MULTIUSER WIRELESS NETWORKS: THE USER COOPERATION PERSPECTIVE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Lifeng Lai,
B.E., M.E.

* * * * *

The Ohio State University

2007

Dissertation Committee:

Hesham El Gamal, Adviser
Philip Schniter
Randolph L. Moses
Andrea Serrani

Approved by

Adviser
Graduate Program in Electrical Engineering
© Copyright by
Lifeng Lai
2007
ABSTRACT

The presence of multiple users in the wireless network offers users with an opportunity to cooperate. Adopting tools from network information theory, this thesis investigates the utilities of user cooperation under various scenarios and examines the validity of cooperation assumption in energy limited networks.

We first consider the construction of cooperation strategies for a wireless network composed of three nodes and limited by half-duplex and total power constraints. We show that noisy feedback plays a crucial role in devising efficient cooperation schemes. We focus on three special cases, namely 1) Relay Channel, 2) Multicast Channel, and 3) Three-way Channel. These special cases are judicially chosen to reflect varying degrees of complexity while highlighting the common ground shared by the different variants of the three node wireless network. For the relay channel, we propose a new cooperation scheme that exploits the wireless feedback gain. This scheme combines the benefits of decode-and-forward and compress-and-forward strategies and avoids the noiseless feedback assumption adopted in earlier works. Our analysis of the achievable rate of this scheme reveals the diminishing feedback gain at both the low and high signal-to-noise ratio regimes. Inspired by the proposed feedback strategy, we identify a greedy cooperation framework applicable to both the multicast and three-way channels. Our performance analysis reveals the asymptotic optimality of the proposed greedy approach and the central role of list source-channel decoding in exploiting the receiver side information in the wireless network setting.
We then establish the utility of user cooperation in facilitating secure wireless communications. In particular, the four-terminal relay-eavesdropper channel is introduced and an outer-bound on the optimal rate-equivocation region is derived. Several cooperation strategies are then devised and the corresponding achievable rate-equivocation region are characterized. Of particular interest is the novel noise forwarding strategy. This strategy is used to illustrate the deaf helper phenomenon, where the relay is able to facilitate secure communications while being totally ignorant of the transmitted messages. Furthermore, the noise forwarding scheme is shown to increase the secrecy capacity in the reversely degraded scenario, where the relay node fails to offer performance gains in the classical setting. The gain offered by the proposed cooperation strategies is then proved theoretically and validated numerically in the Additive White Gaussian Noise (AWGN) channel.

We also examine user cooperation from game theoretic perspective. In many practical scenarios, nodes’ limited power raises doubts on whether each node will be willing to spend its valuable energy in cooperation. Using game theoretic framework, the critical role of altruistic nodes in encouraging cooperation is established, both for small and large scale networks. In a small network, where nodes utilize decode-and-forward scheme to cooperate, we show that a relay node, with appropriate strategy and location, successfully turns the Nash Equilibrium from no-cooperation to full-cooperation. In the large scale network, we show that it is sufficient to have a vanishingly small fraction of the nodes to be altruistic, i.e., relay nodes, in order to ensure full cooperation from all the nodes in the network. An important aspect of our work is that only reward/punishment policies that can be realized on the physical layer are used, and hence, our results establish the achievability of full cooperation without requiring additional incentive mechanisms at the application layer.
To my wife Hongfei, for her longstanding support.
ACKNOWLEDGMENTS

First, I would like to thank my adviser Prof Hesham El Gamal for his guidance and support during the passing three years. His dedication to research, his unique view of research problems, his insightful vision and his inspiration have led me through my Ph.D study, and I hope, will accompany me throughout my future career. I will miss the numerous stimulating discussions through meetings and emails with him. Without his support and encouragement, this thesis would have been impossible.

I was fortunate to collaborate with Dr. Ke Liu during my first year here, which not only led to the first part of this thesis but also taught me what a good researcher should be. In the mean time, I thank Professor Randolph Moses, Professor Philip Schniter and Professor Andrea Serrani for agreeing to be in my thesis committee, for spending their time on reading my thesis and for the wonderful classes they offered.

I thank all the members of the IPS lab for their friendship. The brainstorm time and sports time with them are really fun. I also thank Jeri for the wonderful job on coordinating the IPS lab, which makes it a family and enjoyable place to work.

Last but not least, I thank my parent, my wife and two sisters for their love and support all through my life.
VITA

May, 1980 ................................. Born - Lishui, China

2001 ................................. B.E., Electrical Engineering,
                    Zhejiang University,
                    Hangzhou, China

2004 ................................. M.E., Electrical Engineering,
                    Zhejiang University,
                    Hangzhou, China

2004-present .......................... Distinguished University Fellow,
                    The Ohio State University,
                    Columbus, USA

PUBLICATIONS

Research Publications


**FIELDS OF STUDY**

Major Field: Electrical Engineering

Studies in:

Comm. and Signal Proc.  Prof. Hesham El Gamal
Comm. and Signal Proc.  Prof. Randolph Moses
Comm. and Signal Proc.  Prof. Philip Schniter
Control Theory         Prof. Andrea Serrani
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapters</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Cooperative Communications</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Cooperation for Secure Communications</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Cooperation: Game Theoretic Perspective</td>
<td>9</td>
</tr>
<tr>
<td>1.4 Organization and Contributions of This Thesis</td>
<td>11</td>
</tr>
<tr>
<td>2. Cooperation Strategies for a Three-node Wireless Network</td>
<td>14</td>
</tr>
<tr>
<td>2.1 The Three Node Wireless Network</td>
<td>15</td>
</tr>
<tr>
<td>2.2 The Relay Channel</td>
<td>20</td>
</tr>
<tr>
<td>2.3 The Multicast Channel</td>
<td>33</td>
</tr>
<tr>
<td>2.4 The Three-way Channel</td>
<td>39</td>
</tr>
<tr>
<td>2.4.1 Multicast with Side-information</td>
<td>42</td>
</tr>
<tr>
<td>2.4.2 Multicast Scheduler</td>
<td>47</td>
</tr>
<tr>
<td>2.5 Summary</td>
<td>51</td>
</tr>
</tbody>
</table>

viii
3. Cooperation for Secrecy .................................................. 52
   3.1 The Relay-Eavesdropper Channel .................................. 52
   3.2 Outer Bound and Cooperation Strategies ......................... 55
   3.3 Examples .................................................................. 60
      3.3.1 The Deaf Helper Phenomenon ................................. 60
      3.3.2 The Reversely Degraded Relay-Eavesdropper Channel .... 61
      3.3.3 The AWGN Channel .............................................. 62
   3.4 Summary .................................................................. 69

4. Cooperation in Energy Limited Networks .............................. 70
   4.1 Network Model .......................................................... 71
   4.2 The Small Network Scenario ....................................... 74
      4.2.1 The Cooperation Scheme ....................................... 75
      4.2.2 Equilibrium Analysis ............................................ 78
      4.2.3 Fading Channels ................................................ 83
      4.2.4 Numerical Results ................................................. 84
   4.3 Large Random Network .............................................. 86
   4.4 Summary .................................................................. 93

5. Conclusions .................................................................. 95

Appendices:

A. Proof of Theorems .......................................................... 99
   A.1 Proof of Theorem 2 ..................................................... 99
      A.1.1 Discrete Memoryless Channel ............................... 99
   A.2 Proof of Lemma 3 ....................................................... 105
   A.3 Proof of Lemma 4 ....................................................... 108
   A.4 Proof of Theorem 5 ..................................................... 108
   A.5 Proof of Theorem 7 ..................................................... 109
   A.6 Proof of Theorem 8 ..................................................... 109
   A.7 Proof of Theorem 10 ................................................... 110
   A.8 Proof of Lemma 12 .................................................... 111
      A.8.1 Source coding ...................................................... 111
      A.8.2 Channel coding ................................................... 111
      A.8.3 Decoding ........................................................... 112
      A.8.4 Calculation of Probability of Error ......................... 113
   A.9 Proof of Theorem 14 ................................................... 116

ix
A.10 Proof of Theorem 15 ........................................... 117
A.11 Proof of Theorem 16 ........................................... 119
A.12 Proof of Theorem 17 ........................................... 121
A.13 Proof of Theorem 18 ........................................... 125
A.14 Proof of Theorem 19 ........................................... 130
A.15 Proof of Theorem 20 ........................................... 134
A.16 Proof of Lemma 26 ........................................... 140
A.17 Proof of Theorem 27 ........................................... 142

Bibliography ............................................................. 144


LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>78</td>
</tr>
</tbody>
</table>

The game payoff matrix with one relay node.
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The relay channel.</td>
</tr>
<tr>
<td>1.2</td>
<td>Shannon’s model for security system.</td>
</tr>
<tr>
<td>1.3</td>
<td>Wyner’s wiretap model.</td>
</tr>
<tr>
<td>1.4</td>
<td>The relay-eavesdropper channel.</td>
</tr>
<tr>
<td>2.1</td>
<td>An illustration of the three-node (half-duplex) wireless network. Each node may be interested in a subset or all the observation variables distributed across the network.</td>
</tr>
<tr>
<td>2.2</td>
<td>The operation sequence of the half-duplex relay channel with noisy feedback.</td>
</tr>
<tr>
<td>2.3</td>
<td>A geometric representation of FB-, DF- and CF-relay schemes. The solid lines are for $R_{DF}$ in (2.8), the dash-dotted for $R_{FB}$ in (2.15), and the dashed for the upperbound of $R_{CF}$ in (2.19). The various endpoints in the figure are (a) $C(h_{13}^2P_1^{(2)} + 2r_{12}h_{13}h_{23}\sqrt{P_1^{(2)}P_2^{(2)} + h_{23}^2P_2^{(2)}})$, (b) $C(h_{13}^2P_1^{(1)})$, (c) $C((1 - r_{12}^2)h_{13}^2P_1^{(2)})$, (d) $C(h_{12}^2P_1^{(1)})$, (e) $\alpha C((h_{13}^2 + h_{12}^2)P_1^{(1)}) + (1 - \alpha)C(h_{12}^2P_1^{(3)})$, (f) $\alpha C(h_{13}^2P_1^{(1)})$, (g) $C((h_{13}^2 + h_{12}^2)P_1^{(1)})$, (h) $C(h_{13}^2P_1^{(2)})$, and (i) $C(h_{13}^2P_1^{(2)} + h_{23}^2P_2^{(2)})$.</td>
</tr>
<tr>
<td>2.4</td>
<td>The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.</td>
</tr>
<tr>
<td>2.5</td>
<td>The ratio of achievable rate of various schemes to the cut-set upper-bound, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.</td>
</tr>
<tr>
<td>2.6</td>
<td>The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, SNR = 0 dB.</td>
</tr>
</tbody>
</table>
2.7 The position of the relay where different scheme achieves the largest achievable rate, $P = 3$ dB, $\beta = 4$. .................................................. 33

2.8 The achievable rate of various schemes in the multicast channel, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB. .................................................. 37

2.9 The ratio of the achievable rates of various schemes to the non-cooperation scheme, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB. ....................... 38

2.10 The achievable rate of various schemes in the multicast channel, $h_{12} = 0$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB. .................................................. 39

2.11 The achievable rate of various schemes in the multicast channel, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and SNR = 1.8 dB. .............................. 40

2.12 The genie-aided bound in node-1, in which node-2 and node-3 can fully cooperate with each other. ................................. 41

2.13 The bandwidth expansion factor of various schemes in the multicast channel with side-information, $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 19.5$ dB, $H(S_1|S_2) = 0.9, H(S_1|S_3) = 0.3$. .............................................. 47

2.14 The energy required per source observation of various schemes in the multicast channel with side-information, $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 19.5$ dB, $H(S_1|S_2) = 0.9, H(S_1|S_3) = 0.3$. ................................. 48

2.15 The ratio of the energy required per source observation of the list source-channel decoding scheme to the genie aid bound. ......................... 49

2.16 The ratio of the energy required per source observation of the broadcast with degraded information set based multicast with optimal scheduler to the genie aid bound. ................................. 50

3.1 The relay eavesdropper channel. ........................................ 54

3.2 The rate region of the compound MACs of the relay eavesdropper channel for a fixed input distribution $p(x_1)p(x_2)$. ............................... 56

3.3 The achievable perfect secrecy rate of the proposed schemes in the Gaussian relay eavesdropper channel. ............................... 66
3.4 The network topology. .................................................. 67

3.5 The achievable perfect secrecy capacity for various schemes in the Gaussian relay eavesdropper channel with phase fading. .......................... 68

4.1 A small wireless network with two selfish senders, one altruistic node and two receivers. .................................................. 74

4.2 The TDMA data frame. .................................................. 74

4.3 The cooperation scheme in the small network: a) both the relay and node 2 keep receiving, lasting $\mu T$ b) both node 1 and the relay transmit, lasting $(1 - \mu)T$. .................................................. 77

4.4 The operation sequence when node 1 helps node 3 and $h_{31} > h_{3r}$, (a) node 1 and the relay keep listening, lasting $\mu'^{3,1} T$ (b) node 1 helps transmitting, lasting $\mu'^{3,2} T$ part of the time (c) both node 1 and the relay help, lasting $(1 - \mu'^{3,1} - \mu'^{3,2})T$. .................................................. 79

4.5 The region corresponding to the relay node positions where full-cooperation is the only equilibrium for the selfish nodes. .......................... 85

4.6 Comparison between relay strategies (dumb relaying refers to the case where the relay always helps). .................................................. 86

4.7 Comparison between relay strategies in the channel with phase fading. .......................... 87

4.8 Our large random network with altruistic nodes. .......................... 88
CHAPTER 1

INTRODUCTION

Recent years have witnessed a rapid development and huge success of wireless communication systems. Compared with the traditional point-to-point communication scenario [1], one striking feature of modern wireless communication systems is the presence of multiple users in the network. This unique feature brings both opportunities and challenges to the design of communication systems. On one side, the presence of multiple users in the network offers an opportunity for users to cooperate with each other to get a better performance in terms of channel capacity, energy consumption, etc. On the other side, the simultaneous transmissions of multiple users cause interference to each other in the shared wireless medium. Furthermore, if the transmitted messages are confidential, the presence of other users in the network brings security threat.

Same as the point-to-point communication, network information theory, the study of multiuser communications from information theoretic perspective, was initiated by Shannon [2, 3]. Major progress was made during the 1970s. For example, the capacity region of the Multiple Access Channel (MAC), a good model for the cellular uplink where there are several transmitters communicating with one receiver, was characterized in [4, 5]. The Gaussian Broadcast Channel (BC) with single antenna, a good model for the cellular downlink where one transmitter transmits messages to several receivers, was characterized
After this period, though several other advances were made [7–12], network information theory was hardly the focus of the community. In recent years, driven by the fast-developing wireless communication systems, the community has regained interest in the study of multiuser communications [13–17].

In this thesis, we focus on one facet of multiuser communications: user-cooperation. As a benefit offered by the presence of multiple users in the network, user-cooperation has a long history. Packet forwarding, a primitive form of user-cooperation, was used in China more than 2000 years ago. At that time, Chinese used sequences of fire and smoke beacons relayed by beacon-tower on the Great Wall to send urgent military information over long distance. Packet forwarding is now an indispensable part of modern communication systems, e.g., Internet and ad-hoc networks. These years, more advanced user-cooperation techniques, such as Decode-and-Forward, Compress-and-Forward (we will describe these schemes in detail in the sequel) were devised. The basic idea behind these schemes is that the users in the network pool and share their resources together to create a virtual Multiple Input Multiple Output (MIMO) system [18], hoping to reap the same performance gain as MIMO. Depending on the application, the performance metric measuring each cooperation scheme varies. Most, if not all, of the performance metric considered in the literature can be roughly categorized into throughput (e.g., [16,19,20] and references therein), reliability (e.g., [17,21,22] and references therein) and energy efficiency (e.g., [23–25] and references therein), or combinations of them. Albeit the large body of works, the understanding of user-cooperation is far from completion. A lot of fundamental problems are left open. To name a few: 1) Given a performance metric, what is the optimal cooperation scheme? 2) Are there new applications of user-cooperation? 3) Is user-cooperation a valid assumption in networks? Put in other words, will the users in the network cooperate with each other,
if they are selfish? The answer to these problems will provide insightful guidance for the proper design of modern networks.

In this thesis, we make several progress on these open problems using tools from network information theory. More specifically, we consider 1) the cooperation strategy design in a three-terminal wireless network, 2) the cooperation strategy design for secure communications and 3) the validation of the cooperation assumption in an energy-limited network from game-theoretic perspective.

In the rest of this chapter, we briefly introduce related works and existing results. Motivations for the problems studied in this thesis are then discussed. Finally, main contributions and structure of this thesis are outlined.

1.1 Cooperative Communications

![Relay Channel Diagram](image)

Figure 1.1: The relay channel.

As the building block of cooperative communications, the relay channel, which consists of one source-destination pair and one relay node whose sole purpose is to help the transmission between the source-destination pair as shown in Figure 1.1, is introduced to the information theory society by [26]. Substantial progress in the relay channel was made in the classical paper [19], where an upper-bound on the information rate was derived and two cooperation schemes, namely Decode and Forward (DF), Compress and Forward (CF),
were developed. In DF cooperation, the relay node first decodes the source message and then starts aiding the destination node in decoding. When the source-relay link is very noisy, requiring the relay node to decode the message before starting to help the destination may, in fact, adversely affect performance. The CF strategy avoids this drawback by asking the relay to “compress” its observations and send them to the destination. In this approach, Wyner-Ziv source compression is employed by the relay to allow the destination node to obtain a (noisy) copy of the relay observations. Recent developments focus on generalizing these two schemes to various network scenarios. For example, [20, 27] consider the network with multiple relay nodes, [21] investigates the cooperation in MAC where the multiple access users serve as relay node for each other, [16, 28] study BC where the receivers help to relay information for each other, [17, 22] analyze the schemes in the outage limited scenario. [16] gives a comprehensive tutorial of the recent developments.

While these two schemes are widely used, except for some special cases to the achiev-
able rates of these schemes have not been shown to be the capacity of the relay channel [16]. Specifically, the achievable rates of these two schemes do not match the existing upper-bound. Finding the capacity of the relay channel, by either increasing the achievable rates or tightening the upper-bound derived in [19], has been under investigation for more than two decades. The insight gained in the relay channel will have critical value for the design of cooperation schemes for other channels. But unfortunately, the capacity of the general relay channel is still among the many open problems of network information theory.

In this part of work, we aim to increase the achievable rate of the relay channel. More specifically, we propose a novel cooperation strategy for the relay channel by exploiting the destination feedback. Our scheme combines the benefits of both the DF and CF strategies.
and avoids the noiseless feedback assumption adopted in earlier works. We also generalize the insights gained from the relay channel to other channels (the multicast channel and the three-way channel\(^2\)) and construct cooperation strategies which have larger achievable rates than that of the existing schemes for these channels.

1.2 Cooperation for Secure Communications

![Shannon's model for security system.](image)

Due to the broadcast nature of wireless networks, we need to pay attention to the security of the transmitted confidential messages. Shannon started the study of the security system from information theoretic perspective and introduced the notion of information theoretic secrecy in [2]. The model in [2] assumed that the transmission is noiseless, and used a key \(K\), known at the source and destination but not the wiretapper, to protect the confidential message \(M\), as shown in Figure 1.2. Under this model \(Y = Z = X = f(M, K)\), where \(f\) is the encoder function, which could be stochastic. The destination uses its knowledge of the key \(K\) to decode the message, i.e., \(\hat{M} = g(X, K)\). In [2], the security system is called to be perfectly secure if \(I(X; M) = 0\). This condition says that the received signal \(Z\),

\(^2\)We will give rigorous definitions of these channels in the sequel.
which is the same as $X$, at the wiretapper does not provide any additional information about
the transmitted message $M$ to the wiretapper. Shannon proved a pessimistic result that in
order to achieve perfectly secure, one needs $H(K) \geq H(M)$. Roughly speaking, this con-
dition says that the length of the secret key should be larger than or equal to the length of
the message. The distribution of the secret key for the information theoretic secrecy in this
model is a big challenge. Hence, in the security systems widely used nowadays, such as
the public key system [31], the concept of computational security is adopted. The main
idea is to assume that the wiretapper has limited computation ability, and hence through
properly designed algorithms based on some difficult mathematics problems, one can then
reduce the amount of keys needed to be distributed, while preventing the wiretapper from
decrypting the transmitted messages in a reasonable amount of time. These computational
security systems are not satisfactory due to the following reasons 1) the difficulty of the
mathematics problems have not been proved, 2) the computation resources available to the
wiretapper keep increasing.

Figure 1.3: Wyner’s wiretap model.

Taking the transmission uncertainty into consideration, Wyner introduced the wiretap
channel in [32]. In the three-terminal wiretap channel as shown in Figure 1.3, a source
wishes to transmit confidential messages to a destination while keeping the messages as
secret as possible from a wiretapper. The wiretapper is assumed to have an unlimited
computation ability and to know the coding/decoding scheme used in the main (source-destination) channel. Under the assumption that the source-wiretapper channel is a degraded version of the main channel, Wyner characterized the trade-off between the throughput of the main channel and the level of ignorance of the message at the wiretapper using the rate-equivocation region concept. Loosely speaking, the equivocation rate measures the residual ambiguity about the transmitted message at the wiretapper after seeing the channel output. If the equivocation rate at the wiretapper is arbitrarily close to the information rate, the transmission is called perfectly secure. Csiszár and Körner extended this work to the broadcast channel with confidential messages, where the source sends common information to both the destination and the wiretapper, and confidential messages are sent only to the destination [33].

![Diagram](image)

Figure 1.4: The relay-eavesdropper channel.

In this thesis, we follow Wyner’s approach. Our work here is motivated by the fact that if the wiretapper channel is less noisy than the main channel$^3$, the perfect secrecy capacity of the channel is zero [33]. In this case, it is infeasible to establish a secure link under

$^3$The source-wiretapper channel is said to be less noisy than the source-receiver channel, if for every $V \rightarrow X \rightarrow Y Z$, $I(V; Z) \geq I(V; Y)$, where $X$ is the signal transmitted by the source, $Y, Z$ are the received signal of the receiver and the wiretapper respectively.
Wyner’s wiretap channel model. Our main idea is to exploit user cooperation in facilitating the transmission of confidential messages from the source to the destination. More specifically, we consider a four-terminal relay-eavesdropper channel as shown in Figure 1.4, where a source (S) wishes to send messages to a destination (D) while leveraging the help of a relay (R) node to hide those messages from the eavesdropper (E).

As mentioned in Section 1.1, the relay channel without security constraints was studied under various scenarios [16, 17, 19–22, 26, 34, 35]. In most of these works, cooperation strategies were constructed to increase the throughput and/or reliability function. In this part of work, we identify a novel role of the relay node in establishing a secure link from the source to the destination. The relay channel with confidential messages was studied in [36, 37], where the relay node acts both as an eavesdropper and a helper. In the model of [37], the source sends common messages to the destination using the help of the relay node, but also sends private messages to the destination while keeping them secret from the relay. In contrast with [37], the relay node in our work acts as a trusted “third-party” whose sole goal is to facilitate secure communications (imposing an additional security constraint on the relay node is also considered). The idea of using a “third-party” to facilitate secure communications also appeared in [38]. Contrary to our work, which considers noisy channels, [38] focused on the generation of common random secret keys at two nodes under the assist of a third-party using a noiseless public discussion channel. The users then use the secret key to establish a secure link between the source-destination pair. Other recent works on secure communications investigated multiple access channel with confidential messages [39, 40], the multiple access channel with a degraded wiretapper [41–43],
fading broadcast channel with confidential messages [44–50] and MIMO secure communications [51]. In summary, it appears that our relay-eavesdropper model is fundamentally different from the models considered in all previous works.

1.3 Cooperation: Game Theoretic Perspective

The utility of cooperative communications in many relevant practical scenarios has been established in [14, 16, 17, 19, 22–24, 34, 52]. These physical layer studies assume that each user is willing to expend energy in sending packets for other users. This assumption is reasonable in a network with a central controller who has the ability to enforce the optimal cooperation strategy on the different wireless users. The popularity of ad-hoc networks and the increased programmability of wireless devices, however, raises serious doubts on the validity of this assumption, and hence, motivates investigations on the impact of user selfishness on the performance of wireless networks.

One important thrust in these efforts focuses on designing high-level protocols that prevent users from misbehaving and/or provide incentives for cooperation. To prevent misbehavior, several protocols based on reputation propagation have been proposed in the literature, e.g., [53, 54]. Other works have used ideas from micro-economy to construct protocols that reward cooperation. In [55], for example, a protocol based on virtual currency is proposed. Overall, these protocols are based on ideas rooted in game theory [56, 57], but, in most cases, are not derived from the equilibrium perspective and are hard to analyze, due to the complicated underlying network models.

Another thrust of research analyzes the impact of user selfishness from a game theoretic perspective, e.g., [58, 59]. Since the problem is typically too involved, several simplifications to the network model are usually made to facilitate analysis and allow for extracting
insights. For example, in [58], the wireless nodes are assumed to be interested in maximizing energy efficiency. At each time slot, a certain number of nodes are randomly chosen and assigned to serve as relay nodes on the source-destination route. The authors derive a Pareto optimal operating point and show that a certain variant of the well known TIT-FOR-TAT algorithm converges to this point. In [58], the authors assume that the transmission of each packet costs the same energy and each session uses the same number of relay nodes. Another example is [59], which studies the Nash Equilibrium (NE) of packet forwarding in a static network by taking the network topology into consideration. More specifically, the authors assume that the transmitter/receiver pairs in the network are always fixed and derive the equilibrium conditions for both cooperative and non-cooperative strategies. Similar to [58], the cost of transmitting each packet is assumed fixed. It is worth noting that most, if not all of, the works in this thrust utilize the repeated game formulation, where cooperation among users is sustainable by credible punishment for deviating from the cooperation point.

In this part of work, we adopt a game-theoretic approach to analyze the impact of user selfishness on the energy efficiency of wireless networks, and study how to stimulate user-cooperation in the wireless network with limited energy. The energy efficiency of ad hoc networks has been studied in [23, 24] under the assumption that all the nodes are interested in minimizing the overall energy consumption of the network. In the wireless setting, the transmission power must scale polynomially with the transmitter-receiver distance in order to guarantee reliable decoding of the information. In the no-cooperation scenario, where the wireless users are not willing to relay packets for each other, each user must reach its destination directly. It is then easy to see that this requires significantly higher energy levels, as compared with the globally optimal solution characterized in [23, 24].
This observation motivates our work where we consider wireless networks consisting of both selfish nodes, interested in minimizing their own energy consumption, and altruistic nodes, interested in minimizing the overall energy expenditure of the network. In practice, those altruistic nodes may correspond to wireless relay nodes endowed with more processing powers and tasked with fostering cooperation among the wireless users. Unlike the traditional approach, these relay nodes are not assumed to have any control on the behavior of the selfish nodes directly. The main problem we try to answer is under what conditions user cooperation becomes a sustainable NE, hence, it is of users’ own interest to cooperate.

1.4 Organization and Contributions of This Thesis

In the following, we outline the structure and contributions of this thesis.

Chapter 2 considers cooperation strategy design in a three-terminal wireless network. In this chapter, we propose a novel cooperation strategy for the relay channel with feedback. Our scheme combines the benefits of both DF and CF strategies and avoids the noiseless feedback assumption adopted in earlier works. Our analysis of the achievable rate of the proposed strategy reveals the diminishing gain of feedback in the asymptotic scenarios of low and large signal-to-noise ratio. Inspired by the feedback strategy for the relay channel, we construct a greedy cooperation strategy for the multicast scenario. Motivated by the greedy approach, we show that the weak receiver is led to help the strong receiver first. Based on the same greedy motivation, the strong user starts to assist the weak receiver after successfully decoding the transmitted codeword. We compute the corresponding achievable rate and use it to establish the significant gains offered by this strategy, as compared

The notions of weak and strong receivers will be defined rigorously in the sequel.
with the non-cooperative scenario. Motivated by the sensor networks application, we identify the three-way channel model as a special case of our general formulation. In this model, the three nodes observe correlated data streams and every node wishes to communicate its observations to the other two nodes. Our proposed cooperation strategy in this scenario consists of three stages of multicast with side information, where the multicasting order is determined by a low complexity greedy scheduler. In every stage, we use a cooperation strategy obtained as a generalization of the greedy multicast approach. This strategy highlights the central role of list source-channel decoding in exploiting the side information available at the receivers. By contrasting the minimum energy required by the proposed strategy with the genie-aided and non-cooperative schemes, we establish its superior performance. We identify the greedy principle as the basis for constructing efficient cooperation strategies in the three considered scenarios. Careful consideration of other variants of the three node network reveals the fact that such principle carries over with slight modifications.

Chapter 3 focuses on the cooperation strategy design for secure communications. Towards this end, several cooperation strategies for the relay-eavesdropper channel are constructed and the corresponding achieved rate-equivocation regions are characterized. An outer-bound on the optimal rate-equivocation region is also derived. The proposed schemes are shown to achieve a positive perfect secrecy rate in several scenarios where the secrecy capacity in the absence of the relay node is zero. Quite interestingly, we establish the deaf-helper phenomenon where the relay can help while being totally ignorant of the transmitted message from the source. Furthermore, we show that the relay node can aid in the transmission of confidential messages in some settings where classical cooperation fails to offer performance gains, e.g., the reversely degraded relay channel. Finally, we observe that
the proposed Noise-Forwarding (NF) is intimately related with the multiple access channel with security constraints, as evident in the sequel.

Chapter 4 studies user cooperation from game-theoretic perspective. More specifically, we study a network consisting of both selfish and altruistic nodes and investigate under what conditions full-cooperation would become NE of the selfish nodes, hence validate the assumption that the users will cooperate. Our work establishes the sub-optimality of traditional relaying strategies, which ignore the game theoretic aspect of the problem, and characterizes forwarding strategies that allow the relay nodes to change the cost function of the other selfish nodes such that full cooperation emerges at the equilibrium point. Interestingly, as shown in the sequel, full cooperation becomes a NE in networks where only a vanishingly small fraction of the total number of nodes are altruistic. One distinguishing aspect of our work is that we confine our strategies to the physical layer and avoid introducing elements, like virtual currency, which may add significant complexity to the higher layers.

In Chapter 5, we offer some concluding remarks. Appendix A collects all the proofs of this dissertation.
CHAPTER 2

COOPERATION STRATEGIES FOR A THREE-NODE WIRELESS NETWORK

The booming wireless network applications have sparked a renewed interest in network information theory. Despite the recent progress [14–16, 60–62], developing a unified theory for network information flow remains an elusive task. Even for the three-terminal relay channel, the capacity is still unknown. More specifically, the achievable rates of the existing schemes for the relay channel do not match the upper-bound. In this chapter, we consider, perhaps, the most simplified scenario of wireless networks. The network under consideration is composed of only three nodes and limited by the half-duplex and total power constraints. We first consider the cooperation strategy design for the relay channel by exploiting noisy feedback from the destination. We construct a novel relay strategy that has a larger achievable rate than that of the existing ones under certain conditions and thus narrow down the gap to the upper-bound. We then generalize the insight gained from the relay channel to other instantiations of the three-node wireless network, and devise cooperation schemes that outperform the existing schemes for these channels. Overall, the value of noisy feedback on constructing efficient cooperation schemes is established.
2.1 The Three Node Wireless Network

Figure 2.1 illustrates a network consisting of three nodes each observing a different source. In the general case, the three sources can be correlated. Nodes are interested in obtaining a subset or all the source variables at the other nodes. To achieve this goal, nodes are allowed to coordinate and exchange information over the wireless channel. Different variants of this problem have been previously considered in the literature. For example, [7, 63] investigate the capacity region for general multi-terminal networks, [8] studies the multi-terminal source coding problem, whereas [64] discusses the cost-distortion tradeoff in a similar setup. In this chapter, we focus on the half-duplex wireless setting and assume that each node generates only one source sequence. This source sequence must be reconstructed losslessly at one of the two other nodes, or both of them, with (or without) receiver side-information. Although this model is not the most general, it encompasses many important wireless communication scenarios as argued in the sequel. Mathematically, the three node wireless network studied in this chapter consists of following elements:

1. The three sources $S_i, i = 1, 2, 3$, drawn i.i.d. from certain known joint distribution $p(s_1, s_2, s_3)$ over a finite set $\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3$. We denote by $S_i^K$ the length-$K$ discrete source sequence $S_i(1), \ldots, S_i(K)$ at the $i$-th node. Throughout the sequel, we use capital letters to refer to random variables and small letters for realizations.

2. We consider the discrete-time additive white Gaussian noise (AWGN) channel. At time instant $n$, node $j$ receives

$$Y_j(n) = \sum_{i \neq j} h_{ij} X_i(n) + Z_j(n) \quad (2.1)$$
where $X_i(n)$ is the transmitted signal by node-$i$ and $h_{ij}$ is the channel coefficient from node $i$ to $j$. To simplify the discussion, we assume the channel coefficients are symmetric, i.e., $h_{ij} = h_{ji}$. These channel gains are assumed to be known a-priori at the three nodes. We also assume that the additive zero-mean Gaussian noise is spatially and temporally white and has the same unit variance ($\sigma^2 = 1$).

