Lehigh University
Lehigh Preserve

Theses and Dissertations

2013

Multiuser Wireless Networks: Cooperation and Physical-Layer Security

KYATSANDRA NAGANANDA
Lehigh University

Follow this and additional works at: http://preserve.lehigh.edu/etd

Recommended Citation
NAGANANDA, KYATSANDRA, "Multiuser Wireless Networks: Cooperation and Physical-Layer Security" (2013). *Theses and Dissertations.* Paper 1261.

This Dissertation is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

MULTIUSER WIRELESS NETWORKS:

COOPERATION AND PHYSICAL-LAYER SECURITY

by

KYATSANDRA G NAGANANDA

A Dissertation

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

ELECTRICAL ENGINEERING

LEHIGH UNIVERSITY

JANUARY 2013

Copyright

Kyatsandra G Nagananda, ©2013

All rights reserved.

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Kyatsandra G Nagananda
Multiuser Wireless Networks: Cooperation and Physical-Layer Security

---

Date

Shalinee Kishore, Dissertation Director, Chair

Accepted Date

Committee Members

Shalinee Kishore

Tiffany Jing Li

Parv Venkitasubramaniam

Eugene Perevalov

Acknowledgements

As the formal part of my education comes to an end, I feel saddened but equally proud to have been privileged to belong to a community that stands on the pillars of academic tradition and creativity. So, time has come for me to acknowledge the people who have played significant roles in not only helping me achieve this, but who have played a greater role in my development as a person.

I begin by thanking my mentor and guide Professor Shalinee Kishore who was instrumental in bringing me to Lehigh University in the fall of 2009. She has not only been superlatively kind but also an inspiring mentor. By allowing me the freedom to pursue research problems, she emphasized the importance of treating research as a way of life. I thank her for her never ending patience, and for her calming influence. Last, but not the least, I thank her for the financial support without which this work would never have been possible.

I would like to thank members of my dissertation committee, Professor Tiffany Jing Li, Professor Parv Venkitasubramaniam and Professor Eugene Perevalov, for their constructive criticism and invaluable inputs which shaped this work. Special thanks go to Professor Venkitasubramaniam for offering several courses that influenced my thinking, and for his care and concern towards my personal well-being. Thanks also go to Professor Rick Blum, whose curiosity and eagerness to pursue research problems deeply impacted my thinking. I wish to thank Professor Meghanad Wagh, with whom I have discussed every possible topic under the sun.

I would like to take this opportunity to express my gratitude to Professor G.

V. Anand at the Indian Institute of Science (IISc), who not only taught me the art of technical writing, but also to present research results in an elegant manner. Sincere thanks go to Professor Chandra R. Murthy at IISc, who took me into his team at a crucial moment of my research career. For all the inadequacies that I have perceived in myself, Chandra could accept me for who I was. I will always remember Chandra as a devoted researcher, but more importantly for his magnanimous personality.

On a long journey, one needs a catalyst to keep the process well-tuned. For me, this catalyst came in the form of a text book on information theory by Professor Raymond Yeung at the Chinese University of Hong Kong, which pulled me deeper into the beautiful world of information theory. The joy Prof. Yeung brings into readers of his book, his grace and innate humility have been an inspiration for me. An icing on the cake came when Prof. Yeung agreed to host me as a postdoctoral research fellow in his group. My sincere thanks go to Prof. Yeung for being a role-model for young researchers in this field.

This work would not have been possible without an incredibly strong support network around me, and I would like to take this opportunity to thank my family and friends. I begin by expressing my deepest gratitude to my maternal and paternal grandparents, Sri. C. Gopal Rao, Smt. Ratna, Sri. K. Seetharamiah and Smt. Susheelamma, for imbibing fundamental values and perspective in the younger generation. I thank my maternal uncles, Dr. C. Raghavendra Rao and Dr. C. Gundu Rao, and their families, whose professional accomplishments despite numerous struggles still remain benchmarks for my future goals. I would like to also thank my paternal uncles, Dr. K.S. Janakiram and Dr. K.S. Arun Kumar, and their families, who have placed education above every possible distraction life offers. Thanks are also due to my innumerable cousins and the latest entries into the family. A note of thanks to Mr. Vijendra Kuroodi and his family for encouraging me during various stages of the graduate program. I can never repay the love and affection my extended family have showered on me.

My life would be incomplete without friends. My childhood friends Sriram,

Srinath, Sandeep, Mukunda, and my freinds during the undergraduate program have all had a significant role in my life and my development as a person. Special thanks go to Sandeep and Srinath, for having helped me to settle in when I first came up to Lehigh; and for the endless and mindless conversations that we have had, and continue to have. May peace be with them and their young families.

A note of thanks to all my friends at Lehigh who have been like a family to me: Shubham, Vallisha, Raghu Srinivas, Anand, Abhishek, Shah Rukh, Akshaya, Sonam, Ranjan, Raja Ramakrishna, and Sambhawa. A special note of thanks to Jeff Roquen, who helped keep me sane and grounded at all times. Thanks also go to Chen Chen and Gary Xiong for their friendship and positive influence. A mighty thanks to my friends at IISc: Abhay, Bharath, Sanjeev, Chandru, Deepa, Harshan, Ranjitha, Karthik, Krishna Chaithanya, Ganesan, Partha, Venu, and Nagesha. The friendship and experience I got from them during my tenure at IISc was enriching and thoroughly enjoyable. I knew I could always look to them for help, which they selflessly offered. May peace be with all of them.

I thank all the janitors, security personal, and the building management and library staff at Lehigh and IISc, for setting up the environment conducive for students to pursue their dreams. It amazed me to see them put their heart and soul into work, be it in the freezing winters of Bethlehem, or dusty summers of Bangalore. I have taken so many things in life for granted, but was humbled to see them work hard for these very things. To a large extent, they are my real heros.

I am truly grateful to the Almighty for blessing me with wonderful parents, giving me the best of education, the best of health and the mind to achieve whatever little I have been able to accomplish. I thank Him for putting me in the fortunate section, compared to a large fraction of the population who are less fortunate.

Finally, I would like to mention that I have no formal control on the prospective use of results presented in this thesis. However, I sincerely wish and hope that potential applications resulting out this work are directed towards peaceful purposes.

Dedicated to my parents,

Sri. K. S. Gurukumar and Smt. C. Sudha,

my love and life in this long and weirdly wonderful arduous journey. Without them this journey would never have begun. In the end, whatever I do is for them.

Contents

List of Tables

List of Figures

Abstract

With rapid advances in computational algorithms and silicon technology, there has been a dramatic rise in research endeavors in all areas of wireless communications - physical layer, medium access methods, networks and systems. Information theory provides governing laws for communications systems by establishing fundamental limitations to reliable communications, and aids in leveraging this understanding into engineering solutions for wireless networks. Recently, with increasing emphasis on efficient utilization of radio-frequency spectrum and growing interest in providing wireless services with higher data rates, cooperative-communications has been proposed as a key enabling technology for next generation wireless networks. Although user-cooperation has proven benefits, the broadcast nature of the wireless medium exposes problems related to information security, by facilitating malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, *etc.* In this work, we analyze the impact of cooperation and information security on the fundamental performance limits of four multiuser networks: (i) interference networks; (ii) relay networks; (iii) broadcast networks; and (iv) Z-networks.

First, we consider a three-user interference channel to obtain novel insights into the role of cooperation and interference management on its performance limits. We consider three cooperation schemes - (i) cumulative message-sharing; (ii) primary-only message-sharing; and (iii) cognitive-only message-sharing, and employ different rate-splitting policies for interference management. As a case study, we consider the Gaussian channels, present several corollaries to enlarge the rate regions and derive outer bounds to obtain more insights.

Next, we explore the performance limits of the joint problem of cooperation and information security over a four node discrete memoryless relay network comprising a sender-destination pair, a relay node and an unauthorized eavesdropper. We consider two communication scenarios: In the first scenario, the relay aids

transmissions from the source to the destination. In the second scenario, the relay is considered to be malicious, constraining the source to keep its message confidential from the relay node. In both scenarios, the relay is (i) opportunistic in the sense that, it utilizes the communication opportunity to transmit its own message to the destination and (ii) constrained to secure its communication from the external eavesdropper.

Then, we derive the fundamental limits of three classes of two-user state-dependent discrete memoryless broadcast channels, with noncausal side-information at the encoder. The first class of channels comprises a sender broadcasting two independent messages to two non-cooperating receivers; for channels of the second class, each receiver is given the message it need not decode; and the third class comprises channels where the sender is constrained to keep each message confidential from the unintended receiver.

Lastly, we derive inner and outer bounds on the capacity region of the cognitive Z-channel, comprising two sender-receiver pairs with no cross-talk channel gain between one of the sender-receiver pairs. The non-cognitive sender has two messages, each message transmitted to the intended destination. The cognitive transmitter has one message intended to its pairing receiver; in addition, it has noncausal knowledge of the messages and the corresponding codewords of the non-cognitive sender.

Chapter 1

Introduction

Existing spectrum allocation policies have led to imbalanced use of the radio frequency (RF) spectrum. For example, there is increased congestion in certain frequency bands in or near densely populated urban centers, while there exist large amounts of unused bands in rural areas. To deal with such examples of inefficient spectrum usage, several techniques have been proposed, including in-band sharing, the use of low power radios, multi-modal and ultra-wideband radios, logistical changes to licensing policies for spectrum leasing and trading, and the clever use of unlicensed spectrum. In the U.S., the Federal Communications Commissions (FCC) Spectrum Policy Task Force has been actively involved in optimizing the use of RF spectrum, avoiding and minimizing interference, and in creating a framework for the design of short and long range frequency allocations to ensure greater spectrum efficiency and flexibility.

One popular mechanism to improve the efficiency of the RF spectrum is coordination or cooperation among users, which has gained momentum since the advent of cognitive radio (CR) technology [1], [2]. From an engineering point of view, CR technology broadly falls into the category of user-cooperation diversity [3] - [5], where certain communicating nodes exchange information in a particular manner so as to improve the spectral efficiency of the overall communications system. In the CR context, cognitive terminals rely on the broadcast nature of wireless medium to cooperatively accommodate transmissions from non-cognitive or pri-

mary users. Although the broadcast medium, when accessed by such cognitive, cooperative terminals, enables improved spectral efficiency, it also exposes problems related to information security. That is, the broadcast nature of wireless networks facilitates malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, *etc*.

Motivated by the growing demand for improving spectrum efficiency and providing information security in wireless networks, we consider in this work an *information-theoretic* viewpoint of these two issues in three basic wireless channels, which form building blocks for higher order wireless networks. Specifically, we derive inner and outer bounds on the capacity region of (i) interference channels; (ii) relay channels; and (iii) broadcast channels, under user-cooperation and information-security constraints. We note that, in the information theory literature, the terms cooperation and cognition have been used interchangeably, while confidentiality is also referred to as information-theoretic/wireless physical-layer security [6].

1.1 Interference channels with transmitter cooperation

In this work, we consider the case of *three-user* CR interference channels, where two (or one) CRs and one (or two) primary user(s) communicate with three respective receivers. The transmitters are allowed to cooperate in a unidirectional manner via noncausal message sharing mechanism. The following points summarize the theme of this work:

1. *Message-sharing mechanism:* The first interesting observation we make is that there are multiple ways in which the two-user CR channel can be extended to the three user CR channel, depending on the message sharing mechanism employed. We consider three intuitive schemes, namely (i) cumulative message sharing (CuMS); (ii) primary-only message sharing (PrMS); and (iii) cognitive-only message sharing (CoMS).

2. *Interference management:* Growing network-size presents issues related to interference management. To deal with interference in this three-user channel, we use rate-splitting, which was first reported in [7] to enlarge the achievable rate region for the classical two-user interference channel. The main idea behind rate-splitting is to encode part of the message at a possibly low rate, so that an unintended receiver can decode the interference caused to it by performing joint decoding of part of the interference and its own data. To highlight the benefits and drawbacks of rate-splitting, we define five cognitive channel models, two each for CuMS and PrMS, and one for CoMS, which correspond to different rate-splitting strategies. The different types of message-sharing mechanisms and rate-splitting strategies will be made precise in the next section.

3. *Achievable rate regions:* We derive an achievable rate region for each of the five models by considering first the discrete memoryless version of the channel. To this end, we employ the technique of combining Gel'fand-Pinsker's (GP) binning principle [8] and superposition coding [9]. As a result, we illustrate the generality of the techniques employed here, and provide useful and novel insights into the rate regions and their characterization.

4. *Gaussian channel case:* We specialize the achievable rate regions to the important special case of Gaussian CR channel; this enables comparisons of the different rate regions both analytically and through simulations. It also leads to the development of corollaries that help in enlarging the achievable rate regions in the Gaussian case. Finally, we derive some outer bounds to measure the optimality of the proposed coding scheme for our channel models. Inner bounds are derived using dirty-paper coding [10], while results from the multiple antenna broadcast channels [11] and duality between multiple access and broadcast channels [12] are exploited to derive outer bounds.

The results of this work have appeared in [13] - [16]. Note that, noncausal message-sharing is a very strong assumption leading to skepticism towards the practical realization of such networks. However, the results presented here not only provide useful performance limits, but also throw more light on the possibilities of various cognition schemes and related interference management issues.

1.2 Opportunistic relay channels

In this work, we consider a four node wireless relay network comprising a sender-destination pair, a relay node and an unauthorized external eavesdropper. We consider two communication scenarios:

1. In the first scenario, the relay aids transmissions from the source to the destination.

2. In the second scenario, the relay is considered to be malicious, constraining the source to keep its message confidential from the relay node.

In both scenarios, the relay is (i) *opportunistic* in the sense that, it utilizes the communication opportunity to transmit its own message to the destination and (ii) constrained to secure its communication from the external eavesdropper. Furthermore, we assume the eavesdropper to be geographically located outside the transmission range, and hence remains oblivious to the transmissions of the main sender. This channel model is practically well motivated and provides a basis to jointly address the issues of cooperation, cognition and information security. Also note that, in the second scenario, the relay node tries to decode the message of the main sender (although unsuccessfully); therefore, the model conforms to the classical relay setting.

We state and prove channel coding theorems, and derive a set of achievable rates for these two communication scenarios by considering the discrete memoryless model of the channel. Standard techniques - block Markov superposition coding [17], binning [18], backward decoding [19] and simultaneous decoding [7]

- are employed to prove the coding theorem. Stochastic encoders [20] are used to satisfy confidentiality constraints. Outer bounds on the capacity region are derived using auxiliary random variables for single-letter characterization. We also discuss some of the advantages and drawbacks of our coding strategy in comparison to those in the existing literature, which provides interesting insights into the relative merits of the methods employed in this work for obtaining the capacity bounds. The results of this work can be found in [21].

1.3 State-dependent broadcast channels

In this work, our main goal is to analyze the impact of side information and confidentiality constraints on the information theoretic performance limits of broadcast channels (BC). To this end, we derive capacity bounds on the following three classes of two-user discrete memoryless BC, with noncausal side-information, for *e.g.*, fading in the wireless medium, interference caused by neighboring nodes in the network, *etc.* at the encoder:

1. Class I: A sender broadcasts two independent messages to two non-cooperating receivers. An inner bound for this class of channels was derived by Steinberg and Shamai in [22], by extending Marton's achievability scheme [23] to include noncausal side-information at the encoder. However, in this work, we extend Marton's achievability scheme and use results from the second moment method [24] to derive an inner bound. Our proof is simpler and generalizes well to derive an inner bound for channels of Class III (described below). An outer bound is derived employing the procedure used to prove the converse theorem for GP's channels with random parameters [8]. The bounds are shown to be tight for individual rate constraints, but can be improved upon for the sum-rate. An example for Class I channels is a base station transmitting to two mobile receivers, in the presence of *a priori* known interference from a transmitter located in the vicinity of the base station.

2. Class II: A sender broadcasts two independent messages to two receivers, with each receiver having *a priori* knowledge of the message it need not decode. An example of this scenario is full-duplex communications between two nodes, aided by a relay. The relay node broadcasts the messages to the terminals, with each terminal knowing its own message. Class II channels are also addressed in [25], where an inner bound matching our result was derived; however, there was only an outline for deriving an outer bound. In this work, an inner bound is derived by extending the method proposed by Kramer and Shamai in [26], to include transmitter side-information for BC where each receiver has knowledge of the other's message; our proof is much simpler than the one presented in [25]. Furthermore, our outer bounds are derived using arguments from the proof of converse for GP's channel which is within a fixed gap away from the achievable region, where the gap is independent of the distribution characterizing this class of channels.

3. Class III: A sender broadcasts two independent messages to two receivers, such that each message is kept confidential from the unintended receiver. The achievability theorem is proved by employing the technique used to derive an inner bound for Class I channels, in conjunction with a stochastic encoder to satisfy confidentiality constraints. The technique to derive outer bounds hinges on the confidentiality requirements. We also derive a genie-aided outer bound, where a genie gives a receiver the message it need not decode, while the other receiver computes the equivocation treating this message as side-information. We also suggest a tighter outer bound for the sum rate of this class of channels. As an example for this class of channels, we can extend the example considered for Class I channels, with the additional constraint of keeping each message ignorant from the unintended receiver.

For all the three classes of channels, Csiszár's sum identity [27] plays a central role in establishing the capacity outer bounds. The results in this work demonstrate that, owing to rate-penalties for dealing with side-information and satisfy-

ing confidentiality constraints, the rate region for channels of Class III is smaller than that for Class I, which is further smaller compared to the classical two-user BC. Note that, the comparisons are presented primarily to illustrate the role of side information and secrecy constraints on the achievable rates. However, since each model is characterized by its own probability distribution, these comparisons should be made with caution. The results of this work can be found in [28], [29].

1.4 Z-channels with cooperation

In this work, we consider the cognitive Z-channel which is a combination of the classic multiple access channel and the two-user broadcast channel. Essentially, there are two sender-receiver pairs with the first sender having a message intended to its pairing receiver; the second sender has two messages - one intended to the receiver of the first transmitter, while the second message is intended to its pairing receiver. Furthermore, in our model, the encoder of the first transmitter (cognitive) has noncausal knowledge of the message sets and the corresponding codewords of the second (non-cognitive). We consider both the discrete memoryless and the Gaussian versions of this channel model to derive lower and upper bounds on its capacity region.

For the discrete memoryless version of the channel we employ Marton's achievability scheme for the classic two-user broadcast channel [23] at the non-cognitive encoder. At the cognitive encoder, we introduce a generalization of the technique used to prove the coding theorem for the Gel'fand-Pinsker channel with noncausal side-information at the encoder [8]. Outer bounds are derived by employing the technique used to prove the converse theorem for channels with side-information in conjunction with Nair-El Gamal's technique used to obtain the outer bounds for the classic two-user broadcast channels [30]. For the Gaussian channel case, the above mentioned techniques are translated to the corresponding Gaussian model using standard transformations. The resulting achievable regions and outer bounds are then numerically evaluated and plotted to reveal interesting

observations.

1.5 Organization of the thesis

The remainder of the thesis is organized as follows:

1. In Chapter 2, we introduce the three-user cognitive interference channel, and derive inner and outer bounds for three different message-sharing schemes under various rate-splitting scenarios.

2. In Chapter 3, we consider the opportunistic relay channels and derive capacity bounds for secure and reliable communications over the discrete memoryless model of the channel.

3. In Chapter 4, we present results for three different models of state-dependent broadcast channels. Here again, we consider the discrete memoryless version of the channel.

4. In Chapter 5, we consider the Z-channel with degraded message sets, and present inner and outer bounds on the capacity region of this channel by considering both the discrete memoryless and Gaussian channel models.

5. In Chapter 6, we propose two problems dealing with cooperation and physical-layer security over wireless networks, and relegate this to future work.

Chapter 2

Interference channels with transmitter cooperation

In this chapter, achievable rate regions and outer bounds are derived for three-user interference channels where the transmitters cooperate in a unidirectional manner via a noncausal message-sharing mechanism. The three-user channel facilitates different ways of message-sharing between the primary and secondary (or cognitive) transmitters. Three natural extensions of unidirectional message-sharing from two users to three users are introduced: (i) Cumulative message sharing; (ii) primary-only message sharing; and (iii) cognitive-only message sharing. To emphasize the notion of interference management, channels are classified based on different rate-splitting strategies at the transmitters. The techniques of superposition coding and Gel'fand-Pinsker's binning are employed to derive an achievable rate region for each of the cognitive interference channels. The results are specialized to the Gaussian channel, which enables a visual comparison of the achievable rate regions through simulations and help us achieve some additional rate points under extreme assumptions. We also provide key insights into the role of rate-splitting at the transmitters as an aid to better interference management at the receivers.

2.1 Related work

Besides identifying the underlay, overlay and interweave CR network paradigms mentioned above, [31] explored some of the fundamental capacity limits and associated transmission strategies for CR wireless networks. In [32], [33], an achievable rate region was derived for the two-user discrete memoryless *genie-aided* CR channel. An outer bound was proposed for the corresponding Gaussian channel by allowing *bidirectional* cooperation between the transmitters in a noncausal manner, resulting in a multiple antenna (MIMO) broadcast channel whose capacity region is well-known [11]. In [34], terms like *dumb* and *smart* antennas were introduced to refer to primary and cognitive senders, respectively. Inner and outer bounds were derived for the general discrete memoryless channel, along with capacity results for some special cases. The capacity region was also derived for the Gaussian CR channel under a *weak interference* assumption. In [35], the Gaussian CR channel was presented and capacity results were derived for the low interference regime where the primary receiver uses single-user decoding to achieve the capacity, while in the high interference regime joint code design at the two transmitters and multiuser decoding at the primary receiver was shown to be optimal to maximize the jointly achievable rates for the primary and cognitive users. In [36], an achievable rate region was derived for the two-user CR interference channel, where only the CR transmitter employs rate-splitting. In the high interference regime, the region presented in [36] subsumes the ones derived in [34] and [35]. In [37], the capacity regions were established for several two-sender, two-receiver channels with partial transmitter cooperation: compound multiple access channels (MAC) with common information; compound MACs with conferencing; and interference channels with unidirectional cooperation under strong interference assumptions. Capacity bounds for two-user interference channels with cognitive and partially cognitive transmitters were reported in [38] - [43], while [44] - [47] concerns interference channels with common information.

The most recent results on the cognitive channel include [48] - [50], where a

new achievable rate region for the two-user CR channel has been derived encompassing all the previous ones, with capacity results for a few classes of channels. The above mentioned references employ a combination of the coding scheme proposed by Han and Kobayashi for the interference channel [7], the binning technique proposed by Gel'fand and Pinsker (GP) for coding over channels with random parameters [8], superposition coding proposed first for the broadcast channel [9] and *dirty-paper* coding [10] for Gaussian channels with noncausal interference at the encoder.

2.2 System model and preliminaries

The three-user discrete memoryless cognitive interference channel is described by the tuple $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \mathcal{P}, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3)$, where the notation is as follows. For $k = 1, 2, 3$,

1. senders and receivers are denoted by $\mathcal{S}_k$ and $\mathcal{R}_k$, respectively;
2. finite sets $\mathcal{X}_k$ and $\mathcal{Y}_k$ denote the channel input and output alphabets, respectively;
3. random variables $X_k \in \mathcal{X}_k$ and $Y_k \in \mathcal{Y}_k$ are the inputs and outputs of the channel respectively; and
4. $\mathcal{P}$ denotes the finite set of conditional probabilities $p\left(y_1, y_2, y_3 | x_1, x_2, x_3\right)$, when $(x_1, x_2, x_3) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$ are transmitted and $(y_1, y_2, y_3) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ are obtained by the receivers.

The channels are assumed to be memoryless. As in the classical three-user interference channel, the messages at the senders are denoted $m_k \in \mathcal{M}_k = \{1, \ldots, M_k\}$; $\mathcal{M}_k$ being a finite set with M_k elements. The messages are assumed to be independently generated.

2.2.1 Message-Sharing Mechanisms

We now describe the message-sharing mechanisms considered in this work.

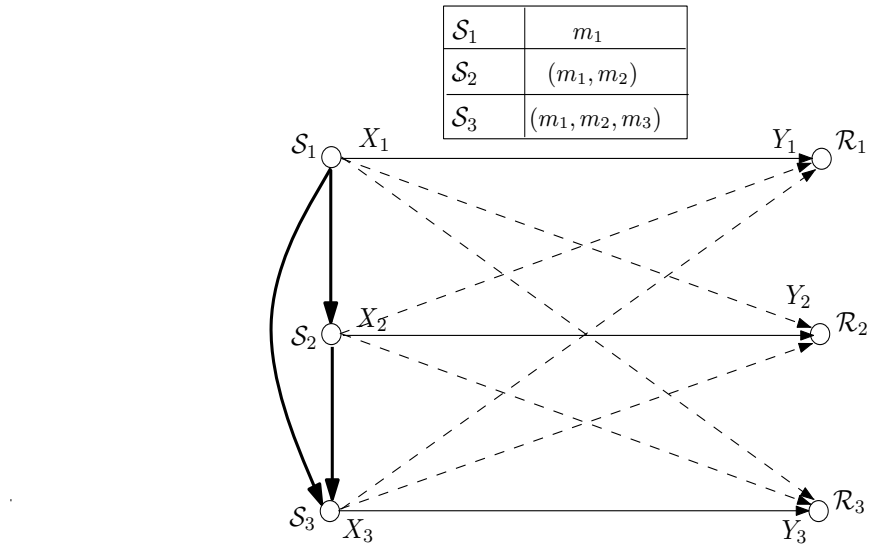

Figure 2.1: Three-user CR channel with CuMS

1. In the case of cumulative message-sharing (CuMS), sender $\mathcal{S}_2$ has noncausal knowledge of the message m_1 and the corresponding codewords of the primary sender, $\mathcal{S}_1$. Sender $\mathcal{S}_3$ has noncausal knowledge of the message m_1 of the primary transmitter as well as the message m_2 of $\mathcal{S}_2$, and their respective codewords. A schematic of CuMS is shown in Fig. 2.1.

2. In the case of primary-only message-sharing (PrMS), senders $\mathcal{S}_2$ and $\mathcal{S}_3$ have noncausal knowledge of the message m_1 and the corresponding codewords of the primary sender, $\mathcal{S}_1$. There is no message-sharing mechanism between $\mathcal{S}_2$ and $\mathcal{S}_3$ themselves. See Fig. 2.2 for a channel schematic.

3. In the case of cognitive-only message-sharing (CoMS), sender $\mathcal{S}_3$ has noncausal knowledge of messages m_1 and m_2, and the corresponding codewords of senders, $\mathcal{S}_1$ and $\mathcal{S}_2$. There is no message-sharing mechanism between the $\mathcal{S}_1$ and $\mathcal{S}_2$. A channel schematic for CoMS is shown in Fig. 2.3.

An $\left(M_1, M_2, M_3, n, P_e^{(n)}\right)$ code exists for these channels, if there exists the following encoding functions:

$$f_1 : \mathcal{M}_1 \mapsto \mathcal{X}_1^n, \qquad f_1' : \mathcal{M}_1 \mapsto \mathcal{X}_1^n, \qquad f_1'' : \mathcal{M}_1 \mapsto \mathcal{X}_1^n$$

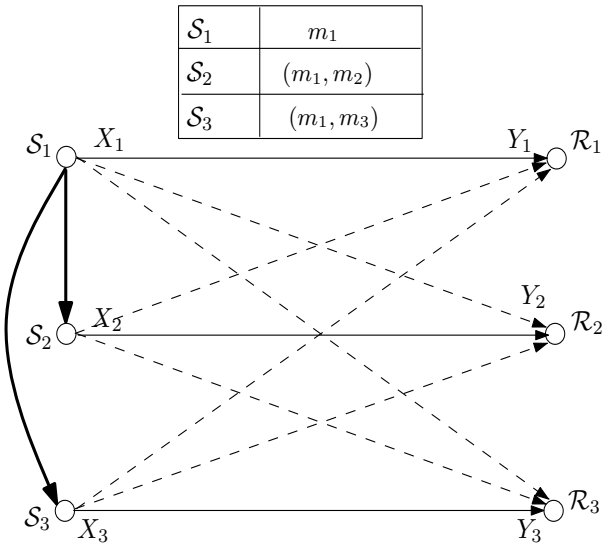

Figure 2.2: Three-user CR channel with PrMS

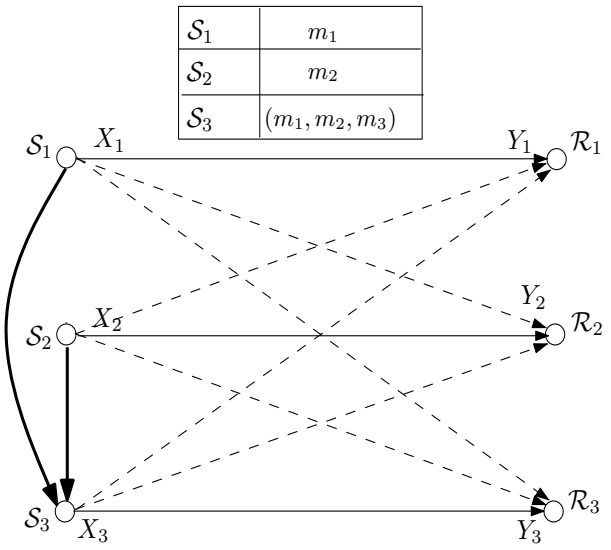

Figure 2.3: Three-user CR channel with CoMS

$$f_2:\ \mathcal{M}_1 \times \mathcal{M}_2 \mapsto \mathcal{X}_2^n, \quad f_2':\ \mathcal{M}_1 \times \mathcal{M}_2 \mapsto \mathcal{X}_2^n \quad f_2'':\ \mathcal{M}_2 \mapsto \mathcal{X}_2^n$$
$$f_3:\ \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3 \mapsto \mathcal{X}_3^n, \quad f_3':\ \mathcal{M}_1 \times \mathcal{M}_3 \mapsto \mathcal{X}_3^n, \quad f_3'':\ \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3 \mapsto \mathcal{X}_3^n$$

and the following decoding functions, for $k = 1, 2, 3$:

$$g_k:\ \mathcal{Y}_k^n \mapsto \mathcal{M}_k, \quad g_k':\ \mathcal{Y}_k^n \mapsto \mathcal{M}_k, \quad g_k'':\ \mathcal{Y}_k^n \mapsto \mathcal{M}_k,$$

such that the decoding error probability $\max\left\{P_{e,1}^{(n)}, P_{e,2}^{(n)}, P_{e,3}^{(n)}\right\}$ is $\leq P_e^{(n)}$. $P_{e,k}^{(n)}$ is the average probability of decoding error computed using:

$$P_{e,k}^{(n)} = \frac{1}{M_1 M_2 M_3} \sum_{m_1,m_2,m_3} p\left[\hat{m}_k \neq m_k | \left(m_1, m_2, m_3\right) \text{ sent}\right]; k = 1, 2, 3.$$

f_k (or g_k) correspond to the encoders (or decoders) used by channels with CuMS, f_k' (or g_k') correspond to the encoders (or decoders) used by channels with PrMS and f_k'' (or g_k'') correspond to the encoders (or decoders) used by channels with CoMS.

We define two channels denoted $\mathcal{C}_{\text{CuMS}}^t$, two channels denoted $\mathcal{C}_{\text{PrMS}}^t$ and one channel denoted $\mathcal{C}_{\text{CoMS}}$; $t = 1, 2$. A non-negative rate triple (R_1, R_2, R_3) is achievable for each of the channels, if there exists a sequence of $\left(2^{\lceil nR_1 \rceil}, 2^{\lceil nR_2 \rceil}, 2^{\lceil nR_3 \rceil}, n, P_e^{(n)}\right)$ codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. The capacity region for the channels is the closure of the set of all achievable rate triples (R_1, R_2, R_3).

2.2.2 Rate-Splitting Strategies

In [7], it has been shown that the achievable rate region for the classical two-user interference channel can be enlarged by rate-splitting. Here, each transmitter splits its message into a *public part* and *private part*. The former is decodable at all receivers, and hence unintended receivers can use it to cancel part of the interference. The latter is only decodable at the intended receivers. In the three-user scenario, however, many more rate-splitting strategies exist compared to the two-user case. For example, sender $\mathcal{S}_1$ can perform rate-splitting in one of

the following four ways: (i) it can encode a part of its message such that both unintended receivers, $\mathcal{R}_2$ and $\mathcal{R}_3$, can decode the sub-message; (ii) encode a part of the message such that $\mathcal{R}_2$ can decode it but not $\mathcal{R}_3$; (iii) encode a part of the message such that $\mathcal{R}_3$ can decode it but not $\mathcal{R}_2$; and finally, (iv) encode in a manner such that the sub-message is not decodable at either $\mathcal{R}_2$ or $\mathcal{R}_3$ (i.e., decodable only at the $\mathcal{R}_1$). In this work, we consider the following rate-splitting strategies:

1. In $\mathcal{C}^1_{\text{CuMS}}$ and $\mathcal{C}^1_{\text{PrMS}}$, the senders encode part of their respective messages at a rate such that it can be reliably decoded by all receivers. The other part of the message is encoded to ensure that it is decodable at the intended or pairing receiver. The other receivers do not attempt do decode this part of the message.

2. In $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$, one part of the message is encoded such that the intended receiver can decode it, and the other receivers treat it as noise. The other part is encoded such that it can be decoded at the intended receiver and the receiver $\mathcal{R}_1$, and the unintended receiver treats it as noise.

3. In $\mathcal{C}_{\text{CoMS}}$, sender $\mathcal{S}_3$ encodes one part of the message at a rate such that all receivers can decode it, while the other part is encoded at a rate such that it can be decoded at its pairing receiver, $\mathcal{R}_3$ (and the other receivers treat it as noise). There is no rate-splitting at $\mathcal{S}_1$ and $\mathcal{S}_2$.

Note that, regardless of the manner in which rate-splitting is performed, $\mathcal{R}_t$ should always be able to reliably decode the codewords from $\mathcal{S}_t$, $t = 1, 2, 3$. We consider the above described rate-splitting strategies for the following reasons:

1. To better understand the role of rate-splitting as a mechanism for interference management at the receivers, especially with growing network-size (for example, from two-user to three-user CR channels).

2. To demonstrate the increasing difficulty in characterizing theoretical limits with the number of rate-splits at the encoder. The number of probability-

of-error terms at a decoder increases exponentially with the number of rate-splits at the encoder (both pairing and non-pairing encoders). Due to this, characterizing the rate region becomes cumbersome, leading also to difficulties in quantifying the performance through simulation results.

3. To demonstrate the effect of reduction in network-size on the rate region characterization. Specifically, we show that the achievable rate regions for the three-user CR channel reduces to known results in the two-user case, corresponding to the rate-splitting strategies employed, when the network-size is scaled down from three-users to two-users.

Sub-message	Rate	Description
$m_{10} \in \{1, ..., 2^{nR_{10}}\}$	R_{10}	Rate achieved: $\mathcal{S}_1 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$m_{11} \in \{1, ..., 2^{nR_{11}}\}$	R_{11}	Rate achieved: $\mathcal{S}_1 \to \mathcal{R}_1$
$m_{20} \in \{1, ..., 2^{nR_{20}}\}$	R_{20}	Rate achieved: $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$m_{21} \in \{1, ..., 2^{nR_{21}}\}$	R_{21}	Rate achieved: $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2)$
$m_{22} \in \{1, ..., 2^{nR_{22}}\}$	R_{22}	Rate achieved: $\mathcal{S}_2 \to \mathcal{R}_2$
$m_{30} \in \{1, ..., 2^{nR_{30}}\}$	R_{30}	Rate achieved: $\mathcal{S}_3 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$m_{31} \in \{1, ..., 2^{nR_{31}}\}$	R_{31}	Rate achieved: $\mathcal{S}_3 \to (\mathcal{R}_1, \mathcal{R}_3)$
$m_{33} \in \{1, ..., 2^{nR_{33}}\}$	R_{33}	Rate achieved: $\mathcal{S}_3 \to \mathcal{R}_3$
$m_1 \in \{1, ..., 2^{nR_1}\}$	R_1	Rate achieved: $\mathcal{S}_1 \to \mathcal{R}_1$
$m_2 \in \{1, ..., 2^{nR_2}\}$	R_2	Rate achieved: $\mathcal{S}_2 \to \mathcal{R}_2$

Table 2.1: Achievable rates and their description. For ex., R_{11} is the rate achieved between $\mathcal{S}_1$ and $\mathcal{R}_1$, while R_{21} is the rate achieved between $\mathcal{S}_2$, and $\mathcal{R}_2$, $\mathcal{R}_1$, *etc.* The last two rows correspond to the channel $\mathcal{C}_{\text{CoMS}}$, wherein the senders $\mathcal{S}_1$ and $\mathcal{S}_2$ do not employ rate-splitting.

The notation for describing the achievable rates of these sub-messages and their respective description is tabulated in Table 2.1. The decoding capabilities of receivers, resulting from rate-splitting at the transmitters, are summarized in Tables 2.2, 2.3 and 2.4. We also introduce auxiliary random variables defined on fi-

Receiver	Decoding capability
$\mathcal{R}_1$	m_{10}, m_{11}, m_{20}, m_{30}
$\mathcal{R}_2$	m_{10}, m_{20}, m_{22}, m_{30}
$\mathcal{R}_3$	m_{10}, m_{20}, m_{30}, m_{33}

Table 2.2: Effect of rate-splitting on the decoding capability of receivers for the channels $\mathcal{C}^1_{\text{CuMS}}$, $\mathcal{C}^1_{\text{PrMS}}$. For ex., receiver $\mathcal{R}_2$ can decode messages m_{10}, m_{20}, m_{22}, m_{30}

Receiver	Decoding capability
$\mathcal{R}_1$	m_{11}, m_{21}, m_{31}
$\mathcal{R}_2$	m_{21}, m_{22}
$\mathcal{R}_3$	m_{31}, m_{33}

Table 2.3: Effect of rate-splitting on the decoding capability of receivers for the channels $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^2_{\text{PrMS}}$. For ex., receiver $\mathcal{R}_3$ can decode messages m_{31}, m_{33}

Receiver	Can decode
$\mathcal{R}_1$	m_1, m_{31}
$\mathcal{R}_2$	m_2, m_{31}
$\mathcal{R}_3$	m_{31}, m_{33}

Table 2.4: Effect of rate-splitting on the decoding capability of receivers for the channel $\mathcal{C}_{\text{CoMS}}$. For *e.g.* the receiver denoted $\mathcal{R}_2$ can decode messages m_2 and m_{31}. Note that, there is no rate-splitting at the senders $\mathcal{S}_1$ and $\mathcal{S}_2$.

nite sets and tabulate them in Table 2.5. Depending on the rate-splitting strategy employed by the senders, only a subset of these sub-messages, their corresponding rates, and the corresponding auxiliary random variables will be used to derive an achievable rate region for each channel model.

2.2.3 Channel Modification

Rate-splitting necessitates modification of the channels $\mathcal{C}^t_{\text{CuMS}}$, $\mathcal{C}^t_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}; t = 1, 2$. Here, we explicitly show the modification for one channel ($\mathcal{C}^2_{\text{CuMS}}$); the modification for the other channel models is similar. Referring to the rate-splitting strategy for the channel $\mathcal{C}^2_{\text{CuMS}}$, the messages at the three senders in the modified channel can be written as:

Variable	Description
$W_0 \in \mathcal{W}_0$	Public Information: $\mathcal{S}_1 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$W_1 \in \mathcal{W}_1$	Private Information: $\mathcal{S}_1 \to \mathcal{R}_1$
$U_0 \in \mathcal{U}_0$	Public Information: $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$U_1 \in \mathcal{U}_1$	Public information: $\mathcal{S}_2 \to (\mathcal{R}_1, \mathcal{R}_2)$
$U_2 \in \mathcal{U}_2$	Private information: $\mathcal{S}_2 \to \mathcal{R}_2$
$V_0 \in \mathcal{V}_0$	Public information: $\mathcal{S}_3 \to (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3)$
$V_1 \in \mathcal{V}_1$	Public information: $\mathcal{S}_3 \to (\mathcal{R}_1, \mathcal{R}_3)$
$V_3 \in \mathcal{V}_3$	Private information: $\mathcal{S}_3 \to \mathcal{R}_3$

Table 2.5: Auxiliary Random variables and their description. For *e.g.*, U_1 denotes public information from $\mathcal{S}_2$ decodable at $\mathcal{R}_1$ and $\mathcal{R}_2$

$$\text{Sender 1: } m_{11} \in \mathcal{M}_{11} = \{1, \ldots, M_{11}\},$$
$$\text{Sender 2: } m_{21} \in \mathcal{M}_{21} = \{1, \ldots, M_{21}\},\ m_{22} \in \mathcal{M}_{22} = \{1, \ldots, M_{22}\},$$
$$\text{Sender 3: } m_{31} \in \mathcal{M}_{31} = \{1, \ldots, M_{31}\},\ m_{33} \in \mathcal{M}_{33} = \{1, \ldots, M_{33}\},$$

with all messages being defined on sets with finite number of elements. Note that, there is no rate-splitting at sender $\mathcal{S}_1$, but for consistency in notation we write m_1 as m_{11}.

We define an $\left(M_{11}, M_{21}, M_{22}, M_{31}, M_{33}, n, P_e^{(n)}\right)$ code for the modified channel as a set of M_{11} codewords for $\mathcal{S}_1$, $M_{11}M_{21}M_{22}$ codewords for $\mathcal{S}_2$, and $M_{11}M_{21}M_{22}M_{31}M_{33}$ codewords for $\mathcal{S}_3$, such that the average probability of decoding error is less than $P_e^{(n)}$. We call a tuple $(R_{11}, R_{21}, R_{22}, R_{31}, R_{33})$ achievable if there exists a sequence of $\left(2^{\lceil nR_{11}\rceil}, 2^{\lceil nR_{21}\rceil}, 2^{\lceil nR_{22}\rceil}, 2^{\lceil nR_{31}\rceil}, 2^{\lceil nR_{33}\rceil}, n, P_e^{(n)}\right)$ codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. Here, R_{11} corresponds to R_1. The capacity region for the modified channel is the closure of the set of all achievable rate tuples $(R_{11}, R_{21}, R_{22}, R_{31}, R_{33})$. It can be shown that if the rate tuple $(R_{11}, R_{21}, R_{22}, R_{31}, R_{33})$ is achievable for the modified channel, then the rate triple $(R_{11}, R_{21}+R_{22}, R_{31}+R_{33})$ is achievable for the channel $\mathcal{C}^2_{\text{CuMS}}$ (see [7, Corollary 2.1]). In a similar fashion, the remaining channel models can be appropriately modified; the details are omitted to avoid repetition.

