DISTRIBUTED SYSTEMS FOR MULTIPLE ACCESS AND RELAY BROADCAST UNDER SECRECY

Submitted in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

by

Krishnamoorthy Iyer
Roll No. 08407617

Supervisors:
Prof. Bikash Kumar Dey
Prof. Sibi Raj B. Pillai

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY - BOMBAY
MUMBAI - 400 076

July 2018

Thesis Approval

The Thesis entitled

Distributed Systems for Multiple Access and Relay Broadcast under Secrecy

by

Krishnamoorthy Iyer
(Roll No. 08407617)

is approved for the degree of

Doctor of Philosophy

Prof. Andrew Thangaraj

Kumar Appaiah
Prof. Kumar Appaiah

Bikash Kumar Dey
Prof. Bikash Kumar Dey

Prof. Sibi Raj Pillai

Arup ranjan Bhattacharyya
Prof. Arup R. Bhattacharyya

Date: July 24, 2018
Place: IIT Bombay, Mumbai

DISTRIBUTED SYSTEMS FOR MULTIPLE ACCESS
AND
RELAY BROADCAST UNDER SECRECY

By

Krishnamoorthy Iyer

A Thesis Submitted to
Indian Institute of Technology Bombay
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Approved:

Prof. Bikash Kumar Dey
Prof. Sibi Raj B. Pillai
Thesis Advisors

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY - BOMBAY
MUMBAI - 400 076

July 2018

Dedication

To my amazing and wonderful mother and father, my greatest sources of inspiration and strength.

Declaration

I declare that this written submission represents my own ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

....

Krishnamoorthy Iyer
Roll No. 08407617

Abstract

Distributed systems are becoming increasingly important due to the ubiquitous deployment of wireless devices and sensor networks. This trend is expected to continue with future Internet of Things (IoT) applications and 5G networks. The participating terminals in such networks have varying demands on throughput and data security. In an uplink cellular system, the throughput under distributed multiple access is an important Quality of Service (QoS) metric. On the other hand, a downlink system provides its users an opportunity to wiretap unintended information flows. Preserving the secrecy of each data link is of paramount importance in this case.

In this thesis, we investigate the communication rates and schemes for some wireless network models. Multiple access channels (MAC), broadcast channels (BC), and relay channels are very fundamental models for wireless systems.

We first consider a block fading multiple access channel with the receiver having complete channel state information (CSI), whereas the transmitters have the CSI only of their respective links. The distributed nature of encoder CSI may lead to outage, unless appropriate communication schemes are devised. The long term average sum-rate while ensuring no outage in each block is known as the adaptive sum-capacity; characterizing this is an open problem in literature. The thesis characterizes the adaptive sum-capacity of several practical fading models, under the assumption of identical fading distributions. Furthermore, the effect of additional quantized CSI on other links is also analyzed, and the sum-capacity identified for some important models.

Data security is the second prominent aspect discussed in this thesis. A downlink wireless channel provides an opportunity for eavesdropping by unintended users. While data can be secured using crypto systems, we can demand the more stringent information theoretic security. To analyze communication rates under secrecy requirements, the thesis first considers a two user discrete memoryless broadcast channel where each receiver shares separate secret keys with the transmitter. We propose an achievable scheme and outer bound, under the availability of different secret key rates on each data link. The rest of the thesis considers secure communication in the absence of secret keys.

We consider a relay channel with an eavesdropper. We give an achievable rate that is the same as that achieved in an earlier work using backward decoding. Our achievable scheme uses block Markov encoding over a super-block consisting of many smaller blocks, and sliding window decoding. Unlike backward decoding, sliding window decoding incurs a smaller decoding delay of two blocks rather than the full super-block. Though this scheme is generalized in the next model we consider, both the current simple model and the simple compact rate expression for this special case are of independent interest.

We finally consider a generalization of the above model to a relay broadcast channel (RBC) with two receivers. Each receiver needs to be sent an independent message, and each message is to be kept secret from the unintended receiver. The relay is used as a broadcast channel in its own right. For the first time, we give an achievable rate region. The encoding is an extension of the encoding scheme for the earlier relay system with an eavesdropper, and uses a block Markov encoding over a super-block and sliding window decoding.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Overview

Sixty-eight years have passed since the publication of Claude Shannon's seminal paper "A Mathematical Theory of Communication" Shannon [1], which was motivated by the engineering problem of transmitting a message (i.e. information) reliably over a telephone wire connecting users – referred to as transmitter (aka sender) and receiver – at two different locations. Physical phenomena associated with the wire itself were usually – and often – not under the control of either the transmitter or the receiver. As a consequence, wires tended to randomly distort messages – a phenomenon referred to as "noise". Engineering intuition indicated that the "noisier" the connecting link (i.e. wire), the lower the maximum rate of reliable information transmission. Shannon's mathematical formulation enabled a tractable treatment of these intuitions, and provided communications engineers, especially coding theorists,[1] with guidelines aka "the Shannon limit"[2] on what targets (in terms of rate, probability of error etc.) were and were not possible in a given scenario.

Technological advances in the meantime have (largely) replaced wires as the physical medium with the surrounding air/ atmosphere itself[3] as the physical medium. Nowadays, wireless telecommunications networks have become ubiquitous. With the advent of the Internet of Things (IoT), they are likely to become even more so. Noise in wireless networks

[1] Analogously, the development of the theoretical principles of thermodynamics in the nineteenth century provided engineers involved in the construction of heat engines with the notions of energy and (thermodynamic) entropy. Both these (then novel) concepts set absolute constraints on what was theoretically achievable. By enabling serious practitioners to recognize that *perpetuum mobile* aficionados were chasing a chimera, the development of thermodynamics freed the field of cranks and charlatans, and more positively, gave engineers meaningful targets to aspire to.

[2] The history of the "Shannon limit" is a fascinating subject. In the early days of the subject, many coding theorists believed that there existed a practical limit that was different from – and lower than – the Shannon limit. See Aftab, Kim, Cheung, Thakkar, Yeddanapudi [2].

[3] In poetic contexts, the (upper) atmosphere itself is sometimes referred to more evocatively as "the ether". But I have not seen the phrase used in engineering contexts (such as the present). Methinks, perhaps we should.

is no less an issue in comparison with wired networks. But because of its broadcast nature, in wireless networks, other issues such as secrecy and (unintended) interference from other users have also taken centre-stage, sometimes in solo acts, sometimes in unison. Deliberately causing interference i.e. jamming is also an issue, especially in military applications. In a network consisting of multiple (i.e. ≥ 3) users, as opposed to a single transmitter-receiver scenario, network issues such as partial and/or distributed availability of information at the nodes also come to the fore.

From the day of its birth, Shannon's information theory provided a rich mathematical framework for addressing questions related to the transmission (and also, as it turned out, the storage) of information. Information-theoretic inequalities have turned out to be of mathematical interest in their own right. See the introductory textbooks by Cover and Thomas [3, Chapter 17] and by Yeung [4, Chapter 12], [5, Chapters 12 through 14], and also the paper Dembo, Cover, Thomas [6]. Indeed, information theory has been found to have profound connections to a variety of other mathematical disciplines such as:

- Probability:
  - Concentration of Measure [7, 8],
  - Proof of the Central Limit Theorem [9],
  - Relating Conditional Expectation and Conditional Mutual Information [10],
- Group Theory [4, Chapter 16],
- Combinatorics [11] and Graph Theory [12],
- Combinatorial Optimization [13],

Information theory has found applications to many fields other than telecommunications engineering, for example:

- Estimation Theory [14, 15, 16],
- Quantum Information Theory [17] and Quantum Computation [18],[4]
- Computational Neuroscience [19],[5] [20, Chapter 4],
- Statistics [21], [3, Chapter 11],
- Machine Learning (ML): The delightful textbook by David MacKay [22] describes connections between information theory and ML. The field of computer vision and

[4] The cognoscenti colloquially refer to [18] as "Ike and Mike".

[5] [19] sparked the author's interest in information theory.

pattern recognition[6] (CVPR) has also benefited from information-theoretic ideas, as evidenced by Escolano, Suau, Bonev [23]. In the field of ML, the "Information Bottleneck Method", a generalization of rate-distortion theory, was developed by Tishby, Pereira, and Bialek [24][7] to study large data sets, and shows promise in the (ML sub-) field of deep learning Tishby and Zaslavsky [25]. See also the PhD thesis by Slonim [26].

Shannon's original formulation can be seen as a theory of information flow between two points connected by a telephone wire. Network Information Theory can be seen as a natural generalization and successor to Shannon's formulation, and provides the mathematical underpinnings of the theory of information flow over wireless telecommunication networks. The bible of the subject is the textbook by Gamal and Kim [27].

Information transmission over a wire provided the motivation for many problems in information theory in its early days. Likewise, wireless networks nowadays provide motivation for and are a rich source of problems that can be studied using the tools of network information theory.*This thesis studies network information theory problems motivated by wireless networks. We believe that our results will be of interest to two communities:*[8]

- *Wireless communications engineers interested in information-theoretic aspects of their subject.*
- *(Network) Information theorists looking to collaborate with and/or convey their insights to wireless communications engineers.*

A number of works have been published in this area. We refer to a small but significant subset of examples. The seminal paper by Xie and Kumar [28], which (re-)introduced[9] sliding window decoding (see subsection 1.3.2) has the phrases 'wireless communication' and 'network information theory' in its title! Avestimehr, Diggavi, Tse [30], which proposed a deterministic approach to the problem of maximum rate of information flow over a wireless network with an arbitrary number of relay nodes and a single source and destination, has the term 'wireless' and the phrase 'network information flow' in its title. Vaze's

[6] Pattern Recognition is traditionally considered a sub-field of ML.

[7] William Bialek is a theoretical physicist whose motivation was the analysis of multi-electrode *spike train* data. "Spikes" are sudden changes in the voltage of a brain cell (aka neuron) with respect to its surroundings. A "spike train" is a temporal sequence of spikes. Spike (trains) constitute the brain's (primary) signaling mechanism. The afore-mentioned recordings obtain data from the order of dozens of neurons. Making sense of the data is a central problem in neuroscience, and the problem of how exactly information is encoded in a spike train is referred to as "the neural code". See [19].

[8] The intersection set is non-empty – the same person may, and often does, wear two hats!

[9] To the best of our knowledge, Carleial [29], in his 1982 paper on the multiple access channel with generalized feedback, was the first to introduce sliding window decoding.

recent textbook [31] has the phrases 'wireless networks' and 'information-theoretic' in its title. Kannan, Raja, Viswanath [32] and Raja and Viswanath [33] have recently proposed approaches for (broadcast-)relay and relay networks that are inspired by issues arising in wireless networks/communication.

First, we give a two paragraph telegraphic[10] *summary of the thesis contents and motivation.*

- *Distributed Information Processing in Wireless Networks:* Chapter 2 is the sole chapter dedicated to a problem of distributed information processing, when decisions regarding transmission rates to a destination node have to be made at two spatially separate source nodes when the relevant information about channel conditions is present only partially at each source node. This is a natural problem to consider in any network, and we have considered it in the context of a multiple access channel (henceforth shortened to MAC).

- *Mutual Secrecy in Wireless Networks:* Chapters 3–5 are motivated by an issue arising due to the broadcast nature of wireless telecommunications networks. Air as the physical medium over which information is sent has many technological and economic[11] advantages. But air does have the drawback that it enables eavesdropping, which thus forms a major security headache. When multiple users, each with legitimate message requirements, attempt to eavesdrop on messages not intended for them, the problem complexity and richness increases and so do the complexity and richness of the solutions proposed i.e. achievable schemes. These chapters study different aspects of the problem of mutual secrecy requirements in increasingly more complex scenarios. Chapter 3 studies the problem of mutual secrecy over a broadcast channel with each destination possessing a secure dedicated link to the transmitter which can be used to share a secret key. Chapters 4 and 5 replace the secret key by a relay.

Before describing the chapter contents in more detail (for which, see section 1.2 below), we briefly digress. The next two items are almost – but not quite – identical to the items above. But note the subtle change of emphasis between the corresponding items – indicated by the crossed out phrases.

[10] The telegraph was an early communications device that was used to transmit messages without the physical exchange of an object bearing the message. The adjective formed by adding the suffix '-ic' seems apt, given the thesis contents.

[11] Air is free, but bandwidth is not!

- *Distributed Information Processing:* ~~*in Wireless Networks:*~~

  Nodes being required to make decisions in the presence of partial information is a standard trope in networks.[12] Anantharam and Borkar [34] is a short, highly readable classic that uses information theory to formulate a distributed zero-sum game over a network, and develops a useful insight regarding the use of common randomness in distributed network control.

  A recent PhD thesis awarded in the area of distributed information processing over networks – with possible applications to network control – is Cuff [35]. Cuff was partially motivated by [34]. To solve these and related problems, Cuff and co-workers [35, 36] introduced the information-theoretic concepts of *empirical coordination* and *strong coordination.*[13] Chou, Bloch, Kliewer [37], Obead, Vellambi, Kliewer [38], and Cervia, Luzzi, Treust, Bloch [39] contain more recent work on strong coordination. [40] discusses an application of strong coordination in the context of a two-way communication via a relay using the (see below) OSRB techniques developed by Yassaee, Aref, Gohari [41].

  The classic textbook by Nancy Lynch [42] discusses distributed algorithms, a (sub-)field of Computer Science that studies algorithmic approaches to distributed information processing.

- ~~*Mutual*~~ *Secrecy in Wireless Networks:* Liang, Poor, and Shamai's monograph [43] discusses information theoretic security at the physical layer, as does Bloch and Barros's book [44]. See also the special issue Debbah, Gamal, Poor, Shamai [45]. [46], by the authors of [43], confines its scope more narrowly to physical layer security in *broadcast channels.* Both Bassily, Ekrem, He, Tekin, Xie, Bloch Ulukus, Yener [47] and Mukherjee, Fakoorian, Huang, Swindlehurst [48] contain summaries of recent ($\leq$ 5 years ago) advances in the area of physical layer security in (multiuser) wireless networks. The monograph by Narayan and Tyagi [49] discusses secrecy key generation, randomness extraction, and secure function computation – among other topics – in multiterminal networks via public discussion in the presence of an eavesdropper.

  Expectedly, many PhD theses have studied information-theoretic security in wireless networks, such as: Bassily [50], Gabry [51], He [52], Hou [53], Nagananda [54] and

[12] Not just telecommunication networks, computer networks and queueing networks also face similar issues.

[13] Empirical coordination is obtained if the empirical joint distribution (aka joint type) of the network actions is "close" to the desired joint distribution. It can be seen as a restatement of rate distortion theory in the context of source coding over a graphical network [27, Chapter 20]. Strong coordination is a strictly stronger requirement, and is obtained when the joint distribution of the *sequence* of actions is "close" to the target distribution that is itself obtained as the product of i.i.d copies of a desired distribution. Strong coordination has applications in cooperative game theory.

> Perron [55]. With the exception of Gabry [51], the aforementioned theses are largely concerned with standard information-theoretic questions such as achievable schemes and outer bounds. Gabry [51] explores the rich interplay between game-theory, cooperation and secrecy in wireless communications. (The definition of secrecy used in [51] is *weak secrecy*, an information-theoretic notion. This is the notion of secrecy used in this thesis).

The author has also benefited from the work and the insights contained in the following PhD theses (and related papers) by Chia [56], Hou [53], Luo [57], Wu [58] and Zhong's MS thesis [59]. Hou [53] considers noisy network based (NNW) coding schemes (which can be considered to belong to CF family of schemes). These incur large delay and it is not quite clear that they give a rate improvement, as the work of Wu and co-workers' [58, 60] shows. One recent work that promises to revolutionise the study of secrecy in network information theory[14] is the recently developed framework "output statistics of random binning" (shortened to OSRB) of Yassaee, Aref, and Gohari [41]. Their techniques directly give *strong* secrecy[15] and what is more, secrecy arises almost trivially from their techniques.

Luo [57], Wu [58], Zhong [59] discuss CF based schemes for relay channels, which we have not considered in this thesis. But see the concluding chapter 6.

To summarise, this thesis studies constrained information transmission over wireless telecommunication networks. We now describe the chapter contents in more detail in the next section.

1.2 Chapter Contents

- Chapter 2 studies information flow over a two-user multiple access channel (MAC).[16] The channel conditions are variable. In the problem's simplest formulation, each transmitter is only aware of her own channel, while the receiver is aware of the entire channel. The information relevant to making decisions at the transmitters about transmission rates is distributed, and hence we refer to this as a distributed multiple access channel with individual CSI (at the transmitters).

  The main ideas on which our achievable scheme is based are the 'alpha-midpoint' strategy which is a distributed rate-allocation strategy. *In this context, the adjective 'distributed' indicates that no central coordination is required, and the transmitters use*

[14] Not just secrecy, their work is also applicable to problems of strong coordination.

[15] See Bloch and Barros [44, Chapter 4] for a definition. Strong secrecy, as the name indicates, is a strictly stronger requirement and implies weak secrecy. The achievable schemes we develop only guarantee weak secrecy.

[16] 'Uplink' in wireless networks.

only local (and partial) information about channel conditions to make decisions about transmission rates. The implementation of the strategy is by means of a low-complexity rate-splitting scheme and an associated novel successive single-user decoding procedure. In many systems of interest, the introduced techniques improve over the performance of conventional centralised scheduling schemes.

Portions of this work were presented at ISITA 2012 Iyer, Pillai, Dey [61]. See also the journal paper Iyer, Pillai, Dey [62]. Our work has been extended in Sreekumar, Dey, Pillai [63], and Deshpande, Dey, Pillai [64].

- Chapter 3 studies the dual of the MAC, namely the broadcast channel (BC).[17] Wireless networks face the problem of securing the messages of their intended receivers from eavesdroppers. It is often the case that two users, each with legitimate message requirements, attempt to eavesdrop on messages not intended for them. The problem was first studied by Liu, Maric, Spasojevic, Yates [65].

  As Perron, Diggavi, Telatar [66] point out, even a small amount of shared key between a source-destination pair, if it can be kept unconditionally secret from the eavesdropper, can enhance the secrecy of the system. With this as motivation, we generalize [65] and study an extension arising from the use of secret keys, inspired by Kang and Liu [67]. We also present a (partial) converse.

  The achievability scheme for this model was presented at NCC 2016 Iyer [68].

- The next two chapters 4 and 5, study the same problem of (mutual) secrecy but we dispense with the secret keys and introduce a relay trusted by *all*[18] parties. *The introduction of even a single relay leads to a four-node network,*[19] *from which a very rich set of scenarios arise.* We call our model the relay broadcast channel with mutual secrecy requirements.

  Dai, Yu, and Ma [70] have studied the same model, but we believe that our achievable scheme constitutes an improvement. *They do not use the relay as a broadcast channel in its own right. They employ backward decoding. Lastly, a close analysis of their scheme reveals that the secret messages are being transmitted via the direct source-to-destination links. In the DF based schemes they consider, this is problematic, as DF based schemes are used precisely because the source-to-destination links are weak*

[17] 'Downlink' in wireless networks.

[18] Dispensing with this assumption leads to another set of models. See subsection 1.3.3.

[19] Ekrem and Ulukus [69] have studied the rate-equivocation region for a three-node network – what they have also called a relay broadcast channel – with mutual secrecy requirements between the destination nodes. See subsection 1.3.1.

compared to the source-to-relay links, which is why the relay is called upon to decode and then forward the message.

- Chapters 4 and 5. Chapter 5 addresses the same problem that Chapter 3 does, but with the secret keys replaced by a (trusted) relay.

  In both these chapters, we assume a "strong" relay scenario i.e. the source-to-relay link is stronger than (both) source-to-destination link(s). Consequently, the relay can decode the legitimate user's message (in chapter 4) and both users' messages (in chapter 5). In these situations, decode-forward (DF) is the preferred option Behboodi and Piantanida [71].

- From an expository point of view, we have found it convenient – in chapter 4 – to describe a scenario with a single legitimate receiver and an external eavesdropper, a problem that was studied by Lai and Gamal [72, Theorem 2] and also by Yuksel and Erkip [73]. The primary contributions of this chapter are as follows:

  – The same pure secrecy rate as obtained by [72] is obtainable via the use of sliding window decoding (SlideWin henceforth). This drastically reduces decoding delay and makes it uniform.

  – By varying a parameter, we can obtain both regular (the codebook sizes at the transmitter and relay are identical) (as in [72]) as well as irregular (the codebook sizes at the transmitter and relay are different) encoding schemes.[20]

  – The requirement of pure secrecy imposes a constraint on the packing of the relay codebook bins – what we have termed the *randomization requirement.*

  – The multi-block equivocation calculation will also provide clues for a similar calculation in the next chapter 5.

- Chapter 5 builds on the insights gained to consider a four-node relay broadcast channel with a trusted relay and a mutual secrecy scenario where both receivers have legitimate message requirements but each also attempts to snoop on the other.

  But for DF to be effective, we also require that the relay-receiver link be stronger than the relay-eavesdropper link [75, Chapter 7*: Ekrem and Ulukus]* . This may mislead one into assuming that a trusted relay cannot be used to simultaneously assist in creating

[20] This was discovered by us independently, but was first noticed by Razaghi and Yu [74]. The application to secrecy is novel, to the best of our knowledge. As [74] point out, unlike in irregular encoding, the relay message rate can be flexible and this can enable higher DF rates in multiple relay networks. Studying the impact of multiple relays would constitute a natural generalization of our model, and would be worth exploring. See the concluding chapter 6.

mutual secrecy between two receivers, both with legitimate message requirements. In chapter 5, we present an achievable scheme that – by using insights from (see below) Liu, Maric, Spasojevic, Yates [65] and Zhao and Chung [76] – shows that this naive intuition is incorrect.

The achievable scheme we present uses elements from three distinct models, and our model can be seen as a generalization of all of these:

- Lai and Gamal [72, Theorem 2] employ block Markov coding, regular encoding, backward decoding and coherent transmission between the transmitter and relay. We use block Markov coding and coherent transmission but use irregular SlideWin.
- Liu, Marić, Spasojević, Yates [65] considered a broadcast channel with mutual secrecy requirements. The primary idea is the notion of double random binning which enables simultaneous non-zero secrecy rates to the receivers in a general broadcast channel. Both chapter 3 and chapter 5 can be seen as generalizations of [65]'s model.
- Zhao and Chung [76] studied the same topology that we have. Both receivers have independent message requirements but there is no secrecy requirement, either individual or mutual. Kramer, Gastpar, Gupta [77] had also developed an achievable scheme for the (four-node) relay broadcast channel but [76] pointed out that their work does not use the relay as a broadcast channel in its own right. Their choice is likely to lead to suboptimal performance in wireless scenarios. Zhao and Chung [76]'s achievable scheme enables coherent transmission between the transmitter and the relay, *with both being used as broadcast channels in their own right.*
- Note that the achievable region for the broadcast channel with mutual secrecy requirements [65] uses double random binning and is not a simple generalization of the achievable region for the broadcast channel, which uses Marton coding [27, Chapter 8], [78]. In like manner, our achievable rate region is not a simple generalization of that of Zhao and Chung [76].

- Chapter 6: Conclusion and Future Work. In the chapters 4 and 5, we have considered DF based schemes, indicated in the "strong" relay scenario, where the source-to-relay link is stronger than the source-to-destination link, as a consequence of which the relay can completely decode the destinations' intended messages. We have some preliminary results on the "weak" relay scenario – the source-to-relay link is weaker than one or both of the source-to-destination(s) link(s) – and so the relay cannot completely decode the destinations' intended messages. In these scenarios, CF based schemes are indicated [27, Chapter 16, Section 16.7].

1.3 Recurring Themes and Literature Survey

1.3.1 The phrase "relay broadcast chanel" (RBC henceforth)

Kramer, Gastpar, Gupta [77, Theorem 2] used the phrase "broadcast relay channel" for the four-node network topology considered in chapter 5. No secrecy requirements, either individual or mutual, were imposed. The relay was a dedicated relay. Rate splitting was used to increase the achievable rate region – as was also done by Behboodi and Piantanida (see below). The relay was used only to increase the common message rate and thus in effect was used as a point-to-point channel. Zhao and Chung [76] pointed out that the relay was not used as a broadcast channel in its own right in [77], and this formed the starting point for our work in chapter 5.

Liang and Veeravalli [79] introduced and presented inner bounds[21] for two three-node networks that they termed:

- the "partially cooperative RBC" [79, Theorem 1] where one user (the "better" user) helps the other user by sending relay signals. Regular encoding/DF/sliding window decoding was used.

- the "fully cooperative RBC" [79, Theorem 9, 10], where both destinations also have relay links. The first inner bound (Theorem 9) is obtained via DF at one node while the other node facilitates[22] and switching between the roles of the two destination users, while the second inner bound (Theorem 10) employs DF at one destination node and CF at the other destination node.

No secrecy requirements (mutual or individual) were imposed in either of the above models.[23]

Liang and Kramer [82] studied both of the above three-node channels [82, Fig. $1(a, b)$] as well as a four-node [82, Fig. $1(c)$] channel, all of which were termed as "relay broadcast channels". *Our relay broadcast channel is the third of these,* referred to in their paper as dedicated-relay broadcast channel. No secrecy requirements were imposed in [82], and the achievable schemes considered were pure DF based schemes, and no CF based schemes were explored.

[21] They also presented outer bounds, but we are not concerned with that here.

[22] That is, it chooses the codeword that achieves the best rate region.

[23] The CF decoding techniques used in their paper use Cover and Gamal [80]'s schemes, and improvements may be possible by applying techniques due to [81].

To the best of our knowledge, the earliest paper considering mutual secrecy in a three-node network other than [65] was by Ekrem and Ulukus [69]. Ekrem and Ulukus's work was alluded to in a footnote in section 1.2. Despite the conflicting requirement of mutual secrecy, a counterintuitive result of the paper was to show that cooperation between the destination nodes was possible by means of a one-sided cooperative link [69, Theorem 1 which is a special case of Theorem 4] or a two-sided cooperative link [69, Theorem 5]. Ekrem and Ulukus [69] also discuss the important notions of jamming and "peeling off", first described in Tannious and Nosratinia [83] in the context of a relay channel with private messages. Note that if neither of the cooperating links are present, the model simplifies to a broadcast channel with mutual secrecy, first studied by Liu, Maric, Spasojevic, Yates [65]. However, the achievable schemes in [69] employ CF[24] and use decoding techniques that have been superseded by the recent work by Luo, Gohary, and Yanikomeroglu [81]. See the concluding chapter 6 on future work.

The phrase "broadcast relay channel" has been used by Behboodi and Piantanida to describe five-node networks with two relays and two destination nodes [85], [86], [87], [71]. No secrecy requirements, either individual or mutual were imposed. Both DF as well as CF based schemes were considered. The destinations employ backward decoding. However, the authors' view that backward decoding provides "better performance" than sliding window – presumably in terms of rates/rate regions – seems to be not the case.

1.3.2 Sliding Window Decoding

A recurring theme in all the chapters involving a relay is the need to use sliding window decoding, henceforth referred to as SlideWin (decoding). The decoding delay of SlideWin is much less compared to backward decoding. Wireless communication engineers – unlike ivory tower information theorists – live in the real world, and reducing decoding delay is always welcome.

Carleial [29] studied a multiple access channel with generalized feedback and was the first, to the best of our knowledge, to introduce sliding window decoding.

Sliding window decoding (SlideWin) was (re-)introduced by Xie and Kumar [28] in the context of wireless networks and elaborated by them in [88]. This was an excellent example of a wireless communication *engineering* problem suggesting a fundamental improvement to techniques in network information *theory.* As [28] pointed out, SlideWin necessitates the use of independently generated codebooks in distinct blocks.

Work by Chong, Motani, Garg [89, 90, 91] and by Hou and Kramer [92, Sec *IV*] has

[24] Bloch and Thangaraj [84] consider a three-node network where the relay has a secret message, and forwards a common message to the destination by means of DF.

shown that SlideWin occurs no rate penalty in comparison with backward decoding, calling into question Behboodi and Pintanida's choice of backward decoding in the sequence of papers above. Note that [92] discuss SlideWin in the context of pure CF, whereas Chong, Motani, Garg in their aforementioned sequence of papers consider a mixed DF-CF scheme for the canonical relay channel. Behboodi and Piantanida [85, 86, 87] consider a two receiver two relay setup.

Note that the phrase "sliding window decoding" has been used in the literature to refer to CF based schemes in ways that are not in keeping with the spirit of the phrase as used in DF based contexts. In SlideWin as used in DF contexts, see [27, Chapter 18] a list of possible relay codewords is intersected with a possible list of transmitter codewords in the previous block, and exactly one message index must belong to both lists. *This means that the relay codeword is decoded uniquely and simultaneously with the previous block transmitter codeword in these DF based schemes.*

Now consider the CF based scheme discussed in Lai and Gamal [72, Theorem 4, Comment 2]. They refer to their decoding scheme for CF as sliding window decoding. In their scheme, the relay channel codeword is decoded first (this gives rise to a bottleneck[25]). This in turn determines the WZ bin used in the previous block. The compression sequence in the previous block must belong to this bin. The compression sequence is either uniquely and sequentially decoded (as in [72, Theorem 3, 4], and also in the original CF formulation [80]) or jointly and nonuniquely decoded (as in [27, p. 400], Dabora and Servetto [93], El-Gamal, Mohseni, and Zahedi [94]).

If the spirit of the SlideWin definition for DF is to be maintained, then of the CF based schemes, the one presented by [53] and [95] both of which do not involve WZ binning, and the scheme presented by Luo in the thesis [57] make the cut.

A recent beautiful[26] paper that uses sliding window decoding in the context of a relay network (with multiple relays) is work by Yassaee and Aref [96], ideas from which are utilised by [57] in developing their scheme.

Tang's MS thesis [97], and [98] contain a discussion of SlideWin in the context of partial DF in a relay network consisting of a single source and single destination with multiple relays.

1.3.3 Relay Assumption(s)

The simplest scenario is a *single* and *dedicated* relay trusted by *all* parties. We have considered this in chapter 5.

[25] This bottleneck disappears in the work of Luo and co-workers [81].

[26] And recondite!

1. By 'dedicated' is meant that the relay has no private/secret messages to transmit and/or receive. Another possibility is that the relay has private messages it would like to transmit to a third party, or if the relay is itself the intended recipient of a private message, as in Tannious and Nosratinia [83],[27] or a secret message, as in Ekrem and Ulukus [69] and Bloch and Thangaraj [84].

2. Another complication arises if the relay is not trusted by one or more receiving nodes. An untrusted relay cannot assist in DF based schemes.

   To the best of our knowledge, Oohama [99] and Oohama and Watanabe [100] were the first to consider a relay channel with an untrusted relay acting as a wiretapper (aka eavesdropper).

   In an insightful sequence of papers, He and Yener [101], [102], [103], [104], [105] showed that an untrusted relay – or a collection of untrusted relays – could nonetheless be used to increase secrecy rates via CF based schemes. Ekrem and Ulukus [69] used these insights and showed in the three-node network (see subsection 1.3.1), where, despite the mutual secrecy requirement, cooperation can increase achievable secrecy rates.

   More recent work, such as Zewail, Nafea, Yener [106] considers multi-terminal networks with an untrusted relay.

   A natural extension of the model studied in chapter 5 would be if one of the two destination nodes distrusts the relay. In that case, we would be forced to use CF based techniques for that destination.

3. An opportunistic relay was studied in Nagananda [107] in the context of physical layer security. Luo, Gohary, Yanikomeroglu [81] studied an opportunistic relay channel with no secrecy requirements. [107], who dubs the relay as a 'cognitive' relay, uses DF based techniques. [81] use CF based techniques.

A running theme through Chapter 5 (pure DF) is the need to *treat the relay as a broadcast channel in its own right*, an insight developed in the work of Zhao and Chung [76]. Note that Pillai [108] had earlier considered – in the context of a Gaussian relay broadcast channel – the use of the relay as a broadcast channel.

1.3.4 Canonical Relay Channel: CF based schemes

The canonical relay channel has still not been fully understood, as the inner and outer bounds do not match in general [27, Chapter 16].

[27] This paper introduced the important notion of "peeling off".

Thesis chapters 4 and 5 both employ DF based achievable schemes[28] to achieve secrecy.[29] These DF based schemes were first described by Cover and Gamal in their classic paper [80, Theorem 1]. In the same paper, Cover and Gamal [80, Theorem 6] also developed achievable schemes based on compress-forward (CF).

We do not discuss CF based schemes that achieve secrecy for the relay channel in this thesis. See concluding chapter 6 on future work. But we briefly review the fascinating history of the increasingly more sophisticated CF based achievable schemes of the past thirty-eight years.

- Cover and Thomas [80, Theorem 6]'s CF scheme – the first such – involved Wyner-Ziv binning of the compression codewords at the relay. The relay codeword in a block encoded the WZ bin index of the previous block. To decode a block's message, the receiver followed a three-step process of successively decoding the relay codeword in the next block, which gave the WZ bin index of the current block. Using the received sequence as side information, the receiver successively decoded the compression codeword in the current block. Next, the message was decoded. *The first two steps of this process gave rise to bottlenecks that subsequent work has shown to be unnecessary.*

- Dabora and Servetto [93] (see also Gamal and Kim [27, Chapter 16]), as before, decoded the relay codeword to obtain the WZ bin. But they replaced the remaining two steps with joint decoding of the message and (nonuniquely) the compression index within the WZ bin. Because the compression codeword is not required to be decoded uniquely, the corresponding bottleneck could be dispensed with. *However, the achievable rates remained unchanged* [27, Appendix 16C].

- El-Gamal, Mohseni, and Zahedi [94] performed essentially the same decoding as above in the context of a Gaussian relay channel.

- Kramer and Hou [109] and Zhong, Haija, Vu [95] and Zhong MS thesis [59, Theorems 6, 7 and remarks immediately following] showed that one could dispense with WZ binning and encode the compression index directly as the relay codeword index, and employ sliding window (joint) decoding. *For the canonical relay channel, in the CF case, this would incur no rate penalty.*[30]

[28] DF based schemes are preferred in the "strong" relay scenario [71].

[29] Perfect Weak Secrecy.

[30] This is a slightly surprising result, in view of Kim, Skoglund, and Caire [110], who showed that sequential decoding of the compression index followed by the message incurs rate loss if WZ binning is not performed. The conclusion is that joint decoding is crucial for CF to achieve optimal performance if WZ binning is not employed. See also [59, Chapter 4, Remark 3].

- Luo, Gohary, Yanikomeroglu [81], and also Luo (PhD thesis) [57] re-introduced WZ binning, and demonstrated the superiority of a decoding scheme developed by them in the context of multi-relay networks. Their decoding scheme for the canonical relay channel – as described in [81] – consisted of forming possible lists of compression sequences both before and after the message in the block was decoded. (The compression sequence was not decoded uniquely). As described in their 2012 paper [81], the second list was used to reduce the search space of the relay codewords in the next block – this enabled the constraint on the size of the relay channel codebook to be replaced by a constraint on the size of the compression codebook that acted as a proxy to determine the relay codeword uniquely. *However, for the canonical relay channel, the achievable rates remained unchanged.*[31]

- *Note that noisy network coding (NNW coding) can be considered to be a kind of CF based scheme, with nonunique decoding of the compression indices.* There are two main kinds of NNW schemes:

  - short message NNW coding [92]
  - long message NNW coding [27, Chapter 18]

  The work of Wu and co-workers' [58], [111], [112] indicates that NNW based schemes do not give any advantage – in terms of rate at least – over other CF based schemes. Further, compared to standard CF, long message NNW entails a large coding delay.

- Note that other variants of CF have also been proposed. For example, Cover and Kim (for the deterministic relay channel) [113], for *primitive*[32] relay channels [116], Razaghi and Yu (generalized hash-forward strategy) [117]. We will not review these here, as this thesis is mainly concerned with DF based strategies.

1.3.5 Relay-Eavesdropper Channel

Note that we have looked at DF schemes for the relay-eavesdropper channel. But for the sake of completeness, we briefly review the following three achievable schemes proposed for the relay-eavesdropper channel, which use CF based and mixed schemes:

[31] In the PhD thesis, Luo [57] showed that the compression sequence could also be decoded uniquely in regimes of interest.

[32] Mondelli, Hassani, Urbanke [114] call relay channels with orthogonal receiver components as *primitive* relay channels. Thus the channel is described by $P_{Y_{rec,1},Y_{rel}|X_{tr}}P_{Y_{rec,2}|X_{rel}}$. Kim [115] uses the term *primitive* relay channel to refer to a special case of the foregoing, namely, channels with a separate and noiseless relay to receiver link. The $P_{Y_{rec,2}|X_{rel}}$ link is replaced by a noiseless bit-pipe. What is the common to both is that the transmitter connects to the relay and the receiver via a broadcast channel, namely $P_{Y_{rec,1},Y_{rel}|X_{tr}}$.

- Xu, Ding, Dai [118] have applied noisy network coding (NNW)[33] to the relay-eavesdropper channel. The work of [58] indicates that NNW does not give any advantages over CF based schemes for the canonical relay channel (or for single-source single-destination channels with multiple relays). It seems unlikely that it could give improvements in achievable rates compared to CF based schemes for the relay eavesdropper channel.
- Sonee, Salimee, Salmasizadeh [119] have considered a mixed DF + CF based scheme for the relay-eavesdropper channel.
- Xu, Ding, Dai [120] have considered a different hybrid scheme for the relay-eavesdropper, one which also uses backward decoding.

1.3.6 Miscellaneous Remarks

Chen [121] considered a wireless broadcast network with multiple relays which also act as destinations. The security clearances of the relay-cum-destination nodes form a hierarchy. Under the crucial assumption that the security clearance directly reflects the strength of the link to the source, they describe an achievable scheme (for the DF case) that uses regular encoding, SlideWin and superposition encoding. Regular encoding and SlideWin together is likely not an optimal choice, as Razaghi and Yu [122] indicate. Furthermore, superposition coding seems an odd choice for a (general) broadcast network, given that Marton coding has been known since the time of, well, Marton [123], [27, Chapter 8]. While, for the specific network they consider, their conclusions may be valid, their more general conclusion that DF is limited in guaranteeing physical layer security seems premature.

We have not considered outer bounds for the relay broadcast channel with mutual secrecy in the model considered in chapter 5. Outer bounds for the four-node relay broadcast channel – no secrecy constraints imposed, and the paper refers to it as "broadcast relay channel" – were presented in Salehkalaibar, Ghabeli, Aref [124]. Outer bounds have also been presented by Dai, Yu, Ma [70].