3. We consider half-duplex nodes that cannot transmit and receive simultaneously using the same degree of freedom. Without loss of generality, we split the degrees of freedom available to each node in the temporal domain, so that, at each time instant $n$, a node-$i$ can either transmit (T-mode, $Y_i(n) = 0$) or receive (R-mode, $X_i(n) = 0$), but never both. Due to the half-duplex constraint, at any time instant, the network nodes are divided into two groups: the T-mode nodes (denoted by $T$) and the R-mode nodes ($R$). A partition $(T, R)$ is called a network state.
4. Let $P^{(l)}_i$ denote the average transmit power at the $i$-th node during the $m_l$ network state. We adopt a short-term power constraint such that the total power of all the T-mode nodes at any network state is limited to $P$, that is,

$$\sum_{i \in T_l} P^{(l)}_i \leq P, \quad \forall m_l.$$  \hspace{1cm} (2.2)

This short term constraint avoids large peak powers and simplifies the power allocation algorithm (at the expense of a possible performance loss).

5. We associate with node-$i$ an index set $I_i$, such that $j \in I_i$ indicates that node-$i$ is interested in obtaining $S_j$ from node-$j$ ($j \neq i$).

6. At node-$i$, a causal joint source-channel encoder converts a length-$K$ block of source sequence into a length-$N$ codeword. The encoder output at time $n$ is allowed to depend on the received signal in the previous $n-1$ instants, i.e.,

$$X_i(n) = f_i(n, S^K_i, Y_i^{n-1}).$$  \hspace{1cm} (2.3)

In the special case of a separate source-channel coding approach, the encoder decomposes into:

- A source encoder $f_{si}$ maps $S^K_i$ into a node message $W_i$, i.e., $W_i = f_{si}(S^K_i)$, $W_i \in [1, M_i]$.

- A channel encoder $f_{ci}(n)$ encodes the node message into a channel input sequence $X_i(n) = f_{ci}(n, W_i, Y_i^{n-1})$.

7. At node-$i$, decoder $d_i$ estimates the source variables indexed by $I_i$

$$\{\hat{S}^{K}_{ij}\} = d_i(Y^N_i, S^K_i), \quad \forall j \in I_i.$$  \hspace{1cm} (2.4)
where $\hat{S}_{ij}^K$ denotes the estimation of $S_j^K$ at node $i$. In the case of a separate source-channel coding scheme, decoder $d_i$ consists of the following:

- A channel decoder $d_{ci}$, $\hat{W}_{ij} = d_{ci}(Y_i^N)$.
- A source decoder $d_{si}$, $\hat{S}_{ij}^K = d_{si}(\hat{W}_{ij}, S_i^K)$.

8. A decoding error is declared if any node fails to reconstruct its intended source variables correctly. Thus, the joint error probability can be expressed as

$$P_{e}^{N,K} = \text{Prob}\left\{ \bigcup_{j \in I_i, i=1,2,3} \{ \hat{S}_{ij}^K \neq S_j^K \} \right\}. \quad (2.5)$$

In the case of a separate coding scheme, the error probability of the channel coding scheme is

$$P_{e}^{N} = \text{Prob}\left\{ \bigcup_{j \in I_i, i=1,2,3} \{ \hat{W}_{ij} \neq W_j \} \right\}. \quad (2.6)$$

9. An efficient cooperation strategy should strive to maximize the achievable rate given by $KH(S_1, S_2, S_3)$, where $N$ is the minimum number of channel uses necessary to satisfy the network requirements. For a fixed $H(S_1, S_2, S_3)$, this optimization is equivalent to minimizing the bandwidth expansion factor $\tau = \frac{N}{K}$. Due to a certain additive property, using the bandwidth expansion factor will be more convenient in the three-way channel scenario. A bandwidth expansion factor $\tau$ is said to be achievable if there exists a series of source-channel codes with $N, K \to \infty$ but $\frac{N}{K} \to \tau$, such that $P_{e}^{N,K} \to 0$. In the feedback-relay and multicast channel, minimizing the bandwidth expansion factor reduces to the more conventional concept of maximizing the rate given by $R = \frac{\log M}{N}$, where $M$ is the size of message set at the source node.

Throughout this thesis, $\log$ is based on 2.

$^5$The bandwidth expansion factor terminology is motivated by the real time application where the bandwidth of the channel must be $N/K$ times the bandwidth of the source process.
Throughout the sequel we will use the shorthand notation

\[ C(x) = \frac{1}{2} \log (1 + x). \]  

The model described here encompasses many important network communication scenarios with a wide range of complexity, controlled by various configurations of the index sets and the sources. From this perspective, the relay channel represents the simplest situation where one node serves as the relay for the other source-destination pair, e.g., \( S_2 = S_3 = \emptyset, I_1 = I_2 = \emptyset \) and \( I_3 = \{1\} \). If we enlarge the index set \( I_2 = \{1\} \), meaning node-2 now is also interested in obtaining the source message, then the problem becomes the multicast channel. Furthermore, if the two receivers (node-2 and 3) in the multicast case have additional observations, i.e., \( S_2 \) and \( S_3 \), which are correlated with the source variable \( S_1 \), then the problem generalizes to the so-called multicast with side information scenario. We refer to the most complex scenario as the three-way channel. In this scenario, the three sources are correlated and every node attempts to reconstruct the other two sources, i.e., \( I_i = \{1, 2, 3\} \setminus \{i\} \). While it is easy to envision other variants of the three node network, we decide to limit ourselves to these special cases. This choice stems from our belief that other scenarios do not add further insights to our framework. For example, another variant of the feedback-relay channel would allow the relay to observe its own side information. Careful consideration of this case, however, shows that our analysis in Section 2.2 extends to this case with only slight modifications. Similarly, inspired by our modular approach for the three-way channel, one can decompose the multiple-access channel with correlated sources into two stages of feedback-relay channels with side information.
2.2 The Relay Channel

Our formulation for the three node network allows for a more realistic investigation of the relay channel with feedback. In this scenario, node-1 is designated as the source node, node-3 the destination, and node-2 the relay. Since there is only one source in this case, one can easily see that maximizing the achievable rate $R$ from source to destination is equivalent to minimizing the bandwidth expansion factor. Contrary to previous works on the relay channel, we allow the destination to transmit over the noisy wireless channel and investigate the achievable rates in this context.

Before proceeding to our scenario of interest, we review briefly the available results on the AWGN relay channel. In a recent work [16], Kramer et al. present a comprehensive overview of existing cooperation strategies, and the corresponding achievable rates, for full-duplex/half-duplex relay channels. In our work, we focus on two classes of cooperation strategies, namely 1) Decode and Forward (DF) and 2) Compress and Forward (CF) strategies.

In DF cooperation, the relay node first decodes the source message and then starts aiding the destination node in decoding. More specifically, the transmission cycle is divided into two stages. In the first stage, which occupies a fraction $t$ of the total time, the source node sends common messages to both the relay and the destination node. Typically more information is sent in this stage than can be decoded by the destination node. Having successfully decoded the source message in this stage, the relay node uses the second stage to help the destination resolve its uncertainty about the transmitted codeword. During the second stage, a new message is also sent to the destination node from the source node, along with the information from the relay. When the source-relay link is very noisy, one can argue that requiring the relay node to decode the message before starting to help the destination
may, in fact, adversely affect performance. The CF strategy avoids this drawback by asking the relay to “compress” its observations and send it to the destination. In this approach, Wyner-Ziv source compression is employed by the relay to allow the destination node to obtain a (noisy) copy of the relay observations. Similar to the DF strategy, the transmission cycle is divided into two stages. During the first stage, both the relay and the destination listen to the source node. The relay then quantizes its observations and sends the quantized data to the destination node during the second stage. In general, the correlation between the relay observations and the destination observations can be exploited by the Wyner-Ziv coding to reduce the data rate at the relay node. During the second stage, new information is also sent by the source that further boosts the total throughput. Here we omit the detailed proofs and refer the interested readers to the relevant works ([16, 17, 19, 62, 65–69]). We note, however, that the statement of the results allows for employing optimal power allocation policies to maximize the throughput.

**Lemma 1** The achievable rate of the DF and CF strategies are given by

\[ R_{DF} = \max_{t, r_{12}, P_i^{(2)}} \min \left\{ tC \left( h_{12}^2 P \right) + (1 - t)C \left( (1 - r_{12}^2) h_{13}^2 P_1^{(2)} \right) ; tC \left( h_{13}^2 P \right) + (1 - t)C \left( h_{13}^2 P_1^{(2)} + 2r_{12}h_{13}h_{23}\sqrt{P_1^{(2)} P_2^{(2)} + h_{23}^2 P_2^{(2)}} \right) \right\}. \]  

(2.8)

\[ R_{CF} = \max_{t, P_i^{(2)}} tC \left( h_{13}^2 + \frac{h_{12}^2}{1 + \sigma_2^2} \right) P + (1 - t)C \left( h_{13}^2 P_1^{(2)} \right). \]  

(2.9)

where

\[ \sigma_2^2 = \frac{(h_{12}^2 + h_{13}^2)P + 1}{(h_{13}^2 P + 1) \left( 1 + \frac{h_{23}^2 P_2^{(2)}}{h_{13}^2 P_1^{(2)} + 1} \right) - 1}, \]  

(2.10)

and

\[ P_1^{(2)} + P_2^{(2)} = P. \]  

(2.11)
One can leverage the feedback, from the destination to the relay, to further increase the achievable rate [19]. The capacity of the full-duplex relay channel with **noiseless** feedback, where the exact received signal at the destination is available to the relay, is known [19]. Applying the coding scheme of [19] to the half-duplex relay channel, one can get the following expression for capacity of the half-duplex relay channel with noiseless feedback [70]

\[
R_{UB} = \max_{t,r_{12},P_i^{(0)}} \min \left\{ tC \left( (h_{12}^2 + h_{13}^2)P_1^{(1)} \right) + (1-t)C \left( (1-r_{12}^2)h_{13}^2 P_1^{(2)} \right); tC \left( h_{13}^2 P_1^{(1)} \right) + (1-t)C \left( h_{13}^2 P_1^{(2)} + 2r_{12}h_{13}h_{23}\sqrt{P_1^{(2)}P_2^{(2)}} + h_{23}^2 P_2^{(2)} \right) \right\}.
\]

(2.12)

![Figure 2.2: The operation sequence of the half-duplex relay channel with noisy feedback.](image)

In the following, we present a cooperation strategy for the relay channel with **noisy** feedback. Our model for the noisy feedback represents a more faithful model for the wireless environments. In a nutshell, the proposed strategy combines the DF and CF strategies to overcome the bottleneck of a noisy source-relay channel. In this FeedBack (FB) approach, the destination first assists the relay in decoding via CF cooperation. After decoding, the relay starts helping the destination via a DF configuration. Due to the half-duplex constraint, every cycle of transmission is divided into the following three stages (as shown in Fig. 2.2).
• The first state lasts for a fraction $\alpha t$ of the cycle ($0 \leq t, \alpha \leq 1$). In this stage, both the relay and the destination listen to the source, thus we set $X_2 = X_3 = 0$. We refer to the network state in this stage as $m_1$.

• The feedback stage lasts for a fraction $(1 - \alpha)t$ of the cycle. In this stage, the relay listens to both the destination and the source, so $X_2 = 0$. Since the destination is not yet able to completely decode the source message, it sends to the relay node a Wyner-Ziv compressed version of its observations. We refer to the network state in this stage as $m_3$.

• The final stage lasts for a fraction $(1 - t)$ of the cycle. Having obtained the source information, the relay is now able to help the destination node in decoding the source message. In this stage $X_3 = 0$. We refer to the network state in this stage as $m_2$.

The time-division parameters $t$ and $\alpha$ control the relative duration of each network state. In particular, $t$ represents the total time when the relay node is in the receive mode. The feedback parameter $\alpha$ controls the amount of feedback, i.e., a $(1 - \alpha)$ fraction of the total relay listening time is dedicated to feedback. Here, we stress that this formulation for a relay channel with feedback represents a “realistic” view that attempts to capture the constraints imposed by the wireless scenario (as opposed to the noiseless feedback mentioned above). The feedback considered here simply refers to transmission from the destination to the relay over the same (noisy) wireless channel. Using random coding arguments we obtain the following achievable rate for the proposed feedback scheme.
Theorem 2 The achievable rate of the noisy feedback scheme for Discrete Memoryless Channel is

\[
R_{FB} = \sup_{\alpha, t, p(x_1|m_1), p(x_1,x_3|m_3), p(x_1,x_2|m_2), p(\hat{y}_3)} \min \left\{ \alpha t I(X_1; Y_2, \hat{Y}_3|m_1) + (1 - \alpha) t I(X_1; Y_2 X_3, m_3) + (1 - t) I(X_1; Y_3 X_2, m_2) ; \
\alpha t I(X_1; Y_3|m_1) + (1 - t) I(X_1, X_2; Y_3|m_2) \right\},
\]

subject to

\[
(1 - \alpha) I(X_3; Y_2|m_3) \geq \alpha I(Y_3; \hat{Y}_3|Y_2, m_1).
\]

Proof: Please refer to Appendix A.1.

The gain leveraged from the noisy feedback can be seen in the increased rate \(\alpha t I(X_1; Y_2, \hat{Y}_3|m_1)\) that the relay can decode after combining the signal from the source and the feedback signal from the destination, where (2.14) captures the constraint on the amount of information the destination can send to the relay. The result for the Gaussian channel now follows.

Lemma 3 The achievable rate of the noisy feedback scheme in the Gaussian channel is given by

\[
R_{FB} = \max_{\alpha, t, r_{12}, P_i^{(i)}} \min \left\{ \alpha C \left( \frac{h_{13}^2}{1 + \sigma_3^2} + h_{12}^2 P_1^{(1)} \right) + (1 - \alpha) t C \left( h_{12}^2 P_1^{(3)} \right) + (1 - t) C \left( 1 - r_{12} h_{13}^2 P_1^{(2)} \right) ; \alpha t C \left( h_{13}^2 P_1^{(1)} \right) + (1 - t) C \left( h_{13}^2 P_1^{(2)} + 2 r_{12} h_{13} h_{23} \sqrt{P_1^{(2)} P_2^{(2)}} + h_{23}^2 P_2^{(2)} \right) \right\}
\]

where

\[
\sigma_3^2 = \frac{\left( h_{12}^2 + h_{13}^2 P_1^{(1)} + 1 \right)}{\left( h_{12}^2 P_1^{(1)} + 1 \right) \left( 1 + \frac{h_{13}^2 P_3^{(3)}}{h_{12}^2 P_1^{(1)} + 1} \right)^{\frac{1-\alpha}{\alpha}} - 1}
\]
and \( r_{12} \) is the correlation between \( X_1, X_2 \) during state \( m_2 \). In the proposed strategy, the total power constraint specializes to

\[
P^{(1)}_1 = P, \quad P^{(2)}_1 + P^{(2)}_2 = P, \quad P^{(3)}_1 + P^{(3)}_3 = P. \quad (2.17)
\]

**Proof:** Please refer to Appendix A.2. \( \square \)

Armed with Lemmas 1 and 3, we can now contrast the performance of the DF, CF, and FB strategies. Our emphasis is to characterize the fundamental properties of the feedback scheme and quantify the gain offered by it under different assumptions on the channel gains and total power. The relay-off performance, i.e., \( R_{ro} = C(h^2_{13}P) \), serves as a lower bound on the achievable rate. In fact, the relay-off benchmark can be viewed as a special case of the three cooperative schemes. For example, setting \( P^{(2)}_2 = 0 \) and \( t = 0 \) effectively reduces both DF and CF strategies to the relay-off case. Therefore, one can conceptually describe the order of containment of various schemes as “relay-off \( \subset \) DF \( \subset \) FB” and “relay-off \( \subset \) CF”. As for the performance upper bounds, the cut-set bounds [19] give rise to 1) a multi-transmitter rate \( R_{(1,2)3} = C((h^2_{13} + h^2_{23})P) \) corresponding to perfect cooperation between the source and relay nodes (note that with the total power constraint \( P_1 + P_2 = P \), the optimal power for node-1 is \( P_1 = \frac{h^2_{12}P}{h^2_{12} + h^2_{23}} \)); and 2) a multi-receiver rate \( R_{1-(2,3)} = C((h^2_{13} + h^2_{12})P) \) corresponding to perfect cooperation between the relay and destination nodes.

The achievable rate of the DF strategy, i.e., \( R_{DF} \), enjoys an intuitive geometric interpretation: each expression within the min operator is a linear segment in the parameter \( t \in [0, 1] \) (see (2.8)). Hence, the optimal time \( t \), assuming the other variables remain fixed, can be simply determined by the intersection point of the two associated line segments, as illustrated in Fig. 2.3. On the other hand, \( R_{FB} \) and \( R_{CF} \) are characterized by more complicated expressions due to the dependency of \( \sigma^2_3 \) and \( \sigma^2_2 \) upon the time-division parameters.
Our next result finds upper bounds on $R_{FB}$ and $R_{CF}$ which allow for the same simple line-crossing interpretation as $R_{DF}$.

**Lemma 4** The achievable rate of the feedback scheme is upper bounded by

$$R_{FB} \leq \max_{\alpha,t,r_{12},P_{1}^{(j)}} \min \left\{ \alpha t C\left(h_{12}^{2}P_{1}^{(1)}\right) + (1 - \alpha) t C\left(h_{12}^{2}P_{1}^{(3)} + h_{23}^{2}P_{2}^{(3)}\right) 
+ (1 - t) C\left((1 - r_{12})h_{13}^{2}P_{1}^{(2)}\right) \right\}.$$  

(2.18)
The achievable rate of compress-and-forward is bounded by

\[
R_{CF} \leq \max_{t,P_1^{(i)}} \left\{ \min_{t,P_1^{(i)}} \left( tC\left(h_{13}^2 P_1^{(1)}\right) + (1-t)C\left(h_{21}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)}\right) \right) ;
\right.

\left. tC\left((h_{13}^2 + h_{12}^2) P_1^{(1)}\right) + (1-t)C\left(h_{13}^2 P_1^{(2)}\right) \right\}.
\]

(2.19)

**Proof:** Please refer to Appendix A.3.

Comparing (2.15) and (2.18) sheds more light on the intuition behind the upper bound on \(R_{FB}\). The upper bound replaces the actual rate received at the relay node with the maximum rate corresponding to the following two stages. In the first stage, only the source is transmitting and the second stage is a multiple access channel where the source and the destination play the roles of transmitting users. Fig. 2.3 compares the performance of the three schemes. For example, when \(h_{12}^2 \leq h_{13}^2\), the intersection point corresponding to decode-forward would fall below the flat line \(C(h_{13}^2 P)\) associated with the relay-off rate.

More rigorously, we have the following statement.

**Theorem 5**

1. If \(h_{12}^2 \leq h_{13}^2\) then \(R_{DF} \leq R_{ro}\).

2. If \(h_{23}^2 \leq h_{13}^2\) then \(R_{CF} \leq R_{ro}\).

3. If \(h_{23}^2 \leq h_{12}^2\) then \(R_{FB} \leq R_{DF}\).

**Proof:** Please refer to Appendix A.4.

Theorem 5 reveals the fundamental impact of channel coefficients on the performance of the different cooperation strategies. In particular, the DF strategy is seen to work well with a “strong” source-relay link. If, at the same time, the relay-destination link is stronger, then one may exploit feedback, i.e., \(\alpha \neq 1\), to improve performance. The next result shows that the feedback gain is diminishing in certain asymptotic scenarios.
Theorem 6 1. As $h_{12}$ increases, both DF- and FB-scheme approach the optimal beamforming benchmark, while CF-scheme is limited by a sub-optimal rate

$$C(\max\{h_{13}^2, h_{23}^2\}P).$$

2. As $h_{23}$ increases, both CF- and FB-scheme approach the optimal multi-receiver benchmark, while DF-scheme only approaches to a sub-optimal rate

$$C(\max\{h_{12}^2, h_{13}^2\}P).$$

The proof of Theorem 6 is a straightforward limit computation, and hence, is omitted for brevity. So far we have kept the total power $P$ constant. But in fact, the achievable rate as a function of $P$ offers another important dimension to the problem. First, we investigate the low power regime. In this case we study the slope $S$ of the achievable rate with respect to $P$ (i.e., $R \sim \frac{1}{2}(\log e)SP$). This slope determines the minimum energy per bit [71] according to the relationship

$$\left(\frac{E_b}{N_0}\right)_{\text{min}} = 1.44 - 10\log_{10}(S). \quad (2.20)$$

We observe that (2.20) suffers from a 3 dB loss due to our real channel model (as opposed to the complex channel model [71]).

Theorem 7 Let $f_1(\theta, r_{12}, h_{13}, h_{23}) = h_{13}^2 \cos^2 \theta + 2r_{12}h_{13}h_{23} \cos \theta \sin \theta + h_{23}^2 \sin^2 \theta$ and $f_2(\theta, r_{12}, h_{13}) = (1 - r_{12}^2)h_{13}^2 \cos^2 \theta$ be a shorthand notation, then

1. When $h_{12}^2 \geq h_{13}^2$

$$S_{DF} = \max_{\theta, r_{12}} \frac{f_1(\theta, r_{12}, h_{13}, h_{23})h_{12}^2 - f_2(\theta, r_{12}, h_{13})h_{13}^2}{f_1(\theta, r_{12}, h_{13}, h_{23}) + h_{12}^2 - f_2(\theta, r_{12}, h_{13}) - h_{13}^2}. \quad (2.21)$$

and

$$\frac{(h_{13}^2 + h_{23}^2)h_{12}^2}{h_{23}^2 + h_{12}^2} \leq S_{DF} \leq \frac{(h_{13}^2 + h_{23}^2)h_{12}^2 - h_{13}^4}{h_{23}^2 + h_{12}^2 - h_{13}^2}. \quad (2.22)$$
2. $S_{CF} = h_{13}^2$ with $t_{opt} \to 1$.

3. $S_{FB} = S_{DF}$ with $\alpha_{opt} \to 1$.

Proof: Please refer to Appendix A.5.

It follows from Theorem 7 that given $h_{12}^2 \geq h_{13}^2$, DF cooperation delivers a larger slope than the relay-off, i.e.,

$$S_{DF} - h_{13}^2 = \frac{h_{13}^2 (h_{12}^2 - h_{13}^2)}{h_{23}^2 + h_{12}^2} \geq 0. \quad (2.23)$$

In this case, the signal the relay receives is better than the signal the destination receives.

If the relay is off, the transmission rate is dictated by the smaller channel gain $h_{13}^2$. In DF cooperation, the source can send information at the rate that only the relay can decode ($h_{12}^2 \geq h_{13}^2$). After the relay decodes, the relay can help the source to send information to the destination. However, CF cooperation does not yield any gain in the low power regime.

Similarly, we see that the CF stage of the proposed FB becomes useless, and hence, the scheme reduces to the DF approach in the low power regime. The reason lies in the fact that for small $P$, the channel output is dominated by the noise, and hence, the compression algorithm inevitably operates on the noise, resulting in diminishing gains.

We next quantify the power offset of the three schemes in the high power regime, that is, to characterize $R \sim \frac{1}{2} \log P + \frac{1}{2} G$ as $P \to \infty$ [72].

**Theorem 8** Following the same shorthand notations as in Theorem 7, we obtain

1. Given $h_{12}^2 \geq h_{13}^2$, 
   $$G_{DF} = \max_{\theta, r_{12}} \log \frac{f_1(\theta, r_{12}, h_{13}, h_{23}) \cdot \log h_{12}^2}{f_2(\theta, r_{12}, h_{13}, h_{23})} - \log \frac{f_2(\theta, r_{12}, h_{13}, h_{23})}{h_{13}^2} \geq 0. \quad (2.24)$$

2. 
   $$G_{CF} = \max_{t, \theta} t \log \left( \frac{h_{13}^2}{1 + \sigma_2^2(\infty)} \right) + (1 - t) \log (h_{13}^2 \cos^2 \theta) \quad (2.25)$$
where
\[
\sigma_2^2(\infty) = \frac{h_{12}^2 + h_{13}^2}{h_{13}^2 \left( 1 + \frac{h_{23}^2}{h_{13}^2} \tan^2 \theta \right)^{\frac{1}{1-\alpha}}}.
\] (2.26)

3. \( G_{FB} = G_{DF} \) with \( \alpha_{opt} \to 1 \).

**Proof:** Please refer to Appendix A.6.

Theorem 8 reveals the fact that strict feedback \( (\alpha \neq 1) \) does not yield a gain in high power regime. The reason for this behavior can be traced back to the half-duplex constraint. When \( \alpha \neq 1 \), the destination spends a fraction \( (1 - \alpha)t \) of time transmitting to the relay, which cuts off the time in which it would have been listening to the source in non-feedback schemes. Such a time loss reduces the pre-log constant, which cannot be compensated by the cooperative gain when \( P \) becomes large. The quantity \( G_{DF} - G_{CF} \) determines the relative order of the DF-scheme and CF-scheme in the high SNR region. In general, this quantity depends on the channel gains and can be computed using numerical methods. For example, when \( h_{12} = 3 \) dB, \( h_{13} = 0 \) dB, \( h_{23} = 10 \) dB, we find that \( G_{DF} - G_{CF} = 1.24 \) dB.

We conclude this section with simulation results that validate our theoretical analysis. Figure 2.4 reports the achievable rate of various schemes, when \( h_{12} = 2.55 \) dB, \( h_{13} = 0 \) dB, and \( h_{23} = 23 \) dB. This corresponds to the case when the source-relay channel is a little better than the source-destination channel, and the relay-destination channel is quite good. This is the typical scenario when feedback results in a significant gain, as demonstrated in the figure. Figure 2.5 reports the ratio \( r = R_i/R_{UB} \) of the achievable rate \( R_i \) of each scheme to the achievable rate of the scheme with noiseless feedback at various SNR value when \( h_{12} = 2.55 \) dB, \( h_{13} = 0 \) dB, and \( h_{23} = 23 \) dB. As we can see, as SNR increases the achievable rate of the FB-scheme converges to that of the DF-scheme, which confirms
Figure 2.4: The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.

Figure 2.5: The ratio of achievable rate of various schemes to the cut-set upper-bound, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.
Figure 2.6: The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, SNR = 0 dB.

our claim in Theorem 8. Figure 2.6 reports the achievable rates of various schemes, when $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, and $P = 0$ dB, as we vary the relay-destination channel gain $h_{23}$. We can see that as the relay-destination channel becomes better, the advantage of feedback increases. Fig. 2.7 shows the position of the relay where each relay scheme achieves the highest achievable rate. Here we put a source node at (0,0) and a destination node at (1,0) and change the position of the relay node. The channel condition between nodes is given by $h_{ij} = \frac{1}{d_{ij}}$. Overall, we can see that the proposed FB cooperation scheme combines the benefits of both the DF and CF cooperation strategies, and hence, attains the union of the asymptotic optimality properties of the two strategies. On the other hand, the gain offered by feedback seems to be limited to certain operating regions, as defined by the channel gains, and diminishes in either the low or high power regime.

32
2.3 The Multicast Channel

The relay channel, considered in the previous section, represents the simplest example of a three-node wireless network. Another example can be obtained by requiring node-2 to decode the message generated at node-1. This corresponds to the multicast scenario. Similar to the relay scenario, we focus on maximizing the achievable rate from node-1 to both node-2 and 3, without any loss of generality. The half-duplex and total power constraints, adopted here, introduce an interesting design challenge. To illustrate the idea,
suppose node-2 decides to help node-3 in decoding. In this case, not only does node-2 compete with the source node for transmit power, but it also sacrifices its listening time for the sake of helping node-3. It is, therefore, not clear \textit{a-priori} if the network would benefit from this cooperation. In the following, we answer this question in the affirmative and further propose a greedy cooperation strategy that is asymptotically optimal.

In a recent work [73], the authors considered another variant of the multicast channel and established the benefits of receiver cooperation in this setup. The fundamental difference between the two scenarios is that, in [73], the authors assumed the existence of a dedicated link between the two receivers. This dedicated link was used by the \textit{strong} receiver to help the \textit{weak} receiver in decoding through a DF strategy. As expected, such a cooperation strategy was shown to strictly enlarge the achievable rate region [73]. In our work, we consider a more representative model of the wireless network in which all communications take place over the same channel, subject to the half-duplex and total power constraints. Despite these constraining assumptions, we still demonstrate the significant gains offered by receiver cooperation. Inspired by the feedback-relay channel, we further construct a greedy cooperation strategy that significantly outperforms the DF scheme [73] in many relevant scenarios.

In the non-cooperative scenario, both node-2 and node-3 will listen all the time, and hence, the achievable rate is given by

\[ C_{\text{non-coop}} = C(\min\{h_{12}^2, h_{13}^2\}P). \]  

(2.27)

Due to the half-duplex constraint, time is valuable to both nodes, which makes them selfish and unwilling to help each other. Careful consideration, however, reveals that such a greedy approach will lead the nodes to cooperate. The enabling observation stems from the feedback strategy proposed for the relay channel in which the destination was found
to get a higher achievable rate if it sacrifices some of its receiving time to help the relay. Motivated by this observation, our strategy decomposes into three stages, without loss of generality we assume $h_{12}^2 > h_{13}^2$, 1) $m_1$ lasting for a fraction $\alpha t$ of the frame during which both receivers listen to node-1; 2) $m_3$ occupying $(1 - \alpha)t$ fraction of the frame during which node-3 sends its compressed signal to node-2; and 3) $m_2$ (the rest $1 - t$ fraction) during which node-1 and 2 help node-3 finish decoding. One major difference between the multicast and relay scenarios is that in the second stage the source cannot send additional (new) information to node-3, for it would not be decoded by node-2, thus violating the multicast requirement that both receivers obtain the same source information. Here, we observe that the last stage of cooperation, in which node-2 is helping node-3, is still motivated by the greedy approach. The idea is that node-1 will continue transmitting the same codeword until both receivers can successfully decode. It is, therefore, beneficial for node-2 to help node-3 in decoding faster to allow the source to move on to the next packet in the queue.

**Lemma 9** The achievable rate of the greedy strategy based multicast scheme is given by

$$R_g = \max_{\alpha,t,P_{(i)}} \min \left\{ \alpha t C \left( \frac{h_{13}^2}{1 + \sigma_4^2} + h_{12}^2 P \right) + (1 - \alpha)t C \left( h_{12}^2 P_{1(3)} \right) \right\}$$

$$\quad \text{(2.28)}$$

where

$$\sigma_4^2 = \frac{(h_{12}^2 + h_{13}^2) P + 1}{(h_{12}^2 P + 1) \left( 1 + \frac{h_{23}^2 P_{1(3)}^{(3)}}{h_{12}^2 P_{1(3)}^{(3)} + 1} \right) \left( \frac{1 - \alpha}{\alpha} - 1 \right)}.$$  

$$\text{(2.29)}$$

The “max” operator is taken over the total power constraint.

The proof follows in the footsteps of the proof of Lemma 3, the only difference is that now the source cannot send new information to the relay in the $m_2$ state.
We observe that the DF multicast scheme corresponds to the special case of $\alpha = 1$, which has a rate

$$R_{DF} = \max_t \min \left\{ tC(h_{12}^2 P); \ tC(h_{13}^2 P) + (1 - t)C \left( (h_{13}^2 + h_{23}^2) P \right) \right\}. \quad (2.30)$$

The cut-set upperbounds give rise to the two following benchmarks: beam-forming $R_{(1,2)-3} = C((h_{13}^2 + h_{23}^2) P)$ and multi-receiver $R_{1-(2,3)} = C((h_{13}^2 + h_{12}^2) P)$. Similar to the relay channel scenario, we examine in the following the asymptotic behavior of the greedy strategy as a function of the channel coefficients and available power.

**Theorem 10**

1. The greedy cooperative multicast scheme strictly increases the multicast achievable rate (as compared to the non-cooperative scenario).

2. The greedy strategy approaches the beam-forming benchmark as $h_{12}$ increases, i.e.,

$$\lim_{h_{12} \to \infty} R_g = C((h_{13}^2 + h_{23}^2) P). \quad (2.31)$$

3. The greedy strategy approaches the multi-receiver benchmark as $h_{23}$ increases, i.e.,

$$\lim_{h_{23} \to \infty} R_g = C((h_{12}^2 + h_{13}^2) P). \quad (2.32)$$

4. As $P \to 0$, the slope of the greedy strategy achievable rate is given by

$$S_g = \frac{h_{12}^2 (h_{23}^2 + h_{13}^2)}{h_{12}^2 + h_{23}^2}. \quad (2.33)$$

5. As $P \to \infty$, the SNR gain $G_g = G_{\text{non-coop}} = \log h_{13}^2$ with $t_{\text{opt}} \to 1$.

**Proof:** Please refer to Appendix A.7.

Parts 2), 3) demonstrate the asymptotic optimality of the greedy multicast as the channel gains increase (the proof follows the same line as that of Theorem 6). On the other hand, we
see that the large-power asymptotic of the multicast channel differs significantly from that of the relay channel. In the relay case (Theorem 8), the contribution of feedback diminishes \(G_{FB} = G_{DF}\) in this asymptotic scenario, but cooperation was found to be still beneficial, that is \(G_{DF} > \log h_{13}^2\). To the contrast, the gain of receiver cooperation in the multicast channel disappears as \(P\) increases. This is because, unlike the relay scenario, at least one receiver must cut its listening time in any cooperative multicast scheme due to the half-duplex constraint. Such a reduction induces a pre-log penalty in the rate, which results in substantial loss that cannot be compensated by cooperation as \(P \to \infty\), and hence, the greedy strategy reduces to the non-cooperative mode automatically.

