Lehrstuhl für Theoretische Informationstechnik
Fakultät für Elektrotechnik und Informationstechnik
Technische Universität München

Secure Communication Over Wiretap Channels with Multiple Receivers and Active Jammers

M. S. S. Ahmed Mansour, M.Sc.

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technische Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs
(Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Wolfgang Kellerer
Prüfer der Dissertation:
1. Prof. Dr.-Ing. Dr. rer. nat. Holger Boche
2. Prof. Dr.-Ing. Rafael Schaefer

Die Dissertation wurde am 21.11.2018 bei der Technische Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 29.01.2020 angenommen.

Abstract

Communication systems nowadays are not only required to deliver a reliable data transmission, but they also need to provide various secrecy requirements. Establishing a secure communication is not an easy task, specially over wireless channels whose open nature allows anyone to have access to the transmitted signals. In order to overcome this exposure problem, high-level cryptographic techniques have been used to encrypt the transmitted information. These techniques are very efficient as long as only limited computational power is available at the eavesdropping nodes. Thus, with the rapid improvement in the fields of number theory and digital design, these techniques are becoming obsolete. In the recent years, physical layer secrecy also known as *information theoretic security* has rose as a prominent alternative technique to achieve secure communication instead of the classical cryptographic techniques.

The basic concepts of information theoretic security were outlined by Shannon in [1]. Afterwards, Wyner formulated the problem of secure communication over the degraded wiretap channel in [2]. Wyner defined an information theoretic quantity known as *weak secrecy* and used it to measure the ignorance of the eavesdropper about the transmitted information. He then proposed the concept of wiretap random codes and showed that they can achieve the secrecy capacity of the degraded wiretap channel under the weak secrecy constraint. Many researchers followed Wyner's line of work and extended his investigation to more complex and real life communication scenarios. However, most of these works suffered from two main limitations: Most of the results established in these works are only valid under the weak secrecy constraint, despite of its inadequacy compared to the recently introduced secrecy measures such as strong secrecy and effective secrecy. This is because the extension of wiretap random codes to complex communication scenarios under the constraints of strong secrecy or effective secrecy is a very challenging task. The second limitation is that the majority of these works investigated wiretap channels that only suffer from passive eavesdropping attacks and totally ignored channels under active jamming attacks.

In the first part of this thesis, we investigate secure communication over multi-user wiretap channels under the *strong secrecy* constraint. We start our investigation by showing that although the multi-user setup is more challenging in general, it exploits an additional dimension for secrecy measures. In particular, multi-user wiretap channels can either be investigated under a conservative secrecy measure known as *joint secrecy*, where the different users do not trust each other; or a more relaxed one known as *individual secrecy* which is based on mutual trust among different users. We then focus our investigation over two classes of multi-user wiretap channels. The first class is the two-receiver wiretap broadcast channel with degraded message sets and message cognition. For this class, we derive a state of the art strong secrecy coding scheme

that combines the concepts of Marton coding, wiretap random encoding and indirect decoding for both the joint and individual secrecy measures. We further showed that our coding scheme establishes the secrecy capacity of less-noisy channels. Our result is the first one to prove the optimality of indirect decoding for a secrecy scenario. The second class that we investigate is the degraded multi-receiver wiretap broadcast channel. For this class, we establish both the joint and individual secrecy capacity under the strong secrecy constraints. For both classes we showed that individual secrecy can utilize the concept of mutual trust to achieve a larger capacity region compared to the joint secrecy measure.

In the second part of this thesis, we consider a communication scenario in which the channel undergoes two simultaneous attacks: passive eavesdropping and active jamming. We start our investigation by pointing out that this scenario is equivalent to problem of secure communication over wiretap channel with imperfect channel state information. We then limit our investigation to the class of arbitrarily varying wiretap channels in which the channel state varies from one channel use to the other due to the manipulation induced by the active jammer. In previous literature, two main coding schemes (deterministic and correlated random) have been used to establish a reliable and secure communication over AVWCs. However, both coding schemes suffer from various drawbacks making them unable to achieve a positive secrecy rate for some scenarios. In order to overcome some of these drawbacks, we offer an alternative coding scheme based on the principle of list decoding. We then derive a full characterization for the list secrecy capacity of arbitrarily varying wiretap channels and establish some interesting results for the continuity and additivity behavior of the capacity function. Finally, we try to bring the two parts of the thesis together by investigating a communication scenario in which public and confidential messages are transmitted over an arbitrarily varying wiretap channel.

Acknowledgments

In the name of Allah, the Gracious, the Merciful who said in the Holy Quran: And when your Lord proclaimed: "If you give thanks, I will grant you increase" [Surah 14, Ayah 7]. The prophet Muhammad –peace and blessings be upon him– has once said: "He who does not thank people, does not thank Allah" [Ahmad, Tirmidhi]. Based on these two quotes which I strongly believe in, I find myself in a deep need to express my thanks and honest gratitude to a lot of people.

Thus, I would like to express my honest and sincere gratitude to my supervisor Prof. Dr.-Ing. Holger Boche who has kindly accepted me as a PhD. student at the chair of Theoretical Information Theory in the Technical University of Munich. Prof. Boche was more than a supervisor to me as I always felt that he was like a father to me. His unfaltering supervision and motivation has helped me to keep going despite all the obstacles I faced during my Ph.D. Also, I would like to thank Prof. Dr.-Ing. Rafael Schaefer for his continuous support and guidance during the bad and the good times. I consider Prof. Schaefer as an older brother and I would like him to know that he has been an important source of inspiration for me. Many thanks go to my colleagues at the chair of Theoretical Information Theory for the fun and entertaining days we spent together.

Next, I would like to thank my father and my mother who have made me the man I am now. No matter how much I do for them, I will not be able to repay their debt. Also, I would like to thank my wife Dr. Mariam Abdelghaffar who despite all the stress and pressure I put on her, she was always there for me as a loving wife and a dear friend. Last but not least, I am grateful to my dear family and friends in Germany and Egypt alike for their emotional support in good times and bad.

Contents

1 Introduction **1**

2 Secure Communication Over Wiretap Channels **7**

2.1 Information Theoretic Security 7

2.1.1 Shannon's Ciphering System 8

2.1.2 Degraded Wiretap Channel 9

2.1.3 General Wiretap Channel with Public and Confidential Services 10

2.1.4 The Dilemma of How to Measure Secrecy 12

2.2 Secrecy Analysis for Wiretap Random Codes 14

2.2.1 Code Concept and Construction 14

2.2.2 Virtual Receiver Technique 15

2.2.3 Variational Distance Measure 17

2.2.4 Information Divergence Approximation 20

2.3 Secrecy Capacity for Wiretap Channels with Shared Key 23

2.3.1 Channel Model and Coding Theorem 23

2.3.2 Achievability Proof 25

2.3.3 Converse Proof 28

2.4 Two-Receiver Wiretap Broadcast Channels 30

2.4.1 Motivation and Channel Model 30

2.4.2 Public-Confidential Degraded Message Sets 32

2.4.3 Secure Broadcasting of Individual Message Sets 34

2.4.4 Superposition Wiretap Random Codes 37

2.4.5 Marton-Coding Wiretap Random Codes 39

2.4.6 Indirect Decoding for Wiretap Random Codes 41

2.4.7 Gaussian Wiretap Broadcast Channels 43

3 Wiretap Broadcast Channels with Degraded Message Sets and Message Cognition **47**

3.1 System Model and Secrecy Criteria 47

3.1.1 Motivation: Secure Bidirectional Relaying 47

3.1.2 A Comprehensive Channel Model 49

3.1.3 Secrecy Criteria: Joint Versus Individual 51

3.1.4 Individual Secrecy in Shannon's Ciphering System 53

3.2 Joint-Strong Secrecy for Wiretap Marton Codes 55

3.2.1 A Universal Achievable Rate Region 55

3.2.2 Coding Scheme: Encoding and Decoding 58

3.2.3 Encoding and Decoding Error Analysis 61

3.2.4 Strong Secrecy Analysis 69

3.2.5 Time Sharing and Rate Splitting 76
3.3 Individual-Strong Secrecy Rate Regions 79
3.3.1 Individual Secrecy Coding Techniques 79
3.3.2 A Universal Individual-Strong Secrecy Coding Scheme 85
3.3.3 Receivers With Statistical Advantage 91
3.4 Secrecy Capacity Regions and Outer Bounds 94
3.4.1 Multi-Letter and Single-Letter Outer Bounds 95
3.4.2 Capacity Regions for Less-Noisy Channels 99
3.4.3 Optimality of Indirect Decoding 105
3.5 Discussion 110

4 Multi-Receiver Wiretap Broadcast Channels 113
4.1 Secrecy in Multi-Receiver WBC 113
4.1.1 Individual Secrecy Without Message Cognition 113
4.1.2 Channel Model and Secrecy Criteria 114
4.1.3 Rate Regions: Joint Vs Individual 118
4.2 Secrecy Capacity Regions for the Degraded Multi-Receiver WBC 121
4.2.1 Joint Secrecy: From Weak to Strong 121
4.2.2 Individual-Strong Secrecy Capacity Region 123
4.2.3 Eavesdropper Degradedness Order 129
4.2.4 Multi-Receiver Gaussian SISO WBC 131
4.2.5 Multi-Receiver Degraded Gaussian MIMO WBC 134
4.3 Achievable Rate Regions for the Two-Receiver DM-WBC 137
4.3.1 Motivation and Main Results 137
4.3.2 Wiretap Marton-Coding with a Superposition Variable 140
4.3.3 Marton Wiretap Coding Under Individual Secrecy 147

5 Wiretap Channels With Active Adversaries 153
5.1 Channel Models With Imperfect CSI 153
5.1.1 Motivation: Active Jamming 154
5.1.2 Compound Wiretap Channels 154
5.1.3 Arbitrary Varying Wiretap Channels 156
5.2 Deterministic and Correlated Random Secrecy Capacity for AVWCs . . 160
5.2.1 Coding Schemes: Basic Definitions 160
5.2.2 Correlated Random Secrecy Capacity: Robustification Technique 164
5.2.3 Coding Techniques for Deterministic Codes 167
5.2.4 Analytical Properties of the Capacity Functions 169
5.3 The Secrecy Capacity of AVWCs Under List Decoding 172
5.3.1 List Decoding for AVWCs: Basic Definitions 173
5.3.2 Coding Theorem and Important Results 175
5.3.3 Direct Coding Approach 180
5.3.4 Elimination of Correlation Approach 184
5.3.5 List Codes With Finite List Size 186
5.3.6 Continuity, Additivity and Super-Activation 191
5.4 Public-Confidential Capacity Regions for AVWCs 195
5.4.1 Motivation and Basic Definitions 196
5.4.2 An Auxiliary Channel Model 198
5.4.3 Coding Theorems: Achievability and Converse 202
5.5 Discussion 205

6 Conclusion and Future Work **209**

A Fundamental Tools in Information Theory **215**
A.1 Entropy, Mutual Information and Divergence 215
A.2 Discrete Memoryless Channels 216
A.3 Channel Comparison and Basic Lemmas 218
A.4 Types and Typical Sequences 221
A.5 Useful Bounding Lemmas 223

List of Figures **227**

Bibliography **229**

Chapter 1

Introduction

No one can deny the impact induced by the advances in the field of communication and information technologies on the life of human beings. It dates back to the year 1870 when Alexander Graham Bell invented the first telephone. Within a few decades, telephone lines have grown to be used by millions of people across the globe. Phones enabled people to connect with each other wherever they are. The technology improved gradually until we reached the first wireless cellular network in 1979. Mobile networks did not only allow verbal communication between people, but it also allowed them to exchange data and multimedia content and all of that is on the move [3]. Parallel to the development of mobile networks, communication systems took a huge leap with the development of the internet of things (IoT), which enabled the interconnection of smart devices over the internet. Humanity nowadays is at the verse of a new evolution in the field of communication which is the **Tactile Internet**. The International Telecommunication Union (ITU) defines the Tactile Internet as an internet network that combines ultra low latency with extremely high availability, reliability and security [4]. It is believed that the Tactile Internet will have a revolutionary effects on different aspects of our lives such as education, healthcare and energy. This is because, It will add a new dimension to the human-to-machine interaction by enabling tactile and haptic sensations.

The previous historical overview shows that it has become very crucial for communication systems to be able to provide a reliable and secure information transmission. In fact, there exists a various security issues in communication networks nowadays such as confidentiality, integrity, authentication, and non-repudiation [5]. Each of these issues considers a certain secrecy requirement for a given communication scenario. The main focus of this thesis is the investigation of the confidentiality of a transmission within a given communication scenario. Thus, whenever the term secure communication is mentioned, it will be in the context of confidentiality, unless stated otherwise. In principal, a secure communication between two nodes in a network must fulfill two requirements simultaneously:

1. Reliability: This requirement assures that the transmitted messages are guaranteed to reach their destination complete and uncorrupted [6].
2. Secrecy: This requirement assures that malicious nodes are not able to interpret the transmitted information [7].

These two requirements simply imply that only legitimate receiving nodes can successfully decode the messages transmitted over the network.

Communication systems are built based on a layered approach, where each layer is used to deliver a certain task. In current technologies, secure communication is established by separating the reliability and secrecy requirements. In particular, reliability is integrated on the lower layers, while secrecy is considered as a part of the upper layers [8]. This is because most of the techniques used to establish secrecy in current communication systems are based on the classical computational cryptographic algorithms for Encryption and Decryption. This implies that secrecy in these techniques completely ignore the physical characteristics of the channel. Moreover, these techniques are only secure, if the eavesdropper is of limited computational complexity or has limited resources. These two observation have raised a very important question in the last decade about the possibility to deploy both the reliability and secrecy requirements on the physical layer of a communication system [9]. This question is related to another question that was raised in the seventies about the capability of achieving secrecy by only exploiting the physical characteristics of the channel such as the randomness of the channel.

The answer to the previous question can be interpreted from the concept of **physical layer security** introduced by Shannon in [1]. Shannon showed that secure communication can be achieved by using a secret key shared between the transmitter and the receiver, as long as the entropy of this key is greater than or equal to the entropy of the message to be transmitted. Although Shannon's work inspired most of the classical cryptographic techniques known nowadays, it has also lit the first spark to the usage of the physical properties of the channel to achieve secrecy. It was then Wyner in [2], who laid the actual foundations of physical layer security. Wyner introduced the model of the degraded wiretap channel that consists of three nodes: a transmitter, a legitimate receiver and an eavesdropper, where the eavesdropper observes a noisy version of the legitimate receiver observation. Wyner showed that for such channel, secure transmission is still achievable in the absence of a secret key by exploiting the noisiness of the channel. He also introduced the term **secrecy capacity** and defined it as the maximum rate at which information can be sent from the transmitter to the legitimate receiver, while keeping it secret from the eavesdropper.

Wyner's work was revolutionary because he was able to show that the secrecy requirement is guaranteed regardless of the computational capabilities of the eavesdropper. He also built on Shannon's work and formulated an information theoretic quantity known as **weak secrecy** that can be used to measure the level of ignorance of the eavesdropper about the confidential information transmitted over the network. Through the years, researchers aimed to extend Wyner's work by considering more complex and practical communication scenarios under various secrecy requirements. In this thesis, we investigate two popular and very interesting extensions for Wyner's work. In particular, we investigate the problem of secure communication over two classes of wiretap channels: The first is the wiretap channel with multiple legitimate receivers, while the second is the wiretap channel with active jammers.

In the seventies and eighties most of the work that built upon Wyner's work only considered a secure communication scenario over a channel with two receiving nodes only (a legitimate receiver and an eavesdropper), cf [10, 11]. During this period of

time, such limitation was understandable, specially because the concept of wiretap channels was relatively new. However, with the rapid growth in communication systems and in particular wireless systems that contains a lot of potential users, it has become more crucial to investigate secure communication scenarios that involves multiple receiving nodes. In order to address this issue, two extensions for the wiretap channel were introduced in [12]. The first extension considered a wiretap channel with two legitimate receivers and an eavesdropper, while the second extension considered a wiretap channel with one legitimate receiver and two eavesdroppers. In this thesis, our investigation focuses on the first extension, which is motivated by the the following reasons: This model is of high practical significance because it simulates the problem of secure communication over broadcast channels. This relation to broadcast channels allows us to investigate real life communication scenarios where public and confidential messages are transmitted together. One of the most interesting features of this model is that it can be investigated with respect to a relatively new secrecy criterion known as **individual secrecy**.

The notation of individual secrecy originates from the concept of mutual trust between the different legitimate receivers. This differs from the conservative notation of secrecy known as **joint secrecy**, which was used in most works that addressed secure communication over broadcast channels. The differences between these two secrecy criteria motivates us to examine the implications of using each criterion. More precisely, we study the distinctive coding schemes that can be used for each criterion and the effect of both secrecy criteria on the secrecy capacity of the investigated scenarios. In doing so, we develop some new coding schemes that can adapt to different secrecy constraints. Moreover, we develop our investigation under the strict secrecy measure known as **strong secrecy** instead of the weak secrecy measure which have been used in most of previous literature. Our investigation allows us to achieve a better understanding to the problem of secure communication over multi-receiver wiretap channels. It also enables us to know the limits and the capabilities of the different coding techniques.

Beside the multi-receiver wiretap channel, we consider another extension to Wyner's work by examining secure communication over a class of channels known as the **arbitrary varying wiretap channel** (AVWC). This line of work is motivated by the fact that communication systems can suffer from two basic classes of attacks: passive attacks and active attacks. In an active attack, a malicious node aims to disrupts the communication by manipulating the physical characteristics of the channel. Alternatively, a passive attack simulates a scenario in which a malicious node only eavesdrop on the transmitted information. Although the model of AVWCs was introduced in [13] to simulate a wiretap channel with an imperfect channel state information, it was shown in [14] that AVWCs can be used to model communication scenarios over wiretap channels with active jammers.

Secure communication over AVWCs has captured a lot of attention recently because it simulates a very general and real life communication scenario. However, it was shown that constructing a secure coding scheme for this class of channels is very challenging. Different form the normal wiretap channel, the coding difficulties for AVWCs are due to the reliability requirement and not the secrecy one. This is because the impact of

active jamming on the communication link between the transmitter and the legitimate receiver can be catastrophic making the establishment of reliable communication using classical coding schemes impossible. In this thesis, we try to solve this issue by providing an alternative coding scheme that can overcome the major drawbacks of the conventional coding schemes for AVWCs.

The investigation of secure communication over wiretap channels is a very challenging task. In this thesis, we try to address some of these difficulties and discuss various techniques that can be used to solve them. We believe that the results established in this thesis have solved some of the major issues of the investigated scenarios. We also hope that this work can allow a general reader to achieve a better understanding to the problem of secure communication over wiretap channels.

Text Organization

This thesis is structured as follows: in the second chapter, a brief review for the problem of secure communication over wiretap channels is provided. In Chapter 3, we focus our attention to a class of wiretap channels with two legitimate receivers known as the two-receiver wiretap channel with degraded message sets and message cognition. For this class of channels, we investigate a secure communication scenario under the joint and individual secrecy criteria. For both criteria, we derive a general achievable rate region that combines various state of the art coding techniques. In Chapter 4, we drop the message cognition assumption and extend our investigation to a wiretap channel with k legitimate receivers. We establish the joint and individual secrecy capacity for the class of degraded channels. In Chapter 5, we consider the problem of secure communication over AVWCs. We present a full characterization for the secrecy capacity of AVWCs under list decoding. Our result implies that list decoding can overcome the drawbacks of other conventional coding schemes. Finally, we provide a conclusion to summarize the most important results of the thesis, then give a short outlook on possible future work.

Notation

In this thesis, random variables are denoted by capital letters and their realizations by the corresponding lower case letters, while bold letters are used to denote matrices. Additionally, calligraphic letters are used to denote sets, while fraktur letters are used to denote a set of sets.

$\mathbb{N}$	set of natural numbers
$[\![a, b]\!]$	represent the set of natural numbers between a and b
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of non-negative real numbers
$\mathbb{R}^n$	set of length-n vectors over $\mathbb{R}$, where $n \in \mathbb{N}$
X^n	denotes the sequence of random variables $(\mathrm{X}_1, \ldots, \mathrm{X}^n)$
X_i	denotes the ith random variable in the sequence X^n, for $i \in [\![1, n]\!]$
$\tilde{\mathrm{X}}^i$	denotes the sequence of random variables $(\mathrm{X}_i, \ldots, \mathrm{X}^n)$
P_X	denotes a probability distribution over the random variable X

$\mathcal{T}_\epsilon^n(\mathrm{P_X})$	denotes the ϵ-strongly typical set of length n with respect to $\mathrm{P_X}$
$\mathbb{P}[\theta]$	denotes the probability of the event θ
$\mathrm{P_X}(x)$	denotes the probability $\mathbb{P}[\mathrm{X} = x]$
$\mathcal{P}(\mathcal{X})$	denotes the set of all possible probability distributions on X
$\mathcal{P}_0^n(\mathcal{X})$	denotes the set of types for all sequences $x^n \in \mathcal{X}^n$
$\mathbb{E}[\theta]$	denotes the expectation of the event θ
$\mathbb{H}[\mathrm{X}]$	denotes the Shannon's entropy of the random variable X
$\mathbb{I}[\mathrm{X};\mathrm{Y}]$	denotes the mutual information between X and Y
$\mathrm{X} - \mathrm{Y} - \mathrm{Z}$	denotes a Markov chain over X, Y and Z in this order
$\mathbf{X} \succeq \mathbf{0}$	implies that $\mathbf{X}$ is a positive semi-definite matrix
$\mathbf{X} \succ \mathbf{0}$	implies that $\mathbf{X}$ is a positive definite matrix
$\lvert\mathbf{X}\rvert$	denotes the determinant of the matrix $\mathbf{X}$
$\lvert\mathcal{X}\rvert$	denotes the cardinality of the set $\mathcal{X}$
$D_{\mathrm{KL}}(\mathrm{P_X}\lVert\mathrm{Q_X})$	denotes the Kullback-Leibler divergence between $\mathrm{P_X}$ and $\mathrm{Q_X}$
$\lVert\mathrm{P_X} - \mathrm{Q_X}\rVert$	denotes the total variation distance between $\mathrm{P_X}$ and $\mathrm{Q_X}$

Abbreviations

Through out this thesis, the following acronyms are used.

$\mathbb{LHS}$	Left Hand Side
$\mathbb{RHS}$	Right Hand Side
DM	Discrete Memoryless
DMC	Discrete Memoryless Channel
DWC	Degraded Wiretap Channel
WC	Wiretap Channel
KWC	Wiretap Channel with Key
BC	Broadcast Channel
WBC	Wiretap Broadcast Channel
DMS	Discrete Memoryless Source
AEP	Asymptotic Equipartition Property
SISO	Single Input Single Output
MIMO	Multiple Input Multiple Output
CSI	Channel State Information
MMSE	Minimum Mean Square Error
MAWC	Multiple Access Wiretap Channel
WRC	Wiretap Random Coding
SKC	Secret Key Encoding
GWBC	Gaussian Wiretap Broadcast Channel
AVC	Arbitrarily Varying Channel
AVWC	Arbitrarily Varying Wiretap Channel
CAVWC	Compound Arbitrarily Varying Wiretap Channel
CQC	Classical Quantum Channel

Chapter 2

Secure Communication Over Wiretap Channels

This chapter reviews the problem of secure communication over wiretap channels where perfect channel state information (CSI) is available at all channel nodes. First, the concept of information theoretic security is introduced and the problem of measuring secrecy is addressed. Next, the mathematical techniques used to analyze coding schemes for wiretap channels are briefly discussed. Finally, an overview over the results established for secure communication over the two-receiver wiretap broadcast channel is presented.

2.1 Information Theoretic Security

Communication systems nowadays are required to provide both a reliable and secure data transmission. The problem of reliable communication over the different layers of the networking architecture have been extensively studied over the last few decades. On the other hand, the problem of secure communication was mainly considered on the higher layers. That is why most of the techniques used to establish secure communication in the current communication systems are based on classical computational cryptographic algorithms. Although these techniques have been very useful in the past, they are becoming more insecure recently [8]. This because most of these techniques are based on the assumption of limited computational power at the eavesdropping nodes. Thus, with the recent advances in number theory and the improvements in the digital design field, this assumption is no longer valid.

Information theoretic security also known as *physical layer secrecy* has captured a lot of attention over the last few years [5]. In particular, it seems to be a a good replacement for the conventional cryptographic techniques because it does not impose any assumptions on the computational power of the eavesdroppers. One of the main features of information theoretic security is that it derives very precise expressions for the amount of information leaked to the eavesdropper in terms of information theoretic quantities and as a function of the communication channel [9]. This feature gives the physical layer security a huge advantage over the conventional cryptographic techniques which do not have a precise metric to measure their strengths where a cryptographic protocol is said to be secure based on its ability to overcome a group of attacks.

In this section, a brief overview about the field of information theoretic security is presented. It starts by introducing Shannon's original work in [1] on secure communication. It then highlights the model of the degraded wiretap channel introduced by Wyner in [2] and the more general form investigated by Csiszár and Körner in [10]. Finally, the dilemma of evaluating the secrecy of a communication in terms of information theoretic measures is addressed.

2.1.1 Shannon's Ciphering System

Shannon was the first one to investigate the problem of secure communication from an information theoretic perspective in [1]. He considered a communication scenario that consists of a transmitter (Alice), a legitimate receiver (Bob) and an eavesdropper (Eve), where Alice communicates with Bob over a noiseless channel which is overheard by Eve as illustrated in Fig. 2.1. Now, given a confidential message $m \in \mathcal{M}$, Alice should be able to inform Bob about m, while keeping Eve completely ignorant. This ignorance reflects the secrecy level of the communication link between Alice and Bob.

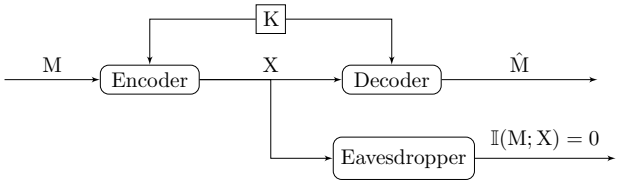


Figure 2.1: Shannon's ciphering system

Shannon used the conditional entropy $\mathbb{H}(\mathrm{M}|\mathrm{X})$ to evaluate the secrecy of the transmission, where M is a random variable distributed over the confidential message set $\mathcal{M}$, while X is a random variable that represents the information transmitted over the channel and observed by the eavesdropper. The term $\mathbb{H}(\mathrm{M}|\mathrm{X})$ is currently known as the eavesdropper's *equivocation* and has been widely used as a reliable secrecy measure. This is because the equivocation represents the uncertainty of Eve about the confidential message M after intercepting the transmitted signal X [15].

Shannon showed that secure communication in terms of prefect secrecy – $\mathbb{H}(\mathrm{M}|\mathrm{X}) = \mathbb{H}(\mathrm{M})$ – can only be achieved using a secret key shared between the transmitter and the legitimate receiver, while keeping the key hidden from the eavesdropper. Shannon's coding scheme is based on the concept of one time pad as follows: Given a confidential message $m \in \mathcal{M}$ and a shared secret key $k \in \mathcal{K}$, Alice uses an encoder $E : \mathcal{M} \times \mathcal{K} \to \mathcal{X}$ to produce the codeword $x = m \otimes k$. At the legitimate receiver, Bob uses x and k along with the decoder $\varphi : \mathcal{X} \times \mathcal{K} \to \mathcal{M}$ to extract the confidential message $m = x \otimes k$. Now, even if Eve knows the structure of the encoder E and the decoder φ, it can not extract any information about m using only x .

Although Shannon's investigation showed that secrecy can be achieved with low complexity, it suffers from some limitations [16]. One of the major issues is that Shannon showed that in order to achieve perfect secrecy, the entropy of the shared secret key must be greater than or equal to the entropy of the confidential message as follows:

$$\mathbb{H}(\mathrm{K}) \geq \mathbb{H}(\mathrm{M}). \tag{2.1}$$

This implies that in order to establish secure communication based on Shannon's model, the transmitter must generate and store secret keys whose length is equivalent to the length of the confidential messages. It also indicates that the transmitter needs to find a secure channel over which it can share the key with the legitimate receivers. Not only that but each of the generated keys can only be used once. Another limitation is that Shannon constructed his model for a noiseless channel between the transmitter and the legitimate receiver while in reality noise is always there. However, this assumption have become acceptable over the years, due to the development of powerful algorithms that can establish a reliable communication over noisy channels with an arbitrary small decoding error probability [17, 18].

2.1.2 Degraded Wiretap Channel

Roughly speaking, secure communication involves two main tasks: reliability and secrecy. Reliability implies that the legitimate receivers are capable of decoding the transmitted message without errors. While secrecy assures that no one aside from the legitimate receivers can acquire any information about the transmitted message by eavesdropping on the communication link. Shannon separated these two tasks in his model as he assumed a noiseless channel where both the legitimate receiver and the eavesdropper observe the exact signal sent by the transmitter. This assumption deprived Shannon from using the noisiness of the channel as a supporting factor to induce secrecy and was the main reason that Shannon argued that without a shared secret key, secure communication can not be established [19].

In order to understand how the presence of noise can be advantageous for the establishment of secure communication, Wyner introduced the model of the degraded wiretap channel in [2] which is illustrated in Fig. 2.2. Differently from Shannon, Wyner considered a discrete memoryless channel (DMC) in which the transmitter encodes the confidential message $m \in \mathcal{M}$ in to a codeword $x^n \in \mathcal{X}^n$ of length n, which is then sent to the legitimate receiver over a noisy channel $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ such that the receiver observes the noisy codeword y^n. Additionally, the eavesdropper observes the codeword z^n through a DMC $V : \mathcal{Y} \to \mathcal{P}(\mathcal{Z})$. This implies that the eavesdropper's observation is a degraded (noisier) version of the legitimate receiver observation [20].

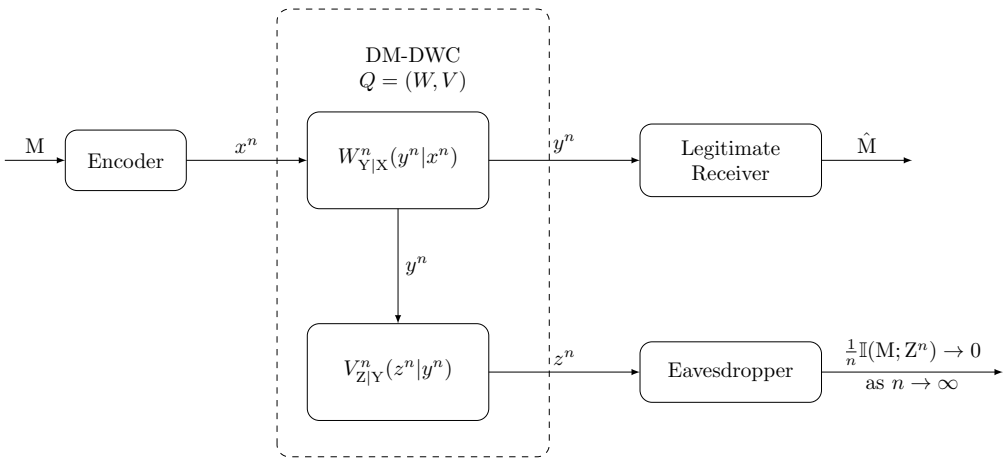


Figure 2.2: Discrete memoryless degraded wiretap channel

Moreover, Wyner used a new secrecy constraint known as the weak secrecy criterion rather than using the perfect secrecy constraint introduced by Shannon. Instead of having the eavesdropper's equivocation to be exactly equal to the entropy of the confidential message, Wyner introduced a new quantity called the eavesdropper's *equivocation rate* given by $\frac{1}{n}\mathbb{H}(\mathrm{M}|\mathrm{Z}^n)$ and requested this quantity to be arbitrarily close to the entropy rate of the confidential message $\frac{1}{n}\mathbb{H}(\mathrm{M})$ for sufficiently large codeword length n. Wyner's target was to investigate the existence of a coding scheme that satisfies the weak secrecy constraint and assures at the same time that the probability of error at the legitimate receiver is arbitrary small for sufficiently large n.

Wyner was not only able to prove the existence of such coding scheme, but he also managed to prove that the codes that apply such coding scheme can achieve the maximum transmission rate possible under the defined reliability and secrecy constraints. Wyner called this maximum rate the secrecy capacity and presented the following coding theorem:

Theorem 2.1. *[2] The secrecy capacity of the discrete memoryless degraded wiretap channel $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V : \mathcal{Y} \to \mathcal{P}(\mathcal{Z})$ is given by the following expression:*

$$C(W, V) = \max_{\mathrm{P_X}} [\mathbb{I}(\mathrm{X};\mathrm{Y}) - \mathbb{I}(\mathrm{X};\mathrm{Z})]. \tag{2.2}$$

where the maximum is taken over all possible input distributions $\mathrm{P_X}$.

The previous theorem confirms Shannon's argument that secure communication cannot be achieved over noiseless channels without secret keys. This is because if $\mathrm{Y} = \mathrm{Z}$ which implies that both the legitimate receiver and the eavesdropper obtain the same observation, then the secrecy capacity vanishes. Although Theorem 2.1 was established by Wyner for degraded wiretap channels, it was shown in [10] that it holds for less noisy and more capable wiretap channels [21] as well. The achievability proof of the previous coding theorem is based on the class of codes introduced by Wyner known as **wiretap random codes**. In Section 2.2, the construction and the secrecy analysis of such codes are introduced. Furthermore, a weak converse based on Fano's inequality [22] is derived for a more general case in Section 2.3.

2.1.3 General Wiretap Channel with Public and Confidential Services

Wyner's investigation of the degraded wiretap channel raised a lot of questions regarding secure communication over noisy channels. Wyner restricted his investigation to the scenario where the eavesdropper's observation is a degraded version of the legitimate receiver's observation. Thus, it was not clear whether secure communication is still achievable over a general wiretap channel where the legitimate receiver does not have a statistical advantage over the eavesdropper. Moreover, Wyner dealt with the eavesdropper as an intruder to the system that should be kept completely ignorant about the transmitted information. However, in some communication scenarios, the eavesdropper is a legitimate part of the system that should have access to some of the transmitted information while keeping him ignorant about other information [23]. These two problems encouraged Csiszár and Körner to investigate the problem of public and confidential communication over the general wiretap channel in [10].

Csiszár and Körner considered a DMC (W, V) given by $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$ that consists of one transmitter and two receivers as illustrated in Fig. 2.3. The transmitter aims to send two messages simultaneously: a common public message $m_p \in \mathcal{M}_p$ and a confidential message $m_c \in \mathcal{M}_c$ by encoding them into the codeword x^n. The first receiver observes the noisy codeword y^n and use it to decode both the public and confidential messages. The other receiver observes the noisy codeword z^n and use it only to decode the public message. The coding scheme should simultaneously assure that the second receiver can not infer any information about the confidential message.

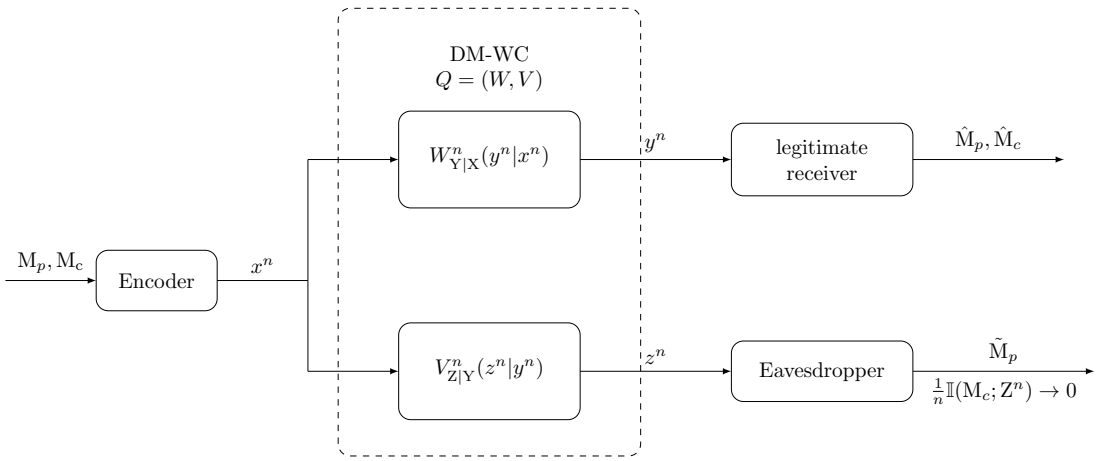


Figure 2.3: Wiretap channel with public and confidential messages

Csiszár and Körner managed to show that for sufficiently large n, there exists a coding scheme that satisfies the weak secrecy constraint, i.e. $\frac{1}{n}\mathbb{H}(\mathrm{M}_c|\mathrm{Z}^n) \approx \frac{1}{n}\mathbb{H}(\mathrm{M}_c)$, while assuring an arbitrary small decoding error probability for $(\mathrm{M}_p, \mathrm{M}_c)$ at the first receiver and for M_p at the second receiver. They also showed that the constructed coding scheme establishes the public-secrecy capacity of the channel and presented the following coding theorem:

Theorem 2.2. *[10] The public-secrecy capacity region of the discrete memoryless wiretap channel (W, V) is given by the union over the set of all rate pairs $(R_p, R_c) \in \mathbb{R}^2_+$ that satisfy:*

$$R_p \leq \min[\mathbb{I}(\mathrm{U}; \mathrm{Y}), \mathbb{I}(\mathrm{U}; \mathrm{Z})] \tag{2.3}$$

$$R_c \leq \mathbb{I}(\mathrm{V}; \mathrm{Y}|\mathrm{U}) - \mathbb{I}(\mathrm{V}; \mathrm{Z}|\mathrm{U}), \tag{2.4}$$

where the union is taken over all probability distributions $\mathrm{P}_{\mathrm{UVX}}$ on the random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X})$, such that $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}, \mathrm{Z})$ forms a Markov chain. Further it suffices to have $|\mathcal{U}| \leq |\mathcal{X}| + 3$ and $|\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3$.

The previous result implies that secure communication over the general wiretap channel is possible as long as there exists at least one joint distribution $\mathrm{P}_{\mathrm{UVX}}$ such that $\mathbb{I}(\mathrm{V}; \mathrm{Y}|\mathrm{U}) > \mathbb{I}(\mathrm{V}; \mathrm{Z}|\mathrm{U})$. If such distribution does not exist, then the eavesdropper's channel is less noisy than the legitimate receiver's channel which consequently implies that the secrecy capacity is zero [16]. If the public message is ignored, the following corollary follows:

Corollary 2.1. *[5] The secrecy capacity of the discrete memoryless wiretap channel (W, V) is given by the following expression:*

$$C(W, V) = \max_{\mathrm{P_{UX}}} [\mathbb{I}(\mathrm{U};\mathrm{Y}) - \mathbb{I}(\mathrm{U};\mathrm{Z})], \tag{2.5}$$

where the maximum is taken over all possible joint distributions $\mathrm{P_{UX}}$ *and it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

It is important to point out that the secrecy capacity of the general wiretap channel is defined in terms of the auxiliary random variable U originated from the prefix channel $Q : \mathcal{U} \to \mathcal{X}$ and not in terms of X directly. This is not the case for the degraded wiretap channel where the secrecy capacity is defined in terms of the random variable that represents the channel input X as shown in Theorem 2.1.

2.1.4 The Dilemma of How to Measure Secrecy

One of the main issues in secure communication is how to evaluate the ignorance of the eavesdropper about a transmitted confidential message [24]. Some cryptographic techniques evaluate the secrecy of a transmission by the ability of the eavesdropper to decode the confidential message. Thus as long as the decoding error probability of the confidential message at the eavesdropper is bounded away from zero then the communication is secure [25]. However, this evaluation depends on the computational power and the decoding capabilities of the eavesdropper [26]. It also implies that although the eavesdropper is not able to decode the whole message, he might be able to extract some information about the transmission.

In [1] Shannon suggested using an information theoretic measure known as the eavesdropper's equivocation and is given by $\mathbb{H}(\mathrm{M}|\mathrm{Z}^n)$. This conditional entropy is an estimate for the eavesdropper's uncertainty about the confidential message M given it's observation Z^n. Shannon requested a very strict secrecy condition known as perfect secrecy where the eavesdropper's equivocation must be equivalent to the entropy of the confidential message. Over the years, the equivocation has been replaced by another information theoretic measure known as the information leakage. This measure is given by the mutual information between the confidential message and the eavesdropper observation. Thus, the perfect secrecy constraint can be formulated as follows:

$$\mathbb{I}(\mathrm{M};\mathrm{Z}^n) = 0. \tag{2.6}$$

The previous condition implies that eavesdropper's observation Z^n is statistically independent of the confidential message M. This means that the perfect secrecy constraint in (2.6) forces the decoding error probability at the eavesdropper to be close to one, regardless of its decoding capabilities [27].

Although perfect secrecy appears to be a convenient measure to evaluate the secrecy of a transmission, it has been really difficult to find coding schemes that fulfill it. In [2], Wyner suggested the usage of a more practical secrecy constraint known as the weak secrecy criterion. This criterion requires that the rate of information leaked to the eavesdropper about the confidential message vanishes as n approaches infinity. This requirement can be formulated as follows:

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M};\mathrm{Z}^n) = 0. \tag{2.7}$$

The weak secrecy notation has become a very popular measure for evaluating secrecy, specially because one can easily follow Wyner's technique introduced in [2] to construct wiretap random codes that fulfill the weak secrecy constraint. However, one can argue that the weak secrecy criterion is an untrustworthy measure for secure communication. This is because the weak secrecy constraint in (2.7) does not necessarily imply that the information leakage decays with n, but it only assures that the information leakage grows at most sub-linearly with n. This means that under the strong secrecy criterion, the eavesdropper might be able to extract some information about the confidential message as highlighted in [16, Example 3.3].

The previous discussion advocates the need a secrecy measure which can overcome the practical issues of perfect secrecy and at the same time is more strict than the weak secrecy criterion. From a mathematical point of view, this secrecy measure should have an asymptotic statistical independence between the confidential message M and the eavesdropper's observations Z^n. This leads to the strong secrecy criterion which is formulated as follows:

$$\lim_{n\to\infty} \mathbb{I}(\mathrm{M};\mathrm{Z}^n) = 0. \tag{2.8}$$

Differently from weak secrecy, the strong secrecy constraint strictly implies that the amount of information about the confidential message M leaked to the eavesdropper decays as n goes to infinity. In some literature, the strong secrecy criterion is defined using various distance measures between the joint distribution $\mathrm{P}_{\mathrm{MZ}^n}$ and the product of its marginals $\mathrm{P}_{\mathrm{M}} \cdot \mathrm{P}_{\mathrm{Z}^n}$. Among the most used distance measure are the Kullback-Leibler divergence as in [28] and the total variation distance as in [29].

Similar to the perfect secrecy, the strong and weak secrecy constraints are not only information theoretic measures but they also have an operational meaning. This operational meaning is related to the decoding error probability at the eavesdropper. Although both measures guarantee a high probability of decoding error at the eavesdropper, they differ in the speed at which this probability converges to one. Using Fano's inequality, one can show that the rate of convergence of the error probability under the weak secrecy criterion is $\mathcal{O}(1)$. On the other hand, it has been shown in [30], that the strong secrecy criterion provides an exponential rate of convergence.

Recently a new secrecy measure known as effective secrecy was introduced in [31]. Effective secrecy does not only investigate the ability of the eavesdropper to extract information about the confidential message, but it also consider an additional operational meaning for secrecy called *stealth*. Stealth determines the ability of the eavesdropper to detect the presence of meaningful communication. The mathematical formulation of effective secrecy is expressed using the Kullback-Leibler divergence as follows:

$$\lim_{n\to\infty} D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{MZ}^n} \| \mathrm{P}_{\mathrm{M}} \mathrm{Q}_{\mathrm{Z}}^n) = 0, \tag{2.9}$$

where $\mathrm{P}_{\mathrm{MZ}^n}$ is the joint distribution on the random variables M and Z^n, P_{M} is the marginal distribution of $\mathrm{P}_{\mathrm{MZ}^n}$ on M, while $\mathrm{Q}_{\mathrm{Z}}^n$ is the distribution expected to be observed by the eavesdropper when the transmitter is sending useless information. In order to understand the implication of the effective secrecy requirement in (2.9), we expand the divergence as follows:

$$D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{MZ}^n} \| \mathrm{P}_{\mathrm{M}} \mathrm{Q}_{\mathrm{Z}}^n) = \mathbb{I}(\mathrm{M};\mathrm{Z}^n) + D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{Z}^n} \| \mathrm{Q}_{\mathrm{Z}}^n). \tag{2.10}$$

The previous expansion captures the essence of effective secrecy as follows: In addition to fulfilling the strong secrecy obligation, effective secrecy enforce an additional constraint due to the second term in (2.10). This constraint provides what is called stealth communication as it reflects hiding the presence of useful communication from the eavesdropper.

2.2 Secrecy Analysis for Wiretap Random Codes

In this section, the concept of wiretap random codes is introduced by highlighting its basic definition and a simple way to construct this class of codes for the general discrete memoryless wiretap channel (DMWC). In addition, the main techniques used to carry out the secrecy analysis for such codes under the different secrecy constraints are investigated.

2.2.1 Code Concept and Construction

Consider a DM-WC that consists of a transmitter, a legitimate receiver and an eavesdropper a shown in Fig. 2.3, where only a confidential message M is transmitted over the channel instead of the public-confidential pair $(\mathrm{M}_p, \mathrm{M}_c)$. The channel is defined by the stochastic matrix $Q : \mathcal{X} \to \mathcal{P}(\mathcal{Y} \times \mathcal{Z})$, where $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are finite discrete input and output alphabets at the transmitter, the legitimate receiver and the eavesdropper respectively. It can also be defined in terms of two independent DMCs: the legitimate channel $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and the eavesdropper channel $V : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$. For a block code of length n, an input sequence $x^n \in \mathcal{X}^n$, and output sequences $y^n \in \mathcal{Y}^n$ and $z^n \in \mathcal{Z}^n$, we have

$$W^n_{\mathrm{Y}|\mathrm{X}}(y^n|x^n) = \prod_{i=1}^{n} W_{\mathrm{Y}|\mathrm{X}}(y_i|x_i) \qquad \text{and} \qquad V^n_{\mathrm{Z}|\mathrm{X}}(z^n|x^n) = \prod_{i=1}^{n} V_{\mathrm{Z}|\mathrm{X}}(z_i|x_i) \tag{2.11}$$

Definition 2.1. A $(2^{nR}, n)$ wiretap random code $\mathcal{C}$ for the general DM-WC $Q = (W, V)$ consists of: a confidential message set $\mathcal{M} = [\![1, 2^{nR}]\!]$, a stochastic encoder at the transmitter

$$E : \mathcal{M} \to \mathcal{P}(\mathcal{X}^n) \tag{2.12}$$

which maps a confidential message $m \in \mathcal{M}$ to a codeword $x^n(m) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m)$, and a deterministic decoder at the legitimate receiver

$$\varphi : \mathcal{Y}^n \to \mathcal{M} \cup \{?\} \tag{2.13}$$

that transforms the channel observation at the legitimate receiver $y^n \in \mathcal{Y}^n$ to a certain confidential message $\hat{m} \in \mathcal{M}$ or an error message.

We assume that the code $\mathcal{C}$ is known to the transmitter, the legitimate receiver and the eavesdropper. It was shown that deterministic encoder in which each confidential message $m \in \mathcal{M}$ is mapped to a fixed codeword $x^n \in \mathcal{X}^n$ is usually insufficient to establish secure communication [32]. This is not the case for decoding where a stochastic decoder can only reduce the rate of reliable communication, while having no effect on secrecy [16, Remark 3.3].

Definition 2.1 implies that the main difference between wiretap random codes and classical deterministic codes is the encoding process. This because Eq. (2.12) implies that for every confidential message m, there exists a group of valid codewords, such that when m is to be transmitted, the encoder selects one of these codewords based on the conditional distribution $E_{\mathrm{X}^n|\mathrm{M}=m}$ and transmits it. In order to construct such encoder, Wyner suggested using a normal deterministic encoder a long with a randomization message set $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$ as follows: Given a confidential message $m \in \mathcal{M}$, the encoder chooses a randomization message $m_r \in \mathcal{M}_r$ uniformly at random then uses the deterministic encoder $E_d : \mathcal{M} \times \mathcal{M}_r \to \mathcal{X}^n$ to produce the codeword $x^n(m, m_r)$ and transmits it.

The previous encoding technique leads to a slight modification to the decoding function in (2.13). Instead of directly mapping the channel observation y^n to the corresponding confidential message, the legitimate receiver uses a decoder $\varphi : \mathcal{Y}^n \to \mathcal{M} \times \mathcal{M}_r$ to decode both the confidential the randomization messages. Based on the channel coding theorem, the described coding scheme is successful with high probability as long as

$$R + R_r \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}). \tag{2.14}$$

The previous equation implies that instead of using the full rate of the channel between the transmitter and the legitimate receiver to transmit only the confidential message, part of this rate is used to transmit the randomization message. This means that the maximum rate available for the confidential message is $\mathbb{I}(\mathrm{X}; \mathrm{Y}) - R_r$.

Wyner showed that R_r needs to be roughly $\mathbb{I}(\mathrm{X}; \mathrm{Z})$ which is the full rate of the channel between the transmitter and the eavesdropper cf. Theorem 2.1. An intuitive explanation for the value of R_r is based on the following arguments:

- R_r is the rate of the randomization message which was introduced to realize a stochastic encoder E in (2.12) by utilizing a deterministic encoder E_d.
- The aim of the stochastic encoder is to confuse the eavesdropper and make sure that it cannot extract any information about the confidential message by assigning different codewords for the same confidential message.
- A possible way to induce this confusion is to jam all the resources available at the eavesdropper – the full rate of its channel – by the randomization message. Thus,

$$R_r \approx \mathbb{I}(\mathrm{X}; \mathrm{Z}). \tag{2.15}$$

Although the previous explanation seems to be logical, it is not enough to establish a coding theorem. It is also not clear which secrecy measure is satisfied by fulfilling the constraint in (2.15). That is why, a step by step mathematical analysis that starts with a certain secrecy measure and ends up with a constraint on the randomization rate R_r is always needed. This analysis is usually known as the secrecy analysis of the coding scheme.

2.2.2 Virtual Receiver Technique

The virtual receiver method is one of the most famous techniques used to carry out the secrecy analysis of wiretap random codes under the weak secrecy criterion. Although this technique is very popular, one can not identify the first time it was introduced.

Nevertheless, one can show that the secrecy analysis carried out by Wyner in [2], Massey in [33], in addition to Csiszár and Körner in [10] might all be interpreted as an application of the virtual receiver technique.

The main idea of this technique is to assume the existence of a virtual receiver that observes the same channel output of the eavesdropper Z^n while having access to the confidential message M through an error free link. With M as a side information and given Z^n, this virtual receiver aims to decode the randomization message M_r reliably. Afterwords, a series of transformation is applied to the weak secrecy constraint to establish a relation between it and the decoding error probability of M_r at the virtual receiver. In order to present the exact steps of the secrecy analysis, it is important to introduce the underlying coding scheme:

- **Codebook $\mathcal{C}$:** Fix an input distribution $\mathrm{P_X}$ and use it to generate the codewords $x^n(m, m_r)$ for $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$ by generating the symbols $x_i(m, m_r)$ for $i \in [\![1, n]\!]$ independently according to $\mathrm{P_X}$. It is important to point out that the constructed codebook $\mathcal{C}$ is just a possible realization of the random variable C that represents all possible codebooks.
- **Encoding:** Given a confidential message $m \in \mathcal{M}$, the encoder chooses a randomization message $m_r \in \mathcal{M}_r$ uniformly at random then transmits the corresponding codeword $x^n(m, m_r)$.
- **Legitimate Receiver Decoding:** Given the received sequence y^n, it founds the unique message pair $(\hat{m}, \hat{m}_r) \in \mathcal{M} \times \mathcal{M}_r$ such that $(x^n(\hat{m}, \hat{m}_r), y^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_X} W_{\mathrm{Y|X}})$.
- **Virtual Receiver Decoding:** Given the received sequence z^n and the transmitted confidential message m, it founds the unique message $\tilde{m}_r \in \mathcal{M}_r$ such that $(x^n(m, \tilde{m}_r), z^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_X} V_{\mathrm{Z|X}})$.
- **Reliability Analysis:** For a codebook $\mathcal{C}$, one can define the average decoding error probability as follows:

$$\bar{\mathbf{P}}_e(\mathcal{C}) \triangleq \mathbb{P}[(\mathrm{M}, \mathrm{M}_r) \neq (\hat{\mathrm{M}}, \hat{\mathrm{M}}_r) \text{ or } \mathrm{M}_r \neq \tilde{\mathrm{M}}_r], \tag{2.16}$$

  where M and M_r are uniformly distributed random variable over the message sets $\mathcal{M}$ and $\mathcal{M}_r$ respectively, while $\hat{\mathrm{M}}$, $\hat{\mathrm{M}}_r$ and $\tilde{\mathrm{M}}_r$ are the random variables that represents the output of the decoders at the legitimate receiver and the virtual receiver respectively. According to the standard joint-typicality decoding arguments [34, 35], $\mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] \leq f_n(\epsilon)$ where $\lim_{n\to\infty} f_n(\epsilon) = 0$, if

$$R + R_r \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}) - \delta_1(\epsilon) \quad \text{and} \quad R_r \leq \mathbb{I}(\mathrm{X}; \mathrm{Z}) - \delta_2(\epsilon). \tag{2.17}$$

Concerning the **"Secrecy Analysis"**, it is enough to show that under the rate constraints in (2.17), the following holds:

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}\big[\mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{C})\big] = 0. \tag{2.18}$$

This because based on the selection lemma (Lemma A.4), if the previous condition is true then there exists a certain codebook realization $\mathcal{C}$ for which the weak secrecy constraint in (2.7) is satisfied. The expectation in (2.18) can be further simplified as:

$$\mathbb{E}\big[\mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{C})\big] = \mathbb{E}\big[\mathbb{I}(\mathrm{M}\mathrm{M}_r;\ \mathrm{Z}^n|\mathrm{C}) - \mathbb{I}(\mathrm{M}_r; \mathrm{Z}^n|\mathrm{M}\mathrm{C})\big]$$

$$\begin{aligned}
&\overset{(a)}{=} \mathbb{E}\big[\mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n|\mathrm{C}) - \mathbb{H}(\mathrm{M}_r|\mathrm{MC}) + \mathbb{H}(\mathrm{M}_r|\mathrm{MZ}^n\mathrm{C})\big] \\
&\overset{(b)}{\leq} \mathbb{E}\big[\mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n) - \mathbb{H}(\mathrm{M}_r) + \mathbb{H}(\mathrm{M}_r|\mathrm{MZ}^n\mathrm{C})\big] \\
&\overset{(c)}{=} \mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n) - \mathbb{H}(\mathrm{M}_r) + \mathbb{E}\big[\mathbb{H}(\mathrm{M}_r|\tilde{\mathrm{M}}_r\mathrm{C})\big] \qquad (2.19)
\end{aligned}$$

where (a) follows from the codebook structure which implies that X^n is fully identified by M and M_r; (b) follows because $\mathrm{C} - \mathrm{X}^n - \mathrm{Z}^n$ forms a Markov chain and the fact that M_r is independent of M and C; while (c) follows by taking the terms independent of C out of the expectation and applying the decoding function at the virtual receiver. The three terms in (2.19) can be further bounded as follows:

$$\mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n) = n\mathbb{I}(\mathrm{X};\mathrm{Z}) \qquad (2.20a)$$

$$\mathbb{H}(\mathrm{M}_r) = nR_r \qquad (2.20b)$$

$$\mathbb{E}\big[\mathbb{H}(\mathrm{M}_r|\tilde{\mathrm{M}}_r\mathrm{C})\big] \leq 1 + \mathbb{E}[nR_r\,\bar{\mathbf{P}}_e(\mathrm{C})], \qquad (2.20c)$$

where Eq. (2.20a) follows because $(\mathrm{X}^n, \mathrm{Z}^n)$ is i.i.d. according to $\mathrm{P}_\mathrm{X}V_{\mathrm{Z}|\mathrm{X}}$; Eq. (2.20b) follows because M_r is uniformly distributed over the set $\mathcal{M}_r$; while Eq. (2.20c) follows from a relaxed version of Fano's inequality (Lemma A.2). Using the rate constraint in (2.17) with equality along with the fact that $\mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] \leq f_n(\epsilon)$, it follows that

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{C})\big] &\leq \lim_{n\to\infty} \delta_2(\epsilon) + \frac{1}{n} + R_r \cdot f_n(\epsilon) \\
&= \delta_2(\epsilon). \qquad (2.21)
\end{aligned}$$

In the construction of the code ϵ can be chosen to be relatively small, for example $\epsilon = n^{-\frac{1}{3}}$ cf. Theorem A.5, such that $\lim_{n\to\infty} \delta_1(\epsilon), \delta_2(\epsilon), f_n(\epsilon) = 0$. This implies that using a randomization rate R_r as given by (2.17) fulfills the weak secrecy constraint in (2.7) and this completes the secrecy analysis of the wiretap random code $\mathcal{C}$ under the weak secrecy constraint using the virtual receiver technique.

2.2.3 Variational Distance Measure

This section investigates the secrecy analysis of wiretap random code under the strong secrecy criterion. In general, two main concepts have been used to show that a certain coding scheme meets the strong secrecy criterion given in (2.8) [36]. The first concept is based on applying additional transformation to a weak secrecy coding scheme with the aim of strengthening the secrecy measure. One of the most popular techniques that apply this concept is the secret-key agreement technique introduced in [37]. The second concept is to apply powerful mathematical tools that can establish the strong secrecy results directly without deriving the weak secrecy results first. Among the techniques that apply the second concept are the graph-coloring technique presented in [38], the information-spectrum methods addressed in [39, 40] and the resolvability approach investigated in [28, 41]. This section considers a technique that applies the second concept and is based on Devetak's approach introduced in [42] and was used for various communication scenarios over wiretap channels cf. [29, 30].

Before presenting the exact steps of the secrecy analysis, it is essential to point out the main features of the coding scheme. Consider a coding scheme like the one introduced in the previous section but with two main exceptions:

- **Codebook and Encoding:** A sequence x^n generated by the input distribution $\mathrm{P_X}$ is included in the codebook, if and only if x^n belongs to the typical set $\mathcal{T}^n_{\epsilon_1}(\mathrm{P_X})$. This leads to a new input distribution as follows:

$$\tilde{\mathrm{P}}_{\mathrm{X}^n}(x^n) = \begin{cases} \frac{\mathrm{P}^n_{\mathrm{X}}(x^n)}{\mathrm{P}^n_{\mathrm{X}}(T^n_{\epsilon_1}(\mathrm{P_X}))} & \text{if } x^n \in \mathcal{T}^n_{\epsilon_1}(\mathrm{P_X}) \\ 0 & \text{otherwise.} \end{cases} \tag{2.22}$$

  Now, use $\tilde{\mathrm{P}}_{\mathrm{X}^n}(x^n)$ to generate the codewords $x^n(m, m_r)$ for $(m, m_r) \in \mathcal{M} \times \mathcal{M}_r$. Further, the encoder works exactly as the one in the previous section.
- **Decoding and Reliability:** No virtual receiver is considered and the decoder at the legitimate receiver works as the one in the previous section but with a new constant for the typicality set $\epsilon_2 \geq \epsilon_1$. Thus, for a code $\mathcal{C}$, the average decoding error probability is given by:

$$\bar{\mathbf{P}}_e(\mathcal{C}) \triangleq \mathbb{P}[(\mathrm{M}, \mathrm{M}_r) \neq (\hat{\mathrm{M}}, \hat{\mathrm{M}}_r)]. \tag{2.23}$$

  Based on the standard joint-typicality decoding arguments [34,35], one can show that $\mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] \leq f_n(\epsilon_1, \epsilon_2)$, where $\lim_{n\to\infty} f_n(\epsilon_1, \epsilon_2) = 0$, if

$$R + R_r \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}) - \delta(\epsilon_1, \epsilon_2). \tag{2.24}$$

The **"Secrecy Analysis"** of the modified coding scheme is based on a very interesting property. This property is that under a certain constraint on R_r, there exists a measure σ on the eavesdropper observation set $\mathcal{Z}^n$ such that:

$$\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}} - \sigma\right\| \leq 2^{-n\beta}, \tag{2.25}$$

where $\beta > 0$ and $\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}}$ is the distribution induced by the codebook on Z^n conditioned on the confidential message M. To find the required constraint on the randomization rate that proves the validity of Eq. (2.25), the following definitions are needed:

$$\forall z^n \in \mathcal{Z}^n \quad \tilde{\sigma}(z^n) \triangleq \mathbb{E}\left[V^n_{\mathrm{Z}^n|\mathrm{X}^n}(z^n|\mathrm{X}^n(m, m_r)) \times \mathbb{1}_{\mathcal{T}^n_{\epsilon_2}(\mathrm{P_X}V_{\mathrm{Z}|\mathrm{X}}|\mathrm{X}^n(m,m_r))}(z^n)\right] \tag{2.26a}$$

$$\forall \beta_n > 0, \quad \mathcal{F} \triangleq \left\{z^n \in \mathcal{Z}^n : z^n \in \mathcal{T}^n_{\epsilon_3}(\mathrm{P_Z}) \text{ and } \tilde{\sigma}(z^n) \geq \beta_n |\mathcal{T}^n_{\epsilon_3}(\mathrm{P_Z})|^{-1}\right\}, \tag{2.26b}$$

where $\epsilon_3 > \epsilon_2$ and $\mathrm{P_Z}$ is the marginal distribution of $\mathrm{P_X}V_{\mathrm{Z}|\mathrm{X}}$ on Z. Additionally, for any set $\mathcal{A} \subset \mathcal{Z}^n$, the indicator function $\mathbb{1}_{\mathcal{A}}$ is defined as follows: $\mathbb{1}_{\mathcal{A}}(z^n) = 1$, if $z^n \in \mathcal{A}$, otherwise $\mathbb{1}_{\mathcal{A}}(z^n) = 0$. Further, the definitions in (2.26) are used to describe the measure σ as follows:

$$\forall z^n \in \mathcal{Z}^n \quad \sigma(z^n) \triangleq \mathbb{E}\left[\tilde{V}^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m, m_r))\right] \tag{2.27}$$

$$\triangleq \mathbb{E}\left[V^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m, m_r)) \times \mathbb{1}_{\mathcal{T}^n_{\epsilon_2}\left(\mathrm{P_X}V_{\mathrm{Z}|\mathrm{X}}|\mathrm{X}^n(m,m_r)\right)}(z^n) \times \mathbb{1}_{\mathcal{F}}(z^n)\right] \tag{2.28}$$

$$= \tilde{\sigma}(z^n) \times \mathbb{1}_{\mathcal{F}}(z^n). \tag{2.29}$$

Now, based on the structure of the underlying coding scheme and the definition in (2.28), one can use the triangle inequality to bound the total variation distance in (2.25) for every confidential message $m \in \mathcal{M}$ as follows:

$$\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}}(\cdot|m) - \sigma(\cdot)\right\| \leq \left\|\frac{1}{|\mathcal{M}_r|}\sum_{m_r} \tilde{V}^n_{\mathrm{Z}|\mathrm{X}}(\cdot|x^n(m, m_r)) - \sigma(\cdot)\right\|$$

$$+\left\|\frac{1}{|\mathcal{M}_r|}\sum_{m_r} V^n_{\mathrm{Z}|\mathrm{X}}(\cdot|x^n(m,m_r)) \times \mathbb{1}_{\mathcal{T}^n_{\epsilon_2}(\mathrm{P_X V_{Z|X}}|x^n(m,m_r))}(\cdot)\Big(1-\mathbb{1}_{\mathcal{F}}(\cdot)\Big)\right\|$$
$$+\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}}(\cdot|m) - \frac{1}{|\mathcal{M}_r|}\sum_{m_r} V^n_{\mathrm{Z}|\mathrm{X}}\big(\,\cdot\,|x^n(m,m_r)\big) \times \mathbb{1}_{\mathcal{T}^n_{\epsilon_2}(\mathrm{P_X} V_{\mathrm{Z}|\mathrm{X}}|x^n(m,m_r))}(\cdot)\right\| \tag{2.30}$$

Denote the three terms summed up in the $\mathbb{RHS}$ of the previous equation by A_1, A_2 and A_3 respectively. In order to bound A_1, one can apply the Chernoff-Hoeffding bound given in Lemma A.3 to the i.d.d. random variables $\tilde{V}^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m,m_r))$ as follows:

$$\begin{aligned}
\mathbb{P}\left[\frac{1}{|\mathcal{M}_r|}\sum_{m_r}\tilde{\mathrm{V}}^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m,m_r)) \notin [(1\pm\beta_n)\sigma(z^n)]\right] &\overset{(a)}{\leq} 2\exp\left(-|\mathcal{M}_r|\cdot\frac{\beta_n^2\sigma(z^n)}{2b\ln 2}\right)\\
&\overset{(b)}{\leq} 2\exp\left(-|\mathcal{M}_r|\cdot\frac{\beta_n^2 2^{n(\mathbb{H}(\mathrm{Z}|\mathrm{X})-\delta(\epsilon_2))}\sigma(z^n)}{2\ln 2}\right)\\
&\overset{(c)}{\leq} 2\exp\left(-|\mathcal{M}_r|\cdot\frac{\beta_n^2 2^{n(\mathbb{H}(\mathrm{Z}|\mathrm{X})-\delta(\epsilon_2))}\beta_n 2^{-n(\mathbb{H}(\mathrm{Z})+\delta(\epsilon_3))}}{2\ln 2}\right)\\
&\overset{(d)}{=} 2\exp\left(-\frac{2^{n(R_r-\mathbb{I}(\mathrm{X};\mathrm{Z})-\delta(\epsilon_2,\epsilon_3)-3\beta)}}{2\ln 2}\right),
\end{aligned} \tag{2.31}$$

where (a) follows from the definition of $\sigma(z^n)$ in (2.27), (b) follows from the properties of typical sequences which implies that $\tilde{V}^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m,m_r)) \in [0, 2^{-n(\mathbb{H}(\mathrm{Z}|\mathrm{X})-\delta(\epsilon_2))}]$, (c) follows form the definition of $\mathcal{F}$ in (2.26b) and the fact that $|\mathcal{T}^n_{\epsilon_3}(\mathrm{P_Z})| \leq 2^{n(\mathbb{H}(\mathrm{Z})+\delta(\epsilon_3))}$; while (d) follows from the definition of $\mathcal{M}_r$ and by letting $\beta_n = 2^{-n\beta}$. It follows directly that as n approaches infinity, Eq. (2.31) approaches zero doubly-exponential, if

$$R_r \geq \mathbb{I}(\mathrm{X};\mathrm{Z}) + \delta(\epsilon_2,\epsilon_3) + 4\beta. \tag{2.32}$$

This implies that under the previous constraint, there exists a codebook realization $\mathcal{C}$, such that $A_1 \leq \beta_n$. On the other hand, one can notice that A_3 can be upper-bounded by the probability of the atypical set of Z conditioned on X. Moreover, A_2 can be bounded by the probability of the complement of the intersection of the conditional typical set of Z given X and the new defined set $\mathcal{F}$. This implies that:

$$A_3 \leq \delta_n(\epsilon_1,\epsilon_2), \qquad \text{and} \qquad A_2 \leq 2\beta_n + \delta_n(\epsilon_1,\epsilon_2), \tag{2.33}$$

where $\delta_n(\epsilon_1,\epsilon_2)$ is as defined in Theorem A.6. A detailed analysis for the calculation of the previous bounds is given in [29, Section IV]. Based on the properties of the total variation distance and the fact that M is uniformly distributed over the set $\mathcal{M}$, it follows directly that:

$$\begin{aligned}
\left\|\mathrm{P}_{\mathrm{M}\mathrm{Z}^n} - \mathrm{P}_{\mathrm{M}}\mathrm{P}_{\mathrm{Z}}\right\| &\leq 2\cdot\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}} - \sigma\right\|\\
&\leq 2(A_1+A_2+A_3) \leq 2^{-n\gamma},
\end{aligned} \tag{2.34}$$

for some $\gamma > 0$. Finally, by applying the selection lemma (Lemma A.4) along with Lemma A.6, one can show that there exists a code realization $\mathcal{C}$ with the rate constraints in (2.24) and (2.32), such that:

$$\lim_{n\to\infty}\bar{\mathbf{P}}_e(\mathcal{C}) = 0 \qquad \text{and} \qquad \lim_{n\to\infty}\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathcal{C}) = 0. \tag{2.35}$$

This completes the strong secrecy analysis of the constructed wiretap random code. The previous analysis showed that strengthening the secrecy constraint from the weak secrecy to the strong one comes at no cost with respect to the secrecy capacity of the general wiretap channel. This is because both criteria require nearly the same amount of randomization which is equivalent to the full rate of the eavesdropper channel $\mathbb{I}(\mathrm{X};\mathrm{Z})$.

2.2.4 Information Divergence Approximation

One of the most universal problems in information theory is the approximation of a product distribution. In particular, for a target distribution $\mathrm{Q}_{\mathrm{X}^n}$, what is the minimum rate required to generate a distribution $\mathrm{P}_{\mathrm{X}^n}$ such that $d(\mathrm{P}_{\mathrm{X}^n}, \mathrm{Q}_{\mathrm{X}^n})$ is small, for for some distance measure $d(\cdot,\cdot)$. This problem was first considered by Wyner in [43], where he identified the smallest rate required to approximate a product distribution with respect to the normalized information divergence (Kullback-Leibler divergence). Han and Verdú proved in [44] that the rate derived by Wyner is necessary and sufficient to approximate a marginal distribution of a given joint distribution with respect to the total variation distance. These results was used to construct resolvability codes capable of establishing secure communication over wiretap channels cf. [40,45]. In [46], Hou and Kramer showed that Wyner's result can be extended to the unnormalized information divergence and use this result to derive the secrecy capacity for wiretap channel under the newly introduced secrecy measure known as effective secrecy [31].

The previous discussion implies that the effective secrecy capacity of the general wiretap channel is equivalent to both the strong and weak secrecy capacities. This is because the smallest rate required to approximate a target distribution at the eavesdropper can be interpreted as the minimum randomization rate required to confuse the eavesdropper. Additionally, one can notice that the distance measure used by Wyner in [43] is strongly related to the weak secrecy constraint in (2.7), while the distance measure used by Han and Verdú in [44] is somehow related to the strong secrecy constraint in (2.8) as highlighted by Lemma A.6. In order to prove this equivalence, the following coding scheme is demonstrated.

Consider a coding scheme like the one introduced in Section 2.2.2, but without the virtual receiver. Since the encoding and decoding functions are identical to the one in Section 2.2.2, it follows directly that $\mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] = \mathbb{E}\left[\mathbb{P}[(\mathrm{M},\mathrm{M}_r) \neq (\hat{\mathrm{M}},\hat{\mathrm{M}}_r)|\mathrm{C}]\right] \leq f_n(\epsilon)$ where $\lim_{n\to\infty} f_n(\epsilon) = 0$, if

$$R + R_r \leq \mathbb{I}(\mathrm{X};\mathrm{Y}) - \delta(\epsilon). \tag{2.36}$$

Before deriving the step by step **"Secrecy Analysis"**, one needs to highlight some important features for the coding scheme. For a fixed codebook realization $\mathcal{C}$ generated using the input distribution P_{X}, the encoding function induces a new distribution over the codewords as follows:

$$\forall x^n(m,m_r) \in \mathcal{C} \qquad \mathrm{Q}_{\mathrm{X}^n}(x^n(m,m_r)) = \frac{1}{|\mathcal{M}||\mathcal{M}_r|}. \tag{2.37}$$

Obviously, $\mathrm{Q}_{\mathrm{X}^n}$ differs from $\mathrm{P}_{\mathrm{X}}^n$. This implies that one can define two distributions over the eavesdropper's observation set $\mathcal{Z}^n$. The first is the marginal distribution of

the joint distribution $\mathrm{Q}_{\mathrm{MZ}^n}$ induced by the code on the confidential message M and the eavesdropper's observation Z^n and is given by:

$$\mathrm{Q}_{\mathrm{MZ}^n}(m, z^n) = \frac{1}{|\mathcal{M}|} \sum_{m_r} \frac{1}{|\mathcal{M}_r|} V^n_{\mathrm{Z}|\mathrm{X}}(z^n|x^n(m, m_r)). \tag{2.38}$$

The second distribution is the one expected to be observed by the eavesdropper when no useful information is transmitted and is given by:

$$\begin{aligned} \mathrm{P}^n_{\mathrm{Z}}(z^n) &= \sum_{x^n} \mathrm{P}^n_{\mathrm{XZ}}(x^n, z^n) \\ &\triangleq \sum_{x^n} \mathrm{P}^n_{\mathrm{X}}(x^n) V^n_{\mathrm{Z}|\mathrm{X}}(z^n|x^n). \end{aligned} \tag{2.39}$$

Based on Eq. (2.38) and Eq. (2.39), one can show the existence of a codebook $\mathcal{C}$ that fulfills the effective secrecy constraint in (2.9), if the following holds:

$$\lim_{n\to\infty} \mathbb{E}[D_{\mathrm{KL}}(\mathrm{Q}_{\mathrm{MZ}^n} \| \mathrm{Q}_{\mathrm{M}} \mathrm{P}^n_{\mathrm{Z}})] = 0, \tag{2.40}$$

where Q_{M} is the marginal distribution of $\mathrm{Q}_{\mathrm{MZ}^n}$ on the confidential message M and the expectation is taken over the random variable C. Using the resolvability proof technique developed in [45, Lemma 11], one can show that Eq. (2.40) holds under a certain condition on the randomization rate R_r as follows:

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}[D_{\mathrm{KL}}(\mathrm{Q}_{\mathrm{MZ}^n} \| \mathrm{Q}_{\mathrm{M}} \mathrm{P}^n_{\mathrm{Z}})] &= \mathbb{E}_{\mathrm{C}} \left[\sum_m \sum_{z^n} \mathrm{Q}_{\mathrm{MZ}^n}(m, z^n) \log \frac{\mathrm{Q}_{\mathrm{MZ}^n}(m, z^n)}{\mathrm{Q}_{\mathrm{M}}(m) \cdot \mathrm{P}^n_{\mathrm{Z}}(z^n)} \right] \\
&\overset{(a)}{=} \mathbb{E}_{\mathrm{C}} \left[\sum_m \sum_{z^n} \frac{1}{|\mathcal{M}|} \sum_{m_r} \frac{1}{|\mathcal{M}_r|} V^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m, m_r)) \log \frac{\sum_{l=1}^{|\mathcal{M}_r|} \mathrm{V}^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m, l))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(z^n)} \right] \\
&= \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\mathrm{C}} \left[\sum_{z^n} V^n_{\mathrm{Z}|\mathrm{X}}(z^n|\mathrm{X}^n(m, m_r)) \log \frac{\sum_{l=1}^{|\mathcal{M}_r|} V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, l))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} \right] \\
&\overset{(b)}{=} \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n} \left[\mathbb{E}_{\tilde{\mathrm{C}}} \left[\log \left(\frac{\sum_{l=1}^{|\mathcal{M}_r|} V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, l))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} \right) \right] \right] \\
&\overset{(c)}{\leq} \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n} \left[\log \left(\mathbb{E}_{\tilde{\mathrm{C}}} \left[\frac{\sum_{l=1}^{|\mathcal{M}_r|} V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, l))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} \right] \right) \right] \\
&= \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n} \left[\log \left(\frac{V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, m_r))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} + \sum_{l \neq m_r}^{|\mathcal{M}_r|} \right.\right. \\
&\qquad \left.\left. \sum_{x^n(m,l)} \frac{\mathrm{P}^n_{\mathrm{XZ}}(x^n(m, l), \mathrm{Z}^n)}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} \right) \right] \\
&\overset{(d)}{=} \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n} \left[\log \left(\frac{V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, m_r))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} + \frac{|\mathcal{M}_r| - 1}{|\mathcal{M}_r|} \right) \right] \\
&\leq \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_m \sum_{m_r} \mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n} \left[\log \left(\frac{V^n_{\mathrm{Z}|\mathrm{X}}(\mathrm{Z}^n|\mathrm{X}^n(m, m_r))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(\mathrm{Z}^n)} + 1 \right) \right]
\end{aligned}$$

$$= \frac{1}{|\mathcal{M}||\mathcal{M}_r|} \sum_{m} \sum_{m_r} \sum_{x^n(m,m_r)} \sum_{z^n} \mathrm{P}^n_{\mathrm{XZ}}(x^n(m,m_r), z^n) \log\left(\frac{V^n_{\mathrm{Z|X}}(z^n|x^n(m,m_r))}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(z^n)} + 1\right)$$

$$\overset{(e)}{=} \sum_{x^n} \sum_{z^n} \mathrm{P}^n_{\mathrm{XZ}}(x^n, z^n) \log\left(\frac{V^n_{\mathrm{Z|X}}(z^n|x^n)}{|\mathcal{M}_r| \cdot \mathrm{P}^n_{\mathrm{Z}}(z^n)} + 1\right). \tag{2.41}$$

(a) follows due to the definition of $\mathrm{Q}_{\mathrm{MZ}^n}(m, z^n)$ in (2.38) and the fact that Q_{M} is a uniform distribution over $\mathcal{M}$. (b) follows by splitting the expectation over the codebook random variable C for a given (m, m_r) into two expectations: an inner expectation taken over a new random variable $\tilde{\mathrm{C}}$ given by the codewords $\mathrm{X}^n(m, l)$ for all $l \in \mathcal{M}_r \backslash \{m_r\}$ and an outer expectation taken over the random variables $\hat{\mathrm{C}}$ given by the codewords $\mathrm{X}^n(m, m_r)$. Additionally, it follows that $\mathbb{E}_{\hat{\mathrm{C}},\mathrm{Z}^n}[f(\mathrm{X}^n(m, m_r), \mathrm{Z}^n)] \triangleq \sum_{x^n(m,m_r)} \sum_{z^n} \mathrm{P}^n_{\mathrm{XZ}}(x^n(m, m_r), z^n) f(x^n(m, m_r), z^n)$. (c) follows by applying Jensen's inequality [47] to the expectation over $\tilde{\mathrm{C}}$ and the fact that the logarithm is a concave function. Finally (d) follows from (2.39), while (e) follows from the independence of the inner sum on the realization of the messages (m, m_r). The summation in (2.41) can be divided into a summation over the jointly typical sequences $(A_{\mathcal{T}})$ and a summation over the atypical ones $(A_{\mathcal{T}^c})$. Each of these two summations can be individually bounded as follows:

$$A_{\mathcal{T}^c} \overset{(a)}{\leq} \sum_{(x^n,z^n) \notin \mathcal{T}^n_\epsilon(\mathrm{P}_{\mathrm{XZ}})} \mathrm{P}^n_{\mathrm{XZ}}(x^n, z^n) \log\left(\frac{1}{\mu^n_{\mathrm{Z}}} + 1\right)$$

$$\overset{(b)}{\leq} \delta_n(\epsilon) \cdot n \cdot \log\left(\frac{1}{\mu_{\mathrm{Z}}} + 1\right), \tag{2.42}$$

where (a) follows because for all $z^n \in \mathcal{Z}^n$, it holds that $\mathrm{P}^n_{\mathrm{Z}}(z^n) \geq \mu^n_{\mathrm{Z}}$ given that $\mu_{\mathrm{Z}} \triangleq \min_{c \in supp(\mathrm{P}_{\mathrm{Z}})} \mathrm{P}_{\mathrm{Z}}(c)$; while (b) follows from the properties of atypical sequences as highlighted in Theorem A.5. On the other hand, the summation over the jointly typical sequences can be bounded as:

$$A_{\mathcal{T}} \overset{(a)}{\leq} \sum_{(x^n,z^n) \in \mathcal{T}^n_\epsilon(\mathrm{P}_{\mathrm{XZ}})} \mathrm{P}^n_{\mathrm{XZ}}(x^n, z^n) \log\left(\frac{2^{-n(1-\epsilon)\mathbb{H}(\mathrm{Z|X})}}{|\mathcal{M}_r| \cdot 2^{-n(1+\epsilon)\mathbb{H}(\mathrm{Z})}} + 1\right)$$

$$\leq \log\left(2^{-n[R_r - \mathbb{I}(\mathrm{X};\mathrm{Z}) - 2\tilde{\delta}(\epsilon)]} + 1\right)$$

$$\leq \log(e) \cdot 2^{-n[R_r - \mathbb{I}(\mathrm{X};\mathrm{Z}) - 2\tilde{\delta}(\epsilon)]}, \tag{2.43}$$

where (a) follows from the consistency of joint typicality along with the properties of typical sequences in Theorem A.5 and conditionally typical sequences in Theorem A.6. Eq. (2.42) and Eq. (2.43) imply that the effective secrecy constraint in (2.40) holds, as long as

$$R_r \geq \mathbb{I}(\mathrm{X};\mathrm{Z}) + 3\tilde{\delta}(\epsilon). \tag{2.44}$$

Finally, by applying the selection lemma (Lemma A.4), one can show that there exists a code realization $\mathcal{C}$ with the rate constraints in (2.36) and (2.44), that satisfies the required reliability and secrecy constraints. This result shows that even effective secrecy comes at no cost for the secrecy capacity of the general wiretap channel.

2.3 Secrecy Capacity for Wiretap Channels with Shared Key

In this section, secure communication scenario that combines the concept of Shannon's ciphering system and the general wiretap channel is investigated. The investigation of this scenario aims to highlight the basic coding schemes for secure communication and the various ways of combining them. It also serves as a good introduction for the more complex models investigated in this thesis. The previous communication scenario was first introduced in [48, 49] with the aim of investigating the rate-distortion problem of wiretap channels with shared secret key. It also appeared as a consequence of studying the problem of secure communication over wiretap channels with secure feedback link in [50–52]. The first secrecy capacity result was established in [48] but only for the degraded and less noisy wiretap channels.

2.3.1 Channel Model and Coding Theorem

Consider a general wiretap channel similar to the one in Fig. 2.3, but with an additional feature which is the existence of a secret key shared between the transmitter and the legitimate receiver, while being concealed from the eavesdropper as shown in Fig. 2.4. The shared secret key K is uniformly distributed over the finite discrete alphabet $\mathcal{K}$. The task is to develop a reliable and secure coding scheme that can be used to transmit

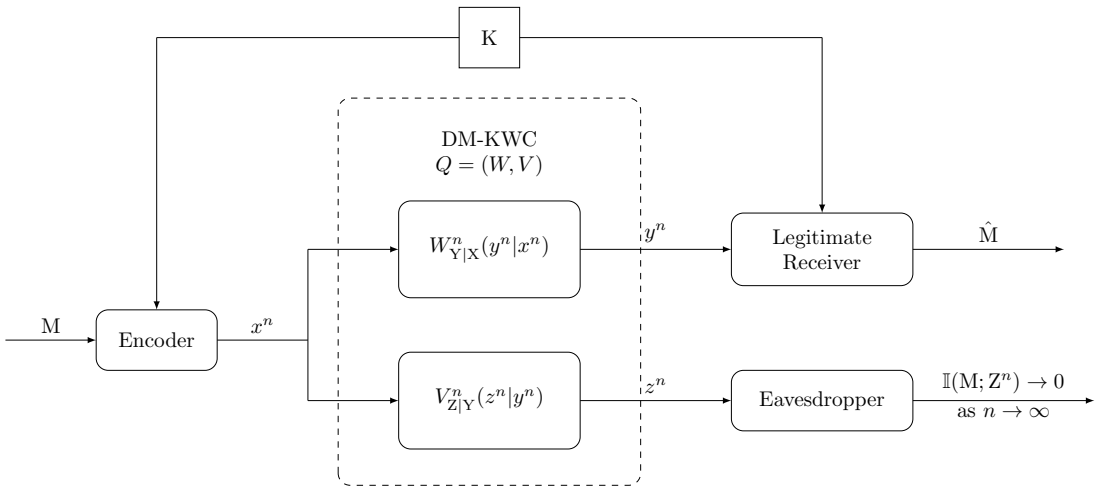


Figure 2.4: Wiretap channel with shared secret key

a confidential message M uniformly distributed over a finite and discrete alphabet $\mathcal{M}$ over the wiretap channel (W, V).

Definition 2.2. A $(2^{nR}, n)$ code $\mathcal{C}$ for the wiretap channel with shared secret key consists of: a confidential message set $\mathcal{M} = [\![1, 2^{nR}]\!]$, a secret key set $\mathcal{K} = [\![1, 2^{nR_k}]\!]$, a stochastic encoder at the transmitter

$$E : \mathcal{M} \times \mathcal{K} \to \mathcal{P}(\mathcal{X}^n) \tag{2.45}$$

which maps a confidential message $m \in \mathcal{M}$ along with a realization of the secret key $k \in \mathcal{K}$ to a codeword $x^n(m, k) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m, k)$, and a deterministic decoder at the legitimate receiver

$$\varphi : \mathcal{Y}^n \times \mathcal{K} \to \mathcal{M} \tag{2.46}$$

that transforms the channel observation at the legitimate receiver $y^n \in \mathcal{Y}^n$ along with the key realization k to a certain confidential message.

The code $\mathcal{C}$ is known to the transmitter, the legitimate receiver and the eavesdropper. It is also assumed that perfect channel state information (CSI) is available at the three nodes. This implies that the three nodes know the channel statistics a head of time. In order to measure the secrecy performance of the code the strong secrecy criterion in (2.8) is investigated. Moreover, the average decoding error probability is used to evaluate the reliability performance of $\mathcal{C}$ as follows:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}) &= \mathbb{P}[\mathrm{M} \neq \hat{\mathrm{M}}|\mathcal{C}] = \frac{1}{|\mathcal{M}|}\sum_m \mathbf{P}_e(m|\mathcal{C}) \\ &= \frac{1}{|\mathcal{M}||\mathcal{K}|}\sum_m\sum_k\sum_{x^n}\sum_{y^n:\varphi(y^n,k)\neq m} W^n(y^n|x^n)E(x^n|m,k), \end{aligned} \tag{2.47}$$

where M is a random variable that represents the confidential message given to the encoder, while $\hat{\mathrm{M}}$ is a random variable that represents the output of the decoder. Indeed, one can use an alternative secrecy measure like the ones introduced in Section 2.1.4 or a different reliability measure such as the maximum decoding error probability given by:

$$\tilde{\mathbf{P}}_e(\mathcal{C}) = \max_m \mathbf{P}_e(m|\mathcal{C}). \tag{2.48}$$

However, throughout this thesis the average error probability and the strong secrecy criterion will be the main reliability and secrecy measures considered.

Definition 2.3. A non negative number $R \in \mathbb{R}_+$ is a strong secrecy achievable rate for the wiretap channel (W, V) with shared secret key, if for all $\eta, \lambda, \tau > 0$ there is an $n(\eta, \lambda, \tau) \in \mathbb{N}$ such that for all $n > n(\eta, \lambda, \tau)$ there exists a sequence of codes $\{\mathcal{C}\}_n$ that satisfies the following constraints:

$$\frac{1}{n}\log|\mathcal{M}| \geq R - \eta \tag{2.49a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda \tag{2.49b}$$

$$\mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathcal{C}) \leq \tau \tag{2.49c}$$

The strong secrecy capacity $C(W, V)$ is given by the supremum of all achievable strong secrecy rates R.

Theorem 2.3. *The strong secrecy capacity of the discrete memoryless wiretap channel (W, V) with shared secret key* K *of rate R_k is given by the following expression:*

$$C(W,V) = \max_{\mathrm{P_{UVX}}} \min\Big[\big[\mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})\big]^+ + R_k, \mathbb{I}(\mathrm{V};\mathrm{Y})\Big], \tag{2.50}$$

where the maximum is taken over all possible input distributions $\mathrm{P_{UVX}}$ *defined over the random variables* U, V, X *such that* U − V − X *forms a Markov chain. Moreover,* $[a]^+$ *is defined as the maximum between* 0 *and* a*. Further, it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 1$ *and* $|\mathcal{V}| \leq |\mathcal{X}|^2 + 3|\mathcal{X}| + 2$.

The previous coding theorem was first established in [53] but only under the weak secrecy constraint. However, based on the results of the previous section one can easily shows that strengthening the secrecy requirement to strong secrecy or the effective

secrecy criterion comes at no cost. Additionally, it was shown that the previous coding theorem is also valid for the maximum error probability reliability constraint. This because for DMCs, where perfect CSI is available, strengthening the reliability constraint from the average to the maximum decoding error probability has no effect on the capacity [54, 55].

2.3.2 Achievability Proof

In this section, the proof of the direct part of Theorem 2.3 which is also known as the achievability proof is presented. The proof is based on the combination of the two fundamental coding schemes for secure communication: secret key encoding and wiretap random coding. Now, for a given wiretap channel (W, V) and a fixed input distribution $\mathrm{P_{UVX}} = \mathrm{P_{UV}} \cdot \mathrm{P_{X|V}}$, there exist three coding scenarios:

Scenario 1: If $\mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) \leq \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})$, then one needs to prove the achievability of any rate that satisfies the following:

$$R \leq \min[R_k, \mathbb{I}(\mathrm{V};\mathrm{Y})]. \tag{2.51}$$

This is done by using the secret key encoding scheme as follows:

1. Codebook Generation $\mathcal{C}_1$: Let $R = R_k$, then construct a new message set $\mathcal{M}_\otimes = [\![1, 2^{nR_\otimes}]\!]$ by *xoring* the corresponding elements of $\mathcal{M}$ and $\mathcal{K}$, i.e. $m_\otimes = m \otimes k$. Next, generate the codewords $v^n(m_\otimes)$ for $m_\otimes \in \mathcal{M}_\otimes$ by generating the symbols $v_i(m_\otimes)$ with $i \in [\![1, n]\!]$ independently at random according to $\mathrm{P_V}$.

2. Encoding E_1: Given the confidential message $m \in \mathcal{M}$ and the key $k \in \mathcal{K}$, the encoder selects the codeword $v^n(m \otimes k)$ then generates a codeword x^n independently at random according to $\prod_{i=1}^n \mathrm{P_{X|V}}(\cdot|v_i(m \otimes k))$ and transmits it.

3. Decoding φ_1: Upon receiving y^n, the decoder determines a unique codeword $v^n(\hat{m}_\otimes)$ which is jointly typical with y^n. It then uses $\hat{m}_\otimes$ and its knowledge of the key to produce an estimate for the confidential message as follows: $\hat{m} = \hat{m}_\otimes \otimes k$.

4. Reliability Analysis: In order to show that this coding scheme satisfies the reliability constraint in (2.49b), it is enough to prove that:

$$\mathbb{P}[\mathrm{M}_\otimes \neq \hat{\mathrm{M}}_\otimes|\mathcal{C}_1] \leq \lambda_n, \tag{2.52}$$

where $\lim_{n\to\infty} \lambda_n = 0$. According to the channel coding theorem (Theorem A.3) and the standard joint-typicality decoding arguments [34], the constraint in (2.52) holds if:

$$R = R_\otimes \leq \mathbb{I}(\mathrm{V};\mathrm{Y}) - \delta(\epsilon), \tag{2.53}$$

where $\epsilon > 0$ is a constant used in the typicality decoding. Additionally, the equality of R and $R_\otimes$ follows from the codebook generation. As highlighted in the previous section, one can choose ϵ in a way such that $\lim_{n\to\infty} \delta(\epsilon) = 0$.

5. Secrecy Analysis: Since the secret key encoding scheme is based on the principle of one time pad (Shannon's ciphering system), then in order to fulfill the secrecy constraint in (2.49c), it is enough to have:

$$\mathbb{H}(\mathrm{M}) \leq \mathbb{H}(\mathrm{K}) \qquad \Longrightarrow \qquad R \leq R_k, \tag{2.54}$$

where the later follows because both M and K are uniformly distributed on their respective sets. In fact the condition in (2.54) does not only fulfill the strong secrecy criterion but it also assures perfect secrecy.

Now, by combining Eq. (2.53) and Eq. (2.54) then taking the limit as $n \to \infty$, which implies that $\delta(\epsilon) = 0$, it follows that any rate R that satisfies the constraint in (2.51) is achievable.

