{E} \leq 1$	32
2.2.3 Evaluation of \mathcal{T}_5 when $\frac{A}{C} + \frac{B}{E} > 1$	33

2.3	Insufficiency of Four-Session Network Coding Scheme	46
3	Performance Analysis in Cellular Relay Network	53
3.1	Simulation Assumptions	55
3.1.1	Deployment Model	55
3.1.2	Power and Bandwidth Settings	56
3.1.3	Propagation Model	57
3.1.4	Single-Hop and Multi-Hop Users	58
3.1.5	Channel Model	59
3.1.6	Fading Model	61
3.2	Three-Phase Mode	61
3.2.1	Resource and Interference Management	62
3.2.2	Maximum Equal-Rate Throughput	66
3.2.3	Performance of Routing	72
3.2.4	Network Coding Gain for Individual User	75
3.2.5	Average Network Coding Gain	77
3.3	Four-Phase Mode	83
3.3.1	Resource and Interference Management	84
3.3.2	Maximum Equal-Rate Throughput	89
3.3.3	Performance of Routing	116
3.3.4	Network Coding Gain for Individual User	125
3.3.5	Average Network Coding Gain	127
II	Physical-Layer Network Coding	135
4	Capacities of Two-Way Relay Channels	136
4.1	Notation	139
4.2	Full-Duplex Two-Way Relay Channel	140

4.3	Full-Duplex Discrete Memoryless Two-Way Relay Channel without Feedback	141
4.3.1	Definitions	142
4.3.2	An Outer Bound	144
4.4	Full-Duplex Discrete Memoryless Two-Way Relay Channel with Feedback	146
4.4.1	Benefit of Feedback	146
4.4.2	Definitions	149
4.4.3	An Outer Bound	153
4.4.4	Looseness of the Outer Bound	153
4.5	Full-Duplex Gaussian TRC with feedback	157
4.5.1	Definitions	157
4.5.2	An Outer Bound	161
4.6	Half-Duplex Two-Way Relay Channels	162
4.6.1	Definitions	162
4.6.2	Discrete Memoryless Two-Way Relay Channel	163
4.6.3	Gaussian Two-Way Relay Channel	166
4.6.4	Bandlimited Gaussian Two-Way Relay Channel	168
4.7	Half-Duplex Gaussian Two-Way Relay Channel with Varying Background Noise	169
4.7.1	Definitions	170
4.7.2	An Outer Bound	176
4.7.3	Half-Duplex Bandlimited Gaussian Two-Way Relay Channel with Varying Background Noise	176
5	Performance Bounds in Cellular Relay Network	179
5.1	Simulation Assumptions	182
5.2	Strategy 1	183
5.2.1	Resource and Interference Management	184

5.2.2	Performance Upper Bound for Individual User	186
5.2.3	System Performance Upper Bound	189
5.3	Strategy 2	192
5.3.1	Resource and Interference Management	192
5.3.2	Performance Upper Bound for Individual User	195
5.3.3	System Performance Upper Bound	196
5.4	Strategy 3	199
5.4.1	Resource and Interference Management	199
5.4.2	Performance Upper Bound for Individual User	203
5.4.3	System Performance Upper Bound	206
6	Conclusion	209
A	Proofs for Chapter 2	213
B	Proofs for Chapter 4	238
	Bibliography	267

Chapter 1

Introduction

Summary

An introduction of the two-way relay channel (TRC) is given and the assumptions of the TRC are presented. Two examples are given to illustrate the benefit of symbol-level network coding and physical-layer network coding on the TRC. The methodologies for investigating the benefits of symbol-level network coding and physical-layer network coding on the TRC as well as on a cellular relay network are described.

In a traditional cellular architecture where the base stations communicate with the mobile users directly, the data rates of the users at the cell edge are severely limited due to strong interference from neighboring cells and large propagation loss from the serving base station. One way to improve the throughput for these cell-edge users is by placing relay nodes at different locations in the cell. Since the relay nodes typically have higher and stronger antennas than the mobile users, they have better channels than the mobile users to

the base station. In addition, the mobile users that are close to some relay nodes can receive signals from the relay nodes with high signal to interference plus noise ratio (SINR) due to close proximity. Therefore, the installation of relay nodes generally improves the overall throughput of the system [1–3].

For a given mobile user and its associated base station, the introduction of a relay node virtually creates a *two-way relay channel* (TRC) [4], in which the two terminals corresponding to the mobile user and the base station exchange their messages with the help of the middle relay. Several achievable rate regions for the TRC have been obtained in [4–11]. However, the transmission schemes proposed in all the above studies are far from practical due to their high complexity.

Therefore, we focus on investigating practical coding schemes for the TRC. Since it is difficult to sufficiently isolate a received wireless signal from a simultaneously transmitted wireless signal in the same frequency spectrum, we assume that all the nodes in the TRC are half-duplex, which means they cannot transmit and receive information at the same time. Let t_1 and t_2 denote the two terminals and r denote the relay in the TRC. In this thesis, we investigate practical network coding schemes which consist of two phases as follows. Node t_1 and node t_2 transmit messages to r through the Multiple Access Channel (MAC) in the first phase and r transmits messages to node t_1 and node t_2 through the Broadcast Channel (BC) in the second phase. Let α and β denote the fractions of time allocated to the MAC phase and the BC phase respectively where α and β are two real numbers such that $0 \leq \alpha, \beta \leq 1$ and $\alpha + \beta \leq 1$. Suppose node t_1 and node t_2 exchange information in n time slots.

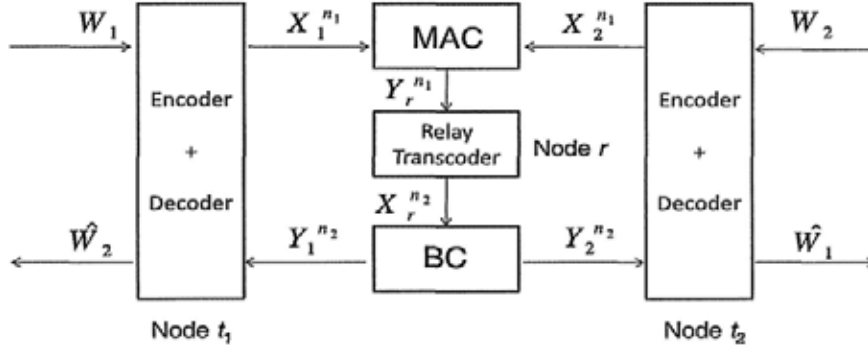


Figure 1.1: A half-duplex TRC.

Let $\lfloor a \rfloor$ denote the largest integer less than or equal to a real number a . To simplify analysis, we assume that the first $n_1 = \lfloor n\alpha \rfloor$ time slots are allocated to the first phase and the last $n_2 = \lfloor n\beta \rfloor$ time slots are allocated to the second phase. Then, the variables involved in the communication through the TRC are:

- $W_i \in \{1, 2, \dots, M_i\}$: messages of node t_i , $i = 1, 2$,
- $X_i^{n_1} = [X_{i,1} X_{i,2} \dots X_{i,n_1}]^T$: codewords transmitted by node t_i , $i = 1, 2$,
- $Y_r^{n_1} = [Y_{r,1} Y_{r,2} \dots Y_{r,n_1}]^T$: channel outputs at relay r ,
- $X_r^{n_2} = [X_{r,1} X_{r,2} \dots X_{r,n_2}]^T$: codewords transmitted by relay r ,
- $Y_i^{n_2} = [Y_{i,1} Y_{i,2} \dots Y_{i,n_2}]^T$: channel outputs at node t_i , $i = 1, 2$,
- $\hat{W}_i \in \{1, 2, \dots, M_i\}$: message estimates of message W_i , $i = 1, 2$.

Node t_1 and node t_2 choose messages W_1 and W_2 independently according to the uniform distribution, and they declare the message estimates \hat{W}_2 and \hat{W}_1 respectively after n time slots. The transmissions of messages in the half-duplex TRC are shown in Figure 1.1.

Definition 1 The average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = Pr\{\hat{W}_1 \neq W_1\}$ and $P_{e,2}^n = Pr\{\hat{W}_2 \neq W_2\}$ respectively.

Definition 2 A rate pair (R_1, R_2) is (α, β) -achievable if there exists a sequence of transmission schemes lasting n time slots that allocate the first $n_1 = \lfloor n\alpha \rfloor$ time slots to the first phase and the last $n_2 = \lfloor n\beta \rfloor$ time slots to the second phase with $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$. A rate R is called an equal-rate throughput if (R, R) is (α, β) -achievable for some α and β .

By Definition 2, an equal-rate throughput is a rate that can be simultaneously achieved in both directions of the TRC.

1.1 Symbol-Level Network Coding

Network coding, first studied by Ahlswede *et al.* [12], reveals that if coding is applied at the nodes in a network rather than routing alone, the network capacity can be increased. The use of network coding also increases the maximum equal-rate throughput for some TRCs. The following example in [13] is given to illustrate the benefit of network coding for a TRC.

Example 1 Consider a TRC consisting of two terminal nodes t_1 and t_2 and a relay node r . Suppose only one of the following three types of transmission can take place in each time slot:

Type 1: Node t_1 transmits a bit b_1 to r and r can decode b_1 correctly.

Type 2: Node t_2 transmits a bit b_2 to r and r can decode b_2 correctly.

Type 3: Node r broadcasts a bit b_r to both node t_1 and node t_2 , and both node t_1 and node t_2 can decode b_r correctly.

Suppose the two terminals want to exchange one bit in two phases where the first phase consists of Type 1 and Type 2 transmissions and the second phase consists of Type 3 transmissions. Then, we observe that any routing scheme that completes the exchange requires at least four time slots and therefore the maximum equal-rate throughput achievable by all routing schemes is $1/4$. Consider the following network coding scheme: Node t_1 transmits a bit b_1 to r using a Type 1 transmission in the first time slot and node t_2 transmits a bit b_2 to r using a Type 2 transmission in the second time slot. Node r receives b_1 and b_2 in the first and second time slots respectively and broadcasts bit $b_1 \oplus b_2$ to node t_1 and node t_2 using a Type 3 transmission in the third time slot, where \oplus is the XOR operation between two bits. Node t_1 can decode b_2 correctly by performing $b_1 \oplus (b_1 \oplus b_2) = b_2$ and node t_2 can decode b_1 correctly by performing $b_2 \oplus (b_1 \oplus b_2) = b_1$. Since only three time slots are required by the network coding scheme for the information exchange, the equal-rate throughput achievable by the network coding scheme is $1/3$, which is greater than the maximum equal-rate throughput achievable by all routing schemes.

Practical capacity-approaching coding schemes for a point-to-point bandlimited Gaussian channel such as Turbo codes [14] and LDPC codes [15] are well understood. Therefore, we study practical network coding schemes on the TRC by modeling the TRC as a *three-node point-to-point relay network* consisting of four independent point-to-point bandlimited Gaussian channels as shown in Figure 1.2, where the MAC and the BC consist of two point-to-point

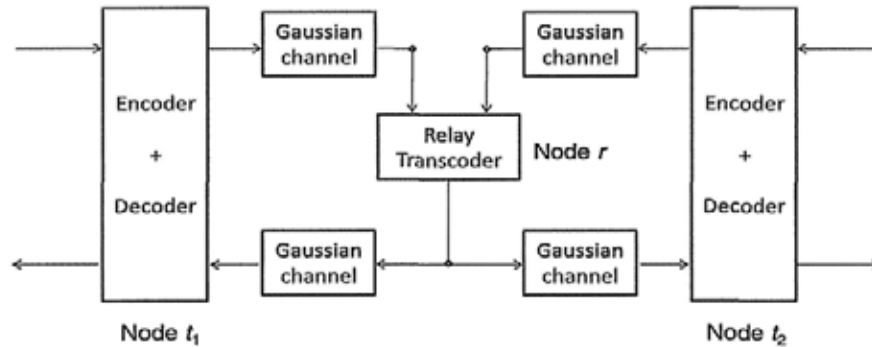


Figure 1.2: A half-duplex TRC modeled as a network of point-to-point Gaussian channels.

bandlimited Gaussian channels and two point-to-point bandlimited Gaussian channels with a common input respectively.

So far, we have discussed network coding at the *symbol level*, which means that the relay decodes the messages from the two terminals before broadcasting any information based on the decoded messages. In the following section, we will discuss another type of network coding which is different from symbol-level network coding.

1.2 Physical-Layer Network Coding

Zhang *et al.* [16] proposed a transmission scheme on a Gaussian TRC called *physical-layer network coding* (PNC), in which the relay does not decode the messages from the two terminals. Instead, the two terminals of the TRC simultaneously transmit their messages to the relay in the MAC phase. The signal received at the relay is the sum of the signals transmitted by the two terminals contaminated by noise. Based on the received signal, the relay broadcasts information to the two terminals in the BC phase. In [16], they showed that PNC can potentially achieve higher maximum equal-rate throughput than

symbol-level network coding for some Gaussian TRC. The following example illustrates the benefit of PNC over symbol-level network coding for a TRC.

Example 2 Consider a TRC consisting of two terminal nodes t_1 and t_2 and a relay node r . Suppose only one of the following four types of transmission can take place in each time slot:

Type 0: Node t_1 transmits a bit b_1 to r and node t_2 transmits a bit b_2 to r simultaneously. Node r can decode $b_1 \oplus b_2$ correctly.

Type 1: Node t_1 transmits a bit b_1 to r and r can decode b_1 correctly.

Type 2: Node t_2 transmits a bit b_2 to r and r can decode b_2 correctly.

Type 3: Node r broadcasts a bit b_r to both node t_1 and node t_2 , and both node t_1 and node t_2 can decode b_r correctly.

Type 1, 2 and 3 transmissions in this example are the same as those in Example 1. Type 0 transmissions are redundant for all routing schemes and all symbol-level network coding schemes because for any routing scheme and any symbol-level network coding scheme, relay r decodes the messages of node t_1 and node t_2 before forwarding the messages to the two terminals. Consequently, the results in Example 1 continue to apply in this example and the results are stated as follows:

1. The maximum equal-rate throughput achievable by all routing schemes is $1/4$.
2. There exists a symbol-level network coding scheme that achieves the equal-rate throughput $1/3$.

Consider the following PNC scheme: Node t_1 and node t_2 transmit two bits b_1 and b_2 respectively using a Type 0 transmission in the first time slot. Node r receives $b_1 \oplus b_2$ in the first time slot and broadcasts $b_1 \oplus b_2$ to node t_1 and node t_2 using a Type 3 transmission in the second time slot. Node t_1 can decode b_2 correctly by performing $b_1 \oplus (b_1 \oplus b_2) = b_2$ and node t_2 can decode b_1 correctly by performing $b_2 \oplus (b_1 \oplus b_2) = b_1$. Since the information exchange is completed in two time slots, the equal-rate throughput achievable by the transmission scheme is $1/2$, which is better than the equal-rate throughput achievable by all routing schemes as well as the symbol-level network coding scheme described in Example 1.

In order to obtain a theoretical outer bound on the capacity region achievable by PNC, we will study different models of TRC, including the discrete memoryless TRC, the Gaussian TRC and the bandlimited Gaussian TRC.

1.3 Contributions

This thesis consists of two parts. Part I is devoted to the study of symbol-level network coding. In this part, we study practical symbol-level network coding schemes on the three-node point-to-point relay network described in Section 1.1. We obtain several rate regions achievable by the network coding schemes on the three-node point-to-point relay network and show theoretically that the use of network coding rather than routing alone always enlarges the achievable rate region for the network. In particular, the use of network coding increases the maximum equal-rate throughput in the three-node point-to-point relay network.

Then, we study practical symbol-level network coding schemes

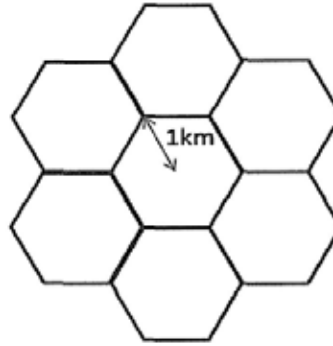


Figure 1.3: Seven hexagonal cells.

for a cellular relay network which consists of seven hexagonal cells with 1km radius as shown in Figure 1.3. Each cell contains a base station, multiple users and multiple relays. The users are classified into two types: *single-hop users* and *multi-hop users*. Each single-hop user communicates with the base station directly, while each multi-hop user communicates with the base station through a relay. We model each single-hop user and its associated base station as a *two-node point-to-point system* where each point-to-point channel in the system is a bandlimited Gaussian channel. On the other hand, we model each multi-hop user, its associated relay and its associated base station together as a three-node point-to-point relay network. In other words, we model the cellular relay network as a collection of two-node point-to-point systems and three-node point-to-point relay networks.

Since the transmitting nodes in the same cell and different cells can interfere with each other in the cellular relay network, each three-node point-to-point relay network in the cellular relay network experiences more interference than an isolated three-node point-to-point relay network. Consequently, the benefit of network coding in the cellular relay network becomes unclear. This motivates us to

investigate the benefit of network coding in such a network.

We propose several practical symbol-level network coding schemes for the cellular relay network and investigate by MATLAB simulation their performance, which cannot be analyzed theoretically due to the large number of nodes. Our simulation results show that the use of these network coding schemes can improve the average maximum equal-rate throughput over all users.

Part II is devoted to the study of physical-layer network coding (PNC). In this part, we study several models of TRC including the discrete memoryless TRC, the Gaussian TRC and the bandlimited Gaussian TRC, and prove an outer bound on the capacity region of each of the TRC models. In particular, the outer bound on the capacity region of the bandlimited Gaussian TRC is a theoretical outer bound on the capacity region achievable by PNC.

Then, we obtain performance bounds of PNC in a cellular relay network. Such a network is identical to the one discussed before except that we model each multi-hop user, its associated relay and its associated base station together as a three-node network. In other words, we model the cellular relay network as a collection of two-node point-to-point systems (each consisting of a single-hop user and a base station) and three-node networks (each consisting of a single-hop user, a relay and a base station). Each two-node point-to-point system consists of two bandlimited Gaussian channels and each three-node network consists of a bandlimited Gaussian TRC. As in the case of symbol-level network coding, since the transmitting nodes in the same cell and different cells can interfere with each other in the cellular relay network, each three-node network in the cellular relay network experiences more interference than an isolated three-

node network. Consequently, the benefit of PNC in the cellular relay network becomes unclear. This motivates us to obtain performance bounds of PNC in the network.

Since the performance bounds of PNC cannot be analyzed theoretically due to the large number of nodes, we conduct the investigation by MATLAB simulation. Our simulation results show that the average maximum equal-rate throughput over all users under every PNC strategy investigated is generally worse than the average equal-rate throughput over all users under some routing strategy. This is possibly due to larger interference among the nodes under the PNC strategies compared with the routing strategy.

□ End of chapter.

Part I

Symbol-Level Network Coding

Chapter 2

Three-Node Relay Network

Summary

We model the two-way relay channel (TRC) as a three-node point-to-point relay network, on which we study two types of practical symbol-level network coding schemes which are called four-session and five-session network coding schemes. A five-session scheme is more general than a four-session scheme. The rate regions achievable by four-session schemes and five-session schemes are obtained in Section 2.1 and Section 2.2 respectively. Section 2.1.3 shows that the use of four-session network coding schemes rather than routing schemes enlarges the achievable rate region, in particular increases the maximum equal-rate throughput. In addition, Section 2.3 gives a criterion that determines when the rate regions achievable by four-session schemes and five-session schemes are equal.

In this chapter, we propose practical symbol-level network coding schemes that enable the exchange of independent messages between the two terminals of the TRC. Practical capacity-approaching cod-

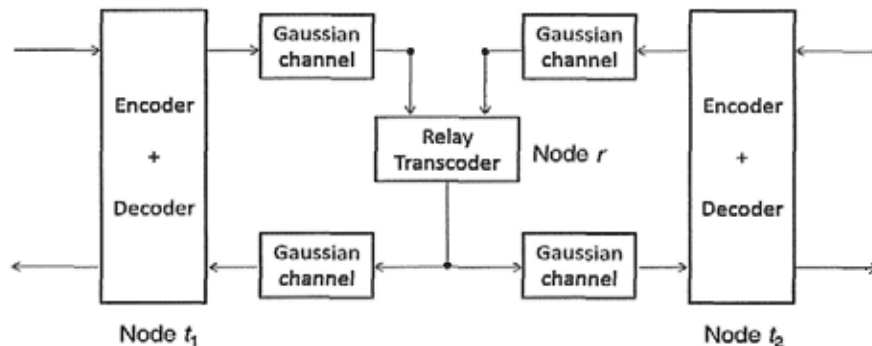


Figure 2.1: A half-duplex TRC modeled as a network of point-to-point Gaussian channels.

ing schemes for a point-to-point bandlimited Gaussian channel such as Turbo codes [14] and LDPC codes [15] are well understood. As discussed in Section 1.1, we therefore study practical network coding schemes on the TRC by modeling the TRC as a *three-node point-to-point relay network* consisting of four independent point-to-point bandlimited Gaussian channels as shown in Figure 1.2, which is reproduced in Figure 2.1 for convenience. The MAC and the BC consist of two point-to-point bandlimited Gaussian channels and two point-to-point bandlimited Gaussian channels with a common input respectively.

Let (u, v) denote the bandlimited Gaussian channel from u to v in the three-node point-to-point relay network where u and v are two adjacent nodes. For comprehensive information-theoretic treatments of Gaussian channels, we refer the reader to [17–19]. One important result for Gaussian channels that we will frequently use is that the channel capacity of a bandlimited Gaussian channel is

$$W \log_2(1 + \sigma) \text{ bits per second,}$$

where W is the bandwidth available for transmission and σ is the

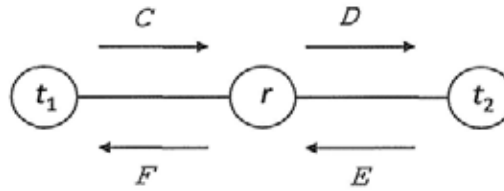


Figure 2.2: The capacities of the links in the three-node point-to-point relay network.

signal-to-noise ratio of the channel. Let $C_{(u,v)}$ denote the channel capacity of the bandlimited Gaussian channel (u, v) . Let C , D , E and F denote $C_{(t_1,r)}$, $C_{(r,t_2)}$, $C_{(t_2,r)}$ and $C_{(r,t_1)}$ respectively. This is illustrated in Figure 2.2.

It is well known from the channel coding theorem that an information rate can be achieved asymptotically by some transmission scheme with arbitrarily small probability of error if and only if it is less than or equal to the channel capacity. Therefore, for any transmission link (u, v) , u can transmit at rate R reliably to v if and only if $R \leq C_{(u,v)}$.

The BC of the three-node point-to-point relay network consists of two point-to-point bandlimited Gaussian channels with a common input. Without loss of generality, we assume in the rest of this chapter that the noise power of channel (r, t_1) is less than the noise power of channel (r, t_2) in the BC phase, which implies that

1.

$$D \leq F; \quad (2.1)$$

2. any transmission scheme that transmits information reliably from r to node t_2 can also transmit information reliably from r to node t_1 .

Note that channels (r, t_1) and (r, t_2) share a common input (cf. Fig-

ure 2.1). It then follows that r can broadcast information reliably from r to both node t_1 and node t_2 at rate R if and only if

$$R \leq \min\{D, F\} = D.$$

We call D the *broadcast capacity*.

2.1 Four-Session Network Coding Scheme

As physical-layer network coding will not be discussed until Part II, symbol-level network coding is simply referred to as network coding throughout Part I. In this section, we study the practical network coding schemes that consist of the following four types of transmissions:

Type 1: Node t_1 transmits messages reliably to r at rate C .

Type 2: Node t_2 transmits messages reliably to r at rate E .

Type 3: Node r broadcasts messages reliably to both node t_1 and node t_2 at rate D , the broadcast capacity.

Type 4: Node r transmits messages reliably to node t_1 at rate F .

Without loss of generality, we assume that the transmissions in the network consist of four sessions such that the transmissions in the i^{th} session are of Type i for $i = 1, 2, 3, 4$. We call the transmission scheme described above a *four-session network coding scheme* or simply a *four-session scheme*.

2.1.1 Achievable Rate Region \mathcal{T}_4

Definition 3 A vector $[p_1 \ p_2 \ p_3 \ p_4]$ is called a four-session allocation if p_1, p_2, p_3 and p_4 are non-negative real numbers such that

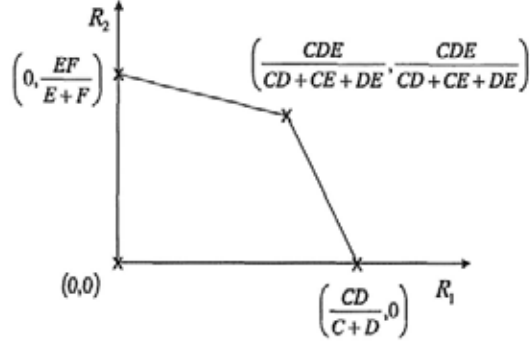
$$p_1 + p_2 + p_3 + p_4 \leq 1. \quad (2.2)$$

Definition 4 Let τ be a four-session scheme on the three-node point-to-point relay network and $s = [p_1 \ p_2 \ p_3 \ p_4]$ be a four-session allocation. The transmission scheme τ is called a transmission scheme with configuration s if under the scheme, a fraction p_i of the time is allocated to Type i transmissions for $i = 1, 2, 3, 4$.

Definition 5 Let s be a four-session allocation, and R_1 and R_2 be two non-negative real numbers. An information rate pair (R_1, R_2) is achievable by a four-session scheme τ if under the scheme, node t_1 and node t_2 exchange independent messages such that node t_1 can transmit messages reliably to node t_2 at an average rate higher than or equal to R_1 and node t_2 can transmit messages reliably to node t_1 at an average rate higher than or equal to R_2 . The rate pair (R_1, R_2) is said to be s -achievable if (R_1, R_2) is achievable by some four-session scheme τ with configuration s .

Definition 6 The four-session achievable information rate region, denoted by \mathcal{T}_4 , is the set

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \mid \begin{array}{l} (R_1, R_2) \text{ is achievable by} \\ \text{some four-session scheme.} \end{array} \right\}.$$

Figure 2.3: \mathcal{T}_4^* .

2.1.2 Evaluation of \mathcal{T}_4

Let \mathcal{T}_4^* denote

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, \\ R_2 \geq 0, \\ R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{CEF+DEF-CDE}{CD(E+F)} \right), \\ R_2 \leq E - R_1 \left(\frac{CE+DE}{CD} \right). \end{array} \right. \right\}, \quad (2.3)$$

which is shown in Figure 2.3, and \mathcal{S}_4 denote

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, p_4 \geq 0, \\ p_1 + p_2 + p_3 + p_4 \leq 1, \\ R_1 \leq \min\{p_1 C, p_3 D\}, \\ R_2 \leq \min\{p_2 E, p_3 D + p_4 F\}. \end{array} \right. \right\}. \quad (2.4)$$

We will show in this subsection that $\mathcal{T}_4 = \mathcal{S}_4 = \mathcal{T}_4^*$.

Lemma 1 Let $s = [p_1 \ p_2 \ p_3 \ p_4]$ be a four-session allocation. If

(R_1, R_2) is s -achievable, then

$$R_1 \leq \min\{p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_2E, p_3D + p_4F\}.$$

Proof: Suppose (R_1, R_2) is s -achievable. Under any four-session transmission scheme τ with configuration s ,

1. the highest average rate that node t_1 can transmit messages reliably to r in the first session is p_1C ;
2. the highest average rate that node t_2 can transmit messages reliably to r in the second session is p_2E ;
3. the highest average rate that r can broadcast messages reliably to both node t_1 and node t_2 in the third session is p_3D ;
4. the highest average rate that r can transmit messages reliably to node t_1 in the fourth session is p_4F .

Therefore, (R_1, R_2) must satisfy the rate constraints induced by the two-source point-to-point network in Figure 2.4(a), where the introduction of the dummy node r' captures the broadcast nature of Type 3 transmissions. We refer the reader to [18, p.415] for the use of the dummy broadcast node. Since node t_1 and node t_2 exchange independent messages in such a way that they eventually possess the same set of messages after the information exchange, the two-source network coding problem for the network in Figure 2.4(a) is equivalent to the single-source multicast problem for the network in Figure 2.4(b), where node t_1 and node t_2 are the receivers and s is the source node of the network. We refer the reader to [18, p.459]

for the equivalence of the two problems. Applying the cut-set bound in [18, p.429] for four different cuts of the network in Figure 2.4(b), we obtain

$$R_1 + R_2 \leq R_2 + p_1 C,$$

$$R_1 + R_2 \leq R_2 + p_3 D,$$

$$R_1 + R_2 \leq R_1 + p_2 E$$

and

$$R_1 + R_2 \leq R_1 + p_3 D + p_4 F.$$

Therefore,

$$R_1 \leq p_1 C,$$

$$R_1 \leq p_3 D,$$

$$R_2 \leq p_2 E$$

and

$$R_2 \leq p_3 D + p_4 F,$$

and the lemma follows. \blacksquare

Lemma 2 Suppose $s = [p_1 \ p_2 \ p_3 \ p_4]$ is an allocation and (R_1, R_2) is a non-negative pair such that

$$R_1 \leq \min\{p_1 C, p_3 D\}$$

and

$$R_2 \leq \min\{p_2 E, p_3 D + p_4 F\}.$$

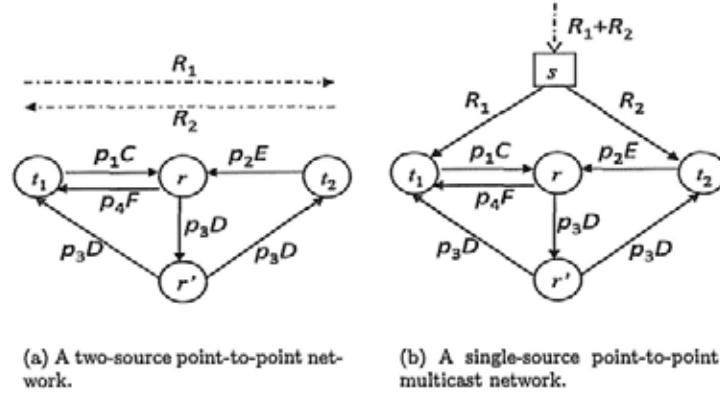


Figure 2.4: Two equivalent network coding problems.

Then, there exists an allocation $s' = [p'_1 \ p'_2 \ p'_3 \ p'_4]$ such that

$$R_1 = p'_1 C = p'_3 D$$

and

$$R_2 \leq \min\{p'_2 E, p'_3 D + p'_4 F\}.$$

Proof: Please refer to Appendix A. ■

Lemma 3 $\mathcal{S}_4 \subset \mathcal{T}_4^*$.

Proof: Please refer to Appendix A. ■

Lemma 4 $\mathcal{T}_4 \subset \mathcal{S}_4 \subset \mathcal{T}_4^*$.

Proof: We have shown that $\mathcal{S}_4 \subset \mathcal{T}_4^*$ in Lemma 3. It remains to show that $\mathcal{T}_4 \subset \mathcal{S}_4$. Suppose (R_1, R_2) is in \mathcal{T}_4 . Then, (R_1, R_2) is s -achievable for some allocation $s = [p_1 \ p_2 \ p_3 \ p_4]$. Using Lemma 1, we obtain

$$R_1 \leq \min\{p_1 C, p_3 D\}$$

and

$$R_2 \leq \min\{p_2E, p_3D + p_4F\},$$

which implies that (R_1, R_2) is in \mathcal{S}_4 (cf. (2.4)). \blacksquare

Proposition 5 \mathcal{T}_4^* is the convex hull of $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ and $(\frac{CD}{C+D}, 0)$.

Proof: Please refer to Appendix A. \blacksquare

Lemma 6 $\mathcal{T}_4 \supset \mathcal{T}_4^*$. In particular, the four extreme points of \mathcal{T}_4^* , which are $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ and $(\frac{CD}{C+D}, 0)$, are achievable by some four-session schemes that involve linear operations only.

Proof: We only need to show that the four extreme points of \mathcal{T}_4^* are achievable by some four-session network coding schemes. Then any other point in \mathcal{T}_4^* can be achieved by time sharing of these schemes. The rate pair $(0, 0)$ is trivially achievable by any transmission scheme. We will propose three transmission schemes τ_1 , τ_2 and τ_3 and show that they achieve the rate pairs $(0, \frac{EF}{E+F})$, $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ and $(\frac{CD}{C+D}, 0)$ respectively.

To facilitate understanding, the allocations $[p_1 \ p_2 \ p_3 \ p_4]$ of τ_1 , τ_2 and τ_3 and the average transmission rates p_1C , p_2E , p_3D , p_4F for Type 1, 2, 3 and 4 transmissions respectively are displayed in Table 2.1 followed by the descriptions of the schemes.

Under scheme τ_1 , the average rates of the messages sent from node t_2 to node r and from node r to node t_1 are both $EF/(E+F)$. Node t_2 transmits its messages to r in the second session and r forwards the messages to node t_1 in the fourth session. Therefore, the rate pair $(0, \frac{EF}{E+F})$ is achievable by scheme τ_2 .

τ_1				τ_2				τ_3			
p_1	0	p_1C	0	p_1	$\frac{DE}{(CD+CE+DE)}$	p_1C	$\frac{CDE}{(CD+CE+DE)}$	p_1	$\frac{D}{(C+D)}$	p_1C	$\frac{CD}{(C+D)}$
p_2	$\frac{F}{(E+F)}$	p_2E	$\frac{EF}{(E+F)}$	p_2	$\frac{CD}{(CD+CE+DE)}$	p_2E	$\frac{CDE}{(CD+CE+DE)}$	p_2	0	p_2E	0
p_3	0	p_3D	0	p_3	$\frac{CE}{(CD+CE+DE)}$	p_3D	$\frac{CDE}{(CD+CE+DE)}$	p_3	$\frac{C}{(C+D)}$	p_3D	$\frac{CD}{(C+D)}$
p_4	$\frac{E}{(E+F)}$	p_4F	$\frac{EF}{(E+F)}$	p_4	0	p_4F	0	p_4	0	p_4F	0

Table 2.1: The allocations of τ_1 , τ_2 and τ_3 and the average transmission rates for Type 1, 2, 3 and 4 transmissions.

Under scheme τ_2 , the average rates of the messages sent from node t_1 to node r , from node t_2 to node r , from node r to node t_1 and from node r to node t_2 are all $\frac{CDE}{(CD+CE+DE)}$. Node t_1 transmits its messages to r in the first session and node t_2 transmits its messages to r in the second session. In the third session, relay r performs XOR operations between the messages from node t_1 and the messages from node t_2 bit by bit and broadcasts the resultant messages to both node t_1 and node t_2 . In addition, node t_1 can recover the messages of node t_2 by performing XOR operations between its own messages and the messages from r bit by bit. Similarly, node t_2 can recover the messages of node t_1 by performing XOR operations between its own messages and the messages from r bit by bit. Consequently, the rate pair $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ is achievable by scheme τ_2 .

Under scheme τ_3 , the average rates of the messages sent from node t_1 to node r and from node r to node t_2 are both $\frac{CD}{(C+D)}$. Node t_1 transmits its messages to r in the first session and r forwards the messages to node t_2 in the third session. Therefore, the rate pair $(\frac{CD}{C+D}, 0)$ is achievable by scheme τ_3 . \blacksquare

Theorem 1 $\mathcal{T}_4 = \mathcal{S}_4 = \mathcal{T}_4^*$.

Proof: It follows directly from Lemma 4 and Lemma 6. ■

The achievable rate region \mathcal{T}_4 is the same as \mathcal{T}_4^* by Theorem 1 and is shown in Figure 2.3. The extreme points $(0, \frac{EF}{E+F})$, $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ and $(\frac{CD}{C+D}, 0)$ of \mathcal{T}_4 correspond to the rate pairs achievable by schemes τ_1 , τ_2 and τ_3 in Lemma 6 respectively. In addition, for any rate pair in \mathcal{T}_4 , Lemma 6 provides a procedure to construct a network code that achieves the rate pair. The network code involves linear operations only since τ_1 , τ_2 and τ_3 involve linear operations only. Consequently, the network code is practical due to its low complexity.

2.1.3 Advantage of Network Coding

For any routing scheme in the three-node point-to-point relay network, the relay receives the messages from node t_1 followed by forwarding the messages to node t_2 , and similar procedures are performed in the opposite direction. Let p_1 , p_2 , p_3 and p_4 denote the fractions of time allocated to the transmission links (t_1, r) , (t_2, r) , (r, t_2) and (r, t_1) respectively. It is readily observed that the transmission rate from node t_1 to node t_2 and the transmission rate from node t_2 to node t_1 of the optimal routing scheme are $\min\{p_1C, p_3D\}$ and $\min\{p_2E, p_4F\}$ respectively. Therefore, the set of rate pairs achievable by the routing schemes on the three-node point-to-point

relay network is

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, p_4 \geq 0, \\ p_1 + p_2 + p_3 + p_4 \leq 1, \\ R_1 \leq \min\{p_1 C, p_3 D\}, \\ R_2 \leq \min\{p_2 E, p_4 F\}. \end{array} \right. \right\}. \quad (2.5)$$

Following similar procedures for proving that (2.3) and (2.4) are equal in Section 2.1.2, we can obtain that (2.5) is equal to

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, \\ R_2 \geq 0, \\ R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{EF(C+D)}{CD(E+F)} \right). \end{array} \right. \right\}, \quad (2.6)$$

which is shown in Figure 2.5.

The rate region achievable by four-session schemes was obtained in the previous section and it is shown in Figure 2.3. It is readily observed by comparing Figure 2.5 with Figure 2.3 that the use of network coding always enlarges the achievable rate region. Using the fact that the maximum equal-rate pair achievable by four-session routing schemes lies on the straight line connecting the points $(\frac{CD}{C+D}, 0)$ and $(0, \frac{EF}{E+F})$, we obtain that the maximum equal-rate pair achievable by four-session routing schemes is $\frac{CDEF}{CDE+CDF+CEF+DEF}$. Since the maximum equal-rate pair in \mathcal{T}_4 is $\frac{CDE}{CD+CE+DE}$ (cf. Fig-

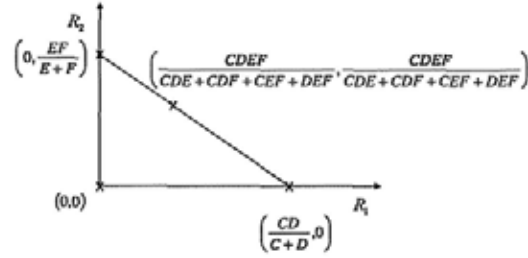


Figure 2.5: The set of rate pairs achievable by four-session routing schemes.

ure 2.3) and

$$\begin{aligned}
 & \frac{CDE}{CD + CE + DE} - \frac{CDEF}{CDE + CDF + CEF + DEF} \\
 &= CDE \left(\frac{CDE + CDF + CEF + DEF - CDF - CEF - DEF}{(CD + CE + DE)(CDE + CDF + CEF + DEF)} \right) \\
 &= \frac{C^2 D^2 E^2}{(CD + CE + DE)(CDE + CDF + CEF + DEF)} \\
 &> 0,
 \end{aligned}$$

network coding rather than routing alone increases the maximum equal-rate throughput in the three-node point-to-point relay network.

Since the maximum equal-rate pair in \mathcal{T}_4 is upper bounded by $\min\{\frac{CD}{C+D}, \frac{EF}{E+F}\}$ (cf. Figure 2.3), the network coding gain in the maximum equal-rate throughput is upper bounded by

$$\begin{aligned}
 & \frac{\min\{\frac{CD}{C+D}, \frac{EF}{E+F}\} - \frac{CDEF}{CDE + CDF + CEF + DEF}}{\frac{CDEF}{CDE + CDF + CEF + DEF}} \\
 &= \min \left\{ \frac{\frac{CD}{C+D} - \frac{CDEF}{CDE + CDF + CEF + DEF}}{\frac{CDEF}{CDE + CDF + CEF + DEF}}, \frac{\frac{EF}{E+F} - \frac{CDEF}{CDE + CDF + CEF + DEF}}{\frac{CDEF}{CDE + CDF + CEF + DEF}} \right\} \\
 &= \min \left\{ \frac{CD(E+F)}{EF(C+D)}, \frac{EF(C+D)}{CD(E+F)} \right\} \\
 &\leq 1.
 \end{aligned}$$

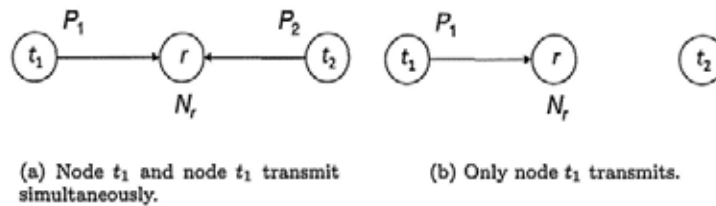


Figure 2.6: Two different conditions.

2.2 Five-Session Network Coding Scheme

In Section 2.1, we studied the four-session network coding schemes where each scheme consists of four types of transmissions. In this section, in addition to the four types of transmissions, we consider *Type 0 transmissions* – node t_1 and node t_2 simultaneously transmit at full power to node r . When node t_1 and node t_2 simultaneously transmit, we assume that the interference incurred on one node by another node behaves as a Gaussian noise. This is the worst-case assumption [18, p.290] [20]. Let $A = W \log_2(1 + P_1/(N_r + P_2))$ and $B = W \log_2(1 + P_1/(N_r + P_1))$ be the channel capacities of the bandlimited Gaussian channels (t_1, r) and (t_2, r) respectively when node t_1 and node t_2 simultaneously transmit, where W , P_1 , P_2 and N_r are the bandwidth available for transmission, the full power of node t_1 , the full power of node t_2 and the noise power at node r respectively. We call A and b the *relay-access capacity of (t_1, r)* and the *relay-access capacity of (t_2, r)* respectively.

Both A and C are the channel capacities of (t_1, r) but they are calculated under the conditions shown in Figure 2.6(a) and Figure 2.6(b) respectively. Since $C = W \log_2(1 + P_1/N_r)$ and $A =$

$W \log_2(1 + P_1/(N_r + P_2))$, it then follows that

$$A < C. \quad (2.7)$$

Similarly,

$$B < E. \quad (2.8)$$

In this section, we study the practical network coding schemes that consist of the following five types of transmissions:

Type 0: Node t_1 and node t_2 transmit messages reliably to r simultaneously at rate A and rate B respectively.

Type 1: Node t_1 transmits messages reliably to r at rate C .

Type 2: Node t_2 transmits messages reliably to r at rate E .

Type 3: Node r broadcasts messages reliably to both node t_1 and node t_2 at rate D , the broadcast capacity.

Type 4: Node r transmits messages reliably to node t_1 at rate F .

Without loss of generality, we assume that the transmissions in the network consist of five sessions such that the transmissions in the $(i+1)^{\text{th}}$ session are of Type i for $i = 0, 1, 2, 3, 4$. We call the transmission scheme described above a *five-session network coding scheme* or simply a *five-session scheme*. Since every four-session scheme is identical to a five-session scheme that assigns no time for Type 0 transmissions (cf. Section 2.1), a five-session scheme is more general than a four-session scheme.

2.2.1 Achievable Rate Region \mathcal{T}_5

Definition 7 A vector $[p_0 \ p_1 \ p_2 \ p_3 \ p_4]$ is called a *five-session allocation* if p_0, p_1, p_2, p_3 and p_4 are non-negative real numbers such

that

$$p_0 + p_1 + p_2 + p_3 + p_4 \leq 1. \quad (2.9)$$

Definition 8 Let τ be a five-session scheme on the three-node point-to-point relay network and $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$ be an allocation. The transmission scheme τ is called a transmission scheme with configuration s if under the transmission scheme, a fraction p_i of the time is allocated to Type i transmissions for $i = 0, 1, 2, 3, 4$.

Definition 9 Let s be a five-session allocation, and R_1 and R_2 be two non-negative real numbers. An information rate pair (R_1, R_2) is achievable by a five-session transmission scheme τ if under the scheme, node t_1 and node t_2 exchange independent messages such that node t_1 can transmit messages reliably to node t_2 at an average rate higher than or equal to R_1 and node t_2 can transmit messages reliably to node t_1 at an average rate higher than or equal to R_2 . The rate pair (R_1, R_2) is said to be s -achievable if (R_1, R_2) is achievable by some five-session scheme τ with configuration s .

Definition 10 The five-session achievable information rate region, denoted by \mathcal{T}_5 , is the set

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \mid \begin{array}{l} (R_1, R_2) \text{ is achievable by} \\ \text{some five-session scheme.} \end{array} \right\}.$$

Lemma 7 Let $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$ be an allocation. If (R_1, R_2) is s -achievable, then

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\}.$$

Proof: Suppose (R_1, R_2) is s -achievable. Under any five-session transmission scheme τ with configuration s ,

1. the highest average rate that node t_1 can transmit messages reliably to r in the first session is p_0A ;
2. the highest average rate that node t_2 can transmit messages reliably to r in the first session is p_0B ;
3. the highest average rate that node t_1 can transmit messages reliably to r in the second session is p_1C ;
4. the highest average rate that node t_2 can transmit messages reliably to r in the third session is p_2E ;
5. the highest average rate that r can broadcast messages reliably to both node t_1 and node t_2 in the fourth session is p_3D ;
6. the highest average rate that r can transmit messages reliably to node t_1 in the fifth session is p_4F .

As in the case of four-session transmission scheme, (R_1, R_2) must satisfy the rate constraints induced by the point-to-point network in Figure 2.7(a). Since node t_1 and node t_2 exchange independent messages in such a way that they eventually possess the same set of messages after the information exchange, the two-source network coding problem for the network in Figure 2.7(a) is equivalent to the single-source multicast problem for the network in Figure 2.7(b), where node t_1 and node t_2 are the receivers and s is the source node of the network. Applying the cut-set bound in [18, p.429] to four

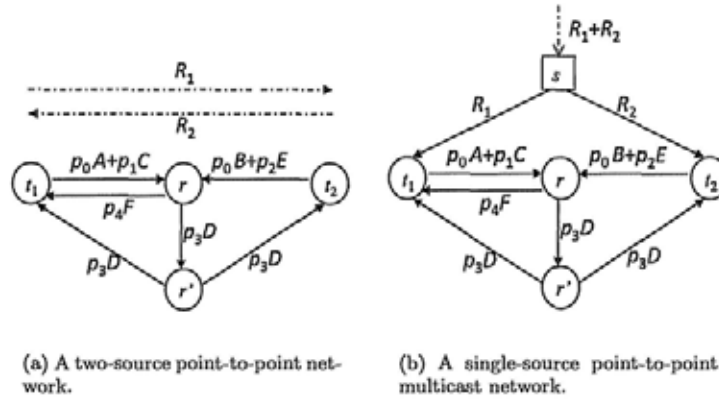


Figure 2.7: Two equivalent network coding problems.

different cuts of the network in Figure 2.7(b), we obtain

$$R_1 + R_2 \leq R_2 + p_0A + p_1C,$$

$$R_1 + R_2 \leq R_2 + p_3D,$$

$$R_1 + R_2 \leq R_1 + p_0B + p_2E$$

and

$$R_1 + R_2 \leq R_1 + p_3D + p_4F.$$

Therefore,

$$R_1 \leq p_0A + p_1C,$$

$$R_1 \leq p_3D,$$

$$R_2 \leq p_0B + p_2E$$

and

$$R_2 \leq p_3D + p_4F,$$

and the lemma follows. □

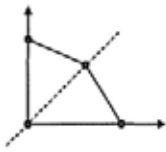
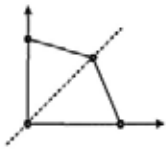
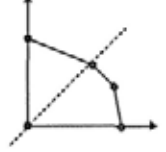
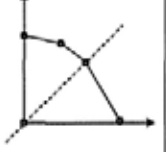
Scenario	$\frac{A}{C} + \frac{B}{E} \leq 1$	$\frac{A}{C} + \frac{B}{E} > 1$		
		$A = B$	$A > B$	$A < B$
\mathcal{T}_5				

Table 2.2: \mathcal{T}_5 under different scenarios.

The main result in this section is an evaluation of \mathcal{T}_5 . The evaluation of \mathcal{T}_5 in the rest of this section is analogous to the evaluation of \mathcal{T}_4 in Section 2.1, but the former is extremely tedious. To facilitate understanding, we present \mathcal{T}_5 briefly in Table 2.2 before proving the results in detail. Every rate region in Table 2.2 represents \mathcal{T}_5 under the scenario specified above the region. The dotted line in each region is the line of equation $y = x$. All the rate regions in Table 2.2 have the same x -intercept and y -intercept. In addition, each region under the scenario $A/C + B/E > 1$ is strictly larger than the region under the scenario $A/C + B/E \leq 1$.

2.2.2 Evaluation of \mathcal{T}_5 when $\frac{A}{C} + \frac{B}{E} \leq 1$

Lemma 8 *If $\frac{A}{C} + \frac{B}{E} \leq 1$, $\mathcal{T}_5 \subset \mathcal{S}_4 = \mathcal{T}_4$.*

Proof: Please refer to Appendix A. ■

Theorem 2 *If $\frac{A}{C} + \frac{B}{E} \leq 1$, $\mathcal{T}_5 = \mathcal{T}_4$.*

Proof: Since a five-session scheme is more general than a four-session scheme, it follows from Definition 6 and Definition 10 that $\mathcal{T}_5 \supset \mathcal{T}_4$. The theorem then follows from Lemma 8. ■

2.2.3 Evaluation of \mathcal{T}_5 when $\frac{A}{C} + \frac{B}{E} > 1$

We have fully characterized \mathcal{T}_5 in Theorem 2 when $\frac{A}{C} + \frac{B}{E} \leq 1$. Therefore, we assume that

$$\frac{A}{C} + \frac{B}{E} > 1 \quad (2.10)$$

in the rest of this subsection.

Lemma 9 *Suppose $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$ is an allocation and (R_1, R_2) is a non-negative pair such that*

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\}.$$

Then, there exists an allocation $s' = [p'_0 \ p'_1 \ p'_2 \ p'_3 \ p'_4]$ such that

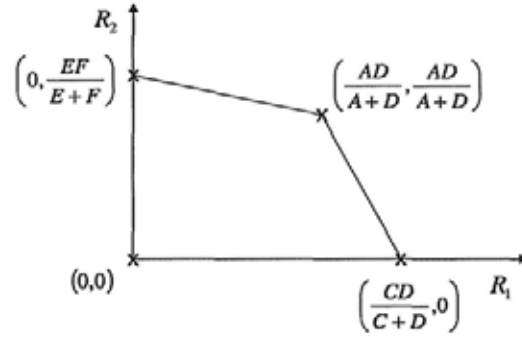
$$R_1 = p'_0A + p'_1C = p'_3D$$

and

$$R_2 \leq \min\{p'_0B + p'_2E, p'_3D + p'_4F\}.$$

Proof: Please refer to Appendix A. ■

We let $\mathcal{T}_{5,A=B}$, $\mathcal{T}_{5,A<B}$ and $\mathcal{T}_{5,A>B}$ denote \mathcal{T}_5 under the scenarios $A = B$, $A < B$ and $A > B$ respectively.

Figure 2.8: $\mathcal{T}_{5,A=B}^*$.

Evaluation of $\mathcal{T}_{5,A=B}$

Let $\mathcal{T}_{5,A=B}^*$ denote

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, \\ R_2 \geq 0, \\ R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-ADF}{AD(E+F)} \right), \\ R_2 \leq \frac{AC}{C-A} - R_1 \left(\frac{AC+AD}{D(C-A)} \right). \end{array} \right. \right\}, \quad (2.11)$$

which is shown in Figure 2.8, and $\mathcal{S}_{A=B}$ denote

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ p_0 \geq 0, p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, p_4 \geq 0 \\ p_0 + p_1 + p_2 + p_3 + p_4 \leq 1, \\ R_1 \leq \min\{p_0 A + p_1 C, p_3 D\}, \\ R_2 \leq \min\{p_0 A + p_2 E, p_3 D + p_4 F\}. \end{array} \right. \right\}. \quad (2.12)$$

We will show that $\mathcal{T}_{5,A=B} = \mathcal{S}_{A=B} = \mathcal{T}_{5,A=B}^*$.

Lemma 10 $\mathcal{S}_{A=B} \subset \mathcal{T}_{5,A=B}^*$.

Proof: Please refer to Appendix A. ■

Lemma 11 $\mathcal{T}_{5,A=B} \subset \mathcal{S}_{A=B} \subset \mathcal{T}_{5,A=B}^*$.

