Coding Schemes for the Two-Way Relay Channels

Peng Zhong

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

August 2012

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

Abstract

In modern transmission networks, relay plays an important role for cooperative strategies. Several relaying strategies, such as decode-forward, compress-forward and amplify-forward, have been proposed for relay channels and networks. However, the capacity for the general relay channel and network is still unknown. In this thesis, we propose several relay schemes for different relay models.

In the first part of the thesis, we propose novel partial decode-forward (PDF) schemes for the two-way relay channel with direct link. Different from pure decode-forward, each user divides its message into two parts and the relay decodes only one part of each. The relay then generates its codeword as a function of the two decoded parts and forwards to the two users. We propose PDF schemes for both the full- and half-duplex modes. Analysis and simulation show that if for one user, the direct link is stronger than the user-to-relay link, while for the other, the direct link is weaker, then PDF can achieve a rate region strictly larger than the time-shared region of pure decode-forward and direct transmission for both full- and half-duplex modes.

The second part of the thesis is based on noisy network coding, which is recently proposed for the general multi-source network by Lim, Kim, El Gamal and Chung. This scheme builds on compress-forward (CF) relaying but involves three new ideas, namely no Wyner-Ziv binning, relaxed simultaneous decoding and message repetition. In this part, using the one-way and two-way relay channel as the underlining example, we analyze the impact of each of these ideas on the achievable rate region of relay networks.

In the third part of the thesis, we propose two coding schemes combining decode-forward (DF) and noisy network coding (NNC) with different flavors. The first is a combined DF-NNC scheme for the one-way relay channel which includes both DF and NNC as special cases by performing rate splitting, partial block Markov encoding and NNC. The second combines two different DF strategies and layered NNC for the two-way relay channel. Analysis and simulation show that both proposed schemes supersede each individual scheme and take full advantage of both DF and NNC.

Abrégé

Dans les réseaux de transmission modernes, les relais jouent un rôle important dans les stratégies coopératives. Plusieurs stratégies de relai, telles que decode-forward, compress-forward et amplify-forward, ont été proposées pour les canaux et réseaux à relais. Cependant, la capacité du canal à relai général et de tels réseaux reste toujours inconnue. Dans cette thèse, nous proposons plusieurs stratégies de relai pour différents modèles.

Dans un premier temps, nous proposons de nouvelles stratégies de decode-forward partiel (PDF) pour le canal à relai bidirectionnel avec lien direct. A la différence du decode-forward classique, chaque utilisateur divise son message en deux parties, mais le relai ne décode que l'une d'entre elles pour chacun. Le relai génère alors un mot de code en fonction de ces deux parties décodées et les transmet aux deux utilisateurs. Nous proposons une stratgie PDF à la fois pour les liaisons half- et full-duplex. Comme le montrent les analyses et simulations réalisées, si, pour l'un des utilisateurs, le lien direct est meilleur que le lien utilisateur-relai alors que, pour l'autre utilisateur, le lien direct est plus faible, dans ce cas, la stratégie PDF permet d'accroître strictement la région des débits atteignables par rapport à la région atteinte par le partage de temps avec la stratégie decode-forward classique et la transmission directe, à la fois pour les liaisons half- et full-duplex.

La deuxième partie de cette thèse s'intéresse au codage de réseau avec bruit, qui a été abordé récemment pour les réseaux multi-sources génériques par Lim, Kim, El Gamal et Chung. Cette stratégie se base sur le relayage par compress-forward (CF), mais utilise trois nouvelles idées, à savoir le binning de Wyner-Ziv, le décodage simultané moins contraignant et la répétition de message. Dans cette partie, nous prenons pour exemple les canaux à relai mono- et bidirectionnels, et nous analysons l'impact de chacune de ces idées sur la région des débits atteignables pour les réseaux à relais.

Dans la troisième partie de cette thèse, nous proposons deux stratégies de codage qui combinent le decode-forward (DF) et le codage de réseau avec bruit (NNC), avec différentes nuances. La première est une stratégie combinée DF-NNC pour le canal à relai monodirectionnel, pour laquelle DF et NNC représentent des cas particuliers par partage de débit, de même que lencodage partiel en bloc de Markov et NNC. La deuxième stratégie combine deux stratégies DF différentes au codage NNC en couches pour le canal à relai bidirectionnel. Les analyses et les simulations montrent que les deux stratégies proposées remplacent chaque stratégie individuelle et prennent pleinement avantage des stratégies DF et NNC.

Acknowledgments

Foremost, I would like to express my gratitude to my supervisor, Prof. Mai Vu. She gave me the opportunity to join her lab and work with her. Over the past two years, I have gained a deep respect for Mai's attitudes and principles about research. She taught me how to do research and how to dig into a problem. Her comments and feedback on my research work benefited me a lot. Without her patient guidance and support, this thesis would have never been written.

I would also like to thank my lab members, including Ahmad, Fanny, Zhuohua, Yao and Jun. Special thanks to Ahmad for his help in my "multiuser communication" course. As course teaching assistant, he always showed enough patience in solving our problems and teaching us many useful methods. I also enjoyed discussing research problems with Zhuohua. Also, thanks for his help in some administrative issues. It has been a great pleasure to work with him. Special thanks go to Fanny for translating the abstract into French.

I am very lucky to meet so many friends at McGill. They helped me a lot with my study and life. It is hard to imagine life at McGill without them. Special thanks to my roommate Xiaofan. He shared his thoughts about his research to me. It helps me to learn a lot of new concepts outside my field. Also, thanks Nancy for proof checking my English writing.

Finally, I would like to express my endless gratitude to my family. My parents instilled a passion for science in me when I was still young. They spare no effort to support me studying abroad. Without their love and support, I would never have been able to study at McGill.

Contents

List of Figures

List of Tables

List of Acronyms

DF	Decode-Forward
CF	Compress-Forward
HF	Hash-Forward
EHF	Extended Hash-Forward
PDF	Partial Decode-Forward
RC	Relay Channel
TWRC	Two-Way Relay Channel
NNC	Noisy Network Coding
LNNC	Layered Noisy Network Coding
MARC	Multiple Access Relay Channel
BRC	Broadcase Relay Channel
SNR	Signal-to-Noise Ratio
AWGN	Additive White Gaussian Noise
DM	Discrete meroryless

Chapter 1

Introduction

1.1 Background

With the increasing size of communication networks, cooperative transmission is becoming more and more important. For example, in a wireless network, the transmitted message from a node is heard not only by its intended receiver, but also by other neighbour nodes. Those neighbour nodes can use the received signals to help transmission. They bring a cooperative transmission by acting as relays.

The relay channel (RC) first introduced by van der Meulen consists of a source aiming to communicate with a destination with the help of a relay (called single relay channel or one-way relay channel). In [1], Cover and El Gamal propose the fundamental decode-forward (DF), compress-forward (CF) schemes for the one-way relay channel. In DF, the relay decodes the message from the source and forwards it to the destination. In CF, the relay compresses received signal and forwards the compression index. Although a combination of these schemes achieve capacity of several types of channels, none of them are optimal in general. We will discuss more details in the literature review section.

The one-way relay channel can be extended to the two-way relay channel (TWRC) in which a relay helps two users exchange messages. Two types of TWRC exist: one without a direct link between the two users, a model suitable for wired communication, and one with the direct link, more suitable for wireless communication. In this thesis, we focus on the TWRC with direct link between the two users, also called the full TWRC. TWRC is a practical channel model for wireless communication systems. For example, a dedicated relay station has been proposed in 4G wireless standards to help the mobile and base station exchange messages. The decode-forward and compress-forward schemes can be generalized to the two-way relay channel, such as in [2] and [3].

More generally, relay channels can be extended to relay networks, in which each node wishes to send a message to some destinations while also acting as a relay for others. In [4], decode-forward and compress-forward are studied in relay networks. In [5], Lim, Kim, El Gamal and Chung propose a noisy network coding scheme based on compress-forward for the general relay network. More details on those works will be discussed in literature review section.

Although relay channels and networks have drawn growing attention, the capacity region of relay network is still unknown. What is the optimal coding scheme that achieves the capacity region? In this thesis, we propose and analyze several coding schemes for the relay channels. These schemes are steps towards understanding the optimal coding.

1.2 Literature Review

A number of coding schemes have been proposed for relay channels and networks. Some basic relaying strategies include amplify-forward, decode-forward and compress-forward. In this section, we will first review works on decode-forward and compress-forward. Most of our works in this thesis are based on these two schemes. After that, we will briefly review a new relaying strategy called compute-forward.

Before our literature review, we introduce two transmission modes: full-duplex transmission and half-duplex transmission. In full-duplex transmission, each node can transmit and receive at the same time; whereas for half-duplex transmission, each node can only either transmit or receive at each time. In this section, unless otherwise specified, the transmission mode is full-duplex.

1.2.1 Decode-Forward

In this part, we review related works on decode-forward for relay channels and relay networks. We divide the discussion into two parts. The first part is on single-source, single destination relay networks. The second part is on multi-source, multi-destination relay networks such as the two-way relay channel.

Single-source single destination relay networks

- In [1], Cover and El Gamal propose a decode-forward scheme for the one-way relay channel. The source uses block Markov superposition encoding. The relay decodes the message and sends its random binning index. The destination performs successive decoding. The following rate is achievable with DF:

$$R \leq \min\{I(X, X_r; Y), I(X; Y_r|X_r)\} \tag{1.1}$$

  for some $p(x, x_r)$. They also propose a partial decode-forward scheme in which the message is split into two parts, and the relay only decodes one part of them. It achieves the same rate either as decode or as direct transmission (without using the relay) for the Gaussian channel.

- In [6], Willems and van der Meulen introduces a backward decoding in which decoding at the receiver is done backwards after all blocks are received. It achieves the same rates as that in [1] for the discrete memoryless channel.

Multi-source multi-destination relay networks

- In [2], Rankov and Wittneben apply decode-forward to the two-way relay channel. In their proposed DF scheme, the two users perform partial block Markov encoding, and the relay sends a superposition of the codewords for the two decoded messages in each block.

- A different DF strategy is proposed in [3] by Xie, in which the users encode independently with the relay without block Markovity, and the relay sends a codeword for the random binning of the two decoded messages. These two DF schemes [2] [3] do not include each other in general.

- In [4], Kramer, Gastpar and Gupta extend decode-forward to several classes of relay networks, including single-source, single-destination, multi-relay network, multiple access relay channel (MARC) and broadcast relay channel (BRC). Sliding-window decoding is performed at the destinations.

- Decode-forward has also been applied to the half-duplex two-way relay channel. In [7], three full decode-forward protocols are proposed which has 2, 3 or 4 phases, in which the 4-phase protocol contains the 2- and 3-phase ones as special cases and achieves the largest rate region. In [8], these authors extend the protocols to a mixed relaying strategy which combines CF in one direction and DF in the other.

1.2.2 Compress-Forward

In this part, we review related works on compress-forward (CF) strategies for relay channels and networks. We divide the discussion into three parts. The first part is on single-source, single-destination relay networks. The second part is on some variants of the CF scheme. The third part is on multi-source multi-destination relay networks.

Single-source single-destination relay networks

In the following works, the source and relay encoding are similar. At each block, the source sends a different message; the relay first compresses its received signal then uses Wyner-Ziv binning to reduce the forwarding rate. The differences are mainly in the decoding at the destination by either performing successive or joint decoding.

- Compress-forward is originally proposed for the 3-node single-relay channel (also called the one-way relay channel) by Cover and El Gamal in [1]. The source sends a new message at each block using independent codebooks. The relay compresses its noisy observation of the source signal and forwards the bin index of the compression to the destination using Wyner-Ziv coding [9]. A 3-step sequential decoding is then performed at the destination. At the end of each block, the destination first decodes the bin index, and then decodes the compression index within that bin, and at last uses this compression index to decode the message sent in the previous block. The following rate is achievable with the 3-step sequential decoding CF scheme:

$$R \leq I(X;Y,\hat{Y}_r|X_r) \tag{1.2}$$

subject to

$$I(X_r;Y) \geq I(\hat{Y}_r;Y_r|X_r,Y).$$

for some $p(x)p(x_r)p(\hat{y}_r|y_r,x_r)p(y,y_r|x,x_r)$.

- El Gamal, Mohseni, and Zahedi put forward a 2-step decoding CF scheme in [10]. The source and relay perform the same encoding as that in [1]. The destination, however, decodes in 2 sequential steps. At the end of each block, it decodes the bin index first, and then decodes the message for some compression indices within that bin instead of decoding the compression index precisely. With this 2-step decoding CF scheme, the following rate is achievable:

$$R \leq \min\{I(X, X_r; Y) - I(\hat{Y}_r; Y_r|X, X_r, Y), I(X; Y, \hat{Y}_r|X_r)\} \tag{1.3}$$

  for some $p(x)p(x_r)p(\hat{y}_r|y_r, x_r)p(y, y_r|x, x_r)$. It has been shown [10] [11] that this 2-step decoding CF achieves the same rate as the original 3-step decoding CF in (1.2) but has a simpler representation.

- Kramer, Gastpar, and Gupta extend the 3-step decoding CF scheme to the single-source, single-destination and multiple-relay network in [4]. The relays can also cooperate with each other to transmit the compression bin indices by partially decoding these bin indices.

- Chong, Motani and Garg propose two coding schemes for the one-way relay channel combining decode-forward and compress-forward in [12]. Similar to the original combined scheme in [1], the source splits its message into two parts and the relay decode-forwards one part and compress-forwards the other. The destination, however, performs backward decoding either successively or simultaneously. These two strategies achieve higher rates than the original combined strategy in [1] for certain parameters of the Gaussian relay channel.

Variants of compress-forward

Several variants of the CF scheme have been proposed for the relay channel.

- Cover and Kim propose a hash-forward (HF) scheme for the deterministic relay channel in [13], in which the relay hashes (randomly bins) its observation directly without compression and forwards the bin index to the destination. HF achieves the capacity of the deterministic relay channel. Kim then proposes an extended hash-forward (EHF) scheme in [14] which allows the destination to perform list decoding of the source messages for the general non-deterministic case.

- Razaghi and Yu introduce in [15] a generalized hash-forward (GHF) relay strategy which allows the relay to choose a description of a general form rather than direct hashing (binning) of its received signal, but with a description rate on the opposite regime of Wyner-Ziv binning. The destination then performs list decoding of the description indices. GHF achieves the same rate as the original CF for the one-way relay channel but have been shown to exhibit advantage for multi-destination networks by allowing different description rates to different destinations [16].

- Recently a new notion of quantize-forward or CF without binning emerges [5] [17] in which the relay compresses its received signal but forwards the compression index directly without using Wyner-Ziv binning. We discuss this idea in more details in the next few paragraphs.

Multi-source multi-destination relay networks

Relatively fewer works have applied CF to the general multi-source multi-destination relay network.

- Rankov and Wittneben applied the 3-step decoding CF scheme to the two-way relay channel (TWRC) in [2], in which two users wish to exchange messages with the help of a relay. The encoding and decoding are similar to those in [1].

- Recently, Lim, Kim, El Gamal and Chung put forward a noisy network coding scheme [5] for the general multi-source noisy network. This scheme involves three key new ideas. The first is message repetition, in which the same message is sent multiple times over consecutive blocks using independent codebooks. Second, each relay does not use Wyner-Ziv binning but only compresses its received signal and forwards the compression index directly. Third, each destination performs simultaneous decoding of the message based on signals received from all blocks without uniquely decoding the compression indices. Noisy network coding simplifies to the capacity-achieving network coding for the noiseless multicast network. Compared to the original CF, it achieves the same rate for the one-way relay channel and achieves a larger rate region when applied to multi-source networks such as the two-way relay channel. However, it also brings more delay in decoding because of message repetition.

- In [18], Lim, Kim, El Gamal and Chung propose an improved NNC scheme termed ”layered noisy network coding” (LNNC). The relay compresses its observation into two layers: one is used at both destinations, while the other is only used at one destination.

- In [19], Ramalingam and Wang propose a superposition NNC scheme for restricted relay networks, in which source nodes cannot act as relays, by combining decode-forward and noisy network coding and show some performance improvement over NNC. Their scheme, however, does not include DF relaying rate because of no block Markov encoding.

Analysis of compress-forward schemes

With the above variants and developments on CF relaying, some works have analyzed the different ideas in compress-forward.

- Kim, Skoglund and Caire [20] show that without Wyner-Ziv binning at the relay, using sequential decoding at the destination incurs rate loss in the one-way relay channel. The amount of rate loss is quantified specifically in terms of the diversity-multiplexing tradeoff for the fading channel.

- Wu and Xie demonstrate in [21] that for single-source, single-destination and multiple-relay networks, using the original CF encoding with Wyner-Ziv binning of [1], there is no improvement on the achievable rate by joint decoding of the message and compression indices. To maximize the CF achievable rate, the compression rate should always be chosen to support successive decoding.

- Wu and Xie then propose in [22] for the single-source, single-destination and multiple-relay network a scheme that achieves the same rate as noisy network coding [5] but with the simpler classical encoding of [1] and backward decoding. The backward decoding involves first decoding the compression indices then successively decoding the messages backward. It requires, however, extending the relay forwarding times for a number of blocks without sending new messages, which causes an albeit small but non-vanishing rate loss.

- Kramer and Hou discuss in [17] a short-message quantize-forward scheme without message repetition or Wyner-Ziv binning but with joint decoding of the message and compression index at the destination. It also achieves the same rate as the original CF and noisy network coding for the one-way relay channel.
- Recently, Hou and Kramer in [23] propose a short message noisy network coding for multiple sources relay network. It transmits independent short messages in blocks rather than using long message repetitive encoding and uses backward decoding. It is shown to achieve the same rates as noisy network coding.

1.2.3 Compute-Forward

A new relaying strategy called compute-forward was recently proposed in [24], in which the relay decodes linear functions of transmitted messages. Nested lattice code [25] is used to implement compute-forward in Gaussian channels, since it ensures the sum of two codewords is still a codeword. Compute-forward has been shown to outperforms DF in moderate SNR regimes but is worse at low or high SNR [24]. Compute-forward can be naturally applied in two-way relay channels as the relay now receives signal containing more than one message. In [26], nested lattice codes were proposed for the Gaussian separated TWRC with symmetric channel, i.e. all source and relay nodes have the same transmit powers and noise variances. For the more general separated AWGN TWRC case, compute-forward coding with nested lattice code can achievable rate region within 1/2 bit of the cut-set outer bound [27] [28]. For the full AWGN TWRC, a scheme based on compute-forward, list decoding and random binning technique is proposed in [29]. This scheme achieves rate region within 1/2 bit of the cut-set bound in some cases.

In [30], we propose a combined decode-forward and compute-forward scheme for the two-way relay channel. The combined scheme uses superposition coding of both Gaussian and lattice codes to allow the relay to decode the Gaussian parts and compute the lattice parts. This scheme can also achieve new rates and outperform both decode-forward and compute-forward separately.

1.3 Outline and Contributions

This section outlines the thesis and summarizes main contributions.

Chapter 2

This chapter introduces channel models that will be used in the thesis, including the one-way relay channel, two-way relay channel and general relay network. For each of them, both discrete memoryless and Gaussian models will be discussed.

Chapter 3

In [3], a decode-forward scheme is proposed for the full-duplex two-way relay channel. In [7], a 4-phase decode-forward scheme is proposed for the half-duplex two-way relay channel. However, similar to the case in the one-way relay channel, both of them cannot include the direct transmission rate region. This motivates us to propose a partial decode-forward scheme for the two-way relay channel. This chapter is organized as follows.

In the first part of this chapter, we propose a partial decode-forward scheme for full-duplex TWRC. Each user divides its message into two parts and the relay decodes only one part. Numerical results have shown that partial decode-forward outperforms pure decode-forward and direct transmission in general. Moreover, we provide the analytical conditions for when partial decode-forward achieves new rates outside the time-shared region of pure decode-forward and direct transmission. As the second part of this thesis, we propose a partial decode-forward scheme for the 4-phase transmission protocol, in which each user divides its message into two parts and the relay only decodes one part of each message. The relay then generates its codeword as a function of the decoded parts and forwards to users. This scheme outperforms both the pure DF scheme in [7] and direct transmission.

Contents in this chapter have been published as [30] [31]:

- P. Zhong and M. Vu, "Decode-forward and compute-forward coding schemes for the two-way relay channel," in *IEEE Info. Theory Workshop (ITW)*, Oct. 2011.
- P. Zhong and M. Vu, "Partial decode-forward coding schemes for the Gaussian two-way relay channel," in *IEEE Int'l Conf. on Comm. (ICC)*, Oct. 2012.

Chapter 4

Noisy network coding is proposed in [5]. It is based on compress-forward and includes three new ideas, namely no Wyner-Ziv binning, relaxed simultaneous decoding and message repetition. Although achieving larger rate region than compress-forward, it brings an infinite delay because of message repetition. This motivates us to propose a compress-forward scheme without Wyner-Ziv binning and analyze the impact of each ideas in relay networks. This chapter is organized as follows.

We first derive the achievable rate using CF without binning (also called quantize-forward) but with joint decoding of both the message and compression index for the one-way relay channel. It achieves the same rate as the original CF in [1] [10]. Compared with the original CF, it simplifies relay operation since Wyner-Ziv binning is not needed, but increases decoding complexity at the destination since joint decoding instead of successive decoding is required. Compared with noisy network coding, it achieves the same rate while having much less encoding and decoding delay.

In the second part, we extend CF without binning to the two-way relay channel and derive its achievable rate region. The scheme achieves a larger rate region than the original CF [2]. With binning and successive decoding, the compression rate is constrained by the weaker of the links from relay to two users. But without binning, this constraint is relaxed. However, CF without binning generally achieves smaller rate region than noisy network coding [5]. In CF without binning, the decoding of the compression index imposes constraints on the compression rate. In noisy network coding, the destinations do not decode the compression index explicitly, thus removing these constraints.

In the third part, using the two-way relay channel as the underlining example, we analyze the effect of each of the three new ideas in noisy network coding for the general multi-source multi-destination relay networks.

Contents in this chapter have been published/submitted as [32] [33]:

- P. Zhong and M. Vu, "Compress-forward without Wyner-Ziv binning for the one-way and two-way relay channels," in *49th Annual Allerton Conf. on Comm., Control, and Computing*, Sept. 2011.

- P. Zhong, A. A. A. Haija, and M. Vu, "On compress-forward without Wyner-Ziv binning for relay networks," *submitted to IEEE Trans. on Info. Theory. Arxiv preprint arXiv:1111.2837*, 2011.

Chapter 5

In this chapter, we first propose a combined DF-NNC scheme for the one-way channel. Different from [19], our proposed scheme performs block Markov encoding and hence encompasses both DF relaying and NNC as special cases. It outperforms the combined decode-forward and compress-forward scheme in [1] under certain channel parameters, and achieves the same rate as the backward decoding strategies in [12] for the Gaussian relay channel. We then propose a combined DF-LNNC scheme for the TWRC. This scheme also includes partial block Markov encoding and, in addition, performs layered NNC. Analysis and numerical results show that this scheme outperforms each individual scheme in [2,3,18] and also the combined scheme in [19].

Contents in this chapter have been published as [34] [1]:

- P. Zhong and M. Vu, "Combined decode-forward and layered noisy network coding schemes for relay channels," submitted to *IEEE Int'l Symp. on Info. Theory (ISIT)*, July 2012.

Chapter 6

This chapter concludes this thesis and discusses potential future work.

[1] This thesis also contains a correction to the result published in Theorem 1 of [34], as described in detail in Chapter 5 Section 5.2.

Chapter 2

Channel Models

In this chapter, we introduce various channel models which will be discussed in our thesis. Those channel models includes the one-way relay channel, two-way relay channel and relay networks. For each channel, we introduce its discrete memoryless model and Gaussian model respectively. We also introduce notations used in this thesis which are similar to those in [11].

2.1 One-Way Relay Channel Model

2.1.1 Discrete Memoryless One-Way Relay Channel Model

The discrete memoryless one-way relay channel (DM-RC) is denoted by $(\mathcal{X}\times\mathcal{X}_r, p(y, y_r|x, x_r), \mathcal{Y}\times\mathcal{Y}_r)$, as in Figure 2.1. Sender X wishes to send a message M to receiver Y with the help of the relay (X_r, Y_r). We consider a full-duplex channel in which all nodes can transmit and receive at the same time.

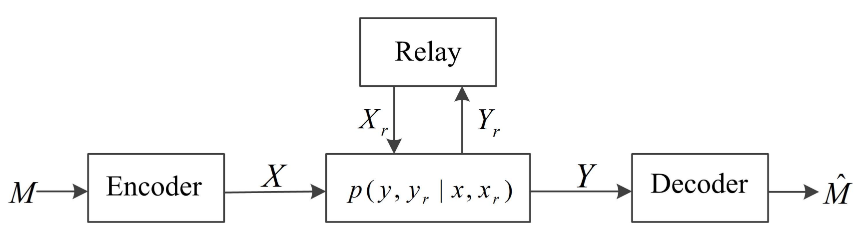

Fig. 2.1 Discrete memoryless one-way relay channel model.

A $(2^{nR}, n, P_e)$ code for a DM-RC consists of: a message set $\mathcal{M} = [1 : 2^{nR}]$; an encoder

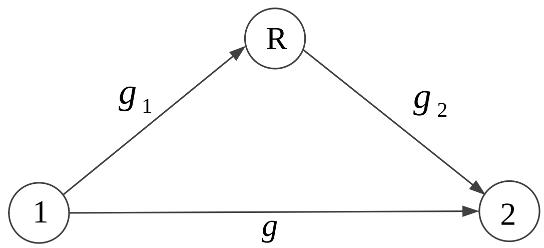

Fig. 2.2 Gaussian one-way relay channel model.

that assigns a codeword $x^n(m)$ to each message $m \in [1:2^{nR}]$; a relay encoder that assigns at time $i \in [1:n]$ a symbol $x_{ri}(y_r^{i-1})$ to each past received sequence $y_r^{i-1} \in \mathcal{Y}_r^{i-1}$; a decoder that assigns a message $\hat{m}$ or an error message to each received sequence $y^n \in \mathcal{Y}^n$. The average error probability is $P_e = \Pr\{\hat{M} \neq M\}$. The rate R is said to be achievable for the DM-RC if there exists a sequence of $(2^{nR}, n, P_e)$ codes with $P_e \to 0$. The supremum of all achievable rates is the capacity of the DM-RC.

2.1.2 Gaussian One-Way Relay Channel Model

As in Figure 2.2, the Gaussian one-way relay channel can be modeled as

$$Y = gX + g_2X_r + Z, \quad Y_r = g_1X + Z_r, \tag{2.1}$$

where $Z, Z_r \sim \mathcal{N}(0,1)$ are independent Gaussian noises, and g, g_1, g_2 are the corresponding channel gains.

2.2 Two-Way Relay Channel Model

2.2.1 Discrete Memoryless Two-Way Relay Channel Model

The discrete memoryless two-way relay channel (DM-TWRC) is denoted by $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_r, p(y_1, y_2, y_r|x_1, x_2, x_r), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_r)$, as in Figure 2.3. Here x_1 and y_1 are the input and output signals of user 1; x_2 and y_2 are the input and output signals of user 2; x_r and y_r are the input and output signals of the relay.