Scenario 2: If $\mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) \geq \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})$ and $R_k - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \leq \mathbb{I}(\mathrm{U};\mathrm{Y})$, then one needs to prove the achievability of any rate that satisfies the following:

$$R \leq \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) + R_k. \tag{2.55}$$

This follows due to Eq. (2.50) and the fact that $\mathbb{I}(\mathrm{V};\mathrm{Y}) = \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Y})$. The achievability proof of this condition is based on the wiretap random encoding scheme as follows:

1. Codebook Generation $\mathcal{C}_2$: Let $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$ be a new message set and let M_r be a random variable uniformly distributed over this set. Randomly generate a sequence u^n by generating the symbols u_i with $i \in [\![1, n]\!]$ independently at random according to P_{U}. Next, construct the codewords $v^n(m, k, m_r)$ for $m \in \mathcal{M}$, $k \in \mathcal{K}$ and $m_r \in \mathcal{M}_r$ by generating the symbols $v_i(m, k, m_r)$ independently at random according to $\mathrm{P}_{\mathrm{V}|\mathrm{U}}(\cdot|u_i)$.

2. Encoding E_2: Given the confidential message $m \in \mathcal{M}$ and the key $k \in \mathcal{K}$, the encoder selects a message m_r uniformly at random then generates a codeword x^n independently at random according to $\prod_{i=1}^{n} \mathrm{P}_{\mathrm{X}|\mathrm{V}}(\cdot|v_i(m, k, m_r))$ and transmits it.

3. Decoding φ_2: Upon receiving y^n, the decoder uses it along with the shared key k to determine a unique codeword $v^n(\hat{m}, k, \hat{m}_r)$ which is jointly typical with y^n conditioned on the sequence u^n.

4. Reliability Analysis: In order to show that this coding scheme satisfies the reliability constraint in (2.49b), it is enough to prove that:

$$\mathbb{P}[(\mathrm{M}, \mathrm{M}_r) \neq (\hat{\mathrm{M}}, \hat{\mathrm{M}}_r)|\mathcal{C}_2] \leq \lambda_n, \tag{2.56}$$

where $\lim_{n\to\infty} \lambda_n = 0$. According to the channel coding theorem and the standard joint-typicality decoding arguments, the constraint in (2.56) holds if:

$$R + R_r \leq \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) - \delta(\epsilon_1, \epsilon_2), \tag{2.57}$$

where $\epsilon > 0$ is a constant used during the typicality decoding. Additionally, one can assure that $\lim_{n\to\infty} \delta(\epsilon) = 0$.

5. Secrecy Analysis: Based on the strong secrecy and the effective secrecy analysis of wiretap random codes presented in the previous section, one can show that the secrecy constraint in (2.49c) is fulfilled as long as:

$$R_k + R_r \geq \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) + \tilde{\delta}(\tau_n), \tag{2.58}$$

where $\lim_{n\to\infty} \tilde{\delta}(\tau_n) = 0$. It worth mentioning that Eq. (2.58) implies that R_k is considered as a part of the randomization rate used to confuse the eavesdropper.

Now, by combining Eq. (2.57) and Eq. (2.58) then taking the limit as $n \to \infty$, it follows directly that any rate R that satisfies the constraint in (2.55) is achievable.

Scenario 3: If $\mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) \geq \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})$ and $R_k - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \geq \mathbb{I}(\mathrm{U};\mathrm{Y})$, then one needs to prove the achievability of any rate that satisfies the following:

$$R \leq \mathbb{I}(\mathrm{V};\mathrm{Y}). \tag{2.59}$$

This follows due to Eq. (2.50) and the fact that $\mathbb{I}(\mathrm{V};\mathrm{Y}) = \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Y})$. The achievability proof of this condition is based on combining the two techniques: secret key encoding and wiretap random encoding scheme along with concept of rate splitting and superposition encoding as follows:

1. Rate Splitting: Divide the confidential message set and the secret key set into two parts as follows: $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ and $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2$. Each of the new constructed message sets can be written as $[\![1, 2^{nR_a}]\!]$, where a is an indicator that identifies each set. Additionally, one needs to make sure that this division satisfies the following rate constraint: $R_1 = R_{k_1}$.

2. Codebook Generation $\mathcal{C}_3$: Construct a new message set $\mathcal{M}_\otimes = [\![1, 2^{nR_\otimes}]\!]$, where $R_\otimes = R_1$. The construction of this message set is done by *xoring* the elements of $\mathcal{M}_1$ and $\mathcal{K}_1$. Moreover, Let $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$ be a randomization message set with M_r be a uniformly distributed random variable over it. Next, randomly generate the cloud center codewords $u^n(m_\otimes)$ for $m_\otimes \in \mathcal{M}_\otimes$ by generating the symbols $u_i(m_\otimes)$ with $i \in [\![1, n]\!]$ independently at random according to P_U. For each cloud center codeword $u^n(m_\otimes)$, generate the satellite codewords $v^n(m_\otimes, m_2, k_2, m_r)$ for $m_2 \in \mathcal{M}_2$, $k_2 \in \mathcal{K}_2$ and $m_r \in \mathcal{M}_r$ by generating the symbols $v_i(m_\otimes, m_2, k_2, m_r)$ independently at random according to $\mathrm{P}_{\mathrm{V}|\mathrm{U}}(\cdot|u_i(m_\otimes))$.

3. Encoding E_3: Given the confidential message $m = (m_1, m_2) \in \mathcal{M}$ and the key $k = (k_1, k_2) \in \mathcal{K}$, the encoder selects a message m_r uniformly at random then generates a codeword x^n independently at random according to $\prod_{i=1}^n \mathrm{P}_{\mathrm{X}|\mathrm{V}}(\cdot|v_i(m_\otimes, m_2, k_2, m_r))$, where $m_\otimes = m_1 \otimes k_1$ then it transmits x^n.

4. Decoding φ_3: Upon receiving y^n, the decoder uses it along with the shared key $k = (k_1, k_2)$ to determine a unique pair of codewords $\big(u^n(\hat{m}_\otimes), v^n(\hat{m}_\otimes, \hat{m}_2, k_2, \hat{m}_r)$ which is jointly typical with y^n. Next, it uses the estimated $\hat{m}_\otimes$ and k_1 to produce the message $\hat{m}_1 = \hat{m}_\otimes \otimes k_1$.

5. Reliability Analysis: In order to show that this coding scheme satisfies the reliability constraint in (2.49b), it is enough to prove that:

$$\mathbb{P}[(\mathrm{M}_\otimes, \mathrm{M}_2, \mathrm{M}_r) \neq (\hat{\mathrm{M}}_\otimes, \hat{\mathrm{M}}_2, \hat{\mathrm{M}}_r)|\mathcal{C}_3] \leq \lambda_n, \tag{2.60}$$

where $\lim_{n\to\infty} \lambda_n = 0$. According to the channel coding theorem, the principle of superposition encoding [56] and the standard joint-typicality decoding arguments, the constraint in (2.60) holds if:

$$R_1 = R_{k_1} = R_\otimes \leq \mathbb{I}(\mathrm{U};\mathrm{Y}) - \delta_1(\epsilon) \tag{2.61a}$$

$$R_2 + R_r \leq \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) - \delta_2(\epsilon), \tag{2.61b}$$

where $\epsilon > 0$ is a constant used during the typicality decoding. Additionally, one can assure that $\lim_{n\to\infty} \delta_1(\epsilon), \delta_2(\epsilon) = 0$.

6. Secrecy Analysis: Due to the splitting of the confidential message set into two parts, the secrecy constraint in (2.49c) can be redefined under this coding scheme as follows:

$$\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathcal{C}_3) = \mathbb{I}(\mathrm{M}_1;\mathrm{Z}^n|\mathcal{C}_3) + \mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1\mathcal{C}_3) \leq \tau_n \tag{2.62}$$

Based on the secrecy analysis of the secret key encoding scheme, the first term in (2.62) vanishes, as long as:

$$\mathbb{H}(\mathrm{M}_1) \leq \mathbb{H}(\mathrm{K}_1) \qquad \Longrightarrow \qquad R_1 \leq R_{k_1}, \tag{2.63}$$

On the other hand, based on the strong secrecy and effective secrecy analysis of wiretap random codes presented in the previous section, the second term in (2.62) is less than τ_n, if:

$$R_{k_2} + R_r \geq \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) + \tilde{\delta}(\tau_n), \tag{2.64}$$

where $\lim_{n\to\infty} \tilde{\delta}(\tau_n) = 0$. This implies that the rate constraints in (2.63) and (2.64) assures the fulfillment of the secrecy constraint in (2.62).

Now, by combining Eq. (2.61), Eq. (2.63) and Eq. (2.64) followed by applying the Fourier-Motzkin elimination procedure [57] then taking the limit as $n \to \infty$, it follows directly that any rate R that satisfies the constraint in (2.59) is achievable.

It is important to point out that the achievability proof of Theorem 2.3 presented in [53] is a little bit different from the one discussed here. In [53], the authors used some of the properties of the mutual information to transform the capacity expression in (2.50) into an equivalent expression. Then, they prove the achievability of the equivalent expression using only wiretap random codes. Although their technique is totally valid, it does not bring the different secrecy coding schemes to the light. Alternatively, the presented achievability proof shows that for a fixed input distribution, a given secret key and the CSI, one can select the optimal coding scheme that achieves the secrecy capacity. This coding scheme might be only secret key encoding like the first scenario or only wiretap encoding as in the second scenario or a combination of the two techniques as in the third scenario. This phenomena is of high importance because it clarifies the strengths of each coding scheme. Additionally, the presented achievability proof emphasizes the importance of having a perfect CSI available at the transmitter because the selection of the optimal coding scheme depends on this information.

2.3.3 Converse Proof

In this section, the proof of the converse part of Theorem 2.3 is presented. The aim of the converse proof is to show that the rate in (2.50) is in fact the maximum possible rate that can be achieved under the required reliability and secrecy constraints. In particular, there are two classes of converse proofs: a weak converse and a strong converse. The main difference between weak and strong converse is the rate at which the reliability and secrecy constraints are violated if the rate exceeds the capacity [6, 58]. In this thesis, only weak converses are investigated and presented.

Now, let R be an achievable strong secrecy rate for the wiretap channel with shared secret key (W, V) and let $\lambda, \tau > 0$. For n sufficiently large, assume that there exists a $(2^{nR}, n)$ code $\mathcal{C}$ such that:

$$nR \leq \mathbb{H}(\mathrm{M}|\mathcal{C}), \qquad \bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n = 2^{-n\lambda}, \qquad \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathcal{C}) \leq \tau_n = 2^{-n\tau}. \tag{2.65}$$

Based on Fano's inequality (Lemma A.2) and the fact that the decoding function of the code $\mathcal{C}$ at the legitimate receiver is not only supplied by the observations Y^n but

with the secret key K as well, it follows that

$$\mathbb{H}(\mathrm{M}|\mathrm{Y}^n\mathrm{K}\mathcal{C}) \leq \mathbb{H}_2(\bar{\mathbf{P}}_e(\mathcal{C})) + nR \cdot \bar{\mathbf{P}}_e(\mathcal{C}) = \Gamma(\lambda_n), \tag{2.66}$$

where one can easily show that $\lim_{n\to\infty} \Gamma(\lambda_n) = 0$. In order to simplify the notation, the conditioning on the code is ignored in next steps. Thus, it follows immediately by combining the constraints in (2.65) and (2.66) that

$$\begin{aligned}
nR &\leq \mathbb{H}(\mathrm{M}) - \mathbb{H}(\mathrm{M}|\mathrm{Y}^n\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n) + \Gamma(\lambda_n, \tau_n) \\
&\overset{(a)}{=} \mathbb{H}(\mathrm{M}|\mathrm{K}) - \mathbb{H}(\mathrm{M}|\mathrm{Y}^n\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n) + \Gamma(\lambda_n, \tau_n) \\
&\overset{(b)}{=} \mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{K}) + \mathbb{I}(\mathrm{K};\mathrm{Z}^n|\mathrm{M}) + \Gamma(\lambda_n, \tau_n) \\
&\overset{(c)}{\leq} \mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{K}) + \mathbb{H}(\mathrm{K}|\mathrm{M}) + \Gamma(\lambda_n, \tau_n) \\
&\overset{(d)}{=} \mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{K}) + nR_k + \Gamma(\lambda_n, \tau_n) \\
&\leq [\mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{K})]^+ + nR_k + \Gamma(\lambda_n, \tau_n),
\end{aligned} \tag{2.67}$$

where $\Gamma(\lambda_n, \tau_n) = \Gamma(\lambda_n) + \tau_n$. (a) follows from the properties of the entropy along with the fact that M and K are independent. (b) follows from the chain rule of the mutual information. (c) follows from the relation between the mutual information and entropy. (d) follow because K is uniformly distributed over the set $\mathcal{K}$. Moreover, one can use Eq. (2.66) to derive another upper-bound on R by ignoring the secrecy constraint in (2.65) as follows:

$$\begin{aligned}
nR &\leq \mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) + \Gamma(\lambda_n) \\
&\overset{(a)}{\leq} \mathbb{I}(\mathrm{MK};\mathrm{Y}^n) + \Gamma(\lambda_n),
\end{aligned} \tag{2.68}$$

where (a) follows from the chain rule of the mutual information and the fact that $\mathbb{I}(\mathrm{K};\mathrm{Y}^n) \geq 0$. At the moment, let $\mathrm{K} \triangleq \mathrm{U}$ and $(\mathrm{M},\mathrm{K}) \triangleq \mathrm{V}$. Hence, by merging the two upper-bounds in Eq. (2.67) and Eq. (2.68) after dividing the $\mathbb{RHS}$ of each of them by n and taking the limit as n approaches infinity, the following bound is established:

$$R \leq \max_{\mathrm{P_{UV}}} \max_{\mathrm{P_{X^n|UV}}} \lim_{n\to\infty} \frac{1}{n} \min\Big[[\mathbb{I}(\mathrm{V};\mathrm{Y}^n|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}^n|\mathrm{U})]^+ + nR_k, \mathbb{I}(\mathrm{V};\mathrm{Y}^n)\Big], \tag{2.69}$$

where the convergence of the limit is guaranteed by the Fekete's lemma [59]. The previous bound is known as a Multi-letter upper-bound because it is given in terms of the n-dimensional random variables Y^n and Z^n. One can notice that the multi-letter bound in (2.69) matches the rate region in (2.50) applied to the n-fold product of the wiretap channel with secret key (W^n, V^n). This implies that the bound in (2.69) defines a multi-letter description for the secrecy capacity of the channel. The problem of multi-letter descriptions is that they might not be efficiently computable if the limit does not converge, that is why a single-letter description for the capacity is always more desirable. Nevertheless, it has been shown that multi-letter descriptions are extremely valuable as they can establish many important features for the capacity [60].

In order to prove the single-letter converse for the capacity expression in (2.50), the following random variables need to be introduced: For $i \in [\![1, n]\!]$, let $\mathrm{U}_i \triangleq (\mathrm{K}, \mathrm{Y}^{i-1}, \tilde{\mathrm{Z}}^{i+1})$ and $\mathrm{V}_i \triangleq (\mathrm{M}, \mathrm{U}_i)$, where for a given sequence A^n, the following convention holds:

$\mathrm{A}^i \triangleq (\mathrm{A}_1, \ldots, \mathrm{A}_i)$ and $\tilde{\mathrm{A}}^i \triangleq (\mathrm{A}_i, \ldots, \mathrm{A}_n)$. Now, consider the bound in (2.67), it can be reformulated as follows:

$$\begin{aligned}
R &\leq \frac{1}{n}\left[\mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{K}) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{K})\right]^+ + R_k + \tilde{\Gamma}(\lambda_n, \tau_n) \\
&\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{I}(\mathrm{M};\mathrm{Y}_i|\mathrm{K}\mathrm{Y}^{i-1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{K}\tilde{\mathrm{Z}}^{i+1})\right]^+ + R_k + \tilde{\Gamma}(\lambda_n, \tau_n) \\
&\overset{(b)}{=} \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{I}(\mathrm{M};\mathrm{Y}_i|\mathrm{K}\mathrm{Y}^{i-1}\tilde{\mathrm{Z}}^{i+1})) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{K}\mathrm{Y}^{i-1}\tilde{\mathrm{Z}}^{i+1})\right]^+ + R_k + \tilde{\Gamma}(\lambda_n, \tau_n) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{I}(\mathrm{V}_i;\mathrm{Y}_i|\mathrm{U}_i) - \mathbb{I}(\mathrm{V}_i;\mathrm{Z}_i|\mathrm{U}_i)\right]^+ + R_k + \tilde{\Gamma}(\lambda_n, \tau_n), \qquad (2.70)
\end{aligned}$$

where $\tilde{\Gamma}(\lambda_n, \tau_n) = \frac{1}{n}\Gamma(\lambda_n, \tau_n)$. (a) follows from the mutual information chain rule; while (b) follows due to the Csiszár and Körner sum identity [10, Lemma 7]. Using the same technique, one can rewrite the bound in (2.68) as follows:

$$\begin{aligned}
R &\leq \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{MK};\mathrm{Y}_i|\mathrm{Y}^{i-1}) + \tilde{\Gamma}(\lambda_n) \\
&\overset{(a)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{V}_i;\mathrm{Y}_i) + \tilde{\Gamma}(\lambda_n), \qquad (2.71)
\end{aligned}$$

where $\Gamma(\tilde{\lambda}_n) = \frac{1}{n}\Gamma(\lambda_n)$ and (a) follows due to the chain rule of the mutual information. Finally, by merging Eq. (2.70) and Eq. (2.71), followed by introducing a time sharing random variable T independent of all others and uniformly distributed over $[\![1, n]\!]$, then taking the limit as n approaches infinity such that $\tilde{\Gamma}(\lambda_n)$ and $\tilde{\Gamma}(\lambda_n, \tau_n)$ vanishes, the following bound is established:

$$R \leq \max_{\mathrm{P}_{\mathrm{UVX}}} \min\Big[\left[\mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})\right]^+ + R_k, \mathbb{I}(\mathrm{V};\mathrm{Y})\Big], \qquad (2.72)$$

where $\mathrm{U} = (\mathrm{U}_\mathrm{T}, \mathrm{T})$, $\mathrm{V} = \mathrm{V}_\mathrm{T}$, $\mathrm{Y} = \mathrm{Y}_\mathrm{T}$ and $\mathrm{Z} = \mathrm{Z}_\mathrm{T}$. Since the previous bound matches the acheivable bound established in the previous section, the converse proof of Theorem 2.3 is complete.

2.4 Two-Receiver Wiretap Broadcast Channels

This section considers the problem of secure communication over the discrete memoryless broadcast channel (DM-BC) with two legitimate receivers and an eavesdropper. At first, the channel model is introduced in addition to the motivation behind this model. Next, different communication scenarios over this channel model are presented along with the most important results established for each scenario. Moreover, the main coding techniques used to establish these results are investigated briefly. Finally, the model of the corresponding Gaussian channel is discussed.

2.4.1 Motivation and Channel Model

The investigation of complicated secure communication scenarios over the wiretap channel have captured a lot of attention in the last few years. This investigation was

inspired by the work of Csiszár and Körner in [10] which considered the transmission of public and confidential messages over the two-receiver DM-BC. This is because it generalized Wyner's model of secure communication over the degraded wiretap channel to a more real life scenario. Since then, a lot of effort have been directed to extend the results in [10] for various communication services (public and confidential) [61–63] and with respect to different secrecy setups [29]. Nevertheless, most of these extensions only considered a channel with two receiving nodes only (a legitimate receiver and an eavesdropper). This was a little bit unsatisfying; because with the rapid growth of networks, it has become crucial to examine how common and confidential services can be integrated over multi-receiver channels. In this section, the first step is taken in this direction by investigating secure communication over a DM-BC with two legitimate receivers and an eavesdropper.

Let $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ be discrete input and output alphabets for the DM-WBC (Q) with two legitimate receivers and an eavesdropper given by the transition matrix $Q : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z})$. Thus, for the input and output sequences $x^n \in \mathcal{X}^n$, $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$ and $z^n \in \mathcal{Z}^n$ of length n, the DMC is given by:

$$Q^n_{\mathrm{Y_1Y_2Z|X}}(y_1^n, y_2^n, z^n|x^n) = \prod_{i=1}^{n} Q_{\mathrm{Y_1Y_2Z|X}}(y_{1_i}, y_{2_i}, z_i|x_i). \tag{2.73}$$

In general, the receiving nodes are not supposed to cooperate during decoding. This implies that one can interpret (Q) as three independent DMCs from the transmitter to the first and second legitimate receivers as well as the eavesdropper as follows: $W^1 : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_1)$, $W^2 : \mathcal{X} \to \mathcal{P}(\mathcal{Y}_2)$ and $V : \mathcal{X} \to \mathcal{P}(Z)$. It is important to highlight that the investigation of secure communication over this channel model is done under the assumption of the availability of perfect CSI for the three DMCs (W^1, W^2, V) at the transmitter. Throughout the rest of the thesis, this channel model will be called as the two-receiver discrete memoryless wiretap broadcast channel (DM-WBC), where it is assumed that the word wiretap directly brings an additional receiver (eavesdropper) to the system.

In this thesis, two main secrecy scenarios will be considered over the two-receiver DM-WBC: The first scenario considers the transmission of a common confidential message to both legitimate receivers, while the second consider an individual confidential message for each legitimate receiver. For both scenarios, the integration of public services along with the confidential ones is considered as well. The main task of investigating these scenarios is to derive outer and inner bounds on the capacity region. While trying to achieve this target, researches found out that the straightforward extension of the single-user coding schemes is only optimal for some special multi-user channels [12] where for the general case more complex coding schemes are needed. Moreover, the reliability and secrecy analysis of the multi-user coding schemes are far more complicated than the single-user ones. Furthermore, in order to derive upper-bounds that assures the optimality of these coding schemes, new converse techniques are required.

In the following subsections, the two mentioned communication scenarios are briefly introduced along with their connections to the non-secrecy domain and the most important results established for each model in literature. Moreover, the basic concepts of the achievablity coding schemes used to derive these results are introduced. In

particular, the encoding-decoding mechanism for each scheme is addressed including a shorten version of the corresponding reliability and secrecy analysis. It is important to highlight that most of the results presented in literature for the two-receiver DM-WBC were established under the weak secrecy criterion.

2.4.2 Public-Confidential Degraded Message Sets

In this section, a communication scenario that investigates the transmission of a common public message M_0 and a confidential message M_1 over a DM-WBC with two legitimate receivers and an eavesdropper as shown in Fig. 2.5 is considered. The common public message needs to be decoded by the three receiving nodes: both legitimate receivers and the eavesdropper as well. On the other hand, the confidential message needs to be only decoded by the two legitimate receivers while keeping it secret from the eavesdropper. This scenario can be interpreted as an extension of the Csiszár and Körner's wiretap channel introduced in [10], where an additional legitimate receiver is introduced to the system.

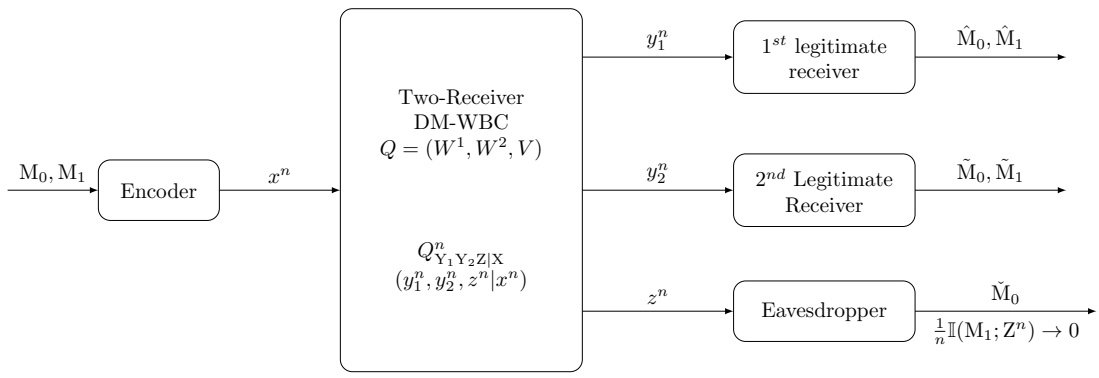


Figure 2.5: Two-receiver DM-WBC with public and confidential messages

The communication scenario modeled in Fig. 2.5 reflects a real life communication setup over DM-BCs in which public and confidential services need to be integrated together. The eavesdropper in this setup is not an intruding receiving node, but it is part of the system and some of the information transmitted over the channel are in fact addressed to it. In fact, having the eavesdropper to be part of the system is more convenient specially under the assumption that perfect CSI for all receiving nodes (legitimate and eavesdropper) is available at the transmitter. This model was first investigated in [64] under the following reliability and secrecy constraints:

$$\lim_{n\to\infty} \mathbb{P}\left[(\mathrm{M}_0, \mathrm{M}_1) \neq (\hat{\mathrm{M}}_0, \hat{\mathrm{M}}_1) \text{ or } (\mathrm{M}_0, \mathrm{M}_1) \neq (\tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_1) \text{ or } \mathrm{M}_0 \neq \check{\mathrm{M}}_0 | \mathcal{C}\right] = 0 \tag{2.74}$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_1; \mathrm{Z}^n|\mathcal{C}) = 0. \tag{2.75}$$

Eq. (2.75) implies that the investigation of this model in [64] and the extended version in [12] only considered the weak secrecy measure.

Non-Secrecy Connection: The communication scenario in Fig. 2.5 can be linked to the non-secrecy communication model of the DM-BC with degraded message sets which was initially introduced by Körner and Marton in [65]. In this work the transmission of two degraded message sets over a DM-BC with two receivers was considered.

It was shown that superposition encoding is optimal in the sense that it establishes the transmission capacity for the general DM-BC. In [66,67], the investigation was extended to the three-receiver DM-BC with two degraded message sets, where a common message is transmitted to all three receivers, while a private message is transmitted to only two of them. Unfortunately, the transmission capacity of this model is still an open problem. However, it was shown that the straightforward extension of the Körner and Marton inner bound given by [68, Bound 1] is optimal for many special cases. Moreover, it was shown in [68] that the concept of indirect decoding introduced by Nair and El Gamal in [69] does not yield any region better than Körner and Marton's inner-bound. It is important to highlight here that the Csiszár and Körner's encoding scheme introduced in [10] is in fact an adaptation of the Körner and Marton's coding scheme introduced in [65] to the secrecy domain.

Important Results: Since the transmission capacity of the corresponding non-secrecy model is still unknown, one should expect that the secrecy capacity of the communication scenario in Fig. 2.5 is still unknown as well. Nevertheless, various acheivable regions have been suggested, among them is the straightforward extension of the Csiszár and Körner's inner-bound given by the following theorem:

Theorem 2.4. *[12] An achievable weak secrecy rate region for transmitting a common public message and a confidential message over the DM-WBC with two legitimate receivers and an eavesdropper is given by the set of all rate pairs* $(R_0, R_1) \in \mathbb{R}_+^2$ *that satisfy*

$$\begin{aligned} R_0 &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\ R_1 &\leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}), \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\right], \end{aligned} \tag{2.76}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{UX}} = \mathrm{P}_{\mathrm{U}}\mathrm{P}_{\mathrm{X|U}}$*, such that* $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Moreover, if both legitimate receivers are less noisy than the eavesdropper, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Z}$ *and* $\mathrm{Y}_2 \succeq \mathrm{Z}$*, the established region is in fact the secrecy capacity region.*

Although for the non-secrecy scenario indirect decoding does not have any advantage over the extended version of the Körner and Marton's inner-bound, the situation is different for the secrecy scenario. In [64], Chia and El Gamal presented a counter example that illustrated the superiority of indirect decoding over the extended version of Csiszár and Körner's inner-bound. They suggested a more general acheivable region, which is presented in the following theorem:

Theorem 2.5. *[12] An achievable weak secrecy rate region for transmitting a common public message and a confidential message over the DM-WBC with two legitimate receivers and an eavesdropper is given by the set of all rate pairs* $(R_0, R_1) \in \mathbb{R}_+^2$ *that satisfy*

$$\begin{aligned} R_0 &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\ R_1 &\leq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}), \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{U})\right] \\ R_0 + R_1 &\leq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}), \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{U})\right] \\ R_0 + 2R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) - 2\mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) \\ R_0 + 2R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}) - 2\mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) \end{aligned} \tag{2.77}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{UV_0V_1V_2X}} = \mathrm{P}_{\mathrm{U}}\mathrm{P}_{\mathrm{V_0|U}}\mathrm{P}_{\mathrm{V_1V_2|V_0}}$ $\mathrm{P}_{\mathrm{X|V_0V_1V_2}}$ *such that* $\mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Moreover, the probability distribution must satisfy the following constraint:*

$$\mathbb{I}(\mathrm{V_1V2}; \mathrm{Z}|\mathrm{V}_0) \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Z}|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_2; \mathrm{Z}|\mathrm{V}_0) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0). \tag{2.78}$$

The previous region was established based on a coding scheme that combines superposition encoding [65], Marton coding [70], indirect decoding [71] along with the concept of wiretap random codes [2, 10]. Some of these coding techniques will be addressed in more details in the next sections. One can also shows that the rate region given in Theorem 2.4 follows as a special case from the one given in Theorem 2.5.

Common Confidential Message: Consider a simpler version of the communication scenario in Fig. 2.5, where only the common confidential message M_1 is to be transmitted over the channel. Thus, the reliability constraint in Eq. (2.74) changes to:

$$\lim_{n\to\infty} \mathbb{P}\left[\mathrm{M}_1 \neq \hat{\mathrm{M}}_1 \text{ or } \mathrm{M}_1 \neq \tilde{\mathrm{M}}_1 | \mathcal{C}\right] = 0, \tag{2.79}$$

while the secrecy constraint in Eq. (2.75) remains unchanged. This simpler scenario can be viewed as an extended version of the secrecy scenario introduced by Wyner in [2] to the DM-WBC with two legitimate receivers and an external eavesdropper. Since the common message secure broadcasting is a special case of the communication scenario in Fig. 2.5, where $\mathcal{M}_0 = \emptyset$, the results established in Theorem 2.5 and Theorem 2.4 can be easily extended to this setup by setting the random variable as follows: $\mathrm{U} = \emptyset$. Moreover, the following corollary is presented to spot the light on the usage of the concept of indirect decoding for secure communication.

Corollary 2.2. *[12] An achievable weak secrecy rate region for transmitting a common confidential message over the DM-WBC with two legitimate receivers and an eavesdropper is given by the set of all rates* $R \in \mathbb{R}_+$ *that satisfy*

$$R \leq \min\left[\mathbb{I}(\mathrm{X}; \mathrm{Y}_1) - \mathbb{I}(\mathrm{X}; \mathrm{Z}), \mathbb{I}(\mathrm{V}; \mathrm{Y}_2) - \mathbb{I}(\mathrm{V}; \mathrm{Z})\right], \tag{2.80}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{VX}} = \mathrm{P}_{\mathrm{V}}\mathrm{P}_{\mathrm{X|V}}$*, such that* $\mathrm{V} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The problem of secure broadcasting of a common message over a DM-WBC with two legitimate receiver and an eavesdropper is very important because it illustrates an important relation between the secrecy and non-secrecy domain. One can observe that this problem is just the secrecy adaptation of the public common message broadcasting over a two-receiver DM-BC introduced in [72]. The transmission capacity of a common public message over the multi-receiver DM-BC was established in [72]. Nevertheless, the secrecy capacity of the common confidential message problem even for the two-receiver DM-WBC is still an unsolved problem. This result is somehow surprising, however it highlights the fact that establishing the secrecy capacity for communication scenarios over two-receiver DM-WBC is a very difficult task.

2.4.3 Secure Broadcasting of Individual Message Sets

This section investigates the secure transmission of two confidential messages M_1 and M_2 over a DM-WBC with two legitimate receivers and an eavesdropper as shown

in Fig. 2.6. Each legitimate receiver is only interested in one message, while both message need to be kept secret from the eavesdropper. This secrecy scenario was first considered in [73, 74] under the following reliability and secrecy constraints:

$$\lim_{n\to\infty} \mathbb{P}\left[\mathrm{M}_1 \neq \hat{\mathrm{M}}_1 \text{ or } \mathrm{M}_2 \neq \tilde{\mathrm{M}}_2 | \mathcal{C}\right] = 0 \tag{2.81}$$

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{I}(\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n | \mathcal{C}) = 0. \tag{2.82}$$

The investigation was then extended to the multi-receiver degraded DM-WBC in [75, 76]. Some earlier work also considered a version of the problem over parallel WBCs in [77] and fading WBCs in [77, 78].

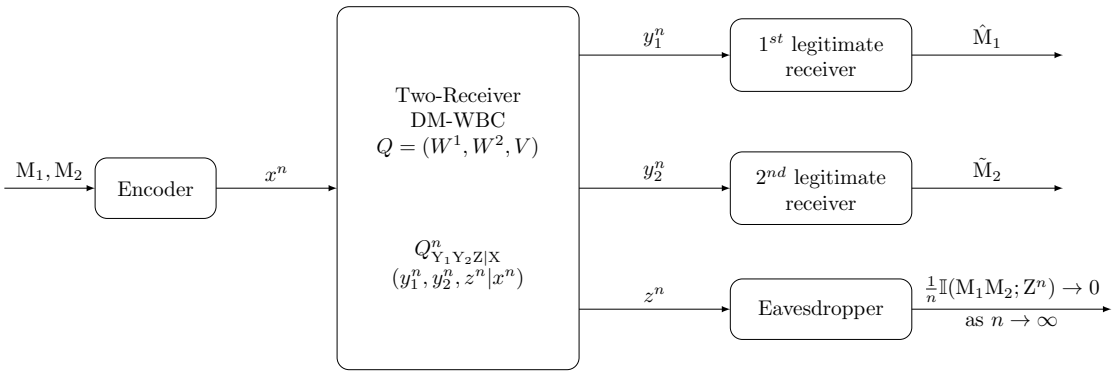


Figure 2.6: Two-receiver DM-WBC with two individual confidential messages

Non-Secrecy Connection: Obviously, this secrecy model is a direct extension of the classical two-receiver DM-BC with two individual public messages [20]. Although this broadcasting problem have been addressed in many literature cf. [56] and references therein, its transmission capacity is still unknown. The best acheivable rate region is based on a coding scheme that combines the basics of superposition encoding introduced in [79] and the principles of Marton encoding proposed in [70] as follows:

Theorem 2.6. *[35, Proposition 8.1] An achievable rate region for transmitting two individual public messages over the DM-BC with two receivers is given by the set of all rate pair* (R_1, R_2) *that satisfy*

$$\begin{aligned} R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1) \\ R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2) \\ R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2; \mathrm{Y}_2|\mathrm{V}_0) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) \\ R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_1; \mathrm{Y}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) \end{aligned} \tag{2.83}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2\mathrm{X}} = \mathrm{P}_{\mathrm{V}_0}\mathrm{P}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\,\mathrm{P}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}$ *such that* $\mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2)$ *forms a Markov chain.*

This previous rate region is tight for all classes of DM-BCs for which the transmission capacity is known such as the degraded DM-BC [80, 81], the less noisy DM-BC [21, 82], the more capable DM-BC [83] and many other channels cf. [56]. It was argued in [35] that a rate region established using Marton coding only is not optimal in general and a superposition variable is usually need to assure the optimality.

3. Important Results: Since the capacity of transmitting two public individual messages over the DM-BC is still an open problem, one should expect that deriving a secrecy capacity result for the DM-WBC is much harder. Although this is true, researchers suggested various achievable rate regions such as the one introduced in [73] and is given by the following theorem:

Theorem 2.7. *An achievable weak secrecy rate region for transmitting a two individual confidential messages over the DM-WBC with two legitimate receivers and an eavesdropper is given by the set of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy*

$$\begin{aligned} R_1 &\leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_1;\mathrm{Z}) \\ R_2 &\leq \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_2;\mathrm{Z}) \\ R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2), \end{aligned} \tag{2.84}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{V}_1\mathrm{V}_2\mathrm{X}} = \mathrm{P}_{\mathrm{V}_1\mathrm{V}_2}\mathrm{P}_{\mathrm{X}|\mathrm{V}_1\mathrm{V}_2}$ *such that* $(\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous region follows by extending the concept of wiretap random codes to Marton coding. The work in [73] did not only provide the previous achievable rate region, but it proved the following secrecy capacity result:

Theorem 2.8. *The weak secrecy capacity region for transmitting two individual confidential message over the full degraded DM-WBC with two legitimate receivers and an eavesdropper, where full degraded implies that* $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ *forms a Markov chain is given by the convex hull of the closure of all rate pairs* (R_1, R_2) *that satisfy:*

$$\begin{aligned} R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\ R_2 &\leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}), \end{aligned} \tag{2.85}$$

for random variables with joint probability distribution $\mathrm{P}_{\mathrm{UX}} = \mathrm{P}_{\mathrm{U}}\mathrm{P}_{\mathrm{X}|\mathrm{U}}$ *such that* $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous region shows that superposition wiretap random codes are optimal for the degraded two-receiver DM-WBC and Marton coding is not needed. This result was extended to the multi-receiver degraded DM-WBC in [75, 76]. It is important to highlight here that the secrecy capacity region in Theorem 2.8 can not be derived from the achievable region given in Theorem 2.7. This because as illustrated by Theorem 2.6, Marton-coding are only optimal for the two-receiver DM-BC if they are constructed with a superposition variable.

Extended Scenario: In [84] a communication scenario that combined the transmission of degraded message sets along with public individual message sets and confidential individual messages sets over a DM-WBC with two legitimate receivers and an eavesdropper was investigated. A general achievable region was established along with the capacity region for two special scenarios of the degraded two-receiver DM-WBC. The first scenario is when no confidential message is intended to the first legitimate receiver (stronger receiver), while the second scenario is when no public message is intended for the second legitimate receiver. Surprisingly, the secrecy capacity was not established for the general scenario, even under the degradedness assumption.

2.4.4 Superposition Wiretap Random Codes

Superposition encoding is a famous coding technique in which the transmitted information is structured in overlapping layers [85]. The concept of superposition appears to be very suitable for the transmission of degraded message sets where there exists some sort of order for the decoding requirements of each receiver [65]. Moreover, superposition can also be useful for the scenarios in which there exists an order for the decoding capabilities of each receiver such as degraded and less noisy channels [80,82]. The work of Csiszár and Körner in [10] was the first one to combine the concept of superposition encoding with wiretap random codes. However, this work only applied the wiretap random codes on the upper layer of the superposition codes only. It was shown in [75] that it can be extended to the various layers of superposition codes. In order to highlight this result, the achievability proof of Theorem 2.8 is presented.

1. Codebook Generation $\mathcal{C}$: Fix an input distribution P_{UX}, then let $\mathcal{M}_{r_1} = [\![1, 2^{nR_{r_1}}]\!]$ and $\mathcal{M}_{r_2} = [\![1, 2^{nR_{r_2}}]\!]$ be two randomization message sets. Next, generate the cloud centers codewords $u^n(m_2, m_{r_2})$, for $m_2 \in \mathcal{M}_2$ and $m_{r_2} \in \mathcal{M}_{r_2}$ by generating the symbols $u_i(m_2, m_{r_2})$ with $i \in [\![1, n]\!]$ independently at random according to P_{U}. Finally, for every codeword $u^n(m_2, m_{r_2})$ construct the satellite codewords $x^n(m_2, m_{r_2}, m_1, m_{r_1})$ for $m_1 \in \mathcal{M}_1$ and $m_{r_1} \in \mathcal{M}_{r_1}$ by generating the symbols $x_i(m_2, m_{r_2}, m_1, m_{r_1})$ independently at random according to $\mathrm{P}_{\mathrm{X|U}}(\cdot|u_i(m_2, m_{r_2}))$.

2. Encoding E: Given the confidential messages pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, the encoder selects two randomization messages m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. After that, it transmits the corresponding satellite codeword $x^n(m_2, m_{r_2}, m_1, m_{r_1})$.

3. First Legitimate Receiver Decoder φ_1: Upon receiving y_1^n, the decoder determines a unique messages quadruple $(\hat{m}_2, \hat{m}_{r_2}, \hat{m}_1, \hat{m}_{r_1})$ such that:

$$\Big(u^n(\hat{m}_2, \hat{m}_{r_2}), x^n(\hat{m}_2, \hat{m}_{r_2}, \hat{m}_1, \hat{m}_{r_1}), y_1^n\Big) \in \mathcal{T}_\epsilon^n\Big(\mathrm{P}_{\mathrm{UX}} W^1_{\mathrm{Y_1|X}}\Big).$$

The previous decoding strategy implies that the first legitimate receiver does not only decode its intended message m_1, but it also decodes the other confidential message m_2. Not to mention that, it decodes both randomization messages m_{r_1} and m_{r_2} as well.

4. Second Legitimate Receiver Decoder φ_2: This receiver is only interested in the second confidential message and consequently it aims to find the corresponding cloud center codeword. Thus, upon receiving y_2^n, the decoder determines a unique pair $(\tilde{m}_2, \tilde{m}_{r_2})$ such that for $W^2_{\mathrm{Y_2|U}} = \sum_x W^2_{\mathrm{Y_2|X}} \mathrm{P}_{\mathrm{X|U}}$, it holds that:

$$\Big(u^n(\tilde{m}_2, \tilde{m}_{r_2}), y_2^n\Big) \in \mathcal{T}_\epsilon^n\Big(\mathrm{P}_{\mathrm{U}} W^2_{\mathrm{Y_2|U}}\Big).$$

5. Reliability Analysis: Based on the definitions of the decoders at the two legitimate receivers, one can evaluate the reliability performance of this coding scheme using the following average error probability:

$$\bar{\mathbf{P}}_e(\mathcal{C}) = \mathbb{P}\Big[(\mathrm{M}_1, \mathrm{M}_2, \mathrm{M}_{r_1}, \mathrm{M}_{r_2}) \neq (\hat{\mathrm{M}}_1, \hat{\mathrm{M}}_2, \hat{\mathrm{M}}_{r_1}, \hat{\mathrm{M}}_{r_2}) \text{ or } (\mathrm{M}_2, \mathrm{M}_{r_2}) \neq (\tilde{\mathrm{M}}_2, \tilde{\mathrm{M}}_{r_2})|\mathcal{C}\Big]. \tag{2.86}$$

In order to fulfill the reliability constraint in (2.81), it is enough to show that $\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n$,where $\lim_{n\to\infty} \lambda_n = 0$. Now, based on the channel coding theorem, the principle of superposition encoding and the standard joint-typicality decoding arguments, one can show that the bound on the average error probability holds, if

$$R_2 + R_{r_2} \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_1) - \delta_1(\epsilon) \tag{2.87a}$$
$$R_1 + R_{r_1} \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \delta_1(\epsilon) \tag{2.87b}$$
$$R_2 + R_{r_2} \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \delta_2(\epsilon). \tag{2.87c}$$

Moreover, the typicality constant ϵ can be selected in a way such that $\delta_1(\epsilon), \delta_2(\epsilon) \to 0$ as $n \to \infty$. Furthermore, the bound in (2.87a) can be neglected because it is already included in (2.87c) due to the fact that the channel is fully degraded, i.e. $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ forms a Markov chain which implies that $\mathbb{I}(\mathrm{U};\mathrm{Y}_1) \geq \mathbb{I}(\mathrm{U};\mathrm{Y}_2)$.

6. Virtual Receiver Decoder φ_3: This receiver is introduced to facilitate the secrecy analysis of this coding scheme. The idea is to extend the virtual receiver 's technique presented in section 2.2.2 to the superposition encoding principle as follows: Given z^n and the transmitted confidential messages (m_1, m_2), the decoder determines a unique randomization messages pair $(\check{m}_{r_1}, \check{m}_{r_2})$ such that:

$$\Big(u^n(m_2, \check{m}_{r_2}), x^n(m_2, \check{m}_{r_2}, m_1, \check{m}_{r_1}), z^n\Big) \in \mathcal{T}_\epsilon^n\Big(\mathrm{P}_{\mathrm{UX}} V_{\mathrm{Z}|\mathrm{X}}\Big).$$

7. Secrecy Analysis: The first step in the secrecy analysis is to derive the required rate constraints that assure a vanishing average error probability at the virtual receiver's decoder. Let this error probability be defined as follows:

$$\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) = \mathbb{P}\left[(\mathrm{M}_{r_1}, \mathrm{M}_{r_2}) \neq (\check{\mathrm{M}}_{r_1}, \check{\mathrm{M}}_{r_2})|\mathcal{C}\right] \tag{2.88}$$

Using the standard joint-typicality decoding arguments, one can show that $\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) \leq \tau_n$,where $\lim_{n\to\infty} \tau_n = 0$ if:

$$R_{r_2} \leq \mathbb{I}(\mathrm{U};\mathrm{Z}) - \delta_3(\epsilon) \tag{2.89a}$$
$$R_{r_1} \leq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) - \delta_3(\epsilon), \tag{2.89b}$$

where again ϵ can be selected in a way such that $\lim_{n\to\infty} \delta_3(\epsilon) = 0$. Now, based on the structure of the codebook, the information leakage of the two confidential messages can be reformulated as follows:

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_1\mathrm{M}_2;\mathrm{Z}^n|\mathcal{C}) &= \mathbb{I}(\mathrm{M}_1\mathrm{M}_2\mathrm{M}_{r_1}\mathrm{M}_{r_2};\mathrm{Z}^n|\mathcal{C}) - \mathbb{I}(\mathrm{M}_{r_1}\mathrm{M}_{r_2};\mathrm{Z}^n|\mathrm{M}_1\mathrm{M}_2\mathcal{C}) \\
&\overset{(a)}{=} \mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n|\mathcal{C}) - \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\mathrm{M}_1\mathrm{M}_2\mathcal{C}) + \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\mathrm{M}_1\mathrm{M}_2\mathrm{Z}^n\mathcal{C}) \\
&\overset{(b)}{=} n\mathbb{I}(\mathrm{X};\mathrm{Z}) - n(R_{r_1} + R_{r_2}) + \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\check{\mathrm{M}}_{r_1}\check{\mathrm{M}}_{r_2}\mathcal{C}) \\
&\overset{(c)}{\leq} 2n\delta_3(\epsilon) + 1 + n\tau_n(R_{r_1} + R_{r_2}).
\end{aligned} \tag{2.90}$$

(a) follows due to the codebook structure which indicates that X^n is fully identified by the random variables $\mathrm{M}_1, \mathrm{M}_2, \mathrm{M}_{r_1}$ and M_{r_2}; (b) follows because M_{r_1} and M_{r_2} are uniformly distributed over their corresponding sets along with the definition of the virtual receiver decoder and the fact that$(\mathrm{X}^n, \mathrm{Z}^n)$ is i.i.d. according to $\mathrm{P}_{\mathrm{X}} V_{\mathrm{Z}|\mathrm{X}}$; while (c) follows due to a relaxed version of Fano's inequality [22] in addition to applying the rate constraints in (2.89a) and (2.89b) with equality, keeping in mind that

$\mathbb{I}(\mathrm{X};\mathrm{Z}) = \mathbb{I}(\mathrm{U};\mathrm{Z}) + \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})$. Recalling the weak secrecy constraint in (2.82) yields the following:

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_1\mathrm{M}_2;\mathrm{Z}^n|\mathcal{C}) \leq \lim_{n\to\infty} 2\delta_3(\epsilon) + \frac{1}{n} + \tau_n(R_{r_1} + R_{r_2}) = 0. \tag{2.91}$$

This implies that the previous coding scheme along with the rate constraints in (2.89) fulfill of the weak secrecy measure in (2.82).

Finally, by applying the Fourier-Motzkin elimination procedure [57] to the rate constraints in (2.87) and (2.89) followed by taking the limit as $n \to \infty$, it follows directly that any rate pair (R_1, R_2) that satisfies the constraints in (2.85) is achievable.

2.4.5 Marton-Coding Wiretap Random Codes

As highlighted in the previous section, superposition encoding is based on some hierarchy regarding the decoding requirements or the decoding capabilities of the receivers. This implies that for the general two-receiver DM-BC where an individual message is transmitted to each receiver, superposition might not be a suitable encoding technique. This idea encouraged Katalin Marton to introduce a new coding scheme for the general DM-BC, which is now known as Marton-Coding [70]. This coding scheme is based on a combination of the coding techniques used in [79, 86] together with the random coding technique used to prove source coding theorems in rate-distortion theory [87]. In [88] El-Gamal and Meulen presented a simpler proof for the concept of Marton-Coding. Further, El-Gamal argued in [35] that Marton codes with a superposition variable given by Theorem 2.6 is the best inner bound to the capacity region of the DM-BC.

In [73] the principle of Marton-Coding was combined with the concept of wiretap random codes to derive an achievable region for the general two-receiver DM-WBC. This section investigates this new class of codes by presenting a detailed proof for Theorem 2.7 as follows:

1. Codebook Generation $\mathcal{C}$: Fix an input distribution $\mathrm{P}_{\mathrm{V}_1\mathrm{V}_2\mathrm{X}}$, then let $\mathcal{M}_{r_1} = [\![1, 2^{nR_{r_1}}]\!]$ and $\mathcal{M}_{r_2} = [\![1, 2^{nR_{r_2}}]\!]$ be two randomization message sets. Moreover, let $\mathcal{M}_{t_1} = [\![1, 2^{nR_{t_1}}]\!]$ and $\mathcal{M}_{t_2} = [\![1, 2^{nR_{t_2}}]\!]$ be two additional message sets needed for the Marton encoding. Next, generate the codewords $v_1^n(m_1, m_{r_1}, m_{t_1})$ for $m_1 \in \mathcal{M}_1$, $m_{r_1} \in \mathcal{M}_{r_1}$ and $m_{t_1} \in \mathcal{M}_{t_1}$ by generating the symbols $v_{1i}(m_1, m_{r_1}, m_{t_1})$ with $i \in [\![1, n]\!]$ independently at random according to $\mathrm{P}_{\mathrm{V}_1}$. At the same time, generate the codewords $v_2^n(m_2, m_{r_2}, m_{t_2})$ for $m_2 \in \mathcal{M}_2$, $m_{r_2} \in \mathcal{M}_{r_2}$ and $m_{t_2} \in \mathcal{M}_{t_2}$ by generating the symbols $v_{2i}(m_2, m_{r_2}, m_{t_2})$ independently at random according to $\mathrm{P}_{\mathrm{V}_2}$.

2. Encoding E: Given the confidential messages pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, the encoder selects two randomization messages m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. Then, it searches for a message pair (m_{t_1}, m_{t_2}) such that:

$$\Big(v_1^n(m_1, m_{r_1}, m_{t_1}), v_2^n(m_2, m_{r_2}, m_{t_2})\Big) \in \mathcal{T}_\epsilon^n\Big(\mathrm{P}_{\mathrm{V}_1\mathrm{V}_2}\Big).$$

Based on the properties of typical sequences cf. Theorem A.7, with a probability that approaches one as n approaches infinity, the encoder succeeds in finding a typical pair as long as:

$$R_{t_1} + R_{t_2} \geq \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2) + \delta_t(\epsilon), \tag{2.92}$$

where ϵ can be selected in away such that $\lim_{n\to\infty}\delta_t(\epsilon)=0$. In case more than one jointly typical pair exist, the encoder chooses one of them uniformly at random. Finally, for the selected codewords pair (v_1^n, v_2^n), the encoder generates a sequence x^n independently at random according to $\prod_{i=1}^n \mathrm{P}_{\mathrm{X}|\mathrm{V}_1\mathrm{V}_2}(\cdot|v_{1i}v_{2i})$ and transmits it.

3. First Legitimate Receiver Decoder φ_1: Upon receiving y_1^n, the decoder determines a unique messages triple $(\hat{m}_1, \hat{m}_{r_1}, \hat{m}_{t_1})$ such that for $W^1_{\mathrm{Y}_1|\mathrm{V}_1}=\sum_{x,v_2} W^1_{\mathrm{Y}_1|\mathrm{X}}\mathrm{P}_{\mathrm{X}|\mathrm{V}_1\mathrm{V}2}$, it holds that:

$$\left(v_1^n(\hat{m}_1, \hat{m}_{r_1}, \hat{m}_{t_1}), y_1^n\right) \in \mathcal{T}_\epsilon^n\left(\mathrm{P}_{\mathrm{V}_1}W^1_{\mathrm{Y}_1|\mathrm{V}_1}\right).$$

4. Second Legitimate Receiver Decoder φ_2: Upon receiving y_2^n, the decoder determines a unique messages triple $(\tilde{m}_2, \tilde{m}_{r_2}, \tilde{m}_{t_2})$ such that for $W^2_{\mathrm{Y}_2|\mathrm{V}_2}=\sum_{x,v_1} W^2_{\mathrm{Y}_2|\mathrm{X}}\mathrm{P}_{\mathrm{X}|\mathrm{V}_1\mathrm{V}2}$, it holds that:

$$\left(v_2^n(\tilde{m}_2, \tilde{m}_{r_2}, \tilde{m}_{t_2}), y_2^n\right) \in \mathcal{T}_\epsilon^n\left(\mathrm{P}_{\mathrm{V}_2}W^2_{\mathrm{Y}_2|\mathrm{V}_2}\right). \tag{2.93}$$

5. Reliability Analysis: Based on the definitions of the decoders at the two legitimate receivers and the encoding mechanism, one can evaluate the reliability performance of this coding scheme using the following average error probability:

$$\bar{\mathbf{P}}_e(\mathcal{C}) = \mathbb{P}\left[(\mathrm{M}_1, \mathrm{M}_{r_1}, \mathrm{M}_{t_1}) \neq (\hat{\mathrm{M}}_1, \hat{\mathrm{M}}_{r_1}, \hat{\mathrm{M}}_{t_1}) \text{ or } (\mathrm{M}_2, \mathrm{M}_{r_2}, \mathrm{M}_{t_2}) \neq (\tilde{\mathrm{M}}_2, \tilde{\mathrm{M}}_{r_2}, \tilde{\mathrm{M}}_{t_2})|\mathcal{C}\right]. \tag{2.94}$$

One can observe that if $\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n$, where $\lim_{n\to\infty}\lambda_n = 0$, then the reliability constraint in (2.81) follows directly. Now, using the reliability analysis of Marton codes presented in [88] along with the standard joint-typicality decoding arguments, one can show that $\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n$ holds, if

$$R_1 + R_{r_1} + R_{t_1} \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1) - \delta_1(\epsilon) \tag{2.95a}$$

$$R_2 + R_{r_2} + R_{t_2} \leq \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2) - \delta_2(\epsilon) \tag{2.95b}$$

As highlighted before, the typicality constant ϵ can be selected in a way such that $\lim_{n\to\infty}\delta_1(\epsilon),\ \delta_2(\epsilon)=0$.

6. Virtual Receiver Decoder φ_3: This decoder is introduced as a necessary part for the secrecy analysis. It works as follows: Given z^n and the transmitted confidential messages (m_1, m_2), the decoder determines a unique messages quadruple $(\check{m}_{r_1}, \check{m}_{r_2}, \check{m}_{t_1}, \check{m}_{t_2})$ such that for $V_{\mathrm{Z}|\mathrm{V}_1\mathrm{V}_2}=\sum_x V_{\mathrm{Z}|\mathrm{X}}\mathrm{P}_{\mathrm{X}|\mathrm{V}_1\mathrm{V}_2}$, it holds that:

$$\left(v_1^n(\check{m}_1, \check{m}_{r_1}), v_2^n(\check{m}_2, \check{m}_{r_2}), z^n\right) \in \mathcal{T}_\epsilon^n\left(\mathrm{P}_{\mathrm{V}_1\mathrm{V}_2}V_{\mathrm{Z}|\mathrm{V}_1\mathrm{V}_2}\right).$$

7. Secrecy Analysis: As usual, the first step in the secrecy analysis using the virtual receiver technique is to derive the bounds on the rates, such that the probability of error at the virtual receiver's decoder vanishes. Thus, one needs to define the following average error probability:

$$\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) = \mathbb{P}\left[(\mathrm{M}_{r_1}, \mathrm{M}_{r_2}, \mathrm{M}_{t_1}, \mathrm{M}_{t_2}) \neq (\check{\mathrm{M}}_{r_1}, \check{\mathrm{M}}_{r_2}, \check{\mathrm{M}}_{t_1}, \check{\mathrm{M}}_{t_2})|\mathcal{C}\right] \tag{2.96}$$

Based on the standard joint-typicality decoding arguments and the ides used to prove [12, Lemma 1], it follows that $\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) \leq \tau_n$, where $\lim_{n\to\infty}\tau_n = 0$, as long as

$$R_{r_1} + R_{t_1} \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Z}\mathrm{V}_2) - \delta_3(\epsilon) \tag{2.97a}$$

$$R_{r_2} + R_{t_2} \leq \mathbb{I}(\mathrm{V}_2; \mathrm{Z}\mathrm{V}_1) - \delta_3(\epsilon) \tag{2.97b}$$
$$R_{r_1} + R_{r_2} + R_{t_1} + R_{t_2} \leq \mathbb{I}(\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}) + \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2) - \delta_3(\epsilon), \tag{2.97c}$$

where again ϵ can be selected in a way such that $\lim_{n\to\infty} \delta_3(\epsilon) = 0$. Now, based on the structure of the codebook and the encoding procedure, the information leakage of the two confidential messages can be reformulated as follows:

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) &= \mathbb{I}(\mathrm{M}_1\mathrm{M}_2\mathrm{M}_{r_1}\mathrm{M}_{r_2}; \mathrm{Z}^n|\mathcal{C}) - \mathbb{I}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}; \mathrm{Z}^n|\mathrm{M}_1\mathrm{M}_2\mathcal{C}) \\
&\overset{(a)}{\leq} \mathbb{I}(\mathrm{V}_1^n\mathrm{V}_2^n; \mathrm{Z}^n|\mathcal{C}) - \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\mathrm{M}_1\mathrm{M}_2\mathcal{C}) + \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\mathrm{M}_1\mathrm{M}_2\mathrm{Z}^n\mathcal{C}) \\
&\overset{(b)}{\leq} n\mathbb{I}(\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}) + n\gamma_n - n(R_{r_1} + R_{r_2}) + \mathbb{H}(\mathrm{M}_{r_1}\mathrm{M}_{r_2}|\check{\mathrm{M}}_{r_1}\check{\mathrm{M}}_{r_2}\mathcal{C}) \\
&\overset{(c)}{\leq} n\delta_3(\epsilon) + n\delta_t(\epsilon) + n\gamma_n + 1 + n\tau_n(R_{r_1} + R_{r_2}).
\end{aligned} \tag{2.98}$$

(a) follows due to the codebook structure; (b) follows by using the properties of DMCs as shown in [73, Section III], where $\lim_{n\to\infty} \gamma_n = 0$; while (c) follows from Fano's inequality in addition to applying the rate constraints in (2.92) and (2.97c) with equality. The previous analysis implies that the previous coding scheme fulfills the weak secrecy constraint in (2.82) as follows:

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) \leq \lim_{n\to\infty} \delta_3(\epsilon) + \delta_t(\epsilon) + \gamma_n + \frac{1}{n} + \tau_n(R_{r_1} + R_{r_2}) = 0. \tag{2.99}$$

One should notice that the rate constraints in (2.92) and (2.97c) with equality directly implies that $\mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{Z}) = 0$. This, indicates that the rate constraints in (2.97a) and (2.97b) can be expressed as follows:

$$R_{r_1} + R_{t_1} \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Z}) - \delta_3(\epsilon) \tag{2.100a}$$
$$R_{r_2} + R_{t_2} \leq \mathbb{I}(\mathrm{V}_2; \mathrm{Z}) - \delta_3(\epsilon) \tag{2.100b}$$

Now, by applying the Fourier-Motzkin elimination procedure to the rate constraints in (2.95), (2.100), along with using Eq. (2.92) and Eq. (2.97c) with equality, then taking the limit as $n \to \infty$, it follows directly that any rate pair (R_1, R_2) that satisfies the constraints in (2.84) is achievable.

2.4.6 Indirect Decoding for Wiretap Random Codes

Unlike superposition encoding and Marton-Coding, indirect decoding is a technique that can be used by the decoders at the receiving ends and not during encoding. The principle of indirect decoding was introduced by Nair and El-Gamal in [71]. They considered a three-receiver DM-BC with two degraded message sets such that a common message is sent to all three receivers, while a private message is sent to only one receiver. They showed that for this scenario indirect decoding establishes an inner bound which is strictly larger than the straightforward extension of Körner and Marton inner bound [65]. In [12], El-Gamal and Chia extended the usage of indirect decoding to the problem of transmitting a confidential message over a two-receiver DM-WBC. They showed that for such scenario indirect decoding is better than the the straightforward extension of Csiszár and Körner inner bound in [10]. In order to investigate the deployment of indirect decoding in wiretap random codes, a detailed proof for Corollary 2.2 is presented as follows:

1. Codebook Generation $\mathcal{C}$: Fix an input distribution P_{VX} then let $\mathcal{M}_{r_1} = [\![1, 2^{nR_{r_1}}]\!]$ and $\mathcal{M}_{r_2} = [\![1, 2^{nR_{r_2}}]\!]$ be two randomization message sets. Next, construct the codewords $v^n(m, m_{r_1})$ for $m \in \mathcal{M}$ and $m_{r_1} \in \mathcal{M}_{r_1}$ by generating the symbols $v_i(m, m_{r_1})$ with $i \in [\![1, n]\!]$ independently at random according to P_{V}. Finally, for every codeword $v^n(m, m_{r_1})$ construct the codewords $x^n(m, m_{r_1}, m_{r_2})$ where $m_{r_2} \in \mathcal{M}_{r_2}$ by generating the symbols $x_i(m, m_{r_1}, m_{r_2})$ independently at random according to $\mathrm{P}_{\mathrm{X|V}}(\cdot|v_i(m, m_{r_1}))$.

2. Encoding E: Given the confidential message $m \in \mathcal{M}$, the encoder selects two randomization messages m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. Then, it transmits the codeword $x^n(m, m_{r_1}, m_{r_2})$.

3. First Legitimate Receiver Decoder φ_1: This decoder demonstrates the main concept behind indirect decoding as follows: Instead of decoding the confidential message m and the first randomization message m_{r_1} directly from the corresponding codeword v^n, it decodes them indirectly from the codeword x^n in such a way: Upon receiving y_1^n, the decoder determines a unique pair $(\hat{m}, \hat{m}_{r_1})$ such that:

$$\left(v^n(\hat{m}, \hat{m}_{r_1}), x^n(\hat{m}, \hat{m}_{r_1}, m_{r_2}), y_1^n\right) \in \mathcal{T}_\epsilon^n\left(\mathrm{P}_{\mathrm{VX}} W^1_{\mathrm{Y_1|X}}\right),$$

for some $m_{r_2} \in \mathcal{M}_{r_2}$. This definition captures the fact that this decoder does not aim to decode m_{r_2} and is only interested in decoding m and m_{r_1}.

4. Second Legitimate Receiver Decoder φ_2: This decoder on the other hand decodes m and m_{r_1} directly from the corresponding codeword v^n as follows: Upon receiving y_2^n, the decoder determines a unique pair $(\tilde{m}, \tilde{m}_{r_1})$ such that for $W^2_{\mathrm{Y_2|V}} = \sum_x W^2_{\mathrm{Y_2|X}} \mathrm{P}_{\mathrm{X|V}}$, it holds that:

$$\left(v^n(\tilde{m}, \tilde{m}_{r_1}), y_2^n\right) \in \mathcal{T}_\epsilon^n\left(\mathrm{P}_{\mathrm{V}} W^2_{\mathrm{Y_2|V}}\right).$$

5. Reliability Analysis: Based on the decoding mechanism at the two legitimate receivers and the encoding function, the reliability performance of this coding scheme can be evaluated using the following average error probability:

$$\bar{\mathbf{P}}_e(\mathcal{C}) = \mathbb{P}\left[(\mathrm{M}, \mathrm{M}_{r_1}) \neq (\hat{\mathrm{M}}, \hat{\mathrm{M}}_{r_1}) \text{ or } (\mathrm{M}, \mathrm{M}_{r_1}) \neq (\tilde{\mathrm{M}}, \tilde{\mathrm{M}}_{r_1}) | \mathcal{C}\right]. \tag{2.101}$$

One can notice that if $\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n$, where $\lim_{n\to\infty} \lambda_n = 0$, then the reliability constraint in (2.79) is satisfied. Using the standard joint-typicality decoding arguments along with the reliability analysis of the indirect decoding used in [12, Section IV-A], one can show that $\bar{\mathbf{P}}_e(\mathcal{C})$ vanishes as n approaches infinity, if

$$R + R_{r_1} + R_{r_2} \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_1) - \delta_1(\epsilon) \tag{2.102a}$$

$$R + R_{r_1} \leq \mathbb{I}(\mathrm{V}; \mathrm{Y}_2) - \delta_2(\epsilon) \tag{2.102b}$$

6. Virtual Receiver Decoder φ_3: Given z^n and the transmitted confidential message m, the decoder determines a unique randomization messages pair $(\check{m}_{r_1}, \check{m}_{r_2})$ such that:

$$\left(v^n(m, \check{m}_{r_1}), x^n(m, \check{m}_{r_1}, \check{m}_{r_2}), z^n\right) \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{VX}} V_{\mathrm{Z|X}}).$$

7. Secrecy Analysis: The first step in the secrecy analysis is to derive the upper-bounds on the randomization rates for which the following average decoding error probability vanishes:

$$\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) = \mathbb{P}\left[(\mathrm{M}_{r_1}, \mathrm{M}_{r_2}) \neq (\check{\mathrm{M}}_{r_1}, \check{\mathrm{M}}_{r_2}) | \mathcal{C}\right] \tag{2.103}$$

One can observe that from the prospective of the virtual receiver, the encoding scheme can be interpreted as a superposition encoding as follows: The first randomization message m_{r_1} is encoded in the cloud center codeword v^n, while the second randomization message is encoded in the satellite codeword x^n. This implies that using the standard joint-typicality arguments along with the reliability analysis of superposition codes cf. [16], one can show that $\bar{\mathbf{P}}_e^{\mathrm{VR}}(\mathcal{C}) \leq \tau_n$, where $\lim_{n\to\infty} \tau_n = 0$ as long as:

$$R_{r_1} \leq \mathbb{I}(\mathrm{V};\mathrm{Z}) - \delta_3(\epsilon) \tag{2.104a}$$
$$R_{r_2} \leq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{V}) - \delta_3(\epsilon). \tag{2.104b}$$

On the other hand, based on the codebook generation and the encoding mechanism, the information leakage of the confidential message to the eavesdropper can be bounded as follows:

$$\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathcal{C}) \leq 2n\delta_3(\epsilon) + 1 + n\tau_n(R_{r_1} + R_{r_2}). \tag{2.105}$$

The previous bound follows using similar steps like the ones used to establish the bound in (2.90) by letting $\mathrm{M} = (\mathrm{M}_1, \mathrm{M}_2)$ in addition to using the rate constraints in (2.104) with equality. Since ϵ can be selected in a way such that $\lim_{n\to\infty} \delta_3(\epsilon) = 0$, the previous bound directly implies that the weak secrecy measure in (2.75) is satisfied.

Now, by applying the Fourier-Motzkin elimination procedure to the reliability constraints in (2.81) and the secrecy rate constraints in (2.104) with equality followed by taking the limit as $n \to \infty$, one can show that any confidential rate R that satisfies the bound in (2.80) is achievable.

2.4.7 Gaussian Wiretap Broadcast Channels

The additive Gaussian white noise channel is a DMC in which the channel output Y is simply the addition of the channel input X and a scalar noise N, where N is a zero-mean Gaussian random variable with variance σ independent from the channel input X [89]. The channel capacity of a Gaussian channel is finite only if the noise variance is non-zero and the channel input has a finite power constraint, otherwise Gaussian channels have an infinite capacity [90]. Gaussian channels are a very common model for various communication scenarios in real life specially wireless communication. There are two main classes of Gaussian channels: The first is known as the *scalar* Gaussian channel in which the transmitter as well as the receiver have one antenna; This class is also known as the Gaussian single-input single-output (SISO) channel. The second class is the Gaussian multiple-input multiple-output (MIMO) channel and is characterized by the presence of multiple antennas at the transmitting and receiving nodes.

In [11], secure communication over the Gaussian SISO wiretap channel was investigated. The authors managed to establish the secrecy capacity as follows: The achievability follows by letting the channel input X that describes the secrecy capacity of the degraded wiretap channel in [2, Theorem 2] to be a Gaussian random variable with variance P, where P is the total power available at the transmitter. On the other hand, the converse follows by using the entropy-power inequality [91, 92] to prove the optimality of Gaussian signalling. Now, consider the two-receiver Gaussian SISO-WBC given by:

$$\mathrm{Y}_1 = \mathrm{X} + \mathrm{N}_1, \qquad \mathrm{Y}_2 = \mathrm{X} + \mathrm{N}_2, \qquad \mathrm{Z} = \mathrm{X} + \mathrm{N}_z. \tag{2.106}$$

N_1, N_2 and N_z are zero-mean Gaussian random variables, whose variances are given by σ_1^2, σ_2^2 and σ_z^2 respectively. Moreover, the channel input X is subject to a power constraint such that $\mathbb{E}[\mathrm{X}^2] \leq P$. The secrecy capacity region of the previous channel under the weak secrecy constraint in (2.82) is given by the following theorem:

Theorem 2.9. *The weak secrecy capacity region of the two-receiver Gaussian SISO-WBC is given by the union of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy*

$$R_2 \leq f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_2^2}\right) - f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_\mathrm{z}^2}\right) \tag{2.107a}$$

$$R_1 \leq f\left(\frac{\alpha P}{\sigma_1^2}\right) - f\left(\frac{\alpha P}{\sigma_\mathrm{z}^2}\right) \tag{2.107b}$$

where the union is taken over all values of $\alpha \in [0, 1]$*, such that* $\bar{\alpha} = 1 - \alpha$ *and* P *is the maximal power available at the transmitter. Moreover, the function* $f(a)$ *is defined as follows:* $f(a) \triangleq \frac{1}{2}\log(1 + a)$.

The previous capacity region was established in [93, 94] for the multi-receiver Gaussian SISO-WBC and not only for the two-receiver case. The prove is based on the fact that the Gaussian SISO-WBC is a special case of degraded DM-WBC along with the optimality of Gaussian signaling for the class of Gaussian channels. The sketch of the proof is given in the next few lines.

1. Achievability: follows as a special case from Theorem 2.8 by selecting (U, X) to be jointly Gaussian, where $\mathrm{X} = \mathrm{U} + \mathrm{V}$ can be viewed as the summation of two independent zero-mean Gaussian random variables U and V, with respective variances $\bar{\alpha}P$ and αP.

2. The converse: follows by showing that Gaussian signalling for this channel is optimal. This step was a little bit challenging because the converse techniques used in [95] and [96] for the two-receiver Gaussian SISO-BC cannot be extended in easily to the wiretap case. This is because the entropy-power inequality [91, 92] which is the main tool of these techniques is not sufficient alone to solve the optimization problem in the presence of the eavesdropper. In [94] Ekrem and Ulukus suggested two possible solutions to overcome this issue: The first solution utilizes the relation between the minimum-mean-square-error (MMSE) and the mutual information [97, 98], while the second one is based on the relation between Fisher information and the differential entropy established via the De Bruijn identity [91, 92]. These two solution are somehow the same because the MMSE and fisher information have a one-to-one relation in the sense that one of them determines the other one [99]. On the other hand the converse proof suggested in [93] used a subsequent variation of the entropy-power inequality known as the Costa's entropy-power inequality [100].

In [94], Ekrem and Ulukus did not only establish the secrecy capacity of the two-receiver Gaussian SISO-WBC, but they extended their results to the general two-receiver Gaussian MIMO-WBC as well. In order to establish the capacity region of this channel, they consider two special cases of Gaussian MIMO-WBC at first. The first class is known as the two-receiver degraded Gaussian MIMO-WBC and is given by:

$$\mathbf{Y}_1 = \mathbf{X} + \mathbf{N}_1, \qquad \mathbf{Y}_2 = \mathbf{X} + \mathbf{N}_2, \qquad \mathbf{Z} = \mathbf{X} + \mathbf{N}_\mathrm{z}, \tag{2.108}$$

where $\mathbf{X}$, $\mathbf{Y}_1$, $\mathbf{N}_1$, $\mathbf{Y}_2$, $\mathbf{N}_2$, $\mathbf{Z}$ and $\mathbf{N}_Z$ are column vectors of length m, where m is the number of antennas available at the transmitter and each receiver. The channel input $\mathbf{X}$ is subject to a covariance constraint $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S}$, where $\mathbf{S} \succ \mathbf{0}$. $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_Z$ are zero-mean Gaussian random vectors, whose covariance matrices are given by $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_Z$, such that

$$\mathbf{0} \prec \mathbf{\Sigma}_1 \preceq \mathbf{\Sigma}_2 \preceq \mathbf{\Sigma}_Z. \tag{2.109}$$

The previous semi-definite ordering of the noise covariance matrices implies that $\mathbf{X} - \mathbf{Y}_1 - \mathbf{Y}_2 - \mathbf{Z}$ forms a Markov chain. This means that the two-receiver degraded Gaussian MIMO-WBC belongs to the class of two-receiver degraded DM-WBCs. The second class of channels is known as the two-receiver aligned Gaussian MIMO-WBC. This channel is also characterized by the same number of antennas at the transmitter and each receive but the noise covariance matrices do not have to satisfy any positive semi-definite order. The general two-receiver Gaussian MIMO-WBC simply drops the constraint on the number of antennas in the aligned class.

Ekrem and Ulukus used a very systematic technique to establish the secrecy capacity of the general two-receiver Gaussian MIMO-WBC as follows: First, they used the secrecy capacity of the two-receiver degraded DM-WBC given by Theorem 2.8 along with the converse technique used to prove the optimality of Gaussian signaling for the two-receiver Gaussian SISO-WBC to establish the secrecy capacity of the two-receiver degraded Gaussian MIMO-WBC as follows:

Theorem 2.10. *The weak secrecy capacity region of the two-receiver degraded Gaussian MIMO-WBC is given by the union of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy*

$$R_2 \leq \frac{1}{2}\log\frac{|\mathbf{S}+\mathbf{\Sigma}_2|}{|\mathbf{K}_1+\mathbf{\Sigma}_2|} - \frac{1}{2}\log\frac{|\mathbf{S}+\mathbf{\Sigma}_z|}{|\mathbf{K}_1+\mathbf{\Sigma}_z|} \tag{2.110a}$$

$$R_1 \leq \frac{1}{2}\log\frac{|\mathbf{K}_1+\mathbf{\Sigma}_2|}{|\mathbf{\Sigma}_1|} - \frac{1}{2}\log\frac{|\mathbf{K}_1+\mathbf{\Sigma}_z|}{|\mathbf{\Sigma}_z|} \tag{2.110b}$$

where the union is taken over all selections of $\mathbf{0} \preceq \mathbf{K}_1 \preceq \mathbf{S}$*, and* $\mathbf{S}$ *is a covariance constraint related to the total available at the transmitter. Moreover, the function* $|\mathbf{A}|$ *is simply the determinant of the matrix* $\mathbf{A}$*.*

The previous capacity region was also established in [93, 101] using a vector generalization of Costa's entropy-power inequality [100]. The next step in establishing the secrecy capacity of the general Gaussian MIMO-WBC is to generalize the region in Theorem 2.10 to the aligned case by using the channel enhancement technique [102]. In this step, Ekrem and Ulukus showed that dirty-paper coding with stochastic encoding is the optimal coding scheme. Finally, the secrecy capacity region of the two-receiver aligned Gaussian MIMO-WBC is extended to the general case using by some limiting arguments as shown in [102, 103], where again dirty-paper coding with stochastic encoding is proved to be optimal. It is worth mentioning that Ekrem and Ulukus manged to extend their results to the general multi-receiver Gaussian MIMO-WBC in [94].

The integration of public and confidential services over Gaussian MIMO-BC was considered in [104]. The model considered a communication scenario over a two-receiver Gaussian MIMO-BC with one common public message and two confidential messages.

The common message is intended for both receivers, while each of the confidential messages is intended only for one of the receivers and must be kept secret from the other one. The capacity region of this model was established in [104, 105] for the general two-receiver Gaussian MIMO-BC using similar arguments as the one discussed above.

Chapter 3

Wiretap Broadcast Channels with Degraded Message Sets and Message Cognition

This chapter investigates the problem of secure communication over a two-receiver DM-WBC with degraded message sets and message cognition. First, we introduce our model and discuss different reliability and secrecy requirements. In particular, we consider two secrecy constraints: a conservative constraint known as the joint secrecy and a more relaxed one known as the individual secrecy. For these two constraints, we derive a general achievable rate region that combines two main coding principles: superposition encoding and Marton-Coding along with the concept of indirect decoding under the strong secrecy criterion. Finally, we establish the secrecy capacity region for some classes of the two-receiver less noisy DM-WBC.

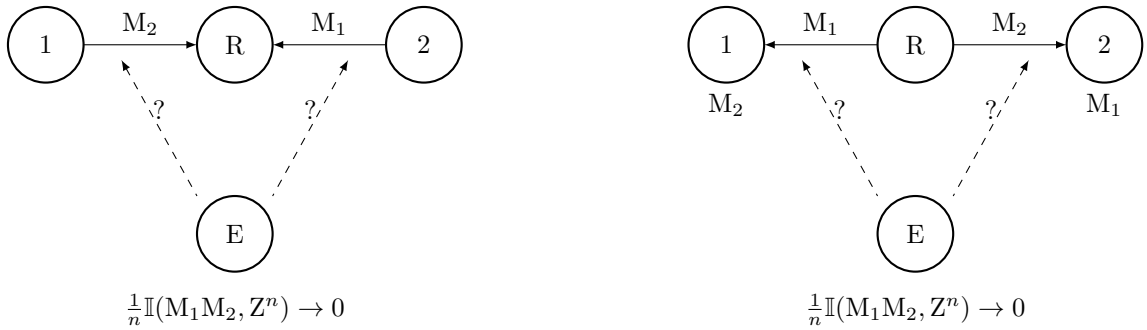
3.1 System Model and Secrecy Criteria

In this section, we introduce the problem of secure broadcasting over a two-receiver DM-WBC with degraded message sets and message cognition. We start by highlighting the problem of secure bidirectional relaying [106], which represents a real life communication scenario for our model. We then present a formal description for our model along with the required communication tasks. Finally, we discuss two different secrecy measures known as joint secrecy and individual secrecy. These two measures will be used to evaluate the secrecy of the communication over our model.

3.1.1 Motivation: Secure Bidirectional Relaying

Consider the problem of two-phase secure bidirectional relaying in a three-node network with an external eavesdropper, as shown in Fig. 3.1. In the first phase, known as the multiple access phase, node 1 and 2 transmit their messages to the relay while keeping the eavesdropper unable to intercept any information about the transmission. This problem corresponds to the problem of secure communication over the discrete memoryless multiple access wiretap channel (DM-MAWC) and was investigated in [107–110], where the latter discusses different secrecy criteria. In the succeeding phase, known as the broadcasting phase, the relay (node 3) re-encodes the messages and broadcasts them to the intended nodes, while assuring that the eavesdropper can not interpret the

transmitted information. This problem is somehow connected to the problem of secure communication over a two-way wiretap channel discussed in [111–113]. It can also be considered as a secrecy extension to the problem of public two-phase bidirectional relaying in a three-node network, where a relay node establishes a bidirectional communication between two other nodes using a decode-and-forward protocol [114–117].



(a) Multiple-Access Phase (b) Broadcasting Phase

Figure 3.1: Secure bidirectional relaying in three-nodes network

Our main focus in this chapter will be the broadcasting phase which was first studied in [116] without any secrecy constraint. It was shown that the transmission capacity region is given by the set of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following:

$$R_1 \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_1) \qquad \text{and} \qquad R_2 \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_2) \tag{3.1}$$

The previous capacity region follows directly from the capacity result of public broadcasting of a common message over a two-receiver DM-BC due to the equivalence of these two problems. This equivalence arises from the fact that each receiver is cognizant of its own message from the multiple access phase and can use it as an additional side information for decoding. Thus, a combined message set $\mathcal{M}_1 \times \mathcal{M}_2$ can be interpreted as a common message from the prospective of each receiver. In literature, the bidirectional broadcasting phase is also known as the DM-BC with receiver side information [116], or the DM-BC with full message cognition [118].

Secrecy was first introduced to the setup of bidirectional broadcasting in [119] under the following reliability and secrecy constraints:

$$\lim_{n\to\infty} \mathbb{P}\left[\mathrm{M}_1 \neq \hat{\mathrm{M}}_1 \text{ or } \mathrm{M}_2 \neq \tilde{\mathrm{M}}_2 | \mathcal{C}\right] = 0 \tag{3.2}$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n | \mathcal{C}) = 0. \tag{3.3}$$

Two achievable rate regions were established: The first region is based on the concept of wiretap random codes and it follows from the straightforward extension of the Csiszár and Körner's inner-bound. This region is given by the set of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy:

$$R_i \leq \mathbb{I}(\mathrm{V}; \mathrm{Y}_i) - \mathbb{I}(\mathrm{V}; \mathrm{Z}) \qquad i = 1, 2 \tag{3.4}$$

On the other hand, the second region is established using the concept of secret key encoding and is given by the set of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy:

$$R_1 = R_2 \leq \min\left[\mathbb{I}(\mathrm{X}; \mathrm{Y}_1), \mathbb{I}(\mathrm{X}; \mathrm{Y}_2)\right]. \tag{3.5}$$

The main idea behind the previous rate region is that the transmitter (relay) uses the fact that each receiver knows its own transmitted message. Thus, each message can be interpreted as a shared secret key for the other one. Beside the two achievable regions, an upper-bound on the secrecy capacity was established in [119]. However, the technique used for this derivation was not correct.

The achievable rate region in Eq. 3.5 has raised a very interesting question. This is because, although the eavesdropper can not extract any information about the confidential messages M_1 and M_2, one can easily show that the coding scheme used to establish this region does not fulfill the secrecy constraint in (3.3). In fact, it satisfies a more tolerant secrecy requirement given by the following:

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_1;\mathrm{Z}^n|\mathcal{C}) = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathcal{C}) = 0. \tag{3.6}$$

This observation raised many questions regarding the differences between the secrecy criterion in (3.3) and the one in (3.6). Additionally, it showed that although the problem of common message broadcasting and bidirectional broadcasting are equivalent, their corresponding secrecy scenarios are not. Moreover, this observation gave us a huge motivation to investigate and study the different secrecy constraints that can appear for the multi-receiver DM-WBC, and how these secrecy constraints can lead to different secrecy capacity regions.

3.1.2 A Comprehensive Channel Model

In this chapter, we consider the broadcasting phase of the secure bidirectional relaying in a three-node network with an external eavesdropper, however with an additional feature. The relay does not only transmit the individual confidential messages received in the multiple access phase, but it also transmits an additional common confidential message to both legitimate receivers and a common public message for all three receiving nodes: the two legitimate receivers and the eavesdropper [1]. This channel model combines the problem of public-confidential degraded message sets discussed in Section 2.4.2 along with the broadcasting phase of the secure bidirectional relaying problem. That is why we come up with the name: The two-receiver DM-WBC with degraded message sets and message cognition.

Let $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ be finite discrete input and output sets. Then, for input and output sequences $x^n \in \mathcal{X}^n$, $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$ and $z^n \in \mathcal{Z}^n$ of length n, the two-receiver DM-WBC: $Q = (W^1, W^2, V)$ is given by the probability transition matrix in (2.73). As indicated by Fig. 3.2, Our model considers four different messages sets. The first set is denoted by $\mathcal{M}_c = [\![1, 2^{nR_c}]\!]$ and it contains the common messages for all three receivers. The second set is denoted by $\mathcal{M}_0 = [\![1, 2^{nR_0}]\!]$ and contains the confidential common messages for the first and second legitimate receivers. The last two sets contain the individual confidential messages $\mathcal{M}_1 = [\![1, 2^{nR_1}]\!]$ and $\mathcal{M}_2 = [\![1, 2^{nR_2}]\!]$. Further, we assume a full message cognition at Y_1 and Y_2, such that Y_1 is cognizant of the entire message M_2 and Y_2 of the entire message M_1.

[1] Although the third receiver (Z) is part of our model and not an external user, we will refer to it in the rest of this chapter as an eavesdropper. From this point, we will sometimes refer to the different receivers by their respective channel outputs interchangeably.

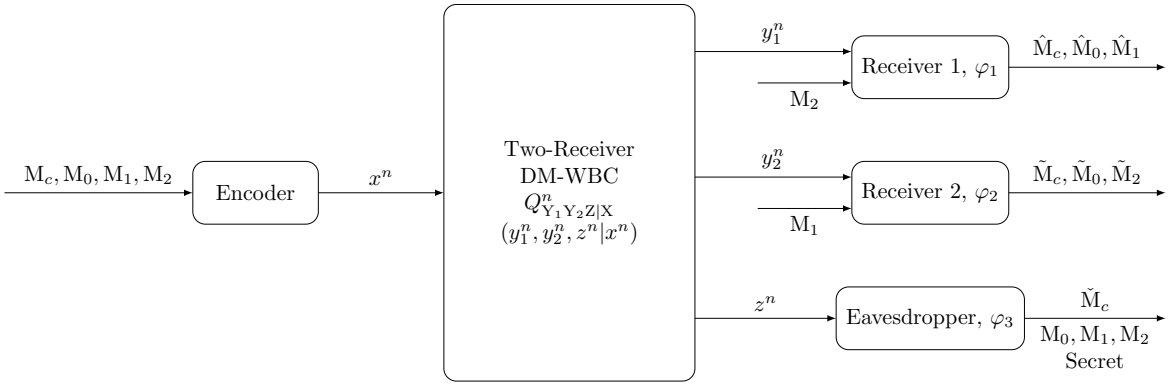


Figure 3.2: Two-receiver DM-WBC with degraded message sets and message cognition

It is important to point out that our model generalizes different works on the two-receiver DM-WBC and the three-receiver DM-BC as follows :

- If we let $\mathcal{M}_1 = \mathcal{M}_2 = \emptyset$, our model reduces to the problem of transmitting public and confidential degraded message sets over the two-receiver DM-WBC investigated in [12].
- If we let $\mathcal{M}_c = \mathcal{M}_0 = \emptyset$, our model reduces to the broadcasting phase of the secure bidirectional relaying problem in a three-node network with an external eavesdropper investigated in [12].
- If we neglected the secrecy requirements by assuming the confidential messages to be public one, our model reduces to the problem of three-receiver DM-BC with degraded message sets and full message cognition investigated in [118].
- If we let $\mathcal{M}_1 = \mathcal{M}_2 = \emptyset$ and assumed that $\mathcal{M}_0$ is a public message set instead of a confidential one, our model reduces to the famous problem of three-receiver DM-BC with two degraded message sets investigated in [68].

Definition 3.1. A $(2^{nR_c}, 2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code $\mathcal{C}$ for the two-receiver DM-WBC with degraded message sets and message cognition consists of: four independent message sets $\mathcal{M}_c$, $\mathcal{M}_0$, $\mathcal{M}_1$ and $\mathcal{M}_2$; a stochastic encoding function at the relay node

$$E : \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{P}(\mathcal{X}^n),$$

which maps a common message $m_c \in \mathcal{M}_c$, a confidential message triple $(m_0, m_1, m_2) \in \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$ into a codeword $x^n(m_c, m_0, m_1, m_2) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m_c, m_0, m_1, m_2)$, and three deterministic decoders, one for each node

$$\begin{aligned}
&\varphi_1 : \mathcal{Y}_1^n \times \mathcal{M}_2 \to \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1 \cup \{?\} \\
&\varphi_2 : \mathcal{Y}_2^n \times \mathcal{M}_1 \to \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_2 \cup \{?\} \\
&\varphi_3 : \mathcal{Z}^n \to \mathcal{M}_c \cup \{?\}
\end{aligned}$$

which maps each channel observation at the respective node along with the available side information into the corresponding required messages or an error message (?).

The previous definition considers a standard block code $\mathcal{C}$ of arbitrary but fixed length n. It is fair to assume that $\mathcal{C}$ is known to the transmitter, the two legitimate receivers and the eavesdropper. Additionally, we assume that M_c, M_0, M_1 and M_2 are independent and uniformly-distributed random variables over their respective message sets.

We also assume that perfect channel state information (CSI) is available at all nodes. This implies that the transmitter knows the channel statistics a head of time.

Now, in order to make sure that a code $\mathcal{C}$ is suitable for our model, we need to define some reliability and secrecy constraints to evaluate the performance of $\mathcal{C}$. In the next few lines, we consider only the reliability constraint, whereas the secrecy constraint is addressed in detail in the next subsection. Consider the average decoding error probability of the code $\mathcal{C}$ given by:

$$\bar{\mathbf{P}}_e(\mathcal{C}) \triangleq \mathbb{P}\Big[(\hat{\mathrm{M}}_c, \hat{\mathrm{M}}_0, \hat{\mathrm{M}}_1) \neq (\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_1) \text{ or } (\tilde{\mathrm{M}}_c, \tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_2) \neq (\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_2) \text{ or } \check{\mathrm{M}}_c \neq \mathrm{M}_c\Big], \tag{3.7}$$

where $(\hat{\mathrm{M}}_c, \hat{\mathrm{M}}_0, \hat{\mathrm{M}}_1)$, $(\tilde{\mathrm{M}}_c, \tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_2)$ and $\check{\mathrm{M}}_c$ are the estimated messages at Y_1, Y_2 and Z, respectively. Our target is to develop an achievablity coding scheme based on the previous code definition such that $\bar{\mathbf{P}}_e(\mathcal{C})$ is arbitrary small. It is important to mention here that since we assume the availability of perfect CSI at the transmitter, then all the achievability results established under the average error probability constraint in Eq. (3.7) are also valid for the maximum decoding error probability requirement [54].

3.1.3 Secrecy Criteria: Joint Versus Individual

Consider a multiple access wiretap channel that consists of two transmitters, a legitimate receiver and an eavesdropper. Each transmitter aims to send an individual confidential message to the legitimate receiver while keeping the eavesdropper ignorant about it. During the investigation of this problem in [110], it was argued that one can measure the ignorance of the eavesdropper about the confidential messages using two secrecy criteria: The first criterion considers the information leakage of each message to the eavesdropper separately, while the second one considers the mutual information leakage of both messages to the eavesdropper.

Aside from the previous scenario, most of the work on secure communication over DM-WBCs has only considered the concept of mutual information leakage to evaluate the ignorance of the eavesdropper about the confidential messages. Using the mutual information leakage might be convincing for the scenarios where only a common confidential message is transmitted over a DM-WBC. However, for the DM-WBCs where the legitimate receivers are interested in different confidential messages, it might be more interesting to consider other secrecy criteria. In fact, one needs to understand the practical implications of each secrecy criterion, and how to construct coding schemes that satisfy the constraints of each criterion. It is even more important to investigate whether changing the secrecy criterion affects the secrecy capacity region or not.

Now, back to our channel model given in Fig. 3.2, we need to construct a code $\mathcal{C}$ that can transmit the three confidential messages M_0, M_1 and M_2 to the respective legitimate receivers, while keeping the eavesdropper completely ignorant about those messages. In order to measure the level of ignorance of the eavesdropper, or in other words to evaluate the secrecy performance of $\mathcal{C}$, we consider two secrecy criteria as follows:

1. Joint Secrecy: This criterion requires the information leakage of the confidential messages intended for one legitimate receiver given the confidential messages intended only for the other legitimate receiver to be small. For our model, this requirement can be expressed as follows:

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) \leq \tau_{1n} \quad \text{and} \quad \mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1) \leq \tau_{2n}, \\ \text{where } \lim_{n\to\infty} \tau_{1n}, \tau_{2n} = 0. \tag{3.8}$$

This criterion indicates that if –by some genie-aided way– the eavesdropper managed to acquire the individual confidential message of one legitimate receiver, the secrecy of the confidential messages intended for the other legitimate receiver will not be affected. This implies that the secrecy of the confidential messages of each legitimate receiver is assured, even if the other legitimate receiver is compromised. The joint secrecy can also be interpreted as a secrecy criterion in which the legitimate receivers do not trust each other such that each receiver requires its confidential messages to be protected independently from the confidential messages of the other ones.

2. Individual Secrecy: This criterion requires the information leaked to the eavesdropper about the confidential messages intended for each legitimate receiver to be small. For our model, this requirement can be expressed as follows:

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n) \leq \tau_{1n} \quad \text{and} \quad \mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n) \leq \tau_{2n}, \\ \text{where } \lim_{n\to\infty} \tau_{1n}, \tau_{2n} = 0. \tag{3.9}$$

The previous definition implies that the individual secrecy simply ignores the conditioning imposed by the joint secrecy criterion. Since, the confidential messages are assumed to be independent which implies that $\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n) \leq \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2)$, it follows directly that the individual secrecy criterion is not strict as the joint criterion. In fact, the individual secrecy criterion is based on some sort of mutual trust between the legitimate receivers. This mutual trust is reflected in the coding schemes that adhere to the individual secrecy requirements.