Proof: Suppose $A = B$. Since $\mathcal{S}_{A=B} \subset \mathcal{T}_{5,A=B}^*$ by Lemma 10, it remains to show that $\mathcal{T}_{5,A=B} \subset \mathcal{S}_{A=B}$. Suppose $(R_1, R_2) \in \mathcal{T}_{5,A=B}$. Then, (R_1, R_2) is s -achievable for some allocation $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$. Using Lemma 7, we obtain

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0A + p_2E, p_3D + p_4F\},$$

which implies that (R_1, R_2) is in $\mathcal{S}_{A=B}$ (cf. (2.12)). ■

Proposition 12 $\mathcal{T}_{5,A=B}^*$ is the convex hull of $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{AD}{A+D}, \frac{AD}{A+D})$ and $(\frac{CD}{C+D}, 0)$.

Proof: Please refer to Appendix A. ■

Lemma 13 $\mathcal{T}_{5,A=B} \supset \mathcal{T}_{5,A=B}^*$.

Proof: We only need to show that the four extreme points of $\mathcal{T}_{5,A=B}^*$ are achievable by some five-session network coding schemes. Then any other point in $\mathcal{T}_{5,A=B}^*$ can be achieved by time sharing of these schemes. Using Lemma 6, we obtain three four-session schemes τ_0 , τ_1 and τ_2 that involve linear operations only such that they achieve the rate pairs $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ respectively. Therefore, $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ are in $\mathcal{T}_{5,A=B}$. We will propose a transmission schemes τ_3 and show that it achieves the rate pair $(\frac{AD}{A+D}, \frac{AD}{A+D})$.

τ_3			
p_0	$D/(A+D)$	p_0A	$AD/(A+D)$
p_1	0	p_1C	0
p_2	0	p_2E	0
p_3	$A/(A+D)$	p_3D	$AD/(A+D)$
p_4	0	p_4F	0

Table 2.3: The allocation of τ_3 and its corresponding p_0A , p_1C , p_2E , p_3D and p_4F .

To facilitate understanding, the allocation $[p_0 p_1 p_2 p_3 p_4]$ of τ_3 and its corresponding p_0A , p_1C , p_2E , p_3D and p_4F are displayed in Table 2.3. Under scheme τ_3 , the average transmission rates of (t_1, r) in the first session, (t_2, r) in the first session, (t_1, r) in the second session, (t_2, r) in the third session, (r, t_1) in the fourth session, (r, t_2) in the fourth session and (r, t_1) in the fifth session are p_0A , p_0A , p_1C , p_2E , p_3D , p_3D and p_4F respectively.

Under scheme τ_3 , the average rates of the messages sent from node t_1 to node r , from node t_2 to node r , from node r to node t_1 and from node r to node t_2 are all $AD/(A+D)$. Node t_1 and node t_2 simultaneously transmit their messages to r in the first session. In the fourth session, relay r performs XOR operations between the messages from node t_1 and the messages from node t_2 bit by bit and broadcasts the resultant messages to both node t_1 and node t_2 . In addition, node t_1 can recover the messages of node t_2 by performing XOR operations between its own messages and the messages from r bit by bit. Similarly, node t_2 can recover the messages of node t_1 by performing XOR operations between its own messages and the messages from r bit by bit. Consequently, the rate pair $(\frac{AD}{A+D}, \frac{AD}{A+D})$

is achievable by scheme τ_3 . \blacksquare

Theorem 3 $\mathcal{T}_{5,A=B} = \mathcal{S}_{A=B} = \mathcal{T}_{5,A=B}^*$.

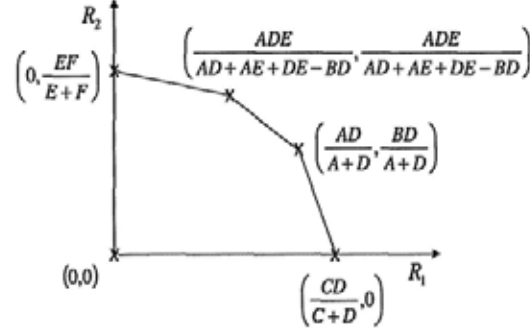
Proof: It follows directly from Lemma 11 and Lemma 13. \blacksquare

The achievable rate region $\mathcal{T}_{5,A=B}$ is the same as $\mathcal{T}_{5,A=B}^*$ by Theorem 3 and is shown in Figure 2.8. The extreme points $(0, \frac{EF}{E+F})$, $(\frac{CD}{C+D}, 0)$ and $(\frac{AD}{A+D}, \frac{AD}{A+D})$ of $\mathcal{T}_{5,A=B}$ correspond to the rate pairs achievable by schemes τ_1 , τ_2 and τ_3 in Lemma 13 respectively. In addition, for any rate pair in $\mathcal{T}_{5,A=B}$, Lemma 13 provides a procedure to construct a network code that achieves the rate pair. The network code involves linear operations only since τ_1 , τ_2 and τ_3 involve linear operations only. Consequently, the network code is practical due to its low complexity.

Evaluation of $\mathcal{T}_{5,A>B}$

Let $\mathcal{T}_{5,A>B}^*$ denote

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, \\ R_2 \geq 0, \\ R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-BDF}{AD(E+F)} \right), \\ R_2 \leq E - R_1 \left(\frac{AE+DE-BD}{AD} \right), \\ R_2 \leq \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right). \end{array} \right. \right\}, \quad (2.13)$$

Figure 2.9: $\mathcal{T}_{5,A>B}^*$.

which is shown in Figure 2.9, and \mathcal{S}_5 denote

$$\left\{ (R_1, R_2) \in \mathbb{R}^2 \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ p_0 \geq 0, p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, p_4 \geq 0 \\ p_0 + p_1 + p_2 + p_3 + p_4 \leq 1, \\ R_1 \leq \min\{p_0 A + p_1 C, p_3 D\}, \\ R_2 \leq \min\{p_0 B + p_2 E, p_3 D + p_4 F\}. \end{array} \right. \right\}. \quad (2.14)$$

We will show that $\mathcal{T}_{5,A>B} = \mathcal{S}_5 = \mathcal{T}_{5,A>B}^*$.

Lemma 14 $\mathcal{S}_5 \subset \mathcal{T}_{5,A>B}^*$.

Proof: Please refer to Appendix A. ■

Lemma 15 $\mathcal{T}_{5,A>B} \subset \mathcal{S}_5 \subset \mathcal{T}_{5,A>B}^*$.

Proof: Suppose $A > B$. Since $\mathcal{S}_5 \subset \mathcal{T}_{5,A>B}^*$ by Lemma 14, it remains to show that $\mathcal{T}_{5,A>B} \subset \mathcal{S}_5$. Suppose $(R_1, R_2) \in \mathcal{T}_{5,A>B}$. Then, (R_1, R_2) is s -achievable for some allocation $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$. Using Lemma 7, we obtain

$$R_1 \leq \min\{p_0 A + p_1 C, p_3 D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\},$$

which implies that (R_1, R_2) is in \mathcal{S}_5 (cf. (2.14)). \blacksquare

Proposition 16 $\mathcal{T}_{5,A>B}^*$ is the convex hull of $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$, $(\frac{AD}{A+D}, \frac{BD}{A+D})$ and $(\frac{CD}{C+D}, 0)$.

Proof: Please refer to Appendix A. \blacksquare

Lemma 17 $\mathcal{T}_{5,A>B} \supset \mathcal{T}_{5,A>B}^*$.

Proof: We only need to show that the five extreme points of $\mathcal{T}_{5,A>B}^*$ are achievable by some five-session network coding schemes. Then any other point in $\mathcal{T}_{5,A>B}^*$ can be achieved by time sharing of these schemes. Using Lemma 6, we obtain three four-session schemes τ_0 , τ_1 and τ_2 that involve linear operations only such that they achieve the rate pairs $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ respectively. Therefore, $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ are in $\mathcal{T}_{5,A>B}$. We will propose two transmission schemes τ_3 and τ_4 and show that they achieve the rate pairs $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ and $(\frac{AD}{A+D}, \frac{BD}{A+D})$ respectively.

To facilitate understanding, the allocations $[p_0 p_1 p_2 p_3 p_4]$ of τ_3 and τ_4 and their corresponding p_0A , p_0B , p_1C , p_2E , p_3D and p_4F are displayed in Table 2.4. Under schemes τ_3 and τ_4 , the average transmission rates of (t_1, r) in the first session, (t_2, r) in the first session, (t_1, r) in the second session, (t_2, r) in the third session, (r, t_1) in the fourth session, (r, t_2) in the fourth session and (r, t_1) in the fifth session are p_0A , p_0B , p_1C , p_2E , p_3D , p_3D and p_4F respectively.

Under scheme τ_3 , the average rates of the messages sent from node t_1 to node r , from node r to node t_2 , from node t_2 to node r

τ_3				τ_4			
p_0	$DE/(AD+AE+DE-BD)$	p_0A	$ADE/(AD+AE+DE-BD)$	p_0	$D/(A+D)$	p_0A	$AD/(A+D)$
		p_0B	$BDE/(AD+AE+DE-BD)$			p_0B	$AB/(A+D)$
p_1	0	p_1C	0	p_1	0	p_1C	0
p_2	$(A-B)D/(AD+AE+DE-BD)$	p_2E	$(A-B)DE/(AD+AE+DE-BD)$	p_2	0	p_2E	0
p_3	$AE/(AD+AE+DE-BD)$	p_3D	$ADE/(AD+AE+DE-BD)$	p_3	$A/(A+D)$	p_3D	$AD/(A+D)$
p_4	0	p_4F	0	p_4	0	p_4F	0

Table 2.4: The allocations of τ_3 and τ_4 and their corresponding p_0A , p_0B , p_1C , p_2E , p_3D and p_4F .

and from node r to node t_1 are all $ADE/(AD + AE + DE - BD)$. Node t_1 transmits its messages to r in the first session and node t_2 transmits its messages to r in the first and third session. In the fourth session, relay r performs XOR operations between the messages from node t_1 and the messages from node t_2 bit by bit and broadcasts the resultant messages to both node t_1 and node t_2 . In addition, node t_1 can recover the messages of node t_2 by performing XOR operations between its own messages and the messages from r bit by bit. Similarly, node t_2 can recover the messages of node t_1 by performing XOR operations between its own messages and the messages from r bit by bit. Consequently, the rate pair $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ is achievable by scheme τ_3 .

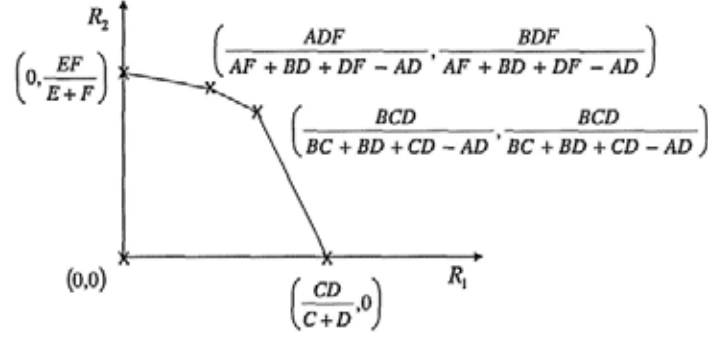
Under scheme τ_4 , the average rates of the messages sent from node t_1 to node r and from node r to node t_2 are both $AD/(A + D)$. The average rates of the messages sent from node t_2 to node r and node r to node t_1 are $BD/(B + D)$ and $AD/(A + D)$ respectively. Since $A > B$, we can split the messages of node t_1 into two messages, denoted by X_1 and X_2 , such that the average rates of X_1 and X_2 are $BD/(B + D)$ and $(A - B)/(B + D)$ respectively. Let Y denote

the messages of node t_2 of rate $BD/(B+D)$. Node t_1 and node t_2 transmit their messages to r in the first session. In the fourth session, relay r performs XOR operations between X_1 from node t_1 and Y from node t_2 bit by bit and broadcasts the resultant messages, denoted by Z , together with X_2 to both node t_1 and node t_2 . In addition, node t_1 can recover Y by performing XOR operations between X_1 , its own messages, and Z from r bit by bit. Similarly, node t_2 can recover X_1 by performing XOR operations between Y , its own messages, and Z from r bit by bit. Consequently, the rate pair $(\frac{AD}{A+D}, \frac{BD}{A+D})$ is achievable by scheme τ_4 . ■

Theorem 4 $\mathcal{T}_{5,A>B} = \mathcal{S}_5 = \mathcal{T}_{5,A>B}^*$.

Proof: It follows directly from Lemma 15 and Lemma 17. ■

The achievable rate region $\mathcal{T}_{5,A>B}$ is the same as $\mathcal{T}_{5,A>B}^*$ by Theorem 4 and is shown in Figure 2.9. The extreme points $(0, \frac{EF}{E+F})$, $(\frac{CD}{C+D}, 0)$, $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ and $(\frac{AD}{A+D}, \frac{BD}{A+D})$ of $\mathcal{T}_{5,A>B}$ correspond to the rate pairs achievable by schemes τ_1 , τ_2 , τ_3 and τ_4 in Lemma 17 respectively. In addition, for any rate pair in $\mathcal{T}_{5,A>B}$, Lemma 17 provides a procedure to construct a network code that achieves the rate pair. The network code involves linear operations only since τ_1 , τ_2 , τ_3 and τ_4 involve linear operations only. Consequently, the network code is practical due to its low complexity.

Figure 2.10: $\mathcal{T}_{5,A<B}^*$.Evaluation of $\mathcal{T}_{5,A<B}$

Let $\mathcal{T}_{5,A<B}^*$ denote

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, \\ R_2 \geq 0, \\ R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-BDF}{AD(E+F)} \right), \\ R_2 \leq \frac{BCF}{BC+CF-AF} - R_1 \left(\frac{BCF+BDF-BCD}{D(BC+CF-AF)} \right), \\ R_2 \leq \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right). \end{array} \right. \right\}, \quad (2.15)$$

which is shown in Figure 2.10. We will show that $\mathcal{T}_{5,A<B} = \mathcal{S}_5 = \mathcal{T}_{5,A<B}^*$.

Lemma 18 $\mathcal{S}_5 \subset \mathcal{T}_{5,A<B}^*$.

Proof: Please refer to Appendix A. ■

Lemma 19 $\mathcal{T}_{5,A<B} \subset \mathcal{S}_5 \subset \mathcal{T}_{5,A<B}^*$.

Proof: Suppose $A < B$. Since $\mathcal{S}_5 \subset \mathcal{T}_{5,A<B}^*$ by Lemma 18, it remains to show that $\mathcal{T}_{5,A<B} \subset \mathcal{S}_5$. Suppose $(R_1, R_2) \in \mathcal{T}_{5,A<B}$. Then,

(R_1, R_2) is s -achievable for some allocation $s = [p_0 p_1 p_2 p_3 p_4]$.

Using Lemma 7, we obtain

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\},$$

which implies that (R_1, R_2) is in \mathcal{S}_5 (cf. (2.14)). \blacksquare

Proposition 20 $\mathcal{T}_{5,A<B}^*$ is the convex hull of $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$, $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ and $(\frac{CD}{C+D}, 0)$.

Proof: Please refer to Appendix A. \blacksquare

Lemma 21 $\mathcal{T}_{5,A<B} \supset \mathcal{T}_{5,A<B}^*$.

Proof: We only need to show that the five extreme points of $\mathcal{T}_{5,A<B}^*$ are achievable by some five-session network coding schemes. Then any other point in $\mathcal{T}_{5,A<B}^*$ can be achieved by time sharing of these schemes. Using Lemma 6, we obtain three four-session schemes τ_0 , τ_1 and τ_2 that involve linear operations only such that they achieve the rate pairs $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ respectively. Therefore, $(0, 0)$, $(0, \frac{EF}{E+F})$ and $(\frac{CD}{C+D}, 0)$ are in $\mathcal{T}_{5,A<B}$. We will propose two transmission schemes τ_3 and τ_4 and show that they achieve the rate pairs $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$ and $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ respectively.

To facilitate understanding, the allocations $[p_0 p_1 p_2 p_3 p_4]$ of τ_3 and τ_4 and their corresponding p_0A , p_0B , p_1C , p_2E , p_3D and p_4F are displayed in Table 2.5. Under schemes τ_3 and τ_4 , the average

τ_3				τ_4			
p_0	$DF/(AF+BD+DF-AD)$	p_0A	$ADF/(AF+BD+DF-AD)$	p_0	$CD/(BC+BD+CD-AD)$	p_0A	$ACD/(BC+BD+CD-AD)$
		p_0B	$BDF/(AF+BD+DF-AD)$			p_0B	$BCD/(BC+BD+CD-AD)$
p_1	0	p_1C	0	p_1	$D(B-A)/(BC+BD+CD-AD)$	p_1C	$CD(B-A)/(BC+BD+CD-AD)$
p_2	0	p_2E	0	p_2	0	p_2E	0
p_3	$AF/(AF+BD+DF-AD)$	p_3D	$ADF/(AF+BD+DF-AD)$	p_3	$BC/(BC+BD+CD-AD)$	p_3D	$BCD/(BC+BD+CD-AD)$
p_4	$(B-A)D/(AF+BD+DF-AD)$	p_4F	$(B-A)DF/(AF+BD+DF-AD)$	p_4	0	p_4F	0

Table 2.5: The allocations of τ_3 and τ_4 and their corresponding p_0A , p_0B , p_1C , p_2E , p_3D and p_4F .

transmission rates of (t_1, r) in the first session, (t_2, r) in the first session, (t_1, r) in the second session, (t_2, r) in the third session, (r, t_1) in the fourth session, (r, t_2) in the fourth session and (r, t_1) in the fifth session are p_0A , p_0B , p_1C , p_2E , p_3D , p_3D and p_4F respectively.

Under scheme τ_3 , the average rates of the messages sent from node t_1 to node r and from node r to node t_2 are both $ADF/(AF + BD + DF - AD)$. The average rates of the messages sent from node t_2 to node r and from node r to node t_1 are both $BDF/(AF + BD + DF - AD)$. Since $B > A$, we can split the messages of node t_2 into two messages, denoted by Y_1 and Y_2 , such that the average rates of Y_1 and Y_2 are $ADF/(AF + BD + DF - AD)$ and $(B - A)DF/(AF + BD + DF - AD)$ respectively. Let X denote the messages of node t_1 of rate $ADF/(AF + BD + DF - AD)$. Node t_1 and node t_2 transmit their messages to r in the first session. In the fourth session, relay r performs XOR operations between X from node t_1 and Y_1 from node t_2 bit by bit and broadcasts the resultant messages, denoted by Z , to both node t_1 and node t_2 . In the fifth session, relay r forwards Y_2 to node t_1 . In addition, node t_1 can recover Y_1 by performing XOR operations between X , its own messages, and Z bit by bit. Similarly, node t_2 can recover X by performing XOR operations between Y_1 , its own messages, and Z bit by

bit. Consequently, the rate pair $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$ is achievable by scheme τ_3 .

Under scheme τ_4 , the average rates of the messages sent from node t_1 to node r , from node r to node t_2 , from node t_2 to node r and from node r to node t_1 are all $BCD/(BC + BD + CD - AD)$. Node t_1 transmits its messages to r in the first and second sessions. Node t_2 transmits its messages to r in the first session. In the fourth session, relay r performs XOR operations between the messages from node t_1 and the messages from node t_2 bit by bit and broadcasts the resultant messages to both node t_1 and node t_2 . In addition, node t_1 can recover the messages of node t_2 by performing XOR operations between its own messages and the messages from r bit by bit. Similarly, node t_2 can recover the messages of node t_1 by performing XOR operations between its own messages and the messages from r bit by bit. Consequently, the rate pair $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ is achievable by scheme τ_4 . ■

Theorem 5 $\mathcal{T}_{5,A<B} = \mathcal{S}_5 = \mathcal{T}_{5,A<B}^*$.

Proof: It follows directly from Lemma 19 and Lemma 21. ■

The achievable rate region $\mathcal{T}_{5,A<B}$ is the same as $\mathcal{T}_{5,A<B}^*$ by Theorem 5 and is shown in Figure 2.10. The extreme points $(0, \frac{EF}{E+F})$, $(\frac{CD}{C+D}, 0)$, $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$ and $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ of $\mathcal{T}_{5,A<B}$ correspond to the rate pairs achievable by schemes τ_1 , τ_2 , τ_3 and τ_4 in Lemma 21 respectively. In addition, for any rate pair in $\mathcal{T}_{5,A<B}$, Lemma 21 provides a procedure to construct a network code that achieves the rate pair. The network code involves linear operations only since τ_1 , τ_2 , τ_3 and τ_4 involve linear

operations only. Consequently, the network code is practical due to its low complexity.

2.3 Insufficiency of Four-Session Network Coding Scheme

Since the four-session schemes do not allow simultaneous transmissions of the two terminal nodes in the three-node point-to-point relay network, they are simpler implementations than the five-session schemes. Therefore, we investigate the criterion that determines the sufficiency of the four-session schemes for achieving all the rate pairs in \mathcal{T}_5 . Under the condition $\frac{A}{C} + \frac{B}{E} \leq 1$, the four-session achievable rate region \mathcal{T}_4 is equal to \mathcal{T}_5 by Theorem 2, which implies that the four-session schemes are sufficient for achieving all the rate pairs in \mathcal{T}_5 . We will show that \mathcal{T}_4 is strictly smaller than \mathcal{T}_5 under the condition $\frac{A}{C} + \frac{B}{E} > 1$ in the following theorem. Recall that $\mathcal{T}_{5,A=B}$, $\mathcal{T}_{5,A<B}$ and $\mathcal{T}_{5,A>B}$ denote the region \mathcal{T}_5 under the cases $A = B$, $A < B$ and $A > B$ respectively with the assumption that $\frac{A}{C} + \frac{B}{E} > 1$.

Theorem 6 $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A=B}$, $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A>B}$ and $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A<B}$.

Proof: Assume $\frac{A}{C} + \frac{B}{E} > 1$. Since a five-session scheme is more general than a four-session scheme, it follows from Definition 6 and Definition 10 that

$$\mathcal{T}_4 \subset \mathcal{T}_5$$

in general, which implies that

$$\mathcal{T}_4 \subset \mathcal{T}_{5,A=B},$$

$$\mathcal{T}_4 \subset \mathcal{T}_{5,A>B}$$

and

$$\mathcal{T}_4 \subset \mathcal{T}_{5,A < B}.$$

In the following, we show that \mathcal{T}_4 is a proper subset of each of $\mathcal{T}_{5,A=B}$, $\mathcal{T}_{5,A > B}$ and $\mathcal{T}_{5,A < B}$. The maximum equal-rate throughput of $\mathcal{T}_{5,A=B}$ and \mathcal{T}_4 are $\frac{AD}{A+D}$ and $\frac{CDE}{CD+CE+DE}$ respectively. Since

$$\begin{aligned} & \frac{AD}{A+D} - \frac{CDE}{CD+CE+DE} \\ &= \frac{D(ACD + ACE + ADE - ACE - CDE)}{(A+D)(CD+CE+DE)} \\ &= \frac{D(ACD + ADE - CDE)}{(A+D)(CD+CE+DE)} \\ &= \frac{CD^2E(A/C + A/E - 1)}{(A+D)(CD+CE+DE)} \\ &> 0, \end{aligned}$$

it then follows that the rate pair $(\frac{AD}{A+D}, \frac{AD}{A+D})$ is outside \mathcal{T}_4 . Thus $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A=B}$.

The maximum equal-rate throughput of $\mathcal{T}_{5,A > B}$ and \mathcal{T}_4 are $\frac{ADE}{AD+AE+DE-BD}$ and $\frac{CDE}{CD+CE+DE}$ respectively. Since

$$\begin{aligned} & \frac{ADE}{AD+AE+DE-BD} - \frac{CDE}{CD+CE+DE} \\ &= \frac{DE(ACD + ACE + ADE - ACD - ACE - CDE + BCD)}{(AD+AE+DE-BD)(CD+CE+DE)} \\ &= \frac{DE(ADE + BCD - CDE)}{(AD+AE+DE-BD)(CD+CE+DE)} \\ &= \frac{CD^2E^2(A/C + B/E - 1)}{(AD+AE+DE-BD)(CD+CE+DE)} \\ &> 0, \end{aligned}$$

it then follows that the rate pair $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ is outside \mathcal{T}_4 . Thus $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A > B}$.

The maximum equal-rate throughput of $\mathcal{T}_{5,A<B}$ and \mathcal{T}_4 are $\frac{BCD}{BC+BD+CD-AD}$ and $\frac{CDE}{CD+CE+DE}$ respectively. Since

$$\begin{aligned} & \frac{BCD}{BC+BD+CD-AD} - \frac{CDE}{CD+CE+DE} \\ &= \frac{CD(BCD+BCE+BDE-BCE-BDE-CDE+ADE)}{(BC+BD+CD-AD)(CD+CE+DE)} \\ &= \frac{CD(ADE+BCD-CDE)}{(BC+BD+CD-AD)(CD+CE+DE)} \\ &= \frac{C^2D^2E(A/C+B/E-1)}{(BC+BD+CD-AD)(CD+CE+DE)} \\ &> 0, \end{aligned}$$

it then follows that the rate pair $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ is outside \mathcal{T}_4 . Thus $\mathcal{T}_4 \subsetneq \mathcal{T}_{5,A<B}$. ■

Theorem 7 *The rate region achievable by five-session schemes is strictly larger than the rate region achievable by four-session schemes if and only if $\frac{A}{C} + \frac{B}{E} > 1$.*

Proof: It follows directly from Theorem 2 and Theorem 6. ■

In the rest of this section, we use Theorem 7 to show that for a five-session scheme, the Type 0 transmission is redundant if the noise power is low, and the Type 0 transmission is useful if the noise power is high. Consider a three-node point-to-point relay network such that

- the received power at r from node t_1 is $P_1 > 0$ when node t_1 transmits;
- the received power at r from node t_2 is $P_2 > 0$ when node t_2 transmits;
- the received background noise power at r is $N_r > 0$.

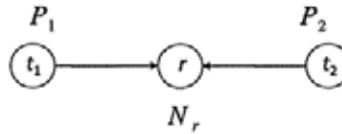


Figure 2.11: A three-node point-to-point relay network.

This is illustrated in Figure 2.11.

For a fixed W , let

$$c(\sigma) = W \log_2(1 + \sigma)$$

be the channel capacity (in bits per second) of a bandlimited Gaussian channel where W is the bandwidth available for transmission and σ is the signal-to-noise ratio of the channel. We assume that the interference incurred on one node by another node behaves as a Gaussian noise. Recall that C and D denote the channel capacities of (t_1, r) and (t_2, r) respectively, and A and B denote the relay-access capacities of (t_1, r) and (t_2, r) respectively. Since the received signal power at r and the noise power at r are P_1 and N_r respectively when only node t_1 transmits, it follows that the corresponding signal-to-noise ratio of (t_1, r) is P_1/N_r , which implies that the channel capacity of (t_1, r) , denoted by C , is equal to $c(P_1/N_r)$. Similarly, the channel capacity of (t_2, r) , denoted by E , is equal to $c(P_2/N_r)$.

Since the received signal power from node t_1 at r and the noise power at r are P_1 and $N + P_2$ respectively when node t_1 and node t_2 both transmit, it follows that the corresponding signal-to-noise ratio of (t_1, r) is $P_1/(N_r + P_2)$, which implies that the relay-access capacity of (t_1, r) , denoted by A , is equal to $c(P_1/(N_r + P_2))$. Similarly, the relay-access capacity of (t_2, r) , denoted by b , is equal to $c(P_2/(N_r + P_1))$.

Let

$$\begin{aligned} f(N_r) &= \frac{A}{C} + \frac{B}{E} \\ &= \frac{\log_2(1 + P_1/(N_r + P_2))}{\log_2(1 + P_1/N_r)} + \frac{\log_2(1 + P_2/(N_r + P_1))}{\log_2(1 + P_2/N_r)} \\ &= \frac{\ln(1 + P_1/(N_r + P_2))}{\ln(1 + P_1/N_r)} + \frac{\ln(1 + P_2/(N_r + P_1))}{\ln(1 + P_2/N_r)} \end{aligned}$$

be a continuous function on $(0, \infty)$. We will show that $f(N_r)$ is strictly increasing in the following proposition.

Proposition 22 *The function $f(N_r)$ is strictly increasing on $(0, \infty)$.*

Proof: For $a, b > 0$, let

$$g_{a,b}(N_r) = \frac{\ln(1 + a/(N_r + b))}{\ln(1 + a/N_r)}$$

be a continuous function on $(0, \infty)$. Since

$$f(N_r) = g_{P_1, P_2}(N_r) + g_{P_2, P_1}(N_r),$$

it suffices to show that $g_{a,b}(N_r)$ is strictly increasing on $(0, \infty)$.

In the rest of the proof, we will show that

$$g'_{a,b}(N_r) > 0$$

on $(0, \infty)$, which will then imply that $g_{a,b}(N_r)$ is strictly increasing on $(0, \infty)$. Using standard differentiation techniques, we obtain

$$\begin{aligned} g'_{a,b}(N_r) &= \frac{a \ln \left(\frac{(1 + \frac{a}{N_r + b})^{(N_r + b)(N_r + a + b)}}{(1 + \frac{a}{N_r})^{N_r(N_r + a)}} \right)}{N_r(N_r + a)(N_r + b)(N_r + a + b) \left(\ln \left(1 + \frac{a}{N_r} \right) \right)^2}. \quad (2.16) \end{aligned}$$

Since

$$\begin{aligned}
& \ln \left(\frac{\left(1 + \frac{a}{N_r+b}\right)^{(N_r+b)(N_r+a+b)}}{\left(1 + \frac{a}{N_r}\right)^{N_r(N_r+a)}} \right) \\
& > \ln \left(\frac{\left(1 + \frac{a}{N_r+b}\right)^{(N_r+b)(N_r+a)}}{\left(1 + \frac{a}{N_r}\right)^{N_r(N_r+a)}} \right) \\
& = a(N_r + a) \ln \left(\frac{\left(1 + \frac{1}{(N_r/a+b/a)}\right)^{(N_r/a+b/a)}}{\left(1 + \frac{1}{N_r/a}\right)^{N_r/a}} \right) \\
& > a(N_r + a) \ln 1 \\
& = 0
\end{aligned}$$

where the last inequality follows from the fact that $(1 + 1/x)^x$ is strictly increasing on $(0, \infty)$, it then follows from (2.16) that

$$g'_{a,b}(N_r) > 0$$

on $(0, \infty)$, which implies that $g_{a,b}(N_r)$ is strictly increasing on $(0, \infty)$. ■

By Proposition 22, $f(N_r)$ is a strictly increasing continuous function on $(0, \infty)$. In addition,

$$\lim_{N_r \rightarrow 0} f(N_r) = 0$$

and

$$\begin{aligned}
& \lim_{N_r \rightarrow \infty} f(N_r) \\
&= \lim_{N_r \rightarrow \infty} \left(\frac{\ln(1 + P_1/(N_r + P_2))}{\ln(1 + P_1/N_r)} + \frac{\ln(1 + P_2/(N_r + P_1))}{\ln(1 + P_2/N_r)} \right) \\
&\stackrel{(a)}{=} \lim_{N_r \rightarrow \infty} \left(\frac{N_r(N_r + P_1)}{(N_r + P_2)(N_r + P_1 + P_2)} + \frac{N_r(N_r + P_2)}{(N_r + P_1)(N_r + P_1 + P_2)} \right) \\
&= 2,
\end{aligned}$$

where (a) follows from L'Hospital's rule. It then follows from the intermediate value theorem that for any value $c \in (0, 2)$, there exists exactly one $N_r \in (0, \infty)$ such that $f(N_r) = c$. In particular, there exists a unique N^* such that $f(N^*) = 1$. Since $f(N_r)$ is strictly increasing,

$$0 < f(N_r) \leq 1$$

for all $N_r \in (0, N^*]$, which implies from Theorem 2 that the four-session schemes for the three-node point-to-point relay network are sufficient for achieving all the rate pairs in \mathcal{T}_5 . However,

$$1 < f(N_r) < 2$$

for all $N_r \in (N^*, \infty)$, which implies from Theorem 6 that the four-session schemes for the three-node point-to-point relay network are insufficient for achieving all the rate pairs in \mathcal{T}_5 . Roughly speaking, the Type 0 transmission in a five-session scheme is redundant if the noise power is low, and the Type 0 transmission is useful if the noise power is high.

□ End of chapter.

Chapter 3

Performance Analysis in Cellular Relay Network

Summary

We investigate by simulation the benefit of network coding in a cellular relay network consisting of multiple users, multiple relays and multiple base stations. The simulation scenarios and assumptions for the cellular relay network are presented in Section 3.1. We investigate the benefit of network coding under two transmission modes in the network: the three-phase mode and the four-phase mode, which are described in Section 3.2 and Section 3.3 respectively. The four-phase mode subsumes the three-phase mode. Under each transmission mode, we propose several practical network coding schemes on the cellular relay network, and our simulation results show that the use of these network coding schemes can improve the average maximum equal-rate throughput over all users.

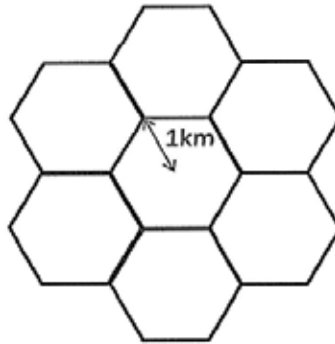


Figure 3.1: Seven hexagonal cells.

In the previous chapter, we propose four-session and five session network coding schemes on the three-node point-to-point relay network and obtain several rate regions achievable by the coding schemes. We show theoretically that the use of network coding rather than routing alone always enlarges the achievable rate region. In particular, the use of network coding increases the maximum equal-rate throughput.

In this chapter, we study practical network coding schemes for a cellular relay network which consists of seven hexagonal cells with 1km radius as shown in Figure 3.1. Each cell contains a base station, multiple users and multiple relays. Since the transmitting nodes in the same cell and different cells can interfere with each other in the cellular relay network, each three-node point-to-point relay network in the cellular relay network experiences more interference than an isolated three-node point-to-point relay network. Consequently, the benefit of network coding in the cellular relay network becomes unclear. This motivates us to investigate the benefit of network coding in such a network.

This chapter is organized as follows. We first present the simulation scenarios and assumptions for the cellular relay network. The simulation scenarios and assumptions in this thesis are established according to Qualcomm's standards and are widely used in industry. Next, we propose several practical symbol-level network coding schemes for the cellular relay network and investigate by MATLAB simulation their performance, which cannot be analyzed theoretically due to the large number of nodes. Finally, we present our simulation results, which show that the use of these network coding schemes can improve the average maximum equal-rate throughput over all users.

3.1 Simulation Assumptions

3.1.1 Deployment Model

We consider a cellular relay network consisting of seven hexagonal cells with 1km radius as shown in Figure 3.1.

One base station is situated at the center of each cell. Each cell has m relays, where m is a parameter of the simulation, and these relays are placed into each cell one by one. Each relay is added to the cell according to the uniform distribution over the cell. The position is then checked to see whether the following two constraints are satisfied:

- The minimum distance between the currently added relay and any previously added relay in the same cell is 35m.
- The minimum distance between the currently added relay and the base station in the same cell is 200m.

If so, the relay is accepted to the cell. Otherwise, repeat the process

until the two constraints above are satisfied. After all the relays are placed into the seven cells, we place n users into each cell, where n is a parameter of the simulation, one by one. Each user is added to the cell according to the uniform distribution over the cell. The position is then checked to see whether the following two constraints are satisfied:

- The minimum distance between the currently added user and any relay in the same cell is 35m.
- The minimum distance between the currently added user and the base station in the same cell is 35m.

If so, the user is accepted to the cell. Otherwise, repeat the process until the two constraints above are satisfied. The antennas of each base station, each relay and each user are omni-directional.

We only investigate the throughput for the users in the central cell. The nodes in other cells are used to simulate the interference experienced by the nodes in the central cell. We will vary n and m in order to investigate the benefit of network coding in the network with different numbers of users and relays.

3.1.2 Power and Bandwidth Settings

Since this thesis serves as a baseline study for cellular relay networks, we do not consider any power control at the nodes in the network. Every node in the network transmits at full power when it transmits information. The full power of each user, base station and relay are 23dBm (0.2W), 43dBm (20W) and 30dBm (1W) respectively. The one-sided power spectral density of the thermal noise is -174dBm/Hz ($10^{-17.4}$ mW/Hz). Every transmission uses the same frequency spectrum of bandwidth 10MHz.

3.1.3 Propagation Model

The propagation model used in the simulation is the COST-Hata-Model in [21], which states that the loss in dB for any signal propagation between a base station and a user separated from each other by a distance of d kilometers is

$$46.3 + 33.9 \log_{10} f - 13.82 \log_{10} h_b - a(h_u) \\ + (44.9 - 6.55 \log_{10} h_b) \log_{10} d \quad (3.1)$$

with

$$a(h_u) = (1.1 \log_{10} f - 0.7)h_u - (1.56 \log_{10} f - 0.8),$$

where f is the carrier frequency in MHz, and h_b and h_u are the antenna heights of the base station and the user in meter respectively.

We assume that the carrier frequency is 2000MHz and the antenna heights of a base station, a relay and a user are 30 meters, 10 meters and 1 meter respectively. In the context of the COST-Hata-Model, each relay has two different roles in the communication process. When it communicates with the base station, it plays the role of a “user”. When it communicates with a user, it plays the role of a “base station”. For the signal propagation loss between a relay and a base station or between a relay and a user, the formula (3.1) is applied for calculating the propagation loss according to the role of the relay. After some calculations, we obtain the following five types of propagation loss:

1. The power loss in dB for any signal propagation between a base station and a user separated from each other by a distance of d meters is $33.5 + 35.2 \log_{10}(d)$;

2. The power loss in dB for any signal propagation between a relay and a user separated from each other by a distance of d meters is $30.8 + 38.4 \log_{10}(d)$;
3. The power loss in dB for any signal propagation between a base station and a relay separated from each other by a distance of d meters is $7.2 + 35.2 \log_{10}(d)$;
4. The power loss in dB for any signal propagation between two base stations separated from each other by a distance of d meters is $-51.5 + 35.2 \log_{10}(d)$;
5. The power loss in dB for any signal propagation between two users separated from each other by a distance of d meters is $24.9 + 44.9 \log_{10}(d)$.

3.1.4 Single-Hop and Multi-Hop Users

Before any data transmission occurs, we introduce a classification stage that divides the users into two types: *single-hop users* and *multi-hop users*. The single-hop users communicate with the base station directly and the multi-hop users communicate with the base station through a relay.

The classification stage is described as follows. Every base station and relay in the network simultaneously transmits an arbitrary message. During this time, every user u keeps silent and computes the signal-to-noise plus interference ratio (SNIR) of channel (v, u) for every transmitting node v in the cell containing u . Then user u chooses to communicate with the local base station b either directly or through a local relay based on the SNIRs. If (b, u) has the largest SNIR, user u will choose to communicate with the base station di-

rectly. If (r, u) has the largest SNIR for a local relay r , user u will choose to communicate with the base station through relay r . The users which communicate with the base station directly are called single-hop users and the users which communicate with the base station through a relay are called multi-hop users. The relays which serve at least one user are said to be *active*.

3.1.5 Channel Model

We model the cellular relay network as a collection of two-node point-to-point systems and three-node point-to-point relay networks in such a way that each single-hop user and its associated base station are modeled as a two-node point-to-point system, and each multi-hop user, its associated relay and its associated base station are modeled as a three-node point-to-point relay network. Each two-node point-to-point system consists of two independent bandlimited Gaussian channels as shown in Figure 3.2. Each three-node point-to-point relay network is the same as the three-node point-to-point relay network defined at the beginning of Chapter 2 and it is shown in Figure 3.3. We assume that every single-hop user exchanges independent messages with the base station through the corresponding two-node point-to-point system and every multi-hop user exchanges independent messages with the base station through the corresponding three-node point-to-point relay network.

Let (u, v) denote a point-to-point channel from node u to node v in the cellular relay network, where (u, v) is either a bandlimited Gaussian channel in a two-node point-to-point system or a bandlimited Gaussian channel in a three-node point-to-point relay network. We assume that the interference incurred on one node by other

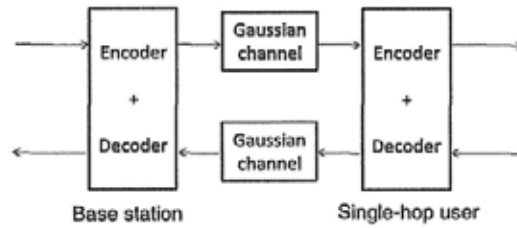


Figure 3.2: A two-node point-to-point system.

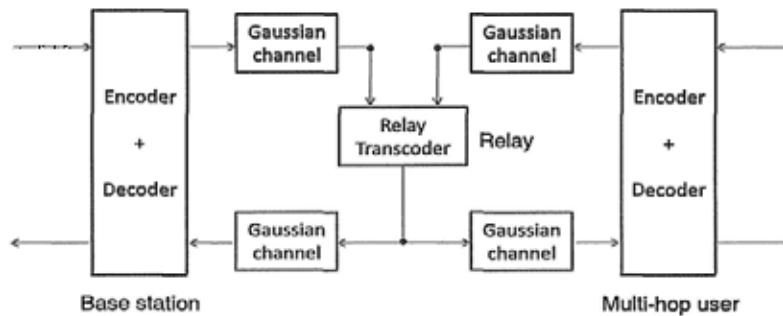


Figure 3.3: A three-node point-to-point relay network.

nodes behaves as independent Gaussian noise. In reality, the signal-to-noise plus interference ratio (SNIR) of a communication channel is upper bounded due to hardware constraints. Therefore, as a conservative assumption, we impose an upper bound on the SNIR as follows: If the calculated SNIR of (u, v) is greater than 20dB, the SNIR is set to 20dB. Since the channel capacity of a bandlimited Gaussian channel is

$$W \log_2(1 + \sigma) \text{ bits per second}$$

where W is the bandwidth available for transmission and σ is the signal-to-noise ratio at the receiver, the capacity of (u, v) is assumed to be

$$W \log_2(1 + \min\{\sigma, 100\}) \text{ bits per second}$$

where σ is the calculated SNIR of (u, v) and $\min\{\sigma, 100\}$ is the SNIR of (u, v) after an upper bound of 20dB has been imposed. Since every transmission uses a frequency spectrum of bandwidth 10MHz and the one-sided power spectral density of the thermal noise is $10^{-17.4}$ mW/Hz it follows that

$$W = 10^7 \text{Hz}$$

and

$$\sigma = \frac{\text{received signal power at } v \text{ from } u}{10^{-20.4}W + \text{interference power at } v \text{ from nodes other than } u}.$$

3.1.6 Fading Model

For simplicity, we do not consider fast fading in the network. Slow fading in the network is modeled by a static factor in such a way that for any communication channel between two nodes involving a user, the SNIR of the channel is decreased by a factor of 2dB.

3.2 Three-Phase Mode

In this section, we propose a transmission mode for the cellular relay network called the *three-phase mode*, under which the transmissions in the network are organized into three phases. We will investigate by simulation the average equal-rate throughput over all users under two scenarios: each user uses an optimal network coding scheme and each user uses an optimal routing scheme.

3.2.1 Resource and Interference Management

To facilitate discussion, we classify the transmissions in each two-node point-to-point system as follows:

Class 1: The base station transmits messages to the single-hop user.

Class 2: The single-hop user transmits messages to the base station.

Similarly, we classify the transmissions in each three-node point-to-point relay network as follows:

Class 1: The base station transmits messages to the relay.

Class 2: The multi-hop user transmits messages to the relay.

Class 3: The relay transmits messages to the base station and the multi-hop user.

The transmissions in each of the seven cells are organized into three phases, denoted by Phase 1, Phase 2 and Phase 3. The transmissions in the seven cells are synchronized in such a way that the seven cells have the same periods for Phase 1, Phase 2 and Phase 3. The fractions of time allocated for the three phases are fixed and we let p_1 , p_2 and p_r denote the fractions of time allocated for Phase 1, Phase 2 and Phase 3 transmissions respectively. We call (p_1, p_2, p_r) a *three-phase tuple*. The transmissions in the three phases are described as follows.

Phase 1: In each cell, the phase is divided into n equal-length time slots which are allocated to the n users in the cell. In each time slot, if the user is a single-hop user, the communication from the base station to the user is conducted over the corresponding

two-node system by Class 1 transmissions. If the user is a multi-hop user, the communication from the base station to the user through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 2: In each cell, all the users simultaneously transmit messages throughout the entire phase. For each single-hop user, the communication from the user to the base station is conducted over the corresponding two-node system by Class 2 transmissions. For each multi-hop user, the communication from the user to the base station through its associated relay is conducted over the corresponding three-node network by Class 2 transmissions.

Phase 3: In each cell, all the active relays simultaneously transmit throughout the entire phase. For each multi-hop user u , the communication between u and the base station b through its associated relay r is as follows: Let $k \geq 1$ denote the number of multi-hop users that choose to communicate with b through r and let t_1, t_2, \dots, t_{k-1} denote the users other than u that communicate with b through r . This is illustrated in Figure 3.4. The phase is divided into k equal-length time slots, which are allocated to the k users served by relay r . Note that the value of k depends on the relay. In the time slot allocated to user u , the communication between u and b through r is conducted over the corresponding three-node network by Class 3 transmissions.

In each of the phases above, each listening station in a two-node system or a three-node network regards unintended signals as independent Gaussian noise.

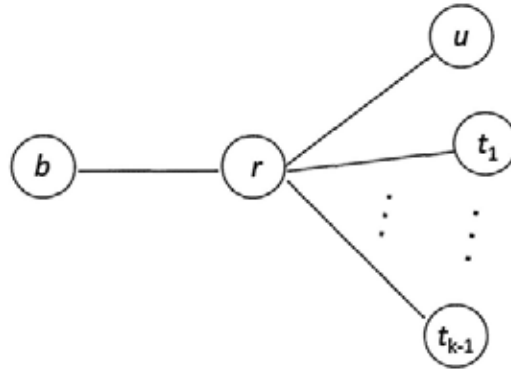


Figure 3.4: A relay r which serves k multi-hop users.

Given a three-phase tuple (p_1, p_2, p_r) , the time allocations for each two-node system and each three-node network in Phase 1 are as follows:

1. For each two-node system, the fraction of time allocated for Class 1 transmissions is p_1/n .
2. For each three-node network, the fraction of time allocated for Class 1 transmissions is p_1/n .

The time allocations for each two-node system and each three-node network in Phase 2 are as follows:

1. For each two-node system, the fraction of time allocated for Class 2 transmissions is p_2 .
2. For each three-node network, the fraction of time allocated for Class 2 transmissions is p_2 .

The time allocations for each two-node system and each three-node network in Phase 3 are as follows:

1. Each two-node system is idle.

	Two-node system	Three-node network
Phase 1	p_1/n	p_1/n
Phase 2	p_2	p_2
Phase 3	0	p_r/k

Table 3.1: The fractions of time allocated to each two-node point-to-point system and each three-node point-to-point network under the three-phase mode.

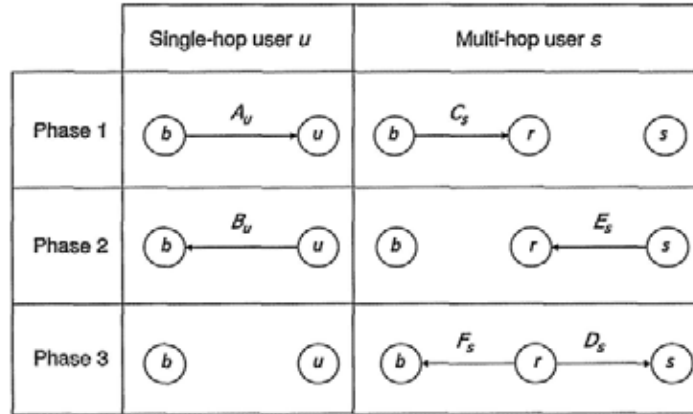


Table 3.2: The capacities of a two-node point-to-point system and a three-node point-to-point network under the three-phase mode.

- For each three-node network, the fraction of time allocated for Class 3 transmissions is p_r/k , where $k \geq 1$ is the number of multi-hop users served by the relay.

To facilitate understanding, the fractions of time allocated to each two-node point-to-point system as well as each three-node point-to-point network for Phase 1, Phase 2 and Phase 3 transmissions are shown in Table 3.1.

For each single-hop user u and the base station b , let A_u denote the capacity of (b, u) in Phase 1 and B_u denote the capacity of (u, b) in Phase 2. This is illustrated on the left side of Table 3.2.

For each multi-hop user s which communicates with the base station b through its associated relay r , let C_s , D_s , E_s and F_s denote the capacities of (b, r) in Phase 1, (r, s) in Phase 3, (s, r) in Phase 2 and (r, b) in Phase 3 respectively. This is illustrated on the right side of Table 3.2. As discussed at the beginning of Chapter 2, relay r can broadcast information reliably from r to both node b and node s at rate R in Phase 3 if and only if $R \leq \min\{D_s, F_s\}$. We call $\min\{D_s, F_s\}$ the *broadcast capacity*.

3.2.2 Maximum Equal-Rate Throughput

Fix a three-phase tuple (p_1, p_2, p_r) . For every user, we are interested in the maximum equal-rate throughput achievable by network coding schemes. For each single-hop user u , the fractions p_1/n and p_2 of the time are allocated to the corresponding two-node point-to-point system for Phase 1 and Phase 2 transmissions respectively. Therefore, the maximum throughput from the base station to u is $p_1 A_u/n$ and the maximum throughput from u to the base station is $p_2 B_u$, which implies that the maximum equal-rate throughput for u is $\min\{p_1 A_u/n, p_2 B_u\}$, which is trivially achievable by some routing scheme.

For each multi-hop user s which communicates with the base station through its associated relay r , we will prove an upper bound on the equal-rate throughput of the corresponding three-node point-to-point relay network. We will also present a network coding scheme that achieves this upper bound. We call this upper bound the *maximum equal-rate throughput for s* and it is shown in Table 3.3, where $k \geq 1$ is the number of multi-hop users served by r .

The techniques involved in proving the bound in Table 3.3 are

Case	Maximum equal-rate throughput of s
$D_s \leq F_s$	$\min\{p_1 C_s / n, p_2 E_s, p_r D_s / k\}$
$D_s > F_s$	$\min\{p_1 C_s / n, p_2 E_s, p_r F_s / k\}$

Table 3.3: The maximum equal-rate throughput for s .

similar to the ones used in the proof of Lemma 1 in Section 2.1.2. The details are as follows. Fix a multi-hop user s and its corresponding three-node point-to-point relay network. Let b and r denote the base station and the relay respectively in the three-node network. Consider the following two cases:

Case 1: $D_s \leq F_s$

The broadcast capacity of the three-node network is $\min\{D_s, F_s\} = D_s$. Therefore, r can broadcast messages reliably to both the base station and s at a rate less than or equal to D_s . Without loss of generality, we assume that only one of the following four types of transmissions can take place in a transmission scheme.

Type 1: The base station transmits messages reliably to r at a rate less than or equal to C_s .

Type 2: Multi-hop user s transmits messages reliably to r at a rate less than or equal to E_s .

Type 3: Relay r broadcasts messages reliably to both the base station and s at a rate less than or equal to D_s .

Type 4: Relay r transmits messages reliably to the base station at a rate less than or equal to F_s but strictly greater than D_s .