[33] NNW can be considered to belong to the CF family of achievable schemes

Chapter 2

Sum-Capacity of Distributed MACs under Individual CSI

2.1 Overview

In wireless communications, a fading multiple access channel (MAC) is typically used to model the uplink communication. Conventional MACs assume a centralized system, where the transmission rate and power are chosen centrally for every fading vector realization. On the other hand, there is considerable interest in the performance of distributed multiple access systems, where the lack of global channel state information (CSI) demands novel communication strategies. We consider a block-fading MAC where each transmitter is aware only of its own link CSI, which we term as the *individual CSI MAC*. The receiver has access to the full CSI of all links. This model was recently introduced in the information theory literature, and naturally leads to a distributed access system with several applications. An important utility of interest for this model is known as the *power controlled adaptive sum-capacity*, whose evaluation is an open problem. This is the main subject of the current chapter.

We present the power-controlled adaptive sum-capacity of a wide class of popular fading MAC models. In particular, we characterize the sum-capacity when the statistics of the channel are identical across users. The proposed schemes also allow a low complexity successive cancellation decoding using rate-splitting. Furthermore, the optimal schemes are extended to situations in which each transmitter has additional finite-rate partial CSI on the link quality of others.

2.2 Introduction

In a multiple access channel (MAC) many transmitters communicate to a single receiver using a shared medium. With its natural applications in wireless communications, the so called fading MAC with additive white Gaussian noise (AWGN) is one of the popular MAC models. In here, the channel from each user to the receiver is modeled by a multiplicative fading channel. For most parts of this chapter, we consider fading MACs with AWGN.

In order to find the rate-tuples at which reliable communication is possible over the fading MAC model, it is important to make assumptions about the amount of channel knowledge available at the transmitters and the receiver. It is natural to assume that the receiver has access to the fading coefficients, by means of pilot-aided channel estimation. In other words, the receiver has full CSI. On the other hand, the same is not true about the transmitter. We consider a MAC model where each transmitter is fully aware of its own fading coefficient, but that of no other. We call this the *individual CSI MAC*. The model was introduced in Gamal and Kim [27], Hwang, Malkin, Gamal, Cioffi [125] for its practical utility. See [27, Chapter 23] for more details. Towards the latter sections of this chapter, we relax this assumption and equip the transmitter with finite-rate partial CSI of other links.

We consider a slow fading model as in [27], which is modeled by block fading: the fading coefficients remain constant for a block of channel uses over which the codewords last. This models the practical assumption of coding within the coherence time of a channel. The transmitters, thus, are not allowed to take advantage of the ergodic nature of the fading process during coding, but may employ adaptive power and rate allocations over blocks of channel uses. This particular situation is motivated by systems involving occasional (opportunistic) access to a shared medium, such as in a cognitive radio or a sensor network with a star topology. Here, multiple users wish to communicate their data to the receiver over the awarded time slot in a fair but distributed fashion. These systems may even lack the global user coordination information to schedule users. However, some limited coordination information can be made available or gleaned from the network. For example, the total number of active users participating in a given slot can be assumed as the common knowledge.

It is natural to look for within-block coding in these systems and demand that communication in each block be outage-free, while allowing for adaptively controlling the power and transmission-rates based on the available channel knowledge [27]. The word *outage-free* signifies that the chosen rate-vector across users is inside the instantiated MAC capacity region for that block. This notion will be made more precise later. The employed power-adaptation strategy should also respect the corresponding average transmit power constraint at each of the users.

There is considerable literature on multiaccess fading channels with instantaneous CSI. The Shannon capacity of a fading Gaussian MAC (GMAC) with CSI available only at the receiver is evaluated rigorously in Shamai and Wyner [126]. The optimal power control strategies to achieve capacity for the case of complete channel state information at the transmitters (CSIT) are given in Knopp and Humblet [127] and Tse and Hanly [128]. Coming to partial side information at the transmitters, Das and Narayan [129] gives the ergodic capacity region of a fading MAC under very general notions of CSI at the transmitters. These notions can be specialized to nearly all practical scenarios including individual transmitter CSI. Our work differs from Das and Narayan [129] in two ways. First, achieving the ergodic Shannon capacity region in [129] requires codewords which span a large number of fading realizations, whereas our system demands within block coding. These are fundamentally different problems. Secondly, the characterization of the capacity region in [129] is given as a complex optimization over power allocation functions, even numerical solutions are hard to compute. On the other hand, in the setup that we consider, the optimal power allocation and the sum-capacity can be explicitly determined in many interesting cases. Alternate notions of capacity motivated by different practical scenarios have also been investigated: delay-limited capacity for the fading MAC is dealt in Tse and Hanly [128], while Effros, Goldsmith, Liang [130] defines the notions of expected capacity and capacity with outage for information unstable single-user channels. Other related works which consider partial CSI in a fading MAC setup are Cemal and Steinberg [131], where non-causal CSI is considered, and its generalization and unification with causal CSI in Jafar [132]. The partial CSI models also have interesting connections to random access models, see Minero, Franceschetti, Tse [133] for a recent account. In another related work, Niesen, Erez, Shah, Wornell [134] considers rateless coding in a distributed MAC set-up. In the rateless coding setup, communication takes place in rounds of variable lengths. Each round goes on till the receiver figures out all the transmitted messages, and the end of a round is signaled by a feedback beacon signal. In contrast, our model has fixed slot-durations, and there is no feedback signal. However, drawing motivation from [134], we construct a rate-splitting strategy with successive cancellation decoder for our MAC setup too.

The model that we consider here, i.e. block fading MAC with individual CSI is applicable in several situations Gamal and Kim [27], Qin and Berry [135], Adireddy and Tong [136], Hwang, Seong, Cioffi [137]. In particular, a GMAC with binary fade values can model random access systems, where the non-zero fade value indicates the presence of a packet. The binary fading model was generalized in Hwang, Malkin, Gamal, Cioffi [125] to discrete memoryless MACs, with the state information of each link available only at the respective encoder, and at the decoder. An additional requirement that the probability of error remains small for every state realization (block) was also enforced. Clearly, the underlying assumption of a sufficiently large block-length allows the construction of near-capacity achieving coding strategies, with error-probabilities exponentially small in the blocklength. Averaging the possible communication rates over different state realizations will result in the *adaptive capacity region* (a term coined in [27]) in the presence of individual CSI. The major challenge here is in choosing the rate-vectors in a distributed fashion, without having access to the global CSI. [125] obtained a single letter characterization for the adaptive capacity region and specialized this to the GMAC setting. These results formulate the adaptive capacity region as an optimization over distributed outage-free rate allocation functions. However, the explicit evaluation of the optimal allocation is left as an open problem, while numerical techniques are used to demonstrate the results for two state channels.

In a GMAC, the users can also adapt their power in addition to the rates [27]. The maximal sum-throughput in this case is called the *power controlled adaptive sum-capacity* (see [27, Section 23.5.2]), which is the maximal empirical average of the sum-rates achieved in each block. In [125] and [27], finding the general power-controlled adaptive capacity region is posed as a convex problem, and the sum-capacity is numerically determined for two-state fading MACs using convex programming techniques. To emphasize the difficulty, evaluating the adaptive sum-capacity of the popular Rayleigh fading MAC model is mentioned as an open problem in [125]. We present an explicit solution to this problem in the current chapter (see also the conference versions [64], [61]), in terms of a closed-form water-filling formula, along with several other interesting results and extensions. Our main contributions are summarized as follows:

- We introduce a simple, distributed rate allocation policy called the 'alpha-midpoint' strategy for the Gaussian multiple-access block-fading channel with individual CSI (Section 2.4).

- The alpha-midpoint strategy achieves the power controlled adaptive sum-capacity under a water-filling power allocation when the channel statistics are identical across users (Theorem 5).

- We propose a low-complexity rate-splitting scheme that allows the alpha-midpoint strategy to be implemented through a novel successive single user decoding. This is surprising, given the lack of centralized CSI and coordination. The highlight here is a unconventional decoder which *peels* off layers of data (called virtual users) in a greedy fashion (Theorem 11).

- When the users are identical, i.e. they have the same channel statistics and average powers, we study the impact of additional finite-rate partial CSI of the other links. The power controlled adaptive sum-capacity is computed when the additional finite rate CSI of each link is generated using identical quantizers (Theorem 12).

- Our schemes do not need any coordination overhead in scheduling users, but can still achieve better performance than conventional centralized scheduling schemes like TDMA in many systems of interest. We present an example model with two senders and three receivers where a natural adaptation of our mid-point scheme shows superior sum-rates when compared to TDMA (Section 2.7).

The organization of the chapter is as follows. Section 2.3 introduces the system model and some notations, and also defines the notion of **power-controlled adaptive sum-capacity**, which is our utility of interest. In Section 2.4 we present an intuitive communication strategy utilizing the available individual CSI, called the *midpoint strategy* and then generalize it to the **alpha-midpoint strategy**. This scheme, under a water-filling power allocation, is shown to achieve the power controlled adaptive sum-capacity of a wide class of fading MACs. Section 2.5 proposes a rate-splitting scheme with low complexity successive cancellation decoding, which can achieve near optimal rates for the individual CSI MAC. Extensions of the sum-capacity results to the case where additional partial finite-rate side information on the other links is available at the transmitters is given in 2.6. We illustrate the superiority of the proposed schemes over conventional schemes by an example in Section 2.7. Section 2.8 concludes the chapter with a discussion of the results and possible extensions.

2.3 System Model and Definitions

Consider L users communicating with a single receiver. These users transmit real-valued signals X_i, encountering real-valued fades H_i. If Y is the value of the received signal at a (discrete) time instant we have

$$Y = \sum_{i=1}^{L} H_i X_i + Z, \tag{2.1}$$

where Z is an independent Gaussian noise process of unit variance. The fading space $\mathcal{H}_i$ of the i-th user is the set of values taken by H_i, and the joint fading space $\mathcal{H}$ is the set of values taken by the joint fading state $\bar{H} = (H_1, H_2, \cdots, H_L)$. We will adhere to this convention of representing vector quantities of user-wise parameters, like rate, power, channel state realization etc with an overbar symbol. The notation $\mathbb{E}[X]$ represents the expectation of random variable X. Small case letters are used for the realizations of a random variable.

Consider a slow-fading model, where each channel coefficient stays constant within a block and varies across blocks in an i.i.d fashion. While we demand *reliable* communication within each block, the utility of interest here is the average sum-throughput, or average sum-capacity, where the average is over different fading realizations or blocks. A more precise definition of our utility is given later in this section. In Gaussian channels, the rate-expressions usually take a logarithmic form, and all logarithms in this chapter are expressed to the base of 2.

We assume that the (stationary and ergodic) fading processes H_i are independent, and their distributions are known to all the transmitters and the receiver. In addition, we have *individual* CSIT, i.e. each transmitter knows its own channel fading coefficient H_i but that of no other. The receiver knows all the fading coefficients. The transmitters have individual average power constraints. i.e.

$$\mathbb{E}\left[P_i(H)\right] = \int_h P_i(h) d\Psi_i(h) \leq P_i^{avg}, \ 1 \leq i \leq L, \tag{2.2}$$

where $\Psi_i(\cdot)$ is the cumulative channel law (cdf) of user i. The users can adapt the rate (and power) according to their own channel conditions. Apart from the notation changes, our model and objectives are similar in spirit to the those presented in [27] (see Section 23.5). In fact, the terminology *power-controlled adaptive sum-capacity* is borrowed from [27], which we explicitly compute in the present chapter.

The adaptive nature of communication naturally leads to the following notion of a power-rate strategy.

Definition 1. *A power-rate strategy is a collection of mappings $(P_i, R_i) : \mathcal{H}_i \longmapsto \mathbb{R}^+ \times \mathbb{R}^+$; $i = 1, 2, \cdots, L$. Thus, in the fading state H_i, the i^{th} user expends power $P_i(H_i)$ and employs a codebook of rate $R_i(H_i)$.*

The component functions P_i and R_i in the above definition will be referred to power allocation and rate allocation strategies respectively.

Let $C_{MAC}(\bar{h}, \bar{P}(\bar{h}))$ denote the capacity region of a Gaussian multiple-access channel with fixed channel gains of $\bar{h} = h_1, \cdots, h_L$ and respective power allocations
$\bar{P}(\bar{h}) = (P_1(h_1), \cdots, P_L(h_L))$. We know that,

$$C_{MAC}(\bar{h}, \bar{P}(\bar{h})) = \left\{ \bar{R} \in \{\mathbb{R}^+\}^L : \forall S \subseteq \{1, 2, \cdots, L\}, \sum_{i \in S} R_i \leq \frac{1}{2} \log \left(1 + \sum_{i \in S} |h_i|^2 P_i(h_i) \right) \right\} \tag{2.3}$$

Definition 2. *We call a power-rate strategy as* feasible *if it satisfies the average power constraints for each user i.e.* $\forall i \in \{1, 2, \cdots, L\}, \quad \mathbb{E}_{H_i} P_i(H_i) \leq P_i^{avg}$.

Definition 3. *A power-rate strategy is termed as* outage-free *if it never results in outage i.e.* $\forall \bar{h} \in \mathcal{H}, (R_1(h_1), \cdots, R_L(h_L)) \in C_{MAC}(\bar{h}, \bar{P}(\bar{h}))$.

In a practical system, an outage-free strategy will ensure a small error probability for large block lengths. A large block length for which fading remains constant will require a large channel coherence time in turn. However, such an outage-free requirement is not only common in works on block-fading partial CSI models [27], [125], but also used in many practical wireless systems. Let Θ_{MAC} denote the collection of all feasible power-rate strategies which are outage-free. Let us now specialize the definitions to the case of identical channel statistics, i.e. the cdf of each user is $\Psi(h)$. For any strategy $\theta \in \Theta_{MAC}$, the throughput is

$$\begin{aligned} T_\theta = \sum_{i=1}^{L} \mathbb{E} R_i^\theta(H_i) &= \sum_{i=1}^{L} \int_h R_i^\theta(h)\, d\Psi(h) \\ &= \int_h d\Psi(h) \left(\sum_{i=1}^{L} R_i^\theta(h) \right), \end{aligned} \tag{2.4}$$

where the superscript θ is used to identify the feasible power-rate strategy employed. i.e. $R_i^\theta(h)$ is the rate allocated to user i while observing fading coefficient h. The corresponding transmit power is denoted as $P_i^\theta(h)$.

Definition 4. *The* ***power controlled adaptive sum-capacity*** *is the maximum (average) throughput achievable, i.e.* $C_{sum}(\Psi) = \max_{\theta \in \Theta_{MAC}} T_\theta$.

One of the main results of the chapter is the computation of the power-controlled adaptive sum-capacity for several popular fading models. In the special case of a single user channel ($L = 1$), the adaptive sum-capacity is well known, as it becomes a full CSI model. We denote the single user-capacity with an average power constraint of P^a as $C_1(\Psi, P^a)$, which can be evaluated using a water-filling formula (see Tse and Viswanath [138] for a recent account),

$$C_1(\Psi, P^a) = \frac{1}{2} \int d\Psi(h) \log(1 + |h|^2 P^*(h)), \tag{2.5}$$

where

$$P^*(h) = \left(\frac{1}{\lambda} - \frac{1}{|h|^2} \right)^+ \text{ and } \int d\Psi(h) P^*(h) = P^a. \tag{2.6}$$

The single user water-filling formula is considered to be in *closed form* for all practical purposes. Our results for the MAC with distributed CSI also take the form of similar water-filling formulas. Thus, we will re-use the notation $C_1(\Psi, P^a)$ several times in this chapter. Let us now focus on computing the quantity $C_{sum}(\Psi)$ for $L > 1$.

2.4 Power Controlled Adaptive Sum-capacity ($C_{sum}(\Psi)$)

Consider a $L-$user distributed MAC with individual CSI, and identical link statistics across users. The main result of this section is to compute the power-controlled adaptive sum-capacity $C_{sum}(\Psi)$ for arbitrary $\Psi(\cdot)$. We first state the result and then explain its structure and implications, before providing the proof.

Theorem 5. *Given independent and identically distributed channels according to the c.d.f $\Psi(h)$,*

$$C_{sum}(\Psi) = C_1\left(\Psi, \sum_{i=1}^{L} P_i^{avg}\right). \tag{2.7}$$

Before presenting the proof, which is done in Section 2.4.2, it is instructive to study the structure of the result. The result states that $C_{sum}(\Psi)$ is same as the capacity of a single user channel with cdf $\Psi(\cdot)$ and average power $\sum_{i=1}^{L} P_i^{avg}$. It has an element of surprise in the first look, as if there is some degeneracy in the problem statement. While this is not the case, the single user result essentially comes from the fact that communication has to be outage-free in every block. The result is re-stating that the worst joint distribution of the MAC fading-states is a highly correlated one, in which the same fading value is observed across users.

In order to achieve $C_{sum}(\Psi)$, we propose a distributed strategy called the **alpha-midpoint strategy**. The ideas behind this scheme can be clearly motivated by considering a special case of the GMAC model, one with *identical users.*

2.4.1 Identical Users and Mid-point Strategy

Assume that the users are identical, i.e. users have identical channel statistics and the same average powers. We will deal with unequal average powers in the next subsection. Under the individual CSI model, consider a scheme where user i, on observing channel state h_i in a block, imagines that every other channel state is also h_i. Under this symmetric MAC assumption, user i can choose an operational rate in a distributed fashion. A natural question now is whether such a distributed choice at each user will lead to a valid rate-tuple within the actual MAC capacity region for that block. It turns out that a careful choice of rates can achieve this objective, which we call the *mid-point strategy.* The mid-point rate allocation is explained now.

Consider a block in which the fading vector is $h_1, \cdots, h_L$, and let the respective powers be $P_1(h_1), \cdots, P_L(h_L)$. We will optimize over the power choice later. Now, the mid-point rate allocation for user i is,

$$R_i^{mid}(h_i) = \frac{1}{2L}\log\left(1 + L|h_i|^2 P_i(h_i)\right). \tag{2.8}$$

Lemma 6. *The midpoint rate strategy is outage-free, i.e. $\forall \bar{h} \quad \bar{R}_i^{mid} \in C_{MAC}(\bar{h}, \bar{P})$.*

Proof. The lemma follows directly from the concavity of the logarithm function, i.e. $\forall S \subset$

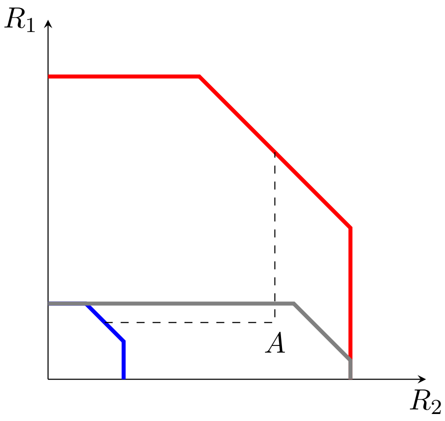

Figure 2.1: The users 1 and 2 construct the innermost and outermost MAC capacity regions respectively. The intermediate pentagon is the instantiated (actual) MAC region and A denotes the operating point.

$\{1, 2, \cdots, L\}$,

$$\begin{aligned}\sum_{i \in S} R_i^{mid}(h_i) &= \frac{1}{2L} \sum_{i \in S} \log\left(1 + L|h_i|^2 P_i(h_i)\right) \\ &\leq \frac{1}{2|S|} \sum_{i \in S} \log\left(1 + |S||h_i|^2 P_i(h_i)\right) \\ &\leq \frac{1}{2} \log\left(1 + \sum_{i \in S} |h_i|^2 P_i(h_i)\right).\end{aligned}$$

□

Suppose we choose a power allocation scheme where each user water-fills over its own channel gain, i.e.

$$P_i(h_i) = \left(\frac{1}{\lambda_i} - \frac{1}{|h_i|^2}\right)^+, \tag{2.9}$$

where λ_i is chosen such that $\mathbb{E}_{H_i} P_i(H_i) = P_i^{avg}$. While the motivation behind this choice may not be immediately clear, this is indeed an optimal choice for identical users, a special case of the result proved in the next subsection. The optimal sum-rate for a two user MAC with individual CSI and identical users is shown in Figure 2.2. For comparison, we also show the sum-capacity in the presence of full CSIT. It is well known that the optimal scheme in the presence of complete CSIT is opportunistic TDMA (O-TDMA) where only the *better* user transmits [127]. The throughputs of the two models are plotted for identical normalized Rayleigh fading links, and observations corrupted with AWGN of unit gain. The users are also assumed to have the same average powers.

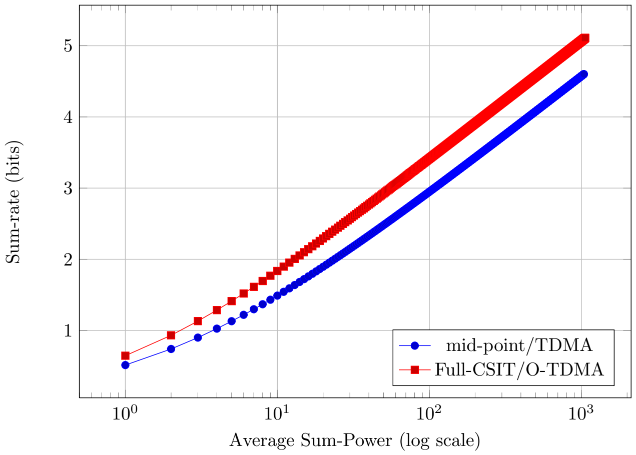

Figure 2.2: Midpoint strategy vs the full CSI rate

2.4.2 Unequal Average Powers (Proof of Theorem 5)

We now prove Theorem 5 in two steps. First, we construct an upperbound to the power controlled adaptive sum capacity $C_{sum}(\Psi)$. The second step generalizes the mid-point strategy to construct an achievable scheme which meets the proposed upper-bound. **An Upperbound:** We will convert our MAC with individual CSI to an equivalent single user channel for evaluating an upper-bound. To this end, consider a single link with cdf $\Psi(h)$. For a non-negative variable $\tilde{P}$, let $\Theta_s(\tilde{P})$ be the collection of all single-user power allocations $P_s(h), \forall h$ over this link such that

$$\int P_s(h) d\Psi(h) = \tilde{P}. \tag{2.10}$$

Let $P_{sum} = \sum_{i=1}^{L} P_i^{avg}$, and also recall the definition of throughput in (2.4).

Lemma 7. *The throughput T_θ obeys,*

$$T_\theta \leq C_1(\Psi, P_{sum}), \forall \theta \in \Theta_{MAC}.$$

Proof.

$$T_\theta \overset{(a)}{\leq} \frac{1}{2} \int_h d\Psi(h) \log \left(1 + |h|^2 \sum_{i=1}^{L} P_i^\theta(h) \right) \tag{2.11}$$

$$\leq \max_{\Theta_s(P_{sum})} \frac{1}{2} \int d\Psi(h) \log \left(1 + |h|^2 P_s(h) \right). \tag{2.12}$$

Here (a) follows from (2.4), by applying the sum-rate upper bound on a MAC with received signal power $\sum_i |h|^2 P_i^\theta(h)$. The second inequality results from relaxing the individual power constraints to a single average sum-power constraint.

It is clear that water-filling of the inverse fading gains is the optimal strategy in a point to point fading channel under an average power constraint. Thus the last expression above is indeed $C_1(\Psi, P_{sum})$. □

Alpha-midpoint strategy: Our achievable scheme is named as **alpha-midpoint strategy**. As the name implies, this is a generalization of the midpoint scheme that we introduced in the previous sub-section. Let $\bar{\alpha}$ be a vector of non-negative values with $\sum_i \alpha_i = 1$. In alpha-midpoint strategy, the rate chosen by user i while encountering a fading coefficient of h_i is,

$$R_i^{\bar{\alpha}}(h_i) = \alpha_i \frac{1}{2} \log\left(1 + |h_i|^2 \frac{P_i(h_i)}{\alpha_i}\right), \tag{2.13}$$

where $P_i(h_i)$ is the transmitted power, chosen such that

$$\int P_i(h) d\Psi(h) = P_i^{avg}.$$

Lemma 8. *The alpha-midpoint strategy is outage-free.*

Proof. For any $S \subseteq \{1, 2, \cdots, L\}$,

$$\begin{aligned} \sum_{i \in S} R_i^{\bar{\alpha}}(h_i) &= \sum_{i \in S} \alpha_i \frac{1}{2} \log\left(1 + |h_i|^2 \frac{P_i(h_i)}{\alpha_i}\right) && (2.14) \\ &\leq \frac{1}{2} \log\left(1 + \sum_{i \in S} |h_i|^2 P_i(h_i)\right), && (2.15) \end{aligned}$$

by concavity of the logarithm. Clearly the chosen rate-tuple across users is within $C_{MAC}(\bar{h}, \bar{P}(\bar{h}))$ for every block, ensuring that there is no outage. □

We now show the optimality of alpha-midpoint schemes.

Lemma 9.

$$\max_{\theta \in \Theta_{MAC}} T_\theta = C_1(\Psi, P_{sum}).$$

Proof. We will specialize our alpha-midpoint strategy to achieve $C_1(\Psi, P_{sum})$. To this end, choose for $1 \leq i \leq L$,

$$\alpha_i = \frac{P_i^{avg}}{\sum_{i=1}^{L} P_i^{avg}} \text{ and } P_i(h) = \alpha_i P^*(h), \tag{2.16}$$

where $P^*(h)$ is given in (2.6), with P_a replaced by P_{sum}. Notice that,

$$\begin{aligned} \int P_i(h) d\Psi(h) &= \alpha_i \int P^*(h) d\Psi(h) \\ &= \alpha_i P_{sum} \\ &= P_i^{avg}. \end{aligned}$$

Furthermore, by (2.13)

$$\begin{aligned}
\sum_{i=1}^{L} \int R_i^{\bar{\alpha}}(h_i) d\Psi(h_i) &= \sum_{i=1}^{L} \frac{\alpha_i}{2} \int \log\left(1 + |h|^2 \frac{P_i(h_i)}{\alpha_i}\right) d\Psi(h) \\
&= \sum_{i=1}^{L} \frac{\alpha_i}{2} \int \log(1 + |h|^2 P^*(h)) d\Psi(h) \\
&= \frac{1}{2} \int \log(1 + |h|^2 P^*(h)) d\Psi(h) \left(\sum_{i=1}^{L} \alpha_i\right) \\
&= C_1(\Psi, P_{sum}).
\end{aligned}$$

This proves Lemma 9, thus also completing the proof of Theorem 5. □

We have shown that the alpha mid-point schemes achieve the power controlled sum-capacity when the channel statistics are identical across users. Notice that the proposed scheme needs little coordination information to schedule users other than a modest frame-synchronization.

2.5 Rate Splitting and Successive Decoding

The alpha-midpoint strategy proposed in this chapter has a fairly simple structure. However, notice that the previous section employed a joint decoder to recover the input data. In this section, we show that these strategies can be implemented by low complexity successive decoding architectures. To this end, we present an asymptotically optimal rate-splitting strategy that replaces the joint decoder with LN_v successive single-user decoders, where N_v is a parameter, signifying the number of layers per user.

Our approach is motivated by the work of [134], where the goal is to demonstrate low-complexity schemes closely approximating the achievable rates. There are some crucial differences between the current setup and [134]. The latter considers a rateless scheme, where a communication round lasts till every participating user gets decoded at the receiver. Thus, the round duration is determined by the number of active participants, with variable length codes required in different rounds of communication. Each round is terminated by a feedback link from the receiver, which announces the next round via a beacon. In contrast, our scenario requires that the communication occur within a fixed block or time slot, and there is no assumption of any feedback link or beacon. Furthermore, though we employ a rate-splitting encoder which converts each user to a set of *virtual users* as in [134], our decoding strategy is highly unconventional. More specifically, the decoder peels off virtual users in a greedy fashion without any pre-defined order, and surprisingly, manages to decode all users when the distributed rate allocations are based on the alpha-midpoint strategy. To the author's knowledge, such a novel rate-splitting decoder does not appear elsewhere in literature.

We will first construct rate-splitting schemes for *identical users*, i.e. the users have the same average power and identical channel statistics. Extensions to arbitrary average powers is done in a separate subsection. We will write the received signal power for user i, i.e. $P_i|h_i|^2$ as simply γ_i, throughout this section. Assume that the users have different (received) powers $\gamma_1, \gamma_2, \cdots, \gamma_L$. For simplicity, we will assume that the additive noise is of unit variance. The values of γ_i may change with each block of communication depending on the individual fading conditions. Each user is *unaware* of the fade values as well as the transmit powers of the rest of the users and, consequently, the interference they may cause.

The encoding and decoding are done as: each user splits itself into N_v virtual users and apportions its power, perhaps unequally, among these users. Each user can be visualized as a 'stack' of virtual users. More specifically, user i, having received power γ_i, splits its data stream by allotting a power/rate pair (γ_i^l, r_i^l) to the l^{th} virtual user, such that $\sum_{l=1}^{N_v} \gamma_i^l = \gamma_i$. We will index the virtual users as (i,l), where $1 \leq i \leq L$ and $1 \leq l \leq N_v$. For decoding, a successive cancellation based single-user decoder which decodes one of the virtual users, treating all other virtual users which are yet to be decoded as Gaussian noise is employed, see [139] for more details of rate-splitting and multiple access.

To determine the transmission-rate, transmitter i assumes that all other users are also at (received) power γ_i and imagines identical power/rate splitting strategies across all users. It then chooses the rates r_i^l by considering all the other virtual users in the same and lower layers as interference, i.e.,

$$r_i^l = \frac{1}{2} \log \left(1 + \frac{\gamma_i^l}{1 + (L-1)\gamma_i^l + L \sum_{j=1}^{l-1} \gamma_i^j} \right). \tag{2.17}$$

Let us denote $(1 + (L-1)\gamma_i^l + L \sum_{j=1}^{l-1} \gamma_i^j)$ as EI_i^l, i.e. the *estimated* interference for virtual user (i,l). However, in reality, the interference encountered from the undecoded virtual users while decoding (i,l) is substantially different from EI_i^l. We call the perceived interference as the *actual interference*, and represent it by AI_i^l. The pre-requisite for decoding the virtual user (i,l) is that

$$AI_i^l \leq EI_i^l.$$

A layer by layer decoding in the order of index (i,l) cannot always guarantee the above condition, as the virtual users are not chosen according to the global channel conditions. Surprisingly, it turns out that this lack of knowledge can be compensated by not strictly adhering to an ordered layer by layer decoding. In particular, the receiver retains the freedom to decode the topmost hitherto undecoded layer of *any* transmitter, irrespective of the number of layers which were already decoded. It is, in fact, this freedom that allows the transmitters to choose the virtual rates without knowledge of interference from the other users. For the simplicity of exposition, let us assume that the blocklengths are arbitrarily large.

Lemma 10. *Assuming layer-wise rate allocation as per (2.17), it is always possible to find a virtual user with the actual interference lower than the estimated interference.*

Proof. We prove this by induction. Assume that layers (virtual users) above l_k have been decoded for the k^{th} transmitter. Choose:

$$\kappa = \arg\max_k \sum_{j=1}^{l_k} \gamma_k^j.$$

For user κ, the remaining interference for decoding layer l_κ is

$$\begin{aligned}
1 + \sum_{j=1}^{l_\kappa - 1} \gamma_\kappa^j + \sum_{k\neq\kappa} \sum_{j=1}^{l_k} \gamma_k^j \\
&= 1 + \sum_{k=1}^{L} \sum_{j=1}^{l_k} \gamma_k^j - \gamma_\kappa^{l_\kappa} \\
&\leq 1 + L \sum_{j=1}^{l_\kappa} \gamma_\kappa^j - \gamma_\kappa^{l_\kappa} \\
&= 1 + \sum_{j=1}^{l_\kappa - 1} \gamma_\kappa^j + (L-1) \sum_{j=1}^{l_\kappa} \gamma_\kappa^j .
\end{aligned}$$

The inequality follows directly from the choice of κ. Thus the actual interference is less than the expected interference, and this virtual user can be correctly decoded. In other words, the user with the 'best' received SNR can always be chosen for decoding. □

Theorem 11. *As $N_v \to \infty$ and $\gamma_k^l \to 0, \forall k, l$, the rate achieved by each user approaches the respective midpoint rate given in* (2.8)*.*

Proof. Using (2.17), we have

$$R_i = \sum_{j=1}^{N_v} \frac{1}{2} \log \left(1 + \frac{\gamma_i^l}{1 + (L-1)\gamma_i^l + L \sum_{j=1}^{l-1} \gamma_i^j} \right).$$

Under the given conditions, we can use the same method as in Lemma 1 of [134] to show that:

$$\begin{aligned}
\lim_{N_v \to \infty} R_i &= \frac{1}{2} \lim_{N_v \to \infty} \sum_{l=1}^{N_v} \frac{\gamma_i^l}{1 + (L-1)\gamma_i^l + L \sum_{j=1}^{l-1} \gamma_i^j} \\
&= \frac{1}{2} \int_0^{\gamma_i} \frac{dy}{1 + Ly} = \frac{1}{2L} \log\left(1 + L\gamma_i\right).
\end{aligned}$$

□

Computational results in [134] show that only a moderate number of virtual users N_v are sufficient to yield good performance. For completeness, the increase in achievable-rate with respect to the number of virtual layers is shown in Figure 2.3.

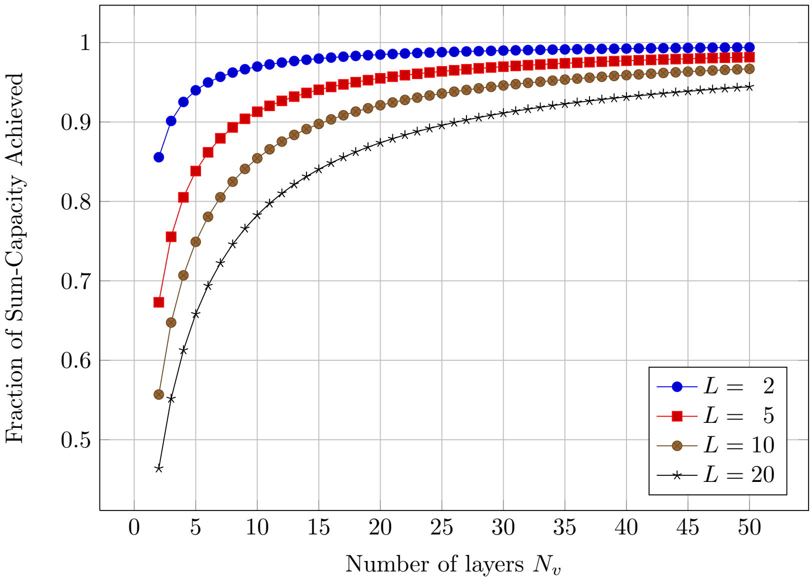

Figure 2.3: The fractional sum-capacity achieved Vs N_v for L users

2.5.1 Unequal Average Powers

We will construct two levels of splitting in the presence of unequal average powers. In particular, we first split user k into N_k pseudo users, in such a way that each pseudo user has an identical[1] average transmit power constraint of P_v, irrespective of the user index k. Thus,

$$N_k P_v = P_k^{avg}, 1 \leq k \leq L.$$

Evaluating the maximal average rate for the $L' = \sum_{k=1}^{L} N_k$ virtual users under the midpoint strategy of (2.8) will also yield $C_1(\Psi, P_{sum})$. To see this, notice that the total-rate obtained by the N_k layers of user k is

$$N_k \frac{1}{L'} \int \frac{1}{2} \log(1 + |h|^2 P^*(h)) d\Psi(h),$$

where $P^*(h)$ is the single-user water-filling allocation with an average power of $L' P_v = \sum_{k=1}^{L} P_i^{avg} = P_{sum}$. Notice that N_k/L' is nothing but the α_k in (2.16), proving that the above strategy can achieve the same rates as the alpha-midpoint scheme. Furthermore, since

[1] sometimes the transmit-power of users may not be commensurate, however we can choose a *slightly* lower power level for some of the users, with negligible loss of performance.

the midpoint rates are achievable by single user decoding techniques [64], [61] alpha-midpoint rates can also be achieved by low complexity successive cancellation based decoding schemes.

2.6 Finite-rate CSI on Other Links

Up to this point, we have assumed only individual CSI. Let us now study the effect of additional partial information about the other links. To keep things simple, we consider *identical users*, i.e. each user has fading cdf $\Psi(h)$ and an average power constraint of P^{avg}. The first case that we consider is 1 bit of additional partial CSI, i.e. each transmitter gets one bit of information from every other link, in addition to its own individual CSI. As usual, the individual fading components are assumed to be independently chosen across users. The additional CSI made available to others is only a function of the individual fading coefficients. Thus, the model captures situations where the extra bit is obtained through transmitter cooperation or cribbing. It is crucial that the receiver has no say on the partial CSI. If the receiver decides the conveyed bit, then the throughput is same as that of full CSI, as in Knopp and Humblet [127].

The partial CSI contains link quality information: let us assume it to be chosen from the set $\{G, B\}$, where G stands for good, and B for bad. A natural separation between G and B is a link gain threshold. In particular, the partial CSI bit $\hat{h}_k$ of transmitter k is

$$\hat{h}_k = \begin{cases} G \text{ if } |h_k| \geq h_T \\ B \text{ otherwise,} \end{cases} \tag{2.18}$$

for some fixed positive threshold h_T. Thus, $\hat{h}_k = G\mathbb{1}_{\{|h_k| \geq h_T\}} + B\mathbb{1}_{\{|h_k| < h_T\}}$, where $\mathbb{1}_{\{\cdot\}}$ is the indicator function. By slight abuse of notation, we will say that link j is in state G (and call it good user), and denote the probability of that event by $\mu(G)$. Using the same token, $1 - \mu(G) = \mu(B)$. Let C_{PSI} be the maximum attainable throughput with 1 bit additional CSI on each of the other links, along with individual CSI.