It is important to point out that in previous literature cf. [84], the joint secrecy criterion is defined based on the concept of requesting the mutual information leakage of all confidential messages to the eavesdropper to be small. For our model, this can be expressed as follows:

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2;\mathrm{Z}^n) \leq \tau_n \quad \text{where} \quad \lim_{n\to\infty} \tau_n = 0. \tag{3.10}$$

Although the previous requirement seems a little bit different from the one in Eq. (3.8), one can easily show that for some $\tau_n \geq \tau_{1n} + \tau_{2n}$ both constraints are equivalent. We start by showing that the constraints in (3.8) directly lead to the one in (3.10):

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2;\mathrm{Z}^n) &= \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) + \mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n) \\
&\overset{(a)}{\leq} \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) + \mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1) \\
&\overset{(b)}{\leq} \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1) \\
&\leq \tau_{1n} + \tau_{2n} \leq \tau_n,
\end{aligned}$$

where (a) follows because M_1 and M_2 are independent which implies that $\mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n) \leq \mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1)$; while (b) follows because $\mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1) \leq \mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1)$.

On the other hand, if Eq. (3.10) holds, it follows directly that $\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) \leq \tau_n$ and $\mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n|\mathrm{M}_1) \leq \tau_n$, which means that the constraints in (3.10) are true. However, we find it more convenient to define the requirements of the joint secrecy criterion using Eq. (3.8), because this definition allows us to identify the immunity of the joint secrecy criterion to compromised receivers. It also captures the main difference between joint secrecy and individual secrecy.

Before, we move to the next section, we need to highlight a very important point. Most of the previous work that investigated secure communication over multi-receiver DM-WBC has only considered the joint secrecy criterion with respect to the weak secrecy measure cf. Section 2.1.4. However, in this chapter, we focus on constructing coding schemes for the two-receiver DM-WBC that adhere to the notation of strong secrecy, as illustrated by the joint secrecy constraints in (3.8) and the individual secrecy constraints in (3.9). The results established in this chapter were published in [120–122]. In a parallel and independent work, secure communication over the two-receiver DM-WBC with receiver side information was investigated under the individual secrecy criterion, but only with respect to the weak secrecy measure cf. [123–126].

3.1.4 Individual Secrecy in Shannon's Ciphering System

In the previous section, we introduced the individual secrecy criterion, and addressed how it differs from the conservative joint secrecy criterion. However, we only considered one of two possible notations used to address individual secrecy. In this section, we introduce the other notation and illustrate through an example why we prefer the notation introduced earlier in Eq. (3.9).

Consider a secrecy criterion that requires the sum of the information leakages of each confidential message to the eavesdropper to be small. By applying this definition to our model, we reach the following expression:

$$\mathbb{I}(\mathrm{M}_0;\mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_1;\mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n) \leq \tau_n \quad \text{where} \quad \lim_{n\to\infty} \tau_n = 0. \tag{3.11}$$

The previous constraint is a direct extension of the concept of individual secrecy introduced in [110] for the multiple access wiretap channel. Nevertheless, the previous constraint is only equivalent to the individual secrecy constraint in (3.9) if and only if $\mathrm{M}_0 = \emptyset$, but in general they are not the same. In particular, one can show that the constraint in (3.9) is stronger than this one. This is because Eq. (3.9) directly implies Eq. (3.11), while the opposite is not true.

The difference between these two constraints lies in the interpretation of the individuality as follows: The constraint in (3.9) considers individuality from the perspective of transmission flows, where as long as the confidential messages are intended to the same legitimate receiver, they can be considered as one information entity. On the other hand, the constraint in (3.11) addresses the individuality according to the description of the given confidential messages sets. In the following lines, we use Shannon's ciphering system to show why addressing individual secrecy with respect to different

messages might be misleading, and that it is more consistent to interpret individuality with respect to different transmission flows.

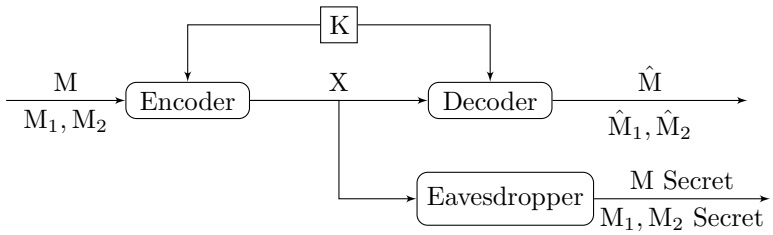


Figure 3.3: Individual secrecy for Shannon's ciphering system

We consider the scenario given by Fig. 3.3. Shannon studied this model under the following secrecy constraint:

$$\mathbb{I}(M; X) = 0. \tag{3.12}$$

He proved that this requirement is achieved if $\mathbb{H}(M) \leq \mathbb{H}(K)$, where K is a random variable that represents the secret key shared between the transmitter and the receiver. Assuming that we have a secret key such that $\mathbb{H}(K) = \frac{1}{2}\mathbb{H}(M)$, we can construct the following coding strategy. First, we divide M into two messages M_1 and M_2, such that $\mathbb{H}(M_1) = \mathbb{H}(M_2) = \mathbb{H}(K)$. We then construct a new secret key $\tilde{K}$ by concatenating K and M_1. Now the encoder outputs $X = M \otimes \tilde{K}$, which is equivalent to the concatenation of $M_1 \otimes K$ and $M_2 \otimes M_1$. The decoder works in the following order; it first extracts $\hat{M}_1$ from the first part of X by *xoring* it with the shared secret key K, then it use $\hat{M}_1$ to extract $\hat{M}_2$ from the second part of X.

Using this technique, we can overcome the problem of short secret keys, however we need to understand its drawbacks. Aside form the problem of error progression that arises form using the estimated $\hat{M}_1$ to decode M_2, this technique does not fulfill the secrecy constraint in (3.12). However, it fulfills the following secrecy constraint:

$$\mathbb{I}(M_1; X) + \mathbb{I}(M_2; X) = 0. \tag{3.13}$$

In general, we can extend this coding technique for keys with smaller entropy by dividing the message M into smaller messages of the same entropy as the given key such that $M = \prod_{i=1}^{L} M_i$. We can show that, the previous technique grants a certain secrecy level, at which the sum of the information leakage of the small messages to the eavesdropper is small.

The difference between the two secrecy criteria in the previous example is related to how to address the secrecy of information transmitted to a single user; whether it should be protected as a one big entity, or it can be divided into smaller parts, where each part is protected separately. This issue is identical to the problem of identifying the individual secrecy and whether individuality means different users (transmission flows) or different messages. That is why, we preferred the individual secrecy constraint in (3.9) because it assures that the whole information transmitted to a certain user is protected as one big entity. In our opinion, this is a more consistent and meaningful notation.

3.2 Joint-Strong Secrecy for Wiretap Marton Codes

In this section, we investigate secure communication over the two-receiver DM-WBC with degraded message sets and message cognition in Fig. 3.2 under the joint-strong secrecy criterion. In this investigation, we establish one of the main results of this thesis which is a general achievable rate region that combines two main coding principles: superposition encoding and Marton-Coding along with the concept of indirect decoding under the joint-strong secrecy criterion. Our result is among the first ones to establish a full Marton-Coding secrecy rate region that satisfies the joint-strong secrecy criterion.

3.2.1 A Universal Achievable Rate Region

In order to present our main result, we first need a formal definition for the achievable rate region of the communication scenario in Fig. 3.2 under the joint-strong secrecy criterion. Thus, we present the following definition:

Definition 3.2. A rate quadruple $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ is a joint-strong secrecy achievable rate quadruple for the two-receiver DM-WBC, i.e. $Q = (W^1, W^2, V)$, if for all $\eta_i > 0$ where $i \in \{c, 0, 1, 2\}$, $\lambda > 0$ and $\tau_j > 0$ where $j \in \{1, 2\}$, there is an $n(\eta_i, \lambda, \tau_j) \in \mathbb{N}$ such that for all $n > n(\eta_i, \lambda, \tau_j)$ there exists a sequence of codes $\{\mathcal{C}\}_n$ according to Definition 3.1 that satisfies the following constraints:

$$\frac{1}{n}\log|\mathcal{M}_i| \geq R_i - \eta_i \qquad \text{for} \qquad i \in \{c, 0, 1, 2\} \tag{3.14a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda \tag{3.14b}$$

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathrm{M}_2\mathcal{C}) \leq \tau_1 \qquad \text{and} \qquad \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathrm{M}_1\mathcal{C}) \leq \tau_2, \tag{3.14c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by Eq. (3.7). The joint-strong secrecy capacity region $C_J(Q)$ is given by the set of all achievable joint-strong secrecy rate quadruples (R_c, R_0, R_1, R_2).

In order to show that a rate quadruple (R_c, R_0, R_1, R_2) is achievable under the joint-strong secrecy criterion, it is enough to show that there exists a random coding scheme with random variable C such that:

$$\frac{1}{n}\log|\mathcal{M}_i| = R_i \tag{3.15a}$$

$$\lim_{n\to\infty} \mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] = 0 \tag{3.15b}$$

$$\lim_{n\to\infty} \mathbb{E}[\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathrm{C})] = 0, \tag{3.15c}$$

where $i \in \{c, 0, 1, 2\}$. The validity of the previous argument follows from Definition 3.2, the selection lemma (Lemma A.4) in addition to the equivalency of the joint secrecy constraints in (3.8) and (3.10). We now present the main result of this section:

Theorem 3.1. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition is given by the set of all rate quadruples $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ that satisfy*

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U}; \mathrm{Z}|\mathrm{T}), \mathbb{I}(\mathrm{U}; \mathrm{Y}_1|\mathrm{T}), \mathbb{I}(\mathrm{U}; \mathrm{Y}_2|\mathrm{T})\Big] \tag{3.16a}$$

$$R_0 + R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Z}|\mathrm{U}\ \mathrm{T}) \tag{3.16b}$$

$$R_0 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}|\mathrm{U}\ \mathrm{T}) \tag{3.16c}$$

$$\begin{aligned} 2R_0 + R_1 + R_2 \leq\ & \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{U}\ \mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0\ \mathrm{T}) \\ & - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{U}\ \mathrm{T}) \end{aligned} \tag{3.16d}$$

for random variables with joint probability distribution $\mathrm{Q}_{\mathrm{T}}\ \mathrm{Q}_{\mathrm{U}|\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{U}}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{T} - \mathrm{U} - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous rate region is similar to the achievable rate region established in [12, Theorem 2] for the two-receiver DM-WBC with public-confidential degraded message sets under the weak secrecy measure. In fact, one can show that if $R_1 = R_2 = 0$ and $\mathrm{T} = \emptyset$, the two rate regions are identical. The equality of the two regions arises from the fact that both regions are based on the same coding principles: superposition encoding, Marton-Coding and indirect decoding. In particular, this equality implies a very important result: Strengthening the secrecy criterion from weak to strong for the two-receiver DM-WBC comes at no cost with respect to any achievable rate region derived using these coding techniques. Next, we present two achievable rate regions that play an important role in establishing Theorem 3.1.

Proposition 3.1. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition is given by the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U}; \mathrm{Z}), \mathbb{I}(\mathrm{U}; \mathrm{Y}_1), \mathbb{I}(\mathrm{U}; \mathrm{Y}_2)\Big] \tag{3.17a}$$

$$R_0 + R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Z}|\mathrm{U}) \tag{3.17b}$$

$$R_0 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) \tag{3.17c}$$

for random variables with joint distribution $\mathrm{Q}_{\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{U}}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that:* $\mathrm{U} - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous proposition serves as the main pillar for the proof of Theorem 3.1. Thus, it is important to point out how the rate region in (3.17) and the one in (3.16) are related.

- Both rate regions are defined with respect to the same joint distribution on the random variables $(\mathrm{U}, \mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2, \mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$. The only difference is that for the rate region in (3.16), this joint distribution is generated conditioned on the probability distribution Q_{T} of the time sharing random variable T.
- If we let $\mathrm{T} = \emptyset$, we find that both rate regions in (3.16) and (3.17) have the same bounds on the rates intended to the first legitimate receiver, i.e. $(R_0 + R_1)$ cf. (3.16b) and (3.17b).
- If we let $\mathrm{T} = \emptyset$, we find that the bounds on the rates intended to the second legitimate receiver, i.e. $(R_0 + R_2)$ in (3.17c) are tighter than the ones in (3.16c). The previous statement follows because:

$$\begin{aligned} \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) &= \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2; \mathrm{Z}|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1|\mathrm{V}_0\mathrm{Z}) \\ &\overset{(a)}{=} \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1|\mathrm{V}_0\mathrm{Z}) \\ &\overset{(b)}{\geq} \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}|\mathrm{U}), \end{aligned}$$

  where (a) follows from the chain rule of the mutual information and the definition of the joint distribution on $(\mathrm{U}, \mathrm{V}_0, \mathrm{V}_2, \mathrm{Z})$, while (b) follows from the non-negativity of the mutual information.

- In spite of the previous two points, one can actually show that at $\mathrm{T} = \emptyset$, both rate regions in (3.17) and (3.16) have the same bounds on the sum rates intended to both legitimate receivers, i.e. $(2R_0 + R_1 + R_2)$. This is because

$$\begin{aligned}\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) &= \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_2;\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1)\\ &\overset{(a)}{=} \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1|\mathrm{V}_0),\end{aligned}$$

  where (a) follows from the chain rule of the mutual information and the definition of the joint distribution on $(\mathrm{U}, \mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2, \mathrm{Z})$.

The previous comparison implies that rate region given in (3.17) is a sub-region of the full rate region in (3.16). In particular, one can argue that the rate region in (3.17) is only optimal for the first legitimate receiver because it has the same bounds as the one in (3.16) on the rates intended to this receiver.

Proposition 3.1 implies that there exists a random coding scheme that can be used to generate a code $\mathcal{C}_1$ according to Definition 3.1, such that the rates of $\mathcal{C}_1$ are bounded according to (3.17) and $\mathcal{C}_1$ satisfies the reliability and secrecy constraints given by (3.14b) and (3.14c) respectively. This directly suggests the existence of an alternative coding scheme that can be used to generate a code $\mathcal{C}_2$ according to Definition 3.1 which satisfies the reliability and secrecy constraints given by (3.14b) and (3.14c) respectively, while its rates are bounded according to the next proposition.

Proposition 3.2. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition is given by the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\Big] \tag{3.18a}$$

$$R_0 + R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2\mathrm{Z}|\mathrm{V}_0) \tag{3.18b}$$

$$R_0 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{U}) \tag{3.18c}$$

for random variables with joint distribution $\mathrm{Q}_{\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{U}}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that:* $\mathrm{U} - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

It is obvious that Proposition 3.2 is a direct consequence for Proposition 3.1 since it follows by interchanging the roles of the first and the second legitimate receivers. This implies that the four remarks on the relation between the rate regions in (3.16) and (3.17) can also be used in a way to describe the relation between the rate regions in (3.16) and (3.18) by interchanging the statements on the first and second legitimate receivers. We also need to keep in mind that the following is true:

$$\begin{aligned}\mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2\mathrm{Z}|\mathrm{V}_0) &\geq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U})\\ \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2\mathrm{Z}|\mathrm{V}_0) &= \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0)\end{aligned}$$

Thus, it follows that the rate region in (3.18) is also a sub-region of the full rate region in (3.16). In fact, it has the same bounds on the rates intended to the second legitimate receiver as the full region in (3.16).

In previous literature, two approaches have been used to prove the achievability of a rate region that utilizes the principle of Marton-Coding. The first approach is a direct

technique in which the full Marton-Coding rate region is derived directly from the reliability analysis of the coding scheme. This approach was introduced by El-Gamal and Meulen in [88] and has been extended to different secrecy scenarios as in [12, 84]. The second approach built upon the original work of Marton in [70]. This approach uses an indirect technique in which the achievability of two smaller rate regions is established at first. Then, by using the principle of time sharing between these two regions, the achievability of the full Marton-Coding rate region can be derived.

In this thesis, we follow Marton's original approach to prove Theorem 3.1 instead of El-Gamal and Meulen's approach. This is because Marton's approach is more suitable for our strong secrecy analysis techniques. Our proof will adhere to the following algorithm: We start by presenting a detailed proof for Proposition 3.1. This directly implies the validity of Proposition 3.2 as well. We then use the concept of rate splitting and time sharing to combine the two rate regions in Proposition 3.1 and 3.2 leading to the achievability of the full rate region of Theorem 3.1.

3.2.2 Coding Scheme: Encoding and Decoding

The main aim of this section is to introduce the coding scheme required to establish the achievability of the rate region given in Proposition 3.1. We mainly address the technique used to generate our codebook and how the encoder and the three decoders defined in Definition 3.1 work. Our coding scheme and in particular our encoder is inspired by the likelihood encoder introduced in [127]

Let $n \in \mathbb{N}$ be an arbitrary but fixed code block length and $\epsilon > 0$ be an arbitrary constant. Moreover, let $\mathrm{Q}_{\mathrm{UV_0V_1V_2X}} \in \mathcal{P}(\mathcal{U} \times \mathcal{V}_0 \times \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{X})$ be a fixed input probability distribution, and recall the probability transition matrix $Q_{\mathrm{Y_1Y_2Z|X}}$ given in (2.73) that defines the two-receiver DM-WBC. Furthermore, we assume that quantities of the form 2^{nR}, where $R \in \mathbb{R}_+$ are integers.

1. Message Sets: We consider the following sets: the set of common messages $\mathcal{M}_c = [\![1, 2^{nR_c}]\!]$, the set of confidential common messages $\mathcal{M}_0 = [\![1, 2^{nR_0}]\!]$, two sets of confidential individual messages $\mathcal{M}_1 = [\![1, 2^{nR_1}]\!]$ and $\mathcal{M}_2 = [\![1, 2^{nR_2}]\!]$ Three sets of randomization messages $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$, $\mathcal{M}_{r_1} = [\![1, 2^{nR_{r_1}}]\!]$ and $\mathcal{M}_{r_2} = [\![1, 2^{nR_{r_2}}]\!]$, and an additional message set $\mathcal{M}_t = [\![1, 2^{nR_t}]\!]$ needed for Marton-coding. Additionally, we use $\mathcal{M} = \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$ to abbreviate the set of all confidential messages, such that $\mathcal{M} = [\![1, 2^{nR}]\!]$, where $R = R_0 + R_1 + R_2$.

2. Random Codebook $\mathcal{C}$: Let $\mathrm{C}_c \triangleq \{\mathrm{U}^n(m_c)\}$ for $m_c \in \mathcal{M}_c$ be a random public common message codebook that consists of 2^{nR_c} i.i.d. random codewords $\mathrm{U}^n(m_c)$, where each codeword is constructed by generating the symbols $\mathrm{U}_i(m_c)$ with $i \in [\![1, n]\!]$ independently according to Q_{U}. A realization of C_c is denoted by $\mathcal{C}_c \triangleq \{u^n(m_c)\}$.

For a fixed public message codebook $\mathcal{C}_c$ and for every $m_c \in \mathcal{M}_c$, let $\mathrm{C}_0(m_c) \triangleq \{\mathrm{V}_0^n(m_c, m, m_r)\}$ where $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$ be a random codebook for the confidential messages that consists of $2^{n(R+R_r)}$ conditionally independent random codewords $\mathrm{V}_0^n(m_c, m, m_r)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{0_i}(m_c, m, m_r)$ independently at random according to $\mathrm{Q}_{\mathrm{V_0|U}}(\cdot|u_i(m_c))$. A realization of $\mathrm{C}_0(m_c)$ is denoted by $\mathcal{C}_0(m_c) \triangleq \{v_0^n(m_c, m, m_r)\}$.

For a fixed confidential messages codebook $\mathcal{C}_0(m_c)$ and for every $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$, let $\mathrm{C}_1(m_c, m, m_r) \triangleq \{\mathrm{V}_1^n(m_c, m, m_r, m_{r_1})\}$ be a confidential random codebook that consists of $2^{nR_{r_1}}$ conditionally independent random codewords $\mathrm{V}_1^n(m_c, m, m_r, m_{r_1})$, where each codeword is constructed by generating the symbols $\mathrm{V}_{1_i}(m_c, m, m_r, m_{r_1})$ independently at random according to $\mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0}(\cdot|v_{0_i}(m_c, m, m_r))$.

Similarly, for a fixed confidential messages codebook $\mathcal{C}_0(m_c)$ and for every $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$, let $\mathrm{C}_2(m_c, m, m_r) \triangleq \{\mathrm{V}_2^n(m_c, m, m_r, m_{r_2}, m_t)\}$ be a confidential random codebook that consists of $2^{n(R_{r_2}+R_t)}$ conditionally independent random codewords $\mathrm{V}_2^n(m_c, m, m_r, m_{r_2}, m_t)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{2_i}(m_c, m, m_r, m_{r_2}, m_t)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_2|\mathrm{V}_0}(\cdot|v_{0_i}(m_c, m, m_r))$.

The total random codebook is denoted by $\mathrm{C} = \{\mathrm{C}_c, \mathrm{C}_0, \mathrm{C}_1, \mathrm{C}_2\}$, while $\mathcal{C} = \{\mathcal{C}_c, \mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2\}$ denotes one possible realization of this codebook selected from the total set of codebooks given by $\mathfrak{C}$. The probability of generating a certain realization $\mathcal{C} \in \mathfrak{C}$ is given by:

$$\begin{aligned}\mathrm{G}(\mathcal{C}) = &\prod_{\substack{(m_c^2, m^2, m_r^2, m_{r_2}, m_t) \in \\ \mathcal{M}_c \times \mathcal{M} \times \mathcal{M}_r \times \mathcal{M}_{r_2} \times \mathcal{M}_t}} \mathrm{Q}_{\mathrm{V}_2|\mathrm{V}_0}^n\Big(v_2^n(m_c^2, m^2, m_r^2, m_{r_2}, m_t)|v_0^n(m_c^2, m^2, m_r^2)\Big) \\ &\prod_{\substack{(m_c^1, m^1, m_r^1, m_{r_1}) \in \\ \mathcal{M}_c \times \mathcal{M} \times \mathcal{M}_r \times \mathcal{M}_{r_1}}} \mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0}^n\Big(v_1^n(m_c^1, m^1, m_r^1, m_{r_1})|v_0^n(m_c^1, m^1, m_r^1)\Big) \\ &\prod_{\substack{(m_c^0, m, m_r) \in \\ \mathcal{M}_c \times \mathcal{M} \times \mathcal{M}_r}} \mathrm{Q}_{\mathrm{V}_0|\mathrm{U}}^n\Big(v_0^n(m_c^0, m, m_r)|u^n(m_c^0)\Big) \prod_{m_c \in \mathcal{M}_c} \mathrm{Q}_{\mathrm{U}}^n\big(u^n(m_c)\big) \qquad (3.19)\end{aligned}$$

3. Encoder E: For a fixed codebook realization $\mathcal{C}$, given a message pair (m_c, m), where $m = (m_0, m_1, m_2)$, the transmitter chooses three randomization messages m_r, m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_r$, $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. Then, it chooses an index $m_t \in \mathcal{M}_t$ based on the following likelihood encoder:

$$\begin{aligned}&E^{(\mathrm{LE})}\big(m_t|m_{r_2}, v_0^n(m_c, m, m_r), v_1^n(m_c, m, m_r, m_{r_1})\big) \\ &\quad = \frac{\mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}^n\big(v_1^n(m_c, m, m_r, m_{r_1})|v_0^n(m_c, m, m_r), v_2^n(m_c, m, m_r, m_{r_2}, m_t)\big)}{\sum_{j \in \mathcal{M}_t} \mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}^n\big(v_1^n(m_c, m, m_r, m_{r_1})|v_0^n(m_c, m, m_r), v_2^n(m_c, m, m_r, m_{r_2}, j)\big)}. \qquad (3.20)\end{aligned}$$

Finally, for the selected triple $v_0^n(m_c, m, m_r)$, $v_1^n(m_c, m, m_r, m_{r_1})$ and $v_2^n(m_c, m, m_r, m_{r_2}, m_t)$, the encoder generates a codeword x^n independently at random according to the conditional distribution $\prod_{i=1}^n \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0,\mathrm{V}_1,\mathrm{V}_2}^n\big(\,\cdot\,|v_0^n(m_c, m, m_r), v_1^n(m_c, m, m_r, m_{r_1}), v_2^n(m_c, m, m_r, m_{r_2}, m_t)\big)$ and transmits it.

4. First Legitimate Decoder φ_1: Given y_1^n and its own message m_2, the decoder searches for the unique quadruple $(\hat{m}_c, \hat{m}_0, \hat{m}_1, \hat{m}_r) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_r$; for which there exists an index $\hat{m}_{r_1} \in \mathcal{M}_{r_1}$ such that:

$$\Big(u^n(\hat{m}_c), v_0^n(\hat{m}_c, \hat{m}, \hat{m}_r), v_1^n(\hat{m}_c, \hat{m}, \hat{m}_r, \hat{m}_{r_1}), y_1^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1\mathrm{Y}_1}),$$

where $\hat{m} = (\hat{m}_0, \hat{m}_1, m_2)$. If such unique quadruple exists, the decoder outputs the triple $(\hat{m}_c, \hat{m}_0, \hat{m}_1)$. Otherwise, the decoder outputs an error message (?) . It is important to point out that this decoder does not aim to determine the unique message m_{r_1} that was transmitted. Nevertheless, the decoding role includes the typicality of the sequence $v_1^n(\hat{m}_c, \hat{m}, \hat{m}_r, \hat{m}_{r_1})$, this is because the previous decoder inherits the concept of indirect decoding.

5. Second Legitimate Decoder φ_2: Given y_2^n and its own message m_1, the decoder searches for the unique quadruple $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2, \tilde{m}_r) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_2 \times \mathcal{M}_r$; for which there exist two indices $(\tilde{m}_{r_2}, \tilde{m}_t) \in \mathcal{M}_{r_2} \times \mathcal{M}_t$ such that:

$$\Big(u^n(\tilde{m}_c), v_0^n(\tilde{m}_c, \tilde{m}, \tilde{m}_r), v_2^n(\tilde{m}_c, \tilde{m}, \tilde{m}_r, \tilde{m}_{r_2}, \tilde{m}_t), y_2^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UV_0V_2Y_2}}),$$

where $\tilde{m} = (\tilde{m}_0, m_1, \tilde{m}_2)$. If such unique quadruple exists, the decoder outputs the triple $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2)$. If not, it outputs the error message (?). The previous decoding role implies that this decoder also applies the concept of indirect decoding by including the sequence $v_2^n(\tilde{m}_c, \tilde{m}, \tilde{m}_r, \tilde{m}_{r_2}, \tilde{m}_t)$ in the typicality condition although it is not interested in finding the unique pair (m_{r_2}, m_t) that was transmitted.

6. Third Eavesdropper Decoder φ_3: Given z^n, the decoder searches for the unique message $\check{m}_c \in \mathcal{M}_c$ such that:

$$\big(u^n(\check{m}_c), z^n\big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UZ}})$$

If this unique message exists, the decoder outputs it. Otherwise, the decoder outputs the error message (?).

7. Induced Joint Distribution: The encoding function E along with the decoding functions $(\varphi_1, \varphi_2, \varphi_3)$ defined with respect to the codebook $\mathcal{C} \in \mathfrak{C}$ compose a $(2^{nR_c}, 2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code[2] of Definition 3.1. For every codebook $\mathcal{C}$, the following joint probability distribution is induced:

$$\begin{aligned}
&\mathrm{P}^{\mathcal{C}}\Big(m_c, m, m_r, m_{r_1}, m_{r_2}, u^n, v_0^n, v_1^n, m_t, v_2^n, x^n, y_1^n, y_2^n, z^n, \hat{m}_c, \hat{m}_0, \hat{m}_1, \tilde{m}_c, \tilde{m}_0, \tilde{m}_2, \check{m}_c\Big) \\
&\quad = 2^{-n(R_c+R+R_r+R_{r_1}+R_{r_2})} \mathbb{1}_{\{u^n = u^n(m_c)\}} \mathbb{1}_{\{v_0^n = v_0^n(m_c, m, m_r)\}} \mathbb{1}_{\{v_1^n = v_1^n(m_c, m, m_r, m_{r_1})\}} \\
&\quad \times E^{(\mathrm{LE})}\big(m_t | m_{r_2}, v_0^n(m_c, m, m_r), v_1^n(m_c, m, m_r, m_{r_1})\big) \mathbb{1}_{\{v_2^n = v_2^n(m_c, m, m_r, m_{r_2}, m_t)\}} \\
&\quad \times \mathrm{Q}^n_{\mathrm{X|V_0V_1V_2}}(x^n | v_0^n, v_1^n, v_2^n)\, Q^n_{\mathrm{Y_1Y_2Z|X}}(y_1^n, y_2^n, z^n | x^n) \mathbb{1}_{\{(\hat{m}_c, \hat{m}_0, \hat{m}_1) = \varphi_1(y_1^n, m_2)\}} \\
&\quad \times \mathbb{1}_{\{(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2) = \varphi_2(y_2^n, m_1)\}} \mathbb{1}_{\{\check{m}_c = \varphi_3(z^n)\}}
\end{aligned} \tag{3.21}$$

Additionally, if we take the random codebook generation process into our consideration, we end up with the subsequent distribution:

$$\mathrm{P} = \mathrm{G}(\mathcal{C}) \cdot \mathrm{P}^{\mathcal{C}}, \tag{3.22}$$

where $\mathrm{G}(\mathcal{C})$ is the probability of obtaining a certain codebook $\mathcal{C}$ given by (3.19). In the upcoming sections, we present the reliability and secrecy analysis of our coding

[2] We slightly abuse the notation by using $\mathcal{C}$ to denote both the codebook and the whole code which includes the codebook itself along with the encoding and decoding functions. We favor this notation for its simplicity and remind the reader that the codebook uniquely defines the code.

scheme. In this analysis, the expectation expression $\mathbb{E}_{\mathrm{C}}[\cdot]$ will come up regularly. It is important to point out that such expression is calculated with respect to the distribution P given by (3.22).

8. Fourier-Motzkin Elimination: In order to derive the achievable rate region that corresponds to our coding scheme, we need to derive the necessary rate bounds. The reliability and secrecy analysis presented in the next two subsections show that our coding scheme satisfies the requirements in (3.15b) and (3.15c), as long as:

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) + \delta_3(\epsilon) \\
R_0 + R_1 + R_r + R_{r_1} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}) + \delta_1(\epsilon) \\
R_0 + R_2 + R_r + R_{r_2} + R_t &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) + \delta_2(\epsilon) \\
R_c + R_0 + R_1 + R_r + R_{r_1} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \delta_1(\epsilon) \\
R_c + R_0 + R_2 + R_r + R_{r_2} + R_t &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) + \delta_2(\epsilon) \\
R_t &\geq \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon,\beta) \\
R_{r_2} + R_t &\geq \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) + \delta_4(\epsilon,\gamma) \\
R_{r_1} &\geq \mathbb{I}(\mathrm{V}_1;\mathrm{Z}|\mathrm{V}_0) + \delta_4(\epsilon,\gamma) \\
R_r &\geq \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) + \delta_4(\epsilon,\gamma),
\end{aligned} \tag{3.23}$$

where ϵ, β and γ can be selected in a way such that as n approaches infinity, the corresponding δ-functions approaches zero. Now, if we apply the Fourier-Motzkin elimination procedure [57], then take the limit as $n \to \infty$, we prove the achievability of any rate quadruple (R_c, R_0, R_1, R_2) satisfying the constraints in (3.17).

3.2.3 Encoding and Decoding Error Analysis

In this section, we present the detailed reliability analysis required to derive some of the rate constraints given by (3.23). Before we present our analysis, we need to introduce some necessary definitions. Given the sequences $u^n \in \mathcal{U}^n$, $v_0^n \in \mathcal{V}_0^n$ and $v_1^n \in \mathcal{V}_1^n$, let $\tilde{\mathrm{C}}_2(v_0^n) \triangleq \{\tilde{\mathrm{V}}_2^n(v_0^n, m_{r_2}, m_t)\}$ be a collection of i.i.d. random codewords of length n generated according to $\mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0}(\cdot|v_0^n)$, for every $m_{r_2} \in \mathcal{M}_{r_2}$ and $m_t \in \mathcal{M}_t$. Additionally, let $\tilde{\mathrm{C}}_2 \triangleq \tilde{\mathrm{C}}_2(v_0)$ be a random codebook generated for all $v_0^n \in \mathcal{V}_0^n$. This implies that the probability of obtaining a certain codebook realization $\tilde{\mathcal{C}}_2 \in \tilde{\mathfrak{C}}_2$ is given by:

$$\mathrm{G}(\tilde{\mathcal{C}}_2) = \prod_{v_0^n \in \mathcal{V}_0^n} \prod_{(m_{r_2}, m_t) \in \mathcal{M}_{r_2} \times \mathcal{M}_t} \mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0}(\tilde{v}_2^n(v_0^n, m_{r_2}, m_t)|v_0^n). \tag{3.24}$$

For the given sequences: (u^n, v_0^n, v_1^n) and a certain codebook realization $\tilde{\mathcal{C}}_2(v_0^2)$, a codeword $\tilde{v}_2^n(v_0^n, m_{r_2}, m_t) \in \mathcal{V}_2^n$ is selected as follows: First a message m_{r_2} is chosen uniformly at random from $\mathcal{M}_{r_2}$, then an index $m_t \in \mathcal{M}_t$ is drawn randomly according to the following likelihood encoder:

$$\tilde{E}^{\mathrm{LE}}(m_t|m_{r_2}, v_0^n, v_1^n) = \frac{\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n|v_0^n, \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\big)}{\sum_{j \in \mathcal{M}_t} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n|v_0^n, \tilde{v}_2^n(v_0^n, m_{r_2}, j)\big)}. \tag{3.25}$$

For a fixed codebook realization $\tilde{\mathcal{C}}_2$, this encoding scheme defines a probability distribution on the random variables $(\mathrm{M}_{r_2}, \mathrm{M}_t, \tilde{\mathrm{V}}_2^n)$ conditioned on the sequences (u^n, v_0^n, v_1^n) as follows:

$$\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}(m_{r_2}, m_t, \tilde{v}_2^n, |u^n, v_0^n, v_1^n) = 2^{-nR_{r_2}}\, \tilde{E}^{\mathrm{LE}}(m_t|m_{r_2}, v_0^n, v_1^n)\, \mathbb{1}_{\{\tilde{v}_2^n = \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\}}. \tag{3.26}$$

Furthermore, assume that the sequences (u^n, v_0^n, v_1^n) are generated by the input distribution $\mathrm{Q}^n_{\mathrm{UV_0V_1}}$ and that the codeword $\tilde{v}_2^n(v_0^n, m_{r_2}, m_t) \in \tilde{\mathcal{C}}_2(v_0^n)$ chosen by the previous encoding scheme is transmitted over the DMC $Q_{\mathrm{Z|V_0V_1V_2}}$ given by:

$$Q_{\mathrm{Z|V_0V_1V_2}}(z|v_0, v_1, v_2) \triangleq \sum_{x\in\mathcal{X}} \sum_{y_1\in\mathcal{Y}_1} \sum_{y_2\in\mathcal{Y}_2} \mathrm{Q}_{\mathrm{X|V_0V_1V_2}}(x|v_0, v_1, v_2)\, Q_{\mathrm{Y_1Y_2Z|X}}(y_1, y_2, z|x). \tag{3.27}$$

Then, for a certain codebook $\tilde{\mathcal{C}}_2$ we can define $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ as the distribution induced by the whole system as follows:

$$\begin{aligned}\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}(u^n, v_0^n, v_1^n, m_{r_2}, m_t, \tilde{v}_2^n, z^n) &= \mathrm{Q}^n_{\mathrm{UV_0V_1}}(u^n, v_0^n, v_1^n)\, 2^{-nR_{r_2}} \tilde{E}^{\mathrm{LE}}(m_t|m_{r_2}, v_0^n, v_1^n) \\ &\quad \times \mathbb{1}_{\{\tilde{v}_2^n = \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\}}\, Q^n_{\mathrm{Z|V_0V_1V_2}}(z^n|v_0^n, v_1^n, v_2^n).\end{aligned} \tag{3.28}$$

Moreover, if we consider the generation process of the random codebook $\tilde{\mathrm{C}}_2$, we can define a probability distribution over the random variables $(\tilde{\mathrm{C}}, \mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \mathrm{M}_{r_2}, \mathrm{M}_t, \tilde{\mathrm{V}}_2^n, \mathrm{Z}^n)$ as follows: $\tilde{\mathrm{P}} \triangleq \mathrm{G}(\tilde{\mathcal{C}}_2) \cdot \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$, where $\mathrm{G}(\tilde{\mathcal{C}}_2)$ is the probability of generating a certain codebook $\tilde{\mathcal{C}}_2$ given by (3.24), while $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ is given by (3.28). Beside $\tilde{\mathrm{P}}$, we will need another probability distribution on the random variables $(\tilde{\mathrm{C}}, \mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \mathrm{M}_{r_2}, \mathrm{M}_t, \tilde{\mathrm{V}}_2^n, \mathrm{Z}^n)$ defined as follows: $\Gamma = \mathrm{G}(\tilde{\mathcal{C}}_2) \cdot \Gamma^{\tilde{\mathcal{C}}_2}$, where $\Gamma^{\tilde{\mathcal{C}}_2}$ is given by:

$$\begin{aligned}\Gamma^{\tilde{\mathcal{C}}_2}(u^n, v_0^n, m_{r_2}, m_t, \tilde{v}_2^n, v_1^n, z^n) &= \mathrm{Q}^n_{\mathrm{UV_0}}(u^n, v_0^n)\, 2^{-n(R_{r_2}+R_t)}\, \mathbb{1}_{\{\tilde{v}_2^n = \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\}} \\ &\quad \times \mathrm{Q}^n_{\mathrm{V_1|V_0V_2}}(v_1^n|v_0^n, \tilde{v}_2^n)\, Q^n_{\mathrm{Z|V_0V_1V_2}}(z^n|v_0^n, v_1^n, v_2^n).\end{aligned} \tag{3.29}$$

We will refer to $\Gamma^{\tilde{\mathcal{C}}_2}$ as the ideal code distribution, because one can argue that the process of producing the sequences $(u^n, v_0^n, v_1^n, \tilde{v}_2^n)$ according to $\Gamma^{\tilde{\mathcal{C}}_2}$ is identical to generating these sequences using the joint input distribution $\mathrm{Q}^n_{\mathrm{UV_0V_1V_2}}$. At first (u^n, v_0^n) are generated using $\mathrm{Q}^n_{\mathrm{UV_0}}$, then a sequence $\tilde{v}_2^n$ is selected uniformly at random from the codebook $\tilde{\mathcal{C}}_2(v_0^n)$ which is constructed using $\mathrm{Q}^n_{\mathrm{V_2|V_0}}(\cdot|v_0^n)$, finally a sequence v_1^n is generated conditioned on v_0^n and the selected sequence $\tilde{v}_2^n$ using $\mathrm{Q}^n_{\mathrm{V_1|V_0V_2}}$. It is important to highlight that the two distributions $\tilde{\mathrm{P}}$ and Γ will play an important role not only in our reliability analysis, but in our secrecy analysis as well.

We are now ready to proceed with our reliability analysis, where we show that the coding scheme defined in Section 3.2.2 and denoted by the random code C satisfies the reliability constraint in (3.15b). We start by defining the average decoding error probability for a given code realization $\mathcal{C} \in \mathfrak{C}$ as follows:

$$\begin{aligned}\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \triangleq \mathbb{P}\Big[&(\hat{\mathrm{M}}_c, \hat{\mathrm{M}}_0, \hat{\mathrm{M}}_1, \hat{\mathrm{M}}_r) \neq (\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_1, \mathrm{M}_r) \text{ or } (\tilde{\mathrm{M}}_c, \tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_2, \tilde{\mathrm{M}}_r) \neq \\ &(\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_2, \mathrm{M}_r) \text{ or } \check{\mathrm{M}}_c \neq \mathrm{M}_c\Big].\end{aligned} \tag{3.30}$$

We then observe that the previous error probability is greater than or equal to the one in (3.7), i.e. $\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \geq \bar{\mathbf{P}}_e(\mathcal{C})$. Our target is to derive the rate constraints under which $\mathbb{E}_{\mathrm{C}}[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})]$ approaches zero as n approaches infinity. Due to the symmetricity of the code with respect to the uniformly distributed variables $(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{M}_{r_2})$, we may restrict our analysis to the case where $(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{M}_{r_2}) = (1, 1, 1, 1, 1)$, where $\mathrm{M} = (\mathrm{M}_0, \mathrm{M}_1, \mathrm{M}_2)$. Thus, while calculating the probability of any event $\mathcal{A}$ with respect to P, we use the following notation:

$$\mathbb{P}^1_{\mathrm{P}}(\mathcal{A}) = \mathbb{P}_{\mathrm{P}}(\mathcal{A}|\mathrm{M}_c = 1, \mathrm{M} = 1, \mathrm{M}_r = 1, \mathrm{M}_{r_1} = 1, \mathrm{M}_{r_2} = 1). \tag{3.31}$$

Due to the structure of our encoding function and the fact that the decoders at the three receivers are based on the arguments of joint typicality, we need to consider two classes of errors:

1- Encoding Error: This type of error occurs if the codeword v_2^n chosen by the likelihood encoder and the codewords (u^n, v_0^n, v_1^n) selected uniformly at random are not jointly typical. With this in mind, we define the following event:

$$\mathcal{E} = \Big\{ \big(\mathrm{U}^n(\mathrm{M}_c), \mathrm{V}_0^n(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r), \mathrm{V}_1^n(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{M}_{r_1}), \mathrm{V}_2^n(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{M}_{r_2}, \mathrm{M}_t)\big) \notin \mathcal{T}_{\bar{\epsilon},\mathrm{inp}}^n \Big\},$$

where $\mathcal{T}_{\bar{\epsilon},\mathrm{inp}}^n \triangleq \mathcal{T}_{\bar{\epsilon}}^n(\mathrm{Q}_{\mathrm{UV_0V_1V_2}})$ and $\bar{\epsilon} \in (0, \epsilon)$, where ϵ is the typicality constant used by the decoding functions. Based on the symmetricity argument, it follows that $\mathbb{E}_{\mathrm{C}}[\mathbb{P}_{\mathrm{P}}(\mathcal{E})] = \mathbb{E}_{\mathrm{C}}[\mathbb{P}_{\mathrm{P}}^1(\mathcal{E})]$. Thus, we only consider the latter expression as follows:

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}\big[\mathbb{P}_{\mathrm{P}}^1(\mathcal{E})\big] &= \mathbb{E}_{\mathrm{C}}\Big[\mathbb{P}_{\mathrm{P}}\Big(\big(\mathrm{U}^n(1), \mathrm{V}_0^n(1,1,1), \mathrm{V}_1^n(1,1,1,1), \mathrm{V}_2^n(1,1,1,1,\mathrm{M}_t)\big) \notin \mathcal{T}_{\bar{\epsilon},\mathrm{inp}}^n\Big)\Big] \\
&= \mathbb{E}_{\mathrm{C}}\Bigg[\sum_{m_t,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n} \mathbb{1}_{\{(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n)\notin\mathcal{T}_{\bar{\epsilon},\mathrm{inp}}^n\}}\, \mathbb{1}_{\left\{\left(\mathrm{U}^n(1),\mathrm{V}_0^n(1,1,1),\mathrm{V}_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} \\
&\qquad \times E^{\mathrm{LE}}(m_t|1,\bar{v}_0^n,\bar{v}_1^n)\, \mathbb{1}_{\{\mathrm{V}_2^n(1,1,1,1,m_t)=\bar{v}_2^n\}}\Bigg] \\
&\overset{(a)}{=} \mathbb{E}_{\mathrm{C}_{c,0,1}}\Bigg[\sum_{m_t,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n} \mathbb{1}_{\{(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n)\notin\mathcal{T}_{\bar{\epsilon},\mathrm{inp}}^n\}}\, \mathbb{1}_{\left\{\left(\mathrm{U}^n(1),\mathrm{V}_0^n(1,1,1),\mathrm{V}_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} \\
&\qquad \times \mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}}\Big[E^{\mathrm{LE}}(m_t|1,\bar{v}_0^n,\bar{v}_1^n)\, \mathbb{1}_{\left\{\mathrm{V}_2^n(1,1,1,1,m_t)=\bar{v}_2^n\right\}}\Big]\Bigg],
\end{aligned} \tag{3.32}$$

where $\mathrm{C}_{c,0,1} \triangleq \{\mathrm{C}_c, \mathrm{C}_0, \mathrm{C}_1\}$. Step (a) follows from the law of total expectation [128] by dividing the expectation over C into an outer expectation over $\mathrm{C}_{c,0,1}$ and an inner one over C_2 conditioned on $\mathrm{C}_{c,0,1}$. For a fixed $(m_t, \bar{u}^n, \bar{v}_0^n, \bar{v}_1^n, \bar{v}_2^n)$ and a certain codebook realization $\mathcal{C} \in \mathfrak{C}$, we redefine the likelihood encoder as follows:

$$E^{\mathrm{LE}}(m_t|1,\bar{v}_0^n,\bar{v}_1^n) = \begin{cases} 0 & \text{if } \bar{u}^n \neq u^n(1) \text{ or } \bar{v}_0^n \neq v_0^n(1,1,1) \text{ or } \bar{v}_1^n \neq v_1^n(1,1,1,1) \\ \text{as given by Eq. (3.20)} & \text{otherwise.} \end{cases} \tag{3.33}$$

Due to the second indicator function in (3.32), the previous formulation does not affect the calculation of the expression in (3.32). This because the indicator function implies that the inner expectation only contributes to the total term when the sequences $(\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n)$ are identical to the codewords $\big(u^n(1), v_0^n(1,1,1), v_1^n(1,1,1,1)\big)$ selected by a certain codebook realization $\mathcal{C}_{c,0,1}$. Thus, we have

$$\begin{aligned}
&\mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}=\mathcal{C}_{c,0,1}}\Big[E^{\mathrm{LE}}(m_t|1,\bar{v}_0^n,\bar{v}_1^n)\, \mathbb{1}_{\left\{\mathrm{V}_2^n(1,1,1,1,m_t)=\bar{v}_2^n\right\}}\Big] \overset{(a)}{=} \mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}=\mathcal{C}_{c,0,1}} \\
&\qquad \times \Big[\mathbb{1}_{\left\{\left(\mathrm{u}^n(1),\mathrm{v}_0^n(1,1,1),\mathrm{v}_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} E^{\mathrm{LE}}\big(m_t|1,\bar{v}_0^n,\bar{v}_1^n\big)\mathbb{1}_{\left\{\mathrm{V}_2^n(1,1,1,1,m_t)=\bar{v}_2^n\right\}}\Big] \\
&\overset{(b)}{\leq} \mathbb{E}_{\mathrm{C}_2|\mathrm{U}^n(1)=\bar{u}^n,\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n,\mathrm{V}_1^n(1,1,1,1)=\bar{v}_1^n}\Big[E^{\mathrm{LE}}\big(m_t|1,\bar{v}_0^n,\bar{v}_1^n\big)\mathbb{1}_{\left\{\mathrm{V}_2^n(1,1,1,1,m_t)=\bar{v}_2^n\right\}}\Big|\mathrm{C}_2(1,1,1)\Big] \\
&\overset{(c)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}\big(m_t,\bar{v}_2^n|1,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n\big)\Big]
\end{aligned} \tag{3.34}$$

where $\tilde{\mathrm{C}}_2$ is the random codebook introduced in the beginning of this section and $\tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}$ is its corresponding induced probability given by (3.26). Step (a) follows from (3.33); (b) follows by removing the indicator function and because the expression in the expectation only depends on the codewords $\mathrm{U}^n(1)$, $\mathrm{V}_0^n(1,1,1)$ and $\mathrm{V}_1^n(1,1,1,1)$ in addition to the codebook $\mathrm{C}_2(1,1,1)$ and not on the entire codebook C; while (c) follows from the definition of $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ in (3.26). Now, by substituting Eq. (3.34) in Eq. (3.32), we have

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}\big[\mathbb{P}_{\mathrm{P}}^1(\mathcal{E})\big] &= \mathbb{E}_{\mathrm{C}_{c,0,1}}\Bigg[\sum_{m_t,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n} \mathbb{1}_{\{(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}\}} \\
&\qquad \times \mathbb{1}_{\left\{\left(\mathrm{U}^n(1),\mathrm{V}_0^n(1,1,1),\mathrm{V}_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}\big(m_t,\bar{v}_2^n|1,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n\big)\Big]\Bigg] \\
&\overset{(a)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Bigg[\sum_{m_t,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n} \mathbb{1}_{\{(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}\}} \tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}\big(m_t,\bar{v}_2^n|1,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n\big) \\
&\qquad \times \mathbb{E}_{\mathrm{C}_{c,0,1}}\Bigg[\mathbb{1}_{\left\{\left(\mathrm{U}^n(1),\mathrm{V}_0^n(1,1,1),\mathrm{V}_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}}\Bigg]\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Bigg[\sum_{m_t,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n} \mathbb{1}_{\{(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n,\bar{v}_2^n)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}\}} \tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}\big(m_t,\bar{v}_2^n|1,\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n\big)\,\mathrm{Q}^n_{\mathrm{UV_0V_1}(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)}\Bigg] \\
&\overset{(c)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\mathbb{P}_{\tilde{\mathrm{P}}}\Big(\big(\mathrm{U}^n,\mathrm{V}_0^n,\mathrm{V}_1^n,\tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n,1,\mathrm{M}_t)\big)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}|\tilde{\mathrm{C}}_2\Big)\Big],
\end{aligned} \tag{3.35}$$

where (a) follows because $\tilde{\mathrm{C}}_2$ and $\mathrm{C}_{c,0,1}$ are independent; (b) is due to the following property of the indicator function: $\mathbb{E}_{\mathrm{P}}[\mathbb{1}_{\mathcal{A}}] = \mathbb{P}_{\mathrm{P}}(\mathcal{A})$ [129] and the fact that the random codebook $\mathrm{C}_{c,0,1}$ was generated using the input distribution $\mathrm{Q}_{\mathrm{UV_0V_1}}$; while (c) follows from the definition of $\tilde{\mathrm{P}}$ and the established notation $\mathbb{P}_{\tilde{\mathrm{P}}}$. In order to evaluate the expression in (3.35), we need to carry out some further steps. First, we recall the probability distribution Γ introduced in (3.29). Based on the weak law of large numbers [130] and the properties of strongly typical sequences in Theorem A.5, for every $m_{r_2} \in \mathcal{M}_{r_2}$, we have:

$$\mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\mathbb{P}_{\Gamma}\Big(\big(\mathrm{U}^n,\mathrm{V}_0^n,\mathrm{V}_1^n,\tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n,m_{r_2},\mathrm{M}_t)\big)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}|\tilde{\mathrm{C}}_2\Big)\Big] \leq \delta_n^{\mathrm{inp}}(\epsilon), \tag{3.36}$$

$$\text{where } \delta_n^{\mathrm{inp}}(\epsilon) = 2\cdot|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\cdot e^{-2n\epsilon^2\mu_{\mathrm{inp}}^2}. \tag{3.37}$$

The value μ_{inp} is given by the minimum probability over the whole support of the input distribution $\mathrm{Q}_{\mathrm{UV_0V_1V_2}}$. One can easily show that $\delta_n^{\mathrm{inp}}(\epsilon)$ approaches zero exponentially fast as n approaches infinity.

Next, we highlight an important property for the total variation distance introduced in [131, Property 1]: Let γ,η be two probability measures on a measurable space $(\mathcal{H},\mathcal{L})$ and $f:\mathcal{H}\to\mathbb{R}$ be a measurable function bounded by $b\in\mathbb{R}$. It follows that:

$$|\mathbb{E}_\gamma f - \mathbb{E}_\eta f| \leq b\cdot\|\gamma-\eta\|. \tag{3.38}$$

In order to use the previous property, we let $f_n:\mathcal{U}^n\times\mathcal{V}_0^n\times\mathcal{V}_1^n\times\mathcal{V}_2^n\to\mathbb{R}$ be defined as follows $f_n(u^n,v_0^n,v_1^n,\tilde{v}_2^n) \triangleq \mathbb{1}_{\{(u^n,v_0^n,v_1^n,\tilde{v}_2^n)\notin\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}\}}$ and observe that f_n is bounded by

$b = 1$. Moreover, we use the property of the expectation of an indicator function: $\mathbb{E}_{\mathrm{P}}[\mathbb{1}_{\mathcal{A}}] = \mathbb{P}_{\mathrm{P}}(\mathcal{A})$. Based on these two arguments, we have

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}\big[\mathbb{P}^1_{\mathrm{P}}(\mathcal{E})\big] &= \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\mathbb{E}_{\tilde{\mathrm{P}}}\Big[f_n\big(\mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n, 1, \mathrm{M}_t)\big)\big|\tilde{\mathrm{C}}_2\Big]\Big] \\
&\leq \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\mathbb{E}_{\Gamma}\Big[f_n\big(\mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n, 1, \mathrm{M}_t)\big)\big|\tilde{\mathrm{C}}_2\Big]\Big] \\
&+ \mathbb{E}_{\tilde{\mathrm{C}}_2}\Bigg[\Bigg|\mathbb{E}_{\tilde{\mathrm{P}}}\Big[f_n\big(\mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n, 1, \mathrm{M}_t)\big)\big|\tilde{\mathrm{C}}_2\Big] \\
&\qquad - \mathbb{E}_{\Gamma}\Big[f_n\big(\mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n, \tilde{\mathrm{V}}_2^n(\mathrm{V}_0^n, 1, \mathrm{M}_t)\big)\big|\tilde{\mathrm{C}}_2\Big]\Bigg|\Bigg] \\
&\overset{(a)}{\leq} \delta_n^{\mathrm{inp}}(\epsilon) + \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n|\tilde{\mathrm{C}}_2} - \Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n|\tilde{\mathrm{C}}_2}\right\|\right] \\
&\leq \delta_n^{\mathrm{inp}}(\epsilon) + \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\right\|\right], \qquad (3.39)
\end{aligned}$$

where (a) follows by letting $m_{r_2} = 1$ in (3.36), along with applying the property of total variation distance in (3.38). It is important to point out here that the last step where we include the random variable Z^n in calculating the total variation distance is not actually needed for the encoding error analysis. However, we include it because it is a crucial part for our secrecy analysis.

We are now one step away from finalizing our encoding error analysis as we only need to bound the total variation distance in (3.39). In order to do so, we need the following relations:

$$\Gamma^{\tilde{\mathcal{C}}_2}_{\mathrm{M}_t|\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} = \tilde{E}^{\mathrm{LE}}_{\mathrm{M}_t|\mathrm{M}_{r_2}\mathrm{V}_0^n\mathrm{V}_1^n} = \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}_{\mathrm{M}_t|\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} \tag{3.40a}$$

$$\Gamma^{\tilde{\mathcal{C}}_2}_{\tilde{\mathrm{V}}_2^n|\mathrm{M}_t\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} = \mathbb{1}_{\{\tilde{\mathrm{V}}_2^n = \tilde{v}_2^n(\mathrm{V}_0^n, \mathrm{M}_{r_2}, \mathrm{M}_t)\}} = \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}_{\tilde{\mathrm{V}}_2^n|\mathrm{M}_t\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} \tag{3.40b}$$

$$\Gamma^{\tilde{\mathcal{C}}_2}_{\mathrm{Z}^n|\tilde{\mathrm{V}}_2^n\mathrm{M}_t\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} = \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2} = \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}_{\mathrm{Z}^n|\tilde{\mathrm{V}}_2^n\mathrm{M}_t\mathrm{M}_{r_2}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n} \tag{3.40c}$$

The relations in (3.40b) and (3.40c) follow directly from the definitions of $\Gamma^{\tilde{\mathcal{C}}_2}$ in (3.29) and $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ in (3.26). On the other hand, the relation in (3.40a) follows because for every $(u^n, v_0^n, v_1^n, m_{r_2}, m_t) \in \mathcal{U}^n \times \mathcal{V}_0^n \times \mathcal{V}_0^n \times \mathcal{M}_{r_2} \times \mathcal{M}_t$, we have

$$\begin{aligned}
\Gamma^{\tilde{\mathcal{C}}_2}(m_t|m_{r_2}, u^n, v_0^n, v_1^n) &= \frac{\Gamma^{\tilde{\mathcal{C}}_2}(m_t, m_{r_2}, u^n, v_0^n, v_1^n)}{\Gamma^{\tilde{\mathcal{C}}_2}(m_t, u^n, v_0^n, v_1^n)} \\
&= \frac{\sum_{\tilde{v}_2^n} \mathrm{Q}^n_{\mathrm{UV}_0}(u^n, v_0^n)\, 2^{-n(R_{r_2}+R_t)}\, \mathbb{1}_{\{\tilde{v}_2^n = \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\}} \times \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0\mathrm{V}_2}(v_1^n|v_0^n, \tilde{v}_2^n)}{\sum_{\tilde{v}_2^n, j \in \mathcal{M}_t} \mathrm{Q}^n_{\mathrm{UV}_0}(u^n, v_0^n)\, 2^{-n(R_{r_2}+R_t)}\, \mathbb{1}_{\{\tilde{v}_2^n = \tilde{v}_2^n(v_0^n, m_{r_2}, j)\}} \times \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0\mathrm{V}_2}(v_1^n|v_0^n, \tilde{v}_2^n)} \\
&= \frac{\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n|v_0^n, \tilde{v}_2^n(v_0^n, m_{r_2}, m_t)\big)}{\sum_{j \in \mathcal{M}_t} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n|v_0^n, \tilde{v}_2^n(v_0^n, m_{r_2}, j)\big)} \\
&\overset{(a)}{=} \tilde{E}^{\mathrm{LE}}(m_t|m_{r_2}, v_0^n, v_1^n) \overset{(b)}{=} \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}(m_t|m_{r_2}, u^n, v_0^n, v_1^n),
\end{aligned}$$

where (a) follows from the definition of the likelihood encoder cf. Eq. (3.25) and (b) follows from (3.26). We now turn to the total variation distance expression in (3.39), and observe that it can be reformulated as follows:

$$
\begin{aligned}
&\mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\right\|\right] \\
&\qquad\leq \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{M}_{r_2}\mathrm{M}_t\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{M}_{r_2}\mathrm{M}_t\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\right\|\right] \\
&\qquad\overset{(a)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{M}_t\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\mathrm{M}_{r_2},\tilde{\mathrm{C}}_2}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{M}_t\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\mathrm{M}_{r_2},\tilde{\mathrm{C}}_2}\right\|\right] \\
&\qquad\overset{(b)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n|\mathrm{M}_{r_2},\tilde{\mathrm{C}}_2}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n|\mathrm{M}_{r_2},\tilde{\mathrm{C}}_2}\right\|\right] \\
&\qquad\overset{(c)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n|\mathrm{M}_{r_2}=1,\tilde{\mathrm{C}}_2}\right\|\right]
\end{aligned}
\tag{3.41}
$$

where (a) follows because for every $m_{r_2} \in \mathcal{M}_{r_2}$ and $\tilde{\mathcal{C}}_2 \in \tilde{\mathfrak{C}}_2$, we have $\Gamma^{\tilde{\mathcal{C}}_2}(m_{r_2}) = \tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}(m_{r_2}) = 2^{-nR_{r_2}}$; (b) follows due to the relations established in (3.40); while (c) follows from the definition of $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ in (3.28), along with the symmetricity of $\Gamma^{\tilde{\mathcal{C}}_2}$ with respect to M_{r_2}. Now, for $\beta > 0$ and $0 < \bar{\epsilon} < \epsilon$, we define the following δ-functions:

$$
\delta_n^{\mathrm{inp}}(\bar{\epsilon},\epsilon) = 2\cdot|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\cdot\exp\left(-2n\left(\frac{\epsilon-\bar{\epsilon}}{1+\bar{\epsilon}}\right)^2\mu_{\mathrm{inp}}^2\right) \tag{3.42}
$$

$$
\delta_0(\epsilon,\beta) = \epsilon\log\big(|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\big) + 5\beta \tag{3.43}
$$

$$
\delta_n(\epsilon,\beta) = \big(1-n\delta_n^{\mathrm{inp}}(\epsilon)\big)\cdot\big(5\delta_n^{\mathrm{inp}}(\bar{\epsilon},\epsilon) + 7\cdot 2^{-n\beta}\big) \tag{3.44}
$$

Based on the result established in [132, Corollary VII.5] and the analysis technique used in Section 2.2.3, one can show that:

$$
\text{if}\quad R_t \geq \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon,\beta), \tag{3.45}
$$

$$
\text{then}\quad \mathbb{E}_{\tilde{\mathrm{C}}_2}\left[\left\|\mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1}-\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n|\mathrm{M}_{r_2}=1,\tilde{\mathrm{C}}_2}\right\|\right] \leq \delta_n(\epsilon,\beta). \tag{3.46}
$$

Based on the previous analysis, we can argue that as long as the rate constraint in (3.45) is satisfied, the following holds:

$$
\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}[\mathbb{P}_{\mathrm{P}}(\mathcal{E})] \leq \lim_{n\to\infty}\delta_n^{\mathrm{inp}}(\epsilon) + \delta_n(\epsilon,\beta) = 0 \tag{3.47}
$$

This implies that the rate constraint in (3.45) assures that the sequences generated and selected by our coding scheme are jointly typical with a probability approaching one as n approaches infinity. This is a very interesting result because in the proof of [12, Theorem 2], a coding scheme was used where the codebook generation method and the decoding functions are nearly identical to ours, but with a completely different encoding technique. In particular, the encoding function in [12] works as follows: Given the messages $(m_c, m, m_r, m_{r_1}, m_{r_2})$, the encoder searches for an index $m_t \in \mathcal{M}_t$, such that the corresponding triple $v_0^n(m_c, m, m_r)$, $v_1^n(m_c, m, m_r, m_{r_1})$, and $v_2^n(m_c, m, m_r, m_{r_2}, m_t)$ are jointly typical. Based on the typicality result in Theorem A.7, the encoder succeeds in finding such index with a probability that goes to 1 exponentially fast, if

$$
R_t > \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0). \tag{3.48}
$$

Surprisingly, the constraint in (3.45) required by our coding scheme is identical to the one in (3.48) which was derived in [12], despite the fact that each coding scheme utilizes a different encoding technique.

2- Decoding Errors: This type of errors occur if the decoders decide on a set of messages different from the ones which were originally transmitted. This will happen if the transmitted messages do not fulfill the decoding rule or if a set of messages different from the transmitted one satisfies the decoding rule. With this in mind, we define the following events:

$$\begin{aligned}
\mathcal{D}_1(m_c, m, m_r, m_{r_1}) &= \left\{ \left(\mathrm{U}^n(m_c), \mathrm{V}_0^n(m_c, m, m_r), \mathrm{V}_1^n(\ldots, m_{r_1}), \mathrm{Y}_1^n \right) \in \mathcal{T}_{\epsilon,1}^n \right\} \\
\mathcal{D}_2(m_c, m, m_r, m_{r_2}, \mathrm{M}_t) &= \left\{ \left(\mathrm{U}^n(m_c), \mathrm{V}_0^n(m_c, m, m_r), \mathrm{V}_2^n(\ldots, m_{r_2}, \mathrm{M}_t), \mathrm{Y}_2^n \right) \in \mathcal{T}_{\epsilon,2}^n \right\} \\
\mathcal{D}_3(m_c) &= \left\{ \left(\mathrm{U}^n(m_c), \mathrm{Z}^n \right) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UZ}}) \right\}, \qquad (3.49)
\end{aligned}$$

where $\mathcal{T}_{\epsilon,1}^n = \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UV_0V_1Y_1}})$ and $\mathcal{T}_{\epsilon,2}^n = \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UV_0V_2Y_2}})$. Based on the definitions of the decoding functions, the expectation of the average decdoing error probability in (3.30) over the codebook ensemble can be bounded as follows:

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}\left[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})\right] &\overset{(a)}{=} \mathbb{E}_{\mathrm{C}}\Bigg[\mathbb{P}_{\mathrm{P}}^1\Bigg(\mathcal{E} \cup \left\{ \bigcup_{m_{r_1}} \mathcal{D}_1(1,1,1,m_{r_1}) \right\}^c \cup \left\{ \bigcup_{m_{r_2}} \mathcal{D}_2(1,1,1,m_{r_2},\mathrm{M}_t) \right\}^c \cup \\
&\qquad \cup \mathcal{D}_3^c(1) \cup \left\{ \bigcup_{m_c \neq 1} \mathcal{D}_3^c(m_c) \right\} \left\{ \bigcup_{\substack{(m_c,m,m_r)\neq(1,1,1)\\ m_{r_1}}} \mathcal{D}_1(m_c,m,m_r,m_{r_1}) \right\} \\
&\qquad \cup \left\{ \bigcup_{\substack{(m_c,m,m_r)\neq(1,1,1)\\ m_{r_2}}} \mathcal{D}_2(m_c,m,m_r,m_{r_2},\mathrm{M}_t) \right\} \Bigg)\Bigg] \\
&\overset{(b)}{\leq} \mathbb{E}_{\mathrm{C}}\Bigg[\mathbb{P}_{\mathrm{P}}^1(\mathcal{E}) + \underbrace{\mathbb{P}_{\mathrm{P}}^1\left(\mathcal{E}^c \cap \mathcal{D}_1^c(1,1,1,1)\right)}_{P_{e_0}^1} + \underbrace{\mathbb{P}_{\mathrm{P}}^1\left(\mathcal{E}^c \cap \mathcal{D}_2^c(1,1,1,1,\mathrm{M}_t)\right)}_{P_{e_0}^2} + \underbrace{\mathbb{P}_{\mathrm{P}}^1\left(\mathcal{E}^c \cap \mathcal{D}_3^c(1)\right)}_{P_{e_0}^3} \\
&\quad + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{\substack{(m_c,m,m_r)\neq(1,1,1),\\ m_{r_1}}} \mathcal{D}_1(m_c,m,m_r,m_{r_1})\Bigg)}_{P_{e_3}^1} + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{\substack{(m,m_r)\neq(1,1),\\ m_{r_1}}} \mathcal{D}_1(1,m,m_r,m_{r_1})\Bigg)}_{P_{e_2}^1} \\
&\quad + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{m_c\neq 1} \mathcal{D}_1(m_c,1,1,1)\Bigg)}_{P_{e_1}^1} + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{\substack{(m_c,m,m_r)\neq(1,1,1),\\ m_{r_1},m_t}} \mathcal{D}_2(m_c,m,m_r,m_{r_2},m_t)\Bigg)}_{P_{e_3}^2} \\
&\quad + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{\substack{(m,m_r)\neq(1,1),\\ m_{r_1},m_t}} \mathcal{D}_2(m_c,m,m_r,m_{r_2},m_t)\Bigg)}_{P_{e_2}^2} + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{m_c\neq 1} \mathcal{D}_2(m_c,1,1,1,\mathrm{M}_t)\Bigg)}_{P_{e_1}^2} \\
&\quad + \underbrace{\mathbb{P}_{\mathrm{P}}^1\Bigg(\bigcup_{m_c\neq 1} \mathcal{D}_3(m_c)\Bigg)}_{P_{e_1}^3}\Bigg], \qquad (3.50)
\end{aligned}$$

where (a) follows due to the symmetricity of the code with respect to the messages M_c, M, M_r, M_{r_1} and M_{r_2}, while (b) follows from Boole's inequality known as the union

bound cf. [133]. Next, we derive the necessary rates constraints such that the the average decoding error probability given by (3.50) vanishes as n approaches infinity. Thus, we have $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[\bar{\mathbf{P}}_e(\mathrm{C})\big] \leq \lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})\big] = 0$. We proceed as follows:

1. Based on our encoding analysis, it follows that $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[\mathbb{P}^1_{\mathrm{P}}(\mathcal{E})\big] = 0$ as long as the rate constraint in (3.45) holds.
2. Based on the properties of conditionally typical sequences given in Theorem A.6, it follows that $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[P^1_{e_0}+P^2_{e_0}+P^3_{e_0}\big] = 0$. In particular, for $0<\hat{\epsilon}<\epsilon$ each of these terms can be upper-bounded by the following function:

$$\delta^{\mathrm{Q}}_n(\hat{\epsilon},\epsilon) = 2\cdot|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\cdot|\mathcal{Y}_1|\cdot|\mathcal{Y}_2|\cdot|\mathcal{Z}|\exp\left(-2n\left(\frac{\epsilon-\hat{\epsilon}}{1+\hat{\epsilon}}\right)^2\mu^2_{\mathrm{Q}}\right), \quad (3.51)$$

   where μ_{Q} is the minimum probability of an event over the support of the joint distribution $\mathrm{Q}_{\mathrm{UV_0V_1V_2Y_1Y_2Z}}$.
3. Based on the properties of jointly typical sequences given in Theorem A.7, $\mathbb{E}_{\mathrm{C}}\big[P^1_{e_3}\big]$ can be bounded as follows:

$$\mathbb{E}_{\mathrm{C}}\big[P^1_{e_3}\big] \leq \sum_{m_{r_1}}\sum_{(m_c,m,m_r)\neq(1,1,1)} 2^{-n\left(\mathbb{I}(\mathrm{V_0V_1};\mathrm{Y_1})-\delta_1(\epsilon)\right)} \quad (3.52)$$

   where $\delta_1(\epsilon) = 2\epsilon\log\big(|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{Y}_1|\big)$. The previous bound follows because for any $(m_c,m,m_r)\neq(1,1,1)$ and $m_{r_1}\in\mathcal{M}_{r_1}$, the corresponding codewords $\mathrm{U}^n(m_c)$, $\mathrm{V}^n_0(m_c,m,m_r)$ and $\mathrm{V}^n_1(m_c,m,m_r,m_{r_1})$ are independent of the observation at the first legitimate receiver Y^n_1. Since M_2 is available as a side information at the first legitimate receiver, Eq. (3.52) directly implies that $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[P^1_{e_3}\big] = 0$, as long as:

$$R_c+R_0+R_1+R_r+R_{r_1} \leq \mathbb{I}(\mathrm{V_0V_1};\mathrm{Y_1})-\delta_1(\epsilon). \quad (3.53)$$

4. Based on similar arguments, $\mathbb{E}_{\mathrm{C}}\big[P^2_{e_3}\big]$ can be bounded as follows:

$$\mathbb{E}_{\mathrm{C}}\big[P^2_{e_3}\big] \leq \sum_{m_{r_2},m_t}\sum_{(m_c,m,m_r)\neq(1,1,1)} 2^{-n\left(\mathbb{I}(\mathrm{V_0V_2};\mathrm{Y_1})-\delta_2(\epsilon)\right)}, \quad (3.54)$$

   where $\delta_2(\epsilon) = 2\epsilon\log\big(|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_2|\cdot|\mathcal{Y}_2|\big)$. Similarly, since M_1 is available as a side information at the second legitimate receiver, Eq. (3.54) directly implies that $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[P^2_{e_3}\big] = 0$, as long as:

$$R_c+R_0+R_2+R_r+R_{r_2}+R_t \leq \mathbb{I}(\mathrm{V_0V_2};\mathrm{Y_2})-\delta_2(\epsilon). \quad (3.55)$$

5. Based on the conditional version of the results presented in Theorem A.7, $\mathbb{E}_{\mathrm{C}}\big[P^1_{e_2}\big]$ can be bounded as follows:

$$\mathbb{E}_{\mathrm{C}}\big[P^1_{e_2}\big] \leq \sum_{m_{r_1}}\sum_{(m,m_r)\neq(1,1)} 2^{-n\left(\mathbb{I}(\mathrm{V_0V_1};\mathrm{Y_1}|\mathrm{U})-\delta_1(\epsilon)\right)}, \quad (3.56)$$

   The previous bound follows because for any $m_c=1$, $(m,m_r)\neq(1,1)$ and $m_{r_1}\in\mathcal{M}_{r_1}$, the corresponding codewords $\mathrm{V}^n_0(1,m,m_r)$ and $\mathrm{V}^n_1(1,m,m_r,m_{r_1})$ are independent of Y^n_1 while both of them are generated conditioned on the codeword $\mathrm{U}^n(1)$. This implies that $\lim_{n\to\infty}\mathbb{E}_{\mathrm{C}}\big[P^1_{e_2}\big] = 0$, as long as:

$$R_0+R_1+R_r+R_{r_1} \leq \mathbb{I}(\mathrm{V_0V_1};\mathrm{Y_1}|\mathrm{U})-\delta_1(\epsilon). \quad (3.57)$$

Applying the same arguments to the second legitimate receiver directly implies that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[P_{e_2}^2\big] = 0$, as long as:

$$R_0 + R_2 + R_r + R_{r_2} + R_t \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) - \delta_2(\epsilon). \tag{3.58}$$

6. Finally, using a similar argument to the one used to bound $P_{e_3}^1$ in (3.52), we can show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[P_{e_1}^1 + P_{e_1}^2 + P_{e_1}^3\big] = 0$ as long as:

$$\begin{aligned} R_c &\leq \min\Big(\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \delta_1(\epsilon), \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \delta_2(\epsilon), \mathbb{I}(\mathrm{U};\mathrm{Z}) - \delta_3(\epsilon)\Big) \\ &\overset{(a)}{\leq} \mathbb{I}(\mathrm{U};\mathrm{Z}) - \delta_3(\epsilon). \end{aligned} \tag{3.59}$$

where $\delta_3(\epsilon) = 2\epsilon \log\big(|\mathcal{U}| \cdot |\mathcal{Z}|\big)$, while (a) follows from (3.53) and (3.55). The previous analysis shows that under the rate constraints in (3.45), (3.53), (3.55), (3.57), (3.58) and (3.59), our coding scheme satisfies the reliability constraint in (3.15b).

3.2.4 Strong Secrecy Analysis

In this section, we present a detailed strong secrecy analysis for our coding scheme. In particular, we derive the rest of the rate bounds given in (3.23). These rate bounds assure that our random code C defined in Section 3.2.2 satisfies the joint-strong secrecy constraint in (3.15c). We start by arguing that for $\tau > 0$, it is enough to show that:

$$\mathbb{E}_{\mathrm{C}}\left[\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{C})\right] \overset{(a)}{\leq} \mathbb{E}_{\mathrm{C}}\left[\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{M}_c\mathrm{C})\right] \leq 2^{-n\tau/2}, \tag{3.60}$$

where (a) follows because M and M_c are independent. Based on the relation between the mutual information and the total variation distance given in Lemma A.6, one can show that the bound in (3.60) holds, if:

$$\mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{M}\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}} - \mathrm{P}_{\mathrm{M}|\mathrm{M}_c,\mathrm{C}} \times \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}\big\|\Big] \leq 2 \cdot 2^{-n\tau}, \tag{3.61}$$

where P is the whole distribution on the random code C given by (3.22). Moreover, since M is uniformly-distributed over the set $\mathcal{M}$ and is independent of the common message M_c and the random codebook C, Eq. (3.61) holds, if:

$$\mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}} - \check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}\big\|\Big] \leq 2^{-n\tau}, \tag{3.62}$$

where $\check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}$ is the probability distribution observed by the eavesdropper when the common message M_c is the only useful information transmitted over the channel using the random code C. The definition of such distribution follows from $\check{\mathrm{P}} = \mathrm{G}(\mathcal{C}) \cdot \check{\mathrm{P}}^{\mathcal{C}}$, where $\mathrm{G}(\mathcal{C})$ is as defined in (3.19), while $\check{\mathrm{P}}^{\mathcal{C}}$ is defined as follows:

$$\check{\mathrm{P}}^{\mathcal{C}}\big(m_c, u^n, z^n\big) = 2^{-nR_c}\, \mathbb{1}_{\{u^n = u^n(m_c)\}}\, Q^n_{\mathrm{Z}|\mathrm{U}}(z^n|u^n), \tag{3.63}$$

The previous definition implies that $\check{\mathrm{P}}^{\mathcal{C}}$ is only a function in the common message codebook $\mathcal{C}_c$, while being independent of $\mathcal{C}_{0,1,2}$. Moreover, $Q^n_{\mathrm{Z}|\mathrm{U}}$ can be interpreted as the DMC from $\mathcal{U}^n$ to $\mathcal{Z}^n$ given by:

$$Q^n_{\mathrm{Z}|\mathrm{U}}(z^n|u^n) \triangleq \sum_{v_0^n \in \mathcal{V}_0^n} \sum_{v_1^n \in \mathcal{V}_1^n} \sum_{v_2^n \in \mathcal{V}_2^n} \mathrm{Q}^n_{\mathrm{V}_0|\mathrm{U}}(v_0^n|u^n) Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}(z^n|v_0^n, v_1^n, v_2^n). \tag{3.64}$$

Before we proceed with our secrecy analysis, we need to introduce some useful random codes and their induced joint distributions. Our first random code is defined as follows: Given two random sequences $u^n \in \mathcal{U}^n$ and $v_0^n \in \mathcal{V}_0^n$ generated according to the input distribution $\mathrm{Q}^n_{\mathrm{UV}_0}$, let $\ddot{\mathrm{C}}_1(v_0^n) \triangleq \{\ddot{\mathrm{V}}_1^n(v_0^n, m_{r_1})\}$ be a collection of i.i.d. random codewords of length n generated according to $\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}(\cdot|v_0^n)$, for every $m_{r_1} \in \mathcal{M}_{r_1}$. Additionally, let $\ddot{\mathrm{C}}_1 \triangleq \ddot{\mathrm{C}}_1(v_0^n)$ be a random codebook generated for all $v_0^n \in \mathcal{V}_0^n$. This implies that the probability of obtaining a certain codebook realization $\ddot{\mathcal{C}}_1 \in \ddot{\mathfrak{C}}_1$ is:

$$\mathrm{G}(\ddot{\mathcal{C}}_1) = \prod_{v_0^n \in \mathcal{V}_0^n} \prod_{m_{r_1} \in \mathcal{M}_{r_1}} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}(\ddot{v}_1^n(v_0^n, m_{r_1})|v_0^n). \tag{3.65}$$

The encoding function works as follows: For a fixed codebook realization $\ddot{\mathcal{C}}_1$ and given the sequences (u^n, v_0^n), we choose an index m_{r_1} uniformly at random from the set $\mathcal{M}_{r_1}$ and transmit the corresponding codeword $\ddot{v}_1^n(v_0^n, m_{r_1}) \in \mathcal{V}_1^n$ over the DMC $Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}$:

$$Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(z^n|v_0^n, v_1^n) \triangleq \sum_{v_2^n \in \mathcal{V}_2^n} \mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0\mathrm{V}_1}(v_2^n|v_0^n, v_1^n)\, Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}(z^n|v_0^n, v_1^n, v_2^n). \tag{3.66}$$

The codebook generation process along with the encoding scheme and the DMC in (3.66) induce a joint probability distribution over the random variables $\big(\ddot{\mathrm{C}}_1, \mathrm{U}^n, \mathrm{V}_0^n, \mathrm{M}_{r_1}, \ddot{\mathrm{V}}_1^n, \mathrm{Z}^n\big)$ as follows: $\ddot{\Gamma} \triangleq \mathrm{G}(\ddot{\mathcal{C}}_1) \cdot \ddot{\Gamma}^{\ddot{\mathcal{C}}_1}$, where $\ddot{\Gamma}^{\ddot{\mathcal{C}}_1}$ is given by:

$$\ddot{\Gamma}^{\ddot{\mathcal{C}}}\big(u^n, v_0^n, m_{r_1}, \ddot{v}_1^n, z^n\big) = \mathrm{Q}^n_{\mathrm{UV}_0}(u^n, v_0^n)\, 2^{-nR_{r_1}}\, \mathbb{1}_{\{\ddot{v}_1^n = \ddot{v}_1^n(v_0^n, m_{r_1})\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(z^n|v_0^n, \ddot{v}_1^n). \tag{3.67}$$

In contrast to $\ddot{\Gamma}$, we define a probability distribution related to the random code C as follows: $\ddot{\mathrm{P}} \triangleq \mathrm{G}(\mathcal{C}) \cdot \ddot{\mathrm{P}}^{\mathcal{C}}$, where $\ddot{\mathrm{P}}^{\mathcal{C}}$ is given by:

$$\begin{aligned}\ddot{\mathrm{P}}^{\mathcal{C}}\big(m_c, m, m_r, u^n, v_0^n, m_{r_1}, v_1^n, z^n\big) &= 2^{-n(R_c+R+R_r)}\, \mathbb{1}_{\{u^n = u^n(m_c)\}}\, \mathbb{1}_{\{v_0^n = v_0^n(m_c, m, m_r)\}} \\ &\quad \times 2^{-nR_{r_1}}\, \mathbb{1}_{\{v_1^n = v_1^n(m_c, m, m_r, m_{r_1})\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(z^n|v_0^n, v_1^n).\end{aligned} \tag{3.68}$$

We need to point out that although we define $\ddot{\mathrm{P}}$ over the whole random codebook C, the distribution $\ddot{\mathrm{P}}^{\mathcal{C}}$ depends only on the codebooks $\mathcal{C}_{c,0,1}$, while being independent of $\mathcal{C}_2$. We also need to specify that $\ddot{\mathrm{P}}^{\mathcal{C}}$ can be interpreted as the multiplication of the marginal distribution of $\mathrm{P}^{\mathcal{C}}$ in (3.22) over the random variables $\big(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n\big)$ and the DMC $Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}$.

Next, we consider another random code defined as follows: Given a random sequence $u^n \in \mathcal{U}^n$ generated by $\mathrm{Q}^n_{\mathrm{U}}$, let $\hat{\mathrm{C}}_0(u^n) \triangleq \{\hat{\mathrm{V}}_0^n(u^n, m, m_r)\}$ be a collection of i.i.d. random codewords of length n generated according to $\mathrm{Q}^n_{\mathrm{V}_0|\mathrm{U}}(\cdot|u^n)$, for every $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$. Additionally, let $\hat{\mathrm{C}}_0 \triangleq \hat{\mathrm{C}}_0(u^n)$ be a random codebook generated for all $u^n \in \mathcal{U}^n$. This implies that the probability of obtaining a certain codebook realization $\hat{\mathcal{C}}_0 \in \hat{\mathfrak{C}}_0$ is:

$$\mathrm{G}(\hat{\mathcal{C}}_0) = \prod_{u^n \in \mathcal{U}^n} \prod_{(m, m_r) \in \mathcal{M} \times \mathcal{M}_r} \mathrm{Q}^n_{\mathrm{V}_0|\mathrm{U}}(\hat{v}_0^n(u^n, m, m_r)|u^n). \tag{3.69}$$

The encoding function works as follows: For a fixed codebook realization $\hat{\mathcal{C}}_0$ and given pair (u^n, m), we choose an index m_r uniformly at random from the set $\mathcal{M}_r$ and transmit

the corresponding codeword $\hat{v}_0^n(u^n, m, m_r) \in \mathcal{V}_0^n$ over the DMC $Q^n_{\mathrm{Z}|\mathrm{V}_0}$:

$$Q^n_{\mathrm{Z}|\mathrm{V}_0}(z^n|v_0^n) \triangleq \sum_{v_1^n \in \mathcal{V}_1^n} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}(v_1^n|v_0^n)\, \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0,\mathrm{V}_1}(z^n|v_0^n, v_1^n). \tag{3.70}$$

The codebook generation process along with the encoding scheme and the DMC in (3.70) induce a joint probability distribution over the random variables $\big(\hat{\mathrm{C}}_0, \mathrm{U}^n, \mathrm{M}, \mathrm{M}_r, \hat{\mathrm{V}}_0^n, \mathrm{Z}^n\big)$ as follows: $\hat{\Gamma} \triangleq \mathrm{G}(\hat{\mathcal{C}}_0) \cdot \hat{\Gamma}^{\hat{\mathcal{C}}_0}$, where $\hat{\Gamma}^{\hat{\mathcal{C}}_0}$ is given by:

$$\hat{\Gamma}^{\hat{\mathcal{C}}}\big(u^n, m, m_r, \hat{v}_0^n, z^n\big) = \mathrm{Q}^n_{\mathrm{U}}(u^n)\, 2^{-n(R+R_r)}\, \mathbb{1}_{\{\hat{v}_0^n = \hat{v}_0^n(u^n, m, m_r)\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0}(z^n|\hat{v}_0^n). \tag{3.71}$$

Similar to $\ddot{\mathrm{P}}$, we define a probability distribution $\hat{\mathrm{P}}$ that resembles $\hat{\Gamma}$, but with respect to the random code C as follows: $\hat{\mathrm{P}} \triangleq \mathrm{G}(\mathcal{C}) \cdot \ddot{\mathrm{P}}^{\mathcal{C}}$, where $\hat{\mathrm{P}}^{\mathcal{C}}$ is given by:

$$\hat{\mathrm{P}}^{\mathcal{C}}\big(m_c, m, m_r, u^n, v_0^n, z^n\big) = 2^{-n(R_c+R+R_r)}\, \mathbb{1}_{\{u^n = u^n(m_c), v_0^n = v_0^n(m_c, m, m_r)\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0}(z^n|v_0^n). \tag{3.72}$$

One should notice that, although we define $\hat{\mathrm{P}}$ over the whole random codebook C, the distribution $\hat{\mathrm{P}}^{\mathcal{C}}$ only depends on the codebooks $\mathrm{C}_{c,0}$ and is independent of $\mathrm{C}_{1,2}$. Additionally, $\hat{\mathrm{P}}^{\mathcal{C}}$ can be interpreted as the multiplication of the marginal distribution of $\mathrm{P}^{\mathcal{C}}$ in (3.22) over the random variables $\big(\mathrm{M}_c, \mathrm{M}, \mathrm{M}_r, \mathrm{U}^n, \mathrm{V}_0^n, \mathrm{V}_1^n\big)$ and the DMC $Q^n_{\mathrm{Z}|\mathrm{V}_0}$.