In addition, we assume that Phase 1 consists of Type 1 transmissions, Phase 2 consists of Type 2 transmissions, and Phase 3

consists of Type 3 and Type 4 transmissions. Since the fractions of time allocated to the three-node network for Phase 1 and Phase 2 transmissions are p_1/n and p_2 respectively, the fractions of time allocated for Type 1 and Type 2 transmissions are p_1/n and p_2 respectively. Let $\alpha_s \geq 0$ and $\beta_s \geq 0$ denote the fractions of time allocated for Type 3 and Type 4 transmissions respectively. Since the fraction of time allocated to the three-node network for Phase 3 transmissions is p_r/k , we assume that

$$\alpha_s + \beta_s = p_r/k.$$

For any fixed α_s , the highest average rates for Type 1, 2, 3 and 4 transmissions are p_1C_s/n , p_2E_s , α_sD_s and $(p_r/k - \alpha_s)F_s$ respectively. Therefore, the equal-rate throughput, denoted by R , must satisfy the rate constraints induced by the two-source point-to-point network in Figure 3.5(a), which are equivalent to the rate constraints induced by the single-source point-to-point multicast network in Figure 3.5(b) (cf. the proof of Lemma 1 in Section 2.1.2). Applying the cut-set bound for four different cuts as in the proof of Lemma 1, we obtain

$$R \leq p_1C_s/n,$$

$$R \leq \alpha_sD_s,$$

$$R \leq p_2E_s$$

and

$$R \leq \alpha_sD_s + (p_r/k - \alpha_s)F_s.$$

Since $0 \leq \alpha_s \leq p_r/k$, it follows that an upper bound on the

equal-rate throughput is

$$\begin{aligned}
& \max_{0 \leq \alpha_s \leq p_r/k} \min\{p_1 C_s/n, p_2 E_s, \alpha_s D_s, \alpha_s D_s + (p_r/k - \alpha_s) F_s\} \\
&= \min\{p_1 C_s/n, p_2 E_s, \max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, \alpha_s D_s + \\
&\quad (p_r/k - \alpha_s) F_s\}\} \\
&= \min\{p_1 C_s/n, p_2 E_s, p_r D_s/k\}, \tag{3.2}
\end{aligned}$$

where the last equality follows from the fact that

$$\max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, \alpha_s D_s + (p_r/k - \alpha_s) F_s\} = p_r D_s/k.$$

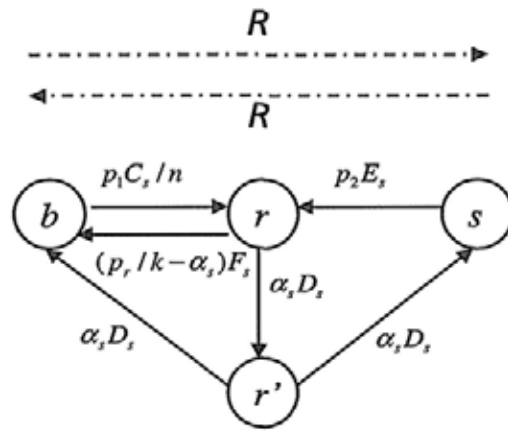
We now construct a network coding scheme for s that achieves the upper bound on the equal-rate throughput in (3.2). A fraction p_1/n of the time is allocated to the corresponding three-node point-to-point relay network for Phase 1 transmissions and during this time, the base station transmits messages to r at rate

$$\min \left\{ C_s, np_2 E_s/p_1, \frac{np_r D_s}{p_1 k} \right\} \leq C_s.$$

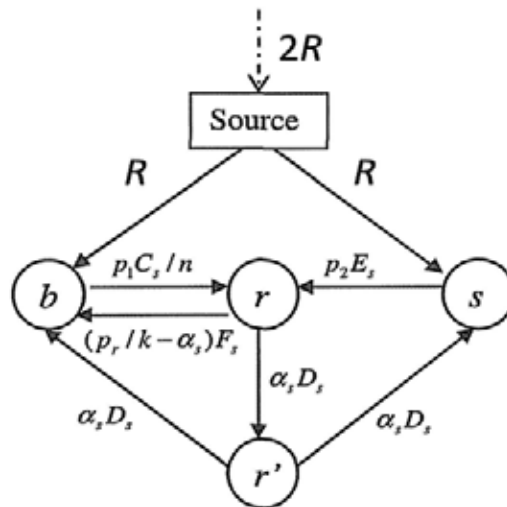
A fraction p_2 of the time is allocated to the three-node network for Phase 2 transmissions and during this time, s transmits messages to r at rate

$$\min \left\{ \frac{p_1 C_s}{p_2 n}, E_s, \frac{p_r D_s}{p_2 k} \right\} \leq E_s.$$

A fraction p_r/k of the time is allocated to the three-node network for Phase 3 transmissions and during this time, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the



(a) A two-source point-to-point network.



(b) A single-source point-to-point multicast network.

Figure 3.5: Two equivalent network coding problems.

resultant messages to both the base station and s at rate

$$\min \left\{ \frac{kp_1C_s}{np_r}, \frac{kp_2E_s}{p_r}, D_s \right\} \leq D_s,$$

where $k \geq 1$ is the number of multi-hop users served by the relay. Under the scheme, the average rates of the messages sent from the base station to the relay, from s to the relay, from the relay to the base station and from the relay to s are all

$$\min\{p_1C_s/n, p_2E_s, p_rD_s/k\}.$$

In addition, the base station can recover the messages of s by performing XOR operations between its own messages and the messages from the relay. Similarly, s can recover the messages of the base station by performing XOR operations between its own messages and the messages from the relay. Consequently, the equal-rate throughput

$$\min\{p_1C_s/n, p_2E_s, p_rD_s/k\}$$

is achievable by the scheme.

Case 2: $D_s > F_s$

The broadcast capacity is $\min\{D_s, F_s\} = F_s$. Following similar procedures in Case 1, we can prove that

$$\min\{p_1C_s/n, p_2E_s, p_rF_s/k\}$$

is an upper bound on the equal-rate throughput and construct a network coding scheme that achieves this upper bound.

Combining the two cases, the maximum equal-rate throughput for s

is

$$\min\{p_1 C_s/n, p_2 E_s, p_r D_s/k, p_r F_s/k\}$$

and we have presented a network coding scheme that achieves this equal-rate throughput. We call this network coding scheme the *XOR scheme*.

3.2.3 Performance of Routing

Fix a three-phase tuple (p_1, p_2, p_r) . For every user, we are interested in the maximum equal-rate throughput achievable by routing schemes. For each single-hop user u , the maximum equal-rate throughput achievable by routing schemes is $\min\{p_1 A_u/n, p_2 B_u\}$ as discussed in the previous subsection.

For each multi-hop user s that uses a routing scheme to communicate with the base station through its associated relay r , let α_s denote the fraction of time allocated for forwarding the messages of the base station from r to s and let β_s denote the fraction of time allocated for forwarding the messages of s from r to the base station during Phase 3. Let $k \geq 1$ denote the number of multi-hop users served by relay r . Since the fractions p_1/n , p_2 and p_r/k of the time are allocated to the corresponding three-node point-to-point relay network for Phase 1, Phase 2 and Phase 3 transmissions respectively, it follows that for every fixed $\alpha_s, \beta_s \geq 0$ that satisfy $\alpha_s + \beta_s \leq p_r/k$, the maximum throughput from the base station to s is $\min\{p_1 C_s/n, \alpha_s D_s\}$ and the maximum throughput from s to the base station is $\min\{p_2 E_s, \beta_s F_s\}$. Therefore, the maximum

equal-rate throughput achievable by routing schemes is

$$\begin{aligned} & \max_{\substack{\alpha_s, \beta_s \geq 0 \\ \alpha_s + \beta_s = p_r/k}} \min\{\min\{p_1 C_s/n, \alpha_s D_s\}, \min\{p_2 E_s, \beta_s F_s\}\} \\ &= \max_{\substack{\alpha_s, \beta_s \geq 0 \\ \alpha_s + \beta_s = p_r/k}} \min\{p_1 C_s/n, \alpha_s D_s, p_2 E_s, \beta_s F_s\}. \end{aligned} \quad (3.3)$$

Consequently, the maximum equal-rate throughput achievable by routing schemes is

$$\begin{aligned} & \max_{0 \leq \alpha_s \leq p_r/k} \min\{p_1 C_s/n, \alpha_s D_s, p_2 E_s, (p_r/k - \alpha_s) F_s\} \\ &= \min\{p_1 C_s/n, p_2 E_s, \max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, (p_r/k - \alpha_s) F_s\}\}. \end{aligned} \quad (3.4)$$

In the rest of this subsection, we will show that

$$\max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, (p_r/k - \alpha_s) F_s\} = \frac{p_r D_s F_s}{k(D_s + F_s)},$$

which will then imply that (3.4) is equal to

$$\min\left\{p_1 C_s/n, p_2 E_s, \frac{p_r D_s F_s}{k(D_s + F_s)}\right\}.$$

Let $\bar{\alpha}_s$ denote

$$\arg \max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, (p_r/k - \alpha_s) F_s\}. \quad (3.5)$$

We claim that

$$\bar{\alpha}_s D_s = (p_r/k - \bar{\alpha}_s) F_s$$

and verify the claim as follows. For any $0 \leq \alpha_s \leq p_r/k$ that satisfies

$$\alpha_s D_s > (p_r/k - \alpha_s) F_s,$$

there exists an $0 \leq \alpha_s^* < \alpha_s$ such that

$$\alpha_s D_s^* = (p_r/k - \alpha_s^*) F_s$$

and

$$\alpha_s D_s > \alpha_s^* D_s = (p_r/k - \alpha_s^*) F_s > (p_r/k - \alpha_s) F_s,$$

which implies that

$$\min\{\alpha_s D_s, (p_r/k - \alpha_s) F_s\} < \min\{\alpha_s^* D_s, (p_r/k - \alpha_s^*) F_s\}. \quad (3.6)$$

Similarly, for any $0 \leq \alpha_s \leq p_r/k$ that satisfies

$$\alpha_s D_s < (p_r/k - \alpha_s) F_s,$$

there exists an $\alpha_s < \alpha_s^* \leq p_r/k$ such that

$$\alpha_s^* D_s = (p_r/k - \alpha_s^*) F_s$$

and

$$\alpha_s D_s < \alpha_s^* D_s = (p_r/k - \alpha_s^*) F_s < (p_r/k - \alpha_s) F_s,$$

which implies that

$$\min\{\alpha_s D_s, (p_r/k + \alpha_s) F_s\} < \min\{\alpha_s^* D_s, (p_r/k - \alpha_s^*) F_s\}. \quad (3.7)$$

It then follows from (3.6) and (3.7) that

$$\bar{\alpha}_s D_s = (p_r/k - \bar{\alpha}_s) F_s$$

(cf. (3.5)), which implies that

$$\bar{\alpha}_s = \frac{p_r}{k(D_s + F_s)}. \quad (3.8)$$

Therefore

$$\begin{aligned} & \max_{0 \leq \alpha_s \leq p_r/k} \min\{\alpha_s D_s, (p_r/k - \alpha_s) F_s\} \\ &= \bar{\alpha}_s D_s \\ &= \frac{p_r D_s}{k(D_s + F_s)}, \end{aligned}$$

which implies from (3.4) that the maximum equal-rate throughput achievable by routing schemes is

$$\min \left\{ p_1 C_s/n, p_2 E_s, \frac{p_r D_s F_s}{k(D_s + F_s)} \right\}. \quad (3.9)$$

We call the routing scheme that achieves the equal-rate throughput (3.9) the *optimal routing scheme*.

3.2.4 Network Coding Gain for Individual User

For each single-hop user u , since u communicates with the base station directly, there is no network coding gain in the maximum equal-rate throughput. The maximum equal-rate throughput for u has been shown in Section 3.2.2 to be

$$\min\{p_1 A_u/n, p_2 B_u\}.$$

For each multi-hop user s that communicates with the base station b through its associated relay r , the equal-rate throughput achievable by the XOR scheme has been shown in Section 3.2.2 to be

$$\min\{p_1 C_s/n, p_2 E_s, p_r D_s/k, p_r F_s/k\}.$$

The equal-rate throughput of the optimal routing scheme for s has

been shown in Section 3.2.3 to be

$$\min \left\{ p_1 C_s/n, p_2 E_s, \frac{p_r D_s F_s}{k(D_s + F_s)} \right\}.$$

We have shown in Section 3.2.2 that the equal-rate throughput achievable by the XOR scheme is an upper bound on the equal-rate throughput achievable by network coding schemes, including the optimal routing scheme as a special case. Therefore,

$$\begin{aligned} & \min \left\{ p_1 C_s/n, p_2 E_s, \frac{p_r D_s F_s}{k(D_s + F_s)} \right\} \\ & \leq \min \{ p_1 C_s/n, p_2 E_s, p_r D_s/k, p_r F_s/k \}. \end{aligned} \quad (3.10)$$

Since

$$\frac{p_r D_s F_s}{k(D_s + F_s)} < p_r D_s/k$$

and

$$\frac{p_r D_s F_s}{k(D_s + F_s)} < p_r F_s/k,$$

it follows that equality holds in (3.10) if and only if

$$\min \{ p_1 C_s/n, p_2 E_s \} \leq \frac{p_r D_s F_s}{k(D_s + F_s)}.$$

In other words, the network coding gain in the maximum equal-rate throughput for s is none if and only if either $p_1 C_s$, the average capacity of (b, r) , or $p_2 E_s$, the average capacity of (s, r) is the bottleneck of the two-way communication. Equivalently, there is network coding gain in the maximum equal-rate throughput for s if and only if the average capacities of both channels (b, r) and (s, r) are not the bottleneck of the two-way communication.

3.2.5 Average Network Coding Gain

The network coding gain for each user in the cellular relay network has been discussed in the previous subsection. In this subsection, we consider the average equal-rate throughput over all users under the following two scenarios:

Scenario 1: Each single-hop user u in the cellular relay network communicates with the base station directly with equal-rate throughput $\min\{p_1 A_u/n, p_2 B_u\}$. Each multi-hop user s in the network uses the optimal routing scheme with the equal-rate throughput $\min\left\{p_1 C_s/n, p_2 E_s, \frac{p_r D_s F_s}{k(D_s + F_s)}\right\}$ (cf. Section 3.2.3).

Scenario 2: Each single-hop user u in the cellular relay network communicates with the base station directly with equal-rate throughput $\min\{p_1 A_u/n, p_2 B_u\}$. Each multi-hop user in the cellular relay network uses the XOR scheme, which depends on the values of D_s and F_s (cf. Section 3.2.2). The equal-rate throughput achievable by the XOR scheme is $\min\{p_1 C_s/n, p_2 E_s, p_r D_s/k, p_r F_s/k\}$ (cf. Table 3.3).

We are interested in the average equal-rate throughput over all users under Scenario 1 and Scenario 2 for every $n = 1, 4, 10, 20, 30$ (n is the number of users in each cell), every $m = 0, 2, 5, 10, 20, 30$ (m is the number of relays in each cell) and every three-phase tuple (p_1, p_2, p_r) in the set

$$\left\{ \begin{array}{l} (p_1, p_2, p_r) \in \mathbb{R}^3 \\ \left. \begin{array}{l} p_1 \geq 0, p_2 \geq 0, p_r \geq 0, \\ p_1 + p_2 + p_r = 1, \\ p_1, p_2 \text{ and } p_r \text{ are multiples of } 1/64. \end{array} \right\} \right\}. \quad (3.11)$$

The three-phase tuples in (3.11) lie on the plane

$$p_1 + p_2 + p_r = 1.$$

When $n = 1$, the cellular relay network resembles low-density cellular networks. When $n = 4$ or 10 , the cellular relay network resembles medium-density cellular networks. When $n = 20$ or 30 , the cellular relay network resembles high-density cellular networks.

We generate 6000 different networks for each n and each m in our simulation. In each network, we calculate the average equal-rate throughput over the users in the central cell under Scenario 1 and Scenario 2 for every (p_1, p_2, p_r) according to the formulae in Section 3.2.4. Then, we calculate the sample mean of the 6000 average equal-rate throughput under Scenario 1 and Scenario 2 for each n , each m and each (p_1, p_2, p_r) . Using standard arguments, we obtain that a 95% confidence interval for each sample mean μ is $(0.99\mu, 1.01\mu)$.

For each n , each m and each (p_1, p_2, p_r) , the equal-rate throughput for a single-hop user under Scenario 1 and Scenario 2 are the same. In addition, it follows from Section 3.2.4 that for a multi-hop user, the equal-rate throughput under Scenario 1 is less than or equal to the equal-rate throughput under Scenario 2. Consequently, the sample mean of the average equal-rate throughput under Scenario 1 is less than or equal to the sample mean of the average equal-rate throughput under Scenario 2.

For each n and each m , let $\pi_{m,n}$ denote the three-phase tuple in (3.11) that maximizes the sample mean of the average equal-rate throughput under Scenario 1, and the $\pi_{m,n}$'s are displayed in Table 3.4. For each n and each m , let $\theta_{m,n}$ denote the three-phase

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64, 0)	(19/64, 44/64, 1/64)	(18/64, 38/64, 8/64)	(19/64, 30/64, 15/64)	(20/64, 22/64, 22/64)	(20/64, 20/64, 24/64)
$n = 4$	(29/64, 35/64, 0)	(28/64, 32/64, 4/64)	(27/64, 29/64, 8/64)	(27/64, 23/64, 14/64)	(27/64, 18/64, 19/64)	(26/64, 15/64, 23/64)
$n = 10$	(31/64, 33/64, 0)	(30/64, 30/64, 4/64)	(30/64, 26/64, 8/64)	(30/64, 21/64, 13/64)	(29/64, 15/64, 20/64)	(28/64, 13/64, 23/64)
$n = 20$	(30/64, 34/64, 0)	(30/64, 30/64, 4/64)	(31/64, 25/64, 8/64)	(30/64, 20/64, 14/64)	(29/64, 15/64, 20/64)	(28/64, 12/64, 24/64)
$n = 30$	(29/64, 35/64, 0)	(30/64, 30/64, 4/64)	(31/64, 24/64, 9/64)	(30/64, 19/64, 15/64)	(29/64, 14/64, 21/64)	(28/64, 12/64, 24/64)

Table 3.4: Three-phase tuple $\pi_{m,n}$ maximizing average equal-rate throughput under Scenario 1.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64, 0)	(19/64, 42/64, 3/64)	(19/64, 37/64, 8/64)	(22/64, 29/64, 13/64)	(23/64, 23/64, 18/64)	(23/64, 22/64, 19/64)
$n = 4$	(29/64, 35/64, 0)	(28/64, 32/64, 4/64)	(29/64, 28/64, 7/64)	(29/64, 24/64, 11/64)	(28/64, 19/64, 17/64)	(28/64, 16/64, 20/64)
$n = 10$	(31/64, 33/64, 0)	(31/64, 30/64, 3/64)	(31/64, 26/64, 7/64)	(31/64, 21/64, 12/64)	(29/64, 16/64, 19/64)	(29/64, 13/64, 22/64)
$n = 20$	(30/64, 34/64, 0)	(31/64, 30/64, 3/64)	(31/64, 25/64, 8/64)	(31/64, 20/64, 13/64)	(29/64, 15/64, 20/64)	(29/64, 12/64, 23/64)
$n = 30$	(29/64, 35/64, 0)	(31/64, 29/64, 4/64)	(32/64, 24/64, 8/64)	(31/64, 19/64, 14/64)	(29/64, 15/64, 20/64)	(28/64, 12/64, 24/64)

Table 3.5: Three-phase tuple $\theta_{m,n}$ maximizing average equal-rate throughput under Scenario 2.

tuple in (3.11) that maximizes the sample mean of the average equal-rate throughput under Scenario 2, and the $\theta_{m,n}$'s are displayed in Table 3.5.

Under the three-phase mode with three-phase tuple $\pi_{m,n}$, we call the sample mean of the average equal-rate throughput under Scenario 1 the *best routing throughput for n users and m relays*. Under the three-phase mode with three-phase tuple $\theta_{m,n}$, we call the sample mean of the average equal-rate throughput under Scenario 2 the *best XOR throughput for n users and m relays*. The best routing throughput for each n and each m is displayed in Table 3.6 and

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9136	3.9460	4.1817	4.9604	5.5314
$n = 4$	1.2556	1.2601	1.3212	1.4458	1.5893	1.6805
$n = 10$	0.4210	0.4424	0.4831	0.5296	0.5820	0.6139
$n = 20$	0.1704	0.1868	0.2041	0.2254	0.2526	0.2672
$n = 30$	0.0994	0.1106	0.1211	0.1349	0.1526	0.1628

Table 3.6: Best routing throughput for n users and m relays (10^6 bits per second).

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9477	4.2514	4.7112	5.7895	6.4766
$n = 4$	1.2556	1.3032	1.3840	1.5174	1.6685	1.7610
$n = 10$	0.4210	0.4529	0.4950	0.5435	0.5981	0.6307
$n = 20$	0.1704	0.1895	0.2075	0.2297	0.2575	0.2723
$n = 30$	0.0994	0.1120	0.1229	0.1370	0.1551	0.1654

Table 3.7: Best XOR throughput for n users and m relays (10^6 bits per second).

the best XOR throughput for each n and each m is displayed in Table 3.7.

Under the case $m = 0$, there is no relay in the cellular relay network, which implies that all the users in the network are single-hop users. Therefore, the best routing throughput and the best XOR throughput for each n are the same, and they are indeed the average equal-rate capacities over all users in the network without relay. Comparing the column $m = 0$ with the other columns in Table 3.6, we observe that the best routing throughput over all users when the number of relays exceeds 10 is always greater than the average equal-rate capacities over all users when there is no relay.

Similarly, we observe in Table 3.7 that the best XOR throughput over all users when the number of relays exceeds 5 is always greater than the average equal-rate capacities over all users when there is no relay.

Table 3.6 and Table 3.7 show that for each n , both the best routing throughput and the best XOR throughput generally increase as m increases, which implies that adding more relays to the network improves both the best routing throughput and the best XOR throughput. For each m , both the best routing throughput and the XOR throughput decrease as n increases, which agrees with the fact that adding more users to the network results in

1. less time allocated to each user for Phase 1 transmissions;
2. larger interference incurred by the users on each receiving node during Phase 2 (cf. Table 3.1 and Table 3.2 in Section 3.2.1).

The percentage of multi-hop users for each n and each m is shown in Table 3.8. For each n , the percentage of multi-hop users increases as m increases. In other words, adding more relays to the network results in larger proportion of multi-hop users, which agrees with the fact that an increase in relay number increases the probability for a user to choose to communicate with the base station through a relay. For each m , the percentage of multi-hop users does not change much as n varies, which implies that the percentage of multi-hop users depends on the relay density regardless of the user density.

Table 3.6 and Table 3.7 show that the use of network coding rather than routing alone increases the average maximum equal-rate throughput over all users in the cellular relay network. For each n and each m , we call the percentage gain by which the best XOR throughput is greater than the best routing throughput the *average*

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	0%	8.17%	18.80%	31.63%	52.25%	64.83%
$n = 4$	0%	7.93%	17.55%	32.29%	52.70%	65.62%
$n = 10$	0%	7.62%	18.22%	32.98%	52.73%	65.33%
$n = 20$	0%	7.91%	17.94%	32.46%	52.96%	65.46%
$n = 30$	0%	7.89%	18.11%	32.43%	52.48%	65.40%

Table 3.8: Percentage of multi-hop users.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	0%	0.87%	7.74%	12.66%	16.71%	17.09%
$n = 4$	0%	3.42%	4.75%	4.95%	4.98%	4.79%
$n = 10$	0%	2.37%	2.46%	2.63%	2.78%	2.74%
$n = 20$	0%	1.49%	1.69%	1.87%	1.94%	1.91%
$n = 30$	0%	1.27%	1.42%	1.61%	1.67%	1.60%

Table 3.9: Average network coding percentage gain over all users.

network coding percentage gain over all users, which is shown in Table 3.9. For each n and each $m \neq 0$, we call the average network coding percentage gain over all users divided by the percentage of multi-hop users the *average network coding percentage gain over multi-hop users*, which is shown in Table 3.10.

Table 3.9 and Table 3.10 show that for each m , both the average network coding gain over all users and the average network coding gain over multi-hop users generally decrease as n increases. Table 3.9 shows that the highest average network coding gain over all users for each $n \neq 1$ is attained at $m = 20$. Table 3.10 shows that for each n , the average network coding gain over multi-hop users

	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	10.66%	41.17%	40.03%	31.99%	26.36%
$n = 4$	43.15%	27.05%	15.34%	9.46%	7.30%
$n = 10$	31.12%	13.52%	7.98%	5.27%	4.19%
$n = 20$	18.78%	9.42%	5.77%	3.67%	2.91%
$n = 30$	16.13%	7.83%	4.97%	3.18%	2.45%

Table 3.10: Average network coding percentage gain over multi-hop users.

generally decreases as m increases. For low-density networks, i.e. $n = 1$, the average network coding gain over all users can be greater than 17% and the average network coding gain over multi-hop users can be greater than 41%. For medium-density networks, i.e. $n = 4$ or 10, the average network coding gain over all users can be greater than 4% and the average network coding gain over multi-hop users can be greater than 43%. For high-density networks, i.e. $n = 20$ or 30, the average network coding gain over all users can be greater than 1% and the average network coding gain over multi-hop users can be greater than 18%.

The four-phase mode proposed in the following section subsumes the three-phase mode. It will be shown that the benefit of network coding under the four-phase mode compared with the three-phase mode is more obvious.

3.3 Four-Phase Mode

In this section, we consider a new phase in addition to the three phases introduced in the previous section and propose a new transmission mode for the cellular relay network called *four-phase mode*.

If no time is allocated for the new phase under the four-phase mode, the transmissions under the four-phase mode are equivalent to the transmissions under the three-phase mode. Therefore, the four-phase mode subsumes the three-phase mode. We will investigate by simulation the average equal-rate throughput over all users under two scenarios: each user uses an optimal network coding scheme and each user uses an optimal routing scheme.

3.3.1 Resource and Interference Management

To facilitate discussion, we classify the transmissions in each two-node point-to-point system as follows:

Class 1: The base station transmits messages to the single-hop user.

Class 2: The single-hop user transmits messages to the base station.

Similarly, we classify the transmissions in each three-node point-to-point relay network as follows:

Class 1: The base station transmits messages to the relay.

Class 2: The multi-hop user transmits messages to the relay.

Class 3: The relay transmits messages to the base station and the multi-hop user.

The transmissions in each of the seven cells are organized into four phases, denoted by Phase 1, Phase 2, Phase 3 and Phase 4. The transmissions in the seven cells are synchronized in such a way that the seven cells have the same periods for Phase 1, Phase 2, Phase 3 and Phase 4. The fractions of time allocated for the four phases are

fixed and we let q_1, q_2, q_3 and q_4 denote the fractions of time allocated for Phase 1, Phase 2, Phase 3 and Phase 4 transmissions respectively. We call (q_1, q_2, q_3, q_4) a *four-phase tuple*. The transmissions in the four phases are described as follows.

Phase 1: In each cell, the phase is divided into n equal-length time slots which are allocated to the n users in the cell. In each time slot, if the user is a single-hop user, the communication from the base station to the user is conducted over the corresponding two-node system by Class 1 transmissions. If the user is a multi-hop user, the communication from the base station to the user through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 2: In each cell, all the users simultaneously transmit messages throughout the entire phase. For each single-hop user, the communication from the user to the base station is conducted over the corresponding two-node system by Class 2 transmissions. For each multi-hop user, the communication from the user to the base station through its associated relay is conducted over the corresponding three-node network by Class 2 transmissions.

Phase 3: In each cell, all the active relays simultaneously transmit throughout the entire phase. For each multi-hop user u , the communication between u and the base station b through its associated relay r is as follows: Let $k \geq 1$ denote the number of multi-hop users that choose to communicate with b through r and let t_1, t_2, \dots, t_{k-1} denote the users other than u that communicate with b through r . This is illustrated in Figure 3.6. The phase is divided into k equal-length time slots, which are

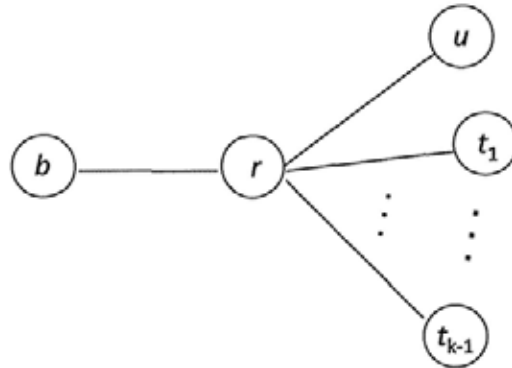


Figure 3.6: A relay r which serves k multi-hop users.

allocated to the k users served by relay r . Note that the value of k depends on the relay. In the time slot allocated to user u , the communication between u and b through r is conducted over the corresponding three-node network by Class 3 transmissions.

Phase 4: For each cell that contains ω multi-hop users, the phase is divided into ω equal-length time slots which are allocated to the ω multi-hop users. In each time slot, the communication between the multi-hop user and the base station through the relay is conducted over the corresponding three-node network by Class 3 transmissions.

In each of the phases above, each listening station in a two-node system or a three-node network regards unintended signals as independent Gaussian noise.

Given a four-phase tuple (q_1, q_2, q_3, q_4) , the time allocations for each two-node system and each three-node network in Phase 1 are as follows:

1. For each two-node system, the fraction of time allocated for

Class 1 transmissions is q_1/n .

2. For each three-node network, the fraction of time allocated for Class 1 transmissions is q_1/n .

The time allocations for each two-node system and each three-node network in Phase 2 are as follows:

1. For each two-node system, the fraction of time allocated for Class 2 transmissions is q_2 .
2. For each three-node network, the fraction of time allocated for Class 2 transmissions is q_2 .

The time allocations for each two-node system and each three-node network in Phase 3 are as follows:

1. Each two-node system is idle.
2. For each three-node network, the fraction of time allocated for Class 3 transmissions is q_3/k , where $k \geq 1$ is the number of multi-hop users served by the relay.

The time allocations for each two-node system and each three-node network in Phase 4 are as follows:

1. Each two-node system is idle.
2. For each three-node network, the fraction of time allocated for Class 3 transmissions is q_4/ω , where $\omega \geq 1$ is the number of multi-hop users in the cell.

To facilitate understanding, the fractions of time allocated to each two-node point-to-point system as well as each three-node point-to-point network for Phase 1, Phase 2, Phase 3 and Phase 4 transmissions are shown in Table 3.11.

	Two-node system	Three-node network
Phase 1	q_1/n	q_1/n
Phase 2	q_2	q_2
Phase 3	0	q_3/k
Phase 4	0	q_4/ω

Table 3.11: The fractions of time allocated to each two-node point-to-point system and each three-node point-to-point network under the four-phase mode.

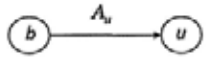
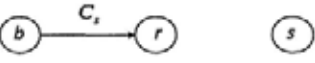
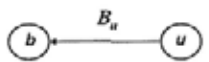
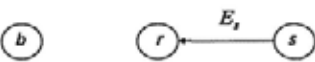


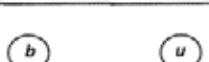
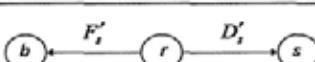
	Single-hop user u	Multi-hop user s
Phase 1		
Phase 2		
Phase 3		
Phase 4		

Table 3.12: The capacities of a two-node point-to-point system and a three-node point-to-point network under the four-phase mode.

For each single-hop user u and the base station b , let A_u denote the capacity of (b, u) in Phase 1 and B_u denote the capacity of (u, b) in Phase 2. This is illustrated on the left side of Table 3.12.

For each multi-hop user s which communicates with the base station b through its associated relay r , let C_s and E_s denote the capacities of (b, r) in Phase 1 and (s, r) in Phase 2 respectively. We let D_s and F_s denote the capacities of (r, s) and (r, b) in Phase 3 respectively and let D'_s and F'_s denote the capacities of (r, s) and (r, b) in Phase 4 respectively. This is illustrated on the right side of Table 3.12. As discussed at the beginning of Chapter 2, relay r can broadcast information reliably from r to both node b and node s at rate R in Phase 3 if and only if $R \leq \min\{D_s, F_s\}$, and relay r can broadcast information reliably from r to both node b and node s at rate R' in Phase 4 if and only if $R' \leq \min\{D'_s, F'_s\}$. We call $\min\{D_s, F_s\}$ and $\min\{D'_s, F'_s\}$ the *broadcast capacity in Phase 3* and the *broadcast capacity in Phase 4* respectively.

In the three-phase mode with any three-phase tuple (p_1, p_2, p_r) (cf. Section 3.2.1), the three-phase mode can be viewed as the four-phase mode with four-phase tuple $(p_1, p_2, p_r, 0)$. Therefore, the four-phase mode subsumes the three-phase mode.

3.3.2 Maximum Equal-Rate Throughput

Fix a four-phase tuple (q_1, q_2, q_3, q_4) . For every user, we are interested in the maximum equal-rate throughput achievable by network coding schemes. For each single-hop user u , the fractions q_1/n and q_2 of the time are allocated to the corresponding two-node point-to-point system for Phase 1 and Phase 2 transmissions respectively. Therefore, the maximum throughput from the base station to u is

$q_1 A_u/n$ and the maximum throughput from u to the base station is $q_2 B_u$, which implies that the maximum equal-rate throughput for u is $\min\{q_1 A_u/n, q_2 B_u\}$, which is trivially achievable by some routing scheme.

For each multi-hop user s which communicates with the base station through its associated relay r , we will prove an upper bound on the equal-rate throughput of the corresponding three-node point-to-point relay network. We will also present a network coding scheme that achieves this upper bound. We call this upper bound the *maximum equal-rate throughput for s* and it is shown in Table 3.13, where $k \geq 1$ is the number of multi-hop users served by r and $\omega \geq 1$ is the number of multi-hop users in the cell containing s .

The techniques involved in proving the upper bounds in Table 3.13 are similar to the ones used in the proof of Lemma 1 in Section 2.1.2. The details are as follows. Fix a multi-hop user s and its corresponding three-node point-to-point relay network. Let b and r denote the base station and the relay respectively in the three-node network. Consider the following four cases:

Case		Maximum equal-rate throughput of s
$D_s \leq F_s$ and $D'_s \leq F'_s$		$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 D_s + q_4 D'_s}{k} + \frac{q_4 D'_s}{\omega}\right\}$
$D_s > F_s$ and $D'_s > F'_s$		$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 F_s + q_4 F'_s}{k} + \frac{q_4 F'_s}{\omega}\right\}$
$D_s \leq F_s$ and $D'_s > F'_s$	$D_s D'_s + F_s F'_s - D'_s F_s \geq 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 D_s + q_4 F'_s}{k} + \frac{q_4 F'_s}{\omega}\right\}$
	$D_s D'_s + F_s F'_s - D'_s F_s < 0$ $\omega q_3 D_s - k q_4 F'_s \leq 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{\omega q_3 F_s (D'_s - F'_s) + k q_4 D'_s F'_s}{k \omega D'_s}\right\}$
	$\omega q_3 D_s - k q_4 F'_s > 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{k q_4 D'_s (F_s - D_s) + \omega q_3 D_s F_s}{k \omega F_s}\right\}$
$D_s > F_s$ and $D'_s \leq F'_s$	$D_s D'_s + F_s F'_s - D_s F'_s \geq 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 F_s + q_4 D'_s}{k} + \frac{q_4 D'_s}{\omega}\right\}$
	$D_s D'_s + F_s F'_s - D_s F'_s < 0$ $\omega q_3 D_s - k q_4 F'_s \geq 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{k q_4 F'_s (D_s - F_s) + \omega q_3 D_s F_s}{k \omega D_s}\right\}$
	$\omega q_3 D_s - k q_4 F'_s < 0$	$\min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \frac{\omega q_3 D_s (F'_s - D'_s) + k q_4 D'_s F'_s}{k \omega F'_s}\right\}$

 Table 3.13: The maximum equal-rate throughput for s .

Case 1: $D_s \leq F_s$ and $D'_s \leq F'_s$

The broadcast capacities in Phase 3 and Phase 4 are $\min\{D_s, F_s\} = D_s$ and $\min\{D'_s, F'_s\} = D'_s$ respectively. Therefore, r can broadcast messages reliably to both the base station and s at a rate less than or equal to D_s and D'_s in Phase 3 and Phase 4 respectively. Without loss of generality, we assume that only one of the following six types of transmissions can take place in a transmission scheme.

Type 1: The base station transmits messages reliably to r at a rate less than or equal to C_s .

Type 2: Multi-hop user s transmits messages reliably to r at a rate less than or equal to E_s .

Type 3: Relay r broadcasts messages reliably to both the base station and s at a rate less than or equal to D_s .

Type 4: Relay r transmits messages reliably to the base station at a rate less than or equal to F_s but strictly greater than D_s .

Type 5: Relay r broadcasts messages reliably to both the base station and s at a rate less than or equal to D'_s .

Type 6: Relay r transmits messages reliably to the base station at a rate less than or equal to F'_s but strictly greater than D'_s .

In addition, we assume that Phase 1 consists of Type 1 transmissions, Phase 2 consists of Type 2 transmissions, Phase 3 consists of Type 3 and Type 4 transmissions, and Phase 4 consists of Type 5 and Type 6 transmissions. Since the fractions of time allocated to the three-node network for Phase 1 and

Phase 2 transmissions are q_1/n and q_2 respectively, the fractions of time allocated for Type 1 and Type 2 transmissions are q_1/n and q_2 respectively. Let $\alpha_s \geq 0$ and $\beta_s \geq 0$ denote the fractions of time allocated for Type 3 and Type 4 transmissions respectively. Since the fraction of time allocated to the three-node network for Phase 3 transmissions is q_3/k , we assume that

$$\alpha_s + \beta_s = q_3/k.$$

Let $\alpha'_s \geq 0$ and $\beta'_s \geq 0$ denote the fractions of time allocated for Type 5 and Type 6 transmissions respectively. Since the fraction of time allocated to the three-node network for Phase 4 transmissions is q_4/ω , we assume that

$$\alpha'_s + \beta'_s = q_4/\omega.$$

For any fixed α_s and α'_s , the highest average rates for Type 1, 2, 3, 4, 5 and 6 transmissions are q_1C_s/n , q_2E_s , α_sD_s , $(q_3/k - \alpha_s)F_s$, $\alpha'_sD'_s$ and $(q_4/\omega - \alpha'_s)F'_s$ respectively. Therefore, the equal-rate throughput, denoted by R , must satisfy the rate constraints induced by the two-source point-to-point network in Figure 3.7(a), which are equivalent to the rate constraints induced by the single-source point-to-point multicast network in Figure 3.7(b) (cf. the proof of Lemma 1 in Section 2.1.2). Applying the cut-set bound for four different cuts as in the proof of Lemma 1, we obtain

$$\begin{aligned} R &\leq q_1C_s/n, \\ R &\leq \alpha_sD_s + \alpha'_sD'_s, \\ R &\leq q_2E_s \end{aligned}$$

and

$$R \leq \alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s.$$

Since $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$, it follows that an upper bound on the equal-rate throughput is

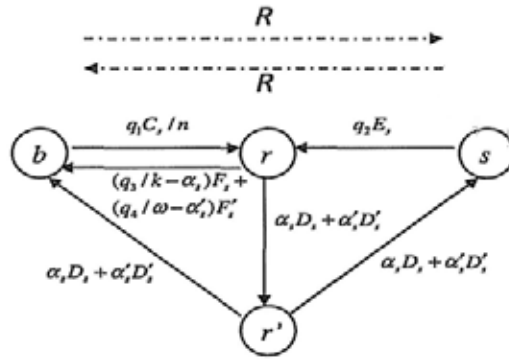
$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{q_1 C_s/n, q_2 E_s, \alpha_s D_s + \alpha'_s D'_s, \alpha_s D_s + \alpha'_s D'_s, \\ & \quad (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\ & = \min\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, \\ & \quad \alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}\} \\ & = \min\{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 D'_s/\omega\}, \end{aligned} \quad (3.12)$$

where the last equality follows from the fact that

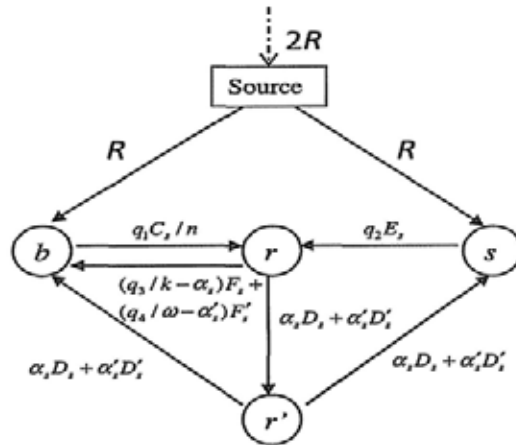
$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, \alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s) F_s \\ & \quad + (q_4/\omega - \alpha'_s) F'_s\} \\ & = q_3 D_s/k + q_4 D'_s/\omega. \end{aligned}$$

We now construct a network coding scheme for s that achieves the upper bound on the equal-rate throughput in (3.12). To simplify notation, we let $\tilde{\alpha}_s = q_3/k$ and $\tilde{\alpha}'_s = q_4/\omega$. A fraction q_1/n of the time is allocated to the corresponding three-node point-to-point relay network for Phase 1 transmissions and during this time, the base station transmits messages to r at rate

$$\min\{C_s, nq_2 E_s/q_1, n\tilde{\alpha}_s D_s/q_1 + n\tilde{\alpha}'_s D'_s/q_1\} \leq C_s.$$



(a) A two-source point-to-point network.



(b) A single-source point-to-point multicast network.

Figure 3.7: Two equivalent network coding problems.

A fraction q_2 of the time is allocated to the three-node network for Phase 2 transmissions and during this time, s transmits messages to r at rate

$$\min \left\{ \frac{q_1 C_s}{q_2 n}, E_s, \tilde{\alpha}_s D_s / q_2 + \tilde{\alpha}'_s D'_s / q_2 \right\} \leq E_s.$$

A fraction $\tilde{\alpha}_s = q_3/k$ of the time is allocated to the three-node network for Phase 3 transmissions and during this time, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station and s at rate

$$\min \left\{ \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) \left(\frac{q_1 C_s}{n \tilde{\alpha}_s} \right), \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) q_2 E_s / \tilde{\alpha}_s, D_s \right\} \leq D_s.$$

A fraction $\tilde{\alpha}'_s = q_4/\omega$ of the time is allocated to the three-node network for Phase 4 transmissions and during this time, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station and s at rate

$$\min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) \left(\frac{q_1 C_s}{n \tilde{\alpha}'_s} \right), \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) q_2 E_s / \tilde{\alpha}'_s, D'_s \right\} \leq D'_s.$$

Under the scheme, the average rates of the messages sent from the base station to the relay and from s to the relay are both

$$\min \{ q_1 C_s / n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s \}.$$

The average rate of the messages sent from the relay to the

base station and s in Phase 3 and Phase 4 combined is

$$\begin{aligned} & \min \left\{ \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) \left(\frac{q_1 C_s}{n} \right), \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) q_2 E_s, \tilde{\alpha}_s D_s \right\} \\ & + \min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) \left(\frac{q_1 C_s}{n} \right), \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s} \right) q_2 E_s, \right. \\ & \left. \tilde{\alpha}'_s D'_s \right\} \\ & = \min \{ q_1 C_s / n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s \}, \end{aligned}$$

where the equality follows from the fact that for all non-negative real numbers a , b , c and d such that $c + d > 0$,

$$\begin{aligned} & \min \{ a, b, c + d \} \\ & = \min \left\{ \frac{ac}{c+d}, \frac{bc}{c+d}, c \right\} + \min \left\{ \frac{ad}{c+d}, \frac{bd}{c+d}, d \right\}, \quad (3.13) \end{aligned}$$

which can be verified by considering all the three cases

$\min \{ a, b, c + d \} = a$, $\min \{ a, b, c + d \} = b$ and $\min \{ a, b, c + d \} = c + d$. Consequently, the average rates of the messages sent from the base station to the relay, from s to the relay, from the relay to the base station and from the relay to s are all $\min \{ q_1 C_s / n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s \}$. In addition, the base station can recover the messages of s by performing XOR operations between its own messages and the messages from the relay. Similarly, s can recover the messages of the base station by performing XOR operations between its own messages and the messages from the relay. Therefore, the equal-rate throughput

$$\begin{aligned} & \min \{ q_1 C_s / n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s \} \\ & = \min \{ q_1 C_s / n, q_2 E_s, q_3 D_s / k + q_4 D'_s / \omega \} \end{aligned}$$

is achievable by the scheme.

Case 2: $D_s > F_s$ and $D'_s > F'_s$

The broadcast capacities in Phase 3 and Phase 4 are $\min\{D_s, F_s\} = F_s$ and $\min\{D'_s, F'_s\} = F'_s$ respectively. Let $(\tilde{\alpha}_s, \tilde{\alpha}'_s)$ denote the argument of

$$\max_{\substack{(\alpha_s, \alpha'_s) \\ 0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s)F_s + (q_4/\omega - \alpha'_s)F'_s, (q_3/k - \alpha_s)F_s + (q_4/\omega - \alpha'_s)F'_s\}.$$

Following similar procedures in Case 1, we can show that

$$(\tilde{\alpha}_s, \tilde{\alpha}'_s) = (0, 0)$$

and

$$\begin{aligned} & \min\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + \\ & (q_3/k - \alpha_s)F_s + (q_4/\omega - \alpha'_s)F'_s, (q_3/k - \alpha_s)F_s + \\ & (q_4/\omega - \alpha'_s)F'_s\}\} \\ & = \min\{q_1 C_s/n, q_2 E_s, q_3 F_s/k + q_4 F'_s/\omega\} \end{aligned} \quad (3.14)$$

is an upper bound on the equal-rate throughput achievable by network coding schemes. In addition, we can construct a network coding scheme that achieves this upper bound.

Case 3: $D_s \leq F_s$ and $D'_s > F'_s$

The broadcast capacities in Phase 3 and Phase 4 are $\min\{D_s, F_s\} = D_s$ and $\min\{D'_s, F'_s\} = F'_s$ respectively. Therefore, r can broadcast messages reliably to both the base station and s at a rate less than or equal to D_s and F'_s in Phase 3 and Phase 4

respectively. Without loss of generality, we assume that only one of the following six types of transmissions can take place in a transmission scheme.

Type 1: The base station transmits messages reliably to r at a rate less than or equal to C_s .

Type 2: Multi-hop user s transmits messages reliably to r at a rate less than or equal to E_s .

Type 3: Relay r broadcasts messages reliably to both the base station and s at a rate less than or equal to D_s .

Type 4: Relay r transmits messages reliably to the base station at a rate less than or equal to F_s but strictly greater than D_s .

Type 5: Relay r broadcasts messages reliably to both the base station and s at a rate less than or equal to F'_s .

Type 6: Relay r transmits messages reliably to s at a rate less than or equal to D'_s but strictly greater than F'_s .

In addition, we assume that Phase 1 consists of Type 1 transmissions, Phase 2 consists of Type 2 transmissions, Phase 3 consists of Type 3 and Type 4 transmissions, and Phase 4 consists of Type 5 and Type 6 transmissions. Since the fractions of time allocated to the three-node network for Phase 1 and Phase 2 transmissions are q_1/n and q_2 respectively, the fractions of time allocated for Type 1 and Type 2 transmissions are q_1/n and q_2 respectively. Let $\alpha_s \geq 0$ and $\beta_s \geq 0$ denote the fractions of time allocated for Type 3 and Type 4 transmissions respectively. Since the fraction of time allocated to the three-node network for Phase 3 transmissions is q_3/k , we

assume that

$$\alpha_s + \beta_s = q_3/k.$$

Let $\alpha'_s \geq 0$ and $\beta'_s \geq 0$ denote the fractions of time allocated for Type 5 and Type 6 transmissions respectively. Since the fraction of time allocated to the three-node network for Phase 4 transmissions is q_4/ω , we assume that

$$\alpha'_s + \beta'_s = q_4/\omega.$$

For any fixed α_s and α'_s , the highest average rates for Type 1, 2, 3, 4, 5 and 6 transmissions are $q_1 C_s/n$, $q_2 E_s$, $\alpha_s D_s$, $(q_3/k - \alpha_s) F_s$, $(q_4/\omega - \alpha'_s) F'_s$ and $\alpha'_s D'_s$ respectively. Therefore, the equal-rate throughput, denoted by R , must satisfy the rate constraints induced by the two-source point-to-point network in Figure 3.8(a), which are equivalent to the rate constraints induced by the single-source point-to-point multicast network in Figure 3.8(b) (cf. the proof of Lemma 1 in Section 2.1.2). Applying the cut-set bound for four different cuts as in the proof of Lemma 1, we obtain

$$\begin{aligned} R &\leq q_1 C_s/n, \\ R &\leq \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \\ R &\leq q_2 E_s \end{aligned}$$

and

$$R \leq \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s.$$

Since $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$, it follows that an

upper bound on the equal-rate throughput is

$$\begin{aligned}
 & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{q_1 C_s/n, q_2 E_s, \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \\
 & \quad \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\
 & = \min\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + \\
 & \quad (q_4/\omega - \alpha'_s) F'_s, \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}\}.
 \end{aligned} \tag{3.15}$$

We will simplify (3.15) and construct a network coding scheme that achieves the equal-rate throughput (3.15). Let $(\bar{\alpha}_s, \bar{\alpha}'_s)$ denote

$$\begin{aligned}
 & \arg \max_{\substack{(\alpha_s, \alpha'_s): 0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \\
 & \quad \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}.
 \end{aligned} \tag{3.16}$$

We claim that

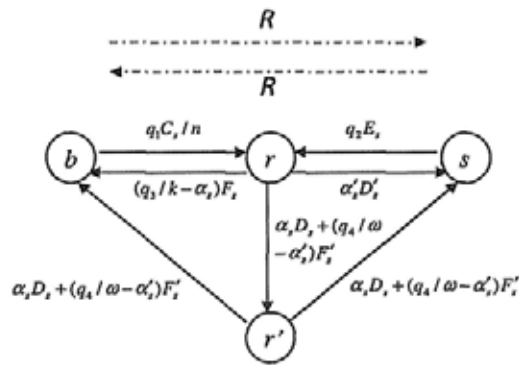
$$\begin{aligned}
 & \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s + (q_4/\omega - \bar{\alpha}'_s) F'_s \\
 & = \bar{\alpha}_s D_s + (q_3/k - \bar{\alpha}_s) F_s + (q_4/\omega - \bar{\alpha}'_s) F'_s
 \end{aligned}$$

and verify the claim as follows. For any $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$ that satisfy

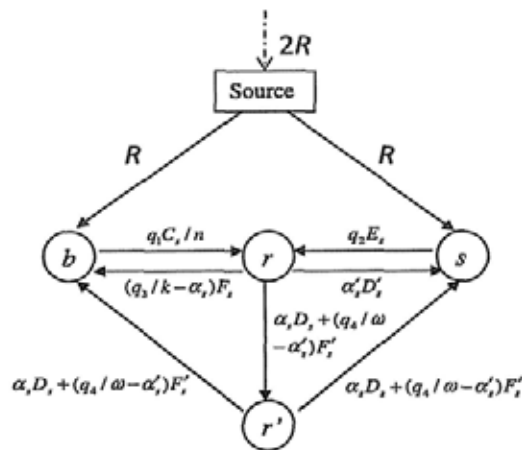
$$\begin{aligned}
 & \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s \\
 & > \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,
 \end{aligned}$$

it follows that

$$\alpha'_s D'_s > (q_3/k - \alpha_s) F_s.$$



(a) A two-source point-to-point network.



(b) A single-source point-to-point multicast network.

Figure 3.8: Two equivalent network coding problems.

Then, there exist $\hat{\alpha}_s = \alpha_s$ and $0 \leq \hat{\alpha}'_s < \alpha'_s$ such that

$$\hat{\alpha}'_s D'_s = (q_3/k - \hat{\alpha}_s) F_s,$$

which implies that

$$\begin{aligned} & \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s \\ & > \hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\ & = \hat{\alpha}_s D_s + (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\ & > \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s \end{aligned}$$

where the first inequality follows from $D'_s > F'_s$, which then implies

$$\begin{aligned} & \min\{\alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \alpha_s D_s + \\ & \quad (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\ & < \min\{\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s + (q_4/\omega - \hat{\alpha}'_s) F'_s, \\ & \quad \hat{\alpha}_s D_s + (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s\}. \quad (3.17) \end{aligned}$$

Similarly, for any $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$ that satisfy

$$\begin{aligned} & \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s \\ & < \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s, \end{aligned}$$

it follows that

$$\alpha'_s D'_s < (q_3/k - \alpha_s) F_s.$$

Then, there exist $\alpha_s < \hat{\alpha}_s \leq q_3/k$ and $\alpha'_s = \hat{\alpha}'_s$ such that

$$\hat{\alpha}'_s D'_s = (q_3/k - \hat{\alpha}_s) F_s,$$

which implies that

$$\begin{aligned}
 & \alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s \\
 & < \hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\
 & = \hat{\alpha}_s D_s + (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\
 & \leq \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,
 \end{aligned}$$

where the second inequality follows from $D_s \leq F_s$, which then implies that

$$\begin{aligned}
 & \min\{\alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \\
 & \quad \alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\
 & < \min\{\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s + (q_4/\omega - \hat{\alpha}'_s) F'_s, \\
 & \quad \hat{\alpha}_s D_s + (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s\}. \quad (3.18)
 \end{aligned}$$

It then follows from (3.17) and (3.18) that

$$\begin{aligned}
 & \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s \\
 & = \tilde{\alpha}_s D_s + (q_3/k - \tilde{\alpha}_s) F_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s. \quad (3.19)
 \end{aligned}$$

Therefore, the equal-rate throughput upper bound in (3.15) can be simplified to

$$\min\{q_1 C_s/n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s\}. \quad (3.20)$$

In addition, it follows from (3.19) that

$$\tilde{\alpha}'_s = \frac{q_3 F_s}{k D'_s} - \frac{\tilde{\alpha}_s F_s}{D'_s}. \quad (3.21)$$

Therefore,

$$\begin{aligned}
 & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + (q_4/\omega - \alpha'_s) F'_s, \alpha_s D_s + \\
 & \quad (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\
 & \stackrel{(a)}{=} \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega \\ \alpha'_s D'_s = (q_3/k - \alpha_s) F_s}} (\alpha_s D_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s) \\
 & \stackrel{(b)}{=} \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \frac{q_3 F_s}{k D'_s} - \frac{\alpha_s F_s}{D'_s} \leq q_4/\omega}} \left(\alpha_s D_s + \left(\frac{q_3}{k} - \alpha_s \right) F_s + \left(\frac{q_4}{\omega} - \right. \right. \\
 & \quad \left. \left. \left(\frac{q_3 F_s}{k D'_s} - \frac{\alpha_s F_s}{D'_s} \right) \right) F'_s \right) \\
 & = \max_{\{0, \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s}\} \leq \alpha_s \leq q_3/k} \left(\alpha_s \left(D_s - F_s + \frac{F_s F'_s}{D'_s} \right) + \frac{q_3 F_s}{k} \right. \\
 & \quad \left. + \frac{q_4 F'_s}{\omega} - \frac{q_3 F_s F'_s}{k D'_s} \right), \tag{3.22}
 \end{aligned}$$

where (a) follows from (3.19) and (b) follows from eliminating α'_s by the equation

$$\alpha'_s D'_s = (q_3/k - \alpha_s) F_s.$$

Under each of the following two subcases, we will find $(\tilde{\alpha}_s, \tilde{\alpha}'_s)$ through investigating (3.22) and construct a network coding scheme that achieves the equal-rate throughput (3.20).