![Figure 2.8: The achievable rate of various schemes in the multicast channel, \(h_{12} = 0.4\) dB, \(h_{13} = 0\) dB, and \(h_{23} = 23\) dB.](image)

Figure 2.8 compares the achievable rate of the various multicast schemes where the DF cooperation strategy is shown to outperform the non-cooperation scheme. It is also
shown that optimizing the parameter $\alpha$ provides an additional gain (Note $R_{DF}$ in the figure corresponds to $\alpha = 1$). Figure 2.9 reports the ratio $r = R_i/C_{non-coop}$ of the achievable rate $R_i$ of each scheme to the channel capacity of the non-cooperation multicast channel, where we observe that the gain offered by cooperation decreases as the SNR increases. Figure 2.10 reports the achievable rate of the three schemes when $h_{12} = h_{13}$. In this case, it is easy to see that DF strategy yields exactly the same performance as the non-cooperative strategy. On the other hand, as illustrated in the figure, the proposed greedy strategy is still able to offer a sizable gain. Figure 2.11 illustrates the fact that the gain of greedy strategy increases as $h_{23}$ increases. The non-cooperation scheme is not able to exploit the interreceiver channel, and hence, its achievable rate corresponds to a flat line. The DF scheme
Figure 2.10: The achievable rate of various schemes in the multicast channel, $h_{12} = 0$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB.

can benefit from the inter-receiver channel, but its maximum rate is limited by $C(h_{12}^2 P)$, whereas the greedy strategy approaches a rate $R_g = C((h_{12}^2 + h_{13}^2) P)$ as $h_{23} \to \infty$.

2.4 The Three-way Channel

Arguably the most demanding instantiation of the three-node network is the three-way channel [7, 26, 74]. To satisfy the three-way channel requirements, every node needs to transmit its message to the other two nodes and receive their messages from them. Due to the half duplex constraint, these two tasks cannot be completed simultaneously. Take node-1 as an example and consider the transmission of a block of observations $S_1^K$ to the
other two nodes using $N_t$ channel uses. To obtain a lower bound on the bandwidth expansion factor, we assume that node-2 and node-3 can fully cooperate, from a joint source-channel coding perspective, which converts the problem into a point-to-point situation. Then node-1 only needs to randomly divide its source sequences into $2^{K H(S_1|S_2,S_3)}$ bins and transmit the corresponding bin index [8,75,76]. With $N_t$ channel uses, the information rate is $\frac{K H(S_1|S_2,S_3)}{N_t}$. The channel capacity between node-1 and the multi-antenna node-2, 3 is $C((h_{12}^2 + h_{13}^2)P)$. In order to decode $S_1^K$ at node-2, 3 with a vanishingly small error probability, the following condition must be satisfied [77] [78],

$$\frac{K H(S_1|S_2,S_3)}{N_t} \leq C((h_{12}^2 + h_{13}^2)P).$$
Similarly, with full cooperation between node-2 and node-3, the following condition is needed to ensure the decoding of the sequence $S_2^K, S_3^K$ at node-1 with a vanishingly small error probability,

$$\frac{K H(S_2, S_3|S_1)}{N_r} \leq C((h_{12}^2 + h_{13}^2)P).$$

These two genie-aided bounds at node-1 imply that the minimum bandwidth expansion factor required for node-1 is $\tau_{1,\text{gen}} = \frac{H(S_1|S_2, S_3) + H(S_2, S_3|S_1)}{C((h_{12}^2 + h_{13}^2)P)}$. Similarly, we can obtain the corresponding genie-aided bounds for node-2 and node-3. To satisfy the requirement for all these three nodes, the minimum bandwidth expansion factor for this half-duplex three-way channel is therefore

$$\tau_{\text{gen}} \geq \max_{i=1,2,3} \tau_{i,\text{gen}}. \quad (2.34)$$

Figure 2.12: The genie-aided bound in node-1, in which node-2 and node-3 can fully cooperate with each other.

At this point, it is not clear whether the genie-aided bound in (2.34) is achievable. Moreover, finding the optimal cooperation strategy for the three-way channel remains an
elusive task. However, inspired by our greedy multicast strategy, we propose in the following a modular cooperation approach composed of three cooperative multicast with side information stages, which achieves a near-optimal performance.

2.4.1 Multicast with Side-information

To simplify the presentation, without sacrificing any generality, we assume that node-1 is the source and node-2 and 3 are provided with the side information $S_2$ and $S_3$, respectively. A related work appears in the paper [79], where the authors consider the broadcast channel with arbitrarily correlated sources. In that paper, the sender has two correlated messages to send to the two users.

Before presenting our greedy cooperation strategy, we briefly discuss the non-cooperative scenario to establish a performance benchmark where the two receive nodes are not allowed to communicate. For the convenience of exposition, we assume that $H(S_1|S_2) > H(S_1|S_3)$. In the nested binning approach of [80], a source sequence $s^K_1$ is randomly assigned to one of $2^{KH(S_1|S_2)}$ bins. This is the low-level indexing sufficient for node-2 to decode with side-information $S_2$. These indices are then (randomly) divided into $2^{KH(S_1|S_3)}$ equal-sized groups, which corresponds to the random binning approach for node-3. Therefore, a source sequence $s^K_1$ is associated with an index-pair $(b, c)$, where $b \in [1, 2^{KH(S_1|S_3)}]$ is the group index and $c \in [1, 2^{K(H(S_1|S_2)-H(S_1|S_3))}]$ identifies the bin index within a group. Given side-information $S_3$ (more correlated with the source), node-3 needs only the group index $b$ to recover the source sequence. But the low-level bin index is necessary for node-2 to decode. In summary, the above nested binning scheme permits the source node to send $(b, c)$ to node-2 while only $b$ to node-3. Such a structured message is called the degraded information set in [81] where $b$ is the “common” information for both receivers and $c$ the
“private” information required by only one of the two receivers. Using the capacity region given by [81], we get the following benchmark for the non-cooperative multicast with side-information.

**Lemma 11** For non-cooperative multicast with side-information, the achievable bandwidth expansion factor $\tau = N/K$ based on nested binning source coding and degraded information set broadcasting is given by

1. if $h_{12}^2 < h_{13}^2$,
   \[
   H(S_1|S_2) \leq \tau C(h_{12}^2 P). \tag{2.35}
   \]

2. if $h_{12}^2 > h_{13}^2$,
   \[
   H(S_1|S_2) - H(S_1|S_3) \leq \tau C(\gamma h_{12}^2 P),
   
   H(S_1|S_3) \leq \tau C\left(\frac{(1 - \gamma) h_{13}^2 P}{1 + \gamma h_{13}^2 P}\right). \tag{2.36}
   \]

for some $\gamma$.

Now, we are ready to describe our greedy cooperation approach (which is not necessarily optimal). In our scheme, each node calculates the expected bandwidth expansion factor assuming no receiver cooperation, $\tau_{ex,i} = H(S_1|S_i)/C_i$, where $C_i$ denotes the link capacity $C(h_{1i}^2 P)$. The receive node with the smaller $\tau_{ex}$ is deemed as the *strong* node, and hence, will decode first. To better describe the proposed approach, we consider first the simple case where node-3 does not help node-2. We randomly bin the sequences $S_{1}^{K}$ into $2^{KH(S_1|S_2)}$ bins and denote the bin index by $w \in [1, 2^{KH(S_1|S_2)}]$. We further denote by $f_{s1}$ the mapping function $w = f_{s1}(s_{1}^{K})$. We then independently generate another bin index $b$ for every sequence $S_{1}^{K}$ by picking $b$ uniformly from $\{1, 2, \ldots, 2^{KR}\}$, where $R$ is to be determined later. Let $B(b)$ be the set of all sequences $S_{1}^{K}$ allocated to bin $b$. Thus, every source
sequence has two bin indices \( \{ w, b \} \) associated with it. A full cooperation cycle is divided into two stages, where we refer to the network state in these two stages as \( m_1 \) and \( m_2 \), respectively. In the first stage using for \( N_1 \) channel uses, node-1 sends the message \( w \) to node-2 using a capacity achieving code. This stage is assumed to last for \( N_1 \) channel uses. At the end of this state, node-2 can get a reliable estimate \( \hat{w} = w \) if the condition 
\[
KH(S_1|S_2) \leq N_1 C(h_{12}^2 P)
\]
is satisfied. Next, node-2 searches in the bin specified by \( \hat{w} \) for the one and only one \( \hat{s}_{K1}^K \) that is typical with \( s_{K1}^K \). If none exists, decoding error is declared, otherwise, \( \hat{s}_{K1}^K \) is the decoding sequence. During this stage, node-3 computes a list \( \ell (y_{3,m_1}) \) such that if \( w' \in \ell (y_{3,m_1}) \) then \( \{ x_{1,m_1}(w'), y_{3,m_1} \} \) are jointly typical. A key point of our scheme is that node-3 does not attempt to decode \( w \), but rather proceeds to decoding \( s_{K1}^K \) directly. After node-2 decodes \( s_{K1}^K \) correctly, it knows the pair \( \{ w, b \} \), and hence, in the second stage node-2 and node-1 cooperate to send the message \( b \) to node-3. At the end this stage, if the parameters are appropriately chosen, node-3 can decode \( b \) correctly. Node-3 then searches in the bin \( B(b) \) for the one and only one \( \hat{s}_{K1}^K \) that is jointly typical with \( s_{K1}^K \) and that \( f_{s1}(\hat{s}_{K1}^K) \in \ell (y_{3,m_1}) \).

**Lemma 12** With the proposed scheme, both node-2, 3 can decode \( S_{K1}^K \) with a vanishingly small probability of error if \( \tau_0 = N_0 / K, \tau_1 = N_1 / K \) satisfy the following conditions 
\[
\begin{align*}
H(S_1|S_2) & \leq \tau_1 C(h_{12}^2 P), \\
H(S_1|S_3) & - \frac{\min\{C(h_{13}^2 P), C(h_{12}^2 P)\}}{C(h_{12}^2 P)} H(S_1|S_2) & \leq \tau_0 C((h_{13}^2 + h_{23}^2) P).
\end{align*}
\]

**Proof:** Please refer to Appendix A.8.

Next, we allow for the weak node-3 to assist the strong node-2 in decoding. The original \( N_1 \) channel uses now split into two parts: 1) state \( m_1 \) occupying \( \alpha N_1 \) channel uses during which both receiver nodes listen to the source node; and 2) state \( m_3 \) of the remaining \( (1 - \alpha) N_1 \) channel uses during which node-3 sends a compressed version of its received
signal to node-2. At the end of the $N_1$ network uses, node-2 decodes the source sequence and then proceeds to facilitate the same list-decoding at the other receiver as described above. The simple case where node-3 does not assist node-2 can be regarded as a special case of the greedy scheme when $\alpha = 1$. Slightly modifying the proof of Lemma 12, we obtain

**Lemma 13** If $\tau_0, \tau_1$ satisfy the following conditions, both node-2,3 will decode $S_1^K$ with vanishingly small probability of error.

\[
H(S_1 | S_2) \leq \tau_1 R_{CF2}(\alpha),
\]

\[
H(S_1 | S_3) - \frac{\alpha \min\{I(X_1; Y_3|m_1); I(X_1; \hat{Y}_3, Y_2|m_1)\} H(S_1 | S_2)}{R_{CF2}(\alpha)} \leq \tau_0 C((h_1^2 + h_2^2)P).
\]

(2.38)

Where $R_{CF2}(\alpha)$ is the achievable rate of compress-forward scheme for the following relay channel: node-1 acts as the source, node-3 the relay that spends $1 - \alpha$ part of the time in helping destination using CF scheme, and node-2 the destination. The symbol $\hat{Y}_3$ stands for the compressed version of the received signal at node-3 ($Y_3$).

Unfortunately, the expressions for the achievable bandwidth expansion factors do not seem to allow for further analytical manipulation. In order to shed more light on the relative performance of the different schemes, we introduce the minimum energy per source observation metric. Given the total transmission power $P$, the bandwidth expansion factor $\tau$ translates to the energy requirement per source observation as

\[
E(P) = \tau(P)P = \frac{N(P)P}{K}.
\]

(2.39)

Let $E_1(P)$ denotes the energy per source symbol for the benchmark based on broadcast with a degraded information set and $E_2(P)$ for the proposed cooperative multicast scheme.
It is easy to see that both \( E_1(P) \) and \( E_2(P) \) are non-increasing function of \( P \), and hence, approach their minimal values as \( P \to 0 \), that is
\[
E_{i,m} = \lim_{P \to 0} E_i(P) \quad \text{for } i \in \{1, 2\}. \tag{2.40}
\]

Under the assumption that \( \tau_{ex,2} < \tau_{ex,3} \) and using Lemmas 11 and 13, one obtain:

**Theorem 14**

1. **Broadcast with degraded information set:**

   When \( H(S_1|S_2) > H(S_1|S_3) \),
   \[
   E_{1,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \left( \frac{1}{h_{13}^2} - \frac{1}{h_{12}^2} \right)^+ H(S_1|S_3) \right). \tag{2.41}
   \]

   When \( H(S_1|S_2) < H(S_1|S_3) \),
   \[
   E_{1,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_3)}{h_{13}^2} + \left( \frac{1}{h_{12}^2} - \frac{1}{h_{13}^2} \right)^+ H(S_1|S_2) \right). \tag{2.42}
   \]

   Here \( x^+ = \max\{x, 0\} \).

2. **Greedy strategy:**

   \[
   E_{2,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1|S_2)}{(h_{13}^2 + h_{23}^2)h_{12}^2} \right). \tag{2.43}
   \]

3. \( E_{2,m} < E_{1,m} \).

**Proof:** Please refer to Appendix A.9.

Combined with the results in Section 2.3, this result argues strongly for receiver cooperation in the multicast scenario even under the stringent half-duplex and total power constraints. Finally, Figures 2.13 and 2.14 validate our theoretical claims.
Figure 2.13: The bandwidth expansion factor of various schemes in the multicast channel with side-information, $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 19.5$ dB, $H(S_1|S_2) = 0.9$, $H(S_1|S_3) = 0.3$.

### 2.4.2 Multicast Scheduler

The second step in the proposed solution for the three-way channel is the design of the scheduler. The **optimal** scheduler will choose the multicast order corresponding to the minimum bandwidth expansion factor among all possible permutations. The following result argues for the efficiency of our proposed cooperation scheme for the three-way channel.

**Theorem 15** The proposed multicast with side information scheme with the optimal scheduler has the following properties (in the three-way channel):

1. It is asymptotically optimal, i.e., approaches the genie-aided bound, when any one of the channel coefficients is sufficiently large.
2. It always outperforms the broadcast with a degraded set based multicast scheme with the optimal scheduler.

Proof: Please refer to Appendix A.10.

To reduce the computational complexity of the scheduler, one can adopt the following greedy strategy. At the beginning of every multicast stage, every node that has not finished multicasting yet will calculate its expected bandwidth expansion factor based on the cooperative scheme for multicast with side-information. The greedy scheduler chooses the node with the least expected bandwidth expansion factor to transmit at this stage. After this node finishes and the side-information is updated, the scheduler computes the expected bandwidth expansion factor for the rest of nodes and selects the one with the least
bandwidth expansion factor to multicast next. In general, this greedy scheduler constitutes a potential source for further sub-optimality. However, it approaches the genie-aided bound in the asymptotic limit when one of the channel gains is sufficiently large. Take $h_{23} \to \infty$ as an example. In this case one can easily verify that, if one of the following conditions is satisfied, then the greedy scheduler will approach the genie-aided bound:

1) $H(S_2|S_1) < \min\{H(S_1|S_2), H(S_1|S_3), H(S_3|S_1)\}$ and $H(S_3|S_1, S_2) < H(S_1|S_2, S_3)$,

2) $H(S_3|S_1) < \min\{H(S_1|S_2), H(S_1|S_3), H(S_2|S_1)\}$ and $H(S_2|S_1, S_3) < H(S_1|S_2, S_3)$.

Figure 2.15: The ratio of the energy required per source observation of the list source-channel decoding scheme to the genie aid bound.

The numerical results in Figures 2.15 and 2.16 validate our claims on the efficiency of the proposed cooperation strategy. These figures compare the minimum energy required per source observation by each scheme. In our experiments, we randomly generate the channel coefficients according to a normalized zero-mean Gaussian distribution. For each source realization, we randomly generate $H(S_1), H(S_2), H(S_3)$ with uniform distribution in $[0, 1]$.
The ratio of minimum energy required per source observation to the genie-aided bound

Figure 2.16: The ratio of the energy required per source observation of the broadcast with degraded information set based multicast with optimal scheduler to the genie aid bound.

(corresponds to binary source). To ensure \(0 \leq H(S_1|S_2) \leq H(S_1)\), we generate \(H(S_1, S_2)\) according to uniform distribution in \([\max\{H(S_1), H(S_2)\}, H(S_1) + H(S_2)]\). Similarly, we generate \(H(S_1, S_3), H(S_2, S_3), H(S_1, S_2, S_3)\) to ensure that all the entropy inequalities are satisfied. Then, we use the entropy equalities to compute all other conditional entropy [82]. For example, \(H(S_1|S_2, S_3)\) is given \(H(S_1|S_2, S_3) = H(S_1, S_2, S_3) - H(S_2, S_3)\). In such way, for each realization, we get different source entropies and correlation patterns that satisfy the appropriate entropy inequalities. For each (source and channel) realization, we then use numerical methods to find the optimal order and greedy order for cooperative scheme, and the corresponding minimum energy required per source observation, namely \(E_{oc}, E_{gc}\). We also find the optimal order for the non-cooperative scheme and the corresponding minimum energy required per source observation \(E_{nc}\). The minimum energy required per source
observation by the genie-aided bound $E_{gen}$ is used as a benchmark. In particular, for each realization, we calculate the ratio of the minimum energy required by the three schemes to the genie aided bound. We repeat the experiment 100000 times and report the histogram of the ratios in the figures. In Fig. 2.15, we see that 94 percent of the time, the proposed cooperative scheme with the greedy scheduler operates within 3 dB of the genie-aided bound. We also see that the performance of greedy scheduler is almost identical to the optimal scheduler. Figure 2.16 shows that the non-cooperative scheme operates more than 3 dB away from the genie-aided bound for 90 percent of the time. Moreover, there is a non-negligible probability, i.e., 8 percent, that this scheme operates 100 dB away from genie aided bound. It is clear that receiver cooperation reduces this probability significantly.

### 2.5 Summary

In this chapter, we have adopted a formulation of the three-node wireless network based on the half-duplex and total power constraints. In this setup, we have proposed a greedy cooperation strategy in which the weak receiver first helps the strong receiver to decode in a CF configuration. After successfully decoding, the strong user starts assisting the weak user in a DF configuration. We have shown that different instantiations of this strategy yield excellent performance in the relay channel with noisy feedback, multicast channel, and three-way channel. Our analysis for the achievable rates in such special cases sheds light on the value of noisy feedback in the relay channel and the need for a list source-channel decoding approach to efficiently exploit receiver side information in the wireless setting.
CHAPTER 3

COOPERATION FOR SECRECY

In this chapter, taking security issues into consideration, we identify a novel role of the relay node in establishing a secure link from the source to the destination. More precisely, we consider the transmission of confidential messages from the source to the destination under the help of a relay node, in the presence of an eavesdropper that intends to intercept the messages. The main motivation for this work is that the secrecy capacity of the wiretap channel is zero, when the wiretapper channel is less noisy than the main channel. Our goal is to construct cooperation schemes that enable the transmission of confidential messages under these disadvantageous situations. Towards this end, we first derive an upper-bound for the rate-equivocation region of the relay-eavesdropper channel, thus characterize the performance limit of this channel. We then devise several cooperation schemes that achieve positive perfect secrecy rate when the secrecy capacity is zero with the absence of the relay node. Several unique features of the relay channel with security constraints are then illustrated by the proposed novel Noise-Forwarding (NF) strategy.

3.1 The Relay-Eavesdropper Channel

We consider a four-terminal discrete channel consisting of finite sets $X_1, X_2, Y, Y_1, Y_2$ and a transition probability distribution $p(y, y_1, y_2|x_1, x_2)$, as shown in Figure 3.1. Here,
\( \mathcal{X}_1, \mathcal{X}_2 \) are the channel inputs from the source and the relay respectively, while \( \mathcal{Y}, \mathcal{Y}_1, \mathcal{Y}_2 \) are the channel outputs at the destination, relay and eavesdropper respectively. We impose the memoryless assumption, i.e., the channel outputs \((y_i, y_{1,i}, y_{2,i})\) at time \(i\) only depend on the channel inputs \((x_{1,i}, x_{2,i})\) at time \(i\). The source needs to transmit messages to the destination under the help of the relay node while keeping the message as secret as possible from the wiretapper. The level of secrecy at the wiretapper is measured by equivocation rate as defined in the sequel.

To fulfill this goal, the source will send the message \(W_1 \in \mathcal{W}_1 = \{1, \cdots, M\}\) to the destination using the \((M, n)\) code consisting: 1) a stochastic encoder \(f_n\) at the source that maps the message \(w_1\) to a codeword \(x_1 \in \mathcal{X}^n_1\), 2) a relay encoder that maps the signals \((y_{1,1}, y_{1,2}, \cdots, y_{1,i-1})\) received before time \(i\) to the channel input \(x_{2,i}\), using the mapping \(\varphi_i: (Y_{1,1}, Y_{1,2}, \cdots, Y_{1,i-1}) \rightarrow X_{2,i}\), 3) a decoding function \(\phi: \mathcal{Y}^n \rightarrow \mathcal{W}_1\). The average error probability of a \((M, n)\) code is defined as

\[
P_e^n = \sum_{w_1 \in \mathcal{W}_1} \frac{1}{M} \text{Pr}\{\phi(y) \neq w_1 | w_1 \text{ was sent}\}.
\] (3.1)

The equivocation rate at the eavesdropper is defined as \(\frac{1}{n} H(W_1 | Y_2)\), which measures the wiretapper’s uncertainty about the transmitted message after seeing channel output. To keep the message secret, the source/destination pair should make this large. Roughly speaking, if the equivocation rate at the eavesdropper is arbitrarily close to the information rate, the system is called perfectly secure, which is rigorously defined as follows.

\footnote{Throughout this chapter, upper-case letter \(X\) denotes a random variable, lower-case letter \(x\) denotes a realization of the random variable, calligraphic letter \(\mathcal{X}\) denotes a finite alphabet set. Boldface letter \(\mathbf{x}\) denotes a vector, \(\{\cdot\}^T\) denotes transpose and \(\{\cdot\}^H\) denotes conjugate transpose. We also let \([x]^+ = \max\{0, x\}\).}
Figure 3.1: The relay eavesdropper channel.

**Definition 1** The rate-equivocation pair \((R_1, R_e)\) is said to be achievable if for any \(\epsilon > 0\), there exists a sequence of codes \((M, n)\) such that for any \(n \geq n(\epsilon)\), we have

\[
R_1 = \frac{1}{n} \log M, \quad (3.2)
\]

\[
P_e^n \leq \epsilon, \quad (3.3)
\]

\[
\frac{1}{n} H(W_1|Y_2) \geq R_e - \epsilon. \quad (3.4)
\]

The rate-equivocation pair characterizes the tradeoff between transmission rate and level of secrecy.

**Definition 2** A perfect secrecy rate \(R_1\) is said to be achievable, if the rate-equivocation pair \((R_1, R_1)\) is achievable.

This definitions says that if there exists a scheme for the transmitter to transmit messages at rate \(R_1\), such that the eavesdropper does not gain any information about the messages from the received signal, then rate \(R_1\) is perfectly secure. The secrecy capacity of the relay-eavesdropper channel is the largest achievable perfect secrecy rate, which is defined as follows.

**Definition 3** The secrecy capacity of the relay-eavesdropper channel is

\[
C_s = \sup \{R_1 : (R_1, R_1) \text{ is achievable}\}.
\]
Note that if $Y_2 = \phi$ (or some other constant), our model reduces to the classical relay channel without security constraints.

### 3.2 Outer Bound and Cooperation Strategies

We first establish an outer-bound on the optimal rate-equivocation region of the relay-eavesdropper channel.

**Theorem 16** *In the relay eavesdropper channel, for any rate-equivocation pair $(R_1, R_e)$ with $P^n_e \to 0$ and the equivocation rate at the eavesdropper larger than $R_e - \epsilon$, there exist some random variables $U \rightarrow (V_1, V_2) \rightarrow (X_1, X_2) \rightarrow (Y, Y_1, Y_2)$, such that $(R_1, R_e)$ satisfies the following conditions

\[
R_1 \leq \min\{I(V_1, V_2; Y), I(V_1; Y_1|V_2)\}, \\
R_e \leq R_1, \\
R_e \leq [I(V_1, V_2; Y|U) - I(V_1, V_2; Y_2|U)]^+.
\]

*Proof:* Please refer to Appendix A.11.

We now turn our attention to constructing cooperation strategies for the relay-eavesdropper channel. Our first step is to characterize the achievable rate-equivocation region of the DF scheme described in Chapter 2. Here, we use the regular coding and backward decoding scheme developed in the classical relay setting [16, 83], with the important difference that each message will be associated with many codewords in order to confuse the eavesdropper.

**Theorem 17** *The rate pairs in the closure of the convex hull of all $(R_1, R_e)$ satisfying

\[
R_1 < \min\{I(V_1, V_2; Y), I(V_1; Y_1|V_2)\}, \\
R_e < R_1, \\
R_e < \left[\min\{I(V_1, V_2; Y), I(V_1; Y_1|V_2)\} - I(V_1, V_2; Y_2)\right]^+,
\]

55
for some distribution \( p(v_1, v_2, x_1, x_2, y_1, y_2) = p(v_1, v_2)p(x_1, x_2|v_1, v_2)p(y_1, y_2, y|x_1, x_2) \), are achievable using the DF strategy.

Hence, for the DF scheme, the following perfect secrecy rate is achievable

\[
R^{(DF)}_s = \sup_{p(v_1, v_2)} \left[ \min \{ I(V_1, V_2; Y), I(V_1; Y_1|V_2) \} - I(V_1, V_2; Y_2) \right]^+.
\] (3.7)

**Proof:** Please refer to Appendix A.12.

The channel between the source and the relay becomes a bottleneck for the DF strategy when it is noisier than the source-destination channel. This motivates our Noise-Forwarding (NF) scheme, where the relay node does not attempt to decode the message but sends codewords that are independent of the source’s message. The enabling observation behind this scheme is that, in the wiretap channel, in addition to its own information, the source should send extra codewords to confuse the wiretapper. In our setting, this task can be accomplished by the relay by allowing it to send independent codewords, which aid in confusing the eavesdropper.

![Figure 3.2: The rate region of the compound MACs of the relay eavesdropper channel for a fixed input distribution \( p(x_1)p(x_2) \).](image)

Our NF scheme transforms the relay-eavesdropper channel into a compound multiple access channel (MAC), where the source/relay to the receiver is the first MAC and source/relay to the eavesdropper is the second one. Figure 3.2 shows the rate region of
these two MACs for a fixed input distribution \( p(x_1)p(x_2) \). In the figure, \( R_1 \) is the codeword rate of the source, and \( R_2 \) is the codeword rate of the relay. We can observe from Figure 2a) that if the relay node does not transmit, the perfect secrecy rate is zero for this input distribution since \( R_1(A) < R_1(C) \). On the other hand, if the relay and the source coordinate their transmissions and operate at point \( B \), we can achieve the equivocation rate \( R_e \), which is strictly larger than zero. In Figure 2b), we can still get a positive perfect secrecy rate by operating at point \( A \) in the absence of the relay. But by moving the operating point to \( B \), we can get a larger secrecy rate. This illustrates the main idea of our NF scheme. The next result establishes the achievable rate-equivocation region for the NF scheme.

**Theorem 18** The rate pairs in the closure of the convex hull of all \((R_1, R_e)\) satisfying

\[
\begin{align*}
R_1 &< I(V_1; Y|V_2), \\
R_e &< R_1, \\
R_e &< [I(V_1; Y|V_2) + \min\{I(V_2; Y), I(V_2; Y_2|V_1)\}] \\
&\quad - \min\{I(V_2; Y), I(V_2; Y_2)\} - I(V_1; Y_2|V_2)]^+, 
\end{align*}
\]

for some distribution \( p(v_1, v_2, x_1, x_2, y_1, y_2) = p(v_1)p(v_2)p(x_1|v_1)p(x_2|v_2)p(y_1, y_2, y|x_1, x_2) \), are achievable using the NF scheme.

Hence, for the NF scheme, the achievable perfect secrecy rate is

\[
R_s^{(NF)} = \sup_{p(v_1)p(v_2)} [I(V_1; Y|V_2) + \min\{I(V_2; Y), I(V_2; Y_2|V_1)\}] \\
- \min\{I(V_2; Y), I(V_2; Y_2)\} - I(V_1; Y_2|V_2)]^+. 
\]

**Proof:** Please refer to Appendix A.13.

The following comments are now in order.
1. The NF scheme is customized to the relay channel with security constraints which make the transmission of codewords that are independent of the source message reasonable. Also, in the NF scheme, the relay node does not need to listen to the source, and hence, this scheme is more suited works for relay nodes limited by the half-duplex constraint [17, 34, 70].

2. In NF cooperation, each user sends independent messages to the destination, which resembles the MAC. Hence, NF cooperation can be adapted to the multiple access eavesdropper channel where the multiple users in the MAC channel can help each other in communicating securely with the destination without listening to each other (note that the results in [41] were limited only to the case where the eavesdropped channel is a degraded version of the channel seen by the destination). Our related results will be reported elsewhere.

Now, we study another cooperation scheme that does not require decoding at the relay: Compress and Forward (CF). The CF cooperation strategy can be viewed as a generalization of NF where, in addition to the independent codewords, the relay also sends a quantized version of its noisy observations to the destination. This noisy version of the relay’s observations helps the destination in decoding the source’s message, while the independent codewords help in confusing the eavesdropper. The following result establishes the achievable rate-equivocation pair in the case when $I(X_1; \hat{Y}_1, Y | X_2) \leq I(X_1; \hat{Y}_1, Y_2 | X_2)$, \textit{i.e.}, the source-eavesdropper channel is better than the source-receiver channel, a situation of particular interest to us.
Theorem 19  The rate pairs in the closure of the convex hull of all \((R_1, R_e)\) satisfying

\[
R_1 < I(X_1; \hat{Y}_1, Y|X_2),
\]
\[
R_e < R_1,
\]
\[
R_e < \left[ R_0 + I(X_1; \hat{Y}_1, Y|X_2) - I(X_1, X_2; Y_2) \right]^+,
\]

subject to

\[
\min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - R_0 \geq I(Y_1; \hat{Y}_1|X_2),
\]

for some distribution \(p(x_1, x_2, y_1, y_2, y, \hat{y}_1) = p(x_1)p(x_2)p(y_1, y_2, y|x_1, x_2)p(\hat{y}_1|y_1, x_2)\), are achievable using CF strategy.

Proof: Please refer to Appendix A.14.

Three comments are now in order.

1. In Theorem 19, \(R_0\) is the rate of pure noise generated by the relay to confuse the eavesdropper, while \(\min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - R_0\) is the part of the rate allocated to send the compressed signal \(\hat{Y}_1\) to help the destination. If we set \(R_0 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}\), this scheme becomes the NF scheme.

2. In order to enable analytical tractability, the coding/decoding scheme used in the proof is slightly different from that of [19]. In [19], the destination uses sliding-window decoding, while our proof uses backward decoding. Hence, the bound for \(R_e\) provided here is a lower-bound for the \(R_e\) achieved by the CF scheme. One may be able to achieve a larger \(R_e\) using exactly the CF scheme proposed in [19]. But, unfortunately, we are not yet able to bound \(R_e\) when sliding-window decoding is used.
3. Compared with CF decoding, the proposed NF strategy enjoys the advantage of simplicity. Also, if one only focuses on the perfect secrecy rate, it is easy to see that these two schemes achieve identical performance. Again, this observation is limited to our lower bound on $R_e$ in Theorem 19.

3.3 Examples

This section discusses several examples that illustrate some unique features of the relay-eavesdropper channel. For simplicity, we only focus on the perfect secrecy rate of various schemes.

3.3.1 The Deaf Helper Phenomenon

The security constraints imposed on the network bring about a new phenomenon which we call the deaf helper phenomenon, where the relay node can still help even it is totally ignorant of the message transmitted from the source. In this setup, we impose an additional security constraint on the relay node.