Now, we are ready to proceed with our secrecy analysis. Let us denote the total variation distance in (3.62) by $\mathcal{L}$. Based on the triangle inequality and the three distributions defined in (3.63), (3.68) and (3.72), it follows that:

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}\big[\mathcal{L}\big] &\leq \mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}} - \ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}}\big\| + \big\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}} - \hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}}\big\| \\
&\qquad + \big\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}} - \check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}\big\|\Big] \\
&\overset{(a)}{=} \mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}} - \ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}}\big\| + \big\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}} - \hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}}\big\| \\
&\qquad + \big\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}} - \check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{C}}\big\|\Big],
\end{aligned} \tag{3.73}$$

where (a) follows due to the symmetricity of the distributions $\mathrm{P}^{\mathcal{C}}$, $\ddot{\mathrm{P}}^{\mathcal{C}}$ and $\hat{\mathrm{P}}^{\mathcal{C}}$ with respect to the messages M_c and M, in addition to the symmetricity of $\check{\mathrm{P}}^{\mathcal{C}}$ with respect to M_c. We denote the three total variation distances summed up in the $\mathbb{RHS}$ of Eq. (3.73) by $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ respectively. We then start by $\mathcal{L}_1$ and observe that $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_1]$ can be bounded as follows:

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}[\mathcal{L}_1] &\leq \mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{M}_r\mathrm{M}_{r_1}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}} - \ddot{\mathrm{P}}_{\mathrm{M}_r\mathrm{M}_{r_1}\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\mathrm{M}=1,\mathrm{M}_c=1,\mathrm{C}}\big\|\Big] \\
&\overset{(a)}{=} \mathbb{E}_{\mathrm{C}}\Bigg[\sum_{m_r, m_{r_1}} 2^{-n(R_r+R_{r_1})} \sum_{\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\, \mathbb{1}_{\{\mathrm{V}_0^n(1,1,m_r)=\bar{v}_0^n\}}\, \mathbb{1}_{\{\mathrm{V}_1^n(1,1,m_r,m_{r_1})=\bar{v}_1^n\}} \\
&\qquad \times \Big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=m_r,\mathrm{M}_{r_1}=m_{r_1},\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n, \bar{v}_1^n)\Big\|\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{\mathrm{C}}\Bigg[\sum_{\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\, \mathbb{1}_{\{\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n\}}\, \mathbb{1}_{\{\mathrm{V}_1^n(1,1,1,1)=\bar{v}_1^n\}}
\end{aligned}$$

$$\times \left\| \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \right\| \Bigg]$$

$$\overset{(c)}{=} \mathbb{E}_{\mathrm{C}_{c,0,1}} \Bigg[\sum_{\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}} \mathbb{1}_{\left\{\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n\right\}} \mathbb{1}_{\left\{\mathrm{V}_1^n(1,1,1,1)=\bar{v}_1^n\right\}} \mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}} \Bigg[$$

$$\times \left\| \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \right\| \Bigg] \Bigg], \tag{3.74}$$

where (a) follows from the definition of $\mathrm{P}^{\mathcal{C}}$ in (3.21) and the definition of $\ddot{\mathrm{P}}^{\mathcal{C}}$ in (3.68); (b) follows due to the symmetricity of $\mathrm{P}^{\mathcal{C}}$ with respect to the messages M_r and M_{r_1}; while (c) follows from the law of total expectation [128] by dividing the expectation as in (3.32). Next, we observe that for the sequences $(\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n)$ and a fixed codebook $\mathcal{C} \in \mathfrak{C}$, the probability distribution $\mathrm{P}^{\mathcal{C}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n}$ is well-defined only if the following condition holds: $\big(\bar{u}^n = u^n(1)$ and $\bar{v}_0^n = v_0^n(1,1,1)$ and $\bar{v}_1^n = v_1^n(1,1,1,1)\big)$. Thus, for the case where $\big(\bar{u}^n \neq u^n(1)$ or $\bar{v}_0^n \neq v_0^n(1,1,1)$ or $\bar{v}_1^n \neq v_2^n(1,1,1,1)\big)$, we assume the following:

$$\mathrm{P}^{\mathcal{C}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n} = \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v_0}^n,\bar{v}_1^n). \tag{3.75}$$

This assumption does not affect the calculation of $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_1]$ because the indicator functions in (3.74) assure that the inner expectation only contributes to the total term when the sequences $(\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n)$ are identical to the codewords $\big(u^n(1), v_0^n(1,1,1), v_1^n(1,1,1,1)\big)$ chosen by a certain codebook realization $\mathcal{C}$. With this in mind, we investigate the inner expectation in (3.74), for a given codebook realization $\mathcal{C}_{c,0,1}$ and fixed sequences $(\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n)$. Thus, we have:

$$\mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}=\mathcal{C}_{c,0,1}} \left[\left\| \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \right\| \right]$$

$$= \mathbb{E}_{\mathrm{C}_2|\mathrm{C}_{c,0,1}=\mathcal{C}_{c,0,1}} \Bigg[\Bigg\| \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}=1,\mathrm{M}_c=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \Bigg\|$$

$$\times \left(\mathbb{1}_{\left\{\left(u^n(1),v_0^n(1,1,1),v_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} + \mathbb{1}_{\left\{\left(u^n(1),v_0^n(1,1,1),v_1^n(1,1,1,1)\right)\neq(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}} \right) \Bigg]$$

$$\overset{(a)}{=} \mathbb{E}_{\mathrm{C}_2|\mathrm{U}^n(1)=u^n(1),\mathrm{V}_0^n(1,1,1)=v_0^n(1,1,1),\mathrm{V}_1^n(1,1,1,1)=v_1^n(1,1,1,1)} \Bigg[\mathbb{1}_{\left\{\left(u^n(1),v_0^n(1,1,1),v_1^n(1,1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\right\}}$$

$$\times \left\| \mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{M}_{r_1}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}_2(1,1,1)} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \right\| \Bigg]$$

$$\overset{(b)}{\leq} \mathbb{E}_{\tilde{\mathrm{C}}_2} \left[\left\| \tilde{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n) \right\| \right], \tag{3.76}$$

where $\tilde{\mathrm{C}}_2$ is the random codebook introduced in Section 3.2.3 and $\tilde{\mathrm{P}}^{\tilde{\mathrm{C}}_2}$ is its induced joint probability given by (3.26). Step (a) follows from (3.75) and the fact that the expectation of the total variation distance depends only on the codewords $\mathrm{U}^n(1)$, $\mathrm{V}_0^n(1,1,1)$ and $\mathrm{V}_1^n(1,1,1,1)$ along with the corresponding codebook $\mathrm{C}_2(1,1,1)$ and not on the entire codebook C; (b) follows by removing the indicator function and because when $v_0^n(1,1,1) = \bar{v}_0^n$ and $\mathcal{C}_2(1,1,1) = \tilde{\mathcal{C}}_2(\bar{v}_0^n)$, the two conditional distributions

$\tilde{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\tilde{\mathrm{C}}_2=\tilde{\mathcal{C}}_2(\bar{v}_0^n)}$ and $\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\ldots,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\mathrm{C}_2(1,1,1)=\mathcal{C}_2(1,1,1)}$ are equivalent. Finally, by substituting Eq. (3.76) in Eq. (3.74), we get

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}[\mathcal{L}_1] &\overset{(a)}{\leq} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Bigg[\sum_{\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n} \mathbb{E}_{\mathrm{C}_{c,0,1}}\Big[\mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\mathbb{1}_{\{\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n\}}\mathbb{1}_{\{\mathrm{V}_1^n(1,1,1,1)=\bar{v}_1^n\}}\Big] \\
&\qquad \times \Big\|\tilde{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n)\Big\|\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Bigg[\sum_{\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n} \mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1}(\bar{u}^n,\bar{v}_0^n,\bar{v}_1^n)\Big\|\tilde{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{V}_1^n=\bar{v}_1^n,\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(\cdot|\bar{v}_0^n,\bar{v}_1^n)\Big\|\Bigg] \\
&\overset{(c)}{=} \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1\mathrm{Z}}\Big\|\Big] \\
&\leq \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\Big\|\Big] + \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1\mathrm{Z}}\Big\|\Big],
\end{aligned} \tag{3.77}$$

where (a) follows because $\tilde{\mathrm{C}}_2$ and $\mathrm{C}_{c,0,1}$ are independent; (b) follows from the properties of the indicator function and the fact that the random codebook $\mathrm{C}_{c,0,1}$ was generated using the input distribution $\mathrm{Q}_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1}$; while (c) follows from the definition of $\tilde{\mathrm{P}}^{\tilde{\mathcal{C}}_2}$ in (3.28). Based on the result established in Section 3.2.3, in particular Eqs. (3.41) and (3.46), we have

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\Big\|\Big] &\leq \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\tilde{\mathrm{P}}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\tilde{\mathrm{V}}_2^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2}\Big\|\Big] \\
&\leq \delta_n(\epsilon,\beta),
\end{aligned} \tag{3.78}$$

as long as the rate constraint in (3.45) is fulfilled. For $\gamma > 0$ and $0 < \tilde{\epsilon} < \epsilon$, let us define the following δ-functions:

$$\delta_n^{\mathrm{Z}}(\tilde{\epsilon},\epsilon) = 2\cdot|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\cdot|\mathcal{Z}|\exp\Big(-2n\Big(\frac{\epsilon-\tilde{\epsilon}}{1+\tilde{\epsilon}}\Big)^2\mu^2_{\mathrm{inp},\mathrm{Z}}\Big) \tag{3.79}$$

$$\delta_4(\epsilon,\gamma) = 2\epsilon\log\big(|\mathcal{U}|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\cdot|\mathcal{Z}|\big) + 5\gamma \tag{3.80}$$

$$\delta_n^{\mathrm{Z}}(\epsilon,\gamma) = \big(1 - n\delta_n^{\mathrm{inp}}(\epsilon)\big)\cdot\big(5\,\delta_n^{\mathrm{Z}}(\tilde{\epsilon},\epsilon) + 7\cdot 2^{-n\gamma}\big), \tag{3.81}$$

where $\mu_{\mathrm{inp},\mathrm{Z}}$ is the minimum probability of an event for the joint distribution $\mathrm{Q}_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2\mathrm{Z}}$. Based on the result established in [132, Corollary VII.5] and the analysis technique used in Section 2.2.3, one can show that:

$$\text{if} \quad R_{r_2} + R_t \geq \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) + \delta_4(\epsilon,\gamma), \tag{3.82}$$

$$\text{then} \quad \mathbb{E}_{\tilde{\mathrm{C}}_2}\Big[\Big\|\Gamma_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{V}_1^n\mathrm{Z}^n|\tilde{\mathrm{C}}_2} - \mathrm{Q}^n_{\mathrm{U}\mathrm{V}_0\mathrm{V}_1\mathrm{Z}}\Big\|\Big] \leq \delta_n(\epsilon,\gamma). \tag{3.83}$$

This implies that the rate constraints in (3.45) and (3.82) assure that $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_1]$ approaches zero as n approaches infinity. We now turn to $\mathcal{L}_2$ and observe that $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_2]$ can be bounded as follows:

$$\mathbb{E}_{\mathrm{C}}[\mathcal{L}_2] \overset{(a)}{\leq} \mathbb{E}_{\mathrm{C}_{c,0,1}}\Big[\Big\|\ddot{\mathrm{P}}_{\mathrm{M}_r\mathrm{U}^n\mathrm{V}_0^n\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}_{c,0,1}} - \hat{\mathrm{P}}_{\mathrm{M}_r\mathrm{U}^n\mathrm{V}_0^n\mathrm{Z}^n|\mathrm{M}=1,\mathrm{M}_c=1,\mathrm{C}_{c,0,1}}\Big\|\Big]$$

$$\overset{(b)}{=} \mathbb{E}_{\mathrm{C}_{c,0,1}}\Bigg[\sum_{\bar{u}^n,\bar{v}_0^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\mathbb{1}_{\left\{\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n\right\}} \times \left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{C}_{c,0,1}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\Bigg]$$

$$\overset{(c)}{=} \mathbb{E}_{\mathrm{C}_{c,0}}\Bigg[\sum_{\bar{u}^n,\bar{v}_0^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\mathbb{1}_{\left\{\mathrm{V}_0^n(1,1,1)=\bar{v}_0^n\right\}}\, \mathbb{E}_{\mathrm{C}_1|\mathrm{C}_{c,0}}\Bigg[\times \left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{C}_{c,0,1}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\Bigg]\Bigg], \tag{3.84}$$

where (a) follows because the two distributions $\ddot{\mathrm{P}}^{\mathcal{C}}$ in (3.68) and $\hat{\mathrm{P}}^{\mathcal{C}}$ in (3.72) are independent of the codebook $\mathcal{C}_2$; (b) follows from the definition of $\ddot{\mathrm{P}}^{\mathcal{C}}$ and $\hat{\mathrm{P}}^{\mathcal{C}}$ along with the symmetricity of $\ddot{\mathrm{P}}^{\mathcal{C}}$ with respect to M_{r_1}; while (c) follows from the law of total expectation. Next, for the sequences $(\bar{u}^n, \bar{v}_0^n, \bar{v}_1^n)$ and a fixed codebook $\mathcal{C} \in \mathfrak{C}$, we assume that:

$$\ddot{\mathrm{P}}^{\mathcal{C}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n} = \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n), \tag{3.85}$$

if $\big(\bar{u}^n \neq u^n(1)$ or $\bar{v}_0^n \neq v_0^n(1,1,1)\big)$. This assumption is based on some arguments similar to the ones used to validate the assumption in (3.75). With this in mind, we investigate the inner expectation in (3.84), for a given codebook realization $\mathcal{C}_{c,0}$ and fixed sequences $(\bar{u}^n, \bar{v}_0^n)$ as follows:

$$\begin{aligned}
&\mathbb{E}_{\mathrm{C}_1|\mathrm{C}_{c,0}=\mathcal{C}_{c,0}}\left[\left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{C}_{c,0,1}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\right] \\
&\overset{(a)}{=} \mathbb{E}_{\mathrm{C}_1|\mathrm{U}^n(1)=u^n(1),\mathrm{V}_0^n(1,1,1)=v_0^n(1,1,1)}\Bigg[\mathbb{1}_{\left\{\left(u^n(1),v_0^n(1,1,1)\right)=(\bar{u}^n,\bar{v}_0^n)\right\}} \times \left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{C}_1(1,1,1)} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\Bigg] \\
&\overset{(b)}{\leq} \mathbb{E}_{\ddot{\mathrm{C}}_1}\left[\left\|\ddot{\Gamma}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\ddot{\mathrm{C}}_1} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\right],
\end{aligned} \tag{3.86}$$

where (a) follows from (3.85) and the fact that the expectation depends only on the codewords $\mathrm{U}^n(1)$ and $\mathrm{V}_0^n(1,1,1)$ along with the corresponding codebook $\mathrm{C}_1(1,1,1)$ and not on the entire codebook $\mathrm{C}_{c,0,1}$; (b) follows because when $v_0^n(1,1,1) = \bar{v}_0^n$ and $\mathcal{C}_1(1,1,1) = \ddot{\mathcal{C}}_1(\bar{v}_0^n)$, the two distributions $\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{M}_r=1,\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\mathrm{C}_1(1,1,1)=\mathcal{C}_2(1,1,1)}$ and $\ddot{\Gamma}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\ddot{\mathrm{C}}_1=\ddot{\mathcal{C}}_1(\bar{v}_0^n)}$ are equivalent. Finally, by substituting Eq. (3.86) in Eq. (3.84), we get

$$\begin{aligned}
\mathbb{E}_{\mathrm{C}}[\mathcal{L}_2] &\overset{(a)}{\leq} \mathbb{E}_{\ddot{\mathrm{C}}_1}\Bigg[\sum_{\bar{u}^n,\bar{v}_0^n} \mathrm{Q}^n_{\mathrm{UV}_0}(\bar{u}^n,\bar{v}_0^n)\left\|\ddot{\Gamma}_{\mathrm{Z}^n|\mathrm{U}^n=\bar{u}^n,\mathrm{V}_0^n=\bar{v}_0^n,\ddot{\mathrm{C}}_1} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{V}_0}(\cdot|\bar{v}_0^n)\right\|\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{\ddot{\mathrm{C}}_1}\left[\left\|\ddot{\Gamma}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{Z}^n|\ddot{\mathrm{C}}_1} - \mathrm{Q}^n_{\mathrm{UV}_0\mathrm{Z}}\right\|\right]
\end{aligned} \tag{3.87}$$

where (a) follows because $\ddot{\mathrm{C}}_1$ and $\mathrm{C}_{c,0}$ are independent and the fact that the random codebook $\mathrm{C}_{c,0}$ was generated using the input distribution $\mathrm{Q}_{\mathrm{UV}_0}$; while (b) follows from the definition of $\ddot{\Gamma}^{\ddot{\mathcal{C}}_1}$ in (3.67). Finally, based on the result established in [132, Corollary

VII.5] and the analysis technique used in Section 2.2.3, one can show that:

$$\text{if} \quad R_{r_1} \geq \mathbb{I}(\mathrm{V}_1;\mathrm{Z}|\mathrm{V}_0) + \delta_4(\epsilon,\gamma), \tag{3.88}$$

$$\text{then} \quad \mathbb{E}_{\ddot{\mathrm{C}}_1}\left[\left\|\ddot{\Gamma}_{\mathrm{U}^n\mathrm{V}_0^n\mathrm{Z}^n|\ddot{\mathrm{C}}_1} - \mathrm{Q}^n_{\mathrm{UV_0Z}}\right\|\right] \leq \delta_n(\epsilon,\gamma). \tag{3.89}$$

This implies that the rate constraint in (3.88) assures that $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_2]$ approaches zero as n approaches infinity. Now, it only remains to bound $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_3]$. Based on the arguments used to bound $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_2]$, we can show that:

$$\begin{aligned}\mathbb{E}_{\mathrm{C}}[\mathcal{L}_3] &\overset{(a)}{\leq} \mathbb{E}_{\mathrm{C}_{c,0}}\left[\left\|\hat{\mathrm{P}}_{\mathrm{U}^n\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{C}_{c,0}} - \check{\mathrm{P}}_{\mathrm{U}^n\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{C}_{c,0}}\right\|\right] \\ &\overset{(b)}{=} \mathbb{E}_{\mathrm{C}_c}\left[\sum_{\bar{u}^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\, \mathbb{E}_{\mathrm{C}_0|\mathrm{C}_c}\left[\left\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{C}_{c,0}} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{U}}(\cdot|\bar{u}^n)\right\|\right]\right],\end{aligned} \tag{3.90}$$

where (a) follows because the two distributions $\hat{\mathrm{P}}^{\mathcal{C}}$ in (3.72) and $\check{\mathrm{P}}^{\mathcal{C}}$ in (3.63) are independent of the codebooks $\mathcal{C}_{1,2}$; while (b) follows from the law of total expectation. Next, for a given sequence $\bar{u}^n$ and a fixed codebook $\mathcal{C} \in \mathfrak{C}$, we assume that:

$$\hat{\mathrm{P}}^{\mathcal{C}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{U}^n=\bar{u}^n} = \mathrm{Q}^n_{\mathrm{Z}|\mathrm{U}}(\cdot|\bar{u}^n), \tag{3.91}$$

if $\bar{u}^n \neq u^n(1)$. The validity of this assumption is based on arguments similar to the ones used to establish (3.75). With this in mind, we can further bound $\mathbb{E}_{\mathrm{C}}[\mathcal{L}_3]$ as follows:

$$\begin{aligned}\mathbb{E}_{\mathrm{C}}[\mathcal{L}_3] &\overset{(a)}{\leq} \mathbb{E}_{\mathrm{C}_c}\left[\sum_{\bar{u}^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\, \mathbb{E}_{\mathrm{C}_0|\mathrm{U}^n(1)}\left[\left\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c=1,\mathrm{M}=1,\mathrm{U}^n=\bar{u}^n,\mathrm{C}_0(1)} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{U}}(\cdot|\bar{u}^n)\right\|\right]\right] \\ &\overset{(b)}{=} \mathbb{E}_{\mathrm{C}_c}\left[\sum_{\bar{u}^n} \mathbb{1}_{\{\mathrm{U}^n(1)=\bar{u}^n\}}\, \mathbb{E}_{\hat{\mathrm{C}}_0}\left[\left\|\hat{\Gamma}_{\mathrm{Z}^n|\mathrm{M}=1,\mathrm{U}^n=\bar{u}^n,\hat{\mathrm{C}}_0} - \mathrm{Q}^n_{\mathrm{Z}|\mathrm{U}}(\cdot|\bar{u}^n)\right\|\right]\right] \\ &\overset{(c)}{=} \mathbb{E}_{\hat{\mathrm{C}}_0}\left[\left\|\hat{\Gamma}_{\mathrm{U}^n\mathrm{Z}^n|\mathrm{M}=1,\hat{\mathrm{C}}_0} - \mathrm{Q}^n_{\mathrm{UZ}}\right\|\right],\end{aligned} \tag{3.92}$$

where (a) and (b) follow using similar arguments to the ones used to establish (3.86) along with Eq. (3.91); while (c) follows from the definition of $\hat{\Gamma}^{\hat{C}_0}$ in (3.71). Again using [132, Corollary VII.5] and the analysis in Section 2.2.3, we can show that:

$$\text{if} \quad R_r \geq \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) + \delta_4(\epsilon,\gamma), \tag{3.93}$$

$$\text{then} \quad \mathbb{E}_{\hat{\mathrm{C}}_0}\left[\left\|\hat{\Gamma}_{\mathrm{U}^n\mathrm{Z}^n|\mathrm{M}=1,\hat{\mathrm{C}}_0} - \mathrm{Q}^n_{\mathrm{UZ}}\right\|\right] \leq \delta_n(\epsilon,\gamma). \tag{3.94}$$

The previous secrecy analysis implies that if the constraints on the rate quadruple $(R_r, R_{r_1}, R_{r_2}, R_t)$ given by (3.45), (3.82), (3.88) and (3.93) are satisfied, then it follows from (3.78), (3.83), (3.89) and (3.94) that:

$$\mathbb{E}_{\mathcal{C}}[\mathcal{L}] \leq \delta_n(\epsilon,\beta) + 3\delta_n(\epsilon,\gamma) \leq 2^{-n\tau}, \tag{3.95}$$

for some $\tau > 0$. This implies that under these rate constraints, our coding scheme satisfies the joint-strong secrecy criterion given in (3.15c). It is important to point out that the strong-secrecy analysis presented in this section is inspired by the concepts and results established in [127].

3.2.5 Time Sharing and Rate Splitting

In this section, we show how the coding scheme used to establish the achievability of the rate regions in Proposition 3.1 and consequently Proposition 3.2 directly implies that the full rate region given by Theorem 3.1 is also achievable. Our proof is based on the principles of rate splitting and time sharing [134] as follows:

Fix an input distribution $\mathrm{Q}_{\mathrm{TUV_0V_1V_2}}$ and let $\pi \in [0,1]$ be a time sharing parameter. Further, let $n = n_1 + n_2 \in \mathbb{N}$ be an arbitrary but fixed code block length such that $n_1 = \pi \cdot n$ and $n_2 = (1-\pi)\cdot n$. It is worth mentioning that we assume $\pi \cdot n$ to be an integer. If not, it is understood as $\lfloor \pi \cdot n \rfloor$ so that $(1-\pi)n$ is $n - \lfloor \pi \cdot n \rfloor$. Next, we construct the following coding scheme.

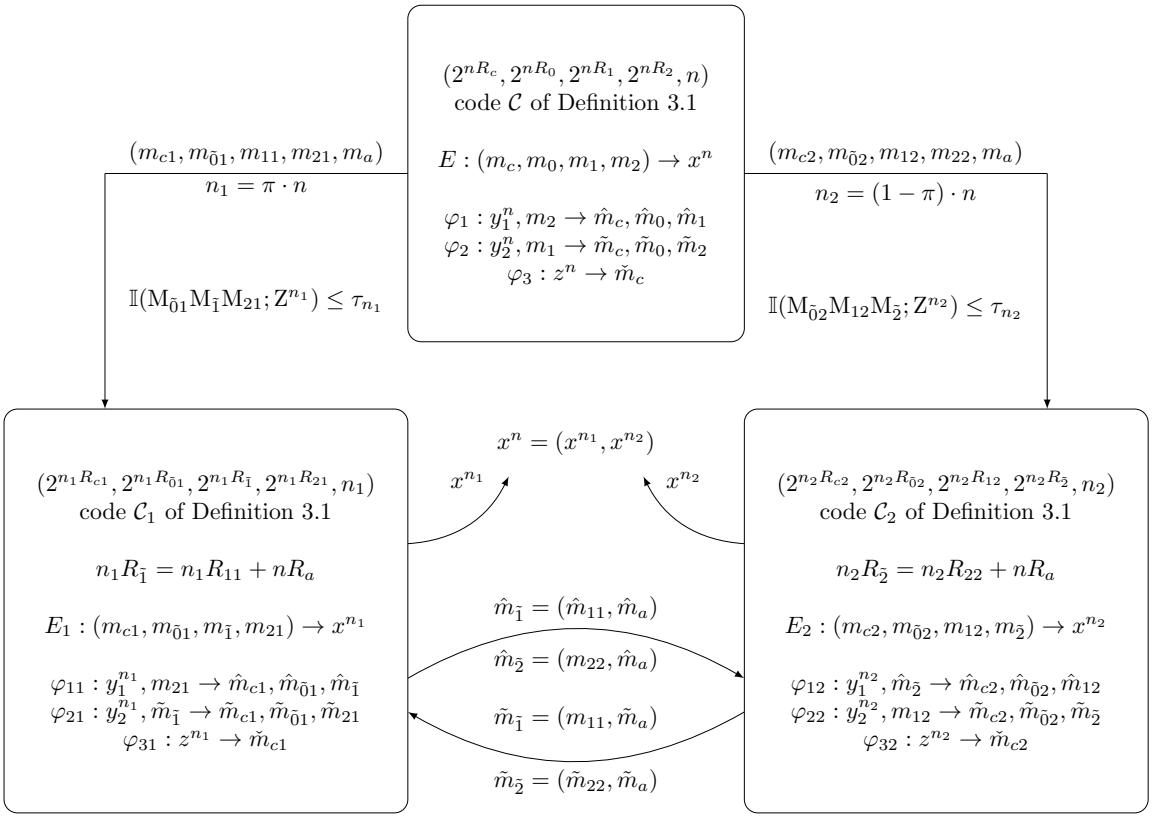


Figure 3.4: Rate Splitting and Time Sharing for the Two-receiver DM-WBC

1. Combined Code $\mathcal{C}$: We start by dividing the common confidential message set $\mathcal{M}_0 = [\![1, 2^{nR_0}]\!]$ into two parts: $\mathcal{M}_{\tilde{0}} = [\![1, 2^{nR_{\tilde{0}}}]\!]$ and $\mathcal{M}_a = [\![1, 2^{nR_a}]\!]$. This division directly implies that:

$$R_0 = R_{\tilde{0}} + R_a. \tag{3.96}$$

We further divide each message set $\mathcal{M}_j = [\![1, 2^{nR_j}]\!]$ for $j \in \{c, \tilde{0}, 1, 2\}$ into two parts: $\mathcal{M}_{j1} = [\![1, 2^{n_1R_{j1}}]\!]$ and $\mathcal{M}_{j2} = [\![1, 2^{n_2R_{j2}}]\!]$. This division directly implies that

$$R_j = \pi \cdot R_{j1} + (1-\pi)\cdot R_{j2}. \tag{3.97}$$

Next, we generate a time sharing sequence $t^n(\pi)$ by generating the symbols $t_i(\pi)$ with $i \in [\![1, n]\!]$ independently according to $\mathrm{Q_T}$. For the messages sets $\mathcal{M}_{j1}$, $\mathcal{M}_a$ and the

code block length n_1, define a code $\mathcal{C}_1$ of Definition 3.1 that adheres to the coding scheme introduced in Section 3.2.2 with some slight modifications:

1. We use the message $m_{\tilde{1}} \in \mathcal{M}_{\tilde{1}} = \mathcal{M}_{11} \times \mathcal{M}_a$ to denote the individual confidential message intended for the first legitimate receiver.
2. We use the message $\tilde{m}_{\tilde{1}} = (m_{11}, \tilde{m}_a) \in \mathcal{M}_{\tilde{1}}$ to denote the side information available at the second legitimate receiver, where $\tilde{m}_a$ is an estimation for the actual message m_a.
3. We use the conditional distribution $\mathrm{Q}_{\mathrm{UV_0V_1V_2|T}}$ to generate the corresponding codebook.

Similarly, use the message sets $\mathcal{M}_{j2}$, $\mathcal{M}_a$ and the code block length n_2 to define a code $\mathcal{C}_2$ of Definition 3.1. The construction of $\mathcal{C}_2$ follows the same scheme used to construct $\mathcal{C}_1$ but after interchanging the subscripts 1 and 2 during the codebook generation as well as the encoding and decoding processes, cf. Fig. 3.4. Finally, for the original message sets $\mathcal{M}_c$, $\mathcal{M}_0$, $\mathcal{M}_1$ and $\mathcal{M}_1$ along with the code block length n, we define the combined code $\mathcal{C}$ as the concatenation of the two codes $\mathcal{C}_1$ and $\mathcal{C}_2$.

2. Encoding: Given the message quadruple $(m_c, m_0, m_1, m_2) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$, the encoder works as follows: First, it finds the corresponding two quadruples: $(m_{c1}, m_{01}, m_{\tilde{1}}, m_{21}) \in \mathcal{M}_{c1} \times \mathcal{M}_{01} \times \mathcal{M}_{\tilde{1}} \times \mathcal{M}_{21}$ and $(m_{c2}, m_{02}, m_{12}, m_{\tilde{2}}) \in \mathcal{M}_{c2} \times \mathcal{M}_{02} \times \mathcal{M}_{12} \times \mathcal{M}_{\tilde{2}}$. The first quadruple is then passed to the encoder of $\mathcal{C}_1$ producing the codeword x^{n_1}, while the second quadruple is passed to the encoder of $\mathcal{C}_2$ producing the codeword x^{n_2}. The codeword x^n is then constructed by concatenating the two codewords x^{n_1} and x^{n_2}, cf. Fig. 3.4. Finally, the encoder transmits x^n.

3. Decoding: At the first legitimate receiver, given the sequence y_1^n and its own message $m_2 = (m_{21}, m_{22})$, the first legitimate decoder works as follows: First, it searches for a unique value $\hat{\pi}$ such that:

$$\left(t^n(\hat{\pi}), y_1^n\right) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{TY}_1}). \tag{3.98}$$

It then uses $\hat{\pi}$ to split y_1^n into two parts $y_1^{\hat{n}_1}$ and $y_1^{n-\hat{n}_1}$, where $\hat{n}_1 = \hat{\pi} \cdot n$. The first part $y_1^{\hat{n}_1}$ along with m_{21} are fed to φ_{11}, which produces the triple $(\hat{m}_{c1}, \hat{m}_{\tilde{0}1}, \hat{m}_{\tilde{1}}) \in \mathcal{M}_{c1} \times \mathcal{M}_{\tilde{0}1} \times \mathcal{M}_{\tilde{1}}$. Afterwords, $y_1^{n-\hat{n}_1}$ along with $\hat{m}_{\tilde{2}} = (m_{22}, \hat{m}_a)$ are fed to φ_{12} which produces the triple $(\hat{m}_{c2}, \hat{m}_{\tilde{0}2}, \hat{m}_{12}) \in \mathcal{M}_{c2} \times \mathcal{M}_{\tilde{0}2} \times \mathcal{M}_{12}$. Finally, the triple $(\hat{m}_c, \hat{m}_0, \hat{m}_1) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1$ is constructed using the two triples produced by φ_{11} and φ_{12}.

At the second legitimate receiver, the decoding process follows the same concept. First, a unique value $\tilde{\pi}$ that satisfy a typicality condition similar to the one in (3.98) is estimated. $\tilde{\pi}$ is then used to split the received sequence y_2^n into two parts $y_2^{\tilde{n}_1}$ and $y_2^{n-\tilde{n}_1}$. Differently from the first legitimate receiver, we first feed $y_2^{n-\tilde{n}_1}$ along with the side information m_{12} to φ_{22}, which produces the triple $(\tilde{m}_{c2}, \tilde{m}_{\tilde{0}2}, \tilde{m}_{\tilde{2}}) \in \mathcal{M}_{c2} \times \mathcal{M}_{\tilde{0}2} \times \mathcal{M}_{\tilde{2}}$. We then feed $y_2^{\tilde{n}_1}$ along with $\tilde{m}_{\tilde{1}} = (m_{11}, \tilde{m}_a)$ to φ_{21}, which outputs the triple $(\tilde{m}_{c1}, \tilde{m}_{\tilde{0}1}, \tilde{m}_{21}) \in \mathcal{M}_{c1} \times \mathcal{M}_{\tilde{0}1} \times \mathcal{M}_{21}$. Finally, we construct $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_2$ using the two estimated triples.

The same concept is used to define the decoding procedure at the eavesdropper using the received sequence z^n as follows: First it searches for a unique value $\check{\pi}$ that satisfy a typicality condition similar to the one in (3.98). It then uses $\check{\pi}$ to split z^n into two

parts. Each part is then fed to the eavesdropper decoders for φ_{31} and φ_{32} producing the message $\check{m}_c = (\check{m}_{c1}, \check{m}_{c2}) \in \mathcal{M}_c$.

4. Reliability and Secrecy Analysis: Using the union bound [133] along with the mechanism of our encoding and decoding functions, the reliability constraint in (3.14b) with respect to the code $\mathcal{C}$ can be bounded as follows:

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \mathbb{P}\Big[(\hat{\Pi} \neq \Pi) \text{ or } (\tilde{\Pi} \neq \Pi) \text{ or } (\check{\Pi} \neq \Pi)\Big] + \bar{\mathbf{P}}_e(\mathcal{C}_1) + \bar{\mathbf{P}}_e(\mathcal{C}_2). \tag{3.99}$$

Based on the convention used to define our messages sets and under the assumption that $\pi \cdot n \in \mathbb{N}$, it follows that:

$$R_\pi \leq \frac{\log n + 1}{n} \tag{3.100}$$

where R_π corresponds to the rate that carries the time sharing parameter π. The previous inequality implies that $\lim_{n\to\infty} R_\pi = 0$, which consequently means that the three receivers can decode the time sharing parameter (π) correctly. Hence, we have

$$\lim_{n\to\infty} \mathbb{P}\Big[(\hat{\Pi} \neq \Pi) \text{ or } (\tilde{\Pi} \neq \Pi) \text{ or } (\check{\Pi} \neq \Pi)\Big] = 0. \tag{3.101}$$

Thus, we only need to consider the decoding errors of $\mathcal{C}_1$ and $\mathcal{C}_2$. Similarly, the joint-strong secrecy requirement with respect to the code $\mathcal{C}$ can be bounded as follows:

$$\begin{aligned}\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) &\overset{(a)}{\leq} \mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}\mathrm{T}) \\ &\overset{(b)}{\leq} \mathbb{I}(\mathrm{M}_{01}\mathrm{M}_{\tilde{1}}\mathrm{M}_{21}; \mathrm{Z}^{n_1}|\mathcal{C}_1) + \mathbb{I}(\mathrm{M}_{02}\mathrm{M}_{12}\mathrm{M}_{\tilde{2}}; \mathrm{Z}^{n_2}|\mathcal{C}_2),\end{aligned} \tag{3.102}$$

where (a) follows because the messages M_0, M_1 and M_2 are independent of the time sharing random variable T; while (b) follows because Z^{n_1} only depends on the codeword X^{n_1} generated by the code $\mathcal{C}_1$ given the messages M_{01}, $\mathrm{M}_{\tilde{1}}$ and M_{21}. The same statement holds for Z^{n_2}, $\mathcal{C}_2$ and the messages $(\mathrm{M}_{02}, \mathrm{M}_{12}, \mathrm{M}_{\tilde{2}})$.

Based on the results established Section 3.2.3 and Section 3.2.4, $\mathcal{C}_1$ satisfies the reliability constraint in (3.15b) and the joint-strong secrecy constraint in (3.15c), as long as the following holds:

$$\begin{aligned}R_{c1} &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_2|\mathrm{T})\Big] \\ R_{\tilde{0}1} + R_{\tilde{1}} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{UT}) \\ R_{\tilde{0}1} + R_{21} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0\mathrm{T}),\end{aligned} \tag{3.103}$$

where $R_{\tilde{1}} = R_{11} + \frac{1}{\pi}R_a$, cf. Fig 3.4. On the other hand, Proposition 3.2 implies that $\mathcal{C}_2$ also satisfies the reliability constraint in (3.15b) and the joint-strong secrecy constraint in (3.15c), as long as the following holds:

$$\begin{aligned}R_{c2} &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_2|\mathrm{T})\Big] \\ R_{\tilde{0}2} + R_{12} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2\mathrm{Z}|\mathrm{V}_0\mathrm{T}) \\ R_{\tilde{0}2} + R_{\tilde{2}} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{UT}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{UT}),\end{aligned} \tag{3.104}$$

where $R_{\tilde{2}} = R_{22} + \frac{1}{1-\pi} R_a$, cf. Fig. 3.4. In order to finalize our prove, we need to highlight two more relations that are based on the way we defined are messages sets. Those two relations are:

$$R_1 + R_a = \pi R_{\tilde{1}} + (1-\pi) R_{12} \quad \text{and} \quad R_2 + R_a = \pi R_{21} + (1-\pi) R_{\tilde{2}} \tag{3.105}$$

Now, if we apply the Fourier-Motzkin elimination procedure [57] to the rate constraints in (3.103) and (3.104) along with the rate splitting equations in (3.96), (3.97) and (3.105), then solve for the rate quadruple (R_c, R_0, R_1, R_2), we can show that the code $\mathcal{C}$ satisfies the reliability constraint in (3.15b) and the joint-strong secrecy constraint in (3.15c), as long as the following holds:

$$\begin{aligned}
R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_2|\mathrm{T})\Big] \\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}\ \mathrm{T}) - (1-\pi)\cdot\mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{Z}\mathrm{T}) \\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{U}\ \mathrm{T}) - \pi\cdot\mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{Z}\mathrm{T}) \\
2R_0 + R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}\ \mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\ \mathrm{T}) \\
&\quad -\mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{U}\ \mathrm{T}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}\ \mathrm{T})
\end{aligned} \tag{3.106}$$

This means that the previous rate region is achievable for all values of the time sharing parameter $\pi \in [0,1]$. This result directly implies the validity of Theorem 3.1 because for any rate quadruple $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ that satisfies the rate constraints given by Theorem 3.1 in Eq. (3.16), there exists a certain π such that the given rate quadruple is included in the rate region defined by the bounds in Eq. (3.106). This completes our achievability proof.

3.3 Individual-Strong Secrecy Rate Regions

In this section, we investigate secure communication over the two-receiver DM-WBC with degraded message sets and message cognition given by Fig. 3.2 under the individual-strong secrecy criterion. We start by investigating different individual secrecy coding schemes for a special case of our model that corresponds to the broadcasting phase of the bidirectional relaying problem in Fig. 3.1. We then establish a general achievable individual secrecy rate region for our full model. Our general rate region combines the principles of superposition encoding and one time pad for Shannon's cipher system, along with the strongly secure Marton-Coding scheme introduced in Section 3.2.2. Finally, we present an additional rate region which is only achievable for a special class of the two-receiver DM-WBC.

3.3.1 Individual Secrecy Coding Techniques

Consider a special case of our communication model in Fig. 3.2, where the common message set $\mathcal{M}_c$ and the common confidential message set $\mathcal{M}_0$ are empty sets, i.e. $\mathcal{M}_c = \mathcal{M}_0 = \emptyset$. This implies that our model simplifies to the broadcasting phase of the bidirectional relaying problem given in Fig. 3.1. We will call this model the two-receiver DM-WBC with receiver side information. Our aim is to investigate secure communication over this model under the individual secrecy criterion. We start by presenting the following definition:

Definition 3.3. A $(2^{nR_1}, 2^{nR_2}, n)$ code $\mathcal{C}$ for the two-receiver DM-WBC with receiver side information consists of: two independent message sets $\mathcal{M}_1$ and $\mathcal{M}_2$, a stochastic encoding function at the transmitter

$$E : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{P}(\mathcal{X}^n)$$

which maps a confidential message pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ into a random codeword $x^n(m_1, m_2) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m_1, m_2)$ and two deterministic decoders, one for each legitimate receiver

$$\varphi_1 : \mathcal{Y}_1^n \times \mathcal{M}_2 \to \mathcal{M}_1 \cup \{?\}$$
$$\varphi_2 : \mathcal{Y}_2^n \times \mathcal{M}_1 \to \mathcal{M}_2 \cup \{?\}$$

which maps each channel observation at the respective node along with the available side information into the corresponding required messages or an error message $\{?\}$.

The previous code definition is simply a special case of our original code given by Definition 3.1. In order to evaluate the reliability and secrecy performance of the code $\mathcal{C}$, we consider the following measures:

$$\bar{\mathbf{P}}_e(\mathcal{C}) \triangleq \mathbb{P}\Big[(\hat{\mathrm{M}}_1) \neq (\mathrm{M}_1) \text{ or } (\tilde{\mathrm{M}}_2) \neq (\mathrm{M}_2)\Big], \tag{3.107}$$

$$\mathbf{L}_i(\mathcal{C}) \triangleq \mathbb{I}(\mathrm{M}_1; \mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}). \tag{3.108}$$

Our target is to construct a coding scheme which assures that the previous two measures vanish as the code block length n approaches infinity. The secrecy constraint in (3.108) adheres to the individual secrecy criterion. This implies that any coding scheme that utilizes the concept of mutual trust between the legitimate receivers can be used. In particular, each message can be used as a secret key for the other one.

The previous problem is closely related to the problem of secure communication over wiretap channel with shared secret key investigated in Section 2.3. This relation implies that the coding schemes used to establish Theorem 2.3 can be modified to derive an individual-strong secrecy achievable rate region for the two-receiver DM-WBC with receiver side information. With this in mind, we present the following coding schemes and their corresponding achievable rate regions:

1. Secret Key Encoding: This scheme utilizes only the techniques of Shannon's ciphering system introduced in [1]. It was used by mistake in [119] as a coding scheme that fulfills the joint secrecy criterion. The achievable rate region for this scheme $\mathcal{R}_1$ is given by the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following:

$$R_1 = R_2 \leq \min\big[\mathbb{I}(\mathrm{X}; \mathrm{Y}_1), \mathbb{I}(\mathrm{X}; \mathrm{Y}_2)\big]. \tag{3.109}$$

1. Codebook Generation $\mathcal{C}$: Fix an input distribution $\mathrm{P_X}$ and let $R_1 = R_2$. Construct a new message set $\mathcal{M}_\otimes = [\![1, 2^{nR_\otimes}]\!]$ by *xoring* the corresponding elements of $\mathcal{M}_1$ and $\mathcal{M}_2$. Next, generate the codewords $x^n(m_\otimes)$ for $m_\otimes \in \mathcal{M}_\otimes$ by generating the symbols $x_i(m_\otimes)$ with $i \in [\![1, n]\!]$ independently at random according to $\mathrm{P_X}$.

2. Encoding E: Given the confidential message pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, the encoder selects the codeword $x^n(m_1 \otimes m_2)$ and transmits it.

3. First Legitimate Decoder φ_1: Upon receiving y_1^n, the decoder determines a unique codeword $x^n(\hat{m}_\otimes)$ which is jointly typical to y_1^n. It then uses $\hat{m}_\otimes$ and the available side information m_2 to produce $\hat{m}_1 = \hat{m}_\otimes \otimes m_2$.

4. Second Legitimate Decoder φ_2: Upon receiving y_2^n, the decoder determines a unique codeword $x^n(\tilde{m}_\otimes)$ which is jointly typical to y_2^n. It then uses $\tilde{m}_\otimes$ and the available side information m_1 to produce $\tilde{m}_2 = \tilde{m}_\otimes \otimes m_1$.

5. Reliability Analysis: In order to evaluate this coding scheme with respect to the reliability measure in (3.107), it is enough to prove that:

$$\lim_{n\to\infty} \mathbb{P}\big[\mathrm{M}_\otimes \neq \hat{\mathrm{M}}_\otimes \text{ or } \mathrm{M}_\otimes \neq \tilde{\mathrm{M}}_\otimes\big] = 0. \tag{3.110}$$

Based on the standard joint-typicality decoding arguments, the reliability constraint in (3.110) holds, as long as:

$$R_1 = R_2 = R_\otimes \leq \min\big[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\big] - \delta(\epsilon), \tag{3.111}$$

where $\epsilon > 0$ is the constant used in the typicality decoding. Additionally, the equality of R_1, R_2 and $R_\otimes$ follows due to the codebook generation. Moreover, ϵ can be chosen in a way such that $\lim_{n\to\infty} \delta(\epsilon) = 0$.

6. Secrecy Analysis: In order to evaluate the secrecy performance of this coding scheme with respect to the secrecy measure in (3.108), we start by considering the information leakage of the first confidential message to the eavesdropper as follows:

$$\begin{aligned}\mathbb{I}(\mathrm{M}_1;\mathrm{Z}^n|\mathcal{C}) &\leq \mathbb{H}(\mathrm{M}_1) - \mathbb{H}(\mathrm{M}_1|\mathrm{Z}^n\mathcal{C}) \\ &\overset{(a)}{\leq} \mathbb{H}(\mathrm{M}_1) - \mathbb{H}(\mathrm{M}_1|\mathrm{M}_\otimes) \overset{(b)}{=} 0,\end{aligned} \tag{3.112}$$

where (a) follows because for a fixed code $\mathcal{C}$, $\mathrm{M}_\otimes - \mathrm{X}^n - \mathrm{Z}^n$ forms a Markov chain; while (b) follows from the principles of one time pad, which implies that M_1 and $\mathrm{M}_\otimes$ are independent as long as $\mathbb{H}(\mathrm{M}_1) = \mathbb{H}(\mathrm{M}_2)$. The previous equality is true because both M_1 and M_2 are uniformly distributed on their respective sets, along with the fact that $R_1 = R_2$. Repeating the same steps for $\mathbb{I}(\mathrm{M}_2;\mathrm{Z}^n|\mathcal{C})$ implies that:

$$\mathbf{L}_i(\mathcal{C}) = 0, \quad \text{as long as} \quad R_1 = R_2. \tag{3.113}$$

This implies that this coding scheme does not only fulfill the individual-strong secrecy criterion but it also assures perfect secrecy. Finally, by combining Eq. (3.111) and Eq. (3.113) then taking the limit as $n \to \infty$, which implies that $\delta(\epsilon) = 0$, it follows that any rate pair (R_1, R_2) that satisfies the constraint in (3.109) is achievable.

2. WRC Superimposed on SKC: This scheme combines the principles of secret key encoding (SKC) [1] and wiretap random coding (WRC) [2] in addition to superposition encoding and rate splitting. It can be interpreted as an adaptation of the coding scheme used to prove Scenario 3 of Theorem 2.3. The achievable rate region for this scheme $\mathcal{R}_2$ is given by the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following:

$$\begin{aligned}R_1 &\leq \big[\mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\big]^+ + \min\big[R_2, \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\big] \\ R_2 &\leq \big[\mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\big]^+ + \min\big[R_1, \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\big],\end{aligned} \tag{3.114}$$

1. Rate Splitting: Divide the two confidential message sets into two parts as follows: $\mathcal{M}_1 = \mathcal{M}_{11} \times \mathcal{M}_{12}$ and $\mathcal{M}_2 = \mathcal{M}_{21} \times \mathcal{M}_{22}$. Each of the new constructed message sets can be written as $[\![1, 2^{nR_a}]\!]$, where a is an indicator that identifies each set. Additionally, we construct a new message set $\mathcal{M}_\otimes = [\![1, 2^{nR_\otimes}]\!]$, where $R_\otimes = R_{11} = R_{21}$ by *xoring* the elements of $\mathcal{M}_{11}$ and $\mathcal{M}_{21}$. In the rate splitting process, we made sure that the following rate constraints are satisfied:

$$R_1 = R_{11} + R_{12} \qquad R_2 = R_{21} + R_{22}$$
$$R_{11} = R_{21} = R_\otimes \leq \min[R_1, R_2]. \tag{3.115}$$

2. Codebook Generation $\mathcal{C}$: Let $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$ be a randomization message set, where M_r is a uniformly distributed random variable over it. Next, fix an input distribution $\mathrm{P_{UX}}$ and randomly generate the cloud center codewords $u^n(m_\otimes)$ for $m_\otimes \in \mathcal{M}_\otimes$ by generating the symbols $u_i(m_\otimes)$ with $i \in [\![1, n]\!]$ independently at random according to $\mathrm{P_U}$. For each cloud center codeword $u^n(m_\otimes)$, generate the satellite codewords $x^n(m_\otimes, m_{12}, m_{22}, m_r)$ for $m_{12} \in \mathcal{M}_{12}$, $m_{22} \in \mathcal{M}_{22}$ and $m_r \in \mathcal{M}_r$ by generating the symbols $x_i(m_\otimes, m_{12}, m_{22}, m_r)$ independently at random according to $\mathrm{P_{X|U}}(\cdot|u_i(m_\otimes))$.

3. Encoding E: Given the confidential message pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, where $m_1 = (m_{11}, m_{12})$ and $m_2 = (m_{21}, m_{22})$, the encoder selects a message m_r uniformly at random then transmits the corresponding codeword $x^n(m_\otimes, m_{12}, m_{22}, m_r)$, where $m_\otimes = m_{11} \otimes m_{21}$.

4. First Legitimate Decoder φ_1: Upon receiving y_1^n, the decoder uses it along with the available side information $m_2 = (m_{21}, m_{22})$ to determine a unique pair of codewords $\big(u^n(\hat{m}_\otimes), x^n(\hat{m}_\otimes,\ \hat{m}_{12}, m_{22}, \hat{m}_r)\big)$, which is jointly typical to y_1^n. Next, it uses the estimated $\hat{m}_\otimes$ and m_{21} to produce the message $\hat{m}_{11} = \hat{m}_\otimes \otimes m_{21}$. Finally, it outputs $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12})$.

5. Second Legitimate Decoder φ_2: Upon receiving y_2^n, the decoder uses it along with the available side information $m_1 = (m_{11}, m_{12})$ to determine a unique pair of codewords $\big(u^n(\tilde{m}_\otimes), x^n(\tilde{m}_\otimes,\ m_{12}, \tilde{m}_{22}, \tilde{m}_r)\big)$, which is jointly typical to y_2^n. Next, it uses the estimated $\tilde{m}_\otimes$ and m_{11} to produce the message $\tilde{m}_{21} = \tilde{m}_\otimes \otimes m_{11}$. Finally, it outputs $\tilde{m}_2 = (\tilde{m}_{21}, \tilde{m}_{22})$.

6. Reliability Analysis: In order to evaluate this coding scheme with respect to the reliability measure in (3.107), it is enough to prove that:

$$\lim_{n\to\infty} \mathbb{P}\big[(\mathrm{M}_\otimes, \mathrm{M}_{12}, \mathrm{M}_r) \neq (\hat{\mathrm{M}}_\otimes, \hat{\mathrm{M}}_{12}, \hat{\mathrm{M}}_r) \text{ or } (\mathrm{M}_\otimes, \mathrm{M}_{22}, \mathrm{M}_r) \neq (\tilde{\mathrm{M}}_\otimes, \tilde{\mathrm{M}}_{22}, \tilde{\mathrm{M}}_r)\big] = 0. \tag{3.116}$$

Based on the principle of superposition encoding and the standard joint-typicality decoding arguments, Eq. (3.116) holds as long as:

$$R_{11} = R_{21} = R_\otimes \leq \min[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)] - \delta_0(\epsilon) \tag{3.117a}$$
$$R_{12} + R_r \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \delta_1(\epsilon) \tag{3.117b}$$
$$R_{22} + R_r \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \delta_2(\epsilon) \tag{3.117c}$$

where $\epsilon > 0$ is the constant used during the typicality decoding. Additionally, the equality of R_{11}, R_{21} and $R_\otimes$ follows due to the codebook generation. Moreover, ϵ can be chosen in a way such that $\lim_{n\to\infty} \delta_0(\epsilon), \delta_1(\epsilon), \delta_2(\epsilon) = 0$.

7. Secrecy Analysis: Due to the splitting of each confidential message set into two parts, the secrecy measure in (3.108) can be reformulated with respect to this coding scheme as follows:

$$\begin{aligned}\mathbf{L}_i(\mathcal{C}) &\overset{(a)}{=} \mathbb{I}(\mathrm{M}_{12};\mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_{12}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{22};\mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_{21};\mathrm{Z}^n|\mathrm{M}_{22}\mathcal{C}) \\ &\overset{(b)}{\leq} \mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_{12}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{21};\mathrm{Z}^n|\mathrm{M}_{22}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{12}\mathrm{M}_{22};\mathrm{Z}^n|\mathcal{C}) \\ &\overset{(c)}{\leq} \mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_{12}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{21};\mathrm{Z}^n|\mathrm{M}_{22}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{12}\mathrm{M}_{22};\mathrm{Z}^n|\mathrm{M}_{\otimes}\mathcal{C}), \end{aligned} \tag{3.118}$$

where (a) follows from the mutual information chain rule; while (b) and (c) follow from the independence of the messages. The first term in (3.118) can be bounded as follows:

$$\begin{aligned}\mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_{12}\mathcal{C}) &\leq \mathbb{H}(\mathrm{M}_{11}) - \mathbb{H}(\mathrm{M}_{11}|\mathrm{Z}^n\mathrm{M}_{12}\mathcal{C}) \\ &\overset{(a)}{\leq} \mathbb{H}(\mathrm{M}_{11}) - \mathbb{H}(\mathrm{M}_{11}|\mathrm{M}_{\otimes}\mathrm{M}_{12}\mathrm{M}_{22}\mathrm{M}_r) \overset{(b)}{=} 0, \end{aligned} \tag{3.119}$$

where (a) follows because for a certain code $\mathcal{C}$, $(\mathrm{M}_{\otimes}, \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{M}_r) - \mathrm{X}^n - \mathrm{Z}^n$ forms a Markov chain; while (b) follows based on the secrecy analysis of the secret key encoding scheme, as long as:

$$\mathbb{H}(\mathrm{M}_{11}) = \mathbb{H}(\mathrm{M}_{21}) \qquad \Longrightarrow \qquad R_{11} = R_{21}. \tag{3.120}$$

Using the same argument, we can show that the second term in (3.118) also vanishes as long as the condition in (3.120) is satisfied. On the other hand, based on the strong secrecy analysis of wiretap random codes presented in the Section 2.2, the third term in (3.118) decays exponentially in n, if:

$$R_r \geq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \tilde{\delta}(\epsilon, \tau), \tag{3.121}$$

where $\tau > 0$. Additionally, τ can be selected in a way such that $\lim_{n\to\infty} \tilde{\delta}(\epsilon, \tau) = 0$. This implies that the rate constraints in (3.120) and (3.121) assure that

$$\lim_{n\to\infty} \mathbf{L}_i(\mathcal{C}) = 0. \tag{3.122}$$

Finally, by combining Eqs. (3.115), (3.117), (3.120) and (3.121), followed by applying the Fourier-Motzkin elimination procedure [57], then taking the limit as $n \to \infty$, it follows directly that any rate pair (R_1, R_2) that satisfies the constraints in (3.114) is achievable.

3. SK Randomization Index for WRC: This scheme also combines the ideas of secret key encoding and wiretap random coding (WRC) along with concept of rate splitting, but from a different perspective. It can be related in some sense to the coding scheme used to prove Scenario 2 of Theorem 2.3. The achievable rate region for this scheme $\mathcal{R}_3$ is given by the set of rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy the following:

$$\begin{aligned} R_1 &\leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1) - \mathbb{I}(\mathrm{X};\mathrm{Z}) + R_2, \mathbb{I}(\mathrm{X};\mathrm{Y}_1)\right] \\ R_2 &\leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_2) - \mathbb{I}(\mathrm{X};\mathrm{Z}) + R_1, \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\right]. \end{aligned} \tag{3.123}$$

The previous rate region is only valid under the following condition: $\mathbb{I}(\mathrm{X};\mathrm{Y}_1) \geq \mathbb{I}(\mathrm{X};\mathrm{Z})$ and $\mathbb{I}(\mathrm{X};\mathrm{Y}_2) \geq \mathbb{I}(\mathrm{X};\mathrm{Z})$. This implies that this coding scheme is only valid for input

distributions $\mathrm{P_X}$ that promote the legitimate receivers to have a statistical advantage over the eavesdropper.

1. Codebook Generation $\mathcal{C}$: We start by applying the same rate splitting procedure introduced for the previous coding scheme, such that the rate constraints in (3.115) are fulfilled. We also let $\mathcal{M}_r = [\![1, 2^{nR_r}]\!]$ be a randomization message set, where M_r is a uniformly distributed random variable over it. Next, fix an input distribution $\mathrm{P_X}$ and randomly generate the codewords $x^n(m_{12}, m_{22}, m_r, m_\otimes)$ for $m_{12} \in \mathcal{M}_{12}$, $m_{22} \in \mathcal{M}_{22}$, $m_r \in \mathcal{M}_r$ and $m_\otimes \in \mathcal{M}_\otimes$ by generating the symbols $x_i(m_{12}, m_{22}, m_r, m_\otimes)$ independently at random according to $\mathrm{P_X}$.

2. Encoding E: Given the confidential message pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, where $m_1 = (m_{11}, m_{12})$ and $m_2 = (m_{21}, m_{22})$, the encoder selects a message m_r uniformly at random, then transmits the corresponding codeword $x^n(m_{12}, m_{22}, m_r, m_\otimes)$.

3. First Legitimate Decoder φ_1: Upon receiving y_1^n, the decoder uses it along with the available side information $m_2 = (m_{21}, m_{22})$ to determine a unique codeword $x^n(\hat{m}_{12}, m_{22}, \hat{m}_r, \hat{m}_\otimes)$ which is jointly typical to y_1^n. Next, it produces the message $\hat{m}_{11} = \hat{m}_\otimes \otimes m_{21}$. Finally, it outputs $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12})$.

4. Second Legitimate Decoder φ_2: Upon receiving y_2^n, the decoder uses it along with the available side information $m_1 = (m_{11}, m_{12})$ to determine a unique codeword $x^n(m_{12}, \tilde{m}_{22}, \tilde{m}_r, \tilde{m}_\otimes)$ which is jointly typical with y_2^n. Next, it produces the message $\tilde{m}_{21} = \tilde{m}_\otimes \otimes m_{11}$. Finally, it outputs $\tilde{m}_2 = (\tilde{m}_{21}, \tilde{m}_{22})$.

5. Reliability Analysis: In order to evaluate this coding scheme with respect to the reliability measure in (3.107), it is enough to prove that the bound in (3.116) holds. Using the standard joint-typicality decoding arguments, we have:

$$R_{12} + R_r + R_\otimes \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_1) - \delta_1(\epsilon) \tag{3.124a}$$

$$R_{22} + R_r + R_\otimes \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_2) - \delta_2(\epsilon). \tag{3.124b}$$

6. Secrecy Analysis: We start by reformulating the secrecy measure in (3.108) with respect to this coding scheme as follows:

$$\mathbf{L}_i(\mathcal{C}) \leq \mathbb{I}(\mathrm{M}_{11}; \mathrm{Z}^n | \mathrm{M}_{12}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{21}; \mathrm{Z}^n | \mathrm{M}_{22}\mathcal{C}) + \mathbb{I}(\mathrm{M}_{12}\mathrm{M}_{22}; \mathrm{Z}^n | \mathcal{C}) \tag{3.125}$$

The previous bound follows based on the same argument used to establish the one in (3.118). Based on the principles of secret key encoding, one can show that the first and second terms in Eq. (3.125) vanish as long as the condition in (3.120) holds. On the other hand, based on the strong secrecy analysis of wiretap random codes presented in the Section 2.2, the third term in (3.125) decays exponentially in n, if:

$$R_r + R_\otimes \geq \mathbb{I}(\mathrm{X}; \mathrm{Z}) + \tilde{\delta}(\epsilon, \tau). \tag{3.126}$$

Finally, by combining the bounds in (3.115), (3.120), (3.124) and (3.126) followed by applying the Fourier-Motzkin elimination procedure, then taking the limit as $n \to \infty$, we prove that any rate pair (R_1, R_2) that satisfies the constraints in (3.123) is achievable.

Before we proceed to the next section, we need to highlight some important points regarding the three discussed coding schemes:

- The rate region in (3.109) established using the first coding scheme follows as a special case from the rate region in (3.114) established using the second coding scheme by letting U = X. This implies that $\mathcal{R}_1 \subset \mathcal{R}_2$.
- Despite the similarities between the second and third coding scheme, the rate region in (3.123) can not be derived as a special case of the one in (3.114). In fact, if the two legitimate receivers are more capable than the eavesdropper, one can show that $\mathcal{R}_2 \subset \mathcal{R}_3$.
- Although the second coding scheme can be used for any input distribution and any channel, it is not optimal in general. On the other hand, the third coding scheme can –in fact– establish bigger rate regions, but only under certain conditions for the input distribution and the channel.
- Although the first coding scheme allows one of the legitimate receivers to reach its maximum achievable transfer rate, it limits the achievable rate of the other legitimate receiver –which might have a better channel– to the rate of the weaker receiver. Differently, the third coding scheme allows both legitimate receivers to reach their maximum transfer rate, but under the condition that their channels are better than the eavesdropper's channel. The second coding scheme is an intermediate solution between these two.
- The condition on the input distribution and the channel in the third coding scheme arises from the decoding procedure. This is because according to the definitions of the decoders, each legitimate receiver does not only decode its intended message, but it has to decode the randomization index as well.

3.3.2 A Universal Individual-Strong Secrecy Coding Scheme

We now investigate secure communication for our full model in Fig. 3.2 under the individual-strong secrecy criterion. We start our investigation by modifying Definition 3.2 as follows:

Definition 3.4. A rate quadruple $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ is an individual-strong secrecy achievable rate quadruple for the two-receiver DM-WBC, i.e. $Q = (W^1, W^2, V)$, if for all $\eta_i > 0$ where $i \in \{c, 0, 1, 2\}$, $\lambda > 0$ and $\tau_j > 0$ where $j \in \{1, 2\}$, there is an $n(\eta_i, \lambda, \tau_j) \in \mathbb{N}$, such that for all $n > n(\eta_i, \lambda, \tau_j)$ there exists a sequence of codes $\{\mathcal{C}\}_n$ according to Definition 3.1 that satisfies the following constraints:

$$\frac{1}{n}\log|\mathcal{M}_i| \geq R_i - \eta_i \qquad \text{for} \qquad i \in \{c, 0, 1, 2\} \tag{3.127a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda \tag{3.127b}$$

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathcal{C}) \leq \tau_1 \qquad \text{and} \qquad \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) \leq \tau_2, \tag{3.127c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by Eq. (3.7). The individual-strong secrecy capacity region $C_I(Q)$ is given by the set of all achievable individual-strong secrecy rate quadruples (R_c, R_0, R_1, R_2).

In order to show that a rate quadruple (R_c, R_0, R_1, R_2) is achievable under the individual-strong secrecy criterion, it is enough to show that there exists a random coding scheme with random variable C such that:

$$\frac{1}{n}\log|\mathcal{M}_i| = R_i \tag{3.128a}$$

$$\lim_{n\to\infty}\mathbb{E}[\bar{\mathbf{P}}_e(\mathrm{C})] = 0 \tag{3.128b}$$

$$\lim_{n\to\infty}\mathbb{E}[\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{C}) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_2;\mathrm{Z}^n|\mathrm{C})] = 0, \tag{3.128c}$$

where $i \in \{c, 0, 1, 2\}$. The validity of the previous argument follows from Definition 3.4 in addition to the selection lemma. We now present our general individual-strong secrecy achievable rate region as follows:

Theorem 3.2. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition is given by the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned}
R_c &\le \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1|\mathrm{T}), \mathbb{I}(\mathrm{U};\mathrm{Y}_2|\mathrm{T})\Big]\\
R_0 + R_1 &\le \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_\otimes\,\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{V}_\otimes\,\mathrm{T})\Big]^+ + R_\otimes\\
R_0 + R_2 &\le \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_\otimes\,\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{V}_\otimes\,\mathrm{T})\Big]^+ + R_\otimes\\
2R_0 + R_1 + R_2 &\le \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_\otimes\,\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_\otimes\,\mathrm{T}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\,\mathrm{T})\\
&\quad - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{V}_\otimes\,\mathrm{T}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{V}_\otimes\,\mathrm{T})\Big]^+ + 2R_\otimes \qquad (3.129)\\
\textit{where}\quad R_\otimes &\le \min\Big[R_1, R_2, \mathbb{I}(\mathrm{V}_\otimes;\mathrm{Y}_1|\mathrm{U}\,\mathrm{T}), \mathbb{I}(\mathrm{V}_\otimes;\mathrm{Y}_2|\mathrm{U}\,\mathrm{T})\Big],
\end{aligned}$$

for random variables with joint probability distribution $\mathrm{Q}_\mathrm{T}\ \mathrm{Q}_{\mathrm{U}|\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_\otimes|\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_\otimes}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{T} - \mathrm{U} - \mathrm{V}_\otimes - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous theorem is based on a coding scheme that combines the principles of superposition encoding, rate splitting, secret key encoding and wiretap Marton coding along with the concept of time sharing. The combination between the secret key encoding and the wiretap Marton coding follows the second coding scheme introduced in Section 3.3.1. In particular, the secret key encoded part of the individual confidential messages is encoded in a separate layer upon which the wiretap Marton coding is superimposed, where the wiretap Marton code is used to encode the common confidential message and the remaining parts of the individual confidential messages.

In order to prove Theorem 3.2, we will use the same algorithm used to prove Theorem 3.1. Thus, we start by presenting the following proposition:

Proposition 3.3. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition is given by the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned}
R_c &\le \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\Big]\\
R_0 + R_1 &\le \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_\otimes) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{V}_\otimes)\Big]^+ + R_\otimes\\
R_0 + R_2 &\le \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_\otimes) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{V}_\otimes) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0)\Big]^+ + R_\otimes
\end{aligned}$$

$$\textit{where} \qquad R_{\otimes} \leq \min\Big[R_1, R_2, \mathbb{I}(\mathrm{V}_{\otimes};\mathrm{Y}_1|\mathrm{U}), \mathbb{I}(\mathrm{V}_{\otimes};\mathrm{Y}_2|\mathrm{U})\Big] \tag{3.130}$$

for random variables with joint probability distribution $\mathrm{Q}_{\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_{\otimes}|\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_{\otimes}}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}$ $Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{U} - \mathrm{V}_{\otimes} - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous proposition plays a major role in the proof of Theorem 3.2 similar to the role played by Proposition 3.1 in proving Theorem 3.1. This is because the validity of Proposition 3.3 directly implies the validity of the following result:

Proposition 3.4. *An achievable individual strong secrecy rate region for the wiretap BC with degraded message sets and message cognition is given by the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\Big] \\ R_0 + R_1 &\leq \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_{\otimes}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{V}_{\otimes}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2\mathrm{Z}|\mathrm{V}_0)\Big]^+ + R_{\otimes} \\ R_0 + R_2 &\leq \Big[\mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_{\otimes}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{V}_{\otimes})\Big]^+ + R_{\otimes} \\ \textit{where} \qquad R_{\otimes} &\leq \min\Big[R_1, R_2, \mathbb{I}(\mathrm{V}_{\otimes};\mathrm{Y}_1|\mathrm{U}), \mathbb{I}(\mathrm{V}_{\otimes};\mathrm{Y}_2|\mathrm{U})\Big] \end{aligned} \tag{3.131}$$

for random variables with joint probability distribution $\mathrm{Q}_{\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_{\otimes}|\mathrm{U}}\ \mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_{\otimes}}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}$ $Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{U} - \mathrm{V}_{\otimes} - (\mathrm{V}_0, \mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

Finally, if we use the rate splitting and time sharing scheme introduced in Section 3.2.5 to combine the rate regions established by Proposition 3.3 and Proposition 3.4, we can prove the achievability of any rate quadruple (R_c, R_0, R_1, R_2) that satisfies the rate constraints in (3.129). Thus, we only need to provide a proof for Proposition 3.3.

Let $n \in \mathbb{N}$ be an arbitrary but fixed code block length and $\epsilon > 0$ be an arbitrary constant. Moreover, let $\mathrm{Q}_{\mathrm{U}\mathrm{V}_{\otimes}\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2\mathrm{X}} \in \mathcal{P}(\mathcal{U} \times \mathcal{V}_{\otimes} \times \mathcal{V}_0 \times \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{X})$ be a fixed input probability distribution, and recall the probability transition matrix $Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$ given in (2.73), that defines the two-receiver DM-WBC. Furthermore, we assume that quantities of the form 2^{nR}, where $R \in \mathbb{R}_+$ are integers.

1. Message Sets: We consider the following sets: $\mathcal{M}_c$, $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_r$, $\mathcal{M}_{r_1}$, $\mathcal{M}_{r_2}$ and $\mathcal{M}_t$, which are defined exactly as in the coding scheme introduced in Section 3.2.2. Further, we divide each confidential individual messages set into two sets as follows: $\mathcal{M}_1 = \mathcal{M}_{11} \times \mathcal{M}_{12}$ and $\mathcal{M}_2 = \mathcal{M}_{21} \times \mathcal{M}_{22}$, where $\mathcal{M}_{ab} = [\![1, 2^{nR_{ab}}]\!]$, for $a, b \in \{1, 2\}$. In this division, we force $\mathcal{M}_{11}$ and $\mathcal{M}_{21}$ to be of the same size and use them to construct a new set $\mathcal{M}_{\otimes} = [\![1, 2^{nR_{\otimes}}]\!]$ by *xoring* the corresponding elements of both sets. Additionally, we use $\mathcal{M} = \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{M}_{22}$ to abbreviate the set of the remaining confidential messages. It is important to note that the message structure forces the following condition:

$$R_{\otimes} = R_{11} = R_{21} \leq \min\big[R_1, R_2\big]. \tag{3.132}$$

We can highlight here that $\mathcal{M}_{11}$ and $\mathcal{M}_{21}$ can be interpreted as the part of the confidential messages, that will be protected against eavesdropping using secret key encoding. On the other hand, $\mathcal{M}$ will be the part of the confidential messages protected by wiretap random coding.

2. Random Codebook $\mathcal{C}$: Let $\mathrm{C}_c \triangleq \{\mathrm{U}^n(m_c)\}$ for $m_c \in \mathcal{M}_c$ be a random public common message codebook as defined in Section 3.2.2. For a fixed public message codebook $\mathcal{C}_c$ and for every $m_c \in \mathcal{M}_c$, let $\mathrm{C}_\otimes(m_c) \triangleq \{\mathrm{V}^n_\otimes(m_c, m_\otimes)\}$ where $m_\otimes \in \mathcal{M}_\otimes$ be a random codebook for the *xored* message that consists of $2^{nR_\otimes}$ conditionally independent random codewords $\mathrm{V}^n_\otimes(m_c, m_\otimes)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{\otimes_i}(m_c, m_\otimes)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_\otimes|\mathrm{U}}(\cdot|u_i(m_c))$. A realization of $\mathrm{C}_\otimes(m_c)$ is denoted by $\mathcal{C}_\otimes(m_c) \triangleq \{v^n_\otimes(m_c, m_\otimes)\}$.

For a fixed *xored* message codebook $\mathcal{C}_\otimes(m_c)$ and for every $m_\otimes \in \mathcal{M}_\otimes$, let $\mathrm{C}_0(m_c, m_\otimes) \triangleq \{\mathrm{V}^n_0(m_c, m_\otimes, m, m_r)\}$ where $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$ be a random codebook for the confidential messages that consists of $2^{n(R+R_r)}$ conditionally independent random codewords $\mathrm{V}^n_0(m_c, m_\otimes,\ m, m_r)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{0_i}(m_c, m_\otimes, m, m_r)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_0|\mathrm{V}_\otimes}(\cdot|v_{\otimes_i}(m_c, m_\otimes))$. A realization of $\mathrm{C}_0(m_c, m_\otimes)$ is denoted by $\mathcal{C}_0(m_c, m_\otimes) \triangleq \{v^n_0(m_c, m_\otimes, m, m_r)\}$.

For a fixed confidential messages codebook $\mathcal{C}_0(m_c, m_\otimes)$ and for every $m \in \mathcal{M}$ and $m_r \in \mathcal{M}_r$, let $\mathrm{C}_1(m_c, m_\otimes, m, m_r) \triangleq \{\mathrm{V}^n_1(m_c, m_\otimes, m, m_r, m_{r_1})\}$ and $\mathrm{C}_2(m_c, m_\otimes, m, m_r) \triangleq \{\mathrm{V}^n_2(m_c, m_\otimes, m, m_r,\ m_{r_2}, m_t)\}$ be two confidential random codebooks that consist of $2^{nR_{r_1}}$ and $2^{n(R_{r_2}+R_t)}$ conditionally independent random codewords respectively.

The total random codebook is denoted by $\mathrm{C} = \{\mathrm{C}_c, \mathrm{C}_\otimes, \mathrm{C}_0, \mathrm{C}_1, \mathrm{C}_2\}$, while $\mathcal{C} = \{\mathcal{C}_c, \mathcal{C}_\otimes, \mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2\}$ denotes one possible realization of this codebook selected from the total set of codebooks given by $\mathfrak{C}$. The probability of generating a certain realization $\mathcal{C} \in \mathfrak{C}$ is given by:

$$\begin{aligned}
\mathrm{G}(\mathcal{C}) = & \prod_{\substack{(m^2_c, m^2_\otimes, m^2, m^2_r, m_{r_2}, m_t)\in \\ \mathcal{M}_c\times\mathcal{M}_\otimes\times\mathcal{M}\times\mathcal{M}_r\times\mathcal{M}_{r_2}\times\mathcal{M}_t}} \mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0}\Big(v^n_2(m^2_c, m^2_\otimes, m^2, m^2_r, m_{r_2}, m_t)|v^n_0(m^2_c, m^2_\otimes, m^2, m^2_r)\Big) \\
& \prod_{\substack{(m^1_c, m^1_\otimes, m^1, m^1_r, m_{r_1})\in \\ \mathcal{M}_c\times\mathcal{M}_\otimes\times\mathcal{M}\times\mathcal{M}_r\times\mathcal{M}_{r_1}}} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}\Big(v^n_1(m^1_c, m^1_\otimes, m^1, m^1_r, m_{r_1})|v^n_0(m^1_c, m^1_\otimes, m^1, m^1_r)\Big) \\
& \prod_{\substack{(m^0_c, m^0_\otimes, m, m_r)\in \\ \mathcal{M}_c\times\mathcal{M}_\otimes\times\mathcal{M}\times\mathcal{M}_r}} \mathrm{Q}^n_{\mathrm{V}_0|\mathrm{U}}\Big(v^n_0(m^0_c, m^0_\otimes, m, m_r)|u^n(m^0_c)\Big) \prod_{m_c\in\mathcal{M}_c} \mathrm{Q}^n_{\mathrm{U}}\big(u^n(m_c)\big) \\
& \prod_{(m^\otimes_c, m_\otimes)\in\mathcal{M}_c\times\mathcal{M}_\otimes} \mathrm{Q}^n_{\mathrm{V}_0|\mathrm{V}_\otimes}\Big(v^n_\otimes(m^\otimes_c, m_\otimes)|u^n(m^\otimes_c)\Big)
\end{aligned} \tag{3.133}$$

3. Encoder E: For a fixed codebook realization $\mathcal{C}$, given a message triple $(m_c, m_\otimes, m)$, where $m_\otimes = m_{11} \otimes m_{21}$ and $m = (m_0, m_{12}, m_{22})$, the encoder chooses three randomization messages m_r, m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_r$, $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. Then, it chooses an index $m_t \in \mathcal{M}_t$ using a likelihood encoder similar to the one in (3.20). Finally, for the selected triple $v^n_0(m_c, m_\otimes, m, m_r)$, $v^n_1(m_c, m_\otimes, m, m_r, m_{r_1})$, and $v^n_2(m_c, m_\otimes, m, m_r, m_{r_2}, m_t)$, the encoder generates a codeword x^n independently at random according to $\prod_{i=1}^n \mathrm{Q}^n_{\mathrm{X}|\mathrm{V}_0,\mathrm{V}_1,\mathrm{V}_2}\big(\cdot\,|v^n_0(m_c, m_\otimes, m, m_r),$ $v^n_1(m_c, m_\otimes, m, m_r,\ m_{r_1}), v^n_2(m_c, m_\otimes, m, m_r, m_{r_2}, m_t)\big)$ and transmits it.

4. First Legitimate Decoder φ_1: Given y^n_1 and its own message $m_2 = (m_{21}, m_{22})$, the decoder searches for the unique pentadruple $(\hat{m}_c, \hat{m}_\otimes, \hat{m}_0, \hat{m}_{12}, \hat{m}_r) \in \mathcal{M}_c \times \mathcal{M}_\otimes \times \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{M}_r$; for which there exists an index $\hat{m}_{r_1} \in \mathcal{M}_{r_1}$ such that:

$$\Big(u^n(\hat{m}_c), v_{\otimes}^n(\hat{m}_c, \hat{m}_{\otimes}), v_0^n(\hat{m}_c, \hat{m}_{\otimes}, \hat{m}, \hat{m}_r), v_1^n(\hat{m}_c, \ldots, \hat{m}_{r_1}), y_1^n\Big) \in \mathcal{T}_{\epsilon}^n(\mathrm{Q}_{\mathrm{UV}_{\otimes}\mathrm{V}_0\mathrm{V}_1\mathrm{Y}_1}),$$

where $\hat{m} = (\hat{m}_0, \hat{m}_{12}, m_{22})$. If such unique pentadruple exists, the decoder uses $\hat{m}_{\otimes}$ and m_{21} to produce the message $\hat{m}_{11} = \hat{m}_{\otimes} \otimes m_{21}$, then it outputs the triple $(\hat{m}_c, \hat{m}_0, \hat{m}_1)$, where $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12})$. Otherwise, the decoder outputs the error message (?). Similar to the first legitimate receiver introduced in Section 3.2.2, this decoder also applies the principle of indirect decoding by including the sequence $v_1^n(\hat{m}_c, \hat{m}_{\otimes}, \hat{m}, \hat{m}_r, \hat{m}_{r_1})$ in the decoding rule despite the fact that it does not aim to find the unique message m_{r_1} that was transmitted.

5. Second Legitimate Decoder φ_2: Given y_2^n and its own message m_1, the decoder searches for the unique pentadruple $(\tilde{m}_c, \tilde{m}_{\otimes}, \tilde{m}_0, \tilde{m}_{22}, \tilde{m}_r) \in \mathcal{M}_c \times \mathcal{M}_{\otimes} \times \mathcal{M}_0 \times \mathcal{M}_{22} \times \mathcal{M}_r$; for which there exist two indices $(\tilde{m}_{r_2}, \tilde{m}_t) \in \mathcal{M}_{r_2} \times \mathcal{M}_t$ such that:

$$\Big(u^n(\tilde{m}_c), v_{\otimes}^n(\tilde{m}_c, \tilde{m}_{\otimes}), v_0^n(\tilde{m}_c, \tilde{m}_{\otimes}, \tilde{m}, \tilde{m}_r), v_2^n(\tilde{m}_c, \ldots, \tilde{m}_{r_2}, \tilde{m}_t), y_2^n\Big) \in \mathcal{T}_{\epsilon}^n(\mathrm{Q}_{\mathrm{UV}_{\otimes}\mathrm{V}_0\mathrm{V}_2\mathrm{Y}_2}),$$

where $\tilde{m} = (\tilde{m}_0, m_{12}, \tilde{m}_{22})$. If such unique pentadruple exists, the decoder uses $\tilde{m}_{\otimes}$ and m_{11} to deduce the message $\tilde{m}_{21} = \tilde{m}_{\otimes} \otimes m_{11}$, then it outputs the triple $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2)$ where $\tilde{m}_2 = (\tilde{m}_{21}, \tilde{m}_{22})$. If such unique pentadruple does not exist, the decoder outputs the error message (?). Again, one should note that this decoder also applies the concept of indirect decoding, as it is not interested in finding the unique pair (m_{r_2}, m_t) that was transmitted.

6. Third Eavesdropper Decoder φ_3: Given z^n, the decoder uses the same decoding function introduced in Section 3.2.2.

7. Induced Joint Distribution: For every codebook $\mathcal{C} \in \mathfrak{C}$, the previous coding scheme induces the following joint probability distribution:

$$\begin{aligned}
&\mathrm{P}^{\mathcal{C}}\Big(m_c, \ldots, m_{r_1}, m_{r_2}, u^n, v_{\otimes}^n, v_0^n, v_1^n, m_t, v_2^n, x^n, y_1^n, y_2^n, z^n, \hat{m}_c, \hat{m}_0, \hat{m}_1, \tilde{m}_c, \tilde{m}_0, \tilde{m}_2, \check{m}_c\Big) \\
&\quad = 2^{-n(R_c+R_{\otimes}+R+R_r+R_{r_1}+R_{r_2})} \mathbb{1}_{\{u^n=u^n(m_c)\}} \mathbb{1}_{\left\{v_{\otimes}^n=v_{\otimes}^n(m_c,m_{\otimes})\right\}} \mathbb{1}_{\left\{v_0^n=v_0^n(m_c,m_{\otimes},m,m_r)\right\}} \\
&\quad \times \mathbb{1}_{\left\{v_1^n=v_1^n(m_c,m_{\otimes},m,m_r,m_{r_1})\right\}} \mathbb{1}_{\left\{v_2^n=v_2^n(m_c,m_{\otimes},m,m_r,m_{r_2},m_t)\right\}} \mathbb{1}_{\left\{(\hat{m}_c,\hat{m}_0,\hat{m}_1)=\varphi_1(y_1^n)\right\}} \\
&\quad \times E^{(\mathrm{LE})}\big(m_t|m_{r_2}, v_0^n(m_c, m_{\otimes}, m, m_r), v_1^n(m_c, m_{\otimes}, m, m_r, m_{r_1})\big) \mathbb{1}_{\{\check{m}_c=\varphi_3(z^n)\}} \\
&\quad \times \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}^n(x^n|v_0^n, v_1^n, v_2^n)\, Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}^n(y_1^n, y_2^n, z^n|x^n) \mathbb{1}_{\left\{(\tilde{m}_c,\tilde{m}_0,\tilde{m}_2)=\varphi_2(y_2^n)\right\}}
\end{aligned} \tag{3.134}$$

Additionally, if we take the random codebook generation process into our consideration, we end up with the subsequent distribution: $\mathrm{P} = \mathrm{G}(\mathcal{C}) \cdot \mathrm{P}^{\mathcal{C}}$, where $\mathrm{G}(\mathcal{C})$ is the probability of obtaining a certain codebook $\mathcal{C}$ given by (3.133).

8. Reliability Analysis: In this paragraph, we derive the rate constraints under which the reliability constraint in (3.128b) is satisfied. We start by defining the average decoding error probability for a given code realization $\mathcal{C} \in \mathfrak{C}$ as follows:

$$\begin{aligned}
\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \triangleq \mathbb{P}\big[&\check{\mathrm{M}}_c \neq \mathrm{M}_c \text{ or } (\hat{\mathrm{M}}_c, \hat{\mathrm{M}}_{\otimes}, \hat{\mathrm{M}}_0, \hat{\mathrm{M}}_{11}, \hat{\mathrm{M}}_r) \neq (\mathrm{M}_c, \mathrm{M}_{\otimes}, \mathrm{M}_0, \mathrm{M}_{11}, \mathrm{M}_r) \text{ or} \\
&(\tilde{\mathrm{M}}_c, \tilde{\mathrm{M}}_{\otimes}, \tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_{22}, \tilde{\mathrm{M}}_r) \neq (\mathrm{M}_c, \mathrm{M}_{\otimes}, \mathrm{M}_0, \mathrm{M}_{22}, \mathrm{M}_r)\big].
\end{aligned} \tag{3.135}$$

We then observe that $\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \geq \bar{\mathbf{P}}_e(\mathcal{C})$ in (3.7). This implies that in order to show that our coding scheme satisfies the reliability constraint in (3.128b), it is enough to show

that $\mathbb{E}_{\mathrm{C}}[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})]$ approaches zero as n approaches infinity. Similar to the reliability analysis presented in Section 3.2.3, we need to consider two types of errors. The first is the encoding error $\mathcal{E}$, which occurs if the sequence v_2^n selected by the likelihood encoder is not jointly typical to the other codewords. Following the exact steps used in Section 3.2.3, we can show that for a typicality constant ϵ and a constant $\beta > 0$, $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\mathbb{P}(\mathcal{E})\big] = 0$, as long as:

$$R_t \geq \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) + \acute{\delta}_0(\epsilon,\beta), \tag{3.136}$$

$$\text{for} \qquad \acute{\delta}_0(\epsilon,\beta) = \epsilon \log\big(|\mathcal{U}|\cdot|\mathcal{V}_\otimes|\cdot|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\big) + 5\beta, \tag{3.137}$$

The second class of errors is the decoding error $\mathcal{D}$, which arises if any of the decoders failed to find a unique set of messages that satisfies the corresponding decoding rule. Again, following the same technique used in Section 3.2.3, keeping in mind the new superposition order used in the codebook generation, we can show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\mathbb{P}(\mathcal{D}|\mathcal{E}^c)\big] = 0$, as long as:

$$\begin{aligned}
R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1) - \acute{\delta}_1(\epsilon), \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \acute{\delta}_2(\epsilon), \mathbb{I}(\mathrm{U};\mathrm{Z}) - \delta_3(\epsilon)\Big] \\
R_\otimes &\leq \min\Big[\mathbb{I}(\mathrm{V}_\otimes;\mathrm{Y}_1|\mathrm{U}) - \acute{\delta}_1(\epsilon), \mathbb{I}(\mathrm{V}_\otimes;\mathrm{Y}_2|\mathrm{U}) - \acute{\delta}_2(\epsilon)\Big] \\
R_0 + R_{11} + R_r + R_{r_1} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_\otimes) - \acute{\delta}_1(\epsilon) \\
R_0 + R_{22} + R_r + R_{r_2} + R_t &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_\otimes) - \acute{\delta}_2(\epsilon).
\end{aligned} \tag{3.138}$$

where $\acute{\delta}_1(\epsilon)$ and $\acute{\delta}_2(\epsilon)$ are defined by modifying the definitions of $\delta_1(\epsilon)$ and $\delta_2(\epsilon)$ in the same way $\acute{\delta}_0(\epsilon,\beta)$ is a modified version of $\delta_0(\epsilon,\beta)$ cf. (3.43) and (3.137). The rate constraints in (3.136) and (3.138) directly imply that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\bar{\mathbf{P}}_e(\mathrm{C})\big] \leq \lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})\big] = 0$. This implies that under these rate constraints our coding scheme fulfills the reliability constraint in (3.128b).

9. Secrecy Analysis: Because of the new message sets structure, the random variable M_1 is now identified as the product of two independent and uniformly distributed random variables M_{11} and M_{12}. Thus, for any code realization $\mathcal{C}$, we have

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathcal{C}) &= \mathbb{I}(\mathrm{M}_0\mathrm{M}_{11}\mathrm{M}_{12};\mathrm{Z}^n|\mathcal{C}) \\
&= \mathbb{I}(\mathrm{M}_0\mathrm{M}_{12};\mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_0\mathrm{M}_{12}\mathcal{C})
\end{aligned} \tag{3.139}$$

The term $\mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_0\mathrm{M}_{12}\mathcal{C})$ represents the information leakage of the confidential message M_{11} to the eavesdropper given M_0 and M_{12}. This confidential message is protected against eavesdropping using secret key encoding. We can show that for every $\mathcal{C} \in \mathfrak{C}$, this term vanishes as follows:

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_{11};\mathrm{Z}^n|\mathrm{M}_0\mathrm{M}_{12}\mathcal{C}) &\overset{(a)}{\leq} \mathbb{H}(\mathrm{M}_{11}) - \mathbb{H}(\mathrm{M}_{11}|\mathrm{Z}^n\mathrm{M}_0\mathrm{M}_{12}\mathcal{C}) \\
&\overset{(b)}{\leq} \mathbb{H}(\mathrm{M}_{11}) - \mathbb{H}(\mathrm{M}_{11}|\mathrm{M}_c\mathrm{M}_\otimes\mathrm{M}_0\mathrm{M}_{12}\mathrm{M}_{22}\mathrm{M}_r\mathrm{M}_{r_1}\mathrm{M}_{r_2}\mathrm{M}_t) \\
&= \mathbb{H}(\mathrm{M}_{11}) - \mathbb{H}(\mathrm{M}_{11}|\mathrm{M}_\otimes) \overset{(c)}{=} 0,
\end{aligned} \tag{3.140}$$

where (a) follows from the independence of the messages; (b) follows because for every code $\mathcal{C}$, the following Markov chain holds $(\mathrm{M}_c, \mathrm{M}_\otimes, \mathrm{M}_0, \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{M}_{r_2}, \mathrm{M}_t) -$

$\mathrm{X}^n - \mathrm{Z}^n$; while (c) follows due to the principles of secret key encoding, and the fact that $\mathbb{H}(\mathrm{M}_{11}) = \mathbb{H}(\mathrm{M}_{21})$, where this equality holds because both M_{11} and M_{21} are uniformly distributed over two sets with the same size cf.(3.132). Using the same arguments, we can show that:

$$\mathbb{I}(\mathrm{M}_{12}; \mathrm{Z}^n | \mathrm{M}_0 \mathrm{M}_{22} \mathcal{C}) = 0, \tag{3.141}$$

where the product of two independent and uniformly distributed random variables M_{21} and M_{22} defines M_2. Now, for a fixed code $\mathcal{C}$, let $\mathbf{L}_i(\mathcal{C}) \triangleq \mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C})$. Then, it follows that:

$$\begin{aligned}
\mathbf{L}_i(\mathcal{C}) &\overset{(a)}{=} \mathbb{I}(\mathrm{M}_0\mathrm{M}_{12}; \mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_{22}; \mathrm{Z}^n|\mathcal{C}) \\
&\overset{(b)}{\leq} 2 \cdot \mathbb{I}(\mathrm{M}_0\mathrm{M}_{12}\mathrm{M}_{22}; \mathrm{Z}^n|\mathcal{C}) \\
&\overset{(c)}{\leq} 2 \cdot \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_{\otimes}\mathrm{M}_c\mathcal{C}),
\end{aligned} \tag{3.142}$$

where (a) follows from (3.140) and (3.141), (b) follows from the chain rule of the mutual information; while (c) follows due to the independence of the messages and by letting $\mathrm{M} \triangleq (\mathrm{M}_0, \mathrm{M}_{12}, \mathrm{M}_{22})$. Eq. (3.142) implies that in order to show that our coding scheme satisfies the individual-strong secrecy constraint in (3.128c), it is enough to show that:

$$\lim_{n\to\infty} \mathbb{E}\big[\mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_{\otimes}\mathrm{M}_c\mathrm{C})\big] = 0. \tag{3.143}$$

Now, we need to highlight that the wiretap Marton code $\mathcal{C}_{0,1,2}$ used for our individual secrecy coding scheme is identical to the one used for our joint secrecy coding scheme, with the exception that in the individual secrecy coding scheme $\mathcal{C}_{0,1,2}$ is constructed conditioned on $\mathrm{V}_{\otimes}^n$ instead of U^n. This implies that using the strong secrecy analysis presented in Section 3.2.4, we can show that Eq. (3.143) holds as long as:

$$\begin{aligned}
R_t &\geq \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) + \acute{\delta}_0(\epsilon, \beta) \\
R_{r_2} + R_t &\geq \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) + \acute{\delta}_4(\epsilon, \gamma) \\
R_{r_1} &\geq \mathbb{I}(\mathrm{V}_1; \mathrm{Z}|\mathrm{V}_0) + \acute{\delta}_4(\epsilon, \gamma) \\
R_r &\geq \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{V}_{\otimes}) + \acute{\delta}_4(\epsilon, \gamma),
\end{aligned} \tag{3.144}$$

where $\gamma > 0$ and $\acute{\delta}_4(\epsilon, \gamma)$ is defined by modifying the definition of $\delta_4(\epsilon, \gamma)$ in (3.80) following the same way used to modify $\delta_0(\epsilon, \beta)$ leading to $\acute{\delta}_0(\epsilon, \beta)$.

10. Fourier-Motzkin Elimination: In order to finalize our achievability proof, we need to highlight that ϵ, β and γ can be selected in a way such that as n approaches infinity, the corresponding $\acute{\delta}$-functions approaches zero. Finally, if we take the limit as $n \to \infty$, then use the Fourier-Motzkin elimination procedure [57] on the rate constraints in (3.132), (3.136), (3.138) and (3.144) while solving for (R_c, R_0, R_1, R_2), we prove the achievability of any rate quadruple (R_c, R_0, R_1, R_2) satisfying the constraints in (3.130). This establish the achievability of the individual-strong secrecy rate region given by Proposition 3.3.

3.3.3 Receivers With Statistical Advantage

In section 3.3.1, we showed that there are two techniques to combine secret key encoding and wiretap encoding. The first technique was adopted by the coding scheme used

to prove Theorem 3.2 along with Proposition 3.3 and 3.4. In this technique, secret key encoding and wiretap encoding are brought together using superposition. Although this technique is valid for any channel, it is not the best solution to use for channels where the two legitimate receivers have a statistical advantage over the eavesdropper such as more capable channels. For these class of channels, it is better to use the second technique where secret key encoding and wiretap encoding are put together in the same layer. We now present the following result:

Theorem 3.3. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC with degraded message sets and message cognition, where the two legitimate receivers Y_1 and Y_2 are more capable than the eavesdropper Z is given by the set of all rate quadruples $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ that satisfy*

$$\begin{aligned} R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2), \mathbb{I}(\mathrm{U};\mathrm{Z})\Big] \\ R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \min\Big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\Big] \\ R_0 + R_2 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \min\Big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\Big] \end{aligned} \tag{3.145}$$

for random variables with joint probability distribution $\mathrm{Q_U}\ \mathrm{Q_{X|U}}\ Q_{\mathrm{Y_1Y_2Z|X}}$, such that $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ forms a Markov chain.

Before we present the proof of the previous theorem, we need to point out that the previous region is only achievable under the more capable assumption, or more precisely, for a probability distribution $\mathrm{Q_{UXY_1Y_2Z}}$ such that: $\mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) \geq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})$ and $\mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) \geq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})$. We now proceed with our proof by presenting the following coding scheme:

Let $n \in \mathbb{N}$ be an arbitrary but fixed code block length and $\epsilon > 0$ be an arbitrary constant. Moreover, let $\mathrm{Q_{UX}} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$ be a fixed input probability distribution and recall the probability transition matrix $Q_{\mathrm{Y_1Y_2Z|X}}$ given in (2.73) that defines the two-receiver DM-WBC. Furthermore, we consider the message sets $\mathcal{M}_c$, $\mathcal{M}$, $\mathcal{M}_\otimes$ and $\mathcal{M}_r$, defined exactly as in the coding scheme used to prove Proposition 3.3. Thus, our new coding scheme inherits the rate constraint in (3.132).

<u>1. Random Codebook $\mathcal{C}$:</u> Let $\mathrm{C}_c \triangleq \{\mathrm{U}^n(m_c)\}$ for $m_c \in \mathcal{M}_c$ be a random public common message codebook as defined in Section 3.2.2. For a fixed public message codebook $\mathcal{C}_c$ and for every $m_c \in \mathcal{M}_c$, let $\mathrm{C}_s(m_c) \triangleq \{\mathrm{X}^n(m_c, m, m_r, m_\otimes)\}$ where $m \in \mathcal{M}$, $m_r \in \mathcal{M}_r$ and $m_\otimes \in \mathcal{M}_\otimes$ be a random codebook for the confidential messages that consists of $2^{n(R+R_r+R_\otimes)}$ conditionally independent random codewords $\mathrm{X}^n(m_c, m, m_r, m_\otimes)$, where each codeword is constructed by generating the symbols $\mathrm{X}_i(m_c, m, m_r, m_\otimes)$ independently at random according to $\mathrm{Q_{X|U}}(\cdot|u_i(m_c))$. A realization of $\mathrm{C}_s(m_c)$ is denoted by $\mathcal{C}_s(m_c) \triangleq \{x^n(m_c, m, m_r, m_\otimes)\}$.

The total random codebook is denoted by $\mathrm{C} = \{\mathrm{C}_c, \mathrm{C}_s\}$, while $\mathcal{C} = \{\mathcal{C}_c, \mathcal{C}_s\}$ denotes one possible realization of this codebook selected from the total set of codebooks given by $\mathfrak{C}$. The probability of generating a certain realization $\mathcal{C} \in \mathfrak{C}$ is given by:

$$\mathrm{G}(\mathcal{C}) = \prod_{\substack{(m_c^0, m, m_r, m_\otimes) \in \\ \mathcal{M}_c \times \mathcal{M}_\otimes \times \mathcal{M} \times \mathcal{M}_r}} \mathrm{Q}^n_{\mathrm{X|U}}\Big(x^n(m_c^0, m, m_r, m_\otimes)|u^n(m_c^0)\Big) \prod_{m_c \in \mathcal{M}_c} \mathrm{Q}^n_{\mathrm{U}}\big(u^n(m_c)\big) \tag{3.146}$$

2. Encoder E: For a fixed codebook realization $\mathcal{C}$, given a message triple $(m_c, m, m_\otimes)$, the encoder first chooses a randomization message m_r uniformly at random from the set $\mathcal{M}_r$. It then finds the corresponding codeword $x^n(m_c, m, m_r m_\otimes)$ and transmits it.

3. First Legitimate Decoder φ_1: Given y_1^n and its own message $m_2 = (m_{21}, m_{22})$, the decoder searches for the unique pentadruple $(\hat{m}_c, \hat{m}_0, \hat{m}_{12}, \hat{m}_r, \hat{m}_\otimes) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{M}_r \times \mathcal{M}_\otimes$, such that:

$$\Big(u^n(\hat{m}_c), x_0^n(\hat{m}_c, \hat{m}, \hat{m}_r, \hat{m}_\otimes), y_1^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UXY}_1}),$$

where $\hat{m} = (\hat{m}_0, \hat{m}_{12}, m_{22})$. If such unique pentadruple exists, the decoder uses $\hat{m}_\otimes$ and m_{21} to produce the message $\hat{m}_{11} = \hat{m}_\otimes \otimes m_{21}$ then it outputs the triple $(\hat{m}_c, \hat{m}_0, \hat{m}_1)$, where $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12})$. Otherwise, the decoder outputs the error message (?).

4. Second Legitimate Decoder φ_2: Given y_2^n and its own message $m_1 = (m_{11}, m_{12})$, the decoder searches for the unique pentadruple $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_{22}, \tilde{m}_r, \tilde{m}_\otimes) \in \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_{22} \times \mathcal{M}_r \times \mathcal{M}_\otimes$; such that:

$$\Big(u^n(\tilde{m}_c), x^n(\tilde{m}_c, \tilde{m}, \tilde{m}_r, \tilde{m}_\otimes), y_2^n) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{UXY}_2}\Big),$$

where $\tilde{m} = (\tilde{m}_0, m_{12}, \tilde{m}_{22})$. If such unique pentadruple exists, the decoder uses $\tilde{m}_\otimes$ and m_{11} to deduce the message $\tilde{m}_{21} = \tilde{m}_\otimes \otimes m_{11}$ then it outputs the triple $(\tilde{m}_c, \tilde{m}_0, \tilde{m}_2)$ where $\tilde{m}_2 = (\tilde{m}_{21}, \tilde{m}_{22})$. If such unique pentadruple does not exist, the decoder outputs the error message (?).