A $(2^{nR_1}, 2^{nR_2}, n, P_e)$ code for a DM-TWRC consists of two message sets $\mathcal{M}_1 = [1:2^{nR_1}]$ and $\mathcal{M}_2 = [1:2^{nR_2}]$, three encoding functions $f_{1,i}, f_{2,i}, f_{r,i}$, $i = 1, \ldots, n$ and two decoding

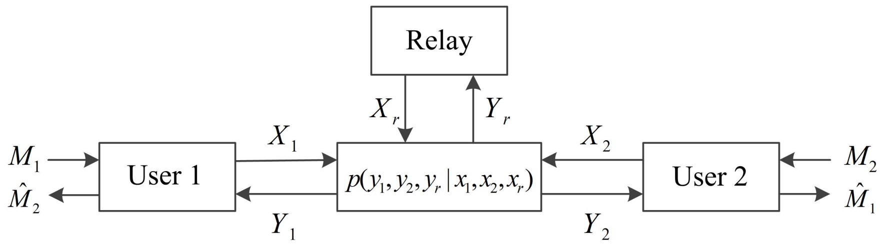

Fig. 2.3 Discrete memoryless two-way relay channel model.

function g_1, g_2 as follows:

$$\begin{aligned}
x_{1,i} &= f_{1,i}(M_1, Y_{1,1}, \ldots, Y_{1,i-1}), \quad i = 1, \ldots, n \\
x_{2,i} &= f_{2,i}(M_2, Y_{2,1}, \ldots, Y_{2,i-1}), \quad i = 1, \ldots, n \\
x_{r,i} &= f_{r,i}(Y_{r,1}, \ldots, Y_{r,i-1}), \quad i = 1, \ldots, n \\
g_1 &: \mathcal{Y}_1^n \times \mathcal{M}_1 \to \mathcal{M}_2 \\
g_2 &: \mathcal{Y}_2^n \times \mathcal{M}_2 \to \mathcal{M}_1
\end{aligned}$$

The average error probability is $P_e = \Pr\{g_1(M_1, Y_1^n) \neq M_2 \text{ or } g_2(M_2, Y_2^n) \neq M_1\}$. A rate pair is said to be achievable if there exists a $(2^{nR_1}, 2^{nR_2}, n, P_e)$ code such that $P_e \to 0$ as $n \to \infty$. The closure of the set of all achievable rates (R_1, R_2) is the capacity region of the two-way relay channel.

2.2.2 Full-Duplex Gaussian Two-Way Relay Channel Model

For the Gaussian two-way relay channel, we consider two transmission modes: full-duplex mode and half-duplex mode. In full-duplex transmission, each node can transmit and receive at the same time; whereas for half-duplex transmission, each node can only either transmit or receive at each time. We first discuss the full-duplex model, and then discuss the half-duplex model in next section.

As in Figure 2.4, the full-duplex Gaussian two-way relay channel can be modeled as:

$$\begin{aligned}
Y_1 &= g_{12}X_2 + g_{1r}X_r + Z_1 \\
Y_2 &= g_{21}X_1 + g_{2r}X_r + Z_2 \\
Y_r &= g_{r1}X_1 + g_{r2}X_2 + Z_r \qquad (2.2)
\end{aligned}$$

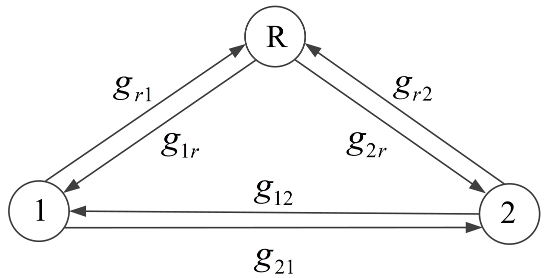

Fig. 2.4 Full-duplex Gaussian two-way relay channel model.

where $Z_1, Z_2, Z_r \sim \mathcal{N}(0,1)$ are independent Gaussian noises. The average input power constraints for user 1, user 2 and the relay are all P. $g_{12}, g_{1r}, g_{21}, g_{2r}, g_{r1}, g_{r2}$ are corresponding channel gains.

2.2.3 Half-Duplex Gaussian Two-Way Relay Channel Model

For the half-duplex mode, each node can only either send or receive at each time. We consider a 4-phase half-duplex Gaussian two-way relay model as in Figure 2.5, as motivated by [7] which shows the best performance out of several protocols.

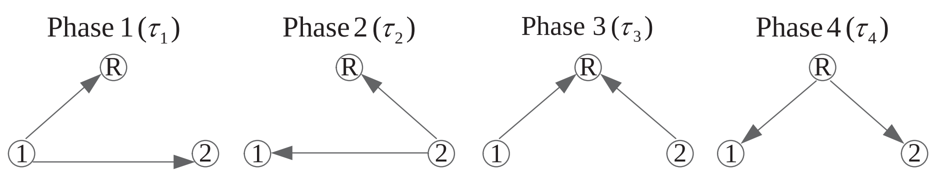

Fig. 2.5 Half-duplex Gaussian two-way relay channel model.

During the 1^{st} phase, user 1 transmits. During the 2^{nd} phase, user 2 transmits. During the 3^{rd} phase, both user 1 and user 2 transmit. During the 4^{th} phase, the relay transmits. Assume all nodes listen while not transmitting. The transmitted signals during each phase can be expressed as

$$\text{Phase 1:} \quad Y_{21} = g_{21}X_{11} + Z_{21}, \quad Y_{r1} = g_{r1}X_{11} + Z_{r1}$$
$$\text{Phase 2:} \quad Y_{12} = g_{12}X_{22} + Z_{12}, \quad Y_{r2} = g_{r2}X_{22} + Z_{r2}$$

$$\text{Phase 3:} \quad Y_{r3} = g_{r1}X_{13} + g_{r2}X_{23} + Z_{r2}$$
$$\text{Phase 4:} \quad Y_{14} = g_{1r}X_r + Z_{14}, \quad Y_{24} = g_{2r}X_r + Z_{24},$$

where X_{ij} represents the transmitted signal of user i during phase j. Y_{ij} represents received signal of user i during phase j. All the noises Z are independently and identically distributed according to $\mathcal{N}(0, 1)$.

2.3 Relay Networks Model

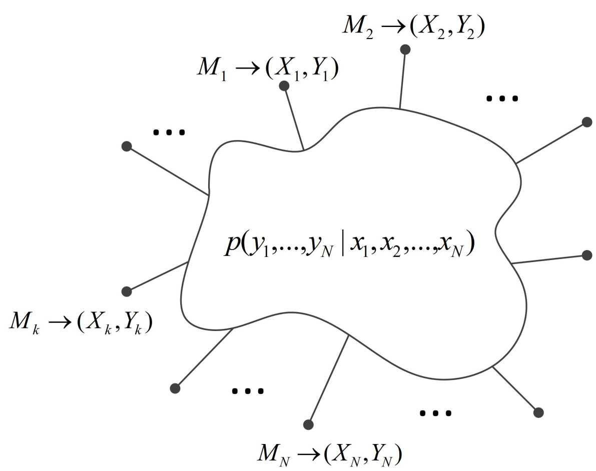

Fig. 2.6 Relay network model.

A N-node discrete memoryless relay network $(\prod_{k=1}^{N} \mathcal{X}_k, p(y_N|x_N), \times_{k=1}^{N} \mathcal{Y}_k)$ is depicted in Figure 2.6. It consists of N sender-receiver alphabet pairs $(\mathcal{X}_k, \mathcal{Y}_k)$, $k \in [1 : N]$ and a collection of conditional pmfs $p(y_1, \cdots y_N|x_1, \cdots x_N)$. Each node $k \in [1 : N]$ wishes to send a message M_k to a set of destination nodes, while also acting as a relay for messages from other nodes.

Chapter 3

Partial Decode-Forward Coding Schemes

3.1 Problem Statement

For the one-way relay channel, Cover and El Gamal propose a decode-forward scheme in [1]. The relay fully decodes the message from the source and forwards it to the destination. However, when the direct link is stronger than the user-to-relay link, decode-forward is not the optimal scheme compared to direct transmission without using the relay. To combine both cases, they also design a partial decode-forward scheme where the relay decides to decode part of the message according to the channel condition.

For the two-way relay channel, there are also two existing coding schemes: direct transmission and full decode-forward scheme. For direct transmission, the relay is not used. For full decode-forward, the relay decodes the whole message of each user then forwards a function of these messages as in [2] [3]. As we will see, similarly to the one-way relay channel, direct transmission achieves strictly larger rate region than decode-forward if for both users, the direct link is stronger than the user-to-relay link. If for both users, the user-to-relay link is sufficiently stronger than the direct link, then decode-forward outperforms direct transmission. However, in cases such that for one user, the user-to-relay link is stronger than the direct link, while for the other user, it's the opposite, then neither existing scheme outperforms the other. This motivates us to put forward the partial decode-forward scheme where the relay only decodes a part of the messages and forwards them. We propose partial decode-forward for both full- and half-duplex two-way relay channels.

3.2 Partial DF for Full-Duplex TWRC

3.2.1 Coding Scheme and Achievable Rate Region for the DM TWRC

In this section, we provide an achievable rate region for the full-duplex TWRC with a partial decode-forward scheme. Each user splits its message into two parts and uses superposition coding to encode them. The relay only decodes one message part of each user and re-encode the decoded message pair together and broadcast. It can either re-encode each message pair separately or divides these message pairs into lists and only encodes the list index, which is similar to the binning technique in [3]. Both strategies achieve the same rate region. The users then decode the message from each other by joint typicality decoding of both the current and previous blocks.

Theorem 1. *The following rate region is achievable for the two-way relay channel with partial decode-forward:*

$$\begin{aligned} R_1 &\leq \min\{I(U_1;Y_r|U_2,X_r) + I(X_1;Y_2|U_1,X_2,X_r), I(X_1,X_r;Y_2|X_2)\} \\ R_2 &\leq \min\{I(U_2;Y_r|U_1,X_r) + I(X_2;Y_1|U_2,X_1,X_r), I(X_2,X_r;Y_1|X_1)\} \\ R_1+R_2 &\leq I(U_1,U_2;Y_r|X_r) + I(X_1;Y_2|U_1,X_2,X_r) + I(X_2;Y_1|U_2,X_1,X_r) \end{aligned} \tag{3.1}$$

for some joint distribution $p(u_1,x_1)p(u_2,x_2)p(x_r)$.

Proof. We use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each.

1) Codebook generation: Fix $p(u_1,x_1)p(u_2,x_2)p(x_r)$. Split each message into two parts: $m_1 = (m_{10}, m_{11})$ with rate (R_{10}, R_{11}), and $m_2 = (m_{20}, m_{22})$ with rate (R_{20}, R_{22}).

- Generate $2^{nR_{10}}$ i.i.d. sequences $u_1^n(m_{10}) \sim \prod_{i=1}^n p(u_{1i})$, where $m_{10} \in [1:2^{nR_{10}}]$. For each $u_1^n(m_{10})$, generate $2^{nR_{11}}$ i.i.d. sequences $x_1^n(m_{11}, m_{10}) \sim \prod_{i=1}^n p(x_{1i}|u_{1i})$, where $m_{11} \in [1:2^{nR_{11}}]$.

- Generate $2^{nR_{20}}$ i.i.d. sequences $u_2^n(m_{20}) \sim \prod_{i=1}^n p(u_{2i})$, where $m_{20} \in [1:2^{nR_{20}}]$. For each $u_2^n(m_{20})$, generate $2^{nR_{22}}$ i.i.d. sequences $x_2^n(m_{22}, m_{20}) \sim \prod_{i=1}^n p(x_{2i}|u_{2i})$, where $m_{22} \in [1:2^{nR_{22}}]$.

- Uniformly throw each message pair (m_{10}, m_{20}) into 2^{nR_r} bins. Let $K(m_{10}, m_{20})$ denote the index of bin.
- Generate 2^{nR_r} i.i.d. sequences $x_r^n(K) \sim \prod_{i=1}^n p(x_{ri})$, where $K \in [1 : 2^{nR_r}]$. If $R_r = R_{10} + R_{20}$, there is no need for binning.

The codebook is revealed to all parties.

2) Encoding: In each block $j \in [1 : b-1]$, user 1 and user 2 transmit $x_1^n(m_{11,j}, m_{10,j})$ and $x_2^n(m_{22,j}, m_{20,j})$ respectively. In block b, user 1 and user 2 transmit $x_1^n(1,1)$ and $x_2^n(1,1)$, respectively.

At the end of block j, the relay has an estimate $(\tilde{m}_{10,j}, \tilde{m}_{20,j})$ from the decoding procedure. It transmits $x_r^n(K(\tilde{m}_{10,j}, \tilde{m}_{20,j}))$ in block $j+1$.

3) Decoding: We explain the decoding strategy at the end of block j.

Decoding at the relay: Upon receiving $y_r^n(j)$, the relay searches for the unique pair $(\tilde{m}_{10,j}, \tilde{m}_{20,j})$ such that

$$\left(u_1^n(\tilde{m}_{10,j}), u_2^n(\tilde{m}_{20,j}), y_r^n(j), x_r^n\left(K(\tilde{m}_{10,j-1}, \tilde{m}_{20,j-1})\right)\right) \in T_\epsilon^n.$$

Following the analysis in multiple access channel, the error probability will go to zero as $n \to \infty$ if

$$\begin{aligned} R_{10} &\le I(U_1; Y_r | U_2, X_r) \\ R_{20} &\le I(U_2; Y_r | U_1, X_r) \\ R_{10} + R_{20} &\le I(U_1, U_2; Y_r | X_r). \end{aligned} \tag{3.2}$$

Decoding at each user: By block j, user 2 has decoded $m_{1,j-2}$. At the end of block j, it searches for a unique message pair $(\hat{m}_{10,j-1}, \hat{m}_{11,j-1})$ such that

$$\left(x_r^n(K(\hat{m}_{10,j-1}, m_{20,j-1})), y_2^n(j), x_{2,j}^n\right) \in T_\epsilon^n$$

and $$\left(u_1^n(\hat{m}_{10,j-1}), x_1^n(\hat{m}_{11,j-1}, \hat{m}_{10,j-1}), y_2^n(j-1), x_r^n(K(m_{1,j-2}, m_{2,j-2})), x_{2,j-1}^n\right) \in T_\epsilon^n.$$

Following joint decoding analysis, the error probability will go to zero as $n \to \infty$ if

$$\begin{aligned} R_{11} &\le I(X_1;Y_2|U_1,X_2,X_r) \\ R_{10}+R_{11} &\le I(X_r;Y_2|X_2)+I(U_1,X_1;Y_2|X_2,X_r)=I(X_1,X_r;Y_2|X_2). \end{aligned} \tag{3.3}$$

Similarly, user 1 can decode $(m_{20,j-1}, m_{22,j-1})$ with error probability goes to zero as $n \to \infty$ if

$$\begin{aligned} R_{22} &\le I(X_2;Y_1|U_2,X_1,X_r) \\ R_{20}+R_{22} &\le I(X_2,X_r;Y_1|X_1). \end{aligned} \tag{3.4}$$

By applying Fourier-Motzkin Elimination to the inequalities in (3.2)-(3.4), the achievable rates in terms of $R_1 = R_{10} + R_{11}$ and $R_2 = R_{20} + R_{22}$ are as given in Theorem 1. □

Remark 1. If $U_1 = X_1, U_2 = X_2$, this region reduces to the decode-forward lower bound in [3]. Therefore, the partial DF scheme contains the DF scheme in [3] as a special case.

3.2.2 Achievable Rate Region Analysis for the Guassian TWRC

Now we apply the proposed partial decode-forward scheme to the full-duplex Gaussian TWRC in (2.2). Using jointly Gaussian codewords, we can derive an achievable rate region as follows.

Theorem 2. *The following rate region is achievable for the Gaussian two-way relay channel.*

$$\begin{aligned} R_1 &\le \min\left\{C\left(\frac{g_{r1}^2\alpha P}{g_{r1}^2\bar{\alpha}P+g_{r2}^2\bar{\beta}P+1}\right)+C(g_{21}^2\bar{\alpha}P), C(g_{21}^2P+g_{2r}^2P)\right\} \\ R_2 &\le \min\left\{C\left(\frac{g_{r2}^2\beta P}{g_{r1}^2\bar{\alpha}P+g_{r2}^2\bar{\beta}P+1}\right)+C(g_{12}^2\bar{\beta}P), C(g_{12}^2P+g_{1r}^2P)\right\} \\ R_1+R_2 &\le C\left(\frac{g_{r1}^2\alpha P+g_{r2}^2\beta P}{g_{r1}^2\bar{\alpha}P+g_{r2}^2\bar{\beta}P+1}\right)+C(g_{21}^2\bar{\alpha}P)+C(g_{12}^2\bar{\beta}P), \end{aligned} \tag{3.5}$$

where $0 \le \alpha, \beta \le 1$ *and* $C(x) = \frac{1}{2}\log(1+x)$.

Achievability follows from Theorem 1 by setting $X_1 = U_1 + V_1$, where $U_1 \sim \mathcal{N}(0, \alpha P_1)$ and $V_1 \sim \mathcal{N}(0, \bar{\alpha} P_1)$ are independent, and by setting $X_2 = U_2 + V_2$, where $U_2 \sim \mathcal{N}(0, \beta P_2)$ and $V_2 \sim \mathcal{N}(0, \bar{\beta} P_2)$ are independent.

Now we analyze and compare above rate region achieved by the proposed partial decode-forward scheme with that achieved by pure decode-forward scheme [3] and direct transmission (without using the relay) for different channel conditions. We first present the achievable rate region of pure decode-forward scheme and direct transmission.

Theorem 3. [3] *The following rate region is achievable for the full-duplex Gaussian two-way relay channel with pure decode-forward scheme:*

$$\begin{aligned} R_1 &\leq \min\left\{C(g_{r1}^2 P), C(g_{21}^2 P + g_{2r}^2 P)\right\} \\ R_2 &\leq \min\left\{C(g_{r2}^2 P), C(g_{12}^2 P + g_{1r}^2 P)\right\} \\ R_1 + R_2 &\leq C(g_{r1}^2 P + g_{r2}^2 P). \end{aligned} \tag{3.6}$$

If the two users only use direct links to exchange message instead of using the relay, the following rate region is achievable:

$$\begin{aligned} R_1 &\leq C(g_{21}^2 P) \\ R_2 &\leq C(g_{12}^2 P). \end{aligned} \tag{3.7}$$

Remark 2. If $\alpha = 1, \beta = 1$, the rate region in (3.5) reduces to the decode-forward lower bound in (3.6). If $\alpha = 0, \beta = 0$, the rate region in (3.5) reduces to the direct transmission lower bound in (3.7). Thus partial decode-forward region always include both decode-forward and direct transmission regions as special cases.

The following theorem compares the rate region of partial decode-forwards with that of pure decode-forwards and direct transmission for different channel cases.

Theorem 4. *Comparing PDF with pure decode-forward and direct transmission (without using the relay), we have the following 4 cases:*
1) PDF can achieve rates strictly outside the time-shared region of DF and direct transmis-

sion if

$$\begin{aligned} &g_{r1}^2 > g_{21}^2 + \min\{g_{21}^2 g_{r2}^2 P, g_{2r}^2\}, \quad g_{12}^2 > g_{r2}^2 \\ or \quad &g_{r2}^2 > g_{12}^2 + \min\{g_{12}^2 g_{r1}^2 P, g_{1r}^2\}, \quad g_{21}^2 > g_{r1}^2 \end{aligned} \tag{3.8}$$

2) PDF achieves the time-shared region of DF and direct transmission if

$$\begin{aligned} g_{21}^2 < g_{r1}^2, \quad g_{12}^2 < g_{r2}^2 \\ C(g_{21}^2 P) + C(g_{12}^2 P) > C(g_{r1}^2 P + g_{r2}^2 P) \end{aligned} \tag{3.9}$$

3) PDF achieves the same rate region as pure DF scheme which is strictly larger than direct transmission if

$$\begin{aligned} g_{21}^2 \leq g_{r1}^2, \quad g_{12}^2 \leq g_{r2}^2 \\ C(g_{21}^2 P) + C(g_{12}^2 P) \leq C(g_{r1}^2 P + g_{r2}^2 P) \end{aligned} \tag{3.10}$$

4) PDF achieves the same rate region as direct transmission which is strictly larger than DF if

$$g_{21}^2 \geq g_{r1}^2, \quad g_{12}^2 \geq g_{r2}^2. \tag{3.11}$$

Proof. See Appendix A.1. □

3.2.3 Discussion and Numerical Examples

Discussion

Some intuition for our proposed partial decode-forward (PDF) scheme can be developed as follows:

- Compared to DF, PDF involves extra superposition encoding which can be easily implemented in practice. It uses joint decoding similar to DF and hence has similar decoding complexity.
- When the user-to-relay link is weaker than the direct link, decoding the whole message at the relay limits the achievable rate. In such a case, partially decoding messages at

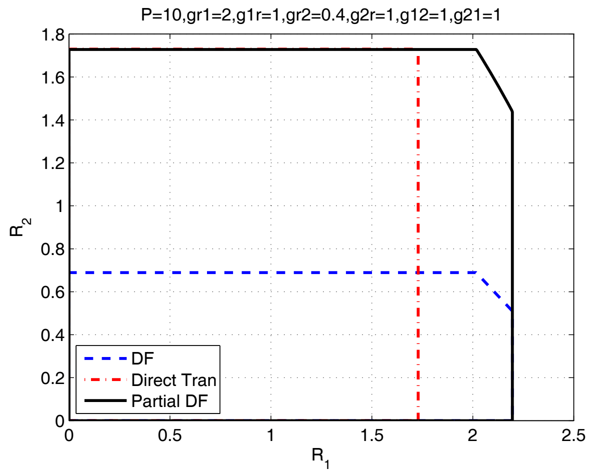

Fig. 3.1 Partial decode-forward achieves rates outside the time-shared region of decode-forward and direct transmission in the full-duplex TWRC.

the relay can relax the constraint and achieve a larger rate region.

- Theorem 4 implies that when both direct links are sufficiently weaker than the user-to-relay links, the relay should fully decode the messages and forward them. When both direct links are stronger than the user-to-relay links, the relay should not be used. If for one user, the direct link is stronger than the user-to-relay link, while for the other one, the direct link is weaker, then the relay should decode only a part of the message from the former.

- Applicability in wireless channels: In the wireless environment, the channel gains fluctuate and can easily cover all cases of Theorem 4. Thus it is useful to know the optimal scheme for each case such that each user can adapt their transmission according to the channel strength.

Numerical examples

For cases 1 and 2 in Theorem 4, we provide each an example. Figure 3.1 shows an example in which partial decode-forward achieves rates outside the time-shared region of decode-forward and direct transmission. Figure 3.2 shows an example where partial decode-forward achieves the time-shared region of decode-forward and direct transmission. In both cases, the proposed PDF scheme outperforms pure DF.

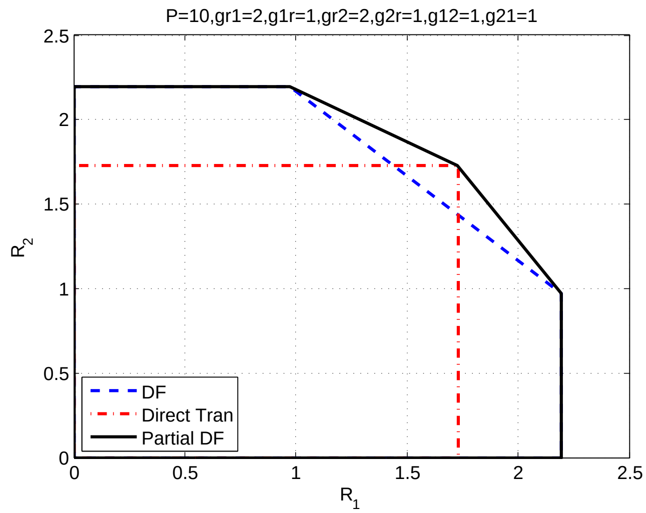

Fig. 3.2 Partial decode-forward achieves time-shared region of decode-forward and direct transmission in the full-duplex TWRC.

3.3 Partial DF for Half-Duplex TWRC

In this section, we design a partial decode-forward scheme for the half-duplex case. The half-duplex mode in which each node can either transmit or receive at each time, is more practical in wireless systems. Moreover, transmissions are performed in independent blocks without block Markovity. Each user can decode the message of the other user at the end of each block without any delay.

Three decode-forward protocols for the half-duplex two-way relay channel have been proposed in [7]. The first protocol divides each block into 2 phases, in which both users transmit in the first phase and the relay transmits in the second phase. The second protocol divides each block into 3 phases, in which user 1 transmits in the first phase, user 2 in the second phase and the relay in the third phase. The third protocol divides each block into 4 phases, in which user 1 transmits in the first phase, user 2 in the second phase, both users transmit in the third phase and the relay transmits in the last phase. All nodes listen while not transmitting. It has been shown that the 4-phase achieves the largest rate region among these three protocols.

We will only discuss a 4-phase partial decode-forward scheme as it outperforms the other two. The main difference between our scheme and the scheme in [7] is that the relay only decodes a part of the messages in our scheme, whereas it decodes the full messages

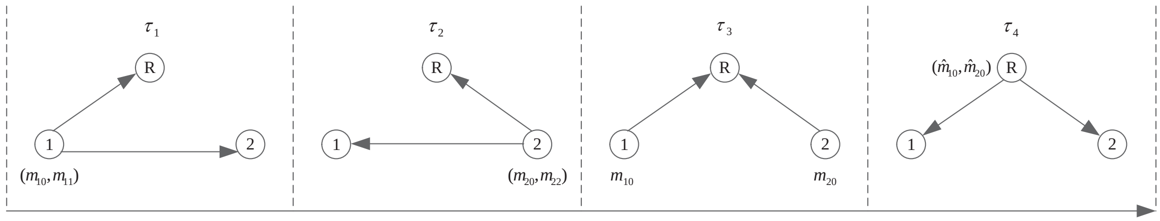

Fig. 3.3 Half-duplex partial decode-forward transmission diagram.

in [7]. When a direct link is stronger than the user-to-relay link, the proposed scheme achieves strictly larger rate region.

3.3.1 Coding Scheme and Achievable Rate Region

Consider the transmission at each block, which is divided into four phases as in Figure 3.3. Each message is divided into two parts for each user. During the 1st phase, user 1 transmits both parts. During the 2nd phase, user 2 transmits both parts. During the 3rd phase, both users transmit only one part of their messages. At the end of the 3rd phase, the relay decodes this part of each message based on the received signals from all first three phases. It then transmits a function of those message parts during the 4th phase. At the end of the 4th phase, user 1 decodes the message of user 2 based on received signals in the 2nd and 4th phases. Similarly for user 2.

User encoding

Let the relative time duration of the phases are τ_1, τ_2, τ_3 and τ_4 respectively, where $\tau_1 + \tau_2 + \tau_3 + \tau_4 = 1$. Let m_1 be the message of user 1 to be sent during a specific block. User 1 divides it into two parts (m_{10}, m_{11}) with rate (R_{10}, R_{11}) and encodes m_{10} and m_{11} by U_1 and V_1 respectively. Then the transmitted signals of user 1 during phase 1 and 3 respectively are as follows.

$$
\begin{aligned}
X_{11} &= \sqrt{\alpha_{11}} U_1(m_{10}) + \sqrt{\beta_{11}} V_1(m_{11}) \\
X_{13} &= \sqrt{\alpha_{13}} U_1(m_{10})
\end{aligned}
$$

where $\alpha_{11}, \beta_{11}, \alpha_{13}$ are corresponding power allocations. Similarly, user 2 divides its message m_2 into two parts (m_{20}, m_{22}) with rate (R_{20}, R_{22}) and encodes m_{20} and m_{22} by U_2 and V_2

respectively. Its transmitted signals in the 2nd and 3rd phases respectively are

$$X_{22} = \sqrt{\alpha_{22}}U_2(m_{20}) + \sqrt{\beta_{22}}V_2(m_{22})$$
$$X_{23} = \sqrt{\alpha_{23}}U_2(m_{20}).$$

Relay operation

Decoding: At the end of the 3rd phase, the relay decodes the messages parts (m_{10}, m_{20}) based on received signals from the 1st, 2nd and 3rd phases by joint decoding.