**Definition 4** A rate $R_s$ is achievable for a deaf helper if for any $\epsilon > 0$, there exists a sequence of codes $(M, n)$ such that for any $n \geq n(\epsilon)$, we have

$$R_s = \frac{1}{n} \log M, \quad P^n_e \leq \epsilon,$$

$$\frac{1}{n} H(W_1|Y_2) \geq R_s - \epsilon, \quad \frac{1}{n} H(W_1|Y_1, X_2) \geq R_s - \epsilon. \quad (3.12)$$

In this case, the signal received by the relay node does not leak any information about the transmitted message $W_1$. This model describes a more conservative scenario where the source does not trust the relay but still wishes to exploit the benefit brought by cooperation. We assume that the relay node is not malicious and, hence, is willing to cooperate with the
Theorem 20 The perfect secrecy rate of the NF scheme with an additional security constraint on the relay node is 
\[
R_s = \max_{p(v_1)p(v_2)} \min \{R_{s1}, R_{s2}\},
\]
where
\[
R_{s1} = \left[ I(V_1; Y|V_2) + \min \{I(V_2; Y), I(V_2; Y_2|V_1)\} - \min \{I(V_2; Y), I(V_2; Y_2)\} \right]^+,
\]
\[
R_{s2} = \left[ I(V_1; Y|V_2) - I(V_1; Y_1|X_2) \right]^+.
\]

Proof: Please refer to Appendix A.15.

3.3.2 The Reversely Degraded Relay-Eavesdropper Channel

In the classical relay channel without security constraints, there exist some scenarios where the relay node does not provide any gain, for example, the reversely degraded relay channel shown in [19]. Here, we focus on this scenario and show that the relay node can still offer a gain in the presence of the eavesdropper.

Definition 5 ([19]) The relay channel is called reversely degraded, if
\[
p(y, y_1|x_1, x_2) = p(y|x_1, x_2) p(y_1|y, x_2).
\]

The following result, borrowed from [19], states the capacity of the classical reversely degraded relay channel.

Theorem 21 (Theorem 2, [19]) The capacity of the reversely degraded relay channel is
\[
C_0 = \max_{x_2} \max_{p(x_1)} I(X_1; Y|x_2).
\]

7If the relay node is malicious, it can then send signals that are dependent with signal received and then could even block the transmission of the main channel.
This result implies that the relay node should send a constant, and hence, does not contribute new information to the destination. In most channel models, the constant sent by the relay does not result in any capacity gain. The question now is whether the same conclusion holds in the presence of an eavesdropper. We first observe that the degradedness of the relay channel implies that DF and CF cooperation will not provide the destination with additional useful information. The relay node, however, can still send codewords independent of the received signal to confuse the eavesdropper, as proposed in the NF scheme. Since we do not require decoding at the relay node in the proof of Theorem 18, the degradedness imposed here does not affect the performance. Hence, we get the following achievable perfect secrecy rate for the reversely degraded relay-eavesdropper channel.

**Corollary 22** The achievable perfect secrecy rate of the reversely degraded relay eavesdropper channel is

\[
R_s = \max_{p(v_1)p(v_2)} \left[ I(V_1; Y|V_2) + \min \{ I(V_2; Y), I(V_2; Y_2|V_1) \} ight] - \min \{ I(V_2; Y), I(V_2; Y_2) \} - I(V_1; Y_2|V_2) + \right].
\]

(3.14)

### 3.3.3 The AWGN Channel

Now we consider the Gaussian relay-eavesdropper channel, where the signal received at each node is

\[
y_j[n] = \sum_{i \neq j} h_{ij} x_i[n] + z_j[n],
\]

where \( h_{ij} \) is the channel coefficient between node \( i \in \{ s, r \} \) and node \( j \in \{ r, w, d \} \), and \( z_j \) is the i.i.d Gaussian noise with unit variance at node \( j \). The source and the relay have average power constraint \( P_1, P_2 \) respectively.

In [84], it was shown that the secrecy capacity of the degraded Gaussian wiretap channel is \([C_M - C_{MW}]^+\), where \( C_M, C_{MW} \) are the capacity of the main channel and wiretap
channel, respectively. This result is also shown to be valid for stochastically degraded channel \[40\]. In our case, if the relay does not transmit, the relay eavesdropper channel becomes a Gaussian eavesdropper channel, which can always be converted into a stochastically degraded channel as done in the Gaussian broadcast channel \[76\]. Applying this result to our case, the secrecy capacity of the Gaussian eavesdropper channel without the relay node is given by 

\[
C(|h_{sd}|^2 P_1) - C(|h_{sw}|^2 P_1)^+ \quad \text{(recall that } C(x) = \frac{1}{2} \log_2(1 + x)) \]

Hence if \(|h_{sw}|^2 \geq |h_{sd}|^2\) and the relay does not transmit, the secrecy capacity is zero, no matter how large \(P_1\) is. On the other hand, as shown later, the relay can facilitate the source-destination pair to achieve a positive perfect secrecy rate under some conditions even when \(|h_{sw}|^2 \geq |h_{sd}|^2\). In the following, we focus on such scenarios.

**DF and NF**

At this point, we do not know the optimal input distribution that maximizes \(R_s^{(DF)}\), \(R_s^{(NF)}\). Here, we let \(V_1 = X_1, V_2 = X_2\) and use a Gaussian input distribution to obtain an achievable lower bound.

For DF cooperation scheme, we let \(X_2 \sim \mathcal{N}(0, P_2), X_{10} \sim \mathcal{N}(0, P)\), where \(\mathcal{N}(0, P)\) is the Gaussian distribution with zero mean and variance \(P\). Also, we let

\[X_1 = c_1 X_2 + X_{10},\]

where \(c_1\) is a constant to be specified later. In this relationship, the novel information is modelled by \(X_{10}\), whereas \(X_2\) represents the part of the signal which the source and the relay cooperate in beamforming towards the destination. To satisfy the average power constraint at the source, we require \(|c_1|^2 P_2 + P \leq P_1\).
Straightforward calculations result in

\[ I(X_1; Y_1|X_2) = C(|h_{sr}|^2 P), \]
\[ I(X_1, X_2; Y) = C(|h_{sd}|^2 + |h_{rd}|^2 P_2 + |h_{sd}|^2 P), \]
\[ I(X_1, X_2; Y_2) = C(|h_{sw}|^2 + |h_{ru}|^2 P_2 + |h_{sw}|^2 P). \]

Hence, we have

\[ R_s^{(OF)} = \max_{c_1, P_1} \left[ \min \left\{ \frac{1}{2} \log \left( \frac{1 + |h_{sr}|^2 P}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2 + |h_{sw}|^2 P} \right), \right. \]
\[ \left. \frac{1}{2} \log \left( \frac{1 + |h_{sd}|^2 + |h_{rd}|^2 P_2 + |h_{sd}|^2 P}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2 + |h_{sw}|^2 P} \right) \right\}^+. \] (3.15)

For NF, we let \( X_1 \sim \mathcal{N}(0, P_1) \), \( X_2 \sim \mathcal{N}(0, P_2) \). Here \( X_1, X_2 \) are independent, resulting in

\[ I(X_1; Y|X_2) = C \left( |h_{sd}|^2 P_1 \right), \]
\[ I(X_1, X_2; Y) - I(X_1, X_2; Y_2) = \frac{1}{2} \log \left( \frac{1 + |h_{sd}|^2 P_1 + |h_{rd}|^2 P_2}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2} \right), \]
\[ I(X_2; Y_2|X_1) + I(X_1; Y|X_2) - I(X_1, X_2; Y_2) = \frac{1}{2} \log \left( \frac{(1 + |h_{ru}|^2 P_2)(1 + |h_{sd}|^2 P_1)}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2} \right). \]

Hence, we have

\[ R_s^{(NF)} = \left[ \min \left\{ C \left( |h_{sd}|^2 P_1 \right), \frac{1}{2} \log \left( \frac{1 + |h_{sd}|^2 P_1 + |h_{rd}|^2 P_2}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2} \right), \right. \]
\[ \left. \frac{1}{2} \log \left( \frac{(1 + |h_{ru}|^2 P_2)(1 + |h_{sd}|^2 P_1)}{1 + |h_{sw}|^2 P_1 + |h_{ru}|^2 P_2} \right) \right\}^+. \] (3.16)

**Amplify and Forward**

In this subsection, we quantify the achievable secrecy rate of Amplify and Forward (AF) cooperation\(^8\). In AF, the source encodes its messages into codewords with length

\(^8\)We did not consider this scheme in the discrete case since, in general, it does not lend itself to a single letter characterization.
ML each, and divides each codeword into \( L \) sub-blocks each with \( M \) symbols, where \( L \) is chosen to be sufficiently large. At each sub-block, the relay sends a linear combination of the received noisy signal of this sub-block so far. For simplicity, we limit our discussion to \( M = 2 \). In this case, the source sends \( X_1(1) \) at the first symbol interval of each sub-block, the relay receives \( Y_1(1) = h_{sr} X_1(1) + Z_1(1) \); At the second symbol interval, the source sends \( \alpha X_1(1) + \beta X_1(2) \), while the relay sends \( \gamma Y_1(1) \). Here \( \alpha, \beta, \gamma \) are chosen to satisfy the average power constraints of the source and the relay. Thus, this scheme allows beam-forming between the source and relay without requiring the relay to fully decode.

Writing the signal received at the destination and the eavesdropper in matrix form, we have

\[
\mathbf{Y} = \mathbf{H}_1 \mathbf{X}_1 + \mathbf{Z}, \quad \mathbf{Y}_2 = \mathbf{H}_2 \mathbf{X}_1 + \mathbf{Z},
\]

(3.17)

where

\[
\mathbf{H}_1 = \begin{bmatrix}
h_{sd} & 0 \\
\beta h_{sd} + \gamma h_{sr} h_{rd} & \alpha h_{sd}
\end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix}
h_{sw} & 0 \\
\beta h_{sw} + \gamma h_{sr} h_{rw} & \alpha h_{sw}
\end{bmatrix},
\]

\[
\mathbf{X}_1 = [X_1(1), X_1(2)]^T, \quad \mathbf{Z} = [Z(1), \gamma h_{rd} Z_1(1) + Z(2)]^T, \quad \mathbf{Z}_2 = [Z_2(1), \gamma h_{rw} Z_1(1) + Z_2(2)]^T, \quad \mathbf{Y} = [Y(1), Y(2)]^T, \quad \mathbf{Y}_2 = [Y_2(1), Y_2(2)]^T.
\]

(3.18)

The channel under consideration can be viewed as an equivalent standard memoryless eavesdropper channel with input \( \mathbf{X}_1 \) and outputs \( \mathbf{Y}, \mathbf{Y}_2 \) at the destination and the eavesdropper respectively. Then, based on the result of [33], an achievable perfect secrecy rate is \( [I(\mathbf{X}_1; \mathbf{Y}) - I(\mathbf{X}_1; \mathbf{Y}_2)]^+ \).
Choosing a Gaussian input with covariance matrix $\mathbb{E}\{XX^H\} = PI$, where $I$ is the identity matrix, we get the following perfect secrecy rate

$$R_s^{(AF)} = \max_{\alpha,\beta,\gamma,P} \left[ \frac{1}{4} \log \left| \frac{\det \{PH_1H_1^H + \mathbb{E}\{ZZ^H\}\} + \gamma h_{rd}}{\det \{\mathbb{E}\{ZZ^H\}\}} \right| \right] +$$

$$- \frac{1}{4} \log \left| \frac{\det \{PH_2H_2^H + \mathbb{E}\{Z_2Z_2^H\}\} + \gamma h_{rw}}{\det \{\mathbb{E}\{Z_2Z_2^H\}\}} \right|$$

$$= \max_{\alpha,\beta,\gamma,P} \left[ \frac{1}{4} \log \left| \frac{\det \{PH_1H_1^H + A\} \det B}{\det \{PH_2H_2^H + B\} \det A} \right| \right], \quad (3.19)$$

where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 + |\gamma h_{rd}|^2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 + |\gamma h_{rw}|^2 \end{bmatrix},$$

and the maximization is over the set of power constraints:

$$(1 + |\alpha|^2 + |\beta|^2)P \leq 2P_1,$$

$$|\gamma|^2(|h_{sr}|^2P + 1) \leq 2P_2. \quad (3.20)$$

Numerical Results

![Figure 3.3: The achievable perfect secrecy rate of the proposed schemes in the Gaussian relay eavesdropper channel.](image)

In this subsection, we give numerical results under two channel models. The first is the real channel where $h_{ij} = d_{ij}^{-\gamma}$, with $d_{ij}$ being the distance between node $i$ and $j$ and $\gamma > 1$.
is the channel attenuation coefficient. In the second model, we assume that each channel
experiences an independent phase fading, that is \( h_{ij} = d_{ij}^{-\gamma}e^{j\theta_{ij}} \), where \( \theta_{ij} \) is uniformly
distributed over \([0, 2\pi)\). We believe that the second model is more practically relevant than
the real channel scenario.

\[ h_{ij} = d_{ij}^{-\gamma}e^{j\theta_{ij}}, \]

\( \theta_{ij} \) is uniformly distributed over \([0, 2\pi)\).

Figure 3.3 shows the achievable perfect secrecy rate of the proposed schemes for the
first channel model. In generating this figure, we use the network topology shown in Fig-
ure 3.4, where we put the source at \((0, 0)\), the destination at \((1, 0)\), the eavesdropper at
\((0, 1)\), and the relay node at \((x, 0)\). We let \( P_1 = 1, P_2 = 8 \). Since \( d_{sd} = d_{sw} \), the perfect
secrecy capacity of the eavesdropper channel without the relay node is zero. But, as shown
in the figure, we can achieve a positive secrecy rate by introducing a relay node. In comput-
ing the upper-bound, we set \( V_1 \sim \mathcal{N}(0, P_1), V_2 \sim \mathcal{N}(0, P_2) \) with a correlation coefficient
\( \rho \), and maximize over \( \rho \in [-1, 1] \). Note that the Gaussian input is not necessarily optimal
for the upper-bound. We can see that, when the relay is near the source, the DF scheme
touches the Gaussian upper-bound. Also, when \( x > 1 \), it is clear that DF cooperation does
not offer any gain, while NF and AF still offer positive rates. Note that when \( x > 1 \), both
d_{sr}, d_{sd} are larger than d_{sw}. The interesting observation here is that though both the des-
tination and relay are in disadvantageous positions compared with the eavesdropper, they
can cooperate with each other and gain some advantage over the eavesdropper. If the relay is at 0, our model is equivalent to the case where the source has two antennas. Note that the upper-bound of the perfect secrecy capacity is zero under this scenario. Hence, increasing the number of transmitting antenna at the source does not increase the secrecy capacity under the real channel model. On the other hand, if there is a relay node at an appropriate position, we can exploit this relay node to establish a secure source-destination link.

![Figure 3.5: The achievable perfect secrecy capacity for various schemes in the Gaussian relay eavesdropper channel with phase fading.](image)

In the second scenario, we assume that before transmission, the source knows the phases $\theta_{sr}, \theta_{sd}, \theta_{rd}$, but does not know $\theta_{sw}, \theta_{rw}$. The random phase will not affect the achievable perfect secrecy rate of NF since it does not depend on beam-forming between the source and relay. But, the rates of DF and AF are different here. In both cases, the source can adjust its phase according to the knowledge of the phase information about $\theta_{sr}, \theta_{sd}, \theta_{rd}$. In this way, the signals of the source and the relay will add up coherently at the destination, but not at the eavesdropper since $\theta_{sw}, \theta_{rw}$ are independent of $\theta_{sd}, \theta_{rd}, \theta_{sr}$. The secrecy rate of DF and AF could then be obtained by averaging (3.15), (3.19) over the random phases. Figure 3.5 shows the achievable perfect secrecy rates of the proposed
strategies for the same setup as the first scenario. Due to the random phases, the achievable
perfect secrecy capacity when the relay is at the same position as the source is not zero any-
more. In this case, it will be beneficial to have multiple transmitting antennas at the source.
Similar to the first scenario, when $x > 1$, DF cooperation does not offer any benefit. But
both NF and AF still enjoy non-zero secrecy rates.

3.4 Summary

In this chapter, the relay-eavesdropper channel was studied. In particular, several co-
operation strategies were proposed and the corresponding achievable performance bounds
were obtained. Furthermore, an outer-bound on the optimal rate-equivocation region for
this channel was developed. Of particular interest is the proposed NF strategy which was
used to illustrate the deaf-helper phenomenon, and to demonstrate the utility of the relay
node in the reversely degraded relay-eavesdropper channel. Overall, our results establish
the critical role of user cooperation in facilitating secure wireless communications and shed
light on the unique feature of the relay-eavesdropper channel.
CHAPTER 4

COOPERATION IN ENERGY LIMITED NETWORKS

In the preceding chapters, we have shown the key role of user-cooperation in increasing network throughput and confidentiality of messages. The utility of user cooperation in many other practical scenarios has also been established by a large body of research works (e.g., [16, 17, 22] and references therein). A fundamental assumption, where these works are based on, is that each user in the network will spend its own valuable energy on sending others’ information. The popularity of ad-hoc networks and the increased programmability of wireless devices, however, raises serious doubts on the validity of this assumption. This chapter develops a game-theoretic framework to study user-cooperation in an energy-limited network. Instead of assuming that users are cooperating, we investigate conditions under which user-cooperation will consist of a Nash Equilibrium (NE)\(^9\) for a network consisting of selfish and rational users. We consider both small and large scale networks. In the small scale network, we show that a relay node with an appropriate relaying strategy successfully turns the NE from non-cooperation to full-cooperation under some network topologies. In the large scale random network, we show that as long as a vanishingly small

\(^9\)Roughly speaking, a NE is a strategy profile where none of the participants can benefit by deviating from this strategy profile. Rigorous definition will be given in the sequel.
fraction of the nodes in the network are altruistic, full-cooperation is a NE of the corresponding game. Under the full-cooperation NE, user-cooperation is not an assumption but the self-interest of the users in the network.

4.1 Network Model

We consider an ad-hoc wireless network with a set of half-duplex wireless nodes \( N \), each located at \( X_i, i \in N \). Our nodes are assumed to have very low rate requirements but a stringent energy budget. Among these \( N = |N| \) wireless nodes, there exists a set of transmitters \( T \subseteq N \) that wish to send packets to different destinations. A time division multiple access (TDMA) approach is used where the nodes will take turns in transmitting their packets (no frequency reuse). This multiple access scheme is known \([23, 24, 85]\) to be optimal from a minimum energy per bit perspective\(^{10}\), the figure of merit considered in this chapter. In the frame assigned to the transmitter \( i \), it will first send a relay request to nodes in the route (we will specify the routing scheme in the sequel) to its destination. Once the relay request is received, transmitter \( j \in T \) will decide whether to accept or reject the relay request and send back its decision. We adopt the discrete additive white Gaussian noise (AWGN) model, i.e., when node \( i \) transmits, the signal received by node \( j \) at time \( t \) is given by

\[
y_j[t] = h_{ij} x_i[t] + z_j[t],
\]

where the channel gain \( h_{ij} \) from node \( i \) to node \( j \) is \( h_{ij}^2 = d_{ij}^{-\gamma} \), \( d_{ij} = |X_i - X_j| \) is the distance between the two nodes, \( \gamma > 2 \) is the channel attenuation coefficient, \( z_j[t] \) is the

\(^{10}\)As pointed out by [85], TDMA is only first-order optimal. If, besides minimum energy per bit, one is also interested with spectral efficiency, superposition schemes like code-division multiple access are more efficient [85].
zero-mean unit-variance Gaussian noise at node $j$. The noise process is assumed to be spatially and temporally white.

Our work focuses on analyzing the decision making process of accepting/rejecting relay requests generated by the different transmitters. In a fully cooperative scenario, each node will accept all relay requests in order to minimize the overall energy consumption of the network. However, the selfishness of the nodes in our network leads to a significantly more complicated process. By accepting a relay request, the individual node is committing to spend some of its valuable energy in forwarding other nodes’ packets. From a selfish perspective, therefore, each node will be tempted to reject the relay requests in order to minimize its own energy consumption. This motivates our formulation which differentiates between two classes of nodes in our network, namely selfish and altruistic nodes. A selfish node is assumed to be interested in minimizing its own energy consumption whereas an altruistic node is interested in minimizing the overall energy expenditure of the network. In practice, those altruistic nodes may correspond to wireless relay nodes tasked with rewarding cooperation in the network. This correspondence justifies our assumption that the altruistic nodes are endowed with more processing power, storage, and monitoring capabilities as compared with the selfish ones. In the sequel, we will refer to the set of altruistic nodes as $U \subseteq N$ and let $\Theta(N) = |U|$ be the number of such nodes.

We model the behavior of selfish nodes as a static non-cooperative game $G = \{M, F, C\}$, where $M = T \setminus U$ is the set of selfish transmitters, $F_i$ is the strategy space of node $i$, and $C_i$ is the cost function of node $i$. The strategy space is given by $F_i = \{p_i : p_i \in [0, 1]\}$, where $p_i$ is the probability that node $i$ will accept a relay request. Here, we observe that our strategy space allows for mixed strategies but does not allow the selfish nodes to offer differentiated service based on the origin of the relay request. In our formulation this
latter feature is reserved for the altruistic nodes. More specifically, an altruistic node will monitor the behavior of the selfish nodes and accept the relay request of selfish node $i$ with probability $g(p_i)$. Under the assumption that the function $g(\cdot)$ is known \textit{a-priori} by all the nodes, one should construct $g(\cdot)$ in order to encourage cooperation between all nodes (i.e., convince node $i$ to work with a higher $p_i$). To simplify our task, we limit ourselves to linear functions of the form $g(p_i) = p_i + (1 - p_i)p_{nc}$, where $p_{nc} \in [0, 1]$. Intuitively, this form can be interpreted as a mixed strategy where the altruistic node \textbf{deterministically} accept the relay requests originated from cooperative nodes (i.e., $g(1) = 1$), but only accepts the requests of non-cooperative node with probability $p_{nc}$ (i.e., $g(0) = p_{nc}$). We note that $p_{nc} = 1$ corresponds to the \textbf{traditional relaying} where the altruistic nodes accept all the relay requests. For simplicity of presentation, and motivated by the results of [23, 24], our cost function corresponds only to the transmission energy\textsuperscript{11}. Therefore, the cost function is given by $C_i = E_{i, tr} + E_{i, re}$ where $E_{i, tr}$ is the energy spent on transmitting the $i^{th}$ node own packets, and $E_{i, re}$ is the energy spent by the $i^{th}$ node on forwarding packets for other transmitters.

\textbf{Definition 6} \textit{A set of strategies $p^* = (p^*_1, ..., p^*_M)$ is said to constitute a Nash Equilibrium (NE) if for any $i \in \mathcal{M}$}

$$C_i(p^*_i, p^\_i) \leq C_i(p'_i, p^\_i), \forall p'_i \in [0, 1],$$ (4.1)

\textit{where $p\_i$ is the strategies of all the players except user $i$.}

Our work seeks to characterize the NE of the aforementioned non-cooperative game in the two extreme cases of small and large networks. This characterization establishes

\textsuperscript{11}As pointed out in [23, 24], the transmission energy is the dominant factor in the energy consumption in the large scale limit. We also note that our conclusions will still hold if the receiving energy is incorporated into our model but the development of the results will be more involved.
rigorously the critical role of wireless altruistic nodes in stimulating cooperation in ad-hoc networks.

4.2 The Small Network Scenario

![Figure 4.1: A small wireless network with two selfish senders, one altruistic node and two receivers.](image)

Aiming towards a succinct development of our ideas, we start in this section with a small network composed of five nodes as shown in Figure 4.1. In this setup, \( \mathcal{N} = \{1, 2, 3, 4, r\} \), \( \mathcal{T} = \{1, 3, r\} \), and \( \mathcal{U} = \{r\} \). In other words, nodes 1, 3 take turns to transmit information to nodes 2, 4 at rate \( R \) and \( r \) is our altruistic relay node which does not have information to send (as illustrated in Figure 4.2). Due to the strict energy limitation imposed on the nodes, we operate in the asymptically low rate regime, i.e., \( R \to 0 \), which is known to be the most energy efficient operating point [86]. The following lemma motivates the idea of using the relay node to stimulate cooperation.

![Figure 4.2: The TDMA data frame.](image)
Lemma 23  In the absence of the relay node $r$, no-cooperation is the only NE of our static game. Moreover, no-cooperation remains the only NE under the repeated game setup if it is known a-priori that the nodes will only interact for a finite number of stages and the nodes know the end of the game.

Proof:  Let’s consider the last stage in the repeated game scenario. In this stage, it is easy to see that the optimal strategy for any selfish node $i \in \mathcal{M}$ is to use adopt $p_i = 0$. The reason is that, for any other $p_i \neq 0$, node $i$ can reduce $E_{i, re}$ without affecting $E_{i, tr}$, and hence, reduce its overall cost by adopting $p_i = 0$. The selfish wireless users can apply the same reasoning backward until the first stage. No cooperation at each stage is, therefore, the only Nash equilibrium.

At this no-cooperation equilibrium, each user will have to use sufficient transmission energy to allow the packet to reach the destination directly. For $\gamma > 2$, this implies a strictly higher energy consumption by each user, as compared to the globally optimal solution. By introducing the node $r$, which adopts the appropriate cooperation strategy, we will show in the following that cooperation emerges at the NE, which translates into a significant reduction in the energy per bit requirements. Note that if the relay always helps, the reasoning in Lemma 23 still applies, and hence, no-cooperation remains the only NE.

4.2.1 The Cooperation Scheme

Before proceeding further, we describe here the underlying cooperation scheme used in this section. Assuming that a certain node decided to cooperate with the source node, the small size of our network allows for using a more sophisticated approach, as compared with simple multi-hop forwarding. We adopt the Decode-Forward (DF) algorithm as described
in the preceding chapters. With a focus on energy efficiency in this chapter, we first calculate the energy efficiency of the DF scheme in the traditional three terminal setting (i.e., one source, one destination, and one relay which is always willing to help) in this section. This setting is sufficient to capture the main idea behind DF cooperation and, moreover, the following section outlines the generalized scenario with multiple helpers. Recall that, in DF, the relay node attempts to decode the information from the source first, then the source and the relay cooperate in transmitting the information to the destination. The basic idea is that signals from the source and the relay will add up coherently at the receiver, and hence, we can reap the beam-forming gain. Without loss of generality, we let node 1 be our source, node 2 be our destination, and node r be our relay.

In the absence of the relay node, the source-destination channel capacity and minimum energy per bit are given by [76]

\[ C = \frac{1}{2} \log(1 + h_{12}^2 P), \quad E_1 = \lim_{C \to 0} \frac{P}{C} = \frac{2 \ln 2}{h_{12}^2}, \]  

(4.2)

where \( h_{12} \) is the source-destination channel gain and \( P \) is the normalized transmit power (assuming a unit variance AWGN process). Throughout the sequel, we will simplify our analysis by assuming that a capacity achieving code is used, i.e., \( R = C \). This assumption will allow for using the minimum energy per bit, e.g., (4.2), as our figure of merit.

Figure 4.3 illustrates the two-stage DF cooperation scheme. Using Gaussian codebook of length \( L \) and rate \( R \), the source sends information with power \( P \) in the first phase (Figure 4.3(a)). The relay node keeps listening until the accumulated mutual information becomes larger than \( LR \), and then tries to decode. The decoding at the relay node will be successful with probability going to 1 as \( L \) increases [22]. After decoding, the relay joins
Figure 4.3: The cooperation scheme in the small network: a) both the relay and node 2 keep receiving, lasting $\mu T$ b) both node 1 and the relay transmit, lasting $(1 - \mu)T$.

in the transmission, while the source keeps sending but with a lower power level in the second phase (Figure 4.3(b)).

More specifically, the relay node will be able to decode the message after receiving $\lceil\mu L\rceil$ symbols [22], where

$$\mu = \lim_{P \to 0} \frac{\log(1 + h_{12}^2 P)}{\log(1 + h_{1r}^2 P)} = \frac{h_{12}^2}{h_{1r}^2}.$$  

After successful decoding, the relay starts sending the decoded information using the same codebook as the source node with power $\alpha_2^2 P$ whereas the source node now lowers its transmit power to $\alpha_1^2 P$ ($\alpha_1 \leq 1$ and $\alpha_r \leq 1$). The signal received at the destination during this phase is

$$y_2[t] = h_{12} \alpha_1 x_1[t] + h_{r2} \alpha_r x_1[t] + z_2[t].$$

Here $\alpha_1, \alpha_r$ are obtained by solving the following optimization problem

$$\min \quad \alpha_1^2 + \alpha_r^2,$$  

$$s.t. \quad h_{12} \alpha_1 + h_{r2} \alpha_r = h_{12}, \alpha_1, \alpha_r \in [0, 1].$$

This constrained optimization minimizes the total transmit power while maintaining the same effective capacity, from the destination perspective. Straightforward calculations
yield the following optimal values for $\alpha_1$, $\alpha_2$, the energy per bit expended by node 1, and the minimum total energy per bit expended by the network (i.e., nodes 1 and $r$), under DF cooperation

$$
\alpha_1 = \frac{h_{1r}^2}{h_{12}^2 + h_{r2}^2}, \quad \alpha_r = \frac{h_{12}h_{r2}}{h_{12}^2 + h_{r2}^2},
$$

$$
E_{1,s} = \frac{2 \ln 2}{h_{1r}^2} + \left(1 - \frac{h_{12}^2}{h_{1r}^2}\right) \frac{2 \ln 2 h_{12}^2}{(h_{12}^2 + h_{r2}^2)^2}, \quad E_t = \frac{2 \ln 2}{h_{1r}^2} + \left(1 - \frac{h_{12}^2}{h_{1r}^2}\right) \frac{2 \ln 2}{h_{12}^2 + h_{r2}^2}. \quad (4.4)
$$

Comparing (4.2) and (4.4) shows that, as expected, the relay node offers gains in the energy consumption for $h_{1r} > h_{12}$ (i.e., setting $h_{1r} = h_{12}$ yields $E_t = E_1$).

### 4.2.2 Equilibrium Analysis

To obtain the equilibrium point of our non-cooperative game, it is sufficient to calculate the cost functions under the pure strategies, i.e., cooperation and no-cooperation, as shown in the Table 4.2.2.

1) (no-cooperation, no-cooperation)

In this case, the relay will help the two nodes with probability $p_{nc}$ ($g(0) = p_{nc}$). Using the results developed in the previous section on DF cooperation, we can see that the average energy per bit required for the two nodes are

$$
C_{1,nn} = (1 - p_{nc})E_1 + p_{nc}E_{1,s}, \quad C_{3,nn} = (1 - p_{nc})E_3 + p_{nc}E_{3,s},
$$

where $E_3, E_{3,s}$ can be obtained from (4.2) and (4.4) by replacing subscripts 1 and 2 with 3 and 4, respectively.
2)(cooperation, no-cooperation)

Let’s consider first a packet transmitted by node 1. Since the relay node always cooperates ($g(1) = 1$) but node 3 doesn’t, this scenario reduces to the classical three terminal DF cooperation described in section 4.2.1, and hence, the energy consumption for node 1 on its packet is $E_{1, tr} = E_{1, s}$ and the energy consumption by node 3 on relaying is $E_{3, re} = 0$.

![Figure 4.4](image-url)

Figure 4.4: The operation sequence when node 1 helps node 3 and $h_{31} > h_{3r}$, (a) node 1 and the relay keep listening, lasting $\mu_{3,1}'T$ (b) node 1 helps transmitting, lasting $\mu_{3,2}'T$ part of the time (c) both node 1 and the relay help, lasting $(1 - \mu_{3,1}' - \mu_{3,2}')T$.