5. Third Eavesdropper Decoder φ_3: Given z^n, the decoder uses the same decoding function introduced in Section 3.2.2.

6. Induced Joint Distribution: For every codebook $\mathcal{C} \in \mathfrak{C}$, the previous coding scheme induces the following joint probability distribution:

$$\begin{aligned}\mathrm{P}^{\mathcal{C}}\Big(m_c, m_\otimes, m, m_r, u^n, x^n, y_1^n, y_2^n, z^n, \hat{m}_c, \hat{m}_0, \hat{m}_1, \tilde{m}_c, \tilde{m}_0, \tilde{m}_2, \check{m}_c\Big) &= 2^{-n(R_c+R+R_r+R_\otimes)} \\ &\times \mathbb{1}_{\{u^n=u^n(m_c)\}} \mathbb{1}_{\{x^n=x^n(m_c,m,m_r,m_\otimes)\}} Q^n_{\mathrm{Y_1Y_2Z|X}}(y_1^n, y_2^n, z^n|x^n) \\ &\times \mathbb{1}_{\{(\hat{m}_c,\hat{m}_0,\hat{m}_1)=\varphi_1(y_1^n)\}} \mathbb{1}_{\{(\tilde{m}_c,\tilde{m}_0,\tilde{m}_2)=\varphi_2(y_2^n)\}} \mathbb{1}_{\{\check{m}_c=\varphi_3(z^n)\}} \qquad (3.147)\end{aligned}$$

Additionally, if we take the random codebook generation process into our consideration, we end up with: $\mathrm{P} = \mathrm{G}(\mathcal{C}) \cdot \mathrm{P}^{\mathcal{C}}$, where $\mathrm{G}(\mathcal{C})$ is given by (3.146).

7. Reliability Analysis: We start by using the average decoding error probability given by (3.135) for a fixed code realization $\mathcal{C} \in \mathfrak{C}$. Based on our encoding function, any satellite codeword x^n is jointly typical to its corresponding cloud center codeword u^n, with probability that goes to one as n approaches infinity. This implies that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\mathbb{P}(\mathcal{E})\big] = 0$, where $\mathcal{E}$ denotes an encoding error. On the other hand, following the same decoding error analysis technique used in Section 3.2.3, we can show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\mathbb{P}(\mathcal{D}|\mathcal{E}^c)\big] = 0$ as long as:

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1) - \grave{\delta}_1(\epsilon), \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \grave{\delta}_2(\epsilon), \mathbb{I}(\mathrm{U};\mathrm{Z}) - \delta_3(\epsilon)\Big]$$

$$R_0 + R_{11} + R_r + R_\otimes \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \grave{\delta}_1(\epsilon)$$
$$R_0 + R_{22} + R_r + R_\otimes \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \grave{\delta}_2(\epsilon). \tag{3.148}$$

where $\grave{\delta}_1(\epsilon) = 2\epsilon \log\big(|\mathcal{U}| \cdot |\mathcal{X}| \cdot |\mathcal{Y}_1|\big)$ and $\grave{\delta}_2(\epsilon) = 2\epsilon \log\big(|\mathcal{U}| \cdot |\mathcal{X}| \cdot |\mathcal{Y}_2|\big)$. The rate constraints in (3.148) directly implies that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\bar{\mathbf{P}}_e(\mathrm{C})\big] \leq \lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\big[\bar{\bar{\mathbf{P}}}_e(\mathrm{C})\big] = 0$. Thus, under these rate constraints, our coding scheme fulfills the reliability constraint in (3.128b).

8. Secrecy Analysis: The first part of the secrecy analysis presented for the coding scheme in Section 3.3.2 implies that in order to show that our coding scheme satisfies the individual-strong secrecy constraint in (3.128c), it is enough to show that:

$$\lim_{n\to\infty} \mathbb{E}\big[\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{M}_c\mathrm{C})\big] = 0, \tag{3.149}$$

where we modified the upper-bound in (3.142) just by conditioning on the common message M_c. Based on the first steps of the strong secrecy analysis in Section 3.2.4, Eq. (3.149) holds if:

$$\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}}\Big[\big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{M},\mathrm{C}} - \check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}\big\|\Big] = 0, \tag{3.150}$$

where $\check{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_c,\mathrm{C}}$ is given by (3.63). Moreover, based on the definition of $\mathrm{P}^{\mathcal{C}}$ in (3.147) and the steps used to bound $\mathcal{L}_3$ in Section 3.2.4, Eq. (3.150) holds as long as:

$$R_r + R_\otimes \geq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \grave{\delta}_4(\epsilon,\gamma), \tag{3.151}$$

where $\gamma > 0$ and $\grave{\delta}_4(\epsilon,\gamma) = 2\epsilon \log\big(|\mathcal{U}| \cdot |\mathcal{X}| \cdot |\mathcal{Z}|\big) + 5\gamma$. Thus, under the rate constraints in (3.132) and (3.151), our coding scheme fulfills the individual-strong secrecy constraint in (3.128c).

9. Fourier-Motzkin Elimination: In order to finalize our achievability proof, we need to highlight that ϵ and γ can be selected in a way, such that as n approaches infinity, the corresponding $\grave{\delta}$-functions approach zero. Finally, if we take the limit as $n \to \infty$, then use the Fourier-Motzkin elimination procedure on the rate constraints in (3.132), (3.148) and (3.151) while solving for (R_c, R_0, R_1, R_2), we prove the achievability of any rate quadruple (R_c, R_0, R_1, R_2) satisfying the constraints in (3.145).

3.4 Secrecy Capacity Regions and Outer Bounds

In this section, we investigate some of the outer-bounds and capacity regions that can be established for the secure communication problem given by Fig. 3.2, under both the joint and individual strong secrecy criteria. We start by establishing some general multi-letter and single-letter upper bounds. We then establish the joint and individual secrecy capacity regions for various setups of less noisy channels. Finally, we establish the joint secrecy capacity region for a class of two-receiver DM-WBC with one degraded relation and one less-noisy relation. In particular, one of the two legitimate receivers is degraded with respect to the other one and at the same time the stronger legitimate receiver is less noisy than the eavesdropper. Moreover, the relation between the weaker legitimate receiver and the eavesdropper is arbitrary. For this class of two-receiver DM-WBC, we show that the principle of indirect decoding is optimal. To the best of our knowledge, this is the first result to demonstrate the optimality of indirect decoding for a secure communication scenario.

3.4.1 Multi-Letter and Single-Letter Outer Bounds

Let (R_c, R_0, R_1, R_2) be an achievable joint-strong secrecy rate quadruple for the two-receiver DM-WBC with degraded message sets and message cognition given by Q in (2.73). Further, let $\lambda, \tau > 0$, and assume that for a sufficiently large $n \in \mathbb{N}$, there exists a $(2^{nR_c}, 2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code $\mathcal{C}$ of Definition 3.1, such that:

$$nR_j = \mathbb{H}(\mathrm{M}_j|\mathcal{C}), \qquad \text{for} \qquad j \in \{c, 0, 1, 2\} \tag{3.152a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n = 2^{-n\lambda} \tag{3.152b}$$

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) \leq \tau_n = 2^{-n\tau}, \tag{3.152c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by (3.7). In order to simplify the notation, the conditioning on the code $\mathcal{C}$ will be ignored in the upcoming steps. Based on Fano's inequality [22] and the fact that the decoding functions of the code $\mathcal{C}$ at the two legitimate receiver are not only supplied by the corresponding observations Y_1^n and Y_2^n, but also by their own transmitted messages M_2 and M_1 respectively, then it follows that:

$$\mathbb{H}(\mathrm{M}_c\mathrm{M}_0\mathrm{M}_1|\mathrm{Y}_1^n\mathrm{M}_2),\ \mathbb{H}(\mathrm{M}_c\mathrm{M}_0\mathrm{M}_2|\mathrm{Y}_2^n\mathrm{M}_1),\ \mathbb{H}(\mathrm{M}_c|\mathrm{Z}^n) \leq \Gamma(\lambda_n) \tag{3.153}$$

where $\Gamma(\lambda_n) = \mathbb{H}_2(\bar{\mathbf{P}}_e(\mathcal{C})) + n(R_c + R_0 + R_1 + R_2) \cdot \bar{\mathbf{P}}_e(\mathcal{C})$. One can easily show that $\Gamma(\lambda_n)$ approaches 0 as n approaches infinity. Using Eq. (3.152a) and Eq. (3.153), we can derive an upper-bound on the common rate R_c as follows:

$$\begin{aligned} nR_c &\leq \mathbb{H}(\mathrm{M}_c) - \mathbb{H}(\mathrm{M}_c|\mathrm{Z}^n) + \Gamma(\lambda_n) \\ &= \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n) + \Gamma(\lambda_n). \end{aligned} \tag{3.154}$$

Next, in order to establish some upper-bounds on the confidential rates (R_0, R_1, R_2), we use Lemma A.7 to modify the joint-strong secrecy constraint in (3.152c) and include conditioning on the common message M_c as follows:

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1\mathrm{M}_2; \mathrm{Z}^n|\mathrm{M}_c) \leq \Gamma(\lambda_n, \tau_n), \tag{3.155}$$

where $\Gamma(\lambda_n, \tau_n) = \Gamma(\lambda_n) + \tau_n$. Now, if we let $\mathrm{M} \triangleq (\mathrm{M}_0, \mathrm{M}_1, \mathrm{M}_2)$ and combine the relations in (3.152a), (3.153) and (3.155), we can derive an upper-bound on the confidential rates $(R_0 + R_1)$ intended to the first legitimate receiver as follows:

$$\begin{aligned} n(R_0 + R_1) &\leq \mathbb{H}(\mathrm{M}_0\mathrm{M}_1) - \mathbb{H}(\mathrm{M}_0\mathrm{M}_1|\mathrm{Y}_1^n\mathrm{M}_2\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n) \\ &\overset{(a)}{=} \mathbb{H}(\mathrm{M}_0\mathrm{M}_1|\mathrm{M}_c\mathrm{M}_2) - \mathbb{H}(\mathrm{M}_0\mathrm{M}_1|\mathrm{Y}_1^n\mathrm{M}_2\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n) \\ &= \mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Y}_1^n|\mathrm{M}_2\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n) \\ &\overset{(b)}{\leq} \mathbb{I}(\mathrm{M}; \mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n), \end{aligned} \tag{3.156}$$

where (a) follows due to the independence of the messages, while (b) follows from the chain rule of the mutual information. Similarly, we can show that following upper-bounds also hold:

$$n(R_0 + R_2) \leq \mathbb{I}(\mathrm{M}; \mathrm{Y}_2^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n) \tag{3.157a}$$

$$n(R_c + R_0 + R_1) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_1^n) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n) \tag{3.157b}$$

$$n(R_c + R_0 + R_2) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_2^n) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + 2\Gamma(\lambda_n, \tau_n). \tag{3.157c}$$

Now, observe that the term $\mathbb{I}(\mathrm{M}_c\mathrm{M};\mathrm{Y}_1^n)$ can be expanded using the mutual information chain rule as $\mathbb{I}(\mathrm{M}_c;\mathrm{Y}_1^n)+\mathbb{I}(\mathrm{M};\mathrm{Y}_1^n|\mathrm{M}_c)$. The same observation holds for the term $\mathbb{I}(\mathrm{M}_c\mathrm{M};\mathrm{Y}_2^n)$. Based on these two relations, we can subtract the bound in (3.156) from the one in (3.157b) and the bound in (3.157a) from the one in (3.157c) to obtain the following:

$$nR_c \leq \min\Big[\mathbb{I}(\mathrm{M}_c;\mathrm{Y}_1^n),\mathbb{I}(\mathrm{M}_c;\mathrm{Y}_2^n)\Big]+2\Gamma(\lambda_n,\tau_n) \tag{3.158}$$

At the moment, let $\mathrm{U}\triangleq\mathrm{M}_c$ and $\mathrm{V}\triangleq(\mathrm{M}_c,\mathrm{M})$. Hence, by merging the upper-bounds in (3.154), (3.156), (3.157a) and (3.158) after dividing the $\mathbb{RHS}$ of all of them by n, and taking the limit as n approaches infinity, keeping in mind that $\lim_{n\to\infty}\Gamma(\lambda_n),\Gamma(\lambda_n,\tau_n)=0$, the following result is established:

Theorem 3.4. *A multi-letter upper-bound for the joint-strong secrecy capacity region of the two-receiver DM-WBC with degraded messages sets and message cognition is given by the union over the set of rate quadruple* (R_c,R_0,R_1,R_2) *that satisfies the following:*

$$\begin{aligned}
R_c &\leq \lim_{n\to\infty}\frac{1}{n}\min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}^n),\mathbb{I}(\mathrm{U};\mathrm{Y}_1^n),\mathbb{I}(\mathrm{U};\mathrm{Y}_2^n)\Big]\\
R_0+R_1 &\leq \lim_{n\to\infty}\frac{1}{n}\big[\mathbb{I}(\mathrm{V};\mathrm{Y}_1^n|\mathrm{U})-\mathbb{I}(\mathrm{V};\mathrm{Z}^n|\mathrm{U})\big]\\
R_0+R_2 &\leq \lim_{n\to\infty}\frac{1}{n}\big[\mathbb{I}(\mathrm{V};\mathrm{Y}_2^n|\mathrm{U})-\mathbb{I}(\mathrm{V};\mathrm{Z}^n|\mathrm{U})\big]
\end{aligned} \tag{3.159}$$

for all possible choices of random variables (U,V)*, such that* $\mathrm{U}-\mathrm{V}-\mathrm{X}^n-(\mathrm{Y}_1^n,\mathrm{Y}_2^n,\mathrm{Z}^n)$ *forms a Markov chain.*

It is important to highlight that the convergence of the limits used to describe the rate region in the previous theorem is guaranteed by the Fekete's lemma [59]. Now, if we recall the rate region given by Proposition 3.1 and let $\mathrm{V}\triangleq\mathrm{V}_0=\mathrm{V}_1=\mathrm{V}_2$, we can argue that the multi-letter bound in (3.159) matches the rate region in (3.17) applied to the n-fold product of the two-receiver DM-WBC Q^n in (2.73). This implies that Theorem 3.4 defines a multi-letter description for the joint-strong secrecy capacity region of the two-receiver DM-WBC with degraded messages sets and message cognition.

We have already mentioned in Section 2.3.3 that a multi-letter description for the capacity region is not enough and a single-letter characterization is usually preferred. In order to derive a single-letter bound, we start by introducing the following: For $i\in[\![1,n]\!]$, let $\mathrm{U}_i\triangleq(\mathrm{M}_c,\tilde{\mathrm{Z}}^{i+1})$, $\mathrm{K}_i^1\triangleq\mathrm{Y}_1^{i-1}$, $\mathrm{K}_i^2\triangleq\mathrm{Y}_2^{i-1}$, $\mathrm{V}_i^1\triangleq(\mathrm{M},\mathrm{U}_i,\mathrm{K}_i^1)$ and $\mathrm{V}_i^2\triangleq(\mathrm{M},\mathrm{U}_i,\mathrm{K}_i^2)$, where for a given sequence A^n, the following conventions hold: $\mathrm{A}^i\triangleq(\mathrm{A}_1,\ldots,\mathrm{A}_i)$ and $\tilde{\mathrm{A}}^i\triangleq(\mathrm{A}_i,\ldots,\mathrm{A}_n)$. The bound in (3.154) can be further formulated as follows:

$$\begin{aligned}
R_c &\leq \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{M}_c;\mathrm{Z}_i|\tilde{\mathrm{Z}}^{i+1})+\Gamma_n(\lambda_n)\\
&\leq \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{U}_i;\mathrm{Z}_i)+\Gamma_n(\lambda_n).
\end{aligned} \tag{3.160}$$

where $\Gamma_n(\lambda_n)=1/n\cdot\Gamma(\lambda_n)$. Next, we consider the confidential rates (R_0+R_1) intended for the first legitimate receiver. From Eq. (3.156), it follows that

$$R_0+R_1\leq\frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1})-\mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{M}_c\tilde{\mathrm{Z}}^{i+1})\Big]+2\Gamma_n(\lambda_n,\tau_n)$$

$$\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + 2\Gamma_n(\lambda_n,\tau_n)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{V}_i^1;\mathrm{Y}_{1i}|\mathrm{U}_i\mathrm{K}_i^1) - \mathbb{I}(\mathrm{V}_i^1;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^1)\Big] + 2\Gamma_n(\lambda_n,\tau_n), \tag{3.161}$$

where $\Gamma_n(\lambda_n,\tau_n) = 1/n\cdot\Gamma(\lambda_n,\tau_n)$. Step (a) follows due to the Csiszár sum identity [10, Lemma 7]. Applying the same arguments to Eq. (3.157a), leads to a similar bound on the confidential rates (R_0+R_2) intended for the second legitimate receiver as follows:

$$R_0+R_2 \le \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{V}_i^2;\mathrm{Y}_{2i}|\mathrm{U}_i\mathrm{K}_i^2) - \mathbb{I}(\mathrm{V}_i^2;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^2)\Big] + 2\Gamma_n(\lambda_n,\tau_n). \tag{3.162}$$

On the other hand, if we consider the sum of the common and the confidential rates $(R_c+R_0+R_1)$ intended for the first legitimate receiver in (3.157b), we obtain

$$R_c+R_0+R_1 \le \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_c;\mathrm{Y}_1^n) + \mathbb{I}(\mathrm{M};\mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{M}_c)\Big] + 2\Gamma_n(\lambda_n,\tau_n)$$
$$\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_c;\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}) + \mathbb{I}(\mathrm{V}_i^1;\mathrm{Y}_{1i}|\mathrm{U}_i\mathrm{K}_i^1) - \mathbb{I}(\mathrm{V}_i^1;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^1)\Big] + 2\Gamma_n(\lambda_n,\tau_n)$$
$$\overset{(b)}{\le} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{U}_i;\mathrm{Y}_{1i}|\mathrm{K}_i^1) + \mathbb{I}(\mathrm{V}_i^1;\mathrm{Y}_{1i}|\mathrm{U}_i\mathrm{K}_i^1) - \mathbb{I}(\mathrm{V}_i^1;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^1)\Big] + 2\Gamma_n(\lambda_n,\tau_n)$$
$$\overset{(c)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{V}_i^1;\mathrm{Y}_{1i}|\mathrm{K}_i^1) - \mathbb{I}(\mathrm{V}_i^1;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^1)\Big] + 2\Gamma_n(\lambda_n,\tau_n), \tag{3.163}$$

where (a) follows as in (3.161); while (b) and (c) follow from the properties of mutual information. Similarly, the same arguments can be used to reformulate the bound on the sum of common and confidential rates $(R_c+R_0+R_2)$ intended for the second legitimate receiver in (3.157c) as follows:

$$R_c+R_0+R_2 \le \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{V}_i^2;\mathrm{Y}_{2i}|\mathrm{K}_i^2) - \mathbb{I}(\mathrm{V}_i^2;\mathrm{Z}_i|\mathrm{U}_i\mathrm{K}_i^2)\Big] + 2\Gamma_n(\lambda_n,\tau_n). \tag{3.164}$$

We now combine the rate constraints established in (3.160) - (3.164) followed by introducing a random variable T uniformly distributed over $[\![1;n]\!]$ and independent of all other random variables. We then let $\mathrm{U} = (\mathrm{U}_\mathrm{T},\mathrm{T})$, $\mathrm{K}^1 = \mathrm{K}_\mathrm{T}^1$, $\mathrm{K}^2 = \mathrm{K}_\mathrm{T}^2$, $\mathrm{V}^1 = \mathrm{V}_\mathrm{T}^1$, $\mathrm{V}^2 = \mathrm{V}_\mathrm{T}^2$, $\mathrm{Y}_1 = \mathrm{Y}_{1\mathrm{T}}$, $\mathrm{Y}_2 = \mathrm{Y}_{2\mathrm{T}}$ and $\mathrm{Z} = \mathrm{Z}_\mathrm{T}$, followed by taking the limit as $n\to\infty$ such that $\Gamma_n(\lambda_n),\Gamma_n(\lambda_n,\tau_n)\to 0$. Accordingly, we reach the following result:

Theorem 3.5. *A single-letter upper-bound for the joint-strong secrecy capacity region of the two-receiver DM-WBC with degraded messages sets and message cognition is given by the union over the set of rate quadruple* (R_c,R_0,R_1,R_2) *that satisfies:*

$$\begin{aligned}
R_c &\le \mathbb{I}(\mathrm{U};\mathrm{Z})\\
R_0+R_1 &\le \mathbb{I}(\mathrm{V}^1;\mathrm{Y}_1|\mathrm{U}\mathrm{K}^1) - \mathbb{I}(\mathrm{V}^1;\mathrm{Z}|\mathrm{U}\mathrm{K}^1)\\
R_0+R_2 &\le \mathbb{I}(\mathrm{V}^2;\mathrm{Y}_2|\mathrm{U}\mathrm{K}^2) - \mathbb{I}(\mathrm{V}^2;\mathrm{Z}|\mathrm{U}\mathrm{K}^2)\\
R_c+R_0+R_1 &\le \mathbb{I}(\mathrm{V}^1;\mathrm{Y}_1|\mathrm{K}^1) - \mathbb{I}(\mathrm{V}^1;\mathrm{Z}|\mathrm{U}\mathrm{K}^1)\\
R_c+R_0+R_2 &\le \mathbb{I}(\mathrm{V}^2;\mathrm{Y}_2|\mathrm{K}^2) - \mathbb{I}(\mathrm{V}^2;\mathrm{Z}|\mathrm{U}\mathrm{K}^2),
\end{aligned} \tag{3.165}$$

for all possible choices of random variables $(\mathrm{U},\mathrm{K}^1,\mathrm{K}^2,\mathrm{V}^1,\mathrm{V}^2)$*, such that* $(\mathrm{U},\mathrm{K}^1,\mathrm{K}^2)-(\mathrm{V}^1,\mathrm{V}^2)-\mathrm{X}-(\mathrm{Y}_1,\mathrm{Y}_2,\mathrm{Z})$ *forms a Markov chain.*

Unfortunately, the previous single-letter characterization does not match our general achievable rate region established in Theorem 3.1. However, this result is expected since the transmission capacity of the corresponding non-secrecy model [118] is still unknown as well.

We now turn to the individual secrecy measure and let (R_c, R_0, R_1, R_2) be an achievable individual-strong secrecy rate quadruple for the two-receiver DM-WBC with degraded message sets and message cognition Q. Further, let $\lambda, \tau > 0$, and assume that for a sufficiently large $n \in \mathbb{N}$, there exists a code $\mathcal{C}$ of Definition 3.1, such that:

$$nR_j = \mathbb{H}(\mathrm{M}_j|\mathcal{C}), \qquad \text{for} \qquad j \in \{c, 0, 1, 2\} \tag{3.166a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n = 2^{-n\lambda} \tag{3.166b}$$

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathcal{C}) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathcal{C}) \leq \tau_n = 2^{-n\tau}. \tag{3.166c}$$

Since the reliability constraint in (3.166b) is identical to the one in (3.152b), then the bound in (3.153) is still valid for the individual secrecy criterion. Additionally, since the rate constraints in (3.166a) are also identical to the ones in (3.152a), then the upper-bound on R_c established in (3.154) is also valid. Moreover, we can use Lemma A.7 to modify the individual secrecy constraint in (3.166c) as follows:

$$\mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathrm{M}_c) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathrm{M}_c) \leq 2\Gamma(\lambda_n, \tau_n), \tag{3.167}$$

where $\Gamma(\lambda_n, \tau_n)$ is as defined before. Next, we let $\hat{\mathrm{M}} \triangleq (\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_2)$ and $\tilde{\mathrm{M}} \triangleq (\mathrm{M}_c, \mathrm{M}_0, \mathrm{M}_1)$, then using the previous inequality along with Eqs. (3.153) and (3.166a), we can bound the confidential rates $(R_0 + R_1)$ as follows:

$$\begin{aligned} n(R_0 + R_1) &\overset{(a)}{\leq} \mathbb{I}(\mathrm{M}; \mathrm{Y}_1^n|\mathrm{M}_c) - \max\Big[\mathbb{I}(\mathrm{M}_0\mathrm{M}_1; \mathrm{Z}^n|\mathrm{M}_c), \mathbb{I}(\mathrm{M}_0\mathrm{M}_2; \mathrm{Z}^n|\mathrm{M}_c)\Big] + 3\Gamma(\lambda_n, \tau_n) \\ &\overset{(b)}{=} \mathbb{I}(\mathrm{M}; \mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + \min\Big[\mathbb{I}(\mathrm{M}_1; \mathrm{Z}^n|\hat{\mathrm{M}}), \mathbb{I}(\mathrm{M}_2; \mathrm{Z}^n|\tilde{\mathrm{M}})\Big] + 3\Gamma(\lambda_n, \tau_n) \\ &\overset{(c)}{\leq} \mathbb{I}(\mathrm{M}; \mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + \min\big[nR_1, nR_2\big] + 3\Gamma(\lambda_n, \tau_n) \end{aligned} \tag{3.168}$$

where (a) follows based on the same arguments used to establish (3.156); (b) follows due to the chain rule of the mutual information; while (c) follows from Eq. (3.166a) which implies that $nR_1 \geq \mathbb{I}(\mathrm{M}_1; \mathrm{Z}^n|\hat{\mathrm{M}})$ and $nR_2 \geq \mathbb{I}(\mathrm{M}_2; \mathrm{Z}^n|\tilde{\mathrm{M}})$. Similarly, we can show that the following bounds hold as well:

$$n(R_0 + R_2) \leq \mathbb{I}(\mathrm{M}; \mathrm{Y}_2^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + \min\big[nR_1, nR_2\big] + 3\Gamma(\lambda_n, \tau_n) \tag{3.169a}$$

$$n(R_c + R_0 + R_1) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_1^n) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + \min\big[nR_1, nR_2\big] + 3\Gamma(\lambda_n, \tau_n) \tag{3.169b}$$

$$n(R_c + R_0 + R_2) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_2^n) - \mathbb{I}(\mathrm{M}; \mathrm{Z}^n|\mathrm{M}_c) + \min\big[nR_1, nR_2\big] + 3\Gamma(\lambda_n, \tau_n). \tag{3.169c}$$

On the other hand, we need to highlight the standard upper-bounds for reliable communication that follows from the constraints in (3.153) and (3.166a) as follows:

$$n(R_0 + R_1) \leq \mathbb{I}(\mathrm{M}; \mathrm{Y}_1^n|\mathrm{M}_c) + \Gamma(\lambda_n) \tag{3.170a}$$

$$n(R_0 + R_2) \leq \mathbb{I}(\mathrm{M}; \mathrm{Y}_2^n|\mathrm{M}_c) + \Gamma(\lambda_n) \tag{3.170b}$$

$$n(R_c + R_0 + R_1) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_1^n) + \Gamma(\lambda_n) \tag{3.170c}$$
$$n(R_c + R_0 + R_2) \leq \mathbb{I}(\mathrm{M}_c\mathrm{M}; \mathrm{Y}_2^n) + \Gamma(\lambda_n). \tag{3.170d}$$

We now let $\mathrm{U} \triangleq \mathrm{M}_c$ and $\mathrm{V} \triangleq (\mathrm{M}_c, \mathrm{M})$ and combine the upper-bounds in (3.154), (3.168), (3.169) and (3.170) after dividing the $\mathbb{RHS}$ of all of them by n. We then take the limit as n approaches infinity, establishing the following result:

Theorem 3.6. *A multi-letter upper-bound for the individual-strong secrecy capacity region of the two-receiver DM-WBC with degraded messages sets and message cognition is given by the union over the set of rate quadruple* (R_c, R_0, R_1, R_2) *that satisfies:*

$$\begin{aligned}
R_c &\leq \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathrm{U}; \mathrm{Z}^n) \\
R_0 + R_1 &\leq \lim_{n\to\infty} \frac{1}{n}\Big[\mathbb{I}(\mathrm{V}; \mathrm{Y}_1^n|\mathrm{U}) - \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U}) + \min\big[nR_1, nR_2, \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U})\big]\Big] \\
R_0 + R_2 &\leq \lim_{n\to\infty} \frac{1}{n}\Big[\mathbb{I}(\mathrm{V}; \mathrm{Y}_2^n|\mathrm{U}) - \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U}) + \min\big[nR_1, nR_2, \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U})\big]\Big] \\
R_c + R_0 + R_1 &\leq \lim_{n\to\infty} \frac{1}{n}\Big[\mathbb{I}(\mathrm{V}; \mathrm{Y}_1^n) - \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U}) + \min\big[nR_1, nR_2, \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U})\big]\Big] \\
R_c + R_0 + R_2 &\leq \lim_{n\to\infty} \frac{1}{n}\Big[\mathbb{I}(\mathrm{V}; \mathrm{Y}_2^n) - \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U}) + \min\big[nR_1, nR_2, \mathbb{I}(\mathrm{V}; \mathrm{Z}^n|\mathrm{U})\big]\Big],
\end{aligned} \tag{3.171}$$

for all possible choices of random variables (U, V)*, such that* $\mathrm{U} - \mathrm{V} - \mathrm{X}^n - (\mathrm{Y}_1^n, \mathrm{Y}_2^n, \mathrm{Z}^n)$ *forms a Markov chain.*

Finally, we recall the ideas used to transform the joint secrecy multi-letter outer-bound in (3.159) to the single letter characterization in (3.165), and apply it to the individual secrecy multi-letter outer-bound in (3.171), we can establish the following result:

Theorem 3.7. *A single-letter upper-bound for the individual-strong secrecy capacity region of the two-receiver DM-WBC with degraded messages sets and message cognition is given by the union over the set of rate quadruple* (R_c, R_0, R_1, R_2) *that satisfies the following:*

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U}; \mathrm{Z}) \\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{V}^1; \mathrm{Y}_1|\mathrm{UK}^1) - \mathbb{I}(\mathrm{V}^1; \mathrm{Z}|\mathrm{UK}^1) + \min\big[R_1, R_2, \mathbb{I}(\mathrm{V}^1; \mathrm{Z}|\mathrm{U})\big] \\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{V}^2; \mathrm{Y}_2|\mathrm{UK}^2) - \mathbb{I}(\mathrm{V}^2; \mathrm{Z}|\mathrm{UK}^2) + \min\big[R_1, R_2, \mathbb{I}(\mathrm{V}^2; \mathrm{Z}|\mathrm{U})\big] \\
R_c + R_0 + R_1 &\leq \mathbb{I}(\mathrm{V}^1; \mathrm{Y}_1|\mathrm{K}^1) - \mathbb{I}(\mathrm{V}^1; \mathrm{Z}|\mathrm{UK}^1) + \min\big[R_1, R_2, \mathbb{I}(\mathrm{V}^1; \mathrm{Z}|\mathrm{U})\big] \\
R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{V}^2; \mathrm{Y}_2|\mathrm{K}^2) - \mathbb{I}(\mathrm{V}^2; \mathrm{Z}|\mathrm{UK}^2) + \min\big[R_1, R_2, \mathbb{I}(\mathrm{V}^2; \mathrm{Z}|\mathrm{U})\big],
\end{aligned} \tag{3.172}$$

for all possible choices of random variables $(\mathrm{U}, \mathrm{K}^1, \mathrm{K}^2, \mathrm{V}^1, \mathrm{V}^2)$*, such that* $(\mathrm{U}, \mathrm{K}^1, \mathrm{K}^2) - (\mathrm{V}^1, \mathrm{V}^2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

3.4.2 Capacity Regions for Less-Noisy Channels

In Section 3.4.1, we established general upper-bounds for the secrecy capacity region of our model in Fig. 3.2 under both the joint and individual secrecy criteria. We noticed that the established single-letter bounds in Theorem 3.5 and 3.7 do not match the achievable bounds provided in Theorem 3.1 and 3.2 respectively. This observation

agrees with the results presented in Section 2.4, which indicate that the transmission and the secrecy capacity of the multi-receiver DM-BC is only known for some special cases, while the general case remains an open problem.

In this section, we will limit our investigation to the class of less-noisy DM-WBC. In particular, we will consider the different less-noisy orders among the three receiving nodes –the two legitimate receivers and the eavesdropper– in our model. Due to the relation between less-noisy channels and degraded channels, our results are also valid for the corresponding degraded channel model. We start by considering the most common scenario for secure communication where the two legitimate receivers are less noisy than the eavesdropper and present the following joint secrecy result:

Theorem 3.8. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the two legitimate receivers are less noisy than the eavesdropper, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Z}$ *and* $\mathrm{Y}_2 \succeq \mathrm{Z}$. *Then, the joint-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\ R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\ R_0 + R_2 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \end{aligned} \tag{3.173}$$

for all possible choices of random variables (U,X), *such that* $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$.

The previous theorem shows that the direct extension of the Csiszár and Körner's inner-bound in [10] is optimal for this less-noisy scenario and the concept of indirect decoding introduced in [12] is not needed. Nevertheless, one can show that the achievability of the previous capacity region follows directly from the rate regions established in Theorem 3.1 as well as Propositions 3.1 and 3.2. In order to confirm this, we let $\mathrm{V}_0 = \mathrm{V}_1 = \mathrm{V}_2 = \mathrm{X}$ and observe that under this assumption the rate region in Proposition 3.1 simplifies to:

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Z}), \mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\Big] \tag{3.174a}$$

$$R_0 + R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \tag{3.174b}$$

$$R_0 + R_2 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) - \cancel{\mathbb{I}(\mathrm{X};\mathrm{XZ}|\mathrm{X})} \tag{3.174c}$$

Due to the properties of the mutual information, it follows that $\mathbb{I}(\mathrm{X};\mathrm{XZ}|\mathrm{X}) = 0$. Additionally, under the less-noisy condition $\mathrm{Y}_1 \succeq \mathrm{Z}$ and $\mathrm{Y}_2 \succeq \mathrm{Z}$, it follows that: $\mathbb{I}(\mathrm{U};\mathrm{Z}) \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_1)$ and $\mathbb{I}(\mathrm{U};\mathrm{Z}) \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2)$. This implies that the bound in (3.174a) simplifies to $R_c \leq \mathbb{I}(\mathrm{U};\mathrm{Z})$. This means that the rate region in (3.174) is equivalent to the one in (3.173).

For the converse proof, one can show that it follows directly from the upper-bound in Theorem 3.5 as follows: First, we start by using the less-noisy relation between the receiving nodes to get rid of the bounds on the sum rates $(R_c + R_0 + R_1)$ and $(R_c + R_0 + R_2)$. Next, we consider the bound on $(R_0 + R_1)$:

$$R_0 + R_1 \overset{(a)}{=} \mathbb{I}(\mathrm{V}^1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{K}^1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}^1;\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{K}^1;\mathrm{Z}|\mathrm{U})$$

$$\begin{aligned}
&\overset{(b)}{\leq} \mathbb{I}(\mathrm{V}^1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}^1;\mathrm{Z}|\mathrm{U}) \\
&\overset{(a)}{=} \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{V}^1) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{V}^1) \\
&\overset{(b)}{\leq} \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}),
\end{aligned} \tag{3.175}$$

where (a) follows because $(\mathrm{U},\mathrm{K}^1) - \mathrm{V}^1 - \mathrm{X} - (\mathrm{Y}_1,\mathrm{Z})$ forms a Markov chain; while (b) follows because $\mathrm{Y}_1 \succeq \mathrm{Z}$. Similarly, we can show that, under the assumption $\mathrm{Y}_2 \succeq \mathrm{Z}$ the bound on $(R_0 + R_2)$ in (3.165) simplifies to:

$$R_0 + R_2 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}), \tag{3.176}$$

Eq. (3.175) and Eq. (3.176) along with the rate constraint on R_c given in (3.165) describe a rate region that matches the one given by (3.173). This concludes our converse. In order to finalize our proof, we need to point out that the cardinality bound on $|\mathcal{U}|$ follows from the Fenchel-Bunt strengthening of the usual Carathéodory's theorem cf. [35, Appendix C] and [135].

Next, we investigate the communication model in Fig. 3.2 under the individual-strong secrecy criterion and the same less-noisy condition. We present the following theorem:

Theorem 3.9. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the two legitimate receivers are less noisy than the eavesdropper, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Z}$ *and* $\mathrm{Y}_2 \succeq \mathrm{Z}$. *Then, the individual-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \min\Big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\Big] \\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \min\Big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\Big]
\end{aligned} \tag{3.177}$$

for all possible choices of random variables (U,X), *such that* $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1,\mathrm{Y}_2,\mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$.

The previous theorem implies that combining the secret key encoding technique with the classical wiretap random encoding in the same layer is optimal for this less-noisy scenario. The achievability proof of the previous theorem follows directly from Theorem 3.3, where under the less noisy condition it follows that:

$$R_c \leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2), \mathbb{I}(\mathrm{U};\mathrm{Z})\Big] = \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{3.178}$$

On the other hand, the converse follows directly from the upper-bound in Theorem 3.7 as follows: First, we ignore the bounds on the sum rates $(R_c + R_0 + R_1)$ and $(R_c + R_0 + R_2)$ in (3.172). Next, we use the implications of the less noisy assumption to reformulate the bounds on $(R_0 + R_1)$ and $(R_0 + R_2)$, as we did in (3.175), while keeping in mind that the data processing inequality implies that: $\mathbb{I}(\mathrm{V}^1;\mathrm{Z}|\mathrm{U}), \mathbb{I}(\mathrm{V}^2;\mathrm{Z}|\mathrm{U}) \leq \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})$. Finally, the cardinality bound on $|\mathcal{U}|$ follows based on the same arguments used in the proof of Theorem 3.8.

We now turn to another class of less-noisy channels where the eavesdropper is less noisy than the two legitimate receiver. This class of less-noisy channels is rarely addressed in secure communication scenario due to the result established in [16, Proposition 3.4], which states that if the legitimate channel is noisier than the eavesdropper channel then the secrecy capacity of the wiretap channel vanishes. Nevertheless, this class of less-noisy channels has very interesting features specially under the individual secrecy criterion. However, for the sake of completeness, we investigate this class of less-noisy channels under the joints secrecy criterion as well.

Theorem 3.10. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the two legitimate receivers are noisier than the eavesdropper, i.e.* $\mathrm{Z} \succeq \mathrm{Y}_1$ *and* $\mathrm{Z} \succeq \mathrm{Y}_2$*. Then, the joint-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c &\leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\right] \\ R_0 &= R_1 = R_2 = 0. \end{aligned} \tag{3.179}$$

for all possible choices of the random variable X.

Although the result presented in [16, Proposition 3.4] was established for wiretap channels with a single legitimate receiver, the previous theorem implies that the result is still valid for wiretap channels with two or more legitimate receivers. The achievability proof of the previous theorem follows directly from the channel coding theorem [136]. Moreover, the converse proof is based on the following arguments: The bound on R_c follows from the transmission capacity broadcasting a public common message over the three-receiver DM-BC [72], which implies that:

$$R_c \leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2), \mathbb{I}(\mathrm{X};\mathrm{Z})\right] \overset{(a)}{=} \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\right], \tag{3.180}$$

where (a) follows because Z is less noisy than both Y_1 and Y_2. On the other hand, the disintegration of the confidential rates (R_0, R_1) is a direct implication of Lemma A.8 and Eq. the (3.156) as follows:

$$R_0 + R_1 \leq \frac{1}{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{M}_c)\Big] + 2\Gamma_n(\lambda_n, \tau_n) \leq 2\Gamma_n(\lambda_n, \tau_n), \tag{3.181}$$

where $\lim_{n\to\infty} \Gamma_n(\lambda_n, \tau_n) = 0$. Similarly, one can use Eq. (3.157a) and Lemma A.8 to show that $R_0 + R_2 = 0$. Combining this bound with the ones in (3.180) and (3.181) completes our converse.

Unlike the joint secrecy criterion, secure communication under the individual secrecy criterion is possible for the class of channels, where the eavesdropper is less noisy than the legitimate receivers. This result is illustrated by the following theorem:

Theorem 3.11. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the two legitimate receivers are noisier than the eavesdropper, i.e.* $\mathrm{Z} \succeq \mathrm{Y}_1$ *and* $\mathrm{Z} \succeq \mathrm{Y}_2$*. Then, the individual-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c + R_1 = R_c + R_2 &\leq \min\left[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\right] \\ R_0 &= 0. \end{aligned} \tag{3.182}$$

for all possible choices of the random variable X.

The previous theorem implies that even under the individual secrecy measure, only the confidential messages that can be protected using secret key encoding (i.e. M_1 and M_2) can be transmitted over a DM-WBC with a less-noisy eavesdropper, while the rate of the confidential message protected only by wiretap random encoding (i.e. R_0) still vanishes. The achievability of Theorem 3.11 can be derived from the achievable rate region established by Proposition 3.3 as follows: We start by letting $\mathrm{V}_\otimes = \mathrm{V}_0 = \mathrm{V}_1 = \mathrm{V}_2 = \mathrm{X}$, while keeping in mind that Z is less noisy than both Y_1 and Y_2. Under these assumptions, the rate region in Eq. (3.130) simplifies to:

$$\begin{aligned} R_c &\leq \min\big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\big] \\ R_0 + R_1 &\leq \min\big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}), \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U})\big] \\ R_0 + R_2 &\leq \min\big[R_1, R_2, \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}), \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U})\big] \end{aligned} \tag{3.183}$$

The previous relations indicate that the two individual confidential rates R_1 and R_2 must be equal. They also imply that the common confidential rate R_0 vanishes. Taking this into consideration, we can argue that the following rate region is also achievable:

$$\begin{aligned} R_c &\leq \min\big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2)\big] \\ R_1 = R_2 &\leq \min\big[\mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}), \mathbb{I}(\mathrm{X};\mathrm{Y}_2|\mathrm{U})\big] \\ R_c + R_1 = R_c + R_2 &\leq \min\big[\mathbb{I}(\mathrm{X};\mathrm{Y}_1), \mathbb{I}(\mathrm{X};\mathrm{Y}_2)\big], \end{aligned} \tag{3.184}$$

where the last bound is a sum rate constraint that follows by combining the condition of the first and second bounds. Finally, if we substitute $\mathrm{U} = \mathrm{X}$ in (3.184), we establish the achievability of the rate region in (3.182).

For the converse, we start by referring to the standard outer-bound for reliable communication established by the channel coding theorem, which implies that:

$$\begin{aligned} R_c + R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) \\ R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_2). \end{aligned} \tag{3.185}$$

With this in mind, we only need to prove that $R_0 = 0$ and $R_1 = R_2$. In order to do so, we recall Lemma A.8 and apply it to the individual secrecy upper-bound on $R_0 + R_1$ in (3.168). We have

$$\begin{aligned} R_0 + R_1 &\leq \frac{1}{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_1^n|\mathrm{M}_c) - \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{M}_c)\Big] + \min\big[R_1, R_2\big] + 3\Gamma_n(\lambda_n, \tau_n) \\ &\leq \min\big[R_1, R_2\big] + 3\Gamma_n(\lambda_n, \tau_n), \end{aligned} \tag{3.186}$$

where $\lim_{n\to\infty}\Gamma_n(\lambda_n, \tau_n) = 0$. The previous bound directly implies that $R_0 = 0$ and that $R_1 \leq R_2$. Similarly, we can apply Lemma A.8 to the individual secrecy upper-bound on $R_0 + R_2$ in (3.169a) to show that $R_2 \leq R_1$. This directly implies that R_1 and R_2 are equivalent. Finally, if we combine this result with the standard bounds in (3.185), our converse is complete.

We now consider our last class of less-noisy channels which is an intermediate scenario between the previous two extreme cases. In this class, the eavesdropper is less noisy than one of the legitimate receiver but noisier than the other one. Without loss of generality, we assume that the following relation holds: $\mathrm{Y}_1 \succeq \mathrm{Z} \succeq \mathrm{Y}_2$. We present the following result:

Theorem 3.12. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the eavesdropper is less noisy than one of the legitimate receiver but noisier than the other one, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Z} \succeq \mathrm{Y}_2$. *Then, the joint-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) \\ R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\ R_0 &= R_2 = 0, \end{aligned} \tag{3.187}$$

for all possible choices of random variables (U,X), *such that* $\mathrm{U} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 2$.

The achievability proof of the previous theorem follows directly from Proposition 3.1. Similar to the proof of Theorem 3.8, we start by letting $\mathrm{V}_0 = \mathrm{V}_1 = \mathrm{V}_2 = \mathrm{X}$ and observe that under the less-noisy assumption of $\mathrm{Y}_1 \succeq \mathrm{Z} \succeq \mathrm{Y}_2$, the rate region in (3.174) simplifies to the one in (3.187) as follows:

$$\begin{aligned} R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) \\ R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\ R_0 + R_2 &\leq 0 \end{aligned} \tag{3.188}$$

On the other hand, the converse proof follows from the upper-bound established in Theorem 3.5 along with the properties of less-noisy channels. We start by pointing out that the vanishment of R_0 and R_2 follows from the same arguments used in the converse proof of Theorem 3.10. Thus, the last bound in the outer rate region in Theorem 3.5 simplifies to:

$$R_c \leq \mathbb{I}(\mathrm{UK}^2;\mathrm{Y}_2) \overset{(a)}{\leq} \mathbb{I}(\mathrm{UK}^1;\mathrm{Y}_2), \tag{3.189}$$

where (a) follows from the definition of the random variables K^1 and K^2 along with the les- noisy lemma established in [137, Lemma 1]. Additionally, the outer-bound on R_1 given by Theorem 3.5 can be reformulated as follows:

$$\begin{aligned} R_1 &\overset{(a)}{=} \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{UK}^1) - \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{V}^1) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{UK}^1) + \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{V}^1) \\ &\overset{(b)}{\leq} \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{UK}^1) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{UK}^1), \end{aligned} \tag{3.190}$$

where (a) follows because $(\mathrm{U},\mathrm{K}^1) - \mathrm{V}^1 - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Z})$ forms a Markov chain; while (b) follows because $\mathrm{Y}_1 \succeq \mathrm{Z}$. Finally, if we combine the bounds in (3.189) and (3.190) then let $\mathrm{U} \triangleq \mathrm{UK}^1$, our converse is complete.

Unfortunately, the individual secrecy capacity region for this class of less-noisy channels is still unknown. We could not even establish the capacity for the corresponding degrading order, i.e. $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Z} - \mathrm{Y}_2$ forms a Markov chain. The reason for this can be summarized in the following points: Since $\mathrm{Z} \succeq \mathrm{Y}_2$, we know that only secret key encoding can be used to establish the confidential rate R_2. On the other hand, we have $\mathrm{Y}_1 \succeq \mathrm{Z}$ which implies that the optimal coding scheme for R_1 is to combine secret key encoding and wiretap random encoding in the same layer. The challenge arises

because we could not find a coding scheme that can achieve these two requirements at the same time. Nevertheless, if we consider a simple version of our model, where the individual confidential messages vanishes, i.e. $\mathcal{M}_1 = \mathcal{M}_2 = \emptyset$, we can establish the following result.

Theorem 3.13. *Consider a two-receiver DM-WBC with a common public message and a common confidential message, such that* $(\mathrm{Y}_1 \succeq \mathrm{Y}_2)$*. Then, both the joint and individual strong-secrecy capacity regions are given by the union over the set of all rates* $(R_c, R_0) \in \mathbb{R}_+^2$ *that satisfy*

$$\begin{aligned} R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_2), \mathbb{I}(\mathrm{U};\mathrm{Z})\Big] \\ R_0 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \end{aligned} \tag{3.191}$$

for all possible choices of random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X})$*, such that* $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$ *and* $|\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3$*.*

The achievability follows from the straightforward extension of the Csiszár-Körner result in [10] using the strong secrecy results established in [29]. It also follows directly from Proposition 3.1 by letting $\mathrm{V}_0 = \mathrm{V}_1 = \mathrm{V}_2 = \mathrm{V}$ leading to the following lower bounds:

$$\begin{aligned} R_c &\leq \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_1), \mathbb{I}(\mathrm{U};\mathrm{Y}_2), \mathbb{I}(\mathrm{U};\mathrm{Z})\Big] \\ R_0 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \\ R_0 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}). \end{aligned} \tag{3.192}$$

Since $\mathrm{Y}_1 \succeq \mathrm{Y}_2$, it implies that $\mathbb{I}(\mathrm{U};\mathrm{Y}_2) \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_1)$ and $\mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) \leq \mathbb{I}(\mathrm{V};\mathrm{Y}_1|\mathrm{U})$. Substituting these two relations in (3.192) leads to the achievability of the region in (3.191). On the other hand, the converse follows directly using the standard techniques in [10], where we only consider the weak legitimate receiver Y_2.

3.4.3 Optimality of Indirect Decoding

In [12], Chia and El Gamal showed by an example that indirect decoding leads to a larger achievable joint secrecy rate region for the two-receiver DM-WBC, as compared to the direct extension of Csiszár and Körner's inner-bound. However, one can notice that for all the cases where the joint secrecy capacity region is known, Csiszár and Körner's inner-bound is –in fact– optimal. This observation encouraged us to investigate whether there exists a specific class of channels, and not just an example where indirect decoding outperforms the classical Csiszár and Körner's inner-bound. In this section, we show that this class of channels does exist. In particular, we show that for this class, indirect decoding is necessary to establish the joint secrecy capacity region of the two-receiver DM-WBC with degraded message sets and message cognition. To the best of our knowledge, this is the first result to demonstrate the optimality of indirect decoding for secure communication in this way.

During our search for this class of channels, we were trying to identify the main premise of indirect decoding. In doing so, we started to raise some questions about the result in Corollary 2.2. Among these questions was the following: Why should it be optimal

for the first legitimate receiver to use the random variable X to decode the common message M, while the second receiver has to use the random variable V instead? In other words, despite the two receivers being interested in the same information, we need to prove that it is optimal for them to use different decoding rules. The only convenient answer for us was that the first legitimate receiver must have some sort of an advantage over the second one. With some investigation, we managed to find two possible scenarios that describe such advantage. The first scenario is highlighted by the following result:

Theorem 3.14. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the second legitimate receiver* Y_2 *is degraded form the first legitimate receiver* Y_1*, i.e.* $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2$ *forms a Markov chain. In addition, the first legitimate receiver* Y_1 *is less noisy than the eavesdropper* Z*, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Z}$*. Then, the joint-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned} R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\ R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\ R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \\ R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) \end{aligned} \tag{3.193}$$

for all possible choices of random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X})$*, such that* $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$ *and* $|\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3$*.*

Before we proceed with the proof of this theorem, we need to highlight some important points. First, one can notice that the previous theorem describes a class of two-receiver DM-WBCs, where the first legitimate receiver is much stronger than the second one. This strength arises from the fact that the second legitimate receiver observation is a degraded version of the first one. In addition, the first legitimate receiver has a statistical advantage over the eavesdropper, while the second one does not, i.e. Y_1 is less noisy than Z, while the relation between Y_2 and Z is arbitrary. Our analysis shows that the two previous conditions are necessary to establish the converse proof.

The second point that we need to highlight is how the rate region in (3.193) is –in fact– an example of indirect decoding and not a simple superposition region. A simple trick to visualize this point is to let $R_c = R_1 = R_2 = 0$ and $\mathrm{U} = \emptyset$. Under these assumptions, the rate region in Theorem 3.14 simplifies to the one given by Corollary 2.2, which is an example of indirect decoding. This point will become clearer as we introduce the achievability proof of the theorem and notice that the messages intended to the first legitimate receiver $(\mathrm{M}_0, \mathrm{M}_1)$ can be decoded from the random variable V. Nevertheless, the first legitimate receiver decodes them indirectly using the random variable X.

We are now ready to proceed with our proof. First, we argue that the achievability is a direct consequence of Theorem 3.1 and can even be established using Proposition 3.1 as well. However, instead of directly using the rate region given by Proposition 3.1, we derive an equivalent rate region using the reliability and secrecy constraints presented in (3.23) as follows:

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}) \\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) \\
R_c + R_0 + R_1 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{U}) \\
R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0)
\end{aligned} \tag{3.194}$$

The equivalency of the rate regions in (3.17) and (3.194) follows from the fact that both regions were derived using the same rate constraints given in (3.23). This equivalency was established in [35] for the general three-receiver DM-BC. Next, we let $\mathrm{V}_2 = \emptyset$, $\mathrm{V}_0 = \mathrm{V}$ and $(\mathrm{V}_0, \mathrm{V}_1) = \mathrm{X}$ and observe that under these assumptions the rate region in Eq. (3.194) simplifies to:

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z}) \\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) - \cancel{\mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0)} \\
R_c + R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) \\
R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) - \cancel{\mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0)}.
\end{aligned} \tag{3.195}$$

One can show that the bound on the sum rate $(R_c + R_0 + R_1)$ is redundant due to the less-noisy relation between Y_1 and Z. This implies that the rate region in (3.195) is equivalent to the one in (3.193). This completes our achievability proof.

For the converse, we start by letting $\mathrm{U}_i \triangleq (\mathrm{M}_c, \mathrm{Y}_2^{i-1}, \tilde{\mathrm{Z}}^{i+1})$, $\mathrm{M} \triangleq (\mathrm{M}_0, \mathrm{M}_1, \mathrm{M}_2)$ and $\mathrm{V}_i \triangleq (\mathrm{M}, \mathrm{U}_i)$. Applying the same procedure used to bound the common rate R_c in Eq. (3.160) leads to:

$$\begin{aligned}
R_c &\leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\mathrm{M}_c\tilde{\mathrm{Z}}^{i+1};\mathrm{Z}_i) + \Gamma_n(\lambda_n) \\
&\leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\mathrm{U}_i;\mathrm{Z}_i) + \Gamma_n(\lambda_n).
\end{aligned} \tag{3.196}$$

Next, we notice that the arguments used to establish the bound in (3.162) are still valid for our scenario. This implies that the confidential rates $(R_0 + R_2)$ intended for the second legitimate receiver are bounded as follows:

$$R_0 + R_2 \leq \frac{1}{n}\sum_{i=1}^{n} \Big[\mathbb{I}(\mathrm{V}_i;\mathrm{Y}_{2i}|\mathrm{U}_i) - \mathbb{I}(\mathrm{V}_i;\mathrm{Z}_i|\mathrm{U}_i)\Big] + \Gamma_n(\lambda_n, \tau_n). \tag{3.197}$$

On the other hand, if we consider the sum of the common and confidential rates $(R_c + R_0 + R_2)$ intended for the second legitimate receiver and use the bound in (3.164), we have

$$\begin{aligned}
R_c + R_0 + R_2 &\leq \sum_{i=1}^{n} \Big[\mathbb{I}(\mathrm{V}_i;\mathrm{Y}_{2i}|\mathrm{Y}_2^{i-1}) - \mathbb{I}(\mathrm{V}_i;\mathrm{Z}_i|\mathrm{U}_i)\Big] + 2\Gamma_n(\lambda_n, \tau_n) \\
&\leq \sum_{i=1}^{n} \Big[\mathbb{I}(\mathrm{V}_i;\mathrm{Y}_{2i}) - \mathbb{I}(\mathrm{V}_i;\mathrm{Z}_i|\mathrm{U}_i)\Big] + 2\Gamma_n(\lambda_n, \tau_n),
\end{aligned} \tag{3.198}$$

where the last inequality follows from the chain rule of mutual information and the definition of V_i. Finally, we consider the confidential rates $(R_0 + R_1)$ intended for the first legitimate receiver. Based on the same arguments that we used to derive the bound in (3.161), we can show that

$$\begin{aligned}
R_0 + R_1 &\leq \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\mathrm{Y}_2^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{M}_c\mathrm{Y}_1^{i-1}\mathrm{Y}_2^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + 2\Gamma_n(\lambda_n,\tau_n)\\
&= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{U}_i\mathrm{Y}_1^{i-1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{U}_i\mathrm{Y}_1^{i-1})\Big] + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(b)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{X}_i;\mathrm{Y}_{1i}|\mathrm{U}_i) - \mathbb{I}(\mathrm{X}_i;\mathrm{Z}_i|\mathrm{U}_i)\Big] + 2\Gamma_n(\lambda_n,\tau_n),
\end{aligned} \tag{3.199}$$

where (a) follows because Y_2 is degraded from Y_1; while (b) follows because Y_1 is less noisy than Z. One can notice that the two properties that define this class of two-receiver DM-WBC are only needed to establish the upper-bound in (3.199). In step (a), we needed to introduce the random variable Y_2^{i-1} into the two mutual information terms; that is where the degradedness relation comes into play. On the other hand, in step (b) we needed to withdraw the term Y_1^{i-1} from the condition of the two mutual information terms, and that is where we needed the less-noisy relation.

We now combine the bounds in (3.196) - (3.199), and introduce a random variable T independent of all others and uniformly distributed over $[\![1;n]\!]$, We then let $\mathrm{U} = (\mathrm{U}_\mathrm{T},\mathrm{T})$, $\mathrm{V} = \mathrm{V}_\mathrm{T}$, $\mathrm{X} = \mathrm{X}_\mathrm{T}$, $\mathrm{Y}_1 = \mathrm{Y}_{1\mathrm{T}}$, $\mathrm{Y}_2 = \mathrm{Y}_{2\mathrm{T}}$ and $\mathrm{Z} = \mathrm{Z}_\mathrm{T}$, followed by taking the limit as $n \to \infty$, such that $\Gamma_n(\lambda_n)$ and $\Gamma_n(\lambda_n,\tau_n) \to 0$, while keeping in mind that $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}_1,\mathrm{Y}_2,\mathrm{Z})$ forms a Markov chain. By doing so, we reach a region that matches the one in (3.193) and this concludes our converse. What remains is to highlight that the cardinality bounds on $|\mathcal{U}|$ and $|\mathcal{V}|$ follow from the Fenchel-Bunt strengthening of the usual Carathéodory's theorem [35, Appendix C].

We have already mentioned in the beginning of this section that there exists two scenarios where indirect decoding outperforms the classical decoding for our secure communication model. In Theorem 3.14, we have presented our first scenario, while the second scenario is introduced in the upcoming theorem:

Theorem 3.15. *Consider a two-receiver DM-WBC with degraded message sets and message cognition, where the second legitimate receiver* Y_2 *is noisier than the first legitimate receiver* Y_1*, i.e.* $\mathrm{Y}_1 \succeq \mathrm{Y}_2$*. In addition, the eavesdropper is degraded from the first legitimate receiver,* $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Z}$ *forms a Markov chain. Then, the joint-strong secrecy capacity region is given by the union over the set of all rate quadruples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$\begin{aligned}
R_c &\leq \mathbb{I}(\mathrm{U};\mathrm{Z})\\
R_0 + R_1 &\leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U})\\
R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})\\
R_c + R_0 + R_2 &\leq \mathbb{I}(\mathrm{V};\mathrm{Y}_2) - \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U})
\end{aligned} \tag{3.200}$$

for all possible choices of random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X})$*, such that* $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$ *and* $|\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3$.

Although the previous theorem describes a class of two-receiver DM-WBCs different from the one described in Theorem 3.14, both classes have the same capacity region. In fact, one can notice that the two classes have something in common: The first legitimate receiver has a statistical advantage over both the second legitimate receiver and the eavesdropper, while the relation between the second legitimate receiver and the eavesdropper is arbitrary. The superiority of the first legitimate receiver is established based on two conditions: a degraded relation and a less noisy relation. The results established in Theorem 3.14 and 3.15 implies that the exact assignment of these two conditions does not matter.

In order to prove Theorem 3.15, we start by pointing out that the achievability of the rate region in (3.193) follows directly from the rate region in (3.194) as long as Y_1 is less noisy than Z, as shown in the proof of Theorem 3.14. Since in the scenario described by Theorem 3.15, we have a much stronger relation as Z is a degraded version from Y_1, then the achievability proof of Theorem 3.15 follows accordingly.

For the converse proof, we have already pointed out that the upper bounds established in (3.196), (3.197) and (3.198) are valid for any general two-receiver DM-WBC and do not require any assumptions on Y_2 or Z. Thus, it remains to show that the bound on $R_0 + R_1$ in (3.197) is also valid under the conditions mentioned in Theorem 3.15. We proceed as follows:

$$\begin{aligned}
R_0 + R_1 &\leq \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{H}(\mathrm{M}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}\mathrm{Z}_i) - \mathbb{H}(\mathrm{M}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}\mathrm{Y}_{1i}\mathrm{Z}_i)\Big] + 2\Gamma_n(\lambda_n,\tau_n)\\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{M};\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}\mathrm{Z}_i) + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(b)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{X}_i;\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_1^{i-1}\tilde{\mathrm{Z}}^{i+1}\mathrm{Z}_i) + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(c)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathrm{X}_i;\mathrm{Y}_{1i}|\mathrm{M}_c\mathrm{Y}_2^{i-1}\tilde{\mathrm{Z}}^{i+1}\mathrm{Z}_i) + 2\Gamma_n(\lambda_n,\tau_n)\\
&\overset{(d)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{X}_i;\mathrm{Y}_{1i}|\mathrm{U}_i) - \mathbb{I}(\mathrm{X}_i;\mathrm{Z}_i|\mathrm{U}_i)\Big] + 2\Gamma_n(\lambda_n,\tau_n),
\end{aligned} \tag{3.201}$$

where (a) follows because Z is degraded from Y_1; (b) follows because $\mathrm{M} - \mathrm{X}_i - \mathrm{Y}_i$ forms a Markov chain; (c) follows from the less-noisy lemma established in [137, Lemma 1] and the fact that $\mathrm{Y}_1 \succeq \mathrm{Z}$; while (d) follows by reversing the first steps while keeping in mind that Z is degraded from Y_1. With this bound, our converse proof is complete. Again we can see that the degraded and the less-noisy condition play a role in deriving only the upper-bound on $R_0 + R_1$.

3.5 Discussion

The results established in this chapter allowed us to achieve a better understanding for the similarities and differences between the joint and individual secrecy criteria. This helped us in identifying the advantages and disadvantages of each secrecy criterion. Moreover, our investigation has lead to a better visualization of the communication model described in Fig. 3.2. Yet, it has raised some further interesting questions. Thus, we think it is important to highlight the following points:

1. Any coding scheme that satisfies the joint secrecy criterion will also satisfy the individual one as well. This implies that any achievable joint secrecy rate region is also an achievable rate region under the individual secrecy criterion. This result is due to the fact that the individual secrecy is a less conservative secrecy measure, compared to the joint one.

2. The joint secrecy criterion is a very conservative secrecy measure. Thus, even if one of the confidential messages is revealed to the eavesdropper in a genie-aided way, one can still show that the other messages remain protected. In order to examine this claim, we investigate the following:

$$\begin{aligned}\mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n\mathrm{M}_2) &= \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{M}_2) + \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2)\\ &\overset{(a)}{=} \mathbb{I}(\mathrm{M}_0\mathrm{M}_1;\mathrm{Z}^n|\mathrm{M}_2) \overset{(b)}{=} \tau_{1n},\end{aligned}$$

where (a) follows because the messages are independent; while (b) follows from the joint secrecy constraint in (3.8). The previous equation shows that the information leakage of M_0 and M_1 to the eavesdropper, when M_2 is known is still small.

3. The previous result does not hold under the individual secrecy criterion, which is based on the mutual trust between the legitimate receivers. This means that if one of the confidential messages is compromised, this might also affects the secrecy of the other one. In order to understand this property, let us consider a special case of our communication model where $\mathcal{M}_0 = \emptyset$ and assume that M_2 is revealed to the eavesdropper. Thus, we have:

$$\mathbb{I}(\mathrm{M}_1;\mathrm{Z}^n\mathrm{M}_2) = \mathbb{H}(\mathrm{M}_1) - \mathbb{H}(\mathrm{M}_1|\mathrm{Z}^n\mathrm{M}_2).$$

Now, if the eavesdropper Z is less noisy than the first legitimate receiver –M_1 is only protected using secret key encoding–, the term $\mathbb{H}(\mathrm{M}_1|\mathrm{Z}^n\mathrm{M}_2)$ will vanish. This is because under the less noisy assumption, the eavesdropper can correctly decode any information which the first legitimate receiver can decode. This implies that M_1 is fully leaked to the eavesdropper, when M_2 is revealed to it. On the other hand, for a general two-receiver DM-WBC the term $\mathbb{H}(\mathrm{M}_1|\mathrm{Z}^n\mathrm{M}_2)$ does not vanish, yet it is smaller than $\mathbb{H}(\mathrm{M}_1)$. This means that a part of M_1 might be leaked to the eavesdropper upon revealing M_2. The size of this part depends on how much the eavesdropper can infer using its received signal Z^n and M_2.

4. Although the previous two points imply that joint secrecy might be more preferable than individual secrecy, the capacity regions established in Theorem 3.8 and 3.9 promote individual secrecy over the joint one. This is because those theorems show

that the individual secrecy criterion provides a larger capacity region compared to the joint one. In fact, for some cases, the individual secrecy criterion can allow for a non vanishing achievable secrecy rate, despite a vanishing joint secrecy capacity region, cf. Theorem 3.10 and 3.11. The difference between the joint and individual capacity regions arises from the fact that individual secrecy allows the usage of secret key encoding beside the standard random wiretap encoding. That is why, the increase in the individual secrecy capacity region is somehow related to the size of the individual messages, which is used for secret key encoding.

5. The previous comparison between the joint and the individual secrecy criteria implies that: The process of choosing among the two secrecy criteria is a simple trade-off between a powerful secrecy measure and a larger capacity region. In particular, the choice is based on whether the legitimate receivers trust one another or not.

6. In Section 3.3.1, we pointed out that the individual secrecy coding scheme that works for all classes of the two-receiver DM-WBC is not optimal for the channels where the legitimate receivers have a statistical advantage over the eavesdropper. This observation showed up while arguing that the achievability proof of Theorem 3.9 can not be derived from the general achievable rate region given in Theorem 3.2 and we need to use Theorem 3.3 instead. This result raises an important question about whether there exists a universal individual secrecy coding scheme, which is optimal in the sense that it establishes the capacity region for all the known cases and is independent of the channel characteristics or not.

7. Another issue that was raised related to individual secrecy is that whether individual secrecy is a general useful secrecy notation that can be investigated for different classes of channels, or it is a unique feature for the DM-WBC with message cognition. This is a very important question. In fact, it is the main motivation behind the work presented in the next chapter.

8. In the beginning of this chapter, we claimed that although the problem of common message broadcasting and individual message broadcasting with message cognition at the receivers are equivalent, their corresponding secrecy scenarios are not. One of the reasons that support our claim is that there exists different classes of the two-receiver DM-WBC, such that the secrecy capacity of the common confidential message broadcasting problem is known, while the secure individual message broadcasting with message cognition is still an unsolved issue. Among these channels are the relative less-noisy channel, i.e. $Y_1 \succeq Y_2$, cf. Theorem 3.13 and the full more capable channel, i.e. Y_1 m.c Y_2 m.c Z.

9. One of the most interesting results that we have established in this chapter is the optimality of indirect decoding for a class of two-receiver DM-WBC with a strong legitimate receiver as shown in Theorems 3.14 and 3.15. In those theorems, we argued that a strong receiver requires a degraded relation and a less noisy relation. Nevertheless, we believe that indirect decoding might be optimal for strong receivers with more relaxed constraints, for example two less noisy relations. However, despite our hard efforts, we could not prove this conjecture.

10. The last point that we would like to discuss is related to our joint-strong secrecy Marton wiretap coding scheme. We already pointed out that the main difference between our coding scheme and the weak secrecy coding scheme introduced in [12] is the encoding function. Nevertheless, both coding schemes resulted in the same constraints on the randomization rates. Thus, one might wonder, whether it is really necessary to use the maximum likelihood encoder in (3.20) to assure that our coding scheme fulfills the strong secrecy measure or the classical Marton-coding scheme suggested by Chia and El Gamal in [12] can be used instead.

Chapter 4

Multi-Receiver Wiretap Broadcast Channels

This chapter discusses the problem of secure communication over a multi-receiver DM-WBC under two different strong secrecy criteria: joint secrecy and individual secrecy. First, we introduce our channel model and demonstrate the effect of varying the secrecy criterion on the achievable rate region for the two-receiver case. Next, we limit our investigation to the degraded multi-receiver DM-WBC and establish the joint and individual strong secrecy capacity regions. The established capacity regions are then used to derive the corresponding capacity regions for the Gaussian SISO and degraded Gaussian MIMO multi-receiver WBC. Furthermore, we present a general achievable rate region for the two-receiver DM-WBC under both the joint and the individual strong secrecy criteria.

4.1 Secrecy in Multi-Receiver WBC

In this section, we introduce the problem of secure communication over the multi-receiver DM-WBC and examine how it differs, depending on the investigated secrecy criterion. This problem was only considered in previous literature under the joint secrecy criterion. That is why, we start this section by showing how individual secrecy is an interesting secrecy criterion that can be applied to any multi-receiver channel and is not a unique feature for the two-receiver DM-WBC with message cognition. We then introduce a communication model that involves the transmission of individual confidential messages over the multi-receiver DM-WBC and investigate the corresponding joint and individual strong secrecy requirements. Finally, we discuss some of the results established for secure communication over the two-receiver DM-WBC. These results illustrate that individual secrecy can provide a larger secrecy capacity region compared to the joint one.

4.1.1 Individual Secrecy Without Message Cognition

In the previous chapter, we showed that relaxing the secrecy constraint from the joint secrecy criterion to the individual one can lead to a larger capacity region for two-receiver DM-WBC with message cognition. The increase in the capacity region arises from the fact that under the individual secrecy criterion, the confidential message of

one receiver –which is available as side information at the other receiver– can be interpreted as a secret key shared between the transmitter and this receiver. Thus, instead of only using wiretap random coding to protect the confidential message as in the joint secrecy case, the individual secrecy criterion allows a combination of wiretap random encoding and secret key encoding. In this combination the secret key encoded message is used as a part of the randomization message required for the wiretap random coding, which consequently leads to a larger capacity region.

The previous discussion advocated that individual secrecy is a unique feature for DM-WBC with receiver side information (message cognition) and is not a general secrecy measure. However, this argument contradicts the fact that the notation of individual secrecy was first introduced in [110] for the DM-MAWC. It goes without saying that the receiving node in the DM-MAWC does not have any additional side information. Nevertheless, one might argue that the individual secrecy notation used for DM-MAWC is different from the one used for the DM-WBC. Although their argument is slightly correct, it ignored the fact that both notation are based on the same concept which is the mutual trust. Of course, the nature of the trusting nodes differs from the DM-MAWC to the DM-WBC: In particular, for the DM-MAWC the mutual trust is among the transmitting nodes, while for the DM-WBC it is among the receiving nodes. However, we are convinced that individual secrecy is in fact a universal secrecy criterion that can be used to investigated for any secure communication scenario.

In addition to our last statement, we need to highlight two important properties for individual secrecy. The first property is that individual secrecy is only meaningful for multi-user communication scenarios. This is because for single-user cases, the individual secrecy criterion and the joint secrecy criterion are equivalent. The second property is related to the dilemma raised in Section 3.1.4 about the interpretation of individuality. Thus, we highlight one more time that in our investigation individuality is addressed with respect to different transmission flows and has nothing to do with the messages structure. Now, the main challenge is to develop coding schemes that can utilize the concept of mutual trust granted by the individual secrecy criterion to achieve bigger rate regions compared to the joint secrecy criterion.

4.1.2 Channel Model and Secrecy Criteria

The multi-receiver DM-WBC consists of a transmitter with an input alphabet $\mathcal{X}$, k legitimate receivers with output alphabets $\mathcal{Y}_j$, where $j \in [\![1;k]\!]$ and an external eavesdropper with output alphabet $\mathcal{Z}$. For the input and output sequences x^n, y_j^n and z^n of length n, the multi-receiver DM-WBC is given by the following transition matrix:

$$Q^n_{\mathrm{Y}_1\ldots\mathrm{Y}_k\mathrm{Z}|\mathrm{X}}(y_1^n,\ldots,y_k^n,z^n|x^n) = \prod_{i=1}^{n} Q_{\mathrm{Y}_1\ldots\mathrm{Y}_k\mathrm{Z}|\mathrm{X}}(x_i,y_{1_i},\ldots,y_{k_i},z_i). \tag{4.1}$$

Again, we assume that the transmitter and all receivers have perfect channel state information (CSI) and limit our investigation to this case. we also assume that the receiving nodes do not cooperate during decoding.

We consider a communication scenario in which the transmitter transmits k confidential messages, where each legitimate receiver is only interested in decoding his corresponding message as shown in the following figure:

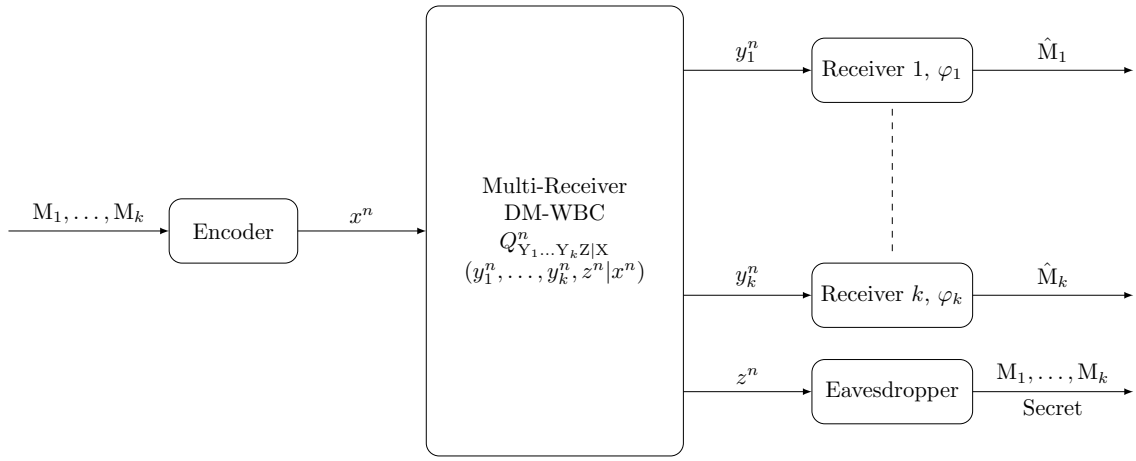


Figure 4.1: Multi-receiver DM-WBC

Definition 4.1. A $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ code $\mathcal{C}$ for the multi-receiver DM-WBC consists of: k independent confidential message sets $\mathcal{M}_1, \ldots, \mathcal{M}_K$, a stochastic encoder at the transmitter

$$E : \mathcal{M}_1 \times \cdots \times \mathcal{M}_k \to \mathcal{P}(\mathcal{X}^n),$$

which maps the k confidential messages $(m_1, \ldots, m_k) \in \mathcal{M}_1 \times \cdots \times \mathcal{M}_k$ into a codeword $x^n(m_1, \ldots, m_k)$ according to the conditional probability $E(x^n|m_1, \ldots, m_k)$, and k decoders, one for each legitimate receiver

$$\varphi_j : \mathcal{Y}_j^n \to \mathcal{M}_j \cup \{?\},$$

for $j \in [\![1; k]\!]$ that maps each channel observation at the respective receiver to the corresponding required message or an error message $\{?\}$.

The previous definition considers a standard block code $\mathcal{C}$ of arbitrary but fixed length n. It is fair to assume that $\mathcal{C}$ is known to the transmitter, the k legitimate receivers and the eavesdropper. We further assume that $\mathrm{M}_1, \ldots, \mathrm{M}_k$ are independent and uniformly distributed random variables over their respective message sets.

In order to evaluate the reliability performance of the code $\mathcal{C}$, we use the average decoding error probability given by:

$$\bar{\mathrm{P}}_e(\mathcal{C}) \triangleq \mathbb{P}\big[\hat{\mathrm{M}}_1 \neq \mathrm{M}_1 \text{ or } \ldots \text{ or } \hat{\mathrm{M}}_k \neq \mathrm{M}_k\big], \tag{4.2}$$

where $\hat{\mathrm{M}}_j$ is the estimated message at the jth legitimate receiver. Our target is to develop a coding scheme that complies with Definition 4.1 such that $\bar{\mathbf{P}}_e(\mathcal{C})$ is arbitrary small. Moreover, it follows from [54] that under the assumption of perfect CSI, all the results established under the average error probability constraint in Eq. (4.2) are also valid for the maximum decoding error probability requirement.

On the other hand, we need the constructed code $\mathcal{C}$ to assure that the k confidential messages are kept hidden from the eavesdropper. In order to measure the ignorance

of the eavesdropper about the confidential messages, we use the two secrecy criteria described in the previous chapter as follows:

1. Joint Secrecy: This criterion requires the leakage of the confidential message of one user to the eavesdropper given the confidential messages of all other users to be small. For our model, this requirement can be expressed as follows:

$$\mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \mathrm{M}_1 \dots \mathrm{M}_{j-1} \mathrm{M}_{j+1} \dots \mathrm{M}_k \mathcal{C}) \leq \tau_{jn}, \quad \text{where} \quad \lim_{n\to\infty} \tau_{jn} = 0. \tag{4.3}$$

This criterion guarantees that the information leaked to the eavesdropper from one user is small even if all the other confidential messages are compromised and known by the eavesdropper. In most of the previous literature, the joint secrecy criterion is defined such that, the mutual leakage of all confidential messages to the eavesdropper is small [73]. This is expressed as follows:

$$\mathbb{I}(\mathrm{M}_1 \dots \mathrm{M}_k; \mathrm{Z}^n | \mathcal{C}) \leq \tau_n \quad \text{where} \quad \lim_{n\to\infty} \tau_n = 0. \tag{4.4}$$

Although this definition is simpler than the one in (4.3), we can show that both definitions are equivalent for some $\tau_n \geq \sum_{j=1}^{k} \tau_{jn}$, using the same arguments introduced in Section 3.1.3. As we mentioned before that we prefer the definition in (4.3) because it provides a better understanding to the interpretation of the relation between the legitimate receivers under the joint secrecy criterion. It also highlights the reason behind the joint secrecy immunity against compromised receivers.

2. Individual Secrecy: This criterion requires the leakage of the confidential message of each user to the eavesdropper to be small without conditioning on the confidential messages of the others users. This requirement can be formulated as follows:

$$\mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \mathcal{C}) \leq \tau_{jn}, \quad \text{where} \quad \lim_{n\to\infty} \tau_{jn} = 0. \tag{4.5}$$

Differently from the conservative joint secrecy constraint in (4.3), the individual secrecy constraint takes the mutual trust between the legitimate receivers into consideration. This allows the legitimate receivers to cooperate in protecting their messages against eavesdropping. For $\tau_n \geq \sum_{j=1}^{k} \tau_{jn}$, the conditions in (4.5) can be combined into one constraint as follows:

$$\sum_{j=1}^{k} \mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \mathcal{C}) \leq \tau_n. \tag{4.6}$$

Although the secrecy constraint in (4.6) inherits the message individuality interpretation for individual secrecy, it still complies with the transmission flow interpretation as well. This is because in the model in Fig 4.1 we do not have a common confidential message, which implies that each individual confidential message simulates a separate information flow. This is different from the communication model investigated in the previous chapter, cf. Fig. 3.2, where we showed in Section 3.1.4 that the two interpretations for the individuality are not the same.

It is worth mentioning that the joint and individual secrecy constraints in (4.3) and (4.5) are defined according to the strong secrecy notation, cf. Section 2.1.4.

Definition 4.2. A rate tuple $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ is a joint-strong secrecy achievable rate tuple for the multi-receiver DM-WBC, if for all η_j where $j \in [\![1;k]\!]$, $\lambda > 0$ and $\tau_j > 0$, there is an $n(\eta_j, \lambda, \tau_j) \in \mathbb{N}$ such that for all $n > n(\eta_j, \lambda, \tau_j)$ there exists a sequence of codes $\{\mathcal{C}\}_n$ of Definition 4.1 that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_j| \geq R_j - \eta_j \tag{4.7a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda \tag{4.7b}$$

$$\mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \mathrm{M}_1 \ldots \mathrm{M}_{j-1} \mathrm{M}_{j+1} \ldots \mathrm{M}_k \mathcal{C}) \leq \tau_j, \tag{4.7c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by Eq. (4.2). The joint-strong secrecy capacity region $C_J(Q)$ is given by the set of all achievable joint-strong secrecy rate tuples $(R_1, \ldots, R_k)$.

In order to show that a rate tuple $(R_1, \ldots, R_k)$ is achievable under the joint-strong secrecy criterion, it is enough to show that there exists a random coding scheme with random variable C such that:

$$\frac{1}{n} \log |\mathcal{M}_j| = R_j \tag{4.8a}$$

$$\lim_{n \to \infty} \mathbb{E}\big[\bar{\mathbf{P}}_e(\mathrm{C})\big] = 0 \tag{4.8b}$$

$$\lim_{n \to \infty} \mathbb{E}\big[\mathbb{I}(\mathrm{M}_1 \ldots \mathrm{M}_k; \mathrm{Z}^n | \mathrm{C})\big] = 0. \tag{4.8c}$$

The validity of the previous argument follows from the Definition 4.2, the selection lemma and the equivalency of the joint secrecy constraints in (4.3) and (4.4).

Similarly, we present the following definition for the achievable rates under the individual secrecy constraint.

Definition 4.3. A rate tuple $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ is an individual-strong secrecy achievable rate tuple for the multi-receiver DM-WBC, if for all η_j where $j \in [\![1;k]\!]$, $\lambda > 0$ and $\tau_j > 0$, there is an $n(\eta_j, \lambda, \tau_j) \in \mathbb{N}$ such that for all $n > n(\eta_j, \lambda, \tau_j)$ there exists a sequence of codes $\{\mathcal{C}\}_n$ of Definition 4.1 that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_j| \geq R_j - \eta_j \tag{4.9a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda \tag{4.9b}$$

$$\mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \mathcal{C}) \leq \tau_j, \tag{4.9c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by Eq. (4.2). The individual-strong secrecy capacity $C_I(Q)$ is given by the set of all achievable individual-strong secrecy rate tuples $(R_1, \ldots, R_k)$.

In order to show that a rate tuple $(R_1, \ldots, R_k)$ is achievable under the individual-strong secrecy criterion, it is enough to show that there exists a random coding scheme with random variable C such that:

$$\frac{1}{n} \log |\mathcal{M}_j| = R_j \tag{4.10a}$$

$$\lim_{n \to \infty} \mathbb{E}\big[\bar{\mathbf{P}}_e(\mathrm{C})\big] = 0 \tag{4.10b}$$

$$\lim_{n \to \infty} \mathbb{E}\Big[\sum_{j=1}^{k} \mathbb{I}(\mathrm{M}_i; \mathrm{Z}^n | \mathrm{C})\Big] = 0. \tag{4.10c}$$

The validity of the previous argument follows from the Definition 4.3, the selection lemma and the equivalency of the individual secrecy constraints in (4.5) and (4.6).

The results established in this chapter were published in [138–140]. In a parallel and independent work, secure communication over the two-receiver DM-WBC was investigated under the joint secrecy criterion in [141, 142] and the individual secrecy criterion in [143, 144]. However, both investigations only considered the weak secrecy measure.

4.1.3 Rate Regions: Joint Vs Individual

In this section, we present some of the rate regions established for the two-receiver DM-WBC under both the joint and individual strong secrecy criteria. These results allow us to see how relaxing secrecy constraint from the joint to the individual secrecy criterion can lead to a larger capacity regions.

Corollary 4.1. *Consider a degraded two-receiver DM-WBC, where* $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ *forms a Markov chain. Then the joint-strong secrecy capacity region is given by the union over the set of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$*, that satisfy*

$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.11a}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}), \tag{4.11b}$$

for all possible choices of random variables (U,X)*, such that* $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$*.*

This region was first established in [73] as mentioned in Section 2.4.3 but under the joint-weak secrecy criterion. In Section 4.3.2, we provide a detailed achievability proof for a more general case showing that strengthening the joint secrecy criterion from weak secrecy to strong secrecy comes at no cost with respect to the secrecy capacity. On the other hand, the converse follows using the standard techniques and procedures for degraded DM-BC and is provided for the multi-receiver case in Section 4.2.1. Next, we present the individual-strong secrecy capacity region for the same channel as follows:

Corollary 4.2. *Consider a degraded two-receiver DM-WBC, where* $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ *forms a Markov chain. The individual-strong secrecy capacity region is given by the union over the set of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$*, that satisfy*

$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.12a}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.12b}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + R_2 \tag{4.12c}$$

for all possible choices of random variables (U,X)*, such that* $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ *forms a Markov chain. Further it suffices to have* $|\mathcal{U}| \leq |\mathcal{X}| + 3$*.*

The achievability proof combines the techniques of wiretap random coding along with secret key encoding and will be provided for the multi-receiver case in Section 4.2.2. It also follows as a special case from the general individual-strong secrecy achievable rate region which will be established in Section 4.3.3. On the other hand, the converse follows by adapting the standard techniques and procedures for degraded DM-BC

to the individual secrecy constraint and will be presented for the multi-receiver case in Section 4.2.1 as well. Now, if we compare the two capacity regions given by the previous two corollaries, we can notice the following:

- Eq. (4.11a) and Eq. (4.12a) imply that both secrecy measures have the same bound on the confidential rate intended to the second legitimate receiver (weaker legitimate receiver).
- The bound on R_2 in (4.12a) implies that even under the individual secrecy criterion, the optimal secrecy encoding scheme for the confidential message M_2 intended for Y_2 is a simple wiretap random coding.
- The two bounds on R_1 in (4.12b) and (4.12c) compared to the one in (4.11b) imply that individual secrecy allows a larger achievable rate for the first legitimate receiver (stronger legitimate receiver).
- The bound in (4.12c) clearly indicate that the optimal secrecy encoding scheme for the confidential message M_1 intended for Y_1 is some sort of combination between wiretap encoding and secret key encoding, where the secret key encoding part is somehow related to the confidential message M_2 intended for Y_2.
- The bound in (4.12b) on the other hand is very interesting as it implies that under the individual secrecy constraint the first legitimate receiver can achieve a secrecy rate larger than the one achieved for the public DM-BC without secrecy constraints, cf. Theorem 2.6. The reason for such behavior is that unlike the public case where the message intended for the first legitimate receiver is encoded in the satellite codeword only, we will see that under the individual secrecy the message intended for the first legitimate receiver is not only encoded in the satellite codeword but also in the cloud center codeword.

In order to provide a better illustration regarding the effect of relaxing the secrecy constraint (from joint to individual) on the secrecy capacity, we considered a very common channel in wireless communication which is the Gaussian SISO-WBC. Since the Gaussian SISO-WBC is an example for the Degraded DM-WBC, we have the following results:

Corollary 4.3. *The joint-strong secrecy capacity region of the two receiver Gaussian SISO-WBC is given by the union of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy*

$$R_2 \leq f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_2^2}\right) - f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_Z^2}\right) \tag{4.13a}$$

$$R_1 \leq f\left(\frac{\alpha P}{\sigma_1^2}\right) - f\left(\frac{\alpha P}{\sigma_Z^2}\right) \tag{4.13b}$$

where the union is taken over all values of $\alpha \in [0, 1]$, such that $\bar{\alpha} = 1 - \alpha$ and the function $f(a)$ is defined as follows: $f(a) \triangleq \frac{1}{2}\log(1 + a)$.

This region was first established in [93, 94] as highlighted in Section 2.4.7 but under the joint-weak secrecy criterion. The achievability follows by selecting the random variables (U, X) in Corollary 4.1 to be jointly Gaussian, where $\mathrm{X} = \mathrm{U} + \mathrm{V}$ can be viewed as the summation of two independent zero-mean Gaussian random variables U and V, with respective variances $\bar{\alpha}P$ and αP. On the other hand, the converse follows due to the optimality of Gaussian signaling.

Corollary 4.4. *The individual-strong secrecy capacity region of the two receive Gaussian SISO-WBC is given by the union of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfy*

$$R_2 \leq f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_2^2}\right) - f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_Z^2}\right) \tag{4.14a}$$

$$R_1 \leq f\left(\frac{\alpha P}{\sigma_1^2}\right) + f\left(\frac{\bar{\alpha}P}{\alpha P + \sigma_Z^2}\right) \tag{4.14b}$$

$$R_1 \leq f\left(\frac{\alpha P}{\sigma_1^2}\right) - f\left(\frac{\alpha P}{\sigma_Z^2}\right) + R_2 \tag{4.14c}$$

where the union is taken over all values of $\alpha \in [0,1]$*, such that* $\bar{\alpha} = 1 - \alpha$ *and the function* $f(a)$ *is defined as follows:* $f(a) \triangleq \frac{1}{2}\log(1+a)$.

The achievability and the converse proofs follow by using the same techniques used to establish the previous corollary. In Section 4.2.4, we will present a detailed proof for the multi-receiver case. We now use the capacity regions in Corollary 4.3 and 4.4 to provide a numerical example that illustrates the advantage of individual secrecy over joint secrecy in terms of the achievable confidential rates.

Consider a two-receiver GWBC with the following parameters: $\sigma_1^2 = 0.05$, $\sigma_2^2 = 0.1$ and $\sigma_Z^2 = 0.15$. Additionally, let the total power achievable at the transmitter be $P = 1$. The normalized confidential rates at different values of the superposition parameter α are plotted in Fig. 4.2.

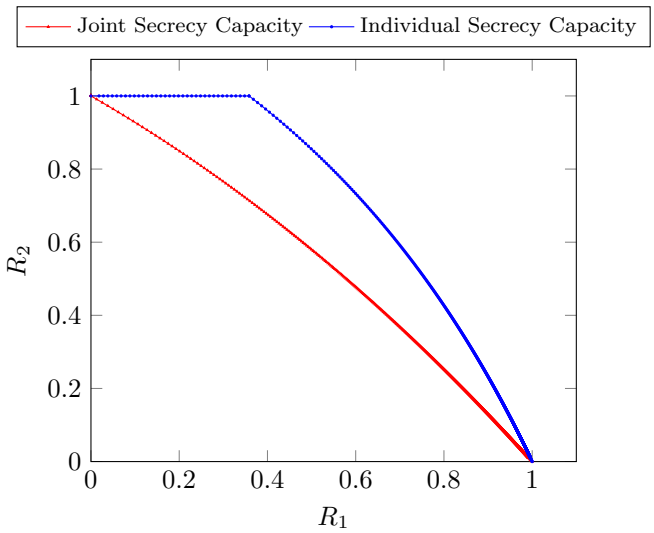


Figure 4.2: Joint and individual secrecy capacity regions of a two-receiver Gaussian SISO-WBC.

Although the previous example clearly shows that the individual secrecy criterion can outperform the joint secrecy criterion in terms of the secrecy capacity region, it is important to keep in mind that the individual secrecy is vulnerable against compromised receivers. Thus, as we mentioned in the previous chapter, The process of choosing among the two secrecy criteria is a simple trade-off between a powerful secrecy measure and a larger capacity region.

4.2 Secrecy Capacity Regions for the Degraded Multi-Receiver WBC

In this section, we limit our investigation to the problem of secure communication over the class of degraded multi-receiver DM-WBC as shown in Fig. 4.3. In this investigation, we establish the second main result of this thesis which is the individual-strong secrecy capacity regions of the degraded, the Gaussian SISO and the degraded Gaussian MIMO multi-receiver WBCs. Moreover, we show that the secrecy capacity region established in [75] under the joint-weak secrecy criterion for the degraded multi-receiver DM-WBC is also valid under the joint-strong secrecy criterion.