Subcase $D_s D'_s + F_s F'_s - D'_s F_s \geq 0$:

Since $D_s - F_s + F_s F'_s / D'_s \geq 0$, it follows that (3.22) is achieved by the largest possible α_s . Since the largest possible α_s for (3.22) is q_3/k , it follows that

$$\tilde{\alpha}_s = q_3/k$$

and

$$\tilde{\alpha}'_s = 0$$

(cf. (3.21)), which implies from (3.20) that an upper bound on the equal-rate throughput is

$$\min\{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 F'_s/\omega\}.$$

Consider the following network coding scheme for s that achieves this equal-rate throughput : A fraction q_1/n of the time is allocated to the corresponding three-node point-to-point relay network for Phase 1 transmissions and during this time, the base station transmits messages to r at rate

$$\min \left\{ C_s, nq_2 E_s/q_1, \frac{nq_3 D_s}{q_1 k} + \frac{nq_4 F'_s}{q_1 \omega} \right\} \leq C_s.$$

A fraction q_2 of the time is allocated to the three-node network for Phase 2 transmissions and during this time, s transmits messages to r at rate

$$\min \left\{ \frac{q_1 C_s}{q_2 n}, E_s, \frac{q_3 D_s}{q_2 k} + \frac{q_4 F'_s}{q_2 \omega} \right\} \leq E_s.$$

A fraction q_3/k of the time is allocated to the three-node network for Phase 3 transmissions and during this time, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station

and s at rate

$$\min \left\{ \left(\frac{q_3 D_s / k}{q_3 D_s / k + q_4 F'_s / \omega} \right) \left(\frac{k q_1 C_s}{n q_3} \right), \right. \\ \left. \left(\frac{q_3 D_s / k}{q_3 D_s / k + q_4 F'_s / \omega} \right) \left(\frac{k q_2 E_s}{q_3} \right), D_s \right\} \\ \leq D_s.$$

A fraction q_4/ω of the time is allocated to the three-node network for Phase 4 transmissions and during this time, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station and s at rate

$$\min \left\{ \left(\frac{q_4 F'_s / \omega}{q_3 D_s / k + q_4 F'_s / \omega} \right) \left(\frac{\omega q_1 C_s}{n q_4} \right), \right. \\ \left. \left(\frac{q_4 F'_s / \omega}{q_3 D_s / k + q_4 F'_s / \omega} \right) \left(\frac{\omega q_2 E_s}{q_4} \right), F'_s \right\} \\ \leq F'_s.$$

Under the scheme, the average rates of the messages sent from the base station to the relay and from s to the relay are both

$$\min\{q_1 C_s / n, q_2 E_s, q_3 D_s / k + q_4 F'_s / \omega\}.$$

The average rate of the messages sent from the relay to the

base station and s in Phase 3 and Phase 4 combined is

$$\begin{aligned} & \min \left\{ \frac{(q_3 D_s/k)(q_1 C_s/n)}{q_3 D_s/k + q_4 F'_s/\omega}, \frac{(q_3 D_s/k)(q_2 E_s)}{q_3 D_s/k + q_4 F'_s/\omega}, q_3 D_s/k \right\} \\ & \quad + \min \left\{ \frac{(q_4 F'_s/\omega)(q_1 C_s/n)}{q_3 D_s/k + q_4 F'_s/\omega}, \frac{(q_4 F'_s/\omega)(q_2 E_s)}{q_3 D_s/k + q_4 F'_s/\omega}, q_4 F'_s/\omega \right\} \\ & = \min \{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 F'_s/\omega\} \end{aligned}$$

(cf. (3.13)). Consequently, the average rates of the messages sent from the base station to the relay, s to the relay, from the relay to the base station and from the relay to s are all $\min\{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 F'_s/\omega\}$. In addition, the base station can recover the messages of s by performing XOR operations between its own messages and the messages from the relay. Similarly, s can recover the messages of the base station by performing XOR operations between its own messages and the messages from the relay. Therefore, the equal-rate throughput

$$\min\{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 F'_s/\omega\}$$

is achievable by the network coding scheme. Since

$$\begin{aligned} & \min\{q_1 C_s/n, q_2 E_s, q_3 D_s/k + q_4 F'_s/\omega\} \\ & = \min\{q_1 C_s/n, q_2 E_s, \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s + (q_4 - \bar{\alpha}'_s) F'_s\}, \end{aligned}$$

it then follows that the network coding scheme achieves the equal-rate throughput (3.20), which is equivalent to (3.15).

Subcase $D_s D'_s + F_s F'_s - D'_s F_s < 0$:

Since $D_s - F_s + F_s F'_s/D'_s < 0$, it follows that (3.22) is achieved by the smallest possible α_s . Since the smallest

possible α_s for (3.22) is

$$\max \left\{ 0, \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} \right\},$$

it follows that

$$\tilde{\alpha}_s = \begin{cases} 0 & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} \leq 0, \\ \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} > 0, \end{cases}$$

which implies from (3.21) that

$$(\tilde{\alpha}_s, \tilde{\alpha}'_s) = \begin{cases} \left(0, \frac{q_3 F_s}{k D'_s} \right) & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} \leq 0, \\ \left(\frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s}, \frac{q_4}{\omega} \right) & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} > 0. \end{cases}$$

Consequently, the equal-rate throughput upper bound (3.20) is

$$\begin{cases} \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 \omega F_s (D'_s - F'_s) + k q_4 D'_s F'_s}{k \omega D'_s} \right\} & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} \leq 0, \\ \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{k q_4 D'_s (F_s - D_s) + \omega q_3 D_s F_s}{k \omega F_s} \right\} & \text{if } \frac{q_3}{k} - \frac{q_4 D'_s}{\omega F_s} > 0, \end{cases}$$

which can be simplified to

$$\begin{cases} \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 \omega F_s (D'_s - F'_s) + k q_4 D'_s F'_s}{k \omega D'_s} \right\} & \text{if } \omega q_3 F_s - k q_4 D'_s \leq 0, \\ \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{k q_4 D'_s (F_s - D_s) + \omega q_3 D_s F_s}{k \omega F_s} \right\} & \text{if } \omega q_3 F_s - k q_4 D'_s > 0. \end{cases} \quad (3.23)$$

Consider the following network coding scheme for s that achieves the equal-rate throughput upper bound in (3.23): A fraction q_1/n of the time is allocated to the corresponding three-node point-to-point relay network for Phase 1 transmissions and during this time, the base station trans-

mits messages to r at rate

$$\begin{aligned} & \min\{C_s, nq_2E_s/q_1, n(\tilde{\alpha}_sD_s + \tilde{\alpha}'_sD'_s + (q_4/\omega - \tilde{\alpha}'_s)F'_s)/q_1\} \\ & \leq C_s. \end{aligned}$$

A fraction q_2 of the time is allocated to the three-node network for Phase 2 transmissions and during this time, s transmits messages to r at rate

$$\min\left\{\frac{q_1C_s}{q_2n}, E_s, (\tilde{\alpha}_sD_s + \tilde{\alpha}'_sD'_s + (q_4/\omega - \tilde{\alpha}'_s)F'_s)/q_2\right\} \leq E_s.$$

A fraction $\tilde{\alpha}_s$ of the time is allocated to the three-node network for the first session of Phase 3 and a fraction $(q_3/k - \tilde{\alpha}_s)$ of the time is allocated to the three-node network for the second session of Phase 3. The transmissions in Phase 3 are as follows. In the first session of Phase 3, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station and s at rate

$$\begin{aligned} & \min\left\{\left(\frac{\tilde{\alpha}_sD_s}{\tilde{\alpha}_sD_s + \tilde{\alpha}'_sD'_s + (q_4/\omega - \tilde{\alpha}'_s)F'_s}\right)\left(\frac{q_1C_s}{n\tilde{\alpha}_s}\right), \right. \\ & \quad \left.\left(\frac{\tilde{\alpha}_sD_s}{\tilde{\alpha}_sD_s + \tilde{\alpha}'_sD'_s + (q_4/\omega - \tilde{\alpha}'_s)F'_s}\right)(q_2E_s/\tilde{\alpha}_s), D_s\right\} \\ & \leq D_s. \end{aligned}$$

In the second session of Phase 3, the relay forwards the

messages of s to the base station at rate

$$\min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) \left(\frac{q_1 C_s}{n(q_3/k - \tilde{\alpha}_s)} \right), \right. \\ \left. \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) \left(\frac{q_2 E_s}{q_3/k - \tilde{\alpha}_s} \right), F'_s \right\} \\ \leq F_s.$$

A fraction $\tilde{\alpha}'_s$ of the time is allocated to the three-node network for the first session of Phase 4 and a fraction $(q_4/\omega - \tilde{\alpha}'_s)$ of the time is allocated to the three-node network for the second session of Phase 4. The transmissions in Phase 4 are as follows. In the first session of Phase 4, the relay forwards the messages of the base station to s at rate

$$\min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) \left(\frac{q_1 C_s}{n\tilde{\alpha}'_s} \right), \right. \\ \left. \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) (q_2 E_s / \tilde{\alpha}'_s), D'_s \right\} \\ \leq D'_s.$$

In the second session of Phase 4, the relay performs XOR operations between the messages of the base station and the messages of s bit by bit and broadcasts the resultant messages to both the base station and s at rate

$$\min \left\{ \left(\frac{(q_4/\omega - \tilde{\alpha}'_s) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) \left(\frac{q_1 C_s}{n(q_4/\omega - \tilde{\alpha}'_s)} \right), \right. \\ \left. \left(\frac{(q_4/\omega - \tilde{\alpha}'_s) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s} \right) \left(\frac{q_2 E_s}{q_4/\omega - \tilde{\alpha}'_s} \right), F'_s \right\} \\ \leq F'_s.$$

Under the scheme, the average rates of the messages sent

from the base station to the relay and from s to the relay are both

$$\min\{q_1 C_s/n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s\}.$$

The average rate of the messages sent from the relay to s in Phase 3 and Phase 4 combined is

$$\begin{aligned} \min & \left\{ \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\ & \left. \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \tilde{\alpha}_s D_s \right\} \\ & + \min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\ & \left. \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \tilde{\alpha}'_s D'_s \right\} \\ & + \min \left\{ \left(\frac{\left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\ & \left. \left(\frac{\left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s \right\} \\ & = \min\{q_1 C_s/n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s\}, \end{aligned}$$

where the equality follows from applying (3.13) twice. The average rate of the messages sent from the relay to the base

station in Phase 3 and Phase 4 combined is

$$\begin{aligned}
& \min \left\{ \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\
& \quad \left. \left(\frac{\tilde{\alpha}_s D_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \tilde{\alpha}_s D_s \right\} \\
& + \min \left\{ \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\
& \quad \left. \left(\frac{\tilde{\alpha}'_s D'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \left(\frac{q_3}{k} - \tilde{\alpha}_s\right) F_s \right\} \\
& + \min \left\{ \left(\frac{\left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) \left(\frac{q_1 C_s}{n} \right), \right. \\
& \quad \left. \left(\frac{\left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s}{\tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s} \right) q_2 E_s, \left(\frac{q_4}{\omega} - \tilde{\alpha}'_s\right) F'_s \right\} \\
& = \min\{q_1 C_s/n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s\},
\end{aligned}$$

where the equality follows from applying (3.13) twice and the fact that

$$\tilde{\alpha}'_s D'_s = (q_3/k - \tilde{\alpha}_s) F_s$$

(cf. (3.19)).

In addition, the base station can recover the messages of s in the first session of Phase 3 and in the second session of Phase 4 by performing XOR operations between its own messages and the messages from the relay. Similarly, s can recover the messages of the base station in the first session of Phase 3 and in the second session of Phase 4 by performing XOR operations between its own messages and the messages from the relay. Consequently,

$$\min\{q_1 C_s/n, q_2 E_s, \tilde{\alpha}_s D_s + \tilde{\alpha}'_s D'_s + (q_4/\omega - \tilde{\alpha}'_s) F'_s\},$$

the equal-rate throughput upper bound in (3.20), is achievable by the network coding scheme.

Combining the two subcases in Case 3, we obtain a network coding scheme that achieves the equal-rate throughput upper bound in (3.20), which is equivalent to (3.15).

Case 4: $D_s > F_s$ and $D'_s \leq F'_s$

The broadcast capacities in Phase 3 and Phase 4 are $\min\{D_s, F_s\} = F_s$ and $\min\{D'_s, F'_s\} = D'_s$ respectively. Following similar procedures in Case 3, we can prove that

$$\begin{aligned}
& \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{q_1 C_s/n, q_2 E_s, \alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s) F_s, \\
& \quad \alpha'_s D'_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\
& = \min\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + \\
& \quad (q_3/k - \alpha_s) F_s, \alpha'_s D'_s + (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}\} \\
& \tag{3.24}
\end{aligned}$$

is an upper bound on the equal-rate throughput, and let $(\tilde{\alpha}_s, \tilde{\alpha}'_s)$ denote the argument of

$$\begin{aligned}
& \max_{\substack{(\alpha_s, \alpha'_s): 0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s + (q_3/k - \alpha_s) F_s, \alpha'_s D'_s + \\
& \quad (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}.
\end{aligned}$$

As in Case 3, we can obtain

$$(\tilde{\alpha}_s, \tilde{\alpha}'_s) = \begin{cases} (0, q_4/\omega) & \text{if } D_s D'_s + F_s F'_s - D_s F'_s \geq 0, \\ \left(\frac{q_4 F'_s}{\omega D_s}, 0\right) & \text{if } D_s D'_s + F_s F'_s - D_s F'_s < 0 \\ & \text{and } \frac{q_4}{\omega} - \frac{q_3 D_s}{k F'_s} \leq 0, \\ \left(\frac{q_3}{k}, \frac{q_4}{\omega} - \frac{q_3 D_s}{k F'_s}\right) & \text{if } D_s D'_s + F_s F'_s - D_s F'_s < 0 \\ & \text{and } \frac{q_4}{\omega} - \frac{q_3 D_s}{k F'_s} > 0, \end{cases} \quad (3.25)$$

and can construct a network coding scheme that achieves the equal-rate throughput upper bound in (3.24), which, by applying (3.25), can be simplified to

$$\begin{cases} \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{q_3 F_s}{k} + \frac{q_4 D'_s}{\omega} \right\} & \text{if } D_s D'_s + F_s F'_s - \\ & D_s F'_s \geq 0, \\ \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{k q_4 F'_s (D_s - F_s) + \omega q_3 D_s F_s}{k \omega D_s} \right\} & \text{if } D_s D'_s + F_s F'_s - \\ & D_s F'_s < 0 \text{ and} \\ & \omega q_3 D_s - k q_4 F'_s \geq 0, \\ \min \left\{ \frac{q_1 C_s}{n}, q_2 E_s, \frac{\omega q_3 D_s (F'_s - D'_s) + k q_4 D'_s F'_s}{k \omega F'_s} \right\} & \text{if } D_s D'_s + F_s F'_s - \\ & D_s F'_s < 0 \text{ and} \\ & \omega q_3 D_s - k q_4 F'_s < 0. \end{cases}$$

In each of the four cases above, we obtain an upper bound on the equal-rate throughput and construct a network coding scheme that achieves the upper bound. We call this network coding scheme the *XOR scheme*.

3.3.3 Performance of Routing

Fix a four-phase tuple (q_1, q_2, q_3, q_4) . For every user, we are interested in the maximum equal-rate throughput achievable by routing schemes. For each single-hop user u , the maximum equal-rate throughput achievable by routing schemes is $\min\{p_1 A_u/n, p_2 B_u\}$ as discussed in the previous subsection.

For each multi-hop user s that uses a routing scheme to communicate with the base station through its associated relay, let α_s denote the fraction of time allocated for forwarding the messages of the base station from the relay to s and let β_s denote the fraction of time allocated for forwarding the messages of s from the relay to the base station during Phase 3. Similarly, let α'_s denote the fraction of time allocated for forwarding the messages of the base station from the relay to s and let β'_s denote the fraction of time allocated for forwarding the messages of s from the relay to the base station during Phase 4. Let $k \geq 1$ denote the number of multi-hop users served by the relay and $\omega \geq 1$ denote the number of multi-hop users in the cell containing s . Since the fractions q_1/n , q_2 , q_3/k and q_4/ω of the time are allocated to the corresponding three-node point-to-point relay network for Phase 1, Phase 2, Phase 3 and Phase 4 transmissions respectively, it follows that for every fixed $\alpha_s, \alpha'_s, \beta_s, \beta'_s \geq 0$ that satisfy

$$\alpha_s + \beta_s \leq q_3/k$$

and

$$\alpha'_s + \beta'_s \leq q_4/\omega,$$

the maximum throughput from the base station to s is

$$\min\{q_1 C_s/n, \alpha_s D_s + \alpha'_s D'_s\}$$

and the maximum throughput from s to the base station is

$$\min\{q_2 E_s, \beta_s F_s + \beta'_s F'_s\}.$$

Therefore, the maximum equal-rate throughput achievable by routing schemes is

$$\begin{aligned} & \max_{\substack{\alpha_s, \alpha'_s, \beta_s, \beta'_s \geq 0 \\ \alpha_s + \beta_s = q_3/k \\ \alpha'_s + \beta'_s = q_4/\omega}} \min\{\min\{q_1 C_s/n, \alpha_s D_s + \alpha'_s D'_s\}, \min\{q_2 E_s, \beta_s F_s + \beta'_s F'_s\}\} \\ & = \max_{\substack{\alpha_s, \alpha'_s, \beta_s, \beta'_s \geq 0 \\ \alpha_s + \beta_s = q_3/k \\ \alpha'_s + \beta'_s = q_4/\omega}} \min\{q_1 C_s/n, \alpha_s D_s + \alpha'_s D'_s, q_2 E_s, \beta_s F_s + \beta'_s F'_s\}. \end{aligned}$$

Consequently, the maximum equal-rate throughput achievable by routing schemes is

$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{q_1 C_s/n, \alpha_s D_s + \alpha'_s D'_s, q_2 E_s, (q_3/k - \alpha_s) F_s + \\ & \quad (q_4/\omega - \alpha'_s) F'_s\} \\ & = \min\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3/k - \alpha_s) F_s \\ & \quad + (q_4/\omega - \alpha'_s) F'_s\}\}. \end{aligned} \tag{3.26}$$

In the rest of this subsection, we simplify (3.26), the maximum equal-rate throughput achievable by routing schemes, as shown in Table 3.14.

Let $(\bar{\alpha}_s, \bar{\alpha}'_s)$ denote

$$\arg \max_{\substack{(\alpha_s, \alpha'_s) \\ 0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\}. \tag{3.27}$$

Case	Maximum equal-rate throughput	
$D_s F'_s \leq D'_s F_s$	$\omega q_3 F_s - k q_4 D'_s \geq 0$	$\min\left\{q_1 C_s, 1/n, q_2 E_s, \frac{q_3 D_s F_s / k + q_4 D'_s F_s / \omega}{D_s + F_s}\right\}$
	$\omega q_3 F_s - k q_4 D'_s < 0$	$\min\left\{q_1 C_s, 1/n, q_2 E_s, \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s}\right\}$
$D_s F'_s > D'_s F_s$	$\omega q_3 D_s - k q_4 F'_s \leq 0$	$\min\left\{q_1 C_s, 1/n, q_2 E_s, \frac{q_3 D_s F'_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s}\right\}$
	$\omega q_3 D_s - k q_4 F'_s > 0$	$\min\left\{q_1 C_s, 1/n, q_2 E_s, \frac{q_3 D_s F_s / k + q_4 D_s F'_s / \omega}{D_s + F_s}\right\}$

Table 3.14: The maximum equal-rate throughput achievable by routing schemes.

We claim that

$$\bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s = (q_3/k - \bar{\alpha}_s) F_s + (q_4/\omega - \bar{\alpha}'_s) F'_s$$

and verify the claim as follows. For any $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$ that satisfy

$$\alpha_s D_s + \alpha'_s D'_s > (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,$$

there exist $0 \leq \hat{\alpha}_s \leq \alpha_s$ and $0 \leq \hat{\alpha}'_s \leq \alpha'_s$ such that

$$\hat{\alpha}_s + \hat{\alpha}'_s < \alpha_s + \alpha'_s$$

and

$$\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s = (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s,$$

which implies that

$$\begin{aligned}
& \alpha_s D_s + \alpha'_s D'_s \\
& > \hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s \\
& = (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\
& > (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,
\end{aligned}$$

which then implies that

$$\begin{aligned}
& \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s\} \\
& < \min\{\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s, (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s\}. \quad (3.28)
\end{aligned}$$

Similarly, for any $0 \leq \alpha_s \leq q_3/k$ and $0 \leq \alpha'_s \leq q_4/\omega$ that satisfy

$$\alpha_s D_s + \alpha'_s D'_s < (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,$$

there exist $\alpha_s \leq \hat{\alpha}_s \leq q_3/k$ and $\alpha'_s \leq \hat{\alpha}'_s \leq q_4/\omega$ such that

$$\hat{\alpha}_s + \hat{\alpha}'_s > \alpha_s + \alpha'_s$$

and

$$\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s = (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s,$$

which implies that

$$\begin{aligned}
& \alpha_s D_s + \alpha'_s D'_s \\
& < \hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s \\
& = (q_3/k - \hat{\alpha}_s) F_s + (q_4/\omega - \hat{\alpha}'_s) F'_s \\
& < (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s,
\end{aligned}$$

which then implies that

$$\begin{aligned} & \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3/k - \alpha_s)F_s + (q_4/\omega - \alpha'_s)F'_s\} \\ & < \min\{\hat{\alpha}_s D_s + \hat{\alpha}'_s D'_s, (q_3/k - \hat{\alpha}_s)F_s + (q_4/\omega - \hat{\alpha}'_s)F'_s\}. \end{aligned} \quad (3.29)$$

It then follows from (3.28) and (3.29) that

$$\bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s = (q_3/k - \bar{\alpha}_s)F_s + (q_4/\omega - \bar{\alpha}'_s)F'_s, \quad (3.30)$$

which implies that

$$\bar{\alpha}'_s = \frac{q_3 F_s / k + q_4 F'_s / \omega - \bar{\alpha}_s (D_s + F_s)}{D'_s + F'_s}. \quad (3.31)$$

Consequently, using (3.27), (3.30) and (3.31), we obtain

$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3 - \alpha_s)F_s + (q_4 - \alpha'_s)F'_s\} \\ & \stackrel{(a)}{=} \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} (\alpha_s D_s + \alpha'_s D'_s) \\ & \quad \alpha_s D_s + \alpha'_s D'_s = (q_3/k - \alpha_s)F_s + (q_4/\omega - \alpha'_s)F'_s \\ & \stackrel{(b)}{=} \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \frac{q_3 F_s / k + q_4 F'_s / \omega - \alpha_s (D_s + F_s)}{D'_s + F'_s} \leq q_4/\omega}} \left(\alpha_s D_s + \right. \\ & \quad \left. \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega - \alpha_s D'_s (D_s + F_s)}{D'_s + F'_s} \right) \\ & = \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ \frac{q_3 F_s / k - q_4 D'_s / \omega}{D'_s + F'_s} \leq \alpha_s \leq \frac{q_3 F_s / k + q_4 F'_s / \omega}{D'_s + F'_s}}} \left(\alpha_s D_s - \right. \\ & \quad \left. \frac{\alpha_s D'_s (D_s + F_s)}{D'_s + F'_s} + \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s} \right) \\ & = \max_{\alpha_s} \left\{ 0, \frac{q_3 F_s / k - q_4 D'_s / \omega}{D'_s + F'_s} \right\} \leq \alpha_s \leq \min \left\{ q_3/k, \frac{q_3 F_s / k + q_4 F'_s / \omega}{D'_s + F'_s} \right\} \\ & \quad \left(\frac{\alpha_s (D_s F'_s - D'_s F_s)}{D'_s + F'_s} + \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s} \right), \end{aligned} \quad (3.32)$$

where (a) follows from (3.30) and (b) follows from eliminating α'_s by the equation

$$\alpha_s D_s + \alpha'_s D'_s = (q_3/k - \alpha_s) F_s + (q_4/\omega - \alpha'_s) F'_s.$$

Under each of the following cases, we will first obtain $(\bar{\alpha}_s, \bar{\alpha}'_s)$ through investigating (3.32), followed by simplifying the maximum equal-rate throughput (3.26) achievable by routing schemes.

Case $D_s F'_s \leq D'_s F_s$:

Since the slope of the function in (3.32) is less than or equal to 0, the maximum equal rate throughput in (3.32) is achieved by the smallest possible α_s . Consider the following two subcases:

Subcase $\omega q_3 F_s - k q_4 D'_s \geq 0$:

Since

$$\begin{aligned} & \max \left\{ 0, \frac{q_3 F_s / k - q_4 D'_s / \omega}{D_s + F_s} \right\} \\ &= \max \left\{ 0, \frac{\omega q_3 F_s - k q_4 D'_s}{k \omega (D_s + F_s)} \right\} \\ &= \frac{q_3 F_s / k - q_4 D'_s / \omega}{D_s + F_s}, \end{aligned}$$

it follows that the smallest possible α_s for (3.32) is $\frac{q_3 F_s / k - q_4 D'_s / \omega}{D_s + F_s}$, which implies that

$$\bar{\alpha}_s = \frac{q_3 F_s / k - q_4 D'_s / \omega}{D_s + F_s}$$

and

$$\bar{\alpha}'_s = q_4 / \omega$$

(cf. (3.31)). It then follows from (3.30) that

$$\begin{aligned}
& \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3 - \alpha_s)F_s + (q_4 - \alpha'_s)F'_s\} \\
&= \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s \\
&= \frac{q_3 D_s F_s / k - q_4 D_s D'_s / \omega}{D_s + F_s} + q_4 D'_s / \omega \\
&= \frac{q_3 D_s F_s / k + q_4 D'_s F_s / \omega}{D_s + F_s},
\end{aligned}$$

which implies from (3.26) that the maximum equal-rate throughput achievable by routing schemes is

$$\min\{q_1 C_s / \pi, q_2 E_s, \frac{q_3 D_s F_s / k + q_4 D'_s F_s / \omega}{D_s + F_s}\}.$$

Subcase $\omega q_3 F_s - k q_4 D'_s < 0$:

Since

$$\begin{aligned}
& \max\left\{0, \frac{q_3 F_s / k - q_4 D'_s / \omega}{D_s + F_s}\right\} \\
&= \max\left\{0, \frac{\omega q_3 F_s - k q_4 D'_s}{k \omega (D_s + F_s)}\right\} \\
&= 0,
\end{aligned}$$

it follows that the smallest possible α_s for (3.32) is 0, which implies that

$$\bar{\alpha}_s = 0$$

and

$$\bar{\alpha}'_s = \frac{q_3 F_s / k + q_4 F'_s / \omega}{D'_s + F'_s}$$

(cf. (3.31)). It then follows from (3.30) that

$$\begin{aligned}
 & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3 - \alpha_s)F_s + (q_4 - \alpha'_s)F'_s\} \\
 &= \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s \\
 &= \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s},
 \end{aligned}$$

which implies from (3.26) that the maximum equal-rate throughput achievable by routing schemes is

$$\min\{q_1 C_s / n, q_2 E_s, \frac{q_3 D'_s F_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s}\}.$$

Case $D_s F'_s > D'_s F_s$:

Since the slope of the function in (3.32) is greater than 0, the maximum equal rate throughput (3.32) is achieved by the largest possible α_s . Consider the following two subcases:

Subcase $\omega q_3 D_s - k q_4 F'_s \leq 0$:

Since

$$\begin{aligned}
 & \min \left\{ q_3/k, \frac{q_3 F_s / k + q_4 F'_s / \omega}{D_s + F_s} \right\} \\
 &= \min \left\{ \frac{\omega q_3 D_s + \omega q_3 F_s}{k\omega(D_s + F_s)}, \frac{\omega q_3 F_s + k q_4 F'_s}{k\omega(D_s + F_s)} \right\} \\
 &= \frac{\omega q_3 D_s + \omega q_3 F_s}{k\omega(D_s + F_s)} \\
 &= q_3/k,
 \end{aligned}$$

it follows that the largest possible α_s for (3.32) is q_3/k , which implies that

$$\bar{\alpha}_s = q_3/k$$

and

$$\bar{\alpha}'_s = \frac{q_4 F'_s / \omega - q_3 D_s / k}{D'_s + F'_s}$$

(cf. (3.31)). It then follows from (3.30) that

$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3 - \alpha_s)F_s + (q_4 - \alpha'_s)F'_s\} \\ &= \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s \\ &= q_3 D_s / k + \frac{q_4 D'_s F'_s / \omega - q_3 D_s D'_s / k}{D'_s + F'_s} \\ &= \frac{q_3 D_s F'_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s}, \end{aligned}$$

which implies from (3.26) that the maximum equal-rate throughput achievable by routing schemes is

$$\min\{q_1 C_s / n, q_2 E_s, \frac{q_3 D_s F'_s / k + q_4 D'_s F'_s / \omega}{D'_s + F'_s}\}.$$

Subcase $\omega q_3 D_s - k q_4 F'_s > 0$:

Since

$$\begin{aligned} & \min\left\{q_3/k, \frac{q_3 F_s / k + q_4 F'_s / \omega}{D_s + F_s}\right\} \\ &= \min\left\{\frac{\omega q_3 D_s + \omega q_3 F_s}{k\omega(D_s + F_s)}, \frac{\omega q_3 F_s + k q_4 F'_s}{k\omega(D_s + F_s)}\right\} \\ &= \frac{\omega q_3 F_s + k q_4 F'_s}{k\omega(D_s + F_s)} \\ &= \frac{q_3 F_s / k + q_4 F'_s / \omega}{D_s + F_s}, \end{aligned}$$

it follows that the largest possible α_s for (3.30) is $\frac{q_3 F_s / k + q_4 F'_s / \omega}{D_s + F_s}$, which implies that

$$\bar{\alpha}_s = \frac{q_3 F_s / k + q_4 F'_s / \omega}{D_s + F_s}$$

and

$$\bar{\alpha}'_s = 0$$

(cf. (3.31)). It then follows from (3.30) that

$$\begin{aligned} & \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3 - \alpha_s)F_s + (q_4 - \alpha'_s)F'_s\} \\ &= \bar{\alpha}_s D_s + \bar{\alpha}'_s D'_s \\ &= \frac{q_3 D_s F_s / k + q_4 D_s F'_s / \omega}{D_s + F_s}, \end{aligned}$$

which implies from (3.26) that the maximum equal-rate throughput achievable by routing schemes is

$$\min\{q_1 C_s / n, q_2 E_s, \frac{q_3 D_s F_s / k + q_4 D_s F'_s / \omega}{D_s + F_s}\}.$$

Under each of the above cases, we have obtained the maximum equal-rate throughput achievable by routing schemes. We call the routing scheme that achieves the maximum equal-rate throughput the *optimal routing scheme*.

3.3.4 Network Coding Gain for Individual User

For each single-hop user u , since u communicates with the base station directly, there is no network coding gain in the maximum equal-rate throughput, which has been shown in Section 3.3.2 to be

$$\min\{q_1 A_u / n, q_2 B_u\}.$$

For each multi-hop user s that communicates with the base station b through its associated relay r , the equal-rate throughput achievable by the XOR scheme has been shown in Section 3.3.2

to be

$$\left\{ \begin{array}{ll}
 \min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq \frac{q_3}{k} \\ 0 \leq \alpha'_s \leq \frac{q_4}{\omega}}} \min\{\alpha_s D_s + \alpha'_s D'_s, \right. & \text{if } D_s \leq F_s \text{ and} \\
 \left. \alpha_s D_s + \alpha'_s D'_s + \left(\frac{q_3}{k} - \alpha_s\right) F_s + \left(\frac{q_4}{\omega} - \alpha'_s\right) F'_s\right\} & D'_s \leq F'_s, \\
 \\
 \min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq \frac{q_3}{k} \\ 0 \leq \alpha'_s \leq \frac{q_4}{\omega}}} \min\{\alpha_s D_s + \alpha'_s D'_s + \right. & \text{if } D_s > F_s \text{ and} \\
 \left. \left(\frac{q_3}{k} - \alpha_s\right) F_s + \left(\frac{q_4}{\omega} - \alpha'_s\right) F'_s, \left(\frac{q_3}{k} - \alpha_s\right) F_s + \left(\frac{q_4}{\omega} - \right. & D'_s > F'_s, \\
 \left. \alpha'_s\right) F'_s\right\} & \\
 \\
 \min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq \frac{q_3}{k} \\ 0 \leq \alpha'_s \leq \frac{q_4}{\omega}}} \min\{\alpha_s D_s + \alpha'_s D'_s + \right. & \text{if } D_s \leq F_s \text{ and} \\
 \left. \left(\frac{q_4}{\omega} - \alpha'_s\right) F'_s, \alpha_s D_s + \left(\frac{q_3}{k} - \alpha_s\right) F_s + \left(\frac{q_4}{\omega} - \alpha'_s\right) F'_s\right\} & D'_s > F'_s, \\
 \\
 \min\left\{\frac{q_1 C_s}{n}, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq \frac{q_3}{k} \\ 0 \leq \alpha'_s \leq \frac{q_4}{\omega}}} \min\{\alpha_s D_s + \alpha'_s D'_s + \right. & \text{if } D_s > F_s \text{ and} \\
 \left. \left(\frac{q_3}{k} - \alpha_s\right) F_s, \alpha'_s D'_s + \left(\frac{q_3}{k} - \alpha_s\right) F_s + \left(\frac{q_4}{\omega} - \alpha'_s\right) F'_s\right\} & D'_s \leq F'_s
 \end{array} \right. \quad (3.33)$$

(cf. (3.12), (3.14), (3.15) and (3.24)), which is simplified in Section 3.3.2 as shown in Table 3.13.

The equal-rate throughput achievable by the optimal routing scheme for s has been shown in Section 3.3.3 to be

$$\min\left\{q_1 C_s/n, q_2 E_s, \max_{\substack{0 \leq \alpha_s \leq q_3/k \\ 0 \leq \alpha'_s \leq q_4/\omega}} \min\{\alpha_s D_s + \alpha'_s D'_s, (q_3/k - \alpha_s) F_s + \right. \\
 \left. (q_4/\omega - \alpha'_s) F'_s\right\} \quad (3.34)$$

(cf. (3.26)), which is simplified in Section 3.3.3 as shown in Table 3.14.

We have shown in Section 3.3.2 that the equal-rate throughput achievable by the XOR scheme is an upper bound on the equal-rate throughput achievable by network coding schemes, including the

optimal routing scheme as a special case. Therefore, the equal-rate throughput achievable by the XOR scheme is larger than or equal to the maximum equal-rate throughput achievable by the optimal routing scheme, which can also be observed by comparing (3.34) with (3.33). We have obtained a necessary and sufficient condition for non-zero network coding gain under the three-phase mode in Section 3.2.4. However, we cannot obtain a similar condition for the four-phase mode due to high complexity.

3.3.5 Average Network Coding Gain

The network coding gain for each user in the cellular relay network has been discussed in the previous subsection. In this subsection, we consider the average equal-rate throughput over all users under the following two scenarios:

Scenario 1: Each single-hop user u in the cellular relay network communicates with the base station directly with equal-rate throughput $\min\{q_1 A_u/n, q_2 B_u\}$. Each multi-hop user s in the cellular relay network uses the optimal routing scheme, which depends on the values of $q_3, q_4, k, \omega, D_s, D'_s, F_s$ and F'_s (cf. Section 3.3.3). The equal-rate throughput achievable by the optimal routing scheme is shown in Table 3.14.

Scenario 2: Each single-hop user u in the cellular relay network communicates with the base station directly with equal-rate throughput $\min\{q_1 A_u/n, q_2 B_u\}$. Each multi-hop user in the cellular relay network uses the XOR scheme, which depends on the values of $q_3, q_4, k, \omega, D_s, D'_s, F_s$ and F'_s (cf. Section 3.3.2). The equal-rate throughput achievable by the XOR scheme is shown in Table 3.13.

We are interested in the average equal-rate throughput over all users under Scenario 1 and Scenario 2 for every $n = 1, 4, 10, 20, 30$ (n is the number of users in each cell), every $m = 0, 2, 5, 10, 20, 30$ (m is the number of relays in each cell) and every four-phase tuple in the set

$$\left\{ (q_1, q_2, q_3, q_4) \in \mathbb{R}^4 \left\{ \begin{array}{l} q_1 \geq 0, q_2 \geq 0, q_3 \geq 0, q_4 \geq 0, \\ q_1 + q_2 + q_3 + q_4 = 1, \\ q_1, q_2, q_3 \text{ and } q_4 \text{ are multiples of } 1/64. \end{array} \right. \right\}. \quad (3.35)$$

The four-phase tuple in (3.35) lie in the three-dimensional space

$$q_1 + q_2 + q_3 + q_4 = 1.$$

When $n = 1$, the cellular relay network resembles low-density cellular networks. When $n = 4$ or 10 , the cellular relay network resembles medium-density cellular networks. When $n = 20$ or 30 , the cellular relay network resembles high-density cellular networks.

We generate 6000 different networks for each n and each m in our simulation. In each network, we calculate the average equal-rate throughput over all users in the central cell under Scenario 1 and Scenario 2 for every (q_1, q_2, q_3, q_4) according to the formulae in Section 3.3.4. Then, we calculate the sample mean of the 6000 average equal-rate throughput under Scenario 1 and Scenario 2 for each n , each m and each (q_1, q_2, q_3, q_4) . Using standard arguments, we obtain that a 95% confidence interval for each sample mean μ is $(0.99\mu, 1.01\mu)$.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64, 0, 0)	(20/64, 44/64, 0, 0)	(18/64, 38/64, 0, 8/64)	(19/64, 30/64, 6/64, 9/64)	(20/64, 22/64, 0, 22/64)	(20/64, 20/64, 0, 24/64)
$n = 4$	(29/64, 35/64, 0, 0)	(28/64, 32/64, 1/64, 3/64)	(29/64, 29/64, 2/64, 6/64)	(28/64, 23/64, 3/64, 10/64)	(27/64, 17/64, 5/64, 15/64)	(27/64, 14/64, 5/64, 18/64)
$n = 10$	(31/64, 33/64, 0, 0)	(30/64, 30/64, 1/64, 3/64)	(31/64, 24/64, 2/64, 7/64)	(31/64, 19/64, 3/64, 11/64)	(30/64, 14/64, 3/64, 17/64)	(29/64, 12/64, 4/64, 19/64)
$n = 20$	(30/64, 34/64, 0, 0)	(30/64, 30/64, 1/64, 3/64)	(32/64, 24/64, 2/64, 6/64)	(32/64, 23/64, 2/64, 12/64)	(31/64, 13/64, 3/64, 17/64)	(30/64, 11/64, 3/64, 20/64)
$n = 30$	(29/64, 35/64, 0, 0)	(31/64, 29/64, 1/64, 3/64)	(33/64, 22/64, 2/64, 7/64)	(33/64, 17/64, 2/64, 12/64)	(31/64, 14/64, 3/64, 16/64)	(30/64, 11/64, 3/64, 20/64)

Table 3.15: Four-phase tuple $\pi_{m,n}^*$ maximizing average equal-rate throughput under Scenario 1.

For each n , each m and each (q_1, q_2, q_3, q_4) , the equal-rate throughput for a single-hop user under Scenario 1 and Scenario 2 are the same. In addition, it follows from Section 3.3.4 that for a multi-hop user, the equal-rate throughput under Scenario 1 is less than or equal to the equal-rate throughput under Scenario 2. Consequently, the sample mean of the average equal-rate throughput under Scenario 1 is less than or equal to the sample mean of the average equal-rate throughput under Scenario 2.

For each n and each m , let $\pi_{m,n}^*$ denote the four-phase tuple in (3.35) that maximizes the sample mean of the average equal-rate throughput under Scenario 1, and the $\pi_{m,n}^*$'s are displayed in Table 3.15. For each n and each m , let $\theta_{m,n}^*$ denote the four-phase tuple in (3.35) that maximizes the sample mean of the average equal-rate throughput under Scenario 2, and the $\theta_{m,n}^*$'s are displayed in Table 3.16.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64, 0, 0)	(19/64, 42/64, 0, 3/64)	(19/64, 37/64, 0, 8/64)	(22/64, 29/64, 0, 13/64)	(23/64, 23/64, 0, 18/64)	(23/64, 22/64, 0, 19/64)
$n = 4$	(29/64, 35/64, 0, 0)	(28/64, 32/64, 0, 4/64)	(29/64, 28/64, 0, 7/64)	(29/64, 23/64, 0, 12/64)	(29/64, 18/64, 0, 17/64)	(28/64, 15/64, 0, 21/64)
$n = 10$	(31/64, 33/64, 0, 0)	(31/64, 30/64, 0, 3/64)	(32/64, 25/64, 0, 7/64)	(32/64, 20/64, 0, 12/64)	(31/64, 15/64, 0, 18/64)	(31/64, 12/64, 0, 21/64)
$n = 20$	(30/64, 34/64, 0, 0)	(31/64, 29/64, 0, 4/64)	(33/64, 23/64, 0, 8/64)	(33/64, 18/64, 0, 13/64)	(32/64, 14/64, 0, 18/64)	(32/64, 12/64, 0, 21/64)
$n = 30$	(29/64, 35/64, 0, 0)	(32/64, 28/64, 0, 4/64)	(33/64, 23/64, 0, 8/64)	(33/64, 18/64, 0, 13/64)	(32/64, 14/64, 0, 18/64)	(31/64, 12/64, 0, 21/64)

Table 3.16: Four-phase tuple $\theta_{m,n}^*$ maximizing average equal-rate throughput under Scenario 2.

Under the four-phase mode with four-phase tuple $\pi_{m,n}^*$, we call the sample mean of the average equal-rate throughput under Scenario 1 the *best routing throughput for n users and m relays*. Under the four-phase mode with four-phase tuple $\theta_{m,n}^*$, we call the sample mean of the average equal-rate throughput under Scenario 2 the *best XOR throughput for n users and m relays*. The best routing throughput for each n and each m is displayed in Table 3.17 and the best XOR throughput for each n and each m is displayed in Table 3.18. The percentage of multi-hop users for each n and each m was shown in Table 3.8, which is reproduced in Table 3.19 for convenience.

Since the three-phase mode with any three-phase tuple (p_1, p_2, p_r) in (3.11) in Section 3.2.5 is equivalent to the four-phase mode with some four-phase tuple $(p_1, p_2, p_r, 0)$ in (3.35), it follows that for each n and each m , the best routing throughput under the three-phase mode obtained in Section 3.2.5 is less than or equal to the best

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9385	3.9460	4.1817	4.9604	5.5314
$n = 4$	1.2556	1.2639	1.3490	1.5127	1.7428	1.9279
$n = 10$	0.4210	0.4462	0.5025	0.5841	0.7029	0.7870
$n = 20$	0.1704	0.1895	0.2204	0.2655	0.3304	0.3733
$n = 30$	0.0994	0.1133	0.1344	0.1655	0.2093	0.2385

Table 3.17: Best routing throughput for n users and m relays (10^6 bits per second).

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9477	4.2514	4.7112	5.7895	6.4766
$n = 4$	1.2556	1.3093	1.4190	1.5960	1.8630	2.0695
$n = 10$	0.4210	0.4575	0.5184	0.6072	0.7372	0.8280
$n = 20$	0.1704	0.1928	0.2257	0.2743	0.3438	0.3893
$n = 30$	0.0994	0.1150	0.1374	0.1705	0.2173	0.2480

Table 3.18: Best XOR throughput for n users and m relays (10^6 bits per second).

routing throughput under the four-phase mode, which can be verified by comparing Table 3.6 with Table 3.17. Similarly, for each n and each m , the best XOR throughput under the three-phase mode obtained in Section 3.2.5 is less than or equal to the best XOR throughput under the four-phase mode, which can be verified by comparing Table 3.7 with Table 3.18.

Under the case $m = 0$, there is no relay in the cellular relay network, which implies that all the users in the network are single-hop users. Therefore, the best routing throughput and the best XOR throughput for each n are the same, and they are indeed the average equal-rate capacities over all users in the network without

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	0%	8.17%	18.80%	31.63%	52.25%	64.83%
$n = 4$	0%	7.93%	17.55%	32.29%	52.70%	65.62%
$n = 10$	0%	7.62%	18.22%	32.98%	52.73%	65.33%
$n = 20$	0%	7.91%	17.94%	32.46%	52.96%	65.46%
$n = 30$	0%	7.89%	18.11%	32.43%	52.48%	65.40%

Table 3.19: Percentage of multi-hop users.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	0%	0.24%	7.74%	12.66%	16.71%	17.09%
$n = 4$	0%	3.59%	5.19%	5.50%	6.90%	7.35%
$n = 10$	0%	2.53%	3.16%	3.96%	4.87%	5.20%
$n = 20$	0%	1.79%	2.45%	3.31%	4.05%	4.27%
$n = 30$	0%	1.49%	2.29%	3.02%	3.80%	3.96%

Table 3.20: Average network coding percentage gain over all users.

	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	2.88%	41.17%	40.03%	31.99%	26.36%
$n = 4$	45.26%	29.56%	17.04%	13.09%	11.19%
$n = 10$	33.27%	17.32%	12.01%	9.23%	7.97%
$n = 20$	22.60%	13.66%	10.21%	7.64%	6.53%
$n = 30$	18.92%	12.64%	9.31%	7.25%	6.06%

Table 3.21: Average network coding percentage gain over multi-hop users.

relay. Comparing the column $m = 0$ with the other columns in Table 3.17, we observe that the best routing throughput over all users when the number of relays exceeds 10 is always greater than the average equal-rate capacities over all users when there is no relay. Similarly, we observe in Table 3.18 that the best XOR throughput over all users when the number of relays exceeds 5 is always greater than the average equal-rate capacities over all users when there is no relay.

Table 3.17 and Table 3.18 show that for each n , both the best routing throughput and the best XOR throughput generally increase as m increases, which implies that adding more relays to the network improves both the best routing throughput and the best XOR throughput. For each m , both the best routing throughput and the XOR throughput decrease as n increases, which agrees with the fact that adding more users to the network results in

1. less time allocated to each user for Phase 1 transmissions;
2. larger interference incurred by the users on each receiving node during Phase 2 (cf. Table 3.11 and Table 3.12 in Section 3.3.1).

Table 3.17 and Table 3.18 show that the use of network coding rather than routing alone increases the average maximum equal-rate throughput over all users in the cellular relay network. For each n and each m , we call the percentage gain by which the best XOR throughput is greater than the best routing throughput the *average network coding percentage gain over all users*, which is shown in Table 3.20. For each n and each $m \neq 0$, we call the average network coding percentage gain over all users divided by the percentage of multi-hop users the *average network coding percentage gain over multi-hop users*, which is shown in Table 3.21.

Table 3.20 and Table 3.21 show that for each m , both the average network coding gain over all users and the average network coding gain over multi-hop users generally decrease as n increases. Table 3.20 shows that for each n , the average network coding gain over all users increases as m increases. Table 3.21 shows that for each n , the average network coding gain over multi-hop users generally decreases as m increases. For low-density networks, i.e. $n = 1$, the average network coding gain over all users can be greater than 17% and the average network coding gain over multi-hop users can be greater than 41%. For medium-density networks, i.e. $n = 4$ or 10, the average network coding gain over all users can be greater than 7% and the average network coding gain over multi-hop users can be greater than 45%. For high-density networks, i.e. $n = 20$ or 30, the average network coding gain over all users can be greater than 4% and the average network coding gain over multi-hop users can be greater than 22%.

□ End of chapter.

Part II

Physical-Layer Network Coding

Chapter 4

Capacities of Two-Way Relay Channels

Summary

We introduce in Sections 4.3, 4.4 and 4.5 three models of the full-duplex two-way relay channel (TRC), namely the discrete memoryless TRC without feedback, the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. We establish in Section 4.3 an outer bound on the capacity region of the discrete memoryless TRC without feedback. For each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback, Knopp [22] stated an outer bound on the capacity region. However, the proofs in [22] are incomplete and we therefore present in Sections 4.4 and 4.5 detailed proofs of the outer bounds stated in [22]. The proofs are nontrivial. Then, we consider the half-duplex discrete memoryless TRC and the half-duplex Gaussian TRC in Sections 4.6 and 4.7 and prove several outer bounds on the capacity regions of the TRCs. In particular, an outer bound on the capacity region achievable by PNC schemes for the bandlimited Gaussian TRC is proved in Section 4.6.4.

In Chapter 2 and Chapter 3, we have studied the benefits of symbol-level network coding on a TRC as well as a cellular relay network. In this chapter, we study the performance of physical-layer network coding on a TRC. Zhang *et al.* [16] showed that PNC can potentially achieve higher maximum equal-rate throughput than symbol-level network coding for some Gaussian TRC. Therefore, we are interested in obtaining an outer bound on the rate region achievable by PNC schemes for the TRC.

We begin the investigation of PNC by studying three models of full-duplex TRC, namely the discrete memoryless TRC without feedback, the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. We establish an outer bound on the capacity region of the discrete memoryless TRC without feedback. For each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback, Knopp [22] stated an outer bound on the capacity region. However, the proofs in [22] are incomplete and we therefore present detailed proofs of the outer bounds stated therein. The proofs are nontrivial.

Then, we extend the results for the full-duplex TRCs to the half-duplex discrete memoryless TRC and the half-duplex Gaussian TRC, and obtain several outer bounds on the capacity regions of the half-duplex TRCs. In particular, an outer bound on the capacity region achievable by PNC schemes for the half-duplex bandlimited Gaussian TRC is obtained.

4.1 Notation

We use $Pr\{E\}$ to represent the probability of an event E . We use a capital letter X to denote a random variable with alphabet \mathcal{X} , and use the small letter x to denote the realization of X . We use $E[X]$ to represent the expectation of a random variable X . We use X^n to denote a random column vector $[X_1, X_2, \dots, X_n]^T$, where the components X_k have the same alphabet. We let $p_X(x)$ and $p_{X^n}(x^n)$ denote the probability mass functions of the discrete random variables X and X^n respectively. We let $p_{Y|X}(y|x)$ denote the conditional probability $Pr\{Y = y|X = x\}$ for any general random variable X and any discrete random variable Y . We let $f_X(x)$ and $f_{X^n}(x^n)$ denote the probability density functions of the random variables X and X^n respectively. We let $f_{Y|X}(y|x)$ denote the probability density function of Y conditioned on the event $\{X = x\}$ for any continuous random variable Y and any general random variable X . For simplicity, we drop the subscript of a notation if there is no ambiguity.

Let $H(X)$ and $H(X|Y) = E[H(X|Y = y)]$ be the entropy of X and the entropy of X conditioned on Y respectively for any discrete random variable X and any general random variable Y . Let $h(X)$ and $h(X|Y) = E[h(X|Y = y)]$ be the differential entropy of X and the differential entropy of X conditioned on Y respectively for any continuous random variable X and any general random variable Y . Let $I(X; Y) = E\left[\log_2\left(\frac{p(x|y)}{p(x)}\right)\right]$ be the mutual information of X and Y , where X is a discrete random variable and Y is either a discrete or a continuous random variable. Let $I(X; Y|Z) = E[I(X; Y|Z = z)]$ be the mutual information of X and Y conditioned on Z for any general random variable Z . The closure of a set S is denoted by \bar{S} and the convex hull of S is denoted

by $\text{conv}(S)$.

4.2 Full-Duplex Two-Way Relay Channel

We assume in Sections 4.2, 4.3, 4.4 and 4.5 that all the nodes in the TRC are full-duplex, which means they can transmit and receive information at the same time. The results for the full-duplex TRCs are extended to half-duplex TRCs in Sections 4.6 and 4.7.