Theorem 12. *For L identical users,*

$$C_{PSI} = C_1\left(\Phi, LP^{avg}\right), \tag{2.19}$$

where the cdf $\Phi(\cdot)$ is such that,

$$d\Phi(h) = d\Psi(h)\left([\mu(B)]^{L-1}\mathbb{1}_{\{h \in B\}} + (1+\zeta)\mathbb{1}_{\{h \in G\}}\right),$$

and the parameter

$$\zeta = \sum_{m=1}^{L-1} \binom{L-1}{m} [\mu(B)]^m [\mu(G)]^{L-1-m} \frac{m}{L-m}. \tag{2.20}$$

Proof. Recall the definition of $C_1(\cdot,\cdot)$ given in (2.5)–(2.6). We explain the proof for $L = 2$, which contains all the essential features. The proof is relegated to appendix .2. □

It is instructive to compare the advantages of 1 bit of extra CSI, which we demonstrate for a two user *identical* Rayleigh fading links of unit second moment, see Figure 2.4. The threshold value h_T for 1−bit CSI was taken as unity.

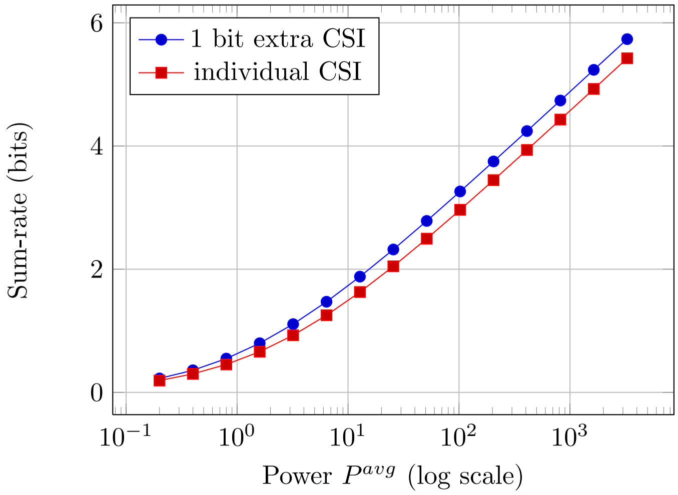

Figure 2.4: Sum-rate improvement by additional CSI

One immediate question is the sensitivity of the results with respect to the fading threshold. Numerical observations suggest that it is not that crucial and the results are robust. In fact, taking the median of the fading distribution seems to be natural choice for many models. Even for moderate power-levels (a few 10s of dB) the performance difference of the median from the the best choice of the threshold is not noticeable in Figure 2.4.

2.6.1 Multiple bits of CSI

We can extend the above 1−bit result to multiple bits of CSI, under the assumption that each user employs an identical quantization scheme to generate partial CSI. More precisely, $n-1$ threshold values $T_1, T_2, \cdots, T_{n-1}$ for the fading gains can be used define the $\log_2 n$ bits of partial CSI about each link. For $0 \le m \le n-1$, define

$$U_{m+1} = \{i : |h_i| \in [T_m, T_{m+1})\}$$

as the set of users belonging to the same fading partition, i.e. they generate identical partial CSI bits. Here we assumed $T_0 = 0$ and $T_n = \infty$. In other words, all users experiencing a absolute fading-value in the threshold bracket $[T_m, T_{m+1}), 0 \le m \le n-1$ will form group U_{m+1}.

The result for one bit partial CSI can now be extended to this setup. In particular, an optimal strategy achieving the maximum average sum-rate schedules the group $m^* = \max\{m : |U_m| > 0\}$ for transmission in each block[2]. All other users remain silent. Now, each user in the group U_m^* transmits at its midpoint rate, by taking $|U_{m^*}|$ users into account, employing the rate allocation in (2.8). It is also clear that we can extend the low complexity rate-splitting technique introduced earlier to the partial CSI setting of this section.

2.7 Fading $W-$Channel

The purpose of this section is to demonstrate that the proposed distributed technique like the alpha-midpoint strategy can perform even better than centralized schemes like TDMA. To exemplify this clearly, we construct a model where significant advantages over TDMA can be demonstrated. The model that we introduce is termed as the *fading $W-$ channel*, it is an adaptation of the *$AWGN$ RAC* network in [133] which has two senders and three receivers. Each of the senders has an independent data stream. One of the three receivers, we call it the central receiver, is interested in the messages from both the senders, while each remaining receiver has a point-to-point link to its corresponding transmitter. The dependencies are pictorially shown in Figure 2.5.

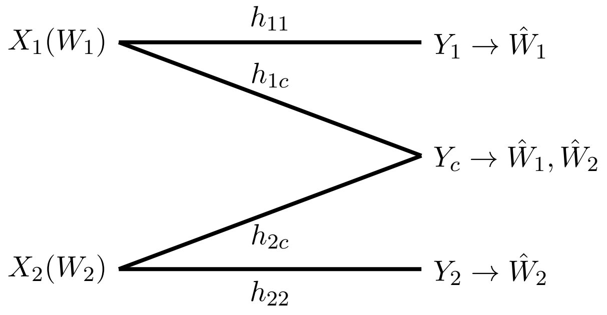

Figure 2.5: Fading $W-$channel with 2 senders and 3-receivers, $W_i, i = 1, 2$ are the messages

The received values are

$$Y_i = H_{ii}X_i + Z_i,\ i = 1, 2$$

and

$$Y_c = H_{1c}X_1 + H_{2c}X_2 + Z_c,$$

where Z_1, Z_2 and Z_c are normalized independent Gaussian random variables. The links H_{ii} and H_{ic} are considered to be block-fading, with full CSIR at each of the receivers. The

[2] $|\cdot|$ denotes cardinality here

coefficients H_{1c} and H_{2c} are independent and identical, the same applies to links H_{11} and H_{22}. We further assume an individual CSIT model, where each transmitter knows the fading coefficients of the links which originate from it. The main differences between our model and the $2-$Sender AWGN RAC in Minero, Franceschetti, Tse [133] are listed below.

- In an AWGN RAC there are two additional private data streams, one between each transmitter-receiver pair $(X_i, Y_i); i = 1, 2$.

- The AWGN RAC model only considers static links whereas we consider fading links with individual CSIT.

For simplicity, we do not consider power-control, and the average transmit power is limited to P in each block of transmissions. A similar MAC model without power control is considered in the context of *adaptive capacity region* in [27]. The objective now is to maximize the sum-rate from the transmitters, ensuring that there is no outage at any of the receivers. We term the optimal sum-rate as the adaptive sum-capacity of the fading $W-$channel and denote it by C^W_{sum}.

Let us first consider TDMA. By the symmetry of the observed channel characteristics, it is sufficient to consider a time-sharing factor of $\frac{1}{2}$. The data-communication problem then decouples into two separate links, where the rate allocation of user i is,

$$R_i^{TDMA}(h_{ii}, h_{ic}) = \frac{1}{4} \log \left(1 + \min\{h_{ii}^2, h_{ic}^2\}\, 2P\right). \tag{2.21}$$

In (2.21), we have used the fact that an active user in TDMA can employ a power of $2P$, while transmitting for half the frame. It turns out that the TDMA sum-rate is inferior to strategies which adapt the mid-point scheme. In particular, let transmitter i choose the mid-point rate to the central receiver, whenever this chosen rate is below the link capacity of h_{ii}. The rate allocation is

$$R_i^{MP}(h_{ii}, h_{ic}) = \min\{\frac{1}{4} \log \left(1 + h_{ic}^2 2P\right), \frac{1}{2} \log(1 + h_{ii}^2 P)\}. \tag{2.22}$$

Clearly, the central receiver will succeed in decoding the data under the mid-point scheme. The point-to-point links will also not face any outage since the respective transmission rates are below their capacity. It turns out that the average sum-rate of this scheme can be strictly better than TDMA. We will state this precisely in the following lemma.

Lemma 13.

$$\mathbb{E}\left[R_i^{TDMA}(H_{ii}, H_{ic})\right] \leq \mathbb{E}\left[R_i^{MP}(H_{ii}, H_{ic})\right].$$

Proof. Rewriting (2.21) as,

$$R_i^{TDMA}(h_{ii}, h_{ic}) = \min\{\frac{1}{4}\log\left(1 + h_{ic}^2 2P\right), \frac{1}{4}\log(1 + h_{ii}^2 2P)\} \tag{2.23}$$

$$\leq \min\{\frac{1}{4}\log\left(1 + h_{ic}^2 2P\right), \frac{1}{2}\log(1 + h_{ii}^2 P)\} \tag{2.24}$$

$$= R_i^{MP}(h_{ii}, h_{ic}), \tag{2.25}$$

where the inequality uses the concavity of logarithm. □

By examining (2.24), the mid-point scheme is strictly superior to TDMA whenever the maximum possible value of h_{ic}^2 is above the minimum possible value of h_{ii}^2. This is a very modest requirement, and our schemes will beat TDMA in most systems of interest. It also turns out that the allocation in (2.22) is in fact optimal, i.e. it achieves the adaptive sum-capacity C_{sum}^W of the fading $W-$channel without powercontrol.

Theorem 14.

$$C_{sum}^W = 2\,\mathbb{E}\left[R_1^{MP}(H_{11}, H_{1c})\right]$$

Proof. The proof employs arguments similar to that of Theorem 5. This is outlined in Appendix .1 for completeness. □

Let us numerically compare the performance of the two schemes. To clearly illustrate the benefits of the scheme, let us assume that the point-to-point links are fixed at $H_{11} = H_{12} = \beta$ almost surely. The MAC fading coefficients H_{1c} and H_{2c} are considered to be normalized Rayleigh distributed. Figure 2.6 compares the sum-rate of TDMA and mid-point schemes for $P = 10$ and $P = 30$, while β is varied. It is seen that the sum-rate saturates for the mid-point scheme for even moderate values of β. The results of this section can be extended to the case

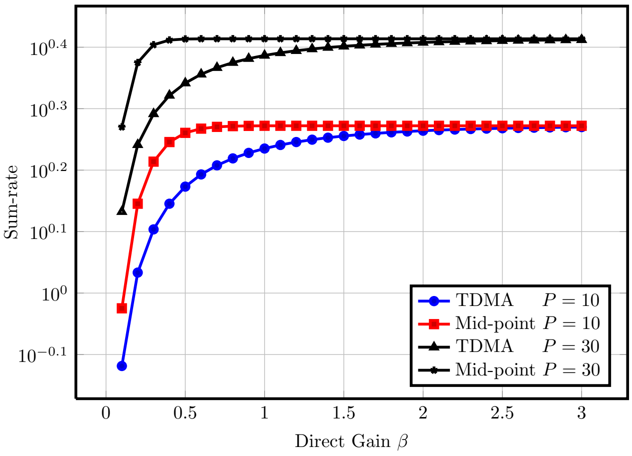

Figure 2.6: Comparing the sum-rates of TDMA and Mid-point for the $W-$Channel

where there is power-control. While this is numerically straightforward, it has less analytical appeal, and we do not pursue it here. More than just the sum-capacity, the $W-$channel result indicates that the proposed mid-point coding schemes will find applications in several contexts, including relaying, interference channels and more complex networks.

2.8 Conclusion

We have presented throughput optimal outage-free communication schemes for a block-fading MAC with identical channel statistics, and distributed (individual) CSI. While the assumption of symmetric channel statistics is very relevant in several situations, we are currently extending our work to asymmetric channel statistics. In the identical channels case, the proposed schemes can be extended to MIMO channels, and also to multi-path channels. The results can further be adapted to the case of limited CSI of each link available at the receivers, and the individual part of these made available at the respective transmitters. Whereas the current chapter focused on the adaptive sum-capacity, this should be considered as a step towards computing the full capacity region, a possible future work. We believe that the proposed strategies have several ingredients which make them suitable for many circumstances. In particular, the results for the fading $W-$channel in 2.7 clearly show the versatility. Extensions of the coding schemes for the individual CSI models to relay channels, interference channels etc are of considerable interest.

In the case of additional partial finite-rate CSI, we computed the sum-capacity under identical users and symmetric CSI. It will be interesting to relax this assumption and compare the performance, in terms of non-identical users or asymmetric CSI. While we limited our treatment here to the block-wise encoding case while ensuring no outage, it will also be interesting to see whether these results can be bridged to the ergodic capacity-achieving schemes.

.1 Proof of Theorem 14

Since user i is only aware of h_{ii} and h_{ic}, its rate allocation $R_i(h_{ii}, h_{ic})$ only depends on these two realizations and the statistics of the other two links. The expected sum-rate is now

$$\begin{aligned}
\mathbb{E}\left[R_1(H_{11}, H_{1c}) + R_2(H_{22}, H_{2c})\right] &= \mathbb{E}R_1(H_{11}, H_{1c}) + \mathbb{E}R_2(H_{11}, H_{1c}) && (26)\\
&= \mathbb{E}R_1(H_{11}, H_{1c}) + R_2(H_{11}, H_{1c}) && (27)\\
&\leq \mathbb{E}\min\{\frac{1}{2}\log(1 + 2H_{1c}^2 P), \log(1 + H_{11}^2 P)\} && (28)\\
&= 2\mathbb{E}R^{MP}\left(R_1(H_{11}, H_{1c})\right) && (29)
\end{aligned}$$

In above, the first equality used the identical distribution of $H_{ii}, i = 1, 2$, and similarly $H_{ic}, i = 1, 2$. The inequality has two components. The first term in the minimum is the sum-rate bound for a MAC with both links H_{1c}. The second term bounds the total rate of two independent point-to-point links, each having a fading value of H_{11}. The last equality is immediate from the mid-point rate allocation of (2.22).

.2 Proof of Theorem 12

We will show the proof for a 2 user system for simplicity. Let $\hat{h}_i$ denote the CSI communicated from user i to all others. User 1 employs a power of $P_1(h_1, \hat{h}_2)$ and user 2 spends $P_2(\hat{h}_1, h_2)$. Let $R_1(h_1, \hat{h}_2)$ and $R_2(\hat{h}_1, h_2)$ be the respective rates chosen. We can express the average sum-rate as,

$$\begin{aligned} R_1 + R_2 = & \int_{G\times G} \Big(R_1(h_1,\hat{h}_2) + R_2(\hat{h}_1,h_2)\Big)\, d\Psi(h_1,h_2) \\ & + \int_{B\times B} \Big(R_1(h_1,\hat{h}_2) + R_2(\hat{h}_1,h_2)\Big)\, d\Psi(h_1,h_2) \\ & + \int_{B\times G} \Big(R_1(h_1,\hat{h}_2) + R_2(\hat{h}_1,h_2)\Big)\, d\Psi(h_1,h_2) \\ & + \int_{G\times B} \Big(R_1(h_1,\hat{h}_2) + R_2(\hat{h}_1,h_2)\Big)\, d\Psi(h_1,h_2). \end{aligned} \tag{30}$$

Consider the first term in the summation of the right hand side. By suitably integrating, it can be written as a single integral,

$$\mu(G)\int_G (R_1(h,G) + R_2(G,h))d\Psi(h) \le \frac{\mu(G)}{2}\int_G \log\left(1 + h^2(P_1(h,G) + P_2(G,h))\right) d\Psi(h), \tag{31}$$

which is the sum-rate bound of the corresponding MAC. Similarly, for the second term,

$$\mu(B)\int_B (R_1(h,B) + R_2(B,h))d\Psi(h) \le \frac{\mu(B)}{2}\int_B \log\left(1 + h^2(P_1(h,B) + P_2(B,h))\right) d\Psi(h). \tag{32}$$

As for the third and fourth terms, the information on who has the better channel is readily available to both parties here. Let us now consider only those channel states $(h_1, h_2) \in \{(G \times B)\bigcup(B \times G)\}$. Let the average power expenditure on these channel states be P_{GB}. Suppose we relax our assumption, and give full CSI to each transmitter whenever one of the links is in state G and the other in B. Furthermore, let us enforce only a average sum-power constraint of P_{GB} in these states. In such a system, only the better user transmits with an appropriate power [127]. This fact can be utilized along with (31) and (32) to equivalently write the maximum throughput as

$$\begin{aligned} J^* = \max \frac{\mu(B)}{2} &\int_B \log\left(1 + h^2(P_1(h,B) + P_2(B,h))\right) d\Psi(h) \\ + \frac{\mu(G)}{2} &\int_G \log\left(1 + h^2(P_1(h,G) + P_2(G,h))\right) d\Psi(h) \\ + \frac{\mu(B)}{2} &\int_G (\log(1 + h^2 P_1(h,B)) + \log(1 + h^2 P_2(B,h))) d\Psi(h), \end{aligned} \tag{33}$$

where the maximization is over $P_1(\cdot,\cdot)$ and $P_2(\cdot,\cdot)$. Furthermore, the original individual power constraint is relaxed to an average sum-power constraint of the form,

$$\begin{aligned} &\mu(B) \int_B (P_1(h,B) + P_2(B,h)) d\Psi(h) + \\ &\quad \mu(G) \int_G (P_1(h,G) + P_2(G,h)) d\Psi(h) + \mu(B) \int_G (P_1(h,B) + P_2(B,h)) d\Psi(h) \le 2P^{avg}. \end{aligned}$$

Notice that our integration now is over just one variable. Let us denote,

$$P_B(h) = \frac{P_1(h,B) + P_2(B,h)}{2} \text{ and } P_G(h) = \frac{P_1(h,G) + P_2(G,h)}{2}.$$

By the concavity of logarithm, the maximization can be bounded in terms of the new variable as

$$\begin{aligned} J^* \le \max \frac{\mu(B)}{2} \int_B \log\left(1 + h^2 2P_B(h)\right) d\Psi(h) + \frac{\mu(G)}{2} \int_G \log\left(1 + h^2 2P_G(h)\right) d\Psi(h) \\ + \frac{\mu(B)}{2} \int_G 2\log(1 + h^2 P_B(h)) d\Psi(h). \end{aligned} \tag{34}$$

The power constraint, in the new notation, is

$$\mu(B) \int_B 2P_B(h) d\Psi(h) + \mu(G) \int_G 2P_G(h) d\Psi(h) + \mu(B) \int_G 2P_B(h) d\Psi(h) \le 2P^{avg}. \tag{35}$$

Let the RHS of (34) be denoted as J^{**}. Further simplification is possible by treating the variable h as one corresponding to a single-user channel with appropriate distribution and an average power of $2P^{avg}$.

Lemma 15. *For 2 identical users with individual cdf $\Psi(\cdot)$, the maximal throughput with partial CSI is $C_1(\Phi, 2P^{avg})$, where*

$$d\Phi(h) = \begin{cases} d\Psi(h)\mu(B) \text{ if } h \in B \\ d\Psi(h)(1 + \mu(B)) \text{ if } h \in G \end{cases} \tag{36}$$

Proof. First we show that

$$C_1(\Phi, 2P^{avg}) \ge J^{**}.$$

For the single user channel $\Phi(h)$, consider two power allocation schemes $\hat{P}$ and $\tilde{P}$ such that

$$\hat{P}(h) = \begin{cases} 2P_B(h), h \in B \\ 2P_G(h), h \in G \end{cases} \tag{37}$$

and

$$\tilde{P}(h) = \begin{cases} 2P_B(h), h \in B \\ P_B(h), h \in G \end{cases} . \tag{38}$$

If we use $\hat{P}$ for a fraction $\frac{\mu(G)}{1+\mu(B)}$ of the times over $\Phi(h)$, and $\tilde{P}$ for the remaining fraction, the throughput is

$$\frac{\mu(B)}{2}\int_B \log(1+h^2P_B(h))d\Psi(h) + \frac{1+\mu(B)}{2}\frac{\mu(G)}{1+\mu(B)}\int_G \log(1+h^2 2P_G(h))d\Psi(h) \\ + \frac{1+\mu(B)}{2}\frac{2\mu(B)}{1+\mu(B)}\int_G \log(1+h^2P_B(h))d\Psi(h), \tag{39}$$

which is indeed J^{**}. Notice that an average power constraint of $2P^{avg}$ is maintained under this allocation. Let us now show that $C_1(\Phi, 2P^{avg})$ is in fact achievable for our MAC with partial CSI model. Let $P'(h)$ be the optimal single-user power allocation for the channel $\Phi(h)$. Consider the following power allocation in (33).

$$P_1(h,G) = P_2(h,G) = 0\,, \forall h \in B\,; \qquad P_1(h,B) = P_2(B,h) = P'(h)\,, \forall h \in G$$

$$P_1(h,G) = P_2(h,G) = \frac{P'(h)}{2}\,, \forall h \in G\,; \qquad P_1(h,B) = P_2(B,h) = \frac{P'(h)}{2}\,, \forall h \in B.$$

The users will choose the midpoint rates whenever both users are either in B or in G. In other cases, only the better user is active. Clearly the power constraints are met and the throughput is indeed $C_1(\Phi, 2P^{avg})$. □

For $L > 2$ users, if there are $K \geq 1$ links in G, only those links with $h_k \in G$ will transmit at their respective $K-$ user midpoint rates. On the other hand, if no links are in G, all L users transmit at their respective $L-$user midpoint rates. The power allocation can be effectively determined by single user water-filling using the cdf $\Phi(h)$ given in Theorem 12.

Chapter 3

Two-Receiver Broadcast Channel with Confidential Messages and Secret Keys

3.1 Motivation

We consider the problem of transmitting confidential messages over a two receiver broadcast channel. Two private messages are to be communicated, one to each of the two receivers. Each message is to be kept secret from the unintended receiver. Secret keys are available at fixed rates between each receiver and the transmitter. Various regimes of the key rates are described and achievable schemes are presented for each. Our schemes involve double random binning, key-dependent codebooks, and a technique called sectioning. Interestingly, double encryption on time-sharing sequences enhances the achievable region in certain regimes. The model subsumes several other models in the literature.

3.2 Introduction

We study a discrete-memoryless broadcast channel (DM-BC) with two receivers. The transmitter needs to send separate messages to each receiver, and the message intended for each receiver is to be kept secret from the other. In addition, each transmitter-receiver pair has a (private) secret key available at a fixed rate. These can assist in achieving secrecy. We present an inner bound to the capacity region for this problem, and also provide an outer bound. Security of the transmitted message is of prime importance in broadcast networks. Even if no external eavesdropper is present, it is sometimes necessary to secure the messages of the receivers against each other. One example of this is the model studied by Liu, Maric, Spasojevic, Yates [65], where secrecy-rate regions were obtained using double-random binning. Xu, Cao, Chen [140] extended this model by requiring the transmission of a common message, and obtained an achievable rate-equivocation region. To the best of our knowledge, Yamamoto [141] was the first to develop a coherent scheme that unified channel coding techniques and the use of secret keys to increase secrecy/equivocation rates for a (degraded/less-noisy) DM-BC. More recently, Kang and Liu [67] extended the result of [141] to a general BC – they introduced the crucial notion of key-dependent codebooks. In both [141] and [67], the legitimate information flow is point-to-point, and the eavesdropper is external. The scheme in [67] also employed, in some regimes, encoding encrypted data inside a sequence decoded by both the eavesdropper as well as the legitimate user. In [65], two legitimate information flows are present, and both receivers also eavesdrop on the others' message. The broadcast model in this chapter is similar to [65], however, the presence of secret keys introduces a degree of freedom that changes the dynamics significantly. In order to achieve (weak) secrecy, the achievable scheme of [65] incurs two separate rate penalties on each individual data stream due to double random binning. We show that the availability of secret keys can be used to progressively dispense with this double penalty. We develop a unified scheme naturally integrating double-random binning [65], key-dependent codebooks [67], and sectioning Ardestanizadeh, Franceschetti, Javidi, Kim [142]. In the extreme cases where there are no secret keys or there is only one legitimate receiver, the region respectively simplifies to the regions of [65] and [67], which can be seen as special cases of our model. As already noted in [67], surprisingly, it is often beneficial to encrypt information into a common sequence, which is decoded by both the receivers. Though the sequence is commonly received, encryption ensures secrecy. Our method of encrypting to a common sequence differs significantly from the scheme of [67, Section IV, Case 2]. We show that a simple, but appropriate, one-time pad (OTP) idea is enough to encrypt the data contained in the common sequence, thereby also simplifying the coding scheme of [67]. In other relevant work, Yin, Pang, Xue, Zhou [143] and Dai, Vinck, Luo, Zhuang [144] considered BCs with common and confidential messages with feedback used to generate secret keys. Schaefer, Khisti, Boche [145] and Schaefer and Khisti [146] respectively study two models with two legitimate receivers and an external eavesdropper, and a secret key shared between each transmitter and legitimate receiver pair. However, there is no confidentiality requirement between the legitimate receivers. The chapter is organized as follows. In Section 3.3, the model is described. In Section 3.4, the main theorem and various cases are enumerated. Section 3.5 details the achievable schemes. In Section 3.6, we present an outer bound.

3.3 The Model

We assume a two-receiver discrete memoryless BC with two confidential messages and two secret keys. The finite sets $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ represent the channel's input and the two output

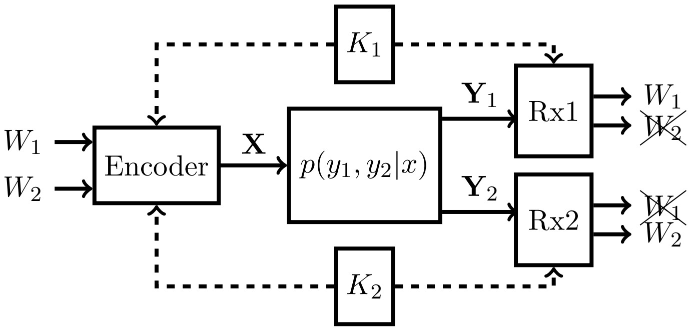

Figure 3.1: Two Receiver Broadcast Channel with Two Confidential Messages and Two Secret Keys

alphabets respectively. The channel is described by the conditional probability distribution $P_{Y_1,Y_2|X}(y_1, y_2|x)$, where RVs $X \in \mathcal{X}$, $Y_1 \in \mathcal{Y}_1$, $Y_2 \in \mathcal{Y}_2$. In addition, we assume the availability of secret keys, denoted by RVs $K_1 \in \mathcal{K}_1$ and $K_2 \in \mathcal{K}_2$, between each respective transmitter-receiver pair 1 and 2 unknown to the other receiver, at rates R_{k_1} and R_{k_2} respectively. The transmitter intends to send an independent message $W_t \in \{1, 2 \ldots, 2^{nR_t}\} \triangleq \mathcal{W}_t$ to the respective receiver $t \in \{1, 2\}$ in n channel uses while ensuring information theoretic secrecy, defined below. The channel is memoryless and without feedback i.e. $\forall \mathbf{x} \in \mathcal{X}^n$, $\mathbf{y}_t \in \mathcal{Y}_t^n$, $t = 1, 2$

$$P(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}) = \prod_{i=1}^{n} P_{Y_1,Y_2|X}(y_{1i}, y_{2i}|x_i)$$

A stochastic encoder is specified by a matrix of conditional probabilities $f(\mathbf{x}|w_1, k_1, w_2, k_2)$, $\forall w_t \in \mathcal{W}_t$, $k_t \in \mathcal{K}_t$, and

$$\sum_{\mathbf{x} \in \mathcal{X}^n} f(\mathbf{x}|w_1, k_1, w_2, k_2) = 1$$

$f(\mathbf{x}|w_1, k_1, w_2, k_2)$ is the probability that the quadruple of messages and keys (w_1, k_1, w_2, k_2) are encoded as the channel input $\mathbf{x}$. The decoding function at the receiver $t = 1, 2$ is a mapping $\phi_t : \mathcal{K}_t \times \mathcal{Y}_t^n \to \mathcal{W}_t$. A $(2^{nR_1}, 2^{nR_2}, 2^{nR_{k_1}}, 2^{nR_{k_2}}, n, P_e^{(n)})$ code for the broadcast channel consists of the encoding function f, decoding functions ϕ_1, ϕ_2, and the error probability defined as

$$P_e^{(n)} \triangleq \max\{P_{e,1}^{(n)}, P_{e,2}^{(n)}\},$$

where for $t = 1, 2$,

$$P_{e,t}^{(n)} = \sum_{w_1,w_2,k_1,k_2} \frac{P[\phi_t(K_t, \mathbf{Y}_t) \neq w_t|(w_1, w_2, k_1, k_2)]}{2^{nR_1} \times 2^{nR_2} \times 2^{nR_{k_1}} \times 2^{nR_{k_2}}}$$

A rate pair (R_1, R_2) is said to be achievable for the broadcast channel with confidential messages and two secret keys at rates (R_{k_1}, R_{k_2}) if, for any $\epsilon_0 > 0$, there exists a $(2^{nR_1}, 2^{nR_2}, 2^{nR_{k_1}}, 2^{nR_{k_2}}, n, P_e^{(n)})$ code which satisfies both

- reliability requirement: $P_e^{(n)} \leq \epsilon_0$
- secrecy constraint: $nR_t - H(W_t|\mathbf{Y}_{\bar{t}}, K_{\bar{t}}) \leq n\epsilon_0, t = 1, 2.$

This definition corresponds to the so-called *weak secrecy-key rate* [65]. We use the notation $\bar{t} \triangleq \{1, 2\} \setminus \{t\}$. We define a class π_{BC} of distributions $P(u, v_1, v_2, x, y_1, y_2)$ that factor as $P(u)P(v_1, v_2|u)P(x|v_1, v_2)P(y_1, y_2|x)$.

3.4 Inner Bound

The main result is presented below.

Theorem 16. *Let $\mathbb{R}_{BC}(\pi_{BC})$ denote the union of all $(R_1, R_2) \in \mathbb{R}_+^2$ satisfying*

$$\begin{aligned} R_1 &\leq I(V_1; Y_1|U) + \min\{R_{k_1} - I(V_1; Y_2|V_2, U) - I(V_1; V_2|U), I(U; Y_1)\} \\ R_2 &\leq I(V_2; Y_2|U) + \min\{R_{k_2} - I(V_2; Y_1|V_1, U) - I(V_2; V_1|U), I(U; Y_2\} \\ R_1 + R_2 &\leq I(V_1; Y_1|U) + I(V_2; Y_2|U) - I(V_1; V_2|U) + \min\{I(U; Y_1), I(U; Y_2)\}. \end{aligned} \tag{3.1}$$

where the union is over all distributions $P(u, v_1, v_2, x, y_1, y_2)$ in π_{BC}. Every rate pair $(R_1, R_2) \in \mathbb{R}_{BC}(\pi_{BC})$ is achievable.

The key rates determine the achievable scheme. We enumerate eight different regimes and present their achievable schemes in Section IV.

- **Case** 1:

$$R_{k_1} \leq I(V_1; Y_2|V_2, U); \;\; R_{k_2} \leq I(V_2; Y_1|V_1, U).$$

- **Case** 2:

$$\begin{aligned} R_{k_2} &\leq I(V_2; Y_1|V_1, U) \\ I(V_1; Y_2|V_2, U) < R_{k_1} &\leq I(V_1; Y_2|V_2, U) + I(V_1; V_2|U). \end{aligned}$$

- **Case** 3: This is identical to Case 2 with the receivers' roles reversed.
- **Case** 4:

$$\begin{aligned} I(V_1; Y_2|V_2, U) < R_{k_1} &\leq I(V_1; Y_2|V_2, U) + I(V_1; V_2|U) \\ I(V_2; Y_1|V_1, U) < R_{k_2} &\leq I(V_2; Y_1|V_1, U) + I(V_2; V_1|U). \end{aligned}$$

- **Case** 5**:**

$$
\begin{aligned}
R_{k_1} &> I(V_1;Y_2|V_2,U) + I(V_1;V_2|U) \\
I(V_2;Y_1|V_1,U) < R_{k_2} &\leq I(V_2;Y_1|V_1,U) + I(V_2;V_1|U).
\end{aligned}
$$

- **Case** 6**:** This is identical to Case 5 with the receivers' roles reversed.
- **Case** 7**:**

$$
\begin{aligned}
R_{k_1} &> I(V_1;Y_2|V_2,U) + I(V_1;V_2|U) \\
R_{k_2} &> I(V_2;Y_1|V_1,U) + I(V_2;V_1|U).
\end{aligned}
$$

- **Case** 8**:** The remaining cases (where key rate mismatch is large) and $5-7$ have achievable schemes etc. similar to Case 4.

3.5 Achievability Schemes

For ease of exposition, we will distinguish between the terms *code* and *codebook*. Also, $R^{\dagger} \triangleq I(V_1;V_2|U) + \epsilon_1'$ where $\epsilon_1' > 0$ is a small positive constant.

3.5.1 Case 1:

The proposed achievable region becomes

$$
\begin{aligned}
R_1 &\leq I(V_1;Y_1|U) - I(V_1;V_2|U) - I(V_1;Y_2|V_2,U) + R_{k_1} \\
R_2 &\leq I(V_2;Y_2|U) - I(V_2;V_1|U) - I(V_2;Y_1|V_1,U) + R_{k_2}
\end{aligned}
$$

Our achievability scheme melds the techniques of *double random binning* [65] and *code consisting of multiple key-dependent codebooks* [67, Section IV, Case 1]. Double binning, in turn, combines Gelfand-Pinsker binning and random binning (to satisfy mutual covering and to confuse the other receiver to maintain perfect secrecy). The employed coding structure is shown below. A joint encoder generates two equivocation codewords $\mathbf{v}_1$ and $\mathbf{v}_2$, one for each message-key pair (W_1, K_1) and (W_2, K_2). The pair $(\mathbf{v}_1, \mathbf{v}_2)$ is stochastically mapped into $\mathbf{x}$. The details follow.

3.5.1.1 Code Construction

Fix $P(u)$, $P(v_1|u)$ and $P(v_2|u)$ as well as $P(x|v_1, v_2)$ and define

$$
\begin{aligned}
R_1' &\triangleq I(V_1;Y_2|V_2,U) - \epsilon_1' - R_{k_1} \\
R_2' &\triangleq I(V_2;Y_1|V_1,U) - \epsilon_1' - R_{k_2}
\end{aligned}
$$

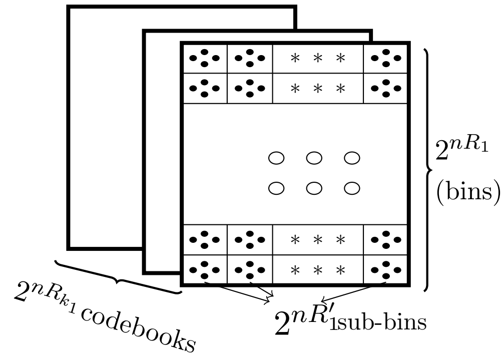

Figure 3.2: Case 1: Code for receiver 1

Randomly generate a sequence $\mathbf{u} \sim P(\mathbf{u}) = \prod_{i=1}^{n} P(u_i)$. For $t = 1, 2$, generate code $\mathcal{C}_t$ with $2^{(R_{k_t}+R_t+R'_t+R^{\dagger})}$ (conditionally) independent sequences $\mathbf{v}_t$ each with probability $P(\mathbf{v}_t|\mathbf{u}) = \prod_{i=1}^{n} P(v_{t,i}|u_i)$ and label them $\mathbf{v}_t(k_t, w_t, s_t, r_t)$ for $k_t \in \{1, \ldots, 2^{nR_{k_t}}\}$, $w_t \in \{1, \ldots, 2^{nR_t}\}$, $s_t \in \{1, \ldots, 2^{nR'_t}\}$, $r_t \in \{1, \ldots, 2^{nR^{\dagger}}\}$. W.l.o.g, $2^{nR_{k_t}}$, 2^{nR_t}, $2^{nR'_t}$, $2^{nR^{\dagger}}$ are considered to be integers.

The code $\mathcal{C}_1$ for receiver 1 consists of $2^{nR_{k_1}}$ codebooks [67, Section IV, Case 1]. Each codebook of $\mathcal{C}_1$ is doubly binned, as in [65], with 2^{nR_1} bins, each containing $2^{nR'_1} = 2^{n[I(V_1;Y_2|V_2,U)-R_{k_1}-\epsilon'_1]}$ sub-bins. Receiver 2's code $\mathcal{C}_2$ is similar. In all codebooks, each sub-bin contains $2^{nR^{\dagger}} = 2^{n[I(V_1;V_2|U)+\epsilon'_1]}$ codewords. Furthermore, each codebook in code $\mathcal{C}_t$ contains $2^{n[I(V_t;Y_t|U)-\epsilon'_1]}$ codewords. Please see Figure 3.2. The sequence $\mathbf{u}$ and code $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2\}$ is commmunicated to all parties.

3.5.1.2 Encoding

Given key pair $(k_1, k_2) \in \mathcal{K}_1 \times \mathcal{K}_2$, the encoder chooses the appropriate codebooks $\mathcal{C}_1(k_1)$ and $\mathcal{C}_2(k_2)$ in the respective codes $\mathcal{C}_1$ and $\mathcal{C}_2$. To send $(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_2$, the transmitter randomly chooses a sub-bin $\mathcal{C}_t(k_t, w_t, s_t)$ from the bin $\mathcal{C}_t(k_t, w_t)$, for $t = 1, 2$. Next, a pair (r_1, r_2) is chosen such that $(\mathbf{v}_1(k_1, w_1, s_1, r_1), \mathbf{v}_2(k_2, w_2, s_2, r_2)) \in A_{\epsilon}^{(n)}(V_1, V_2|\mathbf{u})$, where $A_{\epsilon}^{(n)}(V_1, V_2|\mathbf{u})$ indicates the set of jointly typical sequences $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{u})$ for the given realization $\mathbf{u}$, according to $P_{V_1,V_2|U}$. By mutual covering [27, Chapter 8], such a pair exists with high probability. If more than one jointly typical pair exists, one is randomly chosen. We now employ the stochastic encoder which generates $\mathbf{x} \sim \prod_{i=1}^{n} p(x_i|v_{1i}, v_{2i})$ for transmission. *Note that the* $\mathbf{x}$ *codewords are not part of the code* $\mathcal{C}$. *They are generated at the time of transmission after choosing an appropriate* $(\mathbf{v}_1, \mathbf{v}_2)$ *pair.*

3.5.1.3 Decoding

The decoder t has access to the shared key k_t and so the decoding at decoder t for $t = 1, 2$ has to be done from among $2^{n(R_t+R'_t+R^\dagger)} \approx 2^{n[I(V_t;Y_t|U)]}$ sequences $\mathbf{v}_t$ in the codebook $\mathcal{C}_t(k_t)$, see Fig 3.2. Decoder t chooses w_t such that $(\mathbf{v}_t(k_t, w_t, s_t, r_t), \mathbf{y}_t, \mathbf{u}) \in A_\epsilon^{(n)}(V_t, Y_t, U)$ for some (s_t, r_t), if a unique such w_t exists, else an error is declared.

3.5.1.4 Error Probability Analysis

While the codewords here are quadruply indexed to reflect the codebook index as shown in Fig 3.2, the rest of the details are standard. The error probability analysis similar to [65], with the main difference that the the codewords here are quadruply indexed to reflect the codebook index.