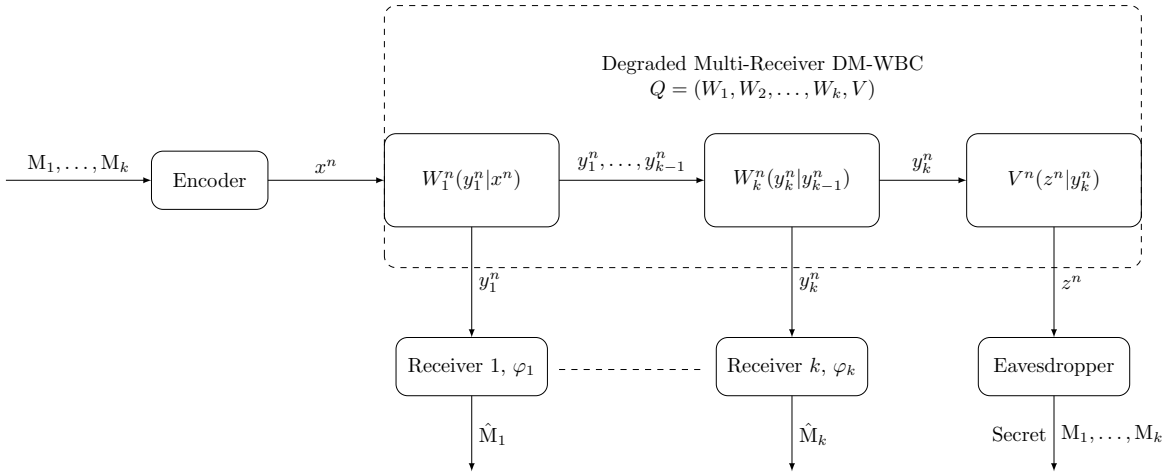


Figure 4.3: Degraded multi-receiver DM-WBC

Before we present our results, we present a quick review about the degraded multi-receiver DM-WBC in Fig. 4.3 and its main features. Degraded multi-receiver DM-WBCs are a special case of the multi-receiver DM-WBC introduced in Section 4.1.2, where the following relation holds:

$$\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \cdots - \mathrm{Y}_k - \mathrm{Z}. \tag{4.15}$$

The previous Markov chain is the main feature of a degraded multi-receiver DM-WBC. The most important implication of this Markov chain is that each legitimate receiver is capable of not only decoding its own message, but also the messages of all the receivers degraded from it. This property plays an important role in establishing the individual-strong secrecy capacity region.

4.2.1 Joint Secrecy: From Weak to Strong

In Section 2.2, we discussed the different secrecy measures used for secure communication over the discrete memoryless wiretap channel. We highlighted that strengthening the secrecy measure from weak to strong secrecy or even effective secrecy does not have an effect on the secrecy capacity of the DM-WC. We will show that the previous statement also holds for the degraded multi-receiver DM-WBC under the joint secrecy criterion. For this, we present the following result:

Theorem 4.1. *The joint-strong secrecy capacity region of the degraded multi-receiver DM-WBC is given by the union over the set of all rate tuples* $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *which satisfy*

$$R_j \leq \mathbb{I}(\mathrm{U}_j; \mathrm{Y}_j | \mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j; \mathrm{Z} | \mathrm{U}_{j+1}), \qquad \text{for} \quad j \in [\![1, k]\!], \tag{4.16}$$

for all possible choices of random variables $(\mathrm{U}_1, \ldots, \mathrm{U}_{k+1})$ *such that* $\mathrm{U}_1 = \mathrm{X}$, $\mathrm{U}_{k+1} = \emptyset$. *Additionally, the following Markov chain holds:* $\mathrm{U}_k - \cdots - \mathrm{X} - \mathrm{Y}_1 - \cdots - \mathrm{Y}_k - \mathrm{Z}$. *Further, it suffices to have* $|\mathcal{U}_j| \leq B_c\big(|\mathcal{U}_{j+1}|\big)(|\mathcal{X}| + 2j - 1)$, *where* $B_c(|\mathcal{A}|)$ *is the upper-bound of the cardinality of the set* $\mathcal{A}$ *and* $B_c(|\mathcal{U}_{k+1}|) \triangleq 1$.

The previous theorem generalizes the joint-strong secrecy capacity region of the degraded two-receiver DM-WBC given by Corollary 4.1. In order to prove the achievability of the rate region in (4.16), we first show that the two-receiver rate region in (4.11) is achievable. We then argue that by extending the coding scheme from two receivers to k receivers, the rate region in (4.16) follows accordingly. We already showed in Section 2.4.4 that a coding scheme that combines the principles of superposition encoding and wiretap random coding can be used to establish the achievability of the rate region in (4.11) for the two-receiver DM-WBC under the joint-weak secrecy criterion. Thus, we only need to show that the constraints on the randomization rates in Eq. (2.89) derived under the weak secrecy analysis are enough to assure strong secrecy as well. This can be done by applying the strong secrecy techniques discussed in Section 2.2.3 and will be shown in details for a more complex coding scheme in Section 4.3.2.

For the converse, we present a simpler proof than the one given in [75]. This proof will help us in establishing the converse proof for the individual secrecy criterion. Our proof is based on the standard converse techniques in addition to the properties of the degraded multi-receiver DM-WBC, in particular the Markov chain in (4.15).

Now, let $(R_1, \ldots, R_k)$ be an achievable joint-strong secrecy rate tuple for the degraded multi-receiver DM-WBC. Furthermore, for $j \in [\![1, k]\!]$, let $\tau_j > 0$ and $\lambda > 0$, then assume that for a sufficiently large $n \in \mathbb{N}$, there exists a $(2^{nR_1}, \ldots, 2^{nR_k}, n)$ code $\mathcal{C}$ of Definition 4.1 such that:

$$nR_j = \mathbb{H}(\mathrm{M}_j | \mathcal{C}), \tag{4.17a}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n = 2^{-n\lambda} \tag{4.17b}$$

$$\mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \ddot{\mathrm{M}}_{j+1} \mathcal{C}) \leq \tau_{jn} = 2^{-n\tau_j}, \tag{4.17c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by (3.7) and $\ddot{\mathrm{M}}_{j+1} \triangleq (\mathrm{M}_{j+1}, \ldots, \mathrm{M}_k)$. In order to simplify the notation, the conditioning on the code $\mathcal{C}$ will be ignored in the upcoming steps. Using the relations in (4.17), we can derive the following bound on the joint secrecy rate intended for each legitimate receiver:

$$\begin{aligned}
R_j &\overset{(a)}{=} \frac{1}{n} \mathbb{H}(\mathrm{M}_j | \ddot{\mathrm{M}}_{j+1}) \\
&\overset{(b)}{\leq} \frac{1}{n} \Big[\mathbb{H}(\mathrm{M}_j | \ddot{\mathrm{M}}_{j+1}) - \mathbb{H}(\mathrm{M}_j | \mathrm{Y}_j^n \ddot{\mathrm{M}}_{j+1}) \Big] + \Gamma_n(\lambda_n) \\
&= \frac{1}{n} \mathbb{I}(\mathrm{M}_j; \mathrm{Y}_j^n | \ddot{\mathrm{M}}_{j+1}) + \Gamma_n(\lambda_n) \\
&\overset{(c)}{\leq} \frac{1}{n} \Big[\mathbb{I}(\mathrm{M}_j; \mathrm{Y}_j^n | \ddot{\mathrm{M}}_{j+1}) - \mathbb{I}(\mathrm{M}_j; \mathrm{Z}^n | \ddot{\mathrm{M}}_{j+1}) \Big] + \Gamma_n(\lambda_n, \tau_{jn}) \qquad (4.18)
\end{aligned}$$

where $\Gamma_n(\lambda_n) = 1/n\ \mathbb{H}_2(\lambda_n) + \lambda_n \sum_{j=1}^{k} R_j$ and $\Gamma_n(\lambda_n, \tau_{jn}) = \Gamma_n(\lambda_n) + \tau_{jn}/n$. One can easily show that $\Gamma_n(\lambda_n, \tau_{jn})$ approaches 0 as n approaches infinity. Step (a) follows due to the independence of the messages; (b) follows by applying Fano's inequality along with the fact that conditioning decreases entropy; while (c) follows from the secrecy constraint in (4.17c). Next, we let $\mathrm{U}_j^i \triangleq (\mathrm{M}_j, \mathrm{Y}_{j-1}^{i-1}, \tilde{\mathrm{Z}}^{i+1}, \mathrm{U}_{j+1}^i)$, where $\mathrm{Y}_0^{i-1} \triangleq \mathrm{U}_{k+1}^i \triangleq \emptyset$, $\mathrm{Y}_j^{i-1} = (\mathrm{Y}_{j1}, \ldots, \mathrm{Y}_{ji})$ and $\tilde{\mathrm{Z}}^{i+1} = (\mathrm{Z}_{i+1}, \ldots, \mathrm{Z}_n)$, then it follows that:

$$
\begin{aligned}
R_j &\overset{(a)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j; \mathrm{Y}_{ji}|\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M}_j; \mathrm{Z}_i|\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + \Gamma_n(\lambda_n, \tau_{jn}) \\
&\overset{(b)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j; \mathrm{Y}_{ji}|\ddot{\mathrm{M}}_{j+1}\dot{\mathrm{Y}}_j^{i-1}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M}_j; \mathrm{Z}_i|\ddot{\mathrm{M}}_{j+1}\dot{\mathrm{Y}}_j^{i-1}\tilde{\mathrm{Z}}^{i+1})\Big] + \Gamma_n(\lambda_n, \tau_{jn}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j; \mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) - \mathbb{I}(\mathrm{M}_j; \mathrm{Z}_i|\mathrm{U}_{j+1}^i)\Big] + \Gamma_n(\lambda_n, \tau_{jn}) \\
&\overset{(c)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{U}_j^i; \mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) - \mathbb{I}(\mathrm{U}_j^i; \mathrm{Z}_i|\mathrm{U}_{j+1}^i)\Big] + \Gamma_n(\lambda_n, \tau_{jn}) \qquad (4.19)
\end{aligned}
$$

where we use the following notation $\dot{\mathrm{Y}}_j^{i-1} \triangleq (\mathrm{Y}_j^{i-1}, \ldots, \mathrm{Y}_{k-1}^{i-1})$. Step (a) follows from the Csiszár sum identity [10, Lemma 7]; (b) follows from the Markov chain in (4.15) which implies that $(\mathrm{Y}_{j+1}, \ldots, \mathrm{Y}_k)$ are degraded from Y_j; while (c) follows because Z_i is degraded from Y_{ji}, which implies that $\mathbb{I}(\mathrm{Y}_{j+1}^{i-1}; \mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) \geq \mathbb{I}(\mathrm{Y}_{j+1}^{i-1}; \mathrm{Z}_i|\mathrm{U}_{j+1}^i)$.

Finally, for the bound in (4.19), we introduce a random variable T independent of all others and uniformly distributed over $[\![1; n]\!]$ and let $\mathrm{U}_j = (\mathrm{U}_j^{\mathrm{T}}, \mathrm{T})$, $\mathrm{Y}_j = \mathrm{Y}_{j\mathrm{T}}$ and $\mathrm{Z} = \mathrm{Z}_{\mathrm{T}}$ for all $j \in [\![1, k]\!]$. We then take the limit as $n \to \infty$, which implies that $\Gamma_n(\lambda_n, \tau_{jn}) \to 0$. This concludes our converse. It only remains to highlight that the cardinality bounds follow from the Fenchel-Bunt strengthening of the Carathéodory's theorem as in [35, Appendix C], while the validity of the Markov chain $\mathrm{U}_k - \cdots - \mathrm{U}_2 - \mathrm{X}$ in the converse follows using the principle of functional dependence graph [145].

4.2.2 Individual-Strong Secrecy Capacity Region

Theorem 4.1 indicates that superposition wiretap encoding is optimal for the degraded multi-receiver DM-WBC. According to the properties of superposition encoding, we know that in order to decode the information encoded in the satellite codeword, one needs to decode the information in the cloud center codeword as well. This implies that for the degraded two-receiver DM-WBC, the first legitimate receiver (strong receiver) does not only decode its own confidential message M_1 but it also decodes the confidential message M_2 intended for the second legitimate receiver (weak receiver).

This observation suggests that M_2 can be interpreted in some sense as a secret key shared between the transmitter and the first legitimate receiver. Thus, under the individual secrecy measure, where the legitimate receivers trust each other, the first legitimate receiver can achieve a higher secrecy rate. This idea motivated us to investigate secure communication over the degraded multi-receiver DM-WBC under the individual secrecy constraint. We present the following result:

Theorem 4.2. *The individual-strong secrecy capacity region of the degraded multi-receiver DM-WBC is given by the union over the set of all rate tuples* $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *that satisfy*

$$R_j \leq \mathbb{I}(\mathrm{U}_j; \mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j; \mathrm{Z}|\mathrm{U}_{j+1}) + \sum_{l=j+1}^{k} R_l \tag{4.20a}$$

$$R_j \leq \mathbb{I}(\mathrm{U}_j; \mathrm{Y}_j|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_{j+1}; \mathrm{Z}) \tag{4.20b}$$

$$\sum_{l=j}^{k} R_l \leq \sum_{l=j}^{k} \mathbb{I}(\mathrm{U}_l; \mathrm{Y}_l|\mathrm{U}_{l+1}) \tag{4.20c}$$

for all possible choices of random variables $(\mathrm{U}_1, \ldots, \mathrm{U}_{k+1})$ *such that* $\mathrm{U}_1 = \mathrm{X}$, $\mathrm{U}_{k+1} = \emptyset$. *Additionally, the following Markov chain holds:* $\mathrm{U}_k - \cdots - \mathrm{X} - \mathrm{Y}_1 - \cdots - \mathrm{Y}_k - \mathrm{Z}$. *Further, it suffices to have* $|\mathcal{U}_j| \leq B_c\big(|\mathcal{U}_{j+1}|\big)(|\mathcal{X}| + 2j - 1)$, *where* $B_c(|\mathcal{A}|)$ *is the upper-bound of the cardinality of the set* $\mathcal{A}$ *and* $B_c(|\mathcal{U}_{k+1}|) \triangleq 1$.

In order to present an achievability proof for the previous coding theorem, we use the following procedure: We start by showing that the for the degraded two-receiver DM-WBC, the individual-strong secrecy capacity region given by Corollary 4.2 is achievable. We then argue that by extending the coding scheme from two receivers to k receivers, the rate region in Eq. (4.20) will follow accordingly. Thus, for now we limit ourselves to the two-receiver channel model and present a coding scheme that can achieve the rate region in (4.12):

Let $n \in \mathbb{N}$ be an arbitrary but fixed code block length and $\epsilon > 0$ be an arbitrary constant. Moreover, let $\mathrm{Q}_{\mathrm{UX}} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})$ be a fixed input probability distribution and recall the probability transition matrix $Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$ given in (4.1) that defines a degraded two-receiver DM-WBC where $\mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ forms a Markov chain.

1. Message Sets: We consider the following sets: two individual confidential message sets $\mathcal{M}_1 = [\![1, 2^{nR_1}]\!]$ and $\mathcal{M}_2 = [\![1, 2^{nR_2}]\!]$ along with two randomization message sets $\mathcal{M}_{r_1} = [\![1, 2^{nR_{r_1}}]\!]$ and $\mathcal{M}_{r_2} = [\![1, 2^{nR_{r_2}}]\!]$. Furthermore, we assume that quantities of the form 2^{nR}, where $R \in \mathbb{R}_+$ are integers.

2. Rate Splitting: We start by dividing each individual message set $\mathcal{M}_j$, for $j \in \{1, 2\}$ into three independent parts $\mathcal{M}_{jl} = [\![1, 2^{nR_{jl}}]\!]$, for $l \in \{1, 2, 3\}$. In this division, we force $\mathcal{M}_{11}$ and $\mathcal{M}_{21}$ to be of the same size and use them to construct $\mathcal{M}_{\otimes_1} = [\![1, 2^{nR_{\otimes_1}}]\!]$ by *xoring* their corresponding elements. We also make sure that $\mathcal{M}_{12}$ and $\mathcal{M}_{22}$ are of the same size and use them to construct $\mathcal{M}_{\otimes_2} = [\![1, 2^{nR_{\otimes_2}}]\!]$. The previous rate splitting procedure implies the following rate constraints:

$$R_{\otimes_1} = R_{11} = R_{21} \quad \text{and} \quad R_{\otimes_2} = R_{12} = R_{22} \tag{4.21a}$$

$$R_{11} + R_{12} \leq R_2. \tag{4.21b}$$

3. Random Codebook $\mathcal{C}$: Using the input distribution Q_{UX}, we generate the cloud centers codewords $u^n(m_2, m_{\otimes_1}, m_{r_1})$ for $m_2 \in \mathcal{M}_2$, $m_{\otimes_1} \in \mathcal{M}_{\otimes_1}$ and $m_{r_1} \in \mathcal{M}_{r_2}$ by generating the symbols $u_i(m_2, m_{\otimes_1}, m_{r_1})$ with $i \in [\![1, n]\!]$ independently according to Q_{U}. For every $u^n(m_2, m_{\otimes_1}, m_{r_1})$, we generate the satellite codewords $x^n(m_2, m_{\otimes_1}, m_{r_1}, m_{13},$

$m_{\otimes_2}, m_{r_2})$ for $m_{13} \in \mathcal{M}_{13}$, $m_{\otimes_2} \in \mathcal{M}_{\otimes_2}$, and $m_{r_2} \in \mathcal{M}_{r_2}$ by generating the symbols $x_i(m_2, m_{\otimes_1}, m_{r_1}, m_{13}, m_{\otimes_2}, m_{r_2})$, independently at random according to $\mathrm{Q}_{\mathrm{X}|\mathrm{U}}(\cdot|u_i(m_2, m_{\otimes_1}, m_{r_1}))$.

4. Encoding E: Given a message pair $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, the encoder deduces the corresponding message triple $(m_{13}, m_{\otimes_1}, m_{\otimes_2})$ then chooses a message pair (m_{r_1}, m_{r_2}) uniformly at random from the sets $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$. Finally, it transmits the chosen codeword $x^n(m_2, m_{\otimes_1}, m_{r_1}, m_{13}, m_{\otimes_2}, m_{r_2})$.

5. First Legitimate Decoder φ_1: Upon receiving y_1^n, it outputs $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12}, \hat{m}_{13})$. First, it finds the unique pair of codewords messages $\big(u^n(\hat{m}_2, \hat{m}_{\otimes_1}, \hat{m}_{r_1}), x^n(\hat{m}_2, \hat{m}_{\otimes_1}, \hat{m}_{r_1}, \hat{m}_{13}, \hat{m}_{\otimes_2}, \hat{m}_{r_2})\big)$ which is jointly typical with y_1^n. Next, it computes the pair $(\hat{m}_{11}, \hat{m}_{12})$ by *xoring* $(\hat{m}_{21}, \hat{m}_{22})$ and $(\hat{m}_{\otimes_1}, \hat{m}_{\otimes_2})$.

6. Second Decoder φ_2: Given y_2^n, it outputs $\tilde{m}_2$ by finding the unique codeword $u^n(\tilde{m}_2, \tilde{m}_{\otimes_1}, \tilde{m}_{r_1})$ which is jointly typical with y_2^n.

7. Reliability Analysis: In order to derive the rate constraints that assure that this coding scheme satisfies the reliability constraint in (4.10b), we start by defining the following average decoding error probability:

$$\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \triangleq \mathbb{P}\Big[(\hat{\mathrm{M}}_2, \hat{\mathrm{M}}_{\otimes_1}, \hat{\mathrm{M}}_{\mathrm{r}_1}, \hat{\mathrm{M}}_{13}, \hat{\mathrm{M}}_{\otimes_2}, \hat{\mathrm{M}}_{\mathrm{r}_2}) \neq (\mathrm{M}_2, \mathrm{M}_{\otimes_1}, \mathrm{M}_{\mathrm{r}_1}, \mathrm{M}_{13}, \mathrm{M}_{\otimes_2}, \mathrm{M}_{\mathrm{r}_2}) \\ \text{or } (\tilde{\mathrm{M}}_2, \tilde{\mathrm{M}}_{\otimes_1}, \tilde{\mathrm{M}}_{\mathrm{r}_1}) \neq (\mathrm{M}_2, \mathrm{M}_{\otimes_1}, \mathrm{M}_{\mathrm{r}_1})\Big]. \tag{4.22}$$

We then observe that $\bar{\bar{\mathbf{P}}}_e(\mathcal{C}) \geq \bar{\mathbf{P}}_e(\mathcal{C})$, cf. (4.2) for $k = 2$. Using the standard joint-typicality decoding arguments along with the fact that Y_2 is degraded from Y_1, we can show that $\lim_{n\to\infty} \bar{\bar{\mathbf{P}}}_e(\mathcal{C}) = 0$ as long as:

$$R_2 + R_{\otimes_1} + R_{r_1} \leq \mathbb{I}(\mathrm{U}; \mathrm{Y}_2) - \delta_2(\epsilon) \tag{4.23a}$$
$$R_{13} + R_{\otimes_2} + R_{r_2} \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}_1|\mathrm{U}) - \delta_1(\epsilon). \tag{4.23b}$$

where $\epsilon > 0$ is the constant used during the typicality decoding. Moreover ϵ can be chosen in a way such that $\lim_{n\to\infty} \delta_1(\epsilon), \delta_2(\epsilon) = 0$. In Section 4.3.3, we present a detailed reliability analysis for a more complex coding scheme that include this coding scheme as a special case.

8. Secrecy Analysis: In order to derive the rate constraints needed to assure that our coding scheme satisfies the individual-strong secrecy constraint in (4.10c), we start arguing that for a fixed code $\mathcal{C}$ the constraint in (4.10c) holds, if:

$$\lim_{n\to\infty} \Big[\mathbb{I}(\mathrm{M}_2\mathrm{M}_{13}; \mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_{11}\mathrm{M}_{12}; \mathrm{Z}^n|\mathrm{M}_{13})\Big] = 0, \tag{4.24}$$

where the previous formulation follows due to the rate splitting procedure used in our coding scheme and the fact that the messages are independent. Based on the strong secrecy analysis of wiretap random codes presented in Section 2.2, the first term in Eq. (4.24) decays exponentially in n, if:

$$R_{\otimes_1} + R_{r_1} \geq \mathbb{I}(\mathrm{U}; \mathrm{Z}) + \tilde{\delta}_1(\epsilon, \gamma) \tag{4.25a}$$
$$R_{\otimes_2} + R_{r_2} \geq \mathbb{I}(\mathrm{X}; \mathrm{Z}|\mathrm{U}) + \tilde{\delta}_2(\epsilon, \gamma), \tag{4.25b}$$

where $\gamma > 0$ can be selected in a way such that $\lim_{n\to\infty} \tilde{\delta}_1(\epsilon,\gamma), \tilde{\delta}_2(\epsilon,\gamma) = 0$. On the other hand, using the concept of secret key encoding, we can prove that the second term in Eq. (4.24) vanishes as follows:

$$\begin{aligned}
\mathbb{I}(\mathrm{M}_{11}\mathrm{M}_{12};\mathrm{Z}^n|\mathrm{M}_{13}) &= \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}|\mathrm{M}_{13}) - \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}|\mathrm{Z}^n\mathrm{M}_{13}) \\
&\overset{(a)}{=} \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}) - \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}|\mathrm{Z}^n\mathrm{M}_{13}) \\
&\overset{(b)}{\leq} \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}) - \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}|\mathrm{M}_{13}\mathrm{M}_{\otimes_1}\mathrm{M}_{\otimes_2}\mathrm{M}_{r_1}\mathrm{M}_{r_2}) \\
&\overset{(c)}{=} \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}) - \mathbb{H}(\mathrm{M}_{11}\mathrm{M}_{12}|\mathrm{M}_{\otimes_1}\mathrm{M}_{\otimes_2}) \overset{(d)}{=} 0 \qquad (4.26)
\end{aligned}$$

where (a) follows because the messages are independent; (b) follows because $(\mathrm{M}_2, \mathrm{M}_{\otimes_1}, \mathrm{M}_{r_1}, \mathrm{M}_{13}, \mathrm{M}_{\otimes_2}, \mathrm{M}_{r_2}) - \mathrm{X}^n - \mathrm{Z}^n$ forms a Markov chain and the fact that under the rate constraint in (4.25), $\mathbb{I}(\mathrm{M}_2\mathrm{M}_{13};\mathrm{Z}^n)$ decays exponentially in n, these two arguments imply that given Z^n and the message M_{13} the best decoder at the eavesdropper can at most decode the messages $(\mathrm{M}_{\otimes_1}, \mathrm{M}_{r_1}, \mathrm{M}_{\otimes_2}, \mathrm{M}_{r_2})$ but not M_2; (c) follows because only $\mathrm{M}_{\otimes_1}$ and $\mathrm{M}_{\otimes_2}$ are related to M_{11} and M_{12}; while (d) follows due to the secret key encoding principle and the fact that the entropy of M_{21} and M_{22} is equivalent to the entropy of M_{11} and M_{12} respectively. This equivalence follows because the random variables that represent each secret key and the corresponding confidential message are uniformly distributed over the same message set cf. Eq. (4.21).

The previous secrecy analysis shows that our coding scheme uses wiretap random coding to protect the confidential messages M_2 and M_{13}, while the confidential messages M_{11} and M_{12} are protected using secret key encoding. Moreover, one can observe that the secret key encoded message $\mathrm{M}_{\otimes_1}$ that contain the confidential message M_{11} is encoded in the cloud center codeword, while $\mathrm{M}_{\otimes_2}$ that contain the confidential message M_{12} is encoded in the satellite codeword. This result is very interesting because both message are only intended to the first legitimate receiver Y_1 (stronger receiver) and in superposition encoding the information intended for the stronger receiver are usually encoded in the satellite codeword only.

9. Fourier-Motzkin Elimination: In order to finalize our achievability proof, we need to confirm that the rate constraint established in (4.21), (4.23) and (4.25) leads to the rate region in (4.12). We start by using the rate constraints in (4.25) with equality and combine them with the rate constraints in (4.23). Thus, we have

$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) - \acute{\delta}_2(\epsilon,\gamma) \qquad (4.27a)$$

$$R_{13} \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) - \acute{\delta}_1(\epsilon,\gamma), \qquad (4.27b)$$

where $\acute{\delta}_2(\epsilon,\gamma) = \delta_2(\epsilon) + \tilde{\delta}_1(\epsilon,\gamma)$ and $\acute{\delta}_1(\epsilon,\gamma) = \delta_1(\epsilon) + \tilde{\delta}_2(\epsilon,\gamma)$. On the other hand, using the rate constraints in (4.25) with equality along with the ones in (4.21a) directly implies that:

$$R_{11} + R_{12} = R_{\otimes_1} + R_{\otimes_2} \leq \mathbb{I}(\mathrm{X};\mathrm{Z}). \qquad (4.28)$$

Finally, we apply the Fourier-Motzkin elimination on the rate constraints in (4.27), (4.28) and (4.21b) to solve for R_2 and R_1, while keeping in mind that R_1 is the summation of R_{11}, R_{12} and R_{13}. We then take the limit as $n \to \infty$, which implies that $\acute{\delta}_1(\epsilon,\gamma), \acute{\delta}_2(\epsilon,\gamma) \to 0$. By doing so, we prove the achievability of any rate pair (R_1, R_2) satisfying the bounds in (4.12).

Now, in order to prove that the rate region given by Theorem 4.2 is achievable, we need to extend the previous coding scheme to the multi-user case. The main approach is to combine the principles of wiretap random coding and secret key encoding as follows: First we divide each confidential message into $k+1$ independent parts. Second we use wiretap random coding only to protect the confidential message of the weakest legitimate receiver Y_k. For any other legitimate receiver Y_j where $j \in [\![1, k-1]\!]$, wiretap random coding is used to protect only one part of the confidential message M_j. The remaining parts of M_j are protected using secret key encoding where the messages of the weaker receivers are used as secret keys. Moreover, the secret key encoded messages for the receiver Y_j can be part of the randomization indexes in its own superposition layer U_j and all lower layers as well. Following such coding scheme establishes the achievability of the rate region in (4.20) because:

- The bounds in (4.20a) and (4.20b) result from the fact that the individual secrecy rate of each legitimate receiver Y_j for $j \in [\![1, k]\!]$ is bounded by the summation of two rates:
  1. The wiretap rate encoded in its layer U_j conditioned on the previous layer U_{j+1}. This rate is given by:
  $$\mathbb{I}(\mathrm{U}_j; \mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j; \mathrm{Z}|\mathrm{U}_{j+1})$$
  2. The secret key encoded rate, where the secret key encoded rate is bounded by minimum of the secrecy rates of the weaker receivers $\sum_{l=j+1}^{k} R_l$ or the randomization rates that can be decoded by this receiver (randomization rates in its layer and all lower layers):
  $$\sum_{l=j}^{k} \mathbb{I}(\mathrm{U}_l; \mathrm{Z}|\mathrm{U}_{l+1}) = \mathbb{I}(\mathrm{U}_j; \mathrm{Z}|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_{j+1}; \mathrm{Z}). \tag{4.29}$$
- For the degraded two-receiver DM-WBC, i.e. $k = 2$, the bounds in (4.20a) and (4.20b) are enough as they directly lead to the rate region in (4.12). However, for the multi-receiver case, we need the third bound in (4.20c) to assure two different requirements:
  1. Any randomization rate that is used to carry a secret key encoded message for a certain user can only be used once. In other words, if the randomization rate $\mathbb{I}(\mathrm{U}_{j+2}; \mathrm{Z}|\mathrm{U}_{j+3})$ is used to carry part of the secret key encoded message of the legitimate receiver Y_{j+1}, it can not be used at the same time to carry part of the secret key encoded message of Y_j.
  2. The total sum rate is bounded by the summation of the information encoded in each layer for the corresponding receiver. This constraint in particular implies that although the individual secrecy rate for one of the legitimate receiver might be larger than the corresponding public rate cf. Theorem 2.6 and Corollary 4.4, the sum rate of the individual secrecy rates can not exceed the sum rate of all public rates.

For the converse, we start by letting $(R_1, \ldots, R_k)$ be an achievable individual-strong secrecy rate tuple for the degraded multi-receiver DM-WBC. Further, for $j \in [\![1, k]\!]$, let $\tau_j > 0$ and $\lambda > 0$, we assume that for a sufficiently large $n \in \mathbb{N}$, there exists a $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code $\mathcal{C}$ of Definition 4.1 that satisfy the following:

$$nR_j = \mathbb{H}(\mathrm{M}_j|\mathcal{C}), \tag{4.30a}$$
$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda_n = 2^{-n\lambda} \tag{4.30b}$$
$$\mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n|\mathcal{C}) \leq \tau_{jn} = 2^{-n\tau_j}, \tag{4.30c}$$

where $\bar{\mathbf{P}}_e(\mathcal{C})$ is given by (3.7). In order to simplify the notation, we ignore the conditioning on the code $\mathcal{C}$ in the rest of the steps. Using the relations in (4.30), we can derive the following bound on the individual secrecy rate intended for each legitimate receiver as follows:

$$\begin{aligned}
R_j &\overset{(a)}{\leq} \frac{1}{n}\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n) + \Gamma_n(\lambda_n) \\
&\overset{(b)}{\leq} \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n) - \mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n)\Big] + \Gamma_n(\lambda_n,\tau_{jn}) \\
&\overset{(c)}{\leq} \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n|\ddot{\mathrm{M}}_{j+1}) - \mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n|\ddot{\mathrm{M}}_{j+1}) + \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n|\mathrm{M}_j)\Big] + \Gamma_n(\lambda_n,\tau_{jn}) \\
&\overset{(d)}{\leq} \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n|\ddot{\mathrm{M}}_{j+1}) - \mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n|\ddot{\mathrm{M}}_{j+1})\Big] + \sum_{l=j+1}^{k} R_l + \Gamma_n(\lambda_n,\tau_{jn}) \\
&\overset{(e)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{U}_j^i;\mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) - \mathbb{I}(\mathrm{U}_j^i;\mathrm{Z}_i|\mathrm{U}_{j+1}^i)\Big] + \sum_{l=j+1}^{k} R_l + \Gamma_n(\lambda_n,\tau_{jn}),
\end{aligned} \tag{4.31}$$

where $\Gamma_n(\lambda_n)$, $\Gamma_n(\lambda_n,\tau_{jn})$, $\ddot{\mathrm{M}}_{j+1}$ and U_j^i are as defined in Section 4.2.1. Step (a) follows from Fano's inequality; (b) follows from the individual secrecy constraint in (4.30c); (c) follows due to the chain rule of mutual information which implies that $\mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n) \geq \mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n|\ddot{\mathrm{M}}_{j+1}) - \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n|\mathrm{M}_j)$; (d) follows because the messages are uniformly distributed over their respective sets, which implies that $n\sum_{l=j+1}^{k} R_l \geq \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n|\mathrm{M}_j)$; while (e) follows by applying the same steps in (4.19).

Next, if we only consider the reliability condition in (4.30b), we can derive another bound for the rate intended to each legitimate receiver as follows:

$$\begin{aligned}
R_j &\overset{(a)}{\leq} \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n|\ddot{\mathrm{M}}_{j+1})\Big] + \Gamma_n(\lambda_n) \\
&\overset{(b)}{\leq} \frac{1}{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n|\ddot{\mathrm{M}}_{j+1}) + \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n)\Big] + \Gamma_n(\lambda_n) \\
&= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_{ji}|\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1}) + \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}_i|\tilde{\mathrm{Z}}^{i+1})\Big] + \Gamma_n(\lambda_n) \\
&\overset{(c)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) + \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}_i|\tilde{\mathrm{Z}}^{i+1}) + \mathbb{I}(\tilde{\mathrm{Z}}^{i+1};\mathrm{Y}_{ji}|\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1})\Big] + \Gamma_n(\lambda_n) \\
&\overset{(d)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) + \mathbb{I}(\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1};\mathrm{Z}_i|\tilde{\mathrm{Z}}^{i+1})\Big] + \Gamma_n(\lambda_n) \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{U}_j^i;\mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) + \mathbb{I}(\mathrm{U}_{j+1}^i;\mathrm{Z}_i)\Big] + \Gamma_n(\lambda_n),
\end{aligned} \tag{4.32}$$

where (a) follows from Fano's inequality in addition to the independence of the messages; (b) follows due to the positivity of the mutual information; (c) follows from the

chain rule of the mutual information; while (d) follows due to the Csiszár sum identity [10, Lemma 7] which implies that $\mathbb{I}(\tilde{\mathrm{Z}}^{i+1};\mathrm{Y}_{ji}|\ddot{\mathrm{M}}_{j+1}\mathrm{Y}_j^{i-1}) = \mathbb{I}(\mathrm{Y}_j^{i-1};\mathrm{Z}_i|\ddot{\mathrm{M}}_{j+1}\tilde{\mathrm{Z}}^{i+1})$.

Finally, we consider the sum of all the confidential rates and derive the following bound:

$$\begin{aligned}
\sum_{l=j}^{k} R_l &\overset{(a)}{\leq} \frac{1}{n}\sum_{l=j}^{K}\mathbb{I}(\mathrm{M}_l;\mathrm{Y}_l^n|\ddot{\mathrm{M}}_{l+1}) + k\cdot\Gamma_n(\lambda_n)\\
&\overset{(b)}{\leq} \frac{1}{n}\Bigg[\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n|\ddot{\mathrm{M}}_{j+1}) + \sum_{l=j+1}^{k}\Big[\mathbb{I}(\mathrm{M}_l;\mathrm{Y}_l^n|\ddot{\mathrm{M}}_{l+1}) - \mathbb{I}(\mathrm{M}_l;\mathrm{Z}^n|\ddot{\mathrm{M}}_{l+1})\Big] + \mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n)\Bigg]\\
&\qquad + k\cdot\Gamma_n(\lambda_n)\\
&\overset{(c)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\Bigg[\mathbb{I}(\mathrm{U}_j^i;\mathrm{Y}_{ji}|\mathrm{U}_{j+1}^i) + \sum_{l=j+1}^{k}\Big[\mathbb{I}(\mathrm{U}_l^i;\mathrm{Y}_{li}|\mathrm{U}_{l+1}^i) - \mathbb{I}(\mathrm{U}_l^i;\mathrm{Z}_i|\mathrm{U}_{l+1}^i)\Big] + \mathbb{I}(\mathrm{U}_{j+1}^i;\mathrm{Z}_i)\Bigg]\\
&\qquad + k\cdot\Gamma_n(\lambda_n)\\
&\overset{(d)}{=} \frac{1}{n}\sum_{i=1}^{n}\sum_{l=j}^{k}\mathbb{I}(\mathrm{U}_l^i;\mathrm{Y}_{li}|\mathrm{U}_{l+1}^i) + k\cdot\Gamma_n(\lambda_n)
\end{aligned} \tag{4.33}$$

where (a) follows from Fano's inequality in addition to the independence of the messages; (b) follows due to the chain rule of the mutual information which implies that $\mathbb{I}(\ddot{\mathrm{M}}_{j+1};\mathrm{Z}^n) = \sum_{l=j+1}^{k}\mathbb{I}(\mathrm{M}_l;\mathrm{Z}^n|\ddot{\mathrm{M}}_{l+1})$; (c) follows by applying the same steps in (4.19) and (4.32); while (d) follows again due to the mutual information chain rule which implies that $\sum_{l=j+1}^{k}\mathbb{I}(\mathrm{U}_l^i;\mathrm{Z}_i|\mathrm{U}_{l+1}^i) = \mathbb{I}(\mathrm{U}_{j+1}^i;\mathrm{Z}_i)$.

Next, we introduce a random variable T independent of all others and uniformly distributed over $[\![1;n]\!]$ to the rate bounds in (4.31), (4.32) and (4.33) then let $\mathrm{U}_j = (\mathrm{U}_j^{\mathrm{T}},\mathrm{T})$, $\mathrm{Y}_j = \mathrm{Y}_{j\mathrm{T}}$ and $\mathrm{Z} = \mathrm{Z}_{\mathrm{T}}$ for all $j \in [\![1,k]\!]$. Finally, if we take the limit as $n\to\infty$, which implies that $\Gamma_n(\lambda_n,\tau_{jn}),\Gamma_n(\lambda_n)\to 0$, our converse is complete. Again, we highlight that the cardinality bounds follow by the Fenchel-Bunt strengthening of the Carathéodory's theorem.

4.2.3 Eavesdropper Degradedness Order

Theorems 4.1 and 4.2 were derived for the degraded multi-receiver DM-WBC in which the eavesdropper is the weakest receiver as illustrated by the Markov chain in (4.15). In general, any degraded multi-receiver DM-WBC –whether it is physically or statistically degraded– is characterized by a Markov chain with a certain degradedness order among the legitimate receivers and the eavesdropper. This degradedness order plays an important role in establishing both the joint and individual secrecy capacity regions as highlighted in Section 4.2.1 and Section 4.2.2.

In the previous chapter particularly in Section 3.4.2, we argued that changing this order affects the individual-strong secrecy capacity region of the two-receive DM-WBC with degraded message sets and message cognition. In fact, we showed that although we can establish the individual secrecy capacity for the two extreme scenarios (eavesdropper is the weakest receiver and eavesdropper is the strongest receiver), the individual

secrecy capacity of the intermediate scenario is still unknown. Thus, it is important to investigate how changing the degradedness order of the eavesdropper affects the joint and individual secrecy capacity regions established in Theorems 4.1 and 4.2 respectively. This investigation is important because it indicates whether the capacity regions in (4.16) and (4.20) can be used to describe the capacity region of practical multi-receiver degraded DM-WBCs, such as the multi-receiver Gaussian SISO WBC.

We start by dividing the k legitimate receivers into two groups. The first group contains the legitimate receivers degraded from the eavesdropper, i.e. the eavesdropper is stronger than those receivers. This group contains the legitimate receivers numbered from d to k, where $1 \leq d \leq k$. On the other hand, the second group contains the remaining legitimate receivers from which the eavesdropper is degraded. The legitimate receivers of this group are numbered from 1 to $d-1$.

1. Joint-Strong Secrecy Capacity Region (Theorem 4.1):

Although this region was established under the condition of having the eavesdropper as the weakest receiver, it can be shown that it is valid for the other scenarios as well.

• *Achievability:* For a general degradedness order, the randomization rate needed to confuse the eavesdropper is bigger than the decoding capability of the legitimate receivers of the first group. In other words, for $j \in [\![d;k]\!]$, we have

$$\mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) \leq \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}). \tag{4.34}$$

Thus, it follows that the achievable joint-strong secrecy rates for the legitimate receivers in this group vanish. Contrarily, nothing changes for the legitimate receivers that belong to the second group of receivers. These two arguments imply that the achievable rate region in (4.16) simplifies to:

$$0 \leq R_j \leq \begin{cases} 0 & \text{for } j \in [\![d;k]\!] \\ \mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}) & \text{for } j \in [\![1;d-1]\!]. \end{cases}$$

• *Converse:* It was shown in [16, Proposition 3.4] that the joint secrecy capacity vanishes if the legitimate receiver is degraded from the eavesdropper. In Section 3.4.2, we extended this result to the less noisy case. This implies that, the confidential rate of any legitimate receiver that belongs to the first group is upper bounded by zero. On the other hand, the converse proof for the confidential rates of the legitimate receivers that belong to the second group follows exactly as in (4.18) and (4.18), where k is replaced by $d-1$.

2. Individual-Strong Secrecy Capacity Region (Theorem 4.2):

Like the joint secrecy case, this region was established under the condition of having the eavesdropper as the weakest receiver. The main challenge is that, based on the results established in the previous chapter, the optimal coding scheme for the degraded two-receiver DM-WBC with message cognition under the individual secrecy criterion varies with the degradedness order of the eavesdropper. However, we show that this result is a unique feature for the degraded two-receiver DM-WBC with message cognition and

does not apply to the current scenario. In doing so, we show that the individual-strong secrecy capacity region in (4.20) is valid for all degraded multi-receiver DM-WBC regardless of the degradedness order of the eavesdropper.

• *Achievability:* Since for $j \in [\![d;k]\!]$, the condition in (4.34) is a property of the channel and is independent of the investigated secrecy criterion, then it follows from (4.20a) that the individual-strong secrecy rates intended for the legitimate receivers that belong to the first group vanish, i.e. $R_j = 0$ for $j \in [\![d;k]\!]$. On the other hand, the legitimate receivers of the second group along with the eavesdropper can be interpreted as a degraded d-receiver DM-WBC where the eavesdropper is the weakest receiver. Thus, for $j \in [\![1;d-1]\!]$, the following region is achievable:

$$\begin{aligned} R_j &\leq \mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}) + \sum_{l=j+1}^{d-1} R_l \\ R_j &\leq \mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_{j+1};\mathrm{Z}) \\ \sum_{l=j}^{d-1} R_l &\leq \sum_{l=j}^{d-1} \mathbb{I}(\mathrm{U}_l;\mathrm{Y}_l|\mathrm{U}_{l+1}). \end{aligned}$$

• *Converse:* We start by the legitimate receivers which belong to the first group. The properties of degraded channel implies that if Y is degraded from Z, then Z is also less noisy than Y. Based on Lemma A.8 and the previous argument, it follows that for $j \in [\![d,k]\!]$, $\mathbb{I}(\mathrm{M}_j;\mathrm{Y}_j^n) \leq \mathbb{I}(\mathrm{M}_j;\mathrm{Z}^n)$. Since the subtraction of these two terms is the first step in deriving an upper bound on R_j as shown in Eq. (4.31), then it follows that R_j is upper-bounded by zero for $j \in [\![d,k]\!]$. On the other hand, the converse of the confidential rates of the legitimate receivers of the second group follows exactly as in Eqs. (4.31), (4.32) and (4.33), where k is replaced by $d-1$.

The previous two arguments assure that Theorems 4.1 and 4.2 establish the characterization of the joint and individual strong secrecy capacity regions for any degraded multi-receiver DM-WBC regardless of the degradedness order of the eavesdropper.

4.2.4 Multi-Receiver Gaussian SISO WBC

The multi-receiver Gaussian SISO WBC is one of the most famous examples for a degraded multi-receiver DM-WBC. The channel model follows directly from the model of the two-receiver Gaussian SISO WBC introduced in Section 2.4.7 as follows:

$$\mathrm{Y}_j = \mathrm{X} + \mathrm{N}_j \qquad \mathrm{Z} = \mathrm{X} + \mathrm{N}_Z, \tag{4.35}$$

where $j \in [\![1,k]\!]$. The channel input X is subject to a power constraint, such that $\mathbb{E}[\mathrm{X}^2] \leq P$. On the other hand, the channel noises $(\mathrm{N}_1,\ldots,\mathrm{N}_k)$ and N_Z are zero-mean Gaussian random variables, whose variances are given by $(\sigma_1^2,\ldots,\sigma_k^2)$ and σ_Z^2 respectively. The relation between these variances defines the degradedness order of the channel in the following sense:

$$\begin{aligned} &\text{if} \quad \sigma_1^2 \leq \cdots \leq \sigma_k^2 \leq \sigma_Z^2, \\ &\text{then} \quad \mathrm{X} - \mathrm{Y}_1 - \cdots - \mathrm{Y}_k - \mathrm{Z} \quad \text{forms a Markov chain.} \end{aligned} \tag{4.36}$$

Based on the results established in the previous subsection, we can use Theorem 4.1 and Theorem 4.2 to evaluate the joint and individual strong secrecy capacity region of the multi-receiver Gaussian SISO WBC for any order among the noises' variances. To do so, we only need to find the optimal joint distribution on the random variables $(\mathrm{X}, \mathrm{U}_2, \ldots, \mathrm{U}_k)$ that trace the boundaries of the capacity regions in (4.16) and (4.20). With this in mind, we present the following result:

Theorem 4.3. *The joint-strong secrecy capacity region of the multi-receiver Gaussian SISO WBC is given by the union over the set of all rate tuples* $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *that satisfy*

$$R_j \leq f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_Z^2}\right). \tag{4.37}$$

On the other hand, the individual-strong secrecy capacity region of the same channel is given by the union over the set of all rate tuples $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *that satisfy*

$$R_j \leq f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_Z^2}\right) + \sum_{l=j+1}^{k} R_l \tag{4.38a}$$

$$R_j \leq f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_j^2}\right) + f\left(\frac{\sum_{i=j+1}^{k} \alpha_i P}{\sum_{i=1}^{j} \alpha_i P + \sigma_Z^2}\right) \tag{4.38b}$$

$$\sum_{l=j}^{k} R_l \leq \sum_{l=j}^{k} f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1} \alpha_i P + \sigma_l^2}\right) \tag{4.38c}$$

where $j \in [\![1, K]\!]$ *and the unions are taken over all values of* $\alpha_j \in [0, 1]$ *such that* $\sum_{i=1}^{k} \alpha_i \leq 1$. *Moreover, the function* $f(a)$ *is defined as follows:* $f(a) \triangleq \frac{1}{2}\log(1 + a)$.

The achievability of the two capacity regions in Theorem 4.3 follows by choosing the random variables $(\mathrm{U}_k, \ldots, \mathrm{U}_2, \mathrm{X})$ in (4.16) and (4.20) respectively to be jointly Gaussian. In particular, for every $j \in [\![1, k]\!]$, we let $\mathrm{U}_j = \mathrm{U}_{j+1} + \mathrm{V}_j$, where $\mathrm{U}_{k+1} = 0$ and V_j is an independent Gaussian random variable with variance α_j. The decoder at a certain receiver Y_i, where $i \in [\![1, k]\!]$, can decode all the random variables V_j for $j \geq i$ because of the order of the variances in (4.36). Additionally the remaining V_j for $j < i$ are interpreted as interfering noise. This scheme will directly lead to the achievability of the rates in (4.37) and (4.38).

Now, for the converse, we only focus on the individual secrecy case as the joint secrecy converse can be established using the same steps. We start with the bound in (4.38a) and consider the kth receiver. We have

$$\begin{aligned}
R_k &\leq \mathbb{I}(\mathrm{U}_k; \mathrm{Y}_k) - \mathbb{I}(\mathrm{U}_k; \mathrm{Z}) \\
&\overset{(a)}{=} \Big[\mathbb{I}(\mathrm{X}; \mathrm{Y}_k) - \mathbb{I}(\mathrm{X}; \mathrm{Z})\Big] - \Big[\mathbb{I}(\mathrm{X}; \mathrm{Y}_k|\mathrm{U}_k) - \mathbb{I}(\mathrm{X}; \mathrm{Z}|\mathrm{U}_k)\Big] \\
&\overset{(b)}{\leq} \left[f\left(\frac{P}{\sigma_k^2}\right) - f\left(\frac{P}{\sigma_Z^2}\right)\right] - \Big[\mathbb{I}(\mathrm{X}; \mathrm{Y}_k|\mathrm{U}_k) - \mathbb{I}(\mathrm{X}; \mathrm{Z}|\mathrm{U}_k)\Big] \\
&\overset{(c)}{=} \left[f\left(\frac{P}{\sigma_k^2}\right) - f\left(\frac{P}{\sigma_Z^2}\right)\right] - \left[f\left(\frac{\bar{\alpha}_k P}{\sigma_k^2}\right) - f\left(\frac{\bar{\alpha}_k P}{\sigma_Z^2}\right)\right] \\
&\overset{(d)}{=} f\left(\frac{\alpha_k P}{\sum_{i=1}^{k-1} \alpha_i P + \sigma_k^2}\right) - f\left(\frac{\alpha_k P}{\sum_{i=1}^{k-1} \alpha_i P + \sigma_Z^2}\right),
\end{aligned} \tag{4.39}$$

where (a) follows by using the chain rule along with the fact that $\mathrm{U}_k - \mathrm{X} - (\mathrm{Y}_k, \mathrm{Z})$ forms a Markov chain; (b) follows because for a Gaussian SISO WBC the term $\mathbb{I}(\mathrm{X};\mathrm{Y}_k) - \mathbb{I}(\mathrm{X};\mathrm{Z})$ is maximized by a Gaussian signaling for the random variable X [11]; while (c) follows because

$$0 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_k|\mathrm{U}_k) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_k) \leq f\left(\frac{P}{\sigma_k^2}\right) - f\left(\frac{P}{\sigma_Z^2}\right).$$

This implies that for any pair $(\mathrm{U}_k, \mathrm{X})$, there exists an $\bar{\alpha}_k \in [0,1]$, such that the equality in (c) holds. On the other hand, (d) follows by letting $\alpha_k = 1 - \bar{\alpha}_k$ and $\bar{\alpha}_k = \sum_{i=1}^{k-1} \alpha_i$. Now, we consider the $(k-1)^{th}$ receiver under the same bound, we have

$$\begin{aligned}
R_{k-1} &\leq \mathbb{I}(\mathrm{U}_{k-1};\mathrm{Y}_{k-1}|\mathrm{U}_k) - \mathbb{I}(\mathrm{U}_{k-1};\mathrm{Z}|\mathrm{U}_k) + R_k \\
&\overset{(a)}{=} \Big[\mathbb{I}(\mathrm{X};\mathrm{Y}_{k-1}|\mathrm{U}_k) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_k)\Big] - \Big[\mathbb{I}(\mathrm{X};\mathrm{Y}_{k-1}|\mathrm{U}_{k-1}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_{k-1})\Big] + R_k \\
&\overset{(b)}{\leq} \left[f\left(\frac{\bar{\alpha}_k P}{\sigma_{k-1}^2}\right) - f\left(\frac{\bar{\alpha}_k P}{\sigma_Z^2}\right)\right] - \Big[\mathbb{I}(\mathrm{X};\mathrm{Y}_{k-1}|\mathrm{U}_{k-1}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_{k-1})\Big] + R_k \\
&\overset{(c)}{=} \left[f\left(\frac{\bar{\alpha}_k P}{\sigma_{k-1}^2}\right) - f\left(\frac{\bar{\alpha}_k P}{\sigma_Z^2}\right)\right] - \left[f\left(\frac{\bar{\alpha}_{k-1} P}{\sigma_k^2}\right) - f\left(\frac{\bar{\alpha}_{k-1} P}{\sigma_Z^2}\right)\right] + R_k \\
&\overset{(d)}{=} f\left(\frac{\alpha_{k-1} P}{\sum_{i=1}^{k-2} \alpha_i P + \sigma_k^2}\right) - f\left(\frac{\alpha_{k-1} P}{\sum_{i=1}^{k-2} \alpha_i P + \sigma_Z^2}\right) + R_k,
\end{aligned} \tag{4.40}$$

where (a) follows from the mutual information chain rule and the fact that $\mathrm{U}_k - \mathrm{U}_{k-1} - \mathrm{X} - (\mathrm{Y}_{k-1}, \mathrm{Z})$ forms a Markov chain; (b) follows because from step (c) in Eq. (4.39), we have:

$$\mathbb{I}(\mathrm{X};\mathrm{Y}_k|\mathrm{U}_k) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_k) = f\left(\frac{\bar{\alpha}_k P}{\sigma_k^2}\right) - f\left(\frac{\bar{\alpha}_k P}{\sigma_Z^2}\right).$$

We know from [94] that under this condition, the expression $\mathbb{I}(\mathrm{X};\mathrm{Y}_{k-1}|\mathrm{U}_k) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}_k)$ is maximized by a joint Gaussian distribution on the pair $(\mathrm{U}_k, \mathrm{X})$; (c) follows based on the same arguments used to establish step (c) in Eq. (4.39); while (d) follows by letting $\alpha_{k-1} = \bar{\alpha}_k - \bar{\alpha}_{k-1}$ and $\bar{\alpha}_{k-1} = \sum_{i=1}^{k-2} \alpha_i$.

Now, if we apply the same steps used to establish (4.40) to the remaining legitimate receivers, we can show that the bound in (4.38a) holds. These calculations establish two additional constraints: the first is $\sum_{i=1}^{k} \alpha_i = 1$, while the second is the following bound:

$$\mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}) \leq f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1} \alpha_i P + \sigma_Z^2}\right). \tag{4.41}$$

The two previous constraints along with Lemma A.9 directly imply that the following power constraint: $\mathbb{E}[\mathrm{U}_j^2] \leq \sum_{i=j}^{k} \alpha_i P$ is true for all $j \in [\![1, k]\!]$. We now focus on the rate constraint in (4.38b) and derive it as follows:

$$\begin{aligned}
R_j &\leq \mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_{j+1};\mathrm{Z}) \\
&\overset{(a)}{=} \mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_j;\mathrm{Z})
\end{aligned}$$

$$
\begin{aligned}
&\overset{(b)}{\leq} f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_Z^2}\right) + \mathbb{I}(\mathrm{U}_j;\mathrm{Z}) \\
&\overset{(c)}{\leq} f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_Z^2}\right) + f\left(\frac{\sum_{i=j}^{k}\alpha_i P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_Z^2}\right) \\
&= f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_j^2}\right) + f\left(\frac{\sum_{i=j+1}^{k}\alpha_i P}{\sum_{i=1}^{j}\alpha_i P+\sigma_Z^2}\right),
\end{aligned}
\tag{4.42}
$$

where (a) follows from the the chain rule of the mutual information and the fact that $\mathrm{U}_{j+1} - \mathrm{U}_j - \mathrm{Z}$ forms a Markov chain; (b) follows based on the same arguments used to establish the rate constraint in (4.38a); while (c) follows because under the power constraints on U_j and X, in addition to the following Markov chain $\mathrm{U}_j - \mathrm{X} - \mathrm{Z}$, the mutual information term $\mathbb{I}(\mathrm{U}_j;\mathrm{Z})$ is maximized by a joint Gaussian distribution on the pair $(\mathrm{U}_j, \mathrm{X})$. Finally, we establish the rate constraint in (4.38c) as follows:

$$
\begin{aligned}
\sum_{l=j}^{k} R_l &\leq \sum_{l=j}^{k} \mathbb{I}(\mathrm{U}_l;\mathrm{Y}_l|\mathrm{U}_{l+1}) \\
&\overset{(a)}{=} \sum_{l=j}^{k}\Big[\mathbb{I}(\mathrm{U}_l;\mathrm{Y}_l|\mathrm{U}_{l+1}) - \mathbb{I}(\mathrm{U}_l;\mathrm{Z}|\mathrm{U}_{l+1})\Big] + \mathbb{I}(\mathrm{U}_j;\mathrm{Z}) \\
&\overset{(b)}{\leq} \sum_{l=j}^{k}\left[f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1}\alpha_i P+\sigma_l^2}\right) - f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1}\alpha_i P+\sigma_Z^2}\right)\right] + \mathbb{I}(\mathrm{U}_j;\mathrm{Z}) \\
&\overset{(c)}{\leq} \sum_{l=j}^{k}\left[f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1}\alpha_i P+\sigma_l^2}\right) - f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1}\alpha_i P+\sigma_Z^2}\right)\right] + f\left(\frac{\sum_{i=j}^{k}\alpha_i P}{\sum_{i=1}^{j-1}\alpha_i P+\sigma_Z^2}\right) \\
&= \sum_{l=j}^{k} f\left(\frac{\alpha_l P}{\sum_{i=1}^{l-1}\alpha_i P+\sigma_l^2}\right),
\end{aligned}
\tag{4.43}
$$

where (a) follows by using the chain rule of the mutual information and the fact that $\mathrm{U}_k - \mathrm{U}_{k-1} - \cdots - \mathrm{U}_j - \mathrm{Z}$ forms a Markov chain; while (b) and (c) follows based on the same arguments used to derive Eq. (4.42). This completes our converse proof.

It is important to point out that comparing the joint-strong secrecy capacity region established in Theorem 4.3 and the joint-weak secrecy capacity established in [94] implies that for the multi-receiver Gaussian SISO WBC strengthening the secrecy criteria from weak secrecy to strong secrecy comes at no cost with respect to the capacity region. However, our results implies that this does not hold, when moving from the strict constraints of the joint secrecy to the loose constraints of individual secrecy.

4.2.5 Multi-Receiver Degraded Gaussian MIMO WBC

In this section, we extend the secrecy capacity results established for the Gaussian SISO WBC to the class of degraded Gaussian MIMO WBC. This class is also a practical example for degraded DM-WBCs. The channel model follows directly from the model of the two-receiver Gaussian MIMO WBC introduced in Section 2.4.7 as follows:

$$
\mathbf{Y}_j = \mathbf{X} + \mathbf{N}_j \qquad \mathbf{Z} = \mathbf{X} + \mathbf{N}_Z, \tag{4.44}
$$

for $j \in [\![1,k]\!]$. The channel input $\mathbf{X}$ is subject to a covariance constraint, such that: $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S}$, where $\mathbf{S} \succ \mathbf{0}$. On the other hand, the noise vectors $(\mathbf{N}_1, \ldots, \mathbf{N}_k, \mathbf{N}_Z)$ are zero-mean Gaussian random vectors, whose covariance matrices are given by $(\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_k, \mathbf{\Sigma}_Z)$. Similar to the Gaussian SISO WBC, the relation between these covariance matrices defines the degradedness order of the channel as follows:

$$\begin{aligned} &\text{if} \quad \mathbf{0} \prec \mathbf{\Sigma}_1 \preceq \mathbf{\Sigma}_2 \preceq \cdots \preceq \mathbf{\Sigma}_k \preceq \mathbf{\Sigma}_Z, \\ &\text{then} \quad \mathrm{X} - \mathrm{Y}_1 - \cdots - \mathrm{Y}_k - \mathrm{Z} \quad \text{forms a Markov chain.} \end{aligned} \tag{4.45}$$

Although it is more common in literature and practically meaningful to consider a sum power constraint, we use a covariance constraint instead because it eases the derivation of the capacity region. Nevertheless, it was shown in [102], that once the capacity region is obtained under a covariance constraint, the capacity region under the sum power constraint follows easily. Now, based on the results established in Section 4.2.3 along with the secrecy capacity regions given in Theorem 4.1 and 4.2, we present the following coding theorem.

Theorem 4.4. *The joint-strong secrecy capacity region of the degraded multi-receiver Gaussian MIMO WBC is given by the union over the set of all rate tuples* $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *that satisfy*

$$R_j \leq \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i + \mathbf{\Sigma}_j\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|}. \tag{4.46}$$

On the other hand, the individual-strong secrecy capacity region of the same channel is given by the union over the set of all rate tuples $(R_1, \ldots, R_k) \in \mathbb{R}_+^k$ *that satisfy*

$$R_j \leq \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i + \mathbf{\Sigma}_j\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|} + \sum_{l=j+1}^{k} R_l \tag{4.47a}$$

$$R_j \leq \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i + \mathbf{\Sigma}_j\right|} + \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j}\mathbf{K}_i + \mathbf{\Sigma}_Z\right|} \tag{4.47b}$$

$$\sum_{l=j}^{k} R_l \leq \sum_{l=j}^{k}\frac{1}{2}\log\frac{\left|\sum_{i=1}^{l}\mathbf{K}_i + \mathbf{\Sigma}_l\right|}{\left|\sum_{i=1}^{l-1}\mathbf{K}_i + \mathbf{\Sigma}_l\right|} \tag{4.47c}$$

where $j \in [\![1,k]\!]$ *and the unions are taken over all possible positive semi-definite matrices* $\mathbf{K}_j \succeq \mathbf{0}$*, such that* $\sum_{i=1}^{k}\mathbf{K}_i \preceq \mathbf{S}$*. Moreover, the function* $|\mathbf{A}|$ *is simply the determinant of the matrix* $\mathbf{A}$*.*

The two secrecy capacity regions in Theorem 4.4 can be achieved using a Gaussian random vector realization for the auxiliary random variables $(\mathrm{U}_k, \ldots, \mathrm{U}_2, \mathrm{X})$ in Theorem 4.1 and Theorem 4.2 respectively. The vectors are constructed recursively as follows: $\mathbf{U}_j = \mathbf{U}_{j+1} + \mathbf{V}_j$, where $\mathbf{V}_j$ is a Gaussian random vector with covariance matrices $\mathbf{K}_j$ and is independent of $\mathbf{U}_{j+1}$. Moreover, $\mathbf{U}_{k+1}$ is assumed to be a zero vector. This means that for every $j \in [\![1,k]\!]$, the vector $\mathbf{V}_j$ contains information about the confidential message M_j. The decoder at a certain receiver $\mathbf{Y}_i$ can decode all vectors

$\mathbf{V}_j$ for $j \geq i$, while handling the remaining $\mathbf{V}_j$ for $j < i$ as noise. This follows due to the order of the covariance matrices of the noise vectors in (4.45).

Now, for the converse, we only focus on the individual secrecy case as the joint secrecy converse can be established by applying similar steps. We start by highlighting an important upper-bound established in [94] for the multi-receiver degraded Gaussian MIMO WBC as follows:

$$\mathbb{I}(\mathrm{U}_j;\mathbf{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathbf{Z}|\mathrm{U}_{j+1}) \leq \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_j\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}, \tag{4.48}$$

where $\mathrm{U}_k - \cdots - \mathrm{U}_2 - \mathbf{X} - \mathbf{Y}_1 - \cdots - \mathbf{Y}_k - \mathbf{Z}$ forms a Markov chain and $\mathbf{K}_j \succeq \mathbf{0}$ are positive semi-definite matrices, such that $\sum_{i=1}^{k}\mathbf{K}_i = \mathbf{S}$, where $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S}$. The establishment of the previous bound uses similar steps to the one used to establish the bound in (4.41) for the Gaussian SISO case, along with the properties of Fisher information matrix. Eq. 4.48 directly implies that the bound in (4.20a) simplifies to the one in (4.47a) for the multi-receiver degraded Gaussian MIMO WBC. Thus, we only need to derive the bounds in (4.47b) and (4.47c). From Eq. (4.20b), we have

$$\begin{aligned}
R_j &\overset{(a)}{\leq} \mathbb{I}(\mathrm{U}_j;\mathbf{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathbf{Z}|\mathrm{U}_{j+1}) + \mathbb{I}(\mathrm{U}_j;\mathbf{Z}) \\
&\overset{(b)}{\leq} \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_j\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|} + \mathbb{I}(\mathbf{U}_j;\mathbf{Z}) \\
&\overset{(c)}{\leq} \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_j\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|} + \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|} \\
&= \frac{1}{2}\log\frac{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_j\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_j\right|} + \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|},
\end{aligned} \tag{4.49}$$

where (a) follows from the the chain rule of the mutual information and the fact that $\mathrm{U}_{j+1} - \mathrm{U}_j - \mathbf{Z}$ forms a Markov chain; (b) follows from (4.48) and the fact that for a degraded Gaussian MIMO BC with the following Markov chain $\mathrm{U}_j - \mathbf{X} - \mathbf{Z}$, the expression $\mathbb{I}(\mathrm{U}_j;\mathbf{Z})$ is maximized by a vector realization $\mathbf{U}_j$ for the auxiliary random variable U_j, such that $\mathbf{U}_j$ and $\mathbf{X}$ are jointly Gaussian. On the other hand, (c) follows from the covariance constraint on $\mathbf{U}_j$ enforced by Lemma A.10. Now, it only remains to derive the bound in (4.47c). From Eq. (4.20c), we have

$$\begin{aligned}
\sum_{l=j}^{k} R_l &\overset{(a)}{\leq} \sum_{l=j}^{k}\Big[\mathbb{I}(\mathrm{U}_l;\mathbf{Y}_l|\mathrm{U}_{l+1}) - \mathbb{I}(\mathrm{U}_l;\mathbf{Z}|\mathrm{U}_{l+1})\Big] + \mathbb{I}(\mathrm{U}_j;\mathbf{Z}) \\
&\overset{(b)}{\leq} \sum_{l=j}^{k}\left[\frac{1}{2}\log\frac{\left|\sum_{i=1}^{l}\mathbf{K}_i+\mathbf{\Sigma}_l\right|}{\left|\sum_{i=1}^{l-1}\mathbf{K}_i+\mathbf{\Sigma}_l\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{l}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{l-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}\right] + \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|}{\left|\sum_{i=1}^{j-1}\mathbf{K}_i+\mathbf{\Sigma}_Z\right|} \\
&= \sum_{l=j}^{k}\frac{1}{2}\log\frac{\left|\sum_{i=1}^{l}\mathbf{K}_i+\mathbf{\Sigma}_l\right|}{\left|\sum_{i=1}^{l-1}\mathbf{K}_i+\mathbf{\Sigma}_l\right|},
\end{aligned} \tag{4.50}$$

where (a) follows by using the chain rule of the mutual information and the fact that $\mathrm{U}_k - \mathrm{U}_{k-1} - \cdots - \mathrm{U}_j - \mathbf{Z}$ forms a Markov chain; while (b) follows based on the same arguments used to establish the bound in (4.49). This completes our converse.

It is important to point out that the joint-strong secrecy capacity region established in Theorem 4.4 can be used to establish the joint-strong secrecy capacity region of the general multi-receiver Gaussian MIMO WBC. This can be done by following the same ideas used in [94], which shows that dirty-paper coding with stochastic encoding is the optimal coding scheme. However, this optimality can not be established under the individual secrecy criterion. This implies that the individual-strong secrecy capacity of the general multi-receiver Gaussian MIMO WBC is still unknown.

4.3 Achievable Rate Regions for the Two-Receiver DM-WBC

In this section, we limit our investigation to the problem of secure communication over a general two-receiver DM-WBC. At first, we derive a general achievable rate region that fulfills the joint-strong secrecy criterion. This region does not only strengthen the region established in [73] to the strong secrecy measure, but it can also be used to recover the joint-strong secrecy capacity region of the degraded two-receiver DM-WBC. Moreover, we modify our coding scheme to adhere to the individual-strong secrecy criterion and establish a general achievable rate region that can also be used to derive the achievability of the individual-strong secrecy capacity region of the degraded two-receiver DM-WBC. Our results have some intuition that even for the general two-receiver DM-WBC, relaxing the secrecy requirement from joint to individual can lead to a larger achievable rate region.

4.3.1 Motivation and Main Results

The general two-receiver DM-WBC was first investigated under the joint-weak secrecy constraint in [73], where an achievable rate region that utilizes the classical wiretap Marton coding was established. The main issue of the region given therein is that it fails to recover the joint-weak secrecy capacity region of the degraded two-receiver DM-WBC, where the optimal coding strategy is a simple superposition wiretap random code. This is because Marton coding is in general not optimal without a superposition variable as highlighted in [35]. This observation encouraged us to derive the following achievable rate region:

Theorem 4.5. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs* $(R_1, R_2) \in \mathbb{R}_+^2$ *that satisfy*

$$R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Z}|\mathrm{T}) \tag{4.51a}$$

$$R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}|\mathrm{T}) \tag{4.51b}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_2; \mathrm{Y}_2|\mathrm{V}_0\mathrm{T}) - R_e \tag{4.51c}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Y}_1|\mathrm{V}_0\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - R_e \tag{4.51d}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{T}) - R_e, \tag{4.51e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{T}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{T})$. *The random variables that define the previous rate region are characterized by the following joint probability distribution:* $\mathrm{Q}_{\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_0|\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$, *such that* $\mathrm{T} - \mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The achievability proof of Theorem 4.5 is based on a coding scheme that utilizes a wiretap Marton code with a superposition variable along with the principles of rate splitting and time sharing. If we let $\mathrm{V}_0 = \emptyset$, the rate region in Theorem 4.5 simplifies to the one established in [73] under the joint-weak secrecy constraint. Moreover, if we let $\mathrm{V}_0 = \mathrm{V}_2 = \mathrm{U}$ and $\mathrm{V}_1 = \mathrm{X}$, the rate region in Theorem 4.5 recover the joint-strong secrecy capacity region of the degraded two-receiver DM-WBC given in Corollary 4.1. This implies that our rate region given by (4.51) is a more general rate region.

In Section 3.4, we showed that for a class of less noisy DM-WBC with message cognition the individual-strong secrecy capacity region is bigger than the joint one. The same result was also established in Section 4.2 for the degraded, Gaussian SISO and degraded Gaussian MIMO multi-receiver WBC. These two observations might suggest that the individual-strong secrecy criterion can in general provide a larger rate region compared to the joint-strong secrecy criterion. On the other hand, one might claim that the increase in the individual secrecy capacity region is a unique feature for these two scenarios. Specially, because for the DM-WBC with message cognition the increase in the individual secrecy capacity originated from the availability of the receiver side information. Similarly, the larger individual capacity regions presented in Section 4.2 are due to the statistical advantage possessed by the legitimate receivers over each other. Thus, in order to show that individual secrecy can outperform joint secrecy for a DM-WBC, where neither receiver side information nor a certain order among the legitimate receivers decoding capabilities exists, we present the following rate region:

Theorem 4.6. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs* $(R_1 = R_{11} + R_{12}, R_2 = R_{21} + R_{22}) \in \mathbb{R}_+^2$ *that satisfy*

$$R_1 + R_{21} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{T}) + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}|\mathrm{T}), R_{21}\right] \tag{4.52a}$$

$$R_2 + R_{11} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{T}) + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}|\mathrm{T}), R_{11}\right] \tag{4.52b}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_0\mathrm{T}) - R_e + R_s \tag{4.52c}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_0\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{T}) - R_e + R_s \tag{4.52d}$$

$$R_1 + R_2 + R_{11} + R_{21} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}|\mathrm{T}) - R_e + R_s, \tag{4.52e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{T}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{T})$, *while* $R_s \triangleq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}|\mathrm{T}), R_{11} + R_{21}\right]$. *The random variables that define the previous rate region are characterized by the following joint probability distribution:* $\mathrm{Q}_{\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_0|\mathrm{T}}\ \mathrm{Q}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$, *such that the following Markov chain holds:* $\mathrm{T} - \mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$.

Besides the coding techniques used to prove Theorem 4.5, the achievability proof of Theorem 4.6 utilizes the concept of secret key encoding as well. This is because under the individual secrecy criterion, the legitimate receivers trust each other which allows the usage of the whole or a part of the confidential message of one receiver as a secret key for the other one. However, in order for this coding technique to work,

each receiver must find a way to acquire a full knowledge about the part of the other receiver's confidential message used as a secret key. This can be done by possessing a prior knowledge about the second user's message as in the DM-WBC with message cognition discussed in Chapter 3 or by obtaining this information from its channel observation as in our case. The detailed achievability proof of this theorem is presented in Section 4.3.3.

It is important to point out that the rate region in Theorem 4.6 recovers the individual-strong secrecy capacity region of the degraded two-receiver DM-WBC given by Corollary 4.2. In order to show this, we start by letting $R_{11} = 0$, $R_{21} = R_2$, $\mathrm{V}_0 = \mathrm{V}_2 = \mathrm{U}$, $\mathrm{V}_1 = \mathrm{X}$ and $\mathrm{T} = \emptyset$. We then observe that under these substitutions, the rate region in (4.52) simplifies to:

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) - \mathbb{I}(\mathrm{X};\mathrm{Z}) + \min\left[\mathbb{I}(\mathrm{X};\mathrm{Z}), R_2\right] \tag{4.53a}$$
$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.53b}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{X};\mathrm{Z}) + \min\left[\mathbb{I}(\mathrm{X};\mathrm{Z}), R_2\right] \tag{4.53c}$$
$$R_1 + 2R_2 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) + \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) - \mathbb{I}(\mathrm{X};\mathrm{Z}) + \min\left[\mathbb{I}(\mathrm{X};\mathrm{Z}), R_2\right], \tag{4.53d}$$

where the rate constraints in (4.52a) and (4.52c) simplify to the one in (4.53a). Moreover, we notice that under these substitutions, the rate constraint in (4.53d) is simply the addition of (4.53a) and (4.53b). Next, we expand each bound in (4.53) that contain a **minimization** into two bounds and make use of the rate constraint on R_2 given by (4.53b) to modify all the bounds on the sum rate $(R_1 + R_2)$ to a bound on R_1 only. By doing so, we reach the following:

$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) - \mathbb{I}(\mathrm{U};\mathrm{Y}_2) + \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.54a}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1) - \mathbb{I}(\mathrm{X};\mathrm{Z}) \tag{4.54b}$$
$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.54c}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.54d}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{X};\mathrm{Z}). \tag{4.54e}$$

We finally use the fact that we are only interested in the degraded two-receiver WBC, i.e. $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2 - \mathrm{Z}$ forms a Markov chain. This implies that $\mathbb{I}(\mathrm{U};\mathrm{Y}_1) \geq \mathbb{I}(\mathrm{U};\mathrm{Y}_2)$. Using this relation along with the fact that $\mathbb{I}(\mathrm{X};\mathrm{Y}_1) = \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Y}_1)$, we can show that the bound in (4.54e) is tighter than the one in (4.54b). This is also true for the bound in (4.54d) and (4.54a). Thus, the rate region in (4.54) simplifies to:

$$R_2 \leq \mathbb{I}(\mathrm{U};\mathrm{Y}_2) - \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.55a}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Z}) \tag{4.55b}$$
$$R_1 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + R_2, \tag{4.55c}$$

where the bound in (4.55c) follows by using the fact that $\mathbb{I}(\mathrm{X};\mathrm{Z}) = \mathbb{I}(\mathrm{X};\mathrm{Z}|\mathrm{U}) + \mathbb{I}(\mathrm{U};\mathrm{Z})$ along with the bound on R_2 in Eq. (4.55b). The rate region in (4.55) is identical to the individual-strong secrecy capacity region of the degraded two-receiver DM-WBC given by Corollary 4.2.

One of the motivations that encouraged us to derive the achievable secrecy rate region in Theorem 4.6 is to compare it with the one in Theorem 4.5. The aim of this comparison is to investigate whether individual secrecy can outperform the joint secrecy or not

for the general two-receiver DM-WBC. However, it is not easy to answer this question using a direct comparison between the rate regions in Theorems 4.5 and 4.6. Thus, we only consider a special case, where the input distribution $\mathrm{Q}_{\mathrm{TV_0V_1V_2X}}$ has the following properties: $\mathrm{T} = \emptyset$, $R_{11} \leq \mathbb{I}(\mathrm{V_0V_1;Z})$, $R_{21} \leq \mathbb{I}(\mathrm{V_0V_2;Z})$ and $R_{11} + R_{21} \leq \mathbb{I}(\mathrm{V_0V_1V_2;Z})$. For such input distribution the achievable joint-strong secrecy rate region in (4.51) simplifies to:

$$R_1 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) - \mathbb{I}(\mathrm{V_0V_1;Z}) \tag{4.56a}$$
$$R_2 \leq \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0V_2;Z}) \tag{4.56b}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) + \mathbb{I}(\mathrm{V_2;Y_2|V_0}) - R_e \tag{4.56c}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_1;Y_1|V_0}) + \mathbb{I}(\mathrm{V_0V_2;Y_2}) - R_e \tag{4.56d}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) + \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0;Z}) - R_e, \tag{4.56e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V_0V_1V_2;Z}) + \mathbb{I}(\mathrm{V_1;V_2|V_0})$. On the other hand, under the same assumptions on the input distribution, the achievable individual-strong secrecy rate region in (4.52) simplifies to:

$$R_1 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) - \mathbb{I}(\mathrm{V_0V_1;Z}) \tag{4.57a}$$
$$R_2 \leq \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0V_2;Z}) \tag{4.57b}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) + \mathbb{I}(\mathrm{V_2;Y_2|V_0}) - R_e + R_{11} + R_{21} \tag{4.57c}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_1;Y_1|V_0}) + \mathbb{I}(\mathrm{V_0V_2;Y_2}) - R_e + R_{11} + R_{21} \tag{4.57d}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) + \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0;Z}) - R_e. \tag{4.57e}$$

It is obvious that the joint secrecy rate constraints in (4.56a), (4.56b) and (4.56e) are identical to the individual ones in (4.57a), (4.57b) and (4.57e). However, the joint secrecy rate constraints in (4.56c) and (4.56d) are tighter than their corresponding individual ones in (4.57c) and (4.57d), unless $R_{11} = R_{21} = 0$. This implies that there exists some scenarios where even for the general two-receiver DM-WBC the individual secrecy rate region can be bigger than the joint one.

4.3.2 Wiretap Marton-Coding with a Superposition Variable

In this section, we present the coding scheme required to establish the achievable rate region given by Theorem 4.5. Our coding scheme utilizes the same approach used in Chapter 3 to prove the validity of Theorem 3.1. Thus, in order to prove the achievability of the rate region in (4.51), we start by validating the following proposition:

Proposition 4.1. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs* $(R_1, R_2) \in R_+^2$ *that satisfy*

$$R_1 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) - \mathbb{I}(\mathrm{V_0V_1;Z}) \tag{4.58a}$$
$$R_2 \leq \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0;Z}) - \mathbb{I}(\mathrm{V_2;V_1Z|V_0}) \tag{4.58b}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_0V_1;Y_1}) + \mathbb{I}(\mathrm{V_2;Y_2|V_0}) - \mathbb{I}(\mathrm{V_0V_1V_2;Z}) - \mathbb{I}(\mathrm{V_1;V_2|V_0}) \tag{4.58c}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V_1;Y_1|V_0}) + \mathbb{I}(\mathrm{V_0V_2;Y_2}) - \mathbb{I}(\mathrm{V_0V_1V_2;Z}) - \mathbb{I}(\mathrm{V_1;V_2|V_0}) \tag{4.58d}$$

for random variables with joint distribution: $\mathrm{Q_{V_0}}\ \mathrm{Q_{V_1V_2|V_0}}\ \mathrm{Q_{X|V_0V_1V_2}}\ Q_{\mathrm{Y_1Y_2Z|X}}$*, such that* $\mathrm{V_0} - (\mathrm{V_1, V_2}) - \mathrm{X} - (\mathrm{Y_1, Y_2, Z})$ *forms a Markov chain.*

By adapting the discussion given in Section 3.2.1 to the rate regions given by Theorem 4.5 and the previous proposition, we can argue that the rate region in (4.58) is a sub-region of the full rate region in (4.51). In fact, the rate region in (4.58) is optimal for the first legitimate receiver in the sense that it achieves the same bounds on the confidential rate R_1 as the one in (4.51). We now present a random coding scheme with random variable C^1 that satisfies the reliability and joint-strong secrecy requirements in (4.8) for $k = 2$, such that the confidential rates R_1 and R_2 are bounded as given by Proposition 4.1.

Let $n \in \mathbb{N}$ be an arbitrary but fixed code block length and $\epsilon > 0$ be an arbitrary constant. Moreover, let $\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2\mathrm{X}} \in \mathcal{P}(\mathcal{V}_0 \times \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{X})$ be a fixed input probability distribution and recall the probability transition matrix $Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$ given in (2.73) that defines the two-receiver DM-WBC. Furthermore, we assume that quantities of the form 2^{nR}, where $R \in \mathbb{R}_+$ are integers.

1. Message sets: We assume that any message set is of the form: $\mathcal{M}_a = [\![1, 2^{nR_a}]\!]$. For our coding scheme, we have the following message sets: two confidential messages sets $\mathcal{M}_1$ and $\mathcal{M}_2$, three randomization messages sets $\mathcal{M}_r$, $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$, along with an additional message set $\mathcal{M}_t$ needed for the construction of the Marton code. Additionally we divide each confidential message into two parts as follows: $\mathcal{M}_1 = \mathcal{M}_{11} \times \mathcal{M}_{12}$ and $\mathcal{M}_2 = \mathcal{M}_{21} \times \mathcal{M}_{22}$. We then use $\mathcal{M}_{11}$ and $\mathcal{M}_{21}$ to construct a virtual common confidential message: $\mathcal{M}_0 \triangleq \mathcal{M}_{11} \times \mathcal{M}_{21}$. Thus, we have

$$\begin{aligned} R_1 &= R_{11} + R_{12} \\ R_2 &= R_{21} + R_{22} \\ R_0 &= R_{11} + R_{21} \end{aligned} \tag{4.59}$$

2. Random Codebook $\mathcal{C}^1$: Let $\mathrm{C}_0^1 \triangleq \{\mathrm{V}_0^n(m_0, m_r)\}$, where $m_0 \in \mathcal{M}_0$ and $m_r \in \mathcal{M}_r$ be a random codebook for the virtual common confidential message $m_0 = (m_{11}, m_{21})$. This codebook consists of $2^{n(R_{11}+R_{21}+R_r)}$ conditionally independent random codewords $\mathrm{V}_0^n(m_0, m_r)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{0_i}(m_0, m_r)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_0}$. A realization of C_0^1 is denoted by $\mathcal{C}_0^1 \triangleq \{v_0^n(m_0, m_r)\}$.

For a fixed codebook $\mathcal{C}_0^1$ and for every $m_0 \in \mathcal{M}_0$ and $m_r \in \mathcal{M}_r$, let $\mathrm{C}_1^1(m_0, m_r) \triangleq \{\mathrm{V}_1^n(m_0, m_r, m_{12}, m_{r_1})\}$, where $m_{12} \in \mathcal{M}_{12}$ and $m_{r_1} \in \mathcal{M}_{r_1}$ be a confidential random codebook that consists of $2^{n(R_{12}+R_{r_1})}$ conditionally independent random codewords $\mathrm{V}_1^n(m_0, m_r, m_{12}, m_{r_1})$, where each codeword is constructed by generating the symbols $\mathrm{V}_{1_i}(m_0, m_r, m_{12}, m_{r_1})$ independently at random according to $\mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0}(\cdot|v_{0_i}(m_0, m_r))$.

Similarly, for a fixed codebook $\mathcal{C}_0^1$ and for every $m_0 \in \mathcal{M}_0$ and $m_r \in \mathcal{M}_r$, let $\mathrm{C}_2^1(m_0, m_r) \triangleq \{\mathrm{V}_2^n(m_0, m_r, m_{22}, m_{r_2}, m_t)\}$ be a confidential random codebook that consists of $2^{n(R_{22}+R_{r_2}+R_t)}$ conditionally independent random codewords $\mathrm{V}_2^n(m_0, m_r, m_{22}, m_{r_2}, m_t)$, where each codeword is constructed by generating the symbols $\mathrm{V}_{2_i}(m_0, m_r, m_{22}, m_{r_2}, m_t)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_2|\mathrm{V}_0}(\cdot|v_{0_i}(m_0, m_r))$.

The total random codebook is denoted by $\mathrm{C}^1 = \{\mathrm{C}_0^1, \mathrm{C}_1^1, \mathrm{C}_2^1\}$, while $\mathcal{C}^1 = \{\mathcal{C}_0^1, \mathcal{C}_1^1, \mathcal{C}_2^1\}$ denotes one possible realization of this codebook selected from the total set of codebooks given by $\mathfrak{C}^1$. The probability of generating a certain realization $\mathcal{C}^1 \in \mathfrak{C}^1$ is given by:

$$\mathrm{G}(\mathcal{C}^1) = \prod_{\substack{(m_0^2,m_r^2,m_{22},m_{r_2},m_t)\in \\ \mathcal{M}_0\times\mathcal{M}_r\times\mathcal{M}_{22}\times\mathcal{M}_{r_2}\times\mathcal{M}_t}} \mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0}\Big(v_2^n(m_0^2,m_r^2,m_{22},m_{r_2},m_t)|v_0^n(m_0^2,m_r^2)\Big)$$
$$\prod_{\substack{(m_0^1,m_r^1,m_{12},m_{r_1})\in \\ \mathcal{M}_0\times\mathcal{M}_r\times\mathcal{M}_{12}\times\mathcal{M}_{r_1}}} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}\Big(v_1^n(m_0^1,m_r^1,m_{12},m_{r_1})|v_0^n(m_0^1,m_r^1)\Big)$$
$$\prod_{\substack{(m_0,m_r)\in \\ \mathcal{M}_0\times\mathcal{M}_r}} \mathrm{Q}^n_{\mathrm{V}_0}\Big(v_0^n(m_0,m_r)\Big) \tag{4.60}$$

3. Encoder E: For a fixed codebook realization $\mathcal{C}^1$, given a message triple (m_0, m_{12}, m_{22}), where $m_0 = (m_{11}, m_{21})$, the transmitter chooses three randomization messages m_r, m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_r$, $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. Then, it chooses an index $m_t \in \mathcal{M}_t$ based on the following likelihood encoder:

$$E^{(\mathrm{LE})}\big(m_t|m_{r_2}, v_0^n(m_0,m_r), v_1^n(m_0,m_r,m_{12},m_{r_1})\big)$$
$$= \frac{\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n(m_0,m_r,m_{12},m_{r_1})|v_0^n(m_0,m_r), v_2^n(m_0,m_r,m_{22},m_{r_2},m_t)\big)}{\sum_{j\in\mathcal{M}_t}\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n(m_0,m_r,m_{12},m_{r_1})|v_0^n(m_0,m_r), v_2^n(m_0,m_r,m_{22},m_{r_2},j)\big)}. \tag{4.61}$$

Finally, for the selected triple $v_0^n(m_0,m_r)$, $v_1^n(m_0,m_r,m_{12},m_{r_1})$, and $v_2^n(m_0,m_r,m_{22},m_{r_2},m_t)$, the encoder generates a codeword x^n independently at random according to the conditional distribution $\prod_{i=1}^n \mathrm{Q}^n_{\mathrm{X}|\mathrm{V}_0,\mathrm{V}_1,\mathrm{V}_2}(\cdot|v_0^n(m_0,m_r), v_1^n(m_0,m_r,m_{12},m_{r_1}), v_2^n(m_0,m_r,m_{22},m_{r_2},m_t))$ and transmits it.

4. First Legitimate Decoder φ_1: Given y_1^n, the decoder searches for the unique quadruple $(\hat{m}_0,\hat{m}_r,\hat{m}_{12},\hat{m}_{r_1}) \in \mathcal{M}_0\times\mathcal{M}_r\times\mathcal{M}_{12}\times\mathcal{M}_{r_1}$; such that:

$$\Big(v_0^n(\hat{m}_0,\hat{m}_r), v_1^n(\hat{m}_0,\hat{m}_r,\hat{m}_{12},\hat{m}_{r_1}), y_1^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{Y}_1}),$$

where $\hat{m}_0 = (\hat{m}_{11},\hat{m}_{21})$. If such unique quadruple exists, the decoder outputs the message $\hat{m}_1 = (\hat{m}_{11},\hat{m}_{12})$. Otherwise, the decoder outputs an error message (?).

5. Second Legitimate Decoder φ_2: Given y_2^n, the decoder searches for the unique quadruple $(\tilde{m}_0,\tilde{m}_r,\tilde{m}_{22},\tilde{m}_{r_2}) \in \mathcal{M}_0\times\mathcal{M}_r\times\mathcal{M}_{22}\times\mathcal{M}_{r_2}$; for which there exist an index $\tilde{m}_t \in \mathcal{M}_t$ such that:

$$\Big(v_0^n(\tilde{m}_0,\tilde{m}_r), v_2^n(\tilde{m}_0,\tilde{m}_r,\tilde{m}_{22},\tilde{m}_{r_2},\tilde{m}_t), y_2^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_2\mathrm{Y}_2}),$$

where $\tilde{m}_0 = (\tilde{m}_{11},\tilde{m}_{21})$. If such unique quadruple exists, the decoder outputs the message $\tilde{m}_2 = (\tilde{m}_{21},\tilde{m}_{22})$. If not, it outputs an error message (?).

6. Induced Joint Distribution: The encoding function E along with the decoding functions (φ_1,φ_2) defined with respect to the codebook $\mathcal{C}^1 \in \mathfrak{C}^1$ compose a $(2^{nR_1}, 2^{nR_2}, n)$ code of Definition[1] 4.1. For every codebook $\mathcal{C}^1$, the following joint probability distribution is induced:

[1]We slightly abuse the notation by using $\mathcal{C}^1$ to denote both the codebook and the whole code which includes the codebook itself along with the encoding and decoding functions. We favor this notation for its simplicity and remind the reader that the codebook uniquely defines the code

$$\begin{aligned}
&\mathrm{P}^{\mathcal{C}^1}\Big(m_0, m_{12}, m_{22}, m_r, m_{r_1}, m_{r_2}, v_0^n, v_1^n, m_t, v_2^n, x^n, y_1^n, y_2^n, z^n, \hat{m}_1, \tilde{m}_2\Big) \\
&\qquad = 2^{-n(R_0+R_{12}+R_{22}+R_r+R_{r_1}+R_{r_2})}\mathbb{1}_{\{v_0^n=v_0^n(m_0,m_r)\}}\mathbb{1}_{\{v_1^n=v_1^n(m_0,m_r,m_{12},m_{r_1})\}} \\
&\qquad \times E^{(\mathrm{LE})}\big(m_t|m_{r_2}, v_0^n(m_0,m_r), v_1^n(m_0,m_r,m_{12},m_{r_1})\big) \\
&\qquad \times \mathrm{Q}^n_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}(x^n|v_0^n,v_1^n,v_2^n)\, Q^n_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}(y_1^n,y_2^n,z^n|x^n)\mathbb{1}_{\{\hat{m}_1=\varphi_1(y_1^n)\}} \\
&\qquad \times \mathbb{1}_{\{\tilde{m}_2=\varphi_2(y_2^n)\}}\mathbb{1}_{\{v_2^n=v_2^n(m_0,m_r,m_{22},m_{r_2},m_t)\}}
\end{aligned} \tag{4.62}$$

Additionally, if we take the random codebook generation process into our consideration, we end up with the subsequent distribution: $\mathrm{P} = \mathrm{G}(\mathcal{C}^1) \cdot \mathrm{P}^{\mathcal{C}^1}$, where $\mathrm{G}(\mathcal{C}^1)$ is the probability of obtaining a certain codebook $\mathcal{C}^1$ given by (4.60).

7. Reliability Analysis: We start by defining the average decoding error probability for a given code realization $\mathcal{C}^1 \in \mathfrak{C}^1$ as follows:

$$\begin{aligned}
\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1) \triangleq \mathbb{P}\Big[&(\hat{\mathrm{M}}_0, \hat{\mathrm{M}}_{12}, \hat{\mathrm{M}}_r, \hat{\mathrm{M}}_{r_1}) \neq (\mathrm{M}_0, \mathrm{M}_{12}, \mathrm{M}_r, \mathrm{M}_{r_1}) \text{ or } (\tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_{22}, \tilde{\mathrm{M}}_r, \tilde{\mathrm{M}}_{r_2}) \\
&\neq (\mathrm{M}_0, \mathrm{M}_{22}, \mathrm{M}_r, \mathrm{M}_{r_2})\Big].
\end{aligned} \tag{4.63}$$

We then observe that the previous error probability is greater than or equal the one in (4.2) for $k = 2$, i.e. $\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1) \geq \bar{\mathbf{P}}_e(\mathcal{C}^1)$. Thus, in order to show that our coding scheme satisfies the reliability constraint in (4.8b), it is enough to show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\left[\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1)\right] = 0$.

