

An Information Theoretic Framework for Two-Way Relay Networks

by

Jonathan Ponniah

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2008

© Jonathan Ponniah 2008

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

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

Abstract

We propose an information theoretic framework for scheduling the transmissions in a two-way multi-hop network. First we investigate some long standing open problems that were encountered during the course of the research. To illustrate their difficulty, we describe some failed attempts at resolving them. We then introduce the two-way one-relay channel. It turns out that the achievable rate region of this network has a nice interpretation, especially when viewed in the context of the open problems examined earlier. Motivated by this observation, we attempt to extend this achievable region to the two-way *two*-relay channel. In the process, we expose a fundamental deadlock problem in which each relay needs to decode before the other in order to enable mutual assistance. Our most important contribution is a resolution to this deadlock problem; we add an additional constraint that ensures some relay can decode at least one message before the other relay. Furthermore, we also introduce several coding schemes to prove that the additional constraint is indeed sufficient. Our schemes also show that information theory provides unique insight into scheduling the transmissions of multi-hop networks.

Acknowledgements

I would like to thank my supervisor Liang for his guidance, encouragement, and unfaltering optimism.

I would like to thank my parents for providing me with food, housing, and transportation - free of charge.

I would like to thank the government of Canada, the university of Waterloo, and the ECE department for all the scholarships and grants that made my graduate studies so comfortable.

Finally, I would like to thank all the graduate students from various programs whom I befriended. You made the tough semesters pass by a little quicker.

Dedication

I dedicate this thesis to my family, especially my littlest sister who never complained when I needed the computer to type it. ;)

Contents

List of Tables	viii
List of Figures	ix
1 Introduction	1
1.1 The Rise (and Fall) of Multi-User Information Theory	3
1.2 An Information Theoretic Revival	4
1.3 The Purpose of this Thesis	5
2 The Problem of Interference	6
2.1 The one-sided interference channel	6
2.1.1 Typical Sequences	8
2.1.2 Mitigating the Interfering Source	9
2.2 The Discrete Memoryless Relay Channel	11
2.3 Summarizing Remarks	14
3 The Two-Way One-Relay Channel	15
3.1 Network Coding and Random Binning	15
3.1.1 Achievable Rates Using Network Coding	16
3.2 Achievable Rates Using Beamforming	18
3.2.1 Comments on Beamforming	19
3.3 A Comparison of Beamforming and Network Coding	20
3.4 Final Remarks on the TWORC	21

4	The Two-Way Two-Relay Channel	22
4.1	The Deadlock Problem	23
4.1.1	The Problem Definition	24
4.1.2	The Problem Resolution	24
4.2	The Main Result	25
4.3	Proof of Theorem 4.2.1	26
4.3.1	Case 1: An Offset Encoding Scheme	27
4.3.2	Case 1: A No-offset Encoding Scheme	29
4.3.3	Case 2: A Multiple-access Scheme	32
4.4	Summarizing Remarks	34
5	The Way Forward	36
	APPENDICES	37
A	Analysis of the One-Sided Interference Channel	38
B	Analysis of Superposition Coding	41
B.1	Generation of Codebooks	41
B.2	Encoding	42
B.3	Decoding	42
B.4	Analysis of Probability of Error	43
	References	44

List of Tables

1.1	Inner and outer bounds for the two-way channel	3
4.1	An offset encoding scheme for case 1	27
4.2	A no-offset encoding scheme for case 1	29
4.3	A multiple-access encoding scheme for case 2	32
B.1	Encoding process for the TWORC	42

List of Figures

1.1	Schematic diagram of a general communication system	2
1.2	Basic Multi-user Channels	4
2.1	The One-Sided Interference Channel	7
2.2	Decoding Strategy for the One-sided Interference Channel	10
2.3	The Discrete Memoryless Relay Channel	11
3.1	The Two-Way One-Relay Channel	15
4.1	The Two-Way Two-Relay Channel	22

Chapter 1

Introduction

Sixty years ago Claude Shannon surprised the world with his paper “A mathematical theory of communication” [1]. In the paper, Shannon argued,

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.

The schematic in figure 1.1 duplicates the type of system considered by Shannon. A significant feature in this scheme is the unidirectional flow of information; messages are transmitted from a single source to a single destination.

Shannon’s most astonishing result was to show that it was possible to transmit information at a positive rate across the noisy channel with a vanishing probability of error. This rate, also known as the capacity of the channel, is *the maximum rate at which information can be sent with as small a frequency of errors or equivocation as desired*. To prove his result, which he called the fundamental theorem for a discrete channel with noise, Shannon introduced several ingenious ideas, including a way of measuring the “surprise” of a random variable from its probability mass function. He termed this measure the entropy of the random variable, and defined it as shown in (1.1).

$$H(X) = - \sum_x p(x) \log p(x) \tag{1.1}$$

Furthermore, Shannon also introduced an associated concept of the mutual information between two random variables, which he defined as the entropy of a

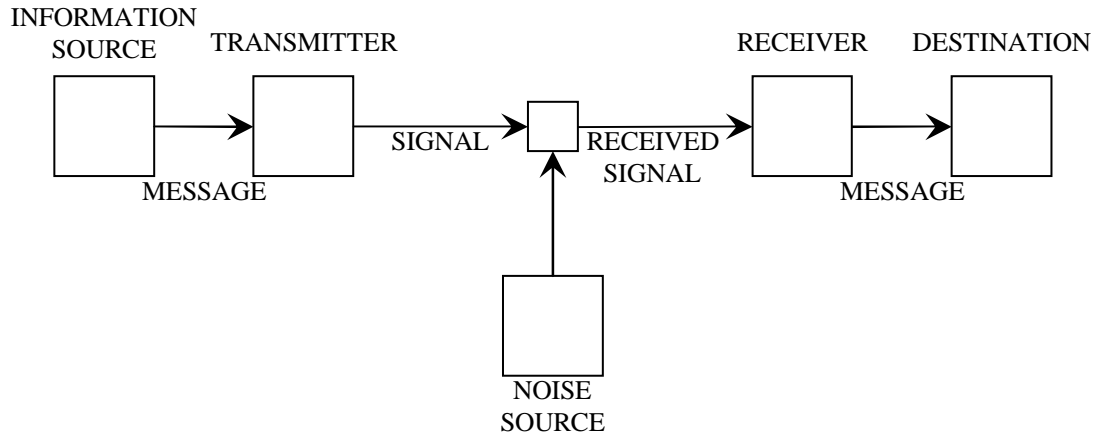


Figure 1.1: Schematic diagram of a general communication system

random variable that is reduced given knowledge of the other. The mutual information of two random variables X and Y is given in (1.2).

$$\begin{aligned}
 I(X;Y) &= H(X) - H(X|Y) \\
 &= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}
 \end{aligned}
 \tag{1.2}$$

The significance of the mutual information in (1.2) is revealed in the fundamental theorem for a discrete channel. For any noisy channel whose dynamics are modeled by the probability transition matrix $p(y|x)$, the channel capacity C is defined as follows,

$$C = \max_{p(x)} I(X;Y)
 \tag{1.3}$$

That is, at all rates below the capacity, information can be transmitted as reliably as desired. But at rates higher than the channel capacity, the probability of error goes to one. The concepts of entropy and mutual information have had a profound impact on the design of modern day communication systems, most notably in the areas of data compression and error correcting codes. Shannon's paper later renamed "*The Mathematical Theory of Communication*" became the foundation of information theory - an area of mathematics that intersects fields ranging from statistical physics to computer science. Shannon himself is often called the father of information theory.

1.1 The Rise (and Fall) of Multi-User Information Theory

Approximately a decade after his famous paper, Shannon decided to revisit the communication model in figure 1.1 that he had originally considered [2]. In his new model, he allowed both nodes to transmit and receive simultaneously.

This two-way channel, now recognized as the first multi-user information theory problem in which multiple sources communicate to multiple destinations, was too difficult for Shannon to solve completely. Instead Shannon was able to determine inner and outer bounds to the capacity region as shown in table 1.1.

The inner bound, or the rates under which communication can be made arbitrarily reliable, closely parallels the result for the one-way channel in (1.3). However, the outer bound, or the rates over which reliable communication is impossible, contains a fundamental difficulty; it permits arbitrary input distributions $p(x_1, x_2)$ whereas the inner bound permits independent input distributions $p(x_1)p(x_2)$. This distinction appears deceptively trivial, but it is actually quite significant. There are channels for which the inner bound is zero; no communication can take place unless the two sources co-operate. Since each source is carrying new information, they have no obvious way of knowing the other's transmission ahead of time, so co-operation is precluded. For the past forty-seven years, Shannon's inner and outer bounds for the two-way channel have remained virtually unchanged. Despite sustained attention, researchers have been unable to either improve the inner bound or tighten the outer bound. This is the longest standing open problem in multi-user information theory.

Several other multi-user channel problems were formulated after the two-way channel. The most significant include the multiple-access channel [3][4], the broadcast channel [5], the relay channel [6], and the interference channel [7][8], all depicted in figure 1.2. Of these, the capacity region is completely known only for the

Table 1.1: Inner and outer bounds for the two-way channel

Inner Bounds	Outer Bounds
For some $p(x_1)p(x_2)$,	For some $p(x_1, x_2)$,
$R_1 < I(X_1; Y_2 X_2)$	$R_1 < I(X_1; Y_2 X_2)$
$R_2 < I(X_2; Y_1 X_1)$	$R_2 < I(X_2; Y_1 X_1)$

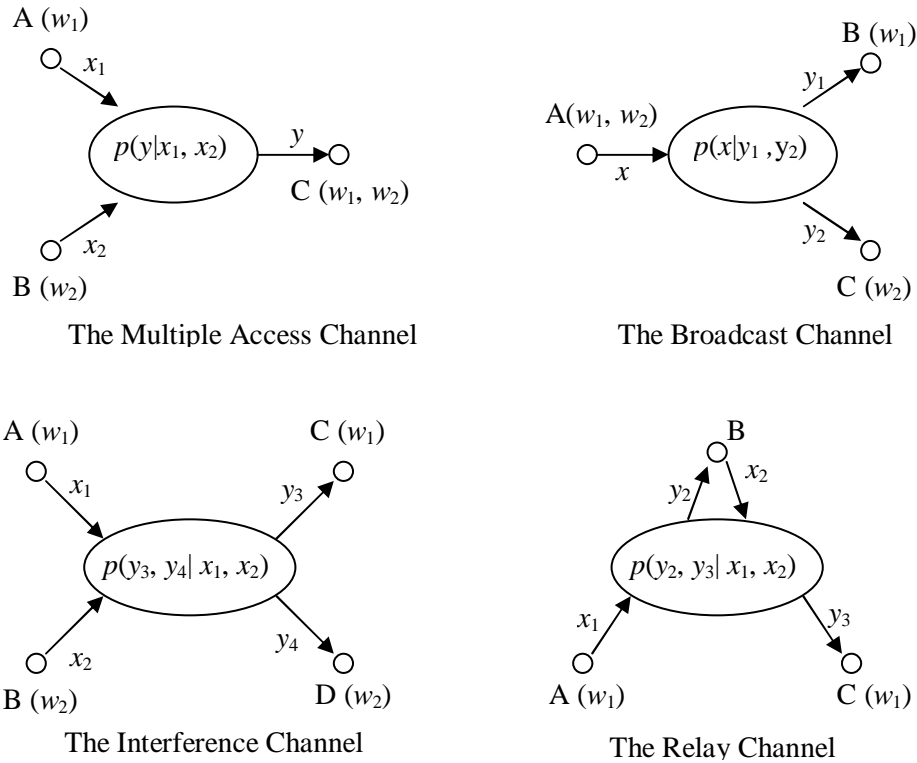


Figure 1.2: Basic Multi-user Channels

multiple-access channel, and then only if the sources are independent.

During the seventies and early eighties the information theory community was preoccupied with fully characterizing the capacity regions of the aforementioned channels. Although there were some prominent advances, a complete characterization remained an elusive and distant goal. Perhaps due to the slow progress, and the computational requirements for application, interest in multi-user information theory subsequently dried up. There was a near two-decade drought during which communication systems and protocols were developed outside of a multi-user information theoretic framework.