2.2.4 Probability Distributions

Here, we present the probability distribution functions which characterize the channels $\mathcal{C}^1_{\text{CuMS}}$, $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^1_{\text{PrMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$. Let $\mathcal{P}^t_{\text{CuMS}}$ denote the set of all joint probability distributions $p^t_{\text{CuMS}}(.);\ t = 1, 2$ respectively, that factor as follows:

$$p^1_{\text{CuMS}}(q, w_0, w_1, x_1, u_0, u_2, x_2, v_0, v_3, x_3, y_1, y_2, y_3) =$$
$$p(q)p(w_0, w_1, x_1|q)p(u_0|w_0, w_1, q)p(u_2|w_0, w_1, q)p(x_2|u_0, u_2, w_0, w_1, q)p(v_0|u_0, u_2, w_0, w_1, q)$$
$$p(v_3|u_0, u_2, w_0, w_1, q)p(x_3|v_0, v_3, u_0, u_2, w_0, w_1, q)p(y_1, y_2, y_3|x_1, x_2, x_3), \quad (2.1)$$

$$p^2_{\text{CuMS}}(q, w, x_1, u_1, u_2, x_2, v_1, v_3, x_3, y_1, y_2, y_3) =$$
$$p(q)p(w, x_1|q)p(u_1|w, q)p(u_2|w, q)p(x_2|u_1, u_2, w, q)p(v_1|u_1, u_2, w, q)p(v_3|u_1, u_2, w, q)$$

$$p(x_3|v_1, v_3, u_1, u_2, w, q)p(y_1, y_2, y_3|x_1, x_2, x_3).$$

(2.2)

Let $\mathcal{P}^t_{\text{PrMS}}$ denote the set of all joint probability distributions $p^t_{\text{PrMS}}(.);\ t = 1, 2$ respectively, that factor as follows:

$$p^1_{\text{PrMS}}(q, w_0, w_1, x_1, u_0, u_2, x_2, v_0, v_3, x_3, y_1, y_2, y_3) =$$
$$p(q)p(w_0, w_1, x_1|q)p(u_0|w_0, w_1, q)p(u_2|w_0, w_1, q)$$
$$p(x_2|u_0, u_2, w_0, w_1, q)p(v_0|w_0, w_1, q)p(v_3|w_0, w_1, q)p(x_3|v_0, v_3, w_0, w_1, q)p(y_1, y_2, y_3|x_1, x_2, x_3),$$

(2.3)

$$p^2_{\text{PrMS}}(q, w, x_1, u_1, u_2, x_2, v_1, v_3, x_3, y_1, y_2, y_3) =$$
$$p(q)p(w, x_1|q)p(u_1|w, q)p(u_2|w, q)$$
$$p(x_2|u_1, u_2, w, q)p(v_1|w, q)p(v_3|w, q)p(x_3|v_1, v_3, w, q)p(y_1, y_2, y_3|x_1, x_2, x_3).$$

(2.4)

Let $\mathcal{P}_{\text{CoMS}}$ denote the set of all joint probability distributions $p_{\text{CoMS}}(.)$ respectively, that factor as follows:

$$p_{\text{CoMS}}(q, w_1, x_1, u_2, x_2, v_0, v_3, x_3, y_1, y_2, y_3) = p(q)p(w_1, x_1|q)p(u_2, x_2|q)p(v_0|w_1, u_2, q)$$
$$p(v_3|w_1, u_2, q)p(x_3|v_0, v_3, u_2, w_1, q)p(y_1, y_2, y_3|x_1, x_2, x_3). \quad (2.5)$$

The lower case letters (q, w, u_2, v_3, *etc.*) are realizations of their corresponding random variables, and note that for notational simplicity, the same letter (p) is used to denote all the different probability distributions above. Here, we only describe an achievable rate region for the channel $\mathcal{C}^2_{\text{CuMS}}$, which is defined by a set of non-negative real numbers $(R_{11}, R_{21}, R_{22}, R_{31}, R_{33})$ that satisfy the following information-theoretic inequalities:

$$R_{11} \leq I(W; U_1, V_1, Y_1|Q), \quad (2.6)$$

$$R_{11} + R_{21} \leq I(W, U_1; V_1, Y_1|Q), \quad (2.7)$$

$$R_{11} + R_{31} \leq I(W, V_1; U_1, Y_1|Q) + I(W; V_1|Q) - I(W, U_1, U_2; V_1|Q), \quad (2.8)$$

$$R_{11} + R_{21} + R_{31} \leq I(W, U_1, V_1; Y_1|Q) + I(W, U_1; V_1|Q) - I(W, U_1, U_2; V_1|Q), \quad (2.9)$$

$$R_{21} \leq I(U_1; U_2, Y_2|Q) - I(W; U_1|Q), \quad (2.10)$$

$$R_{22} \leq I(U_2; U_1, Y_2|Q) - I(W; U_2|Q), \quad (2.11)$$

$$R_{21} + R_{22} \leq I(U_1, U_2; Y_2|Q) + I(U_1; U_2|Q) - I(W; U_1|Q) - I(W; U_2|Q), \quad (2.12)$$

$$R_{31} \leq I(V_1; V_3, Y_3|Q) - I(W, U_1, U_2; V_1|Q), \quad (2.13)$$

$$R_{33} \leq I(V_3; V_1, Y_3|Q) - I(W, U_1, U_2; V_3|Q), \quad (2.14)$$

$$R_{31} + R_{33} \leq I(V_1, V_3; Y_3|Q) + I(V_1; V_3|Q) - I(W, U_1, U_2; V_3|Q) - I(W, U_1, U_2; V_1|Q) \quad (2.15)$$

An achievable rate region for the remaining channels considered in this work are given in [16, Appendices A, B, C].

2.2.5 Achievability Theorem

Theorem 2.2.1. Let $\mathfrak{C}^t_{\text{CuMS}}$(or $\mathfrak{C}^t_{\text{PrMS}}$ or $\mathfrak{C}_{\text{CoMS}}$) denote the capacity region of the channel $\mathcal{C}^t_{\text{CuMS}}$ or $\mathcal{C}^t_{\text{PrMS}}$ or $\mathcal{C}_{\text{CoMS}}$); $t = 1, 2$. Let

$$\mathfrak{R}^t_{\text{CuMS}} = \bigcup_{p^t_{\text{CuMS}}(.) \in \mathcal{P}^t_{\text{CuMS}}} \mathfrak{R}_{\text{CuMS}}(p^t_{\text{CuMS}});$$

$$\mathfrak{R}^t_{\text{PrMS}} = \bigcup_{p^t_{\text{PrMS}}(.) \in \mathcal{P}^t_{\text{PrMS}}} \mathfrak{R}_{\text{PrMS}}(p^t_{\text{PrMS}});$$

$$\mathfrak{R}_{\text{CoMS}} = \bigcup_{p_{\text{CoMS}}(.) \in \mathcal{P}_{\text{CoMS}}} \mathfrak{R}_{\text{CoMS}}(p_{\text{CoMS}}).$$

In the above, $\mathfrak{R}_{\text{CuMS}}(p^t_{\text{CuMS}})$ denotes a set of achievable rates when the channel is characterized by the joint probability distribution function p^t_{CuMS}, and similar definitions apply for the other notations used. The region $\mathfrak{R}^t_{\text{CuMS}}$(or $\mathfrak{R}^t_{\text{PrMS}}$ or $\mathfrak{R}_{\text{CoMS}}$) is an achievable rate region for the channel $\mathcal{C}^t_{\text{CuMS}}$(or $\mathcal{C}^t_{\text{PrMS}}$ or $\mathcal{C}_{\text{CoMS}}$), i.e., $\mathfrak{R}^t_{\text{CuMS}}$(or $\mathfrak{R}^t_{\text{PrMS}}$ or $\mathfrak{R}_{\text{CoMS}}$) $\subseteq \mathfrak{C}^t_{\text{CuMS}}$ (or $\mathfrak{C}^t_{\text{PrMS}}$ or $\mathfrak{C}_{\text{CoMS}}$).

We employ the technique of combining GP's binning principle [8] and superpo-

sition coding [9] to prove the coding theorem and derive a set of achievable rates for each of the channel models. For sake of brevity, we only show the proof of Theorem 2.2.1 for the channel $\mathcal{C}^2_{\text{CuMS}}$ (see Appendix A.1). The proof for the remaining three channels ($\mathcal{C}^1_{\text{CuMS}}$, $\mathcal{C}^1_{\text{PrMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$) can be proved along similar lines.

2.3 The Gaussian Case

In this section, we introduce the Gaussian CR channel to (i) evaluate and plot the rate region for the different channel models considered in this work, (ii) describe several extensions, in the form of corollaries, to the achievable rate regions described above, and (iii) derive outer bounds to help us test the optimality of the coding techniques that we have employed to derive the achievable rate regions.

2.3.1 The Gaussian CR channel

The achievable rate regions described for the discrete memoryless channels can be extended to the Gaussian channels by quantizing the channel inputs and outputs [51]. Let $\mathcal{C}^t_{G,\text{CuMS}}$ denote the cognitive Gaussian channel with cumulative message sharing, $\mathcal{C}^t_{G,\text{PrMS}}$ the cognitive Gaussian channel with primary-only message sharing and $\mathcal{C}^t_{G,\text{CoMS}}$ the cognitive Gaussian channel with cognitive-only message sharing (G for Gaussian, CuMS, PrMS and CoMS are the same as before); $t = 1, 2$. We show the extension for only one of the channel models - from $\mathcal{C}^2_{\text{CuMS}}$ to $\mathcal{C}^2_{G,\text{CuMS}}$.

The cognitive Gaussian channel is described by a discrete-time input $\tilde{X}_k$, a corresponding output $\tilde{Y}_k$, and a random variable $\tilde{Z}_k$ denoting noise at the receiver; $k = 1, 2, 3$. Following the maximum-entropy theorem [52], the input random variable $\tilde{X}_k;\ k = 1, 2, 3$ is assumed to have a Gaussian distribution. The transmitted codeword $\tilde{\mathbf{x}}_k = (\tilde{x}_{k1}, \ldots, \tilde{x}_{kn})$ satisfies the average power constraint given by $\mathbb{E}\{\|\tilde{\mathbf{x}}_k\|^2\} \leq \tilde{P}_k;\ k = 1, 2, 3$, where $\mathbb{E}\{.\}$ is the expectation operator. The zero-mean random variable $\tilde{Z}_k$ is drawn i.i.d from a Gaussian distribution with variance $\tilde{N}_k;\ k = 1, 2, 3$, and is assumed to be independent of the signal $\tilde{X}_k$. The

Gaussian CR channel can be converted to a standard form using invertible transformations [35], [53].

For the channel $\mathcal{C}^2_{G,\text{CuMS}}$, we have W, U_1, U_2, V_1 and V_3 as the random variables (RV) which describe the sources at the transmitters. We also some consider additional RVs - $\tilde{W}$, $\tilde{U}_1$, $\tilde{U}_2$, $\tilde{V}_1$ and $\tilde{V}_3$ - with the following statistics: $\tilde{W} \sim \mathcal{N}(0, P_1)$; $\tilde{U}_1 \sim \mathcal{N}(0, \tau P_2)$, $\tilde{U}_2 \sim \mathcal{N}(0, \bar{\tau} P_2)$, with $\tau + \bar{\tau} = 1$; $\tilde{V}_1 \sim \mathcal{N}(0, \kappa P_3)$, $\tilde{V}_3 \sim \mathcal{N}(0, \bar{\kappa} P_3)$, with $\kappa + \bar{\kappa} = 1$. Further, $W = \tilde{W}$; $U_1 = \tilde{U}_1 + \alpha_1 X_1$, $U_2 = \tilde{U}_2 + \alpha_2 X_1$; $V_1 = \tilde{V}_1 + \alpha_3 X_1 + \beta_1 X_2$, $V_3 = \tilde{V}_3 + \alpha_4 X_1 + \beta_2 X_2$, where the input RV's X_1, X_2 and X_3 are given by $X_1 = \tilde{W}$, $X_2 = \tilde{U}_1 + \tilde{U}_2$ and $X_3 = \tilde{V}_1 + \tilde{V}_3$. Notice that $\tilde{W}$, $\tilde{U}_1$, $\tilde{U}_2$, $\tilde{V}_1$ and $\tilde{V}_3$ are mutually independent. Therefore, $X_1 \sim \mathcal{N}(0, P_1)$, $X_2 \sim \mathcal{N}(0, P_2)$ and $X_3 \sim \mathcal{N}(0, P_3)$. The values of τ and κ are randomly selected from the interval $[0, 1]$. The values of α_1, α_2, α_3, α_4, β_1 and β_2 are repeatedly generated according to $\mathcal{N}(0, 1)$. The channel outputs are

$$
\begin{aligned}
Y_1 &= X_1 + a_{12} X_2 + a_{13} X_3 + Z_1, \\
Y_2 &= a_{21} X_1 + X_2 + a_{23} X_3 + Z_2, \\
Y_3 &= a_{31} X_1 + a_{32} X_2 + X_3 + Z_3,
\end{aligned}
$$

where $Z_1 \sim \mathcal{N}(0, Q_1)$, $Z_2 \sim \mathcal{N}(0, Q_2)$ and $Z_3 \sim \mathcal{N}(0, Q_3)$ are independent additive noise, and Q_1, Q_2 and Q_3 are noise variances when the input-output relations are represented in the standard form. Substituting for X_1, X_2 and X_3, we get,

$$
\begin{aligned}
Y_1 &= \tilde{W} + a_{12}(\tilde{U}_1 + \tilde{U}_2) + a_{13}(\tilde{V}_1 + \tilde{V}_3) + Z_1, \\
Y_2 &= a_{21}\tilde{W} + (\tilde{U}_1 + \tilde{U}_2) + a_{23}(\tilde{V}_1 + \tilde{V}_3) + Z_2, \\
Y_3 &= a_{31}\tilde{W} + a_{32}(\tilde{U}_1 + \tilde{U}_2) + \tilde{V}_1 + \tilde{V}_3 + Z_3,
\end{aligned}
$$

where the interference coefficients a_{12}, a_{13}, a_{21}, a_{23}, a_{31} and a_{32} are assumed to be real and globally known. The rate region $\mathfrak{R}^2_{\text{CuMS}}$ for the channel $\mathcal{C}^2_{\text{CuMS}}$ can be extended to its respective Gaussian channel model by evaluating the mutual information terms. To this end, we construct a covariance matrix given by $\mathbb{E}\{\Theta^T \Theta\}$, where $\Theta = (Y_1,\ Y_2,\ Y_3,\ W,\ U_1,\ U_2,\ V_1,\ V_3)$. The entries of this covariance matrix

are used to compute the differential entropy terms, which are further used to evaluate the mutual information.

Theorem 2.3.1. Let $\Upsilon = (\tau, \kappa, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2)$. For a fixed Υ, let $\mathcal{G}^2_{\text{CuMS}}(\Upsilon)$ be achievable. The rate region $\mathfrak{G}^2_{\text{CuMS}}$ is achievable for the Gaussian channel $\mathcal{C}^2_{G,\text{CuMS}}$ with $\mathfrak{G}^2_{\text{CuMS}} = \bigcup_{\Upsilon} \mathcal{G}^2_{\text{CuMS}}(\Upsilon)$.

Since the computation procedure is cumbersome and lengthy albeit straightforward, we do not provide the proof here. The same procedure is followed to compute the mutual information terms for the remaining channel models - $\mathcal{C}^1_{G,\text{CuMS}}$, $\mathcal{C}^t_{G,\text{PrMS}}; t = 1, 2$, and $\mathcal{C}_{G,\text{CoMS}}$.

2.3.2 Extensions

We state several corollaries in this subsection that help in identifying additional achievable rate points by treating the cognitive transmitters as relays, depending on their knowledge of the other user's message. Also note that we present the achievable rate points as separate corollaries for clarity of presentation; one could state them together as one single result as well.

2.3.2.1 $\mathcal{C}^t_{G,\textbf{CuMS}}$

Corollary 2.3.1. Let $\mathfrak{G}^2_{\text{CuMS}}$ be the set of all points $(R_1, R_{21} + R_{22}, R_{31} + R_{33})$ where $(R_1, R_{21}, R_{22}, R_{31}, R_{33})$ is an achievable rate tuple of Theorem 2.3.1. Then, the convex hull of the region $\mathfrak{G}^2_{\text{CuMS}}$ with the points $(R_1^*, 0, 0)$ and $(0, R_2^*, R_3^*)$ is achievable for the $\mathcal{C}^t_{G,\text{CuMS}}$ model, where

$$
\begin{aligned}
R_1^* &= \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P}_1 + |a_{12}|\sqrt{P}_2 + |a_{13}|\sqrt{P}_3\right)^2}{Q_1}\right), \\
R_2^* &= \frac{1}{2}\log_2\left(1 + \frac{P_2}{Q_2 + |a_{23}|^2 P_3}\right), \\
R_3^* &= \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3}\right).
\end{aligned}
$$

The proof of Corollary 2.3.1 can be found in Appendix A.2. The proofs of the remaining corollaries are omitted as they are similar; the interested reader is referred to [16] for details.

Corollary 2.3.2. Let $\mathfrak{G}^2_{\text{CuMS}}$ be the set of all points $(R_1, R_{21}+R_{22}, R_{31}+R_{33})$, where $(R_1, R_{21}, R_{22}, R_{31}, R_{33})$ is an achievable rate tuple of Theorem 2.3.1. Then the convex hull of the region $\mathfrak{G}^2_{\text{CuMS}}$ with the points $(R_1^*, 0, r)$ and $(0, R_2^*, r)$ are achievable for the $\mathcal{C}^t_{G,\text{CuMS}}$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_1} + |a_{12}|\sqrt{P_2} + |a_{13}|\sqrt{P_3^{\mathcal{S}_1}}\right)^2}{Q_1 + |a_{13}|^2 P_3^{\mathcal{S}_3}}\right),$$
$$R_2^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_2} + |a_{23}|\sqrt{P_3^{\mathcal{S}_2}}\right)^2}{Q_2 + |a_{23}|^2 P_3^{\mathcal{S}_3}}\right),$$
$$r = \frac{1}{2}\log_2\left(1 + \frac{P_3^{\mathcal{S}_3}}{Q_3}\right),$$

where $P_3^{\mathcal{S}_1} = P_3^{\mathcal{S}_2} = P_3 - P_3^{\mathcal{S}_3}$, $\forall\, P_3^{\mathcal{S}_3} \in [0, P_3]$.

Corollary 2.3.3. The convex hull of the region $\mathfrak{G}^2_{\text{CuMS}}$ with the points $(R_1^*, r, 0)$ and $(0, r, R_3^*)$ is achievable for the $\mathcal{C}^t_{G,\text{CuMS}}$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P1} + |a_{12}|\sqrt{P_2^{\mathcal{S}_1}} + |a_{13}|\sqrt{P_3}\right)^2}{Q_1 + |a_{12}|^2 P_2^{\mathcal{S}_2}}\right),$$
$$r = \frac{1}{2}\log_2\left(1 + \frac{P_2^{\mathcal{S}_2}}{Q_2 + |a_{23}|^2 P_3}\right),$$
$$R_3^* = \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3}\right),$$

where $P_2^{\mathcal{S}_2} = (2^{2r} - 1)(Q_2 + |a_{23}|^2 P_3)$, $P_2^{\mathcal{S}_1} = P_2 - P_2^{\mathcal{S}_2}$ and r is the minimum rate that $\mathcal{S}_2$ is guaranteed to achieve.

Corollary 2.3.4. The convex hull of the region $\mathfrak{G}^2_{\text{CuMS}}$ with the points $(0, R_2^*, 0)$ and

$(0, 0, R_3^*)$ is achievable for the $\mathcal{C}^t_{G,\text{CuMS}}$ model, where

$$R_2^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P}_2 + |a_{23}|\sqrt{P_3}\right)^2}{Q_2}\right),$$
$$R_3^* = \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3}\right).$$

The following theorem follows directly from standard time-sharing arguments.

Theorem 2.3.2. The convex hull of the region $\mathfrak{G}^2_{\text{CuMS}}$ with the achievable points in the Corollaries 2.3.1 - 2.3.4 results in an achievable rate region of the $\mathcal{C}^t_{G,\text{CuMS}}$ channel model.

2.3.2.2 $\mathcal{C}^t_{G,\textbf{PrMS}}$

Corollary 2.3.5. Let $\mathfrak{G}^2_{\text{PrMS}}$ be the set of all points $(R_1, R_{21} + R_{22}, R_{31} + R_{33})$ such that $(R_1, R_{21}, R_{22}, R_{31}, R_{33})$ is an achievable rate tuple. Then the convex hull of the region $\mathfrak{G}^2_{\text{PrMS}}$ with the points $(R_1^*, 0, 0)$ and $(0, R_2^*, R_3^*)$ are achievable for the $\mathcal{C}^t_{G,\text{PrMS}}$ model, where

$$R_1^* = \frac{1}{2}\log\left(1 + \frac{\left(\sqrt{P}_1 + |a_{12}|\sqrt{P}_2 + |a_{13}|\sqrt{P}_3\right)^2}{Q_1}\right),$$
$$R_2^* = \frac{1}{2}\log\left(1 + \frac{P_2}{Q_2 + |a_{23}|^2 P_3}\right),$$
$$R_3^* = \frac{1}{2}\log\left(1 + \frac{P_3}{Q_3 + |a_{32}|^2 P_2}\right).$$

Corollary 2.3.6. The convex hull of the region $\mathfrak{G}^2_{\text{PrMS}}$ with the points $(R_1^*, 0, r)$ and $(0, R_2^*, r)$ are achievable for the $\mathcal{C}^t_{G,\text{PrMS}}$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_1} + |a_{12}|\sqrt{P_2} + |a_{13}|\sqrt{P_3^{\mathcal{S}_1}}\right)^2}{Q_1 + |a_{13}|^2 P_3^{\mathcal{S}_3}}\right),$$
$$R_2^* = \frac{1}{2}\log_2\left(1 + \frac{P_2}{Q_2 + |a_{23}|^2 P_3}\right),$$

$$r = \frac{1}{2}\log_2\left(1 + \frac{P_3^{cr2}}{Q_3 + |a_{32}|^2 P_2}\right),$$

where $P_3^{\mathcal{S}_1} = P_3 - P_3^{\mathcal{S}_3}$, $\forall P_3^{\mathcal{S}_3} \in [0, P_3]$.

Corollary 2.3.7. The convex hull of the region $\mathfrak{G}_{\text{PrMS}}^2$ with the points $(R_1^*, r, 0)$ and $(0, r, R_3^*)$ are achievable for the $\mathcal{C}_{G,\text{PrMS}}^t$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_1} + |a_{12}|\sqrt{P_2^{\mathcal{S}_1}} + |a_{13}|\sqrt{P_3}\right)^2}{Q_1 + |a_{12}|^2 P_2^{\mathcal{S}_2}}\right),$$
$$r = \frac{1}{2}\log_2\left(1 + \frac{P_2^{\mathcal{S}_2}}{Q_2 + |a_{23}|^2 P_3}\right),$$
$$R_3^* = \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3 + |a_{32}|^2 P_2}\right),$$

where $P_2^{\mathcal{S}_1} = P_2 - P_2^{\mathcal{S}_2}$, $\forall P_2^{\mathcal{S}_2} \in [0, P_2]$.

The following theorem follows directly from standard time-sharing arguments.

Theorem 2.3.3. The convex hull of the region $\mathfrak{G}_{\text{PrMS}}^2$ with the achievable points in the Corollaries 2.3.5 - 2.3.7 results in an achievable rate region of the $\mathcal{C}_{G,\text{PrMS}}^t$ channel model.

2.3.2.3 $\mathcal{C}_{G,\text{CoMS}}$

Corollary 2.3.8. Let $\mathfrak{G}_{\text{CoMS}}$ be the set of all points $(R_1, R_2, R_{31} + R_{33})$ such that $(R_1, R_2, R_{31}, R_{33})$ is an achievable rate tuple. Then the convex hull of the region $\mathfrak{G}_{\text{CoMS}}$ with the points $(R_1^*, 0, 0)$, $(0, R_2^*, 0)$ and $(0, 0, R_3^*)$ are achievable for the $\mathcal{C}_{G,\text{CoMS}}$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P}_1 + |a_{13}|\sqrt{P}_3\right)^2}{Q_1 + |a_{12}|^2 P_2}\right),$$
$$R_2^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P}_2 + |a_{23}|\sqrt{P}_3\right)^2}{Q_2 + |a_{21}|^2 P_1}\right),$$

$$R_3^* = \frac{1}{2}\log_2\left(1+\frac{P_3}{Q_3}\right).$$

Corollary 2.3.9. The convex hull of the region $\mathfrak{G}_{\text{CoMS}}$ with the points $(R_1^*, 0, r)$, $(0, R_2^*, r)$ and $(0, 0, r)$ are achievable for the $\mathcal{C}_{G,\text{CoMS}}$ model, where

$$R_1^* = \frac{1}{2}\log_2\left(1+\frac{\left(\sqrt{P_1}+|a_{13}|\sqrt{P_3^{\mathcal{S}_1}}\right)^2}{Q_1+|a_{12}|^2 P_2+|a_{13}|^2 P_3^{\mathcal{S}_3}}\right),$$

$$R_2^* = \frac{1}{2}\log_2\left(1+\frac{\left(\sqrt{P_2}+|a_{13}|\sqrt{P_3^{\mathcal{S}_2}}\right)^2}{Q_2+|a_{21}|^2 P_1+|a_{13}|^2 P_3^{\mathcal{S}_3}}\right),$$

$$r = \frac{1}{2}\log_2\left(1+\frac{P_3^{\mathcal{S}_3}}{Q_3}\right),$$

where $P_3^{\mathcal{S}_1} = P_3^{\mathcal{S}_2} = P_3 - P_3^{\mathcal{S}_3}$, $\forall P_3^{\mathcal{S}_3} \in [0, P_3]$.

Again, the following theorem follows directly from standard time-sharing arguments.

Theorem 2.3.4. The convex hull of the region $\mathfrak{G}_{\text{CoMS}}$ with the achievable points in the Corollaries 2.3.8 and2.3.9 results in an achievable rate region of the $\mathcal{C}_{G,\text{CoMS}}$ channel model.

2.3.3 Outer Bounds

For the channel models considered in this work, we derive outer bounds by considering a scenario where the transmitters cooperate in a *bidirectional* manner, i.e., every sender knows the message of every other sender in a noncausal manner. Since bidirectional message sharing is tantamount to having additional information at the transmitters compared to the CR channels, it cannot hurt the capacity. Then, the channel models reduce to a multiple antenna broadcast channel (MIMO-BC) with one sender having three antennas and three receivers with one antenna each. Hence, the capacity region of the MIMO-BC (see [11]) is an

outer bound on our achievable rate regions. We resort to duality results of the broadcast (BC) and the multiple access channels (MAC), reported first in [12] to calculate the capacity of the MIMO-BC.

Let P be the total power constraint for the MIMO-BC and P_1, P_2 and P_3 be the individual power constraint for the MAC. On the MAC channel, the rate achieved by user j is given by

$$R_{\text{MAC},j} = \log_2 \frac{\left| \mathbf{I} + \sum_{i=j}^{K} \mathbf{H}_i^H P_i \mathbf{H}_i \right|}{\left| \mathbf{I} + \sum_{i=j+1}^{K} \mathbf{H}_i^H P_i \mathbf{H}_i \right|}, \tag{2.16}$$

where $|A|$ denotes the determinant of A; and the channel matrices are $\mathbf{H}_1 = [1\ a_{12}\ a_{13}]$, $\mathbf{H}_2 = [a_{21}\ 1\ a_{23}]$ and $\mathbf{H}_3 = [a_{31}\ a_{32}\ 1]$; and $\mathbf{I} + \sum_{i=j+1}^{K} \mathbf{H}_i^H P_i \mathbf{H}_i$ is the interference experienced by the j^{th} user. The MIMO-BC capacity region with power constraint P is equal to the union of capacity regions of the dual MAC, where the union is taken over all individual power constraint, P_1, P_2 and P_3, such that $P = P_1 + P_2 + P_3$. Therefore,

$$C_{\text{BC}}(P, H) = \bigcup_{P_1, P_2, P_3 : \sum_{j=1}^{3} P_j = P} C_{\text{MAC}}(P_1, P_2, P_3; \mathbf{H}^T), \tag{2.17}$$

where $C_{\text{MAC}}(P_1, P_2, P_3; \mathbf{H}^T) = \bigcup_{j \in \{1,2,3\}} R_{\text{MAC},j}$, and $R_{\text{MAC},j}$ is given by (2.16). We thus obtain the capacity region of the MIMO-BC, which forms an outer bound for the channel models considered in this work. Generally, this outer bound tends to be loose, since the MIMO-BC capacity region was obtained by allowing bidirectional (or complete) transmitter cooperation. Nonetheless, these outer bounds provide useful insights into the strengths and weaknesses of the proposed achievable rate regions, as will be shown in the simulation results section, Section 2.4. To the best of our knowledge, this is the first set of outer bounds that have been derived for the three-user Gaussian CR channel. The rates of individual users can be

further bounded depending on the specific channel model.

1. In the case of CuMS, senders $\mathcal{S}_2$ and $\mathcal{S}_3$ have complete knowledge of the $\mathcal{S}_1$'s message and $\mathcal{S}_3$ has knowledge of $\mathcal{S}_2$'s message but not vice-versa. Note that, the rate of $\mathcal{S}_1$ cannot be bounded by the interference-free case where $a_{12} = 0$ and $a_{13} = 0$. This is because unidirectional message sharing enables $\mathcal{S}_2$ and $\mathcal{S}_3$ to transmit the message of $\mathcal{S}_1$, thereby increasing the rate of $\mathcal{S}_1$ beyond what is achievable with the $\mathcal{S}_1$ alone transmitting its message. Hence, rate R_1 can upper bounded as follows.

$$R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{(\sqrt{P_1} + |a_{12}|\sqrt{P_2} + |a_{13}|\sqrt{P_3})^2}{Q_1}\right). \tag{2.18}$$

   Similarly, the rate of $\mathcal{S}_2$ cannot be bounded by the interference free rate, as $\mathcal{S}_3$ can use its knowledge of $\mathcal{S}_2$'s message to enable $\mathcal{S}_2$ increase its rate. Hence, the rate of $\mathcal{S}_2$ can upper bounded as

$$R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{(\sqrt{P_2} + |a_{23}|\sqrt{P_3})^2}{Q_2}\right). \tag{2.19}$$

   Finally, the rate of $\mathcal{S}_3$ can be upper bounded by the interference free case.

$$R_3 \leq \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3}\right). \tag{2.20}$$

2. In the case of PrMS, although $\mathcal{S}_2$ and $\mathcal{S}_3$ have complete knowledge of $\mathcal{S}_1$'s message, they do not have each other's message. Therefore, the bound on the $\mathcal{S}_1$'s rate given by (2.18) remains valid, as the $\mathcal{S}_2$ and $\mathcal{S}_3$ can use their knowledge of $\mathcal{S}_1$'s message to increase its rate. The bound on $\mathcal{S}_3$'s rate is same as in the case of CuMS and is given by (2.20). Lastly, the $\mathcal{S}_2$'s rate can be upper bounded by the interference-free case as follows.

$$R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{P_2}{Q_2}\right). \tag{2.21}$$

3. We upper bound now the sum rate of $\mathcal{S}_2$ and $\mathcal{S}_3$ by allowing full cooperation

between their transmitters and pairing receivers. This results in a point-to-point MIMO channel, whose capacity is expressed as follows.

$$C_{\text{MIMO}} = \max_{i,\sum_i P_i \leq P} \frac{1}{2} \sum_{i=1}^{N} \log_2 \left(1 + \frac{P_i \sigma_i^2}{Q}\right), \tag{2.22}$$

where $\frac{P_i \sigma_i^2}{Q}$ is the signal-to-noise ratio associated with the i^{th} channel, σ_is are the singular values and N represents the number of singular values of the MIMO channel. The optimum power allocation P_i can be obtained by the water-filling algorithm [52].

4. In the case of CoMS, sender $\mathcal{S}_3$ has noncausal knowledge of $\mathcal{S}_1$ and $\mathcal{S}_2$. Therefore, the rates of $\mathcal{S}_1$ and $\mathcal{S}_2$ cannot be bounded by the interference free scenario. The rate of $\mathcal{S}_1$ can be upper bounded as follows:

$$R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{(\sqrt{P_1} + |a_{13}|\sqrt{P_3})^2}{Q_1}\right). \tag{2.23}$$

The rates of $\mathcal{S}_2$ and $\mathcal{S}_3$ can be upper bounded as in (2.19) and (2.20), respectively. To bound the sum rate of $\mathcal{S}_1$ and $\mathcal{S}_2$ we allow full cooperation between the transmitters and pairing receivers, resulting in a point-to-point MIMO channel. The capacity of this channel is given by (2.22).

2.4 Simulation Results and Discussion

We consider a 3-user Gaussian cognitive channel with CuMS, PrMS and CoMS for the simulations. We generate the source and channel symbols as described in Section 2.3. The direct channel gains are $a_{11} = a_{22} = a_{33} = 1$. The interference coefficients $a_{12} = a_{13} = a_{21} = a_{23} = a_{31} = a_{32} = 0.55$. The values of τ and κ are assumed to be randomly selected from the interval $[0, 1]$. The values of α_1, α_2, α_3, α_4, β_1 and β_2 are repeatedly generated according to $\mathcal{N}(0, 1)$. The noise variances $Q_1 = Q_2 = Q_3 = 1$. The transmit powers $P_1 = P_2 = P_3 = 7.8$dB or 10dB, as specified.

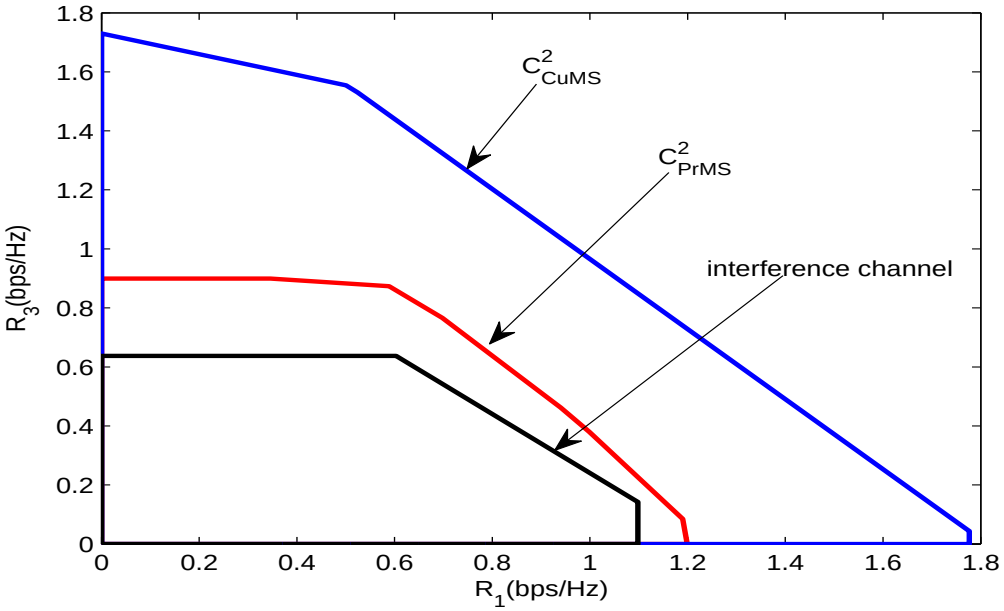

Figure 2.4: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_3$ (R_3) when $\mathcal{S}_2$ is guaranteed to achieve a minimum rate $R_2 = 0.8$ bps/Hz, for $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$ along with the rate region of the corresponding interference channel. The power at the transmitters is 10dB.

2.4.1 Results and Discussion

1. Comparison of the three-user CR channels: The rate region for the channel $\mathcal{C}^2_{\text{CuMS}}$ is obtained following Theorem 2.3.1. Similar procedures are adopted for $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$. We also plot the rate region for the three-user interference channels corresponding to $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$, by considering a simple extension of the Han-Kobayashi scheme [7] to the three-user case. Table 2.3 summarizes the Han-Kobayashi strategy to the three-user case.

   (a) We consider first the achievable rate regions for the channels $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and the three-user interference channel. In Fig. 2.4, we plot the rates of $\mathcal{S}_1$ and $\mathcal{S}_3$, when $\mathcal{S}_2$ achieves a minimum rate of $R_2 = 0.8$ bps/Hz. We notice that, $\mathcal{C}^2_{\text{CuMS}}$ has a bigger rate region than $\mathcal{C}^2_{\text{PrMS}}$. This follows directly from the fact that in $\mathcal{C}^2_{\text{CuMS}}$, the cognitive transmitter $\mathcal{S}_2$ benefits from $\mathcal{S}_3$, while in $\mathcal{C}^2_{\text{PrMS}}$ there is no cooperation between $\mathcal{S}_2$ and $\mathcal{S}_3$. Further, the rate regions for the CR channels are bigger than the corre-

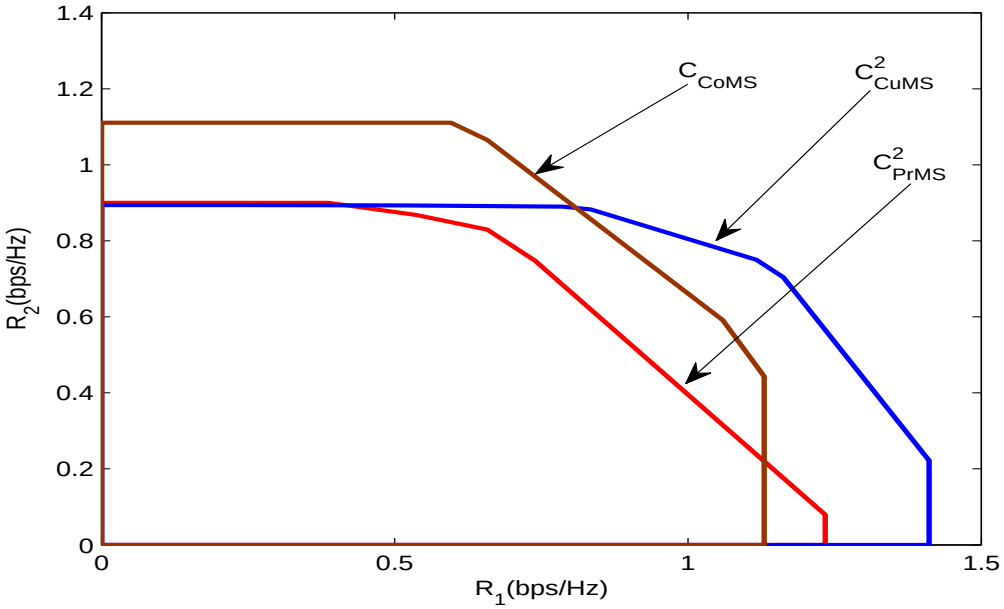

Figure 2.5: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_2$ (R_2) when $\mathcal{S}_3$ is guaranteed to achieve a minimum rate $R_3 = 1$ bps/Hz, for $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$. The power at the transmitters is 10dB.

sponding three-user interference channel, a well-established fact in the classical two-user scenario.

(b) In Fig. 2.5, we plot the rates of $\mathcal{S}_1$ and $\mathcal{S}_2$, when $\mathcal{S}_3$ achieves a minimum rate of $R_3 = 1$ bps/Hz, for the channels $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$. Like in the previous scenarios (Fig. 2.4), $\mathcal{C}^2_{\text{CuMS}}$ has a bigger rate region than $\mathcal{C}^2_{\text{PrMS}}$. We notice that, interestingly, the maximum achievable R_1 for $\mathcal{C}_{\text{CoMS}}$ is smaller than that of $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$. This is due to the message-sharing strategy adopted by $\mathcal{C}_{\text{CoMS}}$, where only $\mathcal{S}_3$ aids the communication of $(\mathcal{S}_1, \mathcal{R}_1)$. We also observe that the maximum achievable R_2 is greater than those of $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$, when one would, at first glance, expect it to be the same as in $\mathcal{C}^2_{\text{CuMS}}$, since in both $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}_{\text{CoMS}}$, $\mathcal{S}_3$ aids the communication of $(\mathcal{S}_2, \mathcal{R}_2)$. The reason can be attributed to the difference in the rate-splitting strategy employed by these channel models (compare Tables 2.3 and 2.4). In case of $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{S}_2$ performs rate-splitting, thereby reducing the effective maximum achiev-

able R_2. But, in case of $\mathcal{C}_{\text{CoMS}}$, $\mathcal{S}_2$ does not employ rate-splitting. This suggests that, similar to the two-user scenario, rate-splitting seems to be less effective than message-sharing. It also suggests that one cannot comment on the superiority of a particular message-sharing scheme compared to another. We also have the following conjecture.

Conjecture 2.4.1. There exists a *tradeoff* between message-sharing mechanisms and rate-splitting strategies. In other words, a particular message-sharing mechanism might not be more beneficial than a specific rate-splitting strategy and vice-versa, in terms of achievable rates on the channel.

For example, consider the following. When a transmitter $\mathcal{S}_1$ shares its message with another transmitter $\mathcal{S}_2$, message splitting by $\mathcal{S}_1$ is not necessary, as it does not have a significant impact on the rates achievable by $\mathcal{S}_1$ and $\mathcal{S}_2$. On the other hand, message splitting by $\mathcal{S}_2$ helps improve the rate achievable by $\mathcal{S}_1$, but does not significantly impact the rate achievable by $\mathcal{S}_2$. Several such instances are possible, and one could argue that performing both rate-splitting and message-sharing would enlarge the overall achievable region. Therefore, a proof of existence and a complete characterization of the *tradeoff* suggested in Conjecture 2.4.1 is an interesting open problem. Also note that, in Fig. 2.5, we have not plotted the rate region of the interference channel. This is because, it is not *fair* to compare the rate region of the three-user interference channel corresponding to the rate-splitting strategy in Table 2.3 with that of $\mathcal{C}_{\text{CoMS}}$ which is based on the rate-splitting strategy in Table 2.4.

In the following, we plot the rate regions, corollaries and outer bounds (see Section 2.3.2) to obtain interesting insights into the achievable rate regions of the different channel models considered in this work.

2. Three-user channels with CuMS (channel $\mathcal{C}^2_{\text{CuMS}}$):

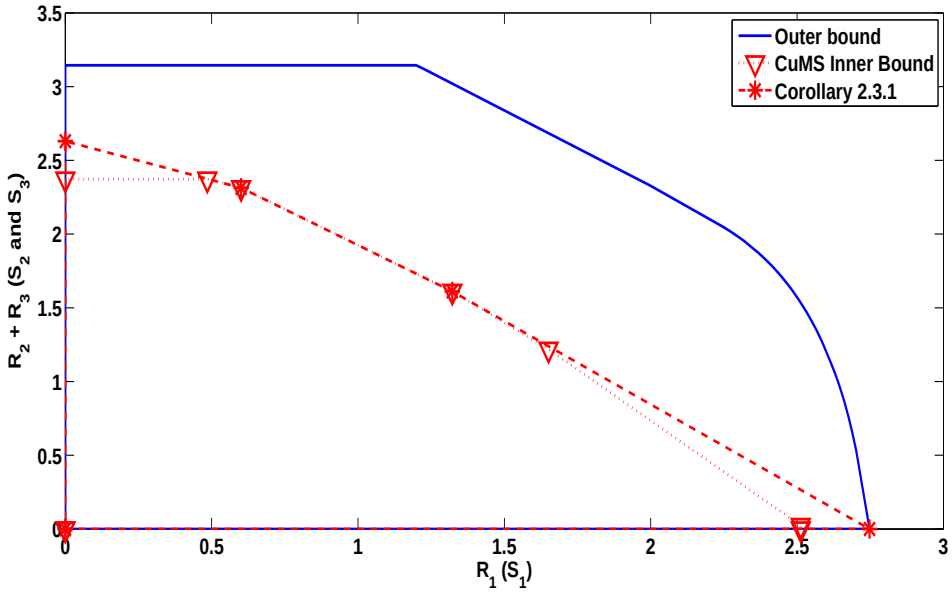

Figure 2.6: Rate of $\mathcal{S}_1$ (R_1) versus the sum rate of $\mathcal{S}_2$ and $\mathcal{S}_3$ ($R_2 + R_3$) for the channel $\mathcal{C}^2_{\text{CuMS}}$. The power at the transmitters is 10dB.

(a) In Fig. 2.6, we plot the rate of sender $\mathcal{S}_1$ (R_1) versus the sum of the rates of $\mathcal{S}_2$ and $\mathcal{S}_3$ (i.e., $R_2 + R_3$) for the channel $\mathcal{C}^2_{\text{CuMS}}$. In the figure, the outer bound, labeled $\mathrm{Outerbound}$, is the intersection of (2.17), (2.18)-(2.20) and (2.22). The innermost region corresponds to the achievable region given in Theorem 2.3.1. The second largest region corresponds to Corollary 2.3.1. Note that our inner bound is for a specific rate-splitting strategy at the transmitters, which the outer bounds do not account for, due to which the outer bounds may be suboptimal and hence loose, for the examples considered here. More insight on the R_2 and R_3 achievable via our scheme, and how it compares with the outer bound, can be obtained from the plots presented later in the discussion.