Encoding: The relay then constructs its transmitted signal in the 4th phase as

$$X_r = \sqrt{\gamma}W(m_{10}, m_{20})$$

where $W(m_{10}, m_{20})$ can be generated as a function (for example, XOR or random binning) of the codewords for (m_{10}, m_{20}).

In the above signals, U_1, V_1, U_2, V_2, W are independent and identically distributed according to $\mathcal{N}(0, 1)$. The power constraints for the two users and the relay are as follows.

$$\begin{aligned} \tau_1(\alpha_{11} + \beta_{11}) + \tau_3\alpha_{13} &= P \\ \tau_2(\alpha_{22} + \beta_{22}) + \tau_3\alpha_{23} &= P \\ \tau_4\gamma &= P. \end{aligned} \tag{3.12}$$

User decoding

At the end of phase 4, user 2 uses joint decoding to decode message $m_1 = (m_{10}, m_{11})$ based on received signals from both the 1st and 4th phases. Similarly, user 1 decodes $m_2 = (m_{20}, m_{22})$ based on received signals from both the 2nd and 4th phases.

Theorem 5. *The following rate region is achievable for the half-duplex Gaussian two-way* *relay channel with partial decode-forward scheme.*

$$R_{10} \le \tau_1 C\left(\frac{g_{r1}^2\alpha_{11}}{g_{r1}^2\beta_{11}+1}\right) + \tau_3 C(g_{r1}^2\alpha_{13}) = I_1 \tag{3.13a}$$
$$R_{20} \le \tau_2 C\left(\frac{g_{r2}^2\alpha_{22}}{g_{r2}^2\beta_{21}+1}\right) + \tau_3 C(g_{r2}^2\alpha_{23}) = I_2 \tag{3.13b}$$
$$R_{11} \le \tau_1 C(g_{21}^2\beta_{11}) = I_3 \tag{3.13c}$$
$$R_{22} \le \tau_2 C(g_{12}^2\beta_{22}) = I_4 \tag{3.13d}$$
$$R_{10}+R_{20} \le \tau_1 C\left(\frac{g_{r1}^2\alpha_{11}}{g_{r1}^2\beta_{11}+1}\right) + \tau_2 C\left(\frac{g_{r2}^2\alpha_{22}}{g_{r2}^2\beta_{21}+1}\right) + \tau_3 C(g_{r1}^2\alpha_{13}+g_{r2}^2\alpha_{23}) = I_5 \tag{3.13e}$$
$$R_{10}+R_{11} \le \tau_1 C(g_{21}^2(\alpha_{11}+\beta_{11})) + \tau_4 C(g_{2r}^2\gamma) = I_6 \tag{3.13f}$$
$$R_{20}+R_{22} \le \tau_2 C(g_{12}^2(\alpha_{22}+\beta_{22})) + \tau_4 C(g_{1r}^2\gamma) = I_7 \tag{3.13g}$$

with power constraints in (3.12), where $\tau_1+\tau_2+\tau_3+\tau_4 = 1$ *and* $C(x) = \frac{1}{2}\log(1+x)$. *By applying Fourier-Motzkin Elimination, the achievable rates in terms of* $R_1 = R_{10}+R_{11}$ *and* $R_2 = R_{20}+R_{22}$ *can be expressed as*

$$\begin{aligned} R_1 &\le \min\{I_1+I_3, I_6\} \\ R_2 &\le \min\{I_2+I_4, I_7\} \\ R_1+R_2 &\le I_3+I_4+I_5. \end{aligned} \tag{3.14}$$

Proof. At the end of the 3rd phase, the relay decodes (m_{10}, m_{20}) based on received signals from the 1st, 2nd and 3rd phases, which can succeed with high probability if (3.13a), (3.13b) and (3.13e) are satisfied. During the 4th phase, the relay sends $X_r(m_{10}, m_{20})$. Based on the received signals in the 1st phase Y_{21} and the 4th phase Y_{41}, user 2 can decode $m_1 = (m_{10}, m_{11})$ with error probability going to zero if (3.13c) and (3.13f) are satisfied. Similarly, user 1 can decode $m_2 = (m_{20}, m_{22})$ with vanishing error if (3.13d) and (3.13g) are satisfied. □

3.3.2 Discussion

Several points can be noted for our proposed PDF scheme as follows:

- For the half-duplex mode, there is no block Markovity. Therefore, encoding and

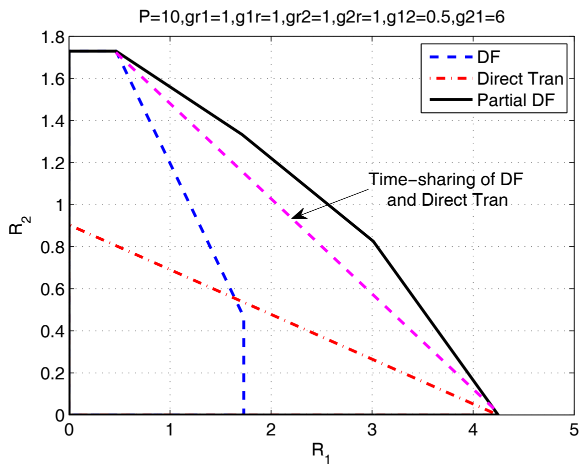

Fig. 3.4 Rate region comparison between partial decode-forward, pure decode-forward and direct transmission for the half-duplex Gaussian TWRC.

decoding are simple and can be done within one block, which is also practical.

- The signaling for each user again involves only 2-part superposition coding which can be easily implemented in practice.
- It is also interesting to find the optimal power allocations and time slot durations to maximize the achievable rates as these are of directly practical value. These can be topics of future work.
- Similar to the full-duplex case, partial decode-forward helps when the direct link is stronger than the user-to-relay link for one user, while is weaker for the other.
- Our proposed half-duplex PDF scheme again includes both the DF scheme in [7] and direct transmission as special cases.

3.3.3 Numerical Comparison

We numerically compare the achievable rate regions of partial decode-forward, pure decode-forward [7] and direct transmission. For direct transmission, we divide each block into two phases, where user 1 transmits in the first phase and user 2 in the second phase. For pure decode-forward, we include power scaling to satisfy the power constraint (3.12), which is different from [7] with fixed power. Hence the DF region here is larger than that in [7]. Figure 3.4 shows that the proposed partial decode-forward scheme achieves strictly larger rate region than the other two schemes with new rates outside the time-shared region of the other two. This result also agrees with the analysis in Theorem 4 that when the direct link is stronger than the user-to-relay link for one user but is weaker for the other, then PDF strictly outperforms the time-sharing of both DF and direct transmission.

Chapter 4

Compress-Forward Without Binning

4.1 Problem Statement

Recently, Lim, Kim, El Gamal and Chung put forward a noisy network coding scheme [5] for the general multi-source noisy network. This scheme involves three key new ideas. The first is message repetition, in which the same message is sent multiple times over consecutive blocks using independent codebooks. Second, each relay does not use Wyner-Ziv binning but only compresses its received signal and forwards the compression index directly. Third, each destination performs simultaneous decoding of the message based on signals received from all blocks without uniquely decoding the compression indices. Compared to the original CF, it achieves the same rate for the one-way relay channel and achieves a larger rate region when applied to multi-source networks such as the two-way relay channel.

However, message repetition in noisy network coding brings an infinite delay. To reduce the delay and analyze the impact of no Wyner-Ziv binning, we propose a compress-forward scheme without Wyner-Ziv binning but with joint decoding of two consecutive blocks. We apply it to the one-way and two-way relay channel, and then extend to the relay network. To analyze the impact of message repetition, we modify the scheme in which the message is repeated only once or twice. We compare their achievable rate regions and analyze the improvements each technique can bring.

4.2 One-Way Relay Channel

4.2.1 Coding Scheme and Achievable Rate

In the original CF scheme [1] [10], the source sends a new message in each block. The relay forwards the bin index of the description of its received signal. The receiver uses successive decoding to decode the bin index first, then decode the message from the sender. Here we analyze a CF scheme in which the relay forwards the index of the description of its received signal directly without binning while the receiver jointly decodes the index and message at the same time. We show that CF without binning can achieve same rate as the original CF scheme with binning.

The encoding and decoding of CF without binning are as follows (also see Table 4.1). We use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each.

Block	...	j	$j+1$	...
X	...	$x^n(m_j)$	$x^n(m_{j+1})$	...
Y_r	...	$\hat{y}_r(k_j\|k_{j-1})$	$\hat{y}_r(k_{j+1}\|k_j)$	...
X_r	...	$x_r^n(k_{j-1})$	$x_r^n(k_j)$	...
Y	...	$\hat{k}_{j-1}, \hat{m}_{j-1}$	$\hat{k}_j, \hat{m}_j$	...

Table 4.1 Encoding and decoding of CF without binning for the one-way relay channel.

Codebook generation

Fix $p(x)p(x_r)p(\hat{y}_r|y_r, x_r)$. We randomly and independently generate a codebook for each block $j \in [1:b]$ as follows.

- Independently generate 2^{nR} sequences $x^n(m_j) \sim \prod_{i=1}^n p(x_i)$, where $m_j \in [1:2^{nR}]$.
- Independently generate 2^{nR_r} sequences $x_r^n(k_{j-1}) \sim \prod_{i=1}^n p(x_{ri})$, where $k_{j-1} \in [1:2^{nR_r}]$.
- For each $k_{j-1} \in [1:2^{nR_r}]$, independently generate 2^{nR_r} sequences $\hat{y}_r^n(k_j|k_{j-1}) \sim \prod_{i=1}^n p(\hat{y}_{ri}|x_{ri}(k_{j-1}))$ where $k_j \in [1:2^{nR_r}]$.

Encoding

The source transmits $x^n(m_j)$ in block j. The relay, upon receiving $y_r^n(j)$, finds an index k_j such that

$$(\hat{y}_r^n(k_j|k_{j-1}), y_r^n(j), x_r^n(k_{j-1})) \in T_\epsilon^{(n)},$$

where $T_\epsilon^{(n)}$ denote the strong ϵ-typical set [11]. Assume that such k_j is found, the relay sends $x_r^n(k_j)$ in block $j+1$.

Decoding

Assume the receiver has decoded k_{j-1} correctly in block j. Then in block $j+1$, the receiver finds a unique pair of $(\hat{m}_j, \hat{k}_j)$ such that

$$\begin{aligned} &(x_r^n(\hat{k}_j), y^n(j+1)) \in T_\epsilon^{(n)} \\ \text{and} \quad &(x^n(\hat{m}_j), x_r^n(\hat{k}_{j-1}), \hat{y}_r^n(\hat{k}_j|\hat{k}_{j-1}), y^n(j)) \in T_\epsilon^{(n)}. \end{aligned}$$

Theorem 6. *Consider a compress-forward scheme in which the relay does not use Wyner-Ziv binning but sends the compression index directly and the destination performs joint decoding of both the message and compression index. The following rate is achievable for the one-way relay channel:*

$$R \leq \min\{I(X, X_r; Y) - I(\hat{Y}_r; Y_r|X, X_r, Y), I(X; Y, \hat{Y}_r|X_r)\} \tag{4.1}$$

subject to

$$I(X_r; Y) + I(\hat{Y}_r; X, Y|X_r) \geq I(\hat{Y}_r; Y_r|X_r) \tag{4.2}$$

for some $p(x)p(x_r)p(\hat{y}_r|y_r, x_r)p(y, y_r|x, x_r)$.

Proof. See Appendix A.2. □

4.2.2 Comparison with Original Compress-Forward Scheme

Theorem 7. *Compress-forward without binning in Theorem 6 achieves the same rate as the original compress-forward scheme for the one-way relay channel, which is:*

$$R \leq \min\{I(X, X_r; Y) - I(\hat{Y}_r; Y_r|X, X_r, Y), I(X; Y, \hat{Y}_r|X_r)\} \tag{4.3}$$

for some $p(x)p(x_r)p(\hat{y}_r|y_r, x_r)p(y, y_r|x, x_r)$.

Proof. To show that the rate region in Theorem 6 is the same as the rate region in Theorem 7, we need to show that constraint (4.2) is redundant. Note that an equivalent characterization of the rate region in Theorem 7 is as follows [1] [10] [11]:

$$R \leq I(X; Y, \hat{Y}_r|X_r) \tag{4.4}$$

subject to

$$I(X_r; Y) \geq I(\hat{Y}_r; Y_r|X_r, Y) \tag{4.5}$$

for some $p(x)p(x_r)p(\hat{y}_r|y_r, x_r)$. Therefore, comparing (4.2) with (4.5), we only need to show that

$$I(\hat{Y}_r; Y_r|X_r, Y) \geq I(\hat{Y}_r; Y_r|X_r) - I(\hat{Y}_r; X, Y|X_r).$$

This is true since

$$\begin{aligned}
I(\hat{Y}_r; Y_r|X_r, Y) &= I(\hat{Y}_r; Y_r, X|X_r, Y) \\
&= I(X; \hat{Y}_r|X_r, Y) + I(Y_r; \hat{Y}_r|X, X_r, Y) \\
&\geq I(Y_r; \hat{Y}_r|X, X_r, Y) \\
&= I(\hat{Y}_r; X, Y, Y_r|X_r) - I(\hat{Y}_r; X, Y|X_r) \\
&= I(\hat{Y}_r; Y_r|X_r) - I(\hat{Y}_r; X, Y|X_r).
\end{aligned}$$

□

Remark 3. If using successive decoding, the rate achieved by CF without binning is strictly less than that with binning. Thus joint decoding is crucial for CF without binning.

Remark 4. Joint decoding does not help improve the rate of the original CF with binning.

Remark 5. Binning technique plays a role of allowing successive decoding instead of joint decoding, thus reduces destination decoding complexity. However, it has no impact on the achievable rate for the one-way relay channel. This effect on decoding complexity is similar to that in decode-forward, in which binning allows successive decoding [1] while no binning requires backward decoding [35].

Remark 6. For the one-way relay channel, CF without binning achieves the same rate region as GHF [15]. However, GHF differs from CF without binning in that the relay still performs binning. In the decoding of GHF, the destination first decode the compression indices into a list, and then use this list to help decode the source message. However, in CF without binning, the receiver performs joint decoding of the compression index and message.

Remark 7. An obvious benefit of this short message CF without binning scheme compared to noisy network coding is short encoding and decoding delay. A few other benefits such as low modulation complexity and potential MIMO gain are also recognized in [17].

Remark 8. For the one-way relay channel, all the following schemes achieve the same rate: the original CF, CF without binning (QF), GHF, and noisy network coding. The difference in achievable rate only appears when applying to a multi-source and multi-destination network.

4.3 Two-Way Relay Channel

4.3.1 Coding Scheme and Achievable Rate Region

In this section, we extend CF without Wyner-Ziv binning but with joint decoding of both the message and compression index to the two-way relay channel. We then compare the achievable rate region with those by the original CF scheme and noisy network coding. Compared with the original CF [2], CF without binning achieves a strictly larger rate region when the channel is asymmetric for the two users. Binning and successive decoding constrains the compression rate to the weaker of the channels from relay to two users. But without binning, this constraint is relaxed. Compared with noisy network coding, CF without binning achieves a smaller rate region. In CF without binning, the users need to

Block	...	j	$j+1$	...
X_1	...	$x_1^n(m_{1,j})$	$x_1^n(m_{1,j+1})$	...
X_2	...	$x_2^n(m_{2,j})$	$x_2^n(m_{2,j+1})$	...
Y_r	...	$\hat{y}_r(k_j\|k_{j-1})$	$\hat{y}_r(k_{j+1}\|k_j)$	...
X_r	...	$x_r^n(k_{j-1})$	$x_r^n(k_j)$	...
Y_1	...	$\hat{k}_{j-1}, \hat{m}_{2,j-1}$	$\hat{k}_j, \hat{m}_{2,j}$	...
Y_2	...	$\hat{k}_{j-1}, \hat{m}_{1,j-1}$	$\hat{k}_j, \hat{m}_{1,j}$	...

Table 4.2 Encoding and decoding of CF without binning for the two-way relay channel.

decode the compression index precisely, which brings an extra constraint on the compression rate. However, this precise decoding is not necessary in noisy network coding.

Before presenting the achievable rate region of CF without binning, we outline its encoding and decoding techniques as follows. In each block, each user sends a new message using an independently generated codebook. At the end of each block, the relay finds a description of its received signal from both users. Then it sends the codeword for the description index at the next block (instead of partitioning the description index into bins and sending the codeword for the bin index as in the original CF scheme). Each user jointly decodes the description index and message from the other user based on signals received in both the current and previous blocks. This decoding technique is different from that in the original decode-forward scheme, in which each user first decodes the bin index from the relay, and then decodes the message from the other user.

Specifically, we use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each (also see Table 4.2).

Codebook generation

Fix joint distribution $p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r, y_r)$. Randomly and independently generate a codebook for each block $j \in [1 : b]$

- Independently generate 2^{nR_1} sequences $x_1^n(m_{1,j}) \sim \prod_{i=1}^n p(x_{1i})$, where $m_{1,j} \in [1 : 2^{nR_1}]$.

- Independently generate 2^{nR_2} sequences $x_2^n(m_{2,j}) \sim \prod_{i=1}^n p(x_{2i})$, where $m_{2,j} \in [1 : 2^{nR_2}]$.
- Independently generate 2^{nR_r} sequences $x_r^n(k_{j-1}) \sim \prod_{i=1}^n p(x_{ri})$, where $k_{j-1} \in [1 : 2^{nR_r}]$.
- For each $k_{j-1} \in [1 : 2^{nR_r}]$, independently generate 2^{nR_r} sequences $\hat{y}_r^n(k_j|k_{j-1}) \sim \prod_{i=1}^n p(\hat{y}_{ri}|x_{ri}(k_{j-1}))$, where $k_j \in [1 : 2^{nR_r}]$.

Encoding

User 1 and user 2 respectively transmit $x_1^n(m_{1,j})$ and $x_2^n(m_{2,j})$ in block j. The relay, upon receiving $y_r^n(j)$, finds an index k_j such that

$$(\hat{y}_r^n(k_j|k_{j-1}), y_r^n(j), x_r^n(k_{j-1})) \in T_\epsilon^{(n)}.$$

Assume that such k_j is found, the relay sends $x_r^n(k_j)$ in block $j+1$.

Decoding

We discuss the decoding at user 1. Assume user 1 has decoded k_{j-1} correctly in block j. Then in block $j+1$, user 1 finds a unique pair of $(\hat{m}_{2,j}, \hat{k}_j)$ such that

$$\begin{aligned} &(x_2^n(\hat{m}_{2,j}), x_r^n(\hat{k}_{j-1}), \hat{y}_r^n(\hat{k}_j|\hat{k}_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)} \\ \text{and} \quad &(x_r^n(\hat{k}_j), y_1^n(j+1), x_1^n(m_{1,j+1})) \in T_\epsilon^{(n)}. \end{aligned} \tag{4.6}$$

Theorem 8. *The following rate region is achievable for the two-way relay channel by using compress-forward without binning but with joint decoding:*

$$\begin{aligned} R_1 &\le \min\{I(X_1; Y_2, \hat{Y}_r|X_2, X_r), I(X_1, X_r; Y_2|X_2) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_2)\} \\ R_2 &\le \min\{I(X_2; Y_1, \hat{Y}_r|X_1, X_r), I(X_2, X_r; Y_1|X_1) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1)\} \end{aligned} \tag{4.7}$$

subject to

$$\begin{aligned} I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1) &\le I(X_r; Y_1|X_1) \\ I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_2) &\le I(X_r; Y_2|X_2) \end{aligned} \tag{4.8}$$

for some $p(x_1)p(x_2)p(x_r)p(y_1, y_2, y_r|x_1, x_2, x_r)p(\hat{y}_r|x_r, y_r)$.

Proof. See Appendix A.3. □

4.3.2 Comparison with Original Compress-Forward Scheme

In this section, we first present the rate region achieved by the original CF scheme for the two way relay channel as in [2]. We then show that CF without binning but with joint decoding can achieve a larger rate region.

We outline the encoding and decoding techniques of the original CF scheme as follows. In each block, each user sends a new message using an independently generated codebook. At the end of each block, the relay finds a description of its received signal from both users. Then it partitions the description index into equal-size bins and sends the codeword for the bin index. Each user applies 3-step successive decoding, in which it first decodes the bin index from the relay, then decodes the compression index within that bin, and at last decodes the message from the other user.

Theorem 9. [Rankov and Wittneben]. *The following rate region is achievable for two-way relay channel with the original compress-forward scheme:*

$$\begin{aligned} R_1 &\le I(X_1; Y_2, \hat{Y}_r|X_2, X_r) \\ R_2 &\le I(X_2; Y_1, \hat{Y}_r|X_1, X_r) \end{aligned} \tag{4.9}$$

subject to

$$\max(I(\hat{Y}_r; Y_r|X_1, X_r, Y_1), I(\hat{Y}_r; Y_r|X_2, X_r, Y_2)) \le \min(I(X_r; Y_1|X_1), I(X_r; Y_2|X_2)) \tag{4.10}$$

for some $p(x_1)p(x_2)p(x_r)p(y_1, y_2, y_r|x_1, x_2, x_r)p(\hat{y}_r|x_r, y_r)$.

We present a short proof of this theorem in Appendix A.4 to show the difference from CF without binning. The proof follows the same lines as in [2], but we also correct an error in the analysis of [2] as pointed out in Remark 30 in Appendix A.4.

Proof. See Appendix A.4. □

Theorem 10. *In the two-way relay channel, the rate region achieved by compress-forward without binning in Theorem 8 is larger than the rate region achieved by the original compress-forward scheme in Theorem 9 when the channel is asymmetric for the two users. The two regions may be equal only if the channel is symmetric, that is the following conditions holds:*

$$\begin{aligned} I(X_r;Y_1|X_1) &= I(X_r;Y_2|X_2) \\ I(\hat{Y}_r;Y_r|X_1,X_r,Y_1) &= I(\hat{Y}_r;Y_r|X_2,X_r,Y_2). \end{aligned} \tag{4.11}$$

Furthermore, (4.11) *is only necessary but may not be sufficient.*

Proof. First, we show that the constraint on the compression rate of Theorem 8 is looser than that of Theorem 9. This is true since from (4.10), we have

$$\begin{aligned} I(X_r;Y_1|X_1) &\geq I(\hat{Y}_r;Y_r|X_1,X_r,Y_1) \\ &= I(\hat{Y}_r;X_2,Y_r|X_1,X_r,Y_1) \\ &\geq I(\hat{Y}_r;Y_r|X_1,X_2,X_r,Y_1) \end{aligned} \tag{4.12}$$

where (4.12) is the right-hand-side of the first term in (4.8). Similar for the other term.

Next we show that (4.9) and (4.10) imply (4.7). From (4.9), we have

$$\begin{aligned} R_2 &\leq I(X_2;Y_1,\hat{Y}_r|X_1,X_r) \\ &= I(X_2;Y_1|X_1,X_r) + I(\hat{Y}_r;X_2|Y_1,X_1,X_r) \\ &= I(X_2,X_r;Y_1|X_1) - I(X_r;Y_1|X_1) + I(\hat{Y}_r;X_2|Y_1,X_1,X_r) \\ &\overset{(a)}{\leq} I(X_2,X_r;Y_1|X_1) - I(\hat{Y}_r;Y_r|X_1,X_r,Y_1) + I(\hat{Y}_r;X_2|Y_1,X_1,X_r) \\ &= I(X_2,X_r;Y_1|X_1) - I(\hat{Y}_r;Y_r|X_1,X_2,X_r,Y_1) \end{aligned}$$

where (a) follows from the constraint of (4.10) in Theorem 9. The equality holds when

$$\begin{aligned} I(X_r;Y_1|X_1) &= \min\{I(X_r;Y_1|X_1), I(X_r;Y_2|X_2)\} \\ &= I(\hat{Y}_r;Y_r|X_1,X_r,Y_1) \\ &= \max(I(\hat{Y}_r;Y_r|X_1,X_r,Y_1), I(\hat{Y}_r;Y_r|X_2,X_r,Y_2)). \end{aligned}$$

Similar for R_1, the equality holds when

$$\begin{aligned} I(X_r;Y_2|X_2) &= \min\{I(X_r;Y_1|X_1), I(X_r;Y_2|X_2)\} \\ &= I(\hat{Y}_r;Y_r|X_2,X_r,Y_2) \\ &= \max(I(\hat{Y}_r;Y_r|X_1,X_r,Y_1), I(\hat{Y}_r;Y_r|X_2,X_r,Y_2)). \end{aligned}$$

The above analysis shows that at the boundary of the compression rate constraint (4.10), the rate region of CF without binning (4.7) is equivalent to that of the original CF (4.9). However, since constraint (4.8) is loser than (4.10), the rate region in Theorem 8 is larger than that in Theorem 9. Only if condition (4.11) holds, the original CF scheme may achieve the same rate region as CF without binning; otherwise, its rate region is strictly smaller. □

Remark 9. For the two-way relay channel, binning and successive decoding constrains the compression rate to the weaker of the channels from relay to two users. But without binning, this constraint is relaxed. Thus CF without binning achieves a larger rate region than the original CF scheme when then channel is asymmetric for the two users.

4.3.3 Comparison with Noisy Network Coding

In this section, we compare the rate region achieved by CF without binning with that by noisy network coding [5] for the two way relay channel. The main differences between these two schemes are as follows. In CF without binning, different messages are sent over different blocks, but in noisy network coding, the same message is sent in multiple blocks using independent codebooks. Furthermore, in noisy network coding, each user performs simultaneous joint decoding of the message based on signals received from all blocks without uniquely decoding the compression indices (i.e. relaxed joint decoding). But in CF without binning, each user jointly decodes both the message and compression index precisely based on signals received from the current and previous blocks.

Theorem 11. [Lim, Kim, El Gamal and Chung]. *The following rate region is achievable for the two-way relay channel with noisy network coding:*

$$\begin{aligned} R_1 &\leq \min\{I(X_1;Y_2,\hat{Y}_r|X_2,X_r), I(X_1,X_r;Y_2|X_2) - I(\hat{Y}_r;Y_r|X_1,X_2,X_r,Y_2)\} \\ R_2 &\leq \min\{I(X_2;Y_1,\hat{Y}_r|X_1,X_r), I(X_2,X_r;Y_1|X_1) - I(\hat{Y}_r;Y_r|X_1,X_2,X_r,Y_1)\} \end{aligned} \tag{4.13}$$

for some $p(x_1)p(x_2)p(x_r)p(y_1, y_2, y_r|x_1, x_2, x_r)p(\hat{y}_r|x_r, y_r)$.

Comparing Theorem 8 with Theorem 11, we find that the rate constraints (4.7) for R_1 and R_2 in CF without binning are the same as those in noisy network coding (4.13). However, CF without binning has an extra constraint on the compression rate (4.8). Therefore, in general, CF without binning achieves a smaller rate region than noisy network coding.

Remark 10. In noisy network coding, the combination of message repetition and relaxed joint decoding is necessary in addition to compress-forward without binning to achieve a better rate region than that in Theorem 8 for the TWRC.