1.2 An Information Theoretic Revival

Today, multi-user information theory is experiencing a renaissance. The old obstacles still remain but a new technological landscape has emerged. In particular, there is a massive consumer demand for wireless broadband connectivity. The internet has become an indispensable aspect of every day life while the possibilities of

unlimited and unfettered access to its content provide tantalizing prospects to consumers and businesses alike. Despite its imperfections, information theory provides some unique insights into the phenomena that occur in wireless systems.

Fortunately, in the last decade, there has been some progress in the information theoretic study of wireless networks: new measures of system performance, such as throughput and transport capacity have been introduced [9][10]; traditional coding techniques have been extended and improved [11]; and new coding techniques have been developed [12]. Finally, there have been triumphs with new perspectives on the classical problems; in 2006 the capacity region of the gaussian MIMO broadcast channel was fully characterized [13][14].

1.3 The Purpose of this Thesis

The purpose of this thesis is to add further information theoretic insights to the design of wireless multi-hop networks. We begin in chapter two by discussing some simple yet fundamental questions that have yet to be answered. These problems were encountered during the course of the research, and to illustrate their difficulty we also describe some failed attempts at resolving them. In chapter three we return to the two-way channel that was introduced earlier in this chapter, except we add a relay to the network. Ostensibly, the additional relay makes the problem more difficult. But interestingly, we gain new insight from this two-way relay channel that does not appear in the two-way channel. In chapter four, we extend the ideas from chapter three to the two-way *two*-relay channel. In the process, we expose a fundamental “deadlock” problem, of which its resolution is the major contribution of this thesis. We hope that the interpretation of this solution will neatly tie together all of the ideas discussed in the previous chapters and illustrate the beauty and practicality of applying information theory to wireless networks.

Chapter 2

The Problem of Interference

There are many unresolved obstacles that prevent information theory from making its full impact on the design of wireless networks. Earlier, we encountered one such fundamental difficulty in the two-way channel where it is not known whether two separate nodes can cooperate when transmitting independent information. Unsurprisingly, this problem resurfaces when we examine the two-way relay channel in chapter three. In this chapter we focus on another challenge; the optimal strategy of dealing with interference. We say that a source causes interference if it transmits a message that cannot be decoded, either by the destination, or by an intermediate node endeavoring to assist the transmission. There are two problems we examine:

- Suppose a receiver is interested in decoding a message from a source. How should the receiver deal with interference from another sender if the receiver knows the codebook of this sender?
- Suppose a relay node would like to help a source transmit its message to a destination. How can the relay help if it knows the codebook but cannot decode the message?

2.1 The one-sided interference channel

Consider the one-sided interference channel depicted in figure 2.1. In this network, there are two transmitters (or sources) and one receiver (destination). The receiver is only interested in obtaining the message from the first transmitter but knows the

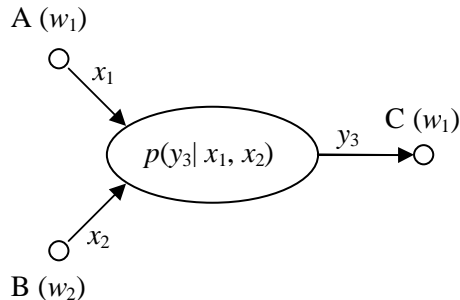


Figure 2.1: The One-Sided Interference Channel

codebooks of both sources. It is known that if the rates lie within the multiple-access region of (2.1)-(2.3), the destination can decode both sources and throw away the message from the second source.

$$R_1 < I(X_1; Y|X_2) \quad (2.1)$$

$$R_2 < I(X_2; Y|X_1) \quad (2.2)$$

$$R_1 + R_2 < I(X_1, X_2; Y) \quad (2.3)$$

The advantage of this strategy is that interference from the second transmitter can be completely removed. Now suppose that R_1 still satisfies (2.1) but R_2 no longer satisfies (2.2). That is, the message of the second source cannot be fully decoded by the receiver. Currently, the state of the art decoding strategy treats the transmission from this source as interference or noise. Under this scheme, the destination can decode sender 1 if its rate, R_1 , satisfies,

$$R_1 < I(X_1; Y) \quad (2.4)$$

We ask the following question: **Can the destination use the codebook from the second source to partially reduce the interference?**

To answer this question we propose a new decoding strategy that uses the codebooks of both sources even if the second source is transmitting at a rate greater than (2.2). We wish to determine if the performance of this scheme is superior to that of treating the source as noise in (2.4). Before proceeding we introduce the concept of “typical sequences” that was pioneered by Shannon in his original 1948 paper.

2.1.1 Typical Sequences

The typical sequences argument is the foundation underlying most information theoretic proofs. The notion of a typical sequence is made precise in definition 2.1.1.

Definition 2.1.1 *The set $A_\epsilon^{(n)}$ of ϵ -typical n -sequences $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$ is defined by*

$$A_\epsilon^{(n)}(Z_1, Z_2, \dots, Z_m) := \left\{ (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) : \left| -\frac{1}{n} \log \text{Prob}(\mathbf{s}) - H(S) \right| < \epsilon, \quad \forall S \subseteq \{Z_1, Z_2, \dots, Z_m\} \right\},$$

where each $\mathbf{z}_i = (z_{i,1}, z_{i,2}, \dots, z_{i,n})$ is a n -vector, $i = 1, 2, \dots, m$, and \mathbf{s} is defined as follows: If $S = (Z_{i_1}, Z_{i_2}, \dots, Z_{i_\ell})$, then $\mathbf{s} = (\mathbf{z}_{i_1}, \mathbf{z}_{i_2}, \dots, \mathbf{z}_{i_\ell})$ and

$$\text{Prob}(\mathbf{s}) = \text{Prob}(\mathbf{z}_{i_1}, \mathbf{z}_{i_2}, \dots, \mathbf{z}_{i_\ell}) = \prod_{t=1}^n p(z_{i_1,t}, z_{i_2,t}, \dots, z_{i_\ell,t})$$

It turns out that typical sequences have important properties that will be exploited in the decoding strategies throughout this thesis. The most relevant ones are given in lemma 2.1.1

Lemma 2.1.1 *For any $\epsilon > 0$, the following hold for sufficiently large n :*

(i) *Let a n -sequence $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m)$ be generated according to*

$$\prod_{t=1}^n p(z_{1,t}, z_{2,t}, \dots, z_{m,t})$$

Then

$$\text{Prob}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) \in A_\epsilon^{(n)}(Z_1, Z_2, \dots, Z_m) \geq 1 - \epsilon.$$

(ii) *Let a n -sequence $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m)$ be generated according to*

$$\prod_{t=1}^n p(z_{1,t} | z_{2,t}, \dots, z_{m-1,t}) p(z_{m,t} | z_{2,t}, \dots, z_{m-1,t}) p(z_{2,t}, \dots, z_{m-1,t})$$

Then

$$\text{Prob}((\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) \in A_\epsilon^{(n)}(Z_1, Z_2, \dots, Z_m)) < 2^{-n(I(Z_1; Z_m | Z_2, \dots, Z_{m-1}) - 6\epsilon)}$$

The proof of lemma 2.1.1 follows immediately from Theorems 14.2.1 and 14.2.3 in [15].

2.1.2 Mitigating the Interfering Source

We are ready to describe a scheme that might reduce the interference from the “unwanted” source. First we define our probabilistic model for the one-sided interference channel, as well as our notions of error probabilities and achievable rates.

Definition 2.1.2 *A one-sided interference channel consists of three alphabets, \mathcal{X}_1 , \mathcal{X}_2 , and \mathcal{Y} and a probability transition function $p(y|x_1, x_2)$.*

Definition 2.1.3 *A $(2^{nR_1}, 2^{nR_2})$ code for the one-sided interference channel consists of*

1. *two sets of integers $W_1 = \{1, 2, \dots, 2^{nR_1}\}$ and $W_2 = \{1, 2, \dots, 2^{nR_2}\}$, called the message sets with $P(W_1 = w_1) = \frac{1}{2^{nR_1}}, P(W_2 = w_2) = \frac{1}{2^{nR_2}}$ for every $w_1 \in \{1, 2, \dots, 2^{nR_1}\}$ and $w_2 \in \{1, 2, \dots, 2^{nR_2}\}$ respectively*
2. *two encoding functions, $X_1 : W_1 \rightarrow \mathcal{X}_1^n$ and $X_2 : W_2 \rightarrow \mathcal{X}_2^n$*
3. *a decoding function $g : \mathcal{Y}_n \rightarrow W_1$*
4. *the average probability of error:*

$$P_e^{(n)} = \frac{1}{2^{nR_1}} \sum_{w_1, w_2} Pr\{g(Y^n) \neq w_1 | W_1 = w_1, W_2 = w_2\}$$

Definition 2.1.4 *A rate R_1 is said to be achievable for the one-sided interference channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \rightarrow 0$ for all R_2 .*

Now consider the following coding scheme. Fix $p(x_1, x_2) = p_1(x_1)p_2(x_2)$ so the sources are independent.

Codebook Generation

Generate 2^{nR_1} independent codewords $\mathbf{X}_1(i)$, $i \in \{1, 2, \dots, 2^{nR_1}\}$, of length n , generating each element i.i.d $\prod_{t=1}^n p(x_{1,t})$. Similarly, generate 2^{nR_2} independent codewords $\mathbf{X}_2(j)$, $j \in \{1, 2, \dots, 2^{nR_2}\}$, of length n , generating each element i.i.d $\prod_{t=1}^n p(x_{2,t})$. These codewords form the codebook, which is revealed to the senders and the receiver.

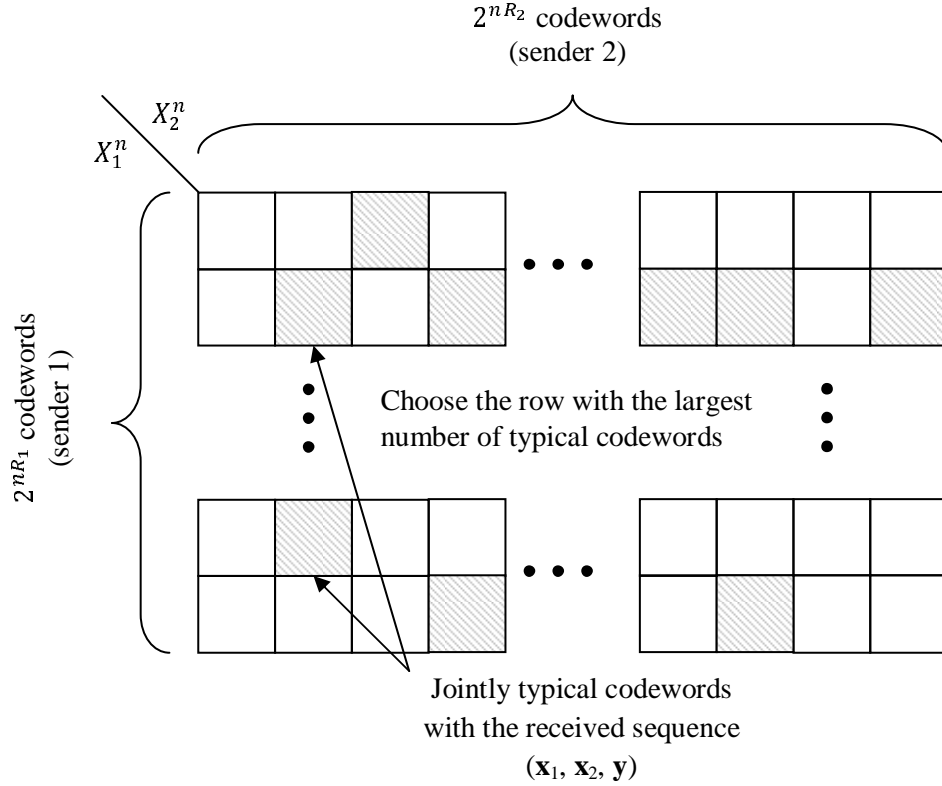


Figure 2.2: Decoding Strategy for the One-sided Interference Channel

Encoding

To send index i , sender 1 sends the codeword $\mathbf{X}_1(i)$. Similarly, to send j , sender 2 sends $\mathbf{X}_2(j)$.