On the other hand, when a packet is generated by node 3, we have two distinct scenarios. When the relay doesn’t cooperate, which happens with probability $1 - p_{nc}$, the energy consumption of the nodes can be written as $\varphi_1, E_3 - \varphi_3$ respectively, where

$$\varphi_1 = \left(1 - \frac{h_{34}^2}{h_{31}^2}\right)\frac{h_{14}^2h_{34}^2}{(h_{34}^2 + h_{14}^2)^2}E_3, \quad \varphi_3 = \left(1 - \frac{h_{34}^2}{h_{31}^2}\right)\frac{2h_{34}^2h_{14}^2 + h_{14}^4}{(h_{34}^2 + h_{14}^2)^2}E_3, \quad (4.5)$$

based on the DF cooperation scheme with node 1 now playing the role of the helper node. If the relay node also cooperates, which happens with probability $p_{nc}$, the energy consumption of the nodes depends on the decoding order of node 1 and the relay nodes. Here, we assume that both $h_{31}, h_{3r}$ are larger than $h_{34}$ since otherwise this scenario is reduced to the three terminal case considered earlier. If $h_{31} > h_{3r}$, the operation sequence is shown in
Figure 4.4. After a

\[
\mu_{3,1} = \lim_{P_3 \to -0} \left( \frac{\frac{1}{2} \log(1 + h_{34}^2 P_3)}{\frac{1}{2} \log(1 + h_{31}^2 P_3)} \right) = \frac{h_{34}^2}{h_{31}^2}
\]

fraction of the frame, node 1 will decode the information and start to help. In this period, the relay node will still keep on listening, and

\[
y_4[t] = h_{14} \alpha_1 x_3[t] + h_{34} \alpha_3 x_3[t] + z_4[t], \quad y_r[t] = h_{1r} \alpha_1' x_3[t] + h_{3r} \alpha_3' x_3[t] + z_r[t],
\]

where \(\alpha_1' = \frac{h_{14} h_{34}}{h_{14}^2 + h_{34}^2}, \alpha_3' = \frac{h_{34}^2}{h_{14}^2 + h_{34}^2}\) are the solution to an optimization problem similar with (4.3). Then, after an additional

\[
\mu_{3,2} = \lim_{P_3 \to -0} \left( \frac{\frac{1}{2} \log(1 + h_{34}^2 P_3) - \frac{\mu_{3,1}'}{2} \log(1 + h_{3r}^2 P_3)}{\frac{1}{2} \log(1 + (h_{1r} \alpha_1' + h_{3r} \alpha_3')^2 P_3)} \right) = \frac{(h_{31}^2 - h_{3r}^2)(h_{14}^2 + h_{34}^2)^2}{(h_{1r} h_{14} + h_{34} h_{3r})^2 h_{31}^2}
\]

fraction of the frame, the relay will be able to decode the message [22, 34], and then join in transmitting. In this period, the signal that node 4 receives is given by

\[
y_4[t] = h_{14} \beta_1 x_3[t] + h_{r4} \beta_2 x_3[t] + h_{34} \beta_3 x_3[t] + z_4[t].
\]

Similar to (4.3), the value of \(\beta_i's\) are obtained as \(\beta_1 = \frac{h_{14} h_{34}}{h_{14}^2 + h_{r4}^2 + h_{34}^2}, \beta_2 = \frac{h_{r4} h_{34}}{h_{14}^2 + h_{r4}^2 + h_{34}^2}, \beta_3 = \frac{h_{34}^2}{h_{14}^2 + h_{r4}^2 + h_{34}^2}\) which correspond to the solution to the following optimization problem

\[
\min \sum_{i=1}^{3} \beta_i^2, \quad \text{s.t.} \quad h_{14} \beta_1 + h_{r4} \beta_2 + h_{34} \beta_3 = h_{34}, \beta_i \in [0, 1].
\]

Hence if \(h_{31} > h_{3r}\), the energy consumptions of the two nodes are

\[
\varphi_1' = \mu_{3,2} \alpha_1^2 E_3 + (1 - \mu_{3,1} - \mu_{3,2}) \beta_1^2 E_3,
\]

\[
\varphi_3' = \mu_{3,1} E_3 + \mu_{3,2} \alpha_3^2 E_3 + (1 - \mu_{3,1} - \mu_{3,2}) \beta_3^2 E_3.
\]

The case where \(h_{31} \leq h_{3r}\) is similar, with the decoding order of node 1 and the relay node interchanged. The energy consumption of the nodes can, therefore, be written as

\[
\varphi_1'' = (1 - \mu_{3,1}'' - \mu_{3,2}'') \beta_1^2 E_3, \quad \varphi_3'' = \mu_{3,1} E_3 + \mu_{3,2} \alpha_3^2 E_3 + (1 - \mu_{3,1}'' - \mu_{3,2}'') \beta_3^2 E_3.
\]
Combining these two cases together, we obtain the following relationship for the energy consumption of the two nodes

\[ \varphi_{1}''' = \varphi_{1}'I(h_{31} > h_{3r}) + \varphi_{1}''I(h_{31} \leq h_{3r}), \quad \varphi_{3}''' = \varphi_{3}'I(h_{31} > h_{3r}) + \varphi_{3}''I(h_{31} \leq h_{3r}) \]

where \( I(\cdot) \) is the indicator function. Finally, the cost functions of the two nodes under the (cooperation, no-cooperation) pure strategy is

\[ C_{1,cn} = E_{1,s} + (1 - p_{nc})\varphi_{1} + p_{nc}\varphi_{1}''', \quad C_{3,cn} = (1 - p_{nc})(E_{3} - \varphi_{3}) + p_{nc}\varphi_{3}''' \]

3)(no-cooperation, cooperation)

For this case, we can follow the same steps as above, and get the cost functions of the nodes

\[ C_{1,nc} = (1 - p_{nc})(E_{1} - \psi_{1}) + p_{nc}\psi_{1}''', \quad C_{3,nc} = E_{3,s} + (1 - p_{nc})\psi_{3} + p_{nc}\psi_{3}''' \]

Here the expressions for \( \psi_{i}, \varphi_{i}''' \) are similar with \( \varphi_{i}, \varphi_{i}''' \) with the roles of node 1 and node 3 interchanged.

4)(cooperation, cooperation)

It’s easy to see that when node 3 transmits, the transmission energy of the nodes are \( \varphi_{1}''', \varphi_{3}''' \) respectively. When node 1 transmits, the transmission energy of the nodes are \( \psi_{1}''', \psi_{3}''' \) respectively. Hence the cost functions are

\[ C_{1,cc} = \psi_{1}'''' + \varphi_{1}'''', \quad C_{3,cc} = \psi_{3}'''' + \varphi_{3}''' \]

**Theorem 24** If \( C_{1,cn} < C_{1,nn} \) or \( C_{3,nc} < C_{3,nn} \), then the relay node will stimulate cooperation (i.e., no-cooperation is not a NE anymore). More strongly, If \( C_{1,cc} < C_{1,nc}, C_{3,cc} < C_{3,cn}, C_{1,en} < C_{1,nn}, C_{3,nc} < C_{3,nn} \), then full cooperation is the only NE.
Proof: If $C_{1, cn} < C_{1, nn}$, node 1 can reduce its cost by deviating from no-cooperation to cooperation. Similarly, if $C_{3, cn} < C_{3, nn}$, node 3 can reduce its cost by deviating from no-cooperation to cooperation. If $C_{1, cc} < C_{1, nc}, C_{3, cc} < C_{3, cn}, C_{1, cn} < C_{1, nn}, C_{3, nc} < C_{3, nn}$, no-cooperation is a dominated strategy, and hence, full-cooperation is the only equilibrium with $p^* = 1$. When pure strategies do not yield a NE, the users will adopt mixed strategies to arrive at an equilibrium point. For our two users matrix form game, we can readily compute those mixed strategies. At the NE, user 1 will choose $p_1^*$ such that user 3 will have the same cost under either cooperation or no-cooperation:

$$C_3^* = (1 - p_1^*)C_{3, nn} + p_1^*C_{3, cn} = (1 - p_1^*)C_{3, nc} + p_1^*C_{3, cc}.$$

User 3 will choose $p_3^*$ in a similar way, hence

$$p_1^* = \frac{C_{3, nn} - C_{3, nc}}{C_{3, cc} - C_{3, cn} + C_{3, nn} - C_{3, nc}}, \quad p_3^* = \frac{C_{1, nn} - C_{1, cn}}{C_{1, cc} - C_{1, nc} + C_{1, nn} - C_{1, cn}}.$$

and $C_1^* = (1 - p_3^*)C_{1, nn} + p_3^*C_{1, cn}$.

Under the equilibrium strategy $(p_1^*, p_3^*)$, the average total energy per bit is $E_t = \frac{1}{2}(C_1^* + C_3^* + E_r^*)$, where $E_r^*$ is the energy consumption of the relay node. Therefore, the altruistic node determines the optimal strategy which corresponds to finding $p_{nc}^* = \arg\min_{p_{nc} \in [0,1]} E_t$. In summary, the main insight from Theorem 24 is that, by adopting the appropriate cooperation strategy, the altruistic node can stimulate cooperation among nodes 1 and 3 under certain conditions on the topology of the networks, which determines the channel gain between every pair of nodes. If we have freedom to choose the altruistic node’s location, we can use Theorem 24 to compute an optimum region, in which the altruistic node can ensure full-cooperation to be the unique NE of this non-cooperative game. Noting that the
conditions in Theorem 24 are polynomial functions of altruistic node’s coordinate, the optimum region, hence, is the intersection of the four regions specified by these polynomial functions, which are easy to calculate.

4.2.3 Fading Channels

In this section, besides attenuation, we assume that the signal also experiences phase fading, i.e. \( h_{ij}^2 = d_{ij}^{-\gamma} e^{i\theta_{ij}} \), where \( \theta_{ij} \) is uniformly distributed over \([0, 2\pi)\) and is independent with everything else. We assume that \( \theta_{ij} \) is only known at the corresponding receiver \( j \) but not at the transmitter \( i \). In the following, we show that the scheme described in Section 4.2.1 becomes multi-hop packet forwarding, a simple and popular scheme in large scale networks [14].

Recall that after receiving \( \lceil LR \rceil \) symbols, the relay can decode successfully and will then send \( x_r \) with power \( \alpha_r^2 P \) to help the source. The signal received at the destination during the second phase is

\[
y_2[t] = h_{12}\alpha_1 x_1[t] + h_{r2}\alpha_r x_r[t] + z_2[t].
\]

Let \( \rho_{1r} \) be the correlation coefficient between \( x_1, x_r \), the capacity is then

\[
C = \max_{\rho_{1r}} \mathbb{E} \left\{ \frac{1}{2} \log \left( 1 + (\alpha_1^2|h_{12}|^2 + \alpha_r^2|h_{r2}|^2) + 2\Re(\rho_{1r}e^{j(\phi_{12}-\phi_{r2})})\alpha_1\alpha_r|h_{12}||h_{r2}|)P \right) \right\}
\]

\[
= \frac{1}{2} \log(1 + (\alpha_1^2|h_{12}|^2 + \alpha_r^2|h_{r2}|^2) P),
\]

(4.11)

where \( \Re(x) \) is the real part of \( x \), since \( \mathbb{E}\{e^{j(\phi_{12}-\phi_{r2})}\} = 0 \) and Jensen’s inequality.

To minimize the energy expenditure while maintaining the same effective capacity from the destination perspective, we solve

\[
\min \quad \alpha_1^2 + \alpha_r^2,
\]

\[
s.t. \quad \alpha_1^2|h_{12}|^2 + \alpha_r^2|h_{r2}|^2 = |h_{12}|^2.
\]

(4.12)
It easy to see that if $h_{r2} > h_{12}$, the solution to (4.12) is $\alpha_1^2 = 0, \alpha_r^2 = |h_{12}|^2/|h_{r2}|^2$, i.e., only the altruistic node relays the information to the destination. On the other hand, if $h_{12} > h_{r2}$, the solution is $\alpha_1^2 = 1, \alpha_r^2 = 0$, i.e., only the source node sends.

In summary, during the second phase, only the node closer to the destination will transmit, a situation similar to the multi-hop packet forwarding with the only difference that, the destination will combine the signal received both during the first and second stage to decode. The analysis in Section 4.2.2 follows naturally and similar conclusion holds, as evident in the numerical part.

4.2.4 Numerical Results

Armed with Theorem 24, we now give various examples to show the effectiveness of the altruistic node.

We first consider the case where there is no phase fading. In this example, we place nodes 1 and 3 (the transmitters) at coordinates $(0, 0.1), (0, -0.1)$, respectively, and nodes 2 and 4 (the receivers) at $(1, 0.2), (1, -0.2)$, respectively. The relay node is placed at position $(x, y)$. In the simulation, we let $\gamma = 3$. Figure 4.5 shows the region corresponding to the optimal location(s) for the relay node, i.e., if the relay node is located in this region full-cooperation becomes the only equilibrium of our non-cooperative game. Figure 4.6 further illustrates the gain offered by the altruistic node as a function of its position. As benchmarks, we use 1) the average transmission energy per bit, $E_{co}$, in the idealistic cooperative network where node 1 and 3 are not selfish and 2) the dumb relaying strategy where the altruistic node always helps the sender (hence, based on Lemma 23, no-cooperation is the only equilibrium among the transmitters.). We also define

$$\zeta_t = 10 \log_{10} \left\{ \frac{E_t(p_{nc})}{(E_1 + E_3)/2} \right\}, \zeta_{co} = 10 \log_{10} \left\{ \frac{E_{co}}{(E_1 + E_3)/2} \right\}$$
Figure 4.5: The region corresponding to the relay node positions where full-cooperation is the only equilibrium for the selfish nodes.

to quantify the gain resulting from cooperation in our non-cooperative game ($\zeta_t$) and the idealistic case ($\zeta_{co}$). In Figure 4.6, we show the gain by letting this node move at the x-axis, that is the position of the relay is $(x, 0)$. We can see that the introduction of the relay node reduces the energy consumption of the network at the equilibrium significantly (e.g., a gain as large as 7 dB). It is also shown that full-cooperation is the NE when the relay node is between $-0.2$ and $0.7$, since the curve of $\zeta_t$ coincides with that of $\zeta_{co}$.

Figure 4.7 shows the numerical result when the channel also experiences phase fading. In generating this figure, we place nodes 1 and 3 at coordinates $(0, 0.1), (0, -0.1)$, respectively, and nodes 2 and 4 (the receivers) at $(1, -0.3), (-1, 0.3)$, respectively. We can see that when the relay is between 0.2 and 0.5, full-cooperation is the NE. Compared with the
Figure 4.6: Comparison between relay strategies (dumb relaying refers to the case where the relay always helps).

no phase fading case, the gain is reduced, since we can only exploit the packet forwarding thus lose the beamforming gain. But, the existence of altruistic node with appropriate strategy still successfully stimulate the cooperation among the nodes.

4.3 Large Random Network

The detailed analysis of the small network scenario shows the effectiveness of the altruistic node in fostering cooperation among the wireless users. This motivates us to investigate whether the same conclusion is valid in a larger scale network or not. Toward this end, we consider ad-hoc networks with large numbers of nodes $N$. We consider the case where these $N$ nodes are randomly distributed on the surface of a sphere of area $A(N)$ according to a uniform distribution as shown in Figure 4.8. We keep the density $\rho$ of the
node as constant, and hence, as the number of nodes in the network $N$ increases, the area of the network $A(N) = N/\rho$ increases accordingly. This corresponds to the extended network model considered in the literature [23, 24]. Without loss of generality, we let $\rho = 1$ resulting in $A(N) = N$. As before, we let $X_i$ be the position of node $i$ and $d_{ij} = |X_i - X_j|$ be the distance between nodes $i$ and $j$. In our model, we allow all the nodes in the set $\mathcal{N}$ to be transmitters, that is $\mathcal{T} = \mathcal{N}$. It is easy to see that the design and analysis of the DF cooperation scheme, adopted in the small network scenario, will become intractable in this large network. Therefore, in the large network, we will limit ourselves to the simplest form of cooperation, namely packet forwarding\textsuperscript{12}. This choice can also be partly justified

\textsuperscript{12}Though sophisticated cooperation scheme may bring some extra gains, packet forwarding is shown [20] to be order optimal in large scale random networks under some specific conditions.

Figure 4.7: Comparison between relay strategies in the channel with phase fading.
by the result in section 4.2.3 where the packet forwarding is also shown to be effective in fostering cooperation. More specifically, the packets propagate in the network in a hop-by-hop fashion. We assume that, at each frame, each node randomly picks another node in the network as its destination and sends a packet to it. Also at the beginning of each transmission, the source node identifies a route to the destination. Let $\Gamma_{ij}$ be the set of nodes in the route from the source $i$ to the destination $j$. We use the routing scheme described in [14]: divide the whole area into small Voronoi cells\(^{13}\), whose size is properly chosen to guarantee that there is at least one node at each cell, and the packet hops from the source to the destination through the cells that have intersection with a line connecting the source and the destination. If a particular cell has multiple nodes, the relay request is assigned randomly to any of the nodes in this cell. Before sending a message to its destination $j$, source $i$ first broadcasts a relay request to the nodes on the route $\Gamma_{ij}$, then node $k$ on this route will decide whether to accept this request or not and sends back a response to the source node. We assume that the nodes always respond\(^{14}\) to the relay requests and are not

\(^{13}\)Given a set of points $a_1, \ldots, a_n$ in the surface of sphere, the Voronoi cell $V(a_i)$ is the set of all points which are closer to $a_i$ than to any of the other $a_j$’s [14].

\(^{14}\)For the sake of simplicity, we ignore the protocol overhead, which can be easily incorporated into our model without changing our main conclusions.
allowed to misbehave later. Furthermore, we assume that if a selfish node accepts a relay request, it will use a sufficient power level to ensure successful decoding only at the next hop on the route (this assumption is consistent with the node selfishness). Therefore, if any node on the route rejects the relay request, the route is interrupted. In this case, the source node will transmit the packet to the destination in one hop (i.e., direct transmission). Clearly, one can envision more sophisticated routing strategies where the source attempts to establish another route when the first route is interrupted. Here, we limit ourselves to the aforementioned routing strategy for the sake of simplicity, and analytical tractability. It also appears that the potential gains from more sophisticated routing strategies will be rather marginal due to the selfishness of the majority of nodes in our network (i.e., every selfish node will not be willing to waste energy in forwarding a packet for a long hop).

Finally, the source node is assumed to have no prior information about the identity of the nodes in the routes (whether a certain node is selfish or altruistic). On the other hand, every node knows the fraction of altruistic nodes present in the network and the functional form of the strategy employed by those nodes (i.e., $g(.)$).

To ensure successful decoding of the transmitted signal at a receiver $d_{ij}$ away from the transmitter, the energy spent on sending one packet (in the low rate regime) is given by $E_{tr}(d_{ij}) = c_1/h_{ij}^2 = c_1d_{ij}^\gamma$, where $c_1$ is a constant. Therefore, the required energy per hop scales polynomially, in the hop length, with order $\gamma$. It follows that, if all the nodes in the network are cooperative and accept all the relay requests, the optimal energy per packet scales as $\Theta\left(\sqrt{N} \log^{(\gamma-1)/2}(N)\right)$ as shown in [23, 24]. This optimal scaling is

The design of cheating-proof mechanisms is out of the scope of this work; relevant works could be found in [87, 88] and references therein.

In this thesis, we use the following Knuth’s asymptotic notations: 1) $f(N) = o(g(N))$ means $\forall c > 0, \exists N_0, \forall N > N_0, f(N) < cg(N)$, 2) $f(N) = \omega(g(N))$ means $\forall c > 0, \exists N_0, \forall N > N_0, g(N) < cf(N)$, 3) $f(n) = \Theta(g(N))$ means $\exists c_1, c_2 > 0, N_0, \forall N > N_0, c_1g(N) \leq f(n) \leq c_2g(N)$.
achieved by the routing scheme used here. On the other hand, if all the nodes are selfish then the result in Lemma 23 will hold, implying that no-cooperation will be the unique NE. At this equilibrium point, each source node must send the packet to the destination in one hop resulting in an average energy per packet which scales as $\Theta(N^{-\gamma/2})$. In the following, we show how altruistic nodes can be utilized to close the huge gap between those two scenarios.

We now recall that there are $\theta(N)$ altruistic nodes among these $N$ nodes, that is $|\mathcal{U}| = \theta(N)$, as shown in Figure 4.8. For simplicity, in the function $g(p_i) = p_i + (1 - p_i)p_{nc}$, we set the parameter $p_{nc}$ to be 0, that is $g(p_i) = p_i$. Let $\hat{\Psi}_i = \{ j : j \neq i, j \in \mathcal{N}, \text{s.t. } \exists k \in \Gamma_{ij}, k \in \mathcal{U} \}$, which means that $\hat{\Psi}_i$ is the set of possible destinations for node $i$ such that the route between it and node $i$ includes at least one altruistic node.

The probability that node $i$ can transmit its packet to the destination $j$ through the relays $\prod_{k \in \Gamma_{ij}} p_k$ (note that, if $k \in \mathcal{U}, p_k = p_i$). Let $i'(j)$ be the first node on the route $\Gamma_{ij}$, then with probability $\prod_{k \in \Gamma_{ij}} p_k$, the energy that node $i$ spends on sending this packet is $c_1 d_{ii'}(j)$. If the relay request is rejected, which happens with probability $1 - \prod_{k \in \Gamma_{ij}} p_k$, node $i$ has to transmit the packet directly to its destination $j$ with an energy expenditure of $c_1 d_{ij}$. Therefore, the expected energy that node $i$ spends on sending its own packet is

$$E_{i, tr} = \frac{1}{N - 1} \left\{ \sum_{j=1}^{N} \left[ \prod_{k \in \Gamma_{ij}} p_k c_1 d_{ii'}(j) + \left( 1 - \prod_{k \in \Gamma_{ij}} p_k \right) c_1 d_{ij} \right] \right\},$$

since node $i$ will choose its destination among other $N - 1$ nodes with equal probability.

Besides spending energy for sending its own packets, node $i$ also needs to relay packets for the other nodes. Let $\Lambda_i = \{ \Gamma_{kj} : i \in \Gamma_{kj}, k, j \in \mathcal{N}, k \neq i, j \neq i \}$ be the set of routes that include node $i$. $\Lambda_i$ is a random set, but at each TDMA super-frame which includes one slot assigned to each source node, $|\Lambda_i|$ is upper-bounded by $N$, since there are at most $N$
relay requests asking for node $i$ to help. Let $\Gamma \in \Lambda_i$ be one of the routes and $\Gamma(i)$ be node $i$’s direct next hop in the route $\Gamma$. We have $E_{i,re} = \mathbb{E}\left\{ \sum_{\Gamma \in \Lambda_i} p_i \prod_{k \in \Gamma} p_k c_1 d_{\Gamma(i)}^{\gamma(i)} \right\}$, here $\mathbb{E}\{\cdot\}$ means the expectation.

Hence the total cost for node $i$ is

$$C_i = \frac{1}{N - 1} \left\{ \sum_{j=1 \atop j \neq i}^{N} \left[ \prod_{k \in \Gamma_{ij}} p_k c_1 d_{ji}^{\gamma(j)} + \left( 1 - \prod_{k \in \Gamma_{ij}} p_k \right) c_1 d_{ij}^{\gamma(i)} \right] \right\} + \mathbb{E}\left\{ \sum_{\Gamma \in \Lambda_i} p_i \prod_{k \in \Gamma} p_k c_1 d_{\Gamma(i)}^{\gamma(i)} \right\}.$$

Now, we proceed to establish a sufficient condition on $\theta(N)$ that ensures full cooperation in the network. Since we consider random networks, our results hold in a probabilistic sense, i.e., we use the notion $w.h.p$ to mean that a certain result is true for any realization of the random network with probability which goes to 1 as the number of nodes in the network increases. First, we need to recall some preliminary results on the topology of random network.

**Lemma 25 (see [14])** Let $r(N)$ be the radius of a disk with area $100 \log N$ in the surface of a sphere. If we divide the total area $A(N) = N$ into Voronoi cells, each contains a disk of area $100 \log N$ and is contained in a disk with radius $2r(N)$, then there exists a sequence $\delta(N) \to 0$, such that $\text{Prob}(\text{every cell contains a node}) \geq 1 - \delta(N)$.

Based on the described routing scheme, the packet will hop from cell to the adjacent cell. Since the radius of each cell, needed to ensure the presence of at least one node, is at most $2r(N)$, the maximum distance of any hop is $8r(N)$. Hence, this lemma shows the existence of a constant $c_2 > 0$ such that the distance of any hop in the route is at most $c_2 \sqrt{\log N}$. That is, for any route $\Gamma_{ij}$, we have

$$d_{ii'(j)} \leq c_2 \sqrt{\log N}, \quad d_{i\Gamma(i)} \leq c_2 \sqrt{\log N}. \quad (4.13)$$

91
Next, we focus on the set of destination nodes which are far away from the source \(i\), since the transmission to these destinations requires large energy expenditure, especially if the node has to transmit the packet directly. Let \(D_i = \{j : j \in \mathcal{N}, d_{ij} \geq \sqrt{N}/4\}\), which is the set of nodes that are more than \(\sqrt{N}/4\) away from node \(i\). Let \(S\) be area of a disc in the sphere with radius \(\sqrt{N}/4\), and \(c_4 = (N - S)/N\). Then it is easy to prove that \(|D_i| = c_4 N + o(N)\) w.h.p. This says that the number of nodes that are at least \(\sqrt{N}/4\) far from the source \(i\) is \(c_4 N\), i.e., has the same order as the number of nodes in the network. Let \(\eta_{ij}\) be the number of hops that a packet has to travel from its source node \(i\) to a destination \(j \in D_i\) via the preferred route assuming that all the relay requests are accepted. Then w.h.p, we have

\[
\eta_{ij} \geq \frac{d_{ij}}{c_2 \sqrt{\log N}} \geq \frac{\sqrt{N}}{4c_2 \sqrt{\log N}},
\]

(4.14)

since the length of each hop is at most \(c_2 \sqrt{\log N}\) and \(d_{ij} \geq \sqrt{N}/4\).

Let \(\hat{D}_i = \{j : j \in D_i, \; \exists k \in \Gamma_{ij}, k \in U\}\) be the set of node destinations (for source node \(i\)), which are at least \(\sqrt{N}/4\) away from node \(i\) and have at least one altruistic node in the routes between them and node \(i\). The following result lower bounds the cardinality of the set \(\hat{D}_i\).

**Lemma 26** For any \(i\),

\[
|\hat{D}_i| \geq c_4 \left( N - N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2 \sqrt{\log N})} \right) + o \left( N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2 \sqrt{\log N})} \right),
\]

w.h.p. In particular, if \(\theta(N) = \omega(\sqrt{N \log N})\), then \(|\hat{D}_i| \geq c_5 N + o(N), \forall i, \) w.h.p.

**Proof:** Please refer to Appendix A.16.

Roughly speaking, Lemma 26 argues that if the number of altruistic nodes in the network scales as \(\omega(\sqrt{N \log N})\), then almost all the routes from source \(i\) to its destinations
that are more than $\sqrt{N}/4$ away from it, will contain at least one altruistic node as one of the hops. This fact allows the altruistic nodes to efficiently enforce the reward/punishment strategy by accepting the relay request from only the cooperative nodes. Based on this observation, the following result establishes an upper bound on the fraction of altruistic nodes needed to ensure full cooperation.

**Theorem 27** If $\theta(N) = \omega(\sqrt{N \log N})$, then full cooperation $p = 1$ is a Nash equilibrium. At this equilibrium, the average energy per packet approaches the optimal scaling law of $\Theta\left(\sqrt{N \log^{(\gamma-1)/2}(N)}\right)$.

**Proof:** Please refer to Appendix A.17. □

Finally, we observe that

$$\lim_{N \to \infty} \theta(N)/N = 0,$$

implying the need for only a vanishingly small fraction of the nodes to be altruistic in order to converge to a full cooperation equilibrium in our energy limited wireless network.

We note that full-cooperation may not be the unique NE. However, operating at this point will minimize the total energy expenditure of the network, and is also compatible with the selfish nature of the users in the network.

### 4.4 Summary

In this chapter, we adopted a game theoretic approach for analyzing the impact of user selfishness on the performance of energy limited ad-hoc network. Our results have established the critical role of altruistic, i.e., relay, nodes in stimulating full cooperation between all nodes. In the small network setting, our numerical results show that the introduction of one relay node, which employs the appropriate cooperation strategy, yields significant
energy savings. In the large network scenario, we have derived an upper bound on the number of altruistic nodes required to ensure full cooperation. This upper bound shows that full cooperation is possible in networks where only a vanishingly small fraction nodes is altruistic. Our results also shed light on the structure of optimal physical layer reward policies and established the strict sub-optimality of the traditional relaying approach, where the relay node offers the same forwarding service to all the nodes in the network.
CHAPTER 5

CONCLUSIONS

In this thesis, we have investigated several aspects of user-cooperation in multi-user wireless networks from network information theoretic perspective.

First, it was shown that noisy feedback can be exploited to construct efficient cooperation schemes. Utilizing the noisy feedback from the destination, we have proposed a novel relay strategy that combines both benefits of DF and CF. This noisy feedback strategy was shown to outperform the existing schemes for certain channel parameters. The noisy feedback cooperation strategy was then generalized to the multicast channel, where both receivers need to decode messages from the source. For the multicast channel, we differentiated the nodes based on their channel conditions. Contrary to the first thought, it was shown that the node with weaker channel condition should help the one with stronger channel first in order to get a larger achievable rate. Taking the side-information at the receivers into consideration, the cooperation scheme developed in the multicast channel was then modified to be a building block of the proposed cooperation scheme for the three-way channel, where every node is required to exchange information with the other nodes. Combined with an efficient scheduler, the proposed cooperation strategy for the three-way channel was shown to be asymptotically optimal when either one of the channel gains is large.
Second, the critical role of user-cooperation in secure communications was established. We have devised several cooperation strategies for the relay-eavesdropper channel to facilitate the transmission of confidential messages. The proposed strategies were shown to provide non-zero perfect secrecy rate, even when the perfect secrecy capacity is zero with the absence of user-cooperation. Unique features of user-cooperation in networks with secrecy constraints were illustrated with various examples. Specifically, the proposed NF scheme was shown to be able to help even it is totally ignorant with the message transmitted from the source, a situation impossible to occur in the classical relay channel. Further, the NF scheme was shown to increase the perfect secrecy rate for the reversely degraded relay channel, while the relay has been proved to provide no gain in the classical relay channel under this situation. Numerical results were also provided for the Gaussian channel to illustrate the analytical results.

Third, in an energy-limited wireless network, full-cooperation was shown to be a NE, provided that a vanishingly small fraction of the nodes are altruistic and employ proper strategies. This result partially justifies the widely-used but never validated assumption that users will spend their valuable energy to help other nodes in the network. The adopted game-theoretic framework assumes that the users are selfish at the very beginning, hence the results are robust to the nodes’ misbehavior, a concern raised by the more and more popular ad-hoc network applications. The traditional relay strategies ignoring the selfish nature of the nodes were also shown to be suboptimal. This result sheds lights on the important difference between the networks with/without central control, thus calls for special strategies to counter the possible selfish behaviors.

Some possible future research directions are as follows.
1. To provide a proper and better guidance for the development of modern wireless networks, the study of cooperation strategies in a larger scale network is crucial. With this respect, the generalization of the greedy cooperation principle would be an interesting direction.

2. Among the many open problems posed by our work on the relay-eavesdropper channel, how to close the gap between the achievable performance and the outer-bound is arguably the most important one. This problem is expected to be challenging since the capacity of the classical relay channel remains unknown. The investigation of the role of feedback in the relay-eavesdropper channel is another interesting problem. In the relay channel without security constraints, noiseless/noisy feedback was shown to be beneficial. On the other hand, in the presence of an eavesdropper, the role and optimal mechanism of feedback is not yet known, since the eavesdropper could also benefit from the feedback signal. Extending our work to a large scale network is expected to be of practical significance.

3. The study of security systems is not limited to the provision of confidentiality for the transmitted messages anymore. The success of security systems also depends on other security primitives and protocols, e.g., data authentication and bit-commitment schemes. But most, if not all, of the available results in this field are not satisfactory in that the results are based on the limited computation ability assumption of the attacker and the intractability of certain mathematical problems [89–102]. With the exponentially increasing computation power available to attacker and the fast developing algorithm design techniques, these assumptions will be obsolete. One striking
feature of these computational models is that transmissions are assumed to be noiseless [103–105]. But the transmission uncertainty (for example, noise and channel fading) is the inherit property of communications, and hence can be exploited to facilitate the provision of security, as evident in the wiretapper channel. Hence taking transmission uncertainty into consideration and deriving information theoretic results for other aspects of security research is a very exciting direction to follow.

4. The game theoretic framework is our first step to understand the behavior of network with selfish nodes. Taking more network parameters, such as channel fading, non-perfect scheduling and spectrum efficiency etc., into consideration are expected to have practical impact. Combining high-level protocols, such as multiple access control layer and application layer, with our physical layer study would be an interesting direction. The efforts in this direction are expected to provide some insightful principles for cross-layer design.
APPENDIX A

PROOF OF THEOREMS

A.1 Proof of Theorem 2

In this thesis, we refer to typical sequences as strong typical sequences (see [16, 19, 76] for details of strong typical sequences).

A.1.1 Discrete Memoryless Channel

Outline

Suppose we want to send i.i.d. source \( w(i), w(i) \in [1, M] \), in which \( M = 2^{NR} \) to the destination. Equally divide these \( 2^{NR} \) messages into \( M_1 = 2^{N\alpha t R_1} \) cells, index the cell number as \( b(i) \). Index the element in every cell as \( d(i) \), \( d(i) \in [1, M_2] \), \( M_2 = 2^{N(1-t)R_2} \). Thus

\[
2^{NR} = 2^{N\alpha t R_1} 2^{N(1-t)R_2}, \tag{A.1}
\]

that is

\[
R = \alpha t R_1 + (1 - t) R_2. \tag{A.2}
\]

The main idea is that the relay and the destination help each other to decode \( b(i) \):
• In the first state $m_1$, the source sends the cell index $b(i)$ to both the relay and the destination. At this time, neither the relay nor the destination can decode this information.

• In the feedback state $m_3$, the destination sends the compressed version of the received noisy signal to the destination. At the same time, the source sends additional information to the relay.

• At the end of the relay receive mode, the relay gets an estimation of $b(i)$, namely $\hat{b}(i)$. Thus in $m_2$, the relay sends its knowledge of $\hat{b}(i)$ to the destination to help it decode $b(i)$. At the same time, the source sends $d(i)$ to the destination.