4.4 Implication for Relay Networks

From the discussion for the one-way and two-way relay channels, we can acquire some implications for the general relay networks. The relay network model is shown in section 2.3. Noisy network coding is proposed for the general relay network in [5]. Three new ideas are used in noisy network coding. One is no Wyner-Ziv binning in relay operation. Another is simultaneous joint decoding of the message over all blocks without uniquely decoding the compression indices. Last is message repetition, in which the same message is sent in multiple blocks using independent codebooks. Next, we discuss the effect of each of these ideas separately.

4.4.1 Implication of no Wyner-Ziv Coding

Generalizing from the two-way relay channel, we can conclude that CF without binning achieves a larger rate region than the original CF scheme for networks with multiple destinations. With binning and successive decoding, the compression rate is constrained by the weakest link from a relay to a destination, as in (A.24). But with joint decoding of the message and compression indices, this constraint is more relaxed since each destination can also use the signals received from other links, including the direct links, to decode the compression indices and provides relays more freedom to choose the compression rates. This explains why for the two-way relay channel, the constraint on compression rate in CF without binning (4.8) is looser than that in the original CF scheme (4.10). Therefore, with joint decoding, Wyner-Ziv binning is not necessary. No binning also simplifies relay operation.

Remark 11. Joint decoding of both the message and compression index is crucial for CF without binning. Without joint decoding, it achieves strictly smaller rate than the original CF with binning for any network.

4.4.2 Implication of Joint Decoding without Explicitly Decoding Compression Indices

A difference between CF without binning and noisy network coding comes from the decoding of the compression indices. For both schemes, the compression rate at a relay is lower bounded by the covering lemma. But in noisy network coding, each destination does not decode the compression indices explicitly, hence there are no additional constraints on the compression rate. However, in CF without binning, each destination decodes the compression indices precisely; this decoding places extra upper bounds on the compression rate, leading to the constraints on compression rate as in (4.8).

The above analysis prompts the question: what if in CF without binning, we also do not decode the compression index precisely, can we achieve the same rate region as noisy network coding? The following analysis shows that the answer is negative. Take the two-way relay channel as an example. In the next few sections, we apply several joint decoding rules to enlarge the rate region.

Relaxed joint decoding of a single message without decoding the compression indices uniquely

Using the same codebook generation and encoding as in CF without binning in Section 4.3.1, but we change the decoding rule in (4.6) to as follows. In block $j+1$, user 1 finds a unique $\hat{m}_{2,j}$ such that

$$\begin{aligned} &(x_2^n(\hat{m}_{2,j}), x_r^n(\hat{k}_{j-1}), \hat{y}_r^n(\hat{k}_j|\hat{k}_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)} \\ \text{and} \quad &(x_r^n(\hat{k}_j), y_1^n(j+1), x_1^n(m_{1,j+1})) \in T_\epsilon^{(n)} \end{aligned} \tag{4.14}$$

for some pair of indices $(\hat{k}_{j-1}, \hat{k}_j)$. With this decoding rule, the error event $\mathcal{E}_{3j}$ which corresponds to wrong compression index only (see Appendix A.3) no longer applies, but other new error events appear in which $\hat{k}_{j-1} \neq 1$ (see the detailed error analysis in Appendix A.5). With decoding rule (4.14), we have following Corollary:

Corollary 1. *For the two-way relay channel, the following rate region is achievable by CF without binning using joint decoding but without decoding the compression index precisely:*

$$R_1 \le I(X_1; Y_2, \hat{Y}_r | X_2, X_r) \tag{4.15a}$$

$$R_1 \le I(X_1, X_r; Y_2 | X_2) - I(\hat{Y}_r; Y_r | X_1, X_2, X_r, Y_2) \tag{4.15b}$$

$$R_1 \le I(X_1, X_r; Y_2 | X_2) - I(\hat{Y}_r; Y_r | X_1, X_2, X_r, Y_2) + I(X_r; Y_2 | X_2) - I(\hat{Y}; Y_2 | X_r) \tag{4.15c}$$

$$R_2 \le I(X_2; Y_1, \hat{Y}_r | X_1, X_r) \tag{4.15d}$$

$$R_2 \le I(X_2, X_r; Y_1 | X_1) - I(\hat{Y}_r; Y_r | X_1, X_2, X_r, Y_1) \tag{4.15e}$$

$$R_2 \le I(X_2, X_r; Y_1 | X_1) - I(\hat{Y}_r; Y_r | X_1, X_2, X_r, Y_1) + I(X_r; Y_1 | X_1) - I(\hat{Y}; Y_1 | X_r) \tag{4.15f}$$

for some $p(x_1)p(x_2)p(x_r)p(y_1, y_2, y_r | x_1, x_2, x_r)p(\hat{y}_r | x_r, y_r)$.

Proof. See Appendix A.5. □

In above rate region, although the new decoding rule (4.14) removes the constraint on compression rate, it brings two new rate constraints (4.15c) (4.15f). Hence the rate region in (4.15) is still smaller than that of noisy network coding in (4.13). These two extra rate constraints come from the block boundary condition when performing joint decoding over two blocks. Specifically, it corresponds to the error event $\mathcal{E}_{7j}$ in Appendix A.5, in which all the compression indices and message are wrong. The boundary condition results from the second decoding rule in (4.14) in which not all input and output signals are involved.

Simultaneous joint decoding of all messages without decoding compression indices uniquely

To reduce the boundary effect, we try simultaneous decoding of all messages over all blocks, again without decoding the compression indices uniquely. We use the same codebook generation and encoding as in CF without binning in Section 4.3.1, but use a different decoding rule at the destination. The destination now jointly decodes all messages based on the signals received in all blocks, without decoding the compression indices explicitly. Specifically, user 1 now finds a unique tuple $(m_{2,1}, \dots m_{2,j}, \dots m_{2,b-1})$ at the end of block b

such that

$$
\begin{aligned}
&(x_2^n(m_{2,1}), x_r^n(1), \hat{y}_r^n(k_1|1), y_1^n(1), x_1^n(m_{1,1})) \in T_\epsilon^{(n)} && \text{(4.16a)}\\
&(x_2^n(m_{2,2}), x_r^n(k_1), \hat{y}_r^n(k_2|k_1), y_1^n(2), x_1^n(m_{1,2})) \in T_\epsilon^{(n)}\\
&\vdots\\
&(x_2^n(m_{2,j}), x_r^n(k_{j-1}), \hat{y}_r^n(k_j|k_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)}\\
&\vdots\\
&(x_2^n(m_{2,b-1}), x_r^n(k_{b-2}), \hat{y}_r^n(k_{b-1}|k_{b-2}), y_1^n(b-1), x_1^n(m_{1,b-1})) \in T_\epsilon^{(n)} && \text{(4.16b)}\\
&(x_2^n(1), x_r^n(k_{b-1}), \hat{y}_r^n(k_b|k_{b-1}), y_1^n(b), x_1^n(1)) \in T_\epsilon^{(n)} && \text{(4.16c)}
\end{aligned}
$$

for some indices $(k_1, \dots k_j, \dots k_{b-1}, k_b)$.

To compare its achievable rate region with noisy network coding, we consider the error event in which $m_{2,1}$ and k_b are wrong while all other messages and indices are right. This error event involves the decoded joint distributions from (4.16a) and (4.16c) as follows.

$$
\begin{aligned}
p(x_1)p(x_2)p(x_r)p(y_1, \hat{y}_r|x_1, x_r)\\
p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1, x_2, x_r)
\end{aligned}
$$

By the packing lemma, the probability of the above error event goes to zero as $n \to \infty$ if

$$
R_2 + R_r \le I(X_2; Y_1, \hat{Y}_r|X_1, X_r) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r) \tag{4.17}
$$

Since $R_r > I(\hat{Y}_r; Y_r|X_r)$ by the covering lemma [11], we obtain

$$
\begin{aligned}
R_2 &\le I(X_2; Y_1, \hat{Y}_r|X_1, X_r) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r) - I(\hat{Y}_r; Y_r|X_r)\\
&= I(X_2; Y_1, \hat{Y}_r|X_1, X_r) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1) \qquad \text{(4.18)}
\end{aligned}
$$

which is tighter than the constraint (4.13) in noisy network coding. Again we see the boundary effect at the last block when k_b is wrong.

Simultaneous joint decoding of all messages but ignoring the last compression index

Since in the last block, each source sends a known message, the last compression index k_b brings no new information. Hence we may choose to omit it in the decoding rule to see if the rate region can be improved. Specifically, user 1 now finds a unique tuple $(m_{2,1}, \dots m_{2,j}, \dots m_{2,b-1})$ at the end of block b such that

$$(x_2^n(m_{2,1}), x_r^n(1), \hat{y}_r^n(k_1|1), y_1^n(1), x_1^n(m_{1,1})) \in T_\epsilon^{(n)} \tag{4.19a}$$
$$(x_2^n(m_{2,2}), x_r^n(k_1), \hat{y}_r^n(k_2|k_1), y_1^n(2), x_1^n(m_{1,2})) \in T_\epsilon^{(n)}$$
$$\vdots$$
$$(x_2^n(m_{2,j}), x_r^n(k_{j-1}), \hat{y}_r^n(k_j|k_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)}$$
$$\vdots$$
$$(x_2^n(m_{2,b-1}), x_r^n(k_{b-2}), \hat{y}_r^n(k_{b-1}|k_{b-2}), y_1^n(b-1), x_1^n(m_{1,b-1})) \in T_\epsilon^{(n)} \tag{4.19b}$$
$$(x_2^n(1), x_r^n(k_{b-1}), y_1^n(b), x_1^n(1)) \in T_\epsilon^{(n)} \tag{4.19c}$$

for some indices $(k_1, \dots k_j, \dots k_{b-1})$. Note that (4.19c) is the only step that is different from (4.16c).

To compare its achievable rate region with noisy network coding, we consider the error event in which $m_{2,1}$ and k_{b-1} are wrong while all other messages and indices are right. This error event involves the decoded joint distributions from (4.19a), (4.19b) and (4.19c) as follows.

$$p(x_1)p(x_2)p(x_r)p(y_1, \hat{y}_r|x_1, x_r)$$
$$p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1, x_2, x_r)$$
$$p(x_1)p(x_2)p(x_r)p(y_1|x_1, x_2)$$

By the packing lemma, the probability of the above error event goes to zero as $n \to \infty$ if

$$R_2 + R_r \le I(X_2; Y_1, \hat{Y}_r|X_1, X_r) + I(X_r; Y_1|X_1, X_2) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r).$$

Since $R_r > I(\hat{Y}_r; Y_r|X_r)$, we obtain

$$\begin{aligned} R_2 &\leq I(X_2; Y_1, \hat{Y}_r|X_1, X_r) + I(X_r; Y_1|X_1, X_2) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r) - I(\hat{Y}_r; Y_r|X_r) \\ &= I(X_2; Y_1, \hat{Y}_r|X_1, X_r) + I(X_r; Y_1|X_1, X_2) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1) \end{aligned} \quad (4.20)$$

Although rate constraint (4.20) is loser than (4.18), it is still tighter than the rate constraint of noisy network coding in (4.13) (we can easily check this for the Gaussian TWRC). Therefore, simultaneous decoding of all messages over all blocks still achieves a smaller rate region than noisy network coding.

Remark 12. Joint decoding without explicitly decoding the compression indices removes the constraints on compression rates but still achieves a smaller rate region than noisy network coding for the general multi-source network.

Remark 13. Simultaneous joint decoding of all messages over all blocks still does not overcome the block boundary limitation, caused by the last block in which no new messages but only the compression indices are sent.

Remark 14. In [22], a technique of appending M blocks for transmission of the compression indices only is proposed to reduce the block boundary effect, but it also reduces the achievable rate by a non-vanishing amount.

The above analysis leads us to discuss the effect of message repetition, in which the same message is transmitted over multiple blocks.

4.4.3 Implication of Message Repetition

Message repetition is performed in noisy network coding by sending the same message in multiple, consecutive blocks using independent codebooks. To understand the effect of message repetition, we use the same decoding rule as in the previous section (decoding the message without explicitly decoding the compression indices) and repeat the message twice. Now each user transmits the same message in every two consecutive blocks. We compare the achievable rate region of this scheme with those of no message repetition and of noisy network coding, in which each message is repeated b times and b approaches to infinity.

Again take the two-way relay channel as an illustrative example. Let each user transmit the same message in every two blocks using independent codebooks. Decoding is performed

Block	...	$2j-1$	$2j$	$2j+1$
X_1	...	$x^n_{1,2j-1}(m_{1,j})$	$x^n_{1,2j}(m_{1,j})$	$x^n_{1,2j+1}(m_{1,j+1})$
X_2	...	$x^n_{2,2j-1}(m_{2,j})$	$x^n_{2,2j}(m_{2,j})$	$x^n_{2,2j+1}(m_{2,j+1})$
Y_r	...	$\hat{y}_r(k_{2j-1}\|k_{2j-2})$	$\hat{y}_r(k_{2j}\|k_{2j-1})$	$\hat{y}_r(k_{2j+1}\|k_{2j})$
X_r	...	$x^n_r(k_{2j-2})$	$x^n_r(k_{2j-1})$	$x^n_r(k_{2j})$
Y_1	...	$\hat{m}_{2,j-1}$ for some $(k_{2j-4}, k_{2j-3}, k_{2j-2})$	–	$\hat{m}_{2,j}$ for some $(k_{2j-2}, k_{2j-1}, k_{2j})$
Y_2	...	$\hat{m}_{1,j-1}$ for some $(k_{2j-4}, k_{2j-3}, k_{2j-2})$	–	$\hat{m}_{1,j}$ for some $(k_{2j-2}, k_{2j-1}, k_{2j})$

Table 4.3 Encoding and decoding of CF without binning but with twice message repetition for the two-way relay channel.

at the end of every two blocks based on signals received from the current and previous two blocks. Specifically, the codebook generation, encoding and decoding are as follows (also see Table 4.3). We use a block coding scheme in which each user sends b messages over $2b+1$ blocks of n symbols each.

Codebook generation

At block $l \in \{2j-1, 2j\}$:

- Independently generate 2^{n2R_1} sequences $x^n_{1,l}(m_{1,j}) \sim \prod_{i=1}^n p(x_{1i})$, where $m_{1,j} \in [1 : 2^{n2R_1}]$.
- Independently generate 2^{n2R_2} sequences $x^n_{2,l}(m_{2,j}) \sim \prod_{i=1}^n p(x_{2i})$, where $m_{2,j} \in [1 : 2^{n2R_2}]$.
- Independently generate 2^{nR_r} sequences $x^n_r(k_l) \sim \prod_{i=1}^n p(x_{ri})$, where $k_l \in [1 : 2^{nR_r}]$.
- For each $k_l \in [1 : 2^{nR_r}]$, independently generate 2^{nR_r} sequences $\hat{y}^n_r(k_{l+1}|k_l) \sim \prod_{i=1}^n p(\hat{y}_{ri}|x_{ri}(k_l))$, where $k_{l+1} \in [1 : 2^{nR_r}]$.

Encoding

In blocks $2j-1$ and $2j$, user 1 transmits $x^n_{1,2j-1}(m_{1,j})$ and $x^n_{1,2j}(m_{1,j})$ respectively. User 2 transmits $x^n_{2,2j-1}(m_{2,j})$ and $x^n_{2,2j}(m_{2,j})$.

In block j, the relay, upon receiving $y_r^n(j)$, finds an index k_j such that

$$(\hat{y}_r^n(k_j|k_{j-1}), y_r^n(j), x_r^n(k_{j-1})) \in T_\epsilon^{(n)}.$$

Assume that such k_j is found, the relay sends $x_r^n(k_j)$ in block $j+1$.

Decoding

We discuss the decoding at user 1. User 1 decodes $m_{2,j}$ at the end of block $2j+1$. Specifically, it finds a unique $\hat{m}_{2,j}$ such that

$$\begin{aligned} (x_{2,2j-1}^n(\hat{m}_{2,j}), x_r^n(\hat{k}_{2j-2}), \hat{y}_r^n(\hat{k}_{2j-1}|\hat{k}_{2j-2}), y_1^n(2j-1), x_{1,2j-1}^n(m_{1,j})) &\in T_\epsilon^{(n)} \\ (x_{2,2j}^n(\hat{m}_{2,j}), x_r^n(\hat{k}_{2j-1}), \hat{y}_r^n(\hat{k}_{2j}|\hat{k}_{2j-1}), y_1^n(2j), x_{1,2j}^n(m_{1,j})) &\in T_\epsilon^{(n)} \\ \text{and} \qquad (x_r^n(\hat{k}_{2j}), y_1^n(2j+1), x_{1,2j+1}^n(m_{1,j+1})) &\in T_\epsilon^{(n)} \end{aligned} \tag{4.21}$$

for some $(\hat{k}_{2j-2}, \hat{k}_{2j-1}, \hat{k}_{2j})$.

Corollary 2. *For the two-way relay channel, the following rate region is achievable by CF without binning and without explicitly decoding the compression indices when repeating each message twice:*

$$R_1 \leq I(X_1; Y_2, \hat{Y}_r|X_2, X_r) \tag{4.22a}$$

$$R_1 \leq I(X_1, X_r; Y_2|X_2) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_2) \tag{4.22b}$$

$$R_1 \leq I(X_1, X_r; Y_2|X_2) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_2) + \frac{1}{2}[I(X_r; Y_2|X_2) - I(\hat{Y}; Y_2|X_r)] \tag{4.22c}$$

$$R_2 \leq I(X_2; Y_1, \hat{Y}_r|X_1, X_r) \tag{4.22d}$$

$$R_2 \leq I(X_2, X_r; Y_1|X_1) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1) \tag{4.22e}$$

$$R_2 \leq I(X_2, X_r; Y_1|X_1) - I(\hat{Y}_r; Y_r|X_1, X_2, X_r, Y_1) + \frac{1}{2}[I(X_r; Y_1|X_1) - I(\hat{Y}; Y_1|X_r)] \tag{4.22f}$$

for some $p(x_1)p(x_2)p(x_r)p(y_1, y_2, y_r|x_1, x_2, x_r)p(\hat{y}_r|x_r, y_r)$.

Proof. See Appendix A.6. □

Comparing the above rate region with that of CF without message repetition in (4.15) and noisy network coding in (4.13), we find that the extra rate constraints (4.15c) and

(4.15f) are relaxed by repeating the message twice. The additional term is divided by 2, which comes from the error event $\mathcal{E}^r_{10j}$ in Appendix A.6, in which the message and all compression indices are wrong. This event also corresponds to a boundary event, but since decoding rule (4.21) spans more blocks, the boundary effect is lessen.

Thus if we repeat the messages b times, this additional term will be divided by b. Taking b to infinity as in noisy network coding completely eliminates the additional terms in (4.22c) and (4.22f). Hence the rate region is increasing with the number of times for message repetition. To achieve the largest rate region, the message repetition times need to be infinity. Therefore, noisy network coding has b blocks decoding delay and b is required to approach infinity.

Remark 15. Message repetition brings more correlation between different blocks. It helps lessen the boundary effect when increasing the repetition times.

Remark 16. Message repetition is necessary for achieving the noisy network coding rate region in multi-source networks. Furthermore, the repetition times need to approach infinity.

Remark 17. This result is different from the single-source single-destination result in [22] which shows backward decoding without message repetition achieves the same rate as noisy network coding, albeit at an expense of extending the relay forwarding times infinitely without actually sending a new message. Thus message repetition appears to be essential in a multi-source network.

Remark 18. Recent result in [23] by Hou and Kramer is different from our conclusion where they show that short message NNC can achieve the same rate as NNC even for the multi-source multi-destination network. This is because there are extra $K(K-1)$ blocks (K is the number of nodes which is fixed for a specific network) via which each node tries to convey the compression index in the last block (block b) to all other nodes. In their decoding, each node first reliably recovers the compression index in the last block, and then uses backward decoding. This broadcasting and decoding of the last block compression index turns out ot be crucial in removing the boundary effect and achieving the same rate as repeating the message as in NNC.

4.5 Gaussian Two-Way Relay Channel

In this section, we focus on the Gaussian two-way relay channel. We compare the rate regions of CF without binning, the original CF and noisy network coding. We also derive the analytical condition for when CF without binning achieves the same rate region or sum rate as noisy network coding.

4.5.1 Achievable Rate Region by CF without Binning

In the Gaussian two-way relay channel model, assume

$$\begin{aligned} &X_1 \sim \mathcal{N}(0,P), X_2 \sim \mathcal{N}(0,P), \\ &X_r \sim \mathcal{N}(0,P), \hat{Z} \sim \mathcal{N}(0,\sigma^2), \end{aligned} \tag{4.23}$$

where X_1, X_2, X_r and $\hat{Z}$ are independent, and $\hat{Y}_r = Y_r + \hat{Z}$. Denote

$$\begin{aligned} R_{11}(\sigma^2) &= C\left(g_{21}^2 P + \frac{g_{r1}^2 P}{1+\sigma^2}\right) \\ R_{12}(\sigma^2) &= C(g_{21}^2 P + g_{2r}^2 P) - C(1/\sigma^2) \\ R_{21}(\sigma^2) &= C\left(g_{12}^2 P + \frac{g_{r2}^2 P}{1+\sigma^2}\right) \\ R_{22}(\sigma^2) &= C(g_{12}^2 P + g_{1r}^2 P) - C(1/\sigma^2), \end{aligned} \tag{4.24}$$

where $C(x) = \frac{1}{2}\log(1+x)$. Then we have following rate region using CF without binning.

Theorem 12. *The following rate region is achievable for the Gaussian two-way relay channel using compress-forward without binning:*

$$\begin{aligned} R_1 &\leq \min\{R_{11}(\sigma^2), R_{12}(\sigma^2)\} \\ R_2 &\leq \min\{R_{21}(\sigma^2), R_{22}(\sigma^2)\} \end{aligned} \tag{4.25}$$

for some $\sigma^2 \geq \max\{\sigma_{c1}^2, \sigma_{c2}^2\}$, *where*

$$\begin{aligned} \sigma_{c1}^2 &= (1+g_{21}^2 P)/(g_{2r}^2 P) \\ \sigma_{c2}^2 &= (1+g_{12}^2 P)/(g_{1r}^2 P), \end{aligned} \tag{4.26}$$

and $R_{11}(\sigma^2), R_{12}(\sigma^2), R_{21}(\sigma^2), R_{22}(\sigma^2)$ *are as defined in (4.24).*

Proof. Applying the rate region in Theorem 8 with the signaling in (4.23), we obtain (4.25). □

4.5.2 Rate Region Comparison with the Original CF Scheme

We now compare the rate regions of CF without binning in Theorem 12 and the original CF in [2] for the Gaussian two-way relay channel. We first present the rate region achieved by the original CF scheme.

Corollary 3. [Rankov and Wittneben]. *The following rate region is achievable for the Gaussian two-way relay channel using the original compress-forward scheme:*

$$\begin{aligned} R_1 &\leq R_{11}(\sigma^2) \\ R_2 &\leq R_{21}(\sigma^2) \end{aligned} \tag{4.27}$$

for some $\sigma^2 \geq \sigma_r^2$, *where*

$$\sigma_r^2 = \max\left\{\frac{1+g_{21}^2P+g_{r1}^2P}{\min\{g_{2r}^2, g_{1r}^2\}P}, \frac{1+g_{12}^2P+g_{r2}^2P}{\min\{g_{2r}^2, g_{1r}^2\}P}\right\} \tag{4.28}$$

and $R_{11}(\sigma^2), R_{21}(\sigma^2)$ *are as defined in (4.24).*

The following result shows the condition for which the original CF achieves the same rate region as CF without binning.

Theorem 13. *The original compress-forward scheme achieves the same rate region as compress-forward without binning for the Gaussian TWRC if and only if*

$$\begin{aligned} g_{1r} &= g_{2r} \\ g_{21}^2 + g_{r1}^2 &= g_{12}^2 + g_{r2}^2. \end{aligned} \tag{4.29}$$

Otherwise the rate region by the original compress-forward scheme is smaller.

Remark 19. Condition (4.29) for the Gaussian TWRC is both sufficient and necessary, and hence is stricter than the result for the DMC case in Theorem 10, which is only necessary.

Proof. Note that both $R_{11}(\sigma^2), R_{21}(\sigma^2)$ in (4.24) are non-increasing and $R_{12}(\sigma^2), R_{22}(\sigma^2)$ are non-decreasing. Let σ_{e1}^2 be the intersection between $R_{11}(\sigma^2)$ and $R_{21}(\sigma^2)$ (similar for σ_{e2}^2) as:

$$\begin{aligned} R_{11}(\sigma_{e1}^2) &= R_{12}(\sigma_{e1}^2) \\ R_{21}(\sigma_{e2}^2) &= R_{22}(\sigma_{e2}^2). \end{aligned}$$

Then we can easily show that

$$\begin{aligned} \sigma_{e1}^2 &= (1 + g_{21}^2 P + g_{r1}^2 P)/(g_{2r}^2 P) \\ \sigma_{e2}^2 &= (1 + g_{12}^2 P + g_{r2}^2 P)/(g_{1r}^2 P). \end{aligned} \tag{4.30}$$

Therefore, $\min\{R_{11}(\sigma^2), R_{12}(\sigma^2)\}$ is maximized when $\sigma^2 = \sigma_{e1}^2$, while $\min\{R_{21}(\sigma^2), R_{22}(\sigma^2)\}$ is maximized when $\sigma^2 = \sigma_{e2}^2$. Noting that $\sigma_r^2 \geq \sigma_{e1}^2, \sigma_r^2 \geq \sigma_{e2}^2$ and $\sigma_{e1}^2 \geq \sigma_{c1}^2, \sigma_{e2}^2 \geq \sigma_{c2}^2$, we conclude that the rate region in Theorem 12 is the same as that in Corollary 3 if and only if $\sigma_r^2 = \sigma_{e1}^2 = \sigma_{e2}^2$. From this equality, we obtain the result in Theorem 13. □

4.5.3 Rate Region Comparison with Noisy Network Coding

We next compare the rate regions of CF without binning and noisy network coding [5] for the Gaussian two-way relay channel. We first present the rate region achieved by noisy network coding.

Corollary 4. [Lim, Kim, El Gamal and Chung]. *The following rate region is achievable for the Gaussian two-way relay channel with noisy network coding:*

$$\begin{aligned} R_1 &\leq \min\{R_{11}(\sigma^2), R_{12}(\sigma^2)\} \\ R_2 &\leq \min\{R_{21}(\sigma^2), R_{22}(\sigma^2)\} \end{aligned} \tag{4.31}$$

for some $\sigma^2 > 0$, where $R_{11}(\sigma^2), R_{12}(\sigma^2), R_{21}(\sigma^2), R_{22}(\sigma^2)$ are defined in (4.24).

The following result shows the condition for which CF without binning achieves the same rate region as noisy network coding.

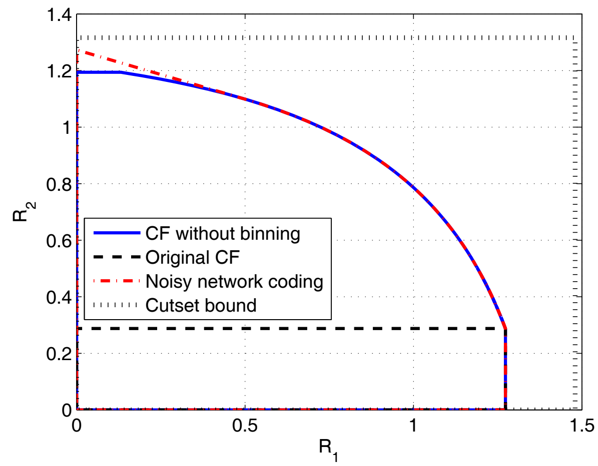

Fig. 4.1 Rate regions for $P = 20, g_{r1} = g_{1r} = 2, g_{r2} = g_{2r} = 0.5, g_{12} = g_{21} = 0.1$.