Decoding

Let $A_\epsilon^{(n)}(i)$ denote the set of typical $(\mathbf{x}_1(i), \mathbf{x}_2, \mathbf{y})$ sequences. The receiver chooses the index i such that $|A_\epsilon^{(n)}(i)| > |A_\epsilon^{(n)}(j)|$ for $i \neq j$. Figure 2.2 gives a visual interpretation of this decoding strategy.

The rows and columns in figure 2.2 correspond to the indices (or codewords) of sender 1 and 2 respectively. The receiver chooses the row that contains the largest number of typical codeword pairs. Intuitively it would seem that if the rate of sender 1 satisfies (2.1) then the corresponding row (or message index) would contain the largest number of typical codewords.

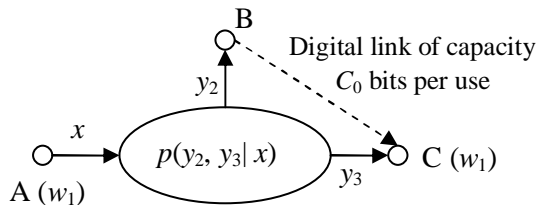


Figure 2.3: The Discrete Memoryless Relay Channel

Probability of Error

The analysis of the probability of error is given in appendix A. The result shows that the probability of error goes to zero if the rate of sender 1 satisfies the inequality in (2.5).

$$R_1 < I(X_1; Y) \tag{2.5}$$

Unfortunately comparing inequalities (2.4) and (2.5) shows that our coding strategy does not perform better than that of treating sender 2 as noise. In fact, the performance remains unchanged. Therefore, we are unable to exploit the knowledge of the second sender’s codebook to reduce its interference.

Comments on the one-sided interference channel

The coding strategy that was suggested for the one-sided interference channel represents a failed attempt by the author to mitigate the interference from a source transmitting above channel capacity. The problem of dealing with “non-decodable” interference is one of the many unsolved mysteries in multi-user information theory. It will reappear later in this thesis.

2.2 The Discrete Memoryless Relay Channel

We are now ready to discuss a second problem that has confounded information theorists for two decades. Consider the seemingly simple discrete memoryless relay channel in figure 2.3. The operation of this channel is made precise in the following definitions.

Definition 2.2.1 *The discrete memoryless relay channel consists of three alphabets, \mathcal{X} , \mathcal{Y}_1 , and \mathcal{Y}_2 and a probability transition function $p(y_1, y_2|x)$. Here Y_1 and Y_2 are conditionally independent and conditionally identically distributed given X . That is $p(y_1, y_2|x) = p(y_1|x)p(y_2|x)$. Also, the channel from Y_1 to Y_2 does not interfere with Y_2 .*

Definition 2.2.2 *A 2^{nR} code for this discrete memoryless channel consists of*

1. *one set of integers $W = \{1, 2, \dots, 2^{nR}\}$ called the message set with $P(W = w) = \frac{1}{2^{nR}}$, for every $w \in \{1, 2, \dots, 2^{nR}\}$*
2. *two encoding functions, $f : W \rightarrow \mathcal{X}^n$ and $f_1 : \mathcal{Y}_1^n \rightarrow 2^{nC_0}$*
3. *a decoding function $g : 2^{nC_0} \times \mathcal{Y}_2^n \rightarrow W$*
4. *the average probability of error $P_e^{(n)} = \frac{1}{2^{nR}} \sum_w Pr\{g(f_1(Y_1^n), Y_2^n) \neq w | W = w\}$*

Definition 2.2.3 *A rate R is said to be achievable for the discrete memoryless channel if there exists a sequence of $((2^{nR}, n)$ codes such that $P_e^{(n)} \rightarrow 0$.*

Now suppose the relay knows the source codebook but is unable to decode the message. In this scenario the only known strategy is called “compress-and-forward”, in which the relay quantizes its observations Y_1^n and sends them over the digital link to the destination. Unfortunately the performance of compress-forward is generally quite poor. We wish to devise a better strategy that makes use of the relay’s knowledge of the sender’s codebook. We ask the following question: **Can the relay use the source codebook to help the destination decode the message?**

This problem is similar to the one-sided interference channel, except the relay is now interested in helping the destination decode the source message, and not in decoding the message for itself.

A Possible Encoding Strategy

We suggest the following encoding scheme. After the source has finished transmitting a message $\mathbf{x}(w)$, $w \in \{1, 2, \dots, 2^{nR}\}$, let the relay make a list of all the possible codewords that are typical with its received sequence, $\mathcal{L}(Y_1^n)$. The relay sends this list to the destination.

Analysis of the Encoding Strategy

Since the relay has a digital link to the destination, we can send at most 2^{nC_0} possible lists. Therefore this scheme is only feasible if the number of lists grows at most exponentially with n . Unfortunately, the actual number of lists will depend on the probabilistic structure of the channel and the transmissions. With the hope of arriving at a general result, we use Stirling's formula to obtain a coarse upper bound to the number of lists. Using the basic properties of typical sequences we know that,

- The expected number of codewords in a specific list, $|\mathcal{L}(Y_1^n)|$ is given by the product of the total number of codewords and the probability that a random codeword is typical with the received sequence. That is,

$$|\mathcal{L}(Y_1^n)| = 2^{nR} 2^{-nI(X;Y_1)}. \quad (2.6)$$

- The number of lists is at most equal to the number of ways of choosing $|\mathcal{L}(Y_1^n)|$ codewords from the total of 2^{nR} in the codebook.

Combining the two items above, we see that the total number of lists is at most,

$$\binom{2^{nR}}{2^{nR} 2^{-nI(X;Y_1)}}$$

This combinatorial can be upper bounded using lemma 2.2.1 below

Lemma 2.2.1 *For $0 < p < 1, q = 1 - p$, such that np is an integer,*

$$\frac{1}{\sqrt{8npq}} \leq \binom{n}{np} 2^{-nH(p)} \leq \frac{1}{\sqrt{\pi npq}}$$

Lemma 2.2.1 corresponds to Lemma 17.5.1 in [15]. Using this lemma we arrive at the following,

$$\binom{2^{nR}}{2^{nR} 2^{-nI(X;Y_1)}} \leq \frac{1}{\sqrt{\pi 2^{nR} pq}} 2^{-2^{nR} H(p)}$$

where,

$$\begin{aligned} p &= 2^{-nI(X;Y_1)}, \\ q &= 1 - 2^{-nI(X;Y_1)} \end{aligned}$$

Letting n become large, we have

$$\frac{1}{\sqrt{\pi 2^{nR} pq}} 2^{-2^{nR} H(p)} \rightarrow \frac{1}{\sqrt{\pi}} 2^{n(I(X;Y_1) 2^{n(R-I(X;Y_1)) - 0.5(R-I(X;Y_1))})} \quad (2.7)$$

The expression in (2.7) shows that if the relay is unable to decode the source (ie. $R > I(X; Y_1)$), the number of lists of typical codewords $\mathcal{L}(Y_1^n)$ can potentially grow at a *doubly* exponential rate with respect to n . As a result we are unable to show that $\binom{2^{nR}}{2^{nR} 2^{-nI(X;Y_1)}} \leq 2^{nC_0}$ for some $C_0 < \infty$. Therefore our coding scheme for this channel may not be feasible.

The discrete memoryless relay channel illustrates the difficulty of determining an appropriate coding strategy when the relay is unable to fully decode the source. A version of this problem was first proposed twenty years ago in [16] has remained unsolved ever since.

2.3 Summarizing Remarks

In this section we encountered two problems that demonstrated the challenges of dealing with a source whose message cannot be decoded, but whose codebook is known. This fundamental difficulty motivates us for the remainder of this thesis to impose conditions that guarantee every node in the network will be able to fully decode the message intended for it. Although it's not possible to prove the necessity of these conditions, our preceding discussions show why alternative strategies have yet to be discovered.

Chapter 3

The Two-Way One-Relay Channel

In this chapter we return to the two-way channel that was introduced in chapter 1, with the addition of a relay to the network. The two-way *one-relay* channel (hereafter the TWORC), depicted in figure 3.1, seems at first glance to be an even more imposing problem than the original two-way channel considered by Shannon. Indeed, the capacity regions for the two-way channel and the relay channel, special cases of the TWORC, have separately been unknown for several decades. It would seem that combining the two networks would merely present more unresolved issues. There is some truth to this assertion, but it also turns out that the TWORC offers its own unique insights that we will encounter shortly. Therefore in some sense, the TWORC is greater than the sum of its parts.

3.1 Network Coding and Random Binning

What is our incentive for studying the TWORC? Simply put, this channel illustrates the advantages of network coding [12], a relatively recent research topic, but one that is expected to bring about fundamental changes to the design of communication systems. The basic idea of network coding can be explained through a simplified

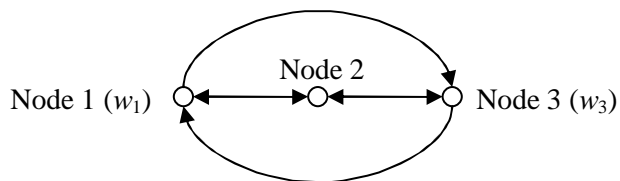


Figure 3.1: The Two-Way One-Relay Channel

version of figure 3.1. Suppose node 1 and node 3 each have one bit of information, b_1 and b_3 respectively, to send each other with the help of a relay node 2. Instead of transmitting each bit separately node 2 can help both destinations simultaneously by sending $b_1 \oplus b_3$. Since node 1 and node 3 know their previously transmitted bits, each can recover the other's bit by computing $(b_1 \oplus b_3) \oplus b_1 = b_3$ and $(b_1 \oplus b_3) \oplus b_3 = b_1$ respectively. An interesting feature of this scheme, is that although only one bit is transmitted by node 2, two different bits are recovered at nodes 1 and 3. This strategy is particularly appealing in a wireless setting, because the relay is exploiting the common channel between itself and nodes 1 and 3 [17]. Furthermore, the XOR operation can easily be implemented with minimal complexity. However, the success of this scheme and network coding in general, depends on the availability of *side information* at each destination, in this case the side information being the previously sent bits.

3.1.1 Achievable Rates Using Network Coding

In this paper, we address a more general framework, where motivated by wireless communications, the TWORC is described through the following definitions.

Definition 3.1.1 *The two-way one-relay channel consists of three input alphabets \mathcal{X}_1 , \mathcal{X}_2 , and \mathcal{X}_3 , three output alphabets \mathcal{Y}_1 , \mathcal{Y}_2 , and \mathcal{Y}_3 , and a probability transition function $p(y_1, y_2, y_3 | x_1, x_2, x_3)$. Assume that each node $i \in \{1, 2, 3\}$ sends $x_{i,t}$ and receives $y_{i,t}$ at time t .*

Definition 3.1.2 *A $(2^{nR_1}, 2^{nR_3})$ code for the two-way one-relay channel consists of*

1. *two sets of integers $W_1 = \{1, 2, \dots, 2^{nR_1}\}$ and $W_3 = \{1, 2, \dots, 2^{nR_3}\}$, called the message sets with $P(W_1 = w_1) = \frac{1}{2^{nR_1}}$, $P(W_3 = w_3) = \frac{1}{2^{nR_3}}$ for every $w_1 \in \{1, 2, \dots, 2^{nR_1}\}$ and $w_3 \in \{1, 2, \dots, 2^{nR_3}\}$ respectively*
2. *two encoding functions, $X_1 : W_1 \rightarrow \mathcal{X}_1^n$ and $X_3 : W_3 \rightarrow \mathcal{X}_3^n$*
3. *a causal relay encoding function, $f_{2,t} : \mathcal{Y}_2^n \rightarrow \mathcal{X}_2^n$ with a one-step time delay to account for the signal processing time so that for all t ,*

$$x_{2,t} = f_{2,t}(y_{2,t-1}, y_{2,t-2}, \dots)$$

4. *two decoding functions $g_1 : \mathcal{Y}_1^n \rightarrow W_3$ and $g_3 : \mathcal{Y}_3^n \rightarrow W_1$*

5. the average probability of error

$$P_e^{(n)} = \frac{1}{2^{nR_1}} \frac{1}{2^{nR_3}} \sum_{w_1, w_3} Pr\{g_1(Y_1^n) \neq w_3 \cup g_3(Y_3^n) \neq w_1 | W_1 = w_1, W_3 = w_3\}$$

Definition 3.1.3 A rate pair (R_1, R_3) is said to be achievable for the two-way one-relay channel if there exists a sequence of $(2^{nR_1}, 2^{nR_3}, n)$ codes such that $P_e^{(n)} \rightarrow 0$.