(b) In Fig. 2.7, we plot the rate of $\mathcal{S}_1$ (R_1), versus that of $\mathcal{S}_2$ (R_2), when $\mathcal{S}_3$ achieves a minimum rate of $R_3 = 0, 1$ and 1.5 bps/Hz. The gap between the inner bound and the outer bound is relatively small. The rate of $\mathcal{S}_2$ does not decrease much as it employs dirty-paper coding to eliminate interference when $\mathcal{S}_1$ and $\mathcal{S}_3$ achieve relatively smaller rates. It can be

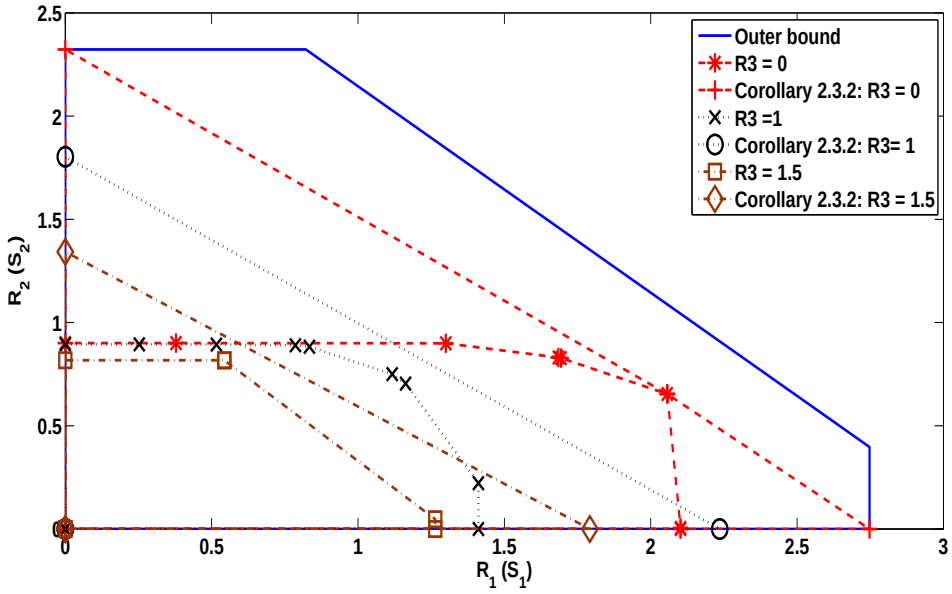

Figure 2.7: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_2$ (R_2) when $\mathcal{S}_3$ is guaranteed to achieve a minimum rate $R_3 = 0, 1$ and 1.5 bps/Hz, for the channel $\mathcal{C}^2_{\text{CuMS}}$. The power at the transmitters is 10dB.

observed that as $\mathcal{S}_3$ achieves higher rates, the achievable rate region of the $\mathcal{S}_1$ and $\mathcal{S}_2$ shrinks. Also, when $R_3 > 0$, the rates achievable using the extensions provided by the corollaries lies completely above the rates achievable by the coding scheme in Section 2.2.5, which is due to the suboptimality of that scheme with respect to the achievable rates of $\mathcal{S}_1$ and $\mathcal{S}_2$ for a fixed R_3. The rate of $\mathcal{S}_1$ has a larger relative reduction compared to that of $\mathcal{S}_2$, yet $\mathcal{S}_1$ achieves a higher rate than $\mathcal{S}_2$, as expected. Figure 2.8 shows a similar plot, but the rate of the $\mathcal{S}_1$ is compared with that of $\mathcal{S}_3$ instead of with $\mathcal{S}_2$. As $\mathcal{S}_2$ achieves a higher and higher rate, the rates of $\mathcal{S}_1$ and $\mathcal{S}_3$ decrease, but the reduction is smaller than that in Fig. 2.7. Note that, in this case, the rate achieved by $\mathcal{S}_3$ matches the outer bound at the corner points when $R_2 = 0$.

3. Three-user channels with PrMS (channel $\mathcal{C}^2_{\text{PrMS}}$):

   (a) In Fig. 2.9, we plot the rate achieved by $\mathcal{S}_1$ versus the sum rate of $\mathcal{S}_2$

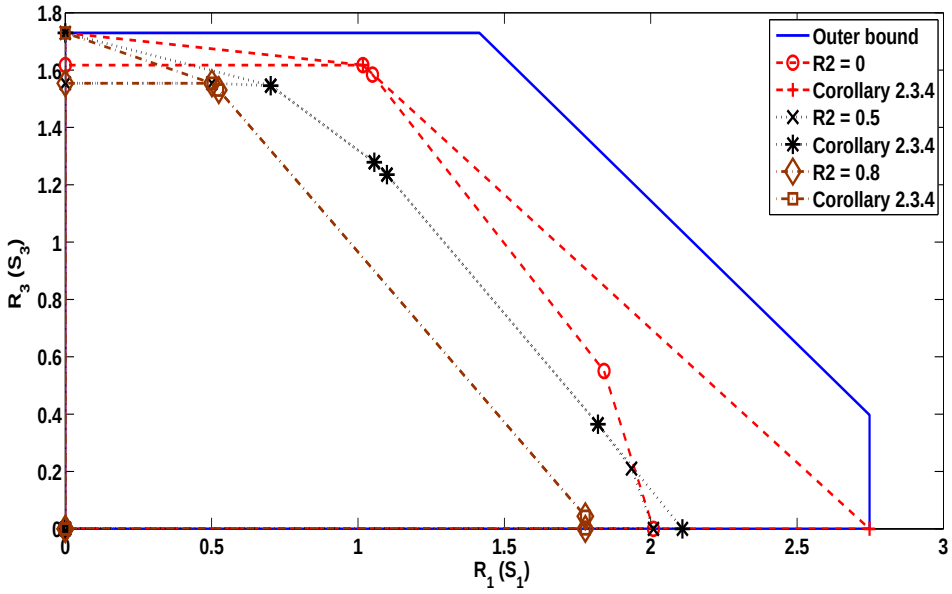

Figure 2.8: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_3$ (R_3) when $\mathcal{S}_2$ is guaranteed to achieve a minimum rate $R_2 = 0, 0.5$ and 0.8 bps/Hz, for the channel $\mathcal{C}^2_{\text{CuMS}}$. The power at the transmitters is 10dB.

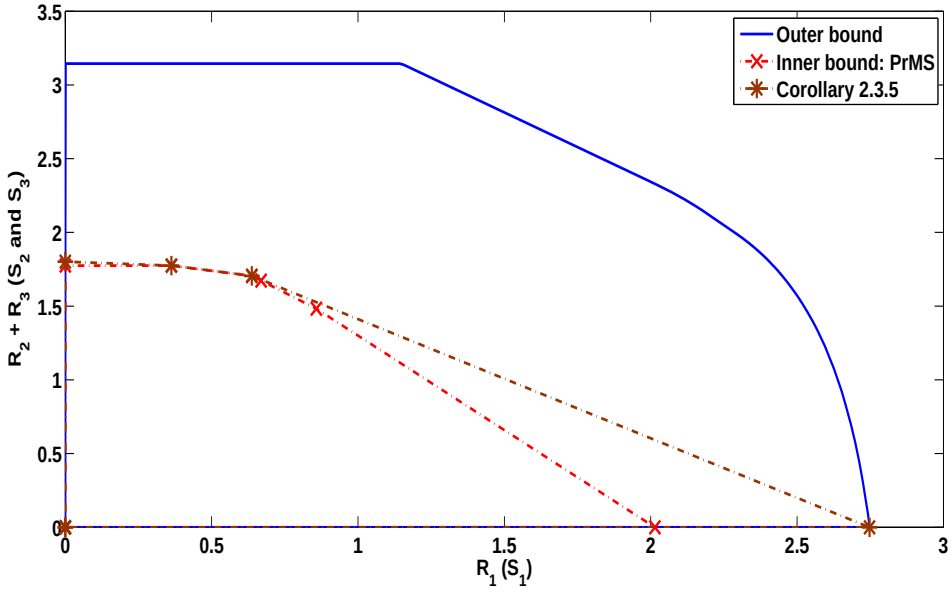

Figure 2.9: Rate of $\mathcal{S}_1$ (R_1) versus the sum rate of $\mathcal{S}_2$ and $\mathcal{S}_3$ ($R_2 + R_3$) for the channel $\mathcal{C}^2_{\text{PrMS}}$. The power at the transmitters is 10dB.

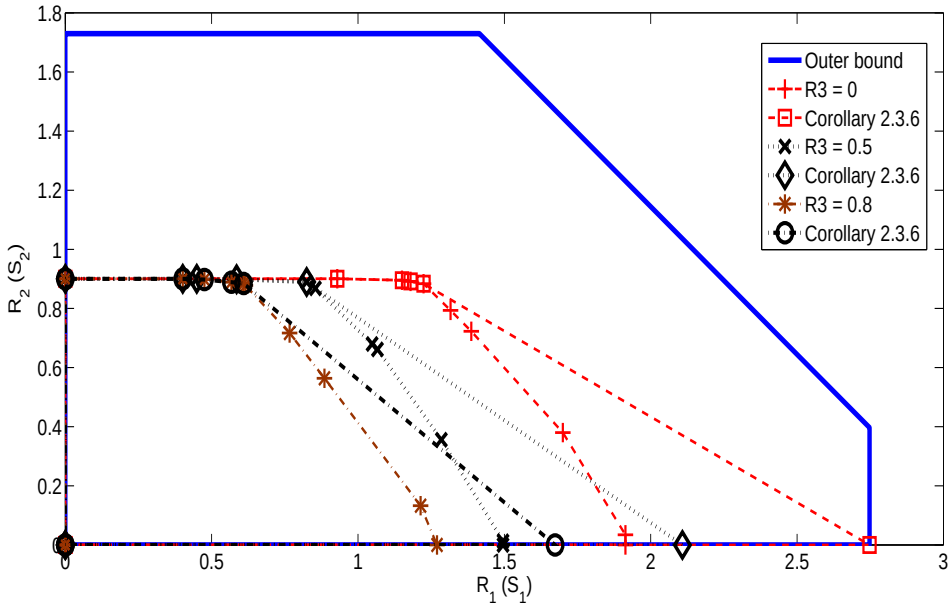

Figure 2.10: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_2$ (R_2) when $\mathcal{S}_3$ is guaranteed to achieve a minimum rate $R_3 = 0, 0.5$ and 0.8 bps/Hz, for the channel $\mathcal{C}^2_{\text{PrMS}}$. The power at the transmitters is 10dB.

and $\mathcal{S}_3$ along with the outer bound. Here, the outer bound is different from the $\mathcal{C}^2_{\text{CuMS}}$ as the cutoff value used to bound R_2 is different for $\mathcal{S}_2$. The plot labeled `Outer bound` is the intersection of the capacity region given by (2.17), (2.18), (2.20) - (2.22). Also shown is the plot of Corollary 2.3.5.

(b) Fig. 2.10 shows the plot of the rate of $\mathcal{S}_1$ versus that of $\mathcal{S}_2$, when $\mathcal{S}_3$ achieves a minimum rates of 0, 0.4 5and 0.8 bps/Hz. Here again, we see that the rates of $\mathcal{S}_1$ and $\mathcal{S}_2$ decrease with increasing rate of $\mathcal{S}_3$. However, the decrease in $\mathcal{S}_2$'s rate is relatively smaller than that of $\mathcal{S}_1$, but $\mathcal{S}_1$ achieves a higher maximum rate compared to $\mathcal{S}_2$.

4. Three-user channel with CoMS:

(a) In Fig. 2.11, we plot the sum rate of senders $\mathcal{S}_1$ and $\mathcal{S}_1$, $R_1 + R_2$, versus the rate of $\mathcal{S}_3$, along with the outer bound and the plot of Corollary 2.3.8. The outer bound is the intersection of (2.17), (2.19), (2.20), (2.22) and

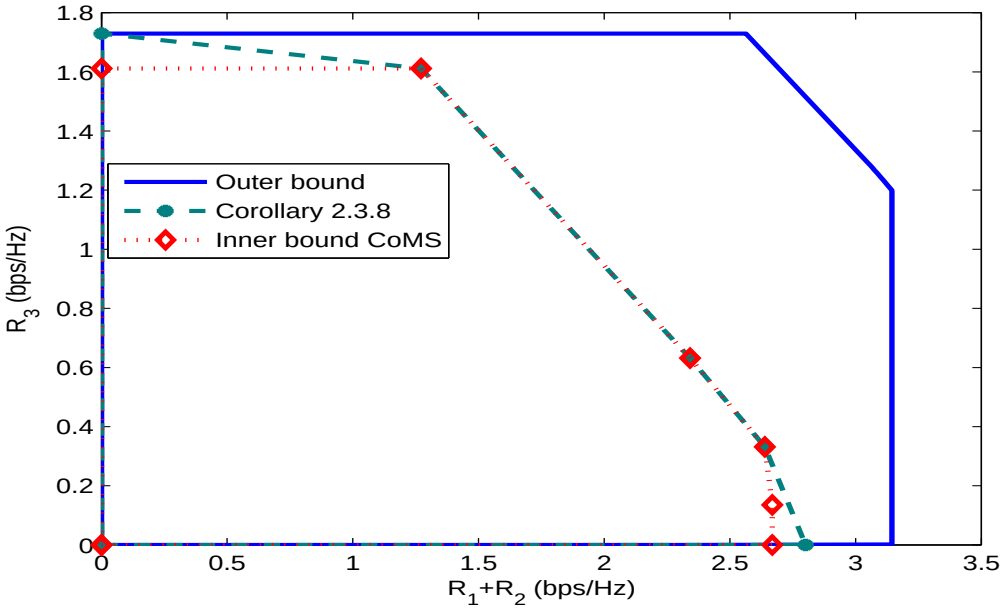

Figure 2.11: Rate of $\mathcal{S}_3$ (R_3) versus the sum rate of $\mathcal{S}_1$ and $\mathcal{S}_2$ ($R_1 + R_2$) for the channel $\mathcal{C}_{\text{CoMS}}$. The power at the transmitters is 10dB.

(2.23).

(b) Figure 2.12 shows the plots of the rates of $\mathcal{S}_1$ and $\mathcal{S}_2$, when $\mathcal{S}_3$ achieves a minimum rate of $0.5, 1$ and 1.5 bps/Hz, along with the plot of Corollary 2.3.9. Here again, we see that the rates of $\mathcal{S}_1$ and $\mathcal{S}_2$ decrease with increasing rate of $\mathcal{S}_3$. However, compared to Fig. 2.10, the reduction in the size of the region is more symmetric i.e., both R_1 and R_2 simultaneously decrease, and roughly speaking, by the same relative amount.

Note 2.4.2. The inner bounds for the C^1_{CuMS} and C^1_{PrMS} have not been plotted here. This is mainly because applying the Fourier-Motzkin elimination procedure on the rate region is a formidable task, given the number of inequalities involved (see [16, Appendices A and B]). This demonstrates practical difficulties involved with rate-splitting, especially with growing network size. Nevertheless, one can expect (i) the achievable rate regions for C^1_{CuMS} and C^1_{PrMS} to be larger than that for C^2_{CuMS} and C^2_{PrMS} and (ii) the gap between the achievable rate region and the outer bound for C^1_{CuMS} and C^1_{PrMS} to be smaller than that to C^2_{CuMS} and C^2_{PrMS}

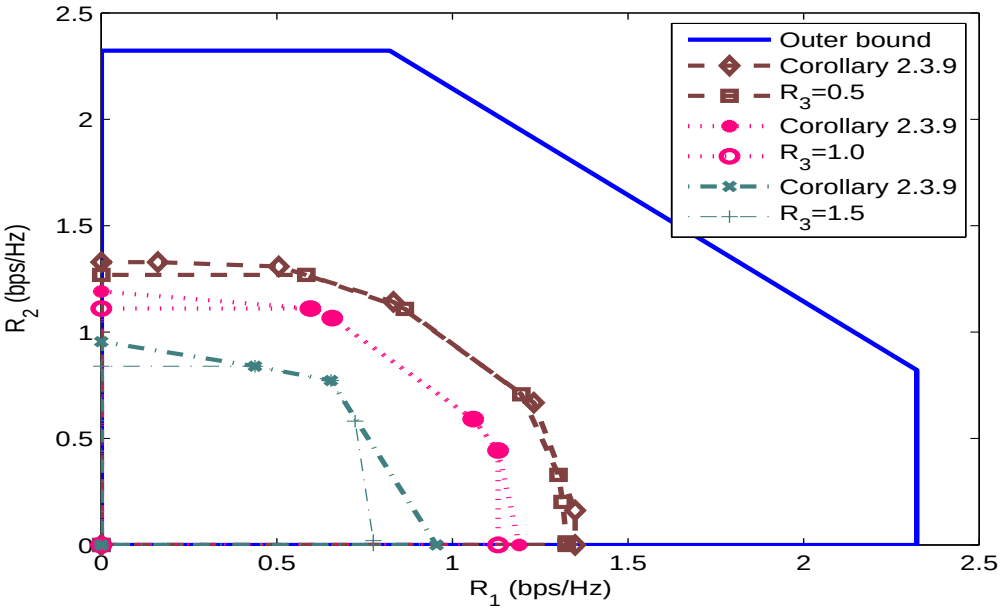

Figure 2.12: Rate of $\mathcal{S}_1$ (R_1) versus the rate of $\mathcal{S}_2$ (R_2) when $\mathcal{S}_3$ is guaranteed to achieve a minimum rate $R_3 = 0.5, 1$ and 1.5 bps/Hz for the channel $\mathcal{C}_{\text{CoMS}}$. The power at the transmitters is 10dB.

respectively, because $\mathcal{S}_1$ also employs rate-splitting strategy in the former case.

As a concluding remark, note that, as mentioned above, there is a gap between the inner and outer bounds in all the cases plotted. There are a couple reasons for this.

1. In the case of C^2_{CuMS} and C^2_{PrMS}, $\mathcal{S}_1$ does not perform rate-splitting, thereby rendering the receivers of $\mathcal{S}_2$ and $\mathcal{S}_3$ vulnerable to interference caused due to $\mathcal{S}_1$'s transmissions. In the case of C_{CoMS}, neither $\mathcal{S}_1$ nor $\mathcal{S}_2$ performs rate-splitting, leading to poor interference management at all the receivers. However, several corollaries were derived based on the idea of allowing senders to dedicate (part of) their power for transmitting the primary sender's message, which resulted in several additional rate points being achievable. And, it was shown that these rate tuples meet the outer bounds at several corner points. A systematic way of expanding the rate region by including the different coding schemes is an open problem, which can be explored by future researchers.

2. The outer bounds were derived by taking the intersection of the capacity region with bidirectional sharing and the individual user rates with unidirectional sharing, and hence have a natural advantage over the purely-unidirectional model assumed in deriving the rate regions. Furthermore, the duality result implicitly assumes that the receivers can successfully decode the interfering signals to a large extent.

Outer bounds can be made tighter by considering discrete memoryless channel models and introducing auxiliary RVs. Existing literature lacks results for tighter outer bounds, except for the most recent work in [48] which are for two-user CR channels not directly applicable to our channel models. From the above discussion, we conclude that, though the three-user channel models considered in this work are logical extensions of the classical two-user scenario, we are able to make interesting observations and draw several inferences on the effect of rate-splitting and message-sharing on *larger* networks. The techniques to analyze two-user networks may carry over to these larger networks, but issues related to interference management via rate-splitting are nontrivial and need further investigation.

2.4.2 Effect of reduction in size of the network

Let us consider the case of removing a transmitter-receiver pair from the three-user CR channel model. In particular, let us assume that $(\mathcal{S}_3, \mathcal{R}_3)$-pair is removed, resulting in a two-user CR channel. We make the following observations:

1. The channels $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$ will now reduce to the model employed in [36]. The achievability scheme results in a rate region which coincides with [36, Theorem 1], which includes the rate regions derived in [34] and [35]. Furthermore, the rate regions derived in [34, Theorem 3.5] and [35, Theorem 4.1] are in fact the capacity regions for the two-user CR channels in the *low-interference* regime. In our three-user channel models $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$, low-interference regime can be considered by letting the auxiliary RVs U_1 and V_1 be constants. However, this does not yield the capacity region (unlike the

two-user scenario) for $\mathcal{C}^2_{\text{CuMS}}$ and $\mathcal{C}^2_{\text{PrMS}}$.

2. The channels $\mathcal{C}^1_{\text{CuMS}}$ and $\mathcal{C}^1_{\text{PrMS}}$ will reduce to the model employed in [32]. However, our achievability scheme results in a slightly larger rate region compared to the one presented in [32, Theorem 1]. This is because of the fact that the rate region of [32, Theorem 1] takes into account *noisy message-sharing* (captured by Eq. $(3)-(5)$ in [32, Theorem 1]), while our problem setup concerns *degraded message sets*.

3. For the channel $\mathcal{C}^1_{\text{CoMS}}$, let us consider removing $(\mathcal{S}_2, \mathcal{R}_2)$-pair. This results in the model employed in [36], which has been addressed in the above discussion.

2.4.3 A recent result on the two-user CR channel

Recently, in [48, Theorem V.1] a new inner bound has been derived for the two-user CR channel which encompasses all of the previously known achievable regions. The technique used to prove their main achievability theorem employs rate-splitting, superposition coding and a sequential binning procedure. However, we notice that their achievability scheme involves a rate-split at both encoders. This does not conform well to some of our channel models, specifically $\mathcal{C}^2_{\text{CuMS}}$, $\mathcal{C}^2_{\text{PrMS}}$ and $\mathcal{C}_{\text{CoMS}}$, where one (or two) encoder(s) do not employ rate-splitting. This suggests that their technique may not be appropriate for our problem setup since, owing to the presence of three transmitters, more than one rate-splitting scheme can be considered leading to a large class of three-user channel models. Furthermore, it remains to be ascertained whether the achievability scheme of [48] generalizes to the three-user channel irrespective of the rate-splitting technique employed by the encoders.

A new outer bound has also been derived in [48, Theorem IV.1], which is looser than previously known outer bounds, but has the advantage of not involving auxiliary RVs. However, we have not investigated outer bounds for the discrete memoryless case, because of the complexity of our problem setup. Instead, we

resorted to obtaining outer bounds for the Gaussian channel model. Deriving tighter inner and outer bounds for the discrete memoryless channel model of our problem setup is challenging and is an interesting open problem.

2.5 Conclusions

We introduced multiuser channels with noncausal transmitter cooperation in the overlay cognitive radio network paradigm and presented three different ways of message sharing which we termed cumulative message sharing (CuMS), primary-only message sharing (PrMS) and cognitive-only message sharing (CoMS). We derived an achievable rate region for each of the channels by employing a combination of superposition and Gel'fand-Pinsker coding techniques. We considered the Gaussian channel model to plot the rate regions and presented some corollaries using which several achievable rate tuples for the Gaussian channel were identified. Later, we derived outer bounds for the Gaussian case by considering bidirectional cooperation between the transmitters, and calculating the capacity region of the resulting Gaussian MIMO broadcast channel using BC-MAC duality results. Simulation results enabled us to compare rate-splitting and message-sharing as a mechanism to improve spectral efficiency. We observed that, while message-sharing is superior to rate-splitting in both two and three-user scenarios, it is not fully clear as to which type of message-sharing mechanism (CuMS, PrMS or CoMS) gives the largest rate region. Open problems include deriving tighter outer bounds; considering rate-constrained cooperation, wherein the cognitive radio estimates the message index transmitted by the primary user in a causal manner; and characterizing the tradeoff between message sharing and rate-splitting in a multi-user cognitive network.

Chapter 3

Opportunistic relay channels

3.1 Introduction

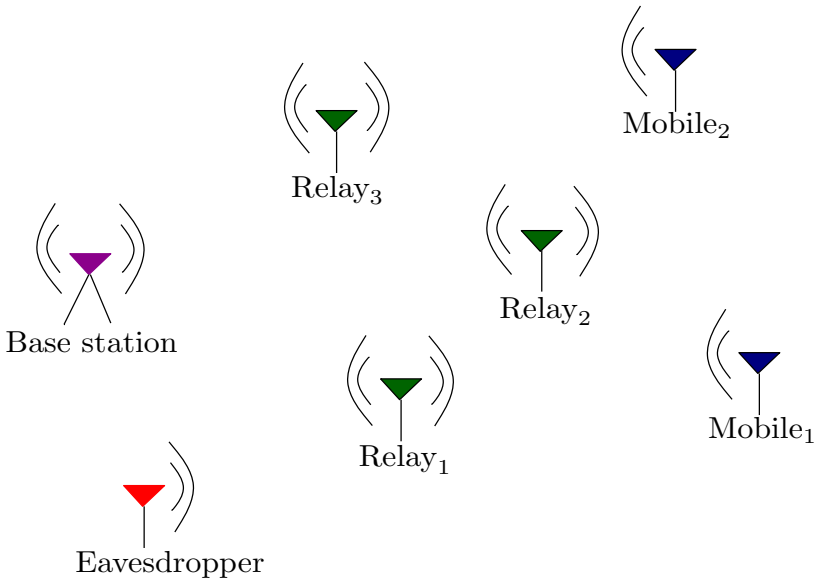

Figure 3.1: A schematic of a relay network aiding cellular infrastructure.

With increasing emphasis on efficient utilization of radio-frequency spectrum and growing interest in providing wireless services with higher data rates, cooperative communications has been proposed as a key enabling technology for next generation wireless networks. User-cooperation is especially popular in multiple node networks, where a node expresses willingness to share its resources (transmit power, computation, *etc.*) with its neighboring nodes with the objective

of improving the overall performance of the network in terms of its throughput. Information-theoretic studies are rigorously pursued to understand the fundamental performance limitations of reliable communications in such cooperative scenarios [54]. In this chapter, we consider the relay network [55], [56], which has emerged as a strong contender to realize cooperation in a wireless infrastructure and derive the performance limits for two different communication scenarios.

We motivate the problems addressed in this chapter through the following example. Consider a typical cellular environment (see Fig. 3.1), in which the remotely located Mobile_1 transmits to the Base station over the broadcast medium using the uplink communication protocols. In order to ensure higher date rates, Mobile_1 enlists the help of relay nodes located at various points in space; relays transmit replicas of the signal of interest, providing diversity and thereby improving spectral efficiency. Relay nodes participating in a cooperative-communication scenario can also use this opportunity to transmit their own data to intended terminals [57] leading to improved spectrum efficiency, and thereby broadly fall into the *cognitive radio network* paradigm. Although user-cooperation and cognition have benefits, the broadcast nature of wireless medium exposes problems related to information security. In the scenario considered in Fig. 3.1, the terminal denoted Eavesdropper can gain unauthorized access to the wireless link between Relay_1 and Base station. That is, the broadcast nature of wireless networks facilitates malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, *etc.*

However, due to geographical separation, it might not be possible for Eavesdropper to hear from Mobile_1, even though we have considered the wireless broadcast medium. In cellular architecture, such scenarios are a commonplace when the eavesdropper is outside the coverage area of the mobile device, but lies in close proximity of an intermediate relay node. Furthermore, the transmit power constraint on the mobile device is a significant factor for Eavesdropper to remain oblivious to the transmissions of Mobile_1. Such an example is admittedly contrived. Nonetheless, it provides a useful model for cases where the mobile device is less

sophisticated to tackle the malicious intent of an eavesdropper, and where the relay node has advanced functionalities to achieve secure communications.

In this chapter, we present an information-theoretic viewpoint of the joint problem of cooperation, cognition and information security/confidentiality over such relay networks. Although the benefits of user-cooperation for secure communications have been reported in the literature (for *e.g.*, see [58] - [60]), to the best of the author's knowledge, this is the first instance where the three issues - cooperation, cognition and confidentiality - are *simultaneously* addressed. It is worthwhile to note that, in the information theory literature, the terms cooperation and cognition have been used interchangeably, but in this chapter we make a clear distinction between the two. We also note that, confidentiality is commonly referred to as information-theoretic/wireless physical-layer security [6].

3.1.1 Communication scenarios

We consider a four node wireless network whose channel schematic is shown in Fig. 3.2, and define the following two communication scenarios:

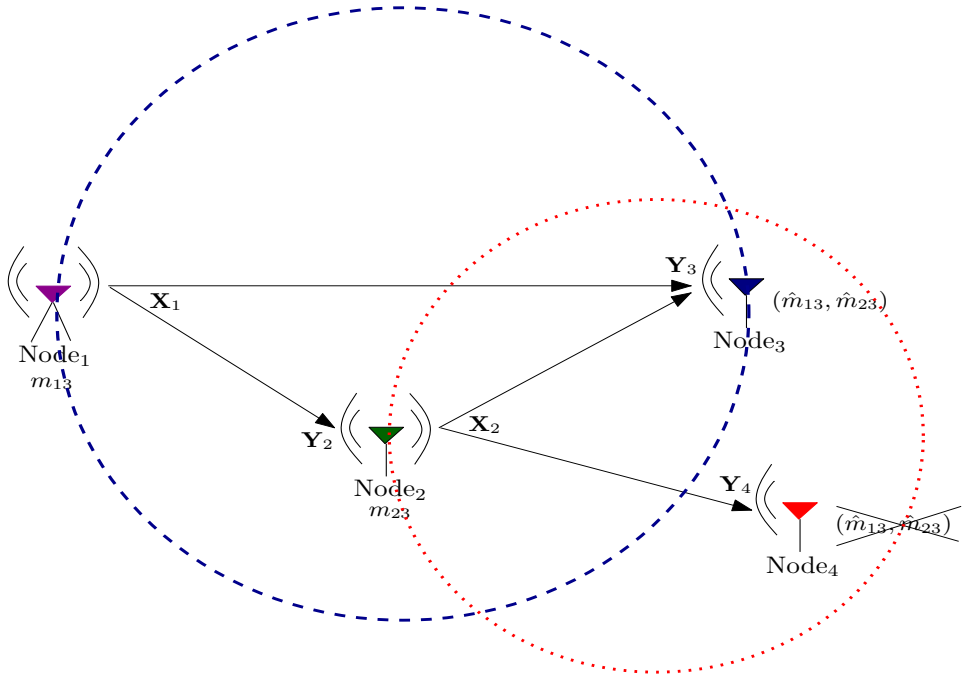

Figure 3.2: Relay network with cooperation, cognition and information security.

1. **Scenario I**, which captures the essence of the following three issues:

   (a) *Cooperation:* Node 1 intends to transmit a message m_{13} to Node 3. Node 2 is used as a cooperative-relay to aid transmissions from Node 1 to Node 3. It employs decode-process-forward mechanism to improve the spectral efficiency of (Node 1, Node 3)-pair.

   (b) *Cognition:* Node 2 also uses this opportunity to transmit its own message m_{23} to Node 3. We term such type of relays as *opportunistic-relays*; they can also be considered as cognitive-relays, since they not only aid other's transmissions, but also better utilize the spectrum by transmitting their own messages.

   (c) *Security:* Transmissions from Node 2 are also received by Node 4, who is considered to be an eavesdropper with malicious intent, unauthorized to participate in the communication scenario. Therefore, Node 2 is constrained to keep the message of Node 1 (m_{13}) and its own message (m_{23}) confidential from Node 4.

   For this scenario, we let the channel to be physically degraded, since the processing to achieve secrecy is performed at the relay.

2. **Scenario II**, with the following setup:

   (a) Node 1 intends to transmit a message m_{13} to Node 3, by deeming Node 2 to be untrustworthy. Therefore, Node 1 is constrained to keep m_{13} secret from Node 2.

   (b) Node 2 transmits its message m_{23} to Node 3, by keeping it confidential from Node 4.

   This scenario is not applicable for physically degraded relay channels, since the secrecy capacity will be zero as pointed out in [61].

Note 3.1.1. 1: In both scenarios, we consider Node 4 to be geographically located outside the transmission range of Node 1, and that it remains oblivious to signals

transmitted by Node 1. Owing to the broadcast nature of the wireless medium, Node 4 can get to hear the transmissions of Node 1. However, if the geographical separation is significant, then the received signal-to-noise ratio at Node 4 due to Node 1 will be negligible; it is reasonable to assume that Node 4 receives a highly corrupted and noisy version of the signal from Node 1, making it difficult for Node 4 to infer anything about Node 1's transmissions. Though this is not the worst case analysis, the assumption is well motivated by real world scenarios (described in the previous paragraph), and provides a basis to jointly address the aforementioned issues. We also note that, in **Scenario II**, Node 2 tries to decode the message of Node 1 (although, unsuccessfully), and the model conforms to the classical relay setting.

3.1.2 Main contribution

We derive capacity bounds for secure and reliable communications for the above two described scenarios. Towards this end, we introduce a novel achievability scheme, namely layered coding, to derive lower bounds on the capacity regions of the two communication scenarios. Outer bounds are derived using auxiliary random variables for single-letter characterization. We compare the layered coding scheme with the noise-insertion strategy, which is prominently used in the existing literature, and explain why layered principle is better suited for opportunistic-relays. We characterize the rate-penalty that incurs for having to deal with security constraints on the messages. We also argue that the channel models presented in this chapter for opportunistic-relays are better - from a practical viewpoint - compared to the classical cognitive radio network model for efficient radio-frequency spectrum utilization.

Note that, we are primarily concerned with establishing the theoretical performance limits of the relay networks considered in this chapter, without dealing with the practical realization of such systems. The interested reader is referred to [62] - [65], where communication infrastructure required for multi-hop relay

networks, along with implementation and performance evaluation of relay-based wireless networks are addressed. It is of interest to note that, for the case of opportunistic-relays (where the relays have their own messages) the physical-layer aspects like modulation and signal design, and medium-access control layer schemes suggested in the above cited references should be appropriately modified. However, these issues are beyond the scope of this chapter.

3.1.3 Organization of the chapter

In Section 3.2, we provide references from the existing literature that closely relate to our chapter. In Section 5.2, we introduce the notation used and provide a mathematical model for the relay network considered in this chapter. In Section 5.3, we describe achievable rate regions and outer bounds on the capacity region of the above described communication scenarios. In Section 3.5, we outline the layered coding principle and provide related discussion. We conclude the chapter in Section 4.6. The proofs of the achievability theorem, outer bounds and a required lemma are relegated to appendices.

3.2 Related work

The problem of confidentiality/security in relay networks has been addressed along various lines in the information theory literature. Capacity bounds for cooperation in wireless networks was presented in [66], where authenticated relay nodes employ noise insertion strategies [67] to achieve secrecy. In [68], an opportunistic selection technique of two relay nodes was presented to secure communications between a source-destination pair from the eavesdropper. The first relay employed a simple decode-and-forward strategy, while the second relay is used to create intentional interference at the eavesdropper, thereby jamming its reception. The relay-eavesdropper channel was considered in [69], where user-cooperation has been exploited to achieve secrecy. In particular, the relay node employed a novel noise-forwarding strategy to confuse the eavesdropper. How-

ever, the relay was considered to be a *deaf-helper*, in the sense that it is totally ignorant of the transmitted messages.

In some other scenarios, the relay was treated as an eavesdropper. For example, in [70], confidential messages were transmitted to a receiver which also served as a relay to its neighboring node. The trade-off between cooperation and secrecy was characterized, along with the rate-equivocation region for the discrete memoryless network. Some jamming strategies as a means to increase secrecy for the Gaussian channel case was also proposed. Likewise, in [71], capacity bounds were derived when cooperation is achieved with the help of untrustworthy relays. Specifically, two models of relay networks with orthogonal components was studied. In the first model, where there is an orthogonal link from the source to the relay, it was shown that cooperating with an untrustworthy relay was never beneficial. However, cooperation was shown to be beneficial in the second model, where there is an orthogonal link from the untrustworthy relay to the destination.

Other noteworthy contributions include cooperation over a two-hop communication network using untrustworthy relays [72], capacity results for orthogonal relay eavesdropper channels [73], cooperative relay broadcast channels [74] and improving wireless security via multiple cooperating relays in the presence of one or more eavesdroppers [75]. Coding for relay channels, where the relay acts as an eavesdropper, was reported in [61].

Most of the results in the existing literature pertain solely to *cooperative* relays, in the sense that the relay nodes aid the communication between a sender-receiver pair, though they may or may not contribute to achieve secrecy. In this chapter, we consider the case where the relay (Node 2) has a private message intended to the destination (Node 3). Hence, these relays can be thought of as *cognitive* relays in addition to being cooperative, in the sense that it can opportunistically utilize the scarce radio-frequency spectrum. A similar model has been considered in [57], without addressing the security issues considered in this chapter.

3.3 System Model and Preliminaries

In this chapter, we only consider the discrete memoryless versions of the relay channels described in Section 3.1. We denote the relay channel characterizing **Scenario I** by C; and the channel characterizing **Scenario II** by C^*. Discrete random variables (RV) defined on finite sets $X_1 \in \mathcal{X}_1$ and $Y_3 \in \mathcal{Y}_3$ denote the input and output at $\mathrm{Node}\ 1$ and $\mathrm{Node}\ 3$, respectively. $X_2 \in \mathcal{X}_2$ and $Y_2 \in \mathcal{Y}_2$ denote the input and output, respectively, at $\mathrm{Node}\ 2$, while the output at $\mathrm{Node}\ 4$ is denoted $Y_4 \in \mathcal{Y}_4$. The channels are assumed to be memoryless and is characterized by the conditional distribution

$$p(y_2^N, y_3^N, y_4^N | x_1^N, x_2^N) = \prod_{n=1}^{N} p(y_{2,n}, y_{3,n}, y_{4,n} | x_{1,n}, x_{2,n}),$$

where N is the number of channel uses. The lower case letters y_2, y_3, y_4, x_1 and x_2 are particular realizations of the corresponding RVs. Further, the channel C will be considered to be physically degraded, so that

$$p(x_1^N, x_2^N, y_2^N, y_3^N, y_4^N) = \prod_{n=1}^{N} p(y_{2,n} | x_{1,n}, x_{2,n}) p(y_{3,n}, y_{4,n} | y_{2,n}, x_{2,n}).$$

To transmit its message, $\mathrm{Node}\ 1$ generates an RV $M_{13} \in \mathcal{M}_{13}$, where $\mathcal{M}_{13} = \{1, \ldots, 2^{NR_{13}}\}$ denotes a set of message indices. Without loss of generality, $2^{NR_{13}}$ is assumed to be an integer, with R_{13} being the transmission rate of $\mathrm{Node}\ 1$. M_{13} denotes the message $\mathrm{Node}\ 1$ intends to transmit to $\mathrm{Node}\ 3$, and is assumed to be independently generated and uniformly distributed over the finite set $\mathcal{M}_{13}$. Integer m_{13} is a particular realization of M_{13} and denotes the message-index. $\mathrm{Node}\ 2$ has message $M_{23} \in \mathcal{M}_{23} = \{1, \ldots, 2^{NR_{23}}\}$ intended to $\mathrm{Node}\ 3$. The symbols M_{23}, $\mathcal{M}_{23}$, m_{23} and R_{23} are similarly notated.

1. For the channel C, a $((2^{NR_{13}}, 2^{NR_{23}}), N, P_e^{(N)})$ code comprises:

   (a) An encoder $f_1 : \mathcal{M}_{13} \to \mathcal{X}_1^N$,

   (b) A set of N relay functions $\{r_n\}_{n=1}^{N}$, such that $x_{2,n} = r_n(Y_{2,1}, \ldots, Y_{2,n-1}, M_{23})$,

$1 \leq n \leq N$. Further, to ensure information security, the relay makes use of a stochastic encoder which is defined by the matrix of conditional probabilities $\phi(x_2^N|\bar{m}_{13}, m_{23})$, such that
$\sum_{x_2^N} \phi(x_2^N|\bar{m}_{13}, m_{23}) = 1$. $\bar{m}_{13}$ is the guess of m_{13} made by the relay. $\phi(x_2^N|\bar{m}_{13}, m_{23})$ denotes the probability that a pair of message-indices $(\bar{m}_{13}, m_{23})$ is encoded as $x_2^N \in \mathcal{X}_2^N$ to be transmitted by the relay.

(c) Two decoders - $g_2 : \mathcal{Y}_2^N \to \mathcal{M}_{13}$, $g_3 : \mathcal{Y}_3^N \to \mathcal{M}_{13} \times \mathcal{M}_{23}$.

The average probability of decoding error for the code, averaged over all codes, is
$P_e^{(N)} = \max\{P_{e,2}^{(N)}, P_{e,3}^{(N)}\}$, where,

$$\begin{aligned} P_{e,2}^{(N)} &= \textstyle\sum_{m_{13}} \frac{1}{2^{NR_{13}}} \mathbf{Pr}\left[g_2(\mathcal{Y}_2^N) \neq m_{13}|m_{13} \text{ sent}\right], \\ P_{e,3}^{(N)} &= \textstyle\sum_{\mathbf{m}} \frac{1}{2^{N[R_{13}+R_{23}]}} \mathbf{Pr}\left[g_3(\mathcal{Y}_3^N) \neq \mathbf{m}|\mathbf{m} \text{ sent}\right], \end{aligned}$$

where $\mathbf{m} = (m_{13}, m_{23})$. A rate pair (R_{13}, R_{23}) is said to be achievable for the channel C, if there exists a sequence of $(2^{NR_{13}}, 2^{NR_{23}}, N, P_e^{(N)})$ codes and any $\epsilon > 0$, such that $P_e^{(N)} \to 0$ as $N \to \infty$ and the following secrecy constraints are satisfied:

$$NR_{13} - H(M_{13}|Y_4) \leq N\epsilon, \tag{3.1}$$

$$NR_{23} - H(M_{23}|Y_4) \leq N\epsilon, \tag{3.2}$$

where $H(\mathrm{x}|\mathrm{y})$ is the conditional entropy of x given y. The capacity region is defined as the closure of the set of all achievable rate tuples (R_{13}, R_{23}).

2. For the channel C*, a $((2^{NR_{13}}, 2^{NR_{23}}), N, P_e^{(N)})$ code comprises:

(a) Two stochastic encoders defined by the matrix of conditional probabilities $\phi^*(x_1^N|m_{13})$ and $\phi^*(x_2^N|m_{23})$, such that $\sum_{x_t^N} \phi^*(x_t^N|m_{t3}) = 1$. $\phi^*(x_t^N|m_{t3})$ denotes the probability that the message-index m_{t3} is encoded as $x_t^N \in \mathcal{X}_t^N; t = 1, 2$.

(b) A decoder - $g_3 : \mathcal{Y}_3^N \rightarrow \mathcal{M}_{13} \times \mathcal{M}_{23}$.

The average probability of decoding error for the code, averaged over all codes, is

$$P_e^{(N)} = \sum_{\mathbf{m}} \frac{1}{2^{N[R_{13}+R_{23}]}} \Pr\left[g_3(\mathcal{Y}_3^N) \neq \mathbf{m} | \mathbf{m} \text{ sent}\right],$$

where $\mathbf{m} = (m_{13}, m_{23})$. A rate pair (R_{13}, R_{23}) is said to be achievable for the channel C*, if there exists a sequence of $(2^{NR_{13}}, 2^{NR_{23}}, N, P_e^{(N)})$ codes and any $\epsilon > 0$, such that $P_e^{(N)} \rightarrow 0$ as $N \rightarrow \infty$ and the following weak-secrecy constraints are satisfied:

$$NR_{13} - H(M_{13}|Y_2) \leq N\epsilon, \tag{3.3}$$

$$NR_{23} - H(M_{23}|Y_4) \leq N\epsilon, \tag{3.4}$$

The secrecy constraint (3.3) signifies that in order to keep the message of the source (m_{13}) confidential from the relay, the source has to transmit at a rate R_{13} which must be less than the conditional entropy $H(M_{13}|Y_2)$. Similarly, the secrecy condition (3.4) states that, in order to keep the message of the relay (m_{23}) confidential from the eavesdropper, the relay has to transmit at a rate R_{23} which must be less than the conditional entropy $H(M_{23}|Y_4)$. The capacity region is defined as the closure of the set of all achievable rate tuples (R_{13}, R_{23}).

In the remainder of this chapter, the following notation is used. For any $\epsilon > 0$, we denote by $A_\epsilon^{(N)}(P_X)$ an ϵ-typical set comprising sequences picked from the distribution $p(x)$.

3.4 Summary of results

In this section, we present achievable rate regions and outer bounds for C and C*. We consider the following auxiliary RVs defined on finite sets: $W \in \mathcal{W}$, $U \in \mathcal{U}$,

$V \in \mathcal{V}$ and $Z \in \mathcal{Z}$. For C, let $\mathcal{P}$ denote the set of all joint probability distributions $p(w, u, v, z, x_1, x_2, y_2, y_3, y_4)$ that is constrained to factor as follows:

$$\begin{aligned} p(w, u, v, z, x_1, x_2, y_2, y_3, y_4) &= p(w, u)p(x_1|w, u)p(v, z|w) \\ &\times p(x_2|w, v, z)p(y_2|x_1, x_2)p(y_3|x_2, y_2)p(y_4|x_2). \end{aligned}$$

The auxiliary RVs serve a dual purpose. On one hand they render the channel causal, while on the other hand they represent the sources to be transmitted when the encoder has side-information to deal with. Note also that, establishing the cardinality bounds of these auxiliary RVs is a tedious task and is not considered in this chapter. The interested reader to referred to [52, Chapter 15] for a brief exposition on the cardinality bounds of auxiliary RVs for simple channel models.