W.l.o.g, assume that the transmitter sends the message pair $(w_1 = 1, w_2 = 1)$ and $(s_1 = 1, s_2 = 1)$ and in addition, the secret keys are $(k_1 = 1, k_2 = 1)$. Consider the (encoding) error event $\mathcal{T}$ that the encoder cannot find an appropriately jointly typical pair, i.e. $\forall r_1, r_2$

$$\mathcal{T} \triangleq \{(\mathbf{V}_1(1,1,1,r_1), \mathbf{V}_2(1,1,1,r_2)) \notin A_\epsilon^{(n)}(V_1, V_2|\mathbf{u})\}.$$

Since $R^\dagger > I(V_1; V_2|U)$, by the mutual covering lemma, [27, Chapter 8], $P\{T\} \leq \delta$ where $\delta > 0$ is small enough for large n. Let us assume that $(\mathbf{V}_1(1,1,1,1), \mathbf{V}_2(1,1,1,1))$ is chosen for tranmission, and define the event

$$\mathcal{T}^c \triangleq \{(\mathbf{V}_1(1,1,1,1), \mathbf{V}_2(1,1,1,1)) \in A_\epsilon^{(n)}(V_1, V_2|\mathbf{u})\}.$$

The decoding error probability at receiver 1 is then bounded as

$$P_{e,1}^{(n)} \leq P\{\mathcal{T}\} + (1 - P\{\mathcal{T}\})[P\{\bigcap_{s_1,r_1} E_1^c(1,1,s_1,r_1)|\mathcal{T}^c\} + \sum_{w_1\neq 1}\sum_{s_1,r_1} P\{E_1(1,w_1,s_1,r_1)|\mathcal{T}^c\}] \tag{3.2}$$

$$\leq P\{\mathcal{T}\} + P\{E_1^c(1,1,1,1)|\mathcal{T}^c\} + \sum_{w_1\neq 1}\sum_{s_1,r_1} P\{E_1(1,w_1,s_1,r_1)|\mathcal{T}^c\} \tag{3.3}$$

where

$$E_t(1, w_t, s_t, r_t) = \{(\mathbf{v}_t(1, w_t, s_t, r_t), \mathbf{y}_t) \in A_\epsilon^{(n)}(V_t, Y_t|\mathbf{u})\}.$$

Since $P\{E_1(1, w_1, s_1, r_1)|\mathcal{T}^c\} \leq 2^{-n[I(V_1;Y_1|U)-\epsilon]}$, by using joint typicality lemma [27], the probability of error can be bounded as

$$P_{e,1}^{(n)} \leq \delta + \epsilon + 2^{nR_1}2^{nR'_1}2^{nR^\dagger}2^{-n[I(V_1;Y_1|U)-\epsilon]}. \tag{3.4}$$

Thus, if $R_1 + R'_1 + R^\dagger < I(V_1; Y_1|U)$ then, $P_{e,1}^{(n)} < \epsilon_0$ for sufficiently large n. Similar calculations for receiver 2 shows that if $R_2 + R'_2 + R^\dagger < I(V_2; Y_2|U)$, then $P_{e,2}^{(n)} \to 0$.

3.5.1.5 Equivocation calculation for Case 1

We prove that secrecy holds. It is worth mentioning that the calculations below have subtle differences from a similar calculation in [67], improving the robustness. We first express the equivocation as

$$H(W_1|\mathbf{Y}_2, K_2) = \sum_{k_2 \in \mathcal{K}_2} P(K_2 = k_2) H(W_1|\mathbf{Y}_2, k_2) \tag{3.5}$$

We will now show that $\forall K_2 = k_2$,

$$H(W_1|\mathbf{Y}_2, k_2) \geq nR_1 - n\tilde{\epsilon}, \tag{3.6}$$

implying secrecy of Receiver 1's messages.

$$\begin{aligned} H(W_1|\mathbf{Y}_2, k_2) \\ &\geq H(W_1|\mathbf{Y}_2, k_2, \mathbf{V}_2, \mathbf{U}) \\ &= H(W_1, \mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) - H(\mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) \\ &= H(W_1, \mathbf{V}_1, \mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) - H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1) \\ &\quad - H(\mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) \\ &= H(W_1, \mathbf{V}_1|k_2, \mathbf{V}_2, \mathbf{U}) + H(\mathbf{Y}_2|k_2, \mathbf{V}_1, \mathbf{V}_2, \mathbf{U}, W_1) \\ &\quad - H(\mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) - H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1) \end{aligned} \tag{3.7}$$

Based on functional dependence graphs, we can show that $\forall K_2 = k_2$, $W_1 \rightarrow (\mathbf{V}_1, \mathbf{V}_2, \mathbf{U}) \rightarrow \mathbf{Y}_2$ forms a Markov Chain. Thus the second term becomes

$$H(\mathbf{Y}_2|k_2, \mathbf{V}_1, \mathbf{V}_2, \mathbf{U}, W_1) = H(\mathbf{Y}_2|k_2, \mathbf{V}_1, \mathbf{V}_2, \mathbf{U})$$

Making the replacement, we now have

$$\begin{aligned} &H(W_1|\mathbf{Y}_2, k_2) \\ &\geq H(W_1, \mathbf{V}_1|k_2, \mathbf{V}_2, \mathbf{U}) + H(\mathbf{Y}_2|k_2, \mathbf{V}_1, \mathbf{V}_2, \mathbf{U}) \\ &\quad - H(\mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) - H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1) \\ &= H(W_1, \mathbf{V}_1|k_2, \mathbf{V}_2, \mathbf{U}) - I(\mathbf{V}_1; \mathbf{Y}_2|k_2, \mathbf{V}_2, \mathbf{U}) - H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1) \\ &. \end{aligned} \tag{3.8}$$

By a calculation analogous to [65, Lemma 2], we can show that

$$H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1) \leq n\epsilon_3', \tag{3.9}$$

where ϵ_3' is small for sufficiently large n. This can be interpreted to mean that there is no uncertainty left in $\mathbf{V}_1$ given $(k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1)$. If there were, the randomness can be included in W_1 to improve the rate R_1.

To compute $H(\mathbf{V}_1|k_2, \mathbf{Y}_2, \mathbf{V}_2, \mathbf{U}, W_1)$, we proceed as follows. Suppose $W_1 = w_1$, Receiver 2 (acting as the eavesdropper) tries to decode $\mathbf{v}_1(k_1, w_1, s_1, r_1)$ based on its received sequence $\mathbf{y}_2$ (of course it already has knowledge of its own key k_2). Since Decoder 2 knows w_1, let $\lambda_{k_2}(w_1)$ denote the average probability of error of decoding the indices (k_1, s_1, r_1) at Receiver 2 (given that its key is k_2). Joint typicality enables us to show that:

Lemma 17. *$\lambda_{k_2}(w_1) \leq \epsilon_0'$ for sufficiently large n.*

Proof. For a given time-sharing sequence $\mathbf{u}$, let $A_\epsilon^{(n)}(P_{V_1,V_2,Y_2|U})$ denote the set of jointly typical sequences $\mathbf{v}_1$ and $(\mathbf{v}_2, \mathbf{y}_2)$ with respect to $P(v_1, v_2, y_2|u)$. For a given $W_1 = w_1$, Decoder 2 chooses (k_1, s_1, r_1) with

$$(\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \in A_\epsilon^{(n)}(P_{V_1,V_2,Y_2|U}),$$

if such a pair (k_1, s_1, r_1) exists and is unique; else an error is declared.

We define the event

$$\hat{E}(k_1, s_1, r_1) = (\mathbf{v}_1(k_1, w_1, s_1, r_1), \mathbf{v}_2, \mathbf{y}_2) \in A_\epsilon^{(n)}(P_{V_1,V_2,Y_2|U})$$

W.l.o.g, we assume that $\mathbf{v}_1(k_1 = 1, w_1, s_1 = 1, r_1 = 1)$ was chosen, and define the event

$$\mathcal{B}_{w_1} = \{\mathbf{v}_1(1, w_1, 1, 1) \text{ chosen}\}$$

Hence

$$\lambda_{k_2}(w_1) \leq P\{\hat{E}^c(k_1 = 1, s_1 = 1, r_1 = 1)|\mathcal{B}_{w_1}\} + \sum_{(k_1,s_1,r_1)\neq(1,1,1)} P\{\hat{E}((k_1, s_1, r_1))|\mathcal{B}_{w_1}\} \tag{3.10}$$

where $\hat{E}^c(k_1 = 1, s_1 = 1, r_1 = 1)$ denotes the event

$$\{(\mathbf{v}_1(1, w_1, 1, 1), \mathbf{v}_2, \mathbf{y}_2) \notin A_\epsilon^{(n)}(P_{V_1,V_2,Y_2|U})\} \tag{3.11}$$

By the joint AEP, $P\{\hat{E}^c(k_1 = 1, s_1 = 1, r_1 = 1)\} \leq \epsilon$ and for $(k_1, s_1, r_1) \neq (1, 1, 1)$,

$$P\{\hat{E}((k_1, s_1, r_1))|\mathcal{B}_{w_1}\} \leq 2^{-n[I(V_1;V_2,Y_2|U)-\epsilon]} \tag{3.12}$$

We upper bound $\lambda_{k_2}(w_1)$ as

$$\epsilon + 2^{nR_{k_1}} 2^{nR_1'} 2^{nR^\dagger} 2^{-n[I(V_1;V_2,Y_2|U)-\epsilon]} \tag{3.13}$$

Now since $R_{k_1} + R_1' + R^\dagger = I(V_1; V_2, Y_2|U)$, we finally have $\lambda_{k_2}(w_1) \leq \epsilon_0'$ where ϵ_0' small for n sufficiently large. □

By Fano's inequality

$$\frac{1}{n}H(\mathbf{V}_1|k_2,\mathbf{Y}_2,\mathbf{V}_2,\mathbf{U},W_1=w_1) \le \frac{1}{n}[1+\lambda_{k_2}(w_1)\log[2^{nR_{k_1}}2^{nR'_1}2^{nR^{\dagger}}]] \overset{\triangle}{=} \epsilon'_3 \tag{3.14}$$

We conclude that

$$\frac{1}{n}H(\mathbf{V}_1|k_2,\mathbf{Y}_2,\mathbf{V}_2,\mathbf{U},W_1) = \frac{1}{n}\sum_{w_1\in\mathcal{W}_1} P(W_1=w_1)H(\mathbf{V}_1|k_2,\mathbf{Y}_2,\mathbf{V}_2,\mathbf{U},W_1=w_1) \le \epsilon'_3 \tag{3.15}$$

The above is an application of the technique from [67] whereby the key index ($K_t, t = 1,2$) introduces an extra degree of randomness which increases equivocation only at the receiver for which the message is not intended. The secret keys ensure that the randomness requirement (the RV that chooses the subbin) for the transmitter-receiver t pair is reduced by the respective key rate R_{k_t} from $I(V_t;Y_{\bar{t}}|V_{\bar{t}},U)$ to $I(V_t;Y_{\bar{t}}|V_{\bar{t}},U) - R_{k_t}$ for $t = 1,2$. We have shown that

$$H(\mathbf{V}_1|k_2,\mathbf{Y}_2,\mathbf{V}_2,\mathbf{U},W_1) \le n\epsilon'_3 \tag{3.16}$$

Substituting the above inequality in (3.8), we get

$$H(W_1|\mathbf{Y}_2,k_2) \ge H(W_1,\mathbf{V}_1|k_2,\mathbf{V}_2,\mathbf{U}) - I(\mathbf{V}_1;\mathbf{Y}_2|k_2,\mathbf{V}_2,\mathbf{U}) - n\epsilon'_3. \tag{3.17}$$

Now, in the first term on the RHS above, the equivocation of $(W_1,\mathbf{V}_1)$ is the logarithm of the total number of cells in the code $\mathcal{C}_1$, which is $2^{n[I(V_1;Y_1|U)+R_{k_1}]}$. Next, we note that in the random codebook $\mathcal{C}$, our coding scheme only requires that we find a codeword $\mathbf{V}_2$ that is jointly typical with $\mathbf{V}_1$, *thus the choice of key k_2 does not play a role.* Conditioning by $\mathbf{V}_2$ causes a reduction by a factor of $2^{I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})}$, to give $2^{n[I(V_1;Y_1|U)+R_{k_1}]}2^{-I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})}$. Taking logs, we get $n[I(V_1;Y_1|U)+R_{k_1}] - I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})$. Thus we have

$$H(W_1|\mathbf{Y}_2,k_2) \ge n[I(V_1;Y_1|U)+R_{k_1}] - I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U}) - I(\mathbf{V}_1;\mathbf{Y}_2|k_2,\mathbf{V}_2,\mathbf{U}) - n\epsilon'_3 \tag{3.18}$$

Now, by a calculation that essentially upper bounds the mutual information between codewords in codebooks (the basic idea comes from [65]), we can obtain the following inequalities:

$$I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U}) \le nI(V_1;V_2|U) + n\epsilon'_2 \tag{3.19}$$

and

$$I(\mathbf{V}_1;\mathbf{Y}_2|k_2,\mathbf{V}_2,\mathbf{U}) \le nI(V_1;Y_2|V_2,U) + n\epsilon'_4 \tag{3.20}$$

Substituting (3.19) and (3.20) in (3.18), we get

$$\begin{aligned} H(W_1|\mathbf{Y}_2,k_2) &\ge n[I(V_1;Y_1|U)+R_{k_1}] - nI(V_1;V_2) - nI(V_1;Y_2|V_2,U) - n(\epsilon'_2+\epsilon'_3+\epsilon'_4) \\ &= nR_1 - n(\epsilon'_2+\epsilon'_3+\epsilon'_4) \end{aligned} \tag{3.21}$$

which gives us (3.6), and thence (3.5), as desired. The equivocation calculation for receiver 2 in this case is similar.

3.5.2 Case 2:

The key idea for this case is to employ the *sectioning* technique of [142]. We have to show that the rate-pairs

$$R_1 \leq I(V_1; Y_1|U) - I(V_1; V_2|U) - I(V_1; Y_2|V_2, U) + R_{k_1}$$
$$R_2 \leq I(V_2; Y_2|U) - I(V_2; V_1|U) - I(V_2; Y_1|V_1, U) + R_{k_2},$$

are achievable. We split the message $W_1 \triangleq (\tilde{W}_1, W_1^{otp})$ and define

$$R'_{k_1} \triangleq I(V_1; Y_2|V_2, U) - \epsilon'_1; \; R^{otp}_{k_1} \triangleq R_{k_1} - R'_{k_1}$$
$$\tilde{R}_1 \triangleq I(V_1; Y_1|U) - I(V_1; V_2|U) - \epsilon'_1$$

Total secure achievable rate for Receiver 1 in this regime consists of $\tilde{W}_1$ at $\tilde{R}_1$ by channel coding techniques, and W_1^{otp} at rate $R_1^{otp} = R_{k_1}^{otp}$.

3.5.2.1 Code Construction

Fix $P(u)$, $P(v_1|u)$, $P(v_2|u)$,$P(x|v_1, v_2)$ and (re-)define

$$R'_1 \triangleq I(V_1; Y_2|V_2, U) - \epsilon'_1 - R'_{k_1} = 0$$

(R'_2 as in Case 1). Randomly generate $\mathbf{u} \sim \prod_{i=1}^n P(u_i)$. The code $\mathcal{C}_1$ for receiver 1 consists of $2^{nR'_{k_1}}$ codebooks. A codebook contains $2^{n\tilde{R}_1}$ bins, each containing $2^{nR^\dagger} = 2^{n[I(V_1;V_2|U)+\epsilon'_1]}$ codewords $\mathbf{v}_1 \sim \prod_{i=1}^n P_{V_1|U}(v_{1i}|u_i)$. The notion of *sub-bins* in Fig 3.2 are replaced by *sections* for user 1 in this case. Each bin is divided evenly into $2^{nR_{k_1}^{otp}}$ sections. If $w_1 = (\tilde{w}_1, w_1^{otp})$, $\tilde{w}_1$ is encoded as the bin index. The pair (w_1^{otp}, k_1^{otp}) picks the section $w_1^{otp} \oplus k_1^{otp}$, as in [142], securing w_1^{otp} by an OTP.

The code $\mathcal{C}_2$ of receiver 2 is identical to that in Case 1. The code-construction for receiver 1 is summarized. Generate a code with $2^{(R'_{k_1}+\tilde{R}_1+R^\dagger)}$ (conditionally) independent sequences $\mathbf{v}_1$ each with probability $P(\mathbf{v}_1|\mathbf{u}) = \prod_{i=1}^n P(v_{1,i}|u_i)$ and label them $\mathbf{v}_1(k'_1, \tilde{w_1}, r_1)$ for $k'_1 \in \{1, \ldots, 2^{nR'_{k_1}}\}$, $\tilde{w_1} \in \{1, \ldots, 2^{n\tilde{R}_1}\}$, $r_1 \in \{1, \ldots, 2^{nR^\dagger}\}$. W.l.o.g, $2^{nR'_{k_1}}$, $2^{nR_{k_1}^{otp}}$,$2^{nR_{k_2}}$,$2^{n\tilde{R}_1}$, 2^{nR_2}, $2^{nR'_2}$, $2^{nR^\dagger}$ are considered to be integers. The sequence $\mathbf{u}$ and code $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2\}$ is commmunicated to all parties.

3.5.2.2 Encoding

Given key pair $(k_1, k_2) \in \mathcal{K}_1 \times \mathcal{K}_2$, where $k_1 = (k'_1, k_1^{otp})$, the encoder chooses the codebooks $\mathcal{C}_1(k'_1)$ and $\mathcal{C}_2(k_2)$. To send $(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_2$, where $w_1 = (\tilde{w}_1, w_1^{otp})$, the encoder chooses the section $w_1^{otp} \oplus k_1^{otp}$ from the bin $\mathcal{C}_1(k'_1, \tilde{w}_1)$. It randomly chooses a sub-bin $\mathcal{C}_2(k_2, w_2, s_2)$ from $\mathcal{C}_2(k_2, w_2)$. Since this contains $2^{n[I(V_1;V_2|U)+\epsilon'_1]}$ codewords, the number of available pairs $(\mathbf{v}_1, \mathbf{v}_2)$ is $\geq 2^{n[I(V_1;V_2|U)+\epsilon_1']}$, so that, by mutual covering [27, Chapter 8], with very high probability, jointly typical pairs exist. One is chosen randomly. Generate $\mathbf{x} \sim \prod_{i=1}^{n} P_{X|V_1,V_2}(x_i|v_{1i}, v_{2i})$ (*stochastic encoding*) and transmit.

3.5.2.3 Decoding

Receiver 1 knows the codebook $\mathcal{C}_1(k_1')$, and so decodes $\mathbf{v}_1$ from among $\approx 2^{n[I(V_1;Y_1|U)]}$ possibilities by joint typicality with $\mathbf{y}_1$ and $\mathbf{u}$. Clearly $\mathbf{v}_1$, and so $w_1 = (\tilde{w}_1, w_1^{otp})$ can be decoded with low error probability. Receiver 2 proceeds as in Case 1.

3.5.2.4 Error Probability Analysis

For receiver 2, the analysis is the same as in Case 1. For receiver 1, analysis similar to Case 1, but the sub-bin index s_1 is not used.

3.5.2.5 Equivocation for Case 2

With the replacements $W_1 \leftarrow \tilde{W}_1$ and $R_1 \leftarrow \tilde{R}_1$, the calculation is similar to Case 1. The message portion w_1^{otp} is secured by OTP, and is perfectly, and hence weakly secure.

Replacing $W_1 \leftarrow \tilde{W}_1$ in the expression (3.7), and by similar steps as in (3.8) – (3.17), along with Markov chain $\tilde{W}_1 \to (\mathbf{V}_1, \mathbf{V}_2, \mathbf{U}) \to \mathbf{Y}_2$ which holds for all $K_2 = k_2$, we get in place of (3.17),

$$H(\tilde{W}_1|\mathbf{Y}_2, k_2) \geq H(\tilde{W}_1, \mathbf{V}_1|\mathbf{V}_2, \mathbf{U}, k_2) - I(\mathbf{V}_1; \mathbf{Y}_2|\mathbf{V}_2, \mathbf{U}, k_2) - n\epsilon_3' \tag{3.22}$$

Expansion of the first term gives

$$H(\tilde{W}_1, \mathbf{V}_1|\mathbf{V}_2, \mathbf{U}, k_2) = H(\tilde{W}_1|\mathbf{V}_2, \mathbf{U}, k_2) + H(\mathbf{V}_1|\mathbf{V}_2, \mathbf{U}, \tilde{W}_1, k_2).$$

Notice that

$$H(\tilde{W}_1|\mathbf{V}_2, \mathbf{U}, k_2) = n\tilde{R}_1. \tag{3.23}$$

To compute the second term, we note that the total number of entries in the code $\mathcal{C}_1$ is $2^{n[I(V_1;Y_1|U)+R_{k_1}']}$. Conditioning by the message $\tilde{W}_1$ causes a reduction of possible transmitted $\mathbf{V}_1$ sequences by a factor of $2^{n\tilde{R}_1}$. Furthermore, conditioning by $\mathbf{V}_2$ causes a further reduction by a factor of $2^{I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})}$. *Note that, as in Case 1, key K_2's value plays no role.* So the remaining number of possible $\mathbf{V}_1$ sequences are

$$\begin{aligned}\tilde{N}_1 &= \frac{2^{n[I(V_1;Y_1|U)+R_{k_1}']}2^{-I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})}}{2^{n[I(V_1;Y_1|U)-I(V_1;V_2|U)-\epsilon_1']}} \\ &= \frac{2^{n[I(V_1;Y_1|U)+I(V_1;Y_2|V_2,U)-\epsilon_1']}2^{-I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U})}}{2^{n[I(V_1;Y_1|U)-I(V_1;V_2|U)-\epsilon_1']}}\end{aligned}$$

Thus

$$\begin{aligned} H(\mathbf{V}_1|\mathbf{V}_2,\mathbf{U},\tilde{W}_1,k_2) &= \log \tilde{N}_1 \\ &= nI(V_1;Y_2|U,V_2) + nI(V_1;V_2|U) - I(\mathbf{V}_1;\mathbf{V}_2|\mathbf{U}) \\ &\geq n\left(I(V_1;Y_2|U,V_2) - \epsilon_2'\right) \end{aligned} \tag{3.24}$$

where the last inequality follows by [65, Lemma 3]. From (3.22) – (3.24), it follows that

$$\begin{aligned} H(\tilde{W}_1|\mathbf{Y}_2,K_2) &= \sum_{k_2\in\mathcal{K}_2} P(K_2=k_2)H(\tilde{W}_1|\mathbf{Y}_2,k_2) \\ &\geq n\tilde{R}_1 + nI(V_1;Y_2|U,V_2) - I(\mathbf{V}_1;\mathbf{Y}_2|\mathbf{V}_2,\mathbf{U},k_2) - n(\epsilon_2'+\epsilon_3') \\ &\geq n\tilde{R}_1 - n\epsilon \end{aligned}$$

where the last inequality again follows from [65, Lemma 3]. The other portion of the message of receiver 1, namely, w_1^{otp} is secured by an OTP, and so is perfectly, and hence weakly secure. Equivocation calculation for receiver 2 in this case is exactly the same as in Case 1.

3.5.3 Case 3:

The achievable region similar to Case 2 with roles of the receivers reversed.

3.5.4 Case 4:

The key idea for this case is double-encryption on time-sharing sequences. The proposed achievable region becomes

$$\begin{aligned} R_1 &\leq I(V_1;Y_1|U) - I(V_1;V_2|U) - I(V_1;Y_2|V_2,U) + R_{k_1} \\ R_2 &\leq I(V_2;Y_2|U) - I(V_2;V_1|U) - I(V_2;Y_1|V_1,U) + R_{k_2} \\ R_1 + R_2 &\leq I(V_1;Y_1|U) + I(V_2;Y_2|U) - I(V_1;V_2|U) \\ &\qquad + \min\{I(U;Y_1),I(U;Y_2)\} \end{aligned} \tag{3.25}$$

3.5.4.1 Overview

We will split the key into three portions. The portion $\mathcal{K}_t^u$ (at rate $R_{k_t}^u$) is used to perform double-encryption on the time-sharing sequence $\mathbf{u}$, the portion $\mathcal{K}_t'$ (at rate R_{k_t}') is used to generate key-dependent codebooks, and the remaining portion $\mathcal{K}_t^{otp}$ (at rate $R_{k_t}^{otp}$) is used for sectioning the bins as in cases 2 and 3. The message is transmitted in three portions, portions W_t^u (at rate R_t^u, carried by $\mathbf{u}$) and W_t^{otp} (at rate R_t^{otp}) are protected by OTP using $\mathcal{K}_t^u$ and $\mathcal{K}_t^{otp}$ respectively, necessitating $R_t^u \leq R_{k_t}^u$ and $R_t^{otp} \leq R_{k_t}^{otp}$. The portion $\tilde{R}_t$ (carried by binned sequences $\mathbf{v}_t$) is transmitted securely by pure channel coding technique. Thus the total secure rate is $R_t = R_t^u + \tilde{R}_t + R_t^{otp}$.

3.5.4.2 Code Construction

The following holds for $t = 1, 2$, as also equations (3.26) to (3.30). Split $W_t \triangleq \left(\tilde{W}_t, W_t^{otp}, W_t^u\right)$ and $\mathcal{K}_t \triangleq \left(\mathcal{K}_t', \mathcal{K}_t^{otp}, \mathcal{K}_t^u\right)$. R_t and R_{k_t} are split as

$$R_t = \left(\tilde{R}_t, R_t^{otp}, R_t^u\right) \text{ s.t } R_t = \tilde{R}_t + R_t^{otp} + R_t^u \tag{3.26}$$

$$R_{k_t} = \left(R_{k_t}', R_{k_t}^{otp}, R_{k_t}^u\right) \text{ s.t } R_{k_t} = R_{k_t}' + R_{k_t}^{otp} + R_{k_t}^u. \tag{3.27}$$

Let us first choose

$$R_{k_t}' = I(V_t; Y_{\bar{t}}|V_{\bar{t}}, U) - \epsilon_1' \tag{3.28}$$

Now a pair (R_1^{otp}, R_2^{otp}) is chosen such that

$$\begin{aligned} &0 \leq R_1^{otp} \leq R_{k_1} - R_{k_1}'; \;\; 0 \leq R_2^{otp} \leq R_{k_2} - R_{k_2}' \\ &0 \leq R_1^{otp} + R_2^{otp} \leq I(V_1; V_2|U) + \epsilon_1'. \end{aligned} \tag{3.29}$$

Let us set $R_{k_t}^{otp} = R_t^{otp}$. Let R_t^u and $R_{k_t}^u$ respectively denote the remaining parts of the message and key rates, which can be empty depending on the choice in (3.29). Generate $2^{n(R_1^u + R_2^u)}$ sequences $\mathbf{u}(l_1, l_2)$, for $t = 1, 2$, $l_t = 0, 1, 2, \ldots, 2^{nR_t^u} - 1$. For each $\mathbf{u}$, generate a satellite code $\mathcal{C}(\mathbf{u}) = \{\mathcal{C}_1(\mathbf{u}), \mathcal{C}_2(\mathbf{u})\}$. $\mathcal{C}_t(\mathbf{u})$ has $2^{nR_{k_t}'}$ codebooks, with $\mathbf{v}_t \sim \prod_{i=1}^n P_{V_t|U}(v_{ti}|u_i)$. A codebook contains $2^{n\tilde{R}_{tt}}$ codewords and is divided into $2^{n\tilde{R}_t}$ bins, each with $2^{nR^\dagger}$ codewords, giving:

$$\tilde{R}_t = \tilde{R}_{tt} - R^\dagger \tag{3.30}$$

Bins are divided evenly into $2^{nR_t^{otp}}$ sections. Every such pair – one each from $\mathcal{C}_1$ and $\mathcal{C}_2$ – contains $\geq 2^{n[I(V_1;V_2|U)+\epsilon_1']}$ $(\mathbf{v}_1, \mathbf{v}_2)$ pairs to satisfy mutual covering [27, Chapter 8], and so:

$$[R^\dagger - R_1^{otp}] + [R^\dagger - R_2^{otp}] \geq I(V_1; V_2|U) + \epsilon_1',$$

which gives (3.29) on simplification.

3.5.4.3 Encoding

(w_1^u, w_2^u) is protected by OTP by picking $\mathbf{u}(w_1^u \oplus k_1^u, w_2^u \oplus k_2^u)$. It searches inside the pair of sections $\mathcal{C}_t(\mathbf{u}(w_1^u \oplus k_1^u, w_2^u \oplus k_2^u), k_t', \tilde{w}_t, w_t^{otp} \oplus k_t^{otp})$ for $t = 1, 2$ for a $(\mathbf{v}_1, \mathbf{v}_2)$ pair that is jointly typical $\in A_{\epsilon_1'}^{(n)}(V_1, V_2|\mathbf{u}(w_1^u \oplus k_1^u, w_2^u \oplus k_2^u))$. By construction, the number of $(\mathbf{v}_1, \mathbf{v}_2)$ pairs $\geq 2^{n[I(V_1;V_2|U)+\epsilon_1']}$, so a jointly typical pair exists with very high probability, by mutual covering, [27, Chapter 8]. The encoder generates and transmits $\mathbf{x} \sim \prod_{i=1}^n P_{X|V_1,V_2}(x_i|v_{1i}, v_{2i})$.

3.5.4.4 Decoding

Receiver t uses simultaneous decoding of $(\mathbf{u}, \mathbf{v}_t)$ by joint typicality with $\mathbf{y}_t$. We concisely describe the error events and their respective rate constraints. For a simpler but similar calculation, see [27, Chapter 8, Theorem 8.4, also Exercise 8.10].

Receiver t incorrectly decodes both $\mathbf{u}$ and $\mathbf{v}_t$, constraining:

$$R_1^u + R_2^u + \tilde{R}_{tt} < I(U, V_t; Y_t) \tag{3.31}$$

Receiver t correctly decodes $\mathbf{u}$ but incorrectly decodes $\mathbf{v}_t$. Since she knows the codebook, this gives the rate constraint:

$$\tilde{R}_{tt} \leq I(V_t; Y_t|U) \tag{3.32}$$

By construction, mutual covering $(\tilde{R}_{11} - \tilde{R}_1) + (\tilde{R}_{22} - \tilde{R}_2) > I(V_1; V_2|U) + \epsilon_1'$ is satisfied for any arbitrary pair of sections inside the respective bins, hence by the bins themselves.

(3.30), (3.31), the definition of $R^\dagger$, and $R_{\bar{t}}^u \geq 0$, give:

$$R_t^u + \tilde{R}_t < I(U, V_t; Y_t) - I(V_1; V_2|U) \tag{3.33}$$

Adding R_t^{otp} to both sides of (3.33) and using (3.29), we obtain:

$$R_t \triangleq R_t^u + \tilde{R}_t + R_t^{otp} < I(U; Y_t) + I(V_t; Y_t|U) - I(V_1; V_2|U) + R_{k_t} - R_{k_t}' \tag{3.34}$$

We can choose the satellite codebook size $\tilde{R}_{tt}$ to be the upper bound (3.32). *This choice of* $\tilde{R}_{tt}$ gives upper bounds via, (3.30), on $\tilde{R}_t$ and, via (3.31), on $R_1^u + R_2^u$:

$$\tilde{R}_{t,max} = I(V_t; Y_t|U) - I(V_1; V_2|U) \tag{3.35}$$

$$(R_1^u + R_2^u)_{max} = \min\{I(U; Y_1), I(U; Y_2)\} \tag{3.36}$$

Since $R_t^u + R_t^{otp} \leq R_{k_t} - R_{k_t}'$, we also have:

$$R_t \triangleq \tilde{R}_t + (R_t^u + R_t^{otp}) < I(V_t; Y_t|U) - I(V_1; V_2|U) + R_{k_t} - R_{k_t}' \tag{3.37}$$

Comparing (3.34) with (3.37), (3.37) is tighter. Substituting for R_{k_t}' gives the individual rate constraint in (3.25). We now upper bound the total secure sum-rate:

$$R_1 + R_2 = (R_1^u + R_2^u) + \tilde{R}_1 + \tilde{R}_2 + (R_1^{otp} + R_2^{otp}) \tag{3.38}$$

Applying (3.29), (3.35) and (3.36) to the RHS of (3.38) gives (3.25):

$$R_1 + R_2 < I(V_1; Y_1|U) + I(V_2; Y_2|U) - I(V_1; V_2|U) + \min\{I(U; Y_1), I(U; Y_2)\} \tag{3.39}$$

3.5.4.5 Error Probability Analysis

The error probability analysis is standard, and there are no surprises.

3.5.4.6 Equivocation Calculation

For $t = 1, 2$, W_t^{otp}, W_t^u are protected by OTPs and are secure. $\tilde{W}_t$ is protected by the scheme developed for $\tilde{W}_1$ of receiver 1 in Case 2.

3.5.4.7 Cases $5, 6, 7$ and Other Cases

In Case 5, an individual $\mathbf{v}_1$ sequence can be used to encode a different message, consequently rates $R_1 > I(V_1; Y_1|U)$ attainable and bottlenecks on the individual rate constraints (3.34), and thence $I(U, V_1; Y_1)$ become active. In Case 6, the same holds for $\mathbf{v}_2$. Note that when $R_{k_1} = R_{k_2} = 0$, the achievable region simplifies to [65], and when $R_2 = 0$, we get [67].

3.5.5 Generalization(s)

The following two generalizations are standard in these scenarios:

- Transmitting a common message.
- Obtaining the rate-equivocation region.

3.6 Outer Bound(s)

Theorem 18. *Let $\mathbb{R}_O(\pi_{BC})$ denote the union of all (R_1, R_2) satisfying*

$$\begin{aligned}
&R_1 \geq 0, R_2 \geq 0\\
&R_1 \leq \min\{\ I(V_1; Y_1|U) + \min\{R_{k_1} - I(V_1; Y_2|U), I(U; Y_1), I(U; Y_2)\},\\
&\qquad\qquad I(V_1; Y_1|V_2, U) + \min\{R_{k_1} - I(V_1; Y_2|V_2, U), I(U, V_2; Y_1), I(U, V_2; Y_2)\}\}\\
&R_2 \leq \min\{\ I(V_2; Y_2|U) + \min\{R_{k_2} - I(V_2; Y_1|U), I(U; Y_1), I(U; Y_2)\},\\
&\qquad\qquad I(V_2; Y_2|V_1, U) + \min\{R_{k_2} - I(V_2; Y_1|V_1, U), I(U, V_2; Y_1), I(U, V_1; Y_1)\}\}
\end{aligned} \tag{3.40}$$

over all distributions P_{U,V_1,V_2,X,Y_1,Y_2} in π_{BC} and auxiliary random variables U, V_1, V_2 satisfying

$$U \rightarrow V_1 \rightarrow X \text{ and } U \rightarrow V_1 \rightarrow X$$

For the broadcast channel with secret keys and with confidential messages, the capacity region

$$\mathbb{C}_{BC} \subseteq \mathbb{R}_O(\pi_{BC})$$

Note that $R_{k_1} = R_{k_2} \overset{\text{set}}{=} 0$ simplifies to the outer bound for the broadcast channel with mutual secrecy requirements studied by [65].

3.6.1 Outer Bound Proofs

Consider R_1 in (3.40). The first term inside the outer minimization corresponds to the receiver 2 attempting to eavesdrop without having decoded its own message, hence no conditioning on V_2. The second term inside the minimization occurs when receiver 2 attempts to decode the message of receiver 1 after decoding its own message, hence the terms are conditioned on V_2. The proof for the first term inside each outer minimization follows closely the associated converse proof in [67]. The proof for the second term inside each outer minimization is more involved, and uses the technique employed in the second outer bound obtained in [65, Section IV-B], where a genie gives receiver 1 the message and key (W_2, K_2), while receiver 2 attempts to evaluate the equivocation with (W_2, K_2) as side information.

We will only prove the bounds for R_1. The corresponding inequality for R_2 follows by symmetry.

3.6.1.1 First Bound: The other receiver attempts to eavesdrop without first decoding its own message/codeword

Unlike in the case of achievability proofs, where we followed the techniques in [65] with appropriate changes due to the presence of secret keys as in [67], here we primarily follow the proof technique in [67], with modifications appropriate to our model. The modifications play an important role in obtaining the second outer bound, as they are suggested by the second bound obtained by [65].

We restate the following inequalities [67, equations (8) and (9)] in terms of our notation where $Y \leftarrow Y_1$ and $Z \leftarrow Y_2$ Note that the inequalities in [67] are themselves taken from Csiszár and Korner's textbook [147, p. 314, equation (3.34)].