Now an immediate question is whether the idea of network coding can be applied to this more general framework, and if so, what are the corresponding achievable rates. It turns out that network coding is a specialized version of random binning, a classical technique in multi-user information theory.

Theorem 3.1.1 For the two-way one-relay problem defined above, fix some $p(x_1)p(x_2)p(x_3)$. For node 1 and 3 to decode each other's message the rate pair (R_1, R_3) must satisfy the following two inequalities:

$$R_1 < I(X_1, X_2; Y_3 | X_3) \tag{3.1}$$

$$R_3 < I(X_2, X_3; Y_1 | X_1) \tag{3.2}$$

Furthermore, for node 2 to decode both messages from node 1 and 3, (R_1, R_3) must also satisfy:

$$R_1 < I(X_1; Y_2 | X_2, X_3) \tag{3.3}$$

$$R_3 < I(X_3; Y_2 | X_1, X_2) \tag{3.4}$$

$$R_1 + R_3 < I(X_1, X_3; Y_2 | X_2) \tag{3.5}$$

A detailed proof of theorem 3.1.1, including a discussion of the relationship between network coding and random binning first appeared in [18].

Comments on the Network Coding Region

Network coding is an example of the “decode-and-forward” strategy for relay networks, in which the relay must fully decode all the messages in order to help. An interesting feature of the decode-and-forward scheme for the TWORC is that there is no interference in the entire network; all signals either carry information that must be decoded or is known a priori.

Network coding is a particularly attractive decode-and-forward scheme because it can easily be implemented using the XOR function as described earlier. Furthermore, its achievable region has a nice interpretation for the following reasons:

Remark 3.1.1 *At nodes 1 and 3, the cut-set bounds are sufficient for each node to decode the other’s message. The cut-set bounds refer to the outer-bounds of the capacity region; no rate pairs outside of the cut-set bounds are achievable.*

Remark 3.1.2 *At node 2, the multiple access region is sufficient for the relay to decode both messages. No rates outside of this region are achievable if the relay is to decode both messages.*

Unfortunately, decode-and-forward schemes in general (and network coding in particular) have their drawbacks as well. First of all, the relay need not decode the messages from the sources in order to help so we cannot prove that (3.3)-(3.5) are necessary. Compress-forward is an example of a scheme that does not fall into the decode-and-forward category. In this scheme, the relay blindly helps the destination by sending quantized versions of its observations. However, this strategy is not desirable; its performance is generally poor, and there are no interesting situations for which it is optimal.

In chapter 2 we described some attempts to find attractive coding strategies that did not require the relay to fully decode the message. Our aim was to exploit the relay’s knowledge of the codebook to mitigate the interference caused by an unknown message. Ultimately, all of our attempts were unsuccessful. In general, there is no desirable alternative to the “decode-forward” strategy in which the relay must decode the message before helping. Therefore the constraints (3.3)-(3.5), though sufficient but not necessary, are still at present required to achieve good performance.

Another limitation of the network coding region is that the input distribution in theorem 3.1.1 must be independent. In practice, this precludes the nodes from beamforming their transmissions; their signals will be non-coherent. However, it is possible using another decode-and-forward scheme called superposition coding to correlate the signals of each node [19]. In the next section, we examine the performance of beamforming in the TWORC, and determine whether there are any improvements over network coding.

3.2 Achievable Rates Using Beamforming

A beamforming coding scheme for the TWORC network was first studied in [20], where, however, it was incorrectly claimed that (3.1)-(3.5) are achievable for any

$p(x_1|x_{2f})p(x_3|x_{2b})p(x_{2f})p(x_{2b})$. The analysis in [20] was limited to the AWGN channel, but erred by not distinguishing the signal power sending superimposed “new” information, from the signal power coherently transmitting “old” information. In addition to recognizing this error, our contribution is the statement of the correct achievable region in theorem 3.2.1 and its proof in appendix B. This region is based on superposition coding and first appeared in [21].

Theorem 3.2.1 *For the two-way one-relay problem defined earlier, fix some $p(u_1)p(u_3)p(x_1|u_1)p(x_2|u_1, u_3)p(x_3|u_3)$. For node 1 and 3 to decode each other’s message the rate pair (R_1, R_3) must satisfy the following two inequalities:*

$$R_1 < I(X_1, X_2; Y_3 | U_3, X_3) \quad (3.6)$$

$$R_3 < I(X_2, X_3; Y_1 | U_1, X_1) \quad (3.7)$$

Furthermore, for node 2 to decode both messages from node 1 and 3, (R_1, R_3) must also satisfy:

$$R_1 < I(X_1; Y_2 | U_1, X_2, X_3) \quad (3.8)$$

$$R_3 < I(X_3; Y_2 | U_3, X_1, X_2) \quad (3.9)$$

$$R_1 + R_3 < I(X_1, X_3; Y_2 | U_1, U_3, X_2) \quad (3.10)$$

3.2.1 Comments on Beamforming

As mentioned earlier, superposition coding (or beamforming) is a type of decode-and-forward scheme. So our previous observation that decode-and-forward may be suboptimal also applies to superposition coding. Furthermore, we can notice other properties about the region in (3.6)-(3.10).

- The transmissions of each node are correlated through U_1 and U_3 . Thus, the relay and each source can beamform, or communicate coherently at signal level.
- We pay a price for this coherence; (3.6) and (3.7) are conditioned on U_1 and U_3 respectively. The relay essentially divides its transmission power between the two messages so only a portion of the total power is used to send the message of interest to each destination (network coding used all the power for each message). However, this power is coherent at signal level with the transmission from the other source (unlike network coding).

- The beamforming region is not as “nice” as the network coding region. The constraints (3.6) and (3.7) do not match the cut-set bounds, and (3.8)-(3.10) does not match the multiple-access region. Therefore remarks 3.1.1 and 3.1.2 do *not* apply with beamforming.

In the next section we compare the performance of beamforming with network coding and draw some important conclusions.

3.3 A Comparison of Beamforming and Network Coding

Of the two decode-and-forward strategies discussed in this chapter, network coding and beamforming, which is superior to the other? To help answer this question, we first examine the outer bounds (or cut-set bounds) of the capacity region for the TWORC.

Theorem 3.3.1 *For the two-way one-relay problem defined earlier, any achievable rate pair (R_1, R_3) must satisfy the following inequalities: For some $p(x_1, x_2, x_3)$*

$$R_1 < I(X_1, X_2; Y_3 | X_3) \tag{3.11}$$

$$R_3 < I(X_2, X_3; Y_1 | X_1) \tag{3.12}$$

and

$$R_1 < I(X_1; Y_2, Y_3 | X_2, X_3)$$

$$R_3 < I(X_3; Y_1, Y_2 | X_1, X_2)$$

The proof of this theorem follows from the standard cut-set bounds in theorem 15.10.1 of [15]. Now we can use this theorem to answer the question posed at the beginning of this section. It turns out that of the two strategies we examined in this chapter, network coding and beamforming, neither strategy is always superior to the other. This conclusion can be justified by the following two remarks.

Remark 3.3.1 *Although some of the cut-set bounds, in particular 3.11 and 3.12 are achievable using network coding, the input distribution of each node must also be independent.*

Remark 3.3.2 *Superposition coding (or beamforming) allows the nodes to coherently transmit their messages, but the correlation carries a price; none of the cut-set bounds in theorem 3.3.1 are achievable.*

To summarize, remarks 3.3.1 and 3.3.2 tell us that in general, the decode-and-forward strategy may not be optimal for the TWORC, even for special cases like the gaussian channel or the degraded channel. To achieve optimality, it is not enough for the relay to decode both sources; each sender must also completely know the relay transmission in advance. But since the relay is carrying new information from the other sender, this requirement may not be feasible. It seems that resolving this dilemma appears to be “two-way channel hard”, or as difficult as finding a solution to the two-way channel in chapter 1. However, in relay networks with a unidirectional flow of information, decode-and-forward has been proven optimal for degraded channels.

3.4 Final Remarks on the TWORC

We now summarize the main points of this chapter.

- The TWORC illustrates the advantages of network coding in the wireless setting. Network coding also leads to a nice achievable region for the TWORC as described in remarks 3.1.1 and 3.1.2.
- Network coding precludes the use of beamforming because the input distributions must be independent. On the other hand, superposition coding includes correlated input distributions, allowing the use of beamforming.
- Beamforming and network coding are both decode-and-forward strategies; the relay must completely decode both source messages in order to help. In general, there are no desirable alternatives to the decode-and-forward scheme. Chapter 2 described some failed attempts at finding such alternatives.
- In general, neither beamforming or network coding is always a superior strategy than the other. Furthermore, the decode-and-forward strategy may not be optimal for the TWORC, even in special cases like the gaussian or degraded channel. Proving optimality appears to be as difficult as finding a solution to the two-way channel.

Chapter 4

The Two-Way Two-Relay Channel

In this chapter we attempt to extend the analysis and insight of the two-way one-relay channel to the two-way *two*-relay channel (hereafter the TWTRC) depicted in figure 4.1. On the surface, it may appear that there is little incentive to carry out this extension. After all, decode-and-forward, the only practical coding strategy presently known, may not be optimal for even specialized cases of the TWTRC. In any event, it is not clear which of the two decode-and-forward strategies, network coding or beamforming, is preferable. These issues reflect the core difficulty of applying information theory to wireless networks; there are so many unresolved questions. Where do we begin?

We begin by tightening our focus. Instead of looking for the final solution, which may never be found, we look for a practical solution that has a simple but insightful interpretation. To that end, we redirect our attention to network coding, a strategy that had two advantages: first, the scheme could be easily implemented

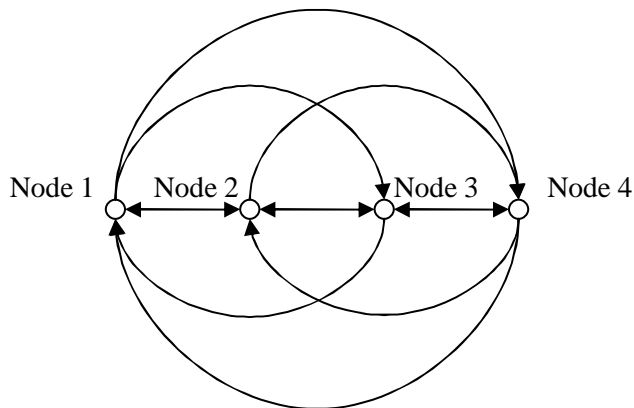


Figure 4.1: The Two-Way Two-Relay Channel

using the XOR function (see section 3.1); secondly, the resulting achievable region had a nice interpretation (see remarks 3.1.1 and 3.1.2). Hence, the two motivating questions of this chapter are as follows:

1. Is it possible to easily extend the network coding scheme of the two-way one-relay channel to the *two*-relay channel, and if so how do we schedule the encoding and decoding at each node?
2. Does the simple interpretation of the network coding achievable region for the two-way one-relay channel (see remarks 3.1.1 and 3.1.2) also apply to the *two*-relay channel?

Lets start by attempting to answer the first question. It does not appear that the network coding scheme of section 3.1 can be trivially extended to the TWTRC. Due to the network model, there is a delay between the time a source begins sending a message bit and the time a relay decodes and retransmits the bit (see definition 3.1.2). With two relays, it is now possible to mix (or XOR) bits from different time intervals. The additional relay introduces a timing problem that was previously absent from the two-way one-relay channel.

4.1 The Deadlock Problem

Since the answer to our first question is not obvious, we will skip to the second question, and return to the first later. Our objective is to determine if the network coding achievable region in theorem 3.1.1 extends nicely to the TWTRC. More specifically, we would like to know if the following rates are achievable in the TWTRC setting.