3.4.1 Achievable region for C

For a given $p(.) \in \mathcal{P}$, an achievable rate region for C is described by the set $\mathfrak{R}_{\text{in}}(p)$, which is defined as the convex-hull of the set of all rate pairs (R_{13}, R_{23}) that simultaneously satisfy (3.5) - (3.7).

$$0 \leq R_{13} \leq \min\{I(U; Y_2|W, V, Z), I(W, U, V; Y_3|Z) - \max[H(W), I(W, V; Y_4)] \tag{3.5}$$

$$0 \leq R_{23} \leq I(Z; Y_3|W, U, V) - I(Z; Y_4), \tag{3.6}$$

$$0 \leq R_{13} + R_{23} \leq I(W, U, V, Z; Y_3) - \max[H(W), I(W, V; Y_4)] - I(Z; Y_4). \tag{3.7}$$

Theorem 3.4.1. Let $\mathfrak{C}$ denote the capacity region of the channel C. Let $\mathfrak{R}_{\text{in}} = \bigcup_{p(.)\in\mathcal{P}} \mathfrak{R}_{\text{in}}(p)$. The region $\mathfrak{R}_{\text{in}}$ is an achievable rate region for C, *i.e.*, $\mathfrak{R}_{\text{in}} \subseteq \mathfrak{C}$.

The proof of Theorem 3.4.1 can be found in Appendix B.1. From the rate region, we clearly see that one has to incur a rate-penalty for having to secure the messages from the eavesdropper. Specifically, the quantities $I(W, V; Y_4)$ in (3.5), $I(Z; Y_4)$ in (3.6), and $I(W, V; Y_4)$ and $I(Z; Y_4)$ in (3.7) signify the rate-penalties $\mathrm{Node}\ 1$ and $\mathrm{Node}\ 2$ incur due to the presence of the eavesdropping $\mathrm{Node}\ 4$. In the absence of an eavesdropper, the rate region reduces the one presented in [57]

for relay channels with private messages. The interesting observation is the case when Node 2 does not have a message for Node 3, *i.e.*, $\mathcal{Z} = \{\emptyset\}$. In this scenario, $R_{13} \leq \min\{I(U;Y_2|W,V,Z), I(W,U,V;Y_3|Z) - \max[H(W), I(W,V;Y_4)]\}$. Comparing this with the achievable rate-equivocation using noise-insertion strategy (see [69]), we notice that the layered coding scheme experiences a *bottleneck* if the channel between Node 1 and Node 2 is noisier than the one between Node 1 and Node 3. In Section 3.5.3, we provide more discussion related to this issue by highlighting the relative merits and demerits of the layered coding principle compared to the noise-insertion strategy.

3.4.2 Outer bounds for C

For a given $p(.) \in \mathcal{P}$, an outer bound for C is described by the set $\mathfrak{R}_{\text{out}}(p)$, which is defined as the convex-hull of the set of all rate pairs (R_{13}, R_{23}) that simultaneously satisfy (3.8) - (3.10).

$$0 \leq R_{13} \leq I(W,U,V;Y_3|Z) - I(V;Y_4), \tag{3.8}$$

$$0 \leq R_{23} \leq I(Z;Y_3|V) - I(Z;Y_4), \tag{3.9}$$

$$0 \leq R_{13} + R_{23} \leq I(W,U,V,Z;Y_3) - I(V;Y_4) - I(Z;Y_4). \tag{3.10}$$

Theorem 3.4.2. Let $\mathfrak{C}$ denote the capacity region of the channel C. Let $\mathfrak{R}_{\text{out}} = \bigcup_{p(.)\in\mathcal{P}} \mathfrak{R}_{\text{out}}(p)$. The region $\mathfrak{R}_{\text{out}}$ is an outer bound for C, *i.e.*, $\mathfrak{C} \subseteq \mathfrak{R}_{\text{out}}$.

The proof of Theorem 3.4.2 can be found in Appendix B.2. The outer bounds are derived utilizing the secrecy constraints (3.1) - (3.2).

For C*, let $\mathcal{P}^*$ denote the set of all joint probability distributions $p(w,u,z,x_1,x_2,y_2,y_3,y_4)$ that is constrained to factor as follows:

$$\begin{aligned} p(w,u,z,x_1,x_2,y_2,y_3,y_4) &= p(w,u)p(x_1|w,u)p(z|w) \\ &\quad \times p(x_2|w,z)p(y_2|x_2)p(y_3|x_2,y_2)p(y_4|x_2). \end{aligned}$$

3.4.3 Achievable region for C^*

Given $p(.) \in \mathcal{P}^*$, an achievable rate region for C^* is described by the set $\mathfrak{R}^*_{\text{in}}(p)$, which is defined as the convex-hull of the set of all rate pairs (R_{13}, R_{23}) that simultaneously satisfy (3.11) - (3.13).

$$\begin{aligned} 0 \leq R_{13} &\leq I(W,U;Y_3|Z) - \max[H(W), I(W,U;Y_2)], && (3.11)\\ 0 \leq R_{23} &\leq I(Z;Y_3|W,U) - I(Z;Y_4), && (3.12)\\ 0 \leq R_{13} + R_{23} &\leq I(W,U,Z;Y_3) - \max[H(W), I(W,U;Y_2)] - I(Z;Y_4). && (3.13) \end{aligned}$$

Theorem 3.4.3. Let $\mathfrak{C}^*$ denote the capacity region of the channel C^*. Let $\mathfrak{R}^*_{\text{in}} = \bigcup_{p(.)\in\mathcal{P}^*} \mathfrak{R}^*_{\text{in}}(p)$. The region $\mathfrak{R}^*_{\text{in}}$ is an achievable rate region for C^*, *i.e.*, $\mathfrak{R}^*_{\text{in}} \subseteq \mathfrak{C}^*$.

The proof of Theorem 4.3.4 can be found in Appendix B.3. In this case, we see that $\mathrm{Node}\ 1$ incurs a rate-penalty for securing the message from $\mathrm{Node}\ 2$ (signified by $\max[H(W), I(W,U;Y_2)]$ in (3.11), while $\mathrm{Node}\ 2$ incurs a penalty $I(Z;Y_4)$ in (3.12) for keeping its message secret from $\mathrm{Node}\ 4$. The benefits of using layered coding over noise-insertion for this scenario is discussed in Section 3.5.3.

3.4.4 Outer bounds for C^*

For a given $p(.) \in \mathcal{P}^*$, an outer bound for C^* is described by the set $\mathfrak{R}^*_{\text{out}}(p)$, which is defined as the convex-hull of the set of all rate pairs (R_{13}, R_{23}) that simultaneously satisfy (3.14) and (3.16).

$$\begin{aligned} 0 \leq R_{13} &\leq I(W,U;Y_3|Z) - I(U;Y_2), && (3.14)\\ 0 \leq R_{23} &\leq I(Z;Y_3|U) - I(Z;Y_4), && (3.15)\\ 0 \leq R_{13} + R_{23} &\leq I(W,U;Y_3|Z) + I(W,Z;Y_3|U) - I(U;Y_2) - I(Z;Y_4). && (3.16) \end{aligned}$$

Theorem 3.4.4. Let $\mathfrak{C}^*$ denote the capacity region of the channel C^*. Let $\mathfrak{R}^*_{\text{out}} = \bigcup_{p(.)\in\mathcal{P}^*} \mathfrak{R}^*_{\text{out}}(p)$. The region $\mathfrak{R}^*_{\text{out}}$ is an outer bound for C^*, *i.e.*, $\mathfrak{C}^* \subseteq \mathfrak{R}^*_{\text{out}}$.

The proof of Theorem 4.3.5 can be found in Appendix B.4.

3.5 Discussion

In this section, we first present a high-level explanation of some well-known coding schemes, namely binning; block Markov superposition coding; and backward decoding to familiarize the reader with these important techniques. Then we present an outline of the coding scheme that we have devised to derive lower bounds on the capacity regions of the two communication scenarios described in the preceding sections. We then discuss our coding strategy in comparison with the noise-forwarding technique that has been used in the existing literature. Finally, we compare our channel model with the classical cognitive radio channel setup.

3.5.1 Some standard coding techniques

In this subsection, we present the salient features of those standard coding principles - binning; block Markov superposition coding; and backward decoding - which forms the basis for the coding technique devised in this chapter.

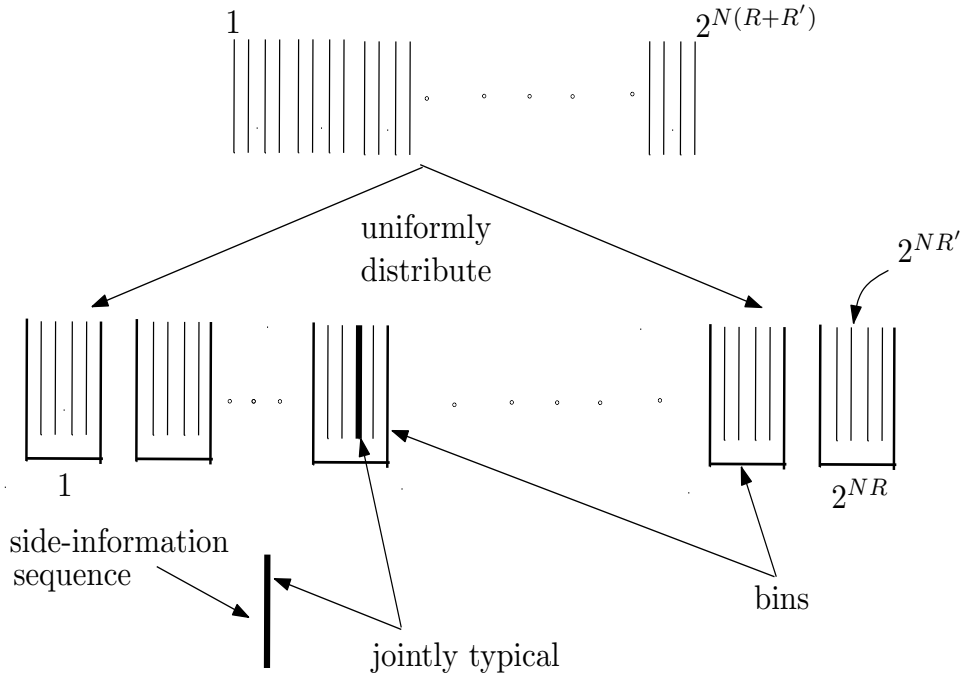

Figure 3.3: Binning principle.

In binning [76], let us suppose there are 2^{NR} messages to be transmitted across the channel. Then generate $2^{N(R+R')}$ independent sequences, and uniformly distribute them into 2^{NR} bins so that each bin comprises $2^{NR'}$ sequences. To transmit a message $k \in \{1, \ldots, 2^{NR}\}$, go to the bin indexed by k and pick a sequence which is jointly typical with a side-information codeword that is available at the encoder in a noncausal/causal manner. A pictorial representation of binning is shown in Fig. 3.3.

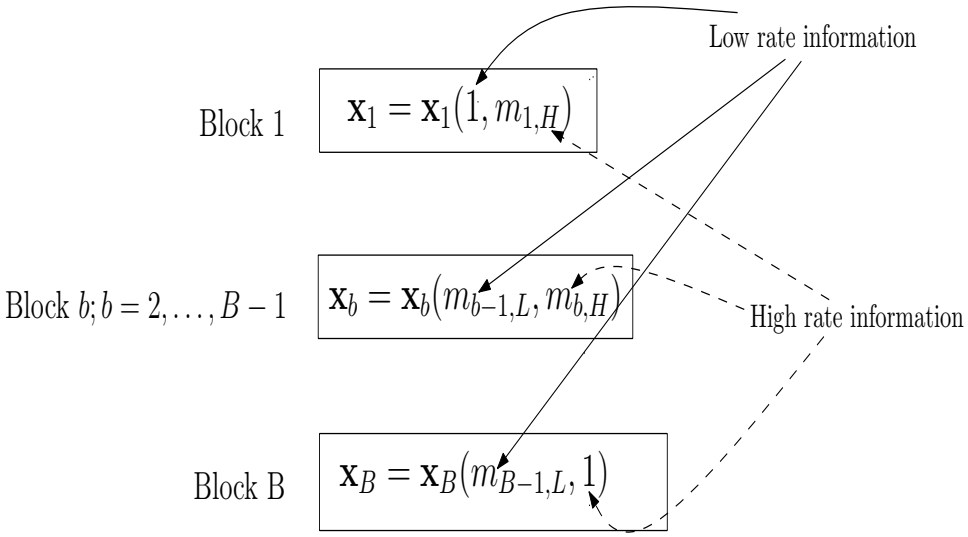

Figure 3.4: Markov superposition coding.

In Markov superposition coding [19], transmissions are in the form of blocks of coded data (see Fig. 3.4). In each block, there are two types of data being encoded. In block b, the message to be transmitted is coded at a rate higher than what the receiver can actually decode. In block $b+1$ the high-rate message of block b is coded at a low rate, upon which is superimposed a new codeword which will be at a higher rate, and this process continues till block B.

In backward decoding [19], all the B blocks of data are accumulated at the decoder, and the codewords are decoded from the *last* block. Furthermore, in block b only the low-rate information is decoded. By doing so, the high-rate infor-

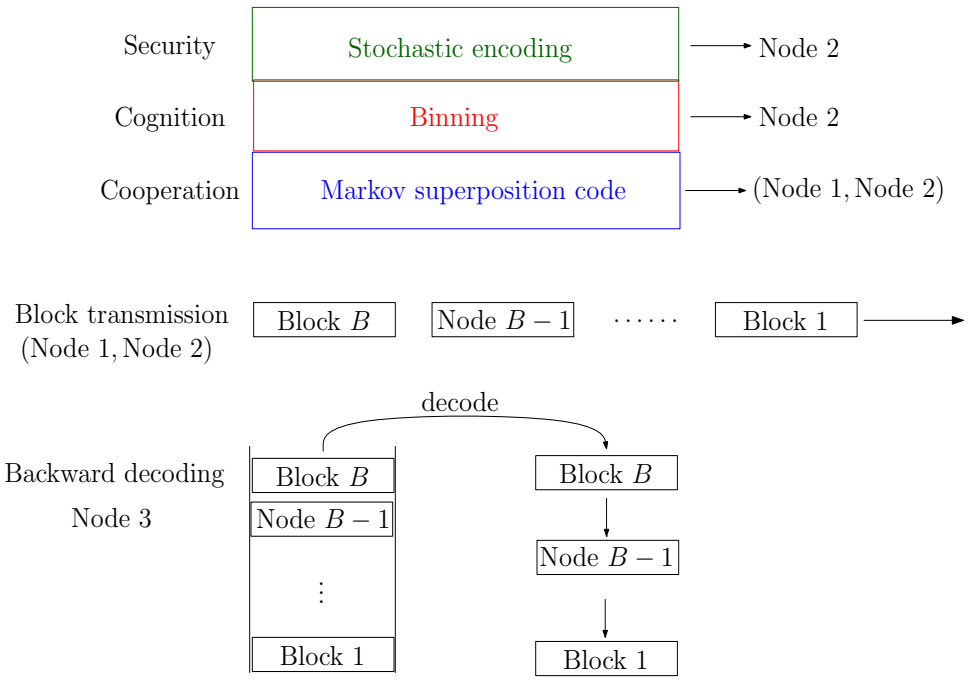

Figure 3.5: Layered coding architecture.

mation of block $b-1$ is automatically decoded (because of the block construction during encoding). And, this process continues till block 2, which also decodes the high-rate message of block 1.

3.5.2 Layered coding

The achievability scheme devised to prove the coding theorems in this chapter is termed "layered coding", which is a combination of block Markov superposition coding [17], binning [18] and stochastic encoding [20]. A schematic of the layered coding architecture is shown in Fig. 3.5. For **Scenario I**, we employ the following layering principle: The bottom most layer is the Markov superposition code, which is devised to realize cooperation between Node 1 and Node 2. Random binning is implemented in the middle layer to enable Node 2 to opportunistically transmit its own messages to Node 3. This is similar to the coding principle employed in the classical cognitive radio channels, where the secondary transmitter uses random binning against known interference from the primary user (for instance, see [32], [35]). The topmost layer incorporates stochastic encoding, using which

Node 2 keeps the messages of Node 1 and its own messages secure from Node 4. For **Scenario II**: Node 1 uses the block Markov superposition code and stochastic encoding, while Node 2 employs binning and stochastic encoding in conjunction with Markov superposition coding to achieve the described rates. Block transmission is used to transmit encoded data, while at the receiver we employ backward [19] and simultaneous decoding [7] techniques to recover the transmitted information.

3.5.3 Comparison with some existing results

1. **Scenario I:**

   (a) This scenario is similar to the models presented in [66], [69] in that the relay plays an active role in confusing the eavesdropper to achieve information confidentiality. In order to achieve this objective, in this chapter we employ a stochastic encoder at the relay node, where we generate additional codewords and *randomly* transmit one of them. Equivocation calculations show that the additional uncertainty confuses the eavesdropper, who cannot decode the received signal with arbitrarily small probability of error. Whereas, in case of [66] and [69], special noise insertion strategies (for example, see [67]) are employed, where the relay node sends codewords independent of the source message to confuse the eavesdropper. Also, in our chapter, the main sender (Node 1) is oblivious to the presence of the eavesdropper (Node 4) and does not play an active role in ensuring secrecy. Therefore, the relay node in our chapter cannot be a *deaf helper* by remaining ignorant of the main sender's message. In fact, in our chapter, the relay node (Node 2) learns the message of Node 1 by decoding the received codewords. As pointed out in [69], this creates a *bottleneck* if the channel between Node 1 and Node 2 is noisier than that between Node 1 and Node 3. In order to circumvent this problem, [69] presented a noise-forwarding strategy where the relay node does not decode the message, but transmits codewords that are

independent of the source's message to confuse the eavesdropper. This is similar to the stochastic encoder that we have employed, where the relay generates additional codewords for every message and transmits one of them at *random*. However, in order to achieve this objective, in our chapter the relay has to decode the message from the main sender which may result in the bottleneck described above.

(b) Since Node 2 needs to decode Node 1's messages, our coding technique is not applicable to relays with half-duplex constraints [77], [78]. Whereas, the noise-forwarding strategy of [69] permits cooperation via half-duplex relays.

(c) Our coding strategy is designed to accommodate for the messages of Node 2 intended to Node 3, along with the need to cooperate with Node 1. Therefore, noise-forwarding strategies may degrade the throughput of the channel between Node 2 and Node 3. In this sense, our strategy better utilizes the spectrum compared to the noise-forwarding strategy.

(d) Referring to Fig. 3.2, we realize the following multiple access channel in our setting: MAC: $(\text{Node } 1, \text{Node } 2) \rightarrow \text{Node } 3$, which models a multiple access eavesdropper channel with Node 2 aiding the transmissions of Node 1 in a causal manner. This resembles the multiple access channel with an eavesdropper setup of [79], but unlike our model the eavesdropper in their case observes a degraded version of the channel seen by the destination.

2. **Scenario II:**

(a) For this scenario, noise-forwarding strategies at Node 2 may not be appropriate, since it has a private message intended to Node 3. This is because, when the channel between Node 2 and Node 3 has a higher gain, noise-insertion at Node 2 might result in the noise being amplified and reducing the received signal-to-noise ratio at Node 3. However, noise-forwarding may be employed by Node 1 to keep its message confidential

from Node 2.

(b) This scenario is similar to the model considered in [70], where the channel model comprises a source broadcasting messages to two receivers. In [70], the message has two parts: A common part and a private part. The receiver of the private part also acts a relay and attempts to keep a part of the received message secret from the other receiver. However, in our chapter, we keep the message of Node 1 fully secret from the relay (Node 2). Further, there is also an external eavesdropper (Node 4) in our setting, whereas in [70], the model has only three communicating nodes without an external eavesdropper. **Scenario II** is also similar to the model considered in [71], where the source enlists the help of a relay who is considered to be untrustworthy. However, in our chapter, Node 1 completely abandons the help from Node 2 who is deemed untrustworthy, by keeping the message fully secret. Note that, Node 2 tries to decode the message from Node 1 so as to keep the relay channel setting fully functional.

(c) For **Scenario II**, MAC: $(\text{Node } 1, \text{Node } 2) \rightarrow \text{Node } 3$ is a multiple access channel with Node 2 eavesdropping on the $(\text{Node } 1, \text{Node } 3)$ link. This scenario is similar to multiple access channels with confidential messages [80], where the transmitting nodes keep their messages confidential from each other. However, in our setup, Node 2 is an untrustworthy relay, which tries to decode the messages of Node 1 on a block-wise/causal basis.

From the above discussion, we can infer that, though the noise-forwarding strategy has several advantages over stochastic encoders, it is not particularly useful for opportunistic/cognitive relays which have private messages to intended destinations. Also, if the link between the relay and the final destination has a higher gain, then noise-forwarding may result in reduced data rates for the main sender-receiver pair.

3.5.4 Comparison with the classical cognitive radio setting

In the existing literature on information theory for cognitive radios, the cognitive terminal which is willing to share its resources (like transmit power, computation, *etc.*) is assumed to have *a priori* knowledge of the messages and codewords of the incumbent primary user. The cognitive node then treats the codewords of the primary as known interference and cancels it out using dirty-chapter coding strategy. It has been shown in [31] - [48] that such a message-sharing mechanism not only aids the primary to achieve better (or higher) data rates, but also increases the overall throughput of the system. The cognitive radio thus 'relays' the message of the primary for better spectrum utilization. Such a model, though clairvoyant, provides reasonable upper bounds for the performance limits for future cooperative networks. On the other hand, the opportunistic-relay considered in this chapter can be thought of as cognitive in a more realistic perspective. In our model, the relay utilizes the cooperative paradigm to opportunistically transmit its own message to the intended destination, thereby improving the spectrum efficiency. Our model is practically more appealing, since the relay nodes do not have *a priori* knowledge of the main sender's message, unlike the classical cognitive radio channel setup.

3.5.5 Connection between the two scenarios

It would be of great interest to construct a unifying model for the two communication channels - **Scenario I** and **Scenario II** - considered in this chapter. A crucial point which acts against such a unifying model is that the resulting model would not be feasible to compare our results with those in the existing literature. This is especially true when one attempts to compare **Scenario II** in the chapter with references [70] and [71] where similar models have been considered, and also when comparing our results with the model presented in [80]. However, the realization of such a unifying model would be a significant next step, and there is much room for further research in this direction.

3.6 Conclusions

We derived achievable rate regions and outer bounds for secure communications over discrete memoryless relay channels. We considered two different communication scenarios over a four node wireless network comprising a source-destination pair, a relay node and a malicious node eavesdropping on the link between the relay and the destination. In both the scenarios, the relay was considered to be opportunistic, in the sense that it had a private message to the destination. To derive inner bounds, we propose the layered coding architecture to simultaneously deal with cooperation, cognition and confidentiality. Auxiliary random variables are used to derive outer bounds to enable single-letter characterization. We pointed out the advantages and drawbacks of layered coding strategy in comparison to those in the existing literature. To the best of the author's knowledge, this is the first instance concerning confidentiality over a relay channel, when the relay has its own message intended to the destination. Gaussian channel models can be considered to plot the rate regions and outer bounds, and is relegated to future chapter.

Chapter 4

State-dependent broadcast channels

4.1 Introduction

The information-theoretic study of broadcast channels (BC) was initiated first by Cover in [76]. In the classical setting, the BC comprises a sender who wishes to transmit k independent messages to k noncooperative receivers. The largest known inner bound on the capacity region when $k = 2$ was derived by Marton [23]. Recently, some ideas were discussed in [82], that is conjectured to lead to a larger inner bound. Capacity outer bounds were presented by Sato in [83] by utilizing the fact that the capacity region of BC depends on the marginal transition probabilities. Nair and El Gamal provided outer bounds for the two-user case [30], based on the results of the more capable BC [84]. Liang *et. al* generalized the outer bounds of [30] by deriving the *New-Jersey* outer bound. Some properties of the *New-Jersey* outer bound were exposed in [85], where it was shown to be equivalent to the computable UVW-bound with bounded cardinalities of the auxiliary random variables.

Several variants of this classical setting have also received considerable attention. One of the most prominent variants is the state-dependent BC with

side-information, where the probability distribution characterizing the channel depends on a state process, and with the channel state made available as side-information at the transmitter, or at the receiver, or at both ends. Capacity inner bounds for the two-user BC with noncausal side-information at the transmitter were derived in [22], where Marton's achievability scheme was extended to state-dependent channels. In [86], inner and outer bounds were derived for the degraded BC with noncausal side-information at the transmitter; the capacity region was derived when side-information was obtained to the encoder in a causal manner. The capacity region for BC with receiver side-information was derived in [87], where a genie provides each receiver with the message it need not decode. To the best of the authors' knowledge, outer bounds for the two-user BC with noncausal side-information at the encoder have not appeared in the literature.

Yet another issue in wireless communications, owing to the broadcast nature of the wireless medium, is related to information security. That is, the broadcast nature of wireless networks facilitates malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, *etc.* An information-theoretic approach to address problems related to security has gained rapid momentum, and is commonly referred to as information-theoretic confidentiality or wireless physical-layer security [6]. An information-theoretic approach to secure broadcasting was inspired by the pioneering work of Csiszár and Körner [88], who derived capacity bounds for the two-user BC, when the sender transmits a private message to receiver 1 and a common message to both receivers, while keeping the private message confidential from receiver 2. Secure broadcasting with a single transmitter and multiple receivers in the presence of an external eavesdropper was considered in [89], where the secrecy capacity region was obtained for several special classes of channels. In [90], capacity bounds were derived for BC where a sender broadcasts two independent messages to two receivers, while keeping each message confidential from the unintended receiver. Capacity results and bounds for Gaussian BC with confidential messages were reported in [91] - [93]. The reader is referred to [94] for a comprehensive review

of physical-layer security in BC. However, to the best of the authors' knowledge, the joint problem of side-information and confidentiality on the BC has not been addressed in the literature.

4.1.1 Main contributions

In this chapter, we aim to provide useful insights into the effect of noncausal side-information at the encoder on (1) the classical two-user BC; (2) the BC with genie-aided receiver side-information; and (3) the BC with confidentiality constraints on the messages. Towards this end, we define three different classes of two-user discrete memoryless BC with noncausal side-information at the encoder. Of particular interest is the Class III channels (described below), which provides a fundamental building block to jointly address side-information and confidentiality in BC.

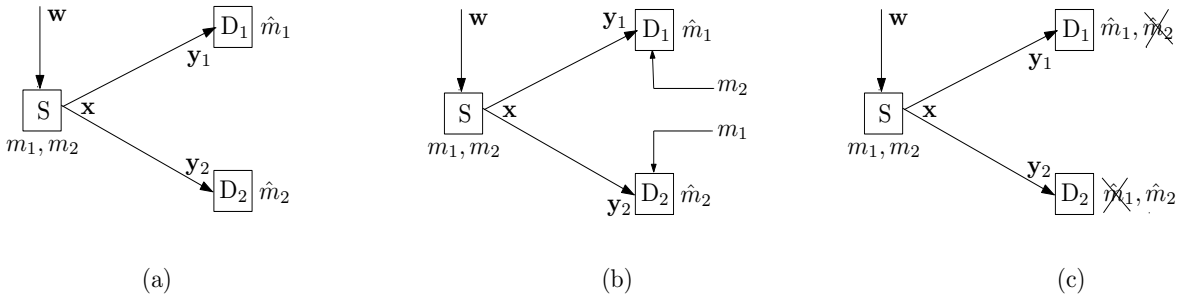

Figure 4.1: State-dependent broadcast channels with side-information at the transmitter: (a) Class I; (b) Class II; and (c) Class III.

1. Class I: A sender broadcasts two independent messages to two non-cooperating receivers (see Fig. 4.1(a)). We derive an inner bound for this class of channels and characterize the rate penalty for dealing with noncausal side-information at the encoder. We are mainly concerned with outer bounds for this class of channels, where we present an explicit single-letter characterization of the sum-rate bound, along with bounds on single-user rates. An example for Class I channels is a base-station transmitting to two mobile receivers, with the base-station having prior knowledge of interference from a transmitter located in its vicinity, *e.g.*, through a backhaul network.

2. Class II: A sender broadcasts two independent messages to two receivers, with each receiver having *a priori* knowledge of the message it need not decode (see Fig. 4.1(b)). An example of this scenario is full-duplex communications between two nodes, aided by a relay. The relay node broadcasts the messages to the terminals, with each terminal knowing its own message. We devise an achievability scheme to derive an inner bound for this class of channels and show that the achievable rate for each user is in fact the maximum rate achievable for a single-user channel with states known *a priori* at the encoder. We also derive an outer bound which is within a fixed gap away from the achievable region, where the gap is independent of the distribution characterizing this class of channels.

3. Class III: A sender broadcasts two independent messages to two receivers, such that each message is kept confidential from the unintended receiver (see Fig. 4.1(c)). To the best of the authors' knowledge, this is the first instance of a study of *simultaneous* impact of side-information and confidentiality constraints on BC. An inner bound for this class of channels is derived employing stochastic encoders to satisfy confidentiality constraints; we characterize the rate penalties for having to deal not only with side-information, but also to satisfy confidentiality constraints. One of the outer bounds is derived by employing a genie, which gives one of the receivers the message it need not decode, while the other receiver computes the equivocation rate treating this message as side-information. We also derive another outer bound, with an explicit characterization of the sum-rate bounds. As an example for this class of channels, we can extend the example considered for Class I channels, with the additional constraint of keeping each message confidential from the unintended receiver.

The remainder of the chapter is organized as follows. In Section 4.2, we introduce the notation used and provide a mathematical model for the discrete memoryless version of the channels considered in this chapter. In Section 4.3,

we summarize the main results of this chapter by describing inner and outer bounds for all the channel models, and provide related discussion. The proofs of the achievability theorems can be found in Section 4.4, while the proofs of the outer bounds are provided in Section 4.5. Finally, we conclude the chapter in Section 4.6. The encoder error analysis is relegated to Appendix C.1.

4.2 System model and notation

The channels belonging to $\mathrm{Class\ I}$, $\mathrm{Class\ II}$ and $\mathrm{Class\ III}$ are denoted C_1, C_2 and C_3, respectively. Calligraphic letters are used to denote finite sets, with a probability function defined on them. N is the number of channel uses, and $\mathrm{n} = 1, \ldots, \mathrm{N}$ denotes the channel index. Uppercase letters denote random variables (RV), while boldface uppercase letters denote a sequence of RVs. The following notation for a sequence of RVs is useful: $\mathbf{Y}_1^{\mathrm{N}} \triangleq (Y_{1,1}, \ldots, Y_{1,N})$; $\mathbf{Y}_1^{\mathrm{n}-1} \triangleq (Y_{1,1}, \ldots, Y_{1,\mathrm{n}-1})$; and $\mathbf{Y}_{1,\mathrm{n}+1}^{\mathrm{N}} \triangleq (Y_{1,\mathrm{n}+1}, \ldots, Y_{1,\mathrm{N}})$. Lowercase letters are used to denote particular realizations of RVs, and boldface lowercase letters denote vectors. The sender is denoted S and the receivers are denoted D_t, where $t = 1, 2$ is the receiver index. Discrete RV $X \in \mathcal{X}$ and $Y_t \in \mathcal{Y}_t$ denote the channel input and outputs, respectively. The encoder of S is supplied with side-information $\mathbf{W} \in \mathcal{W}^{\mathrm{N}}$, in a noncausal manner. The channel is assumed to be memoryless and is characterized by the conditional distribution $p(\mathbf{Y}_1, \mathbf{Y}_2|\mathbf{X}, \mathbf{W}) = \prod_{\mathrm{n}=1}^{\mathrm{N}} p(Y_{1,\mathrm{n}}, Y_{2,\mathrm{n}}|X_{\mathrm{n}}, W_{\mathrm{n}})$. For sake of brevity, in the remainder of this chapter, we use $p(x)$ to denote $p(X = x)$. Unless otherwise stated, $p(\mathbf{x}) = \prod_{\mathrm{n}=1}^{\mathrm{N}} p(x_{\mathrm{n}})$.

To transmit its messages, S generates two RVs $M_t \in \mathcal{M}_t$, where $\mathcal{M}_t = \{1, \ldots, 2^{\mathrm{N}R_t}\}$ denotes a set of message indices. Without loss of generality, $2^{\mathrm{N}R_t}$ is assumed to be an integer, with R_t being the transmission rate intended to D_t. M_t denotes the message S intends to transmit to D_t, and is assumed to be independently generated and uniformly distributed over the finite set $\mathcal{M}_t$. Integer $m_t \in \mathcal{M}_t$ is a particular realization of M_t and denotes the message-index.

Given the conditional distribution characterizing the channel, a

$((2^{\mathrm{N}R_1}, 2^{\mathrm{N}R_2}), \mathrm{N}, P_e^{(\mathrm{N})})$ code for the channels C_1 and C_2 comprises N encoding functions f, such that $\mathbf{X} = \mathbf{f}(m_1, m_2, \mathbf{W})$; for the channel C_3, it comprises a stochastic encoder, which is defined by the matrix of conditional probabilities $\phi(\mathbf{X}|m_1, m_2, \mathbf{W})$, such that $\sum_{\mathbf{X}} \phi(\mathbf{X}|m_1, m_2, \mathbf{W}) = 1$. Here, $\phi(\mathbf{X}|m_1, m_2, \mathbf{W})$ denotes the probability that a pair of message-indices (m_1, m_2) is encoded as $\mathbf{X} \in \mathcal{X}^{\mathrm{N}}$ to be transmitted by S, in the presence of noncausal side-information $\mathbf{W}$. For all channel models, there are two decoders $g_t : \mathcal{Y}_t^{\mathrm{N}} \to \mathcal{M}_t$.

The average probability of decoding error for the code, averaged over all codes, is $P_e^{(\mathrm{N})} = \max\{P_{e,1}^{(\mathrm{N})}, P_{e,2}^{(\mathrm{N})}\}$. A rate pair (R_1, R_2) is said to be achievable for the channel $\mathrm{C_c}; \mathrm{c} = 1, 2, 3$, if there exists a sequence of $((2^{\mathrm{N}R_1}, 2^{\mathrm{N}R_2}), \mathrm{N}, P_e^{(\mathrm{N})})$ codes, such that $\forall \epsilon > 0$ and sufficiently small, $P_e^{(\mathrm{N})} \leq \epsilon$ as $\mathrm{N} \to \infty$. Furthermore, for the channel C_3, the following constraints [95] on the conditional entropy must be satisfied for (R_1, R_2) to be considered achievable:

$$\mathrm{N}R_1 - H(M_1|\mathbf{Y}_2) \leq \mathrm{N}\epsilon, \tag{4.1}$$

$$\mathrm{N}R_2 - H(M_2|\mathbf{Y}_1) \leq \mathrm{N}\epsilon. \tag{4.2}$$

The capacity region is defined as the closure of the set of all achievable rate pairs (R_1, R_2).

4.3 Main results

In this section, we state the achievability and converse theorems for all the channel models considered in this chapter, and provide related discussion. Let $\mathcal{C}_{\mathrm{c}}$ denote the capacity region of the channel $\mathrm{C_c}$; $\mathrm{c} = 1, 2, 3$. We use the following auxiliary RVs defined on finite sets: $U \in \mathcal{U}$, $V_1 \in \mathcal{V}_1$ and $V_2 \in \mathcal{V}_2$.

4.3.1 Class I channels

For the channel C_1, we consider the set $\mathcal{P}_1$ of all joint probability distributions $p_1(.)$ that can be factored as $p(w)p(v_1, v_2|w)p(x|w, v_1, v_2)p(y_1, y_2|x)$. For a given $p_1(.) \in \mathcal{P}_1$,

a lower bound on the capacity region for C_1 is described by the set $\mathcal{R}_{1,\text{in}}(p_1)$, which is defined as the union over all distributions $p_1(.)$ of the convex hull of the set of all rate pairs (R_1, R_2) that simultaneously satisfy (4.3) - (4.5).

$$\begin{aligned} R_1 &\leq I(V_1; Y_1) - I(V_1; W), && (4.3)\\ R_2 &\leq I(V_2; Y_2) - I(V_2; W), && (4.4)\\ R_1 + R_2 &\leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2) - I(V_1, V_2; W), && (4.5) \end{aligned}$$

where V_1 and V_2 are constrained to satisfy the Markov chain $(V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 4.3.1. Let $\mathcal{R}_{1,\text{in}} = \bigcup_{p_1(.)\in\mathcal{P}_1} \mathcal{R}_{1,\text{in}}(p_1)$. Then, $\mathcal{R}_{1,\text{in}} \subseteq \mathcal{C}_1$.

For proof, see Section 4.4.1.

For a given $p_1(.) \in \mathcal{P}_1$, an outer bound for C_1 is described by the set $\mathcal{R}_{1,\text{out}}(p_1)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (4.6) - (4.7).

$$\begin{aligned} R_1 &\leq I(V_1; Y_1) - I(V_1; W), && (4.6)\\ R_2 &\leq I(V_2; Y_2) - I(V_2; W),, && (4.7) \end{aligned}$$

where $(V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 4.3.2. Let $\mathcal{R}_{1,\text{out}} = \bigcup_{p_1(.)\in\mathcal{P}_1} \mathcal{R}_{1,\text{out}}(p_1)$. Then, $\mathcal{C}_1 \subseteq \mathcal{R}_{1,\text{out}}$.

The proof of Theorem 4.3.2 can be found in Section 4.5.1. However, this outer bound does not include a bound on the sum-rates. To explicitly bound the sum-rate, we provide the following alternative outer bound for the channel C_1. We consider the set $\mathcal{P}_1^*$ of all joint probability distributions $p_1^*(.)$ that can be factorized as follows: $p(w)p(u, v_1, v_2|w)p(x|w, u, v_1, v_2)p(y_1, y_2|x)$. For a given $p_1^*(.) \in \mathcal{P}_1^*$, an outer bound for C_1 is described by the set $\mathcal{R}_{1,\text{out}}^*(p_1^*)$, which is defined as the union of all

rate pairs (R_1, R_2) that simultaneously satisfy (4.8) - (4.11).

$$\begin{aligned} R_1 &\leq I(U, V_1; Y_1) - I(V_1; W|U), && (4.8)\\ R_2 &\leq I(U, V_2; Y_2) - I(V_2; W|U), && (4.9)\\ R_1 + R_2 &\leq I(U, V_1; Y_1) - I(V_1; W|U) + I(U, V_2; Y_2|V_1) - I(V_2; W|U, V_1), && (4.10)\\ R_1 + R_2 &\leq I(U, V_2; Y_2) - I(V_2; W|U) + I(U, V_1; Y_1|V_2) - I(V_1; W|U, V_2), && (4.11) \end{aligned}$$

where the following Markov chain is satisfied: $(U, V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 4.3.3. Let $\mathcal{R}^*_{1,\text{out}} = \bigcup_{p_1^*(.)\in\mathcal{P}_1^*} \mathcal{R}^*_{1,\text{out}}(p_1^*)$. Then, $\mathcal{C}_1 \subseteq \mathcal{R}^*_{1,\text{out}}$.

Section 4.5.2 contains the proof of Theorem 4.3.3.

4.3.2 Class II channels

For the channel C_2, we consider the set $\mathcal{P}_2$ of all joint probability distributions $p_2(.)$ of the form $p(w)p(u|w)p(x|w,u)p(y_1,y_2|x)$. For a given $p_2(.) \in \mathcal{P}_2$, a lower bound on the capacity region for C_2 is described by the set $\mathcal{R}_{2,\text{in}}(p_2)$, which is defined as the union over all distributions $p_2(.)$ of the convex-hull of the set of all rate pairs (R_1, R_2) that simultaneously satisfy (4.12) - (4.13).

$$\begin{aligned} R_1 &\leq I(U; Y_1) - I(U; W), && (4.12)\\ R_2 &\leq I(U; Y_2) - I(U; W), && (4.13) \end{aligned}$$

where the Markov chain $U \to (X, W) \to (Y_1, Y_2)$ holds.

Theorem 4.3.4. Let $\mathcal{R}_{2,\text{in}} = \bigcup_{p_2(.)\in\mathcal{P}_2} \mathcal{R}_{2,\text{in}}(p_2)$. Then, $\mathcal{R}_{2,\text{in}} \subseteq \mathcal{C}_2$.

The proof of Theorem 4.3.4 is relegated to Section 4.4.2.

For a given $p_2(.) \in \mathcal{P}_2$, an outer bound for C_2 is described by the set $\mathcal{R}_{2,\text{out}}(p_2)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy

(4.14) - (4.15).

$$R_1 \leq I(U;Y_1) - I(U;W) + H(U), \tag{4.14}$$

$$R_2 \leq I(U;Y_2) - I(U;W) + H(U), \tag{4.15}$$

with $U \to (X,W) \to (Y_1,Y_2)$.

Theorem 4.3.5. Let $\mathcal{R}_{2,\text{out}} = \bigcup_{p_2(.)\in\mathcal{P}_2} \mathcal{R}_{2,\text{out}}(p_2)$. Then, $\mathcal{C}_2 \subseteq \mathcal{R}_{2,\text{out}}$.

The proof of Theorem 4.3.5 can be found in Section 4.5.3.

4.3.3 Class III **channels**

For the channel C_3, we consider the set $\mathcal{P}_3$ of all joint probability distributions $p_3(.)$ that can be written as $p(w)p(u)p(v_1,v_2|w,u)p(x|w,v_1,v_2)p(y_1,y_2|x)$. For a given $p_3(.) \in \mathcal{P}_3$, an inner bound on the capacity region for C_3 is described by the set $\mathcal{R}_{3,\text{in}}(p_3)$, which is defined as the union over all distributions $p_3(.)$ of the convex-hull of the set of all rate pairs (R_1,R_2) that simultaneously satisfy (4.16) - (4.18).

$$R_1 \leq I(V_1;Y_1|U) - \max[I(V_1;Y_2|U,V_2), I(V_1;W|U)], \tag{4.16}$$

$$R_2 \leq I(V_2;Y_2|U) - \max[I(V_2;Y_1|U,V_1), I(V_2;W|U)], \tag{4.17}$$

$$\begin{aligned} R_1+R_2 \leq\ & I(V_1;Y_1|U) + I(V_2;Y_2|U) - I(V_1;Y_2|U,V_2) - I(V_2;Y_1|U,V_1) \\ & - I(V_1;V_2|U) - I(V_1,V_2;W|U), \end{aligned} \tag{4.18}$$

where the following Markov chain is satisfied: $U \to (V_1,V_2) \to (X,W) \to (Y_1,Y_2)$.

Theorem 4.3.6. Let $\mathcal{R}_{3,\text{in}} = \bigcup_{p_3(.)\in\mathcal{P}_3} \mathcal{R}_{3,\text{in}}(p_3)$. Then, $\mathcal{R}_{3,\text{in}} \subseteq \mathcal{C}_3$.

Section 4.4.3 contains the proof of Theorem 4.3.6.

For a given $p_3(.) \in \mathcal{P}_3$, an outer bound for C_3 is described by the set $\mathcal{R}_{3,\text{out}}(p_3)$, which is defined as the union of all rate pairs (R_1,R_2) that simultaneously satisfy

(4.19) - (4.20).

$$R_1 \leq \min[I_1, I_1^*], \tag{4.19}$$

$$R_2 \leq \min[I_2, I_2^*],, \tag{4.20}$$

where $I_1, \ldots, I_2^*$ are given by (4.21) - (4.24), respectively.

$$I_1 \triangleq I(V_1; Y_1|U) - I(V_1; Y_2|U) + H(W|U, V_1), \tag{4.21}$$

$$I_2 \triangleq I(V_2; Y_2|U) - I(V_2; Y_1|U) + H(W|U, V_2), \tag{4.22}$$

$$I_1^* \triangleq I(V_1; Y_1|U, V_2) - I(V_1; Y_2|U, V_2) + H(W|U, V_1, V_2), \tag{4.23}$$

$$I_2^* \triangleq I(V_2; Y_2|U, V_1) - I(V_2; Y_1|U, V_1) + H(W|U, V_1, V_2), \tag{4.24}$$

where $U \to (V_1, V_2) \to (X, W) \to (Y_1, Y_2)$. The expressions (4.23) - (4.24) are obtained by letting a genie give D_1 message M_2, while D_2 computes the equivocation using M_2 as side-information.