For ease of reference, we (re-)derive the following equality

$$\begin{aligned} &H(\mathbf{Y}_{1,1}^{n}) - H(\mathbf{Y}_{2,1}^{n}) = \\ &\sum_{i=1}^{n} [H(\mathbf{Y}_{1i}|\mathbf{Y}_{1,i+1}^{n}, \mathbf{Y}_{2,1}^{i-1}) - H(\mathbf{Y}_{2i}|\mathbf{Y}_{1,i+1}^{n}, \mathbf{Y}_{2,1}^{i-1})] \end{aligned} \tag{3.41}$$

To the LHS of (3.41), add and subtract $H(\mathbf{Y}_{2,1}, \mathbf{Y}^n_{1,2})$ to get

$$\begin{aligned}
&H(\mathbf{Y}^n_{1,1}) - H(\mathbf{Y}^n_{2,1}) = \\
&H(\mathbf{Y}^n_{1,1}) - H(\mathbf{Y}_{2,1}, \mathbf{Y}^n_{1,2}) + H(\mathbf{Y}_{2,1}, \mathbf{Y}^n_{1,2}) - H(\mathbf{Y}^n_{2,1}) = \\
&\big(H(\mathbf{Y}_{1,1}|\mathbf{Y}^n_{1,2}) + \cancel{H(\mathbf{Y}^n_{1,2})}\big) - \big(H(\mathbf{Y}_{2,1}|\mathbf{Y}^n_{1,2})) + \cancel{H(\mathbf{Y}^n_{1,2})}\big) + \\
&\big(H(\mathbf{Y}^n_{1,2}|\mathbf{Y}_{2,1}) + \cancel{H(\mathbf{Y}_{2,1})}\big) - \big(\cancel{H(\mathbf{Y}_{2,1})} + H(\mathbf{Y}^n_{2,2}|\mathbf{Y}_{2,1})\big)
\end{aligned} \tag{3.42}$$

thus obtaining

$$\begin{gathered}
H(\mathbf{Y}^n_{1,1}) - H(\mathbf{Y}^n_{2,1}) = \\
\big(H(\mathbf{Y}_{1,1}|\mathbf{Y}^n_{1,2}) - H(\mathbf{Y}_{2,1}|\mathbf{Y}^n_{1,2}) + \\
H(\mathbf{Y}^n_{1,2}|\mathbf{Y}_{2,1}) - H(\mathbf{Y}^n_{2,2}|\mathbf{Y}_{2,1})
\end{gathered} \tag{3.43}$$

Now, consider the second line of (3.43), namely

$$H(\mathbf{Y}^n_{1,2}|\mathbf{Y}_{2,1}) - H(\mathbf{Y}^n_{2,2}|\mathbf{Y}_{2,1}) \tag{3.44}$$

Note that this resembles what we started with, namely, the LHS of (3.41) with the changes that we have an extra conditioning on $\mathbf{Y}_{2,1}$, and $\mathbf{Y}^n_{1,1} \leftarrow \mathbf{Y}^n_{1,2}$ and $\mathbf{Y}^n_{2,1} \leftarrow \mathbf{Y}^n_{2,2}$. So, in analogy with (3.43), we can write

$$\begin{gathered}
H(\mathbf{Y}^n_{1,2}|\mathbf{Y}_{2,1}) - H(\mathbf{Y}^n_{2,2}|\mathbf{Y}_{2,1}) = \\
\big(H(\mathbf{Y}_{1,2}|\mathbf{Y}^n_{1,3}, \mathbf{Y}_{2,1}) - H(\mathbf{Y}_{2,2}|\mathbf{Y}^n_{1,3}, \mathbf{Y}_{2,1}) + \\
H(\mathbf{Y}^n_{1,3}|\mathbf{Y}_{2,2}, \mathbf{Y}_{2,1}) - H(\mathbf{Y}^n_{2,3}|\mathbf{Y}_{2,2}, \mathbf{Y}_{2,1})
\end{gathered} \tag{3.45}$$

As before, we expand only the second line, namely

$$H(\mathbf{Y}^n_{1,3}|\mathbf{Y}_{2,2}, \mathbf{Y}_{2,1}) - H(\mathbf{Y}^n_{2,3}|\mathbf{Y}_{2,2}, \mathbf{Y}_{2,1}) \tag{3.46}$$

We proceed iteratively. Note that

$$\begin{aligned}
&[H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1})]\Big|_{i=1} \\
&\qquad = H(\mathbf{Y}_{1,1}|\mathbf{Y}^n_{1,2}) - H(\mathbf{Y}_{2,1}|\mathbf{Y}^n_{1,2})
\end{aligned} \tag{3.47}$$

which was the first line on the RHS of (3.43). Similarly

$$\begin{aligned}
&[H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1})]\Big|_{i=2} \\
&\qquad = H(\mathbf{Y}_{1,2}|\mathbf{Y}^n_{1,3}, \mathbf{Y}_{2,1}) - H(\mathbf{Y}_{2,2}|\mathbf{Y}^n_{1,3}, \mathbf{Y}_{2,1})
\end{aligned} \tag{3.48}$$

which was the first line on the RHS of (3.45). Iterating, we finally expand the RHS as

$$\sum_{i=1}^{n} [H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1})]$$

to obtain

$$H(\mathbf{Y}^n_{1,1}) - H(\mathbf{Y}^n_{2,1}) = \sum_{i=1}^{n}[H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1})] \tag{3.49}$$

which is the same as (3.41), which was to be proved. We can also condition on K_2 and on $(W_1, K_1, K_2))$ and derive the following equalities analogously:

$$H(\mathbf{Y}^n_{1,1}|K_2) - H(\mathbf{Y}^n_{2,1}|K_2) = \sum_{i=1}^{n}[H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}, K_2) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}, K_2)].$$

and

$$H(\mathbf{Y}^n_{1,1}|W_1, K_1, K_2) - H(\mathbf{Y}^n_{2,1}|W_1, K_1, K_2) = \sum_{i=1}^{n}[H(\mathbf{Y}_{1i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}, W_1, K_1, K_2) - H(\mathbf{Y}_{2i}|\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}, W_1, K_1, K_2)].$$

We define the auxiliary RVs:

$$U_i \triangleq (\mathbf{Y}^n_{1,i+1}, \mathbf{Y}^{i-1}_{2,1}) \tag{3.50}$$

We also define a time-sharing RV Q, which is independent of everything else, and is uniform on the set $\{1, 2, 3, \ldots, n\}$. With this definition of U and Q, we further define the following RVs:

$$U \triangleq (U_Q, Q), \tilde{U} \triangleq (U, K_2),\ V_1 \triangleq (\tilde{U}, W_1, K_1)$$
$$X \triangleq X_Q,\ Y_1 \triangleq Y_{1Q},\ Y_2 \triangleq Y_{2Q}$$

Note that the Markov chain condition $\tilde{U} \to V_1 \to X \to (Y_1, Y_2)$ is satisfied. Using the above equations and the definitions of the auxiliary RVs, we can show that [147, p. 314, equation (3.34)] $\exists t, t_{K_2}, t_{(W_1,K_1,K_2)} \in \mathbb{R}$ s.t.

$$\frac{1}{n}H(\mathbf{Y}^n_{1,1}) = H(Y_1|U) + t \tag{3.51}$$

$$\frac{1}{n}H(\mathbf{Y}^n_{2,1}) = H(Y_2|U) + t \tag{3.52}$$

$$\frac{1}{n}H(\mathbf{Y}^n_{1,1}|K_2) = H(Y_1|U, K_2) + t_{K_2} = H(Y_1|\tilde{U}) + t_{K_2} \tag{3.53}$$

$$\frac{1}{n}H(\mathbf{Y}^n_{2,1}|K_2) = H(Y_2|U, K_2) + t_{K_2} = H(Y_2|\tilde{U}) + t_{K_2} \tag{3.54}$$

and also

$$\begin{aligned}\frac{1}{n}H(\mathbf{Y}_{1,1}^n|W_1,K_1,K_2) &= H(Y_1|V_1) + t_{(W_1,K_1,K_2)}\\ &= H(Y_1|\tilde{U},V_1) + t_{(W_1,K_1,K_2)} \qquad (3.55)\\ \frac{1}{n}H(\mathbf{Y}_{2,1}^n|W_1,K_1,K_2) &= H(Y_2|V_1) + t_{(W_1,K_1,K_2)}\\ &= H(Y_2|\tilde{U},V_1) + t_{(W_1,K_1,K_2)} \qquad (3.56)\end{aligned}$$

where the last equality in both equations (3.74) and (3.75) above follows due to the Markov Chain condition $\tilde{U} \to V_1 \to Y_1$ and $\tilde{U} \to V_1 \to Y_2$ where

$$0 \leq t \leq \min\{I(U;Y_1), I(U;Y_2)\} \qquad (3.57)$$

$$0 \leq t_{K_2} \leq \min\{I(\tilde{U};Y_1), I(\tilde{U};Y_2)\} \qquad (3.58)$$

$$0 \leq t_{(W_1,K_1,K_2)} \leq \min\{I(V_1;Y_1), I(V_1;Y_2)\} \qquad (3.59)$$

Since the code satisfies the information leakage constraint, namely

$$\begin{aligned}n\mu &\geq I(W_1;\mathbf{Y}_{2,1}^n|K_2)\\ &= I(W_1,K_1;\mathbf{Y}_{2,1}^n|K_2) - I(K_1;\mathbf{Y}_{2,1}^n|W_1,K_2)\\ &= H(\mathbf{Y}_{2,1}^n|K_2) - H(\mathbf{Y}_{2,1}^n|W_1,K_1,K_2)\\ &\qquad - H(K_1|W_1,K_2) + H(K_1|\mathbf{Y}_{2,1}^n,W_1,K_2)\\ &\geq H(\mathbf{Y}_{2,1}^n|K_2) - H(\mathbf{Y}_{2,1}^n|W_1,K_1,K_2) - H(K_1|W_1,K_2)\\ &\overset{\because K_1 \perp\!\!\!\perp (W_1,K_2)}{=}\\ &H(\mathbf{Y}_{2,1}^n|K_2) - H(\mathbf{Y}_{2,1}^n|W_1,K_1,K_2) - H(K_1)\\ &= H(\mathbf{Y}_{2,1}^n|K_2) - H(\mathbf{Y}_{2,1}^n|W_1,K_1,K_2) - nR_{k_1}\\ &= n(H(Y_2|\tilde{U}) + t_{K_2} - H(Y_2|V_1) - t_{(W_1,K_1,K_2)} - R_{k_1})\\ &\overset{\because \tilde{U}\to V_1\to Y_2}{=}\\ &n(H(Y_2|\tilde{U}) + t_{K_2} - H(Y_2|V_1,\tilde{U}) - t_{(W_1,K_1,K_2)} - R_{k_1})\\ &= n(I(V_1;Y_2|\tilde{U}) + t_{K_2} - t_{(W_1,K_1,K_2)} - R_{k_1}) \qquad (3.60)\end{aligned}$$

Therefore

$$\begin{aligned}\cancel{n}\mu &\geq \cancel{n}(I(V_1;Y_2|\tilde{U}) + t_{K_2} - t_{(W_1,K_1,K_2)} - R_{k_1})\\ &\implies t_{K_2} - t_{(W_1,K_1,K_2)} \leq R_{k_1} - I(V_1;Y_2|\tilde{U}) + \mu \qquad (3.61)\end{aligned}$$

We also have

$$t_{K_2} - t_{(W_1,K_1,K_2)} \leq t_{K_2} \leq \min\{I(\tilde{U};Y_1), I(\tilde{U};Y_2)\} \qquad (3.62)$$

Thus, from (3.61) and (3.62), we have

$$t_{K_2} - t_{(W_1,K_1,K_2)} \leq \min\{R_{k_1} - I(V_1;Y_2|\tilde{U}) + \mu, I(\tilde{U};Y_1), I(\tilde{U};Y_2)\} \tag{3.63}$$

Now

$$\begin{aligned}
|\mathcal{W}_1| &= H(W_1) \overset{W_1 \perp\!\!\!\perp (K_1,K_2)}{=} H(W_1|K_1,K_2) \\
&= H(W_1|\mathbf{Y}_{1,1}^n, K_1, K_2) + I(W_1;\mathbf{Y}_{1,1}^n|K_1,K_2) \\
&\overset{\text{Fano}}{\leq} I(W_1;\mathbf{Y}_{1,1}^n|K_1,K_2) + n\epsilon_n \\
&\leq I(W_1,K_1;\mathbf{Y}_{1,1}^n|K_2) + n\epsilon_n \\
&= H(\mathbf{Y}_{1,1}^n|K_2) - H(\mathbf{Y}_{1,1}^n|W_1,K_1,K_2) + n\epsilon_n \\
&= n\big[H(Y_1|\tilde{U}) + t_{K_2} - H(Y_1|V_1) - t_{(W_1,K_1,K_2)} + \epsilon_n\big] \\
&\overset{\tilde{U}\to V_1 \to Y_1}{=} \\
&n\big[H(Y_1|\tilde{U}) + t_{K_2} - H(Y_1|V_1,\tilde{U}) - t_{(W_1,K_1,K_2)} + \epsilon_n\big] \\
&= n\big[I(Y_1;V_1|\tilde{U}) + (t_{K_2} - t_{(W_1,K_1,K_2)}) + \epsilon_n\big] \\
&\leq n\big[I(Y_1;V_1|\tilde{U}) + \min\{R_{k_1} - I(V_1;Y_2|\tilde{U}) + \mu, I(\tilde{U};Y_1), I(\tilde{U};Y_2)\} + \epsilon_n\big]
\end{aligned} \tag{3.64}$$

Together with $H(W_1) = nR_1$, equation (3.64) implies that

$$R_1 \leq I(Y_1;V_1|\tilde{U}) + \epsilon_n + \min\Big\{R_{k_1} - I(V_1;Y_2|\tilde{U}) + \mu, I(\tilde{U};Y_1), I(\tilde{U};Y_2)\Big\} \tag{3.65}$$

which gives (as μ and ϵ_n can be made arbitrarily small)

$$R_1 \leq I(Y_1;V_1|\tilde{U}) + \min\{R_{k_1} - I(V_1;Y_2|\tilde{U}), I(\tilde{U};Y_1), I(\tilde{U};Y_2)\} \tag{3.66}$$

which is the first term inside the outer minimization in (3.40)

3.6.1.2 Second Bound: Where the other receiver decodes its own codeword before eavesdropping

The bound is obtained by considering that:

- a genie gives receiver 1 message-key pair (W_2, K_2)
- receiver 2 attempts to evaluate the equivocation with (W_2, K_2) as side information

This is inspired by [65]. We rewrite the equations/inequalities used in [67, Equations (18) and (22)] employing these insights.

Consider the inequality (note that this follows very closely the corresponding chain of inequalities in [67] with the crucial change of additional conditioning RVs (W_2, K_2)

$$\begin{aligned}|\mathcal{W}_1| &= H(W_1) \\ &= H(W_1|K_1) && \text{Since } W_1 \perp\!\!\!\perp K_1 \\ &= H(W_1|K_1, W_2, K_2) && \text{Since } (W_1, K_1) \perp\!\!\!\perp (W_2, K_2)\end{aligned} \tag{3.67}$$

We interpret the last equation above, (3.67), as a genie giving (W_2, K_2) to receiver 1. We continue

$$\begin{aligned}H(W_1) &= H(W_1|K_1, W_2, K_2) \\ &= H(W_1|K_1, W_2, K_2) - H(W_1|\mathbf{Y}^n_{1,1}, K_1, W_2, K_2) + H(W_1|\mathbf{Y}^n_{1,1}, K_1, W_2, K_2) \\ &= I(W_1; \mathbf{Y}^n_{1,1}|K_1, W_2, K_2) + H(W_1|\mathbf{Y}^n_{1,1}, K_1, W_2, K_2) \\ &\leq I(W_1; \mathbf{Y}^n_{1,1}|K_1, W_2, K_2) + H(W_1|\mathbf{Y}^n_{1,1}, K_1) \\ &\overset{\text{Fano}}{\leq} I(W_1; \mathbf{Y}^n_{1,1}|K_1, W_2, K_2) + n\epsilon_n \\ &\leq I(W_1, K_1; \mathbf{Y}^n_{1,1}|W_2, K_2) + n\epsilon_n \\ &= H(\mathbf{Y}^n_{1,1}|W_2, K_2) - H(\mathbf{Y}^n_{1,1}|W_1, K_1, W_2, K_2) + n\epsilon_n\end{aligned} \tag{3.68}$$

Analogously to (3.53), (3.54), (3.74) and (3.75) it can be s.t. that $\exists t_{(W_2,K_2)}, t_{(W_1,K_1,W_2,K_2)} \in \mathbb{R}$ s.t.

$$\begin{aligned}\frac{1}{n}H(\mathbf{Y}^n_{1,1}|W_2, K_2) &= (1/n)H(\mathbf{Y}^n_{1.1}|K_2)(W_2) \\ &= (H(Y_1|U, K_2) + t_{K_2})(W_2) \\ &= H(Y_1|U, K_2, W_2) + t_{(W_2,K_2)} \\ &= H(Y_1|\tilde{U}, W_2) + t_{(W_2,K_2)}\end{aligned} \tag{3.69}$$

$$\begin{aligned}\frac{1}{n}H(\mathbf{Y}^n_{2,1}|W_2, K_2) &= H(Y_2|U, K_2, W_2) + t_{(W_2,K_2)} \\ &= H(Y_2|\tilde{U}, W_2) + t_{(W_2,K_2)}\end{aligned} \tag{3.70}$$

Define

$$V_2 \triangleq W_2,\ \tilde{V}_2 \triangleq (\tilde{U}, V_2) \tag{3.71}$$

Note that the Markov Chain $\tilde{U} \to \tilde{V}_2 \to X \to (Y_1, Y_2)$ is satisfied. We have

$$\begin{aligned}\frac{1}{n}H(\mathbf{Y}^n_{1,1}|W_2, K_2) = H(Y_1|\tilde{U}, V_2) + t_{(V_2,K_2)} &= H(Y_1|\tilde{V}_2) + t_{(V_2,K_2)} \\ &= H(Y_1|\tilde{V}_2, \tilde{U}) + t_{(V_2,K_2)}\end{aligned} \tag{3.72}$$

$$\begin{aligned}\frac{1}{n}H(\mathbf{Y}^n_{2,1}|W_2, K_2) = H(Y_2|\tilde{U}, V_2) + t_{(V_2,K_2)} &= H(Y_2|\tilde{V}_2) + t_{(V_2,K_2)} \\ &= H(Y_2|\tilde{V}_2, \tilde{U}) + t_{(V_2,K_2)}\end{aligned} \tag{3.73}$$

and also

$$\begin{aligned}
&\frac{1}{n}H(\mathbf{Y}_{1,1}^{n}|W_2,K_2,W_1,K_1)\\
&\qquad = \big(H(Y_1|U)+t\big)(W_1,K_1,W_2,K_2)\\
&\qquad = H(Y_1|U,W_1,K_1,W_2,K_2)+t_{(W_1,K_1,W_2,K_2)}\\
&\qquad = H(Y_1|\tilde{U},V_1,\tilde{V}_2)+t_{(W_1,K_1,W_2,K_2)} \qquad (3.74)\\
&\frac{1}{n}H(\mathbf{Y}_{2,1}^{n}|W_2,K_2,W_1,K_1)\\
&\qquad = \big(H(Y_2|U)+t\big)(W_1,K_1,W_2,K_2)\\
&\qquad = H(Y_2|U,W_1,K_1,W_2,K_2)+t_{(W_1,K_1,W_2,K_2)}\\
&\qquad = H(Y_2|\tilde{U},V_1,\tilde{V}_2)+t_{(W_1,K_1,W_2,K_2)} \qquad (3.75)
\end{aligned}$$

We finally have

$$\begin{aligned}
H(W_1) &\le n\big(H(Y_1|\tilde{U},\tilde{V}_2)+t_{(V_2,K_2)}-H(Y_1|\tilde{U},\tilde{V}_2,V_1)-t_{(W_1,K_1,W_2,K_2)}+\epsilon_n\big)\\
&= n\big(I(V_1;Y_1|\tilde{V}_2,\tilde{U})+t_{(V_2,K_2)}-t_{(W_1,K_1,W_2,K_2)}+\epsilon_n\big) \qquad (3.76)
\end{aligned}$$

Now, since, $H(W_1)=nR_1$, substituting in (3.76), we get

$$R_1 \le I(V_1;Y_1|\tilde{V}_2,\tilde{U})+t_{(V_2,K_2)}-t_{(W_1,K_1,W_2,K_2)}+\epsilon_n \qquad (3.77)$$

Since the information leakage condition is satisfied, and the receiver 2 attempts to evaluate the equivocation with (W_2,K_2) as side information, we can write

$$\begin{aligned}
n\mu &\ge I(W_1;\mathbf{Y}_{2,1}^{n}|W_2,K_2)\\
&= I(W_1,K_1;\mathbf{Y}_{2,1}^{n}|W_2,K_2)-I(K_1;\mathbf{Y}_{2,1}^{n}|W_1,W_2,K_2)\\
&= \big(H(\mathbf{Y}_{2,1}^{n}|W_2,K_2)-H(\mathbf{Y}_{2,1}^{n}|W_1,K_1,W_2,K_2)\big)\\
&\quad -\big(H(K_1|W_2,K_2,W_1)-H(K_1|\mathbf{Y}_{2,1}^{n},W_2,K_2,W_1)\big)\\
&\ge H(\mathbf{Y}_{2,1}^{n}|W_2,K_2)-H(\mathbf{Y}_{2,1}^{n}|W_1,K_1,W_2,K_2)\\
&\qquad\qquad -H(K_1|W_2,K_2,W_1). \qquad (3.78)
\end{aligned}$$

Now, by mutual independence of (W_1,K_1,W_2,K_2), we simplify the last term in (3.78) as

$$H(K_1|W_2,K_2,W_1)=H(K_1) \qquad (3.79)$$

(3.78) and (3.79) together imply that

$$\begin{aligned}
n\mu &\ge H(\mathbf{Y}_{2,1}^{n}|W_2,K_2)-H(\mathbf{Y}_{2,1}^{n}|W_1,K_1,W_2,K_2)-H(K_1)\\
&= n\big(H(Y_2|\tilde{U},\tilde{V}_2)+t_{(V_2,K_2)}-H(Y_2|\tilde{U},\tilde{V}_2,V_1)-t_{(W_1,K_1,W_2,K_2)}-R_{k_1}\big) \qquad (3.80)
\end{aligned}$$

Simplifying gives:

$$\mu \geq I(V_1;Y_2|\tilde{V}_2,\tilde{U}) + t_{(V_2,K_2)} - t_{(W_1,K_1,W_2,K_2)} - nR_{k_1}$$

Rearranging the above gives:

$$t_{(V_2,K_2)} - t_{(W_1,K_1,W_2,K_2)} \leq R_{k_1} - I(V_1;Y_2|\tilde{V}_2,\tilde{U}) + \mu \tag{3.81}$$

We also have

$$t_{(V_2,K_2)} - t_{(W_1,K_1,W_2,K_2)} \leq t_{(V_2,K_2)} \leq \min\{I(\tilde{U},\tilde{V}_2;Y_1), I(\tilde{U},\tilde{V}_2;Y_2)\} \tag{3.82}$$

Inequalities (3.81) and (3.82) together give

$$t_{(V_2,K_2)} - t_{(W_1,K_1,W_2,K_2)} \leq \min\{R_{k_1} - I(V_1;Y_2|\tilde{V}_2,\tilde{U}) + \mu, I(\tilde{U},\tilde{V}_2;Y_1), I(\tilde{U},\tilde{V}_2;Y_2)\} \tag{3.83}$$

On substituting (3.83) into the inequality (3.77), we get

$$R_1 \leq I(V_1;Y_1|\tilde{V}_2,\tilde{U}) + \epsilon_n + \min\{R_{k_1} - I(V_1;Y_2|\tilde{V}_2,\tilde{U}) + \mu, I(\tilde{U},\tilde{V}_2;Y_1), I(\tilde{U},\tilde{V}_2;Y_2)\}$$

which, since ϵ_n, μ can be arbitrarily small, gives

$$R_1 \leq I(V_1;Y_1|\tilde{V}_2,\tilde{U}) + \min\{R_{k_1} - I(V_1;Y_2|\tilde{V}_2,\tilde{U}), I(\tilde{U},\tilde{V}_2;Y_1), I(\tilde{U},\tilde{V}_2;Y_2)\}$$

which is the second term in the outer minimization for R_1 in (3.40).

Chapter 4

Relay Eavesdropper Channel under Decode and Forward

4.1 Introduction

The relay channel models a situation where a causal helper node aids the transmissions from the source to a destination. More specifically, a full duplex relay node listens to a channel from the source, and its transmissions to the destination are a function of the past received symbols. In a Decode-Forward (DF) scheme of operation, the relay first decodes the messages from the source and then cooperates with the source in de-mystifying the uncertainty at the receiver. The DF scheme is known to be optimal when the receiver is physically degraded with respect to the relay [27]. An alternate method is where the relay compresses the received symbols without decoding the actual messages, and then forwards the compressed values to the destination. This is called Compress-Forward (CF) scheme. Vector quantizers are commonly employed to achieve compression in this context. In a Gaussian relay setting, typically used in wireless models, it is popular for the relay to simply scale and retransmit the received values, a technique known as Amplify-Forward (AF). Notice that AF is a version of CF. Conventional CF and DF are also widely used in Gaussian relay models [27].

In this chapter, we study a relay wiretap channel where an eavesdropper named Eve listens to the transmissions in a conventional relay channel. The legitimate receiver is required to decode the encoded message reliably, whereas the message is required to be secret from the eavesdropper. The relay itself is allowed to gain information about the message, and is trusted not to reveal the information to Eve. We present an achievable scheme using multi-block encoding where the relay decodes and forwards (DF), and the legitimate receiver decodes using a sliding window. While, the content of this chapter will be extended in the next chapter to the case of individual private messages to two legitimate receivers, where each receiver attempts to eavesdrop on the unintended messages, the underlying techniques are better explained for a single receiver and eavesdropper.

The relay model with eavesdropper was studied by Lai and Gamal [72, Theorem 2], who employed the so called *regular encoding*, along with a DF scheme. The legitimate receiver performs backward decoding under block Markov encoding, where data is decoded only after all the blocks of received symbols are available. In contrast, the approach here is forward decoding using a sliding window (SW). Though the achieved rate is the same as that in [72], our scheme incurs less decoding delay (see Remark 21) and it also extends naturally to a multiple receiver model in the next chapter.

4.2 System Model and Notation

The system model is depicted in Fig. 5.1. We assume a two receiver discrete memoryless relay eavesdropper channel with a confidential message intended for one of the receivers, with the other acting as an eavesdropper. The finite sets $\mathcal{X}_1, \mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Y}_3$, $\mathcal{Z}$ respectively represent the channel's input at node 1 (the transmitter), at node 2 (the relay), the channel's output at nodes 2, 3 (legitimate receiver aka Bob) and 4 (eavesdropper aka Eve). The channel is described by the conditional probability distribution $P_{Y_2,Y_3,Z|X_1,X_2}$, where RVs $X_i \in \mathcal{X}_i$, $i = 1, 2$ and $Y_i \in \mathcal{Y}_i$, $i = 2, 3$ and $Z \in \mathcal{Z}$. The transmitter intends to send an independent message $W \in \{1, 2, \ldots, 2^{nR}\} \triangleq \mathcal{W}$ to the receiver Rx Y_3 in n channel uses while ensuring information theoretic secrecy, defined below. The channel is memoryless and without feedback i.e. $\forall(\mathbf{x}_1, \mathbf{x}_2) \in \prod_{t=1}^{2} \mathcal{X}_t^n, \mathbf{y}_t \in \mathcal{Y}_t^n$, $t = 2, 3, \mathbf{z} \in \mathcal{Z}^n$,

$$P(\mathbf{y}_2, \mathbf{y}_3, \mathbf{z}|\mathbf{x}_1, \mathbf{x}_2) = \prod_{i=1}^{n} P_{Y_2,Y_3,Z|X_1,X_2}(y_{2i}, y_{3i}, z_i|x_{1i}, x_{2i})$$

Remark 19. *Note that we could also have specified the channel by the marginal transition probabilities $P_{Y_2|X_1}$ and $P_{Y_3,Z|X_1,X_2}$. The availability of joint laws does not change the results of the DF model considered.*

The decoding function at the relay is a mapping $\phi_2 : \mathcal{Y}_2^n \to \mathcal{W} \times \mathcal{S}$. In the preceding, $S \in \{1, 2, \ldots, 2^{n[R_2-R]}\} \triangleq \mathcal{S}$. The decoding function at the legitimate receiver $\equiv Y_3$ is a map $\phi_3 : \mathcal{Y}_3^n \times \mathcal{Y}_3^n \to \mathcal{W} \times \mathcal{S}$.

A $(2^{nR}, n, P_e^{(n)})$ code for the relay eavesdropper channel consists of two (stochastic) encoding functions, two decoding functions ϕ_t, $t = 2, 3$, and the error probability

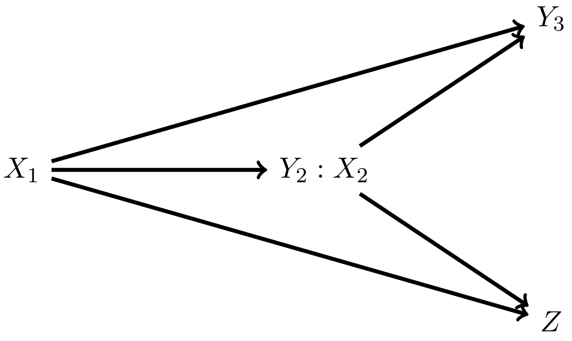

Figure 4.1: Relay eavesdropper channel

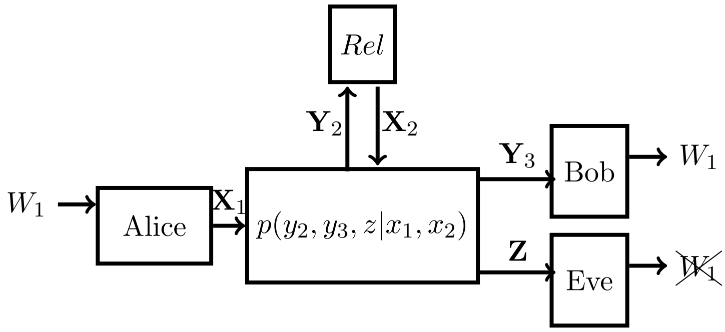

Figure 4.2: Relay-Eavesdropper channel with trusted relay and a confidential message intended for Bob. RV X_2: relay's input to the channel. RV Y_2: the channel's output as seen by the relay

$P_e^{(n)} \triangleq \max\{P_{e,2}^{(n)}, P_{e,3}^{(n)}\}$ where

$$P_{e,2}^{(n)} = \sum_{(w,s)} \frac{\Pr[\phi_2(\mathbf{Y}_2) \neq (w,s)|(w,s)]}{2^{nR} \times 2^{n[R_2-R]}}$$
$$P_{e,3}^{(n)} = \sum_{(w,s)} \frac{\Pr[\phi_3(\mathbf{Y}_3 \times \mathbf{Y}_3) \neq (w,s)|w,s]}{2^{nR} \times 2^{n[R_2-R]}}$$

A secrecy rate R is said to be achievable for the DM relay eavesdropper channel if, for any $\epsilon_0 > 0$, $\exists(2^{nR}, n, P_e^{(n)})$ code s.t. the following requirements are satisfied:

$$\text{reliability: } P_e^{(n)} \leq \epsilon_0$$
$$\text{(weak) secrecy: } n(B-1)R_1 - H(W^{[B-1]}|\mathbf{Z}^{[B]}) \leq n(B-1)\epsilon_0$$

The last constraint is the weak secrecy constraint [65]. We use the notation $\bar{t} \triangleq \{1,2\} \setminus \{t\}$, and $t^{[j]} \triangleq \{t_1, t_2, \ldots, t_j\}$.

4.3 The Achievable Rate

Theorem 20. *[72] Any rate R satisfying*

$$R < \max_{P_{X_1X_2}} \min\big\{I(X_1, X_2; Y_3),\, I(X_1; Y_2|X_2)\big\} - I(X_1, X_2; Z)$$

is achievable.

Notice that the above rate was also achieved by a scheme in [72], which employs a backward decoding scheme. As pointed out in the beginning of the chapter, we consider a forward decoding scheme, which has significantly lower delay in decoding (see Remark 21) than the backward decoding of [72]. Furthermore, the scheme here will be shown to have a natural generalization while broadcasting independent messages to two receivers over a relay channel under a mutual secrecy requirement. We now describe an achievable scheme using block-Markov encoding, and forward decoding using sliding windows over successive blocks.

4.4 Achievable Scheme

For a given distribution $p_{X_1X_2}$, we will show the achievability of a rate

$$R \leq \min\big\{I(X_1, X_2; Y_3),\, I(X_1; Y_2|X_2)\big\} - I(X_1, X_2; Z) - \epsilon, \forall \epsilon > 0. \tag{4.1}$$

We choose two quantities $R_1 > R_2 > R$ such that

$$R_1 = R + I(X_1, X_2; Z) - \epsilon/2 \tag{4.2}$$

$$\text{and } R_2 > R + I(X_2; Z). \tag{4.3}$$

Both conditions arise from the need for perfect secrecy. ϵ is a "small" quantity s.t. $n\epsilon \overset{n\uparrow\infty}{\to} 0$. The equality constraint on R_1 determines the packing $R_1 - R \approx I(X_1, X_2; Z)$ of the transmitter bins, which is identical to that of [72, Theorem 2]. The strict lower bound on R_2 gives in turn a strict lower bound on the packing $R_2 - R$ of a relay bin. Equivalently, we may write $R_2 = R + I(X_2; Z) + \Delta$, $\Delta \in (0, I(X_1; Z|X_2) - \epsilon/2]$. This constraint (which does not appear in [72, Theorem 2] which uses regular encoding) reflects the minimum size of a relay bin to ensure perfect secrecy.

The block Markov transmission scheme involves nB channel uses, over B blocks of n channel uses each. For each block $j \in [1 : B]$, $\mathcal{C}_1^{(j)}$ denotes the code used by the transmitter, $\mathcal{C}_2^{(j)}$ is the code used by the relay, and the pair of these codes is denoted by $\mathcal{C}^{(j)}$. The complete code is denoted by $\mathcal{C} \overset{\Delta}{=} (\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(B)})$. For any j, we denote $\mathcal{C}^{[j]} := (\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(j)})$. Similarly for all relevant random variables, a superscript $[j]$ signifies the collection of random variables till block j. We describe the construction of the code below.

Codebook Generation: The following describes the codebook in an *individual block.* Let R, R_2, R_1 be three non-negative values in the increasing order.

- The relay codebook $\mathcal{C}_2^{(b)}$ in block b is given by

$$\mathcal{C}_2^{(b)} = \left\{ \mathbf{x}_2^{(b)}(m', s') | m' \in \left[2^{nR}\right], s' \in \left[2^{n(R_2-R)}\right] \right\}$$

- Transmitter codebook $\mathcal{C}_1^{(b)}$ in block b is given by

$$\mathcal{C}_1^{(b)} = \left\{ \mathbf{x}_1^{(b)}(m, s, t, m', s') | m, m' \in \left[2^{nR}\right], s, s' \in \left[2^{n(R_2-R)}\right], t \in \left[2^{n(R_1-R_2)}\right] \right\}$$

All the codewords are generated independently, and the components of $\mathbf{x}_1^{(b)}(m, s, t, m', s')$ are generated independently as a satellite of $\mathbf{x}_2^{(b)}(m', s')$ using the conditional distribution $P_{X_1|X_2}$, i.e.

$$\mathbf{x}_1^{(b)}(m, s, t, m', s') \sim \prod_{i=1}^{n} P_{X_1|X_2}(x_{1i}^{(b)} | x_{2i}^{(b)}(m', s')).$$

Here, the code can be thought as the union of satellite codebooks for each relay codeword $\mathbf{x}_2^{(b)}(m', s')$ (m', s' will correspond to the previous block). Each satellite codebook has 2^{nR} bins indexed by m, $2^{n(R_2-R)}$ sub-bins indexed by s in each bin, and $2^{n(R_1-R_2)}$ codewords indexed by t in each sub-bin.

Encoding at the transmitter: Let $m^{(1)}, m^{(2)}, \cdots, m^{(B-1)}$ be the sequence of messages to be encoded. The encoder chooses $s^{(1)}, s^{(2)}, \cdots, s^{(B-1)}$ independently and uniformly at random from $[2^{n(R_2-R)}]$, and also $t^{(1)}, t^{(2)}, \cdots, t^{(B-1)}$ independently and uniformly at random from $[2^{n(R_1-R_2)}]$. In the following, we assume that $m^{(0)} = s^{(0)} = 1$, and this is known to all parties beforehand, including the eavesdropper.

In block $b \in [1 : B]$, the source transmits $\mathbf{x}_1^{(b)}(m^{(b)}, s^{(b)}, t^{(b)}, m^{(b-1)}, s^{(b-1)})$. Here $m^{(B)} = s^{(B)} = t^{(B)} = 1$ is assumed.

Decoding at the relay:

In block $b \in [1 : B-1]$, the relay looks for a unique tuple $(\tilde{m}^{(b)}, \tilde{s}^{(b)})$ and some $\tilde{t}^{(b)}$ that satisfies the joint typicality criterion

$$(\mathbf{x}_1^{(b)}(\widetilde{m}^{(b)}, \widetilde{s}^{(b)}, \widetilde{t}^{(b)}, \widetilde{m}^{(b-1)}, \widetilde{s}^{(b-1)}), \mathbf{x}_2^{(b)}(\widetilde{m}^{(b-1)}, \widetilde{s}^{(b-1)}), \mathbf{y}_2^{(b)}) \in T_\epsilon.$$

Here $\widetilde{m}^{(0)} = \widetilde{s}^{(0)} = 1$, and for $b > 1$, $\widetilde{m}^{(b-1)}, \widetilde{s}^{(b-1)}$ are the relay's decoded values in the previous block. *If the tuple $(\tilde{m}^{(b)}, \tilde{s}^{(b)})$ is not unique, an error is declared.*

Encoding at the relay:

In block 1, the relay transmits $\mathbf{x}_2^{(1)}(1, 1)$. In block $b > 1$, the relay transmits $\mathbf{x}_2^{(b)}(\widetilde{m}^{(b-1)}, \widetilde{s}^{(b-1)})$.

Decoding at the receiver: The *receiver* uses sliding window decoding. To decode m_b ($b \in [1 : B-1]$), it finds a unique $(\hat{m}^{(b)}, \hat{s}^{(b)})$ and some $\hat{t}^{(b)}$ such that

$$(\mathbf{x}_1^{(b)}(\widehat{m}^{(b)}, \widehat{s}^{(b)}, \widehat{t}^{(b)}, \widehat{m}^{(b-1)}, \widehat{s}^{(b-1)}), \mathbf{x}_2^{(b)}(\widehat{m}^{(b-1)}, \widehat{s}^{(b-1)}), \mathbf{y}_2^{(b)})$$

is jointly typical, and also

$$(\mathbf{x}_2^{(b+1)}(\widehat{m}^{(b)}, \widehat{s}^{(b)}), \mathbf{y}_2^{(b+1)})$$

is jointly typical. If there is none or more than one such $(\widehat{m}^{(b)}, \widehat{s}^{(b)})$, then it declares error.

Remark 21.

- *Note that the individual messages are decoded with a delay of two blocks, unlike in backward decoding [72] where the decoding has a delay of B blocks.*
- *$\hat{t}^{(b)}$ is decoded nonuniquely.*
- *Decoding $\hat{m}^{(b)}$ correctly and decoding $\hat{s}^{(b)}$ incorrectly will give the correct message for block b, but will (w.h.p) lead to an error in the next block as the relay codeword will have been decoded incorrectly. Hence this is also considered an error event.*

Probability of decoding error: Let us first analyze the decoding error probability at the relay. By random coding arguments, the condition

$$R_1 \leq I(X_1; Y_2|X_2) - \epsilon_1,\ \epsilon_1 > 0 \tag{4.4}$$

is sufficient for the probability of decoding error at the relay to become arbitrarily small for large enough n.