Based on the definitions of our encoding function and decoding functions, we need to consider two classes of errors:

A- Encoding Error: This error occurs if the codeword v_2^n chosen by the likelihood encoder and the codewords (v_0^n, v_1^n) selected uniformly at random are not jointly typical. For this error, we define the following event:

$$\mathcal{E} = \left\{\Big(\mathrm{V}_0^n(\mathrm{M}_0,\mathrm{M}_r), \mathrm{V}_1^n(\mathrm{M}_0,\mathrm{M}_r,\mathrm{M}_{12},\mathrm{M}_{r_1}), \mathrm{V}_2^n(\mathrm{M}_0,\mathrm{M}_1,\mathrm{M}_{22},\mathrm{M}_{r_2},\mathrm{M}_t)\Big) \notin \mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}}\right\},$$

where $\mathcal{T}^n_{\bar{\epsilon},\mathrm{inp}} \triangleq \mathcal{T}^n_{\bar{\epsilon}}(\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2})$, $\bar{\epsilon} \in (0,\epsilon)$ and $\epsilon > 0$ is the typicality constant used by the decoding functions. By adapting the reliability analysis in Section 3.2.3 to our coding scheme, we can show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\left[\mathbb{P}_{\mathrm{P}}(\mathcal{E})\right] = 0$, as long as:

$$\begin{aligned}
R_t &\geq \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon,\beta), \\
\text{where} \quad \delta_0(\epsilon,\beta) &= \epsilon \log\big(|\mathcal{V}_0|\cdot|\mathcal{V}_1|\cdot|\mathcal{V}_2|\big) + 5\beta \quad \text{for } \beta > 0.
\end{aligned} \tag{4.64}$$

B- Decoding Error: This error occurs if the decoders decide on a set of messages different from the ones which were originally transmitted. Let us denote this class of errors by the event $\mathcal{D}$. Based on the standard joint-typicality decoding argument along with the reliability analysis in Section 3.2.3, we can show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\left[\mathbb{P}_{\mathrm{P}}(\mathcal{D}|\mathcal{E}^c)\right] = 0$, as long as:

$$R_{12} + R_{r_1} \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_0) - \delta_1(\epsilon) \tag{4.65a}$$

$$R_0 + R_{12} + R_r + R_{r_1} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \delta_1(\epsilon) \tag{4.65b}$$

$$R_{22} + R_{r_2} + R_t \leq \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_0) - \delta_2(\epsilon) \tag{4.65c}$$

$$R_0 + R_{22} + R_r + R_{r_2} + R_t \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \delta_2(\epsilon), \tag{4.65d}$$

where $\delta_1(\epsilon) = 2\epsilon \log\big(|\mathcal{V}_0| \cdot |\mathcal{V}_1| \cdot |\mathcal{Y}_1|\big)$ and $\delta_2(\epsilon) = 2\epsilon \log\big(|\mathcal{V}_0| \cdot |\mathcal{V}_2| \cdot |\mathcal{Y}_2|\big)$. The previous analysis implies that under the rate constraints in (4.65) and (4.64), our coding scheme satisfies the reliability constraint in (4.8b) because:

$$\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\big[\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1)\big] \leq \lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\big[\mathbb{P}_{\mathrm{P}}(\mathcal{E})\big] + \mathbb{E}_{\mathrm{C}^1}\big[\mathbb{P}_{\mathrm{P}}(\mathcal{D}|\mathcal{E}^c)\big] = 0. \tag{4.66}$$

8. Secrecy Analysis: Our secrecy analysis adheres to the one used in Section 3.2.4 which inherits the strong secrecy analysis of Marton code presented in [127] along with the strong secrecy results for superposition encoding presented in [29, 31] which were highlighted in Section 2.2.3 and Section 2.2.4. We start by bounding the joint-strong secrecy constraint in (4.8c) for $k = 2$ as follows:

$$\begin{aligned}\mathbb{E}_{\mathrm{C}^1}\left[\mathbb{I}(\mathrm{M}_1\mathrm{M}_2;\mathrm{Z}^n|\mathrm{C}^1)\right] &\overset{(a)}{\leq} \alpha \cdot \mathbb{E}_{\mathrm{C}^1}\Big[\big\|\mathrm{P}_{\mathrm{M}_1\mathrm{M}_2\mathrm{Z}^n|\mathrm{C}^1} - \mathrm{P}_{\mathrm{M}_1\mathrm{M}_2|\mathrm{C}^1} \times \mathrm{P}_{\mathrm{Z}^n|\mathrm{C}^1}\big\|\Big] \\ &\overset{(b)}{\leq} 2\alpha \cdot \mathbb{E}_{\mathrm{C}^1}\Big[\big\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1} - \mathrm{Q}^n_{\mathrm{Z}}\big\|\Big],\end{aligned} \tag{4.67}$$

where (a) follows from the relation between the mutual information and the total variation distance given in Lemma A.6, such that $\alpha > 0$ is a constant that does not depend on n; while (b) follows because M_1 and M_2 are uniformly distributed on their respective sets. Moreover, $\mathrm{Q}^n_{\mathrm{Z}}$ is given by:

$$\mathrm{Q}^n_{\mathrm{Z}}(z^n) \triangleq \sum_{x^n\in\mathcal{X}^n} \sum_{y_1^n\in\mathcal{Y}_1^n} \sum_{y_2^n\in\mathcal{Y}_2^n} \mathrm{Q}^n_{\mathrm{X}}(x^n)\; Q^n_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}(y_1^n, y_2^n, z^n|x^n). \tag{4.68}$$

$\mathrm{Q}^n_{\mathrm{Z}}$ can be interpreted as the probability distribution observed by the eavesdropper when the no useful information is transmitted over the channel. Before we proceed with our secrecy analysis, we need to introduce two additional joint distributions that are very useful in bounding the expression in (4.67). The first one is $\ddot{\mathrm{P}} \triangleq \mathrm{G}(\mathcal{C}^1) \cdot \ddot{\mathrm{P}}^{\mathcal{C}^1}$, where $\mathrm{G}(\mathcal{C}^1)$ is the codebook generating distribution given by (4.60), while $\ddot{\mathrm{P}}^{\mathcal{C}^1}$ is defined as:

$$\begin{aligned}\ddot{\mathrm{P}}^{\mathcal{C}^1}\big(m_0, m_r, v_0^n, m_{12}, m_{r_1}, v_1^n, z^n\big) &= 2^{-n(R_0+R_r)}\, \mathbb{1}_{\{v_0^n = v_0^n(m_0,m_r)\}}\, 2^{-n(R_{12}+R_{r_1})} \\ &\quad \times \mathbb{1}_{\{v_1^n = v_1^n(m_0,m_r,m_{12},m_{r_1})\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}(z^n|v_0^n, v_1^n).\end{aligned} \tag{4.69}$$

The definition of $\ddot{\mathrm{P}}^{\mathcal{C}^1}$ utilizes the same analogy used to define the distribution in Eq. (3.68), where $Q^n_{\mathrm{Z}|\mathrm{V}_0\mathrm{V}_1}$ is the DMC given by (3.66). The second distribution needed is defined as: $\hat{\mathrm{P}} \triangleq \mathrm{G}(\mathcal{C}^1) \cdot \hat{\mathrm{P}}^{\mathcal{C}^1}$, where $\hat{\mathrm{P}}^{\mathcal{C}^1}$ is given by:

$$\hat{\mathrm{P}}^{\mathcal{C}^1}\big(m_0, m_r, v_0^n, z^n\big) = 2^{-n(R_0+R_r)}\, \mathbb{1}_{\{v_0^n = v_0^n(m_0,m_r)\}}\, Q^n_{\mathrm{Z}|\mathrm{V}_0}(z^n|v_0^n). \tag{4.70}$$

Similarly, the previous definition is based on the same analogy used to define the probability distribution in Eq. (3.72), where $Q^n_{\mathrm{Z}|\mathrm{V}_0}$ is the DMC given by (3.70). Now, let us denote the total variation distance term in (4.67) by $\mathcal{L}$, then based on the triangle inequality and the three distributions defined in (4.68), (4.69) and (4.70), it follows that:

$$\begin{aligned}\mathbb{E}_{\mathrm{C}^1}\left[\mathcal{L}\right] &\leq \mathbb{E}_{\mathrm{C}^1}\Big[\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1} - \ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1}\right\| + \left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1} - \hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1}\right\| \\ &\qquad + \left\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_1,\mathrm{M}_2,\mathrm{C}^1} - \mathrm{Q}_{\mathrm{Z}}^n\right\|\Big] \\ &\overset{(a)}{=} \mathbb{E}_{\mathrm{C}}\Big[\left\|\mathrm{P}_{\mathrm{Z}^n|\mathrm{M}_0=1,\mathrm{M}_{12}=1,\mathrm{M}_{22}=1,\mathrm{C}^1} - \ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_0=1,\mathrm{M}_{12}=1,\mathrm{C}^1}\right\| \\ &\qquad + \left\|\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_0=1,\mathrm{M}_{12}=1,\mathrm{C}^1} - \hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_0=1,\mathrm{C}^1}\right\| + \left\|\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{M}_0=1,\mathrm{C}^1} - \mathrm{Q}_{\mathrm{Z}}^n\right\|\Big], \quad (4.71)\end{aligned}$$

where (a) follows from the definition of the messages as: $(\mathrm{M}_1, \mathrm{M}_2) = (\mathrm{M}_0, \mathrm{M}_{12}, \mathrm{M}_{22})$, along with the fact that $\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{C}^1}$ only depends on $(\mathrm{M}_0, \mathrm{M}_{12})$ and is independent of M_{22}, while $\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{C}^1}$ only depends on M_0. Additionally, we make use of the fact that the three distributions $\mathrm{P}_{\mathrm{Z}^n|\mathrm{C}^1}$, $\ddot{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{C}^1}$ and $\hat{\mathrm{P}}_{\mathrm{Z}^n|\mathrm{C}^1}$ are symmetric with respect to the corresponding messages.

Finally, we modify the secrecy analysis derived in Section 3.2.4 to support the structure of our coding scheme. In particular, we apply the same techniques used to bound the $\mathbb{RHS}$ of Eq. (3.73), aiming to bound the $\mathbb{RHS}$ of Eq. (4.71). By doing so, we can show that $\mathbb{E}_{\mathrm{C}^1}\left[\mathcal{L}\right]$ decays exponentially in n, as long as the following rate constraints hold:

$$R_t \geq \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon, \beta) \quad (4.72a)$$
$$R_{r_2} + R_t \geq \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) + \delta_3(\epsilon, \gamma) \quad (4.72b)$$
$$R_{r_1} \geq \mathbb{I}(\mathrm{V}_1; \mathrm{Z}|\mathrm{V}_0) + \delta_3(\epsilon, \gamma) \quad (4.72c)$$
$$R_r \geq \mathbb{I}(\mathrm{V}_0; \mathrm{Z}) + \delta_3(\epsilon, \gamma), \quad (4.72d)$$

where $\gamma > 0$ and $\delta_3(\epsilon, \gamma) = 2\epsilon \log\left(|\mathcal{V}_0| \cdot |\mathcal{V}_1| \cdot |\mathcal{V}_2| \cdot |\mathcal{Z}|\right) + 5\gamma$. The previous result along with Eq. (4.67) imply that under the rate constraints in (4.72), our coding scheme satisfies the joint-strong secrecy constraint in (4.8c) for $k = 2$.

9. <u>Fourier-Motzkin Elimination:</u> We now combine the rate splitting equations in (4.59), along with the reliability rate constraints in (4.64) and (4.65), in addition to the secrecy rate constraints in (4.72), then apply the Fourier-Motzkin Elimination procedure and solve for the confidential rates R_1 and R_2. Moreover, we took the limit as $n \to \infty$ and choose the constants $\epsilon, \beta, \gamma > 0$, such that their corresponding $\delta-$functions approach zero as n approaches infinity. With this analysis, we prove that any rate pair (R_1, R_2) that satisfy the rate constraints in (4.58) is achievable under the joint-strong secrecy criterion; establishing Proposition 4.1.

The previous achievability proof directly suggests that there exists an alternative random coding scheme C^2 which also satisfies the reliability and secrecy constraints given by (4.8b) and (4.8c) respectively, while its rates are bounded according to the following proposition:

Proposition 4.2. *An achievable joint-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs $(R_1, R_2) \in R_+^2$ that satisfy*

$$R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0; \mathrm{Z}) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2\mathrm{Z}|\mathrm{V}_0) \quad (4.73a)$$
$$R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}) \quad (4.73b)$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2; \mathrm{Y}_2|\mathrm{V}_0) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) \quad (4.73c)$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Y}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}) - \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) \quad (4.73d)$$

for random variables with joint distribution: $\mathrm{Q}_{\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

It is obvious that Proposition 4.2 is a direct consequence of Proposition 4.1, as it follows by interchanging the construction procedure of the random codewords V_1^n and V_2^n along with modifying the decoding functions at the first and the second legitimate receivers accordingly.

The last step in the proof of Theorem 4.5 is to use the principles of rate splitting and time sharing to construct a coding scheme similar to the one introduced in Section 3.2.5 to combine the two achievable rate regions given by (4.58) and (4.73). For a time sharing parameter $\pi \in [0, 1]$, we can show that the following rate region is achievable:

$$R_1 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Z}|\mathrm{T}) - (1-\pi) \cdot \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0\mathrm{Z}\mathrm{T}) \tag{4.74a}$$

$$R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Z}|\mathrm{T}) - \pi \cdot \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0\mathrm{Z}\mathrm{T}) \tag{4.74b}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_2; \mathrm{Y}_2|\mathrm{V}_0\mathrm{T}) - R_e \tag{4.74c}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Y}_1|\mathrm{V}_0\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - R_e \tag{4.74d}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1|\mathrm{T}) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2|\mathrm{T}) - \mathbb{I}(\mathrm{V}_0; \mathrm{Z}|\mathrm{T}) - R_e, \tag{4.74e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2; \mathrm{Z}|\mathrm{T}) + \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0\mathrm{T})$. The achievability of the rate region in (4.74) directly implies the validity of Theorem 4.5. This is because for any rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ that satisfies the rate constraints in (4.51), there exists a certain value for the time sharing parameter π such that the given rate pair is included in the rate region in (4.74).

The usage of Marton coding with a superposition variable was first introduced in [146, 147] to establish an achievable rate region for the two-receiver DM-BC with a common message (M_0) and two individual private messages ($\mathrm{M}_1, \mathrm{M}_2$). The idea was to encode the common message M_0 in the superposition variable V_0, while the two private messages M_1 and M_2 are encoded in the classical Marton coding variables V_1 and V_2 respectively. It was then argued in [35, 148] that even for the two-receiver DM-BC without a common message, Marton coding is optimal in the sense that it is tight for all classes of two-receiver DM-BCs for which the transmission capacity is known only in the presence of the superposition variable V_0. This observation seems to be valid for the two-receiver DM-WBC as well. In fact, without the superposition variable V_0, we can not show that the rate region established in Theorem 4.5 recovers the joint-strong secrecy capacity region of the degraded two-receiver DM-WBC given by Corollary 4.1.

The previous discussion was the reason that our coding scheme is based on a Marton wiretap random code with a superposition variable instead of the classical Marton wiretap random code presented in [73], cf. Theorem 2.7. The main issue was to come up with a technique that allows us to include the superposition variable V_0 in our coding scheme without having an actual common confidential message. In order to over come this issue, we thought about two solutions: The first was to modify our model such that we have a real common confidential message M_0, then after we derive our achievable rate region, we can let $R_0 = 0$. This idea is similar to the one used

in [35] for the two-receiver DM-BC without an eavesdropper. The second solution was to utilize the concept of rate splitting to create a virtual common confidential message $\mathrm{M}_0 = \mathrm{M}_{11} \times \mathrm{M}_{21}$ and construct our coding scheme accordingly. Although the second solution forces each legitimate receiver to decode some unneeded information, i.e. Y_1 decodes M_{12} and Y_2 decodes M_{11}, we found that both solutions lead to the same achievable region at the end. Nevertheless, we preferred the second solution because it will be very helpful for the individual secrecy scenario.

4.3.3 Marton Wiretap Coding Under Individual Secrecy

The main feature of the individual secrecy criterion is that the legitimate receivers can cooperate to protect their message against eavesdropping based on the idea of mutual trust. This mutual trust allows us to use the confidential message of one receiver as a secret key for the other one. Thus, we can combine normal wiretap encoding and secret key encoding leading to a larger achievable rate region compared to the joint secrecy criterion where only wiretap encoding is used. However, in order to be able to use secret key encoding, the key must be secretly shared between the transmitter and the respective receiver. Since in our case the key is simply part of the confidential message of another receiver, there must be away for each receiver to acquire a full knowledge about the part of the message of the other receiver used as secret key.

Up until this point, we have seen two ways to acquire this knowledge: The first one was discussed in Section 3.3 for the DM-WBC with message cognition, where each legitimate receiver is fully cognizant of the confidential message of the other receiver prior to the transmission. The second approach was discussed in Section 4.2 for the degraded DM-WBC, where the degradedness property allows each receiver to be able to decode the confidential messages of the weaker receivers. For the general two-receiver DM-WBC, there is no prior knowledge about the confidential messages available at the opposite receiver, neither do the channels of the legitimate receivers exhibit a certain statistical ordering (degradedness or less noisy). Thus, we need another technique to establish a shared secret key.

We start by recalling the discussion at the end of Section 4.3.2. We can notice that the coding scheme used to establish Proposition 4.1 allows each legitimate receiver to acquire some information about part of the confidential message intended to the other receiver. In particular, Y_1 was able to decode M_{21} reliably, Y_2 was able to decode M_{11}. This implies that M_{21} can be used as a secret key for the first legitimate receiver and the same holds for M_{11} and the second legitimate receiver. In this section, we utilize this idea along with the coding scheme introduced in the previous subsection to prove the achievability of the individual-strong secrecy rate region given by Theorem 4.6. We start by validating the following proposition:

Proposition 4.3. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs* $(R_1 = R_{11} + R_{12}, R_2 = R_{21} + R_{22}) \in R_+^2$ *that satisfy*

$$R_1 + R_{21} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}) + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}), R_{21}\right] \tag{4.75a}$$

$$\begin{aligned} R_2 + R_{11} \leq\ & \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1|\mathrm{V}_0\mathrm{Z}) \\ & + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{Z}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0), R_{11}\right] \end{aligned} \tag{4.75b}$$

$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_0) - R_e + R_s \tag{4.75c}$$
$$R_1 + R_2 \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - R_e + R_s \tag{4.75d}$$
$$R_1 + R_2 + R_{11} + R_{21} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}) - R_e + R_s \tag{4.75e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0)$*, while* $R_s \triangleq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}), R_{11} + R_{21}\right]$*. The random variables that define the previous rate region are characterized by the following joint probability distribution:* $\mathrm{Q}_{\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that the following Markov chain holds:* $\mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$*.*

1. Message sets: We assume that any message set is of the form: $\mathcal{M}_a = [\![1, 2^{nR_a}]\!]$. We consider the same message sets used in the proof of Proposition 4.1, with some minor modifications: We divide each confidential message into three parts instead of two as follows: $\mathcal{M}_1 = \mathcal{M}_{11} \times \mathcal{M}_{12} \times \mathcal{M}_{13}$ and $\mathcal{M}_2 = \mathcal{M}_{21} \times \mathcal{M}_{22} \times \mathcal{M}_{23}$. This division implies the following rate constraints:

$$\begin{aligned} R_1 &= R_{11} + R_{12} + R_{13} \\ R_2 &= R_{21} + R_{22} + R_{23} \end{aligned} \tag{4.76}$$

Moreover, in this division we force $\mathcal{M}_{13}$ and M_{21} to be of the same size and use them to construct a new message set $\mathcal{M}_{\otimes_1}$ by *xoring* the corresponding elements of both sets. Similarly, we force $\mathcal{M}_{23}$ and M_{11} to be of the same size and use them to construct a new message set $\mathcal{M}_{\otimes_2}$ in the exact same way. Moreover, we divide the two *xored* message sets into two parts as follows: $\mathcal{M}_{\otimes_1} = \mathcal{M}_{\otimes_{11}} \times \mathcal{M}_{\otimes_{12}}$ and $\mathcal{M}_{\otimes_2} = \mathcal{M}_{\otimes_{21}} \times \mathcal{M}_{\otimes_{22}}$. Based on the structure of these message sets, we have

$$\begin{aligned} R_{13} = R_{21} = R_{\otimes_1} &= R_{\otimes_{11}} + R_{\otimes_{12}} \\ R_{23} = R_{11} = R_{\otimes_2} &= R_{\otimes_{21}} + R_{\otimes_{22}}. \end{aligned} \tag{4.77}$$

Finally, we let $\mathcal{M}_0 \triangleq (\mathcal{M}_{11} \times \mathcal{M}_{21} \times \mathcal{M}_r \times \mathcal{M}_{\otimes_{11}} \times \mathcal{M}_{\otimes_{21}})$, which consequently means that $m_0 = (m_{11}, m_{21}, m_r, m_{\otimes_{11}}, m_{\otimes_{21}})$. It also implies the following:

$$R_0 = R_{11} + R_{21} + R_r + R_{\otimes_{11}} + R_{\otimes_{21}} \tag{4.78}$$

2. Random Codebook $\mathcal{C}^1$: Let $\mathrm{C}_0^1 \triangleq \{\mathrm{V}_0^n(m_0)\}$, be a confidential random codebook that consists of 2^{nR_0} conditionally independent random codewords, where each codeword is constructed by generating the symbols $\mathrm{V}_{0_i}(m_0)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_0}$. A realization of C_0^1 is denoted by $\mathcal{C}_0^1 \triangleq \{v_0^n(m_0)\}$.

For a fixed codebook realization $\mathcal{C}_0^1$ and for every $m_0 \in \mathcal{M}_0$, let $\mathrm{C}_1^1(m_0) \triangleq \{\mathrm{V}_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}})\}$ be a confidential random codebook that consists of $2^{n(R_{12}+R_{r_1}+R_{\otimes_{12}})}$ conditionally independent random codewords, where each codeword is constructed by generating the symbols $\mathrm{V}_{1_i}(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}})$ independently at random according to $\mathrm{Q}_{\mathrm{V}_1|\mathrm{V}_0}(\cdot|v_{0_i}(m_0))$.

Similarly, for a fixed codebook realization $\mathcal{C}_0^1$ and for every $m_0 \in \mathcal{M}_0$, let $\mathrm{C}_2^1(m_0) \triangleq \{\mathrm{V}_2^n(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t)\}$ be a confidential random codebook that consists of $2^{n(R_{22}+R_{r_2}+R_{\otimes_{22}}+R_t)}$ conditionally independent random codewords, where each codeword is constructed by generating the symbols $\mathrm{V}_{2_i}(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t)$ independently at random according to $\mathrm{Q}_{\mathrm{V}_2|\mathrm{V}_0}(\cdot|v_{0_i}(m_0))$.

The total random codebook is denoted by $\mathrm{C}^1 = \{\mathrm{C}_0^1, \mathrm{C}_1^1, \mathrm{C}_2^1\}$, while $\mathcal{C}^1 = \{\mathcal{C}_0^1, \mathcal{C}_1^1, \mathcal{C}_2^1\}$ denotes one possible realization of this codebook. The probability of generating a certain realization $\mathcal{C}^1 \in \mathfrak{C}^1$ is given by:

$$\mathrm{G}(\mathcal{C}^1) = \prod_{\substack{(m_0^2, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t) \in \\ \mathcal{M}_0 \times \mathcal{M}_{22} \times \mathcal{M}_{r_2} \times \mathcal{M}_{\otimes_{22}} \times \mathcal{M}_t}} \mathrm{Q}^n_{\mathrm{V}_2|\mathrm{V}_0}\Big(v_2^n(m_0^2, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t) | v_0^n(m_0^2)\Big)$$
$$\prod_{\substack{(m_0^1, m_{12}, m_{r_1}, m_{\otimes_{12}}) \in \\ \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{M}_{r_1} \times \mathcal{M}_{\otimes_{12}}}} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0}\Big(v_1^n(m_0^1, m_{12}, m_{r_1}, m_{\otimes_{12}}) | v_0^n(m_0^1)\Big) \prod_{m_0 \in \mathcal{M}_0} \mathrm{Q}^n_{\mathrm{V}_0}\Big(v_0^n(m_0)\Big) \quad (4.79)$$

3. Encoder E: For a fixed codebook realization $\mathcal{C}^1$, given a message pair (m_1, m_2), where $m_1 = (m_{11}, m_{12}, m_{13})$ and $m_2 = (m_{21}, m_{22}, m_{23})$, the transmitter first calculates the *xored* messages $m_{\otimes_1} = (m_{\otimes_{11}}, m_{\otimes_{12}})$ and $m_{\otimes_2} = (m_{\otimes_{21}}, m_{\otimes_{22}})$. Then, it chooses three randomization messages m_r, m_{r_1} and m_{r_2} uniformly at random from the sets $\mathcal{M}_r$, $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ respectively. The encoder then chooses an index $m_t \in \mathcal{M}_t$ based on the following likelihood encoder:

$$E^{(\mathrm{LE})}\big(m_t | m_{r_2}, v_0^n(m_0), v_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}})\big)$$
$$= \frac{\mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}}) | v_0^n(m_0), v_2^n(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t)\big)}{\sum_{j \in \mathcal{M}_t} \mathrm{Q}^n_{\mathrm{V}_1|\mathrm{V}_0,\mathrm{V}_2}\big(v_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}}) | v_0^n(m_0), v_2^n(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, j)\big)}. \quad (4.80)$$

Finally, for the selected triple: $v_0^n(m_0)$, $v_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}})$, and $v_2^n(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t)$, the encoder generates a codeword x^n independently at random according to the conditional distribution $\prod_{i=1}^n \mathrm{Q}^n_{\mathrm{X}|\mathrm{V}_0,\mathrm{V}_1,\mathrm{V}_2}(\cdot | v_0^n(m_0), v_1^n(m_0, m_{12}, m_{r_1}, m_{\otimes_{12}}), v_2^n(m_0, m_{22}, m_{r_2}, m_{\otimes_{22}}, m_t))$ and transmits it.

4. First Legitimate Decoder φ_1: Given y_1^n, it outputs $\hat{m}_1 = (\hat{m}_{11}, \hat{m}_{12}, \hat{m}_{13})$ using the following procedure: First, it searches for the unique quadruple $(\hat{m}_0, \hat{m}_{12}, \hat{m}_{r_1}, \hat{m}_{\otimes_{12}}) \in \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{M}_{r_1} \times \mathcal{M}_{\otimes_{12}}$; such that:

$$\Big(v_0^n(\hat{m}_0), v_1^n(\hat{m}_0, \hat{m}_{12}, \hat{m}_{r_1}, \hat{m}_{\otimes_{12}}), y_1^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_1\mathrm{Y}_1}),$$

where $\hat{m}_0 = (\hat{m}_{11}, \hat{m}_{21}, \hat{m}_r, \hat{m}_{\otimes_{11}}, \hat{m}_{\otimes_{21}})$. The decoder then estimates $\hat{m}_{13}$ by *xoring* the messages $\hat{m}_{21}$ and $\hat{m}_{\otimes_1} = (\hat{m}_{\otimes_{11}}, \hat{m}_{\otimes_{12}})$. If the decoder fails to execute one of the previous two steps, it outputs an error message (?).

5. Second Legitimate Decoder φ_2: Given y_2^n, the decoder outputs $\tilde{m}_2 = (\tilde{m}_{21}, \tilde{m}_{22}, \tilde{m}_{23})$ based on the following procedure: First, it searches for the unique quadruple $(\tilde{m}_0, \tilde{m}_{22}, \tilde{m}_{r_2}, \tilde{m}_{\otimes_{22}}) \in \mathcal{M}_0 \times \mathcal{M}_{22} \times \mathcal{M}_{r_2} \times \mathcal{M}_{\otimes_{22}}$; for which there exist an index $\tilde{m}_t \in \mathcal{M}_t$ such that:

$$\Big(v_0^n(\tilde{m}_0), v_2^n(\tilde{m}_0, \tilde{m}_{22}, \tilde{m}_{r_2}, \tilde{m}_{\otimes_{22}}, \tilde{m}_t), y_2^n\Big) \in \mathcal{T}_\epsilon^n(\mathrm{Q}_{\mathrm{V}_0\mathrm{V}_2\mathrm{Y}_2}),$$

where $\tilde{m}_0 = (\tilde{m}_{11}, \tilde{m}_{21}, \tilde{m}_r, \tilde{m}_{\otimes_{11}}, \tilde{m}_{\otimes_{21}})$. The decoder then estimates $\tilde{m}_{23}$ by *xoring* the messages $\tilde{m}_{11}$ and $\tilde{m}_{\otimes_2} = (\tilde{m}_{\otimes_{21}}, \tilde{m}_{\otimes_{22}})$. If the decoder fails to execute any of the previous two steps, it outputs an error message (?).

6. Reliability Analysis: We start by defining the average decoding error probability for a given code realization $\mathcal{C}^1 \in \mathfrak{C}^1$ as follows:

$$\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1) \triangleq \mathbb{P}\Big[(\hat{\mathrm{M}}_0, \hat{\mathrm{M}}_{12}, \hat{\mathrm{M}}_{r_1}, \hat{\mathrm{M}}_{\otimes_{12}}) \neq (\mathrm{M}_0, \mathrm{M}_{12}, \mathrm{M}_{r_1}, \mathrm{M}_{\otimes_{12}}) \text{ or } (\tilde{\mathrm{M}}_0, \tilde{\mathrm{M}}_{22}, \tilde{\mathrm{M}}_{r_2}, \tilde{\mathrm{M}}_{\otimes_{22}}) \neq (\mathrm{M}_0, \mathrm{M}_{22}, \mathrm{M}_{r_2}, \mathrm{M}_{\otimes_{22}})\Big]. \tag{4.81}$$

We then observe that the previous error probability is greater than or equal to the one in (4.2) for $k = 2$, i.e. $\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1) \geq \bar{\mathbf{P}}_e(\mathcal{C}^1)$. Thus, in order to show that our coding scheme satisfies the reliability constraint in (4.10b), it is enough to show that $\lim_{n\to\infty} \mathbb{E}_{\mathrm{C}^1}\left[\bar{\bar{\mathbf{P}}}_e(\mathcal{C}^1)\right] = 0$.

Based on the same arguments used in the reliability analysis of the coding scheme that establishes Proposition 4.1, we can show that the expectation of the encoding and decoding errors for the underlying coding scheme vanishes as $n \to \infty$, as long as the following rate constraints holds:

$$R_t \geq \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon, \beta), \tag{4.82a}$$
$$R_{12} + R_{r_1} + R_{\otimes_{12}} \leq \mathbb{I}(\mathrm{V}_1; \mathrm{Y}_1|\mathrm{V}_0) - \delta_1(\epsilon) \tag{4.82b}$$
$$R_0 + R_{12} + R_{r_1} + R_{\otimes_{12}} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1; \mathrm{Y}_1) - \delta_1(\epsilon) \tag{4.82c}$$
$$R_{22} + R_{r_2} + R_{\otimes_{22}} + R_t \leq \mathbb{I}(\mathrm{V}_2; \mathrm{Y}_2|\mathrm{V}_0) - \delta_2(\epsilon) \tag{4.82d}$$
$$R_0 + R_{22} + R_{r_2} + R_{\otimes_{22}} + R_t \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2; \mathrm{Y}_2) - \delta_2(\epsilon), \tag{4.82e}$$

where the δ-functions are as defined in Section 4.3.2, for the typicality constant ϵ and an additional constant $\beta > 0$.

7. Secrecy Analysis: Because of the way used to divide the confidential message sets, the random variable M_1 is now identified as the product of three independent and uniformly distributed random variables: M_{11}, M_{12} and M_{13}. This feature also applies to M_2 which is the product of the three independent and uniformly distributed random variables M_{21}, M_{22} and M_{23}. Based on the previous argument, we can bound the individual-strong secrecy constraint in (4.10c) for $k = 2$ after dropping the conditioning on C^1 as follows:

$$\begin{aligned}\mathbb{E}_{\mathrm{C}^1}\left[\mathbb{I}(\mathrm{M}_1; \mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_2; \mathrm{Z}^n)\right] &= \mathbb{E}_{\mathrm{C}^1}\left[\mathbb{I}(\mathrm{M}_{11}\mathrm{M}_{12}\mathrm{M}_{13}; \mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_{21}\mathrm{M}_{22}\mathrm{M}_{23}; \mathrm{Z}^n)\right] \\ &\overset{(a)}{\leq} \mathbb{E}_{\mathrm{C}^1}\Big[\mathbb{I}(\mathrm{M}_{11}\mathrm{M}_{12}\mathrm{M}_{21}\mathrm{M}_{22}; \mathrm{Z}^n) + \mathbb{I}(\mathrm{M}_{13}; \mathrm{Z}^n|\mathrm{M}_{11}\mathrm{M}_{12}) \\ &\qquad + \mathbb{I}(\mathrm{M}_{23}; \mathrm{Z}^n|\mathrm{M}_{21}\mathrm{M}_{22})\Big],\end{aligned} \tag{4.83}$$

where (a) follows from the chain rule of the mutual information in addition to the independence of the messages. Based on the secrecy analysis of the coding scheme used to establish Proposition 4.1, it follows that $\mathbb{E}_{\mathrm{C}^1}\left[\mathbb{I}(\mathrm{M}_{11}\mathrm{M}_{12}\mathrm{M}_{21}\mathrm{M}_{22}; \mathrm{Z}^n)\right]$ decays exponentially in n as long as the following rate constraints holds:

$$R_t \geq \mathbb{I}(\mathrm{V}_1; \mathrm{V}_2|\mathrm{V}_0) + \delta_0(\epsilon, \beta) \tag{4.84a}$$
$$R_{r_2} + R_{\otimes_{22}} + R_t \geq \mathbb{I}(\mathrm{V}_2; \mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) + \delta_3(\epsilon, \gamma) \tag{4.84b}$$
$$R_{r_1} + R_{\otimes_{12}} \geq \mathbb{I}(\mathrm{V}_1; \mathrm{Z}|\mathrm{V}_0) + \delta_3(\epsilon, \gamma) \tag{4.84c}$$
$$R_r + R_{\otimes_{11}} + R_{\otimes_{21}} \geq \mathbb{I}(\mathrm{V}_0; \mathrm{Z}) + \delta_3(\epsilon, \gamma), \tag{4.84d}$$

where $\gamma > 0$ and the δ-functions are as defined in the previous subsection. The previous constraints indicate that the *xored* messages are considered as a part of the randomization indexes used to confuse the eavesdropper. On the other hand, for a certain realization $\mathcal{C}^1 \in \mathfrak{C}^1$, we have

$$\begin{aligned}\mathbb{I}(\mathrm{M}_{13};\mathrm{Z}^n|\mathrm{M}_{11}\mathrm{M}_{12}) &\overset{(a)}{=} \mathbb{H}(\mathrm{M}_{13}) - \mathbb{H}(\mathrm{M}_{13}|\mathrm{Z}^n\mathrm{M}_{11}\mathrm{M}_{12}) \\ &\overset{(b)}{\leq} \mathbb{H}(\mathrm{M}_{13}) - \mathbb{H}(\mathrm{M}_{13}|\mathrm{M}_{11}\mathrm{M}_{12}\mathrm{M}_r\mathrm{M}_{r_1}\mathrm{M}_{r_2}\mathrm{M}_{\otimes_1}\mathrm{M}_{\otimes_2}\mathrm{M}_t) \\ &\overset{(c)}{=} \mathbb{H}(\mathrm{M}_{13}) - \mathbb{H}(\mathrm{M}_{13}|\mathrm{M}_{\otimes_1}) \overset{(d)}{=} 0, \end{aligned} \tag{4.85}$$

where (a) follows from the independence of the messages; (b) follows because $(\mathrm{M}_{11}, \mathrm{M}_{21}, \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{M}_{r_2}, \mathrm{M}_{\otimes_1}, \mathrm{M}_{\otimes_2}, \mathrm{M}_t) - \mathrm{X}^n - \mathrm{Z}^n$ forms a Markov chain along with the fact that under the rate constraints in (4.84), there exists a code realization $\mathcal{C}^1$ such that $\mathbb{I}(\mathrm{M}_{21}\mathrm{M}_{22};\mathrm{Z}^n|\mathrm{M}_{11}\mathrm{M}_{12})$ decays exponentially in n. These two arguments imply that given Z^n and the messages $(\mathrm{M}_{11}, \mathrm{M}_{12})$, the best decoder at the eavesdropper can at most decode the messages $(\mathrm{M}_r, \mathrm{M}_{r_1}, \mathrm{M}_{r_2}, \mathrm{M}_{\otimes_1}, \mathrm{M}_{\otimes_2}, \mathrm{M}_t)$ but neither M_{21} nor M_{22}. Step (c) follows because M_{13} is only related to the xored message $\mathrm{M}_{\otimes_1}$; while (d) follows due to the secret key encoding principle and the fact that $\mathbb{H}(\mathrm{M}_{13}) = \mathbb{H}(\mathrm{M}_{21})$, cf. (4.77).

Based on similar arguments, we can show that for a certain realization $\mathcal{C}^1 \in \mathfrak{C}^1$ the following holds:

$$\mathbb{I}(\mathrm{M}_{23};\mathrm{Z}^n|\mathrm{M}_{21}\mathrm{M}_{22}) = 0. \tag{4.86}$$

The previous secrecy analysis implies that under the rate constraints in (4.84), our coding scheme satisfies the individual-strong secrecy constraint in (4.10c) for $k = 2$.

8. Fourier-Motzkin Elimination: We now combine the reliability rate constraints in (4.82) and the secrecy rate constraints in (4.84), then apply the Fourier-Motzkin Elimination procedure and solve for the confidential rates $R_{11} + R_{12}$ and $R_{21} + R_{22}$. Finally, we take the limit as $n \to \infty$ and choose the constants $\epsilon, \beta, \gamma > 0$, such that their corresponding $\delta-$functions approach zero as n approaches infinity. We reach the following:

$$R_{11} + R_{12} + R_{21} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}) \tag{4.87a}$$

$$R_{21} + R_{22} + R_{11} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}) - \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) \tag{4.87b}$$

$$R_{11} + R_{12} + R_{21} + R_{22} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_0) - R_e \tag{4.87c}$$

$$R_{11} + R_{12} + R_{21} + R_{22} \leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - R_e \tag{4.87d}$$

$$2R_{11} + R_{12} + 2R_{21} + R_{22} \leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}) - R_e \tag{4.87e}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0)$. Next, we let $R_r = R_{r_1} = R_{r_2} = 0$ and use the rate constraints in (4.84) with equality to bound the *xored* rates $R_{\otimes_1}$ and $R_{\otimes_2}$ which consequently implies bounding R_{13} and R_{23}, cf. (4.77). Thus, we have:

$$R_{13} = R_{\otimes_1} \leq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}), R_{21}\right] \tag{4.88a}$$

$$R_{23} = R_{\otimes_2} \leq \min\left[\mathbb{I}(\mathrm{V}_0;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_2;\mathrm{V}_1\mathrm{Z}|\mathrm{V}_0) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0), R_{11}\right] \tag{4.88b}$$

$$R_{13} + R_{23} = R_{\otimes_1} + R_{\otimes_2} \leq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}), R_{11} + R_{21}\right] \tag{4.88c}$$

Finally, if we combine the rate constraints in (4.87) and (4.88) along with the rate splitting relation in (4.76), we prove the achievability of any rate pair $(R_1 = R_{11} + R_{21}, R_2 =$

$R_{21} + R_{22}$) that satisfy the rate constraints in (4.75) under the individual-strong secrecy criterion; establishing Proposition 4.3.

The previous achievability proof directly suggests that there exists an alternative random coding scheme C^2 which also satisfies the reliability and secrecy constraints given by (4.10b) and (4.10c) respectively, while its rates are bounded according to the following proposition:

Proposition 4.4. *An achievable individual-strong secrecy rate region for the two-receiver DM-WBC is given by the set of all rate pairs* $(R_1 = R_{11} + R_{12}, R_2 = R_{21} + R_{22}) \in R_+^2$ *that satisfy*

$$\begin{aligned}
R_1 + R_{21} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{Z}) \qquad (4.89\text{a})\\
&\quad + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0\mathrm{Z}) - \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0), R_{21}\right]\\
R_2 + R_{11} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Z}) + \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Z}), R_{11}\right] \qquad (4.89\text{b})\\
R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_2;\mathrm{Y}_2|\mathrm{V}_0) - R_e + R_s \qquad (4.89\text{c})\\
R_1 + R_2 &\leq \mathbb{I}(\mathrm{V}_1;\mathrm{Y}_1|\mathrm{V}_0) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - R_e + R_s \qquad (4.89\text{d})\\
R_1 + R_2 + R_{11} + R_{21} &\leq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1;\mathrm{Y}_1) + \mathbb{I}(\mathrm{V}_0\mathrm{V}_2;\mathrm{Y}_2) - \mathbb{I}(\mathrm{V}_0;\mathrm{Z}) - R_e + R_s \qquad (4.89\text{e})
\end{aligned}$$

where $R_e \triangleq \mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}) + \mathbb{I}(\mathrm{V}_1;\mathrm{V}_2|\mathrm{V}_0)$*, while* $R_s \triangleq \min\left[\mathbb{I}(\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2;\mathrm{Z}), R_{11} + R_{21}\right]$*. The random variables that define the previous rate region are characterized by the following joint probability distribution:* $\mathrm{Q}_{\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{V}_1\mathrm{V}_2|\mathrm{V}_0}\ \mathrm{Q}_{\mathrm{X}|\mathrm{V}_0\mathrm{V}_1\mathrm{V}_2}\ Q_{\mathrm{Y}_1\mathrm{Y}_2\mathrm{Z}|\mathrm{X}}$*, such that* $\mathrm{V}_0 - (\mathrm{V}_1, \mathrm{V}_2) - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Z})$ *forms a Markov chain.*

The previous proposition is a direct consequence of Proposition 4.3, as it follows by interchanging the construction procedure of the random codewords V_1^n and V_2^n along with modifying the decoding functions at the first and the second legitimate receivers accordingly. In order to finalize the proof of Theorem 4.6, we use the rate splitting and time sharing technique introduced in Section 3.2.5 to construct a coding scheme that combines the two achievable rate regions in (4.75) and (4.89).

Chapter 5

Wiretap Channels With Active Adversaries

This chapter considers a communication scenario in which the channel undergoes two different classes of attacks at the same time: a passive eavesdropper and an active jammer. This scenario can be modelled as a secure communication problem over a wiretap channel with imperfect CSI. In particular, we focus our investigation to the class of arbitrarily varying wiretap channels (AVWCs). This class of channels simulate a wiretap channel in which the channel state is manipulated by an active adversary causing it to vary from one channel use to the other in an unknown and arbitrary manner. First, we introduce the attributes of an AVWC and discuss the two main coding schemes (deterministic and correlated random) used in previous literature to establish a reliable and secure communication over AVWCs. We then highlight the drawbacks of these schemes and foster an alternative coding scheme based on the principle of list decoding. We derive a full characterization of the list secrecy capacity of the AVWC, showing that list codes can over come the drawbacks of the earlier coding schemes and simultaneously provide the same secrecy capacity. Moreover, we establish some interesting results for the continuity and additivity behaviour of the list secrecy capacity of AVWCs. Finally, we extend our investigation to a scenario that combines the transmission of public and confidential messages over AVWCs. For this scenario, we derive a multi-letter description for the deterministic and correlated random capacity regions showing the optimality of superposition encoding.

5.1 Channel Models With Imperfect CSI

The previous three chapters of this thesis addressed the problem of secure communication under the following assumption: Perfect channel state information (CSI) is available at the beginning of the transmission and not only for the legitimate channel but for the eavesdropper channel as well. In real life communication scenarios and in particular wireless communication scenarios, it is not an easy task to acquire a perfect CSI. This is because the channel usually varies rapidly over time. The imperfection in the CSI can also originate from inaccurate channel estimation techniques or insufficient feedback schemes [149].

5.1.1 Motivation: Active Jamming

In the secrecy-domain, the acquisition of perfect CSI is more a challenging task. This is because for many wiretap channels, the eavesdropper is an intruder node that does not belong to the network and will consequently deny to share any information about its state [150]. This is not the only reason because: Even under the assumption of perfect cooperation of the eavesdropping node, secure communication suffers from other factors that can lead to an uncertainty in the CSI of the channel. Among these factors is the presence of active adversaries in the network who can maliciously manipulate the channel state. These two reasons motivated a lot of researchers to study the impact of dealing with imperfect CSI on the secrecy capacity of wiretap channels [151, 152].

The model of the wiretap channel introduced in Section 2.1.2 considers a communication scenario in which the channel only suffers from passive attacks. In particular, the confidential message is transmitted over a channel where a non-legitimate receiver eavesdrop over the transmission and aims to extract any information about the confidential message. In real life scenario, there exists other classes of attacks such as active jamming, where an active adversary threatens the secrecy of the communication by maliciously manipulating the channel state from time to time [153]. This observation motivates us to investigate secure communication under a combination of the two classes of attacks at the same time: passive eavesdropping and an active jamming. This scenario is identical to the concept of wiretap channels with imperfect CSI.

In previous literature, two main channel classes have been used to simulate and model a channel with imperfect CSI: Compound channels [72] and Arbitrary varying channels [154]. The main concept behind these two classes is that: Instead of assuming that the exact channel realization is perfectly known, we assume that the actual channel state belongs to a specific uncertainty set of channels known as the channel states set. The basic definitions and characteristics of these two channel classes are addressed in the next sections.

5.1.2 Compound Wiretap Channels

Compound channels are among the simplest non-trivial channels that models a channel with imperfect CSI. The concept of compound channels was first introduced for the broadcast channel in [72, 155] as follows: Let $\mathcal{S}$ be a finite state set, while $\mathcal{X}$ and $\mathcal{Y}$ represent finite input and output alphabets respectively. For every state $s \in \mathcal{S}$, the channel between the transmitter and the receiver is given by the stochastic matrix $W_s : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$. The compound broadcast channel is defined in terms of the families of all possible channel states as follows: $\mathcal{W} = \{W_s : s \in \mathcal{S}\}$. Now, for an input sequence $x^n \in \mathcal{X}^n$ produced by the transmitter, an output sequence $y^n \in \mathcal{Y}^n$ observed by the receiver and a certain channel state s, the discrete memoryless compound broadcast channel is given by:

$$W_s^n(y^n|x^n) = \prod_{i=1}^{n} W_s(y_i|x_i) \tag{5.1}$$

The previous equation reveals the most important feature of a compound channel: Although the channel can take any state from the state set $\mathcal{S}$, the selected state s remains fixed throughout the whole transmission. Moreover, it is assumed that the

transmitter and the receiver know the channel states set $\mathcal{S}$ but possess no information about the exact channel state s used during the transmission. Furthermore, there is not any prior distribution on the channel state set $\mathcal{S}$ that govern the selection of the channel states. It was shown in [72,155] that the capacity of the compound broadcast channel is given by:

$$C(\mathcal{W}) = \max_{\mathrm{P_X}} \min_{s\in\mathcal{S}} \mathbb{I}(\mathrm{X};\mathrm{Y}_s), \tag{5.2}$$

where Y_s is the output of the channel when the channel state s is selected. One can easily show that $C(\mathcal{W})$ is in general less than the transmission capacity of all the DMCs in $\mathcal{W}$, i.e. $C(\mathcal{W}) \leq \min_{s\in\mathcal{S}} C(W_s)$. This implies that the imperfection in the CSI has reduced the transmission capacity of the system. This is because the used coding scheme needs be universal in the sense that it works for all possible channel states.

In [152], the compound wiretap channel was introduced by combining the concepts of compound channels and the classical wiretap channel as follows: Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two finite state sets, then for the finite input and output alphabets $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, the stochastic matrices $W_{s_1} : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V_{s_2} : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$ define the legitimate and eavesdropper channels at a channel state pair $(s_1, s_2) \in \mathcal{S}_1 \times \mathcal{S}_2$. One can define a combined channel state $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2$, such that any state $s \in \mathcal{S}$ is given by the corresponding pair (s_1, s_2). The compound wiretap channel can then be defined in terms of the families of all possible channel states s as follows:

$$\begin{aligned}\mathcal{Q} = (\mathcal{W}, \mathcal{V}) &= \big\{(W_s, V_s) : s \in \mathcal{S}\big\} \\ &= \big\{(W_{s_1}, V_{s_2}) : s_1 \in \mathcal{S}_1 \text{ and } s_2 \in \mathcal{S}_2\big\}.\end{aligned} \tag{5.3}$$

The discrete memoryless compound wiretap channel is defined by following the same procedure in Eq. (5.1). Now, to transmit a confidential message $m_c \in \mathcal{M}_c$, we need to define a coding scheme $\mathcal{C}^{\mathrm{s}}$ with a stochastic encoder $E : \mathcal{M}_c \to \mathcal{P}(\mathcal{X}^n)$ and a deterministic decoder $\varphi : \mathcal{Y}^n \to \mathcal{M}$. The encoder and decoder of this coding scheme should be universal in the sense that they work for all possible channel states $s \in \mathcal{S}$. In order to evaluate the reliability performance of $\mathcal{C}^{\mathrm{s}}$, we define the average decoding error probability for the compound wiretap channel as follows:

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}) = \max_{s\in\mathcal{S}} \frac{1}{|\mathcal{M}_c|} \sum_{m_c\in\mathcal{M}_c} \sum_{x^n\in\mathcal{X}^n} \sum_{y^n:\varphi(y^n)\neq m_c} W_s(y^n|x^n)E(x^n|m_c). \tag{5.4}$$

The coding scheme $\mathcal{C}^{\mathrm{s}}$ is considered reliable, if it assures a small average decoding error probability for all channel states. It is important to highlight here that the results established under the average decoding error probability constraint might not be valid for the maximum decoding error probability [54]. On the other hand, the secrecy performance of $\mathcal{C}^{\mathrm{s}}$ is granted by satisfying the strong secrecy criterion with respect to the information leakage as follows:

$$\max_{s\in\mathcal{S}} \mathbb{I}(\mathrm{M};\mathrm{Z}_s^n) \leq \tau_n, \tag{5.5}$$

where $\lim_{n\to\infty} \tau_n = 0$, while Z_s^n is a random variable associated with the output sequence at the eavesdropper when the channel V_s is selected.

The secrecy capacity of the compound wiretap channel was investigated in various works cf. [30, 152, 156, 157]. Despite these tremendous efforts, finding a single-letter characterization for the secrecy capacity of the general compound wiretap channel has remained an unanswered question. This might be expected because the compound wiretap channel can be interpreted as a normal wiretap channel were a common confidential message is transmitted to multiple receivers and needs to be protected from multiple eavesdroppers. We already highlighted in Chapter 3, that even for the wiretap channel with multiple legitimate receivers and a single eavesdropper, a single-letter characterization for the secrecy capacity is still unknown. Nevertheless, in [30, Proposition 3.8] a multi-letter characterization for the secrecy capacity was established.

Theorem 5.1. *[30] The strong secrecy capacity of the general compound wiretap channel $\mathcal{Q} = (\mathcal{W}, \mathcal{V})$ is given by:*

$$C(\mathcal{Q}) = \lim_{n\to\infty} \frac{1}{n} \max_{\mathrm{P}_{\mathrm{UX}^n}} \big[\min_{s\in\mathcal{S}} \mathbb{I}(\mathrm{U};\mathrm{Y}_s^n) - \max_{s\in\mathcal{S}} \mathbb{I}(\mathrm{U};\mathrm{Z}_s^n) \big], \tag{5.6}$$

where the maximum is taken over all possible input distributions $\mathrm{P}_{\mathrm{UX}^n}$ defined over the random variables U, X^n.

Moreover, it was shown in [30, 152] that for the degraded compound wiretap channel where all channel realization to the eavesdropper are degraded with respect to any channel realization to the legitimate receiver, the multi-letter expression in Theorem 5.1 can be further simplified to establish a single-letter characterization for the secrecy capacity of this class of compound wiretap channels.

Theorem 5.2. *[30, 152] The strong secrecy capacity of the degraded compound wiretap channel is given by:*

$$C(\mathcal{Q}) = \max_{\mathrm{P}_{\mathrm{X}}} \big[\min_{s\in\mathcal{S}} \mathbb{I}(\mathrm{X};\mathrm{Y}_s) - \max_{s\in\mathcal{S}} \mathbb{I}(\mathrm{X};\mathrm{Z}_s) \big], \tag{5.7}$$

where the maximum is taken over all possible input distributions P_{X} such that $\mathrm{X} - \mathrm{Y}_s - \mathrm{Z}_s$ forms a Markov chain.

The previous theorem reflects the effect of the imperfection in the CSI on the secrecy capacity as follows: In order to guarantee a reliable link between the transmitter and the legitimate receiver for all channel states $s \in \mathcal{S}$, the maximum transmission rate should be bounded by the smallest rate among all the channel states, i.e. $\min_{s\in\mathcal{S}} \mathbb{I}(\mathrm{X};\mathrm{Y}_s)$. On the other hand, in order to make sure that the eavesdropper is not capable of inferring any information about the transmitted message, we need to randomize our encoding process to confuse the eavesdropper. This is done by using a randomization message with rate roughly equal to the best channel resources available at the eavesdropper, i.e. $\max_{s\in\mathcal{S}} \mathbb{I}(\mathrm{X};\mathrm{Z}_s)$.

5.1.3 Arbitrary Varying Wiretap Channels

An arbitrary varying channel (AVC) models a sophisticated and more complex class of channels with imperfect CSI. It extends the concept of compound channels by allowing the channel state s to alternate through the transmission. In particular, the channel state of an AVC can vary from one time to another in an arbitrarily manner during the transmission. This behavior simulates a lot of real life communication channels

such as fast fading channels. The concept of arbitrary varying channels was first introduced in [154, 158] as follows: Let $\mathcal{S}$ be a finite state set, $\mathcal{X}$ and $\mathcal{Y}$ be two finite input and output alphabets, such that for a channel state $s \in \mathcal{S}$, the stochastic matrix $W_s : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ defines the relation between the input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{Y}$. In some literature, a different notation was used to define the stochastic matrix that guides the relation between the inputs an outputs of an AVC as follows: $W : \mathcal{X} \times \mathcal{S} \to \mathcal{P}(\mathcal{Y})$. The difference between the previous two notations lies in the interpretation of the channel state set $\mathcal{S}$. The first notation considers s as a constraint enforced by the channel, while the second notation considers s as an additional input to the channel. Nevertheless, one can easily show that the two notation are equivalent.

Although the previous definition seems similar to the definition of compound channels, the extension to the block length channel model clearly mark the difference as follows: The discrete memoryless transition matrix with dimension n that guides the relation between the input sequence $x^n \in \mathcal{X}^n$ and the output sequence $y^n \in \mathcal{Y}^n$ of an AVC is given by:

$$W^n_{s^n}(y^n|x^n) = \prod_{i=1}^{n} W_{s_i}(y_i|x_i), \tag{5.8}$$

where $s^n \in \mathcal{S}^n$ is the channel state sequence selected by the channel. The discrete memoryless (AVC) is defined in terms of the families of all possible channel state sequences for different values of n as follows $\mathfrak{W} = \{W^n(\cdot|\cdot, s^n) : s^n \in \mathcal{S}^n, n = 1, 2, \dots\}$. Similar to the model of compound channels, it is assumed that the transmitter and the receiver only know the channel states set $\mathcal{S}$ but possess no information about the actual channel state sequence s^n selected for the transmission. The model of an AVC can be interpreted from a secrecy perspective as a normal DMC but with an additional active jammer. This jammer aims to disturb the communication on the channel by choosing a certain channel state sequence s^n [159].

Eq. (5.8) implies that response of an AVC depends on the channel state sequence s^n selected by the jammer. Nevertheless, one can argue that the behavior of an AVC should only depend on the number of times each channel state $s \in \mathcal{S}$ is imposed, and not on the order of these states. This argument is due to the memoryless property of the channel and it means that different state sequences might have the same channel response. This observation motivates the introduction of the average channel notation defined as follows: For any probability distribution $q \in \mathcal{P}(\mathcal{S})$, the average channel is given by:

$$\overline{W_q}(y|x) = \sum_{s \in \mathcal{S}} q(s) W_s(y|x) \tag{5.9}$$

In some literature, the average channel is defined over a block length n, such that for any probability distribution $\tilde{q} \in \mathcal{P}(\mathcal{S}^n)$, we have

$$\overline{W^n_{\tilde{q}}}(y^n|x^n) = \sum_{s^n \in \mathcal{S}^n} \tilde{q}(s^n) W^n_{s^n}(y^n|x^n) \tag{5.10}$$

The notation of the average channel establishes some sort of a link between AVCs and compound channels. This is because an AVC with a state sequence $s^n \in \mathcal{S}^n$ can be viewed as a compound channel with a channel state $\tilde{q}$, where $\tilde{q}$ is the average channel for s^n.

Despite the previous argument, it was shown in [160–162] that establishing a reliable communication over an AVC is not as easy as the compound channel. In fact, it was shown in [160] that the classical deterministic codes with a fixed encoder-decoder pair used to establish reliable communication over compound channels fails to achieve this target for a class of AVC known as symmetrizable AVC. For this class of channels, more sophisticated coding schemes like correlated random codes are needed. In correlated random codes, the transmitter and the receiver agree on a family of encoder-decoder pairs before the transmission. Then for every communication on the channel, the transmitter and the receiver coordinate their choice of an encoder-decoder pair based on the outcome of a random experiment shared between them. In [154, 160], the correlated random public transmission capacity of the AVC was established:

$$C^{\mathrm{p}}_{\mathrm{ran}}(\mathfrak{W}) = \max_{\mathrm{P_X}} \min_{q\in\mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X};\bar{\mathrm{Y}}_q), \tag{5.11}$$

where $\bar{\mathrm{Y}}_q$ is a random variable that represents the output of the averaged channel $\bar{W}_q(y|x)$ given by (5.9). In [163], it was shown that deterministic capacity of AVCs exhibits a dichotomy that it is either equivalent to the correlated random capacity or it vanishes. However, the exact condition for a vanishing deterministic capacity was not answered in [163]. In [164], Ericson came up with a sufficient which was shown to be necessary in [165]. This condition is known as the symmetrizablity condition of an AVC.

The previous results implies that one of the main properties that plays an important role in the investigation of reliable communication over AVCs is the concept of symmetrizability. This property describes the ability of an AVC to emulate a valid channel input making it impossible for the decoder to make a correct decision for the actual transmitted codeword. It was first introduced in [164, 165] as follows:

Definition 5.1. An AVC $\mathfrak{W}$ is symmetrizable if there exists an auxiliary channel $\sigma : \mathcal{X} \to \mathcal{P}(\mathcal{S})$ such that

$$\sum_{s\in\mathcal{S}} W_s(y|x)\sigma(s|\tilde{x}) = \sum_{s\in\mathcal{S}} W_s(y|\tilde{x})\sigma(s|x) \tag{5.12}$$

holds for every $x, \tilde{x} \in \mathcal{X}$ and $y \in \mathcal{Y}$.

The condition in (5.12) implies that for a symmetrizable AVC, the jammer can select a certain jamming strategy, i.e. a channel state sequence $s^n \in \mathcal{S}^n$, such that the decoder is confused. In order to understand this concept, the following example for a symmetrizable AVC was given [166]: Consider an AVC, where $\mathcal{S} = \mathcal{X} = \{1, 2\}$, while $\mathcal{Y} = \{1, 2, 3\}$. For $\epsilon \in [0, 1]$, our AVC $\mathfrak{W}$ is defined by the following two stochastic matrices:

$$W_1 := \begin{pmatrix} 0 & 1-\epsilon \\ \epsilon & 0 \\ 1-\epsilon & \epsilon \end{pmatrix} \qquad W_2 := \begin{pmatrix} 1-\epsilon & 0 \\ \epsilon & 1-\epsilon \\ 0 & \epsilon \end{pmatrix}.$$

One can easily show that for every $\epsilon \in [0, 0.5]$, there exists an auxiliary channel $\sigma : \mathcal{X} \to \mathcal{P}(\mathcal{S})$, such that the symmetrizability condition in (5.12) is satisfied. This is because the condition in (5.12) implies the following system of equations:

$$W_1(\cdot|1)\sigma(1|2) + W_2(\cdot|1)\sigma(2|2) = W_1(\cdot|2)\sigma(1|1) + W_2(\cdot|2)\sigma(2|1)$$

This system of linear equations leads to the following auxiliary channel σ: $\sigma(1|1) \triangleq \epsilon/(1-\epsilon)$ and $\sigma(1|2) \triangleq (1-2\epsilon)/(1-\epsilon)$. This implies that for this AVC, the receiver can not decide whether the channel input was $x = 1$ at a channel state $s = 2$ or the channel input was $x = 2$ at a channel state $s = 1$.

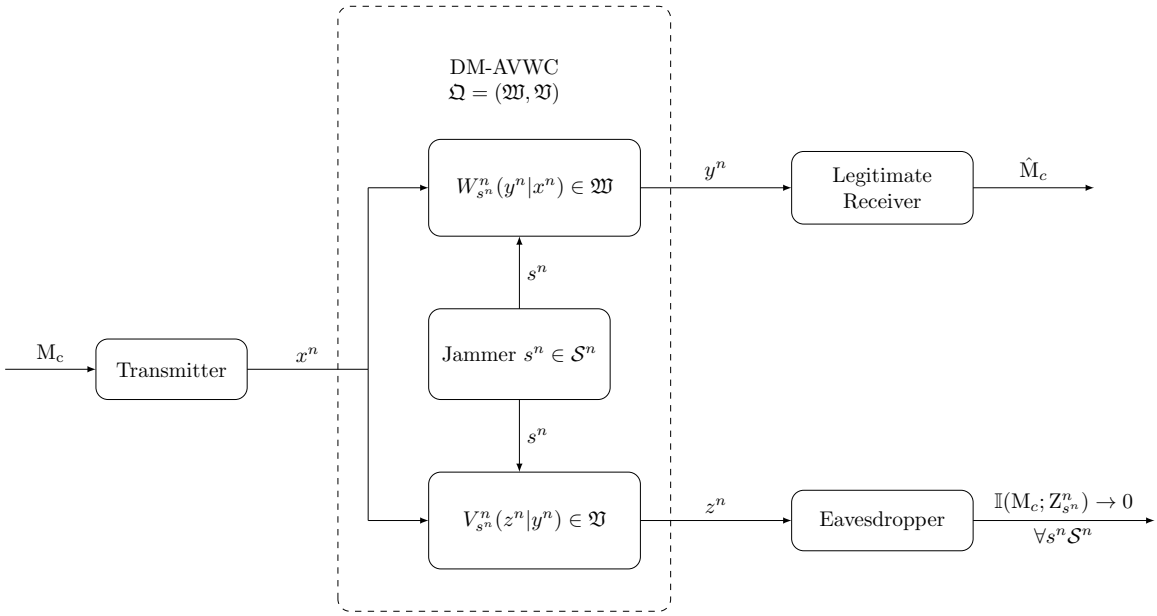


Figure 5.1: Discrete memoryless arbitrary varying wiretap channel

In [13], the concept of AVC was extended to the secrecy domain leading to the arbitrary varying wiretap channel (AVWC). The AVWC can be interpreted as a channel that undergoes two different classes of attacks at the same time: A passive eavesdropper that threatens the secrecy of the communication by eavesdropping upon the transmitted signal and an active jammer that threatens the reliability of the communication by maliciously manipulating the channel state. The model of the AVWC is shown in Fig. 5.1 and is defined as follows: Let $\mathcal{S}$ be a finite state set, while $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ represent finite input and output alphabets. For every state $s \in \mathcal{S}$, the legitimate receiver channel is given by the stochastic matrix $W_s : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, while the eavesdropper channel is given by the stochastic matrix $V_s : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$. Thus, for a state sequence $s^n \in \mathcal{S}^n$ of length n produced by the active jammer, the following discrete memoryless channels can be defined

$$W^n_{s^n}(y^n|x^n) = \prod_{i=1}^{n} W_{s_i}(y_i|x_i) \qquad \text{and} \qquad V^n_{s^n}(z^n|x^n) = \prod_{i=1}^{n} V_{s_i}(z_i|x_i). \tag{5.13}$$

where $x^n \in \mathcal{X}^n$ is the input sequence of the channel, while $y^n \in \mathcal{Y}^n$ and $z^n \in \mathcal{Z}^n$ are output sequences at the legitimate receiver and the eavesdropper respectively. An AVWC is given by the collection of all channels in (5.13) for all possible state sequences s^n as follows:

Definition 5.2. The discrete memoryless AVWC $\mathfrak{Q}$ is denoted by the pair $(\mathfrak{W}, \mathfrak{V})$ and is given by the family of pairs with common input as

$$\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V}) = \big\{(W^n_{s^n}, V^n_{s^n}) : s^n \in \mathcal{S}^n, n = 1, 2, \dots\big\}, \tag{5.14}$$

where $\mathfrak{W}$ represents the AVC between the transmitter and the legitimate receiver, while $\mathfrak{V}$ represents the AVC between the transmitter and the eavesdropper.

An AVWC models a very challenging communication scenario, yet it is of high importance. This is because the model of an AVWC combines various features at the same time:

1. It mimic a very realistic and typical situation for wireless communication where the channel varies rapidly over time.
2. It simulates the most two common attacks in secure communication: passive eavesdropping and active jamming.
3. It includes other important channel models such as compound wiretap channels or classical wiretap channels as special cases.

In the next sections, we will discuss the different coding techniques used to establish a reliable and secure communication over AVWCs. In particular, we will address the secrecy capacity provided by each coding scheme along with the advantages and disadvantages of this scheme.

5.2 Deterministic and Correlated Random Secrecy Capacity for AVWCs

In this section, we highlight the main results established in previous literature for secure communication over AVWCs. We start by introducing a formal definition for the two commonly used coding schemes: deterministic and correlated random codes. We then present the secrecy capacities provided by both coding schemes and discuss their strengths and weaknesses. Finally, we discuss some of the analytical properties of the established capacity functions.

5.2.1 Coding Schemes: Basic Definitions

Consider a communication scenario in which a confidential message M_c is transmitted over an AVWC $\mathfrak{Q}$ as shown in Fig. 5.1. For a given message $m_c \in \mathcal{M}_c$, the transmitter produces a sequence $x^n \in \mathcal{X}^n$ and transmits it. In the mean time, the active jammer chooses a channel state sequence $s^n \in \mathcal{S}^n$ independently from both m_c and x^n. The legitimate receiver observes the sequence $y^n \in \mathcal{Y}^n$ and uses it to predict the transmitted confidential message m_c. On the other hand, the eavesdropper observes the sequence $z^n \in \mathcal{Z}^n$ which should carry no information about the confidential message m_c.

The investigation of the previous communication scenario has captured a lot of attention in the last few years, cf. [167] and references therein. Researchers aimed to develop coding schemes that can deliver two main tasks: It can overcome the different jamming strategies induced by the jammer establishing a reliable communication between the transmitter and the legitimate receiver. It should be able to keep the eavesdropper completely ignorant about the information transmitted. Two main coding techniques have been used to provide these two tasks. The first technique is based on deterministic codes, in which a predefined (fixed) encoder-decoder pair is used through out the whole transmission. The second technique is known as correlated random codes, in which an encoder-decoder pair is selected based on some sort of common randomness shared between the transmitter and the legitimate receiver.

Definition 5.3. A deterministic secrecy code $\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}$ for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ consists of: a set of confidential messages $\mathcal{M}_c$, a fixed stochastic encoder

$$E : \mathcal{M}_c \to \mathcal{P}(\mathcal{X}^n) \tag{5.15}$$

that maps a confidential message $m_c \in \mathcal{M}_c$ to a codeword $x^n(m_c) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m_c)$, and a fixed deterministic decoder

$$\varphi : \mathcal{Y}^n \to \mathcal{M}_c \cup \{?\} \tag{5.16}$$

that maps each channel observation at the legitimate receiver to the corresponding intended message or an error message.

In order to measure the reliability performance of a deterministic secrecy code $\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}$, we can use the average decoding error probability given by:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) &= \max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) \\ &= \max_{s^n \in \mathcal{S}^n} \frac{1}{|\mathcal{M}_c|} \sum_{m_c} \sum_{x^n} \sum_{y^n : \varphi(y^n) \neq m_c} W^n_{s^n}(y^n|x^n) E(x^n|m_c).\end{aligned} \tag{5.17}$$

Alternatively, we can also used the maximum decoding error probability where the expectation over the message set $\mathcal{M}_c$ is replaced by a maximization as follows:

$$\mathbf{P}^{\max}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) = \max_{s^n \in \mathcal{S}^n} \max_{m_c \in \mathcal{M}_c} \sum_{x^n} \sum_{y^n : \varphi(y^n) \neq m_c} W^n_{s^n}(y^n|x^n) E(x^n|m_c). \tag{5.18}$$

On the other hand, the secrecy performance of $\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}$ is measured by investigating the information leakage of the confidential message to the eavesdropper for every state sequence $s^n \in \mathcal{S}^n$ with respect to the strong secrecy criterion as follows:

$$\begin{aligned}\mathbb{L}(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) &= \max_{s^n \in \mathcal{S}^n} \mathbb{L}(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) \\ &= \max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n}|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}),\end{aligned} \tag{5.19}$$

where M_c represents a uniformly distributed random variable over the confidential messages set $\mathcal{M}_c$, while $\mathrm{Z}^n_{s^n}$ is a random variable for the channel observation at the eavesdropper for state sequence s^n.

Definition 5.4. A non-negative number R_c is an achievable strong secrecy rate for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ with respect to the average decoding error criterion, if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\tau, \lambda, \delta)$, there exists a sequence of deterministic codes $(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}})_n$ that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_c| \geq R_c - \delta, \tag{5.20}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) \leq \lambda, \tag{5.21}$$

$$\mathbb{L}(\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}) \leq \tau. \tag{5.22}$$

The deterministic strong secrecy capacity $C^{\mathrm{s}}_{\mathrm{det,avg}}(\mathfrak{Q})$ with respect to the average decoding error criterion is given by the supremum of all achievable rates R_c.

Based on the previous definition, we can use $\mathbf{P}_e^{\max}$ to define the deterministic strong secrecy capacity $C^{\mathrm{s}}_{\mathrm{d,max}}(\mathfrak{Q})$ with respect to the maximum decoding error probability. Unlike the classical wiretap channel where these two capacities are equivalent, it was shown in [163] that for AVC and consequently for AVWC as well, $C^{\mathrm{s}}_{\mathrm{det,avg}}(\mathfrak{Q})$ and $C^{\mathrm{s}}_{\mathrm{det,max}}(\mathfrak{Q})$ might not be the same. In fact, although the deterministic transmission capacity $C^{\mathrm{p}}_{\mathrm{det,avg}}(\mathfrak{W})$ of AVCs with respect to the average decoding error probability was established in [165], the capacity of AVCs under the maximum decoding error probability criterion $C^{\mathrm{p}}_{\mathrm{det,max}}(\mathfrak{W})$ is still unknown.

In [14], it was shown that the incapability of deterministic codes to deal with some class of AVCs extends to AVWCs as well. It was shown that deterministic codes might not be sufficient to establish reliable communication over all AVWCs. In particular, if the AVC $\mathfrak{W}$ between the transmitter and the legitimate receiver is symmetrizable, no reliable communication can be achieved using deterministic codes. Unfortunately, a lot of channels of practical relevance fall into the class of symmetrizable AVCs [168]. For these channels, more sophisticated coding techniques such as correlated random codes are needed.

Definition 5.5. A correlated random secrecy code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ consists of: a set of confidential messages $\mathcal{M}_c$, a set $\mathcal{G}$ of values for the common randomness, a probability distribution P_Γ that guides the selection of a certain realization of the common randomness $\gamma \in \mathcal{G}$, a set of stochastic encoders

$$E^\gamma : \mathcal{M}_c \to \mathcal{P}(\mathcal{X}^n), \tag{5.23}$$

that maps a confidential message $m_c \in \mathcal{M}_c$ to a codeword $x^n(m_c) \in \mathcal{X}^n$ according to the conditional probability $E^\gamma(x^n|m_c)$ and a set of deterministic decoders

$$\varphi^\gamma : \mathcal{Y}^n \to \mathcal{M}_c \cup \{?\}, \tag{5.24}$$

that maps each channel observation at the legitimate receiver to the corresponding intended message or an error message.

The previous definition implies that any correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ can be interpreted as a family of deterministic codes of Definition 5.3 as follows:

$$\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}} = \{\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma) : \gamma \in \mathcal{G}\}. \tag{5.25}$$

Similar to deterministic codes, we can evaluate the reliability performance of a correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ using either the average or the maximum decoding error probability. For the average decoding error, we have:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}) &= \max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}) \\ &= \max_{s^n \in \mathcal{S}^n} \frac{1}{|\mathcal{M}_c|} \sum_{m_c} \sum_{\gamma} \sum_{x^n} \sum_{y^n : \varphi^\gamma(y^n) \neq m_c} W^n_{s^n}(y^n|x^n) E^\gamma(x^n|m_c) P_\Gamma(\gamma).\end{aligned} \tag{5.26}$$

On the other hand, the secrecy performance of $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ can be evaluated with respect to two secrecy criteria. The first criterion is called the mean secrecy criterion and is measured by the investigating the average information leakage of the confidential message to the eavesdropper with respect to the strong secrecy criterion, for every

state sequence $s^n \in \mathcal{S}^n$ over all realizations of the shared common randomness $\gamma \in \mathcal{G}$ as follows:

$$\begin{aligned}\mathbb{L}^{\text{mean}}(\mathcal{C}^{\text{s}}_{\text{ran}}) &= \max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n} | \Gamma \mathcal{C}^{\text{s}}_{\text{ran}}) \\ &= \max_{s^n \in \mathcal{S}^n} \sum_{\gamma} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n}(\gamma) | \mathcal{C}^{\text{s}}_{\text{ran}}) P_{\Gamma}(\gamma), \end{aligned} \tag{5.27}$$

where $\mathrm{Z}^n_{s^n}(\gamma)$ is the channel output at the eavesdropper for a state sequence s^n, when the encoder E^{γ} is used. The second criterion is a more conservative criterion known as the maximum secrecy criterion and is given by

$$\begin{aligned}\mathbb{L}^{\max}(\mathcal{C}^{\text{s}}_{\text{ran}}) &= \max_{s^n \in \mathcal{S}^n} \mathbb{L}^{\max}(s^n | \mathcal{C}^{\text{s}}_{\text{ran}}) \\ &= \max_{s^n \in \mathcal{S}^n} \max_{\gamma \in \mathcal{G}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n}(\gamma) | \mathcal{C}^{\text{s}}_{\text{ran}}). \end{aligned} \tag{5.28}$$

It is important to highlight that both secrecy criteria assume that the eavesdropper has an access to the common randomness shared between the transmitter and the legitimate receiver. However, the mean secrecy criterion puts the distribution P_{Γ} into consideration, while the maximum secrecy criterion considers the information leakage for every realization γ. It is also fair to assume that the eavesdropper has access to the shared common randomness because otherwise the transmitter and the legitimate receiver can use this common randomness to implement alternative secure encoding schemes such as one time pad, where the common randomness is used to create a shared secret key [169, 170].

Definition 5.6. A non-negative number R_c is an achievable strong secrecy rate for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ with respect to the average decoding error criterion and the mean strong secrecy criterion, if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\tau, \lambda, \delta)$, there exists a sequence of correlated random codes $(\mathcal{C}^{\text{s}}_{\text{ran}})_n$ that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_c| \geq R_c - \delta, \tag{5.29}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\text{s}}_{\text{ran}}) \leq \lambda, \tag{5.30}$$

$$\mathbb{L}^{\text{mean}}(\mathcal{C}^{\text{s}}_{\text{ran}}) \leq \tau. \tag{5.31}$$

The correlated random strong mean secrecy capacity $C^{\text{s,mean}}_{\text{ran,avg}}(\mathfrak{Q})$ with respect to the average decoding error criterion is given by the supremum of all achievable mean strong secrecy rates R_c.

The previous definition can be used in combination with the other reliability and secrecy constraints to define the following secrecy capacities: $C^{\text{s,max}}_{\text{ran,avg}}(\mathfrak{Q})$, $C^{\text{s,mean}}_{\text{ran,max}}(\mathfrak{Q})$ and $C^{\text{s,max}}_{\text{ran,max}}(\mathfrak{Q})$. For these capacities, the following convention is used: The superscript is used to identify the selected secrecy criterion, while the subscript is used to identify the selected reliability criterion. From this point, we will mainly focus on the average decoding error probability. Thus, we will omit the subscript that identify the reliability measure used unless we want to stress that a certain result is valid for the maximum decoding error probability.

5.2.2 Correlated Random Secrecy Capacity: Robustification Technique

One of the main tools that was introduced to facilitate the construction of correlated random code for AVWCs is the *robustification* technique introduced by Ahlswede in [171]. Ahlswede used this technique to derive the correlated random transmission capacity of AVCs by establishing a link between correlated random codes for AVCs and deterministic codes for compound channels. This technique is based on the following lemma:

Lemma 5.1. *[171] For a given finite set $\mathcal{S}$ and an integer n, if a function $f : \mathcal{S}^n \to [0,1]$ satisfies*

$$\sum_{s^n \in \mathcal{S}^n} f(s^n) q(s_1) \dots q(s_n) \geq 1 - \epsilon \tag{5.32}$$

for all $q \in \mathcal{P}_0^n(S)$ and some $\epsilon \in [0,1]$, then for all sequences $s^n \in \mathcal{S}^n$, the following holds:

$$\frac{1}{n!} \sum_{\pi \in \Pi_n} f\big(\pi(s^n)\big) > 1 - 3(n+1)^{|\mathcal{S}|}\epsilon, \tag{5.33}$$

where Π_n is the set of all permutations over the set $\{1, 2, \dots, n\}$, such that $\pi(s^n)$ is a new arrangement of the sequence s^n based on the permutation π as follows: $\pi(s^n) = (s_{\pi(1)}, \dots, s_{\pi(n)})$.

In [14], the usage of the robustification technique was extended to the secrecy domain. It provided a way to establish an achievable correlated random secrecy rate for AVWCs based on the deterministic achievable secrecy rate for the compound wiretap channel presented in [30]. The first step in this investigation was to use the average channel notation in (5.10) to define the following discrete memoryless AVWC:

$$\overline{\mathfrak{Q}} = (\overline{\mathfrak{W}}, \overline{\mathfrak{V}}) = \big\{(\overline{W_q^n}, \overline{V_q^n}) : q \in \mathcal{P}(\mathcal{S}^n), n = 1, 2, \dots\big\}. \tag{5.34}$$

Without loss of generality, the channel state q is usually restricted to the set of all types over the alphabet $\mathcal{S}$ with length n, i.e. $q \in \mathcal{P}_0^n(\mathcal{S}^n)$. One should notice that for a fixed block length n, the AVWC $\overline{\mathfrak{Q}}$ simplifies to a compound wiretap channel $\mathcal{Q}$, where the channel state set is given by the set of all types of length n on the state set $\mathcal{S}$ of the corresponding AVWC $\mathfrak{Q}$ as follows:

$$\mathcal{Q} = \overline{\mathfrak{Q}}(n) = \big\{(\overline{W_q^n}, \overline{V_q^n}) : q \in \mathcal{P}_0^n(\mathcal{S}^n)\big\}. \tag{5.35}$$

It was shown in [14, Lemma 3.7] that under the average decoding error probability and the strong secrecy criteria, the secrecy capacity of the AVWC $\mathfrak{Q}$ and $\overline{\mathfrak{Q}}$ are equivalent, regardless of the coding scheme used. This result along with the robustification technique and the achievable deterministic secrecy rate of compound wiretap channel established in [30] were used to derive an achievable correlated random secrecy rate for the AVWC $\mathfrak{Q}$ as follows:

Theorem 5.3. *[14] For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, if there exists a best channel to the eavesdropper, a lower-bound for the correlated random strong secrecy capacity with respect to the average decoding error probability and the mean secrecy criterion is given by*

$$C_{ran,avg}^{s,mean}(\mathfrak{Q}) \geq \max_{\mathrm{P_{UX}}} \Big[\min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{U}; \mathrm{Y}_q) - \max_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{U}; \mathrm{Z}_q) \Big]. \tag{5.36}$$

where Y_q *and* Z_q *are the random variables associated with the output of the averaged channels* $\overline{W_q}$ *and* $\overline{V_q}$ *respectively, while the maximum is taken over all possible input distributions* P_{UX} *defined over the random variables* U *and* X.

One can notice that the previous theorem did not provide an achievable secrecy rate for the general AVWC. This is because, it enforces an additional constraint on the AVC $\mathfrak{V}$ between the transmitter and the eavesdropper. This constraint is the existence of a **best channel** to the eavesdropper. The concept of a best channel was defined as follows:

Definition 5.7. An AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ is said to have a best channel to the eavesdropper, if there exist a channel $V_{q^*} \in \overline{\mathfrak{V}}$, such that all other channels in $\overline{\mathfrak{V}}$ are degraded versions of V_{q^*}. In other words, if the output of any channel $V_q \in \overline{\mathfrak{V}}$ is denoted by Z_q, then it holds that $\mathrm{X} - \mathrm{Z}_{q^*} - \mathrm{Z}_q$ forms a Markov chain for all $q \in \mathcal{P}(S)$.

In order to understand the role played by the best channel constraint, we need to go through the proof of Theorem 5.3. The proof consists of two main steps: In the first step, the result established in [14, Lemma 3.7] is used to transform the coding problem from the AVWC: $\mathfrak{Q}$ to the corresponding AVWC: $\overline{\mathfrak{Q}}$ defined with respect to the averaged channels. For an arbitrary but fixed block length n, the problem of secure communication over the AVWC: $\overline{\mathfrak{Q}}$ simplifies to a secure communication problem over a compound wiretap channel $\mathcal{Q}$ as highlighted by Eq. (5.35). It was shown in [30, Theorem 3.6] that for the compound wiretap channel $\mathcal{Q}$ a deterministic wiretap coding scheme can achieve secrecy rates up to:

$$R \leq \min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{U}; \mathrm{Y}_q) - \max_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{U}; \mathrm{Z}_q). \tag{5.37}$$

One can observe that the previous rate is equivalent to the one given in Theorem 5.3 and this concludes the first part of the achievability proof. The second part of the proof aims to transform the deterministic code constructed for the compound wiretap channel $\mathcal{Q}$ to a correlated random code for the AVWC: $\overline{\mathfrak{Q}}$. This is done by applying the robustification technique given by Lemma 5.1 to the reliability and secrecy constraints. This step lead to two main constraints: The first is the existence of a best channel to the eavesdropper which is a consequence of using the compound component $\mathcal{V}$ to simulate the effect of the original eavesdropper's AVC $\overline{\mathfrak{V}}$. The second constraint is that the amount of shared correlated random between the transmitter and the receiving nodes is $n!$ which corresponds to the number of all possible permutation in the set Π_n. It was then showed in [14, Lemma 3.11] that the amount of correlated randomness can be reduced to the order of $\mathcal{O}(n^2)$.

In [153, Theorem 3], it was shown that the achievable secrecy rate given by Theorem 5.3 establishes the capacity function $C^{\mathrm{s,mean}}_{\mathrm{ran,avg}}(\mathfrak{Q})$ if the AVWC is strongly degraded with independent states and a best channel to the eavesdropper exists. This implies that Theorem 5.3 is an important step in investigating the problem of secure communication over AVWCs. Nevertheless, it is important to highlight that Theorem 5.3 suffered from three main drawbacks:

1. Instead of presenting a full characterization for the correlated random secrecy capacity of AVWCs, Theorem 5.3 only presented an achievable secrecy rate which serves as a lower bound for the capacity.

2. Theorem 5.3 only provided an achievable correlated random coding scheme for the average decoding error probability and the mean information leakage. The possibility of extending the coding scheme to the other reliability and secrecy constraints is not obvious.
3. The last drawback of Theorem 5.3 is that the established achievable secrecy rate is not achievable in general as it requires the availability of a best channel to the eavesdropper.

One might argue that the first and second drawbacks are acceptable. This is because even for the simpler case of the compound wiretap channel a single-letter characterization of the secrecy capacity is still open. However, the last drawback is a serious issue because it raises a lot of speculations about the existence of a universal coding scheme for the AVWC. These drawbacks motivates Wiese, Nötzel and Boche in [159] to present the following result:

Theorem 5.4. *[159] For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, the correlated random strong secrecy capacity with respect to the average decoding error probability and the mean secrecy criterion is given by:*

$$C^{s,mean}_{ran,avg}(\mathfrak{Q}) = \lim_{n\to\infty} \frac{1}{n} \max_{\mathrm{P}_{\mathrm{UX}^n}} \Big[\min_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{U};\mathrm{Y}^n_q) - \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{U};\mathrm{Z}^n_{s^n}) \Big]. \tag{5.38}$$

where $\mathrm{Z}^n_{s^n}$ is the random variable associated with the output of the channel $V^n_{s^n}$, while Y^n_q is the random variable associated with the output of the averaged channel $\overline{W^n_q}$. Moreover, the maximum is taken over all possible input distributions $\mathrm{P}_{\mathrm{UX}^n}$ defined over the random variables U and X^n.

The main idea of the proof of the previous theorem was the introduction of the auxiliary channel known as: the compound arbitrary varying wiretap channel (CAVWC). This channel consists of a compound channel $\mathcal{W} = \{W_q : q \in \mathcal{P}(\mathcal{S}^n)\}$ from the transmitter to the legitimate receiver and an AVC $\mathfrak{V}$ between the transmitter and the eavesdropper. Similar to the proof of Theorem 5.3, the proof of Theorem 5.4 also consists of two steps. The first step is to establish the deterministic secrecy capacity of the CAVWC $(\mathcal{W}, \mathfrak{V})$. This is done by applying the reliability analysis techniques used for the compound channel as in [30]. On the other hand, the secrecy analysis is performed by extending the strong secrecy techniques introduced in Section 2.2.3 to the setup of AVC. The next step of the proof is to apply the robustification technique to transform the constructed deterministic code into a correlated random code for the corresponding AVWC. In Section 5.4, a more general version of the proof will be given in details.

The work in [159] did not only provide the result established in Theorem 5.4, but it also showed that the correlated random secrecy capacity with respect to the mean information leakage is equivalent to the maximum one, i.e. $C^{\mathrm{s,mean}}_{\mathrm{ran,avg}}(\mathfrak{Q}) = C^{\mathrm{s,max}}_{\mathrm{ran,avg}}(\mathfrak{Q})$[1] [159, Theorem 6]. Additionally, it was shown in the proof of [166, Theorem 1] that the expression of the correlated random secrecy capacity in Theorem 5.4 does not change if the sets over which the minimum and the maximum are taken in (5.38) are replaced by different, but related ones. In particular, it was shown that $\max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{U};\mathrm{Z}^n_{s^n})$

[1]From this point, we will refer to the correlated random secrecy capacity of an AVWC as $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$, where it is intuitively implied that the average decoding error probability was used as the reliability measure along with the fact that the exact secrecy criterion used has no effect on the capacity.

can be replaced by $\max_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{U};\mathrm{Z}_q^n)$, while $\min_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{U};\mathrm{Y}_q^n)$ can be replaced by $\min_{\tilde{q}\in\mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{U};\mathrm{Y}_{\tilde{q}}^n)$. This equality follows from the convexity of the mutual information along with the converse proof of [159, Thoerem 6].

Although Theorem 5.4 can be considered as a very promising solution for the problem of secure communication over AVWC, it still suffer from the normal drawbacks of correlated random codes. In particular, correlated random codes can only be used if there exists a shared randomness between the transmitter and the legitimate receiver, where the amount of this randomness grows with the code block length n. In many practical situation, it might be impossible to provide such amount of correlated randomness. Moreover, the reliability analysis of correlated random codes for AVWC is only valid under the assumption of restricted communication between the jammer and the eavesdropper. This assumption is not always true because in some AVWCs the jammer and the eavesdropper are in fact one node.

5.2.3 Coding Techniques for Deterministic Codes

The previous section clearly pointed out that if no common randomness is available, the construction of correlated random codes fails. This implies that in order to establish a secure communication over an AVWC for these scenarios, alternative coding schemes are needed. In [14], deterministic secrecy codes prevailed as an appropriate solution. In particular, it was shown that deterministic codes can achieve the same secrecy rates provided by correlated random codes under some restrictions on the AVC $\mathfrak{W}$ between the transmitter and the legitimate receiver. This result was generalized in [166] leading to the establishment of the deterministic secrecy capacity of AVWCs as follows:

Theorem 5.5. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, the deterministic strong secrecy capacity with respect to the average decoding error probability is given by:*

$$C_{det}^{s}(\mathfrak{Q}) = \begin{cases} 0 & \text{if } \mathfrak{W} \text{ is symmetrizable} \\ C_{ran}^{s}(\mathfrak{Q}) & \text{otherwise,} \end{cases} \tag{5.39}$$

where $\mathfrak{W}$ is the AVC between the transmitter and the legitimate receiver, while $C_{ran}^{s}(\mathfrak{Q})$ is the correlated random strong secrecy capacity.

The previous theorem implies that the dichotomy behavior of the deterministic transmission capacity of AVCs established in [163] extends to the secrecy domain as well. However, the vanishing condition does not depend on the full AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$. Instead, it only depends on the legitimate AVC $\mathfrak{W}$. This result is convenient in some sense because as highlighted by [165, Lemma 1], it is impossible to establish a reliable communication over a symmetrizable AVC using deterministic code. Since the problem of secure communication over AVWCs involves the same reliability constraint required for AVCs, one should expect the symmetrizability of the AVC $\mathfrak{W}$ to be an important factor in characterizing the deterministic secrecy capacity of the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$. On the other hand, the eavesdropper AVC $\mathfrak{V}$ only contributes to the secrecy condition and plays no role in defining the reliability requirement of the system.

Two main coding approaches have been used to establish the achievability of the secrecy capacity given by Theorem 5.5. The first approach is based on the concept of

elimination of correlation introduced by Ahlswede in [163]. Ahlswede use this technique to provide a deterministic coding scheme that can be used to establish a reliable communication over an AVC. The idea of this technique is based on transforming a correlated random code to a deterministic code using the principle of code concatenation. We already know that the main ingredient of a correlated random code is the availability of a random experiment whose output is shared by both the transmitter and the receiving nodes. The output of this random experiment simply inform the receiver about the exact code realization that was used in the encoding process. Thus, if we want to use a correlated random code in the absence of common randomness, we only need to develop an alternative technique to inform the receiver about this information.

Ahlswede used this idea to build the elimination of correlation technique. He suggested to add a short prefix to the actual codeword produced by the correlated random code encoder. This prefix should carry information about the exact code realization used for encoding. Since the number of code realization needed $|\mathcal{G}| = \mathcal{O}(n^2)$ as shown in [163, Corollary 1], the usage of a prefix code to represent an element $\gamma \in \mathcal{G}$ causes no essential loss to the actual transmitted rate. However, the previous technique only works if the receiver is able to decode the transmitted prefix correctly. According to the results established in [165], this is only possible if the AVC: $\mathfrak{W}$ is not symmetrizable. In [166], this approach was extended to the AVWC leading to a simple achievability proof for Theorem 5.5. A more detailed achievability proof that utilizes this technique will be presented in Section 5.3.4 for a more general scenario.

Although the previous approach managed to solve one of the drawbacks of correlated random codes which is the availability of a common randomness, it raises another issue. In Section 5.2.2, it was mentioned that the reliability analysis of correlated random codes is only valid under the assumption that the jammer has no information about the exact encoder-decoder pair used for the transmission. This assumption has led to the enforcement of some restrictions on the communication link between the jammer and the eavesdropper because it was assumed that the eavesdropper knows the exact encoder-decoder pair and is not allowed to share this information with the jammer. Since the elimination of correlation approach utilizes a correlated random code, we need to make sure that the jammer is still unable to acquire any information about the exact code realization used. Unfortunately, the restricted communication assumption is not enough for this case and a stronger assumption is needed. This is because the transmitter encodes the information about the code realization used in the prefix of each transmitted codeword and transmits it over the channel. Thus, in order to make sure that the jammer possesses no information about the selected code realization, it is assumed that the jammer can not decode any received signal whether this signal came from the eavesdropper or from the transmitter.

The previous discussion implied that the elimination of correlation technique is not the optimum way to construct deterministic codes for both AVCs and AVWCs. In [165] Csiszár and Naryan suggested an alternative approach which can be used directly to construct deterministic codes for AVCs. Their approach is based on two main pillars: The first pillar is the generation of some codewords that satisfies a set of typicality properties given in [165, Lemma 3]. Csiszár and Naryan showed that any randomly generated group of codewords will possess the required properties with an arbitrary

probability close to 1. The second pillar of Csiszár and Naryan's technique is a decoding role [165, Definition 3] that assure an unambiguous (unique) output, if the AVC $\mathfrak{W}$ between the transmitter and the receiver is not symmetrizable. In [166], Nötzel, Boche and Wiese managed to extend this approach to the secrecy domain and presented a direct proof for Theorem 5.5. In doing so, they presented [166, Lemma 2] which is an extended version of [165, Lemma 3]. They showed that under some rate constraints, the a randomly selected set of codewords will satisfy the strong secrecy constraint in (5.19) and the original typicality constraints in [165, Lemma 3]. Beside the establishment of [166, Lemma 2], the authors had to deal with other issues related to the constraints that define the vanishing condition of the capacity. In Section 5.3.3, we will present a coding scheme that utilizes the same coding techniques introduced in [166] but for a more general problem.