Theorem 14. *Compress-forward without binning achieves the same rate region as noisy network coding for the Gaussian TWRC if and only if*

$$\sigma_{c1}^2 \leq \sigma_{e2}^2$$
$$\sigma_{c2}^2 \leq \sigma_{e1}^2, \tag{4.32}$$

where $\sigma_{e1}^2, \sigma_{e2}^2$ are defined in (4.30) and $\sigma_{c1}^2, \sigma_{c2}^2$ are defined in (4.26). Otherwise, the rate region by compress-forward without binning is smaller.

Proof. Similar to the proof of Theorem 13, note that both $R_{11}(\sigma^2), R_{21}(\sigma^2)$ are non-increasing and $R_{12}(\sigma^2), R_{22}(\sigma^2)$ are non-decreasing. Also,

$$R_{11}(\sigma_{e1}^2) = R_{12}(\sigma_{e1}^2)$$
$$R_{21}(\sigma_{e2}^2) = R_{22}(\sigma_{e2}^2).$$

Therefore, the constraint in Theorem 12 is redundant if and only if

$$\max\{\sigma_{c1}^2, \sigma_{c2}^2\} \leq \min\{\sigma_{e1}^2, \sigma_{e2}^2\}. \tag{4.33}$$

Since $\sigma_{c1}^2 \leq \sigma_{e1}^2$ and $\sigma_{c2}^2 \leq \sigma_{e2}^2$ always hold, the above condition (4.33) is equivalent to (4.32). $\square$

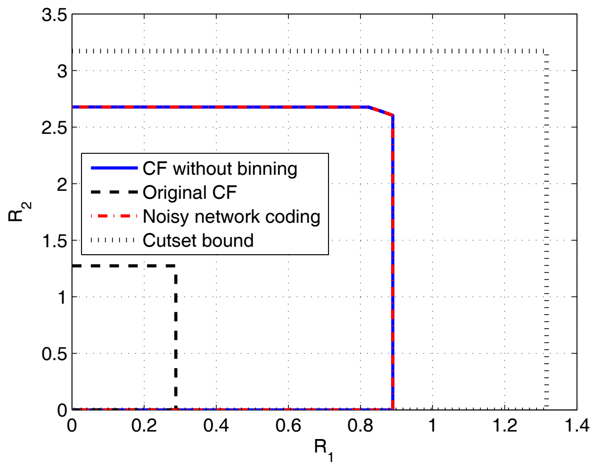

Fig. 4.2 Rate regions for $P = 20, g_{r1} = 0.5, g_{1r} = 2, g_{r2} = 2, g_{2r} = 0.5, g_{12} = g_{21} = 0.1$.

Remark 20. If condition (4.29) holds, then condition (4.32) also holds. Thus when the original CF achieves the same rate region as CF without binning, it also achieves the same rate region as noisy network coding for the Gaussian TWRC.

Figure 4.1 shows an asymmetric channel configuration in which CF without binning achieves a strictly larger rate region than the original CF scheme, but strictly smaller than noisy network coding. Figure 4.2 shows a case that CF without binning achieves the same rate region as noisy network coding and larger than CF with binning.

4.5.4 Sum Rate Comparison with Noisy Network Coding

We notice that in some cases, even though CF without binning achieves smaller rate region than noisy network coding, it still has the same sum rate. Thus we are interested in the channel conditions under which CF without binning and noisy network coding achieve the same sum rate.

Without loss of generality, assume $\sigma_{e1}^2 \geq \sigma_{e2}^2$ as in (4.30). First we analytically derive the optimal σ^2 that maximizes the sum rate of noisy network coding.

Corollary 5. *Let σ_N^2 denotes the optimal σ^2 that maximizes the sum rate of the Gaussian*

TWRC using noisy network coding. Define σ_{N1}^2 and σ_{N2}^2 as follows:

$$\begin{aligned}
&\textit{if } \sigma_{e1}^2 \geq \sigma_g^2 \geq \sigma_{e2}^2,\ \sigma_{N1}^2 = \sigma_g^2;\\
&\textit{if } \sigma_g^2 > \sigma_{e1}^2 \textit{ or } \sigma_g^2 \leq 0,\ \sigma_{N1}^2 = \sigma_{e1}^2;\\
&\textit{if } \sigma_{e2}^2 > \sigma_g^2 > 0,\ \sigma_{N1}^2 = \sigma_{e2}^2;
\end{aligned} \tag{4.34}$$

and

$$\begin{aligned}
&\textit{if } \sigma_{e1}^2 \geq \sigma_g^2 \geq \sigma_{z1}^2,\ \sigma_{N2}^2 = \sigma_g^2;\\
&\textit{if } \sigma_g^2 > \sigma_{e1}^2 \textit{ or } \sigma_g^2 \leq 0,\ \sigma_{N2}^2 = \sigma_{e1}^2;\\
&\textit{if } \sigma_{z1}^2 > \sigma_g^2 > 0,\ \sigma_{N2}^2 = \sigma_{z1}^2,
\end{aligned} \tag{4.35}$$

where

$$\begin{aligned}
\sigma_g^2 &= \frac{g_{r2}^2 P + g_{12}^2 P + 1}{g_{r2}^2 P - g_{12}^2 P - 1},\\
\sigma_{z1}^2 &= \frac{1}{g_{21}^2 P + g_{2r}^2 P}.
\end{aligned}$$

Then,

$$\begin{aligned}
&\textit{if } \sigma_{z1}^2 \leq \sigma_{e2}^2,\ \textit{then } \sigma_N^2 = \sigma_{N1}^2;\\
&\textit{if } \sigma_{z1}^2 > \sigma_{e2}^2,\ \textit{then}\\
&\quad \textit{if } R_{21}(\sigma_{e2}^2) \leq R_{12}(\sigma_{N2}^2) + R_{21}(\sigma_{N2}^2),\ \sigma_N^2 = \sigma_{N2}^2;\\
&\quad \textit{if } R_{21}(\sigma_{e2}^2) > R_{12}(\sigma_{N2}^2) + R_{21}(\sigma_{N2}^2),\ \sigma_N^2 = \sigma_{e2}^2.
\end{aligned} \tag{4.36}$$

Proof. See Appendix A.7.1. □

Based on Theorem 12 and Corollary 5, we now obtain the following result on the conditions for CF without binning to achieve the same sum rate as noisy network coding for the Gaussian TWRC.

Theorem 15. *Compress-forward without binning achieves the same sum rate as noisy network coding for the Gaussian TWRC if and only if $\sigma_{c1}^2 \leq \sigma_N^2$, where σ_{c1}^2 is defined in (4.26) and σ_N^2 is defined in Corollary 5. Otherwise, the sum rate achieved by compress-*

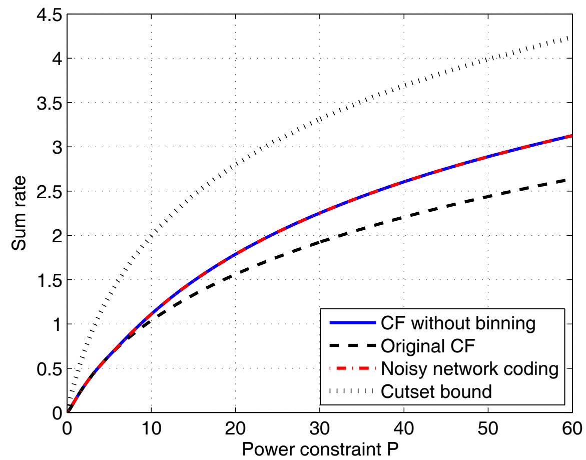

Fig. 4.3 Sum rate for $g_{r1} = g_{1r} = 2, g_{r2} = g_{2r} = 0.5, g_{12} = g_{21} = 0.1$.

forward without binning is smaller.

Proof. According to Corollary 5, the sum rate of noisy network coding is maximized when $\sigma^2 = \sigma_N^2$. According to Theorem 12, for CF without binning, the constraint on σ^2 is $\sigma^2 \geq \sigma_{c1}^2$ (since we assume $\sigma_{c1}^2 \geq \sigma_{c2}^2$). Therefore, compress-forward without binning achieves the same sum rate as noisy network coding if and only if the constraint region $\sigma^2 \geq \sigma_{c1}^2$ contains σ_N^2, which is equivalent to $\sigma_{c1}^2 \leq \sigma_N^2$. □

We can apply the above result to the special case of $g_{r1} = g_{1r}, g_{r2} = g_{2r}, g_{21} = g_{12}$ as follows.

Corollary 6. *If $g_{r1} = g_{1r}, g_{r2} = g_{2r}, g_{21} = g_{12}$, then compress-forward without binning always achieves the same sum rate as noisy network coding.*

Proof. See Appendix A.7.2. □

Figure 4.3 plots the sum rates for the same channel configurations as in Figure 4.1, which shows the sum rates of CF without binning and noisy network coding are the same, even though the rate regions are not.

As confirmed in Figures 4.1, 4.2 and 4.3, CF without binning achieves a larger rate region and sum rate than the original CF scheme in [2] when the channel is asymmetric for the two users. CF without binning achieves the same rate region as noisy network coding when (4.32) is satisfied and achieves the same sum rate for a more relaxed condition. Furthermore, it has less decoding delay which is only 1 instead of b blocks.

Chapter 5

Combined Decode-Forward and Layered Noisy Network Coding

5.1 Problem Statement

We have discussed decode-forward in Chapter 3 and noisy network coding in Chapter 4 for relay channels. Under most channel conditions, these two schemes achieve rate regions which do not include each other. Therefore, we would like to propose a combined scheme which can include both of them as special cases.

In [19], Ramalingam and Wang propose a superposition NNC scheme for restricted relay networks, in which source nodes cannot act as relays, by combining decode-forward and noisy network coding and show some performance improvement over NNC. Their scheme, however, does not include DF relaying rate because of no block Markov encoding. We will first improve this scheme so that it could includes DF relaying rate for the one-way relay channel.

As we have discussed, there are two existing decode-forward schemes [2] [3] for the two-way relay channel. In [2], block Markovity is used at the source to obtain a coherent gain at the cost of reducing power at the relay. In [3], the relay uses its whole power to send their bin index. We could combine these two decode-forward schemes with an improved NNC scheme termed "layered noisy network coding" (LNNC) [18], so that to include the advantages of all schemes.

5.2 Combined DF and NNC Scheme for the One-Way Relay Channel

In this section, we propose a coding scheme combing decode-forward [1] and noisy network coding [5] for the one-way relay channel. The source splits its message into two parts, a common and a private message. The common message is different in each block and is decoded at both the relay and destination as in decode-forward, while the private message is the same for all blocks and is decoded only at the destination as in noisy network coding. The source encodes the common message with block Markovity, and then superimposes the private message on top. The relay decodes the common message at the end of each block and compresses the rest as in NNC. In the next block, it sends a codeword which encodes both the compression index and the decoded common message of the previous block. The destination decodes each common message by forward sliding-window decoding over two consecutive blocks. Then at the end of all blocks, it decodes the private message by simultaneous decoding over all blocks. Our proposed scheme includes both DF relaying and NNC as special cases and outperforms superposition NNC in [19] in that we use block Markov encoding for the common messages, which provides coherency between source and relay and improves the transmission rate.

5.2.1 Coding Scheme and Achievable Rate for the DM One-Way Relay Channel

Theorem 16. [1] *The rate $R = R_{10} + R_{11}$ is achievable for the one-way relay channel by combining decode-forward and noisy network coding*

$$\begin{aligned} R_{10} &\leq \min\{I(Y_r; U|U_r, X_r), I(U, U_r; Y)\} \\ R_{11} &\leq \min\{I(X; Y, \hat{Y}_r|U, U_r, X_r), I(X, X_r; Y|U, U_r) - I(\hat{Y}_r; Y_r|X_r, U, U_r, X, Y)\} \end{aligned} \tag{5.1}$$

for some joint distribution that factors as

$$p(u_r)p(u|u_r)p(x|u, u_r)p(x_r|u_r)p(y, y_r|x, x_r)p(\hat{y}_r|y_r, u, u_r, x_r). \tag{5.2}$$

[1]The result published in [34] contains an error and it should be corrected as in this theorem. See Remark 21 for details.

Proof. We use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each.

1) Codebook generation: Fix a joint distribution as in (5.2). For each block $j \in [1:b]$:

- Independently generate $2^{nR_{10}}$ sequences $u_{r,j}^n(m_{j-1}) \sim \prod_{i=1}^n p(u_{r,i})$, where $m_{j-1} \in [1 : 2^{nR_{10}}]$.
- For each m_{j-1}, independently generate $2^{nR_{10}}$ sequences $u_j^n(m_j|m_{j-1}) \sim \prod_{i=1}^n p(u_i|u_{r,i})$, $m_j \in [1 : 2^{nR_{10}}]$.
- For each (m_{j-1}, m_j), independently generate $2^{nbR_{11}}$ sequences $x_j^n(m|m_j, m_{j-1}) \sim \prod_{i=1}^n p(x_i|\ u_i, u_{r,i})$, $m \in [1 : 2^{nbR_{11}}]$.
- For each m_{j-1}, independently generate $2^{n\hat{R}}$ sequences $x_{r,j}^n(k_{j-1}|m_{j-1}) \sim \prod_{i=1}^n p(x_{r,i}|u_{r,i})$, $k_{j-1} \in [1 : 2^{n\hat{R}}]$.
- For each (m_{j-1}, m_j, k_{j-1}), independently generate $2^{n\hat{R}}$ sequences $\hat{y}_{r,j}^n(k_j|k_{j-1}, m_{j-1}, m_j) \sim \prod_{i=1}^n p(\hat{y}_{r,i}|x_{r,i}, u_{r,i}, u_i)$, $k_j \in [1 : 2^{n\hat{R}}]$.

2) Encoding: In block j, the source sends $x_j^n(m|m_j, m_{j-1})$. Assume that the relay has successfully found compression index k_{j-1} and decoded message m_{j-1} of the previous block, it then sends $x_{r,j}^n(k_{j-1}|m_{j-1})$.

3) Decoding at the relay: At the end of block j, upon receiving $y_{r,j}^n$, the relay finds a $\hat{k}_j$ and a unique $\hat{m}_j$ such that

$$(u_{r,j}^n(m_{j-1}), u_j^n(\hat{m}_j|m_{j-1}), x_{r,j}^n(k_{j-1}|m_{j-1}), \hat{y}_{r,j}^n(\hat{k}_j|k_{j-1}, m_{j-1}, \hat{m}_j), y_{r,j}^n) \in T_\epsilon^{(n)}, \tag{5.3}$$

where $T_\epsilon^{(n)}$ denotes the strong typical set [11]. By the covering lemma and standard analysis, $P_e \to 0$ as $n \to \infty$ if

$$\hat{R} > I(\hat{Y}_r; Y_r|U, U_r, X_r) \tag{5.4}$$

$$\hat{R} + R_{10} \leq I(Y_r; \hat{Y}_r, U|U_r, X_r). \tag{5.5}$$

4) Decoding at the destination: At the end of each block j, the destination finds the unique $\hat{m}_{j-1}$ such that

$$(u_{j-1}^n(\hat{m}_{j-1}|m_{j-2}), u_{r,j-1}^n(m_{j-2}), y_{j-1}^n) \in T_\epsilon^{(n)} \quad \text{and} \quad (u_{r,j}^n(\hat{m}_{j-1}), y_j^n) \in T_\epsilon^{(n)}. \tag{5.6}$$

Following standard analysis, $P_e \to 0$ as $n \to \infty$ if

$$R_{10} \leq I(U;Y|U_r) + I(U_r;Y) = I(U,U_r;Y). \tag{5.7}$$

At the end of block b, it finds the unique $\hat{m}$ such that

$$(u_{r,j}^n(m_{j-1}), u_j^n(m_j|m_{j-1}), x_{r,j}^n(\hat{k}_{j-1}|m_{j-1}), x_j^n(\hat{m}|m_j, m_{j-1}), \hat{y}_{r,j}^n(\hat{k}_j|\hat{k}_{j-1}, m_{j-1}, m_j), y_j^n) \in T_\epsilon^{(n)}$$

for all $j \in [1:b]$ and some vector $\hat{\mathbf{k}}_{\mathbf{j}} \in [1:2^{n\hat{R}}]^b$. As in [5], $P_e \to 0$ as $n \to \infty$ if

$$R_{11} \leq \min\{I_1, I_2 - \hat{R}\}, \quad \text{where} \tag{5.8}$$

$$I_1 = I(X;Y,\hat{Y}_r|U,U_r,X_r)$$

$$I_2 = I(X,X_r;Y|U,U_r) + I(\hat{Y}_r;Y,X|X_r,U,U_r). \tag{5.9}$$

By applying Fourier-Motzkin Elimination to inequalities (5.4)-(5.8), the rate in Theorem 16 is achievable. □

Remark 21. As we have mentioned, there is an error in our previously published paper [34] as follows. When decoding at the destination as in (5.6), we performed the following decoding rules: at the end of each block j, the destination finds the unique $\hat{m}_{j-1}$ such that

$$(u_{j-1}^n(\hat{m}_{j-1}|m_{j-2}), u_{r,j-1}^n(m_{j-2}), x_{r,j-1}^n(k_{j-2}|m_{j-2}), y_{j-1}^n) \in T_\epsilon^{(n)} \quad \text{and} \quad (u_{r,j}^n(\hat{m}_{j-1}), y_j^n) \in T_\epsilon^{(n)}.$$

Following standard analysis, $P_e \to 0$ as $n \to \infty$ if

$$R_{10} \leq I(U;Y|U_r,X_r) + I(U_r;Y). \tag{5.10}$$

However, this is not correct because the destination cannot use the signal of $x_{r,j-1}$ when decoding since it does not know the compression index k_{j-2}. The correct decoding should be as in (5.6) which leads to rate constraint in (5.7).

Remark 22. In relay decoding (5.3), we perform joint decoding of both the message and the compression index. If we use sequential decoding to decode the message first and then to find the compression index, we still get the same rate constraints as in Theorem 16.

Remark 23. In the Numerical Results part, we show that our proposed combined DF and NNC scheme achieves the same rate as the backward decoding strategies in [12]. All of them outperform the original combined DF and CF scheme in [1] under certain channel parameters.

Remark 24. By setting $U_r = X_r, U = X, \hat{Y}_r = 0$, the rate in Theorem 16 reduces to the decode-forward relaying rate [1] as

$$R \leq \min\{I(X;Y_r|X_r), I(X,X_r;Y)\} \tag{5.11}$$

for some $p(x_r)p(x|x_r)p(y,y_r|x,x_r)$. By setting $U = U_r = 0$, it reduces to the NNC rate [5] as

$$R \leq \min\{I(X;Y,\hat{Y}_r|X_r), I(X,X_r;Y) - I(\hat{Y}_r;Y_r|X_r,X,Y)\}$$

for some $p(x)p(x_r)p(y,y_r|x,x_r)p(\hat{y}_r|y_r,x_r)$.

Remark 25. The rate constraints in Theorem 16 are similar to those in superposition NNC (Theorem 1 in [19]), but the code distribution (5.2) is a larger set than that in [19] because of the joint distribution between (x,u,u_r). Hence the achievable rate by the proposed scheme is higher than that in [19]. Specifically, the scheme in [19] does not include the decode-forward relaying rate as in (5.11).

5.2.2 Achievable Rate for the Gaussian One-Way Relay Channel

We now evaluate the achievable rate in Theorem 16 for the Gaussian one-way relay channel as in (2.1).

Corollary 7. *The following rate is achievable for the Gaussian one-way relay channel*

$$R \leq \min\left\{C\left(\frac{g_1^2\beta_1^2}{g_1^2\gamma_1^2+1}\right), C\left(\frac{(g\alpha_1+g_2\alpha_2)^2+g^2\beta_1^2}{g^2\gamma_1^2+g_2^2\beta_2^2+1}\right)\right\}+ C\left(g^2\gamma_1^2+\frac{g_1^2\gamma_1^2\cdot g_2^2\beta_2^2}{g^2\gamma_1^2+g_1^2\gamma_1^2+g_2^2\beta_2^2+1}\right) \tag{5.12}$$

where

$$\alpha_1^2+\beta_1^2+\gamma_1^2 \leq P, \quad \alpha_2^2+\beta_2^2 \leq P. \tag{5.13}$$

To achieve the rate in (5.12), we set

$$\begin{aligned} U &= \alpha_1 S_1 + \beta_1 S_2, \quad & X &= U + \gamma_1 S_3 \\ X_r &= \alpha_2 S_1 + \beta_2 S_4, \quad & \hat{Y}_r &= Y_r + Z' \end{aligned} \tag{5.14}$$

where $S_1, S_2, S_3, S_4 \sim \mathcal{N}(0,1)$ and $Z' \sim \mathcal{N}(0,Q)$ are independent, and the power allocations satisfy constraint (5.13).

5.3 Combined DF and LNNC for the Two-Way Relay Channel

In this section, we propose a combined scheme based on both decode-forward strategies as in [2] [3] and layered noisy network coding [18] for the two-way relay channel. Each user splits its message into three parts: an independent common, a Markov common and a private message. The independent and Markov common messages are encoded differently at the source and are different for each block, both are decoded at both the relay and destination as in decode-forward. The private message is the same for all blocks and is decoded only at the destination as in noisy network coding. Each user encodes the Markov common message with block Markov encoding as in [2], then superimposes the independent common message on top of it without Markovity, and at last superimposes the private message on top of both. The relay decodes the two common messages and compresses the rest into two layers: a common and a refinement layer. In the next block, the relay sends a codeword which encodes the two decoded common messages and two layered compression indices. Then at the end of each block, each user decodes two common messages of the other user by sliding-window decoding over two consecutive blocks. At the end of all blocks, one user uses the information of the common layer to simultaneously decode the private message of the other user, while the other user uses the information of both the common and refinement layers to decode the other user's private message.

5.3.1 Coding Scheme and Achievable Rate for the DM Two-Way Relay Channel

Theorem 17. *Let $\mathcal{R}_1$ denote the set of (R_1, R_2) as follows:*

$$\begin{aligned} R_1 &\leq \min\{I_5, I_{12}\} + \min\{I_5 - I_1, I_{16}\} \\ R_2 &\leq \min\{I_6, I_{14}\} + \min\{I_{17} - I_2, I_{19}\} \\ R_1 + R_2 &\leq \min\{\min\{I_5, I_{12}\} + I_{15} + I_{18} - I_2 + \min\{I_6, I_{14}\}, \\ &\quad I_{10} + I_{15} + I_{18} - I_2, \\ &\quad I_{10} + \min\{I_{15} - I_1, I_{16}\} + \min\{I_{17} - I_2, I_{19}\}\} \\ 2R_1 + R_2 &\leq \min\{I_5, I_{12}\} + I_{15} + I_{18} - I_2 + I_{10} + \min\{I_{15} - I_1, I_{16}\} \end{aligned} \tag{5.15}$$

for some joint distribution

$$\begin{aligned} P^* \triangleq & p(w_1)p(u_1|w_1)p(v_1|w_1, u_1)p(x_1|w_1, u_1, v_1)p(w_2)p(u_2|w_2)p(v_2|w_2, u_2) \\ & p(x_2|w_2, u_2, v_2)p(v_r|w_1, w_2)p(u_r|v_r, w_1, w_2)p(x_r|u_r, v_r, w_1, w_2) \\ & p(\hat{y}_r, \tilde{y}_r|y_r, x_r, u_r, v_r, w_1, w_2, u_1, v_1, u_2, v_2), \end{aligned} \tag{5.16}$$

where I_j are defined in (5.17)-(5.22)*, then $\mathcal{R}_1$ is achievable if user 2 only uses the common layer, while user 1 uses both the common and refinement layers. If the two users exchange decoding layers, they can achieve a corresponding set $\mathcal{R}_2$. By time sharing, the convex hull of $\mathcal{R}_1 \cup \mathcal{R}_2$ is achievable.*

Proof. We use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each.

1) Codebook generation: Fix a joint distribution P^* as in (5.16). Each user $l \in \{1, 2\}$ splits its message into three parts: m_{l0}, m_{l1} and m_{l2}. For each $j \in [1 : b]$ and $l \in \{1, 2\}$

- Independently generate $2^{nR_{l0}}$ sequences $w_{l,j}^n(m_{l0,j-1}) \sim \prod_{i=1}^n p(w_{l,i})$, $m_{l0,j-1} \in [1 : 2^{nR_{l0}}]$.

- For each $m_{l0,j-1}$, independently generate $2^{nR_{l0}}$ sequences $u_{l,j}^n(m_{l0,j}|m_{l0,j-1}) \sim \prod_{i=1}^n p(u_{l,i}|w_{l,i})$, $m_{l0,j} \in [1 : 2^{nR_{l0}}]$.

- For each $m_{l0,j-1}, m_{l0,j}$, independently generate $2^{nR_{l1}}$ sequences $v_{l,j}^n(m_{l1,j}|m_{l0,j}, m_{l0,j-1}) \sim \prod_{i=1}^n p(v_{l,i}|u_{l,i}, w_{l,i})$, $m_{l1,j} \in [1 : 2^{nR_{l1}}]$.

- For each $m_{l0,j-1}, m_{l0,j}, m_{l1,j}$, independently generate $2^{nbR_{l2}}$ sequences $x_{l,j}^n(m_{l2}|m_{l1,j}, m_{l0,j}, m_{l0,j-1}) \sim \prod_{i=1}^n p(x_{l,i}|v_{l,i}, u_{l,i}, w_{l,i})$, $m_{l2} \in [1 : 2^{nbR_{l2}}]$.

- For each $(m_{10,j-1}, m_{20,j-1})$, independently generate $2^{n(R_{11}+R_{21})}$ sequences $v_r^n(K|m_{10,j-1}, m_{20,j-1}) \sim \prod_{i=1}^n p(v_{ri}|w_{1,i}, w_{2,i})$, where $K \in [1 : 2^{n(R_{11}+R_{21})}]$. Map each pair of $(m_{11,j-1}, m_{21,j-1})$ to one K.

- For each vector $\mathbf{m}_{j-1} = (m_{10,j-1}, m_{20,j-1}, m_{11,j-1}, m_{21,j-1})$, independently generate $2^{n\tilde{R}}$ sequences $u_{r,j}^n(t_{j-1}|\mathbf{m}_{j-1}) \sim \prod_{i=1}^n p(u_{r,i}|v_{r,i}, w_{1,i}, w_{2,i})$, $t_{j-1} \in [1 : 2^{n\tilde{R}}]$.

- For each $(t_{j-1}, \mathbf{m}_{j-1})$, independently generate $2^{n\hat{R}}$ sequences $x_{r,j}^n(l_{j-1}|t_{j-1}, \mathbf{m}_{j-1}) \sim \prod_{i=1}^n p(x_{r,i}|u_{r,i}, v_{r,i}, w_{1,i}, w_{2,i})$, $l_{j-1} \in [1 : 2^{n\hat{R}}]$.

- For each $(t_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j)$, independently generate $2^{n\tilde{R}}$ sequences $\tilde{y}_{r,j}^n(t_j|t_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j) \sim \prod_{i=1}^n p(\tilde{y}_{r,i}| u_{r,i}, v_{r,i}, w_{1,i}, w_{2,i}, u_{1,i}, u_{2,i}, v_{1,i}, v_{2,i})$, $t_j \in [1 : 2^{n\tilde{R}}]$.