Theorem 4.3.7. Let $\mathcal{R}_{3,\text{out}} = \bigcup_{p_3(.) \in \mathcal{P}_3} \mathcal{R}_{3,\text{out}}(p_3)$. Then, $\mathcal{C}_3 \subseteq \mathcal{R}_{3,\text{out}}$.

The proof of Theorem 4.3.7 can be found in Section 4.5.4. We also provide the following outer bound for the channel C_3, which explicitly characterizes the sum-rates. Consider the set $\mathcal{P}_3^*$ of all joint probability distributions $p_3^*(.)$ that can be factorized as follows: $p(w)p(u, v_1, v_2|w)p(x|w, u, v_1, v_2)$
$p(y_1, y_2|x)$. For a given $p_3^*(.) \in \mathcal{P}_3^*$, an outer bound for C_3 is described by the set $\mathcal{R}_{3,\text{out}}^*(p_3^*)$, which is defined as the union of all rate pairs (R_1, R_2) that simultaneously satisfy (4.25) - (4.28).

$$R_1 \leq I(U, V_1; Y_1) - I(V_1; W|U) - I(V_1; Y_2), \tag{4.25}$$

$$R_2 \leq I(U, V_2; Y_2) - I(V_2; W|U) - I(V_2; Y_1), \tag{4.26}$$

$$\begin{aligned} R_1 + R_2 \leq\ & I(U, V_1; Y_1) - I(V_1; W|U) + I(U, V_2; Y_2|V_1) \\ & - I(V_2; W|U, V_1) - I(V_1; Y_2), \end{aligned} \tag{4.27}$$

$$R_1 + R_2 \leq I(U, V_2; Y_2) - I(V_2; W|U) + I(U, V_1; Y_1|V_2)$$

$$-I(V_1; W|U, V_2) - I(V_2; Y_1), \tag{4.28}$$

where $(U, V_1, V_2) \to (X, W) \to (Y_1, Y_2)$.

Theorem 4.3.8. Let $\mathcal{R}^*_{3,\text{out}} = \bigcup_{p^*_3(.)\in\mathcal{P}^*_3} \mathcal{R}^*_{3,\text{out}}(p^*_3)$. Then, $\mathcal{C}_3 \subseteq \mathcal{R}^*_{3,\text{out}}$.

The proof of Theorem 4.3.8 can be found in Section 4.5.5.

4.3.4 Discussion

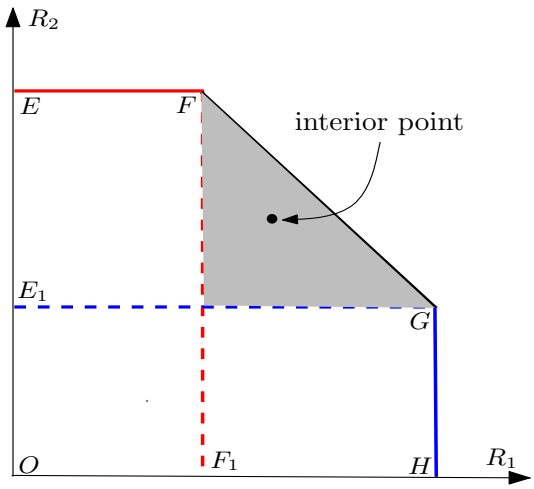

Figure 4.2: Pictorial representation of the rate region for Class I channels.

A pictorial representation of the rate region for the channel C_1 is shown in Fig. 4.2. When $R_2 = 0$, the channel resembles a single-user channel $(\mathrm{S}, \mathrm{D}_1)$ with side-information (the Gel'fand-Pinsker's (GP) channel [8]) and S can transmit at the maximum achievable R_1 given by (4.3), denoted by point the H. At the point H, the maximum achievable R_2 is given by the point $\mathrm{E}_1 \equiv I(V_2; Y_2) - I(V_1; V_2) - I(W; V_2)$; this is obtained by treating the channel $(\mathrm{S}, \mathrm{D}_2)$ as a single-user channel with side-information. Therefore, the rectangle OHGE_1 is achievable. By exchanging R_1 and R_2 and following similar arguments the points E, given by (4.4), and $\mathrm{F}_1 \equiv I(V_1; Y_1) - I(V_1; V_2|U) - I(W; V_1)$ are achievable. Hence, the rectangle OEFF_1 is also achievable. Since the points F and G are shown to be achievable, any point which lies on the line FG can also be achieved by deriving a bound on the binning rates (see (C.2) - (C.4), Appendix C.1). This leads to a sum rate bound given by

(4.5). Finally, owing to convexity of the rate region, any point in the interior of the line FG is also achievable. Therefore, an achievable rate region for C_1 is described by the pentagon OEFGH.

In the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by the convex-hull of the set of all rate pairs (R_1, R_2) that satisfy the following inequalities:

$$R_1 \leq I(V_1; Y_1), \tag{4.29}$$

$$R_2 \leq I(V_2; Y_2), \tag{4.30}$$

$$R_1 + R_2 \leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2). \tag{4.31}$$

For channels of Class II, each bound in (4.12) - (4.13) is the capacity of GP's single-user channel with noncausal side-information. In the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, we get $R_t \leq I(U; Y_t) = I(X; Y_t)$, which represents the capacity region of BC when each receiver is given the message it need not decode [87]. Furthermore, the outer bounds (4.14) - (4.15) is within a fixed gap, $H(U)$, from the achievable region, where $H(U)$ is independent of the distribution characterizing this class of channels.

For Class III channels, the terms $I(V_1; Y_2|U, V_2)$ and $I(V_2; Y_1|U, V_1)$ quantify the rate-penalty for having to deal with confidentiality constraints on the messages, while the terms $I(V_1; W|U)$ and $I(V_2; W|U)$ quantify the rate-penalty for having to deal with side-information. Using a combination of results from GP's channel and wiretap channels with side-information [96], we obtain a pictorial representation of the rate region for the channel C_3 as shown in Fig. 4.3. The arguments used to obtain this schematic are similar to those used for the channel C_1; therefore, we briefly explain the construction of Fig. 4.3. The point A_1 corresponds to the maximum achievable R_1 (when $R_2 = 0$) and is given by (4.16). Exchanging R_1 and R_2 we get the point C_1 given by (4.17). The points $\mathrm{B}_1 \equiv I(V_2; Y_2|U) - I(V_2; Y_1|U, V_1) - \max[I(V_1; V_2|U), I(W; V_2|U)]$ and $\mathrm{D}_1 \equiv I(V_1; Y_1|U) - I(V_1; Y_2|U, V_2) - \max[I(V_1; V_2|U), I(W; V_1|U)]$ are achievable by treating channels

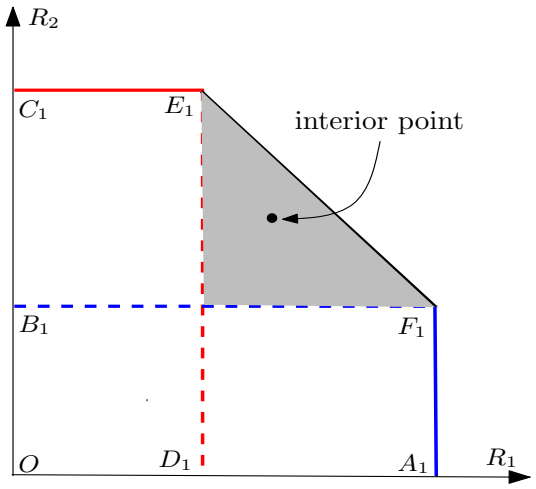

Figure 4.3: Pictorial representation of the rate region for Class III channels.

$(\mathrm{S}, \mathrm{D}_2)$ and $(\mathrm{S}, \mathrm{D}_1)$, respectively, as wiretap channels with side-information. The line $\mathrm{E}_1\mathrm{F}_1$ corresponds to the sum rate bound given by (4.18). Finally, owing to convexity of the rate region, any point in the interior of the line $\mathrm{E}_1\mathrm{F}_1$ is also achievable. Therefore, an achievable rate region for C_3 is described by the pentagon $\mathrm{OA}_1\mathrm{F}_1\mathrm{E}_1\mathrm{C}_1$.

If the confidentiality constraints (4.1) - (4.2) are relaxed, the channel C_3 reduces to the channel C_1, whose rate region is described by (4.3) - (4.5). Further, in the absence of side-information, *i.e.*, $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by (4.29) - (4.31). Lastly, if the encoder satisfies confidentiality constraints in the absence of side-information, the channel C_3 reduces to BC with two independent and confidential messages whose rate region was first characterized by Liu *et. al* [90]. It is described by the convex-hull of the set of all rate pairs (R_1, R_2) that satisfy the following inequalities:

$$R_1 \leq I(V_1; Y_1|U) - I(V_1; Y_2|U) - I(V_1; V_2|U), \tag{4.32}$$

$$R_2 \leq I(V_2; Y_2|U) - I(V_2; Y_1|U) - I(V_1; V_2|U). \tag{4.33}$$

4.3.5 Relation to past work

For Class I channels, an inner bound was presented in [22] by extending Marton's achievability scheme for the classical two-user BC to include noncausal side-information at the encoder. In this chapter, we employ Marton's technique and use results from the second moment method [24] to derive the inner bound which matches with the results presented in [22]. However, our method is simpler and generalizes well for obtaining inner bounds with other channel models, *e.g.*, for channels of Class III considered in this chapter. For the outer bound (specifically, for the sum-rate), we generalize the technique presented in [30], to handle side-information at the encoder. When the side-information constraint is relaxed, our result reduces to the one presented for the classical two-user BC [30].

Class II channels were also addressed in [25], where an inner bound was derived by employing Marton's achievability scheme. An outer bound was also suggested in [25], but without a formal proof. In this chapter, we derive an inner bound by generalizing the method suggested in [87] by incorporating noncausal side-information at the encoder. Our inner bound coincides with the one presented in [25], but once again the proof technique is much simpler. Furthermore, for the outer bounds, we explicitly address the problem of dealing with the two-dimensional rate region with a single auxiliary random variable.

For Class III channels, we show that when the confidentiality constraints are relaxed, our achievable rate region reduces to region presented for the Class I channels, and hence to the one presented in [22]. On the other hand, in the absence of side-information, our achievable region includes an explicit bound on the sum-rate for the two-user BC with confidentiality constraints (a model considered in [90]). This further strengthens the generalization of our proof technique.

4.4 Proofs of achievability theorems

In this section, we prove Theorem 4.3.1, Theorem 4.3.4 and Theorem 4.3.6. For any $\epsilon > 0$, we denote by $A_\epsilon^{(\mathrm{N})}(P_X)$ an ϵ-typical set comprising sequences picked

from the distribution $p(\mathbf{x})$. For all the channel models, the encoder is given an $\epsilon-$typical sequence $\mathbf{W} \in A_\epsilon^{(\mathrm{N})}(P_W)$ in a noncausal manner.

4.4.1 Proof of Theorem 4.3.1

For the channel C_1, generate $2^{\mathrm{N}[R_t+R_t^{'}]}$ independent typical sequences $\mathbf{V}_t(i_t, j_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t}); t = 1, 2$. Here, $i_t \in \{1, \ldots, 2^{\mathrm{N}R_t}\}$; $j_t \in \{1, \ldots, 2^{\mathrm{N}R_t^{'}}\}$. Uniformly distribute $2^{\mathrm{N}[R_t+R_t^{'}]}$ sequences into $2^{\mathrm{N}R_t}$ bins, so that each bin, indexed by i_t, comprises $2^{\mathrm{N}R_t^{'}}$ sequences. To send the message pair $(m_1 = i_1, m_2 = i_2)$, the encoder at S looks for a pair (j_1, j_2) that satisfies the following joint typicality condition: $E_\mathrm{S} \triangleq \{(\mathbf{W}, \mathbf{V}_1(i_1, j_1), \mathbf{V}_2(i_2, j_2)) \in A_\epsilon^{(\mathrm{N})}(P_{W,V_1,V_2})\}$. An error is declared at the encoder of S, if it is not possible to find the $(j_1, j_2)-$pair to satisfy the condition E_S. The encoder error analysis can be found in Appendix C.1. The channel input sequence is $\mathbf{X} \in A_\epsilon^{(\mathrm{N})}(P_{X|W,V_1,V_2})$.

At the destination D_t, the decoder looks for $(\hat{i}_t, \hat{j}_t)$ that satisfies the following joint typicality condition: $E_{\mathrm{D}_t} \triangleq \{(\mathbf{V}_t(\hat{i}_t, \hat{j}_t), \mathbf{Y}_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t,Y_t})\}$. An error is declared at decoder of D_t, if it not possible to find a unique integer $\hat{i}_t$ to satisfy the condition E_{D_t}. From the union of events bound, the probability of decoder error at D_t can be upper bounded as follows: $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \Pr(E_{\mathrm{D}_t}^c|E_\mathrm{S}) + \sum_{\hat{i}_t \neq i_t} \sum_{j_t} \Pr(E_{\mathrm{D}_t}|E_\mathrm{S})$. From the asymptotic equipartition property (AEP) [52], $\forall \epsilon > 0$ and sufficiently small; and for large N, $\Pr(E_{\mathrm{D}_t}^c|E_\mathrm{S}) \leq \epsilon$. Further, for $\hat{i}_t \neq i_t$, $\Pr(E_{\mathrm{D}_t}|E_\mathrm{S}) \leq 2^{-\mathrm{N}[I(V_t;Y_t)-\epsilon]}$. Therefore, we have $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon + 2^{\mathrm{N}[R_t+R_t^{'}]} 2^{-\mathrm{N}[I(V_t;Y_t)-\epsilon]}$, leading us to conclude that, for any $\epsilon_0 > 0$ and sufficiently small; and for large N, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon_0$ if

$$R_t + R_t^{'} < I(V_t; Y_t). \tag{4.34}$$

For the channel C_1, the rate inequalities (4.34) and the bounds on the binning rates $(C.2)$ - $(C.4)$ (see Appendix C.1) are combined to obtain an achievable rate region given by (4.3) - (4.5). This completes the proof of Theorem 4.3.1.

4.4.2 Proof of Theorem 4.3.4

For the channel C_2, we consider the following two cases.

1. When $R_1 \leq R_2$: Generate $2^{\mathrm{N}(R_2+R^*)}$ typical sequences $\mathbf{U}(i,j) \in A_\epsilon^{(\mathrm{N})}(P_U); i \in \{1,\ldots,2^{\mathrm{N}R_2}\}$;
$j \in \{1,\ldots,2^{\mathrm{N}R^*}\}$. Uniformly distribute these sequences into $2^{\mathrm{N}R_2}$ bins, so that each bin comprises $2^{\mathrm{N}R^*}$ sequences. The bins are indexed by i. Define now the following mappings:

$$m_t \in \{1,\ldots,2^{\mathrm{N}R_t}\} \;\; \mapsto \;\; \mathrm{Int}(m_t) \in \{0,\ldots,2^{\mathrm{N}R_2}-1\}; t=1,2,$$

where $\mathrm{Int}(\alpha)$ denotes an integer to represent α. To transmit the message pair (m_1, m_2), compute $\left(\mathrm{Int}(m_1)+\mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right)$. By construction, the bin index
$i \triangleq \mathrm{Int}^{-1}\left(\mathrm{Int}(m_1)+\mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right)$. Given the sequence $\mathbf{W}$, the encoder looks for an integer j to satisfy the following joint typicality condition:

$$(\mathbf{U}(i,j),\mathbf{W}) \in A_\epsilon^{(\mathrm{N})}(P_{W,U}).$$

Finally, $\mathbf{X} \triangleq \mathbf{f}(\mathbf{U}(i,j),\mathbf{W})$ is transmitted in N channel uses.

At receiver D_1, given m_2, the decoder looks for the pair $(\hat{i} \triangleq \hat{m}_1, \hat{j})$ such that the following joint typicality condition is satisfied:

$$E_{\mathrm{D}_1} \triangleq \{(\mathbf{U}(\mathrm{Int}^{-1}\left(\mathrm{Int}(\hat{m}_1)+\mathrm{Int}(m_2) \mod 2^{\mathrm{N}R_2}\right),j),\mathbf{Y}_1) \in A_\epsilon^{(\mathrm{N})}(P_{U,Y_1})\}.$$

From AEP, it can be shown that $\Pr(E_{\mathrm{D}_1}^c) \leq \delta_1; \forall \delta_1 > 0$ and sufficiently small; and for large N, if $R_1+R^* \leq I(U;Y_1)$. Similarly, it can be shown that $\Pr(E_{\mathrm{D}_2}^c) \leq \delta_2; \forall \delta_2 > 0$ and sufficiently small; and for large N, if $R_2 + R^* \leq I(U;Y_2)$. Additionally, by following a procedure similar to the one presented in Appendix C.1, we bound the binning rate as follows: $R^* > I(U;W)$. Therefore,

m_1 (resp. m_2) can be reliably decoded at D_1 (resp. D_2) if

$$R_1 \leq I(U;Y_1) - I(U;W), \tag{4.35}$$

$$R_2 \leq I(U;Y_2) - I(U;W). \tag{4.36}$$

2. When $R_2 \leq R_1$: By symmetry, we get the same rate bounds as in (4.35) and (4.36).

This completes the proof of Theorem 4.3.4.

4.4.3 Proof of Theorem 4.3.6

For the channel C_3, generate a typical sequence $\mathbf{U} \in A_\epsilon^{(\mathrm{N})}(P_U)$, known to all nodes in the network. Generate $2^{\mathrm{N}[R_t+R_t^{'}+R_t^*]}$ independent typical sequences $\mathbf{V}_t(i_t,j_t,k_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t})$; $i_t \in \{1,\ldots,2^{\mathrm{N}R_t}\}$; $j_t \in \{1,\ldots,2^{\mathrm{N}R_t^{'}}\}$; $k_t \in \{1,\ldots,2^{\mathrm{N}R_t^*}\}$. Uniformly distribute $2^{\mathrm{N}[R_t+R_t^{'}+R_t^*]}$ sequences into $2^{\mathrm{N}R_t}$ bins, so that each bin, indexed by i_t, comprises $2^{\mathrm{N}[R_t^{'}+R_t^*]}$ sequences. Uniformly distribute $2^{\mathrm{N}[R_t^{'}+R_t^*]}$ sequences into $2^{\mathrm{N}R_t^{'}}$ sub-bins indexed by (i_t,j_t), so that each sub-bin comprises $2^{\mathrm{N}R_t^*}$ sequences.

To send the message pair (m_1,m_2), S employs a stochastic encoder. In the bin indexed by i_t, *randomly* pick a sub-bin indexed (i_t,j_t). The encoder then looks for a pair (k_1,k_2) that satisfies the following joint typicality condition: $(\mathbf{W},\mathbf{V}_1(i_1,j_1,k_1),\mathbf{V}_2(i_2,j_2,k_2)) \in A_\epsilon^{(\mathrm{N})}(P_{W,V_1,V_2|U})$. The channel input sequence $\mathbf{X} \in A_\epsilon^{(\mathrm{N})}(P_{X|W,V_1,V_2})$ is transmitted in N uses of the channel.

At the destination D_t, given $\mathbf{U}$, the decoder picks k_t that satisfies the following joint typicality condition: $E_{\mathrm{D}_t} \triangleq \{(\mathbf{V}_t(i_t,j_t,k_t),\mathbf{Y}_t) \in A_\epsilon^{(\mathrm{N})}(P_{V_t,Y_t|U})\}$. An error is declared at the decoder of D_t if it not possible to find an integer $\hat{i}_t$ satisfying E_{D_t}. From union of events bound, the probability of decoder error at D_t can be upper bounded as follows: $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \Pr(E_{\mathrm{D}_t}^c|E_{\mathrm{S}}) + \sum_{\hat{i}_t \neq i_t}\sum_{j_t,k_t}\Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}})$. From AEP [52], $\forall \epsilon > 0$ and sufficiently small; and for large N, $\Pr(E_{\mathrm{D}_t}^c|E_{\mathrm{S}}) \leq \epsilon$ and for $\hat{i}_t \neq i_t$, we have $\Pr(E_{\mathrm{D}_t}|E_{\mathrm{S}}) \leq 2^{-\mathrm{N}[I(V_t;Y_t|U)-\epsilon]}$. Therefore, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon + 2^{\mathrm{N}[R_t+R_t^{'}+R_t^*]}2^{-\mathrm{N}[I(V_t;Y_t|U)-\epsilon]}$.

For any $\epsilon_0 > 0$ and sufficiently small; and for large N, $P_{e,\mathrm{D}_t}^{(\mathrm{N})} \leq \epsilon_0$ if

$$R_t + R_t^{'} + R_t^{*} < I(V_t; Y_t|U). \tag{4.37}$$

The equivocation at the decoder of D_2 is calculated by first considering the following lower bound: $H(M_1|\mathbf{Y}_2^{\mathrm{N}}) \geq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}})$. Following the procedure in [90, Section V-B] and using the fact that $M_1 \rightarrow (\mathbf{U}^{\mathrm{N}}, \mathbf{V}_1^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}) \rightarrow \mathbf{Y}_2^{\mathrm{N}}$ forms a Markov chain, we get

$$\begin{aligned} H(M_1|\mathbf{Y}_2^{\mathrm{N}}) \geq H(\mathbf{V}_1^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) - I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{V}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) - H(\mathbf{V}_1^{\mathrm{N}}|M_1, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}, \mathbf{Y}_2^{\mathrm{N}}) - \\ I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{Y}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}). \end{aligned} \tag{4.38}$$

Now, $\forall \epsilon_l > 0; l = 4, \ldots, 10$ and sufficiently small; and for large N, the terms in (4.38) become

$$\begin{aligned} H(\mathbf{V}_1^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) \overset{(a)}{=} \mathrm{N}[R_1 + R_1^{'} + R_1^{*}]; I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{V}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}) \overset{(b)}{=} \mathrm{N}I(V_1; V_2|U) + \mathrm{N}\epsilon_4; \\ H(\mathbf{V}_1^{\mathrm{N}}|M_1, \mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}, \mathbf{Y}_2^{\mathrm{N}}) \overset{(c)}{\leq} \mathrm{N}\epsilon_5; I(\mathbf{V}_1^{\mathrm{N}}; \mathbf{Y}_2^{\mathrm{N}}|\mathbf{U}^{\mathrm{N}}, \mathbf{V}_2^{\mathrm{N}}) \overset{(d)}{=} \mathrm{N}I(V_1; Y_2|U, V_2) + \mathrm{N}\epsilon_6. \end{aligned} \tag{4.39}$$

In (4.39), (a) follows from the codebook construction; (b) and (d) follow from standard techniques (for *e.g.*, see [90, Lemma 3]); and (c) is proved in [90, Lemma 2]. A similar procedure is followed to calculate the equivocation at the decoder at D_1. Finally, the security constraints (4.1) and (4.2) are satisfied by letting

$$R_1^{'} = I(V_1; Y_2|U, V_2) - \epsilon_7; R_1^{*} = I(V_1; V_2|U) - \epsilon_8; \tag{4.40}$$

$$R_2^{'} = I(V_2; Y_1|W, U, V_1) - \epsilon_9; R_2^{*} = I(V_1; V_2|W, U) - \epsilon_{10}. \tag{4.41}$$

For the channel C_3, rate inequalities (4.37), constraints (4.40) - (4.41) and bounds on the binning rates (C.5) - (C.7) (see Appendix C.1) are combined to obtain the rate region described by (4.16) - (4.18). This completes the proof of Theorem 4.3.6.

4.5 Proofs of converse theorems

In this section, we prove Theorem 4.3.2, Theorem 4.3.3, Theorem 4.3.5, Theorem 4.3.7 and Theorem 4.3.8.

4.5.1 Proof of Theorem 4.3.2

For the channel C_1, $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) \\
&\overset{(a)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} [H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) - H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, M_1)] + \mathrm{N}\epsilon \\
&\overset{(c)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}} [H(Y_{1,\mathrm{n}}) - H(Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}, M_1)] + \mathrm{N}\epsilon = \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}; Y_{1,\mathrm{n}}) + \mathrm{N}\epsilon \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}|M_1, \mathbf{Y}_1^{\mathrm{n}-1})] + \mathrm{N}\epsilon \\
&\overset{(d)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(\mathbf{Y}_1^{\mathrm{n}-1}; W_{\mathrm{n}}|M_1, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})] + \mathrm{N}\epsilon \\
&\overset{(e)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(M_1, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}; W_{\mathrm{n}})] + \mathrm{N}\epsilon,
\end{aligned}$$

where (a) follows from Fano's inequality [52], (b) follows from the chain rule, (c) follows from the fact that conditioning reduces entropy, (d) follows from Csiszár's sum identity [27] and (e) is due to the fact that $(M_1, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})$ is independent of W_{n}. We let $V_{1,\mathrm{n}} = (M_1, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1})$ and note that this choice satisfies the Markov chain requirement $V_1 \to (X, W) \to (Y_1, Y_2)$, specified in Section 4.3 for the channel C_1. Thus, we get

$$\mathrm{N}R_1 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{4.42}$$

Proceeding in a similar manner and letting $V_{2,\mathrm{n}} = (M_2, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_2^{\mathrm{n}-1})$, we get

$$\mathrm{N}R_2 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{4.43}$$

4.5.2 Proof of Theorem 4.3.3

For the channel C_1, $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as

$$\begin{aligned} \mathrm{N}R_1 &= H(M_1) = I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) \\ &\overset{(a)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon, \end{aligned}$$

where (a) follows from Fano's inequality. Proceeding in a manner similar to the proof of Theorem 4.3.2 (see Section 4.5.1), and letting $U_{\mathrm{n}} = (\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}, \mathbf{Y}_1^{\mathrm{n}-1}, \mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})$ and $V_{1,\mathrm{n}} = M_1$.

$$\mathrm{N}R_1 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{4.44}$$

Similarly, letting $V_{2,\mathrm{n}} = M_2$, R_2 can be upper bounded as follows:

$$\mathrm{N}R_2 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) + \mathrm{N}\epsilon. \tag{4.45}$$

We next upper bound $R_1 + R_2$ as follows. $\forall \epsilon > 0$ and sufficiently small; and for large N, we have

$$\begin{aligned} \mathrm{N}(R_1 + R_2) &= H(M_1, M_2) = H(M_1) + H(M_2|M_1) \\ &= I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) + I(M_2; \mathbf{Y}_2^{\mathrm{N}}|M_1) + H(M_2|\mathbf{Y}_2^{\mathrm{N}}, M_1) \\ &\overset{(a)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1; Y_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2; Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}}, M_1) + 2\mathrm{N}\epsilon, \end{aligned}$$

where (a) follows from Fano's inequality. Consider

$$\begin{aligned}
\sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1;\mathbf{Y}_{1,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1}) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1,\mathbf{Y}_1^{\mathrm{n}-1};\mathbf{Y}_{1,\mathrm{n}}) \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};\mathbf{Y}_{1,\mathrm{n}}) - \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};\mathbf{Y}_{1,\mathrm{n}}|M_1,\mathbf{Y}_1^{\mathrm{n}-1}) \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{1,\mathrm{n}}) \\
&- I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{1,\mathrm{n}}|M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})] \\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};Y_{1,\mathrm{n}}|M_1,\mathbf{Y}_1^{\mathrm{n}-1}) \\
&\overset{(b)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{1,\mathrm{n}}) \\
&- I(M_1;W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}})] \\
&- \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}};Y_{1,\mathrm{n}}|M_1,\mathbf{Y}_1^{\mathrm{n}-1}) \qquad (4.46)
\end{aligned}$$

Next consider

$$\begin{aligned}
\sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2;Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2,\mathbf{Y}_1^{\mathrm{n}-1};Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1};Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2;\mathbf{Y}_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1};Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2,\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) \\
&\quad - \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{2,\mathrm{n}}|\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1,M_2) \\
&= \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1};Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{2,n}|M_1) \\
&\quad - \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2;W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) \\
&\overset{(c)}{=} \sum_{\mathrm{n}=1}^{\mathrm{N}} I(\mathbf{Y}_1^{\mathrm{n}-1};Y_{2,\mathrm{n}}|\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) + \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2,\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}};Y_{2,\mathrm{n}}|M_1) \\
&\quad - \sum_{\mathrm{n}=1}^{\mathrm{N}} I(M_2;W_{\mathrm{n}}|\mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}},\mathbf{Y}_1^{\mathrm{n}-1},\mathbf{Y}_{2,\mathrm{n}+1}^{\mathrm{N}},M_1) \qquad (4.47)
\end{aligned}$$

where (b) and (c) follow from Csiszár's sum identity. With $U_{\rm n} = (\mathbf{W}_{\rm n+1}^{\rm N}, \mathbf{Y}_1^{\rm n-1}, \mathbf{Y}_{2,\rm n+1}^{\rm N})$; $V_{1,\rm n} = M_1$; and $V_{2,\rm n} = M_2$, from (4.46) and (4.47), we get

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{n=1}^{\mathrm{N}} [I(U_{\rm n}, V_{1,\rm n}; Y_{1,\rm n}) - I(V_{1,\rm n}; W_{\rm n}|U_{\rm n})] \\ &+ \sum_{\rm n=1}^{\mathrm{N}} [I(U_{\rm n}, V_{2,\rm n}; Y_{2,\rm n}|V_{1,\rm n}) - I(V_{2,\rm n}; W_{\rm n}|V_{1,\rm n}, U_{\rm n})] + 2\mathrm{N}\epsilon. \end{aligned} \tag{4.48}$$

Similarly, it can be shown that

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{\rm n=1}^{\mathrm{N}} [I(U_{\rm n}, V_{2,\rm n}; Y_{2,\rm n}) - I(V_{2,\rm n}; W_{\rm n}|U_{\rm n})] \\ &+ \sum_{\rm n=1}^{\mathrm{N}} [I(U_{\rm n}, V_{1,\rm n}; Y_{1,\rm n}|V_{2,\rm n}) - I(V_{1,\rm n}; W_{\rm n}|V_{2,\rm n}, U_{\rm n})] + 2\mathrm{N}\epsilon. \end{aligned} \tag{4.49}$$

4.5.3 Proof of Theorem 4.3.5

For the channel C_2, $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1; \mathbf{Y}_1^{\rm N}) + H(M_1|\mathbf{Y}_1^{\rm N}) \\
&\overset{(a)}{\leq} I(M_1; \mathbf{Y}_1^{\rm N}) + \mathrm{N}\epsilon \overset{(b)}{\leq} I(M_1; \mathbf{Y}_1^{\rm N}, M_2) + \mathrm{N}\epsilon = I(M_1; \mathbf{Y}_1^{\rm N}|M_2) + \mathrm{N}\epsilon \\
&\overset{(c)}{=} \sum_{\rm n=1}^{\mathrm{N}} [H(Y_{1,\rm n}|\mathbf{Y}_1^{\rm n-1}, M_2) - H(Y_{1,\rm n}|\mathbf{Y}_1^{\rm n-1}, M_1, M_2)] + \mathrm{N}\epsilon \\
&\overset{(d)}{\leq} \sum_{\rm n=1}^{\mathrm{N}} [H(Y_{1,\rm n}) - H(Y_{1,\rm n}|\mathbf{Y}_1^{\rm n-1}, M_1, M_2)] + \mathrm{N}\epsilon \\
&= \sum_{\rm n=1}^{\mathrm{N}} I(M_1, M_2, \mathbf{Y}_1^{\rm n-1}; Y_{1,\rm n}) + \mathrm{N}\epsilon \\
&= \sum_{\rm n=1}^{\mathrm{N}} [I(M_1, M_2, \mathbf{Y}_1^{\rm n-1}, \mathbf{W}_{\rm n+1}^{\rm N}; Y_{1,\rm n}) - I(\mathbf{W}_{\rm n+1}^{\rm N}; Y_{1,\rm n}|M_1, M_2, \mathbf{Y}_1^{\rm n-1})] + \mathrm{N}\epsilon \\
&\overset{(e)}{=} \sum_{\rm n=1}^{\mathrm{N}} [I(M_1, M_2, \mathbf{Y}_1^{\rm n-1}, \mathbf{W}_{\rm n+1}^{\rm N}; Y_{1,\rm n}) - I(M_1; W_{\rm n}|M_2, \mathbf{Y}_1^{\rm n-1}, \mathbf{W}_{\rm n+1}^{\rm N})] + \mathrm{N}\epsilon \\
&\overset{(f)}{=} \sum_{\rm n=1}^{\mathrm{N}} [I(M_1, M_2, \mathbf{W}_{\rm n+1}^{\rm N}; Y_{1,\rm n}) - I(M_1, M_2, \mathbf{W}_{\rm n+1}^{\rm N}; W_{\rm n}|\mathbf{Y}_1^{\rm n-1})] + \mathrm{N}\epsilon
\end{aligned}$$

$$\overset{(g)}{\leq} \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(M_1, M_2, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; Y_{1,\mathrm{n}}) - I(M_1, M_2, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}}; W_{\mathrm{n}}) + H(M_1, M_2, \mathbf{W}_{\mathrm{n}+1}^{\mathrm{N}})] + \mathrm{N}\epsilon. \tag{4.50}$$

where (a) follows from Fano's inequality; (b) follows from the data-processing inequality; (c) follows from chain rule; (d) follows from the fact that conditioning reduces entropy; (e) follows from Csiszár's sum identity; (f) is due to the memoryless nature of the channel; and (g) is obtained after simple calculations. We let $U_n \triangleq (M_1, M_2, \mathbf{W}_{n+1}^{\mathrm{N}})$ and note that this choice satisfies the Markov chain requirement $U \rightarrow (X, W) \rightarrow (Y_1, Y_2)$ specified in Section 4.3 for the channel C_2 to get

$$\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(U_n; Y_{1,n}) - I(U_n; W_n) + H(U_n)] + \mathrm{N}\epsilon. \tag{4.51}$$

By symmetry, we get the following bound on R_2:

$$\mathrm{N}R_2 \leq \sum_{n=1}^{\mathrm{N}} [I(U_n; Y_{2,n}) - I(U_n; W_n) + H(U_n)] + \mathrm{N}\epsilon. \tag{4.52}$$

We note that the factor $H(U_n)$ is independent of the distribution characterizing the channel C_2.

4.5.4 Proof of Theorem 4.3.7

For the channel C_3, $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned}
\mathrm{N}R_1 &= H(M_1) = I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) \\
&\overset{(a)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) - I(M_1; \mathbf{Y}_2^{\mathrm{N}}) + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}) - I(M_1; Y_{2,n}|\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\overset{(c)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_2^{n-1}; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}) - I(M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}; Y_{2,n}|\mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon
\end{aligned}$$

$$
\begin{aligned}
&\overset{(d)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1, W_n; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\overset{(e)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) + I(W_n; Y_{1,n}|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) \\
&\quad -I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) + H(W_n|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) \\
&\quad -H(W_n|M_1, Y_{1,n}, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) + H(W_n|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}) \\
&\quad -I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})] + 2\mathrm{N}\epsilon,
\end{aligned}
$$

where (a) is from Fano's inequality, (b) is from confidentiality constraints, (c) and (d) follow from Csiszár's sum identity and (e) is the chain rule for mutual information. Letting $U_n \triangleq (\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})$; and $V_{1,1} = \cdots = V_{1,\mathrm{N}} \triangleq M_1$, where U and V_1 satisfy the Markov chain $U \to V_1 \to X$ specified in Section 4.3 for the channel C_3, we get

$$
\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{1,n}; Y_{1,n}|U_n) + H(W_n|U_n, V_{1,n}) - I(V_{1,n}; Y_{2,n}|U_n)] + 2\mathrm{N}\epsilon. \quad (4.53)
$$

Proceeding in a similar fashion and letting $V_{2,1} = \cdots = V_{2,\mathrm{N}} \triangleq M_2$,

$$
\mathrm{N}R_2 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{2,n}; Y_{2,n}|U_n) + H(W_n|U_n, V_{2,n}) - I(V_{2,n}; Y_{1,n}|U_n)] + 2\mathrm{N}\epsilon. \quad (4.54)
$$

For the channel C_3, we also derive a genie-aided outer bound by letting a hypothetical genie give D_1 message M_2, while D_2 computes the equivocation using M_2 as side-information. $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be upper bounded as follows:

$$
\mathrm{N}R_1 = H(M_1) \leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon \leq H(M_1, M_2|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon
$$

$$
\begin{aligned}
&= H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) + H(M_2|\mathbf{Y}_2^{\mathrm{N}}) + \mathrm{N}\epsilon \leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) + \mathrm{N}\epsilon \\
&\leq H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) - H(M_1|\mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(a)}{\leq} H(M_1|\mathbf{Y}_2^{\mathrm{N}}, M_2) - H(M_1|\mathbf{Y}_1^{\mathrm{N}}, M_2) + \mathrm{N}\epsilon \\
&\leq I(M_1; \mathbf{Y}_1^{\mathrm{N}}|M_2) - I(M_1; \mathbf{Y}_2^{\mathrm{N}}|M_2) + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\overset{(b)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1, \mathbf{Y}_2^{n-1}; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, M_2) - I(M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}; Y_{2,n}|\mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\overset{(c)}{=} \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1, W_n; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + I(W_n; Y_{1,n}|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&= \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + H(W_n|M_1, Y_{n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - H(W_n|M_1, Y_{1,n}, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon \\
&\leq \sum_{n=1}^{\mathrm{N}} [I(M_1; Y_{1,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) + H(W_n|M_1, \mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2) \\
&\quad - I(M_1; Y_{2,n}|\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1}, M_2)] + 2\mathrm{N}\epsilon,
\end{aligned}
$$

where (a) follows since the genie gives D_1 message M_2, (b) and (c) follow from Csiszár's sum identity. Letting $U_n \triangleq (\mathbf{Y}_{1,n+1}^{\mathrm{N}}, \mathbf{Y}_2^{n-1})$, $V_{1,1} = \cdots = V_{1,\mathrm{N}} \triangleq M_1$ and $V_{2,1} = \cdots = V_{2,\mathrm{N}} \triangleq M_2$, where U, V_1 and V_2 satisfy the Markov chains $U \to V_1 \to X$ and $U \to V_2 \to X$ specified in Section 4.3 for the channel C_3, R_1 can be bounded as

$$
\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{1,n}; Y_{1,n}|U_n, V_{2,n}) + H(W_n|U_n, V_{1,n}, V_{2,n}) - I(V_{1,n}; Y_{2,n}|U_n, V_{2,n})] + 2\mathrm{N}\epsilon \tag{4.55}
$$

Similarly,

$$
\mathrm{N}R_1 \leq \sum_{n=1}^{\mathrm{N}} [I(V_{2,n}; Y_{2,n}|U_n, V_{1,n}) + H(W_n|U_n, V_{1,n}, V_{2,n}) - I(V_{2,n}; Y_{1,n}|U_n, V_{1,n})] + 2\mathrm{N}\epsilon \tag{4.56}
$$

For the channel C_3, the outer bound on $R_1 + R_2$ can be made tighter by the following procedure. From (4.19) - (4.20), we see that

$$R_1 + R_2 \leq I_1 + I_2, \tag{4.57}$$

$$R_1 + R_2 \leq I_1^* + I_2^*. \tag{4.58}$$

Therefore,

$$R_1 + R_2 \leq \min[I_1 + I_2^*, I_2 + I_1^*]. \tag{4.59}$$

We show now that the bound (4.59) is a tighter bound than (4.57) and (4.58). It is easy to see that

$$I_1 + I_2 = I_1^* + I_2^* + I(W; V_1|U, V_2) + I(W; V_2|U, V_1).$$

Consider $2(I_1 + I_2) = 2[I_1^* + I_2^* + I(W; V_1|U, V_2) + I(W; V_2|U, V_1)]$, which implies the following:

$$\begin{aligned} \min[I_1 + I_2^*, I_2 + I_1^*] &\leq I_1 + I_2, \\ \min[I_1 + I_2^*, I_2 + I_1^*] &\leq I_1^* + I_2^*. \end{aligned}$$

Therefore, the sum rate bound given by (4.59) is tighter than (4.57) and (4.58).

4.5.5 Proof of Theorem 4.3.8

For the channel C_3, $\forall \epsilon > 0$ and sufficiently small; and for large N, R_1 can be bounded as follows:

$$\begin{aligned} \mathrm{N}R_1 &= H(M_1) = I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + H(M_1|\mathbf{Y}_1^{\mathrm{N}}) \\ &\overset{(a)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) + \mathrm{N}\epsilon \overset{(b)}{\leq} I(M_1; \mathbf{Y}_1^{\mathrm{N}}) - I(M_1; \mathbf{Y}_2^{\mathrm{N}}) + 2\mathrm{N}\epsilon, \end{aligned}$$

where (a) follows from Fano's inequality; and (b) follows from confidentiality constraints. Following the procedure used to prove Theorem 4.3.3 (see Section 4.5.2) and letting $U_{\mathrm{n}} = (\mathbf{W}_{\mathrm{n+1}}^{\mathrm{N}}, \mathbf{Y}_{1}^{\mathrm{n-1}}, \mathbf{Y}_{\mathrm{2,n+1}}^{\mathrm{N}})$ and $V_{1,\mathrm{n}} = M_1$,

$$\mathrm{N}R_1 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) - I(V_{1,\mathrm{n}}; Y_{2,\mathrm{n}}) + 2\mathrm{N}\epsilon. \tag{4.60}$$

Similarly, letting $V_{2,\mathrm{n}} = M_2$, we get

$$\mathrm{N}R_2 \leq \sum_{\mathrm{n}=1}^{\mathrm{N}} I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}}) - I(V_{2,\mathrm{n}}; Y_{1,\mathrm{n}}) + 2\mathrm{N}\epsilon, \tag{4.61}$$

and the following bounds on the sum-rate $R_1 + R_2$:

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{n=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})] \\ + \sum_{\mathrm{n}=1}^{\mathrm{N}} &[I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}|V_{1,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|V_{1,\mathrm{n}}, U_{\mathrm{n}})] - I(V_{1,\mathrm{n}}; Y_{2,\mathrm{n}}) + 2\mathrm{N}\epsilon, \end{aligned} \tag{4.62}$$

$$\begin{aligned}\mathrm{N}(R_1 + R_2) &\leq \sum_{\mathrm{n}=1}^{\mathrm{N}} [I(U_{\mathrm{n}}, V_{2,\mathrm{n}}; Y_{2,\mathrm{n}}) - I(V_{2,\mathrm{n}}; W_{\mathrm{n}}|U_{\mathrm{n}})] \\ + \sum_{\mathrm{n}=1}^{\mathrm{N}} &[I(U_{\mathrm{n}}, V_{1,\mathrm{n}}; Y_{1,\mathrm{n}}|V_{2,\mathrm{n}}) - I(V_{1,\mathrm{n}}; W_{\mathrm{n}}|V_{2,\mathrm{n}}, U_{\mathrm{n}})] - I(V_{2,\mathrm{n}}; Y_{1,\mathrm{n}}) + 2\mathrm{N}\epsilon. \end{aligned} \tag{4.63}$$

A time sharing RV Q, which is uniformly distributed over N symbols and independent of the RVs M_1, M_2, W, U, V_1, V_2, X, Y_1 and Y_2 is introduced for the single letter characterization of the above derived outer bounds. Applying the procedure similar to the one presented in [52, Chapter 15.3.4] on the N-letter expressions obtained in the above stated theorems, we get the outer bounds presented in Section 4.3. This completes the proofs of Theorem 4.3.2, Theorem 4.3.3, Theorem 4.3.5, Theorem 4.3.7 and Theorem 4.3.8.