Fix some $p(x_1)p(x_2)p(x_3)p(x_4)$. For node 1 and 4 to decode their intended messages,

$$R_1 < I(X_1, X_2, X_3; Y_4|X_4) \tag{4.1}$$

$$R_4 < I(X_2, X_3, X_4; Y_1|X_1) \tag{4.2}$$

and for relay node 2 to decode node 1 and 4,

$$R_1 < I(X_1; Y_2|X_2, X_3, X_4) \tag{4.3}$$

$$R_4 < I(X_3, X_4; Y_2|X_1, X_2) \tag{4.4}$$

$$R_1 + R_4 < I(X_1, X_3, X_4; Y_2|X_2) \tag{4.5}$$

and for relay node 3 to decode node 1 and 4,

$$R_1 < I(X_1, X_2; Y_3 | X_3, X_4) \quad (4.6)$$

$$R_4 < I(X_4; Y_3 | X_1, X_2, X_3) \quad (4.7)$$

$$R_1 + R_4 < I(X_1, X_2, X_4; Y_3 | X_3) \quad (4.8)$$

4.1.1 The Problem Definition

The proposed region above appears reasonable at first. Notice that (4.1) and (4.2) correspond to the cut-set bounds (without beamforming). Furthermore (4.3)-(4.5) and (4.6)-(4.8) seem reasonable extensions of the multiple-access region to each relay node. Unfortunately, there is a fundamental difficulty in achieving (4.1)-(4.8). To achieve (4.3)-(4.5), node 3 needs to decode before node 2 in order to help, but the reverse is also needed for (4.6)-(4.8). This “deadlock” problem was identified in [22] when a backward decoding scheme was tried for achieving (4.1)-(4.8).

4.1.2 The Problem Resolution

In this paper, we resolve the deadlock by adding an additional constraint to (4.1)-(4.8) that ensures some relay can decode at least one of the sources before the other relay. This requirement is reasonable, particularly in light of the conclusions in chapter 2; there is no desirable alternative to a coding strategy in which a node fully decodes the message intended for it. So to start the flow of information, it is expected that one of the relays must be able to decode a source message first.

There are two ways in which a coding scheme satisfies the additional constraint: in the first case some relay decodes *one* source before the other relay, and in the second case some relay decodes *both* sources before the other relay. For each case, we will develop coding schemes that can recover the region defined by (4.1)-(4.8). These schemes describe how to schedule the encoding and decoding at each node, thus resolving the first question we posed at the beginning of this chapter. A key ingredient is an offset-encoding strategy developed in [23], that gives more flexibility when combined with sliding-window decoding. As in [18] the schemes presented in this paper also introduce random binning at the relay nodes.

4.2 The Main Result

Consider a network of four nodes 1, 2, 3, 4, with the input-output dynamics modeled by the discrete memoryless channel

$$(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{X}_4, p(y_1, y_2, y_3, y_4 | x_1, x_2, x_3, x_4), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3 \times \mathcal{Y}_4)$$

That is, at any time $t = 1, 2, \dots$, the outputs $y_{1,t}$, $y_{2,t}$, $y_{3,t}$, and $y_{4,t}$ received by the four nodes respectively only depend on the inputs $x_{1,t}$, $x_{2,t}$, $x_{3,t}$, and $x_{4,t}$ transmitted by the four nodes at the same time according to $p(y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t} | x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t})$.

In the two-way two-relay problem, the source nodes 1 and 4 communicate with each other at rates R_1 and R_4 respectively, with the help of the relay nodes 2 and 3. We are interested in the simultaneously achievable rates (R_1, R_4) . The standard definitions of codes and achievable rates are by now familiar and omitted, except a special note that during any time t , each node i can choose its input $x_{i,t}$ based on the past outputs $(y_{i,t-1}, y_{i,t-2}, \dots, y_{i,1})$ it has already received.

Theorem 4.2.1 *For the two-way two-relay problem defined above, any rates (R_1, R_4) satisfying (4.1)-(4.8) are simultaneously achievable, provided that at least one of the following constraints hold:*

$$R_1 < I(X_1; Y_2 | X_2, X_3) \tag{4.9}$$

$$R_4 < I(X_4; Y_3 | X_2, X_3) \tag{4.10}$$

$$R_1 + R_4 < \max\{I(X_1, X_4; Y_2 | X_2, X_3), I(X_1, X_4; Y_3 | X_2, X_3)\} \tag{4.11}$$

Theorem 4.2.1 first appeared in [21]. Observe that by symmetry, it is only necessary to consider two mutually exclusive cases in (4.9)-(4.11).

- **Case 1** in which (4.10) holds. In this case, the rate of node 4 is low enough to ensure node 3 can decode the message despite interference from node 1. More generally, it will be shown that the first case applies when some relay decodes *one* source before the other relay.
- **Case 2** in which the first part of (4.11) holds but neither (4.9) nor (4.10) hold. The throughput is low enough to ensure node 2 can decode without assistance from node 3 but unlike the first case, node 1 does not cause any interference. More generally, it will be shown that the second case applies when some relay decodes *both* sources before the other relay.

It is emphasized that the rate regions of the first and second cases are different. That is, there exist rate pairs (R_1, R_4) that satisfy (4.9) or (4.10) but not (4.11) and vice-versa.

4.3 Proof of Theorem 4.2.1

Each case is examined separately. In the first case we show there are two coding strategies that together, obtain the rate region defined by (4.1)-(4.8) and (4.10): an offset encoding scheme, and a no-offset encoding scheme. In the second case, we use a multiple-access strategy to recover the rate region defined by (4.1)-(4.8) and the first part of (4.11).

Proof For any fixed $p(x_1)p(x_2)p(x_3)p(x_4)$, choose

$$\begin{aligned} R_2 &\geq \max\{I(X_2; Y_1|X_1, X_3), I(X_2; Y_3|X_3, X_4), \\ &\quad I(X_2; Y_4|X_3, X_4)\} \\ R_3 &\geq \max\{I(X_3; Y_1|X_1, X_2), I(X_3; Y_2|X_1, X_2), \\ &\quad I(X_3; Y_4|X_4)\} \end{aligned}$$

We use the Markov block coding argument. Consider B blocks of transmission, each of n transmission slots. For some fixed $K, J \in \mathcal{T}$, a sequence of $B - K$ indices, $w_{1,b} \in \{1, \dots, 2^{nR_1}\}, b = 1, 2, \dots, B - K$ will be sent over from node 1 to node 4 in nB transmission slots, and at the same time, another sequence of $B - J$ indices, $w_{4,b} \in \{1, \dots, 2^{nR_4}\}$ will be sent over from node 4 to node 1.

Generation of Codebooks

For each node $i = 1, \dots, 4$ independently generate 2^{nR_i} i.i.d n -sequences $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,n})$ in \mathcal{X}_i^n according to $p(x_i)$. Index them as $\mathbf{x}_i(w_i), w_i \in \{1, 2, \dots, 2^{nR_i}\}$.

Random Binning

For each relay node $i = 2, 3$, generate 2^{nR_i} bins, indexed by $B_i(k)$ with $k = 1, \dots, 2^{nR_i}$. Independently throw each index pair $(w_1, w_4), w_1 \in \{1, 2, \dots, 2^{nR_1}\}, w_4 \in \{1, 2, \dots, 2^{nR_4}\}$ into the 2^{nR_i} bins according to the uniform distribution. Let $k_i(w_1, w_4)$ be the index of the bin which contains the pair (w_1, w_4) .

Table 4.1: An offset encoding scheme for case 1

Node	Block $b - 3$	Block $b - 2$	Block $b - 1$	Block b
1	$\mathbf{x}_1(w_{1,b-3})$	$\mathbf{x}_1(w_{1,b-2})$	$\mathbf{x}_1(w_{1,b-1})$	$\mathbf{x}_1(w_{1,b})$
2	$\mathbf{x}_2(w_{1,b-5}, w_{4,b-5})$	$\mathbf{x}_2(w_{1,b-4}, w_{4,b-4})$	$\mathbf{x}_2(w_{1,b-3}, w_{4,b-3})$	$\mathbf{x}_2(w_{1,b-2}, w_{4,b-2})$
3	$\mathbf{x}_3(w_{1,b-6}, w_{4,b-4})$	$\mathbf{x}_3(w_{1,b-5}, w_{4,b-3})$	$\mathbf{x}_3(w_{1,b-4}, w_{4,b-2})$	$\mathbf{x}_3(w_{1,b-3}, w_{4,b-1})$
4	$\mathbf{x}_4(w_{4,b-3})$	$\mathbf{x}_4(w_{4,b-2})$	$\mathbf{x}_4(w_{4,b-1})$	$\mathbf{x}_4(w_{4,b})$

4.3.1 Case 1: An Offset Encoding Scheme

Encoding

We use an offset encoding scheme to obtain an achievable rate region consistent with (4.10). The encoding scheme is depicted in table 4.1. It is assumed that a message pair transmitted in block b was decoded at the end of block $b - 1$. An interesting feature of this coding scheme is that node 2 decodes the message from node 1 based on the signals received over two blocks (not one block).

Decoding

1. At the end of each block $b = 3, \dots, B$, node 1 determines the unique index $\hat{w}_{4,b-2}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{X}_{1,b-2}, \mathbf{x}_2(k_2(w_{1,b-4}, w_{4,b-4})), \mathbf{x}_3(k_3(w_{1,b-5}, w_{4,b-3})), \mathbf{x}_4(\hat{w}_{4,b-2}), \mathbf{Y}_{1,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_1) \\ & (\mathbf{X}_{1,b-1}, \mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-4}, \hat{w}_{4,b-2})), \mathbf{Y}_{1,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, Y_1) \\ & (\mathbf{X}_{1,b}, \mathbf{x}_2(k_2(w_{1,b-2}, \hat{w}_{4,b-2})), \mathbf{Y}_{1,b}) \in A_\epsilon^{(n)}(X_1, X_2, Y_1) \end{aligned}$$

2. At the end of each block $b = 2, \dots, B$, node 2 determines the unique index pair $(\hat{w}_{1,b-1}, \hat{w}_{4,b-1})$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-1}), \mathbf{X}_{2,b-1}, \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-2})), \mathbf{x}_4(\hat{w}_{4,b-1}), \mathbf{Y}_{2,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_2) \\ & (\mathbf{X}_{2,b}, \mathbf{x}_3(k_3(w_{1,b-3}, \hat{w}_{4,b-1})), \mathbf{Y}_{2,b}) \in A_\epsilon^{(n)}(X_2, X_3, Y_2) \end{aligned}$$

3. At the end of each block $b = 1, \dots, B$, node 3 determines the unique index pair $(\hat{w}_{1,b-2}, \hat{w}_{4,b})$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-2}), \mathbf{x}_2(k_2(w_{1,b-4}, w_{4,b-4})), \mathbf{X}_{3,b-2}, \mathbf{x}_4(w_{4,b-2}), \mathbf{Y}_{3,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_3) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-2}, w_{4,b-2})), \mathbf{X}_{3,b}, \mathbf{x}_4(\hat{w}_{4,b}), \mathbf{Y}_{3,b}) \in A_\epsilon^{(n)}(X_2, X_3, X_4, Y_3) \end{aligned}$$

4. At the end of each block $b = 4, \dots, B$, node 4 determines the unique index $\hat{w}_{1,b-3}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-3}), \mathbf{x}_2(k_2(w_{1,b-5}, w_{4,b-5})), \mathbf{x}_3(k_3(w_{1,b-6}, w_{4,b-4})), \mathbf{X}_{4,b-3}, \mathbf{Y}_{4,b-3}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-3}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-2})), \mathbf{X}_{4,b-1}, \mathbf{Y}_{4,b-1}) \\ & \in A_\epsilon^{(n)}(X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_3(k_3(\hat{w}_{1,b-3}, w_{4,b-1})), \mathbf{X}_{4,b}, \mathbf{Y}_{4,b}) \in A_\epsilon^{(n)}(X_3, X_4, Y_4) \end{aligned}$$

Analysis of Probability of Error

1. In block b , node 1 can decode $w_{4,b-2}$ with arbitrarily small probability of error if

$$R_4 < I(X_4; Y_1 | X_1, X_2, X_3) + I(X_3; Y_1 | X_1, X_2) + I(X_2; Y_1 | X_1)$$

where the three mutual informations follow from the three typicality checks respectively and their combination leads to (4.2).