Decoding of the message at the receiver can go wrong if some other message index than the intended one satisfies the typicality test. Notice that it is sufficient to recover the indexes (m, s) in each block. *Hence the receiver is able to employ nonunique decoding.* W.l.o.g assume $(m, s) = (1, 1)$. Of the $2^{nR_1} - 2^{n(R_1 - R_2)}$ codewords which have the index $(m, s) \neq (1, 1)$, the chance of each satisfying the typicality tests is at most $2^{-n(I(X_1;Y_3|X_2)-\epsilon_{2n})}2^{-n(I(X_2;Y_3)-\epsilon_{3n})}$, where ϵ_{2n} and ϵ_{3n} goes to zero with n. Using the union bound, the probability of error in decoding can be made to decay exponentially towards zero by taking

$$R_1 \leq I(X_1, X_2; Y_3) - \epsilon_4,\ \ \epsilon_4 \geq 0. \tag{4.5}$$

Equivocation Calculation: In the following, $W^{(b)}$ denotes the random variable for the message in the b-th block, for $b \in [1 : B-1]$. We consider the multi-block equivocation:

$$H(W^{[B-1]}|\mathbf{Z}^{[B]}, \mathcal{C})$$

Here we have explicitly conditioned on the codebook over all blocks, namely $\mathcal{C} \triangleq \{\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(B)}\}$. Note that $\mathbf{X}_2^{(1)} = \mathbf{x}_2^{(1)}(1,1)$ is known to everybody as the code is known.

$$
\begin{aligned}
&H(W^{[B-1]}|\mathbf{Z}^{[B]}, \mathcal{C}) \\
&\quad \geq I(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}; W^{[B-1]}|\mathbf{Z}^{[B]}, \mathcal{C}) \\
&\quad = H(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}|\mathbf{Z}^{[B]}, \mathcal{C}) - H(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}|W^{[B-1]}, \mathbf{Z}^{[B]}, \mathcal{C}) \\
&\quad = H(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}|\mathcal{C}) - I(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}; \mathbf{Z}^{[B]}|\mathcal{C}) - H(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}|W^{[B-1]}, \mathbf{Z}^{[B]}, \mathcal{C}) \qquad (4.6)
\end{aligned}
$$

The first term is bounded as

$$
\begin{aligned}
H(\mathbf{X}_1^{[B]}, \mathbf{X}_2^{[B]}|\mathcal{C}) &\geq H(\mathbf{X}_1^{[B]}|\mathcal{C}) \\
&\geq I(W^{[B]}, S^{[B]}, T^{[B]}; \mathbf{X}_1^{[B]}|\mathcal{C}) \\
&\overset{(a)}{\geq} I(W^{[B]}, S^{[B]}, T^{[B]}; \mathbf{Y}_2^{[B]}|\mathcal{C}) \\
&\overset{(b)}{\geq} H(W^{[B]}, S^{[B]}, T^{[B]}) - n(B-1)\epsilon_{5n} \quad \text{(by Fano's ineq.)} \\
&= n(B-1)(R_1 - \epsilon_1) \qquad (4.7)
\end{aligned}
$$

where $\epsilon_{5n} \to 0$ as $n \to \infty$. Here (a) follows because $(W^{[B]}, S^{[B]}, T^{[B]}) - \mathbf{X}_1^{[B]} - \mathbf{Y}_2^{[B]}$ forms a Markov chain, and (b) follows because the relay decodes $W^{[B]}, S^{[B]}, T^{[B]}$ with a small probability of error.

Now, we proceed to upper bound the second term of (4.6):

$$
\begin{aligned}
I((\mathbf{X}_1, \mathbf{X}_2)^{[B]}; \mathbf{Z}^{[B]}|\mathcal{C}) &= \sum_{j=1}^{B} I((\mathbf{X}_1, \mathbf{X}_2)^{[B]}; \mathbf{Z}_j|\mathbf{Z}^{[j-1]}, \mathcal{C}) \\
&= \sum_{j=1}^{B} H(\mathbf{Z}_j|\mathbf{Z}^{[j-1]}, \mathcal{C}) \\
&\qquad - H(\mathbf{Z}_j|\mathbf{X}_{1,j}, \mathbf{X}_{2,j}, (\mathbf{X}_1, \mathbf{X}_2)^{[B]\setminus\{j\}}, \mathbf{Z}^{[j-1]}, \mathcal{C}) \\
&\leq \sum_{j=1}^{B} H(\mathbf{Z}_j) - H(\mathbf{Z}_j|\mathbf{X}_{1,j}, \mathbf{X}_{2,j}, (\mathbf{X}_1, \mathbf{X}_2)^{[B]\setminus\{j\}}, \mathbf{Z}^{[j-1]}, \mathcal{C}) \\
&\overset{(a)}{=} \sum_{j=1}^{B} H(\mathbf{Z}_j) - H(\mathbf{Z}_j|\mathbf{X}_{1,j}, \mathbf{X}_{2,j}) \\
&= \sum_{j=1}^{B} I(\mathbf{X}_{1,j}, \mathbf{X}_{2,j}; \mathbf{Z}_j) \\
&\overset{(b)}{\leq} nBI(X_1, X_2; Z) + nB\epsilon \qquad (4.8)
\end{aligned}
$$

Here, (a) follows because the channel is a DMC and consequently

$$\mathbf{Z}_j \leftrightarrow (\mathbf{X}_{1,j}, \mathbf{X}_{2,j}) \leftrightarrow \left((\mathbf{X}_1, \mathbf{X}_2)^{[j-1]}, (\mathbf{X}_1, \mathbf{X}_2)^{[j+1:B]}, \mathbf{Z}^{[j-1]}, \mathcal{C}\right)$$

forms a Markov chain, and (b) follows by standard calculations, see for example [65].

Before examining the third term of (4.6), we define the event $\mathcal{R}_j$ to mean that the *relay* has decoded the message in block j correctly. We further define:

$$\mathcal{R}^j \triangleq \cap_{k=1}^{j} \mathcal{R}_k$$

to denote the event that the relay has decoded messages in all blocks up to and including block j correctly. We use $\mathcal{R}^B \equiv \mathcal{R}$ interchangeably and define the RV:

$$\begin{aligned} \mathbb{I}_{\mathcal{R}} &= 0 \text{ if the relay makes a decoding error in some block} \\ &= 1 \text{ if the relay decodes correctly in all } B \text{ blocks.} \end{aligned}$$

Now we proceed to upper bound the third term of (4.6)

$$\begin{aligned} &H((\mathbf{X}_1, \mathbf{X}_2)^{[B]} | \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) \\ &\quad = H((\mathbf{X}_1, \mathbf{X}_2)^{[B]} | \mathbb{I}_{\mathcal{R}}, \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) + I(\mathbb{I}_{\mathcal{R}}; (\mathbf{X}_1, \mathbf{X}_2)^{[B]} | \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) \\ &\quad \overset{(a)}{\leq} H((\mathbf{X}_1, \mathbf{X}_2)^{[B]} | \mathbb{I}_{\mathcal{R}}, \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) + 1 \end{aligned}$$

(a) follows by using the fact that $\mathbb{I}_{\mathcal{R}}$ is a binary random variable. For a tighter and more insightful upper bound on the last term, see the footnote.[1]

So we consider the term:

$$H((\mathbf{X}_1, \mathbf{X}_2)^{[B-1]} | \mathbb{I}_{\mathcal{R}}, \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}).$$

To analyze this term, we study how much uncertainty Eve will have about $(\mathbf{X}_1, \mathbf{X}_2)^{[B-1]}$ if she is provided with $W^{[B-1]}$ in addition to $Z^{[B]}$, under the assumption that the relay decodes correctly.

[1] Let $p_{\mathcal{R}}$ be the probability of correct decoding in all blocks at the relay. Then $\mathbb{I}_{\mathcal{R}} \sim \text{Ber}(p_{\mathcal{R}})$. As block length $n \to \infty$, and recall that $B \overset{\text{say}}{\sim} O(\sqrt{n})$, the Bernoulli RV $\mathbb{I}_{\mathcal{R}}$ which has a probability of success $p_{\mathcal{R}}$ approaches in distribution a Bernoulli RV – call it S for "sure thing"– with probability of success = 1 i.e. $S \sim \text{Ber}(1)$. Mathematically, we have:

$$\mathbb{I}_{\mathcal{R}} \overset{\text{distribution as } n\uparrow\infty}{\longrightarrow} S \implies H(\mathbb{I}_{\mathcal{R}}) \overset{n\to\infty}{\longrightarrow} H(S) = 0$$

and we also have (since conditioning reduces entropy):

$$H(\mathbb{I}_{\mathcal{R}} | \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) \leq H(\mathbb{I}_{\mathcal{R}}) \to 0 \implies H(\mathbb{I}_{\mathcal{R}} | \mathcal{C}, W^{[B-1]}, \mathbf{Z}^{[B]}) \to 0$$

Since as discussed above when considering the term 1, we know that $p_{\mathcal{R}} \to 1$ exponentially fast in n, we can conclude that the same must hold for $H(\mathbb{I}_{\mathcal{R}}) \to 0$ as well.

Eve performs sliding window decoding in a manner analogous to Bob. Since Eve is given $W^{(1)}$, Eve considers the relay bin corresponding to $W^{(1)}$ in the relay codebook intended for receiver 1 in block 2, namely $\mathcal{C}_{21}^{(2)}$. This bin contains

$$2^{n[R_2-R]} = 2^{n[I(X_2;Z)+\Delta]}$$

codewords, and Eve uses its received sequence $\mathbf{Z}_2$ to form a list of possible jointly typical $\mathbf{x}_2^{(2)}$ codewords. By this, it obtains $\approx nI(X_2;Z)$ bits of information and reduces the size of the ambiguity set of $\mathbf{x}_2$ codewords down to

$$\frac{2^{n[I(X_2;Z)+\Delta]}}{2^{nI(X_2;Z)}} = 2^{n\Delta}$$

Each of these surviving $2^{n\Delta}\mathbf{x}_2$ codewords corresponds to a subbin. In block 1, the decoder looks in the appropriate satellite codebook (which is known to all parties beforehand, being the satellite codebook of $\mathbf{X}_2^{(1)}$). The search space of possible $\mathbf{x}_1^{(1)}$ codewords is of size

$$\begin{aligned}&\text{number of surviving } \mathbf{x}_2 \text{ codewords} \times \text{size of a subbin}\\ &\quad = 2^{n\Delta} \times 2^{n[R_1-R_2]}\end{aligned}$$

By our choices,

$$\begin{aligned}&R_1 = R + I(X_1,X_2;Z) - \epsilon/2, \;\; R_2 \overset{\Delta\geq 0}{=} I(X_2;Z)+\Delta\\ &\implies R_1 - R_2 = I(X_1;Z|X_2) - \Delta - \epsilon/2\end{aligned}$$

Substituting, we can calculate the size of the search space as:

$$2^{n\Delta} \times 2^{n[I(X_1;Z|X_2)-\Delta-\epsilon/2]} = 2^{n[I(X_1;Z|X_2)-\epsilon/2]}$$

But knowing $\mathbf{X}_2^{(1)}$ correctly and having received $\mathbf{Z}_1$ means that Eve can obtain $I(X_1;Z|X_2)$ (independent) bits, and this suffices to decode uniquely $\mathbf{X}_1^{(1)}$ and thus also $\mathbf{X}_2^{(2)}$. *This decoding is correct w.h.p.* The process repeats in the next block.

By standard techniques involving Fano's inequality, this can be shown to be upper bounded as

$$H((\mathbf{X}_1,\mathbf{X}_2)^{[B-1]}|\mathbb{I}_{\mathcal{R}},\mathcal{C},W^{[B-1]},\mathbf{Z}^{[B]}) \leq n(B-1)\epsilon_2 \tag{4.9}$$

for some $\epsilon_2 \to 0$. Thus collecting all the expressions from (4.7), (4.8), and (4.9) into (4.6), we obtain

$$\begin{aligned}H(W^{[B-1]}|\mathbf{Z}^{[B]},\mathcal{C}) &\geq n(B-1)(R_1-\epsilon_1) - nBI(X_1,X_2;Z) - n(B-1)\epsilon_2\\ &= n(B-1)\left(R_1 - I(X_1,X_2;Z) - \epsilon_1 - \epsilon_2 - \frac{1}{B-1}\cdot I(X_1,X_2;Z)\right)\\ &= n(B-1)\left(R - \epsilon/2 - \epsilon_1 - \epsilon_2 - \frac{1}{B-1}\cdot I(X_1,X_2;Z)\right) \quad \text{(Using (4.2))}\end{aligned}$$

Since $B \to \infty$ and $\epsilon_1, \epsilon_2 \to 0$ as $n \to \infty$, for large enougn n, we have

$$H(W^{[B-1]}|\mathbf{Z}^{[B]}, \mathcal{C}) \geq n(B-1)(R-\epsilon).$$

for an appropriate $\epsilon > 0$. This completes the proof of secrecy, and thus also the proof of achievability.

Chapter 5

Relay Broadcast Channel under Mutual Secrecy

5.1 Introduction

In this chapter, we study a different kind of extension of Liu, Maric, Spasojevic, Yates [65] than the one studied in Chapter 3. Here we dispense with the secret keys and employ a relay trusted by all parties, one which acts as an extension of the transmitter. As in the previous chapter, we only consider the "strong" relay scenario, where the T-to-R link is stronger than both T-to-D links. As before, DF is indicated – see [71]. However, as pointed out by Ekrem and Ulukus in [75, Chapter 7], for DF to be effective in a relay-eavesdropper channel, the relay-receiver link must be stronger than the relay-eavesdropper link. *This may lead one to naively conclude that a relay cannot be used to obtain mutual secrecy in DF based scenarios.* We show that this intuition is incorrect. The following points are worthy of note:

- For the same reasons as before, we continue to employ sliding window decoding.
- As before, the relay and the receivers employ nonunique decoding.
- We employ *the relay as a broadcast channel in its own right*, as in [76]. Notice that Kramer et al [77] used the relay as a point-to-point channel in their relay broadcast channel. The latter approach is likely to be suboptimal in wireless scenarios, which are inherently broadcast [76].

5.2 System Model

We assume a two receiver discrete memoryless relay broadcast channel (DM-RBC) with two confidential messages. The finite sets $\mathcal{X}_1, \mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Y}_3$, $\mathcal{Y}_4$ respectively represent the channel's input at node 1 (the Tx), at node 2 (the relay), the channel's output at nodes 2, 3 (Rx 1) and 4 (Rx 2). The channel is described by the conditional probability distribution $P_{Y_2,Y_3,Y_4|X_1,X_2}$, where RVs $X_i \in \mathcal{X}_i$, $i = 1, 2$ and $Y_i \in \mathcal{Y}_i$, $i = 2, 3, 4$. The transmitter intends to send an independent message $W_t \in \{1, 2, \ldots, 2^{nR_t}\} \overset{\Delta}{=} \mathcal{W}_t$ to the respective Rx $t \in \{1, 2\}$ in n channel uses while ensuring information theoretic secrecy, defined below. The channel is memoryless and without feedback i.e. $\forall (\mathbf{x}_1, \mathbf{x}_2) \in \prod_{t=1}^{2} \mathcal{X}_t^n, \mathbf{y}_t \in \mathcal{Y}_t^n$, $t = 2, 3, 4$,

$$P(\mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4|\mathbf{x}_1, \mathbf{x}_2) = \prod_{i=1}^{n} P_{Y_2,Y_3,Y_4|X_1,X_2}(y_{2i}, y_{3i}, y_{4i}|x_{1i}, x_{2i})$$

The channel input at the Tx is obtained by passing codewords $\mathbf{u}_t \equiv \mathbf{u}_t(|\mathbf{v}_t)$, $t = 1, 2$ through a stochastic encoder which generates $\mathbf{x}_1 \sim \prod_{i=1}^{n} P_{X_1|U_1,U_2}(x_{1i}|u_{1i}, u_{2i})$. The channel input at the relay is obtained by passing codewords $\mathbf{v}_t$, $t = 1, 2$ through a stochastic encoder which generates $\mathbf{x}_2 \sim \prod_{i=1}^{n} P_{X_2|V_1,V_2}(x_{2i}|v_{1i}, v_{2i})$. (See the encoding for how $\mathbf{u}_t, \mathbf{v}_t$, $t = 1, 2$ are chosen). The sequences $\mathbf{x}_t$, $t = 1, 2$ are not part of code $\mathcal{C}$. They are generated at the time of transmission by choosing an appropriate quadruple $(\mathbf{u}_t, \mathbf{v}_t : t = 1, 2)$. The decoding

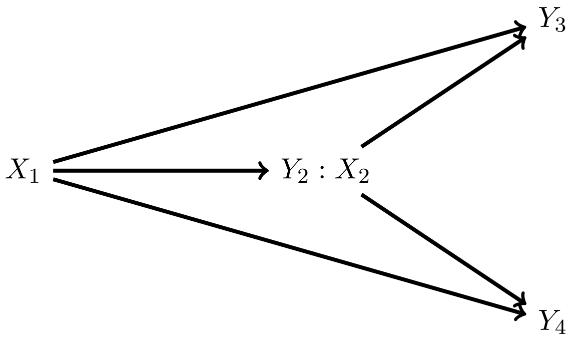

Figure 5.1: Relay broadcast channel

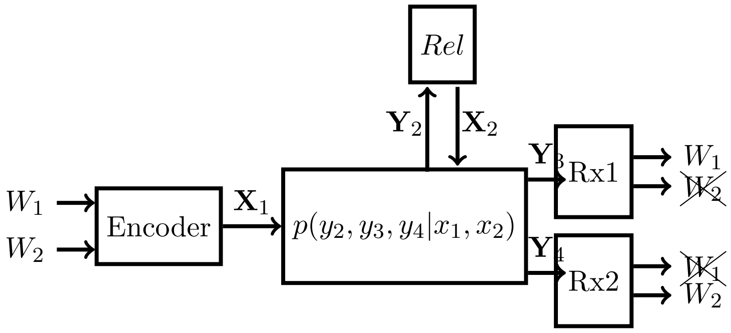

Figure 5.2: Two Receiver Dedicated Relay Broadcast Channel with Trusted Relay (i.e. the relay is an extension of the transmitter, and is trusted by both Rxs) and Two Confidential Messages.

function at the relay is a mapping $\phi_2 : \mathcal{Y}_2^n \to \mathcal{W}_1 \times \mathcal{S}_1 \times \mathcal{W}_2 \times \mathcal{S}_2$. In the preceding, $S_t \in \{1, 2, \ldots, 2^{n[R_{2t}-R_t]}\} \triangleq \mathcal{S}_t$. The decoding function at Rx 1 $\equiv Y_3$ (resp. Rx 2 $\equiv Y_4$) is a map $\phi_3 : \mathcal{Y}_3^n \times \mathcal{Y}_3^n \to \mathcal{W}_1 \times \mathcal{S}_1$ (resp. $\phi_4 : \mathcal{Y}_4^n \times \mathcal{Y}_4^n \to \mathcal{W}_2 \times \mathcal{S}_2$). A $(2^{nR_1}, 2^{nR_2}, n, P_e^{(n)})$ code for the RBC consists of the stochastic encoding functions, three decoding functions ϕ_t, $t = 2, 3, 4$, and the error probability $P_e^{(n)} \triangleq \max\{P_{e,2}^{(n)}, P_{e,3}^{(n)}, P_{e,4}^{(n)}\}$ where

$$P_{e,2}^{(n)} = \sum_{(w_t,s_t)_{t=1,2}} \frac{\Pr[\phi_2(\mathbf{Y}_2) \neq (w_t, s_t)_{t=1,2} | (w_t, s_t)_{t=1,2}]}{2^{nR_1} \times 2^{n[R_{21}-R_1]} \times 2^{nR_2} \times 2^{n[R_{22}-R_2]}}$$

$$\underset{t=1,2}{P_{e,t+2}^{(n)}} = \sum_{(w_t,s_t)} \frac{\Pr[\phi_{t+2}(\mathbf{Y}_{t+2} \times \mathbf{Y}_{t+2}) \neq (w_t, s_t) | w_t, s_t]}{2^{nR_t} \times 2^{n[R_{2t}-R_t]}}$$

A rate pair (R_1, R_2) is said to be achievable for the DM-RBC with confidential messages if, for any $\epsilon_0 > 0$, $\exists(2^{nR_1}, 2^{nR_2}, n, P_e^{(n)})$ code s.t.:

$$\begin{aligned} &P_e^{(n)} \leq \epsilon_0 \text{ reliability requirement} \\ &n(B-1)R_1 - H(W_1^{[B-1]} | \mathbf{Y}_4^{[B]}) \leq n(B-1)\epsilon_0 \\ &n(B-1)R_2 - H(W_2^{[B-1]} | \mathbf{Y}_3^{[B]}) \leq n(B-1)\epsilon_0 \end{aligned}$$

The last two constraints are the weak secrecy constraints [65]. We use the notation $\bar{t} \triangleq \{1, 2\} \setminus \{t\}$, and $t^{[j]} \triangleq \{t_1, t_2, \ldots, t_j\}$.

5.3 Rate Region

Theorem 22. *A (pure secrecy) rate pair (R_1, R_2) is achievable if there exist distributions $P_Q P_{V_1,V_2|Q}, P_{U_1,U_2|V_1,V_2,Q}, P_{X_1|U_1,U_2,Q}, P_{X_2|V_1,V_2,Q}$, so that the following inequalities are satisfied.*

$$\begin{aligned} R_1 &\leq \min\{I(U_1, V_1; Y_3|Q), I(U_1; U_2, V_2, Y_2|V_1, Q)\} - I(U_1, V_1; U_2, V_2, Y_4|Q) \\ R_2 &\leq \min\{I(U_2, V_2; Y_4|Q), I(U_2; U_1, V_1, Y_2|V_2, Q)\} - I(U_2, V_2; U_1, V_1, Y_3|Q) \\ R_1 + R_2 &\leq I(U_1; V_2, Y_2|V_1, Q) + I(U_2; V_1, Y_2|V_2, Q) + I(U_1; U_2|V_1, V_2, Y_2, Q) \\ &\quad - I(U_1, V_1; U_2, V_2, Y_4|Q) - I(U_2, V_2; U_1, V_1, Y_3|Q) \end{aligned}$$

(R_1, R_2) is obtained by Fourier-Motzkin (FM) elimination of the following set of inequalities, followed by convexification using a time-sharing random variable (RV) Q. In the following, if a quantity is doubly subscripted, the first subscript refers to the node index, and the second to the intended Rx. Thus R_{12} refers to the rate of the codebook at the Tx intended for Rx 2.

Theorem 23. *A (pure secrecy) rate pair* (R_1, R_2) *is achievable if* $\exists R_{11}, R_{12}, R_{21}, R_{22}$*, and distributions* $P_{V_1V_2}, P_{U_1U_2|V_1V_2}, P_{X_1|U_1U_2}, P_{X_2|V_1V_2}$*, s.t. the following inequalities are satisfied.*

Decoding Constraints:

$$\begin{aligned}
&R_{11} \le \min\{I(U_1, V_1; Y_3), I(U_1; U_2, V_2, Y_2|V_1)\}\\
&R_{12} \le \min\{I(U_2, V_2; Y_4), I(U_2; U_1, V_1, Y_2|V_2)\}\\
&R_{11} + R_{12} \le I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1; U_2|V_1, V_2, Y_2)
\end{aligned} \tag{5.1}$$

Secrecy Constraints: Packing constraints on the bin sizes at the transmitter and relay due to the secrecy requirements

$$\begin{aligned}
&R_{21} - R_1 > I(V_1; U_2, V_2, Y_4) \overset{\Delta_1 \geq 0}{\equiv} R_{21} - R_1 = I(V_1; U_2, V_2, Y_4) + \Delta_1\\
&R_{22} - R_2 > I(V_2; U_1, V_1, Y_3) \overset{\Delta_2 \geq 0}{\equiv} R_{22} - R_2 = I(V_2; U_1, V_1, Y_3) + \Delta_2\\
&R_{11} - R_1 = I(U_1, V_1; U_2, V_2, Y_4) - \epsilon_{11}\\
&R_{12} - R_2 = I(U_2, V_2; U_1, V_1, Y_3) - \epsilon_{12}
\end{aligned} \tag{5.2}$$

Encoding Constraints: Constraints on the trasmitter subbin sizes

$$\begin{aligned}
&R_{11} - R_{21} > I(U_1; V_2|V_1)\\
&R_{12} - R_{22} > I(U_2; V_1|V_2)\\
&[R_{11} - R_{21}] + [R_{12} - R_{22}] > I(U_1; V_2|V_1) + I(U_2; V_1|V_2) + I(U_1; U_2|V_1, V_2)
\end{aligned} \tag{5.3}$$

Remark 24.

- *From the above, we can obtain:* $R_{1t} \geq R_{2t} \geq R_t, \ \ t = 1, 2$. *This indicates the ordering between the corresponding codebook sizes at the Tx, the relay, and the (pure secrecy) message rate intended for Rxs* 1, 2.
- *In the first decoding two constraints above, the first term inside each* min *arises from the requirement for correct decoding at the respective receivers.*
- *Our achievable scheme uses nonunique decoding at the relay. The second term in each of the first two decoding constraints arises from (nonunique) decoding at the relay, as does the third constraint on the sum-rate* $R_{11} + R_{12}$.

  The following two redundant constraints also arise due to nonunique decoding at the relay:

$$\begin{aligned}
&[R_{11} - R_{21}] + R_{12} \le I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1; U_2|V_1, V_2, Y_2)\\
&R_{11} + [R_{12} - R_{22}] \le I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1; U_2|V_1, V_2, Y_2)
\end{aligned}$$

- *The first two secrecy constraints give a strict lower bound on the relay bin sizes to maintain perfect (weak) secrecy. They are the analogues of the constraint $R_2 - R > I(X_2; Z)$ in the previous chapter. The next two secrecy constraints are (essentially) equality constraints. The quantities ϵ_{1j}, $j = 1, 2$ have the property that $n\epsilon_{1j} \overset{n\uparrow\infty}{\to} 0$. They are the analogues of the constraint $R_1 - R \approx I(X_1, X_2; Z)$ in the previous chapter.*

- *The three encoding constraints arise due to a variant of the mutual covering lemma developed by Zhao and Chung that they have dubbed the "modified mutual covering lemma" [76, Lemma 1]. In [76], there is a further strict lower bounding constraint on the sum-rate of the relay bin sizes. But the analogous constraint on the sum-rate of the relay bin sizes, namely:*

$$[R_{21} - R_1] + [R_{22} - R_2] > I(V_1; V_2)$$

is made redundant by the above individual packing constraints on the relay bin sizes due to secrecy requirements. (See subsection 5.5.1). This recalls how the individual constraint on the bin sizes in the broadcast channel with mutual secrecy [65] makes redundant the constraint on the sum of the bin sizes in Marton coding for a broadcast channel.

5.4 Achievable Scheme

Codebook Generation:

The transmitter employs two codes, $\mathcal{C}_{11}, \mathcal{C}_{12}$, one intended for each receiver. Similarly, the relay has two codes, $\mathcal{C}_{21}, \mathcal{C}_{22}$, one intended for each receiver. Our achievable scheme is over B blocks of n channel uses each. So each code $\mathcal{C}_{ij}$ has B parts: $\mathcal{C}_{ij} = (\mathcal{C}_{ij}^{(1)}, \cdots, \mathcal{C}_{ij}^{(B)})$. We now describe what these codes contain, and how they are generated. The size of the codes used by the relay are given by

$$|\mathcal{C}_{2j}^{(k)}| = 2^{nR_{2j}} \qquad \text{for } j = 1, 2;\ k \in [1 : B].$$

- For blocks $b \in [1 : B]$, the relay codebooks $C_{2j}^{(b)}$ are given by

$$C_{2j}^{(b)} = \left\{ \mathbf{v}_j^{(b)}(m_j', s_j') | m_j' \in \left[2^{nR_j}\right], s_j' \in \left[2^{n(R_{2j} - R_j)}\right] \right\} \qquad \text{for } j = 1, 2.$$

- Transmitter codebooks $\mathcal{C}_{1j}^{(b)}; j = 1, 2$ in block b are given by

$$C_{1j}^{(b)} = \left\{ \mathbf{u}_j^{(b)}(m_j, s_j, t_j, m_j', s_j') | m_j, m_j' \in \left[2^{nR_j}\right], s_j, s_j' \in \left[2^{n(R_{2j} - R_j)}\right], \right.$$
$$\left. t_j \in \left[2^{n(R_{1j} - R_{2j})}\right] \right\}$$

All the codewords are generated independently, and the components of $\mathbf{u}_j^{(b)}(m_j, s_j, t_j, m'_j, s'_j)$ are generated independently as a satellite of $\mathbf{v}_j^{(b)}(m'_j, s'_j)$ using the conditional distribution $P_{U_j|V_j}$, i.e. $\mathbf{u}_j^{(b)}(m_j, s_j, t_j, m'_j, s'_j) \sim \prod_{i=1}^n P_{U_j|V_j}(\mathbf{u}_{ji}^{(b)}|\mathbf{v}_{ji}^{(b)}(m'_j, s'_j))$. Here, the code can be thought as the union of satellite codebooks for each relay codeword $\mathbf{v}_j^{(b)}(m'_j, s'_j)$ (m', s' will correspond to the previous block). Each satellite codebook has 2^{nR_j} bins ($\equiv$ messages) indexed by m_j, $2^{n(R_{2j}-R_j)}$ subbins indexed by s_j in each bin, and $2^{n(R_{1j}-R_{2j})}$ codewords indexed by t_j in each subbin. Each relay codebook has 2^{nR_j} bins indexed by m'_j, and each relay bin has $2^{n[R_{2j}-R_j]}$ codewords, indexed by s'_j. *This is identical to the number of subbins per bin in each satellite codebook, enabling a one-to-one correspondence to be set up.* The following picture illustrates the code construction in each block intended for Rx 1. Rx 2 uses a similar construction. Notice that the intended receiver for the two codebooks in the figure is the same i.e. Rx 1. All the different codebooks to be employed in the B blocks are supplied to all users.

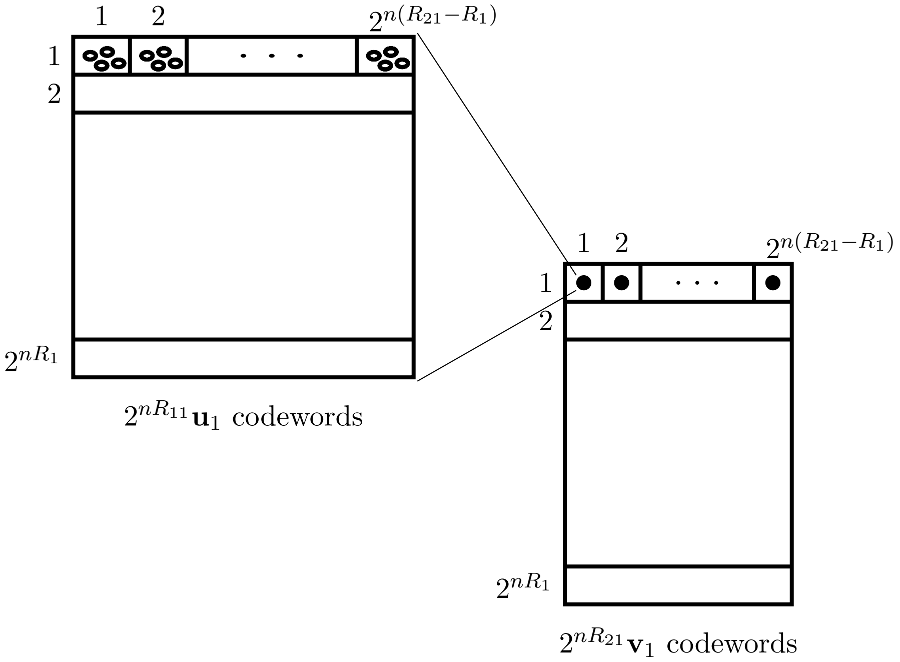
1
2
$2^{n(R_{21}-R_1)}$
1
2
2^{nR_1}
$2^{nR_{11}}\mathbf{u}_1$ codewords
1
2
$2^{n(R_{21}-R_1)}$
1
2
2^{nR_1}
$2^{nR_{21}}\mathbf{v}_1$ codewords

Figure 5.3: Illustrating the Binning Structure

Encoding at the transmitter: To transmit the new message pair $(m_1^{(b)}, m_2^{(b)})$ in block $b \in [1 : B]$, the transmitter first finds a pair of indices $(s_1^{(b)}, s_2^{(b)})$ such that

$$\left(\mathbf{v}_1^{(b+1)}(m_1^{(b)}, s_1^{(b)}), \mathbf{v}_2^{(b+1)}(m_2^{(b)}, s_2^{(b)})\right) \in T_\epsilon. \tag{5.4}$$

Remark 25. • *This step's success is guaranteed w.h.p. by the mutual covering lemma [27, Lemma 8.1] if the sum-rate constraint* $[R_{21} - R_1] + [R_{22} - R_2] > I(V_1; V_2)$. *But* *this strict inequality is a trivial consequence of our choices above w.r.t. the individual relay bin rates.*

- *Note that the transmitter looks inside the appropriate bins in the relay codebooks in the* ***next*** *block. Thus the codebooks must be known at least one block in advance.*

If there is no such pair, then the encoder chooses $(1, 1)$, and if there are more than one, then it chooses the least such pair in the lexicographical order. It then picks a pair $(t_1^{(b)}, t_2^{(b)})$ such that

$$\Big(\mathbf{u}_1^{(b)}(m_1^{(b)}, s_1^{(b)}, t_1^{(b)}, m_1^{(b-1)}, s_1^{(b-1)}), \mathbf{v}_1^{(b)}(m_1^{(b-1)}, s_1^{(b-1)}),\\ \mathbf{u}_2^{(b)}(m_2^{(b)}, s_2^{(b)}, t_2^{(b)}, m_2^{(b-1)}, s_2^{(b-1)}), \mathbf{v}_2^{(b)}(m_2^{(b-1)}, s_2^{(b-1)})\Big) \tag{5.5}$$

is typical where $s_1^{(b-1)}|_{b=1} = m_1^{(b-1)}|_{b=1} = m_2^{(b-1)}|_{b=1} = 1$. Also, $s_2^{(0)} = 1'$ is the least index such that $(v_1^{(1)}(1, 1), v_2^{(1)}(1, 1'))$ is jointly typical. By appropriate relabelling of the indexes we can take $m_2^{(b-1)} = 1$ itself, a convention followed below. The transmitted codeword $\mathbf{x}_1$ is generated from $u_1^{(b)}(m_1^{(b)}, s_1^{(b)}, t_1^{(b)}, m_1^{(b-1)}, s_1^{(b-1)})$ and $u_2^{(b)}(m_2^{(b)}, s_2^{(b)}, t_2^{(b)}, m_2^{(b-1)}, s_2^{(b-1)})$ component-wise using the distribution $\prod_{i=1}^n p(x_{1i}|u_{1i}^{(b)}, u_{2i}^{(b)})$.

Decoding at the relay: Assume that the relay already knows $m_1^{(0)} = m_2^{(0)} = s_1^{(0)} = s_2^{(0)} = 1$. So it assigns the same decoded values: $\widetilde{m}_1^{(0)} = \widetilde{m}_2^{(0)} = \widetilde{s}_1^{(0)} = \widetilde{s}_2^{(0)} = 1$. In block b, the relay chooses the index tuple $(\widetilde{m}_1^{(b)}, \widetilde{s}_1^{(b)}, \widetilde{m}_2^{(b)}, \widetilde{s}_2^{(b)})$ if it is the unique tuple for which

$$\Big(\mathbf{u}_1^{(b)}(\widetilde{m}_1^{(b)}, \widetilde{s}_1^{(b)}, \widetilde{t}_1^{(b)}, \widetilde{m}_1^{(b-1)}, \widetilde{s}_1^{(b-1)}), \mathbf{v}_1^{(b)}(\widetilde{m}_1^{(b-1)}, \widetilde{s}_1^{(b-1)}),\\ \mathbf{v}_2^{(b)}(\widetilde{m}_2^{(b-1)}, \widetilde{s}_2^{(b-1)}), \mathbf{u}_2^{(b)}(\widetilde{m}_2^{(b)}, \widetilde{s}_2^{(b)}, \widetilde{t}_2^{(b)}, \widetilde{m}_2^{(b-1)}, \widetilde{s}_2^{(b-1)}), \mathbf{y}_2^{(b)}\Big) \in T_\epsilon$$

is typical for some $\tilde{t}_1^{(b)}, \tilde{t}_2^{(b)}$. This is the value of the tuple $(m_1^{(b)}, s_1^{(b)}, t_1^{(b)}, m_2^{(b)}, s_2^{(b)}, t_2^{(b)})$ decoded by the relay. *Analogous to the relay-eavesdropper case, where the relay performed nonunique decoding and only the uniqueness of the tuple $(\tilde{m}^{(b)}, \tilde{s}^{(b)})$ was insisted upon, here we only require that the tuple $(\tilde{m}_1^{(b)}, \tilde{s}_1^{(b)}, \tilde{m}_2^{(b)}, \tilde{s}_2^{(b)})$ be unique. The tuple $(\tilde{t}_1^{(b)}, \tilde{t}_2^{(b)})$ need not be.* If no such tuple exists or there are more than one, then the relay chooses $(\widetilde{m}_1^{(b)}, \widetilde{s}_1^{(b)}, \widetilde{m}_2^{(b)}, \widetilde{s}_2^{(b)}) = (1, 1, 1, 1)$.

Encoding at the relay: In block b, the relay transmits $\mathbf{x}_2\left(\mathbf{v}_1^{(b)}(\widetilde{m}_1^{(b-1)}, \widetilde{s}_1^{(b-1)}), \mathbf{v}_2^{(b)}(\widetilde{m}_2^{(b-1)}, \widetilde{s}_2^{(b-1)})\right)$, where $\mathbf{x}_2$ is a stochastic mapping, according to $\prod_{i=1}^n p(x_{2i}|v_{1i}^{(b)}(\widetilde{m}_1^{(b-1)}, \widetilde{s}_1^{(b-1)}), v_{2i}^{(b)}(\widetilde{m}_2^{(b-1)}, \widetilde{s}_2^{(b-1)}))$. Note that $\widetilde{m}_1^{(0)} = \widetilde{m}_2^{(0)} = \widetilde{s}_1^{(0)} = \widetilde{s}_2^{(0)} = 1$ by assumption.