4.6 Conclusions

We presented inner and outer bounds on the capacity region of three classes of two-user discrete memoryless broadcast channels, with noncausal side-information at the encoder. We generalized existing approaches to prove the achievability theorems, and characterized the rate penalties for having to deal with side-information at the encoder. For channels with confidentiality constraints, we showed that rate penalties exist for dealing with both side-information and confidentiality constraints. In the case of outer bounds, we focus on the explicit characterization of the sum-rate bounds. For channels where each receiver has *a priori* knowledge of the message of the other receiver, we showed that the outer bounds are only a factor away from the achievable region, where the factor is independent of the channel distribution.

Chapter 5

Z-channels with cooperation

5.1 Introduction

An information theoretic model of the CR channel comprises the classic $\mathrm{K}-$user interference channel with one (or more) transmitter(s) having *a priori* knowledge of the messages and the corresponding codewords of the others, and is commonly referred to as the interference channel with degraded message sets. For the case with $\mathrm{K} = 2$ users, the best known bounds and in some cases the capacity region were reported in [48], [97]. With $\mathrm{K} = 3$ users, the capacity bounds for three different message sharing schemes first appeared in [13] - [15], while some recent advances were reported in [98], [99].

Yet another popular channel model in multiuser information theory is the 'Z' channel (see Fig. 5.1(a)) [100] - [104], comprising a sender-receiver pair such that one of the senders encodes two independent messages intended to the two receivers, while the other sender only encodes one message without causing interference to the unintended receiver. A special case of the 'Z' channel is the Z-interference channel - the two senders encode independent messages, with each message intended to the pairing receiver, such that one of the transmitters does not interfere with the unintended receiver. Cooperation/cognition on the Z-interference channel was studied in [43] - [107].

In this chapter, we consider a Z-channel with degraded message sets (see

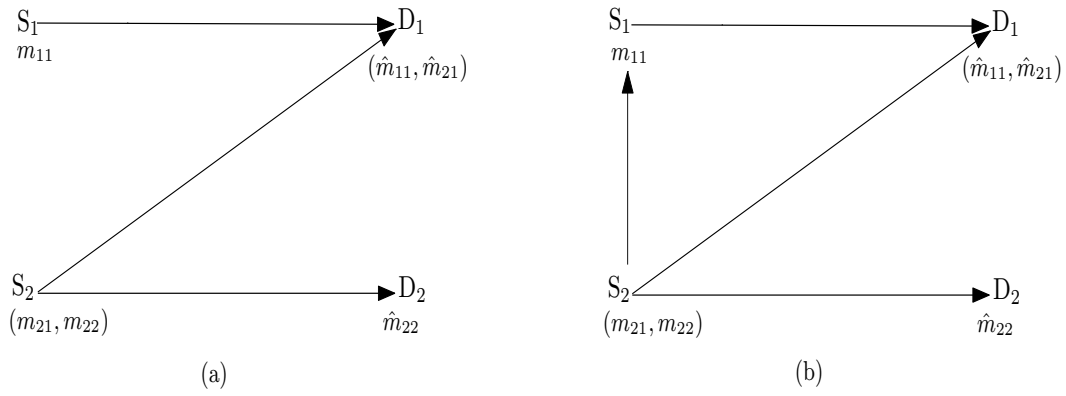

Figure 5.1: (a) Z-channel, (b) Z-channel with degraded message sets.

Fig. 5.1(b)), where sender S_1 (named cognitive) has noncausal knowledge of the messages and the corresponding codewords of S_2 (named primary). This channel model was considered in [108], where the authors employ rate-splitting at S_1 to derive bounds on the capacity region of only the discrete memoryless version of the channel. When the link between S_1-D_1 is noiseless, the bounds were shown to be tight, yielding the capacity region of the channel. In this chapter, we consider first the discrete memoryless channel model without employing rate-splitting at either of the encoders to derive lower and upper bounds on the capacity region. We then extend our results to the Gaussian case, to numerically evaluate and plot these bounds.

To prove our inner bounds, we employ Marton's broadcast code [23] at S_2, and derive an extension of Gel'fand-Pinsker's binning principle [8] at S_1. For the outer bounds, we generalize Nair and El Gamal's technique devised to derive the outer bounds for the classic two-user broadcast channel [30]. Graphical results obtained for the Gaussian case demonstrates the benefits, in terms of achievable rates, of noncausal message sharing between the primary and cognitive users. In addition, we report useful insights obtained by comparing the derived bounds with those for some well-known channel models.

The rest of the chapter is organized as follows. We first consider the discrete memoryless Z-channel with degraded message sets, for which we provide a mathematical model in Section 5.2 and state the achievability and converse theorems

in Section 5.3. In Section 5.4, we introduce the Gaussian channel model, plot the inner and outer bounds, and provide related discussion. We conclude the chapter in Section 5.5.

5.2 Channel Model and Preliminaries

The discrete memoryless Z-channel with degraded message sets is denoted C. For $t = 1, 2$, finite sets $\mathcal{X}_t$ and $\mathcal{Y}_t$ denote the channel input and output alphabets, respectively, random variables (RVs) $X_t \in \mathcal{X}_t$ and $Y_t \in \mathcal{Y}_t$ denote the inputs and outputs, respectively. N is the number of channel uses, and $\mathrm{n} = 1, \ldots, \mathrm{N}$ denotes the channel index. The following notation for a sequence of RVs is useful: $\mathbf{Y}_1^{\mathrm{N}} \triangleq (Y_{1,1}, \ldots, Y_{1,N})$; $\mathbf{Y}_1^{\mathrm{n}-1} \triangleq (Y_{1,1}, \ldots, Y_{1,\mathrm{n}-1})$; and $\mathbf{Y}_{1,\mathrm{n}+1}^{\mathrm{N}} \triangleq (Y_{1,\mathrm{n}+1}, \ldots, Y_{1,\mathrm{N}})$. For sake of brevity, we use $p(x)$ to denote $p(X = x)$. Unless otherwise stated, $p(\mathbf{x}) = \prod_{\mathrm{n}=1}^{\mathrm{N}} p(x_{\mathrm{n}})$. The two transition probabilities $p(\mathbf{y}_1|\mathbf{x}_1, \mathbf{x}_2)$ and $p(\mathbf{y}_2|\mathbf{x}_2)$ characterize the channel when $(X_1, X_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ is transmitted and $(Y_1, Y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ is obtained by the receivers.

To transmit its messages, S_2 generates two RVs $M_{2t} \in \mathcal{M}_{2t}$, where $\mathcal{M}_{2t} = \{1, \ldots, 2^{\mathrm{N}R_{2t}}\}$ denotes a set of message indices. Without loss of generality, $2^{\mathrm{N}R_{2t}}$ is assumed to be an integer, with R_{2t} being the transmission rate intended to D_t. M_{2t} denotes the message S_2 intends to transmit to D_t, and is assumed to be independently generated and uniformly distributed over the finite set $\mathcal{M}_{2t}$. Integer $m_{2t} \in \mathcal{M}_{2t}$ is a particular realization of M_{2t} and denotes the message-index. S_1 generates one RV $M_{11} \in \mathcal{M}_{11}$, where $\mathcal{M}_{11} = \{1, \ldots, 2^{\mathrm{N}R_{11}}\}$ with R_{11} being the transmission rate intended to D_1. Furthermore, S_1 treats the codewords corresponding to m_{21} and m_{22} as non-causally known interference at the encoder.

Given the conditional distribution characterizing the channel, a $((2^{\mathrm{N}R_{11}}, 2^{\mathrm{N}R_{21}}, 2^{\mathrm{N}R_{22}}), \mathrm{N}, P_e^{(\mathrm{N})})$ code for the channel C comprises N encoding functions f_1 and f_2, such that $\mathbf{X}_1 = \mathbf{f}_1(m_{11}, \mathbf{X}_2)$ and $\mathbf{X}_2 = \mathbf{f}_2(m_{21}, m_{22})$; and two decoding function $g_1 : \mathcal{Y}_1^{\mathrm{N}} \mapsto \mathcal{M}_{11} \times \mathcal{M}_{12}$ and $g_2 : \mathcal{Y}_2^{\mathrm{N}} \mapsto \mathcal{M}_2$ such that the average probability of decoding error $P_e^{(\mathrm{N})} \leq \lambda$. The average probability of decoding error for the code,

averaged over all codes, is $P_e^{(\mathrm{N})} = \max\{P_{e,1}^{(\mathrm{N})}, P_{e,2}^{(\mathrm{N})}\}$, where

$$\begin{aligned} P_{e,1}^{(\mathrm{N})} &= \sum_{m_{11},m_{21}} \frac{\Pr\left[g_1(\mathcal{Y}_1^{\mathrm{N}}) \neq (m_{11}, m_{21}) | (m_{11}, m_{21}, m_{22})\right]}{2^{\mathrm{N}[R_{11}+R_{21}]}}, \\ P_{e,2}^{(\mathrm{N})} &= \sum_{m_{22}} \frac{\Pr\left[g_2(\mathcal{Y}_2^{\mathrm{N}}) \neq (m_{22}) | (m_{21}, m_{22})\right]}{2^{\mathrm{N}R_{22}}}. \end{aligned}$$

For the channel C, $(R_{11}, R_{21}, R_{22}) \in \mathbb{R}^+$ is said to be achievable if there exists a $(2^{\lceil \mathrm{n}R_{11} \rceil}, 2^{\lceil \mathrm{n}R_{21} \rceil} 2^{\lceil \mathrm{n}R_{22} \rceil}, \mathrm{N}, P_e^{(\mathrm{N})})$ code such that $P_e^{(\mathrm{N})} \to 0$ as $\mathrm{N} \to \infty$. The capacity region is the closure of the set of all achievable rate triples (R_{11}, R_{21}, R_{22}) and is denoted by $\mathcal{C}$.

5.3 Capacity Bounds for C

For the channel C, let $\mathcal{P}^*$ denote the set of all joint probability distributions $p(.)$, that factor as follows:

$$\begin{aligned} p(w, u, v, x_1, x_2, y_1, y_2) = p(u, v)p(x_2|u, v)p(w|u, v) \\ p(x_1|w, u, v)p(y_1|x_1, x_2)p(y_2|x_2), \end{aligned}$$

where u, v, and w are particular realizations of the auxiliary RVs $U \in \mathcal{U}$, $V \in \mathcal{V}$, and $W \in \mathcal{W}$, respectively, defined on finite sets. For a given $p(.) \in \mathcal{P}^*$, an achievable rate region for C is described by $\mathcal{R}_{\mathrm{in}}(p)$, defined as the set of all rate triples $(R_{11}, R_{21}, R_{22}) \in \mathbb{R}^+$ such that the inequalities (5.1) - (5.5) are simultaneously satisfied:

$$R_{11} \leq I(W; Y_1|U) - I(U, V; W), \tag{5.1}$$

$$R_{21} \leq I(U; Y_1|W) - I(U; V), \tag{5.2}$$

$$R_{11} + R_{21} \leq I(W, U; Y_1) - I(U; V) - I(U, V; W), \tag{5.3}$$

$$R_{22} \leq I(V; Y_2) - I(U; V), \tag{5.4}$$

$$R_{21} + R_{22} \leq I(U; Y_1|W) + I(V; Y_2) - I(U; V). \tag{5.5}$$

Theorem 5.3.1. Let $\mathcal{R}_{\text{in}} = \bigcup_{p(.)\in\mathcal{P}^*} \mathcal{R}_{\text{in}}(p)$. The region $\mathcal{R}_{\text{in}}$ is an achievable rate region for the channel C, *i.e.*, $\mathcal{R}_{\text{in}} \subseteq \mathcal{C}$.

The proof of Theorem 5.3.1 is relegated to the Appendix D.

For a given $p(.) \in \mathcal{P}^*$, an outer bound on the capacity region $\mathcal{C}$ is described by $\mathcal{R}_{\text{out}}(p)$, defined as the set of all rate triples $(R_{11}, R_{21}, R_{22}) \in \mathbb{R}^+$ such that the inequalities (5.6) - (5.11) are simultaneously satisfied:

$$
\begin{aligned}
R_{11} &\leq I(W;Y_1), && (5.6)\\
R_{21} &\leq I(W,U;Y_1), && (5.7)\\
R_{11}+R_{21} &\leq I(W,U,V;Y_1), && (5.8)\\
R_{22} &\leq I(W,V;Y_2), && (5.9)\\
R_{21}+R_{22} &\leq I(W,U;Y_1)+I(V;W,U,Y_2) \\
&\quad -I(V;W,U), && (5.10)\\
R_{21}+R_{22} &\leq I(W,V;Y_2)+I(U;W,V,Y_1) \\
&\quad -I(U;W,V), && (5.11)
\end{aligned}
$$

where W, U and V satisfy the Markov chains: $W \to X_1 \to Y_1$ and $(U,V) \to X_2 \to (Y_1, Y_2)$.

Theorem 5.3.2. Let $\mathcal{R}_{\text{out}} = \bigcup_{p(.)\in\mathcal{P}^*} \mathcal{R}_{\text{out}}(p)$. The region $\mathcal{R}_{\text{out}}$ is an outer bound for the channel C, *i.e.*, $\mathcal{C} \subseteq \mathcal{R}_{\text{out}}$.

The proof of Theorem 5.3.2 follows from the one presented in Chapter 4 and is omitted.

5.4 The Gaussian Case

The Gaussian Z-channel with degraded message sets is denoted by C_{G} and is specified by the following input-output relationship.

$$
Y_1 = h_{11}X_1 + h_{21}X_2 + Z_1, \qquad (5.12)
$$

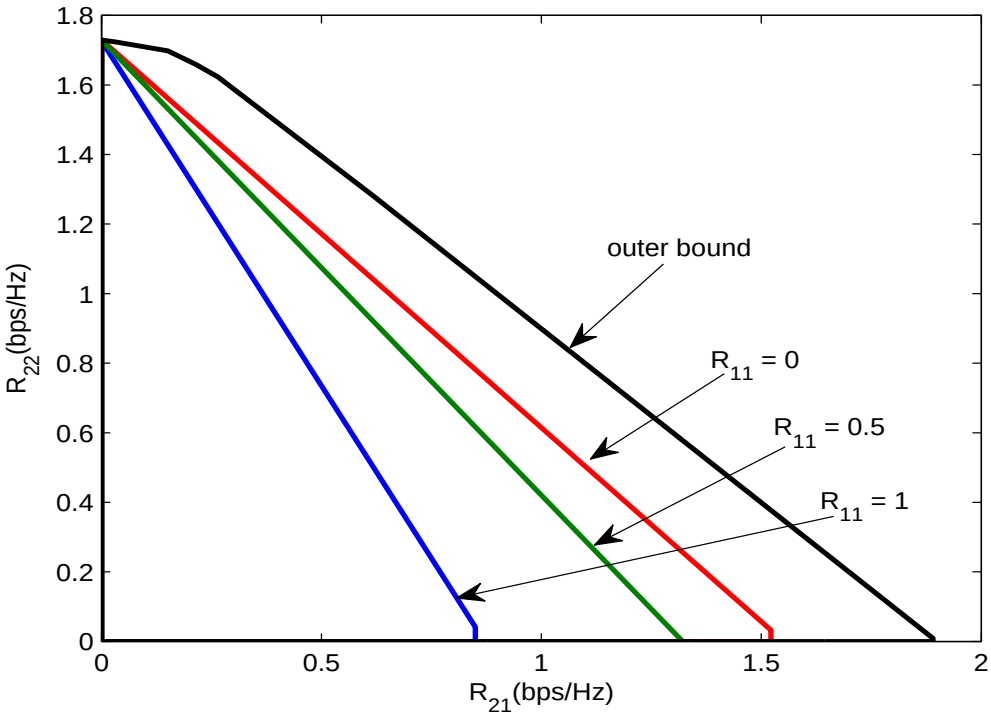

Figure 5.2: R_{21} Vs R_{22} for fixed $R_{11} = 0$ bps/Hz, $R_{11} = 0.5$ bps/Hz and $R_{11} = 1$ bps/Hz along with the outer bound.The power at the transmitters is 10dB.

$$Y_2 = h_{22}X_2 + Z_2, \tag{5.13}$$

where $Z_1 \sim \mathcal{N}(0, Q_1)$ and $Z_2 \sim \mathcal{N}(0, Q_2)$ are drawn i.i.d. The channel coefficients h_{11}, h_{21} and h_{22} are assumed to be real and globally known. For the channel $\mathrm{C_G}$, the RVs W, U and V denote the sources at the transmitters. We also consider the following RVs: $\tilde{W} \sim \mathcal{N}(0, P_1)$; $\tilde{U} \sim \mathcal{N}(0, \tau P_2)$; and $\tilde{V} \sim \mathcal{N}(0, \bar{\tau} P_2)$, where $\tau + \bar{\tau} = 1$. Further, we let $W = \tilde{W} + \alpha X_2$; $U = \tilde{U}$; and $V = \tilde{V} + \beta U$. The input RVs $X_1 = \tilde{W}$ and $X_2 = \tilde{U} + \tilde{V}$, so that $X_1 \sim \mathcal{N}(0, P_1)$ and $X_2 \sim \mathcal{N}(0, P_2)$. τ and $\bar{\tau}$ are randomly selected from the interval $[0, 1]$, while the values α and β are repeatedly generated according to $\mathcal{N}(0, 1)$.

Substituting for X_1 and X_2, we get

$$Y_1 = h_{11}\tilde{W} + h_{21}(\tilde{U} + \tilde{V}) + Z_1, \tag{5.14}$$

$$Y_2 = h_{22}(\tilde{U} + \tilde{V}) + Z_2. \tag{5.15}$$

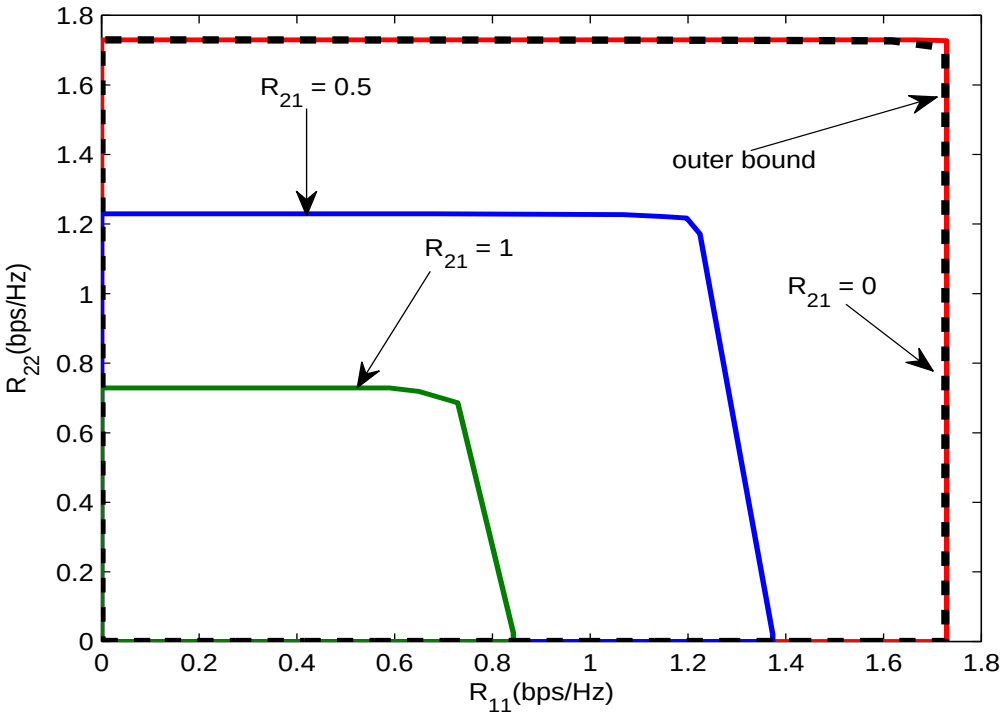

Figure 5.3: R_{11} Vs R_{22} for fixed $R_{21} = 0$ bps/Hz, $R_{21} = 0.5$ bps/Hz and $R_{21} = 1$ bps/Hz along with the outer bound.The power at the transmitters is 10dB.

We construct the following vector $\boldsymbol{\theta} = (Y_1, Y_2, W, U, V)$. The covariance matrix $\Sigma \triangleq \mathbb{E}[\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}]$ is then used to compute the mutual information terms (5.1) - (5.5) and (5.6) - (5.11). Owing to space limitation, we do not show this computation.

The resulting plots of achievable rate regions and outer bounds are shown next. In Fig. 5.2, we plot the rate regions and the outer bound for R_{21} versus R_{22} when R_{11} is promised a constant rate of 0, 0.5 and 1 bps/Hz. As shown, with increasing R_{11}, the achievable region shrinks. Furthermore, it is important to note that there is no change in the maximum achievable rate of R_{22} since the transmissions of S_1 does not interfere with D_2. The shrinkage in the rate region is solely attributed to the reduction in the maximum achievable rate R_{21}. Also note that, the outer bound is only plotted for the case of $R_{11} = 0$ bps/Hz. In Fig. 5.3, we plot the rate regions and the outer bound for R_{11} versus R_{22} when R_{21} is promised a constant rate of 0, 0.5 and 1 bps/Hz. Similar to Fig. 5.2, with increasing R_{21}, the achievable region shrinks. However, unlike Fig. 5.2, the shrinkage is due to the

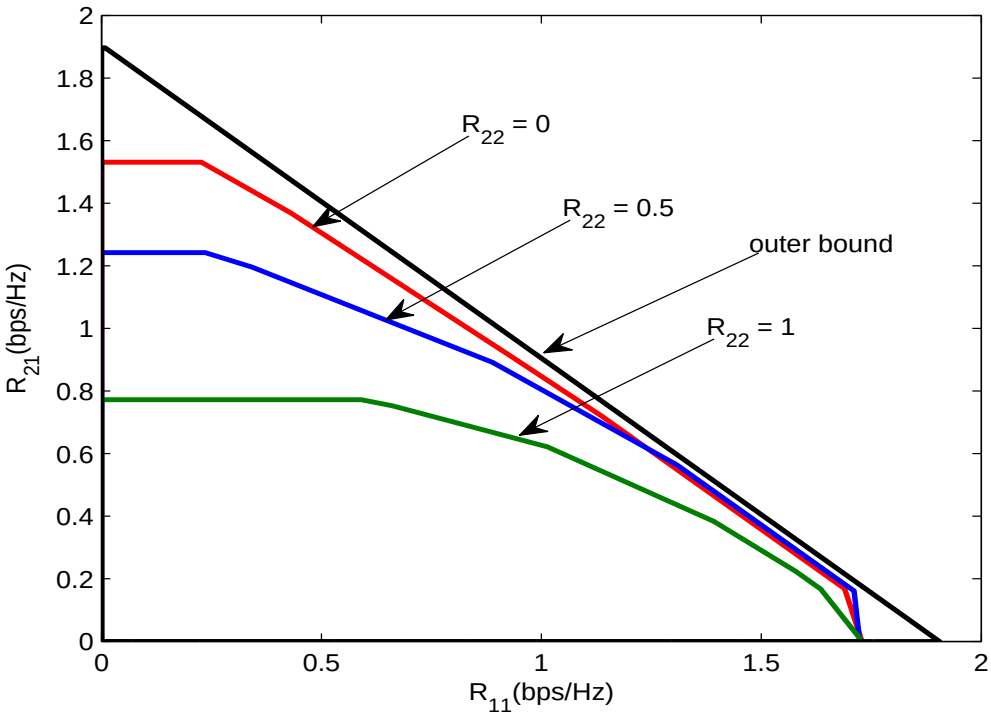

Figure 5.4: R_{11} Vs R_{21} for fixed $R_{22} = 0$ bps/Hz, $R_{22} = 0.5$ bps/Hz and $R_{22} = 1$ bps/Hz along with the outer bound.The power at the transmitters is 10dB.

reduction in the maximum achievable rates of both R_{11} and R_{22} since the rate R_{21} not only acts as interference to D_2, but also penalizes the maximum achievable R_{11}. As before, the outer bound is only plotted for the case of $R_{21} = 0$ bps/Hz, which coincides with the achievable rate region, thereby yielding the capacity region of the channel.

Lastly, in Fig. 5.4, we plot the rate regions and the outer bound for R_{11} versus R_{21} when R_{22} is promised a constant rate of 0, 0.5 and 1 bps/Hz. Similar to Fig. 5.2 and Fig. 5.3, with increasing R_{22}, the achievable region shrinks. Again, the shrinkage is due to the reduction in the maximum achievable rates of both R_{11} and R_{21} since the rate R_{22} acts as interference to both D_1 and D_2. The outer bound is only plotted for the case of $R_{22} = 0$ bps/Hz.

5.5 Conclusions

We studied cooperation on the Z-channel under the assumption that the cognitive transmitter had noncausal knowledge of the message sets and the corresponding codewords of the primary user. We derived lower bounds on the capacity region by employing Marton's coding technique at the non-cognitive encoder; at the cognitive transmitter, we presented a generalization of Gel'fand-Pinsker's binning technique for handling two sequences known *a priori* at the encoder. Plots of the achievable rate regions and outer bounds were presented by considering the Gaussian channel model, and some interesting observations were revealed.

Chapter 6

Future work

In this chapter, we propose two research problems dealing with cooperation and security in wireless networks. First, we plan to establish upper and lower bounds on the capacity region of a broadcast network aided by a layer of relay nodes, such that the encoder is constrained to keep each message confidential from the unintended receiver(s). Second, we plan to work on deriving capacity bounds, and in some case even the capacity regions, for the Z-channel. The Z-channel is especially difficult since it is a combination of the broadcast and multiple access channels. In this chapter, we formally define the problem and suggest some avenues for possible solution of the proposed problems.

6.1 Secure broadcasting with relays

Consider a broadcast network aided by relay nodes, as shown in Fig. 6.1. The main sender, denoted $\mathrm{Base\ Station}$, has k confidential messages intended to k receivers non-cooperating receivers. The encoder of the $\mathrm{Base\ Station}$ is constrained to keep each message secrect/confidential from the unintended receivers. For example, in Fig. 6.1, message m_1 intended to $\mathrm{Receiver}_1$ has to be kept confidential from $\mathrm{Receiver}_2$ and $\mathrm{Receiver}_3$; similarly for messages m_2 and m_3. We plan to establish inner and outer bounds on the capacity region of this channel, by employing a combination of Marton's achievability scheme for the classical broad-

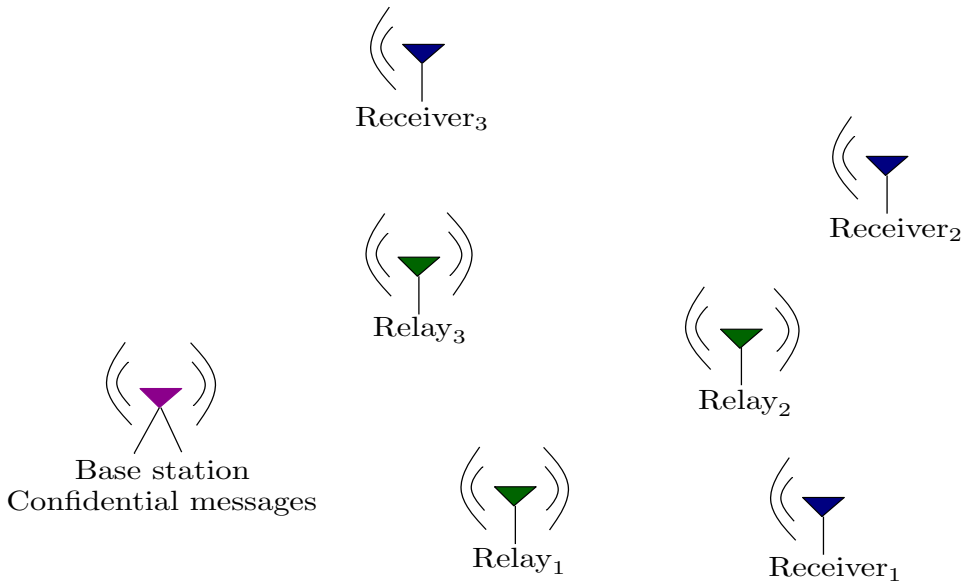

Figure 6.1: Secure broadcasting with relays.

cast channel [23], block Markov superposition coding [17] and backward decoding [19], in conjunction with stochastic encoders to achieve information-theoretic secrecy [20].

6.1.1 Related work

The problem of confidentiality/security in relay networks has been addressed along various lines in the information theory literature. Capacity bounds for cooperation in wireless networks was presented in [66], where authenticated relay nodes employ noise insertion strategies [67] to achieve secrecy. In [68], an opportunistic selection technique of two relay nodes was presented to secure communications between a source-destination pair from the eavesdropper. The first relay employed a simple decode-and-forward strategy, while the second relay is used to create intentional interference at the eavesdropper, thereby jamming its reception. The relay-eavesdropper channel was considered in [69], where user-cooperation has been exploited to achieve secrecy. In particular, the relay node employed a novel noise-forwarding strategy to confuse the eavesdropper. However, the relay was considered to be a *deaf-helper*, in the sense that it is totally

ignorant of the transmitted messages.

An information-theoretic approach to secure broadcasting was inspired by the pioneering work of Csiszár and Körner [88], who derived capacity bounds for the two-user BC, when the sender transmits a private message to receiver 1 and a common message to both receivers, while keeping the private message confidential from receiver 2. In [90], capacity bounds were derived for BC where a sender broadcasts two independent messages to two receivers, while keeping each message confidential from the unintended receiver. Capacity results and bounds for Gaussian BC with confidential messages were reported in [91] - [93]. The reader is referred to [94] for a comprehensive review of physical-layer security in BC.

6.1.2 Proposed methodology

We propose the following coding/decoding scheme:

1. The main sender employs a combination of block Markov superposition coding and Marton's achievability scheme, which further involves GPs' binning principle [8] and the second moment method for bounding the binning rates [24]. It also employs stochastic encoders to achieve secrecy.

2. The relay nodes perform decode-and-forward operations for every block of data it receives from the sender. However, before forwarding, they again perform stochastic binning to ensure full secrecy to respective receivers. Furthermore, during the first block when the relays receive no data from the main sender, they use special noise-insertion strategies [67] to ensure that secrecy is not compromised on any block.

3. The receivers employ backward decoding, where they accumulate all the blocks and start decoding from the last block. Note that, this introduces extensive delays; however, under asymptotic conditions (in the number of blocks), the rates achieved cannot be traded against secrecy.

6.2 Secure broadcasting via distributed coordination

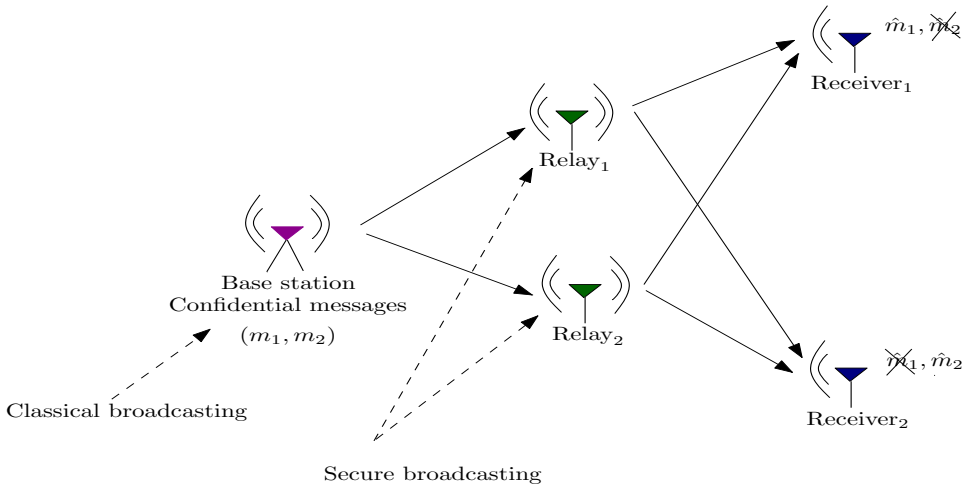

Figure 6.2: Secure broadcasting via distributed coordination.

In this work, we suggest an information-theoretic viewpoint of broadcasting, cooperation and security. Specifically, as shown in Fig. 6.2, it is a variation (in fact, a special case) of the work proposed in Section 6.1. In this model, the Base Station has no direct channel gain to the receivers. As such, it enlists the help of the intermediate relay nodes to improve its throughput to the desired destination. Similar to Section 6.1, we wish to keep the messages confidential from the unintended receiver. However, the notable difference is that the Base Station does not need to perform stochastic encoding since it does not see a channel gain to either of the receivers. Instead, the full responsibility of secure communications is handed over the intermediate relay nodes. That is, all the relay nodes participating in the communication scenario perform stochastic encoding in conjunction with Markov superposition coding to achieve the desired secrecy.

6.2.1 Related work

The channel model considered in Section 6.2 is the broadcast version of the diamond channel first presented in [109], in which a single source transmits to a single destination in the presence of two relay nodes; there is no direct channel gain between the source and the destination. In [110], a special case of the di-

amond channel was considered where the channel between the source and one of the relays is noisy, while the channel between the source and the other relay is noiseless. The capacity region was derived, and is shown to be strictly smaller than the cut-set bound. Another special case of the diamond channel is considered in [111], where the channel between the source and the relays are of finite capacities, while the channel between the relay nodes and the destination is a Gaussian multiple access channel. The capacity region of this channel is established when the separate links to the relays are of the same capacity and the power constraints of the two relays are identical. In [112], the capacity region of the diamond channel is established when the outputs of the two relays are deterministic functions of the source's input, and the channel between the relays and destination is noiseless.

6.2.2 Proposed methodology

For the channel model presented in Fig. 6.2, we propose the following coding scheme:

1. The $\mathrm{Base\ Station}$ employs the standard Marton's achievability scheme [23] for the two-user broadcast channel to transmit the message pair (m_1, m_2) to the relays Relay_1 and Relay_2. Note that, $\mathrm{Base\ Station}$ does not have a direct channel gain to either $\mathrm{Receiver}_1$ or $\mathrm{Receiver}_2$.

2. Relay_k, $k = 1, 2$, then uses a combination of block Markov superposition coding [19], binning [76], and stochastic encoding to not only aid transmissions from $\mathrm{Base\ Station}$, but also to keep m_1 secret from $\mathrm{Receiver}_1$ and m_2 secret from $\mathrm{Receiver}_2$.

3. Lastly, $\mathrm{Receiver}_k$, $k = 1, 2$, employ backward decoding to recover the transmitted information.

Bibliography

[1] J. Mitola, *Cognitive Radio Architecture: The Engineering Foundations of Radio XML*. New York: John Wiley & Sons, Inc, Sep. 2006.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Selected Areas in Comm.*, vol. 23, pp. 201–220, Feb. 2005.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, Part I: System description," *IEEE Trans. Comm.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[4] ——, "User cooperation diversity, Part II: Implementation aspects and performance analysis," *IEEE Trans. Comm.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[5] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[6] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, pp. 355–580, Apr. 2009.

[7] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 49–60, Jan. 1981.

[8] S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[9] T. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.

[10] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[11] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[12] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[13] K. G. Nagananda and C. R. Murthy, "Three-user cognitive channels with cumulative message sharing: An achievable rate region," in *Proc. IEEE Inf. Theory Workshop on Net. and Inf. Theory*, Volos, Greece, Jun. 2009, pp. 291–295.

[14] ——, "Information theoretic results for three-user cognitive radio channels," in *Proc. IEEE Global Telecomm. Conf.*, Hawaii, USA, Nov. 2009, pp. 1–6.

[15] K. G. Nagananda, C. R. Murthy, and S. Kishore, "Achievable rates in three-user interference channels with one cognitive transmitter," in *Proc. IEEE Int. Conf. Signal Process. and Comm.*, Bangalore, India, Jul. 2010, pp. 1–5.

[16] K. G. Nagananda, P. Mohapatra, C. R. Murthy, and S. Kishore, "Multiuser cognitive radio networks: An information theoretic perspective," *Int. Journal Advances in Eng. Sci. App. Mathematics*, 2012, under revision.

[17] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[18] A. E. Gamal and E. C. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 120–122, Jan. 1981.

[19] F. M. J. Willems and E. C. van der Meulen, "The discrete memoryless multiple access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[20] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[21] K. G. Nagananda, "Secure communications over opportunistic-relay channels," *Physical Communication*, 2012.

[22] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, SA, Sep. 2005, pp. 2184–2188.

[23] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[24] N. Alon and J. H. Spencer, *The Probabilistic Method*, 2nd ed. New York: John Wiley, 2000.

[25] T. Oechtering and M. Skoglund, "Coding for the bidirectional broadcast channel with random states known at the encoder," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, S.Korea, Jul. 2009, pp. 2013–2017.

[26] G. Kramer and S. Shamai (Shitz), "Capacity for classes of broadcast channels with receiver side information," in *Proc. IEEE Inf. Theory Workshop*, Tahoe City, CA, Sep. 2007, pp. 313–318.

[27] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.

[28] K. G. Nagananda, C. R. Murthy, and S. Kishore, "State-dependent broadcast channels with noncausal encoder side-information," Jun. 2013, submitted to IEEE Int. Conf. Comm.

[29] ——, "Capacity bounds for state-dependent broadcast channels," *Physical Communication*, 2012, under revision.

[30] C. Nair and A. E. Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-53, pp. 350–355, Jan. 2007.

[31] A. Goldsmith, S. A. Jafar, I. Marić, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[32] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[33] ——, "Limits on communications in a cognitive radio channel," in *IEEE Comm. Magazine*, Jun. 2006, vol. 44, no. 6, pp. 44–49.

[34] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

[35] A. Jovičić and P. Vishwanath, "Cognitive radio: An information theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[36] J. Jiang and Y. Xin, "On the achievable rate regions for interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4707–4712, Oct. 2008.

[37] I. Marić, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, Oct. 2007.

[38] I. Marić, A. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with one cooperating transmitter," *European Trans. Telecomm.*, vol. 19, pp. 405–420, Apr. 2008.

[39] ——, "On the capacity of interference channels with a cognitive transmitter," in *Proc. IEEE Inf. Theory and App. Workshop*, La Jolla, CA, Jan. - Feb. 2007, pp. 268 –273.

[40] ——, "On the capacity of interference channels with a partially-cognitive transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Nise, France, Jun. 2007, pp. 2156 –2160.

[41] ——, "An achievable rate region for interference channels with one cooperating transmitter," in *Proc* 41^{st} *Asilomar Conf. Signals, Syst. and Comp.*, Pacific Grove, CA, Nov. 2007, pp. 888–892.

[42] I. Marić, R. Yates, and G. Kramer, "The capacity region of the strong interference channel with common information," in *Proc.* 43^{rd} *Asilomar Conf. Signals, Syst. and Comp.*, Pacific Grove, CA, Oct. 2005, pp. 1737–1741.

[43] Y. Cao and B. Chen, "Interference channel with one cognitive transmitter," in *Proc.* 42^{nd} *Asilomar Conf. Signals Syst. and Comp.*, Pacific Grove, CA, Oct. 2008, pp. 1593–1597.

[44] J. Jiang, Y. Xin, and H. Garg, "Interference channels with common information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 171–187, Jan. 2008.

[45] Y. Cao, B. Chen, and J. Zhang, "A new achievable rate region for interference channels with common information," in *Proc. IEEE Wireless Comm. Net. Conf.*, Hong Kong, Mar. 2007, pp. 2069–2073.

[46] Y. Cao and B. Chen, "Outer bounds for the capacity region of Gaussian interference channels with common information," in *Proc. IEEE Global Telecomm. Conf.*, Washington DC, Nov. 2007, pp. 1622–1626.

[47] I. Marić, R. D. Yates, and G. Kramer, "The capacity region of the strong interference channel with common information," in 39^{th} *Asilomar Conf. Signals, Syst. and Comp.*, Pacific Grove, CA, Nov. 2005, pp. 1737– 741.

[48] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the discrete memoryless cognitive interference channel and some capacity results," *IEEE Trans*, vol. 57, no. 7, pp. 4087–4109, Jul. 2011.

[49] ——, "The capacity of the semi-deterministic cognitive interference channel and its application to constant gap results for the Gaussian channel," in *Proc. IEEE Int. Conf. Comm.*, Kyoto, Japan, Jun. 2011. [Online]. Available: http://arxiv.org/abs/1009.3083

[50] ——, "Capacity to within 3 bits for a class of Gaussian interference channels with a cognitive relay," in *Proc. Int. Symp. Inf. Theory*, St. Petersburg, Russia, Aug. 2011. [Online]. Available: http://arxiv.org/abs/1102.3225

[51] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, Inc, 1968.

[52] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[53] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[54] G. Kramer, I. Marić, and R. D.Yates, "Cooperative communications," *Found. Trends Net.*, vol. 1, no. 3-4, pp. 271–425, 2006.

[55] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, no. 1, pp. 120–154, 1971.

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

[57] R. Tannious and A. Nosratinia, "Relay channel with private messages," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3777–3785, Oct. 2007.

[58] N. Marina and A. Hjørungnes, "Characterization of the secrecy region of a single relay cooperative system," in *Proc. IEEE Wireless Comm. and Net. Conf.*, Apr. 2010, pp. 1–6.

[59] N. Marina, H. Yagi, and H. Poor, "Improved rate-equivocation regions for secure cooperative communication," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011, pp. 2871–2875.

[60] N. Marina, R. Bose, and A. Hjørungnes, "Increasing the secrecy capacity by cooperation in wireless networks," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Comm.*, Sep. 2009, pp. 1978–1982.

[61] Y. Oohama, "Coding for relay channels with confidential messages," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Qld., Australia, Sep. 2001, pp. 87–89.

[62] H. Yanikomeroglu, "Cellular multihop communications: Infrastructure-based relay network architecture for 4G wireless systems," in *Proc.* 22^{nd} *Biennial Symp. Comm.*, Ontario, Canada, Jun. 2004.

[63] G.-C. Zhang, X.-H. Peng, and X.-Y. Gu, "Implementation and performance evaluation of an experimental wireless relay sensor network," in *Proc. IEEE Int. Conf. High Performance Comp. Comm.*, Washington, DC, USA, 2008, pp. 513–519.

[64] H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMax and LTE-advanced mobile systems," *IEEE Comm. Magazine*, vol. 47, no. 10, pp. 100–105, Oct. 2009.

[65] M. Çinar, "Implementation of relay-based systems in wireless cellular networks," Aug. 2010, M.Sc. thesis, İzmir Institute of Technology.

[66] E. Perron, S. Diggavi, and E. Telatar, "On cooperative secrecy for discrete memoryless relay networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, Jun. 2010, pp. 2573–2577.

[67] ——, "On noise insertion strategies for wireless network secrecy," in *Proc. IEEE Inf. Theory App. Workshop*, San Diego, CA, 2009, pp. 77–84.

[68] I. Krikidis, J. Thompson, and S. Mclaughlin, "Relay selection for secure cooperative networks with jamming," *IEEE Trans. Wireless Comm.*, vol. 8, no. 10, pp. 5003–5011, Oct. 2009.

[69] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[70] M. Bloch and A. Thangaraj, "Confidential messages to a cooperative relay," in *Proc. IEEE Inf. Theory Workshop*, Porto, May 2008, pp. 154–158.

[71] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3807–3827, Aug. 2010.

[72] ——, "Two-hop secure communication using an untrusted relay," *EURASIP J. Wireless Comm. and Net.*, vol. 2009, Article ID 305146, 13 pages, 2009, doi: 10.1155/2009/305146.

[73] V. Aggarwal, L. Sankar, A. R. Calderbank, and H. V. Poor, "Secrecy capacity of a class of orthogonal relay eavesdropper channels," *EURASIP J. Wireless Comm. and Net.*, vol. 2009, Article ID 494696, 14 pages, 2009, doi:10.1155/2009/494696.

[74] E. Ekrem and S. Ulukuş, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 137–155, Jan. 2011.

[75] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.

[76] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[77] L. Lai, K. Liu, and H. El Gamal, "The three-node wireless network: Achievable rates and cooperation strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 805–828, Mar. 2006.

[78] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[79] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5747–5755, Dec. 2008.

[80] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[81] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," *IEEE J. Selected Topics in Signal Process.*, vol. 2, no. 1, pp. 103–117, Feb. 2008.

[82] A. Gohari, A. El Gamal, and V. Anantharam, "On an outer bound and an inner bound for the general broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 540–544.

[83] H. Sato, "An outer bound to the capacity region of broadcast channels (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[84] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 166–169, Mar. 1979.