The TRC consists of two terminal nodes t_1 and t_2 and a relay node r in the middle. Node t_1 and node t_2 do not communicate directly, but communicate through node r using two different channels. In each time slot, node t_1 and node t_2 transmit a symbol to node r through the Multiple Access Channel (MAC), and node r transmits a symbol to node t_1 and node t_2 through the Broadcast Channel (BC). The MAC and the BC do not interfere with each other. Suppose node t_1 and node t_2 exchange information in n time slots. Then, the variables involved in the communication through the TRC are:

- $W_i \in \{1, 2, \dots, M_i\}$: messages of node t_i , $i = 1, 2$,
- $X_i^n = [X_{i,1} \ X_{i,2} \ \dots \ X_{i,n}]^T$: codewords transmitted by node t_i , $i = 1, 2$,
- $Y_r^n = [Y_{r,1} \ Y_{r,2} \ \dots \ Y_{r,n}]^T$: channel outputs at the relay,
- $X_r^n = [X_{r,1} \ X_{r,2} \ \dots \ X_{r,n}]^T$: codewords transmitted by the relay,
- $Y_i^n = [Y_{i,1} \ Y_{i,2} \ \dots \ Y_{i,n}]^T$: channel outputs at node t_i , $i = 1, 2$,
- $\hat{W}_i \in \{1, 2, \dots, M_i\}$: message estimates of message W_i , $i = 1, 2$.

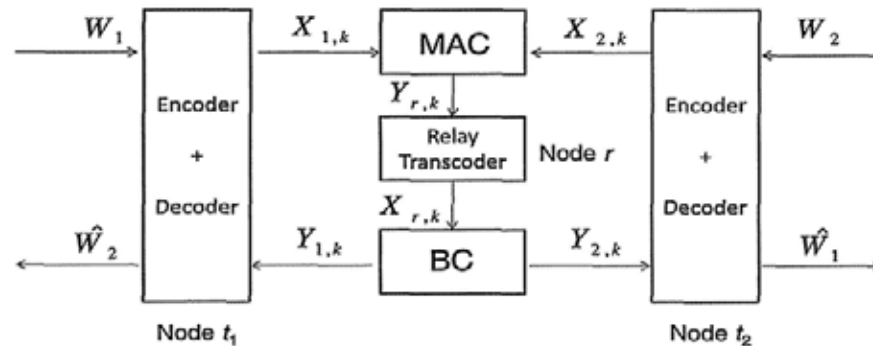


Figure 4.1: A full-duplex discrete memoryless TRC.

Node t_1 and node t_2 choose messages W_1 and W_2 independently according to the uniform distribution, and node t_1 declares the message estimate \hat{W}_2 and node t_2 declares the message estimate \hat{W}_1 after n time slots. The transmissions of messages in the k^{th} time slot are shown in Figure 4.1.

4.3 Full-Duplex Discrete Memoryless Two-Way Relay Channel without Feedback

Unless otherwise specified, the two-way communication in this section is assumed to be restricted in the classical sense [23], meaning that the terminal nodes cannot use their previously received messages for encoding their messages. The TRC described above is said to be used *without feedback*.

In this section, we will present the problem formulation of the discrete memoryless TRC without feedback and establish an outer bound on the capacity region of the discrete memoryless TRC without feedback.

4.3.1 Definitions

Based on message W_1 , node t_i constructs codeword X_i^n for $i = 1, 2$ and transmits $X_{i,k}$ through the MAC in the k^{th} time slot. The received symbol in the k^{th} time slot at r is $Y_{r,k}$ and in the same time slot, r constructs and transmits $X_{r,k}$ through the BC, which depends on $Y_{r,1}, Y_{r,2}, \dots, Y_{r,k-1}$. The received symbol in the k^{th} time slot at node t_i is $Y_{i,k}$ for $i = 1, 2$. After n time slots, t_1 declares \hat{W}_2 to be the transmitted W_2 based on Y_1^n and W_1 , and node t_2 declares \hat{W}_1 to be the transmitted W_1 based on Y_2^n and W_2 .

Definition 11 *The discrete memoryless TRC without feedback consists of six finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, \mathcal{Y}_1, \mathcal{Y}_2$ and \mathcal{Y}_r , one probability mass function $p_1(y_r|x_1, x_2)$ representing the MAC and one probability mass function $p_2(y_1, y_2|x_r)$ representing the BC such that for any integer n , $x_1^n \in \mathcal{X}_1^n$, $x_2^n \in \mathcal{X}_2^n$, $x_r^n \in \mathcal{X}_r^n$, $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$ and $y_r^n \in \mathcal{Y}_r^n$,*

$$p^{(n)}(y_r^n|x_1^n, x_2^n) = \prod_{k=1}^n p_1(y_{r,k}|x_{1,k}, x_{2,k}) \quad (4.1)$$

and

$$p^{(n)}(y_1^n, y_2^n|x_r^n) = \prod_{k=1}^n p_2(y_{1,k}, y_{2,k}|x_{r,k}). \quad (4.2)$$

The discrete memoryless TRC without feedback is denoted by $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, p_1(y_r|x_1, x_2), p_2(y_1, y_2|x_r), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_r)$.

Definition 12 *An (n, M_1, M_2) -code on the channel $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, p_1(y_r|x_1, x_2), p_2(y_1, y_2|x_r), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_r)$ consists of the following:*

1. A message set $\mathcal{W}_1 = \{1, 2, \dots, M_1\}$ at node t_1 and a message set $\mathcal{W}_2 = \{1, 2, \dots, M_2\}$ at node t_2 .

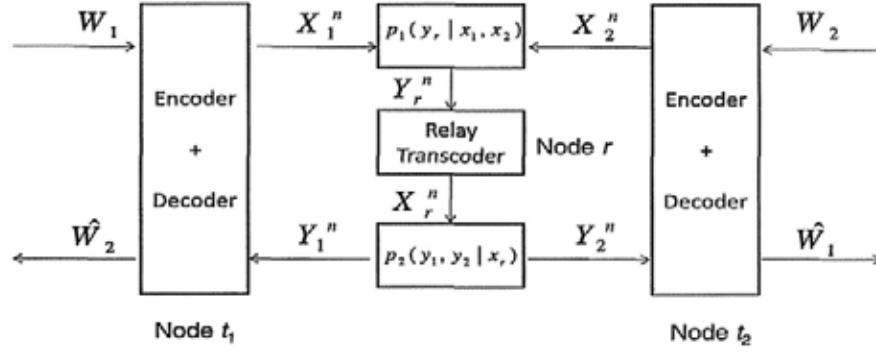


Figure 4.2: A discrete memoryless TRC without feedback.

2. An encoding function $f_1 : \mathcal{W}_1 \rightarrow \mathcal{X}_1^n$ at node t_1 , yielding code-words $x_1^n(1), x_1^n(2), \dots, x_1^n(M_1)$.
3. An encoding function $f_2 : \mathcal{W}_2 \rightarrow \mathcal{X}_2^n$ at node t_2 , yielding code-words $x_2^n(1), x_2^n(2), \dots, x_2^n(M_2)$.
4. A set of n encoding functions $f_{r,k} : \mathcal{Y}_r^{k-1} \rightarrow \mathcal{X}_r$ at r for $k = 1, 2, \dots, n$, where $f_{r,k}$ is the encoding function in the k^{th} time slot.
5. A decoding function $g_1 : \mathcal{W}_1 \times \mathcal{Y}_1^n \rightarrow \mathcal{W}_2$ at node t_1 , where $g_1(w_1, y_1^n) = \hat{w}_2$.
6. A decoding function $g_2 : \mathcal{W}_2 \times \mathcal{Y}_2^n \rightarrow \mathcal{W}_1$ at node t_2 , where $g_2(w_2, y_2^n) = \hat{w}_1$.

From Definition 12, $X_1^n = f_1(W_1)$, $X_2^n = f_2(W_2)$, $X_r^n = f_r(Y_r^n)$, $\hat{W}_1 = g_2(W_2, Y_2^n)$ and $\hat{W}_2 = g_1(W_1, Y_1^n)$ for an (n, M_1, M_2) -code, where f_r is a function completely determined by $f_{r,k}$, $k = 1, 2, \dots, n$. The transmissions of messages in the TRC are illustrated in Figure 4.2.

Definition 13 The average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = Pr\{g_2(W_2, Y_2^n) \neq W_1\}$ and $P_{e,2}^n = Pr\{g_1(W_1, Y_1^n) \neq W_2\}$ respectively.

Definition 14 A rate pair (R_1, R_2) is achievable if there exists a sequence of (n, M_1, M_2) -codes with $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 15 The capacity region \mathcal{R} of the discrete memoryless TRC without feedback is the closure of the set of all achievable rate pairs.

In the next subsection, we investigate the discrete memoryless TRC without feedback and obtain an outer bound for \mathcal{R} .

4.3.2 An Outer Bound

Let $\mathcal{R}^1, \mathcal{R}^2, \mathcal{R}^3$ and \mathcal{R}^4 denote the sets

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq I(X_1; Y_r | X_2), \\ R_2 \leq I(X_2; Y_r | X_1) \\ \text{where } p(x_1, x_2, y_r) = p(x_1)p(x_2)p_1(y_r|x_1, x_2) \\ \text{for some input distribution } p(x_1)p(x_2) \text{ for the} \\ \text{MAC } p_1(y_r|x_1, x_2). \end{array} \right. \right\}, \quad (4.3)$$

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq I(X_1; Y_r | X_2), \\ R_2 \leq I(X_r; Y_1) \\ \text{where } p(x_1, x_2, x_r, y_r, y_1, y_2) = \\ p(x_1)p(x_2)p_1(y_r|x_1, x_2)p(x_r)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1)p(x_2) \text{ for} \\ \text{the MAC } p_1(y_r|x_1, x_2) \text{ and some input} \\ \text{distribution } p(x_r) \text{ for the BC } p_2(y_1, y_2|x_r). \end{array} \right. \right\},$$

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq I(X_r; Y_2) \\ R_2 \leq I(X_2; Y_r | X_1) \\ \text{where } p(x_1, x_2, x_r, y_r, y_1, y_2) = \\ p(x_1)p(x_2)p_1(y_r|x_1, x_2)p(x_r)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1)p(x_2) \text{ for} \\ \text{the MAC } p_1(y_r|x_1, x_2) \text{ and some input} \\ \text{distribution } p(x_r) \text{ for the BC } p_2(y_1, y_2|x_r). \end{array} \right. \right\}$$

and

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq I(X_r; Y_2), \\ R_2 \leq I(X_r; Y_1) \\ \text{where } p(x_r, y_1, y_2) = p(x_r)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_r) \text{ for the} \\ \text{BC } p_2(y_1, y_2|x_r). \end{array} \right. \right\}$$

respectively. Let \mathcal{R}^* denote

$$\overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)} \cap \overline{\text{conv}(\mathcal{R}^3)} \cap \overline{\text{conv}(\mathcal{R}^4)}. \quad (4.4)$$

Theorem 8 $\mathcal{R} \subset \mathcal{R}^*$.

Proof: Please refer to Appendix B. ■

4.4 Full-Duplex Discrete Memoryless Two-Way Relay Channel with Feedback

In this section, we study the discrete memoryless TRC with feedback, which means the terminal nodes can use their previously received messages for encoding their messages. The possible benefit of feedback in the discrete memoryless TRC will be illustrated in Section 4.4.1. Knopp [22] stated an outer bound on the capacity region of the discrete memoryless TRC with feedback. However, the proof in [22] is incomplete. Therefore, we present a detailed proof of the outer bound on the capacity region in [22] in this section, which is organized as follows. Section 4.4.1 illustrates the benefit of feedback for the TRC. Section 4.4.2 presents the problem formulation of the discrete memoryless TRC with feedback. Section 4.4.3 proves the outer bound on the capacity region of the discrete memoryless TRC with feedback. Section 4.4.4 shows that the outer bound for the discrete memoryless TRC with feedback is loose for some TRC.

4.4.1 Benefit of Feedback

The outer bound \mathcal{R}^* in (4.4) only holds for the TRC in which the terminal nodes do not use their previously received messages for encoding their messages. If the terminal nodes can use their previously received messages for encoding their messages, there exists a

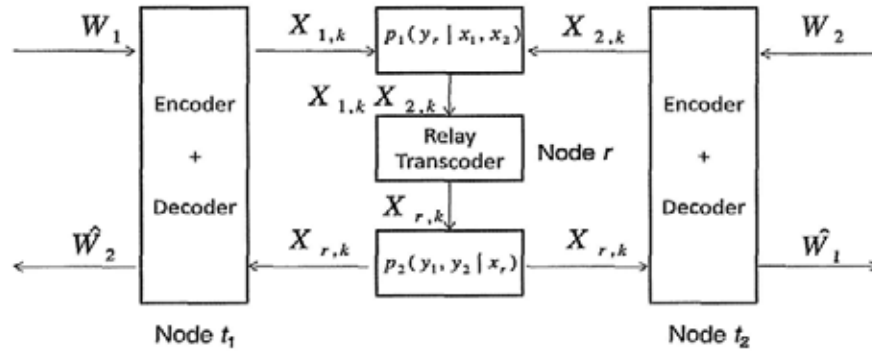


Figure 4.3: Binary multiplying relay channel (BMRC).

discrete memoryless TRC and an achievable rate pair for the TRC such that the rate pair lies outside the outer bound \mathcal{R}^* . Consider the TRC with deterministic MAC $p_1(y_r|x_1, x_2)$ and BC $p_2(y_1, y_2|x_r)$ such that

$$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_r = \mathcal{X}_r = \mathcal{Y}_1 = \mathcal{Y}_2 = 2, \tag{4.5}$$

$$p_1(y_r|x_1, x_2) = \begin{cases} 1 & \text{if } y = x_1x_2, \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

and

$$p_2(y_1, y_2|x_r) = \begin{cases} 1 & \text{if } y_1 = y_2 = x_r, \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

We call the TRC described above the *binary multiplying relay channel* (BMRC) and the transmissions of messages in the BMRC are illustrated in Figure 4.3.

Consider any independent variables X_1 and X_2 input to the MAC with the joint distribution $p(x_1, x_2) = p(x_1)p(x_2)$ where $Pr\{X_1 = 1\} = p$ and $Pr\{X_2 = 1\} = q$. For any $0 \leq \alpha \leq 1$, let $H(\alpha)$ denote

the entropy function of a Bernoulli(α) random variable. Then,

$$\begin{aligned}
 & I(X_1; Y_r | X_2) \\
 & \stackrel{(a)}{=} Pr\{X_2 = 0\}I(X_1; Y_r | X_2 = 0) + Pr\{X_2 = 1\}I(X_1; Y_r | X_2 = 1) \\
 & \stackrel{(b)}{=} Pr\{X_2 = 1\}I(X_1; Y_r | X_2 = 1) \\
 & \stackrel{(c)}{=} qI(X_1; X_1 | X_2 = 1) \\
 & \stackrel{(d)}{=} qH(X_1) \\
 & = qH(p),
 \end{aligned}$$

where

- (a) follows from (4.5),
- (b) follows from the fact that $Pr\{Y_r = 0 | X_2 = 0\} = 1$ (cf. (4.5) and (4.6)),
- (c) follows from (4.5) and (4.6),
- (d) follows from the fact that X_1 and X_2 are independent.

Similarly,

$$I(X_2; Y_r | X_1) = pH(q).$$

Then, \mathcal{R}^1 for the BMRC (cf. (4.3)) becomes

$$\left\{ \begin{array}{l} \cdot \\ (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \middle| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq qH(p), \\ R_2 \leq pH(q) \\ \text{for some } 0 \leq p \leq 1 \text{ and } 0 \leq q \leq 1. \end{array} \right\},$$

which is denoted by \mathcal{K} . It can be shown by standard optimization procedures that the largest value for $qH(p) + pH(q)$ is 1.2339. In

other words, the largest possible sum rate in \mathcal{K} is 1.2339, which implies that the largest possible sum rate in $\overline{\text{conv}(\mathcal{K})}$ is 1.2339. Therefore, the maximum equal-rate pair in $\overline{\text{conv}(\mathcal{K})}$ is less than $(1.2340/2, 1.2340/2) = (0.6170, 0.6170)$. On the other hand, the BMRC can be used as the binary multiplying channel (BMC) [24] if r forwards every received bit to the two terminals. It is shown in [24] that the rate pair $(0.63056, 0.63056)$ is in the capacity region of the BMC and it is achievable by some scheme under which the terminal nodes use their previously received messages for encoding their messages. However, the rate pair $(0.63056, 0.63056)$ is outside $\overline{\text{conv}(\mathcal{K})}$ and therefore outside \mathcal{R}^* . It then follows that using feedback can enlarge the capacity region of the BMRC.

4.4.2 Definitions

In the k^{th} time slot, node t_i constructs and transmits $X_{i,k}$, which is a function of W_i and Y_i^{k-1} for $i = 1, 2$, to node r through the MAC. The received symbol at node r is $Y_{r,k}$. In the same time slot, node r constructs and transmits $X_{r,k}$, which depends on Y_r^{k-1} , to node t_1 and node t_2 through the BC. The received symbol at node t_i is $Y_{i,k}$ for $i = 1, 2$. After n time slots, node t_1 declares \hat{W}_2 to be the transmitted W_2 based on Y_1^n and W_1 , and node t_2 declares \hat{W}_1 to be the transmitted W_1 based on Y_2^n and W_2 .

Definition 16 *The discrete memoryless TRC consists of six finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, \mathcal{Y}_1, \mathcal{Y}_2$ and \mathcal{Y}_r , one probability mass function $p_1(y_r|x_1, x_2)$ representing the MAC and one probability mass function $p_2(y_1, y_2|x_r)$ representing the BC. For any two inputs X_1 and X_2 to the MAC with a joint distribution $p(x_1, x_2)$ and X_r to the BC with a distribution $p(x_r)$, the relationship among X_1, X_2, X_r , the*

output Y_r of the MAC and the outputs Y_1 and Y_2 of the BC satisfies

$$p(x_1, x_2, x_r, y_1, y_2, y_r) = p(x_1, x_2, y_r)p(x_r, y_1, y_2)$$

for all $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, $x_r \in \mathcal{X}_r$, $y_r \in \mathcal{Y}_r$, $y_1 \in \mathcal{Y}_1$ and $y_2 \in \mathcal{Y}_2$, where

$$p(x_1, x_2, y_r) = p(x_1, x_2)p_1(y_r|x_1, x_2)$$

and

$$p(x_r, y_1, y_2) = p(x_r)p_2(y_1, y_2|x_r).$$

The discrete memoryless TRC is denoted by $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, p_1(y_r|x_1, x_2), p_2(y_1, y_2|x_r), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_r)$.

Definition 17 An (n, M_1, M_2) -code on the channel $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r, p_1(y_r|x_1, x_2), p_2(y_1, y_2|x_r), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_r)$ consists of the following:

1. A message set $\mathcal{W}_1 = \{1, 2, \dots, M_1\}$ at node t_1 and a message set $\mathcal{W}_2 = \{1, 2, \dots, M_2\}$ at node t_2 .
2. A set of n encoding functions $f_{1,k} : \mathcal{W}_1 \times \mathcal{Y}_1^{k-1} \rightarrow \mathcal{X}_1$ at node t_1 for $k = 1, 2, \dots, n$, where $f_{1,k}$ is the encoding function in the k^{th} time slot such that $X_{1,k} = f_{1,k}(W_1, Y_1^{k-1})$.
3. A set of n encoding functions $f_{2,k} : \mathcal{W}_2 \times \mathcal{Y}_2^{k-1} \rightarrow \mathcal{X}_2$ at node t_2 for $k = 1, 2, \dots, n$, where $f_{2,k}$ is the encoding function in the k^{th} time slot such that $X_{2,k} = f_{2,k}(W_2, Y_2^{k-1})$.
4. A set of n encoding functions $f_{r,k} : \mathcal{Y}_r^{k-1} \rightarrow \mathcal{X}_r$ at r for $k = 1, 2, \dots, n$, where $f_{r,k}$ is the encoding function in the k^{th} time slot such that $X_{r,k} = f_{r,k}(Y_r^{k-1})$.
5. A decoding function $g_1 : \mathcal{W}_1 \times \mathcal{Y}_1^n \rightarrow \mathcal{W}_2$ at node t_1 such that $g_1(W_1, Y_1^n) = \hat{W}_2$.

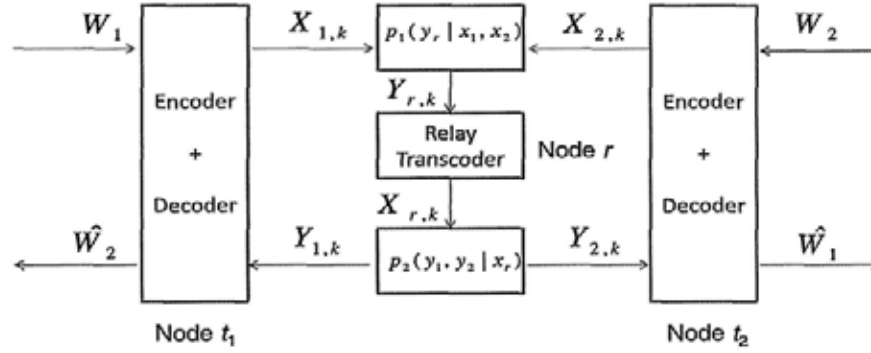


Figure 4.4: A discrete memoryless TRC with feedback.

6. A decoding function $g_2 : \mathcal{W}_2 \times \mathcal{Y}_2^n \rightarrow \mathcal{W}_1$ at node t_2 such that $g_2(W_2, Y_2^n) = \hat{W}_1$.

The transmissions of messages in the TRC are illustrated in Figure 4.4.

The following proposition is cited from Proposition 2.5 of [18] to facilitate discussion.

Proposition 23 For the discrete random variables X , Y and Z , $X \rightarrow Y \rightarrow Z$ forms a Markov Chain if and only if there exists two functions $\chi(x, y)$ and $\varphi(y, z)$ such that

$$p(x, y, z) = \chi(x, y)\varphi(y, z)$$

for all x , y and z where $p(y) > 0$.

Proof: It is contained in Proposition 2.5 of [18]. \blacksquare

Lemma 24 For any (n, M_1, M_2) -code on the discrete memoryless TRC,

$$\begin{aligned} (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\ \rightarrow (X_{1,k}, X_{2,k}) \rightarrow Y_{r,k}, \end{aligned} \quad (4.8)$$

$$\begin{aligned}
 & (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\
 & \quad \rightarrow X_{r,k} \rightarrow (Y_{1,k}, Y_{2,k}), \quad (4.9)
 \end{aligned}$$

$$\begin{aligned}
 & (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^k, Y_2^k, Y_r^{k-1}) \\
 & \quad \rightarrow (X_{1,k}, X_{2,k}) \rightarrow Y_{r,k} \quad (4.10)
 \end{aligned}$$

and

$$\begin{aligned}
 & (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^k) \\
 & \quad \rightarrow X_{r,k} \rightarrow (Y_{1,k}, Y_{2,k}) \quad (4.11)
 \end{aligned}$$

form four Markov Chains for $k = 1, 2, \dots, n$. In addition,

$$p(x_{1,k}, x_{2,k}, y_{r,k}) = p(x_{1,k}, x_{2,k})p_1(y_{r,k}|x_{1,k}, x_{2,k})$$

and

$$p(x_{r,k}, y_{1,k}, y_{2,k}) = p(x_{r,k})p_2(y_{1,k}, y_{2,k}|x_{r,k})$$

for $k = 1, 2, \dots, n$.

Proof: Please refer to Appendix B. ■

Definition 18 For an (n, M_1, M_2) -code on the discrete memoryless TRC, the average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = \Pr\{g_2(W_2, Y_2^n) \neq W_1\}$ and $P_{e,2}^n = \Pr\{g_1(W_1, Y_1^n) \neq W_2\}$ respectively.

Definition 19 A rate pair (R_1, R_2) is achievable if there exists a sequence of (n, M_1, M_2) -codes with $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 20 *The capacity region \mathcal{R}' of the discrete memoryless TRC with feedback is the closure of the set of all achievable rate pairs.*

4.4.3 An Outer Bound

Let

$$\mathcal{R}^{**} = \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min\{I(X_1; Y_r | X_2), I(X_r; Y_2)\}, \\ R_2 \leq \min\{I(X_2; Y_r | X_1), I(X_r; Y_1)\} \\ \text{where } p(x_1, x_2, x_r, y_1, y_2, y_r) = \\ p(x_1, x_2)p(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1, x_2) \\ \text{for the MAC } p_1(y_r|x_1, x_2) \text{ and some} \\ \text{input distribution } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right. \right\}. \quad (4.12)$$

In the following theorem, we prove in detail the outer bound for \mathcal{R}' stated in [22].

Theorem 9 $\mathcal{R}' \subset \mathcal{R}^{**}$.

Proof: Please refer to Appendix B. ■

4.4.4 Looseness of the Outer Bound

The outer bound \mathcal{R}^{**} for the discrete memoryless TRC in (4.12) is not tight in general. Consider the BMRC in Section 4.4.1 again.

Using (4.7), we obtain

$$I(X_r; Y_1) = I(X_r; Y_2) = H(X_r). \quad (4.13)$$

In addition,

$$\begin{aligned} I(X_1; Y_r|X_2) & \\ & \leq H(Y_r) \\ & \stackrel{(a)}{\leq} 1 \\ & = H(1/2) \end{aligned} \quad (4.14)$$

where (a) follows from (4.5). Similarly,

$$I(X_2; Y_r|X_1) \leq H(1/2). \quad (4.15)$$

Then,

$$\mathcal{R}^{**} \stackrel{(a)}{=} \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min\{I(X_1; Y_r|X_2), H(X_r)\}, \\ R_2 \leq \min\{I(X_2; Y_r|X_1), H(X_r)\} \\ \text{where } p(x_1, x_2, x_r, y_1, y_2, y_r) = \\ p(x_1, x_2)p(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1, x_2) \\ \text{for the MAC } p_1(y_r|x_1, x_2) \text{ and some} \\ \text{input distribution } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right. \right\} \quad (4.16)$$

$$\begin{aligned}
& \stackrel{(b)}{=} \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min\{I(X_1; Y_r | X_2), H(X_r)\}, \\ R_2 \leq \min\{I(X_2; Y_r | X_1), H(X_r)\} \\ \text{where } p(x_1, x_2, x_r, y_1, y_2, y_r) = \\ p(x_1, x_2)p(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1, x_2) \text{ for} \\ \text{the MAC } p_1(y_r|x_1, x_2) \text{ and the input} \\ \text{distribution } p(x_r) \text{ is Bernoulli}(1/2) \text{ for} \\ \text{the BC } p_2(y_1, y_2|x_r). \end{array} \right. \right\} \\
& = \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min\{I(X_1; Y_r | X_2), H(1/2)\}, \\ R_2 \leq \min\{I(X_2; Y_r | X_1), H(1/2)\} \\ \text{where } p(x_1, x_2, x_r, y_1, y_2, y_r) = \\ p(x_1, x_2)p(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1, x_2) \text{ for} \\ \text{the MAC } p_1(y_r|x_1, x_2) \text{ and the input} \\ \text{distribution } p(x_r) \text{ is Bernoulli}(1/2) \text{ for} \\ \text{the BC } p_2(y_1, y_2|x_r). \end{array} \right. \right\} \\
& \stackrel{(c)}{=} \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq I(X_1; Y_r | X_2), \\ R_2 \leq I(X_2; Y_r | X_1) \\ \text{where } p(x_1, x_2, y_r) = \\ p(x_1, x_2)p_1(y_r|x_1, x_2) \text{ for} \\ \text{some input distribution} \\ p(x_1, x_2) \text{ for the MAC} \\ p_1(y_r|x_1, x_2). \end{array} \right. \right\}, \quad (4.17)
\end{aligned}$$

where

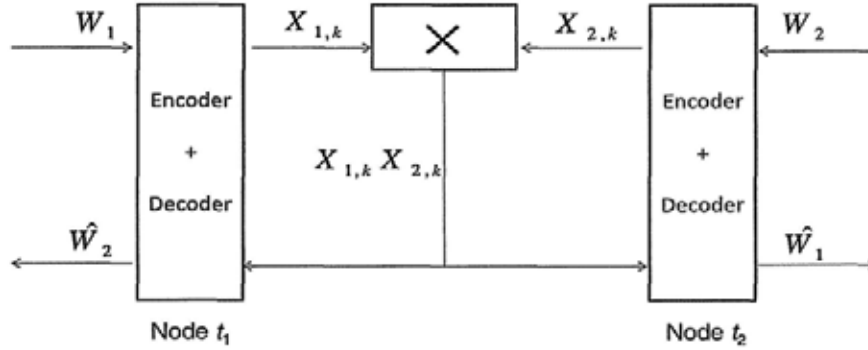


Figure 4.5: Binary multiplying relay channel (BMC).

- (a) follows from (4.12) and (4.13),
- (b) follows from restricting the input distribution for the BC in (4.16) to be Bernoulli(1/2),
- (c) follows from (4.14) and (4.15).

Let G denote the set in (4.17), which is equivalent to the Shannon’s outer bound for the Blackwell multiplying channel (BMC) in [23]. The transmissions of messages in the BMC are illustrated in Figure 4.5.

We prove the looseness of \mathcal{R}^{**} by presenting another outer bound for the BMRC that does not contain \mathcal{R}^{**} . In the BMRC, $Y_{1,k} = Y_{2,k} = X_{r,k}$ with probability 1 and $X_{r,k}$ is a function of $Y_r^{k-1} = [X_{1,1}X_{2,1}, X_{1,2}X_{2,2}, \dots, X_{1,k-1}X_{2,k-1}]$ for any positive integer k . In the BMC, node t_1 and node t_2 receive $X_{1,k}X_{2,k}$ in the k^{th} time slot with probability 1, which implies that node t_1 and node t_2 in the BMC have received $[X_{1,1}X_{2,1}, X_{1,2}X_{2,2}, \dots, X_{1,k}X_{2,k}]$ by the end of the k^{th} time slot. Therefore, the operations performed by the transcoder in any transmission scheme in the BMRC can be imitated identically at the terminal nodes in the BMC, which implies that the terminal nodes in the BMC can construct $X_{r,k}$ in the k^{th} time slot.

It then follows that any transmission scheme in the BMRC can be regarded as a transmission scheme in the BMC. Consequently, the capacity region of the BMRC is a subset of the capacity region of the BMC, which implies that any outer bound for the BMC is an outer bound for the BMRC. Hekstra and Willems [25] obtained an outer bound for the BMC and the maximum equal-rate pair in their outer bound, which is 0.64628, is strictly less than the maximum equal-rate pair in G , which is 0.69424 [23]. Consequently, G is not inside the Hekstra's outer bound for the BMC as well as the BMRC, which implies from (4.17) that \mathcal{R}^{**} is not inside the Hekstra's outer bound for the BMRC.

4.5 Full-Duplex Gaussian TRC with feedback

In this section, we study the Gaussian TRC in which the terminal nodes can use their previously received messages for encoding their messages. Knopp [22] stated an outer bound on the capacity region of the Gaussian TRC. However, the proof in [22] is incomplete. Therefore, we present a detailed proof of the outer bound on the capacity region in [22] in this section. We will present in Section 4.5.1 the problem formulation of the Gaussian TRC and prove in Section 4.5.2 the outer bound on the capacity region of the Gaussian TRC.

4.5.1 Definitions

If $X_{1,k}$ is transmitted at node t_1 and $X_{2,k}$ is transmitted at node t_2 in the k^{th} time slot, the received signal at the relay in the time slot is assumed to be $Y_{r,k} = X_{1,k} + X_{2,k} + Z_{r,k}$ where $Z_{r,k}$ is a Gaussian random variable independent of $(X_{1,k}, X_{2,k})$. If $X_{r,k}$ is transmitted at r

in the same time slot, the received signal at node t_1 in the time slot is assumed to be $Y_{1,k} = X_{r,k} + Z_{1,k}$, where $Z_{1,k}$ is a Gaussian random variable independent of $X_{r,k}$, and the received signal at node t_2 in the time slot is assumed to be $Y_{2,k} = X_{r,k} + Z_{2,k}$, where $Z_{2,k}$ is a Gaussian random variable independent of $X_{r,k}$. For any three codewords $(x_{1,1}, x_{1,2}, \dots, x_{1,n})$, $(x_{2,1}, x_{2,2}, \dots, x_{2,n})$ and $(x_{r,1}, x_{r,2}, \dots, x_{r,n})$ that are transmitted over the channel, we require that $\frac{1}{n} \sum_{k=1}^n x_{1,k}^2 \leq P_1$, $\frac{1}{n} \sum_{k=1}^n x_{2,k}^2 \leq P_2$ and $\frac{1}{n} \sum_{k=1}^n x_{r,k}^2 \leq P_r$ where P_1 , P_2 and P_r represent the maximum power constraints for node t_1 , node t_2 and node r respectively.

Definition 21 *The Gaussian TRC consists of the following:*

1. Three Gaussian random variables $Z_1 = \mathcal{N}(0, N_1)$, $Z_2 = \mathcal{N}(0, N_2)$ and $Z_r = \mathcal{N}(0, N_r)$.
2. A probability density function $f_1(y_r|x_1, x_2)$ defined for all real numbers x_1 , x_2 and y_r representing the MAC and a probability density function $f_2(y_1, y_2|x_r)$ defined for all real numbers x_r , y_1 and y_2 representing the BC such that

$$\begin{aligned}
 f_1(y_r|x_1, x_2) &= f_{Z_r}(y_r - x_1 - x_2|x_1, x_2) \\
 &= f_{Z_r}(y_r - x_1 - x_2)
 \end{aligned} \tag{4.18}$$

and

$$\begin{aligned}
 f_2(y_1, y_2|x_r) &= f_{Z_1, Z_2}(y_1 - x_r, y_2 - x_r|x_r) \\
 &= f_{Z_1, Z_2}(y_1 - x_r, y_2 - x_r).
 \end{aligned} \tag{4.19}$$

In addition, we require that for any two finite discrete random variables X_1 and X_2 with an input distribution $p(x_1, x_2)$ for the MAC and any finite discrete random variable X_r with an input distribution $p(x_r)$ for the BC, the relationship among X_1 , X_2 , X_r , the output Y_r of the MAC and the outputs Y_1 and Y_2 of the BC satisfies

$$\begin{aligned} &Pr\{X_1 = x_1, X_2 = x_2, X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2, Y_r \leq y_r\} \\ &= Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\}Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2\} \end{aligned}$$

for all $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, $x_r \in \mathcal{X}_r$, $y_r \in \mathbb{R}$, $y_1 \in \mathbb{R}$ and $y_2 \in \mathbb{R}$, where

$$Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\} = \int_{-\infty}^{y_r} f_1(v_r|x_1, x_2)p(x_1, x_2)dv_r$$

and

$$\begin{aligned} &Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2\} \\ &= \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f_2(v_1, v_2|x_r)p(x_r)dv_1dv_2. \end{aligned}$$

The Gaussian TRC is denoted by (N_1, N_2, N_r) .

Definition 22 An $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code on the channel (N_1, N_2, N_r) with maximum power constraints P_1 , P_2 and P_r consists of the following:

1. A message set $\mathcal{W}_1 = \{1, 2, \dots, M_1\}$ at node t_1 and a message set $\mathcal{W}_2 = \{1, 2, \dots, M_2\}$ at node t_2 .
2. A set of n encoding functions $\alpha_{1,k} : \mathcal{W}_1 \times \mathcal{Y}_1^{k-1} \rightarrow \mathcal{X}_1$ at node t_1 for $k = 1, 2, \dots, n$, where $\alpha_{1,k}$ is the encoding function in the k^{th} time slot such that $X_{1,k} = \alpha_{1,k}(W_1, Y_1^{k-1})$. Without loss of generality, we assume that \mathcal{X}_1 is finite. In addition, every

codeword x_1^n must satisfy the power constraint

$$\sum_{k=1}^n x_{1,k}^2 \leq nP_1. \quad (4.20)$$

3. A set of n encoding functions $\alpha_{2,k} : \mathcal{W}_2 \times \mathcal{Y}_2^{k-1} \rightarrow \mathcal{X}_2$ at node t_2 for $k = 1, 2, \dots, n$, where $\alpha_{2,k}$ is the encoding function in the k^{th} time slot such that $X_{2,k} = \alpha_{2,k}(W_2, Y_2^{k-1})$. Without loss of generality, we assume that \mathcal{X}_2 is finite. In addition, every codeword x_2^n must satisfy the power constraint

$$\sum_{k=1}^n x_{2,k}^2 \leq nP_2. \quad (4.21)$$

4. A set of n encoding functions $\alpha_{r,k} : \mathcal{Y}_r^{k-1} \rightarrow \mathcal{X}_r$ at r for $k = 1, 2, \dots, n$, where $\alpha_{r,k}$ is the encoding function in the k^{th} time slot such that $X_{r,k} = \alpha_{r,k}(Y_r^{k-1})$. Without loss of generality, we assume that \mathcal{X}_r is finite. In addition, every codeword x_r^n must satisfy the power constraint

$$\sum_{k=1}^n x_{r,k}^2 \leq nP_r. \quad (4.22)$$

5. A decoding function $\beta_1 : \mathcal{W}_1 \times \mathcal{Y}_1^n \rightarrow \mathcal{W}_2$ at node t_1 such that $\beta_1(W_1, Y_1^n) = \hat{W}_2$.
6. A decoding function $\beta_2 : \mathcal{W}_2 \times \mathcal{Y}_2^n \rightarrow \mathcal{W}_1$ at node t_2 such that $\beta_2(W_2, Y_2^n) = \hat{W}_1$.

The transmissions of messages in the Gaussian TRC are illustrated in Figure 4.6.

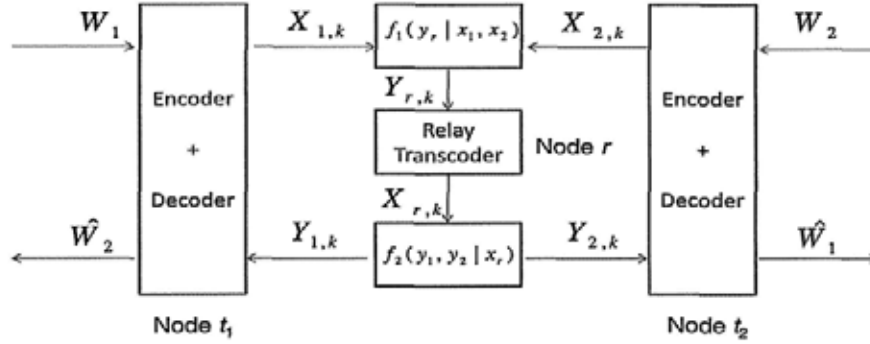


Figure 4.6: Gaussian TRC.

Definition 23 For an $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code on the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r , the average probabilities of decoding error of W_1 and W_2 in (N_1, N_2, N_r) are defined as $P_{e,1}^n = \Pr\{\beta_2(W_2, Y_2^n) \neq W_1\}$ and $P_{e,2}^n = \Pr\{\beta_1(W_1, Y_1^n) \neq W_2\}$ respectively.

Definition 24 A rate pair (R_1, R_2) is Gaussian-achievable if there exists a sequence of $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -codes for the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r such $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1, \lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2, \lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 25 The capacity region \mathcal{R}_G of the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r is the closure of the set of all Gaussian-achievable rate pairs.

4.5.2 An Outer Bound

In the following theorem, we prove in detail the outer bound for \mathcal{R}_G stated in [22].

Theorem 10 *For the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r ,*

$$\mathcal{R}_G \subset \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min \left\{ \frac{1}{2} \log_2(1 + P_1/N_r), \frac{1}{2} \log_2(1 + P_r/N_2) \right\}, \\ R_2 \leq \min \left\{ \frac{1}{2} \log_2(1 + P_2/N_r), \frac{1}{2} \log_2(1 + P_r/N_1) \right\}. \end{array} \right. \right\}.$$

Proof: Please refer to Appendix B. ■

4.6 Half-Duplex Two-Way Relay Channels

In this section and the next section, all the nodes in the TRCs are assumed to be half-duplex. The two terminals transmit messages to the relay through the MAC in the first phase and the relay transmits messages to the two terminals through BC in the second phase. We obtain outer bounds on the capacity regions of the half-duplex TRCs by following similar procedures for proving the outer bounds for the full-duplex TRCs.

4.6.1 Definitions

Let α and β denote the fractions of time allocated to the MAC phase and the BC phase respectively where α and β are two real numbers such that $0 \leq \alpha, \beta \leq 1$ and $\alpha + \beta \leq 1$. Suppose node t_1 and node t_2 exchange information in n time slots. To simplify analysis, we assume that the first $n_1 = \lfloor n\alpha \rfloor$ time slots are allocated to the MAC phase and the last $n_2 = \lfloor n\beta \rfloor$ time slots are allocated to the BC phase. Then, the variables involved in the communication through the TRC are:

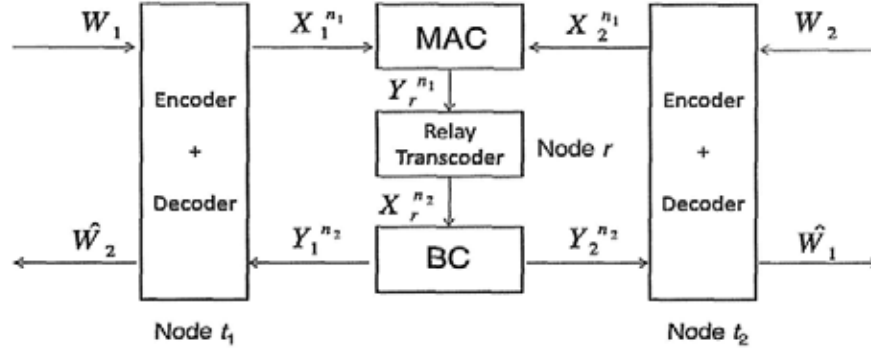


Figure 4.7: A half-duplex TRC.

- $W_i \in \{1, 2, \dots, M_i\}$: messages of node t_i , $i = 1, 2$,
- $X_i^{n_i} = [X_{i,1} X_{i,2} \dots X_{i,n_i}]^T$: codewords transmitted by node t_i , $i = 1, 2$,
- $Y_r^{n_1} = [Y_{r,1} Y_{r,2} \dots Y_{r,n_1}]^T$: channel outputs at the relay,
- $X_r^{n_2} = [X_{r,1} X_{r,2} \dots X_{r,n_2}]^T$: codewords transmitted by the relay,
- $Y_i^{n_2} = [Y_{i,1} Y_{i,2} \dots Y_{i,n_2}]^T$: channel outputs at node t_i , $i = 1, 2$,
- $\hat{W}_i \in \{1, 2, \dots, M_i\}$: message estimates of message W_i , $i = 1, 2$.

Node t_1 and node t_2 choose messages W_1 and W_2 independently according to the uniform distribution, and node t_1 declares the message estimate \hat{W}_2 and node t_2 declares the message estimate \hat{W}_1 after n time slots. The transmissions of messages in the half-duplex TRC are shown in Figure 4.7.

4.6.2 Discrete Memoryless Two-Way Relay Channel

Definition 26 The average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = \Pr\{g_2(W_2, Y_2^{n_2}) \neq W_1\}$ and $P_{e,2}^n = \Pr\{g_1(W_1, Y_1^{n_2}) \neq W_2\}$ respectively.

Definition 27 A rate pair (R_1, R_2) is (α, β) -achievable if there exists a sequence of (n, M_1, M_2) -codes that allocate the first $n_1 = \lfloor n\alpha \rfloor$ time slots to the MAC phase and the last $n_2 = \lfloor n\beta \rfloor$ time slots to the BC phase with $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 28 The (α, β) -capacity region $\underline{\mathcal{R}}$ of the half-duplex discrete memoryless TRC is the closure of the set of all (α, β) -achievable rate pairs.

Let $\underline{\mathcal{R}}^1, \underline{\mathcal{R}}^2, \underline{\mathcal{R}}^3$ and $\underline{\mathcal{R}}^4$ denote the sets

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \alpha I(X_1; Y_r | X_2), \\ R_2 \leq \alpha I(X_2; Y_r | X_1) \\ \text{where } p(x_1, x_2, y_r) = \\ p(x_1)p(x_2)p_1(y_r|x_1, x_2) \\ \text{for some input distribu-} \\ \text{tion } p(x_1)p(x_2) \text{ for the} \\ \text{MAC } p_1(y_r|x_1, x_2). \end{array} \right. \right\}, \quad (4.23)$$

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \alpha I(X_1; Y_r | X_2), \\ R_2 \leq \beta I(X_r; Y_1) \\ \text{where } p(x_1, x_2, x_r, y_r, y_1, y_2) = \\ p(x_1)p(x_2)p_1(y_r|x_1, x_2)p(x_r)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1)p(x_2) \\ \text{for the MAC } p_1(y_r|x_1, x_2) \text{ and some} \\ \text{input distribution } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right. \right\},$$

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \beta I(X_r; Y_2) \\ R_2 \leq \alpha I(X_2; Y_r | X_1) \\ \text{where } p(x_1, x_2, x_r, y_1, y_2) = \\ p(x_1)p(x_2)p_1(y_r|x_1, x_2)p(x_r)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1)p(x_2) \\ \text{for the MAC } p_1(y_r|x_1, x_2) \text{ and some} \\ \text{input distribution } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right. \right\}$$

and

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \beta I(X_r; Y_2), \\ R_2 \leq \beta I(X_r; Y_1) \\ \text{where } p(x_r, y_1, y_2) = \\ p(x_r)p_2(y_1, y_2|x_r) \text{ for} \\ \text{some input distribu-} \\ \text{tion } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right. \right\}$$

respectively. Let $\underline{\mathcal{R}}^*$ denote

$$\overline{\text{conv}(\underline{\mathcal{R}}^1)} \cap \overline{\text{conv}(\underline{\mathcal{R}}^2)} \cap \overline{\text{conv}(\underline{\mathcal{R}}^3)} \cap \overline{\text{conv}(\underline{\mathcal{R}}^4)}. \quad (4.24)$$

Theorem 11 (Counterpart of Theorem 8) $\underline{\mathcal{R}} \subset \underline{\mathcal{R}}^*$.

Proof: Please refer to Appendix B. ■

Let $\underline{\mathcal{R}}^{**}$ denote

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \middle| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min\{\alpha I(X_1; Y_r | X_2), \beta I(X_r; Y_2)\}, \\ R_2 \leq \min\{\alpha I(X_2; Y_r | X_1), \beta I(X_r; Y_1)\} \\ \text{where } p(x_1, x_2, x_r, y_1, y_2, y_r) = \\ p(x_1, x_2)p(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r) \\ \text{for some input distribution } p(x_1, x_2) \\ \text{for the MAC } p_1(y_r|x_1, x_2) \text{ and some} \\ \text{input distribution } p(x_r) \text{ for the BC} \\ p_2(y_1, y_2|x_r). \end{array} \right\}. \quad (4.25)$$

Theorem 12 (Counterpart of Theorem 9)

$$\underline{\mathcal{R}} \subset \underline{\mathcal{R}}^{**}.$$

Proof: Please refer to Appendix B. ■

4.6.3 Gaussian Two-Way Relay Channel

For any three codewords $(x_{1,1}, x_{1,2}, \dots, x_{1,n_1})$, $(x_{2,1}, x_{2,2}, \dots, x_{2,n_2})$ and $(x_{r,1}, x_{r,2}, \dots, x_{r,n_r})$ that are transmitted over the channel, we require that $\frac{1}{n_1} \sum_{k=1}^{n_1} x_{1,k}^2 \leq P_1$, $\frac{1}{n_2} \sum_{k=1}^{n_2} x_{2,k}^2 \leq P_2$ and $\frac{1}{n_r} \sum_{k=1}^{n_r} x_{r,k}^2 \leq P_r$ where P_1 , P_2 and P_r represent the maximum power constraints for node t_1 , node t_2 and node r respectively.

Definition 29 For an $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code on the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1 , P_2 and P_r , the average probabilities of decoding error of W_1 and W_2 in $(N_1, N_2,$

N_r) are defined as $P_{e,1}^n = \Pr\{\beta_2(W_2, Y_2^{n_2}) \neq W_1\}$ and $P_{e,2}^n = \Pr\{\beta_1(W_1, Y_1^{n_2}) \neq W_2\}$ respectively.

Definition 30 A rate pair (R_1, R_2) is (α, β) -Gaussian-achievable if there exists a sequence of $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -codes that allocate the first $n_1 = \lfloor n\alpha \rfloor$ time slots to the MAC phase and the last $n_2 = \lfloor n\beta \rfloor$ time slots to the BC phase of the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r such that $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1$ and $\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2$ such that $\lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 31 The (α, β) -capacity region $\underline{\mathcal{R}}_G$ of the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r is the closure of the set of all (α, β) -Gaussian-achievable rate pairs.

Theorem 13 (Counterpart of Theorem 10) For the Gaussian TRC (N_1, N_2, N_r) with maximum power constraints P_1, P_2 and P_r ,

$$\underline{\mathcal{R}}_G \subset \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \middle| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min \left\{ \frac{\alpha}{2} \log_2(1 + P_1/N_r), \right. \\ \left. \frac{\beta}{2} \log_2(1 + P_r/N_2) \right\}, \\ R_2 \leq \min \left\{ \frac{\alpha}{2} \log_2(1 + P_2/N_r), \right. \\ \left. \frac{\beta}{2} \log_2(1 + P_r/N_1) \right\}. \end{array} \right\}.$$

Proof: Please refer to Appendix B. ■

4.6.4 Bandlimited Gaussian Two-Way Relay Channel

We now consider a half-duplex bandlimited Gaussian TRC with white noise. The MAC output of the bandlimited Gaussian TRC can be described as the convolution

$$Y_r(t) = (X_1(t) + X_2(t) + Z_r(t)) * h(t)$$

where $X_i(t)$ is the signal waveform transmitted by node t_i for $i = 1, 2$, $Z_r(t)$ is the waveform of a white Gaussian noise with two-sided power spectral density $N_r/2$ W/Hz and $h(t)$ is the impulse response of an ideal bandpass filter which cuts out all frequencies greater than W Hz. The BC output of the bandlimited Gaussian TRC at node t_1 and node t_2 can be described as the convolutions

$$Y_1(t) = (X_r(t) + Z_1(t)) * h(t)$$

and

$$Y_2(t) = (X_r(t) + Z_2(t)) * h(t)$$

respectively, where $X_r(t)$ is the signal waveform transmitted by node r , $Z_i(t)$ is the waveform of a white Gaussian noise with two-sided power spectral density $N_i/2$ W/Hz for $i = 1, 2$. Let the bandlimited Gaussian TRC be used over the time interval $[0, T]$ and without loss of generality, the time interval $[0, \alpha T]$ is allocated for the MAC transmissions and the time interval $[\alpha T, (\alpha + \beta)T]$ is allocated for the BC transmissions. The maximum power constraints for $x_1(t)$, $x_2(t)$ and $x_r(t)$ are P_1 , P_2 and P_r respectively for each $x_1(t)$ transmitted at node t_1 , each $x_2(t)$ transmitted at node t_2 and each $x_r(t)$ transmitted at node r . Following similar procedures for deriving the capacity of the bandlimited Gaussian channel from the

capacity of the discrete-time Gaussian channel in [17–19], we can derive the following result from Theorem 13.

Theorem 14 *For the bandlimited Gaussian TRC of bandwidth W with two-sided noise power spectral densities $N_1/2, N_2/2$ and $N_r/2$ and with maximum power constraints P_1, P_2 and P_r , the capacity region is a subset of*

$$\left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min \left\{ \alpha W \log_2 \left(1 + \frac{P_1}{N_r W} \right), \right. \\ \left. \beta W \log_2 \left(1 + \frac{P_r}{N_2 W} \right) \right\}, \\ R_2 \leq \min \left\{ \alpha W \log_2 \left(1 + \frac{P_2}{N_r W} \right), \right. \\ \left. \beta W \log_2 \left(1 + \frac{P_r}{N_1 W} \right) \right\}. \end{array} \right\}.$$

In physical-layer network coding (PNC) [16], transmissions of messages between the two terminals are completed in two stages. In the first stage, the two terminals transmit their waveforms, and the received waveforms at the relay is the sum of their waveforms and a white Gaussian noise. In the second stage, the relay processes the received signal and transmits a waveform to the two terminals based on the processed signal. Theorem 14 provides an outer bound on the achievable rate region of PNC for a bandlimited Gaussian TRC.

4.7 Half-Duplex Gaussian Two-Way Relay Channel with Varying Background Noise

We have assumed in Section 4.6.3 that during the MAC phase in the half-duplex Gaussian TRC, the noise power N_r at relay r re-

mains constant. Similarly, we have assumed in Section 4.6.3 that during the MAC phase in the half-duplex Gaussian TRC, the noise power N_1 at node t_1 remains constant and the noise power N_2 at node t_2 remains constant. In this section, we assume that during the MAC phase, the noise power at r changes from $N_{r,1}$ to $N_{r,2}$. In addition, we assume that during the BC phase, the noise power at node t_1 changes from $N_{1,1}$ to $N_{1,2}$ and the noise power at node t_2 changes from $N_{2,1}$ to $N_{2,2}$.