2. In block b , node 2 can decode the pair $(w_{1,b-1}, w_{4,b-1})$ with arbitrarily small probability of error if

$$\begin{aligned} R_1 & < I(X_1; Y_2 | X_2, X_3, X_4) \\ R_4 & < I(X_4; Y_2 | X_1, X_2, X_3) + I(X_3; Y_2 | X_2) \\ R_1 + R_4 & < I(X_1, X_4; Y_2 | X_2, X_3) + I(X_3; Y_2 | X_2) \\ & = I(X_1, X_3, X_4; Y_2 | X_2) \end{aligned} \tag{4.12}$$

where each inequality corresponds to one of the three ways a message pair can be decoded incorrectly.

3. In block b , node 3 can decode the pair $(w_{1,b-2}, w_{4,b})$ with arbitrarily small probability of error if

Table 4.2: A no-offset encoding scheme for case 1

Node	Block $b - 2$	Block $b - 1$	Block b
1	$\mathbf{x}_1(w_{1,b-2})$	$\mathbf{x}_1(w_{1,b-1})$	$\mathbf{x}_1(w_{1,b})$
2	$\mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-4}))$	$\mathbf{x}_2(k_2(w_{1,b-2}, w_{4,b-3}))$	$\mathbf{x}_2(k_2(w_{1,b-1}, w_{4,b-2}))$
3	$\mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-3}))$	$\mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-2}))$	$\mathbf{x}_3(k_3(w_{1,b-2}, w_{4,b-1}))$
4	$\mathbf{x}_4(w_{4,b-2})$	$\mathbf{x}_4(w_{4,b-1})$	$\mathbf{x}_4(w_{4,b})$

$$\begin{aligned}
 R_1 &< I(X_1; Y_3 | X_2, X_3, X_4) + I(X_2; Y_3 | X_3, X_4) \\
 &= I(X_1, X_2; Y_3 | X_3, X_4) \\
 R_4 &< I(X_4; Y_3 | X_2, X_3) \\
 R_1 + R_4 &< I(X_1; Y_3 | X_2, X_3, X_4) + I(X_2, X_4; Y_3 | X_3) \\
 &= I(X_1, X_2, X_4; Y_3 | X_3)
 \end{aligned}$$

4. In block b , node 4 can decode the message $w_{1,b-3}$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_4 | X_2, X_3, X_4) + I(X_2; Y_4 | X_3, X_4) + I(X_3; Y_4 | X_4)$$

where the three mutual informations lead to (4.1).

4.3.2 Case 1: A No-offset Encoding Scheme

Encoding

We use a no-offset encoding scheme to obtain an achievable region consistent with (4.10). The encoding scheme is depicted in table 4.2.

Decoding

1. At the end of each block $b = 3, \dots, B$, node 1 determines the unique index $\hat{w}_{4,b-2}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{X}_{1,b-2}, \mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-4})), \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-3})), \mathbf{x}_4(\hat{w}_{4,b-2}), \mathbf{Y}_{1,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_1) \\ & (\mathbf{X}_{1,b-1}, \mathbf{x}_2(k_2(w_{1,b-2}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-3}, \hat{w}_{4,b-2})), \mathbf{Y}_{1,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, Y_1) \\ & (\mathbf{X}_{1,b}, \mathbf{x}_2(k_2(w_{1,b-1}, \hat{w}_{4,b-2})), \mathbf{Y}_{1,b}) \in A_\epsilon^{(n)}(X_1, X_2, Y_1) \end{aligned}$$

2. At the end of each block $b = 1, \dots, B$, node 2 determines the unique index pair $(\hat{w}_{1,b}, \hat{w}_{4,b-1})$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(w_{1,b-1}), \mathbf{X}_{2,b-1}, \mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-2})), \mathbf{x}_4(\hat{w}_{4,b-1}), \mathbf{Y}_{2,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_2) \\ & (\mathbf{x}_1(\hat{w}_{1,b}), \mathbf{X}_{2,b}, \mathbf{x}_3(k_3(w_{1,b-2}, \hat{w}_{4,b-1})), \mathbf{Y}_{2,b}) \in A_\epsilon^{(n)}(X_1, X_2, X_3, Y_2) \end{aligned}$$

3. At the end of each block $b = 1, \dots, B$, node 3 determines the unique index pair $(\hat{w}_{1,b-1}, \hat{w}_{4,b})$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-1}), \mathbf{x}_2(k_2(w_{1,b-2}, w_{4,b-3})), \mathbf{X}_{3,b-1}, \mathbf{x}_4(w_{4,b-1}), \mathbf{Y}_{3,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_3) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-1}, w_{4,b-2})), \mathbf{X}_{3,b}, \mathbf{x}_4(\hat{w}_{4,b}), \mathbf{Y}_{3,b}) \in A_\epsilon^{(n)}(X_2, X_3, X_4, Y_3) \end{aligned}$$

4. At the end of each block $b = 3, \dots, B$, node 4 determines the unique index $\hat{w}_{1,b-2}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-2}), \mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-4})), \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-3})), \mathbf{X}_{4,b-2}, \mathbf{Y}_{4,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-2}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-2})), \mathbf{X}_{4,b-1}, \mathbf{Y}_{4,b-1}) \\ & \in A_\epsilon^{(n)}(X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_3(k_3(\hat{w}_{1,b-2}, w_{4,b-1})), \mathbf{X}_{4,b}, \mathbf{Y}_{4,b}) \in A_\epsilon^{(n)}(X_3, X_4, Y_4) \end{aligned}$$

Analysis of Probability of Error

1. In block b , node 1 can decode $w_{4,b-2}$ with arbitrarily small probability of error if

$$R_4 < I(X_4; Y_1 | X_1, X_2, X_3) + I(X_3; Y_1 | X_1, X_2) + I(X_2; Y_1 | X_1)$$

where the combination of the three mutual informations leads to (4.2).

2. In block b , node 2 can decode the pair $(w_{1,b}, w_{4,b-1})$ with arbitrarily small probability of error if

$$\begin{aligned} R_1 &< I(X_1; Y_2 | X_2, X_3) & (4.13) \\ R_4 &< I(X_4; Y_2 | X_1, X_2, X_3) + I(X_3; Y_2 | X_1, X_2) \\ &= I(X_3, X_4; Y_2 | X_1, X_2) \\ R_1 + R_4 &< I(X_4; Y_2 | X_1, X_2, X_3) + I(X_1, X_3; Y_2 | X_2) \\ &= I(X_1, X_3, X_4; Y_2 | X_2) \end{aligned}$$

3. In block b , node 3 can decode the pair $(w_{1,b-1}, w_{4,b})$ with arbitrarily small probability of error if

$$\begin{aligned} R_1 &< I(X_1; Y_3 | X_2, X_3, X_4) + I(X_2; Y_3 | X_3, X_4) \\ &= I(X_1, X_2; Y_3 | X_3, X_4) \\ R_4 &< I(X_4; Y_3 | X_2, X_3) \\ R_1 + R_4 &< I(X_1; Y_3 | X_2, X_3, X_4) + I(X_2, X_4; Y_3 | X_3) \\ &= I(X_1, X_2, X_4; Y_3 | X_3) \end{aligned}$$

4. In block b , node 4 can decode the message $w_{1,b-2}$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_4 | X_2, X_3, X_4) + I(X_2; Y_4 | X_3, X_4) + I(X_3; Y_4 | X_4)$$

where the combination of the three mutual informations leads to (4.1).

The combined rate region obtained from the offset and no-offset coding schemes is given by (4.1)-(4.8) and (4.10). This statement is verified by observing that (4.13) and (4.12) imply (4.5). In other words, we have shown that if (4.10) is true, there are two coding schemes that together recover the region defined by (4.1)-(4.8). In both schemes, node 3 decodes $w_{4,b}$ before node 2, which is consistent with the practical interpretation of case 1.

Table 4.3: A multiple-access encoding scheme for case 2

Node	Block $b - 2$	Block $b - 1$	Block b
1	$\mathbf{x}_1(w_{1,b-2})$	$\mathbf{x}_1(w_{1,b-1})$	$\mathbf{x}_1(w_{1,b})$
2	$\mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-3}))$	$\mathbf{x}_2(k_2(w_{1,b-2}, w_{4,b-2}))$	$\mathbf{x}_2(k_2(w_{1,b-1}, w_{4,b-1}))$
3	$\mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-4}))$	$\mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-3}))$	$\mathbf{x}_3(k_3(w_{1,b-2}, w_{4,b-2}))$
4	$\mathbf{x}_4(w_{4,b-2})$	$\mathbf{x}_4(w_{4,b-1})$	$\mathbf{x}_4(w_{4,b})$

4.3.3 Case 2: A Multiple-access Scheme

Encoding

We present a scheme that obtains an achievable rate region consistent with the first part of (4.11) given that neither (4.9) nor (4.10) holds. The encoding scheme is depicted in table 4.3.

Decoding

1. At the end of each block $b = 3, \dots, B$, node 1 determines the unique index $\hat{w}_{4,b-2}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{X}_{1,b-2}, \mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-4})), \mathbf{x}_4(\hat{w}_{4,b-2}), \mathbf{Y}_{1,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_1) \\ & (\mathbf{X}_{1,b-1}, \mathbf{x}_2(k_2(w_{1,b-2}, \hat{w}_{4,b-2})), \mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-3})), \mathbf{Y}_{1,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, Y_1) \\ & (\mathbf{X}_{1,b}, \mathbf{x}_3(k_3(w_{1,b-2}, \hat{w}_{4,b-2})), \mathbf{Y}_{1,b}) \in A_\epsilon^{(n)}(X_1, X_3, Y_1) \end{aligned}$$

2. At the end of each block $b = 1, \dots, B$, node 2 determines the unique index pair $(\hat{w}_{1,b}, \hat{w}_{4,b})$ that satisfies the joint typicality check:

$$(\mathbf{x}_1(\hat{w}_{1,b}), \mathbf{X}_{2,b}, \mathbf{x}_3(k_3(w_{1,b-2}, w_{4,b-2})), \mathbf{x}_4(\hat{w}_{4,b}), \mathbf{Y}_{2,b}) \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_2)$$

3. At the end of each block $b = 2, \dots, B$, node 3 determines the unique index

pair $(\hat{w}_{1,b-1}, \hat{w}_{4,b-1})$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-1}), \mathbf{x}_2(k_2(w_{1,b-2}, w_{4,b-2})), \mathbf{X}_{3,b-1}, \mathbf{x}_4(\hat{w}_{4,b-1}), \mathbf{Y}_{3,b-1}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_3) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-1}, \hat{w}_{4,b-1})), \mathbf{X}_{3,b}, \mathbf{Y}_{3,b}) \in A_\epsilon^{(n)}(X_2, X_3, Y_3) \end{aligned}$$

4. At the end of each block $b = 3, \dots, B$, node 4 determines the unique index $\hat{w}_{1,b-2}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{x}_1(\hat{w}_{1,b-2}), \mathbf{x}_2(k_2(w_{1,b-3}, w_{4,b-3})), \mathbf{x}_3(k_3(w_{1,b-4}, w_{4,b-4})), \mathbf{X}_{4,b-2}, \mathbf{Y}_{4,b-2}) \\ & \in A_\epsilon^{(n)}(X_1, X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_2(k_2(\hat{w}_{1,b-2}, w_{4,b-2})), \mathbf{x}_3(k_3(w_{1,b-3}, w_{4,b-3})), \mathbf{X}_{4,b-1}, \mathbf{Y}_{4,b-1}) \\ & \in A_\epsilon^{(n)}(X_2, X_3, X_4, Y_4) \\ & (\mathbf{x}_3(k_3(\hat{w}_{1,b-2}, w_{4,b-2})), \mathbf{X}_{4,b}, \mathbf{Y}_{4,b}) \in A_\epsilon^{(n)}(X_3, X_4, Y_4) \end{aligned}$$

Analysis of Probability of Error

1. In block b , node 1 can decode $w_{4,b-2}$ with arbitrarily small probability of error if

$$R_4 < I(X_4; Y_1 | X_1, X_2, X_3) + I(X_2; Y_1 | X_1, X_3) + I(X_3; Y_1 | X_1)$$

where the combination of the three mutual informations leads to (4.2).