Decoding at the receivers: The receiver knows $m_1^{(0)} = m_2^{(0)} = s_1^{(0)} = s_2^{(0)} = 1$. So it assigns the same decoded values: $\widehat{m}_1^{(0)} = \widehat{m}_2^{(0)} = \widehat{s}_1^{(0)} = \widehat{s}_2^{(0)} = 1$. *We assume that the* *receiver* 1 *has correctly decoded* $(\hat{m}_1^{(b-1)}, \hat{s}_1^{(b-1)})$. To decode $m_1^{(b)}$, the receiver 1 at Y_3 performs sliding window decoding and looks for a unique tuple $(\hat{m}_1^{(b)}, \hat{s}_1^{(b)})$ such that

$$\left(\mathbf{v}_1^{(b+1)}(\widehat{m}_1^{(b)},\widehat{s}_1^{(b)}),\mathbf{y}_3^{(b+1)}\right)\in T_\epsilon$$

$$\text{and } \left(\mathbf{u}_1^{(b)}(\widehat{m}_1^{(b)},\widehat{s}_1^{(b)},\widehat{t}_1^{(b)},\widehat{m}_1^{(b-1)},\widehat{s}_1^{(b-1)}),\mathbf{v}_1^{(b)}(\widehat{m}_1^{(b-1)},\widehat{s}_1^{(b-1)}),\mathbf{y}_3^{(b)}\right)\in T_\epsilon$$

for some $\hat{t}_1^{(b)}$. If a unique such $(\hat{m}_1^{(b)}, \hat{s}_1^{(b)})$ is found, then $\hat{m}_1^{(b)}$ is declared as the decoded message, otherwise the decoder declares an error. Receiver 2 $\equiv Y_4$ decodes $m_2^{(b)}$ in a similar manner, with appropriate changes to reflect its received sequences.

Remark 26.

- *$\hat{t}_1^{(b)}$ (resp. $\hat{t}_2^{(b)}$) is decoded nonuniquely by Rx* 1 *(resp. Rx* 2*).*
- *Decoding $\hat{m}_1^{(b)}$ correctly and decoding $\hat{s}_1^{(b)}$ incorrectly will give the correct message for block b, but will (w.h.p) lead to an error in block $b+1$ as the relay codeword will have been decoded incorrectly. Hence this is also considered an error event in the error analysis. Likewise for receiver* 2.

5.5 Probability of Error Calculations

5.5.1 Probability of encoding error:

Let us first consider the probability of success in the encoding step at the transmitter given in (5.5), this can be made arbitrarily close to one for large enough n by choosing the transmitter subbin sizes (in terms of rates):

$$R_{11} - R_{21} > I(U_1; V_2|V_1) \tag{5.6}$$

$$R_{12} - R_{22} > I(U_2; V_1|V_2) \tag{5.7}$$

$$[R_{11} - R_{21}] + [R_{12} - R_{22}] > I(U_1; V_2|V_1) + I(U_2; V_1|V_2) + I(U_1; U_2|V_1, V_2) \tag{5.8}$$

The above requirements follow from the "modified mutual covering lemma" developed by Zhao and Chung [76, Lemma 1] which we reproduce below for ease of reference. (See also Remark 24, bullet point 4).

Lemma 27. *[76, Modified Mutual Covering Lemma] Let $(V_1, V_2, U_1, U_2) \sim p(v_1, v_2, u_1, u_2)$ and $\epsilon_\prime < \epsilon$. Let $(V_1^n, V_2^n) \sim p(v_1^n, v_2^n)$ be a pair of random sequences with $P\{(V_1^n, V_2^n) \in \mathcal{T}_{\epsilon'}^{(n)}\} \overset{n\uparrow\infty}{\to} 1$. Let $U_1^n(m_1), m_1 \in [1 : 2^{nr_1}]$, be pairwise conditionally independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_1|V_1}(u_{1i}|v_{1i})$. Similarly, let $U_2^n(m_2), m_2 \in [1 : 2^{nr_2}]$, be pairwise conditionally independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_2|V_2}(u_{2i}|v_{2i})$. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nr_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nr_2}]\}$ are conditionally independent given (V_1^n, V_2^n). Then, there exists $\delta(\epsilon) > 0 \overset{\epsilon\to 0}{\to} 0$ such that*

$$\lim_{n\to\infty} P\{(V_1^n, V_2^n, U_1^n(m_1), U_2^n(m_2)) \notin \mathcal{T}_\epsilon^{(n)}\ \forall (m_1 \in [1 : 2^{nr_1}], m_2 \in [1 : 2^{nr_2}])\} = 0$$

if

$$\begin{aligned} r_1 &> I(U_1;V_2|V_1) + 3\delta(\epsilon) \\ r_2 &> I(U_2;V_1|V_2) + 3\delta(\epsilon) \\ r_1 + r_2 &> H(U_1|V_1) + H(U_2|V_2) - H(U_1,U_2|V_1,V_2) + \delta(\epsilon) \end{aligned}$$

Furthermore, success in the encoding step (5.4) can be achieved by the condition on the sum of the relay binning rates:

$$[R_{21} - R_1] + [R_{22} - R_2] > I(V_1;V_2) + \epsilon. \tag{5.9}$$

which follows from the standard mutual covering lemma [27, Lemma 8.1]. But note that the individual constraints on the relay binning rates that arise due to secrecy make these redundant.

5.5.2 Probability of decoding error:

Under random coding arguments, assume w.l.o.g that the transmitted indices are $(1,1,1,1)$ for all $b \in [1:B]$. Let us first consider the decoding at the relay. We list three kinds of error events.

- $E_r(e,1)$: The tuple $(m_1,s_1,1,1)$ satisfies the typicality test for $(m_1,s_1) \neq (1,1)$. The number of possibilities is $(2^{nR_{11}} - 2^{n[R_{11}-R_{12}]}) \times 1$, which, in rate, is asymptotically $\overset{n\uparrow\infty}{\to} R_{11}$. *The second factor* 1 *hides a crucial subtlety.* See remark 28 below.

$$\begin{aligned} P(E_r(e,1)) &= \sum_{(\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2,\mathbf{y}_2)\in T_\epsilon} p(\mathbf{u}_1,\mathbf{v}_1)p(\mathbf{y}_2,\mathbf{u}_2,\mathbf{v}_2|\mathbf{v}_1) \\ &\approx 2^{nH(U_1,V_1,U_2,V_2,Y_2)}2^{-nH(U_1,V_1)}2^{-nH(U_2,V_2,Y_2|V_1)} \\ &= 2^{-nI(U_1;V_2,U_2,Y_2|V_1)}. \end{aligned}$$

- $E_r(1,e)$: The tuple $(1,1,m_2,s_2)$ satisfies the typicality test for $(m_2,s_2) \neq (1,1)$. The error probability $P(E_r(1,e))$ can be obtained as above with a change of variables.

- $E_r(e,e)$: The tuple (m_1,s_1,m_2,s_2) satisfies the typicality test for $(m_1,s_1) \neq (1,1)$ and $(m_2,s_2) \neq (1,1)$. The number of possibilities is $(2^{nR_{11}} - 2^{n[R_{11}-R_{12}]})\times$ $(2^{nR_{12}} - 2^{n[R_{12}-R_{22}]})$, which, in rate, is asymptotically $\overset{n\uparrow\infty}{\to} R_{11} + R_{12}$.

$$\begin{aligned} P(E_r(e,e)) &= \sum_{(\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2,\mathbf{y}_2)\in T_\epsilon} p(\mathbf{v}_1,\mathbf{v}_2)p(\mathbf{u}_1|\mathbf{v}_1)p(\mathbf{u}_2|\mathbf{v}_2)p(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2) \\ &\approx 2^{nH(U_1,U_2|Y_2,V_1,V_2)}2^{-nH(U_1|V_1)}2^{-nH(U_2|V_2)} \\ &= 2^{-n(I(U_1;V_2,Y_2|V_1)+I(U_2;U_1,V_1,Y_2|V_2))}. \end{aligned}$$

Using the union bound, we deduce that the error probability of decoding at the relay can be made arbitrarily small by choosing,

$$\begin{aligned}R_{11} &\leq I(U_1; V_2, U_2, Y_2|V_1) - \epsilon\\ R_{12} &\leq I(U_2; V_1, U_1, Y_2|V_2) - \epsilon\\ R_{11} + R_{12} &\leq I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1; U_2|V_2, V_1, Y_2) - \epsilon.\end{aligned}$$

Notice that the decoding employed here did not really care about the indices t_1 and t_2, thus leading to non-unique decoding at the relay.

Remark 28. *In computing the number of possibilities above for error event $E_r(e, 1)$, we have written the second factor as 1. It means that the relay has decoded correctly not just the indices (m_2, s_2) correctly – by assumption – as $(1, 1)$, but also the index t_2 correctly. What happens if (m_2, s_2) is decoded correctly as $(1, 1)$, but the index t_2 is decoded incorrectly within the correct subbin? Then the second factor above becomes $2^{n[R_{12}-R_{22}]} - 1$. In this case, the RHS will be identical to the error event $E_r(e, e)$, but the LHS will have an extra term $[R_{11} - R_{21}]$. This gives rise to:*

$$R_{11} + [R_{12} - R_{22}] \leq I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1 : U_2|V_2, V_1, Y_2) - \epsilon \quad (5.10)$$

This is redundant w.r.t the sum-rate constraint on $R_{11} + R_{12}$. A symmetric (and also redundant) constraint is obtained by switching the roles in the computation of $E_r(1, e)$:

$$[R_{11} - R_{21}] + R_{12} \leq I(U_1; V_2, Y_2|V_1) + I(U_2; V_1, Y_2|V_2) + I(U_1; U_2|V_2, V_1, Y_2) - \epsilon \quad (5.11)$$

Let us now consider decoding at the respective receivers. We will show that we can guarantee successful decoding at receivers by choosing

$$R_{11} \leq I(U_1, V_1; Y_3) - \epsilon \quad (5.12)$$

$$R_{12} \leq I(U_2, V_2; Y_4) - \epsilon \quad (5.13)$$

Using the received sequence in block $b + 1$, Rx 1 performs list decoding to reduce the ambiguity set of possible $\mathbf{v}_1$ sequences from $2^{nR_{21}}$ down to $2^{n[R_{21}-I(V_1;Y_3)]}$. Each of these $\mathbf{v}_1$ codewords corresponds to a subbin in the previous block of size $2^{n[R_{11}-R_{21}]}$. Thus the search space of possible $\mathbf{u}_1$ codewords is now of size $2^{n[R_{21}-I(V_1;Y_3)+R_{11}-R_{21}]} = 2^{n[R_{11}-I(V_1;Y_3)]}$. Assuming that the relay codeword in block b has already been correctly decoded (in the previous decoding step), the information obtained from the other JT condition is $2^{nI(U_1;Y_3|V_1)}$. If $R_{11} - I(V_1; Y_3) < I(U_1; Y_3|V_1)$, equivalently, if $R_{11} < I(U_1, V_1; Y_3)$, then we can decode the $\mathbf{u}_1^{(b)}$ correctly and uniquely w.h.p., and thus also the $\mathbf{v}_1^{(b+1)}$ codeword, as it is wholly determined by $\mathbf{u}_1^{(b)}$. With appropriate changes, a similar calculation is performed by Rx 2.

5.5.3 Constraints on Δ_1 and Δ_2

Since:

$$
\begin{aligned}
R_{11} - R_1 &= [R_{11} - R_{21}] + [R_{21} - R_1] \\
\implies I(U_1, V_1; U_2, V_2, Y_4) - \epsilon_{11} &= [R_{11} - R_{21}] + I(V_1; U_2, V_2, Y_4) + \Delta_1 \\
\text{giving: } R_{11} - R_{21} &= I(U_1; U_2, V_2, Y_4|V_1) - \Delta_1 - \epsilon_{11}
\end{aligned}
$$

$$
\text{Similarly: } R_{12} - R_{22} = I(U_2; U_1, V_1, Y_3|V_2) - \Delta_2 - \epsilon_{12}
$$

In what follows, we drop the quantities ϵ_{1j}, $j = 1, 2$. Now recalling the constraints on the subbin sizes, we have:

$$
\begin{aligned}
I(U_1; U_2, V_2, Y_4|V_1) - \Delta_1 &> I(U_1; V_2|V_1) \\
I(U_2; U_1, V_1, Y_3|V_2) - \Delta_2 &> I(U_2; V_1|V_2)
\end{aligned}
$$

and the sum-rate constraint:

$$
\begin{aligned}
&I(U_1; U_2, V_2, Y_4) - \Delta_1 + I(U_2; U_1, V_1, Y_3) - \Delta_2 \\
&\qquad > I(U_1; V_2|V_1) + I(U_1; V_2|V_1) + I(U_1; U_2|V_1, V_2)
\end{aligned}
$$

After appropriate algebraic manipulations, we obtain:

$$
\begin{aligned}
\Delta_1 &< I(U_1; U_2, Y_4|V_1, V_2) \\
\Delta_2 &< I(U_2; U_1, Y_3|V_1, V_2) \\
\Delta_1 + \Delta_2 &< I(U_1; U_2, Y_4|V_1, V_2) + I(U_2; U_1, Y_3|V_1, V_2) - I(U_1; U_2|V_1, V_2)
\end{aligned}
$$

5.5.4 Equivocation Calculation:

We denote the random variable for the messages in block b for the two receivers by $W_1^{(b)}$ and $W_2^{(b)}$. The codewords chosen by the transmitter and the relay in block b are denoted by $\mathbf{U}_1^{(b)}, \mathbf{U}_2^{(b)}, \mathbf{V}_1^{(b)}, \mathbf{V}_2^{(b)}$ respectively. The random vectors transmitted by the transmitter and the relay in block b are denoted by $\mathbf{X}_1^{(b)}$ and $\mathbf{X}_2^{(b)}$ respectively. We consider the equivocation of the message $W_1^{[B-1]}$ intended for the first receiver given the observation of the second receiver:

$$
H(W^{[B-1]}|\mathbf{Y}_4^{[B]}, \mathcal{C}).
$$

Since the relay always chooses $m_1^{(0)} = m_2^{(0)} = s_1^{(0)} = s_2^{(0)} = 1$, it is known by everybody that

$$
\mathbf{V}_1^{(0)} = \mathbf{v}_1^{(b)}(1,1) \quad \mathbf{V}_2^{(0)} = \mathbf{v}_2^{(b)}(1,1).
$$

We now have,

$$
\begin{aligned}
& H(W_1^{[B-1]}|\mathbf{Y}_4^{[B]},\mathcal{C}) \\
& \geq H(W_1^{[B-1]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathbf{Y}_4^{[B]},\mathcal{C}) \\
& \geq I(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]};W_1^{[B-1]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathbf{Y}_4^{[B]},\mathcal{C}) \\
& = H(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathbf{Y}_4^{[B]},\mathcal{C}) - H(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},W_1^{[B-1]},\mathbf{Y}_4^{[B]},\mathcal{C}) \\
& = H(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) - I(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]};\mathbf{Y}_4^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) \\
& \qquad\qquad - H(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},W_1^{[B-1]},\mathbf{Y}_4^{[B]},\mathcal{C})
\end{aligned} \tag{5.14}
$$

We now bound the first term in (5.14):

$$
\begin{aligned}
H(\mathbf{U}_1^{[B]},\mathbf{V}_1^{[B]}|(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) &= \sum_{j=1}^{B} H((\mathbf{U}_1,\mathbf{V}_1)^{(j)}|(\mathbf{U}_1,\mathbf{V}_1)^{[j-1]},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) \\
&\geq \sum_{j=1}^{B} H(\mathbf{U}_1^{(j)}|\mathbf{V}_1^{(j)},(\mathbf{U}_1,\mathbf{V}_1)^{[j-1]},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) \\
&\overset{(a)}{=} \sum_{j=1}^{B-1} H(\mathbf{U}_1^{(j)}|\mathbf{V}_1^{(j)},(\mathbf{U}_1,\mathbf{V}_1)^{[j-1]},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}), \qquad (5.15) \\
&\overset{(b)}{=} \sum_{j=1}^{B-1} H(\mathbf{U}_1^{(j)},\mathbf{V}_1^{(j+1)}|\mathbf{V}_1^{(j)},(\mathbf{U}_1,\mathbf{V}_1)^{[j-1]},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}),
\end{aligned} \tag{5.16}
$$

where (a) follows because $\mathbf{U}_1^{(B)} = \mathbf{u}_1^{(B)}(1,1,1,m^{(B-1)},s^{(B-1)})$. Step (b) follows because $\mathbf{V}_1^{(j+1)}$ is determined by $\mathbf{U}_1^{(j)}$.

Now consider $j = 1$ term i.e. $H(\mathbf{U}_1^{(1)},\mathbf{V}_1^{(2)}|\mathbf{V}_1^{(1)},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C})$. The number of $\mathbf{v}_1^{(2)}$ sequences typical with a given $(\mathbf{v}_2,\mathbf{u}_2)^{(2)}$ pair is $\approx 2^{n(R_{21}-I(V_1;U_2,V_2))}$. For each possible $\mathbf{v}_1^{(2)}$, the encoder can choose from among $2^{n(R_{11}-R_{21})}$ sequences $\mathbf{u}_1^{(1)}$ (based on the index t). The number of possible choices for $\mathbf{u}_1^{(1)}$ is thus:

$$
2^{n[R_{21}-I(V_1;U_2,V_2)]} \times 2^{n[R_{11}-R_{21}]} = 2^{n[R_{11}-I(V_1;U_2,V_2)]}
$$

But now we note that since $\mathbf{v}_1^{(1)}$ and $(\mathbf{u}_2,\mathbf{v}_2)^{(1)}$ are in the conditioning, we can further reduce the number of possibilities for $\mathbf{u}_1^{(1)}$ by $2^{nI(U_1;U_2,V_2|V_1)}$. Thus we finally have:

$$
\frac{2^{n[R_{11}-I(V_1;U_2,V_2)]}}{2^{nI(U_1;U_2,V_2|V_1)}} = 2^{n[R_{11}-I(U_1,V_1;U_2,V_2)]}
$$

possibilities. In terms of equivocation, this contributes a term $n(R_{11} - I(U_1,V_1;U_2,V_2) - \epsilon)$. Essentially the same argument applies for $j = 2,3,\ldots,B-1$, and so we have

$$
\begin{aligned}
& \sum_{j=1}^{B-1} H(\mathbf{U}_1^{(j)},\mathbf{V}_1^{(j+1)}|\mathbf{V}_1^{(j)},(\mathbf{U}_1,\mathbf{V}_1)^{[j-1]},(\mathbf{U}_2,\mathbf{V}_2)^{[B]},\mathcal{C}) \\
& \qquad = n(B-1)\left(R_{11} - I(U_1,V_1;U_2,V_2) - \epsilon\right)
\end{aligned} \tag{5.17}
$$

Now, we proceed to upper bound the second term of (5.14):

$$
\begin{aligned}
&I(\mathbf{U}_1^{[B]}, \mathbf{V}_1^{[B]}; \mathbf{Y}_4^{[B]} | (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}) \\
&= \sum_{j=1}^{B} I((\mathbf{U}_1, \mathbf{V}_1)^{[B]}; \mathbf{Y}_4^{(j)} | \mathbf{Y}_4^{[j-1]}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}) \\
&= \sum_{j=1}^{B} \Big(H(\mathbf{Y}_4^{(j)} | \mathbf{Y}_4^{[j-1]}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}) - H(\mathbf{Y}_4^{(j)} | (\mathbf{U}_1, \mathbf{V}_1)^{[B]}, \mathbf{Y}_4^{[j-1]}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}) \Big) \\
&\overset{(a)}{\leq} \sum_{j=1}^{B} \Big(H(\mathbf{Y}_4^{(j)} | (\mathbf{U}_2, \mathbf{V}_2)^{(j)}, \mathcal{C}) - H(\mathbf{Y}_4^{(j)} | (\mathbf{U}_1, \mathbf{V}_1, \mathbf{U}_2, \mathbf{V}_2)^{(j)}, \mathcal{C}) \Big) \\
&= \sum_{j=1}^{B} I((\mathbf{U}_1, \mathbf{V}_1)^{(j)}; \mathbf{Y}_4^{(j)} | (\mathbf{U}_2, \mathbf{V}_2)^{(j)}) \\
&\overset{(b)}{\leq} n(B-1)[I(Y_4; U_1, V_1 | U_2, V_2) + \epsilon]
\end{aligned}
$$

for large enough n for any $\epsilon > 0$. Here, (a) follows because

$$
\begin{aligned}
\mathbf{Y}_4^{(j)} &\leftrightarrow (\mathbf{U}_1^{(j)}, \mathbf{V}_1^{(j)}, \mathbf{U}_2^{(j)}, \mathbf{V}_2^{(j)}) \\
&\qquad \leftrightarrow \big((\mathbf{U}_1, \mathbf{V}_1, \mathbf{U}_2, \mathbf{V}_2)^{[j-1]}, (\mathbf{U}_1, \mathbf{V}_1, \mathbf{U}_2, \mathbf{V}_2)^{[j+1:B]}, \mathbf{Y}_4^{[j-1]}, \mathcal{C} \big)
\end{aligned}
$$

forms a Markov chain. Also, in (b) we have a factor $B-1$ and not B because $\mathbf{V}_1^{(1)}$ and $\mathbf{U}_1^{(B)}$ do not contribute. See, for example, [65].

As in the previous chapter, we define the RV:

$$
\begin{aligned}
\mathbb{I}_{\mathcal{R}} &= 0 \text{ if the relay makes a decoding error in some block} \\
&= 1 \text{ if the relay decodes correctly in all } B \text{ blocks.}
\end{aligned}
$$

Now we proceed to upper bound the third term of (5.14)

$$
\begin{aligned}
&H((\mathbf{U}_1, \mathbf{V}_1)^{[B]} | (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) \\
&= H((\mathbf{U}_1, \mathbf{V}_1)^{[B]} | \mathbb{I}_{\mathcal{R}}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) \\
&\quad + I(\mathbb{I}_{\mathcal{R}}; (\mathbf{U}_1, \mathbf{V}_1)^{[B]} | (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) \\
&\overset{(a)}{\leq} H((\mathbf{U}_1, \mathbf{V}_1)^{[B]} | \mathbb{I}_{\mathcal{R}}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) + 1 \\
&\leq H((S_1, T_1)^{[B]} | \mathbb{I}_{\mathcal{R}}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) + 1 \\
&= \sum_{b=1}^{B-1} H(S_1^{(b)}, T_1^{(b)} | S_1^{[b-1]}, T_1^{[b-1]}, \mathbb{I}_{\mathcal{R}}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]}) + 1
\end{aligned}
$$

where (a) follows because the conditional mutual information term is upper bounded by $H(\mathbb{I}_{\mathcal{R}})$ which in turn is ≤ 1, as $\mathbb{I}_{\mathcal{R}}$ is a binary RV.

We will examine how much ambiguity about $(S_1, T_1)^{(b)}$ the receiver with Y_4 will have if it is additionally provided $W_1^{[B-1]}, S_1^{[b-1]}, T_1^{[b-1]}, (\mathbf{U}_2, \mathbf{V}_2)^{[B]}$ and also assuming that the relay decodes correctly. Consider $b = 1$.

1. From $(\mathbf{U}_2, \mathbf{V}_2, \mathbf{Y}_4)^{(2)}$ receiver 2 gets almost $nI(V_1; Y_4, U_2, V_2)$ bits of information about $\mathbf{V}_1^{(2)} = \mathbf{v}_1^{(2)}(W_1^{(1)}, S_1^{(1)})$. Since $W_1^{(1)}$ is already provided to the receiver, its remaining uncertainty about $S_1^{(1)}$ is $\leq n(R_{21} - R_1) - nI(V_1; Y_4, U_2, V_2)$. Here note that $R_{21} - R_1 > I(V_1; Y_4, U_2, V_2)$ by (5.2)
2. The remaining uncertainty in $(S_1^{(1)}, T_1^{(1)})$ is thus $\leq n(R_{21} - R_1) - nI(V_1; Y_4, U_2, V_2) + n(R_{11} - R_{21}) = n(R_{11} - R_1) - nI(V_1; Y_4, U_2, V_2)$. So, there is a list of about $2^{n((R_{11}-R_1)-I(V_1;Y_4,U_2,V_2))}$ pairs $(S_1^{(1)}, T_1^{(1)})$ pairs. This includes the uncertainty in $T_1^{(1)}$.
3. All the possible codewords in the list are from the satellite codebook of $\mathbf{V}_1^{(1)}$. So, a randomly chosen codeword from the list has a probability of about $2^{-nI(U_1;U_2,V_2,Y_4|V_1)}$ of being jointly typical with $(\mathbf{U}_2^{(1)}, \mathbf{V}_2^{(1)}, \mathbf{Y}_4^{(1)})$. So, the number of codewords from the list that are jointly typical with $\mathbf{U}_2^{(1)}, \mathbf{V}_2^{(1)}, \mathbf{Y}_4^{(1)}$ is about
$$2^{n((R_{11}-R_1)-I(V_1;U_2,V_2,Y_4)-I(U_1;U_2,V_2,Y_4|V_1))}.$$
So the uncertainty remaining in $(S_1^{(1)}, T_1^{(1)})$ is
$$\begin{aligned}&\leq n\left((R_{11} - R_1) - I(V_1; U_2, V_2, Y_4) - I(U_1; U_2, V_2, Y_4|V_1) - \epsilon_{15}\right)\\ &\leq n\left((R_{11} - R_1) - I(U_1, V_1; U_2, V_2, Y_4) - \epsilon_{11}\right)\end{aligned}$$
But we have packed each transmitter bin with $R_{11} - R_1 \approx I(U_1, V_1; U_2, V_2, Y_4)$ and by our choice $n\epsilon_{11} \overset{n\uparrow\infty}{\rightarrow} 0$, and so each term can be made as small as desired.
4. In the last block $b = B$, $S_1^{(B)} = T_1^{(B)} = 1$. So there is no uncertainty and the last block does not contribute.

Thus, we have a sum of $B - 1$ terms:

$$\begin{aligned}&H((\mathbf{U}_1, \mathbf{V}_1)^{[B]}|(\mathbf{U}_2, \mathbf{V}_2)^{[B]}, \mathcal{C}, W_1^{[B-1]}, \mathbf{Y}_4^{[B]})\\ &\qquad \leq n(B-1)(R_{11} - R_1 - I(U_1, V_1; Y_4, U_2, V_2) - \epsilon_{11}) + 1\end{aligned}$$

Finally, the upper and lower bounds for terms in (5.14) give:

$$\begin{aligned}H(W_1^{[B-1]}|\mathbf{Y}_4^{[B]}, \mathcal{C}) &\geq n(B-1)(R_{11} - I(U_1, V_1; V_2, U_2) - \epsilon_1)\\ &\quad - n(B-1)(I(Y_4; U_1, V_1|U_2, V_2) + \epsilon_2) - n(B-1)\epsilon_3'\\ &= n(B-1)[R_{11} - I(U_1, V_1; U_2, V_2, Y_4) - \epsilon_5]\\ &\overset{\because R_{11}-R_1=I(U_1,V_1;U_2,V_2,Y_4)}{=} n(B-1)[R_1 - (\epsilon_{1j} + \epsilon_5)]\end{aligned}$$

Chapter 6

Conclusion and Future Work

As mentioned in the introductory chapter, the introduction of even a single relay entails many complexities. In chapters 4 and 5, we have considered "strong" relay scenarios i.e. scenarios where the relay performs DF. We also have some (preliminary) results in the "weak" relay scenario, described below. Here, compress-forward (CF) is the preferred option [27, Chapter 16, Section 16.7]. Our preliminary results, based on combining insights obtained from the work of Luo and co-workers' [81] and Tang, Liu, Spasojevic, Poor [148], and also Wu and Xie (see Wu's PhD thesis [58], for example) seem to indicate that noise forwarding is unnecessary, and possibly suboptimal.

- *Relay-Eavesdropper*

  Chapter 4 studies the relay-eavesdropper channel with a "strong" relay – this means that the relay can decode completely the message intended for the destination, and that the Tx→Rel link is stronger than the Tx→Rx link. What happens if the relay is "weak" i.e. the Tx→Rel link is weaker than the Tx→Rx link? In this case, requiring the relay to decode the message would introduce an unnecessary bottleneck and give a suboptimal solution. The "weak" relay problem in the context of an external eavesdropper was also first studied by Lai and Gamal [72, Theorems $3, 4$]. Each compression sequence was represented by many different relay channel codewords – a technique they referred to as "noise forwarding" (NF). No Wyner-Ziv binning of the compression sequences was performed. Tang, Liu, Spasojevic, Yates [148] (see also [149]) suggested an improvement to the NF technique developed by [72, Theorem 3] by noting that there was no need for the eavesdropper to be able to decode the relay codeword. We take [148] as our starting point.

  For the canonical relay channel using CF (hence "weak" relay scenario), Luo, Gohary, Yanikomeroglu [81, 57] have developed a decoding technique which uses WZ binning. The usual constraint on relay codeword rate that enables it to be decoded uniquely is replaced by a constraint on the compression sequence rate that acts as a proxy for the relay codeword to be decoded uniquely. The work of Wu, Fa, Xie shows that too many compression sequences in a WZ bin may themselves act as "noise" [111], [112], [58]. *Thus there may not be need for separate NF step as used by [72].* We are currently investigating this matter.

Post-Review Update:

The following is a plausible achievable region via a CF scheme without NF obtained after the thesis reviews arrived. The possible enlargement of the achievable rate due to this scheme is of future interest.

We define:

$$WZ^{Bob} \overset{\text{def}}{=} I(\hat{Y}_2; X_1, Y_3|X_2) \text{ and } WZ^{Eve} \overset{\text{def}}{=} I(\hat{Y}_2; X_1, Z|X_2)$$

We use R_1 to denote the achievable secrecy rate. Also, the general constraint $\hat{R} \geq R_2$ holds throughout.

1. Case 1 : $I(X_2;Z) + WZ^{Eve} < I(X_2;Y_3) + WZ^{Bob}$.
   (a) Case $1(a)$: $(I(X_2;Z) <) \, I(X_2;Z) + WZ^{Eve} < I(X_2;Y_3) + I(\hat{Y}_2;Y_3|X_2)$
      i. Case $1(a)(i)$: $I(X_2;Z|X_1) < I(X_2;Y_3) + I(\hat{Y}_2;Y_3|X_2)$. The choices $R_2 = \max\{I(X_2;Y_3), I(X_2;Z|X_1)\}+\epsilon$ and $\hat{R} = I(X_2;Y_3)+I(\hat{Y}_2;Y_3|X_2)-\epsilon > I(X_2;Z) + WZ^{Eve}$, enable the secrecy rate $R_1 = I(X_1;\hat{Y}_2,Y_3|X_2) - I(X_1;Z)$.
      ii. Case $1(a)(ii)$: $I(X_2;Y_3)+I(\hat{Y}_2;Y_3|X_2) < I(X_2;Z|X_1)$. The choices $R_2 = \max\{I(X_2;Y_3), I(X_2;Z)\}+\epsilon$, $\hat{R}$ as before, enable $R_1 = I(X_1;\hat{Y}_2,Y_3|X_2) - [I(X_1,X_2;Z) - R_2]$. *The penalty term can be reduced by choosing* $\epsilon > 0$ *s.t.* $R_2 \uparrow \hat{R}$.
   (b) Case $1(b)$: $I(X_2;Y_3) + I(\hat{Y}_2;Y_3|X_2) < I(X_2;Z) + WZ^{Eve} < I(X_2 : Y_3) + WZ^{Bob}$
      i. Case $1(b)(i)$: $I(X_2;Z|X_1) < I(X_2;Y_3) + WZ^{Bob}$. The choices $R_2 = \max\{I(X_2;Z|X_1), I(X_2;Y_3)\} + \epsilon(> I(X_2;Z))$ and $\hat{R} \in (I(X_2;Z)+WZ^{Eve}, I(X_2;Y_3)+WZ^{Bob})(> I(X_2;Y_3)+I(\hat{Y}_2;Y_3|X_2))$, enable the secrecy rate $R_1 = I(X_1;\hat{Y}_2,Y_3|X_2)+I(X_2;Y_3)+I(\hat{Y}_2;Y_3|X_2) - \max\{I(X_2;Z|X_1), I(X_2;Y_3)\} - I(X_1;Z)$.
      ii. Case $1(b)(ii)$: $I(X_2;Y_3) + WZ^{Bob} < I(X_2;Z|X_1)$. The choices $R_2 > \max\{I(X_2;Y_3), I(X_2;Z)\} + \epsilon$ and $\hat{R} \in (I(X_2;Z) + WZ^{Eve}, I(X_2;Y_3) + WZ^{Bob})$ enable the secrecy rate $R_1 = I(X_1;\hat{Y}_2,Y_3|X_2) + I(X_2;Y_3) + I(\hat{Y}_2;Y_3|X_2) - [I(X_1,X_2;Z)]$.

2. Case 2 : $I(X_2;Y_3) + WZ^{Bob} < I(X_2;Z) + WZ^{Eve}$

    (a) Case 2(a) : $WZ^{Eve} < WZ^{Bob}(\implies I(X_2;Y_3) < I(X_2;Z))$

    i. Case 2$(a)(i)$: $WZ^{Eve} < I(\hat{Y}_2;Y_3|X_2)$. The choices $R_2 > I(X_2;Y_3), \hat{R} - R_2 > WZ^{Eve}, \hat{R} < I(X_2;Y_3) + I(\hat{Y}_2;Y_3|X_2)$ are consistent and enable secrecy rate $R_1 = I(X_1;\hat{Y}_2,Y_3|X_2) - I(X_1;Z|X_2)$.

    ii. Case 2$(a)(ii)$: $I(\hat{Y}_2;Y_3|X_2) < WZ^{Eve} < WZ^{Bob}$. The choices $R_2 > I(X_2;Y_3), \hat{R} - \breve{R} > WZ^{Eve}, \hat{R} < I(X_2;Y_3) + WZ^{Bob}$ enable secrecy rate $R_1 = [I(X_1;Y_3|X_2) - I(X_1;Z|X_2)] + [WZ^{Bob} - WZ^{Eve}]$.

    (b) Case 2(b) : $WZ^{Bob} < WZ^{Eve}$

    i. Case 2$(b)(i)$: $I(X_2;Y_3) < I(X_2;Z)$. The choices $R_2 < I(X_2;Y_3)$ and $\hat{R} > I(X_2;Z) + WZ^{Eve}$ enable the secrecy rate $R_1 = I(X_1;Y_3|X_2) - I(X_1;Z|X_2)$.

    ii. Case 2$(b)(ii)$: $I(X_2;Z) < I(X_2;Y_3)$.

    A. $I(X_2;Z) < I(X_2;Y_3) < I(X_2;Z|X_1)$. The choices $R_2 \in (I(X_2;Z), I(X_2;Y_3))$, and $\hat{R} - R_2 > WZ^{Eve}$ enable the secrecy rate $R_1 = I(X_1;Y_3|X_2) - [I(X_1,X_2;Z) - R_2]$. *The penalty term is reduced by choosing* $\breve{R} \uparrow I(X_2;Y_3)$

    B. $I(X_2;Z) < I(X_2;Z|X_1) < I(X_2;Y_3)$. The choices $R_2 \in (I(X_2;Z|X_1), I(X_2;Y_3))$ and $\hat{R} - R_2$ as before, enable the secrecy rate $R_1 = I(X_1;Y_3|X_2) - I(X_1;Z)$.

In many regimes, the relay both assists Bob and interferes with Eve.
End of Update.

- *Relay-Broadcast with Mutual Secrecy: DF + "Peeling Off" + CF case.*

The relay plays a starring role. It is possible that of two receivers, the relay is "strong" with respect to one (say, Rx 1) and "weak" with respect to the other (say, Rx 2). The relay can perform the following operations on its received sequence. It decodes the message of Rx 1, and then "peels off" (see [150, Theorem 2] and [69, Theorems 1, 4, 5]) the transmitted codeword from its received sequence. It re-encodes the message intended for Rx 1. It then compresses the *peeled off sequence* and performs WZ binning on it. Finally, taking into account the requirements for coherent transmission for the DF receiver and mutual secrecy for both receivers, the relay broadcasts a re-encoded version of the message intended for Rx 1 together with a compressed version of the "peeled off" sequence intended for Rx 2. This has not been done – to the best of our knowledge – even in the case of no secrecy requirements, at least using [76]'s insight that the relay be used as a broadcast channel. We are currently exploring these matters.

Remark 29. ***Peeling off:*** *It seems to us that the papers that have discussed "peeling off" have not described the theoretical justification for the appropriate expressions. We have described what we think is the correct way to think about the process using an application of the functional representation lemma [27, Appendix B].*

Studying the combined effect of a pair of secret keys and a trusted relay would also be a worthwhile exercise. This would, in effect, constitute a generalization of both the models studied in chapter 3 and chapter 5.

Consider Ekrem and Ulukus [69]. This is a three-node network that they have referred to as a relay broadcast channel. They study both a one-sided cooperative link as well as a two-sided cooperative link. It will be interesting to see what the addition of a relay trusted by one (or both destination nodes) has on the achievable region of [69].

Many of the following are natural generalizations of interest in their own right.

- *Common message, rate-equivocation:* A very natural generalization is to consider a common message to both receivers. Further, in our achievable schemes, we have only considered pure secrecy in all the models studied in chapters 3 to 5. It may be that pure secrecy is not desired for one or both receivers, i.e. a portion of one or both receivers' messages is private but not necessarily secret. In that case, we would need to obtain the rate-equivocation region.

- *Multiple relays, with all relays trusted by both destinations:* assisting in creating mutual secrecy through pure DF based techniques. This would be a natural and useful generalization of the work of Bassily and Ulukus [151]. The techniques developed in Razaghi and Yu [122] are applicable to this problem and may need to be extended keeping the constraint of mutual secrecy in mind.

  The (single-)relay broadcast channel using DF considers coherent transmission between two broadcast channels, with both the transmitter and relay treated as broadcast channels in their own right. With multiple relays, an extension of the lemma developed in [76] would be required if all relays perform DF and coherent transmission is desired.