4.7.1 Definitions

Let $\alpha_1, \alpha_2, \beta_1$ and β_2 be real numbers such that $0 \leq \alpha_1, \alpha_2, \beta_1, \beta_2 \leq 1$ and $\alpha_1 + \alpha_2 + \beta_1 + \beta_2 \leq 1$. Let n_1, n_2, n_3 and n_4 denote $\lfloor n\alpha_1 \rfloor, \lfloor n\alpha_2 \rfloor, \lfloor n\beta_1 \rfloor$ and $\lfloor n\beta_2 \rfloor$ respectively. Suppose node t_1 and node t_2 exchange information in n time slots. The MAC is used in two sessions and the BC is also used in two sessions. The first session of MAC consists of the first n_1 time slots for the MAC transmissions and the received signal at r is $X_1 + X_2 + Z_{r,1}$ where $Z_{r,1}$ is a zero-mean Gaussian random variable with variance $N_{r,1}$. The second session of MAC consists of the next n_2 time slots for the MAC transmissions and the received signal at r is $X_1 + X_2 + Z_{r,2}$ where $Z_{r,2}$ is a zero-mean Gaussian random variable with variance $N_{r,2}$. The first session of BC consists of the first n_3 time slots for the BC transmissions and the received signals at node t_1 and node t_2 are $X_r + Z_{1,1}$ and $X_r + Z_{2,1}$ respectively, where $Z_{1,1}$ and $Z_{2,1}$ are two zero-mean Gaussian random variables with variances $N_{1,1}$ and $N_{2,1}$ respectively. The second session of BC consists of the next n_4 time slots for the BC transmissions and the received signals at node t_1 and node t_2 are $X_r + Z_{1,2}$ and $X_r + Z_{2,2}$ respectively, where $Z_{1,2}$ and $Z_{2,2}$ are two

zero-mean Gaussian random variables with variances $N_{1,2}$ and $N_{2,2}$ respectively. Then, the variables involved in the communication through the TRC are:

- $W_i \in \{1, 2, \dots, M_i\}$: messages of node t_i , $i = 1, 2$,
- $X_i^{n_1+n_2} = [X_{i,1} \ X_{i,2} \ \dots \ X_{i,n_1+n_2}]^T$: codewords transmitted by node t_i , $i = 1, 2$,
- $Y_r^{n_1+n_2} = [Y_{r,1} \ Y_{r,2} \ \dots \ Y_{r,n_1+n_2}]^T$: channel outputs at the relay,
- $X_r^{n_3+n_4} = [X_{r,1} \ X_{r,2} \ \dots \ X_{r,n_3+n_4}]^T$: codewords transmitted by the relay,
- $Y_i^{n_3+n_4} = [Y_{i,1} \ Y_{i,2} \ \dots \ Y_{i,n_3+n_4}]^T$: channel outputs at node t_i , $i = 1, 2$,
- $\hat{W}_i \in \{1, 2, \dots, M_i\}$: message estimates of message W_i , $i = 1, 2$.

Node t_1 and node t_2 choose messages W_1 and W_2 independently according to the uniform distribution, and node t_1 declares the message estimate \hat{W}_2 and node t_2 declares the message estimate \hat{W}_1 after n time slots. For any three codewords $(x_{1,1}, x_{1,2}, \dots, x_{1,n_1+n_2})$, $(x_{2,1}, x_{2,2}, \dots, x_{2,n_1+n_2})$ and $(x_{r,1}, x_{r,2}, \dots, x_{r,n_3+n_4})$ that are transmitted over the channel, we require that

$$\begin{aligned} \frac{1}{n_1} \sum_{k=1}^{n_1} x_{1,k}^2 &\leq P_1, \\ \frac{1}{n_2} \sum_{k=n_1+1}^{n_1+n_2} x_{1,k}^2 &\leq P_1, \\ \frac{1}{n_1} \sum_{k=1}^{n_1} x_{2,k}^2 &\leq P_2, \end{aligned}$$

$$\frac{1}{n_2} \sum_{k=n_1+1}^{n_1+n_2} x_{2,k}^2 \leq P_2,$$

$$\frac{1}{n_3} \sum_{k=1}^{n_3} x_{r,k}^2 \leq P_r$$

and

$$\frac{1}{n_4} \sum_{k=n_3+1}^{n_3+n_4} x_{r,k}^2 \leq P_r,$$

where P_1 , P_2 and P_r represent the maximum power constraints for node t_1 , node t_2 and node r respectively.

Definition 32 *The half-duplex Gaussian TRC with varying background noise consists of the following:*

1. Six Gaussian random variables $Z_{1,1} = \mathcal{N}(0, N_{1,1})$, $Z_{1,2} = \mathcal{N}(0, N_{1,2})$, $Z_{2,1} = \mathcal{N}(0, N_{2,1})$, $Z_{2,2} = \mathcal{N}(0, N_{2,2})$, $Z_{r,1} = \mathcal{N}(0, N_{r,1})$ and $Z_{r,2} = \mathcal{N}(0, N_{r,2})$.
2. A probability density function $f_{1,i}(y_r|x_1, x_2)$ for session i , $i = 1, 2$, of MAC defined for all real numbers x_1 , x_2 and y_r representing the MAC with the noise variable $Z_{r,i}$ such that

$$\begin{aligned} f_{1,i}(y_r|x_1, x_2) &= f_{Z_{r,i}}(y_r - x_1 - x_2|x_1, x_2) \\ &= f_{Z_{r,i}}(y_r - x_1 - x_2). \end{aligned}$$

3. A probability density function $f_{2,j}(y_1, y_2|x_r)$ for session j , $j = 1, 2$, of BC defined for all real numbers x_r , y_1 and y_2 represent-

ing the BC with the noise variables $Z_{1,j}$ and $Z_{2,j}$ such that

$$\begin{aligned} f_{2,j}(y_1, y_2 | x_r) &= f_{Z_{1,j}, Z_{2,j}}(y_1 - x_r, y_2 - x_r | x_r) \\ &= f_{Z_{1,j}, Z_{2,j}}(y_1 - x_r, y_2 - x_r). \end{aligned}$$

In addition, we require that for any two finite discrete random variables X_1 and X_2 with an input distribution $p(x_1, x_2)$ for the MAC in session i , $i = 1, 2$, and any finite discrete random variable X_r with an input distribution $p(x_r)$ for the BC in session j , $j = 1, 2$, the relationship among X_1 , X_2 , X_r , the output Y_r of the MAC and the outputs Y_1 and Y_2 of the BC satisfies

$$\begin{aligned} \Pr\{X_1 = x_1, X_2 = x_2, X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2, Y_r \leq y_r\} \\ = \Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\} \Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2\} \end{aligned}$$

for all $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, $x_r \in \mathcal{X}_r$, $y_r \in \mathbb{R}$, $y_1 \in \mathbb{R}$ and $y_2 \in \mathbb{R}$, where

$$\Pr\{X_1 = x_1, X_2 = x_2, Y_r \leq y_r\} = \int_{-\infty}^{y_r} f_{1,i}(v_r | x_1, x_2) p(x_1, x_2) dv_r$$

and

$$\begin{aligned} \Pr\{X_r = x_r, Y_1 \leq y_1, Y_2 \leq y_2\} \\ = \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f_{2,j}(v_1, v_2 | x_r) p(x_r) dv_1 dv_2. \end{aligned}$$

The half-duplex Gaussian TRC with varying background noise is denoted by $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$.

Definition 33 An $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code on the channel $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$ with maximum power constraints P_1, P_2

and P_r consists of the following:

1. A message set $\mathcal{W}_1 = \{1, 2, \dots, M_1\}$ at node t_1 and a message set $\mathcal{W}_2 = \{1, 2, \dots, M_2\}$ at node t_2 .
2. A set of n encoding functions $\alpha_{1,k} : \mathcal{W}_1 \times \mathcal{Y}_1^{k-1} \rightarrow \mathcal{X}_1$ at node t_1 for $k = 1, 2, \dots, n$, where $\alpha_{1,k}$ is the encoding function in the k^{th} time slot such that $X_{1,k} = \alpha_{1,k}(W_1, Y_1^{k-1})$. Without loss of generality, we assume that \mathcal{X}_1 is finite. In addition, every codeword x_1^n must satisfy the power constraints

$$\sum_{k=1}^{n_1} x_{1,k}^2 \leq n_1 P_1$$

and

$$\sum_{k=n_1+1}^{n_1+n_2} x_{1,k}^2 \leq n_2 P_1.$$

3. A set of n encoding functions $\alpha_{2,k} : \mathcal{W}_2 \times \mathcal{Y}_2^{k-1} \rightarrow \mathcal{X}_2$ at node t_2 for $k = 1, 2, \dots, n$, where $\alpha_{2,k}$ is the encoding function in the k^{th} time slot such that $X_{2,k} = \alpha_{2,k}(W_2, Y_2^{k-1})$. Without loss of generality, we assume that \mathcal{X}_2 is finite. In addition, every codeword x_2^n must satisfy the power constraints

$$\sum_{k=1}^{n_1} x_{2,k}^2 \leq n_1 P_2$$

and

$$\sum_{k=n_1+1}^{n_1+n_2} x_{2,k}^2 \leq n_2 P_2.$$

4. A set of n encoding functions $\alpha_{r,k} : \mathcal{Y}_r^{k-1} \rightarrow \mathcal{X}_r$ at r for $k = 1, 2, \dots, n$, where $\alpha_{r,k}$ is the encoding function in the k^{th} time slot such that $X_{r,k} = \alpha_{r,k}(Y_r^{k-1})$. Without loss of generality, we assume that \mathcal{X}_r is finite. In addition, every codeword x_r^n must

satisfy the power constraints

$$\sum_{k=1}^{n_3} x_{r,k}^2 \leq n_3 P_r$$

and

$$\sum_{k=n_3+1}^{n_3+n_4} x_{r,k}^2 \leq n_4 P_r.$$

5. A decoding function $\beta_1 : \mathcal{W}_1 \times \mathcal{Y}_1^{n_3+n_4} \rightarrow \mathcal{W}_2$ at node t_1 such that $\beta_1(W_1, Y_1^{n_3+n_4}) = \hat{W}_2$.
6. A decoding function $\beta_2 : \mathcal{W}_2 \times \mathcal{Y}_2^{n_3+n_4} \rightarrow \mathcal{W}_1$ at node t_2 such that $\beta_2(W_2, Y_2^{n_3+n_4}) = \hat{W}_1$.

Definition 34 For an $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code on the channel $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$ with maximum power constraints P_1, P_2 and P_r , the average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = \Pr\{\beta_2(W_2, Y_2^{n_3+n_4}) \neq W_1\}$ and $P_{e,2}^n = \Pr\{\beta_1(W_1, Y_1^{n_3+n_4}) \neq W_2\}$ respectively.

Definition 35 Fix $\alpha_1, \alpha_2, \beta_1$ and β_2 . A rate pair (R_1, R_2) is $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ -Gaussian-achievable if there exists a sequence of $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -codes for the TRC $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$ with maximum power constraints P_1, P_2 and P_r such that $\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1, \lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2, \lim_{n \rightarrow \infty} P_{e,1}^n = 0$ and $\lim_{n \rightarrow \infty} P_{e,2}^n = 0$.

Definition 36 The capacity region $\mathcal{R}_{\alpha_1, \alpha_2, \beta_1, \beta_2}$ of the half-duplex Gaussian TRC $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$ with maximum power constraints P_1, P_2 and P_r is the closure of the set of all $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ -Gaussian-achievable rate pairs.

4.7.2 An Outer Bound

Theorem 15 For the half-duplex Gaussian TRC $(N_{1,1}, N_{1,2}, N_{2,1}, N_{2,2}, N_{r,1}, N_{r,2})$ with maximum power constraints P_1, P_2 and P_r ,

$$\mathcal{R}_{\alpha_1, \alpha_2, \beta_1, \beta_2} \subset \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \middle| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min \left\{ \frac{\alpha_1}{2} \log_2(1 + P_1/N_{r,1}) + \right. \\ \quad \left. \frac{\alpha_2}{2} \log_2(1 + P_1/N_{r,2}), \right. \\ \quad \left. \frac{\beta_1}{2} \log_2(1 + P_r/N_{2,1}) + \right. \\ \quad \left. \frac{\beta_2}{2} \log_2(1 + P_r/N_{2,2}) \right\}, \\ R_2 \leq \min \left\{ \frac{\alpha_1}{2} \log_2(1 + P_2/N_{r,1}) + \right. \\ \quad \left. \frac{\alpha_2}{2} \log_2(1 + P_2/N_{r,2}), \right. \\ \quad \left. \frac{\beta_1}{2} \log_2(1 + P_r/N_{1,1}) + \right. \\ \quad \left. \frac{\beta_2}{2} \log_2(1 + P_r/N_{1,2}) \right\}. \end{array} \right\}.$$

Proof: It is similar to the proof of Theorem 13 and is therefore omitted. \blacksquare

4.7.3 Half-Duplex Bandlimited Gaussian Two-Way Relay Channel with Varying Background Noise

We now consider a half-duplex bandlimited Gaussian TRC with white noise. The MAC is used in two sessions and the BC is also used in two sessions. The output at node r in session j , $j = 1, 2$, of the MAC can be described as the convolution

$$Y_{r,j}(t) = (X_{1,j}(t) + X_{2,j}(t) + Z_{r,j}(t)) * h(t)$$

where $X_{i,j}(t)$ is the signal waveform transmitted by node t_i , $i = 1, 2$, in session j , $Z_{r,j}(t)$ is the waveform of a white Gaussian noise with two-sided power spectral density $N_{r,j}/2$ W/Hz and $h(t)$ is the impulse response of an ideal bandpass filter which cuts out all frequencies greater than W Hz. The outputs at node t_1 and node t_2 in session j , $j = 1, 2$, of the BC can be described as the convolutions

$$Y_{1,j}(t) = (X_{r,j}(t) + Z_{1,j}(t)) * h(t)$$

and

$$Y_{2,j}(t) = (X_{r,j}(t) + Z_{2,j}(t)) * h(t)$$

respectively, where $X_{r,j}(t)$ is the signal waveform transmitted by node r in session j and $Z_{1,j}(t)$ and $Z_{2,j}(t)$ are the waveforms of white Gaussian noise with two-sided power spectral density $N_{1,j}/2$ W/Hz and $N_{2,j}/2$ W/Hz respectively. Let the bandlimited Gaussian TRC be used over the time interval $[0, T]$. We assume that the time intervals $[0, \alpha_1 T]$ and $[\alpha_1 T, (\alpha_1 + \alpha_2) T]$ are allocated for the first and second sessions of the MAC respectively and the time intervals $[(\alpha_1 + \alpha_2) T, (\alpha_1 + \alpha_2 + \beta_1) T]$ and $[(\alpha_1 + \alpha_2 + \beta_1) T, (\alpha_1 + \alpha_2 + \beta_1 + \beta_2) T]$ are allocated for the first and second sessions of the BC respectively. Suppose the maximum power constraints for $x_{1,1}(t)$ in $[0, \alpha_1 T]$ and $x_{1,2}(t)$ in $[\alpha_1 T, (\alpha_1 + \alpha_2) T]$ are both P_1 and the maximum power constraints for $x_{2,1}(t)$ in $[0, \alpha_1 T]$ and $x_{2,2}(t)$ in $[\alpha_1 T, (\alpha_1 + \alpha_2) T]$ are both P_2 . Suppose the maximum power constraints for $x_{r,1}(t)$ in $[(\alpha_1 + \alpha_2) T, (\alpha_1 + \alpha_2 + \beta_1) T]$ and $x_{r,2}(t)$ in $[(\alpha_1 + \alpha_2 + \beta_1) T, (\alpha_1 + \alpha_2 + \beta_1 + \beta_2) T]$ are both P_r . Following similar procedures for deriving the capacity of the bandlimited Gaussian channel from the capacity of the discrete-time Gaussian channel in [17–19], we can derive the following result from Theorem 15.

Theorem 16 For the half-duplex bandlimited Gaussian TRC with varying background noise of bandwidth W with two-sided noise power spectral densities $N_{1,1}/2$, $N_{1,2}/2$, $N_{2,1}/2$, $N_{2,2}/2$, $N_{r,1}/2$ and $N_{r,2}/2$ and with maximum power constraints P_1 , P_2 and P_r ,

$$\mathcal{R}_{\alpha_1, \alpha_2, \beta_1, \beta_2} \subset \left\{ \begin{array}{l} (R_1, R_2) \\ \in \mathbb{R}^2 \end{array} \left| \begin{array}{l} R_1 \geq 0, R_2 \geq 0, \\ R_1 \leq \min \left\{ \alpha_1 W \log_2 \left(1 + \frac{P_1}{N_{r,1}W} \right) + \right. \\ \quad \alpha_2 W \log_2 \left(1 + \frac{P_1}{N_{r,2}W} \right), \\ \quad \beta_1 W \log_2 \left(1 + \frac{P_r}{N_{2,1}W} \right) + \\ \quad \left. \beta_2 W \log_2 \left(1 + \frac{P_r}{N_{2,2}W} \right) \right\}, \\ R_2 \leq \min \left\{ \alpha_1 W \log_2 \left(1 + \frac{P_2}{N_{r,1}W} \right) + \right. \\ \quad \alpha_2 W \log_2 \left(1 + \frac{P_2}{N_{r,2}W} \right), \\ \quad \beta_1 W \log_2 \left(1 + \frac{P_r}{N_{1,1}W} \right) + \\ \quad \left. \beta_2 W \log_2 \left(1 + \frac{P_r}{N_{1,2}W} \right) \right\}. \end{array} \right\}.$$

□ End of chapter.

Chapter 5

Performance Bounds in Cellular Relay Network

Summary

We investigate by simulation the performance bounds of physical-layer network coding (PNC) in a cellular relay network consisting of multiple users, multiple relays and multiple base stations. We model the cellular relay network as a collection of two-node point-to-point systems and three-node networks, where each two-node point-to-point system consists of two bandlimited Gaussian channels and each three-node network consists of a bandlimited Gaussian two-way relay channel. We investigate by simulation three PNC strategies in Sections 5.2, 5.3 and 5.4 and apply the theoretical outer bounds for PNC in Section 4.6.4 and Section 4.7.3 in the simulation. By comparing our simulation results for PNC in Sections 5.2.3, 5.3.3 and 5.4.3 with our simulation results for symbol-level network coding in Section 3.2.5, we show that the average maximum equal-rate throughput over all users under each PNC strategy is generally worse than the average maximum equal-rate throughput over all users under some routing strategy. This is possibly due to larger interference among the nodes under the PNC strategies compared with the routing strategy.

In Chapter 4, we obtained an outer bound on the rate region achievable by PNC schemes for a half-duplex bandlimited Gaussian TRC. In this chapter, we obtain performance bounds of PNC in a cellular relay network. Such a network is identical to the one discussed in Chapter 3 except that we model each multi-hop user, its associated relay and its associated base station together as a three-node network. In other words, we model the cellular relay network as a collection of two-node point-to-point systems (each consisting of a single-hop user and a base station) and three-node networks (each consisting of a single-hop user, a relay and a base station). Each two-node point-to-point system consists of two bandlimited Gaussian channels and each three-node network consists of a half-duplex bandlimited Gaussian TRC. As in the case of symbol-level network coding, since the transmitting nodes in the same cell and different cells can interfere with each other in the cellular relay network, each three-node network in the cellular relay network experiences more interference than an isolated three-node network. Consequently, the benefit of PNC in the cellular relay network becomes unclear. This motivates us to obtain performance bounds of PNC in the network.

Since the performance bounds of PNC cannot be analyzed theoretically due to the large number of nodes, we conduct the investigation by MATLAB simulation. Three PNC strategies are investigated by simulation and the theoretical outer bounds for PNC in Section 4.6.4 and Section 4.7.3 are applied in the simulation. Our simulation results show that the average maximum equal-rate throughput over all users under each PNC strategy is generally worse than the average maximum equal-rate throughput over all users under some routing strategy. This is possibly due to larger interference

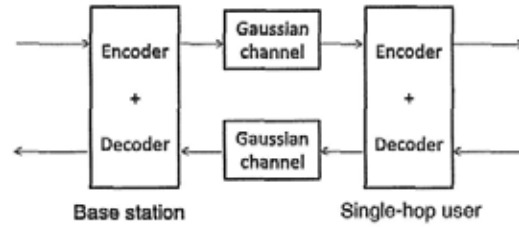


Figure 5.1: A two-node point-to-point system.

among the nodes under the PNC strategies compared with the routing strategy.

5.1 Simulation Assumptions

In this chapter, we use the same simulation assumptions as those in Section 3.1 except for the assumption for the channel model, which is described as follows. We model the cellular relay network as a collection of two-node point-to-point systems and three-node networks in such a way that each single-hop user and its associated base station are modeled as a two-node point-to-point system, and each multi-hop user, its associated relay and its associated base station are modeled as a three-node network. Each two-node point-to-point system consists of two independent bandlimited Gaussian channels as shown in Figure 5.1. Each three-node network consists of a half-duplex bandlimited Gaussian TRC as shown in Figure 5.2. We assume that every single-hop user exchanges independent messages with the base station through the corresponding two-node point-to-point system and every multi-hop user exchanges independent messages with the base station through the three-node network.

Let (u, v) denote a bandlimited Gaussian channel from node u to node v in a two-node point-to-point system. As discussed in Sec-

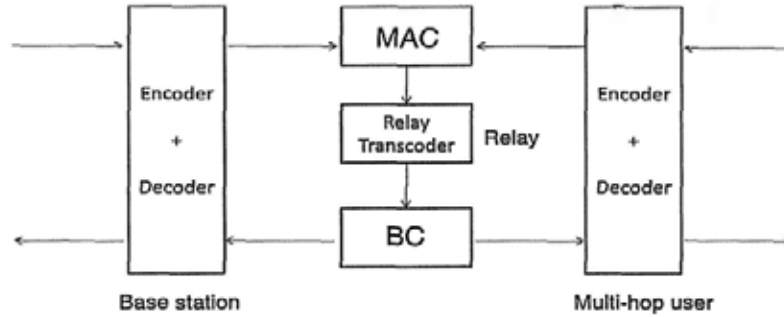


Figure 5.2: A three-node network.

tion 3.1.5, the channel capacity of a bandlimited Gaussian channel is assumed to be

$$W \log_2(1 + \min\{\sigma, 100\}) \text{ bits per second}$$

where σ is the calculated SNIR of (u, v) and $\min\{\sigma, 100\}$ is the SNIR of (u, v) after an upper bound of 20dB has been imposed. Since every transmission uses a frequency spectrum of bandwidth 10MHz and the one-sided power spectral density of the thermal noise is $10^{-17.4}$ mW/Hz, it follows that

$$W = 10^7 \text{ Hz}$$

and

$$\sigma = \frac{\text{received signal power at } v \text{ from } u}{10^{-20.4}W + \text{interference power at } v \text{ from nodes other than } u}$$

5.2 Strategy 1

In this section, we propose a transmission mode for the cellular relay network called *Strategy 1*, under which the transmissions in the network are organized into two phases. We will investigate by

simulation the average upper bound on the equal-rate throughput over all users when each multi-hop user uses a PNC scheme. The theoretical outer bound for PNC in Section 4.6.4 will be applied in the simulation.

5.2.1 Resource and Interference Management

To facilitate discussion, we classify the transmissions in each two-node point-to-point system as follows:

Class 1: The base station transmits messages to the single-hop user.

Class 2: The single-hop user transmits messages to the base station.

Similarly, we classify the transmissions in each three-node network as follows:

Class 1: The base station and the multi-hop user simultaneously transmit messages to the relay.

Class 2: The relay transmits messages to the base station and the multi-hop user.

The transmissions in each of the seven cells are organized into two phases, denoted by Phase 1 and Phase 2. The transmissions in the seven cells are synchronized in such a way that the seven cells have the same periods for Phase 1 and Phase 2. The fractions of time allocated for the two phases are fixed and we let p_1 and p_2 denote the fractions of time allocated for Phase 1 and Phase 2 transmissions respectively. We call (p_1, p_2) a *two-phase tuple*. The transmissions in the two phases are described as follows.

Phase 1: In each cell, the time is divided into n equal-length time slots which are allocated to the n users in the cell. In each time slot, if the user is a single-hop user, the communication from the base station to the user is conducted over the corresponding two-node system by Class 1 transmissions. If the user is a multi-hop user, the communication between the user and the base station through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 2: In each cell, the time is divided into two subphases such that the fractions of time allocated for the first subphase and the second subphase are $p_2(1 - \omega/n)$ and $p_2\omega/n$ respectively, where ω is the number of multi-hop users in the cell. The first subphase is allocated to single-hop users and the second subphase is allocated to multi-hop users. During the first subphase, all the single-hop users simultaneously transmit messages in such a way that the communication from each single-hop user to the base station is conducted over the corresponding two-node system by Class 2 transmissions.

During the second subphase, the time is divided into ω equal-length time slots which are allocated to the ω multi-hop users in the cell. In each time slot, the communication between the multi-hop user and the base station through the relay is conducted over the corresponding three-node network by Class 2 transmissions.

In each of the phases above, each listening station in a two-node system or a three-node network regards unintended signals as independent Gaussian noise.

Given a two-phase tuple (p_1, p_2) , the time allocations for each two-node system and each three-node network in Phase 1 are as follows:

1. For each two-node system, the fraction of time allocated for Class 1 transmissions is p_1/n .
2. For each three-node network, the fraction of time allocated for Class 1 transmissions is p_1/n .

The time allocations for each two-node system and each three-node network in Phase 2 are as follows:

1. For each two-node system, the fraction of time allocated for Class 2 transmissions is $p_2(1 - \omega/n)$.
2. For each three-node network, the fraction of time allocated for Class 2 transmissions is p_2/n .

To facilitate understanding, the fractions of time allocated to each two-node point-to-point system as well as each three-node network for Phase 1 and Phase 2 transmissions are shown in Table 5.1. For each single-hop user u and its associated base station b , the transmissions in the corresponding two-node point-to-point system are shown on the left side of Table 5.2. For each multi-hop user s , its associated relay r and its associated base station b , the transmissions in the corresponding three-node network are shown on the right side of Table 5.2.

5.2.2 Performance Upper Bound for Individual User

For each single-hop user u and its associated base station b , the fractions of time allocated to the corresponding two-node point-to-point system for Phase 1 and Phase 2 transmissions are p_1/n and

	Two-node system	Three-node network
Phase 1	p_1/n	p_1/n
Phase 2	$p_2(1 - \omega/n)$	p_2/n

Table 5.1: The fractions of time allocated to each two-node point-to-point system and each three-node network under Strategy 1.

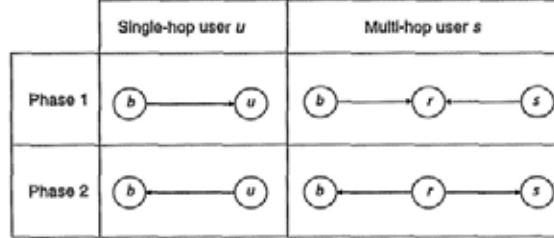


Table 5.2: The transmissions in each two-node point-to-point system and each three-node network under Strategy 1.

$p_2(1 - \omega/n)$ respectively, which implies that the maximum equal-rate throughput for u is $\min\{p_1A_u/n, p_2(1 - \omega/n)B_u\}$ where A_u is the capacity of (b, u) in Phase 1 and B_u is the capacity of (u, b) in Phase 2. To facilitate discussion, we call $\min\{p_1A_u/n, p_2(1 - \omega/n)B_u\}$ the *PNC upper bound for u under Strategy 1*.

For each multi-hop user s which communicates with its associated base station b through its associated relay r , the fractions of time allocated to the corresponding three-node network for Phase 1 and Phase 2 transmissions are p_1/n and p_2/n respectively. By Theorem 14 in Section 4.6.4, the equal-rate throughput achievable by PNC schemes for the three-node network is upper bounded by

$$\min \left\{ \alpha W \log_2 \left(1 + \frac{P_1}{N_r W} \right), \alpha W \log_2 \left(1 + \frac{P_2}{N_r W} \right), \beta W \log_2 \left(1 + \frac{P_r}{N_2 W} \right), \beta W \log_2 \left(1 + \frac{P_r}{N_1 W} \right) \right\} \quad (5.1)$$

where

1. α and β are the fractions of time allocated for Phase 1 and Phase 2 transmissions respectively,
2. P_1 is the signal power of b for Phase 1 transmissions,
3. P_2 is the signal power of s for Phase 1 transmissions,
4. P_r is the signal power of r for Phase 2 transmissions,
5. N_r is the one-sided noise power spectral density at r during Type 1 transmissions,
6. N_1 is the one-sided noise power spectral density at node b during Phase 2 transmissions,
7. N_2 is the one-sided noise power spectral density at node s during Phase 2 transmissions.

We assume that the interference incurred on one node by other nodes behaves as independent Gaussian noise. Since propagation loss and SNIR ceiling are considered in the cellular relay network, the upper bound on the equal-rate throughput in (5.1) becomes

$$\min \left\{ \begin{aligned} &\frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{b \rightarrow r}, 100\}), \\ &\frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{s \rightarrow r}, 100\}), \\ &\frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{r \rightarrow s}, 100\}), \\ &\frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{r \rightarrow b}, 100\}) \end{aligned} \right\}, \quad (5.2)$$

where

1. $\gamma_{b \rightarrow r} = \frac{\text{received signal power at } r \text{ from } b}{10^{-20} 4W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (b, r) for Phase 1 transmissions,

2. $\gamma_{s \rightarrow r} = \frac{\text{received signal power at } r \text{ from } s}{10^{-20} 4W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (s, r) for Phase 1 transmissions,
3. $\gamma_{r \rightarrow s} = \frac{\text{received signal power at } s \text{ from } r}{10^{-20} 4W + \text{interference power at } s \text{ from nodes other than } r}$ is the SNIR of (r, s) for Phase 2 transmissions,
4. $\gamma_{r \rightarrow b} = \frac{\text{received signal power at } b \text{ from } r}{10^{-20} 4W + \text{interference power at } b \text{ from nodes other than } r}$ is the SNIR of (r, b) for Phase 2 transmissions.

We call the upper bound in (5.2) the *PNC upper bound for s under Strategy 1*.

For each multi-hop user, since the transmission mode of Strategy 1 and the three-phase mode in Section 3.2 are different, the PNC upper bound under Strategy 1 and the equal-rate throughput achievable by the optimal routing scheme under the three-phase mode (cf. Section 3.2.4) cannot be compared directly due to different background noise. Therefore, we will compare in the following subsection the performance of PNC under Strategy 1 with the performance of routing under the three-phase mode using our simulation results. Note that we can compare the equal-rate throughput achievable by the optimal routing scheme with the XOR scheme directly for each user in Section 3.2.4 because both schemes are under the same transmission mode.

5.2.3 System Performance Upper Bound

We are interested in the average PNC upper bound over all users under Strategy 1 for every $n = 1, 4, 10, 20, 30$ (n is the number of users in each cell), every $m = 0, 2, 5, 10, 20, 30$ (m is the number of relays in each cell) and every two-phase tuple

$$(p_1, p_2) \in \{(i/64, (64 - i)/64) \mid i = 1, 2, \dots, 63\}. \quad (5.3)$$

The two-phase tuples in (5.3) lie on the line

$$p_1 + p_2 = 1.$$

When $n = 1$, the cellular relay network resembles low-density cellular networks. When $n = 4$ or 10 , the cellular relay network resembles medium-density cellular networks. When $n = 20$ or 30 , the cellular relay network resembles high-density cellular networks.

We generate 6000 different networks for each n and each m in our simulation. In each network, we calculate the average PNC upper bound over the users in the central cell for every (p_1, p_2) according to the formulae in Section 5.2.2. Then, we calculate the sample mean of the 6000 average PNC upper bounds under Strategy 1 for each n , each m and each (p_1, p_2) . Using standard arguments, we obtain that a 95% confidence interval for each sample mean μ is $(0.99\mu, 1.01\mu)$.

For each n and each m , let $\pi_{m,n}^{**}$ denote the two-phase tuple in (5.3) that maximizes the sample mean of the average PNC upper bounds under Strategy 1, and the $\pi_{m,n}^{**}$'s are displayed in Table 5.3. Under Strategy 1 with two-phase tuple $\pi_{m,n}^{**}$, we call the sample mean of the average PNC upper bounds the *average PNC upper bound under Strategy 1 for n users and m relays*, which is displayed in Table 5.4.

Table 5.4 shows that for each m , the average PNC upper bound under Strategy 1 decreases as n increases. In addition, the average PNC upper bound under Strategy 1 for each n decreases as m increases.

Comparing Table 3.6 in Section 3.2.5 with Table 5.4, we observe that the performance of the two-phase PNC schemes under Strategy 1 is generally worse than the performance of the optimal routing

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64)	(20/64, 44/64)	(21/64, 43/64)	(22/64, 42/64)	(23/64, 41/64)	(25/64, 39/64)
$n = 4$	(29/64, 35/64)	(29/64, 35/64)	(28/64, 36/64)	(29/64, 35/64)	(29/64, 35/64)	(29/64, 35/64)
$n = 10$	(31/64, 33/64)	(31/64, 33/64)	(31/64, 33/64)	(31/64, 33/64)	(32/64, 32/64)	(32/64, 32/64)
$n = 20$	(30/64, 34/64)	(30/64, 34/64)	(31/64, 33/64)	(32/64, 32/64)	(33/64, 31/64)	(34/64, 30/64)
$n = 30$	(29/64, 35/64)	(30/64, 34/64)	(31/64, 33/64)	(32/64, 32/64)	(33/64, 31/64)	(35/64, 29/64)

Table 5.3: Two-phase tuple $\pi_{m,n}^{**}$ maximizing the average PNC upper bound.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9500	3.8894	3.6537	3.3105	3.1651
$n = 4$	1.2556	1.2011	1.1217	1.0133	0.8453	0.7383
$n = 10$	0.4210	0.4043	0.3802	0.3452	0.2945	0.2607
$n = 20$	0.1704	0.1636	0.1536	0.1391	0.1189	0.1060
$n = 30$	0.0994	0.0951	0.0880	0.0803	0.0692	0.0613

Table 5.4: Average PNC upper bound under Strategy 1 (10^6 bits per second).

schemes under the three-phase mode. One possible reason for the poor performance of PNC schemes is that the interference from the surrounding cells at the central cell is larger under Strategy 1 compared with the three-phase routing mode.

5.3 Strategy 2

In this section, we propose a transmission mode for the cellular relay network called *Strategy 2*. The only difference between Strategy 1 and Strategy 2 of PNC is that the transmission sequence for single-hop users in the two phases is swapped (the transmission sequence for multi-hop users remains unchanged). We will investigate by simulation the average upper bound on the equal-rate throughput over all users when each multi-hop user uses a PNC scheme. The theoretical outer bound for PNC in Section 4.6.4 will be applied in the simulation.

5.3.1 Resource and Interference Management

To facilitate discussion, we classify the transmissions in each two-node point-to-point system as follows:

Class 1: The base station transmits messages to the single-hop user.

Class 2: The single-hop user transmits messages to the base station.

Similarly, we classify the transmissions in each three-node network as follows:

Class 1: The base station and the multi-hop user simultaneously transmit messages to the relay.

Class 2: The relay transmits messages to the base station and the multi-hop user.

The transmissions in each of the seven cells are organized into two phases, denoted by Phase 1 and Phase 2. The transmissions in the seven cells are synchronized in such a way that the seven cells have the same periods for Phase 1 and Phase 2. The fractions of time allocated for the two phases are fixed and we let p_1 and p_2 denote the fractions of time allocated for Phase 1 and Phase 2 transmissions respectively. We call (p_1, p_2) a *two-phase tuple*. The transmissions in the two phases are described as follows.

Phase 1: In each cell, the time is divided into two subphases such that the fractions of time allocated for the first subphase and the second subphase are $p_1(1 - \omega/n)$ and $p_1\omega/n$ respectively, where ω is the number of multi-hop users in the cell. The first subphase is allocated to single-hop users and the second subphase is allocated to multi-hop users. During the first subphase, all the single-hop users simultaneously transmit messages in such a way that the communication from each single-hop user to the base station is conducted over the corresponding two-node system by Class 2 transmissions.

During the second subphase, the time is divided into ω equal-length time slots which are allocated to the ω multi-hop users in the cell. In each time slot, the communication between the multi-hop user and the base station through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 2: In each cell, the time is divided into n equal-length time slots which are allocated to the n users in the cell. In each time

slot, if the user is a single-hop user, the communication from the base station to the user is conducted over the corresponding two-node system by Class 1 transmissions. If the user is a multi-hop user, the communication between the user and the base station through the relay is conducted over the corresponding three-node network by Class 2 transmissions.

In each of the phases above, each listening station in a two-node system or a three-node network regards unintended signals as independent Gaussian noise.

Given a two-phase tuple (p_1, p_2) , the time allocations for each two-node system and each three-node network in Phase 1 are as follows:

1. For each two-node system, the fraction of time allocated for Class 2 transmissions is $p_1(1 - \omega/n)$.
2. For each three-node network, the fraction of time allocated for Class 1 transmissions is p_1/n .

The time allocations for each two-node system and each three-node network in Phase 2 are as follows:

1. For each two-node system, the fraction of time allocated for Class 1 transmissions is p_2/n .
2. For each three-node network, the fraction of time allocated for Class 2 transmissions is p_2/n .

The fractions of time allocated to each two-node point-to-point system as well as each three-node network for Phase 1 and Phase 2 transmissions are shown in Table 5.5. For each single-hop user u and its associated base station b , the transmissions in the corresponding two-node point-to-point system are shown on the left side

	Two-node system	Three-node network
Phase 1	$p_1(1 - \omega/n)$	p_1/n
Phase 2	p_2/n	p_2/n

Table 5.5: The fractions of time allocated to each two-node point-to-point system and each three-node network under Strategy 2.

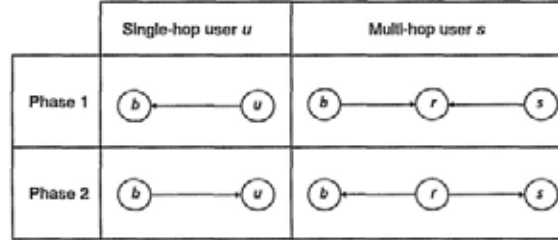


Table 5.6: The transmissions in each two-node point-to-point system and each three-node network under Strategy 2.

of Table 5.6. For each multi-hop user s , its associated relay r and its associated base station b , the transmissions in the corresponding three-node network are shown on the right side of Table 5.6.

5.3.2 Performance Upper Bound for Individual User

For each single-hop user u and its associated base station b , the fraction of time allocated to the corresponding two-node point-to-point system for Phase 1 and Phase 2 transmissions are $p_1(1 - \omega/n)$ and p_2/n respectively, which implies that the maximum equal-rate throughput for u is $\min\{p_1(1 - \omega/n)A_u, p_2B_u/n\}$ where A_u is the capacity of (u, b) in Phase 1 and B_u is the capacity of (b, u) in Phase 2. To facilitate discussion, we call $\min\{p_1(1 - \omega/n)A_u, p_2B_u/n\}$ the *PNC upper bound for u under Strategy 2*.

For each multi-hop user s which communicates with the base station b through its associated relay r , the fraction of time allocated to the corresponding three-node network for Phase 1 and Phase 2

transmissions are p_1/n and p_2/n respectively. Following similar procedures and using similar notations for Strategy 1 of PNC in Section 5.2.2, we can obtain that an upper bound on the equal-rate throughput for u in the network is

$$\min \left\{ \begin{aligned} & \frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{b \rightarrow r}, 100\}), \\ & \frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{s \rightarrow r}, 100\}), \\ & \frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{r \rightarrow s}, 100\}), \\ & \frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{r \rightarrow b}, 100\}) \end{aligned} \right\}, \quad (5.4)$$

where

1. $\gamma_{b \rightarrow r} = \frac{\text{received signal power at } r \text{ from } b}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (b, r) for Phase 1 transmissions,
2. $\gamma_{s \rightarrow r} = \frac{\text{received signal power at } r \text{ from } s}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (s, r) for Phase 1 transmissions,
3. $\gamma_{r \rightarrow s} = \frac{\text{received signal power at } s \text{ from } r}{10^{-20.4}W + \text{interference power at } s \text{ from nodes other than } r}$ is the SNIR of (r, s) for Phase 2 transmissions,
4. $\gamma_{r \rightarrow b} = \frac{\text{received signal power at } b \text{ from } r}{10^{-20.4}W + \text{interference power at } b \text{ from nodes other than } r}$ is the SNIR of (r, b) for Phase 2 transmissions.

As in Strategy 1, the upper bound is obtained by applying Theorem 14 in Section 4.6.4. We call the upper bound in (5.4) the *PNC upper bound for s under Strategy 2*.

5.3.3 System Performance Upper Bound

We are interested in the average PNC upper bound over all users under Strategy 2 for every $n = 1, 4, 10, 20, 30$ (n is the number of

users in each cell), every $m = 0, 2, 5, 10, 20, 30$ (m is the number of relays in each cell) and every two-phase tuple in the set

$$(p_1, p_2) \in \{(i/64, (64 - i)/64) \mid i = 1, 2, \dots, 63\}. \quad (5.5)$$

The two-phase tuples in (5.5) lie on the line

$$p_1 + p_2 = 1.$$

When $n = 1$, the cellular relay network resembles low-density cellular networks. When $n = 4$ or 10 , the cellular relay network resembles medium-density cellular networks. When $n = 20$ or 30 , the cellular relay network resembles high-density cellular networks.

We generate 6000 different networks for each n and each m in our simulation. In each network, we calculate the average PNC upper bound over all users in the central cell for every (p_1, p_2) according to the formulae in Section 5.3.2. Then, we calculate the sample mean of the 6000 average PNC upper bounds under Strategy 2 for each n , each m and each (p_1, p_2) . Using standard arguments, we obtain that a 95% confidence interval for each sample mean μ is $(0.99\mu, 1.01\mu)$.

For each n and each m , let $\theta_{m,n}^{**}$ denote the two-phase tuple in (5.5) that maximizes the sample mean of the average PNC upper bounds under Strategy 2, and the $\theta_{m,n}^{**}$'s are displayed in Table 5.7. Under Strategy 2 with two-phase tuple $\theta_{m,n}^{**}$, we call the sample mean of the average PNC upper bounds the *average PNC upper bound under Strategy 2 for n users and m relays*, which is displayed in Table 5.8.

Table 5.8 shows that for each m , the average PNC upper bound under Strategy 2 decreases as n increases. In addition, the average PNC upper bound under Strategy 2 for each n decreases as m in-

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(45/64, 19/64)	(44/64, 20/64)	(44/64, 20/64)	(42/64, 22/64)	(42/64, 22/64)	(40/64, 24/64)
$n = 4$	(35/64, 29/64)	(36/64, 28/64)	(36/64, 28/64)	(36/64, 28/64)	(37/64, 27/64)	(37/64, 27/64)
$n = 10$	(33/64, 31/64)	(34/64, 30/64)	(33/64, 31/64)	(34/64, 30/64)	(34/64, 30/64)	(34/64, 30/64)
$n = 20$	(34/64, 30/64)	(34/64, 30/64)	(34/64, 30/64)	(33/64, 31/64)	(33/64, 31/64)	(33/64, 31/64)
$n = 30$	(35/64, 29/64)	(34/64, 30/64)	(34/64, 30/64)	(33/64, 31/64)	(33/64, 31/64)	(33/64, 31/64)

Table 5.7: Two-phase tuple $\theta_{m,n}^{**}$ maximizing the average PNC upper bound.

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9382	3.8611	3.6015	3.2128	3.0254
$n = 4$	1.2556	1.1994	1.1174	1.0004	0.8152	0.6976
$n = 10$	0.4210	0.4036	0.3778	0.3391	0.2817	0.2427
$n = 20$	0.1704	0.1630	0.1521	0.1358	0.1120	0.0961
$n = 30$	0.0994	0.0947	0.0869	0.0778	0.0645	0.0545

Table 5.8: Average PNC upper bound under Strategy 2 (10^6 bits per second).

creases. Comparing Table 5.8 with Table 5.4, we observe that the average PNC upper bound under Strategy 2 is always worse than the average PNC upper bound under Strategy 1.

5.4 Strategy 3

In this section, we propose a transmission mode for the cellular relay network called *Strategy 3*, under which the transmissions in the network are organized into three phases. The main difference between Strategy 3 and the previous two strategies is that Strategy 3 uses a three-phase transmission mode instead of a two-phase transmission mode. We will investigate by simulation the average upper bound on the equal-rate throughput over all users when each multi-hop user uses a PNC scheme. The theoretical outer bound for PNC in Section 4.7.3 will be applied in the simulation.

5.4.1 Resource and Interference Management

To facilitate discussion, we classify the transmissions in each two-node point-to-point system as follows:

Class 1: The base station transmits messages to the single-hop user.

Class 2: The single-hop user transmits messages to the base station.

Similarly, we classify the transmissions in each three-node network as follows:

Class 1: The base station and the multi-hop user simultaneously transmit messages to the relay.

Class 2: The relay transmits messages to the base station and the multi-hop user.

The transmissions in each of the seven cells are organized into three phases, denoted by Phase 1, Phase 2 and Phase 3. The transmissions in the seven cells are synchronized in such a way that the seven cells have the same periods for Phase 1, Phase 2 and Phase 3. The fractions of time allocated for the three phases are fixed and we let p_1 , p_2 and p_r denote the fractions of time allocated for Phase 1, Phase 2 and Phase 3 transmissions respectively. We call (p_1, p_2, p_r) a *three-phase tuple*. The transmissions in the three phases are described as follows.

Phase 1: In each cell, the phase is divided into n equal-length time slots which are allocated to the n users in the cell. In each time slot, if the user is a single-hop user, the communication from the base station to the user is conducted over the corresponding two-node system by Class 1 transmissions. If the user is a multi-hop user, the communication between the user and the base station through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 2: In each cell, the time is divided into two subphases such that the fractions of time allocated for the first subphase and the second subphase are $p_2(1 - \omega/n)$ and $p_2\omega/n$ respectively, where ω is the number of multi-hop users in the cell. The first subphase is allocated to single-hop users and the second subphase is allocated to multi-hop users. During the first subphase, all the single-hop users simultaneously transmit messages in such a way that the communication from each single-hop user to the base station is conducted over the corresponding two-

node system by Class 2 transmissions.

During the second subphase, the time is divided into ω equal-length time slots which are allocated to the ω multi-hop users in the cell. In each time slot, the communication between the multi-hop user and the base station through the relay is conducted over the corresponding three-node network by Class 1 transmissions.

Phase 3: In each cell, all the active relays simultaneously transmit throughout the entire phase. For each multi-hop user u , the communication between u and the base station b through its associated relay r is as follows: Let $k \geq 1$ denote the number of multi-hop users that choose to communicate with b through r and let t_1, t_2, \dots, t_{k-1} denote the users other than u that communicate with b through r . This is illustrated in Figure 5.3. The phase is divided into k equal-length time slots, which are allocated to the k users served by relay r . Note that the value of k depends on the relay. In the time slot allocated to user u , the communication between u and b through r is conducted over the corresponding three-node network by Class 2 transmissions.

In each of the phases above, each listening station in a two-node system or a three-node network regards unintended signals as independent Gaussian noise.

Given a three-phase tuple (p_1, p_2, p_r) , the time allocations for each two-node system and each three-node network in Phase 1 are as follows:

1. For each two-node system, the fraction of time allocated for Class 1 transmissions is p_1/n .

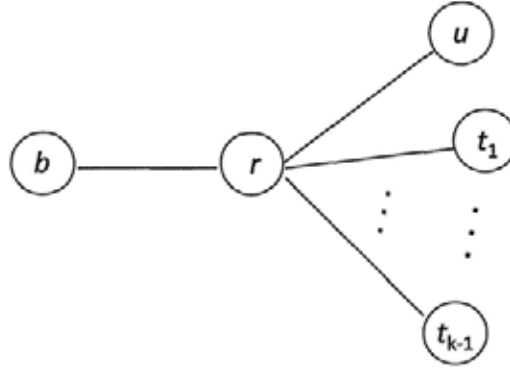


Figure 5.3: A relay r which serves k multi-hop users.

2. For each three-node network, the fraction of time allocated for Class 1 transmissions is p_1/n .

The time allocations for each two-node system and each three-node network in Phase 2 are as follows:

1. For each two-node system, the fraction of time allocated for Class 2 transmissions is $p_2(1 - \omega/n)$.
2. For each three-node network, the fraction of time allocated for Class 1 transmissions is p_2/n .

The time allocations for each two-node system and each three-node network in Phase 3 are as follows:

1. Each two-node system is idle.
2. For each three-node network, the fraction of time allocated for Class 2 transmissions is p_r/k , where $k \geq 1$ is the number of multi-hop users served by the relay.

To facilitate discussion, the fractions of time allocated to each two-node point-to-point system as well as each three-node network for Phase 1, Phase 2 and Phase 3 transmissions are shown in Table 5.9.

	Two-node system	Three-node network
Phase 1	p_1/n	p_1/n
Phase 2	$p_2(1 - \omega/n)$	p_2/n
Phase 3	0	p_r/k

Table 5.9: The fractions of time allocated to each two-node point-to-point system and each three-node network under Strategy 3.

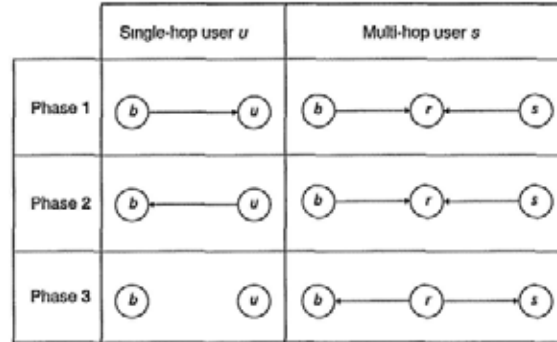


Table 5.10: The transmissions in each two-node point-to-point system and each three-node network under Strategy 3.

For each single-hop user u and its associated base station b , the transmissions in the corresponding two-node point-to-point system are shown on the left side of Table 5.10. For each multi-hop user s , its associated relay r and its associated base station b , the transmissions in the corresponding three-node network are shown on the right side of Table 5.10.

5.4.2 Performance Upper Bound for Individual User

For each single-hop user u and its associated base station b , the fraction of time allocated to the corresponding two-node point-to-point system for Phase 1 and Phase 2 transmissions are p_1/n and $p_2(1 - \omega/n)$ respectively, which implies that the maximum equal-rate throughput for u is $\min\{p_1 A_u/n, p_2(1 - \omega/n)B_u\}$ where A_u is the ca-

capacity of (b, u) in Phase 1 and B_u is the capacity of (u, b) in Phase 2. To facilitate discussion, we call $\min\{p_1 A_u/n, p_2(1 - \omega/n)B_u\}$ the *PNC upper bound for u under Strategy 3*. For each multi-hop user s which communicates with the base station b through a relay r , the fraction of time allocated to the corresponding three-node network for Phase 1, Phase 2 and Phase 3 transmissions are p_1/n , p_2/n and p_r/k respectively, where $k \geq 1$ is the number of multi-hop users served by relay r . By Theorem 16 in Section 4.7.3, the equal-rate throughput achievable by PNC schemes for the three-node network is upper bounded by

$$\min \left\{ \begin{aligned} &\alpha_1 W \log_2 \left(1 + \frac{P_1}{N_{r,1}W} \right) + \alpha_2 W \log_2 \left(1 + \frac{P_1}{N_{r,2}W} \right), \\ &\alpha_1 W \log_2 \left(1 + \frac{P_2}{N_{r,1}W} \right) + \alpha_2 W \log_2 \left(1 + \frac{P_2}{N_{r,2}W} \right), \\ &\beta W \log_2 \left(1 + \frac{P_r}{N_2W} \right), \beta W \log_2 \left(1 + \frac{P_r}{N_1W} \right) \end{aligned} \right\} \quad (5.6)$$

where

1. α_1 , α_2 and β are the fractions of time allocated for Phase 1, Phase 2 and Phase 3 transmissions respectively,
2. P_1 is the signal power of b during Phase 1 and Phase 2 transmissions,
3. P_2 is the signal power of s during Phase 1 and Phase 2 transmissions,
4. P_r is the signal power of r during Phase 3 transmissions,
5. $N_{r,1}$ and $N_{r,2}$ are the one-sided noise power spectral densities at r during Phase 1 and Phase 2 transmissions respectively,
6. N_1 is the one-sided noise power spectral density at node b dur-

ing Phase 3 transmissions,

7. N_2 is the one-sided noise power spectral density at node s during Phase 3 transmissions.

We assume that the interference incurred on one node by other nodes behaves as independent Gaussian noise. Since propagation loss and SNIR ceiling are considered in the cellular relay network, the upper bound on the equal-rate throughput in (5.6) becomes

$$\min \left\{ \begin{aligned} & \frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{b \rightarrow r, 1}, 100\}) + \\ & \frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{b \rightarrow r, 2}, 100\}), \\ & \frac{p_1 W}{n} \log_2(1 + \min\{\gamma_{s \rightarrow r, 1}, 100\}) + \\ & \frac{p_2 W}{n} \log_2(1 + \min\{\gamma_{s \rightarrow r, 2}, 100\}), \\ & \frac{p_r W}{k} \log_2(1 + \min\{\gamma_{r \rightarrow s}, 100\}), \\ & \frac{p_r W}{k} \log_2(1 + \min\{\gamma_{r \rightarrow b}, 100\}) \end{aligned} \right\}, \quad (5.7)$$

where

1. $\gamma_{b \rightarrow r, 1} = \frac{\text{received signal power at } r \text{ from } b}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (b, r) for Phase 1 transmissions,
2. $\gamma_{b \rightarrow r, 2} = \frac{\text{received signal power at } r \text{ from } b}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (b, r) for Phase 2 transmissions,
3. $\gamma_{s \rightarrow r, 1} = \frac{\text{received signal power at } r \text{ from } s}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (s, r) for Phase 1 transmissions,
4. $\gamma_{s \rightarrow r, 2} = \frac{\text{received signal power at } r \text{ from } s}{10^{-20.4}W + \text{interference power at } r \text{ from nodes other than } b \text{ and } s}$ is the SNIR of (s, r) for Phase 2 transmissions,

5. $\gamma_{r \rightarrow s} = \frac{\text{received signal power at } s \text{ from } r}{10^{-20.4}W + \text{interference power at } s \text{ from nodes other than } r}$ is the SNIR of (r, s) for Phase 3 transmissions,
6. $\gamma_{r \rightarrow b} = \frac{\text{received signal power at } b \text{ from } r}{10^{-20.4}W + \text{interference power at } b \text{ from nodes other than } r}$ is the SNIR of (r, b) for Phase 3 transmissions.