2. In block b , node 2 can decode the pair $(w_{1,b}, w_{4,b})$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_2 | X_2, X_3, X_4) \quad (4.14)$$

$$R_4 < I(X_4; Y_2 | X_1, X_2, X_3) \quad (4.15)$$

$$R_1 + R_4 < I(X_1, X_4; Y_2 | X_2, X_3) \quad (4.16)$$

where each inequality corresponds to one of the three ways a message pair can be decoded incorrectly.

3. In block b , node 3 can decode the pair $(w_{1,b-1}, w_{4,b-1})$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_3 | X_2, X_3, X_4) + I(X_2; Y_3 | X_3) \quad (4.17)$$

$$R_4 < I(X_4; Y_3 | X_1, X_2, X_3) + I(X_2; Y_3 | X_3) \quad (4.18)$$

$$\begin{aligned} R_1 + R_4 & < I(X_1, X_4; Y_3 | X_2, X_3) + I(X_2; Y_3 | X_3) \quad (4.19) \\ & = I(X_1, X_2, X_4; Y_3 | X_3) \end{aligned}$$

where each inequality corresponds to one of the three ways a message pair can be decoded incorrectly.

4. In block b , node 4 can decode the message $w_{1,b-2}$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_4 | X_2, X_3, X_4) + I(X_2; Y_4 | X_3, X_4) + I(X_3; Y_4 | X_4)$$

where the combination of the three mutual informations leads to (4.1).

Observe that node 2 decodes both source messages $(w_{1,b}, w_{4,b})$ before node 3, which is consistent with the practical interpretation of case 2. Now from the definition of this case we know that (R_1, R_4) satisfies the following inequalities:

$$R_1 \geq I(X_1; Y_2 | X_2, X_3) \tag{4.20}$$

$$R_4 \geq I(X_4; Y_3 | X_2, X_3) \tag{4.21}$$

$$R_1 + R_4 < I(X_1, X_4; Y_2 | X_2, X_3) \tag{4.22}$$

The constraints (4.20)-(4.22) are sufficient to obtain the region defined by (4.1)-(4.8) using the third coding scheme. This statement is proved by observing that (4.20) and (4.22) imply (4.15), (4.8) and (4.21) imply (4.17), and (4.7) implies (4.18). Furthermore, if either (4.20) or (4.21) does not hold, then the rate pair will be included in case 1.

Therefore it follows from symmetry and the achievable rate regions in case 1 and 2, that one of the constraints (4.9)-(4.11) is sufficient to obtain the region defined by (4.1)-(4.8). Thus theorem 4.2.1 is proved.

4.4 Summarizing Remarks

We started this chapter looking to extend the analysis of the two-way one-relay channel to the two-way *two*-relay channel. In particular, we decided to focus on the network coding strategy because of its convenient implementation and the simple interpretation of its achievable region in the one relay case. There were two questions that motivated our study in this chapter: is it possible to easily extend the achievable region if we add another relay to the network; and how should we schedule the encoding and decoding in the two-relay case. We found insightful answers to both questions.

Our attempt to extend the achievable region to the two-relay case encountered a fundamental deadlock problem in which each relay needed to decode before the other in order to enable mutual assistance. However, we resolved this deadlock problem by adding an additional constraint that ensured some relay could decode at least one message before the other relay. This resolution is quite natural in the light of the conclusions in chapter 2; there is no desirable alternative to the present strategy, in which the relay fully decodes the message intended for the node it is trying to help.

Furthermore, we also proposed some new encoding schemes to prove that the additional constraint was sufficient. These coding schemes described how to schedule the encoding and decoding at each node, proving that information theory provides some unique insight into the operation of wireless multi-hop networks.

Chapter 5

The Way Forward

Information theory has developed enormously since the 1948 paper that started it. The more famous results, particularly in multi-user information theory, showcase some ingenious and creative thinking . Disappointingly, most of these results have yet to be applied in practice. Although multi-user information theory is a research intensive area with recent breakthroughs, most modern day communication protocols are still designed outside of an information theoretic framework. There are probably many reasons for this dichotomy. For one, it's possible that information theory is still ahead of its time; technology has not advanced enough to support the computational requirements.

Another reason could be that researchers have expended too much of effort in finding the final solution to every problem. It is now widely accepted that a final solution is not on the horizon and may never be reached in our lifetime. A more promising endeavor should play to information theory's strengths instead of succumbing to its weaknesses. Perhaps we should refocus on shedding light into the operation of wireless networks, maybe through innovative coding schemes or simple achievable regions with insightful and extendable interpretations. That perspective, in any event, was the motivation underlying the research presented in this thesis.

There is still much work to be done. In the future, we hope to extend our results further to the two-way multiple-level relay channel, and possibly to general networks. We may also revisit some of the "old" problems encountered in chapter 2; sometimes the answers to old problems are found while exploring new opportunities.

APPENDICES

Appendix A

Analysis of Error Probability for the One-Sided Interference Channel

By the symmetry of the random code construction, the probability of error does not depend on which pair of indices is sent by the sources. So, without loss of generality, we can assume that $(i, j) = (1, 1)$ was sent.

Each row and column in figure 2.1 correspond to the index of sender 1 and 2 respectively and their associated codewords. The decoding strategy determines which row in figure 2.1 contains the largest number of codewords that are typical with the received sequence. An error occurs if the “wrong” row contains more typical codewords than the “correct” row (row 1). Since the codebook is generated randomly, by symmetry we only need to evaluate the probability that in a given column, the correct row does not contain a typical codeword pair but some wrong row does. We define the following events:

- Let \mathcal{E}_1 be the event that $(x_1(i), x_1(j))$ belongs to $A_\epsilon^{(n)}$ for $i, j \neq (1, 1)$; for a wrong column, any wrong row contains a codeword pair typical with the received sequence.
- Let \mathcal{E}_2 be the event that $(x_1(1), x_1(j))$ belongs to $A_\epsilon^{(n)}$ for $j \neq 1$; for a wrong column, the correct row contains a codeword pair typical with the received sequence.
- Let \mathcal{E}_3 be the event that in a wrong column the correct row does not contain a typical codeword pair but *at least one* wrong row does.

- Let \mathcal{E}_4 be the event that for the correct column, the correct row does not contain a typical pair but at least one row does.

The decoding strategy will succeed if $P(\mathcal{E}_3 \cup \mathcal{E}_4) \rightarrow 0$ as n becomes large. Using the basic properties of typical sequences in lemma 2.1.1 we make the following observations:

$$P(\mathcal{E}_1) = 2^{-nI(X_1, X_2; Y) - 4\epsilon} \quad (\text{A.1})$$

$$P(\mathcal{E}_2) = 2^{-n(X_2; Y | X_1) - 3\epsilon} \quad (\text{A.2})$$

$$P(\mathcal{E}_4) \leq \epsilon \quad (\text{A.3})$$

Now since,

$$\mathcal{E}_1 = \mathcal{E}_1 \cap (\mathcal{E}_1 \cup \mathcal{E}_2) \quad (\text{A.4})$$

we have

$$P(\mathcal{E}_1) = P(\mathcal{E}_1 \cup \mathcal{E}_2)P(\mathcal{E}_1 | \mathcal{E}_1 \cup \mathcal{E}_2) \quad (\text{A.5})$$

Now,

$$P(\mathcal{E}_1 \cup \mathcal{E}_2) = P(\mathcal{E}_1) + P(\mathcal{E}_2) - P(\mathcal{E}_1)P(\mathcal{E}_2) \quad (\text{A.6})$$

and

$$P(\mathcal{E}_1 | \mathcal{E}_1 \cup \mathcal{E}_2) = \frac{P(\mathcal{E}_1)}{P(\mathcal{E}_1) + P(\mathcal{E}_2) - P(\mathcal{E}_1)P(\mathcal{E}_2)} \quad (\text{A.7})$$

If we substitute (A.1) and (A.2) into (A.6) and (A.7), and let n become large then

$$P(\mathcal{E}_1 | \mathcal{E}_1 \cup \mathcal{E}_2) \rightarrow 2^{-nI(X_1; Y) - 7\epsilon} \quad (\text{A.8})$$

$$P(\mathcal{E}_1 \cup \mathcal{E}_2) \rightarrow 0 \quad (\text{A.9})$$

$$P(\mathcal{E}_2) \rightarrow 0 \quad (\text{A.10})$$

Therefore the probability that in a wrong column, the correct row does not contain a typical codeword pair but at least one wrong row does, is given by,

$$P(\mathcal{E}_3) = [1 - (1 - P(\mathcal{E}_1))^{2^{nR_1} - 1}](1 - P(\mathcal{E}_2)) \quad (\text{A.11})$$

$$= [1 - (1 - P(\mathcal{E}_1 \cup \mathcal{E}_2)P(\mathcal{E}_1 | \mathcal{E}_1 \cup \mathcal{E}_2))^{2^{nR_1} - 1}](1 - P(\mathcal{E}_2)) \quad (\text{A.12})$$

$$\leq e^{-2^{n(I(X_1; Y) - R_1) + 7\epsilon}} \quad (\text{A.13})$$

where (A.12) follows by substituting (A.5) into (A.11). Furthermore (A.13) follows from lemma 10.5.3 in [15], and by substituting the limits in (A.8), (A.9), and (A.10) in (A.12). Finally, the probability that in any column, the correct row does not contain a typical codeword pair but some wrong row does, is upper bounded by

$$P(\mathcal{E}_3 \cup \mathcal{E}_4) \leq P(\mathcal{E}_3) + P(\mathcal{E}_4) \tag{A.14}$$

$$\tag{A.15}$$

If we substitute (A.3) and (A.13) into (A.14) and let n become large, we see that the probability of error goes to zero if

$$R_1 < I(X_1; Y) \tag{A.16}$$

Appendix B

Proof of Theorem 3.2.1

We use a block coding argument. We consider B blocks of transmission, each of n transmission slots. A sequence of $B - 1$ index pairs, $(w_{1,b}, w_{3,b}), b = 1, 2, \dots, B - 1$ will be sent over in nB transmission slots. Note that as $B \rightarrow \infty$, the rate $\frac{nR(B-1)}{nB}$ is arbitrarily close to R for any n .

B.1 Generation of Codebooks

- At each node $i \in \{1, 3\}$ generate at random 2^{nR_i} independent and identically distributed (i.i.d) n -sequences $\mathbf{u}_i(w_i), w_i \in \{1, 2, \dots, 2^{nR_i}\}$, each drawn according to

$$\text{Prob}(\mathbf{u}_i) = \prod_{j=1}^n p(u_i, j)$$

- For each $\mathbf{u}_i(w_i)$, generate 2^{nR_i} conditionally independent n -sequences $\mathbf{x}_i(s_i|w_i), s_i \in \{1, 2, \dots, 2^{nR_i}\}$, drawn independently according to

$$\text{Prob}(\mathbf{x}_i|\mathbf{u}_i(w_i)) = \prod_{j=1}^n p(x_{i,j}|u_{i,j}(w_i)).$$

- At node 2, for each $\mathbf{u}_1(w_1)$ and $\mathbf{u}_3(w_3)$ generate *one* conditionally independent n -sequence $\mathbf{x}_2(w_1, w_3)$, drawn independently according to

$$\text{Prob}(\mathbf{x}_2|\mathbf{u}_1(w_1), \mathbf{u}_3(w_3)) = \prod_{j=1}^n p(x_{2,j}|u_{1,j}(w_1), u_{3,j}(w_3)).$$

This defines the joint codebook for nodes 1,2 and 3 as

$$\mathcal{C}_0 := \{\mathbf{u}_1(w_1), \mathbf{x}_1(s_1|w_1), \mathbf{u}_3(w_3), \mathbf{x}_3(s_3|w_3), \mathbf{x}_2(w_1, w_3)\}$$

Table B.1: Encoding process for the TWORC

block 1	block 2	...	block B
$\mathbf{x}_1(w_{1,1} 1)$	$\mathbf{x}_1(w_{1,2} w_{1,1})$...	$\mathbf{x}_1(w_{1,B} w_{1,B-1})$
$\mathbf{u}_1(1)$	$\mathbf{u}_1(w_{1,1})$...	$\mathbf{u}_1(w_{1,B-1})$
$\mathbf{x}_3(w_{3,1} 1)$	$\mathbf{x}_3(w_{3,2} w_{3,1})$...	$\mathbf{x}_3(w_{3,B} w_{3,B-1})$
$\mathbf{u}_3(1)$	$\mathbf{u}_3(w_{3,1})$...	$\mathbf{u}_3(w_{3,B-1})$
$\mathbf{x}_2(1, 1)$	$\mathbf{x}_2(w_{1,1}, w_{3,1})$...	$\mathbf{x}_2(w_{1,B-1}, w_{3,B-1})$

Repeating the above process *independently* once more, we generate another random codebook \mathcal{C}_1 similar to \mathcal{C}_0 . We will use these two codebooks alternatively as follows: In block $b = 1, 2, \dots, B$, the codebook $\mathcal{C}_{(b \bmod 2)}$ is used. Hence, in any two consecutive blocks, codewords from different blocks are independent. This is a property we will use in the analysis of the probability of error. Before the transmission, the joint codebooks $\mathcal{C}_0, \mathcal{C}_1$ are revealed to all the nodes 1, 2 and 3.