[85] C. Nair, "A note on outer bounds for broadcast channel," in *Proc. Int. Zurich Seminar Comm.*, 2010. [Online]. Available: http://arxiv.org/abs/1101.0640v1

[86] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2867–2877, Aug. 2005.

[87] G. Kramer, "Topics in multi-user information theory," *Found. Trends Comm. Inf. Theory*, vol. 4, no. 4-5, pp. 265–444, 2007. [Online]. Available: http://ee.usc.edu/~gkramer/Papers/kramerNOW07.pdf

[88] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[89] E. Ekrem and S. Ulukuş, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP J. Wireless Comm. and Net.*, vol. 2009, Article ID 824235, 29 pages, 2009, doi: 10.1155/2009/824235.

[90] R. Liu, I. Marić, P. Spasojević, and R. D.Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[91] G. Bagherikaram, A. Motahari, and A. Khandani, "Secrecy capacity region of Gaussian broadcast channel," in *Proc. IEEE* 43^{rd} *Annual Conf. Inf. Sciences Syst.*, Baltimore, MD, Mar. 2009, pp. 152–157.

[92] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235–1249, Mar. 2009.

[93] H. Ly, T. Liu, and Y. Liang, "Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5477–5487, Nov. 2010.

[94] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Physical layer security in broadcast networks," *Security and Comm. Networks,*, vol. 2, no. 3, pp. 227–238, May-Jun. 2009.

[95] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proc.* 19^{th} *Int. Conf. Theory App. Crypt. Tech.*, Bruges, Belgium, 2000, pp. 351–368.

[96] Y. Chen and A. J. Han Vinck, "Wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 395–402, Jan. 2008.

[97] S. Rini, D. Tuninetti, and N. Devroye, "Inner and outer bounds for the gaussian cognitive interference channel and new capacity results," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 820–848, Feb. 2012.

[98] D. Maamari, N. Devroye, and D. Tuninetti, "The sum-capacity of the linear deterministic three-user cognitive interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Boston, MA, 2012.

[99] M. Mirmohseni, B. Akhbari, and M. R. Aref, "Three-user cognitive interference channel: Capacity region with strong interference," *IET Comm.*, 2012, to appear.

[100] S. Vishwanath, N. Jindal, and A. Goldsmith, "The "Z" channel," in *Proc. IEEE Global Telecomm. Conf.*, vol. 3, San Francisco, CA, Dec. 2003, pp. 1726–1730.

[101] N. Liu and S. Ulukuş, "On the capacity region of the Gaussian "Z" channel," in *Proc. IEEE Global Telecomm. Conf.*, vol. 1, Dallas, TX, Nov. 2004, pp. 415–419.

[102] H.-F. Chong, M. Motani, and H. K. Garg, "Capacity theorems for the "Z" channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1348–1365, Apr. 2007.

[103] N. Liu and A. Goldsmith, "Capacity regions and bounds for a class of Z-interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4986–4994, Nov. 2009.

[104] S. Salehkalaibar and M. Aref, "On the capacity region of the degraded Z channel," in *Proc. IEEE Inf. Theory Workshop*, Dublin, Sep. 2010, pp. 1–5.

[105] Y. Cao and B. Chen, "Capacity outer bounds for the cognitive Z channel," in *Proc. IEEE Global Telecomm. Conf.*, Honolulu, HI, Nov. 2009, pp. 1–6.

[106] ——, "The cognitive Z channel," in *Proc.* 43^{rd} *Annual Conf. Inf. Sci. and Syst.*, Baltimore, MD, Mar. 2009, pp. 528–532.

[107] N. Liu, I. Maric, A. Goldsmith, and S. Shamai (Shitz), "Bounds and capacity results for the cognitive Z-interference channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, S. Korea, 2009, pp. 2422–2426.

[108] Y. Cheng and N. Liu, "Bounds and capacity results for the cognitive Z channel," Jun. 2012, accepted at IEEE Int. Conf. Comm.

[109] B. E. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Massachusetts Institute of Technology, Sep. 2001.

[110] W. Kang and S. Ulukuş, "Capacity of a class of diamond channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4955–4960, Aug. 2011.

[111] W. Kang and N. Liu, "The Gaussian multiple access diamond channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 2011, pp. 1499–1503.

[112] R. Tandon and S. Ulukuş, "Diamond channel with partially separated relays," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 644–648.

Appendix A

Proofs for the interference channel

A.1 Proof of Theorem 2.2.1

Here, we present the proof of achievability for the channel $\mathcal{C}^2_{\text{CuMS}}$. The proof is presented in four parts, namely, codebook generation, encoding, decoding and analysis of probabilities of decoding errors at the three receivers. We start with the codebook generation scheme.

A.1.1 Codebook Generation

Let us fix $p(.) \in \mathcal{P}^2_{\text{CuMS}}$. Generate a random time sharing codeword $\mathbf{q}$, of length n, according to the distribution $\prod_{i=1}^{n} p(q_i)$. Generate $2^{nR_{11}}$ independent codewords $\mathbf{W}(j)$, according to $\prod_{i=1}^{n} p(w_i|q_i)$. For every $\mathbf{w}(j)$, generate one codeword $\mathbf{X}_1(j)$ according to $\prod_{i=1}^{n} p(x_{1i}|w_i(j), q_i)$. For $\tau = 1, 2$, generate $2^{n(R_{2\tau}+I(W;U_\tau|Q)+4\epsilon)}$ independent codewords $\mathbf{U}_\tau(l_\tau)$, according to $\prod_{i=1}^{n} p(u_{\tau i}|q_i)$. For every codeword triple $[\mathbf{u}_1(l_1), \mathbf{u}_2(l_2), \mathbf{w}(j)]$, generate one codeword $\mathbf{X}_2(l_1, l_2, j)$ according to $\prod_{i=1}^{n} p(x_{2i}|u_{1i}(l_1), u_{2i}(l_2), w_i(j), q_i)$. Uniformly distribute the $2^{n(R_{2\tau}+I(W;U_\tau|Q)+4\epsilon)}$ codewords $\mathbf{U}_\tau(l_\tau)$ into $2^{nR_{2\tau}}$ bins indexed by $k_\tau \in \left\{1, \ldots, 2^{nR_{2\tau}}\right\}$ such that each bin contains $2^{n(I(W;U_\tau|Q)+4\epsilon)}$ codewords. For $\rho = 1, 3$, generate

$2^{n(R_{3\rho}+I(W,U_1,U_2;V_\rho|Q)+4\epsilon)}$ independent codewords $\mathbf{V}_\rho(t_\rho)$, according to $\prod_{i=1}^{n} p(v_{\rho i}|q_i)$. For every codeword quadruple $[\mathbf{v}_1(t_1), \mathbf{v}_3(t_3), \mathbf{u}_1(l_1), \mathbf{u}_2(l_2), \mathbf{w}(j)]$, generate one codeword $\mathbf{X}_3(t_1, t_3, l_1, l_2, j)$ according to $\prod_{i=1}^{n} p\left(x_{3i}|v_{1i}(t_1), v_{3i}(t_3), u_{1i}(l_1), u_{2i}(l_2), w_i(j), q_i\right)$. Distribute $2^{n(R_{3\rho}+I(W,U_1,U_2;V_\rho|Q)+4\epsilon)}$ codewords $\mathbf{V}_\rho(t_\rho)$ uniformly into $2^{nR_{3\rho}}$ bins indexed by $r_\rho \in \left\{1, \ldots, 2^{nR_{3\rho}}\right\}$ such that each bin contains $2^{n(I(W,U_1,U_2;V_\rho|Q)+4\epsilon)}$ codewords. The indices are given by $j \in \left\{1, \ldots, 2^{nR_{11}}\right\}$, $l_\tau \in \left\{1, \ldots, 2^{n(R_{2\tau}+I(W;U_\tau|Q)+4\epsilon)}\right\}$ and $t_\rho \in \{1, \ldots, 2^{n(R_{3\rho}+I(W,U_1,U_2;V_\rho|Q)+4\epsilon)}\}$.

A.1.2 Encoding & Transmission

Let $A_\epsilon^{(n)}$ be a typical set. We will be using the notation $A_\epsilon^{(n)}$ to describe a typical set over many different random variables, but the definition will be clear from the context.

Let us suppose that the source message vector generated at the three senders is

$(m_{11}, m_{21}, m_{22}, m_{31}, m_{33}) = (j, k_1, k_2, r_1, r_3)$. At the encoders, the first component is treated as the message index and the last four components are treated as the bin indices. $\mathcal{S}_2$ looks for a codeword $\mathbf{u}_1(l_1)$ in bin k_1 and a codeword $\mathbf{u}_2(l_2)$ in bin k_2 such that $(\mathbf{u}_1(l_1), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$ and $(\mathbf{u}_2(l_2), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$, respectively. $\mathcal{S}_3$ looks for a codeword $\mathbf{v}_1(t_1)$ in bin r_1 and a codeword $\mathbf{v}_3(t_3)$ in bin r_3 such that $(\mathbf{v}_1(t_1), \mathbf{u}_1(l_1), \mathbf{u}_2(l_2), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$ and $(\mathbf{v}_3(t_3), \mathbf{u}_1(l_1), \mathbf{u}_2(l_2), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$, respectively. $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ then transmit codewords $\mathbf{x}_1(j)$, $\mathbf{x}_2(l_1, l_2, j)$ and $\mathbf{x}_3(t_1, t_3, l_1, l_2, j)$, respectively, through n channel uses. The transmissions are assumed to be synchronized.

A.1.3 Decoding

Recall that in $\mathcal{C}^2_{\text{CuMS}}$, the primary receiver can decode the public parts of the non-pairing sender's messages, while the secondary receivers can only decode the messages from their pairing transmitters. The three receivers accumulate an n-length channel output sequence: $\mathbf{y}_1$ at $\mathcal{R}_1$, $\mathbf{y}_2$ at $\mathcal{R}_2$ and $\mathbf{y}_3$ at $\mathcal{R}_3$. Decoders

1, 2 and 3 look for all indices $(\hat{j}, \hat{\hat{l}}_1, \hat{\hat{t}}_1)$, $(\hat{l}_1, \hat{l}_2)$ and $(\hat{t}_1, \hat{t}_3)$, respectively, such that $(\mathbf{w}(\hat{j}), \mathbf{u}_1(l_1), \mathbf{v}_1(t_1), \mathbf{y}_1, \mathbf{q}) \in A_\epsilon^{(n)}$, $(\mathbf{u}_1(\hat{l}_1), \mathbf{u}_2(\hat{l}_2), \mathbf{y}_2, \mathbf{q}) \in A_\epsilon^{(n)}$ and $(\mathbf{v}_1(\hat{t}_1), \mathbf{v}_3(\hat{t}_3), \mathbf{y}_3, \mathbf{q}) \in A_\epsilon^{(n)}$. If $\hat{j}$ in all the index triples found are the same, $\mathcal{R}_1$ declares $m_{11} = \hat{j}$, for some l_1 and t_1. If $\hat{l}_1$ in all the index pairs found are indices of codewords $\mathbf{u}_1(\hat{l}_1)$ from the same bin with index $\hat{k}_1$, and $\hat{l}_2$ in all the index pairs found are indices of codewords $\mathbf{u}_2(\hat{l}_2)$ from the same bin with index $\hat{k}_2$, then $\mathcal{R}_2$ determines $(m_{21}, m_{22}) = (\hat{k}_1, \hat{k}_2)$. Similarly, if $\hat{t}_1$ in all the index pairs found are indices of codewords $\mathbf{v}_1(\hat{t}_1)$ from the same bin with index $\hat{r}_1$, and $\hat{t}_3$ in all the index pairs found are indices of codewords $\mathbf{v}_3(\hat{t}_3)$ from the same bin with index $\hat{r}_3$, then $\mathcal{R}_3$ determines $(m_{31}, m_{33}) = (\hat{r}_1, \hat{r}_3)$. Otherwise, the receivers $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ declare an error.

A.1.4 Analysis of the Probabilities of Error

Upper bounds on the probabilities of error events which could happen during encoding and decoding processes are derived using typicality arguments [52]. Here, we only show the analysis of the probability of encoding error at the sender $\mathcal{S}_2$, and the probability of decoding error at the receiver $\mathcal{R}_1$.

We assume that a source message vector $(m_{11}, m_{21}, m_{22}, m_{31}, m_{33}) = (j, k_1, k_2, r_1, r_3)$ is encoded and transmitted. First, let us define the following events:
(i) $E_{jl_1} \triangleq \left\{(\mathbf{W}(j), \mathbf{U}_1(l_1), \mathbf{q}) \in A_\epsilon^{(n)}\right\}$,
(ii) $E_{jl_2} \triangleq \left\{(\mathbf{W}(j), \mathbf{U}_2(l_2), \mathbf{q}) \in A_\epsilon^{(n)}\right\}$,
(iii) $E_{jl_1t_1} \triangleq \left\{(\mathbf{W}(j), \mathbf{U}_1(l_1), \mathbf{V}_1(t_1), \mathbf{Y}_1, \mathbf{q}) \in A_\epsilon^{(n)}\right\}$.
$E_{(.)}^c \triangleq$ complement of the event $E_{(.)}$. Events (i) and (ii) will be used in the analysis of probability of encoding error while the event (iii) will be used in the analysis of probability of decoding error.

A.1.4.1 Probability of Error at the Encoder of $\mathcal{S}_2$

An error is made if (a) the encoder cannot find a $\mathbf{u}_1(l_1)$ in the bin indexed by k_1 such that the event E_{jl_1} occurs or (b) it cannot find a $\mathbf{u}_2(l_2)$ in the bin indexed by k_2 such that the event E_{jl_2} occurs. The probability of encoding error at $\mathcal{S}_2$ can be

bounded as

$$\begin{aligned} P_{e,\mathcal{S}_2} &\leq P\left(\bigcap_{\mathbf{U}_1(l_1)\in \mathbf{bin}(k_1)} (\mathbf{W}(j),\mathbf{U}_1(l_1),\mathbf{q}) \notin A_\epsilon^{(n)}\right) \\ &+ P\left(\bigcap_{\mathbf{U}_2(l_2)\in \mathbf{bin}(k_2)} (\mathbf{W}(j),\mathbf{U}_2(l_2),\mathbf{q}) \notin A_\epsilon^{(n)}\right), \qquad \text{(A.1)} \\ &\leq (1-P(E_{jl_1}))^{2^{n(I(W;U_1|Q)+4\epsilon)}} + (1-P(E_{jl_2}))^{2^{n(I(W;U_2|Q)+4\epsilon)}}, \end{aligned}$$

where $P(.)$ is the probability of an event. Since $\mathbf{q}$ is predetermined, and $\mathbf{w}$ and $\mathbf{u}_1$ are independent given $\mathbf{q}$,

$$\begin{aligned} P(E_{jl_1}) &= \sum_{(\mathbf{w},\mathbf{u}_1,\mathbf{q})\in A_\epsilon^{(n)}} P(\mathbf{W}(j)=\mathbf{w}|\mathbf{q})P(\mathbf{U}_1(l_1)=\mathbf{u}_1|\mathbf{q}) \\ &\geq 2^{n(H(W,U_1|Q)-\epsilon)}2^{-n(H(W|Q)+\epsilon)}2^{-n(H(U_1|Q)+\epsilon)} = 2^{-n(I(W;U_1|Q)+3\epsilon)}. \end{aligned}$$

Similarly, $P(E_{jl_2}) \geq 2^{-n(I(W;U_2|Q)+3\epsilon)}$. Therefore,

$$P_{e,\mathcal{S}_2} \leq (1-2^{-n(I(W;U_1|Q)+3\epsilon)})^{2^{n(I(W;U_1|Q)+4\epsilon)}} + (1-2^{-n(I(W;U_2|Q)+3\epsilon)})^{2^{n(I(W;U_2|Q)+4\epsilon)}}.$$

Now,

$$\begin{aligned} (1-2^{-n(I(W;U_1|Q)+3\epsilon)})^{2^{n(I(W;U_1|Q)+4\epsilon)}} &= e^{2^{n(I(W;U_1|Q)+4\epsilon)}\ln(1-2^{-n(I(W;U_1|Q)+3\epsilon)})} \\ &\leq e^{2^{n(I(W;U_1|Q)+4\epsilon)}(-2^{-n(I(W;U_1|Q)+3\epsilon)})} = e^{-2^{n\epsilon}}. \end{aligned}$$

Clearly, $P_{e,\mathcal{S}_2} \to 0$ as $n \to \infty$.

A.1.4.2 Probability of Error at the Decoder of $\mathcal{R}_1$

There are two possible events which result in errors: (a)The codewords transmitted are not jointly typical i.e., $E^c_{jl_1t_1}$ happens or (b) there exists some $\hat{j} \neq j$ such that $E_{\hat{j}\hat{l}_1\hat{t}_1}$ happens. Note that $\hat{l}_1$ need not equal l_1, and $\hat{t}_1$ need not equal t_1, since $\mathcal{R}_1$ is not required to decode $\hat{l}_1$ and $\hat{t}_1$ correctly. The probability of decoding error

can, therefore, be expressed as

$$P_{e,\mathcal{R}_1}^{(n)} = P\left(E_{jl_1t_1}^c \bigcup \cup_{\hat{j}\neq j} E_{\hat{j}\hat{l}_1\hat{t}_1}\right) \tag{A.2}$$

Applying union of events bound, $(A.2)$ can be written as,

$$P_{e,\mathcal{R}_1}^{(n)} \leq P\left(E_{jl_1t_1}^c\right) + P\left(\cup_{\hat{j}\neq j} E_{\hat{j}\hat{l}_1\hat{t}_1}\right)$$

$$= P\left(E_{jl_1t_1}^c\right) + \sum_{\hat{j}\neq j} P\left(E_{\hat{j}l_1t_1}\right) + \sum_{\hat{j}\neq j,\hat{l}_1\neq l_1} P\left(E_{\hat{j}\hat{l}_1t_1}\right) + \sum_{\hat{j}\neq j,\hat{t}_1\neq t_1} P\left(E_{\hat{j}l_1\hat{t}_1}\right) + \sum_{\hat{j}\neq j,\hat{l}_1\neq l_1,\hat{t}_1\neq t_1} P\left(E_{\hat{j}\hat{l}_1\hat{t}_1}\right).$$

$P\left(E_{\hat{j}l_1t_1}\right), P\left(E_{\hat{j}\hat{l}_1t_1}\right), P\left(E_{\hat{j}l_1\hat{t}_1}\right)$ and $P\left(E_{\hat{j}\hat{l}_1\hat{t}_1}\right)$ can be upper bounded as follows.

$$\begin{aligned}
P\left(E_{\hat{j}l_1t_1}\right) &\leq 2^{-n(I(W;U_1,V_1,Y_1|Q)-3\epsilon)},\\
P\left(E_{\hat{j}\hat{l}_1t_1}\right) &\leq 2^{-n(I(W,U_1;V_1,Y_1|Q)+I(W;U_1|Q)-4\epsilon)},\\
P\left(E_{\hat{j}l_1\hat{t}_1}\right) &\leq 2^{-n(I(W,V_1;U_1,Y_1|Q)+I(W;V_1|Q)-4\epsilon)},\\
P\left(E_{\hat{j}\hat{l}_1\hat{t}_1}\right) &\leq 2^{-n(I(W,U_1,V_1;Y_1|Q)+I(W,U_1;V_1|Q)+I(W;U_1|Q)-5\epsilon)}.
\end{aligned}$$

Substituting these in the probability of decoding error at $\mathcal{R}_1$, we have,

$$\begin{aligned}
P_{e,\mathcal{R}_1}^{(n)} = \epsilon + 2^{nR_{11}}2^{-n(I(W;U_1,V_1,Y_1|Q)-3\epsilon)}\\
+2^{n(R_{11}+R_{21}+I(W;U_1|Q)+4\epsilon)}2^{-n(I(W,U_1;V_1,Y_1|Q)+I(W;U_1|Q)-4\epsilon)} +\\
2^{n(R_{11}+R_{31}+I(W,U_1,U_2;V_1|Q)+4\epsilon)}2^{-n(I(W,V_1;U_1,Y_1|Q)+I(W;V_1|Q)-4\epsilon)} +\\
2^{n(R_{11}+R_{21}+I(W;U_1|Q)+4\epsilon+R_{31}+I(W,U_1,U_2;V_1|Q)+4\epsilon)}2^{-n(I(W,U_1,V_1;Y_1|Q)+I(W,U_1;V_1|Q)+I(W;U_1|Q)-5\epsilon)}.
\end{aligned}$$

$P_{e,\mathcal{R}_1}^{(n)} \to 0$ as $n \to \infty$ if R_{11}, R_{21} and R_{31} satisfy the following constraints:

$$\begin{aligned}
R_{11} &\leq I(W;U_1,V_1,Y_1|Q)\\
R_{11}+R_{21} &\leq I(W,U_1;V_1,Y_1|Q)\\
R_{11}+R_{31} &\leq I(W,V_1;U_1,Y_1|Q) + I(W;V_1|Q) - I(W,U_1,U_2;V_1|Q)
\end{aligned}$$

$$R_{11} + R_{21} + R_{31} \leq I(W, U_1, V_1; Y_1|Q) + I(W, U_1; V_1|Q) - I(W, U_1, U_2; V_1|Q).$$

The analysis of error-probabilities for the sender $\mathcal{S}_3$, and the receivers $\mathcal{R}_2$ and $\mathcal{R}_3$ can be found in [16]. Therefore, we conclude that, the probability of error terms can be made arbitrarily small, if the rate inequalities (2.6)-(2.15) are simultaneously satisfied. This concludes the proof of the achievability of the rate region for $\mathcal{C}^2_{\text{CuMS}}$.

A.2 Proof of Corollary 2.3.1

In the case of $\mathcal{C}^t_{G,\text{CuMS}}$, when senders $\mathcal{S}_2$ and $\mathcal{S}_3$ do not have any message of their own to transmit, they can use their noncausal message knowledge to entirely help sender $\mathcal{S}_1$. The rate tuple $(R_1^*, 0, 0)$ is therefore achievable, where R_1^* is the capacity of the vector channel $(\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3) \to \mathcal{R}_1$, given by

$$R_1^* = \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_1} + |a_{12}|\sqrt{P_2} + |a_{13}|\sqrt{P_3}\right)^2}{Q_1}\right). \tag{A.3}$$

Next, when the rate achieved by sender $\mathcal{S}_1$ is zero, $\mathcal{S}_2$ can cancel the interference from $\mathcal{S}_1$ completely by employing dirty-paper coding. However, due to the the message splitting model assumed here, $\mathcal{R}_2$ sees interference from $\mathcal{S}_3$ regardless of the R_3 achieved, except in the case where $\mathcal{S}_3$ helps $\mathcal{R}_2$ in receiving its message. This case is dealt with in Corollary 2.3.2. Hence, the rate achievable by $(\mathcal{S}_2, \mathcal{R}_2)$ is

$$R_2^* = \frac{1}{2}\log_2\left(1 + \frac{P_2}{Q_2 + |a_{23}|^2 P_3}\right). \tag{A.4}$$

When $R_1 = 0$ and $R_2 = R_2^*$, due to the noncausal knowledge of $\mathcal{S}_1$ and $\mathcal{S}_2$'s messages, $\mathcal{S}_3$ can completely mitigate the effect of interference and achieve the interference free rate, R_3^*, given by

$$R_3^* = \frac{1}{2}\log_2\left(1 + \frac{P_3}{Q_3}\right). \tag{A.5}$$

Hence, the rate tuple $(0, R_2^*, R_3^*)$ is achievable. Finally, the convex hull of the rate region $\mathfrak{G}^2_{\text{CuMS}}$ with these points is achievable by standard time-sharing arguments.

Appendix B

Proofs for the relay channel

B.1 Proof of Theorem 3.4.1

Here, we provide the proof of Theorem 3.4.1. Let $n = 1, \ldots, N$; $j = 1, \ldots, 2^{NR'_{13}}$; and $k = 1, \ldots, 2^{NR'_{23}}$. Generate $2^{NR_{13}}$ $N-$sequences $\mathbf{w}(m'_{13})$ with the n^{th} symbol of every sequence picked i.i.d from the distribution $P(w'_n)$. For every sequence $\mathbf{w}(m'_{13})$, generate

1. $2^{NR_{13}}$ $N-$sequences $\mathbf{u}(m'_{13}, m_{13})$ with the n^{th} symbol of every sequence picked i.i.d from $P(u_n|w_n)$.

2. $2^{NR'_{13}}$ $N-$sequences $\mathbf{v}(m'_{13}, j)$ with the n^{th} symbol of every sequence picked i.i.d from $P(v_n|w_n)$. This resembles $2^{NR_{13}}$ bins, each comprising $2^{NR'_{13}}$ sequences. These bins are indexed by m'_{13}.

Also, generate $2^{N[R_{23}+R'_{23}]}$ $N-$sequences $\mathbf{z}(m_{23}, k)$ with the n^{th} symbol of every sequence picked i.i.d from $P(z_n)$. Uniformly distribute these $2^{N[R_{23}+R'_{23}]}$ sequences into $2^{NR_{23}}$ bins, so that each bin comprises $2^{NR'_{23}}$ sequences. These bins are indexed by m_{23}. For every pair of sequences $(\mathbf{w}, \mathbf{u})$, generate one $N-$ sequence $\mathbf{x}_1(m'_{13}, m_{13})$ with the n^{th} symbol picked i.i.d from $P(x_{1,n}|w_n, u_n)$. For every triplet of sequences $(\mathbf{w}, \mathbf{v}, \mathbf{z})$, generate one $N-$ sequence $\mathbf{x}_2(m'_{13}, j, m_{23}, k)$ with the n^{th} symbol picked i.i.d from $P(x_{2,n}|w_n, v_n, z_n)$.

For notational convenience, the messages are denoted $m'_{13,b}$, $m_{13,b}$ and $m_{23,b}$, and are transmitted in B blocks each with N channel uses; $b = 1, \ldots, B$ is the block index. First, we define the following events:

(a) $E_{2,1} \triangleq \{(\mathbf{w}(1), \mathbf{u}(1, \bar{m}_{13,1}), \mathbf{v}(1, j_1), \mathbf{z}(m_{23,1}, k_1), \mathbf{y}_{2,1}) \in A_\epsilon^{(N)}(P_{W,U,V,Z,Y_2})\}$,

(b) $E_{2,b} \triangleq \{(\mathbf{w}(\bar{m}'_{13,b}), \mathbf{u}(\bar{m}'_{13,b}, \bar{m}_{13,b}), \mathbf{v}(\bar{m}'_{13,b}, j_b), \mathbf{z}(m_{23,b}, k_b), \mathbf{y}_{2,b}) \in A_\epsilon^{(N)}(P_{W,U,V,Z,Y_2})\}$.

1. In block 1,

   (a) Node 1 transmits $\mathbf{x}_{1,1} = \mathbf{x}_1(1, m_{13,1})$,

   (b) Node 2 transmits $\mathbf{x}_{2,1} = \mathbf{x}_2(1, j_1, m_{23,1}, k_1)$.

   In block 1, Node 2 has no information necessary for cooperation. Therefore, its transmits $\mathbf{x}_{2,1}$. Without loss of generality, one can assume such a protocol to exist between Node 2 and Node 3. The resulting loss in rate is negligible as $B \to \infty$. At the end of block 1, Node 2 chooses $\bar{m}_{13,1}$ such that the joint typicality condition $E_{2,1}$ is satisfied. This information will be used by Node 2 to transmit $\mathbf{x}_2(\bar{m}'_{13,2}, j_2, m_{23,2}, k_2)$ in block 2.

2. In block $b = 2, \ldots, B-1$,

   (a) Node 1 transmits $\mathbf{x}_{1,b} = \mathbf{x}_1(m'_{13,b}, m_{13,b})$,

   (b) Node 2 transmits $\mathbf{x}_{2,b} = \mathbf{x}_2(\bar{m}'_{13,b}, j_b, m_{23,b}, k_b)$.

   At the end of block b, Node 2 chooses $\bar{m}_{13,b}$ such that the joint typicality condition $E_{2,b}$ is satisfied. This information will be used by Node 2 to transmit $\mathbf{x}_{2,b+1}$ in block $b+1$. Here, $\bar{m}'_{13,b} = \bar{m}_{13,b-1}$ is obtained by Node 2 in block $b-1$.

3. In block B,

   (a) Node 1 transmits $\mathbf{x}_{1,B} = \mathbf{x}_1(m'_{13,B}, 1)$,

   (b) Node 2 transmits $\mathbf{x}_{2,B} = \mathbf{x}_2(\bar{m}'_{13,B}, j_B, 1, k_B)$.

   Here, $\bar{m}'_{13,B} = \bar{m}_{13,B-1}$ is obtained by Node 2 in block $B-1$.

The encoding and decoding (described later in this Appendix) operations are done on a "block-by-block" basis. In each block, there are n channel uses; on the current channel use, the relay will output a codeword which is a function of the previous $n-1$ symbols of the source and its own message intended to the destination.

We describe now the transmission procedure adopted by Node 2. In block $b+1$, to transmit the message pair $(\bar{m}_{13,b}, m_{23,b+1})$, Node 2 employs a stochastic encoder:

1. *Randomly* choose a sequence $\mathbf{v}(\bar{m}'_{13,b+1}, j_{b+1})$ in the bin indexed $\bar{m}'_{13,b+1}$.
2. In the bin indexed by $m_{23,b+1}$, *randomly* choose a sequence $\mathbf{z}(m_{23,b+1}, k_{b+1})$.

Corresponding to the triplet $(\mathbf{w}, \mathbf{v}, \mathbf{z})$, Node 2 transmits $\mathbf{x}_2(\bar{m}'_{13,b+1}, j_{b+1}, m_{23,b+1}, k_{b+1})$.

We employ a combination of backward and simultaneous decoding techniques to recover the transmitted information, where the decoders accumulate B blocks of data and start decoding from the last block. Before proceeding, we define the following events:

(a) $E_{3,B} \triangleq \{\Big(\mathbf{w}(\hat{m}'_{13,B}), \mathbf{u}(\hat{m}'_{13,B}, 1), \mathbf{v}(\hat{m}'_{13,B}, j_B), \mathbf{z}(m'_{13,b}, 1, k_b), \mathbf{y}_{3,B}\Big) \in A_\epsilon^{(N)}(P_{W,U,V,Z,Y_3})\}$,

(b) $E_{3,b} \triangleq \{\Big(\mathbf{w}(\hat{m}'_{13,b}), \mathbf{u}(\hat{m}'_{13,b}, \hat{m}_{13,b}), \mathbf{v}(\hat{m}'_{13,b}, j_b), \mathbf{z}(m'_{13,b}, \hat{m}_{23,b}, k_b), \mathbf{y}_{3,b}\Big)$
$\in A_\epsilon^{(N)}(P_{W,U,V,Z,Y_3})\}$,

(c) $E_{3,1} \triangleq \{(\mathbf{w}(1), \mathbf{u}(1, m_{13,1}), \mathbf{v}(1, j_1), \mathbf{z}(1, \hat{m}_{23,1}, k_1), \mathbf{y}_{3,1}) \in A_\epsilon^{(N)}(P_{W,U,V,Z,Y_3})\}$.

1. In block B, the decoder of Node 3 looks for the pair $(\hat{m}'_{13,B}, 1)$ that satisfies the joint typicality condition $E_{3,B}$.
2. In block $b = 2, \ldots, B-1$, the decoder of Node 3 first sets $\hat{m}_{13,b} = \hat{m}'_{13,b+1}$ and looks for $(\hat{m}'_{13,b}, \hat{m}_{23,b})$ that satisfies the joint typicality condition $E_{3,b}$.
3. In block 1, the decoder of Node 3 sets $m_{13,1} = \hat{m}'_{13,2}$ and looks for $\hat{m}_{23,1}$ such that $E_{3,1}$ is satisfied.

The message m_{13} was coded at a higher rate, $R'_{13} + R_{13}$. However, during the decoding process, the receiver can only decode the low rate message m'_{13} which

was coded at a rate R'_{13}. Also, note that, there is no difference between m'_{13} and m_{13}, both denote the message indices of the same message set $\mathcal{M}_{13}$; the former is the message coded at a lower rate, while the latter will be coded at a higher rate using superposition principle. This is the essence of Markov superposition coding.

The average error probability at Node 3, averaged over all codes, is given by

$$P_{e,3}^{(N)} = \Pr\left\{\bigcup_b \bigcup_t (\hat{\mathbf{M}}_{t3,b} \neq \mathbf{M}_{t3,b})\right\}; b = 1, \ldots, B; t = 1, 2.$$

We now derive an upper bound for $P_{e,3}^{(N)}$. For notational convenience, we drop the block index b and without loss of generality assume that $(m'_{13}, m_{13}, m_{23}) = (1, 1, 1)$ was transmitted. By the union of events bound, we have

$$P_{e,3}^{(N)} \leq \sum_{b=2}^{B} \Pr(E_3^c) + \sum_{b=2}^{B} \sum_{(\hat{m}'_{13}, \hat{m}_{23}) \neq (1,1)} \sum_{j,k} \Pr(E_3) + \sum_{b=1}^{B-1} \Pr(E_2^c) + \sum_{b=1}^{B-1} \sum_{\bar{m}_{13} \neq 1} \Pr(E_2).$$

From the asymptotic equipartition property [52], $\Pr(E_3^c) \leq \epsilon$ and $\Pr(E_2^c) \leq \epsilon$, $\forall \epsilon > 0$ and sufficiently small for large N. Furthermore, from the properties of jointly typical sequences [52, Section 15.2, pp. 520-524], the probabilities of the individual error events can be upper bounded as follows:

$$\begin{aligned}
\Pr(E_3) &\leq 2^{-N[I(W,U,V;Y_3|Z)-3\epsilon]}, \text{ when } \hat{m}'_{13} \neq 1, \\
\Pr(E_3) &\leq 2^{-N[I(Z;Y_3|W,U,V)-3\epsilon]}, \text{ when } \hat{m}_{23} \neq 1, \\
\Pr(E_3) &\leq 2^{-N[I(W,U,V,Z;Y_3)-4\epsilon]}, \text{ when } (\hat{m}'_{13}, \hat{m}_{23}) \neq (1,1), \\
\Pr(E_2) &\leq 2^{-N[I(U;Y_2|W,V,Z)-4\epsilon]}, \text{ when } \bar{m}_{13} \neq 1.
\end{aligned}$$

Therefore, we have,

$$\begin{aligned}
P_{e,3}^{(N)} \leq\ & 3(B-1)\epsilon + (B-1)2^{N(R_{13}+R'_{13})}2^{-N[I(W,U,V;Y_3|Z)-3\epsilon]} \\
& + (B-1)2^{N(R_{23}+R'_{23})}2^{-N[I(Z;Y_3|W,U,V)-3\epsilon]} \\
& + (B-1)2^{N(R_{13}+R'_{13}+R_{23}+R'_{23})}2^{-N[I(W,U,V,Z;Y_3)-4\epsilon]} + (B-1)2^{NR_{13}}2^{-N[I(U;Y_2|W,V,Z)-4\epsilon]}.
\end{aligned}$$

Finally, $P_{e,3}^{(N)}$ can be made arbitrarily small if, $\forall \epsilon > 0$ and sufficiently small for $N \to \infty$, the following inequalities are simultaneously satisfied:

$$\begin{aligned} R_{13} + R'_{13} &\leq I(W,U,V;Y_3|Z), && \text{(B.1)} \\ R_{23} + R'_{23} &\leq I(Z;Y_3|W,U,V), && \text{(B.2)} \\ R_{13} + R'_{13} + R_{23} + R'_{23} &\leq I(W,U,V,Z;Y_3), && \text{(B.3)} \\ R_{13} &\leq I(U;Y_2|W,V,Z). && \text{(B.4)} \end{aligned}$$

An error is declared at the encoder of Node 2, if (i) it is not able to find a typical sequence $\mathbf{v}(m'_{13}, j) \in A_\epsilon^{(N)}(P_{V|W})$ or (ii) it is not able to find a typical sequence $\mathbf{z}(m_{23}, k) \in A_\epsilon^{(N)}(P_Z)$. Let us define the event $E_1 \triangleq \{\mathbf{v} \in A_\epsilon^{(N)}(P_{V|W})\}$. The probability of encoder error at Node 2 can be bounded as follows:

$$P_{e,2}^{(N)} \leq P\left(\bigcap_j \mathbf{v} \notin A_\epsilon^{(N)}(P_{V|W})\right) \leq (1 - P(E_1))^{2^{N(R'_{13}+4\epsilon)}}.$$

$$\begin{aligned} P(E_1) &= \sum_{A_\epsilon^{(N)}(P_{V|W})} P(\mathbf{W} = \mathbf{w}) P(\mathbf{V} = \mathbf{v}|\mathbf{W} = \mathbf{w}) \\ &\geq 2^{N(H(V|W)-\epsilon)} 2^{-N(H(W)+\epsilon)} 2^{-N(H(V|W)+\epsilon)} = 2^{-N[H(W)+3\epsilon]}. \end{aligned}$$

Therefore,

$$\begin{aligned} P_{e,2}^{(N)} &\leq \left(1 - 2^{-N[H(W)+3\epsilon]}\right)^{2^{N[R'_{13}+4\epsilon]}} \\ &= e^{2^{N[R'_{13}+4\epsilon]} \ln(1-2^{-N[H(W)+3\epsilon]})} \\ &\leq e^{2^{N[R'_{13}+4\epsilon]}(-2^{-N[H(W)+3\epsilon]})} = e^{-2^{N[R'_{13}-H(W)]\epsilon}}. \end{aligned}$$

Clearly, $P_{e,2}^{(N)} \to 0$, $\forall \epsilon > 0$ and sufficiently small for $N \to \infty$ iff $R'_{13} > H(W)$.

An error is also declared at the encoder of Node 2 if it is not able to find a typical sequence $\mathbf{z}(m_{23}, k) \in A_\epsilon^{(N)}(P_Z)$. Let us define the event $E_2 \triangleq \{\mathbf{z} \in A_\epsilon^{(N)}(P_Z)\}$.

The probability of encoder error at Node 2 can be bounded as follows:

$$P_{e,2}^{(N)} \leq P\left(\bigcap_k \mathbf{z} \notin A_\epsilon^{(N)}(P_Z)\right) \leq (1 - P(E_2))^{2^{N(R'_{23}+4\epsilon)}}.$$

$$P(E_2) = \sum_{A_\epsilon^{(N)}(P_Z)} P(\mathbf{Z} = \mathbf{z}) \geq 2^{N(H(Z)-\epsilon)} 2^{-N(H(W)+\epsilon)} = 2^{-2N\epsilon}.$$

Clearly, $P_{e,2}^{(N)} \to 0$, $\forall \epsilon > 0$ and sufficiently small for $N \to \infty$.

We compute now the equivocation at Node 4's receiver and show that the code satisfies confidentiality constraints. In each block $b = 1, \ldots, B$, consider the following equivocation lower bound:

$$\begin{aligned}
H(M_{13}|Y_4^N) &= H(M_{13}, W^N, V^N, Y_4^N) - H(W^N, V^N|M_{13}, Y_4^N) - H(Y_4^N) \\
&= H(M_{13}, W^N, V^N) + H(Y_4^N|M_{13}, W^N, V^N) - H(W^N, V^N|M_{13}, Y_4^N) \\
&- H(Y_4^N) \\
&= H(M_{13}, W^N, V^N) - H(W^N, V^N|M_{13}, Y_4^N) - I(M_{13}, W^N, V^N; Y_4^N) \\
&\overset{(a)}{=} H(M_{13}, W^N, V^N) - H(W^N, V^N|M_{13}, Y_4^N) - I(W^N, V^N; Y_4^N), \\
&\geq H(W^N, V^N) - H(W^N, V^N|M_{13}, Y_4^N) - I(W^N, V^N; Y_4^N),
\end{aligned}$$

where (a) follows from the Markov chain $M_{13} \to W^N \to V^N \to Y_4^N$, with $I(M_{13}; Y_4^N|W^N, V^N) = 0$. Let us consider each term separately.

1. $H(W^N, V^N) = N(R_{13} + R'_{13})$ (see codebook construction),
2. $H(W^N, V^N|M_{13}, Y_4^N) \leq N\epsilon_1$ (see Appendix B.5),
3. $I(W^N, V^N; Y_4^N) \leq NI(W, V; Y_2)$ (using standard techniques).

Therefore, $H(M_{13}|Y_4^N) \geq NR_{13} + NR'_{13} - N\epsilon - NI(W, V; Y_2)$. Let $R'_{13} = I(W, V; Y_2) - \epsilon_2$, $\forall \epsilon_2 > 0$ and sufficiently small for large N. Therefore, we have $H(M_{13}|Y_4^N) \geq NR_{13} - N\epsilon$ and the secrecy constraint (3.1) is satisfied.

Also consider the following:

$$\begin{aligned} H(M_{23}|Y_4^N) &= H(M_{23}, Z^N, Y_4^N) - H(Z^N|M_{23}, Y_4^N) - H(Y_4^N) \\ &= H(M_{23}, Z^N) + H(Y_4^N|M_{23}, Z^N) - H(Z^N|M_{23}, Y_4^N) - H(Y_4^N) \\ &= H(M_{23}, Z^N) - H(Z^N|M_{23}, Y_4^N) - I(M_{23}, Z^N; Y_4^N) \\ &\stackrel{(a)}{=} H(M_{23}, Z^N) - H(Z^N|M_{23}, Y_4^N) - I(Z^N; Y_4^N), \\ &\geq H(Z^N) - H(Z^N|M_{23}, Y_4^N) - I(Z^N; Y_4^N), \end{aligned}$$

where (a) follows from the Markov chain $M_{23} \to Z^N \to Y_4^N$, with $I(M_{23}; Y_4^N|Z^N) = 0$. Let us consider each term separately.

1. $H(Z^N) = N(R_{13} + R_{13}^*)$ (see codebook construction),

2. $H(Z^N|M_{23}, Y_4^N) \leq N\epsilon_1$ (see Appendix B.5),

3. $I(Z^N; Y_4^N) \leq NI(Z; Y_4)$ (using standard techniques).

Therefore, $H(M_{23}|Y_4^N) \geq NR_{23} + NR'_{23} - N\epsilon - NI(Z; Y_4)$. Let $R'_{23} = I(Z; Y_4) - \epsilon_3$, $\forall \epsilon_3 > 0$ and sufficiently small for large N. Therefore, we have $H(M_{23}|Y_4^N) \geq NR_{23} - N\epsilon$ and the secrecy constraint (3.2) is satisfied.

Finally, an achievable rate region for $\mathcal{C}$, given by (3.5) - (3.7), is obtained by substituting for R'_{13} and R'_{23} in (B.1) - (B.3).