We call the upper bound in (5.7) the *PNC upper bound for s under Strategy 3*.

For each multi-hop user, since the transmission mode of Strategy 3 and the three-phase mode in Section 3.2 are different, the PNC upper bound under Strategy 3 and the equal-rate throughput achievable by the optimal routing scheme under the three-phase mode (cf. Section 3.2.4) cannot be compared directly due to different background noise. Therefore, we will compare in the following subsection the performance of PNC under Strategy 3 with the performance of routing under the three-phase mode using our simulation results.

5.4.3 System Performance Upper Bound

We are interested in the average PNC upper bound over all users under Strategy 3 for every $n = 1, 4, 10, 20, 30$ (n is the number of users in each cell), every $m = 0, 2, 5, 10, 20, 30$ (m is the number of relays in each cell) and every three-phase tuple (p_1, p_2, p_r) in the set

$$\left\{ \begin{array}{l} (p_1, p_2, p_r) \in \mathbb{R}^3 \\ \left. \begin{array}{l} p_1 \geq 0, p_2 \geq 0, p_r \geq 0, \\ p_1 + p_2 + p_r = 1, \\ p_1, p_2 \text{ and } p_r \text{ are multiples of } 1/64. \end{array} \right\} \right\}. \quad (5.8)$$

The three-phase tuples in (5.8) lie on the plane

$$p_1 + p_2 + p_r = 1.$$

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	(19/64, 45/64, 0)	(20/64, 44/64, 0)	(18/64, 40/64, 6)	(21/64, 40/64, 3)	(21/64, 40/64, 3)	(23/64, 38/64, 3)
$n = 4$	(29/64, 35/64, 0)	(28/64, 35/64, 1)	(28/64, 35/64, 1)	(28/64, 35/64, 1)	(27/64, 35/64, 2)	(26/64, 35/64, 3)
$n = 10$	(31/64, 33/64, 0)	(30/64, 33/64, 1)	(30/64, 33/64, 1)	(30/64, 33/64, 1)	(30/64, 32/64, 2)	(29/64, 32/64, 3)
$n = 20$	(30/64, 34/64, 0)	(29/64, 34/64, 1)	(30/64, 33/64, 1)	(30/64, 33/64, 1)	(30/64, 32/64, 2)	(30/64, 31/64, 3)
$n = 30$	(29/64, 35/64, 0)	(29/64, 34/64, 1)	(29/64, 34/64, 1)	(30/64, 33/64, 1)	(30/64, 32/64, 2)	(30/64, 31/64, 3)

Table 5.11: Three-phase tuple $\pi_{m,n}^{***}$ maximizing the average PNC upper bound.

When $n = 1$, the cellular relay network resembles low-density cellular networks. When $n = 4$ or 10 , the cellular relay network resembles medium-density cellular networks. When $n = 20$ or 30 , the cellular relay network resembles high-density cellular networks.

We generate 6000 different networks for each n and each m in our simulation. In each network, we calculate the average PNC upper bound over all users in the central cell for every (p_1, p_2, p_r) according to the formulae in Section 5.4.2. Then, we calculate the sample mean of the 6000 average PNC upper bounds under Strategy 3 for each n , each m and each (p_1, p_2, p_r) . Using standard arguments, we obtain that a 95% confidence interval for each sample mean μ is $(0.99\mu, 1.01\mu)$.

For each n and each m , let $\pi_{m,n}^{***}$ denote the three-phase tuple in (5.8) that maximizes the sample mean of the average PNC upper bounds under Strategy 3, and the $\pi_{m,n}^{***}$'s are displayed in Table 5.11. Under Strategy 3 with three-phase tuple $\pi_{m,n}^{***}$, we call the sample mean of the average PNC upper bounds the *average PNC upper bound under Strategy 3 for n users and m relays*, which is displayed in Table 5.12.

Table 5.12 shows that for each m , the average PNC upper bound

	$m = 0$	$m = 2$	$m = 5$	$m = 10$	$m = 20$	$m = 30$
$n = 1$	4.1346	3.9385	3.9341	3.6840	3.3376	3.1649
$n = 4$	1.2556	1.2033	1.1201	1.0059	0.8301	0.7213
$n = 10$	0.4210	0.4054	0.3793	0.3421	0.2877	0.2514
$n = 20$	0.1704	0.1645	0.1537	0.1379	0.1158	0.1014
$n = 30$	0.0994	0.0959	0.0883	0.0796	0.0673	0.0582

Table 5.12: Average PNC upper bound under Strategy 3 (10^6 bits per second).

under Strategy 3 decreases as n increases. In addition, the average PNC upper bound under Strategy 3 for each n decreases as m increases.

Comparing Table 3.6 in Section 3.2.5 with Table 5.12, we observe that the performance of the three-phase PNC schemes under Strategy 3 is generally worse than the performance of the optimal routing schemes under the three-phase mode. One possible reason for the poor performance of PNC schemes is that the interference from the surrounding cells at the central cell is larger under Strategy 3 compared with the three-phase routing mode.

□ End of chapter.

Chapter 6

Conclusion

Summary

The results for the benefit of symbol-level network coding schemes on the three-node point-to-point relay network, a model for the two-way relay channel (TRC), are summarized. The investigation of symbol-level network coding schemes on the cellular relay network by MATLAB simulation is described and our simulation results are summarized. In addition, the outer bound results of the capacity regions of the TRC models investigated in this thesis are summarized. In particular, the outer bound on the capacity region of the bandlimited Gaussian TRC is a theoretical outer bound on the capacity region achievable by physical-layer network coding (PNC). The investigation of PNC schemes on the cellular relay network by MATLAB simulation is described and our simulation results are summarized. Finally, several suggestions for further research are given.

In the first part of this thesis, we model the TRC as a three-node

point-to-point relay network and investigate the performance of two practical symbol-level network coding schemes: the four-session and five-session network coding schemes. The five-session schemes are more general than the four-session schemes, but at the same time they are more complicated. Several achievable rate regions for the four-session and five-session schemes are obtained. We show that the use of network coding rather than routing alone always enlarges the achievable rate region, in particular increases the maximum equal-rate throughput in the three-node point-to-point relay network.

Furthermore, we investigate the benefit of symbol-level network coding in a cellular relay network consisting of multiple users, multiple relays and multiple base stations and present the formulation of the network. We model the cellular relay network as a collection of two-node point-to-point systems and three-node point-to-point relay networks where each point-to-point channel is modeled as a bandlimited Gaussian channel. We propose the low-complexity XOR network coding schemes for the multi-hop users in the cellular relay network and show analytically that the equal-rate throughput achievable by each XOR scheme is always greater than or equal to the maximum equal-rate throughput achievable by all routing schemes. The benefit of the XOR schemes in the cellular relay network is further investigated by MATLAB simulation and our simulation results show that the use of the XOR scheme rather than routing schemes can improve the average maximum equal-rate throughput over all users under the three-phase mode as well as the four-phase mode. For low-density networks, the average network coding gain over all users can be greater than 17% and the average network coding gain over multi-hop users can be greater than 41%. For

medium-density networks, the average network coding gain over all users can be greater than 7% and the average network coding gain over multi-hop users can be greater than 45%. For high-density networks, the average network coding gain over all users can be greater than 4% and the average network coding gain over multi-hop users can be greater than 22%.

In the second part of this thesis, we investigate three models of full-duplex TRC (the discrete memoryless TRC without feedback, the discrete memoryless TRC with feedback and the Gaussian TRC with feedback) and four models of half-duplex TRC (the discrete memoryless TRC, the Gaussian TRC, the bandlimited Gaussian TRC and the bandlimited Gaussian TRC with varying background noise). We prove an outer bound for the discrete memoryless TRC without feedback and show that the use of feedback enlarges the capacity region of the discrete memoryless TRC. Knopp [22] stated an outer bound on the capacity region of each of the discrete memoryless TRC with feedback and the Gaussian TRC with feedback. However, the proofs in [22] are incomplete. We therefore present detailed proofs of the outer bounds stated in [22] and show that the outer bound on the capacity region of some discrete memoryless TRC with feedback is loose. In addition, we prove an outer bound on the capacity region of each of the half-duplex discrete memoryless TRC, the half-duplex Gaussian TRC, the half-duplex bandlimited Gaussian TRC and the half-duplex bandlimited Gaussian TRC with varying background noise. In particular, the outer bound on the capacity region of the half-duplex bandlimited Gaussian TRC is a theoretical outer bound on the capacity region achievable by PNC.

Furthermore, we model a cellular relay network consisting of mul-

multiple users, multiple relays and multiple base stations as a collection of two-node point-to-point systems and three-node networks, where each two-node point-to-point system consists of two bandlimited Gaussian channels and each three-node network consists of a half-duplex bandlimited Gaussian TRC. In order to obtain performance bounds for PNC schemes, we apply the outer bounds for PNC in Section 4.6.4 in the MATLAB simulation. Our simulation results show that the average maximum equal-rate throughput over all users under every PNC strategy investigated is generally worse than the average equal-rate throughput over all users under some routing strategy. This is possibly due to larger interference from the surrounding cells at the central cell under each PNC strategy compared with the routing strategy.

Directions for further research include obtaining tighter outer bounds on the capacity regions of different TRC models, proposing new practical transmission schemes for different TRC models and improving our simulation model so that the simulation results will reflect more accurately the benefit of network coding and PNC in practical cellular relay networks.

□ End of chapter.

Appendix A

Proofs for Chapter 2

Proof of Lemma 2: There exist $p'_1 \leq p_1$ and $p'_3 \leq p_3$ such that

$$R_1 = p'_1 C = p'_3 D.$$

Let $p'_2 = p_2$ and $p'_4 = p_4 + p_3 - p'_3$. Then,

$$\begin{aligned} R_2 &\leq \min\{p_2 E, p_3 D + p_4 F\} \\ &= \min\{p'_2 E, p'_3 D + (p_3 - p'_3) D + p_4 F\} \\ &\stackrel{(a)}{\leq} \min\{p'_2 E, p'_3 D + (p_3 - p'_3) F + p_4 F\} \\ &= \min\{p'_2 E, p'_3 D + p'_4 F\}, \end{aligned}$$

where (a) follows from (2.1). Since

$$\begin{aligned} p'_1 + p'_2 + p'_3 + p'_4 &\leq p_1 + p_2 + p'_3 + (p_4 + p_3 - p'_3) \\ &= p_1 + p_2 + p_3 + p_4 \\ &\leq 1, \end{aligned}$$

it follows that $s' = [p'_1 \ p'_2 \ p'_3 \ p'_4]$ is the desired allocation. ■

Proof of Lemma 3: Since \mathcal{S}_4 is the projection of a 6-dimensional

polyhedron on a 2-dimensional space, it can be shown by Fourier-Motzkin elimination [26] that $\mathcal{S}_4 = \mathcal{T}_4^*$. Since Fourier-Motzkin elimination is very tedious, we present a simpler alternative proof as follows. Suppose (R_1, R_2) is in \mathcal{S}_4 . Then, (R_1, R_2) is a non-negative pair such that

$$R_1 \leq \min\{p_1 C, p_3 D\}$$

and

$$R_2 \leq \min\{p_2 E, p_3 D + p_4 F\}$$

for some allocation $[p_1 \ p_2 \ p_3 \ p_4]$. By Lemma 2, there exists an allocation $[p'_1 \ p'_2 \ p'_3 \ p'_4]$ such that

$$R_1 = p'_1 C = p'_3 D \tag{A.1}$$

and

$$R_2 \leq \min\{p'_2 E, p'_3 D + p'_4 F\}. \tag{A.2}$$

We will show that (R_1, R_2) satisfies the four inequalities defining \mathcal{T}_4^* (cf. (2.3)), which will imply that $(R_1, R_2) \in \mathcal{T}_4^*$. The lemma then follows.

Clearly, (R_1, R_2) satisfies the first two inequalities of \mathcal{T}_4^* .

Consider the following chain of inequalities:

$$\begin{aligned}
R_2 - \frac{EF}{E+F} + R_1 \left(\frac{CEF + DEF - CDE}{CD(E+F)} \right) \\
&= \frac{1}{CD(E+F)} (CDER_2 + CDFR_2 - CDEF + CEFR_1 \\
&\quad + DEFR_1 - CDER_1) \\
&\stackrel{(a)}{\leq} \frac{1}{CD(E+F)} (CDEFp'_4 + CD^2Ep'_3 + CDEFp'_2 - CDEF \\
&\quad + CDEFp'_3 + CDEFp'_1 - CD^2Ep'_3) \\
&= \frac{EF}{E+F} (p'_1 + p'_2 + p'_3 + p'_4 - 1) \\
&\stackrel{(b)}{\leq} 0,
\end{aligned}$$

where (a) follows from (A.1) and (A.2), and (b) follows from (2.2).

Therefore,

$$R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{CEF + DEF - CDE}{CD(E+F)} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned}
R_2 - E + R_1 \left(\frac{CE + DE}{CD} \right) \\
&= \frac{CDR_2 - CDE + CER_1 + DER_1}{CD} \\
&\stackrel{(a)}{\leq} \frac{CDEp'_2 - CDE + CDEp'_3 + CDEp'_1}{CD} \\
&= E(p'_1 + p'_2 + p'_3 - 1) \\
&\stackrel{(b)}{\leq} -p_4E, \\
&\leq 0,
\end{aligned}$$

where (a) follows from (A.1) and (A.2), and (b) follows from (2.2).

Therefore,

$$R_2 \leq E - R_1 \left(\frac{CE + DE}{CD} \right).$$

■

Proof of Proposition 5: We will show that $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$ and $(\frac{CD}{C+D}, 0)$ are all the vertices of \mathcal{T}_4^* . The proposition then follows.

Let

$$L_1 : R_1 = 0,$$

$$L_2 : R_2 = 0,$$

$$L_3 : R_2 = \frac{EF}{E+F} - R_1 \left(\frac{CEF+DEF-CDE}{CD(E+F)} \right)$$

and

$$L_4 : R_2 = E - R_1 \left(\frac{CE+DE}{CD} \right)$$

be the four lines defining the boundary of \mathcal{T}_4^* (cf. (2.3)). The x -intercept and y -intercept of L_3 are $\frac{CDF}{CF+DF-CD}$ and $\frac{EF}{E+F}$ respectively. The x -intercept and y -intercept of L_4 are $\frac{CD}{C+D}$ and E respectively. Since

$$\begin{aligned} & \frac{CDF}{CF+DF-CD} - \frac{CD}{C+D} \\ &= \frac{C^2D^2}{(C+D)(C(F-D)+DF)} \\ &> 0 \end{aligned}$$

(cf. (2.1)) and

$$\begin{aligned} & E - \frac{EF}{E+F} \\ &= \frac{E^2}{E+F} \\ &> 0, \end{aligned}$$

the x -intercept of L_3 is larger than that of L_4 and the y -intercept of L_4 is larger than that of L_3 . By inspecting Figure A.1, we

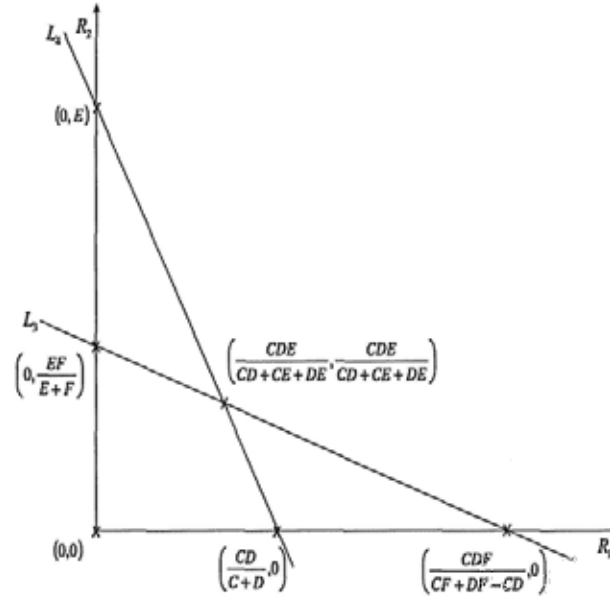


Figure A.1: L_3 and L_4 .

observe that the vertices of \mathcal{T}_4^* are $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{CD}{C+D}, 0)$ and $(\frac{CDE}{CD+CE+DE}, \frac{CDE}{CD+CE+DE})$, the intersection point between L_3 and L_4 .

■

Proof of Lemma 8: Since $\mathcal{S}_4 = \mathcal{T}_4$ by Theorem 1, it remains to show that $\mathcal{T}_5 \subset \mathcal{S}_4$. Suppose (R_1, R_2) is in \mathcal{T}_5 . Then, (R_1, R_2) is s -achievable for some allocation $s = [p_0 \ p_1 \ p_2 \ p_3 \ p_4]$. Using Lemma 7, we obtain

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\}.$$

Let $p'_1 = \frac{p_0A}{C} + p_1$, $p'_2 = \frac{p_0B}{E} + p_2$, $p'_3 = p_3$ and $p'_4 = p_4$ be four

non-negative numbers. Then,

$$R_1 \leq \min\{p'_1 C, p'_3 D\}$$

and

$$R_2 \leq \min\{p'_2 E, p'_3 D + p'_4 F\}.$$

If $\frac{A}{C} + \frac{B}{E} \leq 1$, then

$$\begin{aligned} p'_1 + p'_2 + p'_3 + p'_4 &= p_0 \left(\frac{A}{C} + \frac{B}{E} \right) + p_1 + p_2 + p_3 + p_4 \\ &\leq p_0 + p_1 + p_2 + p_3 + p_4 \\ &\leq 1. \end{aligned}$$

It then follows that $(R_1, R_2) \in \mathcal{S}_4$ (cf. (2.4)). \blacksquare

Proof of Lemma 9: There exist $p'_0 \leq p_0$, $p'_1 \leq p_1$ and $p'_3 \leq p_3$ such that

$$R_1 = p'_0 A + p'_1 C = p'_3 D.$$

Let $p'_2 = p_2 + p_0 - p'_0$ and $p'_4 = p_4 + p_3 - p'_3$. Then,

$$\begin{aligned} R_2 &\leq \min\{p_0 B + p_2 E, p_3 D + p_4 F\} \\ &= \min\{p'_0 B + (p_0 - p'_0) B + p_2 E, p'_3 D + (p_3 - p'_3) D + p_3 F\} \\ &\stackrel{(a)}{\leq} \min\{p'_0 B + (p_0 - p'_0) E + p_2 E, p'_3 D + (p_3 - p'_3) F + p_4 F\} \\ &= \min\{p'_0 B + p'_2 E, p'_3 D + p'_4 F\}, \end{aligned}$$

where (a) follows from (2.1) and (2.8). Since

$$\begin{aligned}
 & p'_0 + p'_1 + p'_2 + p'_3 + p'_4 \\
 & \leq p'_0 + p_1 + (p_2 + p_0 - p'_0) + p'_3 + (p_4 + p_3 - p'_3) \\
 & = p_0 + p_1 + p_2 + p_3 + p_4 \\
 & \leq 1,
 \end{aligned}$$

it then follows that $s' = [p'_0 \ p'_1 \ p'_2 \ p'_3 \ p'_4]$ is the desired allocation. \blacksquare

Proof of Lemma 10: Since $\mathcal{S}_{A=B}$ is the projection of a 7-dimensional polyhedron on a 2-dimensional space, it can be shown by Fourier-Motzkin elimination [26] that $\mathcal{S}_{A=B} = \mathcal{T}_{5,A=B}^*$. Since Fourier-Motzkin elimination is very tedious, we present a simpler alternative proof as follows. Suppose (R_1, R_2) is in $\mathcal{S}_{A=B}$. Then, (R_1, R_2) is a non-negative pair such that

$$R_1 \leq \min\{p_0A + p_3C, p_4D\}$$

and

$$R_2 \leq \min\{p_0A + p_2E, p_3D + p_4F\}$$

for some allocation $[p_0 \ p_1 \ p_2 \ p_3 \ p_4]$. By Lemma 9, there exists an allocation $[p'_0 \ p'_1 \ p'_2 \ p'_3 \ p'_4]$ such that

$$R_1 = p'_0A + p'_1C = p'_3D \quad (\text{A.3})$$

and

$$R_2 \leq \min\{p'_0A + p'_2E, p'_3D + p'_4F\}. \quad (\text{A.4})$$

We will show that (R_1, R_2) satisfies the four inequalities defining $\mathcal{T}_{5,A=B}^*$ (cf. (2.11)), which will imply that $(R_1, R_2) \in \mathcal{T}_{5,A=B}^*$. The

lemma then follows.

Clearly, (R_1, R_2) satisfies the first two inequalities of $\mathcal{T}_{5,A=B}^*$.

Consider the following chain of inequalities:

$$\begin{aligned}
& R_2 - \frac{EF}{E+F} + R_1 \left(\frac{AEF + DEF - ADE - ADF}{AD(E+F)} \right) \\
&= \frac{1}{AD(E+F)} (ADER_2 + ADFR_2 - ADEF + AEFR_1 \\
&\quad + DEFR_1 - ADE R_1 - ADFR_1) \\
&\stackrel{(a)}{\leq} \frac{1}{AD(E+F)} (AD^2 E p'_3 + ADEF p'_4 + ADEF p'_2 \\
&\quad + A^2 DF p'_0 - ADEF + ADEF p'_3 + ADEF p'_0 \\
&\quad + CDEF p'_1 - AD^2 E p'_3 - ACDF p'_1 - A^2 DF p'_0) \\
&= \frac{ADEF(p'_0 + p'_2 + p'_3 + p'_4 - 1) + CDEF p'_1 - ACDF p'_1}{AD(E+F)} \\
&\stackrel{(b)}{\leq} \frac{DF}{AD(E+F)} (-AE p'_1 + CE p'_1 - AC p'_1) \\
&= \frac{p'_1 CEF}{A(E+F)} \left(1 - \frac{A}{C} - \frac{A}{E} \right) \\
&\stackrel{(c)}{\leq} 0,
\end{aligned}$$

where

(a) follows from (A.3) and (A.4),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF + DEF - ADE - ADF}{AD(E+F)} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned}
& R_2 - \frac{AC}{C-A} + R_1 \left(\frac{AC+AD}{D(C-A)} \right) \\
&= \frac{D(C-A)R_2 - ACD + ACR_1 + ADR_1}{D(C-A)} \\
&\stackrel{(a)}{\leq} \frac{1}{D(C-A)} (D(C-A)(p'_0A + p'_2E) - ACD \\
&\quad + ACDp'_3 + ACDp'_1 + A^2Dp'_0) \\
&= \frac{1}{D(C-A)} (ACDp'_0 - A^2Dp'_0 + CDEp'_2 - ADEp'_2 - ACD \\
&\quad + ACDp'_3 + ACDp'_1 + A^2Dp'_0) \\
&= \frac{1}{D(C-A)} (ACD(p'_0 + p'_1 + p'_3 - 1) + CDEp'_2 - ADEp'_2) \\
&\stackrel{(b)}{\leq} \frac{AC(-p'_2 - p'_4) + CEp'_2 - AEp'_2}{C-A} \\
&= \frac{-ACp'_4 - ACp'_2 + CEp'_2 - AEp'_2}{C-A} \\
&\leq \frac{-ACp'_2 + CEp'_2 - AEp'_2}{C-A} \\
&= \frac{CEp'_2 \left(1 - \frac{A}{C} - \frac{A}{E}\right)}{C-A} \\
&\stackrel{(c)}{\leq} 0,
\end{aligned}$$

where

(a) follows from (A.3) and (A.4),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{AC}{C-A} - R_1 \left(\frac{AC+AD}{D(C-A)} \right).$$

■

Proof of Proposition 12: We will show that $(0, 0)$, $(0, \frac{EF}{E+F})$,

$(\frac{AD}{A+D}, \frac{AD}{A+D})$ and $(\frac{CD}{C+D}, 0)$ are all the vertices of $\mathcal{T}_{5,A=B}^*$. The proposition then follows.

Let

$$L_1 : R_1 = 0,$$

$$L_2 : R_2 = 0,$$

$$L_3 : R_2 = \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-ADF}{AD(E+F)} \right)$$

and

$$L_4 : R_2 = \frac{AC}{C-A} - R_1 \left(\frac{AC+AD}{D(C-A)} \right)$$

be the four lines defining the boundary of $\mathcal{T}_{5,A=B}^*$ (cf. (2.11)). The x -intercept and y -intercept of L_3 are $\frac{ADEF}{AEF+DEF-ADE-ADF}$ and $\frac{EF}{E+F}$ respectively. The x -intercept and y -intercept of L_4 are $\frac{CD}{C+D}$ and $\frac{AC}{C-A}$ respectively. Since

$$\begin{aligned} & \frac{ADEF}{AEF+DEF-ADE-ADF} - \frac{CD}{C+D} \\ &= \frac{AD^2EF - CD^2EF + ACD^2E + ACD^2F}{(C+D)(AEF+DEF-ADE-ADF)} \\ &= \frac{CD^2EF(A/C + A/E + A/F - 1)}{(C+D)(AE(F-D) + DF(E-A))} \\ &> 0 \end{aligned}$$

(cf. (2.1), (2.8) and (2.10)) and

$$\begin{aligned} & \frac{AC}{C-A} - \frac{EF}{E+F} \\ &= \frac{ACE + ACF - CEF + AEF}{(C-A)(E+F)} \\ &= \frac{CEF(A/F + A/E + A/C - 1)}{(C-A)(E+F)} \\ &> 0 \end{aligned}$$

(cf. (2.7) and (2.10)), the x -intercept of L_3 is larger than that of L_4 and the y -intercept of L_4 is larger than that of L_3 . By inspecting

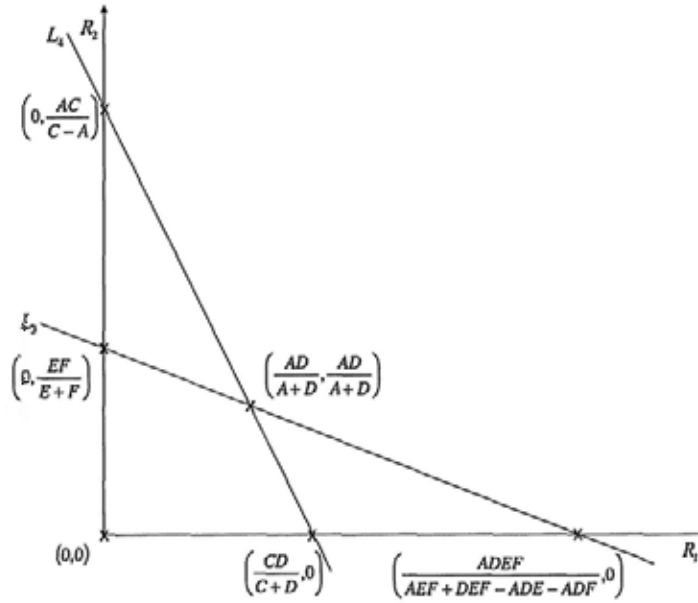


Figure A.2: L_3 and L_4 .

Figure A.2, we observe that the vertices of \mathcal{T}_4^* are $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{CD}{C+D}, 0)$ and $(\frac{AD}{A+D}, \frac{AD}{A+D})$, the intersection point between L_3 and L_4 . ■

Proof of Lemma 14: Since \mathcal{S}_5 is the projection of a 7-dimensional polyhedron on a 2-dimensional space, it can be shown by Fourier-Motzkin elimination [26] that $\mathcal{S}_5 = \mathcal{T}_{5,A>B}^*$. Since Fourier-Motzkin elimination is very tedious, we present a simpler alternative proof as follows. Suppose (R_1, R_2) is in \mathcal{S}_5 . Then, (R_1, R_2) is a non-negative pair such that

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\}$$

for some allocation $[p_0 \ p_1 \ p_2 \ p_3 \ p_4]$. By Lemma 9, there exists an

allocation $[p'_0 p'_1 p'_2 p'_3 p'_4]$ such that

$$R_1 = p'_0 A + p'_1 C = p'_3 D \quad (\text{A.5})$$

and

$$R_2 \leq \min\{p'_0 A + p'_2 E, p'_3 D + p'_4 F\}. \quad (\text{A.6})$$

We will show that (R_1, R_2) satisfies the five inequalities defining $\mathcal{T}_{5,A>B}^*$ (cf. (2.13)), which will imply that $(R_1, R_2) \in \mathcal{T}_{5,A>B}^*$. The lemma then follows.

Clearly, (R_1, R_2) satisfies the first two inequalities of $\mathcal{T}_{5,A>B}^*$.

Consider the following chain of inequalities:

$$\begin{aligned} & R_2 - \frac{EF}{E+F} + R_1 \left(\frac{AEF + DEF - ADE - BDF}{AD(E+F)} \right) \\ &= \frac{1}{AD(E+F)} (ADER_2 + ADFR_2 - ADEF + AEF R_1 \\ &\quad + DEFR_1 - ADE R_1 - BDF R_1) \\ &\stackrel{(a)}{\leq} \frac{1}{AD(E+F)} (AD^2 E p'_3 + ADEF p'_4 + ADEF p'_2 + ABDF p'_0 \\ &\quad - ADEF + ADEF p'_3 + ADEF p'_0 + CDEF p'_1 - AD^2 E p'_3 \\ &\quad - BCDF p'_1 - ABDF p'_0) \\ &= \frac{1}{AD(E+F)} (ADEF(p'_0 + p'_2 + p'_3 + p'_4 - 1) + CDEF p'_1 \\ &\quad - BCDF p'_1) \\ &\stackrel{(b)}{\leq} \frac{DF}{AD(E+F)} (-AE p'_1 + CE p'_1 - BC p'_1) \\ &= \frac{p'_1 CEF}{A(E+F)} \left(1 - \frac{A}{C} - \frac{B}{E} \right) \\ &\stackrel{(c)}{\leq} 0, \end{aligned}$$

where

(a) follows from (A.5) and (A.6),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF + DEF - ADE - BDF}{AD(E+F)} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned} & R_2 - E + R_1 \left(\frac{AE + DE - BD}{AD} \right) \\ &= \frac{ADR_2 - ADE + AER_1 + DER_1 - BDR_1}{AD} \\ &\stackrel{(a)}{\leq} \frac{1}{AD} (ABDp'_0 + ADEp'_2 - ADE + ADEp'_3 + CDEp'_1 \\ &\quad + ADEp'_0 - ABDp'_0 - BCDp'_1) \\ &= \frac{1}{A} (AEp'_0 + AEp'_2 + AEp'_3 - AE + CEp'_1 - BCp'_1) \\ &= \frac{1}{A} (AE(p'_0 + p'_2 + p'_3 - 1) + CEp'_1 - BCp'_1) \\ &\stackrel{(b)}{\leq} \frac{1}{A} (AE(-p'_1 - p'_4) + CEp'_1 - BCp'_1) \\ &\leq \frac{1}{A} (-AEp'_1 + CEp'_1 - BCp'_1) \\ &= \frac{CEp'_1}{A} \left(1 - \frac{A}{C} - \frac{B}{E} \right) \\ &\stackrel{(c)}{\leq} 0, \end{aligned}$$

where

(a) follows from (A.5) and (A.6),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq E - R_1 \left(\frac{AE + DE - BD}{AD} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned} & R_2 - \frac{BC}{C-A} + R_1 \left(\frac{BC + BD}{D(C-A)} \right) \\ &= \frac{D(C-A)R_2 - BCD + BCR_1 + BDR_1}{D(C-A)} \\ &\stackrel{(a)}{\leq} \frac{1}{D(C-A)} (D(C-A)(p'_0 B + p'_2 E) - BCD + BCDp'_3 \\ &\quad + BCDp'_1 + ABDp'_0) \\ &= \frac{1}{D(C-A)} (BCDp'_0 + CDEp'_2 - ABDp'_0 - ADEp'_2 \\ &\quad - BCD + BCDp'_3 + BCDp'_1 + ABDp'_0) \\ &= \frac{1}{D(C-A)} (BCD(p'_0 + p'_1 + p'_3 - 1) + CDEp'_2 - ADEp'_2) \\ &\stackrel{(b)}{\leq} \frac{BC(-p'_2 - p'_4) + CEp'_2 - AEp'_2}{C-A} \\ &= \frac{-BCp'_4 - BCp'_2 + CEp'_2 - AEp'_2}{C-A} \\ &\leq \frac{-BCp'_2 + CEp'_2 - AEp'_2}{C-A} \\ &= \frac{CEp'_2 \left(1 - \frac{A}{C} - \frac{B}{E} \right)}{C-A} \\ &\stackrel{(c)}{\leq} 0, \end{aligned}$$

where

(a) follows from (A.5) and (A.6),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right).$$

■

Proof of Proposition 16: Suppose

$$A > B. \quad (\text{A.7})$$

We will show that $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$, $(\frac{AD}{A+D}, \frac{BD}{A+D})$ and $(\frac{CD}{C+D}, 0)$ are all the vertices of $\mathcal{T}_{5,A>B}^*$. The proposition then follows.

Let

$$L_1 : R_1 = 0,$$

$$L_2 : R_2 = 0,$$

$$L_3 : R_2 = \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-BDF}{AD(E+F)} \right),$$

$$L_4 : R_2 = E - R_1 \left(\frac{AE+DE-BD}{AD} \right),$$

and

$$L_5 : R_2 = \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right)$$

be the five lines defining the boundary of $\mathcal{T}_{5,A>B}^*$ (cf. (2.13)). Let

$\alpha_x = \frac{ADEF}{AEF+DEF-ADE-BDF}$, $\beta_x = \frac{ADE}{AE+DE-BD}$ and $\gamma_x = \frac{CD}{C+D}$ be the x -

intercepts of L_3 , L_4 and L_5 respectively and let $\alpha_y = \frac{EF}{E+F}$, $\beta_y = E$

and $\gamma_y = \frac{BC}{C-A}$ be the y -intercepts of L_3 , L_4 and L_5 respectively.

Since

$$\begin{aligned} & \alpha_x - \beta_x \\ &= \frac{ADEF}{AEF+DEF-ADE-BDF} - \frac{ADE}{AE+DE-BD} \\ &= \frac{A^2 D^2 E^2}{(AE+D(E-B))(AE(F-D)+DF(E-B))} \\ &> 0 \end{aligned}$$

(cf. (2.1) and (2.8)) and

$$\begin{aligned}
 \beta_x - \gamma_x &= \frac{ADE}{AE + DE - BD} - \frac{CD}{C + D} \\
 &= \frac{D(ADE - CDE + BCD)}{(C + D)(AE + DE - BD)} \\
 &= \frac{CD^2E(A/C + B/E - 1)}{(C + D)(AE + D(E - B))} \\
 &> 0
 \end{aligned}$$

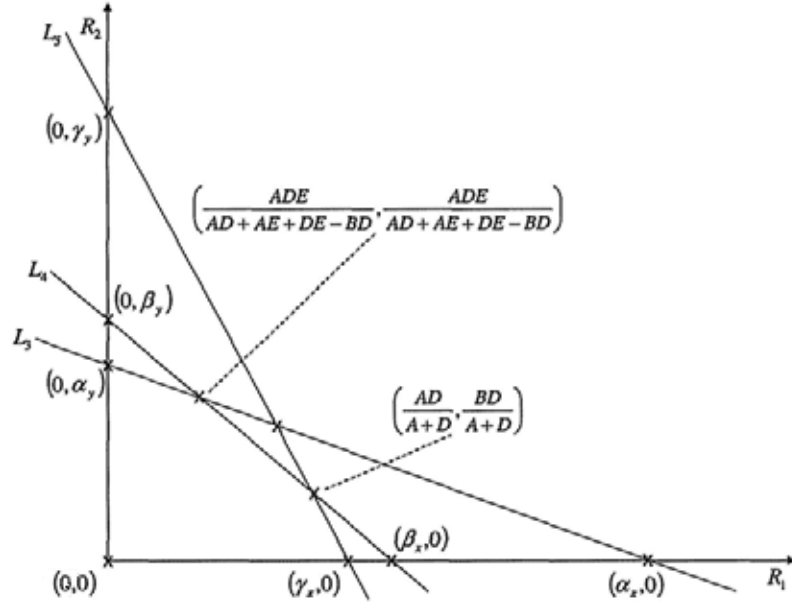
(cf. (2.8) and (2.10)), it follows that $\gamma_x < \beta_x < \alpha_x$. On the other hand, since

$$\begin{aligned}
 \gamma_y - \beta_y &= \frac{BC}{C - A} - E \\
 &= \frac{BC - CE + AE}{C - A} \\
 &= \frac{CE(A/C + B/E - 1)}{C - A} \\
 &> 0
 \end{aligned}$$

(cf. (2.7) and (2.10)) and

$$\begin{aligned}
 \beta_y - \alpha_y &= E - \frac{EF}{E + F} \\
 &= \frac{E^2}{E + F} \\
 &> 0,
 \end{aligned}$$

it follows that $\alpha_y < \beta_y < \gamma_y$. The intersection point between L_3 and L_4 is $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ and the intersection point


 Figure A.3: L_3 , L_4 and L_5 .

between L_4 and L_5 is $(\frac{AD}{A+D}, \frac{BD}{A+D})$. Since

$$\begin{aligned} & \frac{AD}{A+D} - \frac{ADE}{AD+AE+DE-BD} \\ &= \frac{AD^2(A-B)}{(A+D)(AD+AE+DE-BC)} \\ &> 0 \end{aligned}$$

(cf. (A.7)), it follows that $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$ is on the left side of $(\frac{AD}{A+D}, \frac{BD}{A+D})$. By inspecting Figure A.3, we observe that the vertices of $\mathcal{T}_{5,A>B}^*$ are $(0,0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADE}{AD+AE+DE-BD}, \frac{ADE}{AD+AE+DE-BD})$, $(\frac{AD}{A+D}, \frac{BD}{A+D})$ and $(\frac{CD}{C+D}, 0)$. ■

Proof of Lemma 18: Since \mathcal{S}_5 is the projection of a 7-dimensional polyhedron on a 2-dimensional space, it can be shown by Fourier-Motzkin elimination [26] that $\mathcal{S}_5 = \mathcal{T}_{5,A<B}^*$. Since Fourier-Motzkin elimination is very tedious, we present a simpler alternative proof as

follows. Suppose (R_1, R_2) is in \mathcal{S}_5 . Then, (R_1, R_2) is a non-negative pair such that

$$R_1 \leq \min\{p_0A + p_1C, p_3D\}$$

and

$$R_2 \leq \min\{p_0B + p_2E, p_3D + p_4F\}$$

for some allocation $[p_0 p_1 p_2 p_3 p_4]$. By Lemma 9, there exists an allocation $[p'_0 p'_1 p'_2 p'_3 p'_4]$ such that

$$R_1 = p'_0A + p'_1C = p'_3D \quad (\text{A.8})$$

and

$$R_2 \leq \min\{p'_0A + p'_2E, p'_3D + p'_4F\}. \quad (\text{A.9})$$

We will show that (R_1, R_2) satisfies the five inequalities defining $\mathcal{T}_{5,A<B}^*$ (cf. (2.15)), which will imply that $(R_1, R_2) \in \mathcal{T}_{5,A<B}^*$. The lemma then follows.

Clearly, (R_1, R_2) satisfies the first two inequalities of $\mathcal{T}_{5,A<B}^*$.

Consider the following chain of inequalities:

$$\begin{aligned}
& R_2 - \frac{EF}{E+F} + R_1 \left(\frac{AEF + DEF - ADE - BDF}{AD(E+F)} \right) \\
&= \frac{1}{AD(E+F)} (ADER_2 + ADFR_2 - ADEF + AEFR_1 \\
&\quad + DEFR_1 - ADE R_1 - BDFR_1) \\
&\stackrel{(a)}{\leq} \frac{1}{AD(E+F)} (AD^2 E p'_3 + ADEF p'_4 + ADEF p'_2 + ABDF p'_0 \\
&\quad - ADEF + ADEF p'_3 + ADEF p'_0 + CDEF p'_1 - AD^2 E p'_3 \\
&\quad - BCDF p'_1 - ABDF p'_0) \\
&= \frac{1}{AD(E+F)} (ADEF(p'_0 + p'_2 + p'_3 + p'_4 - 1) + CDEF p'_1 \\
&\quad - BCDF p'_1) \\
&\stackrel{(b)}{\leq} \frac{DF}{AD(E+F)} (-AE p'_1 + CE p'_1 - BC p'_1) \\
&= \frac{p'_1 CEF}{A(E+F)} \left(1 - \frac{A}{C} - \frac{B}{E} \right) \\
&\stackrel{(c)}{\leq} 0,
\end{aligned}$$

where

(a) follows from (A.8) and (A.9),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{EF}{E+F} - R_1 \left(\frac{AEF + DEF - ADE - ADF}{AD(E+F)} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned}
R_2 &= \frac{BCF}{BC + CF - AF} + R_1 \left(\frac{BCF + BDF - BCD}{D(BC + CF - AF)} \right) \\
&= \frac{1}{D(BC + CF - AF)} (BCDR_2 + DF(C - A)R_2 - BCDF \\
&\quad + BCFR_1 + BDFR_1 - BCDR_1) \\
&\stackrel{(a)}{\leq} \frac{1}{D(BC + CF - AF)} (BCD^2p'_3 + BCDFp'_4 \\
&\quad + DF(C - A)(p'_0B + p'_2E) - BCDF + BCDFp'_3 + ABDFp'_0 \\
&\quad + BCDFp'_1 - BCD^2p'_3) \\
&= \frac{1}{D(BC + CF - AF)} (BCD^2p'_3 + BCDFp'_4 + BCDFp'_0 \\
&\quad - ABDFp'_0 + CDEFp'_2 - ADEFp'_2 - BCDF + BCDFp'_3 \\
&\quad + ABDFp'_0 + BCDFp'_1 - BCD^2p'_3) \\
&= \frac{BCF(p'_0 + p'_1 + p'_3 + p'_4 - 1) + CEFp'_2 - AEFp'_2}{BC + CF - AF} \\
&\stackrel{(b)}{\leq} \frac{-BCFp'_2 + CEFp'_2 - AEFp'_2}{BC + CF - AF} \\
&= \frac{CEFp'_2 \left(1 - \frac{A}{C} - \frac{B}{E}\right)}{BC + CF - AF} \\
&\stackrel{(c)}{\leq} 0,
\end{aligned}$$

where

(a) follows from (A.8) and (A.9),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{BCF}{BC + CF - AF} - R_1 \left(\frac{BCF + BDF - BCD}{D(BC + CF - AF)} \right).$$

Consider the following chain of inequalities:

$$\begin{aligned}
& R_2 - \frac{BC}{C-A} + R_1 \left(\frac{BC+BD}{D(C-A)} \right) \\
&= \frac{D(C-A)R_2 - BCD + BCR_1 + BDR_1}{D(C-A)} \\
&\stackrel{(a)}{\leq} \frac{1}{D(C-A)} (D(C-A)(p'_0 B + p'_2 E) - BCD + BCDp'_3 \\
&\quad + BCDp'_1 + ABDp'_0) \\
&= \frac{1}{D(C-A)} (BCDp'_0 + CDEp'_2 - ABDp'_0 - ADEp'_2 \\
&\quad - BCD + BCDp'_3 + BCDp'_1 + ABDp'_0) \\
&= \frac{1}{D(C-A)} (BCD(p'_0 + p'_1 + p'_3 - 1) + CDEp'_2 - ADEp'_2) \\
&\stackrel{(b)}{\leq} \frac{BC(-p'_2 - p'_4) + CEp'_2 - AEp'_2}{C-A} \\
&= \frac{-BCp'_4 - BCp'_2 + CEp'_2 - AEp'_2}{C-A} \\
&\leq \frac{-BCp'_2 + CEp'_2 - AEp'_2}{C-A} \\
&= \frac{CEp'_2 \left(1 - \frac{A}{C} - \frac{E}{E}\right)}{C-A} \\
&\stackrel{(c)}{\leq} 0,
\end{aligned}$$

where

(a) follows from (A.8) and (A.9),

(b) follows from (2.9),

(c) follows from (2.10).

Therefore,

$$R_2 \leq \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right).$$

■

Proof of Proposition 20: Suppose

$$A < B. \quad (\text{A.10})$$

We will show that $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$, $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ and $(\frac{CD}{C+D}, 0)$ are all the vertices of $\mathcal{T}_{5,A<B}^*$. The proposition then follows.

Let

$$L_1 : R_1 = 0,$$

$$L_2 : R_2 = 0,$$

$$L_3 : R_2 = \frac{EF}{E+F} - R_1 \left(\frac{AEF+DEF-ADE-BDF}{AD(E+F)} \right),$$

$$L_4 : R_2 = \frac{BCF}{BC+CF-AF} - R_1 \left(\frac{BCF+BDF-BCD}{D(BC+CF-AF)} \right),$$

and

$$L_5 : R_2 = \frac{BC}{C-A} - R_1 \left(\frac{BC+BD}{D(C-A)} \right)$$

be the five lines defining the boundary of $\mathcal{T}_{5,A<B}^*$ (cf. (2.15)). Let $\alpha_x = \frac{ADEF}{AEF+DEF-ADE-BDF}$, $\beta_x = \frac{CDF}{CF+DF-CD}$ and $\gamma_x = \frac{CD}{C+D}$ be the x -intercepts of L_3 , L_4 and L_5 respectively and let $\alpha_y = \frac{EF}{E+F}$, $\beta_y = \frac{BCF}{BC+CF-AF}$ and $\gamma_y = \frac{BC}{C-A}$ be the y -intercepts of L_3 , L_4 and L_5 respectively. Since

$$\begin{aligned} & \alpha_x - \beta_x \\ &= \frac{ADEF}{AEF+DEF-ADE-BDF} - \frac{CDF}{CF+DF-CD} \\ &= \frac{DF(ADEF - CDEF + BCDF)}{(CF+DF-CD)(AEF+DEF-ADE-BDF)} \\ &= \frac{CD^2EF^2(A/C + B/E - 1)}{(DF+C(F-D))(AE(F-D) + DF(E-B))} \\ &> 0 \end{aligned}$$

(cf. (2.1), (2.8) and (2.10)) and

$$\begin{aligned}
 & \beta_x - \gamma_x \\
 &= \frac{CDF}{CF + DF - CD} - \frac{CD}{C + D} \\
 &= \frac{C^2D^2}{(C + D)(DF + C(F - D))} \\
 &> 0
 \end{aligned}$$

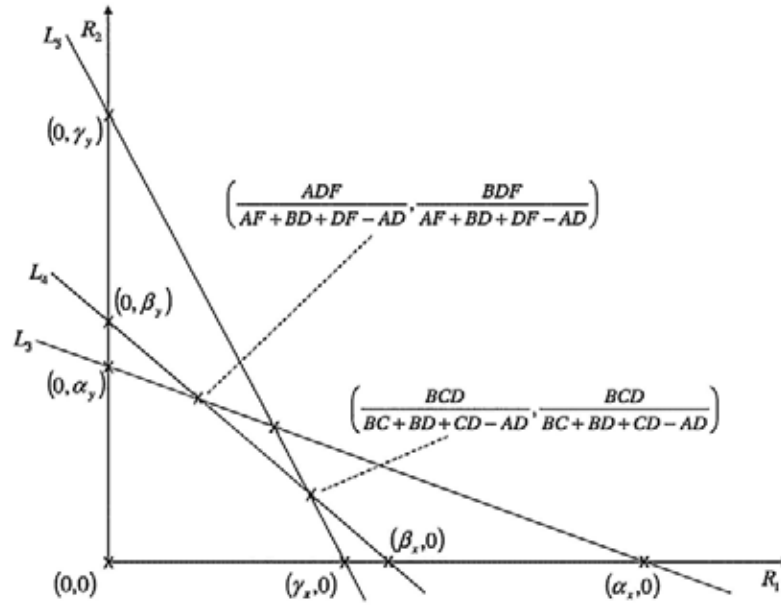
(cf. (2.1)), it follows that $\gamma_x < \beta_x < \alpha_x$. On the other hand, since

$$\begin{aligned}
 & \gamma_y - \beta_y \\
 &= \frac{BC}{C - A} - \frac{BCF}{BC + CF - AF} \\
 &= \frac{B^2C^2}{(C - A)(BC + F(C - A))} \\
 &> 0
 \end{aligned}$$

(cf. (2.7)) and

$$\begin{aligned}
 & \beta_y - \alpha_y \\
 &= \frac{BCF}{BC + CF - AF} - \frac{EF}{E + F} \\
 &= \frac{F(BCF - CEF + AEF)}{(E + F)(BC + CF - AF)} \\
 &= \frac{CEF^2(A/C + B/E - 1)}{(E + F)(BC + F(C - A))} \\
 &> 0
 \end{aligned}$$

(cf. (2.7) and (2.10)), it follows that $\alpha_y < \beta_y < \gamma_y$. The intersection point between L_3 and L_4 is $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$ and the intersection point between L_4 and L_5 is $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$.

Figure A.4: L_3 , L_4 and L_5 .