- For each $(t_j, t_{j-1}, l_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j)$, independently generate $2^{n\hat{R}}$ sequences $\hat{y}_{r,j}^n(l_j|l_{j-1}, t_j, t_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j) \sim \prod_{i=1}^n p(\hat{y}_{r,i}|\tilde{y}_{r,i}, x_{r,i}, u_{r,i}, v_{r,i}, w_{1,i}, w_{2,i}, u_{1,i}, u_{2,i}, v_{1,i}, v_{2,i})$, $t_j \in [1 : 2^{n\tilde{R}}]$.

2) Encoding: In block j, user $l \in \{1, 2\}$ sends $x_{l,j}^n(m_{l2}|m_{l1,j}, m_{l0,j}, m_{l0,j-1})$.

Let $\mathbf{m}_j = (m_{10,j}, m_{20,j}, m_{11,j}, m_{21,j})$. At the end of block j, the relay has decoded $\mathbf{m}_{j-1}, \mathbf{m}_j$. Upon receiving $y_{r,j}^n$, it finds an index pair $(\hat{t}_j, \hat{l}_j)$ such that

$$(\hat{y}_{r,j}^n(\hat{l}_j|l_{j-1}, \hat{t}_j, t_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j), \tilde{y}_{r,j}^n(\hat{t}_j|t_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j), x_{r,j}^n(l_{j-1}|t_{j-1}, \mathbf{m}_{j-1}), \\ u_{r,j}^n(t_{j-1}|\mathbf{m}_{j-1}), w_{1,j}^n, w_{2,j}^n, v_{r,j}^n, u_{1,j}^n, v_{1,j}^n, u_{2,j}^n, v_{2,j}^n, y_{r,j}^n) \in T_\epsilon^{(n)}.$$

According to Lemma 1 in [18], the probability that no such $(\hat{t}_j, \hat{l}_j)$ exists goes to 0 as $n \to \infty$

if

$$\begin{aligned}\tilde{R} &> I(\tilde{Y}_r; X_r, Y_r | U_r, V_r, U_1, V_1, U_2, V_2, W_1, W_2) \triangleq I_1 \\ \tilde{R} + \hat{R} &> I(\tilde{Y}_r; X_r, Y_r | U_r, V_r, U_1, V_1, U_2, V_2, W_1, W_2) + \\ &\quad I(\hat{Y}_r; Y_r | \tilde{Y}_r, X_r, U_r, V_r, U_1, V_1, U_2, V_2, W_1, W_2) \triangleq I_2.\end{aligned} \tag{5.17}$$

The relay then sends $x^n_{r,j+1}(l_j|t_j, \mathbf{m}_j)$ at block $j+1$.

3) Relay decoding: At the end of block j, the relay finds the unique $(\hat{m}_{10,j}, \hat{m}_{20,j}, \hat{m}_{11,j}, \hat{m}_{21,j})$ such that

$$\begin{aligned}(w^n_{1,j}(m_{10,j-1}), u^n_{1,j}(\hat{m}_{10,j}|m_{10,j-1}), v^n_{1,j}(\hat{m}_{11,j}|\hat{m}_{10,j}, m_{10,j-1}), w^n_{2,j}(m_{20,j-1}), u^n_{2,j}(\hat{m}_{20,j}|m_{20,j-1}), \\ v^n_{2,j}(\hat{m}_{21,j}|\hat{m}_{20,j}, m_{20,j-1}), v^n_{r,j}(m_{11,j-1}, m_{21,j-1}|m_{10,j-1}, m_{20,j-1}), y^n_{r,j}) \in T^{(n)}_\epsilon.\end{aligned}$$

As in the multiple access channel, $P_e \to 0$ as $n \to \infty$ if

$$\begin{aligned}R_{11} &\le I(V_1; Y_r | V_r, W_1, U_1, W_2, U_2, V_2) \triangleq I_3 \\ R_{21} &\le I(V_2; Y_r | V_r, W_2, U_2, W_1, U_1, V_1) \triangleq I_4 \\ R_{10} + R_{11} &\le I(U_1, V_1; Y_r | V_r, W_1, W_2, U_2, V_2) \triangleq I_5 \\ R_{20} + R_{21} &\le I(U_2, V_2; Y_r | V_r, W_2, W_1, U_1, V_1) \triangleq I_6 \\ R_{11} + R_{21} &\le I(V_1, V_2; Y_r | V_r, W_1, U_1, W_2, U_2) \triangleq I_7 \\ R_{10} + R_{11} + R_{21} &\le I(U_1, V_1, V_2; Y_r | V_r, W_1, W_2, U_2) \triangleq I_8 \\ R_{20} + R_{11} + R_{21} &\le I(U_2, V_1, V_2; Y_r | V_r, W_1, W_2, U_1) \triangleq I_9 \\ R_{10} + R_{20} + R_{11} + R_{21} &\le \\ &(U_1, V_1, U_2, V_2; Y_r | V_r, W_1, W_2) \triangleq I_{10}.\end{aligned} \tag{5.18}$$

4) User decoding: At the end of block j, user 2 finds the unique $(\hat{m}_{10,j-1}, \hat{m}_{11,j-1})$ such that

$$\begin{aligned}(u^n_{1,j-1}(\hat{m}_{10,j-1}|m_{10,j-2}), v^n_{1,j-1}(\hat{m}_{11,j-1}|\hat{m}_{10,j-1}, m_{10,j-2}), \\ w^n_{1,j-1}, w^n_{2,j-1}, u^n_{2,j-1}, v^n_{2,j-1}, x^n_{2,j-1}, v^n_{r,j-1}, y^n_{2,j-1}) \in T^{(n)}_\epsilon\end{aligned}$$

and $$(w^n_{1,j}(\hat{m}_{10,j-1}), w^n_{2,j}, u^n_{2,j}, v^n_{2,j}, x^n_{2,j}, v^n_{r,j}(\hat{m}_{11,j-1}, m_{21,j-1}|\hat{m}_{10,j-1}, m_{20,j-1}), y^n_{2,j}) \in T^{(n)}_\epsilon.$$

The error probability goes to 0 as $n \to \infty$ if

$$\begin{aligned} R_{11} &\leq I(V_1; Y_2|V_r, W_1, U_1, W_2, U_2, V_2, X_2) + I(V_r; Y_2|W_1, W_2, U_2, V_2, X_2) \triangleq I_{11} \\ R_{10} + R_{11} &\leq I(U_1, V_1; Y_2|V_r, W_1, W_2, U_2, V_2, X_2) + I(W_1, V_r; Y_2|W_2, U_2, V_2, X_2) \\ &= I(W_1, U_1, V_1, V_r; Y_2|W_2, U_2, V_2, X_2) \triangleq I_{12}. \end{aligned} \tag{5.19}$$

Similarly, for vanishing-error decoding at user 1

$$\begin{aligned} R_{21} &\leq I(V_2; Y_1|V_r, W_2, U_2, W_1, U_1, V_1, X_1) + I(V_r; Y_1|W_2, W_1, U_1, V_1, X_1) \triangleq I_{13} \\ R_{20} + R_{21} &\leq I(W_2, U_2, V_2, V_r; Y_1|W_1, U_1, V_1, X_1) \triangleq I_{14}. \end{aligned} \tag{5.20}$$

At the end of last block b, user 2 uses one compression layer to find the unique $\hat{m}_{12}$ such that

$$\begin{gathered} (x_{1,j}^n(\hat{m}_{12}|m_{11,j}, m_{10,j}, m_{10,j-1}), v_{1,j}^n, u_{1,j}^n, w_{1,j}^n, x_{2,j}^n, v_{2,j}^n, u_{2,j}^n, \\ w_{2,j}^n, v_{r,j}^n, u_{r,j}^n(\hat{t}_{j-1}|\mathbf{m}_{j-1}), \tilde{y}_{r,j}^n(\hat{t}_j|\hat{t}_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j), y_{2,j}^n) \in T_\epsilon^{(n)} \end{gathered}$$

for all $j \in [1:b]$ and some vector $\hat{\mathbf{t}}_\mathbf{j} \in [1:2^{n\tilde{R}}]^b$. As in [18], $P_e \to 0$ as $n \to \infty$ if

$$\begin{aligned} R_{12} + \tilde{R} &\leq I(X_1, U_r; Y_2|X_2, W_1, U_1, V_1, W_2, U_2, V_2, V_r) + \\ &\quad I(\tilde{Y}_r; X_1, X_2, Y_2|U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \triangleq I_{15} \\ R_{12} &\leq I(X_1; \tilde{Y}_r, Y_2|X_2, U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \triangleq I_{16} \end{aligned} \tag{5.21}$$

Using both layers, user 1 finds the unique $\hat{m}_{22}$ such that

$$\begin{aligned} &(x_{1,j}^n, v_{1,j}^n, u_{1,j}^n, w_{1,j}^n, x_{2,j}^n(\hat{m}_{22}|m_{21,j}, m_{20,j}, m_{20,j-1}), y_{1,j}^n v_{2,j}^n, u_{2,j}^n, w_{2,j}^n v_{r,j}^n, u_{r,j}^n(\hat{t}_{j-1}|\mathbf{m}_{j-1}), \\ &\quad x_{r,j}^n(\hat{l}_{j-1}|\hat{t}_{j-1}, \mathbf{m}_{j-1}), \tilde{y}_{r,j}^n(\hat{t}_j|\hat{t}_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j), \hat{y}_{r,j}^n(\hat{l}_j|\hat{t}_j, \hat{l}_{j-1}, \hat{t}_{j-1}, \mathbf{m}_{j-1}, \mathbf{m}_j)) \in T_\epsilon^{(n)} \end{aligned}$$

for all $j \in [1:b]$ and some vectors $\hat{\mathbf{t}}_\mathbf{j} \in [1:2^{n\tilde{R}}]^b, \hat{\mathbf{l}}_\mathbf{j} \in [1:2^{n\hat{R}}]^b$. As in [18], $P_e \to 0$ as

$n \to \infty$ if

$$\begin{aligned}
R_{22} + \tilde{R} + \hat{R} &\leq I(X_2, X_r; Y_1 | X_1, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \\
&+ I(\hat{Y}_r; X_1, X_2, Y_1 | \tilde{Y}_r, X_r, U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \\
&+ I(\tilde{Y}_r; X_1, X_2, X_r, Y_2 | U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \triangleq I_{17} \\
R_{22} + \hat{R} &\leq I(X_2, X_r; Y_1, \tilde{Y}_r | X_1, U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \\
&+ I(\hat{Y}_r; X_1, X_2, Y_1 | \tilde{Y}_r, X_r, U_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \triangleq I_{18} \\
R_{22} &\leq I(X_2; \tilde{Y}_r, \hat{Y}_r, Y_1 | X_1, U_r, X_r, W_1, U_1, V_1, W_2, U_2, V_2, V_r) \\
&\triangleq I_{19}.
\end{aligned} \tag{5.22}$$

By applying Fourier-Motzkin Elimination to inequalities (5.17)-(5.22), the rate region in Theorem 17 is achievable. □

Remark 26. By setting $V_1 = W_2 = U_2 = V_2 = X_2 = U_r = V_r = \tilde{Y}_r = 0$, we obtain the rate for the one-way channel in Theorem 16 from the region in Theorem 17.

Remark 27. The proposed combined DF-LNNC scheme includes each schemes in [2, 3, 18] as a special case. Specifically, it reduces to the scheme in [2] by setting $U_1 = X_1, U_2 = X_2, V_r = X_r, V_1 = V_2 = U_r = \hat{Y}_r = \tilde{Y}_r = 0$, to the scheme in [3] by setting $V_1 = X_1, V_2 = X_2, V_r = X_r, W_1 = U_1 = W_2 = U_2 = U_r = \hat{Y}_r = \tilde{Y}_r = 0$, and to the scheme in [18] by setting $W_1 = U_1 = V_1 = W_2 = U_2 = V_2 = V_r = 0$.

Remark 28. In our proposed scheme, the Markov common messages bring a coherent gain between the source and the relay, but they require the relay to split its power for each message because of block Markov superposition coding. For the independent common messages, the relay uses its whole power to send the bin index of them which can then solely represent one message when decoding because of side information on the other message at each destination.

Remark 29. Rate region for the Gaussian TWRC can be obtained by applying Theorem

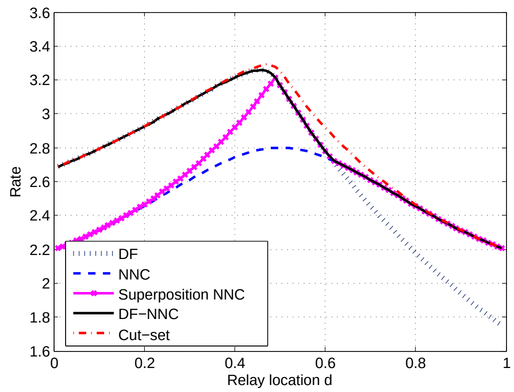

Fig. 5.1 Achievable rate comparison for the one-way relay channel with $P = 10, g_1 = d^{-\gamma/2}, g_2 = (1-d)^{-\gamma/2}, g = 1, \gamma = 3$.

17 with the following signaling:

$$\begin{aligned}
X_1 &= \alpha_1 S_1 + \beta_1 S_2 + \gamma_1 S_3 + \delta_1 S_4 \\
X_2 &= \alpha_2 S_5 + \beta_2 S_6 + \gamma_2 S_7 + \delta_2 S_8 \\
X_r &= \alpha_{31} S_1 + \alpha_{32} S_5 + \gamma_3 S_9 + \beta_3 S_{10} + \delta_3 S_{11} \\
\hat{Y}_r &= Y_r + \hat{Z}_r; \quad \tilde{Y}_r = \hat{Y}_r + \tilde{Z}_r,
\end{aligned} \tag{5.23}$$

where the power allocations satisfy

$$\begin{aligned}
\alpha_1^2 + \beta_1^2 + \gamma_1^2 + \delta_1^2 \leq P, \quad \alpha_2^2 + \beta_2^2 + \gamma_2^2 + \delta_2^2 \leq P, \\
\alpha_{31}^2 + \alpha_{32}^2 + \beta_3^2 + \gamma_3^2 + \delta_3^2 \leq P,
\end{aligned} \tag{5.24}$$

all $S_i \sim \mathcal{N}(0,1)$ and $\hat{Z}_r \sim \mathcal{N}(0,\hat{Q}), \tilde{Z}_r \sim \mathcal{N}(0,\tilde{Q})$ are independent.

5.4 Numerical Results

We numerically compare the performance of the proposed combined schemes with the original DF and NNC. Consider the Gaussian channels as in (2.1) and (2.2). Assume all the nodes are on a straight line. The relay is at a distance d from the source and distance $1-d$ from the destination which makes $g_1 = d^{-\gamma/2}$ and $g_2 = (1-d)^{-\gamma/2}$, where γ is the path loss exponent. Figure 5.1 shows the achievable rate for the one-way relay channel with

d	R_U	R_1	R_2	R_3	R_4	R_5	R_6
0.72	1.9838	1.7061	1.6845	1.7061	1.7061	1.7061	1.7061
0.73	1.9716	1.6881	1.6908	1.6927	1.6984	1.6984	1.6984
0.74	1.9597	1.6703	1.6971	1.6971	1.7012	1.7012	1.7012
0.75	1.9481	1.6529	1.7033	1.7033	1.7052	1.7052	1.7052
0.76	1.9367	1.6538	1.7094	1.7094	1.7099	1.7099	1.7099
0.78	1.9148	1.6022	1.7210	1.7210	1.7210	1.7210	1.7210

Table 5.1 Comparison of achievable rates for $P_1 = 5, P_2 = 1$

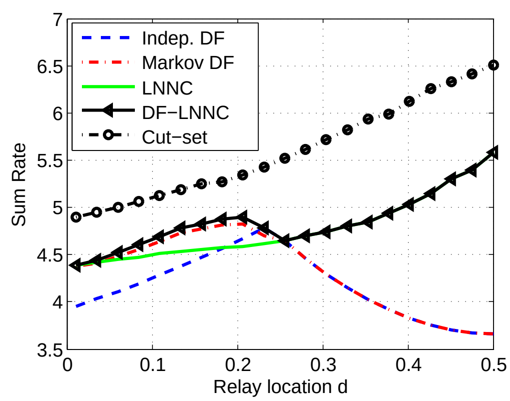

Fig. 5.2 Sum rate for the two-way relay channel with $P = 10, g_{r1} = g_{1r} = d^{-\gamma/2}, g_{r2} = g_{2r} = (1-d)^{-\gamma/2}, g_{12} = g_{21} = 1, \gamma = 3$.

$P = 10, \gamma = 3$. The combined DF-NNC scheme supersedes both the DF and NNC schemes. It can achieve the capacity of the one-way relay channel when the relay is close to either the source or the destination. Table 5.1 shows the rates for the various schemes, i.e., cut-set bound (R_U), decode-forward (R_1), compress-forward (R_2), combined DF and CF of Cover and El Gamal (R_3), the SeqBack decoding in [12] (R_4), the SimBack decoding strategy in [12] (R_5), the combined DF and NNC (R_6). We see that the proposed combined DF and NNC scheme achieves the same rate as the two backward decoding strategies in [12]. All of them outperform the generalized combined DF and CF scheme of Cover and El Gamal in [1] for a certain range of d. Figure 5.2 shows the sum rate for the two-way relay channel with $P = 10, \gamma = 3$. Our proposed scheme achieves larger sum rate than all 3 individual

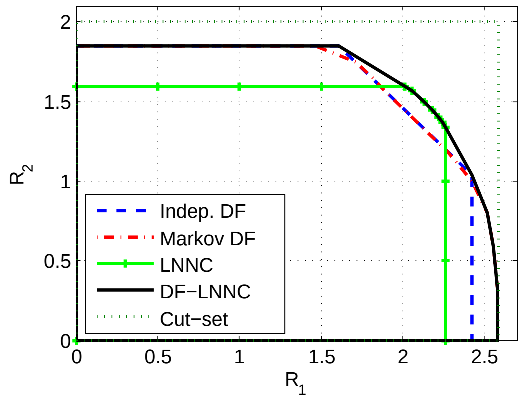

Fig. 5.3 Achievable rate region comparison for the two-way relay channel with $P = 3, g_{r1} = 6, g_{1r} = 2, g_{r2} = 2, g_{2r} = 3, g_{12} = 1, g_{21} = 0.5$.

schemes when the relay is close to either user, while reducing to layered NNC when the relay is close to the middle of the two users. Figure 5.3 shows the achievable rate regions for the Gaussian TWRC using these 4 schemes. The achievable region of our proposed scheme encompasses all 3 individual schemes.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, we have proposed and analyzed several coding schemes for the one-way, two-way relay channels and relay networks. Simulation and analysis show that our proposed bring improvements on existing schemes, such as decode-forward, compress-forward.

In Chapter 3, we have proposed partial decode-forward (PDF) schemes for both full- and half-duplex two-way relay channels. Such schemes have not been considered for the full TWRC (with direct link) before. Each user splits its message into 2 parts and the relay decodes only one. Analysis and simulation both show that when the direct link is stronger than the user-to-relay link for one user but is weaker for the other, partial decode-forward in either full- or half-duplex mode can achieve larger rate region than both pure decode-forward and direct transmission. Thus, unlike in the one-way Gaussian relay channel, partial decode-forward is beneficial in the two-way relay channel. Furthermore, PDF involves only 2-part superposition coding and hence is simple to implement in practice.

In Chapter 4, we have analyzed impacts of the 3 new ideas for compress-forward (CF) in noisy network coding (NNC): no Wyner-Ziv binning, relaxed simultaneous decoding and message repetition. For the one-way relay channel (single-source single-destination), no Wyner-Ziv binning alone without message repetition can achieve the NNC rate at shorter decoding delay. The extra requirement is joint (but not necessarily relaxed) decoding at the destination of both the message and compression index. However, for multi-source multi-destination networks such as the two-way relay channel (TWRC), we find that all 3 techniques together is necessary to achieve the NNC rate region. Under certain channel conditions, CF without Wyner-Ziv binning and without message repetition can achieve the same rate region or sum rate as NNC. Deriving these conditions explicitly for the Gaussian TWRC, we find that the difference in rate regions between CF without binning and NNC is often small. Results also show that these two schemes achieve the same sum rate for a majority of channel configurations. Therefore, CF without binning is more useful in practice because of the short encoding and decoding delay.

In Chapter 5, we have proposed two combined schemes: DF-NNC for the one-way and DF-LNNC for the two-way relay channels. Both schemes perform message splitting, block Markovity encoding, superposition coding and noisy network coding. Each combined scheme encompasses all respective individual schemes (DF and NNC or LNNC) and strictly outperforms superposition NNC in [19]. These are initial results for combining decode-forward and noisy network coding for a multi-source network.

6.2 Future Work

We have proposed combined decode-forward and noisy network coding scheme for the one-way and two-way relay channels. One possible generalization of this work is to extend the combined scheme to general multi-source network where each source also acts as a relay. When generalizing decode-forward to a relay network, a node can decode the message from other nodes after one block or multiple blocks. Different choices for each node brings different achievable rate region. The union of all these achievable rate regions is the final rate region. The difficulty is to find a general way to denote the decoding order and analyze the achievable for fixed decoding order.

We have proposed a partial decode-forward scheme for both full- and half-duplex two-way relay channels. For the full-duplex channel, we are proposing a more general partial decode-forward scheme which combined both block Markov decode-forward and independent decode-forward scheme. We have analyzed the channel conditions for some special cases under which the combined scheme reduces to a single scheme. For the half-duplex channel, we are proposing a 6-phase partial decode-forward scheme. This scheme includes the 4-phase decode-forward scheme as special case. We will have more analysis on this scheme in the future.

Appendix A

Proofs

A.1 Proof of Theorem 4

We prove the first case in Theorem 4. The proofs for other cases are obvious which are omitted here. According to symmetric, we only need to consider the first part in case 1 which is $g_{r1}^2 > g_{21}^2 + \min\{g_{21}^2 g_{r2}^2 P, g_{2r}^2\}, \quad g_{12}^2 > g_{r2}^2$. This is equivalent to any one of the following cases is satisfied:

$$\text{case 1}: \quad C\left(\frac{g_{r1}^2 P}{g_{r2}^2 P + 1}\right) > C(g_{21}^2 P) \tag{A.1}$$

$$\text{case 2}: \quad C(g_{r1}^2 P) > C(g_{21}^2 P + g_{2r}^2 P) \tag{A.2}$$

We now show that for any of the above two cases, partial decode-forward can achieve rates outside the the time-shard region of decode-forward and direct transmission.

Case 1: $C\left(\frac{g_{r1}^2 P}{g_{r2}^2 P+1}\right) > C(g_{21}^2 P)$

We first present an example satisfying the condition of case 1 as shown in Figure A.1. For point X, Let $(R_1(X), R_2(X))$ denote its coordinate. The coordinate of point D is $(R_1(D), R_2(D)) = (C(g_{21}^2 P), C(g_{12}^2 P))$, which is achieved by direct transmission. With decode-forward scheme, we can achieve point A with $R_2(A) = C(g_{r2}^2)$. Since $g_{12}^2 > g_{r2}^2$, we have $R_2(D) > R_2(A)$. Now for the partial decode-forward scheme, set $\alpha = 1, \beta = 0$ in

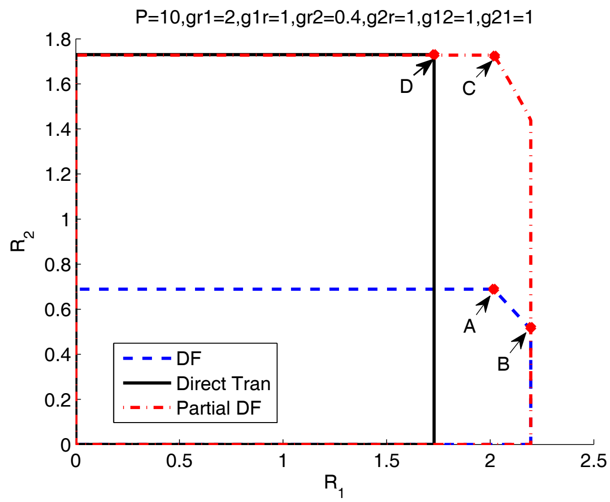

Fig. A.1 Example for case 1.

(3.5), we have

$$\begin{aligned} R_1 &\le \min\left\{C\left(\frac{g_{r1}^2P}{g_{r1}^2P+1}\right), C(g_{21}^2P+g_{2r}^2P)\right\} \\ R_2 &\le \min\{C(g_{12}^2P), C(g_{12}^2P+g_{1r}^2P)\} = C(g_{12}^2P) \\ R_1+R_2 &\le C\left(\frac{g_{r1}^2P}{g_{r2}^2P+1}\right)+C(g_{12}^2P). \end{aligned} \tag{A.3}$$

Therefore, partial decode-forward can achieve point C with coordinate $(R_1(C), R_2(C)) = \left(\min\left\{C\left(\frac{g_{r1}^2P}{g_{r1}^2P+1}\right), C(g_{21}^2P+g_{2r}^2P)\right\}, C(g_{12}^2P)\right)$. Since $C\left(\frac{g_{r1}^2P}{g_{r1}^2P+1}\right) > C(g_{21}^2P)$ and $C(g_{21}^2P+g_{2r}^2P) > C(g_{21}^2P)$, we have $R_1(C) > R_1(D)$. Also note that $R_2(C) = R_2(D) > R_2(A)$. Therefore, point C, achieved by partial decode-forward, is outside the time-shard region of decode-forward and direct transmission.

Case 2: $C(g_{r1}^2P) > C(g_{21}^2P+g_{2r}^2P)$

We also present an example satisfying the condition of case 2 as shown in Figure A.2. We show that in this case partial decode-forward can achieve point E which is outside the time-shared region of decode-forward and direct transmission. Let's assume

$$C\left(\frac{g_{r1}^2P}{g_{r2}^2P+1}\right) < C(g_{21}^2P); \tag{A.4}$$

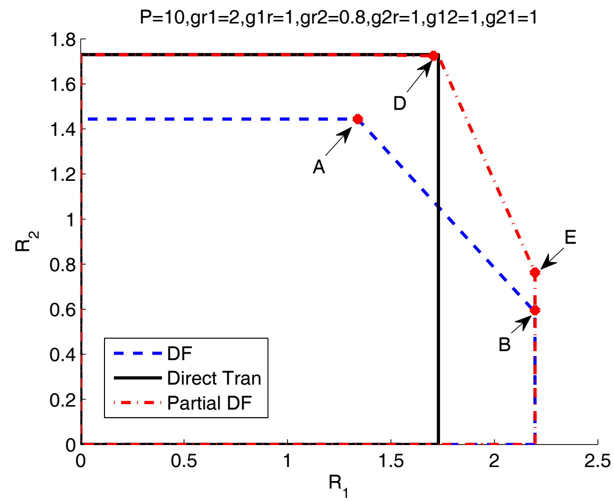

Fig. A.2 Example for case 2.

otherwise, it will come back to case 1. Define the following items:

$$
\begin{aligned}
I_1(\beta) &= C\left(\frac{g_{r1}^2 P}{g_{r2}^2\bar{\beta}P+1}\right) \\
I_2(\beta) &= C\left(\frac{g_{r2}^2 \beta P}{g_{r2}^2\bar{\beta}P+1}\right) + C(g_{12}^2\bar{\beta}P) \\
I_3(\beta) &= C\left(\frac{g_{r1}^2 P + g_{r2}^2\beta P}{g_{r2}^2\bar{\beta}P+1}\right) + C(g_{12}^2\bar{\beta}P)
\end{aligned} \tag{A.5}
$$

We first show some properties about these three functions. Firstly, for $\beta \in [0,1]$, we have

$$
\begin{aligned}
I_2(\beta) &= C\left(\frac{g_{r2}^2 \beta P}{g_{r2}^2\bar{\beta}P+1}\right) + C(g_{12}^2\bar{\beta}P) \\
&= \frac{1}{2}\log\frac{(g_{r2}^2 P+1)(g_{12}^2\bar{\beta}P+1)}{g_{r2}^2\bar{\beta}P+1}
\end{aligned}
$$

Since $g_{12}^2 > g_{r2}^2$, $I_2(\beta)$ is decreasing with the increasing of β. Therefore, $I_2(\beta) \le I_2(0) = C(g_{12}^2 P) < C(g_{12}^2 P + g_{1r}^2 P)$ for $\beta \in [0,1]$. Similarly, $I_3(\beta)$ is decreasing with the increasing of β for $\beta \in [0,1]$. Therefore, $I_3(\beta) \ge I_3(1)$ for $\beta \in [0,1]$.