Random code generation

Fix $p(x_1|m_1), p(x_1, x_3|m_3), p(x_1, x_2|m_2), p(\hat{y}_3)$.

• state $m_1$:

  At the source, generate $2^{\alpha tNR_1}$ i.i.d. length-$\alpha tN$ sequence $x_{1,m_1}$ each with probability $p(x_{1,m_1}) = \prod_{j=1}^{\alpha tN} p(x_{1j}|m_1)$. Label these sequences as $x_{1,m_1}(b)$, where $b \in [1, 2^{\alpha tNR_1}]$ is called the cell index.

• state $m_3$:

  – source node:

    Generate $2^{(1-\alpha)tNR_4}$ i.i.d. length-$(1-\alpha)tN$ codewords $x_{1,m_3}$ with $p(x_{1,m_3}) = \prod_{j=1}^{(1-\alpha)tN} p(x_{1j}|m_3)$. Label these sequences as $x_{1,m_3}(q), q \in [1, 2^{(1-\alpha)tNR_4}]$. Randomly partition the $2^{\alpha tNR_1}$ cell indices $\{b\}$ into $2^{(1-\alpha)tNR_4}$ bins $Q_q$ with $q \in [1, 2^{(1-\alpha)tNR_4}]$.  

100
– destination node:

Generate \(2^{(1-\alpha)tNR_3}\) i.i.d. length-\((1-\alpha)tN\) codewords \(x_{3,m_3}\) with \(p(x_{3,m_3}) = \prod_{j=1}^{(1-\alpha)tN} p(x_{3j}\mid m_3)\). Index them as \(x_{3,m_3}(u)\). Generate \(2^{\alpha tR}\) i.i.d. length-\(\alpha tN\) sequences \(\hat{y}_3(z) = \prod_{j=1}^{\alpha tN} p(\hat{y}_{3j})\). Randomly partition the set \(z \in [1, 2^{\alpha tR}]\) into bins \(U_u, u \in [1, 2^{(1-\alpha)NR_3}]\).

• state \(m_2\):

– relay node:

Randomly generate \(M_0 = 2^{(1-t)NR_0}\) i.i.d. length-\((1-t)N\) sequences \(x_{2,m_2}\) with \(p(x_{2,m_2}) = \prod_{j=1}^{(1-t)N} p(x_{2j}\mid m_2)\). Index them as \(x_{2,m_2}(c), c \in [1, 2^{(1-t)NR_0}]\). Randomly partition the \(2^{\alpha tR_1}\) cell indices into \(2^{(1-t)NR_0}\) bins \(C_c\).

– source node:

Generate \(M_2 = 2^{(1-t)NR_2}\) i.i.d. length-\((1-t)N\) sequences \(x_{1,m_2}\) with \(p(x_{1,m_2}) = \prod_{j=1}^{(1-t)N} p(x_{1j}\mid x_{2j}, m_2)\) for every \(x_{2,m_2}\) sequence. Index them \(x_{1,m_2}(d\mid c), d \in [1, 2^{(1-t)NR_2}]\).

**Encoding**

Partition the source message set into \(2^{\alpha tR_1}\) equal-sized cells. Let \(w(i)\) be the message to be sent in block \(i\). Suppose \(w(i)\) is the \(d(i)\)-th message in cell-\(b(i)\) and the cell index \(b(i)\) is in bin-\(q(i)\) and bin-\(c(i)\) respectively. For brevity we drop the block index \(i\) in the following.

• state \(m_1\):

The source sends \(x_{1,m_1}(b)\).

• state \(m_3\):

101
– The source node knows that the cell index $b$ is in bin-$q$, so it sends $x_{1,m3}(q)$.

– The destination first selects $\hat{y}_3(z)$ that is jointly typical with $y_{3,m1}$. It then sends $x_{3,m3}(u)$ where $z$ is in the bin $U_u$.

• state $m_2$:

– Knowing the cell index $b$ is in bin-$c$, the source node sends the corresponding $x_{1,m3}(d|c)$.

– Using the information received in state $m_1$ and $m_3$, the relay gets an estimation of the cell index $\hat{b}$. Suppose $\hat{b}$ is in bin-$\hat{c}$. Then it sends $x_{2,m3}(\hat{c})$.

**Decoding**

In the following, code length $N$ is chosen sufficiently large.

• at the end of $m_1$:

  The destination has received $y_{3,m1}$ and it decides a sequence $\hat{y}_3(z)$ if $(\hat{y}_3(z), y_{3,m1})$ are jointly typical. There exists such a $z$ with high probability if

  $$\hat{R} \geq I(\hat{Y}_3; Y_3).$$

  (A.3)

• at the end of $m_3$:

  At this stage, only the relay decodes the message.

  – The relay estimates $u$ by looking for the unique $\hat{u}$ such that $(x_{3,m3}(\hat{u}), y_{2,m3})$ are jointly typical. $\hat{u} = u$ with high probability if

    $$R_3 \leq I(X_3; Y_2).$$

    (A.4)
– Knowing \( \hat{u} \), the relay tries to decode \( q \) by selecting the unique \( \hat{q} \) such that
\[
(x_{1,m_3}(\hat{q}), x_{3,m_3}(\hat{u}), y_{2,m_3})
\]
are jointly typical. \( \hat{q} = q \) with high probability if
\[
R_4 \leq I(X_1;Y_2|X_3). \tag{A.5}
\]

– The relay calculates a list \( \ell(y_{2,m_1}) \) such that \( z \in \ell(y_{2,m_1}) \) if \( (\hat{y}_{3,m_1}(z), y_{2,m_1}) \) are jointly typical. Assuming \( u \) decoded successfully at the relay, \( \hat{z} \) is selected if it is the unique \( \hat{z} \in U_u \cap \ell(y_{2,m_1}) \).

Using the same argument as in [19], it can be shown that \( \hat{z} = z \) occurs with high probability if
\[
\alpha t \hat{R} \leq \alpha t I(\hat{Y}_3;Y_2|m_1) + (1 - \alpha) t R_3. \tag{A.6}
\]

– The relay computes another list \( \ell(y_{2,m_1}, \hat{y}_{3,m_1}) \) such that \( b \in \ell(y_{2,m_1}, \hat{y}_{3,m_1}) \) if
\[
(x_{1,m_1}(b), y_{2,m_1}, \hat{y}_{3,m_1})
\]
are jointly typical.

– Finally, the relay declares \( \hat{b} \) is received if it is the unique \( \hat{b} \in Q_q \cap \ell(y_{2,m_1}, \hat{y}_{3,m_1}) \).

Using the same argument as in [19], one can show \( \hat{b} = b \) with high probability if
\[
\alpha t R_1 \leq \alpha t I(X_1;\hat{Y}_3,Y_2|m_1) + (1 - \alpha) t I(X_1;Y_2|X_3, m_3). \tag{A.7}
\]

• at the end of \( m_2 \):

– The destination declares that \( \hat{c} \) was sent from the relay if there exists one and only one \( \hat{c} \) such that \( (x_{2,m_2}(\hat{c}), y_{3,m_2}) \) are jointly typical. Then \( \hat{c} = c \) with high probability if
\[
R_0 \leq I(X_2;Y_3|m_2). \tag{A.8}
\]
After decoding \( \hat{c} \), the destination further declares that \( \hat{d} \) was sent from the source if it is the unique \( \hat{d} \) such that \((x_{1,m_{2}}(\hat{d}), x_{2,m_{2}}(\hat{c}), y_{3,m_{2}})\) are joint typical. Assuming \( c \) decoded correctly, the probability of error of \( \hat{d} \) is small if

\[
R_{2} \leq I(X_{1}; Y_{3}|X_{2}, m_{2}).
\]  

At first, the destination calculates a list \( \ell(y_{3,m_{1}}) \), such that \( b \in \ell(y_{3,m_{1}}) \) if \((x_{1,m_{1}}(b), y_{3,m_{1}})\) are jointly typical. Assuming \( c \) decoded successfully at the destination, \( \hat{b} \) is declared to be the cell index if there is a unique \( \hat{b} \in C_{c} \cap \ell(y_{3,m_{1}}) \). As in [19], the decoding error is small if

\[
\alpha t R_{1} \leq \alpha t I(X_{1}; Y_{3}|m_{1}) + (1 - t) R_{0} \leq \alpha t I(X_{1}; Y_{3}|m_{1}) + (1 - t) I(X_{2}; Y_{3}|m_{2}).
\]

(A.10)

From the cell index \( \hat{b} \) and the message index \( \hat{d} \) within the cell, the destination can recover the source message.

Combining (A.7) and (A.9), we have

\[
R < \alpha t I(X_{1}; \hat{Y}_{3}, Y_{2}|m_{1}) + (1 - \alpha) t I(X_{1}; Y_{2}|X_{3}, m_{3}) + (1 - t) I(X_{1}; Y_{3}|X_{2}, m_{2}).
\]

(A.11)

It follows from (A.10) and (A.9) that

\[
R < \alpha t I(X_{1}; Y_{3}|m_{1}) + (1 - t) I(X_{1}, X_{2}; Y_{3}|m_{2}).
\]

(A.12)

From (A.3) and (A.6), we have the constraint

\[
(1 - \alpha) I(X_{3}; Y_{2}) > \alpha I(\hat{Y}_{3}; Y_{3}|Y_{2}, m_{1}).
\]

(A.13)

Thus if (A.11), (A.12), and (A.13) are satisfied, there exist a channel code that makes the decoding error at destination less than \( \epsilon \).
A.2 Proof of Lemma 3

As mentioned in [16], strong typicality does not apply to continuous random variables in general, but for Gaussian input distributions, one can generalize the Markov lemma along the lines of [106, 107] and thereby the DMC result derived above applies to the Gaussian inputs \((X_1, X_2, X_3, \hat{Y})\). Since \(\hat{Y}_3\) is a degraded version of \(Y_3\), we write \(\hat{Y}_3 = Y_3 + Z'\) where 

\[Z'\text{ is Gaussian noise with variance } \sigma_3^2\text{ (see [62, 108] for a similar analysis).}\]

First, we examine the constraint (A.13) under the Gaussian inputs,

\[
I(X_3; Y_2|m_3) = h(Y_2|m_1) - h(Y_2|X_3, m_1) = \frac{1}{2} \log \left( \frac{1}{1 - \rho_{Y_2,X_3}^2} \right). \tag{A.14}
\]

And

\[
\rho_{Y_2,X_3}^2 = \frac{E^2\{(h_{12}X_1 + h_{23}X_3 + Z_2)X_3\}}{Var(Y_2)Var(X_3)} = \frac{\left( h_{12}r_{13}\sqrt{P_1^{(3)}} + h_{23}\sqrt{P_3^{(3)}} \right)^2}{h_{12}^2P_1^{(3)} + h_{23}^2P_3^{(3)} + \sigma^2 + 2h_{12}h_{23}r_{13}\sqrt{P_1^{(3)}P_3^{(3)}}}. \tag{A.15}
\]

So

\[
I(X_3; Y_2|m_3) = \frac{1}{2} \log \left( \frac{h_{12}^2P_1^{(3)} + h_{23}^2P_3^{(3)} + \sigma^2 + 2h_{12}h_{23}r_{13}\sqrt{P_1^{(3)}P_3^{(3)}}}{h_{12}^2(1 - r_{13}^2)P_1^{(3)} + \sigma^2} \right). \tag{A.16}
\]

We observe that the correlation coefficient \(r_{13} = 0\) because neither the source nor the destination knows the codeword sent by the other during the feedback state. Thus, one has

\[
I(X_3; Y_2|m_3) = \frac{1}{2} \log \left( 1 + \frac{h_{23}^2P_3^{(3)}}{h_{12}^2P_1^{(3)} + \sigma^2} \right) = C\left( \frac{h_{23}^2P_3^{(3)}}{h_{12}^2P_1^{(3)} + \sigma^2} \right). \tag{A.17}
\]
Similarly, one has

\[ I(\hat{Y}_3; Y_3|Y_2, m_1) = h(\hat{Y}_3|m_1) - \frac{1}{2} \log \left( \frac{1}{1 - \rho_{Y_3Y_2}^2} \right) - h(Y_3|m_1) \]

\[ = h(h_{13}X_1 + Z_3 + Z') - h(h_{13}X_1 + Z_3 + Z'|h_{13}X_1 + Z_3) \]

\[ - \frac{1}{2} \log \left( \frac{1}{1 - \rho_{Y_3Y_2}^2} \right) \]  

(\(A.18\))

\[ = \frac{1}{2} \log \left( \frac{h_1^2P_1^{(1)} + \sigma^2 + \sigma_3^2}{\sigma_3^2} \right) - \frac{1}{2} \log \left( \frac{1}{1 - \rho_{Y_3Y_2}^2} \right) \]

where

\[ \rho_{Y_3Y_2}^2 = \frac{E^2(\hat{Y}_3Y_2)}{\text{Var}(\hat{Y}_3)\text{Var}(Y_2)} \]

\[ = \frac{E^2\{(h_{13}X_1 + Z_3 + Z')(h_{12}X_1 + Z_2)\}}{\text{Var}(\hat{Y}_3)\text{Var}(Y_2)} \]

\[ = \frac{(h_{12}h_{13}P_1^{(1)})^2}{(h_{13}P_1^{(1)} + \sigma^2 + \sigma_3^2)(h_{12}P_1^{(1)} + \sigma^2)}. \]  

(A.19)

So

\[ I(\hat{Y}_3; Y_3|Y_2, m_1) = \frac{1}{2} \log \left( 1 + \frac{\sigma^2}{\sigma_3^2} + \frac{1}{\sigma_3^2} \left( \frac{\sigma^2h_{13}P_1^{(1)} + \sigma_3^2}{h_{12}P_1^{(1)} + \sigma^2} \right) \right). \]  

(A.20)

Setting

\[ (1 - \alpha)tI(X_3; Y_2) = \alpha tI(\hat{Y}_3; Y_3|m_1) - \alpha tI(\hat{Y}_3; Y_2|m_1) \]  

(A.21)

to solve for \(\sigma_3^2\)

\[ \sigma_3^2 = \frac{\sigma^2 + \frac{\sigma^2h_{13}P_1^{(1)} + \sigma_3^2}{h_{12}P_1^{(1)} + \sigma^2}}{(1 + \frac{h_{23}P_3^{(3)}}{h_{12}P_1^{(3)} + \sigma^2})\frac{1-\alpha}{\alpha} - 1}. \]  

(A.22)

Next, we examine the achievable rate expression (A.11).

\[ I(X_1; \hat{Y}_3, Y_2|m_1) = \frac{1}{2} \log \left( 1 + \frac{h_{12}^2P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2P_1^{(1)}}{\sigma^2 + \sigma_3^2} \right) = C\left( \frac{h_{12}^2P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2P_1^{(1)}}{\sigma^2 + \sigma_3^2} \right), \]

(I\(X_1; Y_2|X_3, m_3\) = \(h(Y_2|X_3, m_3) - h(Y_2|X_1, X_3, m_3)\)

\[ = \frac{1}{2} \log \left( 1 + \frac{h_{12}^2P_1^{(3)}}{\sigma^2} \right) = C\left( \frac{h_{12}^2P_1^{(3)}}{\sigma^2} \right), \]  

(A.23)

\[ I(X_1; Y_3|X_2, m_2) = h(Y_3|X_2, m_2) - h(Y_3|X_1, X_2, m_2) \]

\[ = \frac{1}{2} \log \left( 1 + \frac{(1 - r_{12}^2)h_{13}^2P_1^{(2)}}{\sigma^2} \right) = C\left( \frac{(1 - r_{12}^2)h_{13}^2P_1^{(2)}}{\sigma^2} \right). \]
Combining them together, we get

\[
\alpha tI(X_1; Y_3, Y_2|m_1) + (1 - \alpha)tI(X_1; Y_2|X_3, m_3) + (1 - t)I(X_1; Y_3|X_2, m_2) \\
= \alpha tC\left(\frac{h_{12}^2 P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2 + \sigma_3^2}\right) + (1 - \alpha)tC\left(\frac{h_{12}^2 P_1^{(3)}}{\sigma^2}\right) + (1 - t)C\left(\frac{(1 - r_{12}^2)h_{13}^2 P_1^{(2)}}{\sigma^2}\right).
\]

(A.24)

Similarly for (A.12), one has

\[
I(X_1; Y_3|m_1) = h(h_{13} X_1 + Z_3|m_1) - h(h_{13} X_1 + Z_3|X_1, m_1) \\
= \frac{1}{2} \log \left(1 + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) = C\left(\frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right),
\]

\[
I(X_1, X_2; Y_3|m_2) = h(h_{13} X_1 + h_{23} X_2 + Z_3|m_2) - h(h_{12} X_1 + h_{23} X_2 + Z_3|X_1, X_2, m_2) \\
= \frac{1}{2} \log \left(1 + \frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{23} r_{12} \sqrt{P_1^{(2)} P_2^{(2)}}}{\sigma^2}\right) \\
= C\left(\frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{23} r_{12} \sqrt{P_1^{(2)} P_2^{(2)}}}{\sigma^2}\right),
\]

(A.25)

which gives rise to

\[
\alpha tI(X_1; Y_3|m_1) + (1 - t)I(X_1, X_2; Y_3|m_2) \\
= \frac{\alpha t}{2} \log \left(1 + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) + \frac{1 - t}{2} \log \left(1 + \frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{23} r_{12} \sqrt{P_1^{(2)} P_2^{(2)}}}{\sigma^2}\right) \\
= \frac{\alpha t}{2} C\left(\frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) + (1 - t)C\left(\frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{23} r_{12} \sqrt{P_1^{(2)} P_2^{(2)}}}{\sigma^2}\right).
\]

(A.26)

Setting the noise variance \(\sigma^2 = 1\), the proof is complete.
We only show the upperbound of $R_{FB}$. The proof for $R_{CF}$ is similar and thus omitted.

Setting shorthand notation $\Delta = 1 + \frac{h_{13}^2 P_1^{(3)}}{h_{12}^2 P_1^{(3)} + 1}$, one has from (2.15) that
\[
1 + \frac{h_{13}^2 P_1^{(1)}}{1 + \sigma_3^2} + h_{12}^2 P_1^{(1)} = \frac{h_{13}^2 P_1^{(1)}}{1 + \frac{h_{12}^2 P_1^{(1)}}{(h_{12}^2 P_1^{(1)} + 1)(\Delta - 1) - 1}} + (1 + h_{12}^2 P_1^{(1)})
\]
\[
= \frac{h_{13}^2 P_1^{(1)}(h_{12}^2 P_1^{(1)} + 1)(\Delta - 1) - 1}{\Delta} + (1 + h_{12}^2 P_1^{(1)})
\]
\[
\leq (1 + h_{12}^2 P_1^{(1)})\Delta - \frac{1}{\alpha}.
\]

Hence,
\[
\alpha t C\left(\frac{h_{13}^2}{1 + \sigma_3^2} + h_{12}^2\right) + (1 - \alpha) t C\left(\frac{h_{13}^2 P_1^{(3)}}{h_{12}^2 P_1^{(3)} + 1}\right)
\]
\[
\leq \alpha t C\left(h_{13}^2 P_1^{(1)}\right) + (1 - \alpha) t C\left(h_{23}^2 P_3^{(3)} + h_{13}^2 P_1^{(3)}\right)
\]
\[
= \alpha t C\left(h_{13}^2 P_1^{(1)}\right) + (1 - \alpha) t C\left(h_{13}^2 P_1^{(3)} + h_{23}^2 P_3^{(3)}\right),
\]
which proves (2.18).

### A.4 Proof of Theorem 5

In view of $R_{DF}$ in (2.8), $h_{12}^2 \leq h_{13}^2$ implies that
\[
R_{DF} \leq t C\left(h_{12}^2 P_1^{(1)}\right) + (1 - t) C\left((1 - r_{12}^2)h_{13}^2 P_1^{(2)}\right) \leq C\left(h_{13}^2 P\right) = R_{ro}
\]
(A.29)

where we have used the total power constraint (2.17). To prove 2), consider the upperbound for $R_{CF}$ in (2.19). Given the total power constraint $P_1^{(2)} + P_2^{(2)} \leq P$, it is easy to verify that $h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} \leq \max\{h_{13}^2, h_{23}^2\} P$. Therefore, the condition $h_{23}^2 \leq h_{13}^2$ implies that
\[
R_{CF} \leq t C\left(h_{13}^2 P_1^{(1)}\right) + (1 - t) C\left(h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)}\right) \leq C\left(h_{13}^2 P\right) = R_{ro}.
\]
(A.30)
The last statement of the theorem can be shown in a similar fashion using the $R_{FB}$ upper-bound in (2.18).
A.5 Proof of Theorem 7

Since \( C(h_{12}^2 P) \geq C(h_{13}^2 P) \) by \( h_{12}^2 \geq h_{13}^2 \), the two line segments in the \( R_{DF} \) expression intersect at some optimal \( t^* \in (0, 1) \) (see Fig. 2.3). The corresponding rate is given by

\[
R_{DF} = \frac{C\left( f_1(\theta, r_{12}, h_{13}, h_{23})P \right) C\left( h_{12}^2 P \right) - C\left( f_2(\theta, r_{12}, h_{13})P \right) C\left( h_{13}^2 P \right)}{C\left( f_1(\theta, r_{12}, h_{13}, h_{23})P \right) + C\left( h_{12}^2 P \right) - C\left( f_2(\theta, r_{12}, h_{13})P \right) - C\left( h_{13}^2 P \right)},
\]

where we have set \( P^{(2)}_1 = P \cos^2 \theta \) and \( P^{(2)}_2 = P \sin^2 \theta \) according to the total power constraint. Taking \( P \to 0 \), the Taylor expansion is sufficient to establish (2.21). To prove the lowerbound in (2.21), note that \( f_1(\theta, r_{12}, h_{13}, h_{23}) \leq (h_{12}^2 + h_{23}^2) \) with equality when \( r_{12} = 1 \) and \( \tan(\theta) = \frac{h_{23}}{h_{13}} \), which, together with \( f_2(\theta, r_{12}, h_{13}) \leq h_{13}^2 \), also proves the upperbound of \( S_{DF} \) in (2.21).

On the other hand, as \( P \to 0 \), it is seen from (2.9) that \( \sigma_2^2 \to \infty \), thus showing \( S_{CF} \leq h_{13}^2 \). The similar behavior holds for the feedback scheme, that is, \( \sigma_3^2 \to \infty \) as \( P \to 0 \), in which case \( S_{FB} \leq S_{DF} \) with the optimal \( \alpha \) approaches 1.

A.6 Proof of Theorem 8

The results for \( G_{DF} \) and \( G_{CF} \) follows from direct computation of large \( P \) limit. We only show the last statement concerning the feedback scheme. As in the case of decode-forward, the line-crossing point gives the optimal \( t \) and the associated rate \( R_{FB} \) is given by

\[
R_{FB} = \frac{C\left( f_1(\theta, r_{12}, h_{13}, h_{23})P \right) - C\left( f_2(\theta, r_{12}, h_{13})P \right) \alpha C\left( h_{13}^2 P \right)}{C\left( f_1(\theta, r_{12}, h_{13}, h_{23})P \right) + A - C\left( f_2(\theta, r_{12}, h_{13})P \right) - \alpha C\left( h_{13}^2 P \right)},
\]

in which

\[
A = \alpha C\left( \left( \frac{h_{13}^2}{1 + \sigma_3^2} + h_{12}^2 \right) P \right) + (1 - \alpha) C\left( h_{12}^2 P \cos^2 \psi \right),
\]

109
where we set \( P_1^{(3)} = P \cos^2 \psi \) and \( P_3^{(3)} = P \sin^2 \psi \). Taking \( P \to \infty \),
\[
\sigma_3^2 \to \frac{h_{12}^2 + h_{13}^2}{h_{12}^2 \left[ \left(1 + \frac{h_{23}^2}{h_{12}^2} \tan^2 \psi \right)^{\frac{1-\alpha}{\alpha}} - 1 \right]} \quad (= \sigma_3^2(\infty)). \tag{A.34}
\]
Denote \( f_3(\psi, \alpha, h_{13}, h_{12}, h_{23}) = \alpha \log \left( \frac{h_{12}^2}{1 + \sigma_3^2(\infty)} + h_{12}^2 \right) + (1 - \alpha) \log h_{12}^2 \cos^2 \psi \), one has
\[
R_{FB} \sim \frac{1}{2} \log P + \frac{1}{4} (1 - \alpha) \log f_2 \cdot \log P + \frac{1}{4} \left[ \log f_1 \cdot \log f_3 - \alpha \log f_2 \cdot \log h_{12}^2 \right] \tag{A.35}
\]
It follows that if \( \alpha < 1 \)
\[
G_{FB} \leq \log f_2 \leq \log h_{13}^2 \quad \text{(relay-off)}, \tag{A.36}
\]
which forces \( \alpha = 1 \), that is, \( G_{FB} = G_{DF} \).

### A.7 Proof of Theorem 10

Here we only prove the part 1) of this theorem. Parts 2) - 5) follow the same lines as the corresponding results in the relay case.

To prove part 1), it suffices to show the statement for \( \alpha = 1 \). The capacity of the multicast channel without cooperation is given by \( R_{\text{non-coop}} = C(\min\{h_{13}^2, h_{12}^2\} P) \). With the assumption that \( h_{12}^2 > h_{13}^2 \), we have \( R_{\text{non-coop}} = C(h_{13}^2 P) \).

Note that the rate expression of (2.30) admits the same line-crossing interpretation as in the relay case. Thus, the intersection determines the optimal rate point. Equate the two terms
\[
tC(h_{12}^2 P_1^{(1)}) = tC(h_{13}^2 P_1^{(1)}) + (1 - t)C((h_{13}^2 + h_{23}^2) P) \tag{A.37}
\]
to solve
\[
t^* = \frac{C((h_{13}^2 + h_{23}^2) P)}{C((h_{13}^2 + h_{23}^2) P) + C(h_{12}^2 P) - C(h_{13}^2 P)}, \tag{A.38}
\]

110
which gives the corresponding rate

\[
R_{DF} = \frac{C((h_{12}^2 + h_{23}^2)P)C(h_{12}^2P)}{C((h_{13}^2 + h_{23}^2)P) + C(h_{12}^2P) - C(h_{13}^2P))}.
\] (A.39)

Therefore, using \(h_{12}^2 > h_{13}^2\), one has

\[
R_{DF} - R_{non-coop} = (1 - t^*)(C((h_{13}^2 + h_{23}^2)P) - C(h_{13}^2P)) > 0,
\] (A.40)

which proves the theorem.

### A.8 Proof of Lemma 12

Here we first prove the result for DMC case, then apply the result to the Gaussian channel.

#### A.8.1 Source coding

Randomly bin all the sequence \(S^K_1\) into \(2^{K(H(S_1|S_2)+\epsilon)}\) bins by independently generating an index \(w\) uniformly distributed on \(\{1, 2, ..., 2^{K(H(S_1|S_2)+\epsilon)}\}\). Let \(f_{s_1}\) be the mapping function, such that \(w = f_{s_1}(s^K_1)\). Independently generate another bin index \(b\) for every sequence \(S^K_1\) by picking \(b\) uniformly from \(\{1, 2, ..., 2^{KR}\}\). Let \(B(b)\) be the set of all sequences \(S^K_1\) allocated to bin \(b\). Thus, every source sequence is associated with two bin indexes \(\{w, b\}\).

#### A.8.2 Channel coding

Random code generation

- At state \(m_1\), generate \(2^{K(H(S_1|S_2)+\epsilon)}\) i.i.d. length-\(N_1\) sequence \(x_{1,m_1}\), each with probability \(p(x_{1,m_1}) = \prod_{j=1}^{N_1} p(x_{1j}|m_1)\), in which \(p(x_{1j}|m_1)\) is the input distribution that maximizes \(I(X_1;Y_2)\). Assign every bin index \(w\) to one sequence \(x_{1,m_1}(w)\), \(w \in [1, 2^{K(H(S_1|S_2)+\epsilon)}]\).
• At state $m_2$, randomly generate $2^{KR}$ i.i.d. length-$N_0 x_{2,m_2}$ at node-2, each with probability $p(x_{2,m_2}) = \prod_{j=1}^{N_0} p(x_{2j|m_2})$. Generate $2^{KR}$ i.i.d. length-$N_0 x_{1,m_2}$ at node-1, each with probability $p(x_{1,m_2}) = \prod_{j=1}^{N_0} p(x_{1j|m_2})$, in which $p(x_{1j|m_2}) = \sum_{x_2} p(x_{1j|x_2})$. And $p(x_{1j|x_2})$ is the input distribution that maximizes $I(X_1, X_2; Y_3)$. Associate every bin index $b$ to one sequence pair $\{x_{1,m_2}(b), x_{2,m_2}(b)\}$.

Coding

Suppose we want to send source sequence $s^K_1(i)$ at block $i$, and $w(i) = f_{s_1}(s^K_1(i))$, $s^K_1(i) \in B(b(i))$. For brevity of notation, we drop block index $i$ in the following.

• State $m_1$:

  Node-1 sends $x_{1,m_1}(w)$.

• State $m_2$:

  - Node-1 knows $s^K_1$ is in $b$, so it sends $x_{1,m_2}(b)$.

  - At the end of state $m_1$, node-2 gets an estimation of $s^K_2$ (details will be given in the following), and suppose $s^K_2$ is in bin $\hat{b}$. Then in state $m_2$ nodes-2 sends the corresponding $x_{2,m_2}(\hat{b})$.

A.8.3 Decoding

At the end of state $m_1$:

• At node-2:

  At first, node-2 looks for the one and only one $\hat{w}$ such that $\{x_{1,m_1}(\hat{w}), y_{2,m_1}\}$ are jointly typical. Then node-2 searches in the bin indexed by $\hat{w}$ for source sequence $s^K_2$ such that $\{s^K_2, \hat{w}\}$ are jointly typical. If it finds only one such sequence, it declares it has received $s^K_2$, otherwise declares an error.
At node-3:
Node-3 calculates a list $\ell(y_{3,m_1})$, such that $w' \in \ell(y_{3,m_1})$ if \{x_{1,m_1}(w'), y_{3,m_1}\} are jointly typical.

At the end of state $m_2$, only node-3 needs to decode:

- Step 1:
  node-3 declares it receives $\hat{b}$, if $\hat{b}$ is the one and only one index such that
  \[
  \{x_{1,m_2}(\hat{b}), x_{2,m_2}(\hat{b}), y_{3,m_2}\}
  \]
  are jointly typical.

- Step 2:
  node-3 searches in the bin $B(\hat{b})$ for the one and only one source sequence $\hat{s}_{31}^K$, such that \{\hat{s}_{31}^K, s_{31}^K\} are jointly typical and $f_{s_1}(\hat{s}_{31}^K) \in \ell(y_{3,m_1})$. If it finds such a unique one, it declares that $\hat{s}_{31}^K$ is the source sequence. Otherwise it declares an error.

### A.8.4 Calculation of Probability of Error

**Node-2**

For node-2 there are following error events:

\[
E_0 = \{(s_1^K, s_2^K) \notin A_e^K\},
\]

\[
E_1 = \{\hat{w} \neq w\},
\]

\[
E_2 = \{\exists s_1': s_1' \neq s_1^K, f_{s_1}(s_1') = w \text{ and } (s_1', s_2^K) \in A_e^K\}.
\]

And

\[
P_{e^{N_1,K}} = P(E_0 \cup E_1 \cup E_2) \leq P(E_0) + P(E_1|E_0^c) + P(E_2|E_0^c, E_1^c).
\]
When $K$ is sufficiently large, using the AEP, $P(E_0) \to 0$. Now consider $P(E_1|E_0^c)$, if channel code rate is less than the capacity, receiver will decode channel code with error probability less than $\epsilon$. Here, there are $2^{K(H(S_1|S_2)+\epsilon)}$ code words, and channel code length is $N_1$, then the rate of channel code is $\frac{K(H(S_1|S_2)+\epsilon)}{N_1}$. Thus for sufficiently large $N_1$ and $K$, $P(E_1|E_0^c) \leq \epsilon$ if

$$
\frac{K(H(S_1|S_2)+\epsilon)}{N_1} < \max_{p(x_1)} I(X_1;Y_2|m_1) = C_2,
$$

which is the same as:

$$
H(S_1|S_2) + \epsilon < \tau_1 C_2.
$$

Because source code rate is $H(S_1|S_2) + \epsilon$, using the same argument as [76], one can get $P(E_2|E_0^c, E_1^c) < \epsilon$, if $K$ is sufficiently large. So if (A.46) is satisfied, and $N_1, K$ are sufficiently large, there exists a source-channel code that make the error probability at node-2

$$
P_{e^{N_1,K}} = P(\hat{s}_{21}^K \neq s_1^K) \leq 3\epsilon.
$$

**Node-3**

For node-3, there are following error events:

$E_0 = \{(s_1^K, s_3^K) \notin A^K_\epsilon\}$,

$E_1 = \{\text{node 2 can not decode successfully}\}$,

$E_2 = \{\hat{b} \neq b\}$,

$E_3 = \{\exists s_1'^K : s_1'^K \neq s_1^K, f_{s_1}(s_1'^K) \in f(y_3|m_1), s_1'^K \in B(\hat{b}), (s_1'^K, s_3^K) \in A^K_\epsilon\}$.