B.2 Proof of Theorem 3.4.2

We prove now Theorem 3.4.2. For any sequence of $\left(\left(2^{NR_{13}}, 2^{NR_{23}}\right), N\right)$ codes such that $P_{e,3}^{(N)} \to 0$ and $P_{e,2}^{(N)} \to 0$ for $N \to \infty$, the probability mass function on the joint ensemble space $M_{13} \times M_{23} \times \mathcal{X}_1^N \times \mathcal{X}_2^N \times \mathcal{Y}_2^N \times \mathcal{Y}_3^N \times \mathcal{Y}_4^N$ is given by

$$\begin{aligned} p(m_{13}, m_{23}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4) = p(m_{13})p(m_{23})p(\mathbf{x}_1|m_{13}) \prod_{n=1}^{N} p(x_{2,n}|m_{23}, y_2^{n-1}) \\ \times p(y_{2,n}|x_{1,n}, x_{2,n})p(y_{3,n}, y_{4,n}|y_{2,n}, x_{2,n}). \end{aligned}$$

For reliable communications, we have from Fano's inequality [52],

$$H(M_{13}|Y_2^N) \leq NR_{13}P_{e,2}^{(N)} + 1 = N\delta_{2,N}, \tag{B.5}$$

$$H(M_{13}, M_{23}|Y_3^N) \leq N(R_{13} + R_{23})P_{e,3}^{(N)} + 1 = N\delta_{3,N}, \tag{B.6}$$

where $\delta_{2,N} > 0$, $\delta_{3,N} > 0$ and sufficiently small for large N. We also use the following: $M_{13} = V_1 = \cdots = V_N$, $M_{23} = Z_1 = \cdots = Z_N$, $Y_3^{n-1} = W_n$ and $Y_2^{n-1} = U_n$. R_{13} can be upper bounded as follows. Consider the following bound on the equivocation obtained from the security constraint (3.1) and Fano's inequality (B.6):

$$\begin{aligned}
NR_{13} &= H(M_{13}) \\
&\leq H(M_{13}|Y_4^N) + N\epsilon \\
&\leq H(M_{13}|Y_4^N) - H(M_{13}|Y_3^N) + N\delta_{3,N} \\
&= H(M_{13}|Y_{4,1}) - H(M_{13}|Y_{3,N}) \\
&\quad + \sum_{n=2}^{N}[H(M_{13}|Y_{4,n})] - \sum_{n=1}^{N-1}[H(M_{13}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{13}|Y_{4,n})] - \sum_{n=1}^{N}[H(M_{13}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{13}; Y_{3,n})] - \sum_{n=1}^{N}[I(M_{13}; Y_{4,n})] + N\delta_{3,N} \\
&\overset{(a)}{\leq} \sum_{n=1}^{N}[H(M_{13}|M_{23}) - H(M_{13}|M_{23}, Y_{3,n})] - \sum_{n=1}^{N}[I(M_{13}; Y_{4,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{13}; Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13}; Y_{4,n})] + N\delta_{3,N} \\
&\overset{(b)}{\leq} \sum_{n=1}^{N}[H(Y_{3,n}|M_{23}) - H(Y_{3,n}|Y_2^{n-1}, Y_3^{n-1}, M_{23}, M_{13})] - \sum_{n=1}^{N}[I(M_{13}; Y_{4,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(Y_2^{n-1}, Y_3^{n-1}, M_{13}; Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13}; Y_{4,n})] + N\delta_{3,N},
\end{aligned}$$

where (a) and (b) follow from the fact that conditioning reduces entropy. There-

fore,

$$NR_{13} \leq \sum_{n=1}^{N} I(W_n, U_n, V_n; Y_{3,n}|Z_n) - I(V_n; Y_{4,n}) + N\delta_{3,N}. \tag{B.7}$$

R_{23} can be upper bounded as follows. Consider the following bound on the equivocation obtained from the security constraint (3.2) and Fano's inequality (B.6):

$$\begin{aligned}
NR_{23} &= H(M_{23}) \\
&\leq H(M_{23}|Y_4^N) + N\epsilon \\
&\leq H(M_{23}|Y_4^N) - H(M_{23}|Y_3^N) + N\delta_{3,N} \\
&= H(M_{23}|Y_{4,1}) - H(M_{23}|Y_{3,N}) \\
&\quad + \sum_{n=2}^{N}[H(M_{23}|Y_{4,n})] - \sum_{n=1}^{N-1}[H(M_{23}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{23}|Y_{4,n})] - \sum_{n=1}^{N}[H(M_{23}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{23}; Y_{3,n})] - \sum_{n=1}^{N}[I(M_{23}; Y_{4,n})] + N\delta_{3,N} \\
&\overset{(a)}{\leq} \sum_{n=1}^{N}[H(M_{23}|M_{13}) - H(M_{23}|M_{13}, Y_{3,n})] - \sum_{n=1}^{N}[I(M_{23}; Y_{4,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{23}; Y_{3,n}|M_{13})] - \sum_{n=1}^{N}[I(M_{23}; Y_{4,n})] + N\delta_{3,N},
\end{aligned}$$

where (a) follows because of the fact that (i) M_{13} and M_{23} are independent and (ii) conditioning reduces entropy. Therefore,

$$NR_{23} \leq \sum_{n=1}^{N} I(Z_n; Y_{3,n}|V_n) - I(Z_n; Y_{4,n}) + N\delta_{3,N}. \tag{B.8}$$

$R_{13} + R_{23}$ can be upper bounded as follows. Consider the following bound on the equivocation obtained from the security constraint (3.1), (3.2) and Fano's

inequalities (B.5), (B.6):

$$
\begin{aligned}
N(R_{13}+R_{23}) &= H(M_{13},M_{23}) \\
&\leq H(M_{13},M_{23}|Y_4^N)+N\epsilon \\
&\leq H(M_{13},M_{23}|Y_4^N)-H(M_{13},M_{23}|Y_3^N)+N\delta_{3,N} \\
&= H(M_{13},M_{23}|Y_{4,1})-H(M_{13},M_{23}|Y_{3,N}) \\
&\quad +\sum_{n=2}^{N}[H(M_{13},M_{23}|Y_{4,n})]-\sum_{n=1}^{N-1}[H(M_{13},M_{23}|Y_{3,n})]+N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{13},M_{23}|Y_{4,n})]-\sum_{n=1}^{N}[H(M_{13},M_{23}|Y_{3,n})]+N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{13}|Y_{4,n})]+\sum_{n=1}^{N}[H(M_{23}|Y_{4,n})]-\sum_{n=1}^{N}[H(M_{13},M_{23}|Y_{3,n})]+N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{13},M_{23};Y_{3,n})]-\sum_{n=1}^{N}[I(M_{13};Y_{4,n})]-\sum_{n=1}^{N}[I(M_{23};Y_{4,n})]+N\delta_{3,N} \\
&\overset{(a)}{\leq} \sum_{n=1}^{N}[H(Y_{3,n})-H(Y_{3,n}|Y_2^{n-1},Y_3^{n-1},M_{13},M_{23})] \\
&\quad -\sum_{n=1}^{N}[I(M_{13};Y_{4,n})]-\sum_{n=1}^{N}[I(M_{23};Y_{4,n})]+N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(Y_2^{n-1},Y_3^{n-1},M_{13},M_{23};Y_{3,n})]-\sum_{n=1}^{N}[I(M_{13};Y_{4,n})] \\
&\quad -\sum_{n=1}^{N}[I(M_{23};Y_{4,n})]+N\delta_{3,N},
\end{aligned}
$$

where (a) follows from the fact that conditioning reduces entropy. Therefore,

$$
N(R_{13}+R_{23}) \leq \sum_{n=1}^{N} I(W_n,U_n,V_n,Z_n;Y_{3,n})-I(V_n;Y_{4,n})-I(Z_n;Y_{4,n})+N\delta_{3,N}. \quad \text{(B.9)}
$$

Finally, a time-sharing RV Q, which is uniformly distributed over N symbols and and independent of M_{13}, M_{23}, $\mathcal{X}_1^N$, $\mathcal{X}_2^N$, $\mathcal{Y}_2^N$,$\mathcal{Y}_3^N$,$\mathcal{Y}_4^N$, can be introduced for the single letter characterization of the above derived outer bounds. Applying the procedure similar to the one presented in [52, Chapter 15.3.4] on (B.7), (B.8) and (B.9), we get the outer bounds on R_{13} and R_{23} as given by (3.8) - (3.10),

respectively.

B.3 Proof of Theorem 4.3.4

Let $n = 1, \ldots, N$; $l = 1, \ldots, 2^{NR^*_{13}}$; and $k = 1, \ldots, 2^{NR'_{23}}$. Generate $2^{NR_{13}}$ $N-$sequences $\mathbf{w}(m'_{13})$ with the n^{th} symbol of every sequence picked i.i.d from the distribution $P(w'_n)$. For every sequence $\mathbf{w}(m'_{13})$, generate $2^{NR^*_{13}}$ $N-$sequences $\mathbf{u}(m'_{13}, l)$ with the n^{th} symbol of every sequence picked i.i.d from $P(u_n|w_n)$. This resembles $2^{NR_{13}}$ bins with each bin comprising $2^{NR^*_{13}}$ sequences. These bins are indexed by m'_{13}.

Generate $2^{N[R_{23}+R'_{23}]}$ independent $N-$sequences $\mathbf{z}(m_{23}, k)$ with the n^{th} symbol of every sequence picked i.i.d from $P(z_n)$. Uniformly distribute these $2^{N[R_{23}+R'_{23}]}$ sequences into $2^{NR_{23}}$ bins, so that each bin comprises $2^{NR'_{23}}$ sequences. These bins are indexed by m_{23}.

For every pair of sequences $(\mathbf{w}, \mathbf{u})$, generate one $N-$ sequence $\mathbf{x}_1(m'_{13}, l)$ with the n^{th} symbol picked i.i.d from $P(x_{1,n}|w_n, u_n)$. For every sequence $\mathbf{z}$, generate one $N-$ sequence $\mathbf{x}_2(m_{23}, k)$ with the n^{th} symbol picked i.i.d from $P(x_{2,n}|z_n)$.

For notational convenience, the messages are denoted $m'_{13,b}$ and $m_{23,b}$, and are transmitted in B blocks each with N channel uses; $b = 1, \ldots, B$ is the block index. Let us first define the following event: $E_{2,b} \triangleq \{(\mathbf{w}(\bar{m}'_{13,b}), \mathbf{u}(\bar{m}'_{13,b}, l_b), \mathbf{z}(m_{23,b}, k_b), \mathbf{y}_{2,b}) \in A^{(N)}_\epsilon(P_{W,U,Z,Y_2})\}$.

1. In block 1,
   (a) Node 1 transmits $\mathbf{x}_{1,1} = \mathbf{x}_1(1, l_1)$,
   (b) Node 2 transmits $\mathbf{x}_{2,1} = \mathbf{x}_2(m_{23,1}, 1)$.
2. In block $b = 2, \ldots, B-1$,
   (a) Node 1 transmits $\mathbf{x}_{1,b} = \mathbf{x}_1(m'_{13,b}, l_b)$, where $m'_{13,b} = l_{b-1}$,
   (b) Node 2 transmits $\mathbf{x}_{2,b} = \mathbf{x}_2(m_{23,b}, k_b)$.
3. In block B,
   (a) Node 1 transmits $\mathbf{x}_{1,B} = \mathbf{x}_1(m'_{13,B}, 1)$,

(b) Node 2 transmits $\mathbf{x}_{2,B} = \mathbf{x}_2(1, k_B)$.

At the end of block b, Node 2 tries to pick a $\bar{m}_{13,b}$ such that the joint typicality condition $E_{2,b}$ is satisfied. Here, $\bar{m}'_{13,b} = \bar{l}_{b-1}$ which Node 2 tries to obtain in block $b-1$. However, we will later show that this condition is indeed not satisfied. We describe now the transmission procedure adopted by Node 1 and Node 2. In block b, Node 1 employs a stochastic encoder by *randomly* picking a sequence $\mathbf{u}(m'_{13,b}, l_b)$. Corresponding to $\mathbf{u}$, Node 1 transmits $\mathbf{x}_1(m'_{13,b}, l_b)$. Likewise, to transmit the message $m_{23,b}$, Node 2 *randomly* chooses a sequence $\mathbf{z}(m_{23,b}, k_b)$ in the bin indexed by $m_{23,b}$. Corresponding to $\mathbf{z}$, Node 2 transmits $\mathbf{x}_2(m_{23,b}, k_b)$.

We employ a combination of backward and simultaneous decoding techniques to recover the transmitted information, where the decoders accumulate B blocks of data and start decoding from the last block. First, we define the following events:

(a) $E_{3,B} \triangleq \{\Big(\mathbf{w}(\hat{m}'_{13,B}), \mathbf{u}(\hat{m}'_{13,B}, 1), \mathbf{z}(1, k_B), \mathbf{y}_{3,B}\Big) \in A_\epsilon^{(N)}(P_{W,U,Z,Y_3})\}$,

(b) $E_{3,b} \triangleq \{\Big(\mathbf{w}(\hat{m}'_{13,b}), \mathbf{u}(\hat{m}'_{13,b}, \hat{l}_b), \mathbf{z}(\hat{m}_{23,b}, k_b), \mathbf{y}_{3,b}\Big) \in A_\epsilon^{(N)}(P_{W,U,Z,Y_3})\}$,

(c) $E_{3,1} \triangleq \{(\mathbf{w}(1), \mathbf{u}(1, m_{13,1}, l_1), \mathbf{z}(\hat{m}_{23,1}, k_1), \mathbf{y}_{3,1}) \in A_\epsilon^{(N)}(P_{W,U,Z,Y_3})\}$.

1. In block B, the decoder of Node 3 looks for a pair $(\hat{m}'_{13,B}, 1)$ that satisfies the joint typicality condition $E_{3,B}$.

2. In block $b = 2, \ldots, B-1$, the decoder of Node 3 first sets $\hat{l}_b = \hat{m}'_{13,b+1}$ and looks for a pair $(\hat{m}'_{13,b}, \hat{m}_{23,b})$ that satisfies the joint typicality condition $E_{3,b}$.

3. In block 1, the decoder of Node 3 sets $l_1 = \hat{m}'_{13,2}$ and looks for $\hat{m}_{23,1}$ such that $E_{3,1}$.

The average error probability at Node 3, averaged over all codes, is given by

$$P_{e,3}^{(N)} = \Pr\left\{\bigcup_b \bigcup_t (\hat{\mathbf{M}}_{t3,b} \neq \mathbf{M}_{t3,b})\right\}; b = 1, \ldots, B; t = 1, 2.$$

The procedure to derive an upper bound for $P_{e,3}^{(N)}$ is similar to that presented in Appendix B.1 and is omitted here for sake of brevity. Finally, $P_{e,3}^{(N)}$ can be made arbitrarily small if, $\forall \epsilon > 0$ and $N \to \infty$, the following inequalities are simultaneously satisfied:

$$R_{13} + R_{13}^* \leq I(W, U; Y_3|Z), \tag{B.10}$$

$$R_{23} + R_{23}' \leq I(Z; Y_3|W, U), \tag{B.11}$$

$$R_{13} + R_{13}^* + R_{23} + R_{23}' \leq I(W, U, Z; Y_3|). \tag{B.12}$$

An error is declared at the encoder of $\mathrm{Node}\ 1$, if it is not able to find a typical sequence $\mathbf{u}(m_{13}', l) \in A_\epsilon^{(N)}(P_{U|W})$. Let us define the event $E_1 \triangleq \{\mathbf{u} \in A_\epsilon^{(N)}(P_{U|W})\}$. The probability of encoder error at $\mathrm{Node}\ 1$, $P_{e,1}^{(N)} \to 0$, $\forall \epsilon > 0$ and sufficiently small for $N \to \infty$ iff $R_{13}^* > H(W)$ (for proof see Appendix B.1).

An error is declared at the encoder of $\mathrm{Node}\ 2$, if it is not able to find a typical sequence $\mathbf{z}(m_{23}, k) \in A_\epsilon^{(N)}(P_Z)$. Let us define the event $E_2 \triangleq \{\mathbf{z} \in A_\epsilon^{(N)}(P_Z)\}$. The probability of encoder error at $\mathrm{Node}\ 2$, $P_{e,2}^{(N)} \to 0$, $\forall \epsilon > 0$ and sufficiently small for $N \to \infty$ (for proof see Appendix B.1).

We compute now the equivocation at $\mathrm{Node}\ 2$'s receiver and show that the code satisfies confidentiality constraints. In each block $b = 1, \ldots, B$, consider the following equivocation lower bound:

$$\begin{aligned}
H(M_{13}|Y_2^N) &= H(M_{13}, W^N, U^N, Y_2^N) - H(W^N, U^N|M_{13}, Y_2^N) - H(Y_2^N) \\
&= H(M_{13}, W^N, U^N) + H(Y_2^N|M_{13}, W^N, U^N) - H(W^N, U^N|M_{13}, Y_2^N) \\
&- H(Y_2^N) \\
&= H(M_{13}, W^N, U^N) - H(W^N, U^N|M_{13}, Y_2^N) - I(M_{13}, W^N, U^N; Y_2^N) \\
&\overset{(a)}{=} H(M_{13}, W^N, U^N) - H(W^N, U^N|M_{13}, Y_2^N) - I(W^N, U^N; Y_2^N), \\
&\geq H(W^N, U^N) - H(W^N, U^N|M_{13}, Y_2^N) - I(W^N, U^N; Y_2^N),
\end{aligned}$$

where (a) follows from the Markov chain $M_{13} \to W^N \to U^N \to Y_2^N$, with

$I(M_{13}; Y_2^N | W^N, U^N) = 0$. Let us consider each term separately.

1. $H(W^N, U^N) = N(R_{13} + R_{13}^*)$ (see codebook construction),

2. $H(W^N, U^N | M_{13}, Y_2^N) \leq N\epsilon_1$ (see Appendix B.5),

3. $I(W^N, U^N; Y_2^N) \leq NI(W, U; Y_2)$ (using standard techniques).

Therefore, $H(M_{13}|Y_2^N) \geq NR_{13} + NR_{13}^* - N\epsilon - NI(W, U; Y_2)$. Let $R_{13}^* = I(W, U; Y_2) - \epsilon_4$, $\forall \epsilon_4 > 0$ and sufficiently small for large N. Therefore, we have $H(M_{13}|Y_2^N) \geq NR_{13} - N\epsilon$ and the secrecy constraint (3.3) is satisfied. The equivocation calculation at Node 4 leads to $H(M_{23}|Y_4^N) \geq NR_{23} - N\epsilon$ with $R_{13}^* = I(Z; Y_4) - \epsilon_5$ $\forall \epsilon_5 > 0$ and sufficiently small for large N. The proof is similar to that of channel $\mathcal{C}$ (see Appendix B.1) and is omitted. Thus, the secrecy constraint (3.4) is satisfied.

Finally, an achievable rate region for $\mathcal{C}^*$, given by (3.11) - (3.13), is obtained by substituting for R_{13}^* and R_{23}' in (B.10) - (B.12).

B.4 Proof of Theorem 4.3.5

We present now the proof of Theorem 4.3.5. For any sequence of $\left(\left(2^{NR_{13}}, 2^{NR_{23}}\right), N\right)$ codes such that $P_{e,3}^{(N)} \to 0$ and $P_{e,2}^{(N)} \to 0$ for $N \to \infty$, the probability mass function on the joint ensemble space $M_{13} \times M_{23} \times \mathcal{X}_1^N \times \mathcal{X}_2^N \times \mathcal{Y}_2^N \times \mathcal{Y}_3^N \times \mathcal{Y}_4^N$ is given by

$$p(m_{13}, m_{23}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4) = p(m_{13})p(m_{23})p(\mathbf{x}_1|m_{13}) \prod_{n=1}^{N} p(x_{2,n}|m_{23}) \times p(y_{2,n}|x_{1,n})p(y_{3,n}, y_{4,n}|x_{1,n}, x_{2,n}).$$

For reliable communications, we have from Fano's inequality [52],

$$H(M_{13}|Y_3^N) \leq NR_{13}P_{e,3}^{(N)} + 1 = N\delta_{3,N}, \tag{B.13}$$

$$H(M_{23}|Y_3^N) \leq NR_{23}P_{e,3}^{(N)} + 1 = N\delta_{3,N}, \tag{B.14}$$

where $\delta_{3,N} > 0$ and sufficiently small for large N. We also use the following to derive the upper bounds: $M_{23} = U_1 = \cdots = U_N$, $M_{23} = Z_1 = \cdots = Z_N$ and $Y_3^{n-1} =$

W_n.

R_{13} can be upper bounded as follows. Consider the following bound on the equivocation obtained from the security constraint (3.3) and Fano's inequality (B.13):

$$\begin{aligned}
NR_{13} &= H(M_{13}) \\
&\leq H(M_{13}|Y_2^N) + N\epsilon \\
&\leq H(M_{13}|Y_2^N) - H(M_{13}|Y_3^N) + N\delta_{3,N} \\
&= H(M_{13}|Y_{2,1}) - H(M_{13}|Y_{3,N}) \\
&\quad + \sum_{n=2}^{N}[H(M_{13}|Y_{2,n})] - \sum_{n=1}^{N-1}[H(M_{13}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{13}|Y_{2,n})] - \sum_{n=1}^{N}[H(M_{13}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{13};Y_{3,n})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] + N\delta_{3,N} \\
&\overset{(a)}{\leq} \sum_{n=1}^{N}[H(M_{13}|M_{23}) - H(M_{13}|M_{23},Y_{3,n})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{13};Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] + N\delta_{3,N} \\
&\overset{(b)}{\leq} \sum_{n=1}^{N}[H(Y_{3,n}|M_{23}) - H(Y_{3,n}|Y_3^{n-1},M_{23},M_{13})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(Y_3^{n-1},M_{13};Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] + N\delta_{3,N},
\end{aligned}$$

where (a) follows from the fact that M_{13} and M_{23} are independent, while (b) follows from the fact that conditioning reduces entropy. Therefore,

$$NR_{13} \leq \sum_{n=1}^{N} I(W_n,U_n;Y_{3,n}|Z_n) - I(U_n;Y_{2,n}) + N\delta_{3,N}. \tag{B.15}$$

R_{23} can be bounded as follows:

$$NR_{23} = H(M_{23})$$

$$
\begin{aligned}
&\leq H(M_{23}|Y_4^N) + N\epsilon \\
&\leq H(M_{23}|Y_4^N) - H(M_{23}|Y_3^N) + N\delta_{3,N} \\
&= H(M_{23}|Y_{4,1}) - H(M_{23}|Y_{3,N}) \\
&\quad + \sum_{n=2}^{N}[H(M_{23}|Y_{4,n})] - \sum_{n=1}^{N-1}[H(M_{23}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[H(M_{23}|Y_{4,n})] - \sum_{n=1}^{N}[H(M_{23}|Y_{3,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{23};Y_{3,n})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{3,N} \\
&\overset{(a)}{\leq} \sum_{n=1}^{N}[H(M_{23}|M_{13}) - H(M_{23}|M_{13},Y_{3,n})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(M_{23};Y_{3,n}|M_{13})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{3,N} \\
&\overset{(b)}{\leq} \sum_{n=1}^{N}[H(Y_{3,n}|M_{13}) - H(Y_{3,n}|Y_3^{n-1},M_{23},M_{13})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{3,N} \\
&= \sum_{n=1}^{N}[I(Y_3^{n-1},M_{23};Y_{3,n}|M_{13})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{3,N},
\end{aligned}
$$

where (a) and (b) follow from the fact that conditioning reduces entropy. Therefore,

$$
NR_{23} \leq \sum_{n=1}^{N} I(W_n,Z_n;Y_{3,n}|U_n) - I(Z_n;Y_{4,n}) + N\delta_{3,N}. \tag{B.16}
$$

An upper bound on $R_{13}+R_{23}$ can be derived by using the fact that M_{13} and M_{23} are independent, and following the procedure used to derive (B.15) and (B.16) . Consider the following bound on the equivocation obtained from the security constraint (3.3), (3.4) and Fano's inequalities (B.13), (B.14):

$$
\begin{aligned}
N(R_{13}+R_{23}) &= H(M_{13},M_{23}) \\
&= H(M_{13}) + H(M_{23}|M_{13}) = H(M_{13}) + H(M_{23}) \\
&\leq H(M_{13}|Y_2^N) + H(M_{23}|Y_4^N) + N\epsilon \\
&\leq H(M_{13}|Y_2^N) - H(M_{13}|Y_3^N) + H(M_{23}|Y_4^N) - H(M_{23}|Y_3^N) + N\delta_{4,N}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{n=1}^{N}[H(M_{13}|Y_{2,n})] - \sum_{n=1}^{N}[H(M_{13}|Y_{3,n})] \\
&\quad + \sum_{n=1}^{N}[H(M_{23}|Y_{4,n})] - \sum_{n=1}^{N}[H(M_{23}|Y_{3,n})] + N\delta_{4,N} \\
&= \sum_{n=1}^{N}[I(M_{13};Y_{3,n})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] \\
&\quad + \sum_{n=1}^{N}[I(M_{23};Y_{3,n})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{4,N} \\
&\leq \sum_{n=1}^{N}[I(M_{13};Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] \\
&\quad + \sum_{n=1}^{N}[I(M_{23};Y_{3,n}|M_{13})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{4,N} \\
&\leq \sum_{n=1}^{N}[I(Y_3^{n-1},M_{13};Y_{3,n}|M_{23})] - \sum_{n=1}^{N}[I(M_{13};Y_{2,n})] \\
&\quad + \sum_{n=1}^{N}[I(Y_3^{n-1},M_{23};Y_{3,n}|M_{13})] - \sum_{n=1}^{N}[I(M_{23};Y_{4,n})] + N\delta_{4,N}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
N(R_{13}+R_{23}) &\leq \sum_{n=1}^{N}[I(W_n,U_n;Y_{3,n}|Z_n)] - \sum_{n=1}^{N}[I(U_n;Y_{2,n})] \\
&\quad + \sum_{n=1}^{N}[I(W_n,Z_n;Y_{3,n}|U_n)] - \sum_{n=1}^{N}[I(Z_n;Y_{4,n})] + N\delta_{4,N}. \qquad \text{(B.17)}
\end{aligned}
$$

Finally, a time-sharing RV Q, which is uniformly distributed over N symbols and and independent of M_{13}, M_{23}, $\mathcal{X}_1^N$, $\mathcal{X}_2^N$, $\mathcal{Y}_2^N$,$\mathcal{Y}_3^N$,$\mathcal{Y}_4^N$, can be introduced for the single letter characterization of the above derived outer bounds. Applying the procedure similar to the one presented in [52, Chapter 15.3.4] on (B.15), (B.16) and (B.17), we get the outer bounds on R_{13} and R_{23} as given by (3.14) - (3.16), respectively.

B.5 Bound on the conditional entropy

Here, we prove that $H(V^N|M_{13},W^N,Y_4^N) \leq N\epsilon$ for any $\epsilon > 0$ and N sufficiently large. Given $M_{13} = m_{13}$, the decoder at $\mathrm{Node}\ 4$ chooses j such that the following

typicality condition is satisfied: $\tilde{E} = \{(\mathbf{w}, \mathbf{v}, \mathbf{y}_4) \in A_\epsilon^{(N)}(P_{W,V,Y_4})\}$. Let $P_{e,4}^{(N)}$ denote the average probability of error of decoding j at Node 4. Therefore, we have

$$P_{e,4}^{(N)} \leq P(\tilde{E}^c | m_{13} \text{ sent}) + \sum_j P(\tilde{E} | m_{13} \text{ sent}),$$

where $\tilde{E}^c \triangleq \{(\mathbf{w}, \mathbf{v}, \mathbf{y}_4) \notin A_\epsilon^{(N)}(P_{V,Y_4|W})\}$. From joint AEP [52], $P(\tilde{E}^c|\tilde{K}) \leq \epsilon$, for $\epsilon > 0$ and sufficiently small for large N. And, $P(\tilde{E}|m_{13} \text{ sent}) \leq 2^{-N[I(W,V;Y_4)-\epsilon]}$. Therefore, $P_{e,4}^{(N)} \leq \epsilon + 2^{NR'_{13}} 2^{-N[I(W,V;Y_4)-\epsilon]}$. But, from equivocation computation, $R'_{13} = I(W,V;Y_4) - \epsilon_2$. Choosing $\epsilon_2 > \epsilon$, we get $P_{e,4}^{(N)} \leq \epsilon$. Next, from Fano's inequality, we have for any $\epsilon_1 > 0$,

$$\frac{1}{N} H(V^N | M_{13} = m_{13}, W^N, Y_4^N) \leq \frac{1}{N}\left[1 + P_{e,4}^{(N)} R'_{13}\right] \leq \frac{1}{N} + \epsilon I(W,V;Y_4) \triangleq \epsilon_1.$$

Finally,

$$\frac{1}{N} H(V^N | M_{13}, W^N, Y_4^N) \leq \frac{1}{N} \sum_{m_{13}} P(M_{13} = m_{13}) H(V^N | M_{13} = m_{13}, W^N, Y_4^N) \leq \epsilon_1.$$

Appendix C

Proofs for the broadcast channels

C.1 Bound on the probability of error using the second moment method

Here, we upper bound the probability of encoder error for the channel C_1, by using results from the second moment method [24]. This method was also employed in [18] and [26, Chap. 7, pp. 354] to provide an alternative proof of Marton's achievability scheme. An error is declared at the encoder of S if it is not possible to find a pair (i_1, i_2) to satisfy the condition $E_{\mathrm{S}} \triangleq \{(\mathbf{W}, \mathbf{V}_1(i_1, j_1), \mathbf{V}_2(i_2, j_2)) \in A_\epsilon^{(\mathrm{N})}(P_{W,V_1,V_2})\}$. Let $P_{e,E_{\mathrm{S}}}$ denote the probability of error at the encoder, *i.e.*, $P_{e,E_{\mathrm{S}}} \triangleq \Pr(E_{\mathrm{S}}^c)$. Let I be an indicator RV that the event E_{S} has occurred. Let $Q = \sum_{j_1,j_2} I$; $\bar{Q} = \mathbb{E}[Q]$; and $\mathrm{Var}[Q] = \mathbb{E}[(Q - \bar{Q})^2]$, where $\mathbb{E}(.)$ denotes the expectation operator. $P_{e,E_{\mathrm{S}}}$ can be upper bounded as follows:

$$P_{e,E_{\mathrm{S}}} = \Pr(Q = 0) \overset{(i)}{\leq} \mathrm{Var}[Q]/\bar{Q}^2, \tag{C.1}$$

where (i) follows from Markov's inequality for non-negative RVs. Consider now

$$\begin{aligned} \bar{Q} &= \sum_{j_1,j_2} \mathbb{E}(I) \geq \sum_{j_1,j_2} (1 - \delta^{(\mathrm{N})}) 2^{-\mathrm{N}[I(V_1;V_2|U)+I(V_1,V_2;W|U)+4\epsilon]} \\ &= (1 - \delta^{(\mathrm{N})}) 2^{-\mathrm{N}[R_1^* + R_2^* - I(V_1;V_2|U) - I(V_1,V_2;W|U) - 4\epsilon]}. \end{aligned}$$

Next, consider $\text{Var}[Q] = \sum_{j_1,j_2} \sum_{j_1',j_2'} \{\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] - \mathbb{E}[I(j_1,j_2)]\mathbb{E}I(j_1',j_2')]\}$. We have the following four cases:

1. If $j_1' \neq j_1$ and $j_2' \neq j_2$, then $I(j_1,j_2)$ and $I(j_1',j_2')$ are independent and $\text{Var}[Q] = 0$.
2. If $j_1' = j_1$ and $j_2' = j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] = \mathbb{E}[I(j_1,j_2)] \leq 2^{-\text{N}[I(V_1;V_2)+I(V_1,V_2;W)-4\epsilon]}$.
3. If $j_1' \neq j_1$ and $j_2' = j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-\text{N}[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_1;V_2,W)-6\epsilon]}$.
4. If $j_1' = j_1$ and $j_2' \neq j_2$, then $\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-\text{N}[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_2;V_1,W)-6\epsilon]}$.

Substituting for $\bar{Q}$ and $\text{Var}[Q]$ in $(C.1)$, we can show that $P(E_{\text{S}}) \leq \delta_{\text{C}_1}^{(\text{N})}$, $\forall \delta_{\text{C}_1}^{(\text{N})} > 0$ and sufficiently small; and for N large, if the following conditions are simultaneously satisfied:

$$R_1^{'} > I(W;V_1) - \epsilon_1, \tag{C.2}$$

$$R_2^{'} > I(W;V_2) - \epsilon_2, \tag{C.3}$$

$$R_1^{'} + R_2^{'} > I(V_1;V_2) + I(V_1,V_2;W) - \epsilon_3. \tag{C.4}$$

Similar analysis is done to bound the binning rates for the channel C_3. The probability of encoder error $P(E_{\text{S}}) \leq \delta_{\text{C}_3}^{(\text{N})}$, $\forall \delta_{\text{C}_3}^{(\text{N})} > 0$ and sufficiently small; and for N large, if the following conditions are simultaneously satisfied:

$$R_1^{*} > I(W;V_1|U) - \epsilon_{11}, \tag{C.5}$$

$$R_2^{*} > I(W;V_2|U) - \epsilon_{12}, \tag{C.6}$$

$$R_1^{*} + R_2^{*} > I(V_1;V_2|U) + I(V_1,V_2;W|U) - \epsilon_{13}. \tag{C.7}$$

Appendix D

Proofs for the Z-channel

D.1 Proof of Theorem 5.3.1

D.1.1 Encoding and transmission

At S_2, generate $2^{\mathrm{N}R'_{21}}$ sequences $\mathbf{u}(m_{21}, m'_{21})$ for every m_{21} and $2^{\mathrm{N}R'_{22}}$ sequences $\mathbf{v}(m_{22}, m'_{22})$ for every m_{22} according to $p(\mathbf{u})$ and $p(\mathbf{v})$, respectively. Here, $m'_{21} \in \{1, \ldots, 2^{\mathrm{N}R'_{21}}\}$; $m'_{22} \in \{1, \ldots, 2^{\mathrm{N}R'_{22}}\}$. For every pair of $(\mathbf{u}, \mathbf{v})$ sequences, generate one sequence $\mathbf{x}_2$ according to $p(\mathbf{x}_2|\mathbf{u}, \mathbf{v})$. Given (m_{21}, m_{22}), the encoder looks for the pair (m'_{21}, m'_{22}) that satisfies the joint typicality condition
$E_{\mathrm{S}_2} \triangleq \{(\mathbf{u}(m_{21}, m'_{21}), \mathbf{v}(m_{22}, m'_{22})) \in A_\epsilon^{\mathrm{N}} P(U, V)\}$. If there is one or more such pairs, the encoder chooses one and transmits the corresponding $\mathbf{x}_2$ in N channel uses. Otherwise, an encoding error is declared at S_2.

At S_1, given the pair (m_{21}, m_{22}), generate $2^{\mathrm{N}R'_{11}}$ sequences $\mathbf{w}(m_{11}, m'_{11})$ for every m_{11} according to $p(\mathbf{w})$; $m'_{11} \in \{1, \ldots, 2^{\mathrm{N}R'_{11}}\}$. For every sequence triple $(\mathbf{w}, \mathbf{u}, \mathbf{v})$, generate one sequence $\mathbf{x}_1$ according to $p(\mathbf{x}_1|\mathbf{w}, \mathbf{u}, \mathbf{v})$. Given the triple (m_{11}, m_{21}, m_{22}), the encoder looks for m'_{11} that satisfies the joint typicality condition
$E_{\mathrm{S}_1} \triangleq \{(\mathbf{w}(m_{11}, m'_{11}), \mathbf{u}(m_{21}, m'_{21}), \mathbf{v}(m_{22}, m'_{22})) \in A_\epsilon^{\mathrm{N}} P(W, U, V)\}$. We note that, the encoder at S_1 is given $(\mathbf{u}(m_{21}, m'_{21}), \mathbf{v}(m_{22}, m'_{22}))$ in a *noncausal* manner, so that it need not look for the pair (m'_{21}, m'_{22}) to satisfy E_{S_1}. Finally, S_1 transmits the corresponding $\mathbf{x}_1$ in N channel uses. Otherwise, an encoding error is declared at

S_1.

D.1.2 Decoding

Since $(m'_{11}, m'_{21}, m'_{22})$ is a function of (m_{11}, m_{21}, m_{22}), it does not matter if the decoders at D_1 and D_2 have knowledge of $(m'_{11}, m'_{21}, m'_{22})$ ahead of time. The decoder at D_1 accumulates an N-sequence $\mathbf{y}_1$ and looks for the index pair $(\hat{m}_{11}, \hat{m}_{21})$ that satisfies the joint typicality condition $E_{\mathrm{D}_1} \triangleq \{(\mathbf{w}(\hat{m}_{11}, m'_{11}), \mathbf{u}(\hat{m}_{21}, m'_{21}), \mathbf{y}_1) \in A_\epsilon^{\mathrm{N}} P(W, U, Y_1)\}$. If there is one or more such pair, the decoder picks one and declares $(\hat{m}_{11}, \hat{m}_{21})$ to be the transmitted message indices. Otherwise, an error is declared at D_1.

At D_2, the decoder accumulates an N-sequence $\mathbf{y}_2$ and looks for the index $\hat{m}_{22}$ that satisfies the joint typicality condition $E_{\mathrm{D}_2} \triangleq \{(\mathbf{v}(\hat{m}_{22}, m'_{22}), \mathbf{y}_2) \in A_\epsilon^{\mathrm{N}} P(V, Y_2)\}$. If there is one or more such indices, the decoder picks one and declares $\hat{m}_{22}$ to be the transmitted message index. Otherwise, an error is declared at D_2.

D.1.3 Analysis of the probability of error

D.1.3.1 Probability of error at S_1 and S_2

An error is declared at S_1, if the encoder cannot find an index m'_{11} that satisfies the joint typicality condition $E_{\mathrm{S}_1} \triangleq \{(\mathbf{w}(m_{11}, m'_{11}), \mathbf{u}(m_{21}, m'_{21}), \mathbf{v}(m_{22}, m'_{22})) \in A_\epsilon^{\mathrm{N}} P(W, U, V)\}$. Let $P_{e,\mathrm{S}_1}^{\mathrm{N}}$ denote the probability of error at the encoder of S_1, *i.e.*, $P_{e,\mathrm{S}_1}^{\mathrm{N}} \triangleq \Pr(E_{\mathrm{S}_1}^c)$. Using the second moment method [24], it can be shown that, $\forall \delta_1, \epsilon_1 > 0$, $P_{e,\mathrm{S}_1}^{\mathrm{N}} \leq \delta_1$ as $\mathrm{N} \to \infty$ if the following constraints are satisfied:

$$R'_{11} > I(U, V; W) - \epsilon_1. \tag{D.1}$$

Similarly, an error is declared at S_2, if the encoder cannot find a pair (m'_{21}, m'_{22}) that satisfies the joint typicality condition
$E_{\mathrm{S}_2} \triangleq \{(\mathbf{u}(m_{21}, m'_{21}), \mathbf{v}(m_{22}, m'_{22})) \in A_\epsilon^{\mathrm{N}} P(U, V)\}$. Let $P_{e,\mathrm{S}_2}^{\mathrm{N}}$ denote the probability of error at the encoder of S_2, *i.e.*, $P_{e,\mathrm{S}_2}^{\mathrm{N}} \triangleq \Pr(E_{\mathrm{S}_2}^c)$. Using the second moment, it can

be shown that, $\forall \delta_2, \epsilon_2, \epsilon_3, \epsilon_4 > 0$, $P_{e,\mathrm{S}_2}^{\mathrm{N}} \leq \delta_2$ as $\mathrm{N} \to \infty$ if the following constraints are satisfied:

$$R'_{21} > I(U;V) - \epsilon_2, \tag{D.2}$$

$$R'_{22} > I(U;V) - \epsilon_3, \tag{D.3}$$

$$R'_{21} + R'_{22} > I(U;V) - \epsilon_4. \tag{D.4}$$

D.1.3.2 Probability of error at D_1 and D_2

An error is declared at D_1, if the encoder cannot find a pair $(\hat{m}_{11}, \hat{m}_{21})$ that satisfies the joint typicality condition $E_{\mathrm{D}_1} \triangleq \{(\mathbf{w}(\hat{m}_{11}, m'_{11}), \mathbf{u}(\hat{m}_{21}, m'_{21}), \mathbf{y}_1) \in A_\epsilon^{\mathrm{N}} P(W, U, Y_1)\}$. Let $P_{e,\mathrm{D}_1}^{\mathrm{N}}$ denote the probability of error at the decoder of D_1, *i.e.*, $P_{e,\mathrm{D}_1}^{\mathrm{N}} \triangleq \Pr(E_{\mathrm{D}_1}^c)$. Using standard arguments, it can be shown that, $\forall \delta_3, \epsilon_5, \epsilon_6, \epsilon_7 > 0$, $P_{e,\mathrm{D}_1}^{\mathrm{N}} \leq \delta_3$ as $\mathrm{N} \to \infty$ if the following constraints are satisfied:

$$R_{11} + R'_{11} \leq I(W;Y_1|U) - \epsilon_5, \tag{D.5}$$

$$R_{21} + R'_{21} \leq I(U;Y_1|W) - \epsilon_6, \tag{D.6}$$

$$R_{11} + R'_{11} + R_{21} + R'_{21} \leq I(W,U;Y_1) - \epsilon_7. \tag{D.7}$$

An error is declared at D_2, if the decoder cannot find an index $\hat{m}_{22}$ that satisfies the joint typicality condition $E_{\mathrm{D}_2} \triangleq \{(\mathbf{v}(\hat{m}_{22}, m'_{22}), \mathbf{y}_2) \in A_\epsilon^{\mathrm{N}} P(V, Y_2)\}$. Let $P_{e,\mathrm{D}_2}^{\mathrm{N}}$ denote the probability of error at the decoder of D_2, *i.e.*, $P_{e,\mathrm{D}_2}^{\mathrm{N}} \triangleq \Pr(E_{\mathrm{D}_2}^c)$. Using standard arguments, it can be shown that, $\forall \delta_4, \epsilon_8 > 0$, $P_{e,\mathrm{D}_2}^{\mathrm{N}} \leq \delta_4$ as $\mathrm{N} \to \infty$ if the following constraints are satisfied:

$$R_{22} + R'_{22} \leq I(V;Y_2) - \epsilon_8. \tag{D.8}$$

Vita

Kyatsandra G. Nagananda was born in Bangalore, India on 13^{th} March, 1981. He earned his Bachelor of Engineering in telecommunications engineering from Visveswaraiah Technological University in 2003, Master of Engineering degree in electrical engineering from Oregon Health Sciences University in 2007. His recent past affiliations include the Indian Institute of Science, Bangalore in the capacity of a research assistant $(2003-2005)$, research associate $(2007-2009)$ and research consultant (May - August, 2012). His research interests include information theory; physical-layer security in wireless networks; communication protocols and security issues in smart grid.

Journal publications

1. K.G. Nagananda, R. Blum and S. Kishore, '**Reducing bandwidth consumption for synchrophasor data transmission on the smart grid**', under preparation.

2. Chen Chen, K.G. Nagananda, G. Xiong, S. Kishore and L. Synder '**A communication based appliance scheduling scheme for consumer-premise energy management systems**', under review, IEEE Transactions on Smart Grid.

3. K.G. Nagananda, C.R. Murthy and S. Kishore, '**Capacity bounds for state dependent broadcast channels**', under review, Elsevier Physical Communications.

4. K.G. Nagananda, '**Secure communications over opportunistic-relay channels**', under review, Elsevier Physical Communications.

5. K.G. Nagananda, P. Mahopatra, C.R. Murthy and S. Kishore '**Multiuser cognitive radio networks: An information theoretic perspective**', under review, International Journal of Advances in Engineering Sciences and Applied Mathematics.

6. K.G. Nagananda and C.R. Murthy , '**On Z-channels with degraded message sets**', under preparation.

7. K.G. Nagananda, G.V. Anand. '**Subspace intersection method of high-resolution bearing estimation in shallow ocean using acoustic vector sensors**', *Signal Processing*, 2009.

Conference publications

1. K.G. Nagananda, Chandra R Murthy and Shalinee Kishore, '**State dependent broadcast channels with noncausal encoder side-information**', submitted to *IEEE International Communications Conference*, June 2013.

2. K.G. Nagananda and Shalinee Kishore, '**A class of three-user multiple access cognitive radio channels: An achievable rate region**', in the *Proceedings of IEEE Global Telecommunications Conference*, December 2010, Florida, U.S.A.

3. K.G. Nagananda, Chandra R Murthy and Shalinee Kishore, '**Achievable rates in three-user interference channels with one cognitive transmitter**', in the *Proceedings of IEEE Signal Processing and Communications Conference*, July 2010, Bangalore, India.

4. K.G. Nagananda, Chandra R. Murthy, '**Information theoretic results for three-user cognitive radio channels**', in the *Proceedings of IEEE Global Communications Conference*, December 2009, Hawaii, U.S.A.

5. K.G. Nagananda, Chandra R. Murthy, '**Three-user cognitive channels with cumulative message sharing: An achievable rate region**', in the *Proceedings of IEEE Information Theory Workshop on Networking and Information Theory*, June 2009, Volos, Greece.

6. K.G. Nagananda, G.V. Anand. '**Subspace intersection method of bearing estimation in shallow ocean using acoustic vector sensors**', in the *Proceedings of* 16^{th} *European Signal Processing Conference*, 2008, Laussane, Switzerland.

7. Chen Chen, K.G. Nagananda, G. Xiong, S. Kishore and L. Synder '**Analysis of a joint access and scheduling scheme for residential energy management controller**', in the *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop*, June 2012, New Jersey, U.S.A