Secondly, for $\beta \in [0,1]$,

$$\begin{aligned}
I_3(\beta) - I_2(\beta) &= C\left(\frac{g_{r1}^2 P + g_{r2}^2 \beta P}{g_{r2}^2 \bar{\beta} P + 1}\right) - C\left(\frac{g_{r2}^2 \beta P}{g_{r2}^2 \bar{\beta} P + 1}\right) \\
&= \frac{1}{2}\log\frac{g_{r1}^2 P + g_{r2}^2 P + 1}{g_{r2}^2 P + 1} \\
&= C\left(\frac{g_{r1}^2 P}{g_{r2}^2 P + 1}\right) \\
&\overset{(a)}{<} C(g_{21}^2 P) \\
&< C(g_{21}^2 P + g_{2r}^2 P).
\end{aligned}$$

where (a) follows from (A.4).

Thirdly, let $I_1(\beta) \geq C(g_{21}^2 P + g_{2r}^2 P)$, we have

$$\beta \geq 1 + \frac{1}{g_{r2}^2 P}\left(1 - \frac{g_{r1}^2}{g_{21}^2 + g_{2r}^2}\right) \triangleq \beta^* \tag{A.6}$$

We can show that $\beta^* \in (0,1)$.

Now, the rate region in (3.6) achieved by decode-forward can be written as

$$\begin{aligned}
R_1 &\leq \min\left\{I_1(1), C(g_{21}^2 P + g_{2r}^2 P)\right\} = C(g_{21}^2 P + g_{2r}^2 P) \\
R_2 &\leq \min\left\{I_2(1), C(g_{12}^2 P + g_{1r}^2 P)\right\} = I_2(1) \\
R_1 + R_2 &\leq I_3(1).
\end{aligned} \tag{A.7}$$

According to the second property, $I_3(1) < I_2(1) + C(g_{21}^2 P + g_{2r}^2 P)$. Therefore, a corner point B with coordinate $(R_1(B), R_2(B)) = (C(g_{21}^2 P + g_{2r}^2 P), I_3(1) - C(g_{21}^2 P + g_{2r}^2 P))$ can be achieved by decode-forward.

Let $\alpha = 1, \beta = \beta^*$, the rate region in (3.5) achieved by partial decode-forward can be written as

$$\begin{aligned}
R_1 &\leq \min\left\{I_1(\beta^*), C(g_{21}^2 P + g_{2r}^2 P)\right\} = C(g_{21}^2 P + g_{2r}^2 P) \\
R_2 &\leq \min\left\{I_2(\beta^*), C(g_{12}^2 P + g_{1r}^2 P)\right\} = I_2(\beta^*) \\
R_1 + R_2 &\leq I_3(\beta^*).
\end{aligned} \tag{A.8}$$

According to the second property, $I_3(\beta^*) < I_2(\beta^*) + C(g_{21}^2 P + g_{2r}^2 P)$. Therefore, a corner point E with coordinate $(R_1(E), R_2(E)) = (C(g_{21}^2 P + g_{2r}^2 P), I_3(\beta^*) - C(g_{21}^2 P + g_{2r}^2 P))$ can be achieved by partial decode-forward. Comparing point B with point E, we see that $R_1(B) = R_1(E), R_1(B) < R_1(E)$ with first property. Therefore, point E, achieved by the partial decode-forward, is outside the time-shard region of decode-forward and direct transmission.

A.2 Proof of Theorem 6

Assume without loss of generality that $m_j = 1$ and $k_{j-1} = k_j = 1$. First define the following two events:

$$\begin{aligned} \mathcal{E}'_{1j}(k_j) &= \big\{(x_r^n(k_j), y^n(j+1)) \in T_\epsilon^{(n)}\big\} \\ \mathcal{E}'_{2j}(m_j, k_j) &= \big\{(x^n(m_j), x_r^n(1), \hat{y}_r^n(k_j|1), y^n(j)) \in T_\epsilon^{(n)}\big\}. \end{aligned}$$

Then the decoder makes an error only if one or more of the following events occur:

$$\begin{aligned} \mathcal{E}_{1j} =& \big\{(\hat{y}_r^n(k_j|1), y_r^n(j), x_r^n(1)) \notin T_\epsilon^{(n)} \text{ for all } k_j \in [1:2^{nR_r}]\big\} \\ \mathcal{E}_{2j} =& \big\{(x_r^n(1), y^n(j+1)) \notin T_\epsilon^{(n)} \text{ or } (x^n(1), x_r^n(1), \hat{y}_r^n(1|1), y^n(j)) \notin T_\epsilon^{(n)}\big\} \\ \mathcal{E}_{3j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(1, k_j) \text{ for some } k_j \neq 1\big\} \\ \mathcal{E}_{4j} =& \big\{\mathcal{E}'_{1j}(1) \text{ and } \mathcal{E}'_{2j}(m_j, 1) \text{ for some } m_j \neq 1\big\} \\ \mathcal{E}_{5j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(m_j, k_j) \text{ for some } m_j \neq 1, k_j \neq 1\big\}. \end{aligned} \tag{A.9}$$

Thus, the probability of error is bounded as

$$P\{\hat{m}_j \neq 1, \hat{k}_j \neq 1\} \leq P(\mathcal{E}_{1j}) + P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) + P(\mathcal{E}_{3j}) + P(\mathcal{E}_{4j}) + P(\mathcal{E}_{5j}).$$

By the covering lemma [11], $P(\mathcal{E}_{1j}) \to 0$ as $n \to \infty$, if

$$R_r > I(\hat{Y}_r; Y_r | X_r). \tag{A.10}$$

By the conditional typicality lemma [36], the second term $P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) \to 0$ as $n \to \infty$.

For the rest of the error events, the decoded joint distribution for each event is as follows.

$$\begin{aligned}
\mathcal{E}'_{1j}(k_j) &: p(x_r)p(y) \\
\mathcal{E}'_{2j}(1,k_j) &: p(x)p(x_r)p(\hat{y}_r|x_r)p(y|x,x_r) \\
\mathcal{E}'_{2j}(m_j,1) &: p(x)p(x_r)p(y,\hat{y}_r|x_r) \\
\mathcal{E}'_{2j}(m_j,k_j) &: p(x)p(x_r)p(\hat{y}_r|x_r)p(y|x_r),
\end{aligned}$$

where $m_j \neq 1, k_j \neq 1$. Using standard joint typicality analysis with the above decoded joint distribution, we can obtain a bound on each error event as follows.

Consider $\mathcal{E}_{3j}$:

$$\begin{aligned}
P(\mathcal{E}_{3j}) &= P(\cup_{k_j \neq 1}(\mathcal{E}'_{1j}(k_j) \cap \mathcal{E}'_{2j}(1,k_j))) \\
&\leq \sum_{k_j \neq 1} P(\mathcal{E}'_{1j}(k_j)) \times P(\mathcal{E}'_{2j}(1,k_j)).
\end{aligned}$$

Note that if $k_j \neq 1$, then

$$\begin{aligned}
&P(\mathcal{E}'_{1j}(k_j)) \leq 2^{-n(I(X_r;Y)-\delta(\epsilon))}; \\
&P(\mathcal{E}'_{2j}(1,k_j)) \\
&= \sum_{(x,x_r,\hat{y}_r,y)\in T_\epsilon^{(n)}} p(x)p(x_r)p(\hat{y}_r|x_r)p(y|x,x_r) \\
&\leq 2^{n(H(X,X_r,\hat{Y}_r,Y)-H(X)-H(X_r)-H(\hat{Y}_r|X_r)-H(Y|X,X_r)-3\delta(\epsilon))} \\
&= 2^{n(H(\hat{Y}_r,Y|X,X_r)-H(\hat{Y}_r|X_r)-H(Y|X,X_r)-3\delta(\epsilon))} \\
&= 2^{n(H(\hat{Y}_r|Y,X,X_r)-H(\hat{Y}_r|X_r)-3\delta(\epsilon))} \\
&= 2^{-n(I(\hat{Y}_r;X,Y|X_r)-3\delta(\epsilon))}.
\end{aligned}$$

Therefore

$$P(\mathcal{E}_{3j}) \leq 2^{nR_r} \cdot 2^{-n(I(X_r;Y)-\delta(\epsilon))} \cdot 2^{-n(I(\hat{Y}_r;X,Y|X_r)-3\delta(\epsilon))}$$

which tends to zero as $n \to \infty$ if

$$R_r \leq I(X_r;Y) + I(\hat{Y}_r;X,Y|X_r). \tag{A.11}$$

Next consider $\mathcal{E}_{4j}$:

$$\begin{aligned} P(\mathcal{E}_{4j}) &= P(\cup_{m_j\neq 1}(\mathcal{E}'_{1j}(1)\cap\mathcal{E}'_{2j}(m_j,1))) \\ &\leq \sum_{m_j\neq 1} P(\mathcal{E}'_{2j}(m_j,1)). \end{aligned}$$

Note that if $m_j \neq 1$, then

$$\begin{aligned} &P(\mathcal{E}'_{2j}(m_j,1)) \\ &= \sum_{(x,x_r,\hat{y}_r,y)\in T_\epsilon^{(n)}} p(x)p(x_r)p(y,\hat{y}_r|x_r) \\ &\leq 2^{n(H(X,X_r,\hat{Y}_r,Y)-H(X)-H(X_r)-H(Y,\hat{Y}_r|X_r)-2\delta(\epsilon))} \\ &= 2^{n(H(\hat{Y}_r,Y|X,X_r)-H(Y,\hat{Y}_r|X_r)-2\delta(\epsilon))} \\ &= 2^{-n(I(X;Y,\hat{Y}_r|X_r)-2\delta(\epsilon))}. \end{aligned}$$

Therefore

$$P(\mathcal{E}_{4j}) \leq 2^{nR}\cdot 2^{-n(I(X;Y,\hat{Y}_r|X_r)-2\delta(\epsilon))}$$

which tends to zero as $n\to\infty$ if

$$R \leq I(X;Y,\hat{Y}_r|X_r). \tag{A.12}$$

Now consider $\mathcal{E}_{5j}$:

$$\begin{aligned} P(\mathcal{E}_{5j}) &= P(\cup_{m_j\neq 1}\cup_{k_j\neq 1}(\mathcal{E}'_{1j}(k_j)\cap\mathcal{E}'_{2j}(m_j,k_j))) \\ &\leq \sum_{m_j\neq 1}\sum_{k_j\neq 1} P(\mathcal{E}'_{1j}(k_j))\times P(\mathcal{E}'_{2j}(m_j,k_j)). \end{aligned}$$

Note that if $m_j \neq 1$ and $k_j \neq 1$, then

$$
\begin{aligned}
& P(\mathcal{E}'_{1j}(k_j)) \leq 2^{-n(I(X_r;Y)-\delta(\epsilon))}; \\
& P(\mathcal{E}'_{2j}(m_j, k_j)) \\
& = \sum_{(x,x_r,\hat{y}_r,y)\in T_\epsilon^{(n)}} p(x)p(x_r)p(\hat{y}_r|x_r)p(y|x_r) \\
& \leq 2^{n(H(X,X_r,\hat{Y}_r,Y)-H(X)-H(X_r)-H(\hat{Y}_r|X_r)-H(Y|X_r)-3\delta(\epsilon))} \\
& = 2^{n(H(\hat{Y}_r,Y|X,X_r)-H(\hat{Y}_r|X_r)-H(Y|X_r)-3\delta(\epsilon))} \\
& = 2^{n(H(Y|X,X_r)+H(\hat{Y}_r|Y,X,X_r)-H(\hat{Y}_r|X_r)-H(Y|X_r)-3\delta(\epsilon))} \\
& = 2^{-n(I(X;Y|X_r)+I(\hat{Y}_r;X,Y|X_r)-3\delta(\epsilon))}.
\end{aligned}
$$

Therefore

$$
P(\mathcal{E}_{5j}) \leq \; 2^{nR}\cdot 2^{nR_r}\cdot 2^{-n(I(X_r;Y)-\delta(\epsilon))}\cdot 2^{-n(I(X;Y|X_r)+I(\hat{Y}_r;X,Y|X_r)-3\delta(\epsilon))}
$$

which tends to zero as $n \to \infty$ if

$$
\begin{aligned}
R + R_r &\leq I(X_r;Y) + I(X;Y|X_r) + I(\hat{Y}_r;X,Y|X_r) \\
&= I(X,X_r;Y) + I(\hat{Y}_r;X,Y|X_r). \qquad \text{(A.13)}
\end{aligned}
$$

Combining the bounds (A.10) and (A.13), we have

$$
\begin{aligned}
R &\leq I(X,X_r;Y) + I(\hat{Y}_r;X,Y|X_r) - I(\hat{Y}_r;Y_r|X_r) \\
&= I(X,X_r;Y) + I(\hat{Y}_r;X,Y|X_r) - I(\hat{Y}_r;Y_r,X,Y|X_r) \\
&= I(X,X_r;Y) - I(\hat{Y}_r;Y_r|X,X_r,Y). \qquad \text{(A.14)}
\end{aligned}
$$

Combining (A.10), (A.11) and (A.12), (A.14), we obtain the result of Theorem 6.

A.3 Proof of Theorem 8

Assume without loss of generality that $m_{1,j} = m_{1,j+1} = m_{2,j} = 1$ and $k_{j-1} = k_j = 1$. First define the following two events:

$$\mathcal{E}'_{1j}(k_j) = \big\{(x_r^n(k_j), y_1^n(j+1), x_1^n(1)) \in T_\epsilon^{(n)}\big\}$$
$$\mathcal{E}'_{2j}(m_{2,j}, k_j) = \big\{(x_2^n(m_{2,j}), x_r^n(1), \hat{y}_r^n(k_j|1), y_1^n(j), x_1^n(1)) \in T_\epsilon^{(n)}\big\}.$$

Then the decoder makes an error only if one or more of the following events occur:

$$\begin{aligned}
\mathcal{E}_{1j} =& \big\{(\hat{y}_r^n(k_j|1), y_r^n(j), x_r^n(1)) \notin T_\epsilon^{(n)} \text{ for all } k_j \in [1:2^{nR_r}]\big\} \\
\mathcal{E}_{2j} =& \big\{(x_r^n(1), y_1^n(j+1), x_1^n(1)) \notin T_\epsilon^{(n)} \text{ or } (x_2^n(1), x_r^n(1), \hat{y}_r^n(1|1), y_1^n(j), x_1^n(1)) \notin T_\epsilon^{(n)}\big\} \\
\mathcal{E}_{3j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(1, k_j) \text{ for some } k_j \neq 1\big\} \qquad \text{(A.15)} \\
\mathcal{E}_{4j} =& \big\{\mathcal{E}'_{1j}(1) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, 1) \text{ for some } m_{2,j} \neq 1\big\} \\
\mathcal{E}_{5j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, k_j) \text{ for some } m_{2,j} \neq 1, k_j \neq 1\big\}.
\end{aligned}$$

Thus, the probability of error is bounded as

$$P\{\hat{m}_{2,j} \neq 1, \hat{k}_j \neq 1\} \leq P(\mathcal{E}_{1j}) + P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) + P(\mathcal{E}_{3j}) + P(\mathcal{E}_{4j}) + P(\mathcal{E}_{5j}).$$

By the covering lemma, $P(\mathcal{E}_{1j}) \to 0$ as $n \to \infty$, if

$$R_r > I(\hat{Y}_r; Y_r | X_r). \qquad \text{(A.16)}$$

By the conditional typicality lemma, the second term $P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) \to 0$ as $n \to \infty$.

For the rest of the error events, the decoded joint distribution for each event is as follows.

$$\begin{aligned}
\mathcal{E}'_{1j}(k_j) &: p(x_1)p(x_r)p(y_1|x_1) \\
\mathcal{E}'_{2j}(1, k_j) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_2, x_r, x_1) \\
\mathcal{E}'_{2j}(m_{2,j}, 1) &: p(x_1)p(x_2)p(x_r)p(y_1, \hat{y}_r|x_r, x_1) \\
\mathcal{E}'_{2j}(m_{2,j}, k_j) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_r, x_1), \qquad \text{(A.17)}
\end{aligned}$$

where $m_{2,j} \neq 1, k_j \neq 1$. Using standard joint typicality analysis with the above decoded joint distribution, we can obtain a bound on each error event as follows.

Consider $\mathcal{E}_{3j}$:

$$\begin{aligned} P(\mathcal{E}_{3j}) &= P(\cup_{k_j \neq 1}(\mathcal{E}'_{1j}(k_j) \cap \mathcal{E}'_{2j}(1, k_j))) \\ &\leq \sum_{k_j \neq 1} P(\mathcal{E}'_{1j}(k_j)) \times P(\mathcal{E}'_{2j}(1, k_j)). \end{aligned}$$

Note that if $k_j \neq 1$, then

$$\begin{aligned} & P(\mathcal{E}'_{1j}(k_j)) \leq 2^{-n(I(X_r;Y_1|X_1)-\delta(\epsilon))}; \\ & P(\mathcal{E}'_{2j}(1, k_j)) \\ & = \sum_{(x_1,x_2,x_r,\hat{y}_r,y_1)\in T_\epsilon^{(n)}} p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_2,x_r,x_1) \\ & \leq 2^{n(H(X_1,X_2,X_r,\hat{Y}_r,Y_1)-H(X_1)-H(X_2)-H(X_r)-H(\hat{Y}_r|X_r)-H(Y_1|X_2,X_r,X_1)-4\delta(\epsilon))} \\ & = 2^{n(H(\hat{Y}_r,Y_1|X_2,X_r,X_1)-H(\hat{Y}_r|X_r)-H(Y_1|X_2,X_r,X_1)-4\delta(\epsilon))} \\ & = 2^{n(H(\hat{Y}_r|Y_1,X_2,X_r,X_1)-H(\hat{Y}_r|X_r)-4\delta(\epsilon))} \\ & = 2^{-n(I(\hat{Y}_r;X_1,X_2,Y_1|X_r)-4\delta(\epsilon))}. \end{aligned}$$

Therefore

$$P(\mathcal{E}_{3j}) \leq 2^{nR_r} \cdot 2^{-n(I(X_r;Y_1|X_1)-\delta(\epsilon))} \cdot 2^{-n(I(\hat{Y}_r;X_1,X_2,Y_1|X_r)-4\delta(\epsilon))}$$

which tends to zero as $n \to \infty$ if

$$R_r \leq I(X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r). \tag{A.18}$$

Next consider $\mathcal{E}_{4j}$:

$$\begin{aligned} P(\mathcal{E}_{4j}) &= P(\cup_{m_{2,j} \neq 1}(\mathcal{E}'_{1j}(1) \cap \mathcal{E}'_{2j}(m_{2,j}, 1))) \\ &\leq \sum_{m_{2,j} \neq 1} P(\mathcal{E}'_{2j}(m_{2,j}, 1)). \end{aligned}$$

Note that if $m_{2,j} \neq 1$, then

$$\begin{aligned}
&P(\mathcal{E}'_{2j}(m_{2,j}, 1)) \\
&= \sum_{(x_1,x_2,x_r,\hat{y}_r,y_1)\in T_\epsilon^{(n)}} p(x_1)p(x_2)p(x_r)p(y_1,\hat{y}_r|x_r,x_1) \\
&\leq 2^{n(H(X_1,X_2,X_r,\hat{Y}_r,Y_1)-H(X_1)-H(X_2)-H(X_r)-H(Y_1,\hat{Y}_r|X_r,X_1)-3\delta(\epsilon))} \\
&= 2^{n(H(\hat{Y}_r,Y_1|X_2,X_r,X_1)-H(Y_1,\hat{Y}_r|X_r,X_1)-3\delta(\epsilon))} \\
&= 2^{-n(I(X_2;Y_1,\hat{Y}_r|X_r,X_1)-3\delta(\epsilon))}.
\end{aligned}$$

Therefore

$$P(\mathcal{E}_{4j}) \leq 2^{nR_2} \cdot 2^{-n(I(X_2;Y_1,\hat{Y}_r|X_r,X_1)-3\delta(\epsilon))}$$

which tends to zero as $n \to \infty$ if

$$R_2 \leq I(X_2; Y_1, \hat{Y}_r|X_r, X_1). \tag{A.19}$$

Now consider $\mathcal{E}_{5j}$:

$$\begin{aligned}
P(\mathcal{E}_{5j}) &= P(\cup_{m_{2,j}\neq 1} \cup_{k_j\neq 1} (\mathcal{E}'_{1j}(k_j) \cap \mathcal{E}'_{2j}(m_{2,j}, k_j))) \\
&\leq \sum_{m_{2,j}\neq 1} \sum_{k_j\neq 1} P(\mathcal{E}'_{1j}(k_j)) \times P(\mathcal{E}'_{2j}(m_{2,j}, k_j)).
\end{aligned}$$

Note that if $m_{2,j} \neq 1$ and $k_j \neq 1$, then

$$\begin{aligned}
&P(\mathcal{E}'_{1j}(k_j)) \leq 2^{-n(I(X_r;Y_1|X_1)-\delta(\epsilon))}; \\
&P(\mathcal{E}'_{2j}(m_{2,j}, k_j)) \\
&= \sum_{(x_1,x_2,x_r,\hat{y}_r,y_1)\in T_\epsilon^{(n)}} p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_r,x_1) \\
&\leq 2^{n(H(X_1,X_2,X_r,\hat{Y}_r,Y_1)-H(X_1)-H(X_2)-H(X_r)-H(\hat{Y}_r|X_r)-H(Y_1|X_r,X_1)-4\delta(\epsilon))} \\
&= 2^{n(H(\hat{Y}_r,Y_1|X_2,X_r,X_1)-H(\hat{Y}_r|X_r)-H(Y_1|X_r,X_1)-4\delta(\epsilon))} \\
&= 2^{n(H(Y_1|X_2,X_r,X_1)+H(\hat{Y}_r|Y_1,X_2,X_r,X_1)-H(\hat{Y}_r|X_r)-H(Y_1|X_r,X_1)-4\delta(\epsilon))} \\
&= 2^{-n(I(X_2;Y_1|X_1,X_r)+I(\hat{Y}_r;X_1,X_2,Y_1|X_r)-4\delta(\epsilon))}.
\end{aligned}$$

Therefore

$$P(\mathcal{E}_{5j}) \leq 2^{nR_2} \cdot 2^{nR_r} \cdot 2^{-n(I(X_r;Y_1|X_1)-\delta(\epsilon))} \cdot 2^{-n(I(X_2;Y_1|X_1,X_r)+I(\hat{Y}_r;X_1,X_2,Y_1|X_r)-4\delta(\epsilon))}$$

which tends to zero as $n \to \infty$ if

$$\begin{aligned} R_2 + R_r &\leq I(X_r;Y_1|X_1) + I(X_2;Y_1|X_1,X_r) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r) \\ &= I(X_2,X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r). \end{aligned} \tag{A.20}$$

Combining the bounds (A.16) and (A.20), we have

$$\begin{aligned} R_2 &\leq I(X_2,X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r) - I(\hat{Y}_r;Y_r|X_r) \\ &= I(X_2,X_r;Y_1|X_1) - I(\hat{Y}_r;Y_r|X_1,X_2,X_r,Y_1). \end{aligned} \tag{A.21}$$

Combining the bounds (A.19) and (A.21), we obtain the rate constraint on R_2 in Theorem 8. Similar for R_1. From (A.16) and (A.18), we obtain constraint (4.8).

A.4 Proof of Theorem 9

We use a block coding scheme in which each user sends $b-1$ messages over b blocks of n symbols each.

Codebook generation

Fix $p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r,y_r)$. We randomly and independently generate a codebook for each block $j \in [1:b]$

- Independently generate 2^{nR_1} sequences $x_1^n(m_{1,j}) \sim \prod_{i=1}^n p(x_{1i})$, where $m_{1,j} \in [1 : 2^{nR_1}]$.
- Independently generate 2^{nR_2} sequences $x_2^n(m_{2,j}) \sim \prod_{i=1}^n p(x_{2i})$, where $m_{2,j} \in [1 : 2^{nR_2}]$.
- Independently generate 2^{nR_r} sequences $x_r^n(q_{j-1}) \sim \prod_{i=1}^n p(x_{ri})$, where $q_{j-1} \in [1 : 2^{nR_r}]$.

- For each $q_{j-1} \in [1 : 2^{nR_r}]$, independently generate $2^{n(R_r+R'_r)}$ sequences $\hat{y}_r^n(q_j, r_j|q_{j-1}) \sim \prod_{i=1}^n p(\hat{y}_{ri}|x_{ri}(q_{j-1}))$. Throw them into 2^{nR_r} bins, where $q_j \in [1 : 2^{nR_r}]$ denotes the bin index and $r_j \in [1 : 2^{nR'_r}]$ denotes the relative index within a bin.

Encoding

User 1 and user 2 transmits $x_1^n(m_{1,j})$ and $x_2^n(m_{2,j})$ in block j separately. The relay, upon receiving $y_r^n(j)$, finds an index pair (q_j, r_j) such that

$$(\hat{y}_r^n(q_j, r_j|q_{j-1}), y_r^n(j), x_r^n(q_{j-1})) \in T_\epsilon^{(n)}.$$

Assume that such (q_j, r_j) is found, the relay sends $x_r^n(q_j)$ in block $j+1$. By the covering lemma, the probability that there is no such (q_j, r_j) tends to 0 as $n \to \infty$ if

$$R_r + R'_r > I(\hat{Y}_r; Y_r|X_r). \tag{A.22}$$

Decoding

At the end of block j, user 1 determines the unique $\hat{q}_{j-1}$ such that

$$(x_r^n(\hat{q}_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)}.$$

Similar for user 2. Both succeed with high probability if

$$R_r \leq \min\{I(X_r; Y_1|X_1), I(X_r; Y_2|X_2)\}. \tag{A.23}$$

Then user 1 uses $y_1^n(j-1)$ to determine the unique $\hat{r}_{j-1}$ such that

$$(\hat{y}_r^n(\hat{q}_{j-1}, \hat{r}_{j-1}|\hat{q}_{j-2}), y_1^n(j-1), x_r^n(\hat{q}_{j-2}), x_1^n(m_{1,j-1})) \in T_\epsilon^{(n)}.$$

Similar for user 2. Both succeed with high probability if

$$R'_r \leq \min\{I(\hat{Y}_r; X_1, Y_1|X_r), I(\hat{Y}_r; X_2, Y_2|X_r)\}. \tag{A.24}$$

Finally, user 1 uses both $y_1^n(j-1)$ and $\hat{y}_r^n(j-1)$ to determine the unique $\hat{m}_{2,j-1}$ such that

$$(x_2^n(\hat{m}_{2,j-1}), \hat{y}_r^n(\hat{q}_{j-1}, \hat{r}_{j-1}|\hat{q}_{j-2}), y_1^n(j-1), x_r^n(\hat{q}_{j-2}), x_1^n(m_{1,j-1})) \in T_\epsilon^{(n)}.$$

Similar for user 2. Both succeed with high probability if

$$\begin{aligned} R_1 &\le I(X_1; Y_2, \hat{Y}_r | X_2, X_r) \\ R_2 &\le I(X_2; Y_1, \hat{Y}_r | X_1, X_r). \end{aligned} \tag{A.25}$$

The constraint (4.10) comes from combining (A.22), (A.23) and (A.24).