$P_{e^{N,K}} = P(E_0 \cup E_1 \cup E_2 \cup E_3) \leq P(E_0) + P(E_1) + P(E_2|E_0^c, E_1^c) + P(E_3|E_0^c, E_1^c, E_2^c)$.

(A.48)
When $K$ is sufficiently large, $P(E_0) \to 0$. And if (A.46) is satisfied, $P(E_1) \leq 3\epsilon$. Now consider $P(E_2|E_0^c, E_1^c)$, the channel code rate is $\frac{KR}{N_0}$. So, $P(E_2|E_0^c, E_1^c) \leq \epsilon$ for sufficiently large $N_0$, if

$$\frac{KR}{N_0} \leq \max_{p(x_1,x_2)} I(X_1, X_2; Y_3) = C_{(1,2)-3}, \quad \text{(A.49)}$$

that is,

$$R \leq \tau_0 C_{(1,2)-3}. \quad \text{(A.50)}$$

Now consider $P(E_3|E_0^c, E_1^c, E_2^c)$:

$$P(E_3|E_0^c, E_1^c, E_2^c)$$

$$= P(\exists s_1'K : s_1'K \neq S_1^K, f_{s_1}(s_1'K) \in \ell(Y_3|m_1), s_1'K \in B(b), (s_1'K, S_3^K) \in A_e^K)$$

$$= \sum_{(s_1^K, s_3^K)} p(s_1^K, s_3^K) P(\exists s_1'K \neq s_1^K, f_{s_1}(s_1'K) \in \ell(Y_3|m_1), s_1'K \in B(b), (s_1'K, s_3^K) \in A_e^K)$$

$$\leq \sum_{(s_1^K, s_3^K)} P(f_{s_1}(s_1'K) \in \ell(Y_3|m_1), s_1'K \in B(b))$$

$$= \sum_{(s_1^K, s_3^K)} p(s_1^K, s_3^K) \sum_{s_1'K \neq s_1^K} P(f_{s_1}(s_1'K) \in \ell(Y_3|m_1)) P(s_1'K \in B(b))$$

$$\leq \sum_{(s_1^K, s_3^K)} p(s_1^K, s_3^K) 2^{-K(H(S_1|S_2)+\epsilon)} ||\ell(Y_3|m_1)|| 2^{-KR||A_e(S_1^K|s_3^K)||}$$

$$\leq 2^{-K(H(S_1|S_2)+\epsilon)} E\{||\ell(Y_3|m_1)||\} 2^{-KR} 2^K(H(S_1|S_3)+\epsilon).$$

Follow the same steps in the [19], one has

$$E\{||\ell(Y_3|m_1)||\} \leq 1 + 2^{K(H(S_1|S_2)+\epsilon)} 2^{-N_1(I(X_1;Y_3|m_1)-7\epsilon)}.$$

So

$$P(E_3|E_0^c, E_1^c, E_2^c)$$

$$\leq 2^{-K(H(S_1|S_2)+\epsilon)} (1 + 2^{K(H(S_1|S_2)+\epsilon)} 2^{-N_1(I(X_1;Y_3|m_1)-7\epsilon)}) 2^{-KR} 2^K(H(S_1|S_3)+\epsilon)$$

$$= 2^{-K(R-\epsilon-(H(S_1|S_2)+\epsilon))} + 2^{-K(R+\frac{N_1}{2}(I(X_1;Y_3|m_1)-7\epsilon)-(H(S_1|S_3)+\epsilon))}.$$
and

\[ R > H(S_1|S_3) + \epsilon - \frac{N_1}{K} (I(X_1;Y_3|m_1) - 7\epsilon) > H(S_1|S_3) + \epsilon - \tau_1 I(X_1;Y_3|m_1), \quad (A.52) \]

and \( K \) is sufficiently large, \( P(E_0|E_1^c,E_2^c) \leq \epsilon \). Together with (A.46) and (A.50), one can get

\[ H(S_1|S_3) + \epsilon - \min\{I(X_1;Y_3|m_1) - 7\epsilon, C_2\}(H(S_1|S_2) + \epsilon) \leq \tau_0 C_{(1,2)} - 3. \quad (A.53) \]

Thus, if both (A.46) and (A.53) are satisfied, there exists a source-channel code that makes the error probability at node-3 \( P_{e}^{N,K} < 6\epsilon \).

Next step is to apply the result to the Gaussian channel. In this case, we have

\[ C_2 = C(h_{12}^2 P), \]

\[ I(X_1;Y_3|m_1) = C(h_{13}^2 P), \quad (A.54) \]

\[ C_{(1,2)} - 3 = C((h_{13}^2 + h_{23}^2) P). \]

Inserting (A.46) to (A.53) and (A.54) completes the proof.

**A.9 Proof of Theorem 14**

Part 1) and 2) of this theorem follow straightforward limit calculation, we only prove part 3).

The assumption \( \tau_{ex,2} < \tau_{ex,3} \) becomes \( \frac{H(S_1|S_2)}{h_{12}^2} < \frac{H(S_1|S_3)}{h_{13}^2} \) when \( P \to 0 \). Under this assumption, there are two different cases corresponding to different cost functions for the benchmark scheme: \( H(S_1|S_2) > H(S_1|S_3) \) and \( H(S_1|S_2) < H(S_1|S_3) \).

When \( H(S_1|S_2) > H(S_1|S_3) \), in which case \( h_{12}^2 > h_{13}^2 \) and

\[ E_{1,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \left( \frac{1}{h_{13}^2} - \frac{1}{h_{12}^2} \right) H(S_1|S_3) \right). \quad (A.55) \]
\[ E_{2,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_2) - h_{23}^2 H(S_1|S_2)}{(h_{13}^2 + h_{23}^2) h_{12}^2} \right) \]

\[ < \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_3) - h_{13}^2 H(S_1|S_3)}{(h_{13}^2 + h_{23}^2) h_{12}^2} \right) \]

\[ < \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_3) - h_{13}^2 H(S_1|S_3)}{h_{13}^2 h_{12}^2} \right) \]

\[ \leq \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \left( \frac{1}{h_{13}^2} - \frac{1}{h_{12}^2} \right)^+ H(S_1|S_3) \right) \]

\[ = E_{1,m}. \]

When \( H(S_1|S_2) < H(S_1|S_3) \),

\[ E_{1,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_3)}{h_{13}^2} + \left( \frac{1}{h_{12}^2} - \frac{1}{h_{13}^2} \right)^+ H(S_1|S_2) \right), \quad (A.57) \]

so

\[ E_{2,m} = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1|S_2)}{(h_{13}^2 + h_{23}^2) h_{12}^2} \right) \]

\[ < \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \frac{h_{12}^2 H(S_1|S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1|S_2)}{h_{13}^2 h_{12}^2} \right) \]

\[ = \frac{2}{\log e} \left( \frac{H(S_1|S_2)}{h_{12}^2} + \left( \frac{1}{h_{13}^2} - \frac{1}{h_{12}^2} \right)^+ H(S_1|S_2) \right) \]

\[ = E_{1,m}. \]

**A.10 Proof of Theorem 15**

For part 1) of this theorem, without loss of generality, we only prove the case when \( h_{23}^2 \to \infty \). In this case, \( \lim_{h_{23}^2 \to \infty} \tau_{2,gen} = \lim_{h_{23}^2 \to \infty} \tau_{3,gen} = 0 \), \( \tau_{gen} = \tau_{1,gen} \). In the following, we will show that the genie-aided bound could be approached using the following multicast order \( 2 \to 3 \to 1 \).

When node-2 multicasts \( S_2^K \) to both node-3 and node-1 using the proposed cooperative multicast with side-information scheme, from Lemma 13 we know it requires

\[ \tau_{2-(3,1)} = \frac{H(S_2|S_3)}{R_{CF_1d3}(\alpha)} + \frac{H(S_2|S_1) - \alpha H(S_2|S_3) \min\{I(X_2;Y_1),I(X_2;Y_1,Y_3)\}}{R_{CF_1d3}(\alpha)} \frac{C((h_{12}^2 + h_{13}^2)P)}{C((h_{12}^2 + h_{13}^2)^+ P)}. \]
$R_{CF_{1d3}}$ means the achievable rate of the following relay channel using Compress-Forward scheme: node-2 is the source, node-1 acts as relay that spends $1 - \alpha$ part of the time in helping destination using CF scheme, and node-3 acts as the destination.

Next consider node-3 multicasts $S^K_3$ to both node-1, node-2. At this time, node-1 already has $S_1, S_2$, thus this step requires

$$\tau_{3-(2,1)} = \frac{H(S_3|S_2)}{R_{CF_{1d2}}(\alpha)} + \frac{H(S_3|S_1, S_2) - \alpha H(S_2|S_3)\min\{I(X_3;Y_1), I(X_3;Y_1,Y_2)\}}{C((h_{12}^2 + h_{13}^2)P)}.$$  

Final step, node-1 multicasts $H(S_1|S_2, S_3)$ to both node-2, node-3 using the greedy multicast scheme developed in the multicast section, this step requires $\tau_{1-(2,3)} = \frac{H(S_1|S_2, S_3)}{R_g}$.

Thus, the total bandwidth expansion factor of this scheme is

$$\tau = \tau_{2-(3,1)} + \tau_{3-(1,2)} + \tau_{1-(2,3)}.$$  

(A.59)

Based on the results on the relay channel and multicast channel $h_{23}^2 \to \infty$, $R_{CF_{1d3}} \to \infty$, $R_{CF_{1d2}} \to \infty$, $\lim_{h_{23}^2 \to \infty} R_g = C((h_{12}^2 + h_{13}^2)P)$. Then

$$\lim_{h_{23}^2 \to \infty} \tau = \frac{H(S_2|S_1) + H(S_3|S_1, S_2) + H(S_1|S_2, S_3)}{C((h_{12}^2 + h_{13}^2)P)} = \frac{H(S_2, S_3|S_1) + H(S_1|S_2, S_3)}{C((h_{12}^2 + h_{13}^2)P)} = \tau_{1,gen} = \tau_{gen}.$$  

(A.60)

To prove the second part of this theorem, without loss of generality, suppose $1 \to 2 \to 3$ is the optimal multicast order for the scheme that uses broadcast with degraded information set. Then, just use the same order for the list source-channel decoding scheme based multicast with side-information. Theorem 14 shows that at every multicast step, the list source-channel decoding scheme outperforms the broadcast with degraded information set. Thus even with this not necessarily optimal order, the list source-channel decoding scheme outperforms the scheme that uses broadcast with degraded information set with optimal order.

118
A.11 Proof of Theorem 16

The proof follows that of [33].

\[ nR_e = H(W_1|Y_2^n) \]  
\[ = H(W_1) - I(W_1; Y_2^n) \]  
\[ = I(W_1; Y^n) - I(W_1; Y_2^n) + H(W_1|Y^n) \]  
\[ \leq \sum_{i=1}^{n} [I(W_1; Y_i|Y^{i-1}) - I(W_1; Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n, \]  
where \( Y^{i-1} = Y(1, \cdots, i-1), Y_{2,i+1}^n = Y_2(i+1, \cdots, n), \) and \( \delta_n \to 0 \) as \( n \to \infty. \) We get this by using the chain rule to expand \( I(W_1; Y^n) \) from \( i = 1 \) and expand \( I(W_1; Y_2^n) \) from \( i = n, \) also we use the Fano’s inequality to bound \( H(W_1|Y^n). \)

We continue

\[ nR_e \leq \sum_{i=1}^{n} [I(W_1; Y_i|Y^{i-1}) - I(W_1; Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n \]  
\[ = \sum_{i=1}^{n} [I(W_1, Y_{2,i+1}^n; Y_i|Y^{i-1}) - I(Y_{2,i+1}^n; Y_i|Y^{i-1}, W_1)] \]  
\[ - I(W_1, Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n) + I(Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n, W_1) \]  
\[ + n\delta_n \]  
\[ = \sum_{i=1}^{n} [I(W_1, Y_{2,i+1}^n; Y_i|Y^{i-1}) - I(W_1, Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n, \]  
since \( \sum_{i=1}^{n} I(Y_{2,i+1}^n; Y_i|Y^{i-1}, W_1) = \sum_{i=1}^{n} I(Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n, W_1), \) which is proved in the lemma 7 of [33]. Now

\[ nR_e \leq \sum_{i=1}^{n} [I(W_1, Y_{2,i+1}^n; Y_i|Y^{i-1}) - I(W_1, Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n \]  
\[ = \sum_{i=1}^{n} [I(Y_{2,i+1}^n; Y_i|Y^{i-1}) + I(W_1; Y_i|Y^{i-1}, Y_{2,i+1}^n)] \]  
\[ - I(Y^{i-1}; Y_{2,i}|Y_{2,i+1}^n) - I(W_1; Y_{2,i}|Y^{i-1}, Y_{2,i+1}^n) \]  
\[ + n\delta_n \]  
\[ = \sum_{i=1}^{n} [I(W_1; Y_i|Y^{i-1}, Y_{2,i+1}^n) - I(W_1; Y_{2,i}|Y^{i-1}, Y_{2,i+1}^n)] + n\delta_n, \]  
\[ \text{(A.61)} \]  
\[ \text{(A.62)} \]  
\[ \text{(A.63)} \]  
\[ \text{(A.64)} \]  
\[ \text{(A.65)} \]  
\[ \text{(A.66)} \]  
\[ \text{(A.67)} \]  
\[ \text{(A.68)} \]  
\[ \text{(A.69)} \]  
\[ \text{(A.70)} \]  
\[ \text{(A.71)} \]  
\[ 119 \]
since $\sum_{i=1}^{n} I(Y_{2,i+1}; Y_i | Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1}; Y_{2,i} Y_{2,i+1}^{n})$, which is also proved in [33].

Now, let $J$ be a random variable uniformly distributed over $\{1, \cdots, n\}$, set $U = JY_{2,i+1}^{n}, V_1 = JY_{2,i+1}^{n} W_1, V_2 = JY_{2,i+1}^{n}, Y_1 = Y_{1,j}, Y_2 = Y_{2,j}, Y = Y_J, X_1 = X_{1,j}, X_2 = X_{2,j}$ we have

$$R_e \leq \frac{1}{n} \sum_{i=1}^{n} [I(W_1; Y_i | Y^{i-1}, Y_{2,i+1}^{n}) - I(W_1; Y_{2,i} | Y^{i-1}, Y_{2,i+1}^{n})] + \delta_n$$

(A.72)

$$- I(W_1, Y_{2,i+1}^{n}; Y_{2,i} | Y^{i-1}, Y_{2,i+1}^{n})] + \delta_n$$

$$= I(V_1, V_2; Y | U) - I(V_1, V_2; Y | U) + \delta_n.$$  (A.73)

Since the channel is memoryless, one can then check that $U \rightarrow (V_1, V_2) \rightarrow (X_1, X_2) \rightarrow (Y, Y_1, Y_2)$ is a Markov chain. In the following, we bound $R_1$.

$$I(W_1; Y) = \sum_{i=1}^{n} I(W_1; Y_i | Y^{i-1})$$

$$= \sum_{i=1}^{n} [H(Y_i | Y^{i-1}) - H(Y_i | W_1, Y^{i-1})]$$

$$\leq \sum_{i=1}^{n} [H(Y_i) - H(Y_i | W_1, Y^{i-1})]$$

$$\leq \sum_{i=1}^{n} [H(Y_i) - H(Y_i | W_1, Y^{i-1}, Y_{2,i+1}^{n})]$$

$$= \sum_{i=1}^{n} I(W_1, Y^{i-1}, Y_{2,i+1}^{n}; Y_i).$$  (A.74)

Hence

$$R_1 \leq \frac{1}{n} I(W_1; Y) \leq I(V_1, V_2; Y).$$  (A.75)
Also

\[ I(W_1; Y) = \sum_{i=1}^{n} I(W_1; Y_i|Y_{i-1}) \]

\[ \leq \sum_{i=1}^{n} I(W_1; Y_i, Y_{i,i}|Y_{i-1}) \]

\[ = \sum_{i=1}^{n} [H(Y_i, Y_{1,i}|Y_{i-1}) - H(Y_i, Y_{1,i}|W_1, Y_{i-1})] \]

\[ \leq \sum_{i=1}^{n} [H(Y_i, Y_{1,i}|Y_{i-1}) - H(Y_i, Y_{1,i}|W_1, Y_{i-1}, Y_{2,i+1})] \]

\[ = \sum_{i=1}^{n} I(W_1, Y_{2,i+1}; Y_i, Y_{1,i}|Y_{i-1}). \quad (A.76) \]

Hence, we have

\[ R_1 \leq \frac{1}{n} I(W_1; Y) = I(V_1; Y, Y_1|V_2). \quad (A.77) \]

So, we have

\[ R_1 \leq \min\{I(V_1, V_2; Y), I(V_1; Y, Y_1|V_2)\}. \quad (A.78) \]

The claim is proved.

### A.12 Proof of Theorem 17

The proof is a combination of the coding schemes of Csiszár et al. [33] and the regular coding and backward decoding scheme in the relay channel [16, 83]. We first replace \( V_1, V_2 \) in Theorem 17 with \( X_1, X_2 \). After proving Theorem 17 with \( V_1, V_2 \) replaced by \( X_1, X_2 \), we then prefix a memoryless channel with input \( V_1, V_2 \) and transmission probability \( p(x_1, x_2|v_1, v_2) \) as reasoned in [33] to finish our proof.

1. **Codebook generation:**
For a given distribution \(p(x_1, x_2)\), we first generate at random \(2^{nR}\) i.i.d \(n\)-sequence at the relay node each drawn according to \(p(x_2) = \prod_{i=1}^{n} p(x_{2,i})\), index them as \(x_2(a), a \in [1, 2^{nR}]\), where \(R = \min\{I(X_1, X_2; Y), I(X_1; Y_1 | X_2)\} - \epsilon_0\). For each \(x_2(a)\), generate \(2^{nR}\) conditionally independent \(n\)-sequence \(x_1(k, a)\), \(k \in [1, 2^{nR}]\) \(\) drawn randomly according to \(p(x_1|x_2(a)) = \prod_{i=1}^{n} p(x_{1,i}|x_{2,i}(a))\). Define \(\mathcal{W} = \{1, \ldots, 2^{n[R-I(X_1,X_2;Y_2)]}\}, \mathcal{L} = \{1, \ldots, 2^{nI(X_1,X_2;Y_2)}\}\) and \(\mathcal{K} = \mathcal{W} \times \mathcal{L} = \{1, \ldots, 2^{nR}\}\).

2. Encoding

We exploit the block Markov coding scheme, as argued in [19], the loss induced by this scheme is negligible as the number of blocks \(B \rightarrow \infty\).

For a given rate pair \((R_1, R_e)\) with \(R_1 \leq R\) and \(R_e \leq R_1\), we give the following coding strategy. Let the message to be transmitted at block \(i\) be \(w_1(i) \in \mathcal{W}_1 = \{1, \ldots, M\}\), where \(M = 2^{nR_1}\).

The stochastic encoder at the transmitter first forms the following mappings.

- If \(R_1 > R - I(X_1, X_2; Y_2)\), then we let \(\mathcal{W}_1 = \mathcal{W} \times \mathcal{J}\), where

\[
\mathcal{J} = \{1, \ldots, 2^{n(R_1 - [R - I(X_1, X_2; Y_2)])}\}.
\]

We let \(g_1\) be the partition that partitions \(\mathcal{L}\) into \(|\mathcal{J}|\) equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message \(w_1(i) = (w(i), j(i)) \rightarrow (w(i), l(i))\), where \(l(i)\) is chosen randomly from the set \(\mathcal{g}_1^{-1}(j(i)) \subset \mathcal{L}\) with uniform distribution.

- If \(R_1 < R - I(X_1, X_2; Y_2)\), the stochastic encoder will choose a mapping \(w_1(i) \rightarrow (w_1(i), l(i))\), where \(l(i)\) is chosen uniformly from the set \(\mathcal{L}\).
Assume that the message \( w_1(i - 1) \) transmitted at block \( i - 1 \) is associated with \((w(i - 1), l(i - 1))\) and the message \( w_1(i) \) intended to send at block \( i \) is associated with \((w(i), l(i))\) by the stochastic encoder at the transmitter. We let \( a(i - 1) = (w(i - 1), l(i - 1)) \) and \( b(i) = (w(i), l(i)) \). The encoder then sends \( x_1(b(i), a(i - 1)) \). The relay has an estimation \( \hat{a}(i - 1) \) (see the decoding part), and thus sends the corresponding codeword \( x_2(\hat{a}(i - 1)) \).

At block 1, the source sends \( x_1(b(1), 1) \), the relay sends \( x_2(1) \).

At block \( B \), the source sends \( x_1(1, a(B - 1)) \), and the relay sends \( x_2(\hat{a}(B - 1)) \).

3. Decoding

At the end of block \( i \), the relay already has an estimation of the \( \hat{a}(i - 1) \), which was sent at block \( i - 1 \), and will declare that it receives \( \hat{a}(i) \), if this is the only pair such that \( (x_1(\hat{a}(i), \hat{a}(i - 1)), x_2(\hat{a}(i - 1)), y_1(i)) \) are jointly typical. Since \( R = \min\{I(X_1; Y_1|X_2), I(X_1, X_2; Y)\} - \epsilon \leq I(X_1; Y_1|X_2) - \epsilon \), then based on the AEP, one has \( \hat{a}(i) = a(i) \) with probability goes to 1.

The destination decodes from the last block, i.e. block \( B \). Suppose that at the end of block \( B - 1 \), the relay decodes successfully, then the destination will declare that \( \hat{a}(B - 1) \) is received, if \( (x_1(1, \hat{a}(B - 1)), x_2(\hat{a}(B - 1)), y) \) are jointly typical. It’s easy to see that if \( R \leq I(X_1, X_2; Y) \), we will have \( \hat{a}(B - 1) = a(B - 1) \) with probability goes to 1, as \( n \) increases.

After getting \( \hat{a}(B - 1) \), the receiver can get an estimation of \( a(i), i \in [1, B - 2] \) in a similar way.

Having \( \hat{a}(i - 1) \), the destination can get the estimation of the message \( w_1(i - 1) \) by letting
1) \( \hat{w}_1(i-1) = (\hat{w}(i-1), j(i-1)) = (\hat{w}(i-1), g_1(\hat{l}(i-1))) \) if \( R_1 > R - I(X_1, X_2; Y_2) \),

2) \( \hat{w}_1(i - 1) = \hat{w}(i - 1) \) if \( R_1 < R - I(X_1, X_2; Y_2) \).

The probability that \( \hat{w}_1(i - 1) = w_1(i - 1) \) goes to one for sufficiently large \( n \).

4. **Equivocation Computation**

\[
H(W_1|Y_2) = H(W_1, Y_2) - H(Y_2)
\]

\[
= H(W_1, Y_2, X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2)
\]

\[
= H(X_1, X_2) + H(W_1, Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2)
\]

\[
\geq H(X_1) + H(Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2).
\]

First, let us calculate \( H(X_1, X_2|W_1, Y_2) \). Given \( W_1 \), the eavesdropper can also do backward decoding as the receiver. At the end of block \( B \), given \( W_1 \), the eavesdropper knows \( w(B - 1) \), hence it will decode \( l(B - 1) \), by letting \( l(B - 1) = \hat{l}(B - 1) \), if \( \hat{l}(B - 1) \) is the only one such that \( (x_1(1, (w(B - 1), \hat{l}(B - 1))), x_2((w(B - 1), \hat{l}(B - 1))))) \) are jointly typical. Since \( l \in [1, 2^{nI(X_1,X_2;Y_2)}] \), we have

\[
\Pr\{((X_1(1, a(i - 1)), X_2(\hat{a}(i - 1)))) \neq (X_1(1, (w(B - 1), \hat{l}(B - 1))), X_2((w(B - 1), \hat{l}(B - 1)))))\} \leq \epsilon_1.
\]

Then based on Fano’s inequality, we have

\[
\frac{1}{n} H(X_1, X_2|W_1 = w_1, Y_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2) \tag{A.80}
\]

Hence, we have

\[
\frac{1}{n} H(X_1, X_2|W_1, Y_2) = \frac{1}{n} \sum_{w_1 \in W_1} p(W_1 = w_1) H(X_1, X_2|W_1 = w_1, Y_2) \leq \epsilon_2, \tag{A.81}
\]

124
when $n$ is sufficiently large.

Since the channel is memoryless, we have $H(Y_2) - H(Y_2|X_1,X_2) \leq nI(X_1,X_2;Y_2) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$ [32].

Now, from the code construction, we have $H(X_1) = nR$ if $R_1 > R - I(X_1,X_2;Y_2)$. In this case, we get $nR_e = H(W_1|Y_2) \geq n(R - I(X_1,X_2;Y_2) - \epsilon_3)$. If $R_1 \leq R - I(X_1,X_2;Y_2)$, $H(X_1) = n(R_1 + I(X_1,X_2;Y_2))$, in this case, we get the perfect secrecy, since

$$nR_e \geq n(R_1 + I(X_1,X_2;Y_2)) - nI(X_1,X_2;Y_2) - n\epsilon_3 \geq n(R_1 - \epsilon_3).$$

The claim is proved.

### A.13 Proof of Theorem 18

As [33], we first prove the result for the case where $V_1,V_2$ in Theorem 18 are replaced with $X_1,X_2$, then prefix a memoryless channel with transition probability $p(x_1|v_1)p(x_2|v_2)$ to finish our proof.

We first consider the case $I(X_1;Y|X_2) < I(X_1;Y_2|X_2)$, i.e., the channel between the source and the eavesdropper is better than the channel between the source and the destination. In this case, we only need to consider $\min\{I(X_2;Y), I(X_2;Y_2)\} = I(X_2;Y_2)$, otherwise, the secrecy rate will be zero. Thus in this case, the last equation in (3.8) changes to $R_e < \left[ I(X_1;Y|X_2) + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - I(X_1,X_2;Y_2) \right]^+.$

1. **Codebook generation:**

   For a given distribution $p(x_1)p(x_2)$, we generate at random $2^{nR_2}$ i.i.d $n$-sequence at the relay node each drawn according to $p(x_2) = \prod_{i=1}^n p(x_{2,i})$, index them as
\( x_2(a), a \in [1, 2^{nR_2}] \), where we set \( R_2 = \min\{I(X_2; Y), I(X_2; Y|X_1)\} - \epsilon \). We also generate random \( 2^n R \) i.i.d \( n \)-sequence at the source each drawn according to \( p(x_1) = \prod_{i=1}^{n} p(x_{1,i}) \), index them as \( x_1(k), k \in [1, 2^n R] \), where \( R = I(X_1; Y|X_2) - \epsilon \). Let

\[
R' = \min\{I(X_2; Y), I(X_2; Y|X_1)\} + I(X_1; Y|X_2) - I(X_1, X_2; Y),
\]

and define \( \mathcal{W} = \{1, \cdots, 2^{nR'}\} \), \( \mathcal{L} = \{1, \cdots, 2^{n(R-R')}\} \) and \( \mathcal{K} = \mathcal{W} \times \mathcal{L} = \{1, \cdots, 2^{nR}\} \).

2. Encoding

For a given rate pair \((R_1, R_e)\) with \( R_1 \leq R, R_e \leq R_1 \), we give the following coding strategy. Let the message to be transmitted at block \( i \) be \( w_1(i) \in \mathcal{W}_1 = [1, M] \), where \( M = 2^{nR_1} \).

The stochastic encoder at the transmitter first forms the following mappings.

- If \( R_1 > R' \), then we let \( \mathcal{W}_1 = \mathcal{W} \times \mathcal{J} \), where \( \mathcal{J} = \{1, 2^{n(R_1-R')}\} \). We let \( g_1 \) be the partition that partitions \( \mathcal{L} \) into \( |\mathcal{J}| \) equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message \( w_1(i) = (w(i), j(i)) \rightarrow (w(i), l(i)) \), where \( l(i) \) is chosen randomly from the set \( g_1^{-1}(j(i)) \subset \mathcal{L} \) with uniform distribution.

- If \( R_1 < R' \), the stochastic encoder will choose a mapping \( w_1(i) \rightarrow (w_1(i), l(i)) \), where \( l(i) \) is chosen uniformly from the set \( \mathcal{L} \).

Suppose the message \( w_1(i) \) intended to send at block \( i \) is associated with \((w(i), l(i))\) by the stochastic encoder at the transmitter. The encoder then sends \( x_1((w(i), l(i))) \). The relay uniformly picks a code \( x_2(a) \) from \( a \in [1, \cdots, 2^{nR_2}] \), and sends \( x_2(a) \).
3. Decoding

At the end of block $i$, the destination declares that $\hat{a}(i)$ is received, if $\hat{a}(i)$ is the only one such that $(x_2(\hat{a}(i)), y)$ are jointly typical. If there does not exist or there exist more than one such sequences, the destination declares an error. Since $R_2 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon \leq I(X_2; Y) - \epsilon$, then based on AEP, we know that the error probability will be less than any given positive number $\epsilon$, when the codeword length $n$ is long enough.

The destination then declares that $\hat{k}$ is received, if $\hat{k}$ is the only one such that $(x_1(\hat{k}), x_2(\hat{a}), y)$ are jointly typical, otherwise declares an error. Since $R = I(X_1; Y|X_2) - \epsilon$, then based on AEP, we know that we will have error probability goes to zero, when $n$ is sufficiently large.

Having $\hat{k}(i)$, the destination can get the estimation of the message $w_1(i)$ by letting

1) $\hat{w}_1(i) = (\hat{\omega}(i), \hat{j}(i)) = (\hat{\omega}(i), g_1(\hat{l}(i)))$, if $R_1 > R'$,

2) $\hat{w}_1(i) = \hat{\omega}(i)$, if $R_1 < R'$.

The probability that $\hat{w}_1(i) = w_1(i)$ goes to one for sufficiently large $n$.

4. Equivocation Computation

\[
H(W_1|Y_2) = H(W_1, Y_2) - H(Y_2)
\]

\[
= H(W_1, Y_2, X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2)
\]

\[
= H(X_1, X_2) + H(W_1, Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2)
\]

\[
\geq H(X_1, X_2) + H(Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2).
\]

Now let’s calculate $H(X_1, X_2|W_1, Y_2)$. Given $W_1$, the eavesdropper can do joint decoding. At any block $i$, given $W_1$, the eavesdropper knows $w(i)$, hence it will
decode $l(i)$ and $a(i)$ sent by the relay, by letting $l(i) = \hat{l}(i), a(i) = \hat{a}(i)$, if $\hat{l}(i), \hat{a}(i)$ are the only one such that $(x_1(w(i), \hat{l}(i)), x_2(\hat{a}(i)), y)$ are jointly typical. Then, since $R_2 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon \leq I(X_2; Y_2|X_1) - \epsilon$, we get

$$\frac{1}{2} \log(|\mathcal{L}|) + R_2 = R + I(X_1, X_2; Y_2) - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}$$

$$- I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon$$

$$\leq I(X_1, X_2; Y_2) - \epsilon,$$  \hspace{1cm} (A.82)

Also, we have $\frac{1}{2} \log(|\mathcal{L}|) < R \leq I(X_1; Y_2|X_2) - \epsilon$.

So $\Pr\{(X_1(w(i), \hat{l}(i)), X_2(\hat{a}(i))) \neq (X_1(w(i), l(i)), X_2(a(i)))\} \leq \epsilon_1$.

Then based on Fano’s inequality, we have

$$\frac{1}{n} H(X_1, X_2|W_1 = w_1, Y_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2)$$  \hspace{1cm} (A.83)

Hence, we have

$$\frac{1}{n} H(X_1, X_2|W_1, Y_2) = \frac{1}{n} \sum_{w_1 \in W_1} p(W_1 = w_1) H(X_1, X_2|W_1 = w_1, Y_2) \leq \epsilon_2,$$

when $n$ is sufficiently large.

Now, $H(Y_2) - H(Y_2|X_1, X_2) \leq nI(X_1, X_2; Y_2) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$.

Also we have $H(X_1, X_2) = H(X_1) + H(X_2)$ since $x_1$ and $x_2$ are independent. If $R_1 > R'$, we have $H(X_1, X_2) = R + R_2 = I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}$.

Combining these, we get $nR_e = H(W_1|Y_2) \geq n(\min\{I(X_2; Y), I(X_2; Y_2|X_1)\} + I(X_1; Y|X_2)) - nI(X_1, X_2; Y_2) - n\epsilon_4$.

On the other hand, if $R_1 < R'$, we have

$$H(X_1) = R_1 + I(X_1, X_2; Y_2) - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\},$$

128
hence we have $H(X_1, X_2) = R_1 + I(X_1, X_2; Y_2) - \epsilon$. We get perfect secrecy rate, since $nR_e = H(W_1|Y_2) \geq nR_1 - n\epsilon_4$.

This case is proved.