Since

$$\begin{aligned}
 & \frac{BCD}{BC+BD+CD-AD} - \frac{ADF}{AF+BD+DF-AD} \\
 &= \frac{D^2(B^2C+BCF-ABC-ABF-ACF+A^2F)}{(AF+BD+DF-AD)(BC+BD+CD-AD)} \\
 &= \frac{D^2(BC(B-A)+CF(B-A)-AF(B-A))}{(AF+BD+DF-AD)(BC+BD+CD-AD)} \\
 &= \frac{D^2(BC(B-A)+F(B-A)(C-A))}{(BD+DF+A(F-D))(BC+BD+D(C-A))} \\
 &> 0
 \end{aligned}$$

(cf. (2.1), (2.7) and (A.10)), it follows that $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$ is on the left side of $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$. By inspecting Figure A.4, we observe that the vertices of $\mathcal{T}_{5,A < B}^*$ are $(0, 0)$, $(0, \frac{EF}{E+F})$, $(\frac{ADF}{AF+BD+DF-AD}, \frac{BDF}{AF+BD+DF-AD})$, $(\frac{BCD}{BC+BD+CD-AD}, \frac{BCD}{BC+BD+CD-AD})$ and $(\frac{CD}{C+D}, 0)$. ■

□ **End of chapter.**

Appendix B

Proofs for Chapter 4

Proof of Theorem 8: Suppose (R_1, R_2) is achievable. By Definition 14, there exists a sequence of (n, M_1, M_2) -codes with

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1 \quad (\text{B.1})$$

and

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2 \quad (\text{B.2})$$

such that

$$\lim_{n \rightarrow \infty} P_{e,1}^n = 0 \quad (\text{B.3})$$

and

$$\lim_{n \rightarrow \infty} P_{e,2}^n = 0. \quad (\text{B.4})$$

Fix n and the corresponding (n, M_1, M_2) -code. Then,

$$(W_1, W_2) \rightarrow (X_1^n, X_2^n) \rightarrow Y_r^n \rightarrow X_r^n \rightarrow Y_1^n \quad (\text{B.5})$$

and

$$(W_1, W_2) \rightarrow (X_1^n, X_2^n) \rightarrow Y_r^n \rightarrow X_r^n \rightarrow Y_2^n \quad (\text{B.6})$$

form two Markov Chains. In addition, if W_2 is fixed in the channel,

$$W_1 \rightarrow X_1^n \rightarrow Y_r^n \rightarrow X_r^n \rightarrow Y_2^n \rightarrow \hat{W}_1 \quad (\text{B.7})$$

forms a Markov Chain. Similarly, if W_1 is fixed in the channel,

$$W_2 \rightarrow X_2^n \rightarrow Y_r^n \rightarrow X_r^n \rightarrow Y_1^n \rightarrow \hat{W}_2 \quad (\text{B.8})$$

forms a Markov Chain. Since W_1 and W_2 are independent, we have

$$\begin{aligned} \log_2 M_1 &= H(W_1) \\ &= H(W_1|W_2) \\ &= H(W_1|W_2, \hat{W}_1) + I(W_1; \hat{W}_1|W_2) \\ &\leq H(W_1|\hat{W}_1) + I(W_1; \hat{W}_1|W_2) \\ &= H(W_1|\hat{W}_1) + \sum_{w_2 \in \mathcal{W}_2} p(w_2) I(W_1; \hat{W}_1|W_2 = w_2) \\ &\leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{w_2 \in \mathcal{W}_2} p(w_2) I(W_1; \hat{W}_1|W_2 = w_2), \quad (\text{B.9}) \end{aligned}$$

where the last inequality follows from the Fano's inequality. Applying the data processing inequality on the Markov Chain in (B.7), we obtain

$$\begin{aligned} &\sum_{w_2 \in \mathcal{W}_2} p(w_2) I(W_1; \hat{W}_1|W_2 = w_2) \\ &\leq \sum_{w_2 \in \mathcal{W}_2} p(w_2) I(X_1^n; Y_r^n|W_2 = w_2) \\ &= I(X_1^n; Y_r^n|W_2) \quad (\text{B.10}) \end{aligned}$$

and

$$\begin{aligned}
& \sum_{w_2 \in \mathcal{W}_2} p(w_2) I(W_1; \hat{W}_1 | W_2 = w_2) \\
& \leq \sum_{w_2 \in \mathcal{W}_2} p(w_2) I(X_r^n; Y_2^n | W_2 = w_2) \\
& = I(X_r^n; Y_2^n | W_2). \tag{B.11}
\end{aligned}$$

Consider the following chain of inequalities:

$$\begin{aligned}
& I(X_1^n; Y_r^n | W_2) \\
& = H(Y_r^n | W_2) - H(Y_r^n | W_2, X_1^n) \\
& \stackrel{(a)}{=} H(Y_r^n | W_2, X_2^n) - H(Y_r^n | W_2, X_1^n, X_2^n) \\
& \leq H(Y_r^n | X_2^n) - H(Y_r^n | W_1, W_2, X_1^n, X_2^n) \\
& \stackrel{(b)}{=} H(Y_r^n | X_2^n) - H(Y_r^n | X_1^n, X_2^n) \\
& \stackrel{(c)}{=} H(Y_r^n | X_2^n) - \sum_{k=1}^n H(Y_{r,k} | X_{1,k}, X_{2,k}) \\
& = \sum_{k=1}^n H(Y_{r,k} | X_2^n, Y_{r,1}, Y_{r,2}, \dots, Y_{r,k-1}) - H(Y_{r,k} | X_{1,k}, X_{2,k}) \\
& \leq \sum_{k=1}^n H(Y_{r,k} | X_{2,k}) - H(Y_{r,k} | X_{1,k}, X_{2,k}) \\
& = \sum_{k=1}^n I(X_{1,k}; Y_{r,k} | X_{2,k}) \tag{B.12}
\end{aligned}$$

where

(a) follows from the fact that X_2^n is a function of W_2 ,

(b) follows from the Markov Chain in (B.6),

(c) follows from (4.1) in Definition 11.

Using (B.9), (B.10) and (B.12), we obtain

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{k=1}^n I(X_{1,k}; Y_{r,k} | X_{2,k}).$$

Dividing both sides by n and letting n go to infinity, it then follows from (B.3) that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}),$$

which implies from (B.1) that

$$R_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}), \quad (\text{B.13})$$

where each $I(X_{1,k}; Y_{r,k} | X_{2,k})$ is attained by some

$$\begin{aligned} & p_{X_{1,k}, X_{2,k}, Y_{r,k}}(x_{1,k}, x_{2,k}, y_{r,k}) \\ &= \sum_{x_{1,i}, x_{2,i}, y_{r,i}, i \neq k} p_{X_1^n, X_2^n, Y_r^n}(x_1^n, x_2^n, y_r^n) \\ &= \sum_{x_{1,i}, x_{2,i}, y_{r,i}, i \neq k} p_{X_1^n}(x_1^n) p_{X_2^n}(x_2^n) p^{(n)}(y_r^n | x_1^n, x_2^n) \\ &= \sum_{x_{1,i}, x_{2,i}, i \neq k} p_{X_1^n}(x_1^n) p_{X_2^n}(x_2^n) \sum_{y_{r,i}, i \neq k} \prod_{j=1}^n p_1(y_{r,j} | x_{1,j}, x_{2,j}) \\ &= p_{X_{1,k}}(x_{1,k}) p_{X_{2,k}}(x_{2,k}) p_1(y_{r,k} | x_{1,k}, x_{2,k}). \end{aligned}$$

On the other hand,

$$\begin{aligned} & I(X_r^n; Y_2^n | W_2) \\ &= H(Y_2^n | W_2) - H(Y_2^n | W_2, X_r^n) \\ &\leq H(Y_2^n) - H(Y_2^n | W_1, W_2, X_1^n, X_2^n, Y_r^n, X_r^n) \\ &= H(Y_2^n) - H(Y_2^n | X_r^n) \end{aligned} \quad (\text{B.14})$$

where the last equality follows from the Markov Chain in (B.6). Using (4.2) in Definition 11, we obtain

$$\begin{aligned}
& H(Y_2^n) - H(Y_2^n | X_r^n) \\
&= H(Y_2^n) - \sum_{k=1}^n H(Y_{2,k} | X_{r,k}) \\
&= \sum_{k=1}^n H(Y_{2,k} | Y_{2,1}, Y_{2,2}, \dots, Y_{2,k-1}) - H(Y_{2,k} | X_{r,k}) \\
&\leq \sum_{k=1}^n H(Y_{2,k}) - H(Y_{2,k} | X_{r,k}) \\
&= \sum_{k=1}^n I(X_{r,k}; Y_{2,k}). \tag{B.15}
\end{aligned}$$

Using (B.9), (B.11), (B.14) and (B.15), we obtain

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{k=1}^n I(X_{r,k}; Y_{2,k}).$$

Dividing both sides by n and letting n go to infinity, it then follows from (B.3) that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{2,k}),$$

which implies from (B.1) that

$$R_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{2,k}), \tag{B.16}$$

where each $I(X_{r,k}; Y_{2,k})$ is attained by some

$$\begin{aligned}
& p_{X_{r,k}, Y_{1,k}, Y_{2,k}}(x_{r,k}, y_{1,k}, y_{2,k}) \\
&= \sum_{x_{r,i}, y_{1,i}, y_{2,i}, i \neq k} p_{X_r^n, Y_1^n, Y_2^n}(x_r^n, y_1^n, y_2^n) \\
&= \sum_{x_{r,i}, y_{1,i}, y_{2,i}, i \neq k} p_{X_r^n}(x_r^n) p^{(n)}(y_1^n, y_2^n | x_r^n) \\
&= \sum_{x_{r,i}, i \neq k} p_{X_r^n}(x_r^n) \sum_{y_{1,i}, y_{2,i}, i \neq k} \prod_{j=1}^n p_2(y_{1,j}, y_{2,j} | x_{r,j}) \\
&= p_{X_{r,k}}(x_{r,k}) p_2(y_{1,k}, y_{2,k} | x_{r,k}).
\end{aligned}$$

Following similar procedures for proving (B.13) and (B.16), we can obtain from (B.2), (B.4), (B.5) and (B.8) that

$$R_2 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{2,k}; Y_{r,k} | X_{1,k}) \quad (\text{B.17})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{1,k}). \quad (\text{B.18})$$

Using (B.13) and (B.17), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^1)}$. Using (B.13) and (B.18), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^2)}$. Using (B.16) and (B.17), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^3)}$. Using (B.16) and (B.18), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^4)}$. Consequently, $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)} \cap \overline{\text{conv}(\mathcal{R}^3)} \cap \overline{\text{conv}(\mathcal{R}^4)}$. The theorem then follows from Definition 15 and the fact that $\overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)} \cap \overline{\text{conv}(\mathcal{R}^3)} \cap \overline{\text{conv}(\mathcal{R}^4)}$ is closed. ■

Proof of Lemma 24: Fix an (n, M_1, M_2) -code on the discrete memoryless TRC and consider any $1 \leq k \leq n$. Figure B.1 shows the dependency graph of the variables $W_1, W_2, X_{1,j}, X_{2,j}, X_{r,j}, Y_{1,j}, Y_{2,j}$ and $Y_{r,j}$ for $j = 1, 2, 3$. We can then extend the dependency

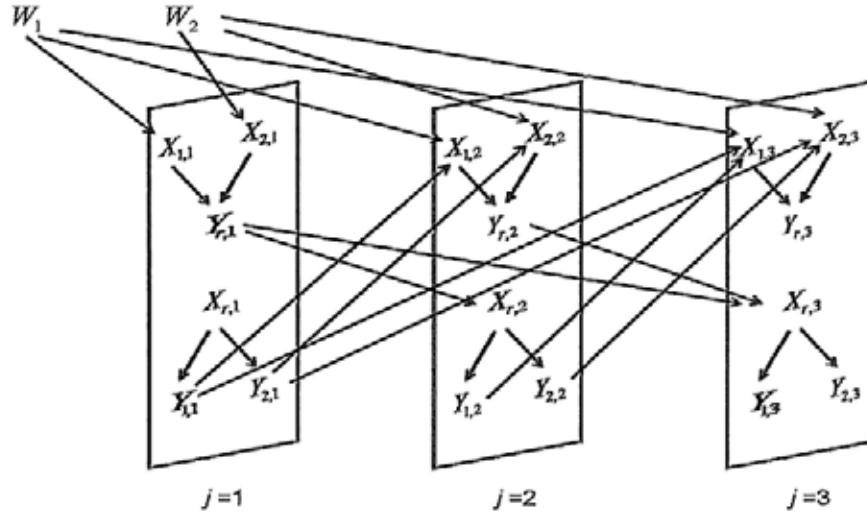


Figure B.1: Dependency graph of the discrete memoryless TRC in three time slots.

graph so that the extended graph includes all the variables by the k^{th} time slot and conclude from the extended graph that for any $w_1 \in \mathcal{W}_1$, $w_2 \in \mathcal{W}_2$, $x_1^k \in \mathcal{X}_1^k$, $x_2^k \in \mathcal{X}_2^k$, $x_r^k \in \mathcal{X}_r^k$, $y_1^k \in \mathcal{Y}_1^k$, $y_2^k \in \mathcal{Y}_2^k$ and $y_r^k \in \mathcal{Y}_r^k$ with $p(w_1, w_2, x_1^{k-1}, x_2^{k-1}, x_r^{k-1}, y_1^{k-1}, y_2^{k-1}, y_r^{k-1}) > 0$,

$$\begin{aligned}
 & p(w_1, w_2, x_1^k, x_2^k, x_r^k, y_1^k, y_2^k, y_r^k) \\
 &= p(w_1, w_2, x_1^{k-1}, x_2^{k-1}, x_r^{k-1}, y_1^{k-1}, y_2^{k-1}, y_r^{k-1}) \\
 & \quad p(x_{1,k}, x_{2,k} | w_1, w_2, x_1^{k-1}, x_2^{k-1}, x_r^{k-1}, y_1^{k-1}, y_2^{k-1}, y_r^{k-1}) \\
 & \quad p_1(y_{r,k} | x_{1,k}, x_{2,k}) \\
 & \quad p(x_{r,k} | w_1, w_2, x_1^{k-1}, x_2^{k-1}, x_r^{k-1}, y_1^{k-1}, y_2^{k-1}, y_r^{k-1}) \\
 & \quad p_2(y_{1,k}, y_{2,k} | x_{r,k}). \tag{B.19}
 \end{aligned}$$

Let \mathcal{U} denote $(W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1})$ to sim-

plify notation. If $p(u) > 0$, then

$$\begin{aligned}
& p(u, x_{1,k}, x_{2,k}, y_{r,k}) \\
&= \sum_{x_{r,k}, y_{1,k}, y_{2,k}} p(u, x_{1,k}, x_{2,k}, x_{r,k}, y_{1,k}, y_{2,k}, y_{r,k}) \\
&\stackrel{(a)}{=} \sum_{x_{r,k}, y_{1,k}, y_{2,k}} p(u) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}) \\
&\quad p(x_{r,k} | u) p_2(y_{1,k}, y_{2,k} | x_{r,k}) \\
&= p(u) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}) \\
&\quad \sum_{x_{r,k}, y_{1,k}, y_{2,k}} p(x_{r,k} | u) p_2(y_{1,k}, y_{2,k} | x_{r,k}) \\
&= p(u) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}) \\
&= p(u, x_{1,k}, x_{2,k}) p_1(y_{r,k} | x_{1,k}, x_{2,k}) \tag{B.20}
\end{aligned}$$

where (a) follows from (B.19). If $p(u) = 0$, then

$$p(u, x_{1,k}, x_{2,k}, y_{r,k}) = p(u, x_{1,k}, x_{2,k}) p_1(y_{r,k} | x_{1,k}, x_{2,k}) = 0. \tag{B.21}$$

It then follows from (B.20) and (B.21) that for all $u \in \mathcal{U}$, $x_{1,k} \in \mathcal{X}_1$, $x_{2,k} \in \mathcal{X}_2$ and $y_{r,k} \in \mathcal{Y}_r$,

$$p(u, x_{1,k}, x_{2,k}, y_{r,k}) = p(u, x_{1,k}, x_{2,k}) p_1(y_{r,k} | x_{1,k}, x_{2,k}), \tag{B.22}$$

which implies from Proposition 23 that (4.8) is a Markov Chain. Similarly, it can be shown that

$$p(u, x_{r,k}, y_{1,k}, y_{2,k}) = p(u, x_{r,k}) p_2(y_{1,k}, y_{2,k} | x_{r,k}), \tag{B.23}$$

which then implies from Proposition 23 that (4.9) is a Markov Chain.

In addition, if $p(u) > 0$, then

$$\begin{aligned}
& p(u, x_{1,k}, x_{2,k}, y_{1,k}, y_{2,k}, y_{r,k}) \\
&= \sum_{x_{r,k}} p(u, x_{1,k}, x_{2,k}, x_{r,k}, y_{1,k}, y_{2,k}, y_{r,k}) \\
&\stackrel{(a)}{=} \sum_{x_{r,k}} p(u) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}) p(x_{r,k} | u) p_2(y_{1,k}, y_{2,k} | x_{r,k}) \\
&\stackrel{(b)}{=} \sum_{x_{r,k}} p(u, x_{r,k}, y_{1,k}, y_{2,k}) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}) \\
&= p(u, y_{1,k}, y_{2,k}) p(x_{1,k}, x_{2,k} | u) p_1(y_{r,k} | x_{1,k}, x_{2,k}), \tag{B.24}
\end{aligned}$$

where (a) follows from (B.19) and (b) follows from (B.23). If $p(u) = 0$,

$$p(u, x_{1,k}, x_{2,k}, y_{1,k}, y_{2,k}, y_{r,k}) = 0. \tag{B.25}$$

Let χ and φ be two functions that

$$\chi(u, x_{1,k}, x_{2,k}, y_{1,k}, y_{2,k}) = \begin{cases} p(u, y_{1,k}, y_{2,k}) p(x_{1,k}, x_{2,k} | u) & \text{if } p(u) > 0, \\ 0 & \text{if } p(u) = 0 \end{cases}$$

and

$$\varphi(x_{1,k}, x_{2,k}, y_{r,k}) = p_1(y_{r,k} | x_{1,k}, x_{2,k}).$$

It then follows from (B.24) and (B.25) that for all $u \in \mathcal{U}$, $x_{1,k} \in \mathcal{X}_1$, $x_{2,k} \in \mathcal{X}_2$, $y_{1,k} \in \mathcal{Y}_1$, $y_{2,k} \in \mathcal{Y}_2$ and $y_{r,k} \in \mathcal{Y}_r$,

$$\begin{aligned}
& p(u, x_{1,k}, x_{2,k}, y_{1,k}, y_{2,k}, y_{r,k}) = \\
& \chi(u, x_{1,k}, x_{2,k}, y_{1,k}, y_{2,k}) \varphi(x_{1,k}, x_{2,k}, y_{r,k}),
\end{aligned}$$

which then implies from Proposition 23 that (4.10) is a Markov Chain. Similarly, it can be shown from (B.19) and (B.22) that (4.11) is a Markov Chain.

The second statement of the lemma follows from summing u on both sides of (B.22) and (B.23). ■

Proof of Theorem 9: Suppose (R_1, R_2) is in \mathcal{R}' . By Definition 19 and Definition 20, there exists a sequence of (n, M_1, M_2) -codes with

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1 \quad (\text{B.26})$$

and

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2 \quad (\text{B.27})$$

such that

$$\lim_{n \rightarrow \infty} P_{e,1}^n = 0 \quad (\text{B.28})$$

and

$$\lim_{n \rightarrow \infty} P_{e,2}^n = 0. \quad (\text{B.29})$$

Fix n and the corresponding (n, M_1, M_2) -code. It then follows from Lemma 24 that

$$\begin{aligned} (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\ \rightarrow (X_{1,k}, X_{2,k}) \rightarrow Y_{r,k}, \end{aligned} \quad (\text{B.30})$$

$$\begin{aligned} (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\ \rightarrow X_{r,k} \rightarrow Y_{2,k} \end{aligned} \quad (\text{B.31})$$

and

$$\begin{aligned} (W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^k) \\ \rightarrow X_{r,k} \rightarrow Y_{2,k} \end{aligned} \quad (\text{B.32})$$

form three Markov Chains. Since W_1 and W_2 are independent, we

have

$$\begin{aligned}\log_2 M_1 &= H(W_1) \\ &= H(W_1|W_2) \\ &= I(W_1; Y_2^n, Y_r^n | W_2) + H(W_1 | Y_2^n, Y_r^n, W_2) \\ &\leq I(W_1; Y_2^n, Y_r^n | W_2) + H(W_1 | Y_2^n, W_2) \\ &\leq I(W_1; Y_2^n, Y_r^n | W_2) + 1 + P_{e,1}^n \log_2 M_1, \quad (\text{B.33})\end{aligned}$$

where the last inequality follows from the Fano's inequality.

Consider the following chain of inequalities:

$$\begin{aligned}
& I(W_1; Y_2^n, Y_r^n | W_2) \\
&= \sum_{k=1}^n I(W_1; Y_{2,k}, Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) \\
&= \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) + I(W_1; Y_{2,k} | W_2, Y_2^{k-1}, Y_r^k) \\
&= \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) + H(Y_{2,k} | W_2, Y_2^{k-1}, Y_r^k) \\
&\quad - H(Y_{2,k} | W_1, W_2, Y_2^{k-1}, Y_r^k) \\
&\stackrel{(a)}{=} \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) + H(Y_{2,k} | W_2, X_{r,k}, Y_2^{k-1}, Y_r^k) \\
&\quad - H(Y_{2,k} | W_1, W_2, Y_2^{k-1}, Y_r^k) \\
&\leq \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) + H(Y_{2,k} | X_{r,k}) \\
&\quad - H(Y_{2,k} | W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^k, Y_1^{k-1}, Y_2^{k-1}, Y_r^k) \\
&\stackrel{(b)}{=} \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) + H(Y_{2,k} | X_{r,k}) - H(Y_{2,k} | X_{r,k}) \\
&= \sum_{k=1}^n I(W_1; Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) \\
&= \sum_{k=1}^n H(Y_{r,k} | W_2, Y_2^{k-1}, Y_r^{k-1}) - H(Y_{r,k} | W_1, W_2, Y_2^{k-1}, Y_r^{k-1}) \\
&\stackrel{(c)}{=} \sum_{k=1}^n H(Y_{r,k} | W_2, X_{2,k}, Y_2^{k-1}, Y_r^{k-1}) - H(Y_{r,k} | W_1, W_2, Y_2^{k-1}, Y_r^{k-1}) \\
&\leq \sum_{k=1}^n H(Y_{r,k} | X_{2,k}) - H(Y_{r,k} | W_1, W_2, X_1^k, X_2^k, X_r^{k-1}, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\
&\stackrel{(d)}{=} \sum_{k=1}^n H(Y_{r,k} | X_{2,k}) - H(Y_{r,k} | X_{1,k}, X_{2,k}) \\
&= \sum_{k=1}^n I(X_{1,k}; Y_{r,k} | X_{2,k}), \tag{B.34}
\end{aligned}$$

where

- (a) follows from the fact that $X_{r,k}$ is a function of Y_r^{k-1} ,
- (b) follows from the Markov Chain in (B.32),
- (c) follows from the fact that $X_{2,k}$ is a function of W_2 and Y_2^{k-1} ,
- (d) follows from the Markov Chain in (B.30).

Let Q be a new timesharing random variable distributed uniformly on $\{1, 2, \dots, n\}$ such that Q is independent of the collection of random variables $\{X_{1,k}, X_{2,k}, Y_{r,k}, X_{r,k}, Y_{1,k}, Y_{2,k} \mid k = 1, 2, \dots, n\}$. Then,

$$\begin{aligned}
 & p_{X_{1,Q}, X_{2,Q}, Y_{r,Q} | Q}(x_1, x_2, y_r | k) \\
 &= p_{X_{1,k}, X_{2,k}, Y_{r,k} | Q}(x_1, x_2, y_r | k) \\
 &\stackrel{(a)}{=} p_{X_{1,k}, X_{2,k}, Y_{r,k}}(x_1, x_2, y_r)
 \end{aligned} \tag{B.35}$$

and

$$\begin{aligned}
 & p_{Q, X_{1,Q}, X_{2,Q}, Y_{r,Q}}(k, x_1, x_2, y_r) \\
 &= p_Q(k) p_{X_{1,Q}, X_{2,Q}, Y_{r,Q} | Q}(x_1, x_2, y_r | k) \\
 &\stackrel{(b)}{=} p_Q(k) p_{X_{1,k}, X_{2,k}, Y_{r,k}}(x_1, x_2, y_r) \\
 &\stackrel{(c)}{=} p_Q(k) p_{X_{1,k}, X_{2,k}}(x_1, x_2) p_1(y_r | x_1, x_2)
 \end{aligned} \tag{B.36}$$

$$= \chi(k, x_1, x_2) \varphi(x_1, x_2, y_r) \tag{B.37}$$

for some functions χ and φ , where

- (a) follows from the construction of Q that Q is independent of $X_{1,k}$, $X_{2,k}$ and $Y_{r,k}$,
- (b) follows from (B.35),
- (c) follows from the second statement of Lemma 24.

Following from the above,

$$\begin{aligned}
& \frac{1}{n} \sum_{k=1}^n I(X_{1,k}; Y_{r,k} | X_{2,k}) \\
& \stackrel{(a)}{=} \sum_{k=1}^n \frac{1}{n} I(X_{1,Q}; Y_{r,Q} | X_{2,Q}, Q = k) \\
& = I(X_{1,Q}; Y_{r,Q} | X_{2,Q}, Q) \\
& = H(Y_{r,Q} | X_{2,Q}, Q) - H(Y_{r,Q} | X_{1,Q}, X_{2,Q}, Q) \\
& \leq H(Y_{r,Q} | X_{2,Q}) - H(Y_{r,Q} | X_{1,Q}, X_{2,Q}, Q) \\
& \stackrel{(b)}{=} H(Y_{r,Q} | X_{2,Q}) - H(Y_{r,Q} | X_{1,Q}, X_{2,Q}) \\
& = I(X_{1,Q}; Y_{r,Q} | X_{2,Q}) \tag{B.38}
\end{aligned}$$

where

(a) follows from (B.35),

(b) follows from the fact that $Q \rightarrow (X_{1,Q}, X_{2,Q}) \rightarrow Y_{r,Q}$ forms a Markov Chain (cf. (B.37) and Proposition 23).

Using (B.33), (B.34) and (B.38), we obtain

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + nI(X_{1,Q}; Y_{r,Q} | X_{2,Q}). \tag{B.39}$$

By symmetry, we obtain

$$\log_2 M_2 \leq 1 + P_{e,2}^n \log_2 M_2 + nI(X_{2,Q}; Y_{r,Q} | X_{1,Q}). \tag{B.40}$$

Since

$$\begin{aligned}
 & p_{X_{1,Q}, X_{2,Q}, Y_{r,Q}}(x_1, x_2, y_r) \\
 &= \sum_{k=1}^n p_{Q, X_{1,Q}, X_{2,Q}, Y_{r,Q}}(k, x_1, x_2, y_r) \\
 &\stackrel{(a)}{=} \sum_{k=1}^n p_Q(k) p_{X_{1,k}, X_{2,k}}(x_1, x_2) p_1(y_r | x_1, x_2) \\
 &= \sum_{k=1}^n \frac{1}{n} p_{X_{1,k}, X_{2,k}}(x_1, x_2) p_1(y_r | x_1, x_2)
 \end{aligned}$$

where (a) follows from (B.36), it then follows that $I(X_{1,Q}; Y_{r,Q} | X_{2,Q})$ and $I(X_{2,Q}; Y_{r,Q} | X_{1,Q})$ are achieved by the distribution $p_{X_{1,Q}, X_{2,Q}, Y_{r,Q}}(x_1, x_2, y_r) = q_n(x_1, x_2) p_1(y_r | x_1, x_2)$ where $q_n(x_1, x_2) = \sum_{k=1}^n \frac{1}{n} p_{X_{1,k}, X_{2,k}}(x_1, x_2)$ is the input distribution for the MAC. Dividing both sides of (B.39) and (B.40) by n and letting n go to infinity, it then follows from (B.28) and (B.29) that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 \leq \liminf_{n \rightarrow \infty} I(X_{1,Q}; Y_{r,Q} | X_{2,Q})$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_2 \leq \liminf_{n \rightarrow \infty} I(X_{2,Q}; Y_{r,Q} | X_{1,Q}),$$

which imply from (B.26) and (B.27) that

$$R_1 \leq \liminf_{n \rightarrow \infty} I(X_{1,Q}; Y_{r,Q} | X_{2,Q}) \quad (\text{B.41})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} I(X_{2,Q}; Y_{r,Q} | X_{1,Q}). \quad (\text{B.42})$$

Consider each distribution of (X_1, X_2) as a point in an $|\mathcal{X}_1| |\mathcal{X}_2|$ -dimensional Euclidean space. Let $\{q_{n_k}(x_1, x_2)\}_{k=1,2,\dots}$ be a convergent subsequence of $\{q_n(x_1, x_2)\}_{n=1,2,\dots}$ with respect to the \mathcal{L}_1 -

distance, where the \mathcal{L}_1 -distance between two distributions $u(x)$ and $v(x)$ on the same discrete alphabet \mathcal{X} is defined as $\sum_{x \in \mathcal{X}} |u(x) - v(x)|$. Since the set of all joint distributions $\{p_{X_1, X_2}(x_1, x_2)\}$ is closed with respect to the \mathcal{L}_1 -distance, there exists a joint distribution $\bar{q}(x_1, x_2)$ such that $\lim_{k \rightarrow \infty} q_{n_k}(x_1, x_2) = \bar{q}(x_1, x_2)$. Since $I(X_1; Y_r | X_2)$ is a continuous functional of $p_{X_1, X_2}(x_1, x_2)$, it then follows from (B.41) and (B.42) that

$$R_1 \leq \liminf_{n \rightarrow \infty} I(X_{1,Q}; Y_{r,Q} | X_{2,Q}) \leq I(X_1; Y_r | X_2) \quad (\text{B.43})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} I(X_{2,Q}; Y_{r,Q} | X_{1,Q}) \leq I(X_2; Y_r | X_1) \quad (\text{B.44})$$

for the distribution $p_{X_1, X_2, Y_r}(x_1, x_2, y_r) = \bar{q}(x_1, x_2)p_1(y_r | x_1, x_2)$.

On the other hand,

$$\begin{aligned} & \log_2 M_1 \\ &= H(W_1) \\ &= H(W_1 | W_2) \\ &= I(W_1; Y_2^n | W_2) + H(W_1 | Y_2^n, W_2) \\ &\leq I(W_1; Y_2^n | W_2) + 1 + P_{e,1}^n \log_2 M_1 \end{aligned} \quad (\text{B.45})$$

where the last inequality follows from the Fano's inequality.

Consider the following chain of inequalities:

$$\begin{aligned}
& I(W_1; Y_2^n | W_2) \\
&= \sum_{k=1}^n I(W_1; Y_{2,k} | W_2, Y_2^{k-1}) \\
&= \sum_{k=1}^n H(Y_{2,k} | W_2, Y_2^{k-1}) - H(Y_{2,k} | W_1, W_2, Y_2^{k-1}) \\
&\leq \sum_{k=1}^n H(Y_{2,k}) - H(Y_{2,k} | W_1, W_2, X_1^{k-1}, X_2^{k-1}, X_r^k, Y_1^{k-1}, Y_2^{k-1}, Y_r^{k-1}) \\
&\stackrel{(a)}{=} \sum_{k=1}^n H(Y_{2,k}) - H(Y_{2,k} | X_{r,k}) \\
&= \sum_{k=1}^n I(X_{r,k}; Y_{2,k}) \tag{B.46}
\end{aligned}$$

where (a) follows from the Markov Chain in (B.31). In addition,

$$\begin{aligned}
& p_{X_{r,Q}, Y_{1,Q}, Y_{2,Q} | Q}(x_r, y_1, y_2 | k) \\
&= p_{X_{r,k}, Y_{1,k}, Y_{2,k} | Q}(x_r, y_1, y_2 | k) \\
&\stackrel{(a)}{=} p_{X_{r,k}, Y_{1,k}, Y_{2,k}}(x_r, y_1, y_2) \tag{B.47}
\end{aligned}$$

and

$$\begin{aligned}
& p_{Q, X_{r,Q}, Y_{1,Q}, Y_{2,Q}}(k, x_r, y_1, y_2) \\
&= p_Q(k) p_{X_{r,Q}, Y_{1,Q}, Y_{2,Q} | Q}(x_r, y_1, y_2 | k) \\
&\stackrel{(b)}{=} p_Q(k) p_{X_{r,k}, Y_{1,k}, Y_{2,k}}(x_r, y_1, y_2) \\
&\stackrel{(c)}{=} p_Q(k) p_{X_{r,k}}(x_r) p_2(y_1, y_2 | x_r) \tag{B.48}
\end{aligned}$$

$$= \chi(k, x_r) \varphi(x_r, y_1, y_2) \tag{B.49}$$

for some functions χ and φ , where

(a) follows from the construction of Q that Q is independent of

$X_{r,k}$, $Y_{1,k}$ and $Y_{2,k}$,

(b) follows from (B.47),

(c) follows from the second statement of Lemma 24.

Following from the above,

$$\begin{aligned}
 \frac{1}{n} \sum_{k=1}^n I(X_{r,k}; Y_{2,k}) &\stackrel{(a)}{=} \sum_{k=1}^n \frac{1}{n} I(X_{r,Q}; Y_{2,Q} | Q = k) \\
 &= I(X_{r,Q}; Y_{2,Q} | Q) \\
 &= H(Y_{2,Q} | Q) - H(Y_{2,Q} | X_{r,Q}, Q) \\
 &\leq H(Y_{2,Q}) - H(Y_{2,Q} | X_{r,Q}, Q) \\
 &\stackrel{(b)}{=} H(Y_{2,Q}) - H(Y_{2,Q} | X_{r,Q}) \\
 &= I(X_{r,Q}; Y_{2,Q}) \tag{B.50}
 \end{aligned}$$

where

(a) follows from (B.47),

(b) follows from the fact that $Q \rightarrow X_{r,Q} \rightarrow Y_{2,Q}$ forms a Markov Chain (cf. (B.49) and Proposition 23).

Using (B.45), (B.46) and (B.50), we obtain

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + nI(X_{r,Q}; Y_{2,Q}). \tag{B.51}$$

By symmetry, we obtain

$$\log_2 M_2 \leq 1 + P_{e,2}^n \log_2 M_2 + nI(X_{r,Q}; Y_{1,Q}). \tag{B.52}$$

Since

$$\begin{aligned}
& p_{X_r, Q, Y_1, Q, Y_2, Q}(x_r, y_1, y_2) \\
&= \sum_{k=1}^n p_{Q, X_r, Q, Y_1, Q, Y_2, Q}(k, x_r, y_1, y_2) \\
&\stackrel{(a)}{=} \sum_{k=1}^n p_Q(k) p_{X_r, k}(x_r) p_2(y_1, y_2 | x_r) \\
&= \sum_{k=1}^n \frac{1}{n} p_{X_r, k}(x_r) p_2(y_1, y_2 | x_r)
\end{aligned}$$

where (a) follows from (B.48), it then follows that $I(X_{r, Q}; Y_{1, Q})$ and $I(X_{r, Q}; Y_{2, Q})$ are achieved by the distribution $p_{X_r, Q, Y_1, Q, Y_2, Q}(x_r, y_1, y_2) = s_n(x_r) p_2(y_1, y_2 | x_r)$ where $s_n(x_r) = \sum_{k=1}^n \frac{1}{n} p_{X_r, k}(x_r)$ is the input distribution for the BC. Dividing both sides of (B.51) and (B.52) by n and letting n go to infinity, it then follows from (B.28) and (B.29) that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 \leq \liminf_{n \rightarrow \infty} I(X_{r, Q}; Y_{2, Q})$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_2 \leq \liminf_{n \rightarrow \infty} I(X_{r, Q}; Y_{1, Q})$$

which imply from (B.26) and (B.27) that

$$R_1 \leq \liminf_{n \rightarrow \infty} I(X_{r, Q}; Y_{2, Q}) \quad (\text{B.53})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} I(X_{r, Q}; Y_{1, Q}). \quad (\text{B.54})$$

Consider each distribution of X_r as a point in an $|X_r|$ -dimensional Euclidean space. Let $\{s_{n_k}(x_r)\}_{k=1, 2, \dots}$ be a convergent subsequence of $\{s_n(x_r)\}_{n=1, 2, \dots}$ with respect to the \mathcal{L}_1 -distance. Since the set of

all distributions $\{p_{X_r}(x_r)\}$ is closed with respect to the \mathcal{L}_1 -distance, there exists a distribution $\bar{s}(x_r)$ such that $\lim_{k \rightarrow \infty} s_{n_k}(x_r) = \bar{s}(x_r)$. Since $I(X_r; Y_2)$ is a continuous functional of $p_{X_r}(x_r)$, it then follows from (B.53) and (B.54) that

$$R_1 \leq \liminf_{n \rightarrow \infty} I(X_{r,Q}; Y_{2,Q}) \leq I(X_r; Y_2) \quad (\text{B.55})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} I(X_{r,Q}; Y_{1,Q}) \leq I(X_r; Y_1) \quad (\text{B.56})$$

for the distribution $p_{X_r, Y_1, Y_2}(x_r, y_1, y_2) = \bar{s}(x_r)p_2(y_1, y_2|x_r)$.

Let

$$\begin{aligned} p_{X_1, X_2, X_r, Y_1, Y_2, Y_r}(x_1, x_2, x_r, y_1, y_2, y_r) \\ = \bar{q}(x_1, x_2)\bar{s}(x_r)p_1(y_r|x_1, x_2)p_2(y_1, y_2|x_r), \end{aligned}$$

i.e. (X_1, X_2, Y_r) and (X_r, Y_1, Y_2) are independent. Then we see that

$$p_{X_1, X_2, Y_r}(x_1, x_2, y_r) = \bar{q}(x_1, x_2)p_1(y_r|x_1, x_2)$$

and

$$p_{X_r, Y_1, Y_2}(x_r, y_1, y_2) = \bar{s}(x_r)p_2(y_1, y_2|x_r).$$

It then follows from (B.43), (B.44), (B.55) and (B.56) that for this choice of $p_{X_1, X_2, X_r, Y_1, Y_2, Y_r}(x_1, x_2, x_r, y_1, y_2, y_r)$,

$$R_1 \leq \min\{I(X_1; Y_r|X_2), I(X_r; Y_2)\}$$

and

$$R_2 \leq \min\{I(X_2; Y_r|X_1), I(X_r; Y_1)\}.$$

■

Proof of Theorem 10: Suppose (R_1, R_2) is in \mathcal{R}_G . By Definitions 24 and 25, there exists a sequence of $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -codes with

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_1}{n} \geq R_1 \quad (\text{B.57})$$

and

$$\lim_{n \rightarrow \infty} \frac{\log_2 M_2}{n} \geq R_2 \quad (\text{B.58})$$

such that

$$\lim_{n \rightarrow \infty} P_{e,1}^n = 0 \quad (\text{B.59})$$

and

$$\lim_{n \rightarrow \infty} P_{e,2}^n = 0. \quad (\text{B.60})$$

Fix n and the corresponding $(n, M_1, M_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_r)$ -code. Following similar procedures for proving (B.33), (B.34), (B.45) and (B.46), we can obtain

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{k=1}^n I(X_{1,k}; Y_{r,k} | X_{2,k}) \quad (\text{B.61})$$

$$\log_2 M_1 \leq 1 + P_{e,1}^n \log_2 M_1 + \sum_{k=1}^n I(X_{r,k}; Y_{2,k}), \quad (\text{B.62})$$

$$\log_2 M_2 \leq 1 + P_{e,2}^n \log_2 M_2 + \sum_{k=1}^n I(X_{2,k}; Y_{r,k} | X_{1,k}), \quad (\text{B.63})$$

and

$$\log_2 M_2 \leq 1 + P_{e,2}^n \log_2 M_2 + \sum_{k=1}^n I(X_{r,k}; Y_{1,k}). \quad (\text{B.64})$$

Dividing both sides of (B.61), (B.62), (B.63) and (B.64) by n and

letting n go to infinity, it then follows from (B.59) and (B.60) that

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 &\leq \liminf_{n \rightarrow \infty} I(X_{1,k}; Y_{r,k} | X_{2,k}), \\ \lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_1 &\leq \liminf_{n \rightarrow \infty} I(X_{r,k}; Y_{2,k}), \\ \lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_2 &\leq \liminf_{n \rightarrow \infty} I(X_{2,k}; Y_{r,k} | X_{1,k})\end{aligned}$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_2 \leq \liminf_{n \rightarrow \infty} I(X_{r,k}; Y_{1,k}),$$

which imply from (B.57) and (B.58) that

$$R_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}), \quad (\text{B.65})$$

$$R_1 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{2,k}). \quad (\text{B.66})$$

$$R_2 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{2,k}; Y_{r,k} | X_{1,k}) \quad (\text{B.67})$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{1,k}). \quad (\text{B.68})$$

Consider the following chain of inequalities:

$$\begin{aligned}I(X_{1,k}; Y_{r,k} | X_{2,k}) &= h(Y_{r,k} | X_{2,k}) - h(Y_{r,k} | X_{1,k}, X_{2,k}) \\ &= h(X_{1,k} + X_{2,k} + Z_{r,k} | X_{2,k}) - h(X_{1,k} + X_{2,k} + Z_{r,k} | X_{1,k}, X_{2,k}) \\ &\stackrel{(a)}{=} h(X_{1,k} + Z_{r,k}) - h(Z_{r,k}) \\ &= h(X_{1,k} + Z_{r,k}) - \frac{1}{2} \log_2 2\pi e N_r \\ &\stackrel{(b)}{\leq} \frac{1}{2} \log_2 (2\pi e (E[X_{1,k}^2] + N_r)) - \frac{1}{2} \log_2 (2\pi e N_r) \\ &= \frac{1}{2} \log_2 (1 + E[X_{1,k}^2]/N_r),\end{aligned}$$

where

- (a) follows from (4.18),
- (b) follows from the fact that $X_{1,k}$ and $Z_{r,k}$ are independent, and the differential entropy of a random variable is upper bounded by the differential entropy of a Gaussian random variable with the same second moment.

Therefore,

$$\begin{aligned}
 & \sum_{k=1}^n \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \\
 & \leq \frac{1}{2} \sum_{k=1}^n \frac{1}{n} \log_2(1 + E[X_{1,k}^2]/N_r) \\
 & \stackrel{(a)}{\leq} \frac{1}{2} \log_2 \left(1 + \sum_{k=1}^n \frac{E[X_{1,k}^2]}{nN_r} \right) \\
 & = \frac{1}{2} \log_2 \left(1 + \frac{E[\sum_{k=1}^n X_{1,k}^2]}{nN_r} \right) \\
 & \stackrel{(b)}{\leq} \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N_r} \right),
 \end{aligned}$$

where

- (a) follows from applying Jensen's inequality to the concave function $\log_2(1+x)$,
- (b) follows from (4.20) and the fact that $\log_2(1+x)$ is increasing in x .

Consequently,

$$\liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N_r} \right),$$

which then implies from (B.65) that

$$R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N_r} \right). \quad (\text{B.69})$$

By symmetry, we obtain from (B.67) that

$$R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_2}{N_r} \right). \quad (\text{B.70})$$

Consider the following chain of inequalities:

$$\begin{aligned} I(X_{r,k}; Y_{2,k}) &= h(Y_{2,k}) - h(Y_{2,k}|X_{r,k}) \\ &= h(X_{r,k} + Z_{2,k}) - h(X_{r,k} + Z_{2,k}|X_{r,k}) \\ &= h(X_{r,k} + Z_{2,k}) - h(Z_{2,k}|X_{r,k}) \\ &\stackrel{(a)}{=} h(X_{r,k} + Z_{2,k}) - h(Z_{2,k}) \\ &= h(X_{r,k} + Z_{2,k}) - \frac{1}{2} \log_2 2\pi e N_2 \\ &\stackrel{(b)}{\leq} \frac{1}{2} \log_2 (2\pi e (E[X_{r,k}^2] + N_2)) - \frac{1}{2} \log_2 (2\pi e N_2) \\ &= \frac{1}{2} \log_2 (1 + E[X_{r,k}^2]/N_2), \end{aligned}$$

where

(a) follows from (4.19),

(b) follows from the fact that $X_{r,k}$ and $Z_{2,k}$ are independent, and the differential entropy of a random variable is upper bounded by the differential entropy of a Gaussian random variable with the same second moment.

Therefore,

$$\begin{aligned}
 & \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{2,k}) \\
 & \leq \frac{1}{2} \sum_{k=1}^n \frac{1}{n} \log_2(1 + E[X_{r,k}^2]/N_2) \\
 & \stackrel{(a)}{\leq} \frac{1}{2} \log_2 \left(1 + \sum_{k=1}^n \frac{E[X_{r,k}^2]}{nN_2} \right) \\
 & = \frac{1}{2} \log_2 \left(1 + \frac{E[\sum_{k=1}^n X_{r,k}^2]}{nN_2} \right) \\
 & \stackrel{(b)}{\leq} \frac{1}{2} \log_2 \left(1 + \frac{P_r}{N_2} \right),
 \end{aligned}$$

where

- (a) follows from applying Jensen's inequality to the concave function $\log_2(1+x)$,
- (b) follows from (4.22) and the fact that $\log_2(1+x)$ is increasing in x .

Consequently,

$$\liminf_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} I(X_{r,k}; Y_{2,k}) \leq \frac{1}{2} \log_2 \left(1 + \frac{P_r}{N_2} \right),$$

which then implies from (B.66) that

$$R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_r}{N_2} \right). \tag{B.71}$$

By symmetry, we obtain from (B.68) that

$$R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_r}{N_1} \right). \tag{B.72}$$

The theorem then follows from (B.69), (B.70), (B.71) and (B.72). ■

Proof of Theorem 11: Suppose $(R_1, R_2) \in \mathcal{R}$. Following similar procedures for proving (B.13), (B.16) (B.17) and (B.18), we can obtain

$$\begin{aligned}
R_1 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{1}{n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{n_1}{n_1 n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \\
&\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{n\alpha}{n_1 n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{\alpha}{n_1} I(X_{1,k}; Y_{r,k} | X_{2,k}), \tag{B.73}
\end{aligned}$$

$$\begin{aligned}
R_1 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{1}{n} I(X_{r,k}; Y_{2,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{n_2}{n_2 n} I(X_{r,k}; Y_{2,k}) \\
&\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{n\beta}{n_2 n} I(X_{r,k}; Y_{2,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{\beta}{n_2} I(X_{r,k}; Y_{2,k}), \tag{B.74}
\end{aligned}$$

$$\begin{aligned}
R_2 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{1}{n} I(X_{2,k}; Y_{r,k} | X_{1,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{n_1}{n_1 n} I(X_{2,k}; Y_{r,k} | X_{1,k}) \\
&\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{n\alpha}{n_1 n} I(X_{2,k}; Y_{r,k} | X_{1,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{\alpha}{n_1} I(X_{2,k}; Y_{r,k} | X_{1,k}) \tag{B.75}
\end{aligned}$$

and

$$\begin{aligned}
R_2 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{1}{n} I(X_{r,k}; Y_{1,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{n_2}{n_2 n} I(X_{r,k}; Y_{1,k}) \\
&\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{n\beta}{n_2 n} I(X_{r,k}; Y_{1,k}) \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{\beta}{n_2} I(X_{r,k}; Y_{1,k}). \tag{B.76}
\end{aligned}$$

Using (B.73) and (B.75), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^1)}$. Using (B.73) and (B.76), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^2)}$. Using (B.74) and (B.75), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^3)}$. Using (B.74) and (B.76), we obtain $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^4)}$. Consequently, $(R_1, R_2) \in \overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)} \cap \overline{\text{conv}(\mathcal{R}^3)} \cap \overline{\text{conv}(\mathcal{R}^4)}$. The theorem then follows from Definition 28 and the fact that $\overline{\text{conv}(\mathcal{R}^1)} \cap \overline{\text{conv}(\mathcal{R}^2)} \cap \overline{\text{conv}(\mathcal{R}^3)} \cap \overline{\text{conv}(\mathcal{R}^4)}$ is closed. ■

Proof of Theorem 12: Suppose $(R_1, R_2) \in \mathcal{R}$. Following similar procedures for proving (B.43) and (B.44), we can obtain

$$R_1 \leq \liminf_{n \rightarrow \infty} \frac{n_1}{n} I(X_{1,Q}; Y_{r,Q} | X_{2,Q}) \leq \alpha I(X_1; Y_r | X_2) \tag{B.77}$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} \frac{n_1}{n} I(X_{2,Q}; Y_{r,Q} | X_{1,Q}) \leq \alpha I(X_2; Y_r | X_1) \tag{B.78}$$

for some distribution $p_{X_1, X_2, Y_r}(x_1, x_2, y_r) = \bar{q}(x_1, x_2) p_1(y_r | x_1, x_2)$. Following similar procedures for proving (B.53) and (B.54), we can obtain

$$R_1 \leq \liminf_{n \rightarrow \infty} \frac{n_2}{n} I(X_{r,Q}; Y_{2,Q}) \leq \beta I(X_r; Y_2) \tag{B.79}$$

and

$$R_2 \leq \liminf_{n \rightarrow \infty} \frac{n_2}{n} I(X_{r,Q}; Y_{1,Q}) \leq \beta I(X_r; Y_1) \quad (\text{B.80})$$

for some distribution $p_{X_r, Y_1, Y_2}(x_r, y_1, y_2) = \bar{s}(x_r) p_2(y_1, y_2 | x_r)$. Let

$$\begin{aligned} p_{X_1, X_2, X_r, Y_1, Y_2, Y_r}(x_1, x_2, x_r, y_1, y_2, y_r) \\ = \bar{q}(x_1, x_2) \bar{s}(x_r) p_1(y_r | x_1, x_2) p_2(y_1, y_2 | x_r), \end{aligned}$$

i.e. (X_1, X_2, Y_r) and (X_r, Y_1, Y_2) are independent. Then we see that

$$p_{X_1, X_2, Y_r}(x_1, x_2, y_r) = \bar{q}(x_1, x_2) p_1(y_r | x_1, x_2)$$

and

$$p_{X_r, Y_1, Y_2}(x_r, y_1, y_2) = \bar{s}(x_r) p_2(y_1, y_2 | x_r).$$

It then follows from (B.77), (B.78), (B.79) and (B.80) that for this choice of $p_{X_1, X_2, X_r, Y_1, Y_2, Y_r}(x_1, x_2, x_r, y_1, y_2, y_r)$,

$$R_1 \leq \min\{\alpha I(X_1; Y_r | X_2), \beta I(X_r; Y_2)\}$$

and

$$R_2 \leq \min\{\alpha I(X_2; Y_r | X_1), \beta I(X_r; Y_1)\}.$$

■

Proof of Theorem 13: Suppose $(R_1, R_2) \in \underline{\mathcal{R}}_G$. Following similar procedures for proving (B.69), we can obtain

$$\begin{aligned} R_1 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_1} \frac{n_1}{n_1 n} I(X_{1,k}; Y_{r,k} | X_{2,k}) \\ &\leq \frac{n_1}{2n} \log_2 \left(1 + \frac{P_1}{N_r} \right) \\ &\leq \frac{\alpha}{2} \log_2 \left(1 + \frac{P_1}{N_r} \right). \end{aligned} \quad (\text{B.81})$$

By symmetry, we obtain

$$R_2 \leq \frac{\alpha}{2} \log_2 \left(1 + \frac{P_2}{N_r} \right). \quad (\text{B.82})$$

Following similar procedures for proving (B.71), we can obtain

$$\begin{aligned} R_1 &\leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{n_2} \frac{n_2}{n_2 n} I(X_{r,k}; Y_{2,k}) \\ &\leq \frac{n_2}{n} \log_2 \left(1 + \frac{P_r}{N_2} \right) \\ &\leq \frac{\beta}{2} \log_2 \left(1 + \frac{P_r}{N_2} \right). \end{aligned} \quad (\text{B.83})$$

By symmetry, we obtain

$$R_2 \leq \frac{\beta}{2} \log_2 \left(1 + \frac{P_r}{N_1} \right). \quad (\text{B.84})$$

The theorem then follows from (B.81), (B.82), (B.83) and (B.84). \blacksquare

□ End of chapter.

Bibliography

- [1] J. N. Laneman, G. W. Wornell and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," in *Proc. IEEE ISIT'01*, 2001.
- [2] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2415–2425, 2003.
- [3] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, 2004.
- [4] B. Rankov, and A. Wittneben, "Achievable rate regions for the two-way relay channel," *IEEE International Symposium on Information Theory*, pp. 1668–1672, Jul. 2006.
- [5] R. Knopp, "Two-Way Radio Networks with a Star Topology," in *Proc. International Zurich Seminar on Communications (IZS)*, Feb. 2006.
- [6] T. J. Oechtering, C. Schnurr, I. Bjelakovic and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 54, pp. 454–458, Jan. 2008.

- [7] C. Schnurr, T. J. Oechtering and S. Stanczak, "Achievable rates for the restricted half-duplex two-way relay channel," in *Proc. Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers (ACSSC)*, Nov. 2007, pp. 1468–1472.
- [8] C. Schnurr, S. Stanczak and T. J. Oechtering, "Achievable rates for the restricted half-duplex two-way relay channel under a partial-decode-and-forward protocol," in *Proc. IEEE Information Theory Workshop (ITW)*, May 2008, pp. 134–138.
- [9] W. Nam, S.-Y. Chung and Y. H. Lee, "Capacity bounds for two-way relay channels," in *Proc. IEEE International Zurich Seminar on Communications*, Mar. 2008, pp. 144–147.
- [10] I.-J. Baik and S.-Y. Chung, "Network Coding for Two-Way Relay Channels using Lattices," in *Proc. IEEE International Conference on Communications (ICC)*, May 2008, pp. 3898–3902.
- [11] A. S. Avestimehr, A. Sezgin and D. N. C. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *Proc. Allerton Conference on Communication, Control, and Computing*, Sep. 2008, pp. 1582–1589.
- [12] R. Ahlswede, N. Cai, S.-Y. R. Li and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, 2000.
- [13] Y. Wu, P. A. Chou and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2005.

- [14] A. G. C. Berrou and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE International Conference on Communications (ICC), Geneva, Switzerland*, vol. 2, 1993, pp. 1064–1070.
- [15] D. J. C. MacKay and R. M. Neal, "Near shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 32, pp. 1645–1646, 1996.
- [16] S. Zhang, S. C. Liew and P. P. Lam, "Physical-layer network coding," in *Proc. ACM Mobicom*, Sep. 2006, pp. 358–365.
- [17] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, New York, 1968.
- [18] R. W. Yeung, *Information Theory and Network Coding*. Springer, 2008.
- [19] T. M. Cover, *Elements of Information Theory*, 2nd ed. Wiley, 2006.
- [20] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 47, pp. 3072–3081, 2001.
- [21] E. Damosso and L. Correira, Ed., *Digital Mobile Radio Towards Future Generation Systems: COST 231 Final Report*. European Commission, 1999.
- [22] R. Knopp, "Two-way wireless communication via a relay station," in *GDR-ISIS meeting*, Mar. 2007.
- [23] C. E. Shannon, "Two-way communication channels," in *Proc. the 4th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1961, pp. 611–644.

- [24] J. P. M. Schalkwijk, "On an extension of an achievable rate region for the binary multiplying channel," *IEEE Trans. Inf. Theory*, vol. 29, pp. 445–448, May 1983.
- [25] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, pp. 44–53, Jan. 1989.
- [26] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, Belmont, Massachusetts, 1997.