- *Multiple relays, with each destination trusting some – possibly distinct– subset:* A pair of relays, each trusted by exactly one receiver. This would also constitute a generalization of the model studied by Behboodi and Piantanida [71] to the problem of mutual secrecy. Note that in this case, each destination has to contend with two eavesdroppers – the other destination and "its" relay. [71] employ backward decoding, which we can dispense with.

If multiple relays are present, some subset may not be trusted by one or more destinations, yet their assistance may prove useful, as demonstrated by He and Yener in a sequence of papers [102], [104], [152], [105]. (This automatically necessitates CF based techniques).

- Wu and Xie [60] have developed a framework with both D-F and C-F nodes. This is a current topic of much interest. See also Behboodi and Piantanida [71]. The introduction of secrecy and/or untrusted relays will add an additional degree of complexity, as an untrusted relay can only employ CF techniques.

- *Side information:* The impact of receiver side information on secrecy in *broadcast channels* is of interest, see, for example, [153]. The presence of a relay that is aware of some or all of the side information would be useful to understand.

- *Cognitive relay and/or relay as recipient:* The relay(s) may itself (themselves) have messages of its (their) own to transmit or to receive. This kind of scenario was first explored by Tannious and Nosratinia [150], and also by Ekrem and Ulukus [69], and more recently, by Nagananda [54, 107].

- It is likely that the techniques developed by Razaghi and Yu [117] and which were developed for the interference channel will also find application in multi-relay networks, as Zhong, Haija, Vu [95] indicate. This holds even more strongly if the two destinations are attempting to eavesdrop on each other's message.

- ≥ 3 *destinations:* Chia and Gamal [56], [154] have studied the 3-receiver broadcast channel with confidential messages. It would be interesting to study the impact of a dedicated relay. Chia and Gamal also have some results on 2-eavesdropper scenarios, which would give insights into the above suggested generalization of Behbodi and Piantanida. *Note that indirect (aka nonunique) decoding is used, and the work of Bidokhti, Diggavi and Prabhakaran [155], [156] indicates that this may be unnecessary*

  In the above, we have considered secrecy. A different kind of generalization of Zhao and Chung [76] – without taking secrecy into account – would be to consider ≥ 3 destinations and a dedicated relay. Note that there are partial results on broadcast channels with 3 destinations in the literature, for example, Chia and El Gamal [157] (which involves secrecy), and Nair and El Gamal [158], [159] (which do not). The introduction of single relay in such a scenario would require studying the interaction between two 3-receiver broadcast channels.

Bibliography

[1] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948. [Online]. Available: http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

[2] Aftab, Cheung, Kim, Thakkar, and Yeddanapudi, “Information Theory: Information Theory And The Digital Age.”

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition.* Wiley-Interscience, 2006.

[4] R. W. Yeung, *A First Course in Information Theory.* Springer, 2006.

[5] ——, *Information Theory and Network Coding.* Springer, 2008.

[6] A. Dembo, T. M. Cover, and J. A. Thomas, “Information Theoretic Inequalities,” *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1501–1518, Nov 1991.

[7] A. Dembo, “Information Inequalities and Concentration of Measure,” *The Annals of Probability*, vol. 25, no. 2, pp. 927–939, 1997. [Online]. Available: http://www.jstor.org/stable/2959616

[8] M. Raginsky and I. Sason, “Concentration of Measure Inequalities in Information Theory, Communications, and Coding,” *Foundations and Trends in Communications and Information Theory*, vol. 10, no. 1-2, pp. 1–246, 2013. [Online]. Available: http://dx.doi.org/10.1561/0100000064

[9] J. V. Linnik, “An Information-Theoretic Proof of the Central Limit Theorem with Lindeberg Conditions,” *Theory of Probability & Its Applications*, vol. 4, no. 3, pp. 288–299, 1959. [Online]. Available: https://doi.org/10.1137/1104028

[10] R. Ahlswede, “The Final Form Of Tao’s Inequality Relating Conditional Expectation And Conditional Mutual Information,” *Advances in Mathematics of Communications*, vol. 1, no. 2, pp. 239–242, 2007. [Online]. Available: http://aimsciences.org/journals/displayArticlesnew.jsp?paperID=2565

[11] M. Madiman, A. W. Marcus, and P. Tetali, “Entropy and set cardinality inequalities for partition-determined functions,” *Random Structures & Algorithms*, vol. 40, no. 4, pp. 399–424, 2012. [Online]. Available: http://dx.doi.org/10.1002/rsa.20385

[12] E. Friedgut, “Hypergraphs, Entropy, and Inequalities,” *The American Mathematical Monthly*, vol. 111, no. 9, pp. 749–760, 2004. [Online]. Available: http://www.jstor.org/stable/4145187

[13] S. Pokutta, “Information Theory and Polyhedral Combinatorics,” in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2015, pp. 1119–1126.

[14] D. Guo, S. Shamai, and S. Verdú, “Mutual Information and Minimum Mean-Square Error in Gaussian Channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[15] D. Guo, “Gaussian Channels: Information, Estimation and Multiuser Detection,” Ph.D. dissertation, Princeton University, 2004.

[16] D. Guo, S. S. (Shitz), and S. Verdú, “The Interplay Between Information and Estimation Measures,” *Foundations and Trends in Signal Processing*, vol. 6, no. 4, pp. 243–429, 2013. [Online]. Available: http://dx.doi.org/10.1561/2000000018

[17] M. M. Wilde, *Quantum Information Theory.* Cambridge University Press, 2017.

[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, 2011.

[19] F. Rieke, D. Warland, R. d. R. van Steveninck, and W. Bialek, *Spikes: Reading the Neural Code.* Bradford, 1996.

[20] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* The MIT Press, 2005.

[21] I. Csiszár and P. Shields, “Information Theory and Statistics: A Tutorial,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 417–528, 2004. [Online]. Available: http://dx.doi.org/10.1561/0100000004

[22] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.

[23] F. Escolano, P. Suau, and B. Bonev, *Information Theory in Computer Vision and Pattern Recognition.* Springer, 2009.

[24] N. Tishby, F. C. Pereira, and W. Bialek, "The Information Bottleneck Method," in *Proceedings of the 37th Allerton Conference on Communication, Control and Computation*, vol. 49, Sept 2001, pp. 368–377.

[25] N. Tishby and N. Zaslavsky, "Deep Learning and the Information Bottleneck Principle," *CoRR*, vol. abs/1503.02406, 2015. [Online]. Available: http://arxiv.org/abs/1503.02406

[26] N. Slonim, "The Information Bottleneck: Theory and Applications," Ph.D. dissertation, Hebrew University, 2002.

[27] A. E. Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge University Press, 2011.

[28] L.-L. Xie and P. R. Kumar, "A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[29] A. Carleial, "Multiple-Access Channels with Different Generalized Feedback Signals," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 841–850, Nov 1982.

[30] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless Network Information Flow: A Deterministic Approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[31] R. Vaze, *Random Wireless Networks: An Information-Theoretic Approach.* Cambridge University Press, 2015.

[32] S. Kannan, A. Raja, and P. Viswanath, "Approximately Optimal Wireless Broadcasting," *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7154–7167, Dec 2012.

[33] A. Raja and P. Viswanath, "Compress-and-Forward Scheme for Relay Networks: Backword Decoding and Connection to Bisubmodular Flows," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5627–5638, Sept 2014.

[34] V. Anantharam and V. S. Borkar, "Common randomness and distributed control: A counterexample," *Systems & Control Letters*, vol. 56, no. 7-8, pp. 568–572, 2007. [Online]. Available: https://doi.org/10.1016/j.sysconle.2007.03.010

[35] P. W. Cuff, "Communication in Networks for Coordinating Behaviour," Ph.D. dissertation, Stanford University, 2009.

[36] P. W. Cuff, H. H. Permuter, and T. M. Cover, “Coordination Capacity,” *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4181–4206, Sept 2010.

[37] R. A. Chou, M. R. Bloch, and J. Kliewer, “Polar Coding for Empirical and Strong Coordination via Distribution Approximation,” in *2015 IEEE International Symposium on Information Theory (ISIT)*, June 2015, pp. 1512–1516.

[38] S. A. Obead, B. N. Vellambi, and J. Kliewer, “Strong Coordination over Noisy Channels: Is Separation Sufficient?” in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 2840–2844.

[39] G. Cervia, L. Luzzi, M. L. Treust, and M. R. Bloch, “Strong Coordination of Signals and Actions over Noisy Channels,” in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 2835–2839.

[40] F. Haddadpour, M. H. Yassaee, A. Gohari, and M. R. Aref, “Coordination via a Relay,” in *Proceedings of the 2012 IEEE International Symposium on Information Theory, ISIT 2012, Cambridge, MA, USA, July 1-6, 2012*, 2012, pp. 3048–3052. [Online]. Available: https://doi.org/10.1109/ISIT.2012.6284121

[41] M. H. Yassaee, M. R. Aref, and A. Gohari, “Achievability Proof via Output Statistics of Random Binning,” *CoRR*, vol. abs/1203.0730, 2012. [Online]. Available: http://arxiv.org/abs/1203.0730

[42] N. Lynch, *Distributed Algorithms.* Morgan Kaufmann, 1996.

[43] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Information Theoretic Security,” *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 45, pp. 355–580, 2009. [Online]. Available: http://dx.doi.org/10.1561/0100000036

[44] M. Bloch and J. Barros, *Physical Layer Security.* Cambridge University Press, 2011.

[45] M. Debbah, H. El-Gamal, H. V. Poor, and S. Shamai (Shitz), “Wireless Physical Layer Security,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, p. 404061, Mar 2010. [Online]. Available: https://doi.org/10.1155/2009/404061

[46] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Physical Layer Security in Broadcast Networks,” *Security and Communication Networks*, vol. 2, no. 3, pp. 227–238, 2009. [Online]. Available: http://dx.doi.org/10.1002/sec.110

[47] R. Bassily, E. Ekrem, X. He, E. Tekin, J. Xie, M. R. Bloch, S. Ulukus, and A. Yener, “Cooperative Security at the Physical Layer: A Summary of Recent Advances,” *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 16–28, Sept 2013.

[48] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1550–1573, Third 2014.

[49] P. Narayan and H. Tyagi, "Multiterminal Secrecy by Public Discussion," *Foundations and Trends in Communications and Information Theory*, vol. 13, no. 2-3, pp. 129–275, 2016. [Online]. Available: http://dx.doi.org/10.1561/0100000072

[50] R. Bassily, "Alignment and Cooperation for Secrecy in Multi-User Channels," Ph.D. dissertation, University of Maryland, College Park, 2011.

[51] F. Gabry, "Cooperation for Secrecy in Wireless Networks," Ph.D. dissertation, KTH Royal Institute of Technology, 2012.

[52] X. He, "Cooperation and Information-Theoretic Security in Wireless Networks," Ph.D. dissertation, The Pennsylvania State University, 2010.

[53] J. Hou, "Coding for Relay Networks and Effective Secrecy for Wiretap Channels," Ph.D. dissertation, Technische Universitat Munchen, 2014.

[54] K. Nagananda, "Multiuser Wireless Networks: Cooperation and Physical-Layer Security," Ph.D. dissertation, Lehigh University, 2013.

[55] E. Perron, "Information-Theoretic Security for Wireless Networks," Ph.D. dissertation, Ecole Polytechnique Federale de Lausanne, 2009.

[56] Y.-K. Chia, "Multi-Terminal Secrecy and Source Coding," Ph.D. dissertation, Stanford University, 2011.

[57] W. Luo, "Enhancing Rates in Relay Channels," Ph.D. dissertation, Carleton University, 2015.

[58] X. Wu, "Coding Schemes for Multiple-Relay Channels," Ph.D. dissertation, University of Waterloo, 2013.

[59] P. Zhong, "Coding Schemes for the Two-Way Relay Channel," Master's thesis, McGill University, 2012.

[60] X. Wu and L. L. Xie, "A Unified Relay Framework With Both D-F and C-F Relay Nodes," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 586–604, Jan 2014.

[61] K. Iyer, S. R. B. Pillai, and B. K. Dey, "Power Controlled Adaptive Sum-Capacity in the Presence of Distributed CSI," in *Proceedings of the International Symposium on Information Theory and its Applications, ISITA 2012, Honolulu, HI, USA, October 28-31, 2012*, 2012, pp. 66–70. [Online]. Available: http://ieeexplore.ieee.org/document/6401024/

[62] ——, "On the Adaptive Sum-Capacity of Distributed Multiple Access with Individual CSI," *IETE Journal of Research*, May 2015. [Online]. Available: http://dx.doi.org/10.1080/03772063.2015.1027306

[63] S. Sreekumar, B. K. Dey, and S. R. B. Pillai, "Distributed Rate Adaptation and Power Control in Fading Multiple Access Channels," *Information Theory, IEEE Transactions on*, vol. 61, no. 10, pp. 5504–5524, Oct 2015.

[64] Y. Deshpande, S. Pillai, and B. Dey, "On the Sum Capacity of Multiaccess Block-Fading Channels with Individual Side Information," in *Information Theory Workshop (ITW), 2011 IEEE*, Oct 2011, pp. 588–592.

[65] R. Liu, I. Marić, P. Spasojević, and R. D. Yates, "Discrete Memoryless Interference and Broadcast Channels With Confidential Messages: Secrecy Rate Regions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2493–2507, June 2008.

[66] E. Perron, S. Diggavi, and E. Telatar, "On Cooperative Wireless Network Secrecy," in *INFOCOM 2009, IEEE*, April 2009, pp. 1935–1943.

[67] W. Kang and N. Liu, "Wiretap Channel with Shared Key," in *Information Theory Workshop (ITW), 2010 IEEE*, Aug 2010, pp. 1–5.

[68] K. Iyer, "Broadcast Channel with Confidential Messages and Secret Keys," in *2016 Twenty Second National Conference on Communication (NCC)*, March 2016, pp. 1–6.

[69] E. Ekrem and S. Ulukus, "Secrecy in Cooperative Relay Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 137–155, Jan 2011.

[70] B. Dai, L. Yu, and Z. Ma, "Relay Broadcast Channel With Confidential Messages," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 410–425, Feb 2016.

[71] A. Behboodi and P. Piantanida, "Cooperative Strategies for Simultaneous and Broadcast Relay Channels," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1417–1443, March 2013.

[72] L. Lai and H. E. Gamal, "The Relay-Eavesdropper Channel: Cooperation for Secrecy," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4005–4019, Sept 2008.

[73] M. Yuksel and E. Erkip, "The Relay Channel with a Wire-tapper," in *2007 41st Annual Conference on Information Sciences and Systems*, March 2007, pp. 13–18.

[74] P. Razaghi and W. Yu, "Parity Forwarding for Multiple-Relay Networks," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 158–173, Jan 2009.

[75] R. Liu and W. Trappe (eds.), *Securing Wireless Communications at the Physical Layer*. Springer, 2010.

[76] L. Zhao and S. Y. Chung, "Marton-Marton Coding for a Broadcast Relay Network," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, July 2013, pp. 1282–1286.

[77] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sept 2005.

[78] K. Marton, "A Coding Theorem for the Discrete Memoryless Broadcast Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[79] Y. Liang and V. V. Veeravalli, "Cooperative Relay Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 900–928, March 2007.

[80] T. Cover and A. E. Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep 1979.

[81] K. Luo, R. H. Gohary, and H. Yanikomeroglu, "A Decoding Procedure for Compress-and-Forward and Quantize-and-Forward Relaying," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2012, pp. 2112–2119.

[82] Y. Liang and G. Kramer, "Rate Regions for Relay Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3517–3535, Oct 2007.

[83] R. Tannious and A. Nosratinia, "Relay Channel With Private Messages," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3777–3785, Oct 2007.

[84] M. Bloch and A. Thangaraj, "Confidential Messages to a Cooperative Relay," in *Information Theory Workshop, 2008. ITW '08. IEEE*, May 2008, pp. 154–158.

[85] A. Behboodi and P. Piantanida, "On the Simultaneous Relay Channel with Informed Receivers," in *2009 IEEE International Symposium on Information Theory*, June 2009, pp. 1179–1183.

[86] ——, "Capacity of a Class of Broadcast Relay Channels," in *2010 IEEE International Symposium on Information Theory*, June 2010, pp. 590–594.

[87] ——, "Broadcasting over the Relay Channel with Oblivious Cooperative Strategy," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, Sept 2010, pp. 1098–1103.

[88] L.-L. Xie and P. R. Kumar, "An Achievable Rate for the Multiple-Level Relay Channel," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1348–1358, April 2005.

[89] H. F. Chong, M. Motani, and H. K. Garg, "Generalized Backward Decoding Strategies for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 394–401, Jan 2007.

[90] ——, "On Achievable Rates for Relay Channels," in *2007 Information Theory and Applications Workshop*, Jan 2007, pp. 431–436.

[91] H. F. Chong and M. Motani, "On Achievable Rates for the General Relay Channel," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1249–1266, March 2011.

[92] J. Hou and G. Kramer, "Short Message Noisy Network Coding with a Decode-Forward Option," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 89–107, Jan 2016.

[93] R. Dabora and S. D. Servetto, "On the Role of Estimate-and-Forward with Time Sharing in Cooperative Communication," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4409–4431, Oct 2008.

[94] A. E. Gamal, M. Mohseni, and S. Zahedi, "Bounds on Capacity and Minimum Energy-Per-Bit for AWGN Relay Channels," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1545–1561, 2006. [Online]. Available: http://dx.doi.org/10.1109/TIT.2006.871579

[95] P. Zhong, A. A. A. Haija, and M. Vu, "On Compress-Forward without Wyner-Ziv Binning for Relay Networks," *CoRR*, vol. abs/1111.2837, 2011. [Online]. Available: http://arxiv.org/abs/1111.2837

[96] M. H. Yassaee and M. R. Aref, “Slepian-wolf coding over cooperative relay networks,” *CoRR*, vol. abs/0910.3509, 2009. [Online]. Available: http://arxiv.org/abs/0910.3509

[97] Y. Tang, “Partial Decode-Forward in Relay Networks,” Master’s thesis, McGill University, 2013.

[98] Y. Tang and M. H. Vu, “A Partial Decode-Forward Scheme For A Network with N Relays,” *2013 47th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2013.

[99] Y. Oohama, “Capacity Theorems for Relay Channels with Confidential Messages,” in *2007 IEEE International Symposium on Information Theory*, June 2007, pp. 926–930.

[100] Y. Oohama and S. Watanabe, “Capacity Results for Relay Channels with Confidential Messages,” *CoRR*, vol. abs/1009.5829, 2010. [Online]. Available: http://arxiv.org/abs/1009.5829

[101] X. He and A. Yener, “On the Equivocation Region of Relay Channels with Orthogonal Components,” in *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, Nov 2007, pp. 883–887.

[102] ——, “Two-hop Secure Communication Using an Untrusted Relay: A Case for Cooperative Jamming,” in *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, Nov 2008, pp. 1–5.

[103] ——, “The Role of an Untrusted Relay in Secret Communication,” in *2008 IEEE International Symposium on Information Theory*, July 2008, pp. 2212–2216.

[104] ——, “End-to-end Secure Multi-hop Communication with Untrusted Relays is Possible,” in *2008 42nd Asilomar Conference on Signals, Systems and Computers*, Oct 2008, pp. 681–685.

[105] ——, “Cooperation With an Untrusted Relay: A Secrecy Perspective,” *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3807–3827, Aug 2010.

[106] A. A. Zewail, M. Nafea, and A. Yener, “Multi-terminal Networks with an Untrusted Relay,” in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014, pp. 895–902.

[107] K. G. Nagananda, “Secure communications over opportunistic-relay channels,” *Physical Communication*, vol. 7, pp. 105–121, 2013. [Online]. Available: https://doi.org/10.1016/j.phycom.2012.11.002

[108] S. R. Bhaskaran, “Gaussian Degraded Relay Broadcast Channel,” *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3699–3709, Aug 2008.

[109] G. Kramer and J. Hou, “Short-Message Quantize-Forward Network Coding,” in *Multi-Carrier Systems Solutions (MC-SS), 2011 8th International Workshop on*, May 2011, pp. 1–3.

[110] T. T. Kim, M. Skoglund, and G. Caire, “Quantifying the Loss of Compress-Forward Relaying without Wyner-Ziv Coding,” *IEEE Trans. Inf. Theor.*, vol. 55, no. 4, pp. 1529–1533, Apr. 2009. [Online]. Available: http://dx.doi.org/10.1109/TIT.2009.2013006

[111] X. Wu and L. L. Xie, “On the Optimal Compressions in the Compress-and-Forward Relay Schemes,” *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2613–2628, May 2013.

[112] X. Wu, G. Fan, and L. L. Xie, “An Optimality-Robustness Tradeoff in the Compress-and-Forward Relay Scheme,” in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, Sept 2010, pp. 1–5.

[113] T. M. Cover and Y. H. Kim, “Capacity of a Class of Deterministic Relay Channels,” in *2007 IEEE International Symposium on Information Theory*, June 2007, pp. 591–595.

[114] M. Mondelli, S. Hamed Hassani, and R. Urbanke, “A New Coding Paradigm for the Primitive Relay Channel,” *ArXiv e-prints*, Jan. 2018. [Online]. Available: http://arxiv.org/abs/1801.03153

[115] Y. H. Kim, “Coding Techniques for Primitive Relay Channels,” in *Forty-Fifth Annual Allerton Conference*, 2007, pp. 129–135.

[116] ——, “Coding Techniques for Primitive Relay Channels,” in *in Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput*, 2007.

[117] P. Razaghi and W. Yu, “Universal Relaying for the Interference Channel,” in *2010 Information Theory and Applications Workshop (ITA)*, Jan 2010, pp. 1–6.

[118] P. Xu, Z. Ding, X. Dai, and K. K. Leung, “An Improved Achievable Secrecy Rate for the Relay-Eavesdropper Channel,” in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, pp. 2440–2445.

[119] A. Sonee, S. Salimi, and Salmasizadeh, “A New Achievable Rate-Equivocation Region for the Relay-Eavesdropper Channel,” in *2010 18th Iranian Conference on Electrical Engineering*, May 2010, pp. 188–193.

[120] P. Xu, Z. Ding, and X. Dai, “A Hybrid Cooperative Coding Scheme for the Relay-Eavesdropper Channel,” *Entropy*, vol. 16, no. 3, p. 18191841, Mar 2014. [Online]. Available: http://dx.doi.org/10.3390/e16031819

[121] L. Chen, “Physical Layer Security for Cooperative Relaying in Broadcast Networks,” in *2011 - MILCOM 2011 Military Communications Conference*, Nov 2011, pp. 91–96.

[122] P. Razaghi and W. Yu, “A Structured Generalization of Decode-and-Forward Strategies for Multiple-Relay Networks,” in *2007 IEEE International Symposium on Information Theory*, June 2007, pp. 271–275.

[123] K. Marton, “A Coding Theorem for the Discrete Memoryless Broadcast Channel,” *IEEE Trans. Inf. Theory*, vol. IT-25, no. 3, pp. 306–311, 1979.

[124] S. Salehkalaibar, L. Ghabeli, and M. R. Aref, “An Outer Bound on the Capacity Region of Broadcast-Relay-Channel,” in *2010 IEEE International Symposium on Information Theory*, June 2010, pp. 599–603.

[125] C.-S. Hwang, M. Malkin, A. El Gamal, and J. M. Cioffi, “Multiple-Access Channels with Distributed Channel State Information,” in *ISIT*, June 2007, pp. 1561 –1565.

[126] S. Shamai and A. D. Wyner, “Information-Theoretic Considerations for Symmetric, Cellular, Multiple-Access Fading Channels – Part I,” *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1877–1894, 1997.

[127] R. Knopp and P. Humblet, “Information Capacity and Power Control in Single-Cell Multiuser Communications,” in *ICC ’95 Seattle*, vol. 1, Jun. 1995, pp. 331 –335.

[128] D. Tse and S. Hanly, “Multiaccess Fading Channels. I. Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities,” *Information Theory, IEEE Trans on*, vol. 44, no. 7, pp. 2796 –2815, Nov. 1998.

[129] A. Das and P. Narayan, “Capacities of Time-Varying Multiple-Access Channels With Side Information,” *Information Theory, IEEE Transactions on*, vol. 48, no. 1, pp. 4 –25, Jan. 2002.

[130] M. Effros, A. J. Goldsmith, and Y. Liang, “Generalizing Capacity: New Definitions and Capacity Theorems for Composite Channels,” *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3069–3087, 2010.

[131] Y. Cemal and Y. Steinberg, “The Multiple-Access Channel With Partial State Information at the Encoders,” *Information Theory, IEEE Transactions on*, vol. 51, no. 11, pp. 3992 – 4003, Nov. 2005.

[132] S. A. Jafar, "Channel Capacity with Causal and Non-Causal State Information– A Unified View," *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5468–5474, 2006.

[133] P. Minero, M. Franceschetti, and D. Tse, "Random Access: An Information-Theoretic Perspective," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 909 –930, Feb. 2012.

[134] U. Niesen, U. Erez, D. Shah, and G. Wornell, "Rateless Codes for the Gaussian Multiple Access Channel," in *IEEE GLOBECOM*, 2006, pp. 1 –5.

[135] X. Qin and R. Berry, "Exploiting Multiuser Diversity for Medium Access Control in Wireless Networks," in *INFOCOM*, vol. 2, 2003, pp. 1084 – 1094.

[136] S. Adireddy and L. Tong, "Exploiting Decentralized Channel State Information for Random Access," *Information Theory, IEEE Transactions on*, vol. 51, no. 2, pp. 537 – 561, Feb. 2005.

[137] C. S. Hwang, K. Seong, and J. M. Cioffi, "Opportunistic p-persistent CSMA in wireless networks," in *Proc. IEEE ICC 2006*, 2006, pp. 183–188.

[138] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[139] B. Rimoldi and R. Urbanke, "A Rate-Splitting Approach to the Gaussian Multiple-Access Channel," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 364 –375, Mar 1996.

[140] J. Xu, Y. Cao, and B. Chen, "Capacity Bounds for Broadcast Channels With Confidential Messages," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4529–4542, Oct 2009.

[141] H. Yamamoto, "Rate-Distortion Theory for the Shannon Cipher System," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 827–835, May 1997.

[142] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y. H. Kim, "Wiretap Channel With Secure Rate-Limited Feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, Dec 2009.

[143] X. Yin, L. Pang, Z. Xue, and Y. Zhou, "Degraded Broadcast Channels with Rate-Limited Feedback," in *2013 8th International Conference on Communications and Networking in China (CHINACOM)*, Aug 2013, pp. 911–916.

[144] B. Dai, A. J. H. Vinck, Y. Luo, and Z. Zhuang, "Capacity Region of Non-degraded Wiretap Channel with Noiseless Feedback," in *2012 IEEE International Symposium on Information Theory Proceedings*, July 2012, pp. 244–248.

[145] R. F. Schaefer, A. Khisti, and H. Boche, "On The Use Of Secret Keys In Broadcast Channels With Receiver Side Information," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 1582–1586.

[146] R. F. Schaefer and A. Khisti, "Secure Broadcasting of a Common Message with Independent Secret Keys," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, March 2014, pp. 1–6.

[147] I. Csiszár and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 1981.

[148] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Interference Assisted Secret Communication," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3153–3167, May 2011.

[149] N. Marina, H. Yagi, and H. V. Poor, "Improved Rate-Equivocation Regions for Secure Cooperative Communication," *CoRR*, vol. abs/1102.3500, 2011. [Online]. Available: http://arxiv.org/abs/1102.3500

[150] R. Tannious and A. Nosratinia, "Relay Channel with Private Messages," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3777–3785, Oct 2007.

[151] R. Bassily and S. Ulukus, "Secure Communication in Multiple Relay Networks Through Decode-and-Forward Strategies," *Journal of Communications and Networks*, vol. 14, no. 4, pp. 352–363, Aug 2012.

[152] X. He and A. Yener, "End-to-end Secure Multi-Hop Communication with Untrusted Relays is Possible," in *2008 42nd Asilomar Conference on Signals, Systems and Computers*, Oct 2008, pp. 681–685.

[153] R. F. Wyrembelski, A. Sezgin, and H. Boche, "Secrecy in Broadcast Channels with Receiver Side Information," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2011, pp. 290–294.

[154] Y. K. Chia and A. E. Gamal, "3-Receiver Broadcast Channels with Common and Confidential Messages," in *2009 IEEE International Symposium on Information Theory*, June 2009, pp. 1849–1853.

[155] S. S. Bidokhti, V. M. Prabhakaran, and S. N. Diggavi, “Is Non-Unique Decoding Necessary?” in *2012 IEEE International Symposium on Information Theory Proceedings*, July 2012, pp. 398–402.

[156] S. S. Bidokhti and V. M. Prabhakaran, “Is Non-Unique Decoding Necessary?” *IEEE Trans. Information Theory*, vol. 60, no. 5, pp. 2594–2610, 2014. [Online]. Available: http://dx.doi.org/10.1109/TIT.2014.2312285

[157] Y.-K. Chia and A. El Gamal, “Three-Receiver Broadcast Channels with Common and Confidential Messages,” *Information Theory, IEEE Transactions on*, vol. 58, no. 5, pp. 2748–2765, May 2012.

[158] C. Nair and A. E. Gamal, “The Capacity Region of a Class of 3-Receiver Broadcast Channels with Degraded Message Sets,” in *2008 IEEE International Symposium on Information Theory*, July 2008, pp. 1706–1710.

[159] ——, “The Capacity Region of a Class of Three-Receiver Broadcast Channels With Degraded Message Sets,” *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4479–4493, Oct 2009.

Publications based on this Thesis

Journal Papers

[J1] K. Iyer, S. R. B. Pillai, and B. K. Dey, “On the Adaptive Sum-Capacity of Distributed Multiple Access with Individual CSI,” *IETE Journal of Research*, May 2015. [Online]. Available: http://dx.doi.org/10.1080/03772063.2015.1027306

Conference Proceedings

[C1] K. Iyer, S. R. B. Pillai, and B. K. Dey, “Power Controlled Adaptive Sum-Capacity in the Presence of Distributed CSI,” in *Proceedings of the International Symposium on Information Theory and its Applications, ISITA 2012, Honolulu, HI, USA, October 28-31, 2012*, 2012, pp. 66–70. [Online]. Available: http://ieeexplore.ieee.org/document/6401024/

[C2] K. Iyer, “Broadcast Channel with Confidential Messages and Secret Keys,” in *2016 Twenty Second National Conference on Communication (NCC)*, March 2016, pp. 1–6.

[C3] K. Iyer, “Two Receiver Relay Broadcast Channel with Mutual Secrecy,” in *International Conference on Signal Processing and Communications, SPCOM 2018, Bangalore, India.*

Acknowledgements

A PhD thesis is the culmination of many years of intellectual effort. Its successful completion has been possible due to the support, guidance, and encouragement I received from many individuals. I would like to acknowledge their contribution and take this opportunity to express my sincerest thanks to them.

It takes a village to raise a child.

First and foremost, I would like to express my deepest gratitude to my advisor Professor Bikash Kumar Dey and co-advisor Professor Sibi Raj B. Pillai.

In the initial three years of my PhD, Professor Dey's patience enabled me to slowly and steadily find my bearings in research. He was extremely supportive of my crediting many Mathematics and Computer Science department courses, and these, together with discussions with him, played a key role in building my confidence and developing my capacity for abstract thinking. He also showed understanding when I felt jaded and gave me the freedom to read and explore very widely. His contributions were always incisive, even when the topic lay outside the main area of his interests.

I will always be in debt to Professor Sibi Raj B. Pillai for involving me in his research project which led to my first conference-cum-journal paper, which in turn led to the S.K. Mitra memorial award, my first. He gave a patient hearing to my often tentative initial attempts at formulating a research problem. His insights greatly contributed to improving the quality of solutions to my research problems.

I would like to thank my academic grandfather, Professor B. Sundar Rajan of the Indian Institute of Science Bangalore who brought Gamal and Kim's pre-publication textbook notes on network information theory to my attention in May 2011. This turned my interest from network coding to network information theory.

I would like to thank the faculty of the departments of Electrical Engineering, Computer Science, and Mathematics for allowing me to attend courses before joining the PhD program and even assigning me unofficial grades. In particular, I would like to acknowledge my gratitude to Professors U.B. Desai, R. Shevgaonkar, Harish Pillai, D. Manjunath, Sadanand Agashe, Shreevardhan A. Soman, Subhasis Chaudhuri all of Electrical Engineering, Professors Amiya K. Pani and Neela Nataraj of Mathematics, and Professors Soumen Chakrabarti and Supratim Biswas of Computer Science.

I would like to thank Professor Shreevardhan A. Soman in particular for giving me the opportunity to work in the PowerAnser lab just prior to joining the PhD program. He gave me the freedom to attend the aforementioned courses in EE, CS and Math

while I was working as a research project staff in his lab.

I do not have words enough to adequately express my gratitude to Professor Raghava Varma of the department of Physics at IIT Bombay whose staunch support, kindness and immense faith in me during a difficult time gave me the moral confidence to pursue doctoral studies. I would also like to thank Professor Jayanta K. Bhattacharjee of the Indian Association for the Cultivation of Science, who encouraged me to continue with academia and obtain a PhD.

I would like to thank Professor Abhay Karandikar for my first introduction to communication engineering through his course on Digital Message Transmission. Thanks are also due to Professor Harish Pillai for his course on Finite Fields.

I have had the good fortune to learn Digital Signal Processing and Wavelets from Professor Vikram Gadre. Since my undergraduate days, Professor Gadre has always been there for me in times of need and has always supported me and demonstrated immense faith in me.

I thank Professor Harihar Narayanan for teaching me Matroids and Professor S. Agashe for teaching me Linear Algebra and Optimal Control, and for being exemplary and inspiring role models. I also owe thanks to Professor Vivek Shripad Borkar for giving me the opportunity to be the TA in a course taught by him, and for showing me that a person can be both a great scientist and a wonderful human being.

I owe special thanks to the following Professors of the CS department for their courses: Sunita Sarawagi for Graphical Models and Structured Learning, Ajit A. Diwan for the Design of Algorithms, and Amitabha Sanyal on Functional Programming. Each of these courses opened the gateway to a new world of ideas.

I would like to thank the following Professors of the department of Mathematics: Anant Shastri for Real Analysis, Sivaramakrishnan Sivasubramanian for Combinatorics, and for DS and Algorithms. Inder Kumar Rana for Measure Theory, Santanu Dey for Functional Analysis, and K. Suresh for Advanced Probability and Alladi Subramanyam for Stochastic Processes, and Bata Krishna Das for Operators on Hilbert Spaces. I would also like to thank Gopal K. Srinivasan and Swapneel Mahajan respectively for their sound advice regarding Real Analysis and Mathematics department courses. As with CS, these courses also introduced me to new worlds of thought.

I thank Professors Nandyala Hemachandra and K. S. Mallikarjuna Rao of the department of Industrial Engineering and Operations Research. Professor Hemachandra for his advanced course on the Applications of Stochastic Models and both Professors Hemachandra and Rao for their course on Networks, Games, and Algorithms.

My special thanks to Professors Madhu Belur and Saravanan Vijayakumaran for generously granting me access to computer resources.

Having thanked my teachers and mentors, I would now like to make a special mention of friends and classmates who have had a positive impact during my studies.

Of my many friends, Shantanu Desai deserves a special mention. He has stood by me since my undergraduate days at IIT Bombay. He has given me moral and psychological support and exemplifies the phrase “A friend in need is a friend indeed.”

From my days in the Essar Corporate house, I would like to thank Brinda Ramanathan, Jyoti Fernandes, and Monica Yadav.

I may not be able to list all who have contributed, but I would like to mention the following who have played a positive role in my journey: Arnab Majumdar, Ankhi Roy, Neha Shah, Himani Bhat, Kavita Lalwani nee Chandwani, Arnab Ray, Prashant Navalkar, Nilesh Bhattad, Viji Paul, Shishir Sahay, Imdad Rizvi, Sandip Kothari, Sanket Lad, Prasad Gawde, Mugdha Nazare, (Prasad and Mugdha are both members of the Art of Living, and my sincere thanks goes to them both in particular and to the AoL in general), Chirag Patel, Shalini Gupta, Vishwesh Vyawahare, Kunal Puri, Rajeev Gajbhiye, Sucheta Dudwadkar, Parakram Majumdar, Hari Priyadarshan, Gautam Shenoy, Amit Verma, Bhushan Virkar, Shraddha Srivastava, Santhana Krishnan, Aditya Maheshwari, Sundeep Rao, Sankeerth Rao, Nalin Pithwa, Aishwarya Babu, Soumya Narayana, Neha Bhargava, Haseen Rahman, Viswanathan Ramachandran, Arijit Dutta, Tejas Bodas and his better half Arti Yardi. Last but not least, Prasanna Kumar for introducing me to Functional Programming and his better half Nanditha Rao. If I have inadvertently missed anyone, please accept my sincere apologies.

I thank Professor Dey's senior PhD student Brijesh Rai and my fellow PhD students Amitalok Budkuley and Jithin for many useful inputs, suggestions and discussions during the course of my studies.

The Bharti Centre for Communication has provided a great lab environment to pursue my PhD. I would like to thank our lab administrator Vibhavari Bhandare for her kindness and sweet-natured efficiency. I would like to express my gratitude to Professor B. G. Fernandes, the current head of our department and also to Professor Abhay Karandikar, the previous HoD. More generally, the academic and non-academic staff, especially the departmental office staff, of the department of Electrical Engineering and my alma mater IIT Bombay.

For providing me with an outlet for the stress of my PhD and contributing immensely to stress relief, I would like to thank dance instructors Conrad Coelho, Ajaz Shaikh, Deepa Kumar, Deepika Gomes, Darryl Anthony Neves, Vineet Bangera, Sahiba Sawhney, Reesha Dhulap, Medha Shah and Yoga instructor Bharat Gajare.

Last but not the least, I thank my family. My mother Lakshmi Vasudevan, my father K. Vasudevan, my brother Sundar Iyer and my aunt Bhagyam Ramani and uncle G. Ramani for their moral, psychological and financial support. My mother has been the greatest source of inspiration, strength, and support throughout my life and career. I am truly blessed and fortunate to have such an amazing mother and superb support system in my family.

V. Krish

....

Krishnamoorthy Iyer
Roll No. 08407617