Now, consider the case $I(X_1; Y|X_2) > I(X_1; Y_2|X_2)$. If min\{\(I(X_2; Y), I(X_2; Y_2)\)\} = $I(X_2; Y)$, then we have min\{\(I(X_2; Y), I(X_2; Y_2|X_1)\)\} = $I(X_2; Y)$, because $I(X_2; Y_2|X_1) > I(X_2; Y_2)$ since $X_1, X_2$ are independent. Under this case, we only need to prove $R_e \leq I(X_1; Y|X_2) - I(X_1; Y_2|X_2)$ are achievable, which can be achieved by letting the code-word rate be $I(X_1; Y|X_2)$ and $R' = I(X_1; Y|X_2) - I(X_1; Y|X_2)$. Now the equivocation rate of the eavesdropper can be calculated as

$$H(W_1|Y_2) \geq H(W_1|Y_2, X_2)$$

$$= H(W_1, Y_2|X_2) - H(Y_2|X_2)$$

$$= H(W_1, Y_2, X_1|X_2) - H(X_1|W_1, Y_2, X_2) - H(Y_2|X_2)$$

$$= H(X_1) + H(Y_2|X_1, X_2) - H(X_1|W_1, Y_2, X_2) - H(Y_2|X_2)$$

since $x_1, x_2$ are independent. This can then be shown to be larger than $n(I(X_1; Y|X_2) - I(X_1; Y_2|X_2) - \epsilon)$.

If min\{\(I(X_2; Y), I(X_2; Y_2)\)\} = $I(X_2; Y_2)$, the last line in (3.8) changes to $R_e < [I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - I(X_1, X_2; Y_2)]^+$, then we can use a coding/decoding scheme similar to the one developed above to show the achievability.

The claim is achieved.
A.14 Proof of Theorem 19

The proof is a combination of the coding scheme of Csiszár et. al. [33] and a revised CF scheme in the relay channel [19].

1. Codebook generation:

We first generate at random $2^{nR} \text{i.i.d } n$-sequence $x_1$ at the source node each drawn according to $p(x_1) = \prod_{j=1}^{n} p(x_{1,j})$, index them as $x_1(k), k \in [1, 2^{nR}]$, with $R = I(X_1; \hat{Y}_1, Y | X_2) - \epsilon$.

Generate at random $2^{nR_2} \text{i.i.d } n$-sequence $x_2$ each with probability $p(x_2) = \prod_{j=1}^{n} p(x_{2,j})$.

Index these as $x_2(s), s \in [1, 2^{nR_2}]$, where $R_2 = \min\{I(X_2; Y), I(X_2; Y_2 | X_1)\} - \epsilon$.

For each $x_2(s)$, generate at random $2^{n(R_2-R_0)} \text{i.i.d } \hat{y}_1$, each with probability $p(\hat{y}_1 | x_2(s)) = \prod_{j=1}^{n} p(\hat{y}_{1,j} | x_{2,j}(s))$. Label these $\hat{y}_1(z, s), z \in [1, 2^{nR}], s \in [1, 2^{nR_2}]$, where we set $\hat{R} = R_2 - R_0$. Equally divide these $2^{nR_2} x_2$ sequences into $2^{nR}$ bins, hence there are $2^{nR_0} x_2$ sequences at each bin. Let $f$ be this mapping, that is $z = f(s)$.

Let $R' = \min\{I(X_2; Y), I(X_2; Y_2 | X_1)\} + I(X_1; \hat{Y}_1, Y | X_2) - I(X_1, X_2; Y_2)$.

Define $\mathcal{W} = \{1, \cdots, 2^{nR'}\}, \mathcal{L} = \{1, \cdots, 2^{n(R-R')}\}$ and $\mathcal{K} = \mathcal{W} \times \mathcal{L} = \{1, \cdots, 2^{nR}\}$.

2. Encoding

We exploit the block Markov coding scheme.

For a given rate pair $(R_1, R_e)$, where $R_1 \leq R, R_e \leq R_1$, we give the following coding strategy. Let the message to be transmitted at block $i$ be $w_1(i) \in \mathcal{W}_1 = [1, M]$, where $M = 2^{nR_1}$. We require $R_1 \leq R$. 

130
The stochastic encoder at the transmitter first forms the following mappings.

- If $R_1 > R'_1$, we let $W_1 = W \times J$, where $J = \{1, 2^{n(R_1-R'_1)}\}$. We let $g_1$ be the partition that partitions $L$ into $|J|$ equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message $w_1(i) = (w(i), j(i)) \rightarrow (w(i), l(i))$, where $l(i)$ is chosen randomly from the set $g_1^{-1}(j(i)) \subseteq L$ with uniform distribution.

- If $R_1 < R'_1$, the stochastic encoder will choose a mapping $w_1(i) \rightarrow (w_1(i), l(i))$, where $l(i)$ is chosen uniformly from the set $L$.

At first consider block $i$, where $i \neq 1, B$, which means it’s not the first or the last block. Assume that the message $w_1(i)$ intended to send at block $i$ is associated with $(w(i), l(i))$ by the stochastic encoder at the transmitter. We let $k(i) = (w(i), l(i))$.

Then the encoder at the source sends $x_1(k(i))$ at block $i$. At the end of block $i - 1$, we assume that $(x_2(s(i - 1)), \hat{y}_1(z(i - 1), s(i - 1)), y_1(i - 1))$ are jointly typical$^{17}$, then we choose $s(i)$ uniformly from bin $z(i - 1)$, and the relay sends $x_2(s(i))$ at block $i$.

When $i = 1$, the source sends $x_1(k(1))$, the relay sends $x_2(1)$. When $i = B$, the source sends $x_1(1)$, the relay sends $x_2(s(B))$.

3. Decoding

First consider the relay node. At the end of block $i$, the relay already has $s(i)^{18}$, it then decides $z(i)$ by choosing $z(i)$ such that $(x_2(s(i)), \hat{y}_1(z(i), s(i)), y_1(i))$ are jointly typical. There exists such $z(i)$, if

$$\hat{R} \geq I(Y_1; \hat{Y}_1|X_2),$$  \hspace{1cm} (A.85)

$^{17}$See the decoding part, such $z(i - 1)$ exists.

$^{18}$At the end of block 1, relay knows $s(i) = 1$, this is the starting point.
and \( n \) is sufficiently large. Choose \( s(i + 1) \) uniformly from bin \( z(i) \).

The destination does backward decoding. The decoding process starts at the last block \( B \), the destination decodes \( s(B) \) by choosing unique \( \hat{s}(B) \) such that \((x_2(\hat{s}(B)), y(B))\) are jointly typical. We will have \( \hat{s}(B) = s(B) \), if

\[
R_2 \leq I(X_2; Y),
\]

(A.86)

and \( n \) is sufficiently large.

Next, the destination moves to the block \( B - 1 \). Now it already has \( s(B) \), hence we also have \( z(B - 1) = f(s(B)) \). It first declares that \( \hat{s}(B - 1) \) is received, if \( \hat{s}(B - 1) \) is the unique one such that \((x_2(\hat{s}(B - 1)), y(B - 1))\) are jointly typical. If (A.86) is satisfied, \( \hat{s}(B - 1) = s(B - 1) \) with high probability. After knowing \( \hat{s}(B - 1) \), the destination gets an estimation of \( \hat{k}(B - 1) \), by picking the unique \( \hat{k}(B - 1) \) such that \((x_1(\hat{k}(B - 1)), \hat{y}_1(z(B - 1), \hat{s}(B - 1)), y(B - 1), x_2(\hat{s}(B - 1)))\) are jointly typical. We will have \( \hat{k}(B - 1) = k(B - 1) \) with high probability, if

\[
R \leq I(X_1; \hat{Y}_1, Y | X_2),
\]

(A.87)

and \( n \) is sufficiently large.

When the destination moves to block \( i \), the destination has \( s(i + 1) \) and hence \( z(i) = f(s(i + 1)) \). It first declares that \( \hat{s}(i) \) is received, by choosing unique \( \hat{s}(i) \) such that \((x_2(\hat{s}(i)), y(i))\) are jointly typical. If (A.86) is satisfied, \( \hat{s}(i) = s(i) \) with high probability. After knowing \( \hat{s}(i) \), the destination declares that \( \hat{k}(i) \) is received, if \( \hat{k}(i) \) is the unique one such that \((x_1(\hat{k}(i)), \hat{y}_1(z(i), \hat{s}(i)), y(i), x_2(\hat{s}(i)))\) are jointly typical. If (A.87) is satisfied, \( \hat{k}(i) = k(i) \) with high probability when \( n \) is sufficiently large.
Having $\hat{k}(i)$, the destination can get the estimation of the message $w_1(i)$ by letting

1) $\hat{w}_1(i) = (\hat{w}(i), \hat{j}(i)) = (\hat{w}(i), g_1(\hat{l}(i)))$, if $R_1 > R - R'$, 2) $\hat{w}_1(i) = \hat{w}(i)$, if $R_1 < R - R'$. The probability that $\hat{w}_1(i) = w_1(i)$ goes to one for sufficiently large $n$.

4. Equivocation Computation

\[
H(W_1|Y_2) = H(W_1, Y_2) - H(Y_2) = H(W_1, Y_2, X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2) = H(X_1, X_2) + H(Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2) \geq H(X_1, X_2) + H(Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2).
\]

Following [32], we will have $H(Y_2) - H(Y_2|X_1, X_2) \leq nI(X_1, X_2; Y_2) + n\delta_n$, where $\delta_n \to 0$ as $n \to \infty$.

Now let’s calculate $H(X_1, X_2|W_1, Y_2)$. Given $W_1$, the eavesdropper can do joint decoding. It does backward decoding. We pick up the story at block $i$, we suppose it already decodes $s(i + 1)$ and hence $z(i) = f(s(i + 1))$. Given $W_1$, the eavesdropper knows $w(i)$, hence it will decode $l(i)$ and $s(i)$ sent by the relay, by letting $l(i) = \hat{l}(i)$, $s(i) = s(i)$, if $\hat{l}(i)$, $s(i)$ are the only ones such that

\[
(x_1(w(i), \hat{l}(i)), x_2(s(i)), y_1(z(i), s(i)), y_2(i))
\]

are jointly typical. Then, if $R_2 \leq I(X_2; Y_2|X_1)$ and (A.87) is satisfied, we have

\[
\frac{1}{2} \log(|\mathcal{L}|) + R_2 = R - R' + R_2 = R - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - I(X_1; Y, Y_2) + I(X_1, X_2, Y_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} \leq I(X_1, X_2; Y_2).
\]

(A.88)
Also, we have

$$\frac{1}{2} \log(|\mathcal{L}|) < R \leq I(X_1; \hat{Y}_1, Y_2|X_2).$$

Thus, we have

$$\Pr\{(X_1(w(i), \hat{I}(i)), X_2(\hat{s}(i)) \neq (X_1(w(i), I(i)), X_2(s(i)))\} \leq \epsilon_1. \quad (A.89)$$

Then based on Fano’s inequality, we have

$$\frac{1}{n} H(X_1, X_2, |W_1 = w_1, Y_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2). \quad (A.90)$$

Hence, we have

$$\frac{1}{n} H(X_1, X_2|W_1, Y_2) = \frac{1}{n} \sum_{w_1 \in W_1} p(W_1 = w_1) H(X_1, X_2|W_1 = w_1, Y_2) \leq \epsilon_2,$$

when $n$ is sufficiently large.

We know $H(X_1, X_2) = H(X_1) + H(X_2|X_1) \geq n(R + R_0)$.

If $R_1 > R'$, we have $H(X_1) = nR$, then we get

$$nR_e = H(W_1|Y_2) \geq n(R_0 + I(X_1; \hat{Y}_1, Y|X_2) - I(X_1, X_2; Y_2) - \epsilon_4).$$

If $R_1 < R'$, we have $H(X_1) = n(R_1 + R - R')$, hence

$$nR_e = H(W_1|Y_2) \geq nR_1 + n(R_0 - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}) - \epsilon_4).$$

The claim is proved.

### A.15 Proof of Theorem 20

The proof follows closely with that of Theorem 18. We first consider the case $I(X_1; Y|X_2) < I(X_1; Y_2|X_2)$, i.e., the channel between the source and the eavesdropper is better than
the channel between the source and the destination. In this case, we only need to consider the case \( \min \{ I(X_2; Y), I(X_2; Y_2) \} = I(X_2; Y_2) \), otherwise, the perfect secrecy rate will be zero. Thus in this case, \( R_{s_1} = \left[ I(X_1; Y | X_2) + \min \{ I(X_2; Y), I(X_2; Y_2 | X_1) \} - I(X_1, X_2; Y_2) \right]^{+} \).

1. Codebook generation:

For a given distribution \( p(x_1)p(x_2) \), we generate at random \( 2^{nR_2} \) i.i.d \( n \)-sequence at the relay node each drawn according to \( p(x_2) = \prod_{i=1}^{n} p(x_{2,i}) \), index them as \( x_2(a), a \in [1, 2^{nR_2}] \). Here we set \( R_2 = \min \{ I(X_2; Y), I(X_2; Y_2 | X_1) \} - \epsilon \). We also generate random \( 2^{nR} \) i.i.d \( n \)-sequence at the source each drawn according to \( p(x_1) = \prod_{i=1}^{n} p(x_{1,i}) \), index them as \( x_1(k), k \in [1, 2^n] \) with \( R = I(X_1; Y | X_2) - \epsilon \).

Let
\[
R_{\text{min}} = \min \{ R_{s_1}, R_{s_2} \}, R_{\text{max}} = \max \{ R_{s_1}, R_{s_2} \},
\]
where \( R_{s_1} = \min \{ I(X_2; Y), I(X_2; Y_2 | X_1) \} + I(X_1; Y | X_2) - I(X_1, X_2; Y_2) \), \( R_{s_2} = I(X_1; Y | X_2) - I(X_1; Y_1 | X_2) \).

We now define
\[
W = \{1, \cdots, 2^{nR_{\text{min}}}\}, \mathcal{L}_1 = \{1, \cdots, 2^{n(R_{\text{max}} - R_{\text{min}})}\}, \mathcal{L}_2 = \{1, \cdots, 2^{n(R - R_{\text{max}})}\}
\]
and \( \mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2, \mathcal{K} = W \times \mathcal{L} \).

2. Encoding

Here, we consider perfect secrecy rate. For a given rate \( R_1 \leq R_{\text{min}} \), we give the following coding strategy to show that for any given \( \epsilon \geq 0 \), the equivocation rate at the eavesdropper and the relay node can be made to be larger or equal \( R_1 - \epsilon \).

Let the message to be transmitted at block \( i \) be \( w_1(i) \in W_1 = [1, M] \), where \( M = 2^{nR_1} \). The stochastic encoder will choose a mapping \( w_1(i) \rightarrow (w_1(i), l_1(i), l_2(i)) \),
where \( l_1(i), l_2(i) \) are chosen uniformly from the set \( L_1, L_2 \) respectively. We write \( l(i) = (l_1(i), l_2(i)) \).

Suppose the message \( w_1(i) \) intended to send at block \( i \) is associated with \( (w(i), l(i)) \) by the stochastic encoder at the transmitter. The encoder then sends \( x_1((w(i), l(i))) \). The relay uniformly picks a code \( x_2(a) \) from \( a \in [1, \cdots, 2^{nR_2}] \), and sends \( x_2(a) \).

3. Decoding

At the end of block \( i \), the destination declares that \( \hat{a}(i) \) is received, if \( \hat{a}(i) \) is the only one such that \( (x_2(\hat{a}(i)), y) \) are jointly typical. If there does not exist or there exist more than one such sequences, the destination declares an error. Since \( R_2 = \min\{I(X_2; Y), I(X_2; Y | X_1)\} \leq I(X_2; Y) - \epsilon \), then based on AEP, we know that the error probability will be less than any given positive number \( \epsilon \), when the codeword length \( n \) is long enough.

The destination then declares that \( \hat{k} \) is received, if \( \hat{k} \) is the only one such that

\[
(x_1(\hat{k}), x_2(\hat{a}), y)
\]

are jointly typical, otherwise declares an error. Since \( R = I(X_1; Y | X_2) - \epsilon \), then based on AEP, we know that we will have error probability goes to zero, when \( n \) is sufficiently large.

Having \( \hat{k}(i) \), the destination can get the estimation of the message \( w_1(i) \) by letting \( \hat{w}_1(i) = \hat{w}(i) \). The probability that \( \hat{w}_1(i) = w_1(i) \) goes to one for sufficiently large \( n \).

4. Equivocation Computation
We first calculate the equivocation rate of the eavesdropper when \( R_{s1} \leq R_{s2} \).

\[
H(W_1|Y_2) = H(W_1, Y_2) - H(Y_2) \\
\quad = H(W_1, Y_2, X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2) \\
\quad = H(X_1, X_2) + H(W_1, Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2) \\
\geq H(X_1, X_2) + H(Y_2|X_1, X_2) - H(X_1, X_2|W_1, Y_2) - H(Y_2).
\]

Now let’s calculate \( H(X_1, X_2|W_1, Y_2) \). Given \( W_1 \), the eavesdropper can do joint decoding. At any block \( i \), given \( W_1 \), the eavesdropper knows \( w(i) \), hence it will decode \( l(i) = (l_1(i), l_2(i)) \) and \( a(i) \) sent by the relay, by letting \( l(i) = \hat{l}(i), a(i) = \hat{a}(i) \), if \( \hat{l}(i), \hat{a}(i) \) are the only one pair such that \( (x_1(w(i), \hat{l}(i)), x_2(\hat{a}(i)), y) \) are jointly typical. Since \( R_2 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon \leq I(X_2; Y_2|X_1) - \epsilon \), we

\[
\frac{1}{2} \log(|\mathcal{L}|) + R_2 = R + I(X_1, X_2; Y_2) - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} \\
\quad - I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon \\
\quad \leq I(X_1, X_2; Y_2) - \epsilon. \tag{A.91}
\]

Also, we have \( \frac{1}{2} \log(|\mathcal{L}|) < R \leq I(X_1; Y_2|X_2) \).

So \( \Pr\{(X_1(w(i), \hat{l}(i)), X_2(\hat{a}(i))) \neq (X_1(w(i), l(i)), X_2(a(i)))\} \leq \epsilon_1 \).

Then based on Fano’s inequality, we have

\[
\frac{1}{n} H(X_1, X_2|W_1 = w_1, Y_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2). \tag{A.92}
\]

Hence, we have

\[
\frac{1}{n} H(X_1, X_2|W_1, Y_2) = \frac{1}{n} \sum_{w_1 \in W_1} p(W_1 = w_1) H(X_1, X_2|W_1 = w_1, Y_2) \leq \epsilon_2,
\]

when \( n \) is sufficiently large.
Now, $H(Y_2) - H(Y_2|X_1, X_2) \leq nI(X_1, X_2; Y_2) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$. Also we have $H(X_1, X_2) = H(X_1) + H(X_2)$ since $x_1$ and $x_2$ are independent. Now $H(X_1) = R_1 + I(X_1, X_2; Y_2) - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}$, hence $H(X_1, X_2) = R_1 + I(X_1, X_2; Y_2) - \epsilon$.

We get $nR_e = H(W_1|Y_2) \geq nR_1 - n\epsilon_4$.

Now we calculate the equivocation rate at the relay node.

$$H(W_1|Y_1, X_2) \geq H(W_1|Y_1, X_2, L_1)$$

$$= H(W_1, Y_1, L_1|X_2) - H(Y_1, L_1|X_2)$$

$$= H(W_1, L_1, Y_1, X_1|X_2) - H(X_1|W_1, L_1, Y_1, X_2) - H(Y_1, L_1|X_2)$$

$$= H(X_1|X_2) + H(W_1, L_1, Y_1|X_1, X_2)$$

$$- H(X_1|W_1, L_1, Y_1, X_2) - H(Y_1, L_1|X_2) \quad \text{(a)}$$

$$\geq H(X_1) + H(Y_1|X_1, X_2) - H(X_1|W_1, L_1, Y_1, X_2)$$

$$- H(L_1) - H(Y_1|X_2),$$

where the first term of (a) comes from the fact that $x_1, x_2$ are independent, and the fourth term comes from the fact that $l_1, x_2$ are independent.

Now, $H(L_1) = n(R_{\text{max}} - R_{\text{min}}), H(Y_1|X_1, X_2) - H(Y_1|X_2) \leq nI(X_1; Y_1|X_2) + n\delta_n$.

Given $w_1, l_1, x_2$, the relay can just choose the $x_1$ in the bin $(w_1, l_1)$ which is jointly typical with $x_2, y_1$. Since $\frac{1}{n}\log(|L_2|) \leq I(X_1; Y_1|X_2)$, we have $\Pr\{X_1 \neq X_1\} \leq \epsilon_2$.

Then based on Fano’s inequality, we have

$$\frac{1}{n} H(X_1|W_1 = w_1, L_1 = l_1, Y_1, X_2 = x_2) \leq \frac{1}{n} + \epsilon_1 I(X_1; Y_1|X_2), \quad (A.93)$$
Hence, we have

\[
\frac{1}{n} H(X_1|W_1, L_1, Y_1, X_2) \leq \epsilon_2,
\]

when \(n\) is sufficiently large.

Also, based on the encoding part, we have

\[
H(X_1) = n(R_1 + I(X_1; X_2) - R_{\text{min}}).
\]

Combining these, we get

\[
H(W_1|Y_1, X_2) \geq n(R_1 + I(X_1; X_2) - R_{\text{min}} - (R_{\text{max}} - R_{\text{min}}) - I(X_1; Y_1|X_2) - \delta_n)
\]

\[
= n(R_1 - \delta_n).
\]

The equivocation rate of the relay and the eavesdropper when \(R_{1s} \geq R_{2s}\) can be calculated similarly, with the only difference that we bound the equivocation rate of the eavesdropper by giving it \(L_1\). This case is proved.

Now, consider the case \(I(X_1; Y|X_2) > I(X_1; Y_2|X_2)\). If \(\min\{I(X_2; Y), I(X_2; Y_2)\} = I(X_2; Y)\), then we have \(\min\{I(X_2; Y), I(X_2; Y_2|X_1)\} = I(X_2; Y)\), because \(I(X_2; Y_2|X_1) > I(X_2; Y_2)\) since \(X_1, X_2\) are independent. Under this case, we only need to prove the case \(R_{s1} = [I(X_1; Y|X_2) - I(X_1; Y_2|X_2)]^+\), which can be achieved by using a scheme similar to the one developed in proving (A.84). If \(\min\{I(X_2; Y), I(X_2; Y_2)\} = I(X_2; Y_2)\), and we only need to consider \(R_{s1} = [I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - I(X_1, X_2; Y_2)]^+\), then we can use a coding/decoding scheme similar to the one developed above to show the achievability.

The claim is achieved.
A.16 Proof of Lemma 26

For any destination \( j \in D_i \), define \( Y_{ij} \) be the Bernoulli random variables, such that

\[
Y_{ij} = 1 \quad \text{if} \quad \forall k \in \Gamma_{ij}, k \notin U, \text{ and } Y_{ij} = 0 \quad \text{otherwise.}
\]

This means that if the route \( \Gamma_{ij} \) only consists of selfish nodes, then \( Y_{ij} = 1 \), otherwise, if there exists at least one altruistic nodes in the route \( \Gamma_{ij} \), \( Y_{ij} = 0 \). Since these \( \theta(N) \) altruistic nodes are distributed in the network randomly, we have

\[
P\{Y_{ij} = 1\} = \frac{\eta_{ij} - 1}{\prod_{k=0}^{\eta_{ij} - 1} \left( \frac{1 - \theta(N)/N}{N - k} \right) < (1 - \theta(N)/N)^{\eta_{ij}} \leq (1 - \theta(N)/N)^{\sqrt{N/(4c_2\sqrt{\log N})}}}
\]

due to (4.14) and the fact that \( 1 - \theta(N)/N < 1 \).

We want to bound the number of routes that only consists of selfish nodes. Let \( Y_i = \sum_{j \in D_i} Y_{ij} \), then

\[
\mathbb{E}\{Y_i\} = \mathbb{E}\left\{\sum_{j \in D_i} Y_{ij}\right\} = \sum_{j \in D_i} \mathbb{E}\{Y_{ij}\} \leq c_4 N (1 - \theta(N)/N)^{\sqrt{N/(4c_2\sqrt{\log N})}}. \quad (A.96)
\]

For any given \( i \), \( Y_{ij} \)'s are not independent. For example, let \( j_1, j_2 \in D_i \) be such that \( \Gamma_{ij_2} \) includes all the nodes in \( \Gamma_{ij_1} \) and some other nodes \( k \in \mathcal{N} \), that is \( \Gamma_{ij_1} \subset \Gamma_{ij_2} \), then if \( y_{ij_1} = 0 \), \( y_{ij_2} \) also equals to 0. The reason is that \( y_{ij_1} = 0 \) means that \( \exists k \in U, s.t. k \in \Gamma_{ij_1} \). Now, since \( \Gamma_{ij_1} \subset \Gamma_{ij_2} \), we have \( k \in \Gamma_{ij_2} \), hence, \( y_{ij_2} = 0 \). Due to this dependence, we couldn’t use the Chernoff bound for i.i.d. Bernoulli random variables directly. Instead, we use similar technique as [109] and define \( \hat{Y}_i = \sum_{j \in D_i} \hat{Y}_{ij} \), where \( \hat{Y}_{ij} \) be the i.i.d. Bernoulli random variables with

\[
\Pr(\hat{Y}_{ij} = 1) = (1 - \theta(N)/N)^{\sqrt{N/(4c_2\sqrt{\log N})}}. \quad (A.97)
\]

We know that if \( \forall y_{ij_1} = 0 \), and \( \Gamma_{ij_1} \subset \Gamma_{ij_2} \), we will have \( y_{ij_2} = 0 \), thus, it is easy to check that

\[
\mathbb{E}\{Y_i^m\} \leq \mathbb{E}\{\hat{Y}_i^m\} \quad (A.98)
\]
for any \( m > 0 \). Hence \( \mathbb{E}\{\exp(\phi Y_i)\} \leq \mathbb{E}\{\exp(\phi \hat{Y}_i)\} \) for any \( \phi > 0 \).

Let \( P(\hat{Y}_i, \delta) = \Pr\{\hat{Y}_i \geq (1 + \delta)E\{\hat{Y}_i\}\}, P(Y_i, \delta) = \Pr\{Y_i \geq (1 + \delta)E\{Y_i\}\} \), by the Chernoff bound for i.i.d. Bernoulli random variables, we have for any \( \delta > 0 \),
\[
P(\hat{Y}_i, \delta) \leq \exp(-\delta^2 \mathbb{E}\{\hat{Y}_i\}/2).
\]

From (A.98), and follows [109], we have
\[
P(Y_i, \delta) \leq P(\hat{Y}_i, \delta) \leq \exp(-\delta^2 \mathbb{E}\{\hat{Y}_i\}/2).
\]

Let \( \delta = 2\sqrt{\log N/E\{\hat{Y}_i\}} \), we have
\[
\Pr\left(\hat{Y}_i \geq E\{\hat{Y}_i\} + 2\sqrt{\log N E\{\hat{Y}_i\}}\right) \leq 1/N^2.
\]

We know that \( |\hat{D}_i| = c_4N - Y_i \), hence combining (A.97)(A.99)(A.100), we have that with probability larger than \( 1 - 1/N^2 \), there are
\[
c_4 \left( N - N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2\sqrt{\log N})} \right) + o\left( N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2\sqrt{\log N})} \right)
\]
destinations in \( \mathcal{D}_i \), whose routes from node \( i \) includes at least one altruistic node.

Using union bound, we have, for \( \forall i \in \mathcal{N} \)
\[
|\hat{D}_i| > c_4 \left( N - N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2\sqrt{\log N})} \right) + o\left( N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2\sqrt{\log N})} \right),
\]
with probability larger than \( 1 - 1/N \).

In particular, if \( \theta(N) = \omega(\sqrt{N \log N}) \),
\[
0 \leq \lim_{N \to \infty} \frac{\mathbb{E}\{Y_i\}}{N} \leq \lim_{N \to \infty} \left[(1 - \theta(N)/N)^{N/\theta(N)}\right]^{\theta(N)/(4c_2\sqrt{N \log N})} = 0,
\]
since \( \lim_{N \to \infty} (1 - \theta(N)/N)^{N/\theta(N)} = 1/e \) and \( \theta(N) = \omega(\sqrt{N \log N}) \). Hence
\[
N(1 - \theta(N)/N)^{\sqrt{N}/(4c_2\sqrt{\log N})} = o(N),
\]
the claim is proved.

141
A.17 Proof of Theorem 27

If \( p = 1 \), then

\[
E_{i,tr} = \frac{1}{N-1} \sum_{j \neq i} c_1 d^\gamma_{ij}, E_{i,re} = \mathbb{E}\left\{ \sum_{\Gamma \in \Lambda_i} c_1 d^\gamma_{i\Gamma(i)} \right\},
\]

since all its relay request will be accepted.

Hence the cost under this point is

\[
C_{i,co} = E_{i,tr} + E_{i,re} \leq c_1 c_2^2 \log^{\gamma/2} N + c_1 N c_2^2 \log^{\gamma/2} N = \Theta(N \log^{\gamma/2} N), \tag{A.102}
\]
due to the bound in (4.13), and at each frame, there are at most \( N \) relay requests for node \( i \) at each frame.

On the other hand if node \( i \) changes its strategy to \( p_i < 1 \), then

\[
E_{i,tr}(p_i) = \frac{1}{N-1} \left( \sum_{j \in \Psi_i \backslash \hat{\Psi}_i} c_1 d^\gamma_{ij} + \sum_{j \in \Psi_i} p_i^{f(\Gamma_{ij})} c_1 d^\gamma_{ij} + \sum_{j \in \Psi_i} (1 - p_i^{f(\Gamma_{ij})}) c_1 d^\gamma_{ij} \right),
\]

where \( f(\Gamma_{ij}) \) is the number of altruistic nodes in the route \( \Gamma_{ij} \), and

\[
E_{i,re}(p_i) = \mathbb{E}\left\{ p_i \sum_{\Gamma \in \Lambda_i} c_1 d^\gamma_{i\Gamma(i)} \right\}, \tag{A.103}
\]

Hence for \( \forall \epsilon > 0, p_i < 1 - \epsilon \), we have

\[
C_i(p_i) = E_{i,tr}(p_i) + E_{i,re}(p_i) \\
\geq \frac{1}{N-1} \sum_{j \in \Psi_i} (1 - p_i^{f(\Gamma_{ij})}) c_1 d^\gamma_{ij} \tag{A.104} \\
\geq \frac{1}{N-1} \sum_{j \in \hat{\Omega}_i} (1 - p_i^{f(\Gamma_{ij})}) c_1 d^\gamma_{ij} \tag{A.105} \\
\geq \epsilon c_1 \frac{c_2 N}{N-1} N^{\gamma/2} / 4^\gamma \tag{A.106} \\
= \Theta(N^{\gamma/2}) > C_{i,co}. \tag{A.107}
\]
(A.104) is true, since we only consider the energy spent on direct transmission and ignore the energy spent on helping. (A.105) is true, since \( \hat{D}_i \subseteq \hat{\Psi}_i \). (A.106) is true, because:

1) \(|\hat{D}_i| \geq c_5 N\) (see lemma 26), 2) \( \forall j \in \hat{D}_i, d_{ij} \geq \frac{\sqrt{N}}{4}\) (definition), and 3) \( 1 - p_i f(\Gamma_{ij}) \geq 1 - p_i \geq \epsilon\), since \( f(\Gamma_{ij}) \geq 1, \forall j \in \hat{D}_i \). (A.107) is true when \( N \) is large, since \( \gamma > 2 \).

Also

\[
\frac{\partial C_i}{\partial p_i} \bigg|_{p_i=1} = -\frac{1}{N-1} \sum_{j \in \hat{\Psi}_i} f(\Gamma_{ij}) c_1 \left( d_{ij}^2 - d_{ii'}^2(j) \right) + \mathbb{E} \left\{ \sum_{\Gamma \in \Lambda_i} c_1 d_{i\Gamma}^2(\iota) \right\}
\leq -\frac{1}{N-1} \sum_{j \in \hat{D}_i} c_1 \left( d_{ij}^2 - d_{ii'}^2(j) \right) + c_1 N c_2^2 \log^{\gamma/2} N \tag{A.108}
\leq -\frac{c_1 c_5 N}{N-1} \left( N^{\gamma/2}/4^{\gamma} - c_2^2 \log^{\gamma/2} N \right) + c_1 N c_2^2 \log^{\gamma/2} N \tag{A.109}
\leq 0, \tag{A.110}
\]

for sufficiently large \( N \).

(A.108) is true since: 1) \( \hat{D}_i \subseteq \hat{\Psi}_i \), 2) \( \forall j \in \hat{\Psi}_i, f(\Gamma_{ij}) \geq 1 \), and 3) at each frame, there are at most \( N \) relay requests for each node. (A.109) is true, because of lemma 26, and \( \forall j \in \hat{D}_i, d_{ij} \geq \sqrt{N}/4 \). (A.110) is true since \( \gamma > 2 \). (A.110) is true for all the nodes in the network w.h.p.

Hence no user will deviate from \( p_i = 1 \), full-cooperation is a nash-equilibrium. At this equilibrium, the average energy per packet scales the same as the cooperative network, which is \( \Theta\left( \sqrt{N} \log^{(\gamma-1)/2}(N) \right) \) given by [23, 24].

143
BIBLIOGRAPHY