Remark 30. We note an error in the proof in [2]. In [2], when user 1 determines the unique $\hat{r}_{j-1}$, it is stated that it succeeds with high probability if

$$R_r' \le I(\hat{Y}_r; Y_1 | X_1, X_r), \tag{A.26}$$

which corresponds to step (A.24) in our analysis. However, this is incorrect since the decoded joint distribution of this error event is $p(x_1)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1,x_r)$. Therefore, the error probability can be bounded as

$$\begin{aligned} P(\mathcal{E}) &= \sum_{(x_1,x_r,\hat{y}_r,y_1)\in T_\epsilon^{(n)}} p(x_1)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1,x_r) \\ &\le 2^{n(H(X_1,X_r,\hat{Y}_r,Y_1)-H(X_1)-H(X_r)-H(\hat{Y}_r|X_r)-H(Y_1|X_1,X_r)-3\delta(\epsilon))} \\ &= 2^{-n(I(\hat{Y}_r;X_1,Y_1|X_r)-3\delta(\epsilon))}, \end{aligned}$$

which tends to zero as as $n \to \infty$ if (A.24) is satisfied instead of (A.26).

A.5 Proof of Corollary 1

The codebook generation and encoding is the same as that in Theorem 8. The decoding rule is changed as follows: in block $j+1$, user 1 finds a unique $\hat{m}_{2,j}$ such that

$$\begin{aligned} &(x_2^n(\hat{m}_{2,j}), x_r^n(\hat{k}_{j-1}), \hat{y}_r^n(\hat{k}_j|\hat{k}_{j-1}), y_1^n(j), x_1^n(m_{1,j})) \in T_\epsilon^{(n)} \\ \text{and} \quad &(x_r^n(\hat{k}_j), y_1^n(j+1), x_1^n(m_{1,j+1})) \in T_\epsilon^{(n)} \end{aligned}$$

for some pair of $(\hat{k}_{j-1}, \hat{k}_j)$. Next we present the error analysis.

Assume without loss of generality that $m_{1,j} = m_{1,j+1} = m_{2,j} = 1$ and $k_{j-1} = k_j = 1$. First define the following two events:

$$\mathcal{E}'_{1j}(k_j) = \big\{(x_r^n(k_j), y_1^n(j+1), x_1^n(1)) \in T_\epsilon^{(n)}\big\}$$
$$\mathcal{E}'_{2j}(m_{2,j}, k_j, k_{j-1}) = \big\{(x_2^n(m_{2,j}), x_r^n(k_{j-1}), \hat{y}_r^n(k_j|k_{j-1}), y_1^n(j), x_1^n(1)) \in T_\epsilon^{(n)}\big\}.$$

Then the decoder makes an error only if one or more of the following events occur:

$$\begin{aligned}
\mathcal{E}_{1j} =& \big\{(\hat{y}_r^n(k_j|1), y_r^n(j), x_r^n(1)) \notin T_\epsilon^{(n)} \text{ for all } k_j \in [1:2^{nR_r}]\big\}\\
\mathcal{E}_{2j} =& \big\{(x_r^n(1), y_1^n(j+1), x_1^n(1)) \notin T_\epsilon^{(n)} \text{ or } (x_2^n(1), x_r^n(1), \hat{y}_r^n(1|1), y_1^n(j), x_1^n(1)) \notin T_\epsilon^{(n)}\big\}\\
\tilde{\mathcal{E}}_{4j} =& \big\{\mathcal{E}'_{1j}(1) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, 1, 1) \text{ for some } m_{2,j} \neq 1\big\}\\
\tilde{\mathcal{E}}_{5j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, k_j, 1) \text{ for some } m_{2,j} \neq 1, k_j \neq 1\big\}\\
\mathcal{E}_{6j} =& \big\{\mathcal{E}'_{1j}(1) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, 1, k_{j-1}) \text{ for some } m_{2,j} \neq 1, k_{j-1} \neq 1\big\}\\
\mathcal{E}_{7j} =& \big\{\mathcal{E}'_{1j}(k_j) \text{ and } \mathcal{E}'_{2j}(m_{2,j}, k_j, k_{j-1}) \text{ for some } m_{2,j} \neq 1, k_j \neq 1, k_{j-1} \neq 1\big\},
\end{aligned}$$

where $\mathcal{E}_{1j}, \mathcal{E}_{2j}$ are the same as in (A.15) in the proof of Theorem 8 (Appendix A.3); $\mathcal{E}_{3j}$ in (A.15) is not an error here. $\tilde{\mathcal{E}}_{4j}$ and $\tilde{\mathcal{E}}_{5j}$ are similar to $\mathcal{E}_{4j}$ and $\mathcal{E}_{5j}$ in (A.15). $\mathcal{E}_{6j}$ and $\mathcal{E}_{7j}$ are new error events.

The probability of error is bounded as

$$P\{\hat{m}_{2,j} \neq 1\} \leq P(\mathcal{E}_{1j}) + P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) + P(\tilde{\mathcal{E}}_{4j}) + P(\tilde{\mathcal{E}}_{5j}) + P(\mathcal{E}_{6j}) + P(\mathcal{E}_{7j}).$$

Similar to the proof of Theorem 8 in Appendix A.3, by the covering lemma, $P(\mathcal{E}_{1j}) \to 0$ as $n \to \infty$, if

$$R_r > I(\hat{Y}_r; Y_r|X_r). \tag{A.27}$$

By the conditional typicality lemma, the second term $P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) \to 0$ as $n \to \infty$.

By the packing lemma, $P(\tilde{\mathcal{E}}_{4j}) \to 0$ as $n \to \infty$ if

$$R_2 \leq I(X_2; Y_1, \hat{Y}_r|X_r, X_1). \tag{A.28}$$

Similarly, $P(\tilde{\mathcal{E}}_{5j}) \to 0$ as $n \to \infty$ if

$$R_2 + R_r \leq I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r). \tag{A.29}$$

For the new error events $\mathcal{E}_{6j}$ and $\mathcal{E}_{7j}$, the decoded joint distributions are as follows.

$$\begin{aligned} \mathcal{E}'_{2j}(m_{2,j}, 1, k_{j-1}) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1) \\ \mathcal{E}'_{2j}(m_{2,j}, k_j, k_{j-1}) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1), \end{aligned} \tag{A.30}$$

where $m_{2,j} \neq 1, k_j \neq 1, k_{j-1} \neq 1$. Using standard joint typicality analysis, we can obtain a bound on each error event as follows. $P(\mathcal{E}_{6j}) \to 0$ as $n \to \infty$ if

$$R_2 + R_r \leq I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r). \tag{A.31}$$

$P(\mathcal{E}_{7j}) \to 0$ as $n \to \infty$ if

$$R_2 + 2R_r \leq I(X_r; Y_1|X_1) + I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r). \tag{A.32}$$

Combining the above inequalities, we obtain the rate region in Corollary 1. Compared to the rate region of noisy network coding in (4.13), the only new constraint is (A.32), which occurs when both compression indices are wrong in addition to a wrong message.

A.6 Proof of Corollary 2

Assume without loss of generality that $m_{1,j} = m_{1,j+1} = m_{2,j} = 1$ and $k_{2j-2} = k_{2j-1} = k_{2j} = 1$. First define the following three events:

$$\begin{aligned} \mathcal{E}'_{1j}(k_{2j}) &= \left\{(x_r^n(k_{2j}), y_1^n(2j+1), x_{1,2j+1}^n(1)) \in T_\epsilon^{(n)}\right\} \\ \mathcal{E}'_{2j}(m_{2,j}, k_{2j}, k_{2j-1}) &= \left\{(x_{2,2j}^n(m_{2,j}), x_r^n(k_{2j-1}), \hat{y}_r^n(k_{2j}|k_{2j-1}), y_1^n(2j), x_{1,2j}^n(1)) \in T_\epsilon^{(n)}\right\} \\ \mathcal{E}'_{3j}(m_{2,j}, k_{2j-1}, k_{2j-2}) &= \left\{(x_{2,2j-1}^n(m_{2,j}), x_r^n(k_{2j-2}), \hat{y}_r^n(k_{2j-1}|k_{2j-2}), y_1^n(2j-1), x_{1,2j-1}^n(1)) \in T_\epsilon^{(n)}\right\}. \end{aligned}$$

Then the decoder makes an error only if one or more of the following events occur:

$$\begin{aligned}
\mathcal{E}_{1j} &= \big\{(\hat{y}_r^n(k_j|1), y_r^n(j), x_r^n(1)) \notin T_\epsilon^{(n)} \text{ for all } k_j \in [1:2^{nR_r}]\big\}\\
\mathcal{E}_{2j} &= \big\{(x_r^n(1), y_1^n(2j+1), x_{1,2j+1}^n(1)) \notin T_\epsilon^{(n)} \text{ or}\\
&\quad (x_{2,2j}^n(1), x_r^n(1), \hat{y}_r^n(1|1), y_1^n(2j), x_{1,2j}^n(1)) \notin T_\epsilon^{(n)} \text{ or}\\
&\quad (x_{2,2j-1}^n(1), x_r^n(1), \hat{y}_r^n(1|1), y_1^n(2j-1), x_{1,2j-1}^n(1)) \notin T_\epsilon^{(n)}\big\}\\
\mathcal{E}_{3j}^r &= \big\{\mathcal{E}'_{1j}(1), \mathcal{E}'_{2j}(m_{2,j},1,1) \text{ and } \mathcal{E}'_{3j}(m_{2,j},1,1) \text{ for some } m_{2,j} \neq 1\big\}\\
\mathcal{E}_{4j}^r &= \big\{\mathcal{E}'_{1j}(1), \mathcal{E}'_{2j}(m_{2,j},1,1) \text{ and } \mathcal{E}'_{3j}(m_{2,j},1,k_{2j-2}) \text{ for some } m_{2,j} \neq 1, k_{2j-2} \neq 1\big\}\\
\mathcal{E}_{5j}^r &= \big\{\mathcal{E}'_{1j}(1), \mathcal{E}'_{2j}(m_{2,j},1,k_{2j-1}) \text{ and } \mathcal{E}'_{3j}(m_{2,j},k_{2j-1},1) \text{ for some } m_{2,j} \neq 1, k_{2j-1} \neq 1\big\}\\
\mathcal{E}_{6j}^r &= \big\{\mathcal{E}'_{1j}(k_{2j}), \mathcal{E}'_{2j}(m_{2,j},k_{2j},1) \text{ and } \mathcal{E}'_{3j}(m_{2,j},1,1) \text{ for some } m_{2,j} \neq 1, k_{2j} \neq 1\big\}\\
\mathcal{E}_{7j}^r &= \big\{\mathcal{E}'_{1j}(k_{2j}), \mathcal{E}'_{2j}(m_{2,j},k_{2j},k_{2j-1}) \text{ and } \mathcal{E}'_{3j}(m_{2,j},k_{2j-1},1) \text{ for some } m_{2,j} \neq 1, k_{2j} \neq 1,\\
&\quad k_{2j-1} \neq 1\big\}\\
\mathcal{E}_{8j}^r &= \big\{\mathcal{E}'_{1j}(k_{2j}), \mathcal{E}'_{2j}(m_{2,j},k_{2j},1) \text{ and } \mathcal{E}'_{3j}(m_{2,j},1,k_{2j-2}) \text{ for some } m_{2,j} \neq 1, k_{2j} \neq 1, k_{2j-2} \neq 1\big\}\\
\mathcal{E}_{9j}^r &= \big\{\mathcal{E}'_{1j}(1), \mathcal{E}'_{2j}(m_{2,j},1,k_{2j-1}) \text{ and } \mathcal{E}'_{3j}(m_{2,j},k_{2j-1},k_{2j-2}) \text{ for some } m_{2,j} \neq 1, k_{2j-1} \neq 1,\\
&\quad k_{2j-2} \neq 1\big\}\\
\mathcal{E}_{10j}^r &= \big\{\mathcal{E}'_{1j}(k_{2j}), \mathcal{E}'_{2j}(m_{2,j},k_{2j},k_{2j-1}) \text{ and } \mathcal{E}'_{3j}(m_{2,j},k_{2j-1},k_{2j-2})\\
&\quad \text{for some } m_{2,j} \neq 1, k_{2j} \neq 1, k_{2j-1} \neq 1, k_{2j-2} \neq 1\big\}
\end{aligned}$$

Similar to the proof of Theorem 8 in Appendix A.3, by the covering lemma, $P(\mathcal{E}_{1j}) \to 0$ as $n \to \infty$, if

$$R_r > I(\hat{Y}_r; Y_r|X_r). \tag{A.33}$$

By the conditional typicality lemma, the second term $P(\mathcal{E}_{2j} \cap \mathcal{E}_{1j}^c) \to 0$ as $n \to \infty$.

Let symbol "$*$" represent the wrong message or compression index. For $l \in \{2j, 3j\}$, similar to (A.17) and (A.30), the joint decoded distributions for the rest of the error events are as follows.

$$\begin{aligned}
\mathcal{E}'_{1j}(*) &: p(x_1)p(x_r)p(y_1|x_1)\\
\mathcal{E}'_{l}(*,1,1) &: p(x_1)p(x_2)p(x_r)p(y_1,\hat{y}_r|x_r,x_1)\\
\mathcal{E}'_{l}(*,1,*) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1)\\
\mathcal{E}'_{l}(*,*,1) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_r,x_1)\\
\mathcal{E}'_{l}(*,*,*) &: p(x_1)p(x_2)p(x_r)p(\hat{y}_r|x_r)p(y_1|x_1).
\end{aligned}$$

Using standard joint typicality analysis, we can obtain a bound on each error event as follows.
$P(\mathcal{E}^r_{3j}) \to 0$ as $n \to \infty$ if

$$2R_2 \leq 2I(X_2;Y_1,\hat{Y}_r|X_r,X_1). \tag{A.34}$$

$P(\mathcal{E}^r_{4j}) \to 0$ as $n \to \infty$ if

$$2R_2 + R_r \leq \ I(X_2;Y_1,\hat{Y}_r|X_r,X_1) + I(X_2,X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r). \tag{A.35}$$

$P(\mathcal{E}^r_{5j}) \to 0$ as $n \to \infty$ if

$$2R_2 + R_r \leq \ I(X_2,X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r) + I(X_2;Y_1|X_1,X_r) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r). \tag{A.36}$$

$P(\mathcal{E}^r_{6j}) \to 0$ as $n \to \infty$ if

$$2R_2 + R_r \leq \ I(X_2;Y_1,\hat{Y}_r|X_r,X_1) + I(X_2,X_r;Y_1|X_1) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r). \tag{A.37}$$

$P(\mathcal{E}^r_{7j}) \to 0$ as $n \to \infty$ if

$$\begin{aligned}
2R_2 + 2R_r \leq{}& I(X_r;Y_1|X_1) + I(X_2,X_r;Y_1|X_1)\\
&+ I(\hat{Y}_r;X_1,X_2,Y_1|X_r) + I(X_2;Y_1|X_1,X_r) + I(\hat{Y}_r;X_1,X_2,Y_1|X_r).
\end{aligned} \tag{A.38}$$

$P(\mathcal{E}_{8j}^r) \to 0$ as $n \to \infty$ if

$$2R_2 + 2R_r \leq 2[I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r)]. \tag{A.39}$$

$P(\mathcal{E}_{9j}^r) \to 0$ as $n \to \infty$ if

$$2R_2 + 2R_r \leq 2[I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r)]. \tag{A.40}$$

$P(\mathcal{E}_{10j}^r) \to 0$ as $n \to \infty$ if

$$2R_2 + 3R_r \leq \ I(X_r; Y_1|X_1) + 2[I(X_2, X_r; Y_1|X_1) + I(\hat{Y}_r; X_1, X_2, Y_1|X_r)]. \tag{A.41}$$

Combining the above inequalities, we obtain the rate region in Corollary 2. The extra rate constraints (4.22c) and (4.22f) come only from combing (A.41) with (A.33). Again we see the boundary effect when the message and all compression indices are wrong.

A.7 Proofs of Corollary 5 and Corollary 6

A.7.1 Proof of Corollary 5

Our objective is to find the optimal σ^2 which maximizes the sum rate

$$s(\sigma^2) \triangleq \max\left\{0, \min\{R_{11}(\sigma^2), R_{12}(\sigma^2)\}\right\} + \max\left\{0, \min\{R_{21}(\sigma^2), R_{22}(\sigma^2)\}\right\}.$$

Recall that we have assumed $\sigma_{e1}^2 \geq \sigma_{e2}^2$, where $\sigma_{e1}^2, \sigma_{e2}^2$ are defined in (4.30). Note that both $R_{11}(\sigma^2), R_{21}(\sigma^2)$ are non-increasing and $R_{12}(\sigma^2), R_{22}(\sigma^2)$ are non-decreasing. Also,

$$\begin{aligned} R_{11}(\sigma_{e1}^2) &= R_{12}(\sigma_{e1}^2) \\ R_{21}(\sigma_{e2}^2) &= R_{22}(\sigma_{e2}^2). \end{aligned}$$

Therefore, the problem of maximizing $s(\sigma^2)$ is equivalent to

$$\begin{aligned} \max \quad & q(\sigma^2) \triangleq \max\{0, R_{12}(\sigma^2)\} + R_{21}(\sigma^2) \\ \text{s.t.} \quad & \sigma^2 \leq \sigma_{e1}^2, \\ & \sigma^2 \geq \sigma_{e2}^2. \end{aligned} \tag{A.42}$$

Note that $R_{12}(\sigma_{z1}^2) = 0$ and $R_{12}(\sigma^2) < 0$ for $\sigma^2 \in (0, \sigma_{z1}^2)$, $R_{12}(\sigma^2) > 0$ for $\sigma^2 \in (\sigma_{z1}^2, \infty)$. Therefore, the optimization problem in (A.42) can be derived into the following two cases:

Case 1: $\sigma_{z1}^2 \leq \sigma_{e2}^2$

In this case, the objective function $q(\sigma^2)$ can be simplified to $q(\sigma^2) = R_{12}(\sigma^2) + R_{21}(\sigma^2)$, which is continuously differentiable for $\sigma^2 \in [\sigma_{e2}^2, \sigma_{e1}^2]$. Thus the optimization problem is equivalent to

$$\begin{aligned} \min \quad & f(\sigma^2) = -R_{12}(\sigma^2) - R_{21}(\sigma^2) \\ \text{s.t.} \quad & c_1(\sigma^2) = \sigma_{e1}^2 - \sigma^2 \geq 0, \\ & c_2(\sigma^2) = \sigma^2 - \sigma_{e2}^2 \geq 0. \end{aligned}$$

Form the Lagrangian as $\mathcal{L}(\sigma^2, \lambda_1, \lambda_2) = f(\sigma^2) - \lambda_1 c_1(\sigma^2) - \lambda_2 c_2(\sigma^2)$. With the KKT conditions, we have:

$$\begin{aligned} \nabla_{\sigma^2} \mathcal{L}(\sigma_{N1}^2, \lambda_1, \lambda_2) = 0, \\ \sigma_{e1}^2 - \sigma_{N1}^2 \geq 0, \\ \sigma_{N1}^2 - \sigma_{e2}^2 \geq 0, \\ \lambda_1, \lambda_2 \geq 0, \\ \lambda_1 c_1(\sigma_{N1}^2) = 0, \\ \lambda_2 c_2(\sigma_{N1}^2) = 0. \end{aligned} \tag{A.43}$$

By solving the above conditions, the optimal σ_{N1}^2 for this case is characterized as in (4.34).

Case 2: $\sigma_{z1}^2 > \sigma_{e2}^2$

In this case, the objective function $q(\sigma^2)$ is no longer continuously differentiable for $\sigma^2 \in [\sigma_{e2}^2, \sigma_{e1}^2]$. To solve this problem, we divide this interval into two parts. When $\sigma^2 \in [\sigma_{e2}^2, \sigma_{z1}^2)$, the objective function is simplified to $q(\sigma^2) = R_{21}(\sigma^2)$ and is maximized at $\sigma^2 = \sigma_{e2}^2$. When $\sigma^2 \in [\sigma_{z1}^2, \sigma_{e1}^2]$, the objective function is simplified to $q(\sigma^2) = R_{12}(\sigma^2) + R_{21}(\sigma^2)$ and is continuously differentiable for $\sigma^2 \in [\sigma_{z1}^2, \sigma_{e1}^2]$. With the KKT conditions, the optimal σ^2 for this case is characterized as σ_{N2}^2 in (4.35). Combining the two intervals, the optimal σ_N^2 for this case can be characterized as in (4.36).

A.7.2 Proof of Corollary 6

From Corollary 5 and Theorem 15, we have:
If $\sigma_{z1}^2 \leq \sigma_{e2}^2$, then the two schemes achieve the same sum rate if and only if at least one of the following three conditions holds:

$$\sigma_{c1}^2 \leq \sigma_{e2}^2 \tag{A.44}$$

$$0 < \sigma_{c1}^2 \leq \sigma_g^2 \tag{A.45}$$

$$\sigma_g^2 \leq 0 \tag{A.46}$$

For the channel configuration of $g_{r1} = g_{1r}, g_{r2} = g_{2r}, g_{21} = g_{12}$, we can show that either (A.45) or (A.46) will hold.

If $\sigma_{z1}^2 > \sigma_{e2}^2$, we can show that $R_{21}(\sigma_{e2}^2) < R_{12}(\sigma_{e1}^2) + R_{21}(\sigma_{e1}^2)$ always holds. Therefore, according to (4.35) and (4.36), the optimal σ_N^2 can be characterized as:

$$\text{if } \sigma_{e1}^2 \geq \sigma_g^2 \geq 0,\ \sigma_N^2 = \sigma_g^2;$$
$$\text{if } \sigma_g^2 > \sigma_{e1}^2 \text{ or } \sigma_g^2 < 0,\ \sigma_N^2 = \sigma_{e1}^2.$$

Since either (A.45) or (A.46) holds and $\sigma_{c1}^2 \leq \sigma_{e1}^2$, we obtain $\sigma_{c1}^2 \leq \sigma_N^2$. According to Theorem 15, the two schemes then achieve the same sum rate.

References

[1] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Info. Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[2] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Int'l Symp. on Info. Theory (ISIT)*. IEEE, 2006, pp. 1668–1672.

[3] L. Xie, "Network coding and random binning for multi-user channels," in *10th Canadian Workshop on Info. Theory (CWIT)*. IEEE, 2007, pp. 85–88.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. on Info. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[5] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. on Info. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[6] F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. on Info. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[7] S. J. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. on Info. Theory*, vol. 54, no. 11, pp. 5235 –5241, Nov. 2008.

[8] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols," *IEEE Trans. on Info. Theory*, vol. 57, no. 10, pp. 6405–6418, Oct. 2011.

[9] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Info. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[10] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. on Info. Theory*, vol. 52, no. 4, pp. 1545–1561, April 2006.

[11] A. E. Gamal and Y.-H. Kim, *Network information theory.* Cambridge University Press, 2011.

[12] H.-F. Chong, M. Motani, and H. K. Garg, "Generalized backward decoding strategies for the relay channel," *IEEE Trans. on Info. Theory*, vol. 53, no. 1, pp. 394–401, Jan. 2007.

[13] T. M. Cover and Y.-H. Kim, "Capacity of a class of deterministic relay channels," in *IEEE Int'l Symp. on Info. Theory (ISIT)*, June 2007, pp. 591–595.

[14] Y.-H. Kim, "Coding techniques for primitive relay channels," in *45th Annual Allerton Conf. on Comm., Control, and Computing*, Sept. 2007.

[15] P. Razaghi and W. Yu, "Universal relaying for the interference channel," in *Info. Theory and Applications Workshop (ITA)*, Feb. 2010.

[16] L. Zhou and W. Yu, "Incremental relaying for the Gaussian interference channel with a degraded broadcasting relay," in *49th Annual Allerton Conf. on Comm., Control, and Computing*, Sept. 2011.

[17] G. Kramer and J. Hou, "Short-message quantize-forward network coding," in *8th Int'l Workshop on Multi-Carrier Systems Solutions (MC-SS)*, May 2011, pp. 1–3.

[18] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Layered noisy network coding," in *IEEE Wireless Net. Coding Conf. (WiNC)*, June 2010.

[19] N. Ramalingam and Z. Wang, "Superposition noisy network coding," in *IEEE Int'l Symp. on Info. Theory (ISIT)*, Aug. 2011.

[20] T. Kim, M. Skoglund, and G. Caire, "Quantifying the loss of compress-forward relaying without Wyner-Ziv coding," *IEEE Trans. on Info. Theory*, vol. 55, no. 4, pp. 1529–1533, April 2009.

[21] X. Wu and L.-L. Xie, "On the optimality of successive decoding in compress-and-forward relay schemes," in *48th Annual Allerton Conf. on Comm., Control, and Computing*, Oct. 2010.

[22] ——, "On the optimal compressions in the compress-and-forward relay schemes," *Arxiv preprint arXiv: 1009.5959v3*, 2011.

[23] J. Hou and G. Kramer, "Short message noisy network coding for multiple sources," in *IEEE Int'l Symp. on Info. Theory (ISIT)*, July 2012.

[24] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *Arxiv preprint arXiv:0908.2119*, 2009.

[25] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. on Info. Theory*, vol. 51, no. 10, pp. 3401–3416, 2005.

[26] K. Narayanan, M. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for bi-directional relaying," in *45th Annual Allerton Conf. on Comm., Control, and Computing*, 2007.

[27] W. Nam, S. Chung, and Y. Lee, "Capacity bounds for two-way relay channels," in *IEEE Int'l Zurich Seminar on Communications.* IEEE, 2008, pp. 144–147.

[28] ——, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Trans. on Info. Theory*, vol. 56, no. 11, pp. 5488–5494, 2010.

[29] Y. Song and N. Devroye, "List decoding for nested lattices and applications to relay channels," in *48th Annual Allerton Conf. on Comm., Control, and Computing.* IEEE, 2010, pp. 1038–1045.

[30] P. Zhong and M. Vu, "Decode-forward and compute-forward coding schemes for the two-way relay channel," in *IEEE Info. Theory Workshop (ITW)*, Oct. 2011.

[31] ——, "Partial decode-forward coding schemes for the Gaussian two-way relay channel," in *IEEE Int'l Conf. on Comm. (ICC)*, June. 2012 to appear.

[32] ——, "Compress-forward without Wyner-Ziv binning for the one-way and two-way relay channels," in *49th Annual Allerton Conf. on Comm., Control, and Computing*, Sept. 2011.

[33] P. Zhong, A. A. A. Haija, and M. Vu, "On compress-forward without Wyner-Ziv binning for relay networks," *submitted to IEEE Trans. on Info. Theory. Arxiv preprint arXiv:1111.2837*, 2011.

[34] P. Zhong and M. Vu, "Combined decode-forward and layered noisy network coding schemes for relay channels," in *IEEE Int'l Symp. on Info. Theory (ISIT)*, July 2012.

[35] F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. on Info. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley, 2006.