B.2 Encoding

The encoding process is depicted in table B.1. At the beginning of each block $b \in \{1, 2, \dots, B\}$, assume that node 2 has correctly decoded the pair $(w_{1,b-1}, w_{3,b-1})$ and sends the following n -sequences from the codebook $\mathcal{C}_{(b \bmod 2)}$ in the block:

$$\mathbf{X}_{2,b} = \mathbf{x}_2(w_{1,b-1}, w_{3,b-1}).$$

Also, in the same block, node 1 and 3 send the following n -sequences from the same codebook $\mathcal{C}_{(b \bmod 2)}$:

$$\mathbf{X}_{1,b} = \mathbf{x}_1(w_{1,b}|w_{1,b-1})$$

$$\mathbf{X}_{3,b} = \mathbf{x}_3(w_{3,b}|w_{3,b-1}).$$

B.3 Decoding

At each node we assume that all previous message estimates are correct.

1. At the end of each block $b = 2, 3, \dots, B$, node 1 determines the unique index $\hat{w}_{3,b-1}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{u}_1(w_{1,b-2}), \mathbf{x}_1(w_{1,b-1}|w_{1,b-2}), \mathbf{u}_3(w_{3,b-2}), \mathbf{x}_3(\hat{w}_{3,b-1}|w_{3,b-2}), \mathbf{x}_2(w_{1,b-2}, w_{3,b-2}), \\ & \quad \mathbf{Y}_{1,b-1}) \in A_\epsilon^{(n)}(U_1, U_3, X_1, X_2, X_3, Y_1) \\ & (\mathbf{u}_1(w_{1,b-1}), \mathbf{x}_1(w_{1,b}|w_{1,b-1}), \mathbf{u}_3(\hat{w}_{3,b-1}), \mathbf{x}_2(w_{1,b-1}, \hat{w}_{3,b-1}), \mathbf{Y}_{1,b}) \\ & \quad \in A_\epsilon^{(n)}(U_1, U_3, X_1, X_2, Y_1) \end{aligned}$$

2. At the end of each block $b = 2, 3, \dots, B$, node 2 determines the unique index pair $(\hat{w}_{1,b}, \hat{w}_{3,b})$ that satisfies the joint typicality check:

$$\begin{aligned} & (\mathbf{u}_1(w_{1,b-1}), \mathbf{x}_1(\hat{w}_{1,b}|w_{1,b-1}), \mathbf{u}_3(w_{3,b-1}), \mathbf{x}_3(\hat{w}_{3,b}|w_{3,b-1}), \mathbf{x}_2(w_{1,b-1}, w_{3,b-1}), \\ & \quad \mathbf{Y}_{2,b}) \in A_\epsilon^{(n)}(U_1, U_3, X_1, X_2, X_3, Y_2) \end{aligned}$$

3. At the end of each block $b = 2, 3, \dots, B$, node 3 determines the unique index $\hat{w}_{1,b-1}$ that satisfies the joint typicality checks:

$$\begin{aligned} & (\mathbf{u}_1(w_{1,b-2}), \mathbf{x}_1(\hat{w}_{1,b-1}|w_{1,b-2}), \mathbf{u}_3(w_{3,b-2}), \mathbf{x}_3(w_{3,b-1}|w_{3,b-2}), \mathbf{x}_2(w_{1,b-2}, w_{3,b-2}), \\ & \quad \mathbf{Y}_{3,b-1}) \in A_\epsilon^{(n)}(U_1, U_3, X_1, X_2, X_3, Y_3) \\ & (\mathbf{u}_1(\hat{w}_{1,b-1}), \mathbf{u}_3(w_{3,b-1}), \mathbf{x}_3(w_{3,b}|w_{3,b-1}), \mathbf{x}_2(\hat{w}_{1,b-1}, w_{3,b-1}), \mathbf{Y}_{1,b}) \\ & \quad \in A_\epsilon^{(n)}(U_1, U_3, X_2, X_2, Y_3) \end{aligned}$$

B.4 Analysis of Probability of Error

1. In block b , node 1 can decode $w_{3,b-1}$ with arbitrarily small probability of error if

$$R_3 < I(X_3; Y_1 | U_1, U_3, X_1, X_2) + I(U_3, X_2; Y_1 | U_1, X_1) \quad (\text{B.1})$$

$$\begin{aligned} & = I(U_3, X_2, X_3; Y_1 | U_1, X_1) \\ & = I(X_2, X_3; Y_1 | U_1, X_1) + I(U_3; Y_1 | U_1, X_1, X_2, X_3) \\ & = I(X_2, X_3; Y_1 | U_1, X_1) \end{aligned} \quad (\text{B.2})$$

where the two mutual informations in (B.1) follow from the two typicality checks, and (B.2) follows from the Markov chain $U_3 - (U_1, X_1, X_2, X_3) - Y_1$

2. In block b , node 2 can decode $(w_{1,b}, w_{3,b})$ with arbitrarily small probability of error if

$$R_1 < I(X_1; Y_2 | U_1, U_3, X_2, X_3) \quad (\text{B.3})$$

$$= I(X_1; Y_2 | U_1, U_3, X_2, X_3) + I(U_3; Y_2 | U_1, X_2, X_3) \quad (\text{B.4})$$

$$= I(U_3, X_1; Y_2 | U_1, X_2, X_3)$$

$$= I(X_1; Y_2 | U_1, X_2, X_3) + I(U_3; Y_2 | U_1, X_1, X_2, X_3) \quad (\text{B.5})$$

$$= I(X_1; Y_2 | U_1, X_2, X_3)$$

$$R_3 < I(X_3; Y_2 | U_1, U_3, X_1, X_2) \quad (\text{B.6})$$

$$= I(X_3; Y_2 | U_1, U_3, X_1, X_2) + I(U_1; Y_2 | U_3, X_1, X_2) \quad (\text{B.7})$$

$$= I(U_1, X_3; Y_2 | U_3, X_1, X_2)$$

$$= I(X_3; Y_2 | U_3, X_1, X_2) + I(U_1; Y_2 | U_3, X_1, X_2, X_3) \quad (\text{B.8})$$

$$= I(X_3; Y_2 | U_3, X_1, X_2)$$

$$R_1 + R_3 < I(X_1, X_3; Y_3 | U_1, U_3, X_2) \quad (\text{B.9})$$

where the three mutual informations in (B.3) (B.6), and (B.9) follow from the three ways a message pair can be decoded incorrectly. Furthermore, (B.4) and (B.5) follow from the Markov chains $U_3 - (U_1, X_2, X_3) - Y_2$ and $U_3 - (U_1, X_1, X_2, X_3) - Y_2$ respectively. In addition, (B.7) and (B.8) follow from the Markov chains $U_1 - (U_3, X_1, X_2) - Y_2$ and $U_1 - (U_3, X_1, X_2, X_3) - Y_2$ respectively

3. In block b , node 3 can decode $w_{1,b-1}$ with arbitrarily small probability of error if

$$R_3 < I(X_1; Y_3 | U_1, U_3, X_2, X_3) + I(U_1, X_2; Y_3 | U_3, X_2) \quad (\text{B.10})$$

$$= I(U_1, X_1, X_2; Y_3 | U_3, X_3)$$

$$= I(X_1, X_2; Y_3 | U_3, X_3) + I(U_1; Y_3 | U_3, X_1, X_2, X_3)$$

$$= I(X_1, X_2; Y_3 | U_3, X_3) \quad (\text{B.11})$$

where the two mutual informations in (B.10) follow from the two typicality checks, and (B.11) follows from the Markov chain $U_1 - (U_3, X_1, X_2, X_3) - Y_3$

References

- [1] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, Oct. 1948. 1
- [2] ———, “Two-way communication channels,” in *Proc. 4th Berkeley Symp. on Mathematical Statistics and Probability*, J. Neyman, Ed., vol. 1, Berkeley, CA, 1961, pp. 611–644. 3
- [3] R. Ahlswede, “Multi-way communication channels,” in *Proc. 2nd International Symposium on Information Theory (Tsahkadsor, Armenian S.S.R)*, Budapest, Hungary, Sep. 1971, pp. 23–52. 3
- [4] H. Liao, “Multiple access channels,” Ph.D. dissertation, Department of Electrical Engineering, University of Hawaii, Honolulu, 1972. 3
- [5] T. M. Cover, “Broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 18, pp. 2–14, Jan. 1972. 3
- [6] E. C. van der Meulen, “Three-terminal communication channels,” *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971. 3
- [7] H. Sato, “Two-user communication channels,” *IEEE Trans. Inf. Theory*, vol. 23, pp. 295–304, May 1977. 3
- [8] A. B. Carleial, “Interference channels,” *IEEE Trans. Inf. Theory*, vol. 24, pp. 60–70, Jan. 1978. 3
- [9] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inf. Theory*, vol. 46, pp. 388–404, Mar. 2000. 5
- [10] L.-L. Xie and P. R. Kumar, “A network information theory for wireless communication: Scaling laws and optimal operation,” *IEEE Trans. Inf. Theory*, vol. 50, pp. 748–767, May 2004. 5

- [11] —, “An achievable rate for the multiple-level relay channel,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 1348–1358, Apr. 2005. 5
- [12] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, Jul. 2000. 5, 15
- [13] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna gaussian broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 1691–1706, Jul. 2003. 5
- [14] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the gaussian multiple-input multiple-output broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 52, pp. 3936–3964, Sep. 2006. 5
- [15] T. M. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: John Wiley and Sons, 2006. 8, 13, 20, 40
- [16] T. M. Cover, “The capacity of the relay channel,” in *Open Problems in Communication and Computation*, T. M. Cover and B. Gopinath, Eds. Springer-Verlag, 1988, pp. 72–73. 14
- [17] Y. Wu, P. A. Chou, and S.-Y. Kung, “Information exchange in wireless networks with network coding and physical-layer broadcast,” in *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2005. 16
- [18] L. L. Xie, “Network coding and random binning for multi-user channels,” in *Proc. 10th Canadian Workshop on Information Theory*, Edmonton, Canada, Jun. 2007, pp. 85–88. 17, 24
- [19] A. El Gamal and T. M. Cover, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Jan. 1979. 18
- [20] B. Rankov and A. Wittneben, “Achievable rate regions for the two-way relay channel,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Seattle, WA, Jul. 2006. 18, 19
- [21] J. Ponniah and L.-L. Xie, “An achievable rate region for the two-way two-relay channel,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, Jul. 2008. 19, 25
- [22] L.-L. Xie and P. R. Kumar, “Multisource, multidestination, multirelay wireless networks,” *IEEE Trans. Inf. Theory*, vol. 53, pp. 3586–3595, Oct. 2007. 24

- [23] L. Sankar, G. Kramer, and N. B. Mandayam, “Offset encoding for multiaccess relay channels,” *IEEE Trans. Inf. Theory*, vol. 53, pp. 3814–3821, Oct. 2007.