It is important to highlight that the deterministic coding scheme presented in [166] to prove the achievability of the deterministic secrecy rate in Theorem 5.5 does not enforce any restrictions on the information shared between the eavesdropper and the jammer. In particular, the eavesdropper is allowed to know the jamming strategy used to manipulate the channel state or even the exact channel state selected by the jammer. At the same time, the jammer is allowed to know any information possessed by the eavesdropper. This result implies that this coding scheme managed to overcome the main drawbacks of the correlated random techniques highlighted in Section 5.2.2. Nevertheless, this coding scheme suffers from the usual drawback of deterministic codes which is the inadequacy of dealing with symmetrizable AVWCs.

5.2.4 Analytical Properties of the Capacity Functions

Over the last decade, there has been a huge controversy in the information theory community about the usability of multi-letter capacity descriptions. In particular, many researchers claimed that multi-letter capacity descriptions are useless because they are inefficiently computable: One might notice that in order to calculate the correlated random secrecy capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ given by (5.38), the limit of a series of convex optimization problems need to be computed. Although this is true, it has turned out that multi-letter capacity descriptions like the one in Theorem 5.4 can be incredibly useful. In particular, it was shown in [172] that multi-letter capacity descriptions can be used to describe some of the analytical properties such as continuity and additivity of the capacity functions. These properties are of high importance as they play an important role in determining the performance of the communication system and its stability.

The investigation of the analytical properties of the Shannon's capacity of a DMC has been widely addressed. In particular, it has been shown that for a DMC: W, the capacity function $C(W)$ with respect to the average and maximum decoding error are continuous and additive [173]. Continuity means that any small change in the channel transition matrix W, only leads to a small change in the capacity. On the other hand, additivity means that for two parallel DMCs W_1 and W_2, the capacity of the overall system $W_1 \otimes W_2$ satisfies the following:

$$C(W_1 \otimes W_2) = C(W_1) + C(W_2). \tag{5.40}$$

The investigation of the continuity and the additivity of a capacity function is very crucial for any communication system. This is because the continuity property guarantees a certain level of stability to the transmission rate of the system [174]. On the other hand, the additivity property indicates that using a joint encoder-decoder pair for two parallel channels does not provide any gain in the capacity compared to individual encoding [175]. However, the investigation of continuity and additivity in classical information theory did not capture a lot of attention in previous literature. This because aside from DMCs, the investigation of these properties appears to be a very difficult task.

In [176], Shannon tried to investigate the additivity of the zero-error capacity of a DMC: $C^0(W)$. He conjectured that the $C^0(W)$ possess the same property of the average and maximum error capacities, i.e. $C^0(W)$ is also additive and satisfy the relation in (5.40). This conjecture was disproved by Haermers in [177], where he managed to construct a counter example for which the bound in (5.40) does not hold. In [178], Alon gave a stronger counterexample showing that the discrepancy between the overall capacity and the sum of the individual capacities can be arbitrarily large. The previous examples implied that the zero-error capacity is super-additive and there exist channels such that

$$C^0(W_1 \otimes W_2) > C^0(W_1) + C^0(W_2). \tag{5.41}$$

Although Alon and Haermers managed to construct explicit examples for which the zero-error capacity is super-additive, a general characterization of the additivity property remains unknown. Moreover, in [178] Alon raised a more interesting question as follow: By how much can the additivity of the zero-error capacity be violated. In particular, given two parallel DMCs W_1 and W_2, where $C^0(W_i) \leq \epsilon$ for $i = 1, 2$, what bounds do we have on the normalized zero-error capacity $\bar{C}^0(W)$ given by:

$$\bar{C}^0(W) = \frac{C^0(W)}{\log_2(\min(|\mathcal{X}|, |\mathcal{Y}|))}, \tag{5.42}$$

where $W = W_1 \otimes W_2$. Alon conjectured that the zero-error capacity is violated in a strong form in the sense that $\bar{C}^0(W)$ is close to 1, i.e. $\bar{C}^0(W) > 1 - \epsilon$. However, it was recently shown in [179] that the normalized zero-error capacity of two parallel channels is close to 1/2, i.e. $\bar{C}^0(W) > 1/2 - \epsilon$. This implied that additivity of the zero-error capacity is not violated in a strong form.

The investigation of the continuity and additivity of the capacity functions of AVCs and AVWCs has recently captured the attention of some researchers for the following reasons: The capacity functions of AVCs and AVWCs depends on the channel states set $\mathcal{S}$, which is controlled by an active adversary. Thus, it is important to investigate whether the jammer has the ability to affect the capacity function severely by a simple manipulation to the channel states set $\mathcal{S}$. The investigation of the continuity behavior is not only related to the stability of the capacity function, but it also plays an important role in determining the computability of the capacity function on Turing's machines [180, 181]. The second reason is related to the argument presented by Ahlswede in [182], where he showed that the characterization of the zero-error capacity is included as a special case in the problem of determining the uncorrelated

deterministic transmission capacity of the AVC under the maximal decoding error criterion. Thus, characterizing the additivity behavior of this capacity function will give an answer to the questions raised by Shannon in [176] and Lovasz in [183] In next few lines, a brief summary about the continuity and additivity properties of the correlated random capacity for AVCs and AVWCs is given.

Before these results are presented, a formal definition for the continuity and additivity behavior of a given capacity function is presented as follows:

Definition 5.8. For a finite AVWC: $\mathfrak{Q}$, a given capacity function $C(\mathfrak{Q})$ is said to be continuous in all finite AVWCs: $\mathfrak{Q}$, if for all sequences of finite AVWCs: $\{\mathfrak{Q}_n\}_{n=1}^{\infty}$ such that:

$$\lim_{n\to\infty} D(\mathfrak{Q}_n, \mathfrak{Q}) = 0, \tag{5.43}$$

where $D(\cdot,\cdot)$ is a distance measure used to describe how well one AVWC can be approximated by the other one, the following holds:

$$\lim_{n\to\infty} C(\mathfrak{Q}_n) = C(\mathfrak{Q}). \tag{5.44}$$

In order to lay out a mathematical definition for the distance measure $D(\cdot,\cdot)$, the concept of total variation distance between the transition probability matrices of two DMCs $W_1, W_2 \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$ given by the next equation was used.

$$d(W_1, W_2) = \max_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} \big|W_1(y|x) - W_2(y|x)\big|. \tag{5.45}$$

This measure was then extend to define the following measure between two AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$ defined with respect to the channel states sets $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively as follows:

$$G(\mathfrak{W}_1, \mathfrak{W}_2) = \max_{s_1\in\mathcal{S}_1} \max_{s_2\in\mathcal{S}_2} d(W_1(\cdot|\cdot, s_1), W_2(\cdot|\cdot, s_2)). \tag{5.46}$$

Although one might argue that G is not symmetric, we can still use it to define the following symmetric distance measure between $\mathfrak{W}_1$ and $\mathfrak{W}_2$ as follows:

$$D(\mathfrak{W}_1, \mathfrak{W}_2) = \max\big\{G(\mathfrak{W}_1, \mathfrak{W}_2), G(\mathfrak{W}_2, \mathfrak{W}_1)\big\}. \tag{5.47}$$

It worth mentioning that $\mathcal{S}_1$ and $\mathcal{S}_2$ can be any arbitrary finite state sets and it is not necessarily to have $|\mathcal{S}_1| = |\mathcal{S}_2|$. Now using the same concept, the distance measure $D(\cdot,\cdot)$ between two AVWCs: $\mathfrak{Q}_1 = (\mathfrak{W}_1, \mathfrak{V}_1)$ and $\mathfrak{Q}_2 = (\mathfrak{W}_2, \mathfrak{V}_2)$ is defined as follows:

$$D(\mathfrak{Q}_1, \mathfrak{Q}_2) = \max\big\{D(\mathfrak{W}_1, \mathfrak{W}_2), D(\mathfrak{V}_1, \mathfrak{V}_2)\big\}. \tag{5.48}$$

On the other hand, in order to investigate the additivity behaviour of a given capacity function, the concept of parallel AVCs was introduced. Given two AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$, let $\mathfrak{W}$ be their parallel combination defined as follows:

$$\mathfrak{W} = \mathfrak{W}_1 \otimes \mathfrak{W}_2 = \big\{W_1(\cdot|\cdot, s_1)\big\}_{s_1\in\mathcal{S}_1} \otimes \big\{W_2(\cdot|\cdot, s_2)\big\}_{s_2\in\mathcal{S}_2}. \tag{5.49}$$

This implies that $\mathfrak{W}$ can be interpreted as an AVC with an input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ and an output alphabet $\mathcal{Y}_1 \times \mathcal{Y}_2$ such that $\mathfrak{W}$ is characterized by the following transition matrix:

$$W(y_1, y_2|x_1, x_2, s_1, s_2) = W_1(y_1|x_1, s_1)W_2(y_2|x_2, s_2), \tag{5.50}$$

where $x_i \in \mathcal{X}_i$, $y_i \in \mathcal{Y}_i$ and $s_i \in \mathcal{S}_i$, for $i = 1, 2$. The concept of parallel AVCs can be easily extended to the secrecy domain to define parallel AVWCs as well.

Definition 5.9. Let $\mathfrak{Q}_1$ and $\mathfrak{Q}_2$ be two finite AVWCs and $\mathfrak{Q} = \mathfrak{Q}_1 \otimes \mathfrak{Q}_2$ be their parallel combination defined using Eqs. (5.49) and (5.50). Then, a given capacity function for these AVWCs is said to be additive if

$$C(\mathfrak{Q}) = C(\mathfrak{Q}_1) + C(\mathfrak{Q}_2), \tag{5.51}$$

while it is said to be super-additive if

$$C(\mathfrak{Q}) > C(\mathfrak{Q}_1) + C(\mathfrak{Q}_2), \tag{5.52}$$

We are now ready to present the main continuity and additivity results established in previous literature for the correlated random capacity function of AVCs and AVWCs.

Theorem 5.6. *[172] The correlated random secrecy capacity of the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ with respect to the average decoding error probability and the maximum secrecy criterion $C^{s,max}_{ran,avg}(\mathfrak{Q})$ is continuous in the two finite AVCs $(\mathfrak{W}, \mathfrak{V})$.*

The previous theorem was established although a single-letter description for $C^{\mathrm{s,max}}_{\mathrm{ran,avg}}(\mathfrak{Q})$ is still unknown. This result reflects the importance of multi-letter capacity description. To the best of our knowledge, the continuity of the correlated random public capacity of AVCs was not investigated in previous literature despite the fact that a single-letter description was established in [154, 158]. Nevertheless, one can use the result established in Theorem 5.6 to prove the following behavior: computable.

Corollary 5.1. *The correlated random public capacity of the AVC $\mathfrak{W}$ under the average decoding error probability $C^{p}_{ran,avg}(\mathfrak{W})$ is continuous in the finite AVC $\mathfrak{W}$.*

It was also surprising to discover that the additivity behavior of the correlated random capacity of parallel AVCs was only investigated recently leading the following result.

Theorem 5.7. *[184] Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two parallel AVCs. Then the correlated random public capacity under the average decoding error probability is additive, i.e. $C^{p}_{ran}(\mathfrak{W}_1 \otimes \mathfrak{W}_2) = C^{p}_{ran}(\mathfrak{W}_1) + C^{p}_{ran}(\mathfrak{W}_2)$.*

Unfortunately, the additivity behavior of the correlated random secrecy capacity for AVWCs is still an open problem. In section 5.3.6, the analytical properties of the deterministic capacity for AVCs and AVWCs established in [185] will be derived as a special case from a more general scenario. In fact, we will show that the deterministic secrecy capacity of AVWCs exhibits an extreme super-additivity behavior known as super-activation. This behavior indicates that the joint encoding-decoding for two parallel AVWCs each with vanishing capacity can lead to a capacity greater than zero. Super-activation has been investigated in the field of quantum information theory in [186] and was believed to be a distinct phenomena for quantum information theory. However, the recent results established in [184, 187] showed that this behavior can occur in the classical domain but it seems to be a unique feature for AVWCs.

5.3 The Secrecy Capacity of AVWCs Under List Decoding

In this section, we present a full characterization for the list secrecy capacity of AVWCs. We start by stating the main concepts of list decoding along with some of the main

results established in previous literature. We then present our main coding theorem and derive its achievability using two different coding schemes. We discuss and compare the two coding schemes, showing that although they both lead to the same list secrecy capacity, each coding scheme has its own preferences. We then derive a special result for the secrecy capacity of AVWCs under the class of list codes with finite list size. Finally, we establish the continuity and additivity behavior for the capacity functions of AVCs and AVWCs under list decoding. The results established in this section was published in [188–192].

5.3.1 List Decoding for AVWCs: Basic Definitions

List codes are a special class of deterministic codes, in which the decoder outputs a list of L possible messages, instead of deciding on exactly one message. The principle of list decoding was first introduced in [193] in order to minimize the gap between upper and lower bounds of the decoding error probability. This idea suggested that list codes can be used for the AVC to overcome the symmetrizability issue. It was then shown in [194–197] that list codes are a very good coding alternative for AVCs as it can overcome the drawbacks of correlated random codes and the classical deterministic codes. We start by presenting the following definition.

Definition 5.10. A public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ with list size L for the AVC: $\mathfrak{W}$ consists of: a set of public messages $\mathcal{M}_p$, a deterministic encoder

$$E : \mathcal{M}_p \to \mathcal{X}^n$$

that maps a message $m_p \in \mathcal{M}_p$ into a codeword $x^n(m_p) \in \mathcal{X}^n$ and a deterministic list decoder given by:

$$\varphi_L : \mathcal{Y}^n \to \mathfrak{P}_L(\mathcal{M}_p) \cup \{?\}$$

that maps each channel observation at the receiver to a list of up to L messages or an error message, where $\mathfrak{P}_L(\mathcal{M}_p)$ is the set of all subsets of $\mathcal{M}_p$ with cardinality at most L.

The reliability performance of $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ is measured in terms of its average decoding error probability given by:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) &= \max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) \\ &= \max_{s^n \in \mathcal{S}^n} \frac{1}{|\mathcal{M}_p|} \sum_{m_p} \sum_{y^n : \varphi_L(y^n) \not\ni m_p} W^n_{s^n}(y^n|x^n(m_p)). \end{aligned} \tag{5.53}$$

Definition 5.11. A non-negative number R_p is an achievable public list rate for the AVC: $\mathfrak{W}$; if for all $\delta, \lambda > 0$ there is an $n(\delta, \lambda) \in \mathbb{N}$, such that for all $n > n(\delta, \lambda)$, there exists a sequence of public list codes $(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}})_n$, such that the following holds:

$$\frac{1}{n} \log \frac{|\mathcal{M}_p|}{L} \geq R_p - \delta, \tag{5.54}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) \leq \lambda. \tag{5.55}$$

The public list capacity $C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}, L)$ with respect to the average decoding error probability is given by the supremum of all achievable rates R_p.

It was shown in [194, 195] that the public list capacity of an AVC $\mathfrak{W}$ exhibits a dichotomy similar to the deterministic capacity of AVCs but with respect to a generalized symmetrizability condition known as L-symmetrizable as follows:

Theorem 5.8. *For an AVC:* $(\mathfrak{W})$*, the public list capacity with respect to the average decoding error probability is characterized by the following:*

$$C^p_{list}(\mathfrak{W}, L) = \begin{cases} 0 & \text{if } \mathfrak{W} \text{ is } L\text{-symmetrizable} \\ C^p_{ran}(\mathfrak{W}) & \text{otherwise.} \end{cases}$$

The concept of L-symmetrizability extends the symmetrizability condition introduced in Definition 5.1. It characterizes the ability of an AVC to emulate not only one valid channel input, but up to L valid channel inputs as follows:

Definition 5.12. An AVC $\mathfrak{W}$ is said to be L-symmetrizable, if there exists an auxiliary channel $\sigma : \mathcal{X}^L \to \mathcal{P}(\mathcal{S})$, such that for every permutation π of the sequence $(1, \ldots, L+1)$

$$\sum_{s\in\mathcal{S}} W_s(y|x_1)\sigma(s|x_2, \ldots, x_{L+1}) = \sum_{s\in\mathcal{S}} W_s(y|x_{\pi(1)})\sigma(s|x_{\pi(2)}, \ldots, x_{\pi(L+1)}) \tag{5.56}$$

holds for every $x^{L+1} \in \mathcal{X}^{L+1}$ and $y \in \mathcal{Y}$.

Similar to the implication of the symmetrizability condition in (5.12), the condition in (5.56) implies that an L-symmetrizable AVC can emulate L valid replicas of the channel input. This implies that for a given observation at the receiver y^n, there exists $L+1$ input sequences that might have been sent. For a given AVC: $\mathfrak{W}$, the largest L for which this AVC is L-symmetrizable is called the order of symmetrizability and is denoted by $L(\mathfrak{W})$. An example for L-symmetrizable AVC was given in [194, Theorem 3]. It is worth mentioning that the condition in (5.56) implies that any L-symmetrizable AVC is an L'-symmetrizable one as well, for all $0 \leq L' \leq L$.

In order to extend the concept of list decoding to fit the model of secure communication over an AVWC, we modify Definition 5.10 as follows:

Definition 5.13. A secrecy list code $\mathcal{C}^{\text{s}}_{\text{list}}$ with list size L for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ consists of: a set of confidential messages $\mathcal{M}_c$, a stochastic encoder

$$E : \mathcal{M}_c \to \mathcal{P}(\mathcal{X}^n) \tag{5.57}$$

that maps a confidential message $m_c \in \mathcal{M}_c$ to a codeword $x^n \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m_c)$, and a deterministic list decoder with list size L

$$\varphi_L : \mathcal{Y}^n \to \mathfrak{P}_L(\mathcal{M}_c) \cup \{?\}$$

that maps a channel observation at the legitimate receiver into a list of up to L messages or an error message, where $\mathfrak{P}_L(\mathcal{M}_c)$ is the set of all subsets of $\mathcal{M}_c$ with cardinality at most L.

One can notice that the main difference between the previous definition and Definition 5.10 is the usage of a stochastic encoder instead of the deterministic one. This is a necessary step for secrecy codes because as highlighted in Section 2.2, secrecy can not

be achieved using normal deterministic encoders. In order to evaluate the reliability performance of $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$, we use the average decoding error probability as follows:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) &= \max_{s^n\in\mathcal{S}^n} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \\ &= \max_{s^n\in\mathcal{S}^n} \frac{1}{|\mathcal{M}_c|}\sum_{m_c}\sum_{x^n}\sum_{y^n:\varphi_L(y^n)\not\ni m_c} W^n_{s^n}(y^n|x^n)E(x^n|m_c). \end{aligned} \tag{5.58}$$

One the other hand, we use the information leakage of the confidential message M_c to the eavesdropper with respect to the strong secrecy criterion to investigate the secrecy performance of the list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ as follows:

$$\begin{aligned}\mathbb{L}(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) &= \max_{s^n\in\mathcal{S}^n} \mathbb{L}(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \\ &= \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c;\mathrm{Z}^n_{s^n}|\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}). \end{aligned} \tag{5.59}$$

Definition 5.14. A non-negative number R_c is an achievable list secrecy rate for the AVWC: $\mathfrak{Q} = (\mathfrak{W},\mathfrak{V})$, if for all $\delta,\lambda,\tau > 0$ there is an $n(\delta,\lambda,\tau)\in\mathbb{N}$, such that for all $n > n(\tau,\lambda,\delta)$, there exists a sequence of secrecy list codes $(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})_n$ with list size L that satisfies the following constraints:

$$\frac{1}{n}\log\frac{|\mathcal{M}_c|}{L} \geq R_c - \delta, \tag{5.60}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq \lambda, \tag{5.61}$$

$$\mathbb{L}(\mathcal{C}^{s}_{\mathrm{list}}) \leq \tau. \tag{5.62}$$

The list secrecy capacity $C^{\mathrm{s}}_{\mathrm{list}}(\mathfrak{W},\mathfrak{V},L)$ is given by the supremum of all achievable secrecy list rates R_c.

5.3.2 Coding Theorem and Important Results

In this section, we present one of the main contribution of this chapter, which is a full characterization of the list secrecy capacity of AVWCs. This result generalizes the deterministic secrecy capacity of AVWCs given by Theorem 5.5.

Theorem 5.9. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W},\mathfrak{V})$, the list strong secrecy capacity with respect to the average decoding error probability is characterized by the following:*

$$C^{s}_{list}(\mathfrak{Q},L) = \begin{cases} 0 & \text{if } \mathfrak{W} \text{ is } L\text{-symmetrizable} \\ C^{s}_{ran}(\mathfrak{Q}) & \text{otherwise,}\end{cases}$$

where $\mathfrak{W}$ is the AVC between the transmitter and the legitimate receiver, while $C^{s}_{ran}(\mathfrak{Q})$ is the correlated random strong secrecy capacity given by (5.38).

The previous capacity characterization implies that for an AVWC: $\mathfrak{Q}$ where the order of symmetrizability of the legitimate AVC: $\mathfrak{W}$ is given by $L(\mathfrak{W})$, a list code with list size $L \geq L(\mathfrak{W}) + 1$ can provide a reliable and secure communication at the rate equivalent to the correlated random secrecy capacity. Theorem 5.9 also implies that list codes can overcome the drawbacks of both uncorrelated and correlated random codes as follows: Given an AVWC $\mathfrak{Q} = (\mathfrak{W},\mathfrak{V})$ where:

1. $\mathfrak{W}$ is symmetrizable with an order of symmetrizability $L(\mathfrak{W}) = 1$, which directly implies that $C^{\mathrm{s}}_{\mathrm{det}}(\mathfrak{Q}) = 0$.
2. No common randomness is shared between the transmitter and the legitimate receiver, which implies that correlated random codes can not be utilized and consequently it follows that $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q}) = 0$.

For this AVWC, we can still achieve a positive secrecy rate equivalent to the correlated random secrecy capacity using list decoding with list size $L \geq 2 = L(\mathfrak{W}) + 1$. The previous argument implies that we can always overcome the different jamming strategies induced by the jammer to disturb reliable communication over an AVWC by using list codes with list size $L \geq L(\mathfrak{W}) + 1$.

Although the previous argument advocates the usage of list codes as the best coding scheme for AVWCs, it was shown in [194, Theorem 3] that even for simple AVWCs, where $|\mathcal{X}| = |\mathcal{Y}| = |\mathcal{S}| = 2$, there exist some jamming strategies that lead to an arbitrarily large order of symmetrizability $L(\mathfrak{W})$. This implies that, for these AVWCs, we need to construct list codes with an arbitrarily large list size, which is not always feasible. Nevertheless, Theorem 5.9 implies that list codes are a very powerful and useful tool for reliable and secure communication over AVWCs.

In order to prove Theorem 5.9, we need an achievability proof and a converse proof. In this section, we will mainly focus on the converse proof and postpone the achievability proof to the following sections. For the converse, we start by presenting a lemma that shows that reliable communication can not be established over an L-symmetrizable AVC using list codes with list size L as follows:

Lemma 5.2. *[194, Lemma 1] [195, Lemma 4] For an AVC: $\mathfrak{W}$, if $\mathfrak{W}$ is L-symmetrizable, then the public list capacity with list size L vanishes, i.e. $C^{p}_{list}(\mathfrak{W}, L) = 0$.*

This lemma generalizes the result established in [165, Lemma 1] for deterministic codes which can be interpreted as list codes with list size $L = 1$. The proof of Lemma 5.2 is based on showing that if the AVC: $\mathfrak{W}$ is L-symmetrizable, then for any public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ with list size L, there exists a state sequence $s^n \in \mathcal{S}^n$ such that the average decoding error probability is bounded away from zero as n approaches infinity. In particular, it follows that:

$$\limsup_{n\to\infty} \bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) \geq \liminf_{n\to\infty} \max_{s^n\in\mathcal{S}^n} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) \geq \frac{1}{L+1}. \tag{5.63}$$

One can easily show that the previous equation is bounded away from zero as long as the list size L is finite. This implies that no list code with list size L exists, such that the average decoding error probability vanishes as $n \to \infty$, which consequently means a zero capacity. The previous argument establishes only one part of the converse proof of Theorem 5.9, which is the vanishing capacity if $\mathfrak{W}$ is L-symmetrizable.

For the second part of the converse, we start by pointing out that the average decoding error probability of any secrecy list code $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})$ is an affine function in the channel. This implies that $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})$ does not change if it is calculated with respect to the generalized channel state space $\mathcal{P}(\mathcal{S}^n)$. Thus for $q \in \mathcal{P}(\mathcal{S}^n)$, we have

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) = \max_{s^n\in\mathcal{S}^n} \frac{1}{|\mathcal{M}_c|} \sum_{m_c} \sum_{x^n} \sum_{y^n:\varphi_L(y^n)\not\ni m_c} W^n_{s^n}(y^n|x^n) E(x^n|m_c)$$

$$= \max_{q\in\mathcal{P}(\mathcal{S}^n)} \frac{1}{|\mathcal{M}_c|} \sum_{m_c}\sum_{x^n}\sum_{y^n:\varphi_L(y^n)\not\ni m_c} W_q^n(y^n|x^n)E(x^n|m_c) \overset{(a)}{\leq} \lambda, \tag{5.64}$$

where (a) follows from Eq. (5.61). If we use Eq. (5.64) along with Fano's inequality, then for every $q \in \mathcal{P}(\mathcal{S}^n)$, the following holds: $\mathbb{H}(\mathrm{M}_c|\mathrm{Y}_q^n) \leq 1 + \lambda \log |\mathcal{M}_c|$. Now, from Eq. (5.60), it follows that:

$$\begin{aligned}
R_c &\leq \frac{1}{n}\big[\log|\mathcal{M}_c| - \log L\big] + \delta \\
&= \frac{1}{n}\big[\mathbb{H}(\mathrm{M}_c) - \log L\big] + \delta \\
&\overset{(a)}{\leq} \frac{1}{n}\Big[\min_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{M}_c;\mathrm{Y}_q^n) + 1 + \lambda\log|\mathcal{M}_c| - \log L\Big] + \delta \\
&\overset{(b)}{\leq} \frac{1}{n}\Big[\min_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{M}_c;\mathrm{Y}_q^n) - \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c;\mathrm{Z}_{s^n}^n)\Big] + \epsilon \\
&\overset{(c)}{\leq} \frac{1}{n}\Big[\min_{q\in\mathcal{P}(\mathcal{S}^n)} \mathbb{I}(\mathrm{U};\mathrm{Y}_q^n) - \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{U};\mathrm{Z}_{s^n}^n)\Big] + \epsilon
\end{aligned} \tag{5.65}$$

where (a) follows by adding and subtracting the two sides of the previous inequality; (b) follows similarly by applying the inequality in (5.62), where $\epsilon = 1/n(\tau + 1 + \lambda \log|\mathcal{M}_c| - \log L) + \delta$; while (c) by defining the auxiliary channel $F : \mathcal{M}_c \to \mathcal{U}$. Now, if we take the limit as $n \to \infty$, which implies that $\epsilon \to 0$, the bound in (5.65) simplifies to the expression of the correlated random secrecy capacity in (5.38) and this completes our converse.

Before we move to the achievability proof of our coding theorem, we need to highlight some of results established in previous literature for AVCs and AVWCs. In particular, we need to highlight the main tools used to define the encoding and decoding functions for a list code $\mathcal{C}_{\text{list}}^s$ of Definition 5.13. We start by the encoding function and notice that $\mathcal{C}_{\text{list}}^s$ requires a stochastic encoder as shown in (5.57). We transform this stochastic encoder to a deterministic one using a randomization message set $\mathcal{M}_r$ as follows:

$$E : \mathcal{M}_c \times \mathcal{M}_r \to \mathcal{X}^n.$$

This implies that for a confidential message $m_c \in \mathcal{M}_c$, the encoder chooses a randomization message $m_r \in \mathcal{M}_r$ uniformly at random then transmits the corresponding codeword $x^n(m_c, m_r)$. The generation of the codewords $x^n(m_c, m_r) \in \mathcal{X}^n$ that define the codebook of our list code $\mathcal{C}_{\text{list}}^s$ is guided by the following lemma:

Lemma 5.3. *For any $L \geq 1$, $\epsilon > 0$ and $\beta > 0$, there exists an $n_0(\epsilon, L)$, such that for all $n \geq n_0(\epsilon, L)$, all message sets $\mathcal{M}_c$, $\mathcal{M}_r$ satisfying $|\mathcal{M}_c||\mathcal{M}_r| \geq L \cdot 2^{n\epsilon}$ and all types $\mathrm{P_X} \in \mathcal{P}_0^n(\mathcal{X})$ satisfying $\min_{x:\mathrm{P_X}(x)>0} \mathrm{P_X}(x) \geq \beta$, there exist codewords $x^n(m_c, m_r) \in \mathcal{T}_\epsilon^n(\mathrm{P_X}) \subset \mathcal{X}^n$, for $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$, such that upon setting $R = \frac{1}{n}\log\frac{|\mathcal{M}_c||\mathcal{M}_r|}{L}$, we have for all $x^n \in \mathcal{T}_\epsilon^n(\mathrm{P_X})$, $s^n \in \mathcal{S}^n$ and the joint type $\mathrm{P_{XX^LS}}$*

$$\Big|\big\{(m_c,m_r) : (x^n(m_c,m_r), s^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_{XS}})\big\}\Big| \leq |\mathcal{M}_c||\mathcal{M}_r|2^{-\frac{n\epsilon}{2}},$$
$$\text{if } \mathbb{I}(\mathrm{X},\mathrm{S}) \geq \epsilon \tag{5.66a}$$

$$\Big|\big\{(m_c,m_r) : (x^n, x^n(m_c,m_r), s^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_{XX_kS}})\big\}\Big| \leq 2^{n(|R-\mathbb{I}(\mathrm{X}_k;\mathrm{XS})|^+ + \epsilon)} \tag{5.66b}$$

$$\left|\left\{(m_c,m_r) : (x^n(m_c,m_r), x^n(\hat{m}_c,\hat{m}_r), s^n) \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{XX}_k\mathrm{S}})\right\}\right| \leq |\mathcal{M}_c||\mathcal{M}_r|2^{-\frac{n\epsilon}{2}},$$
$$\text{if } \mathbb{I}(\mathrm{X};\mathrm{X}_k\mathrm{S}) \geq |R - \mathbb{I}(\mathrm{X}_k;\mathrm{S})|^+ + \epsilon \tag{5.66c}$$

for $k = 1, \ldots, L$ and $(\hat{m}_c, \hat{m}_r) \neq (m_c, m_r)$. Moreover, if $R < \min_k \mathbb{I}(\mathrm{X}_k;\mathrm{S})$, then for $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$ the codewords $x^n(m_c, m_r)$ can be selected to further satisfy

$$\left|\left\{\mathcal{J} : (x^n, x^n_{\mathcal{J}}, s^n) \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{XX}^L\mathrm{S}})\right\}\right| \leq 2^{n\epsilon} \tag{5.67a}$$

$$\left|\left\{(m_c,m_r) : (x^n(m_c,m_r), x^n_{\mathcal{J}}, s^n) \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{XX}^L\mathrm{S}}), (m_c,m_r) \notin \mathcal{J}\right\}\right| \leq |\mathcal{M}_c||\mathcal{M}_r|2^{-\frac{n\epsilon}{2}},$$
$$\text{if } \mathbb{I}(\mathrm{X};\mathrm{X}^L\mathrm{S}) \geq \epsilon \tag{5.67b}$$

where $\mathcal{J} \in \mathfrak{P}_L(\mathcal{M}_c \times \mathcal{M}_r)$ is a set with cardinality L that contains pairs of messages (m_c, m_r), while $x^n_{\mathcal{J}}$ denotes the ordered L-tuple $\left(x^n(m_{c_1}, m_{r_1}), \ldots, x^n(m_{c_L}, m_{r_L})\right)$.

The previous lemma can be viewed as a generalization to the result established in [165, Lemma 3] for deterministic codes of AVC, i.e. $L = 1$. It is also an extension of [195, Lemma 1] used to establish public list codes for AVCs. In order to understand the role played by the Lemma 5.3 in the encoding process of $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$, we need to analyze the statement of the Lemma as follows:

Assumptions: Lemma 5.3 starts by defining a list size $L \geq 1$ that will be used during the decoding process. Then, for some constant $\epsilon > 0$, it defines the code block length n to be sufficiently large, i.e. $n \geq n_0(\epsilon, L)$. Next, Lemma 5.3 addresses the messages needed to be transmitted from the sets $\mathcal{M}_c$ and $\mathcal{M}_r$, where a constraint on the product of the cardinality of the message sets must be fulfilled. Finally, Lemma 5.3 defines a probability distribution P_{X} on $\mathcal{X}$, such that P_{X} belongs to the set of types on $\mathcal{X}^n$ and the minimum non-zero probability of an element x should be greater than or equal $\beta > 0$. This probability distribution is used to generate the random codewords $x^n(m_c, m_r) \in \mathcal{X}^n$

Results: For any sequence $x^n \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{X}})$, a given sequence $s^n \in \mathcal{S}^n$, a fixed type $\mathrm{P}_{\mathrm{XX}^L\mathrm{S}}$ and under the previous assumptions, Lemma 5.3 claims that using the probability distribution P_{X}, we can generate a random codebook $\{x^n(m_c, m_r) \in \mathcal{T}_\epsilon^n(\mathrm{P}_{\mathrm{X}}) : (m_c, m_r) \in \mathcal{M}_c \times \mathcal{M}_r\} \subset \mathcal{X}^n$ such that the conditions in (5.66) and (5.67) are satisfied. The conditions in (5.66) are simple joint typicality constraints and are identical to the ones in [165, Lemma 3], for $L = 1$. On the other hand, the conditions in (5.67) are an extension of the joint typicality arguments to the principle of list decoding.

It was shown in [195] that any randomly generated set of codewords will satisfy the conditions in (5.66) and (5.67) of Lemma 5.3 with probability that approaches one exponentially fast as $n \to \infty$.

Beside Lemma 5.3, we need to define the decoding algorithm for the list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$. To do so, we use the list decoder introduced in [195, Definition 4] as follows:

Definition 5.15. Given a received sequence y^n and the set of codewords $x^n(m_c, m_r)$, for $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$, we let $(m_c, m_r) \in \varphi_L(y^n)$, if and only if for an $\eta \geq 0$, the following holds:

1. There exists an $s^n \in \mathcal{S}^n$ such that: $D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{YXS}} \| W \times \mathrm{P}_{\mathrm{X}} \times \mathrm{P}_{\mathrm{S}}) \leq \eta$, where $\mathrm{P}_{\mathrm{YXS}}$ is the joint type of $(y^n, x^n(m_c, m_r), s^n)$, while P_{X} and P_{S} are the types of $x^n(m_c, m_r)$ and s^n respectively. Additionally, we have
$$(W \times \mathrm{P}_{\mathrm{X}} \times \mathrm{P}_{\mathrm{S}})(y, x, s) = W_s(y|x)\mathrm{P}_{\mathrm{X}}(x)\mathrm{P}_{\mathrm{S}}(s).$$

2. For every choice of L other distinct codewords: $x^n(m_{c_1}, m_{r_1}), \ldots, x^n(m_{c_L}, m_{r_L})$, where $(m_{c_i}, m_{r_i}) \in \mathcal{M}_c \times \mathcal{M}_r$, such that each of them satisfies: $D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{YX}_i\mathrm{S}_i} \| W \times \mathrm{P}_{\mathrm{X}_i} \times \mathrm{P}_{\mathrm{S}_i}) \leq \eta$, for some $s_i^n \in \mathcal{S}^n$, it holds that:

$$\mathbb{I}(\mathrm{XY}; \mathrm{X}^L | \mathrm{S}) \leq \eta, \tag{5.68}$$

where $\mathrm{P}_{\mathrm{YX}_i\mathrm{S}_i}$ is the joint type of $(y^n, x^n(m_{c_i}, m_{r_i}), s_i^n)$, while $\mathrm{P}_{\mathrm{X}_i}$ and $\mathrm{P}_{\mathrm{S}_i}$ are the types of $x^n(m_{c_i}, m_{r_i})$ and s_i^n respectively,

The mutual information in (5.68) is calculated with respect to the joint distribution $\mathrm{P}_{\mathrm{XYX}^L\mathrm{S}}$ given by the joint type of $\big(x^n(m_c, m_r), y^n, x^n(m_{c_1}, m_{r_1}), \ldots, x^n(m_{c_L}, m_{r_L}), s^n\big)$.

The previous list decoder generalizes the decoder introduced in [165, Definition 3] for deterministic codes. It was shown in [195] that the previous decoding role is unambiguous (unique), if η is sufficiently small. In particular, it was shown that if the AVC $\mathfrak{W}$ is not L-symmetrizable, no set of random variables $(\mathrm{X}, \mathrm{X}^L, \mathrm{Y}, \mathrm{S}, \mathrm{S}^L)$ can satisfy the decoding roles in Definition 5.15. With Lemma 5.3 and Definition 5.15, we possess the main pillars needed to realize an encoder-decoder pair for a list code $\mathcal{C}_{\mathrm{list}}^{\mathrm{s}}$ given by Definition 5.13.

We now present some of the results established in previous literature for reliable communication over AVCs using list codes. We start by a lemma that shows that reliable communication at a non-zero rate is possible over an AVC: $\mathfrak{W}$ with order of symmetrizability $L(\mathfrak{W})$ using a list code with list size $L \geq L(\mathfrak{W}) + 1$.

Lemma 5.4. *[194, Lemma 3] For an AVC: $\mathfrak{W}$ with an order of symmetrizability $L(\mathfrak{W})$, let $L \geq L(\mathfrak{W}) + 1$. Then there exists $\sigma > 0$, such that for all rates $R_p \in (0, \sigma)$, there exists a public list code $\mathcal{C}_{list}^p$ with list size L and a public rate R_p, such that for $\zeta > 0$, we have*

$$\bar{\boldsymbol{P}}_e(\mathcal{C}_{list}^p) = \max_{s^n \in \mathcal{S}^n} \bar{\boldsymbol{P}}_e(s^n | \mathcal{C}_{list}^p) \leq 2^{-n\zeta}. \tag{5.69}$$

The main key to the proof of Lemma 5.4 is to show that for a sufficiently small rate R_p and a decoding parameter η, the list decoder in Definition 5.15 can not produce more than L values for any $y^n \in \mathcal{Y}^n$ as long as the AVC $\mathfrak{W}$ is not L-symmetrizable. In [195, Lemma 3], the result in Lemma 5.4 was extended showing that reliable communication is possible as long as the rate is bounded by the minimum mutual information between the channel input X and the average channel output Y_q. The result established in [195, Lemma 3] can be easily extended to AVWCs as follows:

Lemma 5.5. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, where the legitimate AVC: $\mathfrak{W}$ is characterized by an order of symmetrizability $L(\mathfrak{W})$, let $L \geq L(\mathfrak{W}) + 1$ and $\beta > 0$. Then for any $\delta > 0$, $n \geq n_0(\beta, \delta)$ and any type $\mathrm{P}_{\mathrm{X}} \in \mathcal{P}_0^n(\mathcal{X})$ such that $\min_{x:\mathrm{P}_{\mathrm{X}}(x)>0} \mathrm{P}_{\mathrm{X}}(x) \geq \beta$, there exists a secrecy list code $\mathcal{C}_{list}^s$ with list size L such that for $\zeta > 0$, we have:*

$$\min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X}; \mathrm{Y}_q) - 2\delta/3 \geq \frac{1}{n} \log \frac{|\mathcal{M}_c||\mathcal{M}_r|}{L} \geq \min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X}; \mathrm{Y}_q) - \delta,$$
$$\bar{\boldsymbol{P}}_e(\mathcal{C}_{list}^s) = \max_{s^n \in \mathcal{S}^n} \bar{\boldsymbol{P}}_e(s^n | \mathcal{C}_{list}^s) \leq 2^{-n\zeta}.$$

The proof of the previous lemma follows by using a secrecy list code $\mathcal{C}_{\mathrm{list}}^{\mathrm{s}}$ given by the set of codewords $x^n(m_c, m_r)$, for $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$ that satisfy the constraints

in Lemma 5.3 along with the list decoder in Definition 5.15. The proof is based on the original steps used to establish [195, Lemma 3] using a simplified version of Lemma 5.3 and Definition 5.15 that only consider one message set $\mathcal{M}$. The extension of the proof to our setup is straight forward because the proof technique is independent from the way used to enumerate the codewords X^n. In other words, it makes no difference to enumerate the codewords by one index taken from a message set $\mathcal{M}$ or by two indices taken from $\mathcal{M}_c \times \mathcal{M}_r$.

The final result that we need to highlight is an extension of the strong secrecy results discussed in Section 2.2.3 to the problem of secure communication over AVWCs. The following lemma gives a lower bound on the amount of randomization needed to confuse the eavesdropper of an AVWC.

Lemma 5.6. *[166, Lemma 2] For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ and any $\tau, \beta, \epsilon > 0$, there exists a function $\delta(\tau) > 0$ and an integer $n_0(\tau)$, such that for all $n \geq n_0(\tau)$ and type $\mathrm{P_X} \in \mathcal{P}_0^n(\mathcal{X})$ satisfying $\min_{x:\mathrm{P_X}(x)>0} \mathrm{P_X}(x) \geq \beta$, there exist codewords $x^n(m_c, m_r) \in \mathcal{T}_\epsilon^n(\mathrm{P_X}) \subset \mathcal{X}^n$, where $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$, such that for all $s^n \in \mathcal{S}^n$ and $m_c \in \mathcal{M}_c$, the following holds:*

$$\text{if } \frac{\log |\mathcal{M}_r|}{n} \geq \max_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X}; \mathrm{Z}_q) + \tau, \text{ then}$$

$$\left\| \frac{1}{|\mathcal{M}_r|} \sum_{m_r=1}^{|\mathcal{M}_r|} \mathrm{V}_{s^n}(\cdot|x^n(m_c, m_r)) - \mathbb{E}\big[\mathrm{V}_{s^n}(\cdot|\mathrm{X}^n)\big] \right\|_1 \leq 2^{-n\delta(\tau)} \tag{5.70}$$

where $\mathbb{E}[\cdot]$ is the expectation, X^n is distributed according to $\mathbb{P}(\mathrm{X}^n = x^n) := \frac{1}{|T_\mathrm{X}|} \mathbb{1}_{T_\mathrm{X}}(x^n)$ and $\lim_{\tau \to 0} \delta(\tau) = 0$.

The proof of the previous lemma follows from the strong secrecy analysis introduced in Section 2.2.3. Using the triangle inequality along with the relation between the total variation distance and the mutual information established in Lemma A.6, one can easily show that the condition in (5.70) assures that the information leakage of the confidential message to the eavesdropper with respect to the strong secrecy decays exponentially as follows:

$$\max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}_{s^n}^n) \leq 2^{-n\delta(\tau)}.$$

5.3.3 Direct Coding Approach

In this section, we present a detailed achievability proof for Theorem 5.9. The first case of Theorem 5.9 states that if $\mathfrak{W}$ is L-symmetrizable, then $C_{\text{list}}^{\text{s}}(\mathfrak{Q}, L) = 0$, so there is nothing to prove. For the second case, we need to show that if $\mathfrak{W}$ is not L-symmetrizable then $C_{\text{ran}}^{\text{s}}(\mathfrak{Q})$ is achievable. To do so, we will present a direct coding approach that utilizes a secrecy list code of Definition 5.13. Our coding scheme combines the principles used to establish the public list capacity of AVCs in [195] along with the techniques used to establish the deterministic secrecy capacity of AVWCs in [166].

Before we jump to our analysis, we need to introduce a couple of notations at first. For the conditional distributions $\mathrm{P_{A|B}} \in \mathcal{P}(\mathcal{A}|\mathcal{B})$ and $\mathrm{P_{B|C}} \in \mathcal{P}(\mathcal{B}|\mathcal{C})$, we define $\mathrm{P_{A|C}} =$

$\mathrm{P}_{\mathrm{A}|\mathrm{B}} \circ \mathrm{P}_{\mathrm{B}|\mathrm{C}} \in \mathcal{P}(\mathcal{A}|\mathcal{C})$ as follows: $\mathrm{P}_{\mathrm{A}|\mathrm{C}}(a|c) = \sum_{b\in\mathcal{B}} \mathrm{P}_{\mathrm{A}|\mathrm{B}}(a|b)\mathrm{P}_{\mathrm{B}|\mathrm{C}}(b|c)$. Moreover, for the distributions $\mathrm{P}_{\mathrm{A}} \in \mathcal{P}(\mathcal{A})$ and $\mathrm{P}_{\mathrm{B}} \in \mathcal{P}(B)$, we define $\mathrm{P}_{\mathrm{A},\mathrm{B}} = \mathrm{P}_{\mathrm{A}} \otimes \mathrm{P}_{\mathrm{B}} \in \mathcal{P}(\mathcal{A} \times \mathcal{B})$ as follows: $\mathrm{P}_{\mathrm{A},\mathrm{B}}(a, b) = \mathrm{P}_{\mathrm{A}}(a)\mathrm{P}_{\mathrm{B}}(b)$. Finally, for an alphabet $\mathcal{A}$, we define the conditional distribution ${\mathrm{P}_{\mathrm{A}|\mathrm{A}}}^{\mathrm{Id}} \in \mathcal{P}(\mathcal{A}|\mathcal{A})$, such that ${\mathrm{P}_{\mathrm{A}|\mathrm{A}}}^{\mathrm{Id}}(a|\bar{a}) = 1$ if and only if $a = \bar{a}$. Now, We start our analysis by pointing out the main challenge of using a direct coding approach to establish Theorem 5.9.

Coding Problem: Theorem 5.9 shows that the list secrecy capacity of an AVWC: $\mathfrak{Q}$ exhibits a certain dichotomy, where the capacity is either equivalent to the correlated random secrecy capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$, or it is zero. Although this might seems like a normal behavior, it has an alarming consequence. In order to understand such consequence, we need to highlight the following points:

- The condition that guides the dichotomy behavior of the capacity is the order of symmetrizability of the AVC $\mathfrak{W}$ between the transmitter and the legitimate receiver.
- The expression of the correlated random secrecy capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ in (5.38) is calculated with respect to the auxiliary AVWC $\tilde{\mathfrak{Q}} = (\tilde{\mathfrak{W}}, \tilde{\mathfrak{V}})$ instead of the original AVWC $\mathfrak{Q}$. Thus, we have to keep in mind that the AVC that guides the communication between the transmitter and the legitimate receiver is given by: $\tilde{\mathfrak{W}} = (\mathfrak{W} \circ \mathrm{P}_{\mathrm{X}^n|\mathrm{U}})$. This AVC arises from combining the prefix channel $\mathrm{P}_{\mathrm{X}^n|\mathrm{U}} \in \mathcal{P}(\mathcal{X}^n|\mathcal{U})$ and the original AVC $\mathfrak{W}$.
- It was shown in [166, Example 1] that applying a prefix channel can change a non-symmetrizable AVC to a symmetrizable one. This implies that, although $\mathfrak{W}$ is not L-symmetrizable, $\tilde{\mathfrak{W}}$ might be L-symmetrizable.

The previous discussion summarizes the main challenge of using a direct coding approach to establish Theorem 5.9. This is because all the reliable encoding-decoding techniques introduced in the previous section for list codes are only valid if the AVC between the transmitter and the legitimate receiver is not L-symmetrizable. In other words, in order to use the techniques introduced in the previous section, we have to make sure that $\tilde{\mathfrak{W}}$ is not L-symmetrizable.

In order to overcome this issue, we used an approach that redefines the optimization problem used to calculate the expression of the correlated random secrecy capacity in (5.38). We start by transforming the auxiliary random variable U to another one U_n defined over the alphabet $\mathcal{U}_n = \mathcal{X}^n$. We then restrict the calculation of $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ to certain class of prefix channels of the form $\hat{\mathrm{P}}_{\mathrm{X}^n|\mathrm{U}_n} = \mathrm{P}^{\mathrm{Id}}_{\mathrm{X}|\mathrm{X}} \otimes \check{\mathrm{P}}_{\mathrm{X}^{n-1}|\mathrm{U}_{n-1}}$. We then show that using this restricted class of prefix channels to calculate $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ is asymptotically as good as using the full set of prefix channels $\mathcal{P}(\mathcal{X}^n|\mathcal{U}_n)$. The reason for choosing this restricted class of prefix channels is that it assures that the resultant AVC: $\tilde{\mathfrak{W}} = (\mathfrak{W} \circ \hat{\mathrm{P}}_{\mathrm{X}^n|\mathrm{U}_n})$ is not L-symmetrizable as long as $\mathfrak{W}$ is not L-symmetrizable. We now present our coding scheme that consists of 4 main steps:

1) Prefix Channel Optimization: We start by considering the input distributions $\mathrm{P}_{\mathrm{U}_n}$ and the conditional distributions $\mathrm{P}_{\mathrm{X}^n|\mathrm{U}_n}$ arising from the optimization problem in (5.38). Without loss of generality, for every $r \in \mathbb{N}$, let $\mathcal{U}_r = \mathcal{X}^r$ and define the following:

$$C_r := \max_{\mathrm{P}_{\mathrm{U}_r}\in\mathcal{P}(\mathcal{U}_r)} \max_{\mathrm{P}_{\mathrm{X}^r|\mathrm{U}_r}\in\mathcal{P}(\mathcal{X}^r|\mathcal{U}_r)} \Big(\min_{q\in\mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\mathrm{U}_r; \mathrm{Y}^r_q) - \max_{s^r\in\mathcal{S}^r} \mathbb{I}(\mathrm{U}_r; \mathrm{Z}^r_{s^r}) \Big), \tag{5.71}$$

where $W_q^r(y^r|x^r) = \sum_{s^r} q(s^r) W_{s^r}^r(y^r|x^r)$. Then, for an arbitrary but fixed $r \in \mathbb{N}$ and an arbitrary $\epsilon \geq 0$, let $\mathrm{P}^*_{\mathrm{U}_r} \in \mathcal{P}(\mathcal{U}_r)$ and $\mathrm{P}^*_{\mathrm{X}^r|\mathrm{U}_r} \in \mathcal{P}(\mathcal{X}^r|\mathcal{U}_r)$ be such that

$$C_r - \epsilon = \min_{q \in \mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\mathrm{U}_r^*; \mathrm{Y}_q^r) - \max_{s^r \in \mathcal{S}^r} \mathbb{I}(\mathrm{U}_r^*; \mathrm{Z}_{s^r}^r). \tag{5.72}$$

We further let $\tilde{\mathrm{P}}_{\mathrm{U}_{r+1}} \in \mathcal{P}(\mathcal{U}_{r+1})$, be such that $\tilde{\mathrm{P}}_{\mathrm{U}_{r+1}} := \mathrm{P}^*_{\mathrm{U}_r} \otimes \pi$, where $\pi \in \mathcal{P}(\mathcal{X})$ is defined as $\pi(x) := |\mathcal{X}|^{-1}$ and $\tilde{\mathrm{P}}_{\mathrm{X}_{r+1}|\mathrm{U}_{r+1}} \in \mathcal{P}(\mathcal{X}^{r+1}|\mathcal{U}_{r+1})$, be such that $\tilde{\mathrm{P}}_{\mathrm{X}_{r+1}|\mathrm{U}_{r+1}} := \mathrm{P}^*_{\mathrm{X}^r|\mathrm{U}_r} \otimes \sigma$, where $\sigma \in \mathcal{P}(\mathcal{X}|\mathcal{X})$ is defined as $\sigma(x|\bar{x}) = 1$ if and only if $x = \bar{x}$. Then, from (5.71), it holds that

$$\begin{aligned}
C_{r+1} &\geq \min_{q \in \mathcal{P}(\mathcal{S}^{r+1})} \mathbb{I}(\tilde{\mathrm{U}}_{r+1}; \mathrm{Y}_q^{r+1}) - \max_{s^{r+1} \in \mathcal{S}^{r+1}} \mathbb{I}(\tilde{\mathrm{U}}_{r+1}; \mathrm{Z}_{s^{r+1}}^{r+1}) \\
&\overset{(a)}{=} \min_{q \in \mathcal{P}(\mathcal{S}^{r+1})} \mathbb{I}(\mathrm{U}_r^* \mathrm{X}_\pi; \mathrm{Y}_q^{r+1}) - \max_{s^{r+1} \in \mathcal{S}^{r+1}} \mathbb{I}(\mathrm{U}_r^* \mathrm{X}_\pi; \mathrm{Z}_{s^{r+1}}^{r+1}) \\
&\overset{(b)}{=} \min_{q \in \mathcal{P}(\mathcal{S}^{r+1})} \mathbb{I}(\mathrm{U}_r^* \mathrm{X}_\pi; \mathrm{Y}_q^{r+1}) - \max_{s^r \in \mathcal{S}^r} \mathbb{I}(\mathrm{U}_r^*; \mathrm{Z}_{s^r}^r) - \max_{s \in \mathcal{S}} \mathbb{I}(\mathrm{X}_\pi; \mathrm{Z}_{s_{r+1}}) \\
&\overset{(c)}{\geq} \min_{q \in \mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\mathrm{U}_r^*; \mathrm{Y}_q^r) - \max_{s^r \in \mathcal{S}^r} \mathbb{I}(\mathrm{U}_r^*; \mathrm{Z}_{s^r}^r) - \log|\mathcal{X}| \\
&= C_r - \epsilon - \log|\mathcal{X}|
\end{aligned} \tag{5.73}$$

where (a) follows from the definition of $\tilde{\mathrm{P}}_{\mathrm{U}_{r+1}}$ and $\tilde{\mathrm{P}}_{\mathrm{X}_{r+1}|\mathrm{U}_{r+1}}$; (b) follows from the mutual information chain rule and the fact that U_r^* and $\mathrm{Z}_{s_{r+1}}$ are independent as well as X_π and $\mathrm{Z}_{s^r}^r$; while (c) follows because $\mathbb{I}(\mathrm{X}_\pi; \mathrm{Z}_{s_{r+1}}) \leq \log|\mathcal{X}|$. Now, let us investigate the effect of the prefix channel $\tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r}$ on the symmetrizability of the resultant channel $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r})$ as follows:

$$\begin{aligned}
\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r} &= \mathfrak{W}^{\otimes r} \circ (\mathrm{P}^*_{\mathrm{X}_{r-1}|\mathrm{U}_{r-1}} \otimes \sigma) \\
&= (\mathfrak{W}^{\otimes r-1} \circ \mathrm{P}^*_{\mathrm{X}_{r-1}|\mathrm{U}_{r-1}}) \otimes (\mathfrak{W} \circ \sigma) \\
&= (\mathfrak{W}^{\otimes r-1} \circ \mathrm{P}^*_{\mathrm{X}_{r-1}|\mathrm{U}_{r-1}}) \otimes \mathfrak{W},
\end{aligned} \tag{5.74}$$

where the previous relation follows due to the properties of the operators $\circ$ and $\otimes$ along with the definition of the conditional distribution σ. Since the AVC: $\mathfrak{W}$ is not L-symmetrizable, then for every $r \geq 2$, it follows that the AVC: $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r})$ is not L-symmetrizable as well, even if $(\mathfrak{W}^{\otimes r-1} \circ \mathrm{P}^*_{\mathrm{X}_{r-1}|\mathrm{U}_{r-1}})$ is L-symmetrizable.

2) Block Coding: In this point, we describe the construction of a list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ according to Definition 5.13, for the AVWC $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r}, \mathfrak{V}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r})$. We start by generating our codebook as follows: For a code block length $t \in \mathbb{N}$ and $r \in \mathbb{N}\backslash\{1\}$, let $\mathrm{P}_{\mathrm{U}_r} \in \mathcal{P}_0^t(\mathcal{U}_r)$ and use it to generate the set of codewords $u_r^t(m_c, m_r) \in \mathcal{T}_\epsilon^t(\mathrm{P}_{\mathrm{U}_r}) \subset \mathcal{U}_r^t$, for $(m_c, m_r) \in \mathcal{M}_c \times \mathcal{M}_r$ and a typicality constant $\epsilon > 0$. Using the union bound, we can show that the produced codewords satisfy the constraints in Lemma 5.3 and Lemma 5.6 with probability approaches one as $t \to \infty$.

Next, we define our encoding and decoding functions as follows: Given a confidential message $m_c \in \mathcal{M}_c$, the encoder chooses a randomization message $m_r \in \mathcal{M}_r$ uniformly at random then outputs the codeword $u_r^t(m_c, m_r)$. At the legitimate receiver, we use a list decoder φ_L similar to the one given by Definition 5.15, that takes the received sequence y^n and outputs a list of message pairs $\mathcal{J} \in \mathfrak{P}_L(\mathcal{M}_c \times \mathcal{M}_r)$.

3) Reliability and secrecy analysis: Lemma 5.5 implies that for an AVWC, whose legitimate channel is not L-symmetrizable, there exists a list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ with list size L and block length t constructed as described in the previous point which can be used to transmit the messages (m_c, m_r) reliably over the channel $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r})$ as long as:

$$\min_{q\in\mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\mathrm{U}_r;\mathrm{Y}^r_q) - 2\delta/3 \geq \frac{1}{t}\log\frac{|\mathcal{M}_c||\mathcal{M}_r|}{L} \geq \min_{q\in\mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\mathrm{U}_r;\mathrm{Y}^r_q) - \delta, \tag{5.75}$$

where $\delta > 0$. The previous result is only valid if the AVC $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}_r|\mathrm{U}_r})$ is not L-symmetrizable. Moreover, we mean by reliably transmitted that the average decoding error probability of $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ decays exponentially fast. Thus, for $\zeta > 0$ we have

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) = \max_{s^{r\cdot t}\in\mathcal{S}^{r\cdot t}} \bar{\mathbf{P}}_e(s^{r\cdot t}|\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq 2^{-t\zeta}. \tag{5.76}$$

On the other hand, based on the secrecy analysis in [166] that relates the strong secrecy constraint in (5.59) to variations distance in Lemma 5.6, we can show that the constructed list code is asymptotically secure in the strong sense, as long as

$$\max_{s^r\in\mathcal{S}^r} \mathbb{I}(\mathrm{U}_r;\mathrm{Z}^r_{s^r}) + \delta \leq \liminf_{t\to\infty} \frac{1}{t}\log|\mathcal{M}_r| \leq \max_{s^r\in\mathcal{S}^r} \mathbb{I}(\mathrm{U}_r;\mathrm{Z}^r_{s^r}) + 2\delta. \tag{5.77}$$

It is worth mentioning that by asymptotically secure in the strong sense, we mean that the information leakage of the confidential message to the eavesdropper decays exponentially fast. Thus, for $\tau > 0$, we have

$$\max_{s^{r\cdot t}\in\mathcal{S}^{r\cdot t}} \mathbb{I}(\mathrm{M};\mathrm{Z}^{r\cdot t}_{s^{r\cdot t}}) \leq 2^{-t\tau}. \tag{5.78}$$

Now, let $\mathrm{P}_{\mathrm{U}_r} \in \mathcal{P}^t_0(\mathcal{U}_r)$ converges to $\tilde{\mathrm{P}}_{\mathrm{U}_r} \in \mathcal{P}(\mathcal{U}_r)$ as $t \to \infty$, such that $\tilde{\mathrm{P}}_{\mathrm{U}_r} = \tilde{P}^*_{\mathrm{U}_{r-1}} \otimes \pi$, where π is as defined before and $\mathrm{P}^*_{\mathrm{U}_{r-1}} \in \mathcal{P}(\mathcal{U}_{r-1})$ being an optimal choice for the optimization problem in (5.71), then from (5.73), (5.75) and (5.77) we have

$$\begin{aligned}\liminf_{t\to\infty} \frac{1}{t}\log\frac{|\mathcal{M}_c|}{L} &\geq \min_{q\in\mathcal{P}(\mathcal{S}^r)} \mathbb{I}(\tilde{\mathrm{U}}_r;\mathrm{Y}^r_q) - \max_{s^r\in\mathcal{S}^r} \mathbb{I}(\tilde{\mathrm{U}}_r;\mathrm{Z}^r_{s^r}) - 3\delta. \\ &\geq C_{r-1} - \log|\mathcal{X}| - 3\delta.\end{aligned} \tag{5.79}$$

4) Code Transformation: Up until this point, we have shown that we can construct a list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ that can be used to establish a reliable and secure communication over the AVWC $(\mathfrak{W}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}^r|\mathrm{U}_r}, \mathfrak{V}^{\otimes r} \circ \tilde{\mathrm{P}}_{\mathrm{X}^r|\mathrm{U}_r})$. Now, we want to use this code to establish a reliable and secure communication over the original AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$.

Let $\tilde{t} \in \{0, \ldots, r-1\}$, such that for every $n \in \mathbb{N}$, we define $n = t \cdot r + \tilde{t}$. Since we assumed without loss of generality that $\mathcal{U}_r = \mathcal{X}^r$, we can transform any codeword $u^t_r(m_c, m_r)$ into a codeword $x^{t\cdot r}(m_c, m_r)$. Next, we generate the codewords $x^n(m_c, m_r)$ for $m_c \in \mathcal{M}_c$ and $m_r \in \mathcal{M}_r$ by concatenating a dummy codeword $x^{\tilde{t}}$ to the transformed codeword $x^{t\cdot r}(m_c, m_r)$. Using this technique, we are able to transform the list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ constructed on the alphabet $\mathcal{U}_r$ and with block length t into a new list code $\bar{\mathcal{C}}^{\mathrm{s}}_{\mathrm{list}}$ constructed on the alphabet $\mathcal{X}$ and with block length n.

One can easily show that under this transformation mechanism, the newly constructed code $\bar{\mathcal{C}}^{\mathrm{s}}_{\mathrm{list}}$ will have the same reliability and secrecy performance of the original code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$. This implies that we have constructed a list code $\bar{\mathcal{C}}^{\mathrm{s}}_{\mathrm{list}}$ with list size L that can

be used to establish a reliable and secure communication over the original AVWC: $\mathfrak{Q}$, where the rate of this code is given by:

$$\begin{aligned} R_c &= \liminf_{n\to\infty} \frac{1}{n} \log \frac{|\mathcal{M}_c|}{L} \\ &= \liminf_{t\to\infty} \frac{1}{t\cdot r + \tilde{t}} \log \frac{|\mathcal{M}_c|}{L} \\ &\geq \liminf_{t\to\infty} \frac{1}{r}\cdot\frac{1}{t} \log \frac{|\mathcal{M}_c|}{L} \\ &= \frac{1}{r}(C_{r-1} - \log|\mathcal{X}| - 3\delta). \end{aligned} \tag{5.80}$$

Finally, since $\lim_{r\to\infty} \frac{r-1}{r} = 1$, it follows that $C^{\mathrm{s}}_{\mathrm{list}}(\mathfrak{Q}, L) \geq \lim_{r\to\infty} \frac{1}{r} C_r = C^{s}_{\mathrm{ran}}(\mathfrak{Q})$, where we replaced the random variable U_n by a random variable U that takes values in an alphabet $\mathcal{U} = \mathcal{X}^n$. This argument completes our achievability proof.

It is important to point out that the previous coding scheme establishes the achievability of Theorem 5.9 without imposing any restriction on the information shared between the jammer and the eavesdropper. This feature is a consequence of using a pure secrecy list code of Definition 5.13.

5.3.4 Elimination of Correlation Approach

In this section, we present an alternative achievability proof for Theorem 5.9. Our proof is based on the principle of elimination of correlation introduced by Ahlswede in [163]. Similar to the previous section, we will only show that if the AVC $\mathfrak{W}$ between the transmitter and the legitimate receiver is not L-symmetrizable, then one can use list codes to achieve the correlated random secrecy capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$. We consider a coding scheme that combines the techniques used to derive the public list capacity of AVCs in [194] along with the results established in [159] for correlated random secrecy codes.

1- Code Structure: We construct a secrecy list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ of Definition 5.13 with list size L and block length $n = \hat{n} + \tilde{n}$ by concatenating a public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ of Definition 5.10 with list size L and block length $\hat{n}$ and a correlated random secrecy code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ of Definition 5.5 with block length $\tilde{n}$. The idea is to use the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ to encode a helper public message $\mathcal{M}_p$ that will be used to establish some sort of correlated randomness between the transmitter and the legitimate receiver. On the other hand, the correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ is used to encode the actual confidential message $\mathcal{M}_c$. In order for this coding scheme to work, we have to make the set of public messages $\mathcal{M}_p$ transmitted by $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ identical to the set of values of the common randomness $\mathcal{G}$ used by $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$, i.e. $\mathcal{M}_p = \mathcal{G}$.

2- Encoding: Given a confidential message $m_c \in \mathcal{M}_c$, the encoder chooses an index $\gamma \in \mathcal{G}$ uniformly at random and encodes it using the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ producing a codeword $x^{\hat{n}}(\gamma)$. After that m_c and γ are then given to the encoder of the correlated random secrecy code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ which outputs a codeword $x^{\tilde{n}}_{\gamma}(m_c)$. Finally the encoder concatenates the two codewords producing $x^n(\gamma, m_c) = \big(x^{\hat{n}}(\gamma), x^{\tilde{n}}_{\gamma}(m_c)\big)$ and transmits it.

3- Decoding: Having received $y^n = (y^{\hat{n}}, y^{\tilde{n}})$, the decoder works as follows: First it uses $y^{\hat{n}}$ and the decoder of the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ to produce a list of size L, i.e. $\hat{\mathcal{G}} = \varphi_L(y^{\hat{n}}) \in \mathfrak{P}_L(\mathcal{G})$. Next, for every $\hat{\gamma} \in \hat{\mathcal{G}}$, we use $y^{\tilde{n}}$ along with the decoder $\varphi^{\hat{\gamma}}$ of the correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ to find a corresponding confidential message $\hat{m}_c \in \mathcal{M}_c$. Thus, the final decoding output is a list of size L as follows:

$$\left\{\hat{m}_c \in \mathcal{M}_c : \hat{m}_c = \varphi^{\hat{\gamma}}(y^{\tilde{n}})\ \forall \hat{\gamma} \in \varphi_L(y^{\hat{n}})\right\}.$$

4- Reliability Analysis: Based on our coding scheme, we can bound the average decoding error probability of our code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ in terms of the average decoding error probability of $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ given by (5.53) and the average decoding error probability of $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ given by (5.26) as follows:

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq \bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) + \bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}). \tag{5.81}$$

Based on the results presented in Section 5.2.2, we can construct a correlated random secrecy code whose rate R_c is close to $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$, such that $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}})$ decays exponentially fast. The amount of common randomness required to construct such code is quadratic in the block length as highlighted by [14, Lemma 6], i.e. $|\mathcal{G}| = \mathcal{O}(\tilde{n}^2)$. This implies that the public rate R_p needed to be reliably transmitted by the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ is given by:

$$R_p = \frac{1}{\hat{n}} \log |\mathcal{G}| = \frac{2}{\hat{n}} \log \mathcal{O}(\tilde{n}). \tag{5.82}$$

Since the AVC $\mathfrak{W}$ is not L-symmetrizable, Lemma 5.4 implies that $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}})$ decays exponentially fast as long as R_p is small enough. According to (5.82), as $n \to \infty$, R_p vanishes. This implies that $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})$ decays exponentially fast as long as $R_c \leq \lim_{n\to\infty} \frac{\tilde{n}}{n} C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q}) = C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$.

5- Secrecy Analysis: Let Γ be a uniformly distributed random variable over the set $\mathcal{G}$. Then, based on the structure of our coding scheme, the secrecy constraint in (5.59) can be bounded as follows:

$$\begin{aligned}
\max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n} | \mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) &\overset{(a)}{\leq} \max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n} | \mathrm{M}_p \mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \\
&\overset{(b)}{=} \max_{s^{\hat{n}} \in \mathcal{S}^{\hat{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\hat{n}}_{s^{\hat{n}}} | \mathrm{M}_p \mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) + \max_{s^{\tilde{n}} \in \mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}} | \mathrm{Z}^{\hat{n}}_{s^{\hat{n}}} \mathrm{M}_p \mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \\
&\overset{(c)}{=} \max_{s^{\tilde{n}} \in \mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}} | \mathrm{M}_p \mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \\
&\overset{(d)}{=} \max_{s^{\tilde{n}} \in \mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}} | \Gamma \mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}) \\
&\leq \max_{s^{\tilde{n}} \in \mathcal{S}^{\tilde{n}}} \max_{\gamma \in \mathcal{G}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}}(\gamma) | \mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}) \overset{(e)}{\leq} e^{-\tilde{n}\epsilon},
\end{aligned} \tag{5.83}$$

where (a) follows due to the independence of the messages M_p and M_c; (b) follows from the mutual information chain rule; (c) follows because M_c and $\mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}$ are independent along with the fact that $\mathrm{M}_p - \mathrm{X}^{\hat{n}} - \mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}$ forms a Markov chain; $d)$ follows because Γ is identical to M_p and the fact that M_c is only encoded using $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$; (e) follows from the secrecy analysis of the correlated random codes in [159], where $\epsilon > 0$.

The previous coding scheme shows that for an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, where the AVC $\mathfrak{W}$ is not L-symmetrizable, there exists a secrecy list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ with list size L with a confidential rate $R_c = C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$. This completes our achievability proof. It is important to highlight that the coding scheme introduced in this section is only valid under the assumption of restricted communication between the eavesdropper and the jammer. This is because the code concept used in this section inherits the characteristics of the sub-codes used to construct it and in particular the correlated random codes of Definition 5.5.

5.3.5 List Codes With Finite List Size

Theorem 5.9 shows that in order to achieve a reliable and secure communication over an AVWC, using a list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ that satisfies the following requirements:

1. The average decoding error probability $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})$ given by (5.58) decays exponentially fast with the code block length n.
2. The information leakage of the confidential message M_c with respect to the strong secrecy criteria $\mathbb{L}(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})$ given by (5.59) also decays exponentially fast with the code block length n.

We need a list code with list size L that depends on the order of symmetrizability of the AVC: $\mathfrak{W}$ between the transmitter and the legitimate receiver, i.e. $L \geq L(\mathfrak{W}) + 1$. This implies that the capability of list codes to overcome the different jamming strategies induced by the jammer depends on the order of symmetrizability of the AVC $\mathfrak{W}$. This observation raised some speculations about the usability of list codes for AVWCs. The reason for that originates from the results established in [194, Theorem 3], which showed that even for simple AVCs, the order of symmetrizability $L(\mathfrak{W})$ might be arbitrarily large. This implies that, for their corresponding AVWCs, we need list codes with an arbitrarily large list size, which is not always feasible.

The previous issue motivated us to investigate the following problem: Does relaxing the reliability and secrecy constraints affect the value of the list size. In order to answer this question, we consider the following scenario: Given an AVWC $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, where we allow for a small but non-vanishing probability of error λ and a small but non-vanishing information leakage τ, what is the maximum rate that we can achieve using list codes and what is the smallest list size required to achieve such rate. We formalize our investigation using the following definition:

Definition 5.16. For an arbitrary but fixed $\lambda, \tau > 0$, a non negative number R_c is a (λ, τ)-achievable list secrecy rate with list size L for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, if for all $\delta > 0$, there is an $n(\delta) \in \mathbb{N}$, such that for all $n \geq n(\delta)$, there exists a sequence of secrecy list codes $(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}})_n$, that satisfy the following:

$$\frac{1}{n} \log \frac{|\mathcal{M}_c|}{L} \geq R_c - \delta, \tag{5.84}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq \lambda, \tag{5.85}$$

$$\mathbb{L}(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq \tau. \tag{5.86}$$

The supremum of all (λ, τ)-achievable list secrecy rates R_c with list size L is denoted by $C^{\mathrm{s}}_{\mathrm{list}}(\mathfrak{Q}, \lambda, \tau, L)$.

It is important to highlight the difference between this definition and Definition 5.14. In Definition 5.14, we need to find list codes that satisfy the reliability and secrecy constraints in (5.61) and (5.62) for all values of $\lambda, \tau > 0$, while in the previous definition we need to satisfy the reliability and secrecy constraints for certain values of λ and τ. We now present the following result:

Theorem 5.10. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ with correlated random secrecy capacity $C^s_{ran}(\mathfrak{Q})$, let $\lambda, \tau > 0$ be arbitrary but fixed. Then for every secrecy rate $R_c < C^s_{ran}(\mathfrak{Q})$, there exists a finite list size L, such that*

$$R_c < C^s_{list}(\mathfrak{Q}, \lambda, \tau, L).$$

The proof of the previous theorem consists of two main steps. In the first step, we show that there exists a correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ that can achieve the correlated random capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ by utilizing a finite amount of shared common randomness $|\tilde{\mathcal{G}}|$ such that its average error probability $\bar{\mathbf{P}}_e(\tilde{\mathcal{C}}^s_{\mathrm{cr}})$ given by (5.26) is less than or equal to λ, while its maximum information leakage $\mathbb{L}^{\max}(\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}})$ given by (5.28) is less than or equal to τ. In the second step, we show how the constructed correlated random code $\tilde{\mathcal{C}}^s_{\mathrm{cr}}$ can be transformed to a list code using the concept of correlation elimination as shown in Section 5.3.4.

Corollary 5.2. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, every secrecy rate $R_c < C^s_{ran}(\mathfrak{Q})$ is a (λ, τ)-achievable secrecy rate using a list code with list size $L(\lambda)$ given by the smallest natural number that satisfy:*

$$L(\lambda) > \frac{1}{\epsilon\lambda} \ln |\mathcal{S}|,$$

for some constant $\epsilon > 0$ and $\lambda \in (0, 1)$, where $\mathcal{S}$ is the channel state set. Additionally, the information leakage parameter τ can be selected, such that $\lim_{n\to\infty} \tau = 0$.

Theorem 5.10 and Corollary 5.2 provide an answer for the two questions raised at the beginning of this section as follows: Theorem 5.10 implies that we can construct list codes that achieve secrecy rates up to the correlated random secrecy capacity, if we allow for a small but non-vanishing probability of error and information leakage. On the other hand, Corollary 5.2 specifies a sufficient condition for the list size L required for the construction of such codes, which interestingly only depends on the reliability parameter λ and is independent of the secrecy parameter τ.

We now derive the first step required to prove Theorem 5.10. In particular, we construct a correlated random secrecy code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ that only utilizes a finite amount of shared common randomness $|\tilde{\mathcal{G}}|$. We then show that the constructed code can provide a secure communication over the AVWC: $\mathfrak{Q}$ at a rate equivalent to the correlated random secrecy capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ under the following constraints:

1. The maximum information leakage of the confidential message to the eavesdropper given by (5.28) vanishes.
2. The average decoding error probability $\bar{\mathbf{P}}_e(\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{rsn}})$ given by (5.26) does not exceed an arbitrary but fixed value $\lambda \in (0, 1)$.

For this step, we present a lemma that extends the result established in [198, Theorem 3] for correlated random codes with finite resources under the mean information leakage of the confidential message with respect to the weak secrecy criterion.

Lemma 5.7. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, let $\lambda \in (0,1)$ be arbitrary but fixed. There exists a correlated random code $\tilde{\mathcal{C}}^s_{ran}$ that can achieve secrecy rates equivalent to the correlated random secrecy capacity $C^s_{ran}(\mathfrak{Q})$ using a finite set $\tilde{\mathcal{G}}$ of values of a shared common randomness, such that:*

$$\bar{\boldsymbol{P}}_e(\tilde{\mathcal{C}}^s_{ran}) \leq \lambda \tag{5.87}$$

$$\lim_{n\to\infty} \mathbb{L}^{max}(\tilde{\mathcal{C}}^s_{ran}) = 0. \tag{5.88}$$

For an AVWC: $\mathfrak{Q}$, we know from Theorem 5.4 that there exists a correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ that can achieve the correlated random capacity $C^{\mathrm{s}}_{\mathrm{ran}}(\mathfrak{Q})$ given by (5.38) using a correlated randomness source Γ, where cardinality of the set of values of this source $\mathcal{G}$ grows quadratic in the block length n, i.e. $|\mathcal{G}| = \mathcal{O}(n^2)$. We also know that for this code, both the average decoding error probability $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}})$ given by (5.26) and the maximum information leakage $\mathbb{L}^{\max}(\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}})$ given by (5.28) decays exponentially fast as highlighted by the reliability and secrecy analysis in [159].

Now, Consider a correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ constructed by selecting $|\tilde{\mathcal{G}}|$ deterministic codes from the original correlated random code $\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}$ used to prove the achievability of Theorem 5.4, where $\tilde{\mathcal{G}} \subset \mathcal{G}$ is a finite set whose cardinality is independent of n. Now, let $\alpha > 0$ be arbitrary but fixed, then for a fixed channel state sequence $s^n \in \mathcal{S}^n$, we have

$$\begin{aligned}
\mathbb{P}\left[\frac{1}{|\tilde{\mathcal{G}}|}\sum_{\gamma\in\tilde{\mathcal{G}}} \bar{\mathbf{P}}_e(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)) \geq \lambda\right] &= \mathbb{P}\left[\exp\Big(\alpha\sum_{\gamma\in\tilde{\mathcal{G}}} \bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\Big) \geq \exp\Big(\alpha\lambda|\tilde{\mathcal{G}}|\Big)\right] \\
&\overset{(a)}{\leq} \exp\Big(-\alpha\lambda|\tilde{\mathcal{G}}|\Big)\mathbb{E}\left[\exp\Big(\alpha\sum_{\gamma\in\tilde{\mathcal{G}}} \bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\Big)\right] \\
&= \exp\Big(-\alpha\lambda|\tilde{\mathcal{G}}|\Big)\prod_{\gamma\in\tilde{\mathcal{G}}}\mathbb{E}\Big[\exp\Big(\alpha\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\Big)\Big]
\end{aligned} \tag{5.89}$$

where (a) follows using the Markov inequality. If we let $\epsilon > 0$ and use the Taylor expansion of $\exp(t)$, we can bound the expectation term in (5.89) as follows:

$$\begin{aligned}
\mathbb{E}\Big[\exp\Big(\alpha\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\Big)\Big] &= \mathbb{E}\left[\sum_{k=0}^{\infty}\frac{\Big(\alpha\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\Big)^k}{k!}\right] \\
&\overset{(a)}{\leq} \mathbb{E}\left[1 + \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\right] \\
&\overset{(b)}{\leq} 1 + \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\exp(-n\epsilon) \\
&\leq 1 + \exp(-n\epsilon + \alpha)
\end{aligned} \tag{5.90}$$

where (a) follows because $\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big) \leq 1$; while (b) follows as for a fixed channel state sequence s^n the expectation of $\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)$ is a negative exponential in n [166].

Now, if we take all the state sequences $s^n \in \mathcal{S}^n$ into consideration, we reach

$$\begin{aligned}\mathbb{P}\Bigg[\frac{1}{|\tilde{\mathcal{G}}|}\sum_{\gamma\in\tilde{\mathcal{G}}}\bar{\mathbf{P}}_e\big(s^n|\mathcal{C}^{\mathrm{s}}_{\mathrm{det}}(\gamma)\big)\geq\lambda,\,\forall s^n\in\mathcal{S}^n\Bigg]&\\ &\overset{(a)}{\leq}\exp\Big(-\alpha\lambda|\tilde{\mathcal{G}}|\Big)\cdot\Big(1+\exp(-n\epsilon+\alpha)\Big)^{|\tilde{\mathcal{G}}|}\cdot|\mathcal{S}|^n\\ &\overset{(b)}{\leq}\exp\Big(-n\epsilon\lambda|\tilde{\mathcal{G}}|+|\tilde{\mathcal{G}}|\ln 2+n\ln|\mathcal{S}|\Big)\end{aligned}\tag{5.91}$$

where (a) follows due to (5.89) and (5.90); while (b) follows by choosing $\alpha = n\epsilon$. Now, if we make sure that:

$$|\tilde{\mathcal{G}}| > \frac{1}{\epsilon\lambda}\ln|\mathcal{S}|,\tag{5.92}$$

then by (5.91) the average decoding error probability of $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ is smaller than λ with a probability that approaches one exponentially fast. We now turn to the secrecy performance of $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$. Due to the way $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ was constructed, we have

$$\begin{aligned}\max_{s^n\in\mathcal{S}^n}\max_{\gamma\in\tilde{\mathcal{G}}}\mathbb{I}\big(\mathrm{M}_c;\mathrm{Z}^n_{s^n}(\gamma)|\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}\big)&\leq\max_{s^n\in\mathcal{S}^n}\max_{\gamma\in\mathcal{G}}\mathbb{I}\big(\mathrm{M}_c;\mathrm{Z}^n_{s^n}(\gamma)|\mathcal{C}^{\mathrm{s}}_{\mathrm{ran}}\big)\\ &\leq e^{-n\tilde{\epsilon}},\end{aligned}\tag{5.93}$$

where $\tilde{\epsilon} > 0$. The last step follows from the achievability proof of Theorem 5.4. This implies that for an AVWC: $\mathfrak{Q}$, we can construct a correlated random secrecy code that only utilizes a finite set $\tilde{\mathcal{G}}$ of values for a correlated randomness, where the cardinality of $\tilde{\mathcal{G}}$ is given by (5.92), such that the constraints in (5.87) and (5.88) are satisfied with high probability. This completes our proof for Lemma 5.7.

We now move to the second step required to prove Theorem 5.10, where we use the correlated random secrecy code with finite resources $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ constructed in the previous lines to construct a secrecy list code. This construction is based on the concept of elimination of correlation and follows similar steps to one used in Section 5.3.4. We present the following a coding scheme:

1- Code Structure: We construct a list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ of Definition 5.13 with list size L and block length $n = \hat{n} + \tilde{n}$ by concatenating a public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ of Definition 5.10 with list size L and block length $\hat{n}$ and a finite coordination correlated random secrecy code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ with block length $\tilde{n}$. In our construction, we let the set of public messages $\mathcal{M}_p$ encoded by the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ to be identical to the set of values of the correlated randomness $\tilde{\mathcal{G}}$ used by $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$, i.e. $\mathcal{M}_p = \tilde{\mathcal{G}}$. Additionally, we set the list size L as follows: $L = |\tilde{\mathcal{G}}| = |\mathcal{M}_p|$.

2- Encoding: Given a confidential message $m_c \in \mathcal{M}_c$, the encoder chooses an index $\gamma \in \tilde{\mathcal{G}} = \mathcal{M}_p$ uniformly at random and encodes it using the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ producing a codeword $x^{\hat{n}}(\gamma)$. After that m_c and γ are then given to the encoder of the correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ which outputs a codeword $x^{\tilde{n}}_{\gamma}(m_c)$. Finally the encoder concatenates the two codewords producing $x^n(\gamma, m_c) = (x^{\hat{n}}(\gamma), x^{\tilde{n}}_{\gamma}(m_c))$ and transmits it.

3- Decoding: Having received $y^n = (y^{\hat{n}}, y^{\tilde{n}})$, the decoding process is sequential as follows: First $y^{\hat{n}}$ is given to the decoder of the list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ to produce a list of size L, i.e. $\varphi_L(y^{\hat{n}}) = \hat{\mathcal{J}} \in \mathfrak{P}_L(\tilde{\mathcal{G}})$. Then for every $\hat{\gamma} \in \hat{\mathcal{J}}$, we use $y^{\tilde{n}}$ along with the decoder $\tilde{\varphi}^{\hat{\gamma}}$ of the correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ to find a corresponding confidential message $\hat{m}_c \in \mathcal{M}_c$. Thus, our decoder can be expressed as follows:

$$\varphi_L(y^n) = \Big\{\hat{m}_c \in \mathcal{M}_c : \hat{m}_c = \tilde{\varphi}^{\hat{\gamma}}(y^{\tilde{n}})\ \forall \hat{\gamma} \in \varphi_L(y^{\hat{n}})\Big\}.$$

4- Reliability Analysis: Based on our coding scheme, we can bound the average decoding error probability of $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ given by (5.58) in terms of the average decoding error probability of $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ given by (5.53) and the average decoding error probability of $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ given by (5.26) as follows:

$$\bar{\mathbf{P}}_e(\mathcal{C}^s_{\mathrm{list}}) \leq \bar{\mathbf{P}}_e(\mathcal{C}_{\mathrm{list}}) + \bar{\mathbf{P}}_e(\tilde{\mathcal{C}}^s_{\mathrm{cr}}). \tag{5.94}$$

$\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ is a public list code with list size L used to transmit the public message $\gamma \in \tilde{\mathcal{G}}$ where $L = |\tilde{\mathcal{G}}|$. This implies that $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ always makes a correct decoding decision. Thus, the first term in (5.94) vanishes, i.e. $\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}) = 0$. On the other hand, according to the proof of Lemma 5.7 if we choose the amount of shared randomness $\tilde{\mathcal{G}}$ utilized by the correlated random secrecy code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ such that $|\tilde{\mathcal{G}}|$ satisfies the constraint in (5.92), then the second term in (5.94) is less than λ with a probability that goes exponentially to one. Thus, we have

$$\mathbb{P}\Big[\bar{\mathbf{P}}_e(\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) \leq \lambda\Big] \underset{n\to\infty}{\longrightarrow} 1. \tag{5.95}$$

5- Secrecy Analysis: Let $\tilde{\Gamma}$ be a uniformly distributed random variable over the set $\tilde{\mathcal{G}}$. Then, based on the structure of our coding scheme, the secrecy constraint in (5.59) can be bounded as follows:

$$\begin{aligned}
\max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n}|\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}) &\overset{(a)}{\leq} \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^n_{s^n}|\mathrm{M}_p\mathcal{C}^s_{\mathrm{list}}) \\
&\overset{(b)}{=} \max_{s^{\hat{n}}\in\mathcal{S}^{\hat{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}|\mathrm{M}_p\mathcal{C}^s_{\mathrm{list}}) + \max_{s^{\tilde{n}}\in\mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}}|\mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}\mathrm{M}_p\mathcal{C}^s_{\mathrm{list}}) \\
&\overset{(c)}{=} \max_{s^{\tilde{n}}\in\mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}}|\mathrm{M}_p\mathcal{C}^s_{\mathrm{list}}) \\
&\overset{(d)}{=} \max_{s^{\tilde{n}}\in\mathcal{S}^{\tilde{n}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}}}|\tilde{\Gamma}\tilde{\mathcal{C}}^s_{\mathrm{cr}}) \\
&= \max_{s^{\tilde{n}}\in\mathcal{S}^{\tilde{n}}} \frac{1}{|\tilde{\mathcal{G}}|}\sum_{\gamma\in\tilde{\mathcal{G}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}},\gamma}|\tilde{\mathcal{C}}^s_{\mathrm{cr}}) \\
&\leq \max_{s^{\tilde{n}}\in\mathcal{S}^{\tilde{n}}} \max_{\gamma\in\tilde{\mathcal{G}}} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}^{\tilde{n}}_{s^{\tilde{n}},\gamma}|\tilde{\mathcal{C}}^s_{\mathrm{cr}}) \\
&\overset{(e)}{\leq} e^{-\tilde{n}\epsilon},
\end{aligned} \tag{5.96}$$

where (a) follows because M_c and M_p are independent; (b) follows from the mutual information chain rule; (c) follows because M_c and $\mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}$ are independent along with the fact that $\mathrm{M}_p - \mathrm{X}^{\hat{n}} - \mathrm{Z}^{\hat{n}}_{s^{\hat{n}}}$ forms a Markov chain; (d) follows because $\tilde{\Gamma}$ is identical to M_p and the fact that M_c is only encoded using $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$; while (e) follows from the secrecy analysis of the correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$ as in (5.93), where $\epsilon > 0$.

In order to finalize our proof, we need to show that the overall transmission rate of the constructed list code $\mathcal{C}^{\mathrm{s}}_{\mathrm{list}}$ is equivalent to the transmission rate of the correlated random code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{ran}}$. Since $\hat{n} = \mathcal{O}(\log L)$ where L is given by (5.92). This implies that $\hat{n}+n \to n$ as $n \to \infty$, which means that using the public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ as a prefix code on the top of the correlated random secrecy code $\tilde{\mathcal{C}}^{\mathrm{s}}_{\mathrm{cr}}$ does not affect its transmission rate.

5.3.6 Continuity, Additivity and Super-Activation

In this section, we investigate and discuss the continuity and additivity behavior of the public list capacity for AVCs and the list secrecy capacity of AVWCs. For this investigation, we need introduce some basic tools that play an important role in describing the continuity and additivity behavior of the capacity functions. We start by introducing the function $F_L(\mathfrak{W}) : \mathfrak{W} \to \mathbb{R}_+$ as follows:

$$F_L(\mathfrak{W}) = \inf_{\sigma\in\mathbf{CH}(\mathcal{X}^L,\mathcal{S})} \max_{x^{L+1}\in\mathcal{X}^{L+1}} \max_{\pi\in\Pi_{L+1}} \sum_{y\in\mathcal{Y}} \Big| \sum_{s\in\mathcal{S}} W(y|x_1,s)\sigma(s|x_2,\ldots,x_{L+1}) - \sum_{s\in\mathcal{S}} W(y|x_{\pi(1)},s)\sigma(s|x_{\pi(2)},\ldots,x_{\pi(L+1)})\Big|, \tag{5.97}$$

where $\mathbf{CH}(\mathcal{X}^L;\mathcal{S})$ denotes the set of all stochastic matrices $\mathcal{X}^L \to \mathcal{P}(\mathcal{S})$, while Π_{L+1} denotes the set of all permutations on $\{1,2,\ldots,L+1\}$. One can notice that $F_L(\mathfrak{W})$ is somehow related to the L-symmetrizability property of the AVC: $\mathfrak{W}$. In fact, one can easily show that $\mathfrak{W}$ is L-symmetrizable if and only if $F_L(\mathfrak{W}) = 0$, otherwise $F_L(\mathfrak{W}) > 0$. To see this let us consider the case where $L = 1$, Eq. (5.97) simplifies to

$$F_1(\mathfrak{W}) = \min_{\sigma\in\mathbf{CH}(\mathcal{X},\mathcal{S})} \left[\max_{(x,\tilde{x})\in\mathcal{X}^2} \sum_{y\in\mathcal{Y}} \left| \sum_{s\in\mathcal{S}} \mathrm{W}(y|\tilde{x},s)\sigma(s|x) - \mathrm{W}(y|x,s)\sigma(s|\tilde{x}) \right| \right]. \tag{5.98}$$

Now by comparing the symmetrizability condition in (5.12) to Eq. (5.98), we can see that $F_1(\mathfrak{W}) = 0$ if and only if $\mathfrak{W}$ is symmetrizable. It is important to highlight that the function F_L does not only describe whether or not an AVC $\mathfrak{W}$ is symmetrizable, but it also quantifies how "far away" a non-symmetrizable AVC is from the extreme case of being symmetrizable.

In the next few lines, we will use the definition of $F_L(\mathfrak{W})$ given by Eq. (5.97) along with the concept of a distance measure between two AVCs and the principle of parallel AVCs introduced in Section 5.2.4 to introduce some basic lemmas. These lemmas play an important role in characterizing the continuity and additivity behaviour of the public list capacity of AVCs and the secrecy list capacity of AVWCs.

Lemma 5.8. *Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two finite AVCs. Then the following inequalities hold:*

$$F_L(\mathfrak{W}_2) \le 2G(\mathfrak{W}_1,\mathfrak{W}_2) + F_L(\mathfrak{W}_1) \tag{5.99}$$

$$F_L(\mathfrak{W}_1) \le 2G(\mathfrak{W}_2,\mathfrak{W}_1) + F_L(\mathfrak{W}_2) \tag{5.100}$$

$$\big|F_L(\mathfrak{W}_1 - F_L(\mathfrak{W}_2)\big| \le 2D(\mathfrak{W}_1,\mathfrak{W}_2) \tag{5.101}$$

One can easily notice that it is enough to prove the first inequality only as the second inequality follows by interchanging the indices of the first inequality, while the third inequality follows by combining the first and the second one. The detailed proof of the first inequality was given in [199, 200] by exploiting the definitions of the $\mathbb{LHS}$ and the $\mathbb{RHS}$ of the inequality.

Lemma 5.9. *[200] Let $\tilde{\mathfrak{W}}$ be an arbitrary finite AVC and let $\{\mathfrak{W}_n\}_{n=1}^{\infty}$ be a sequence of finite AVCs.*

$$\textit{If} \lim_{n\to\infty} D(\mathfrak{W}_n, \tilde{\mathfrak{W}}) = 0, \ \textit{then} \lim_{n\to\infty} F_L(\mathfrak{W}_n) = F_L(\tilde{\mathfrak{W}}).$$

The proof follows immediately from Lemma 5.8. It is important to highlight that Lemma 5.9 implies that $F_L(\mathfrak{W})$ is continuous function in $\mathfrak{W}$ as long as the state set $\mathcal{S}$ is finite. We need to highlight one last result related to the concept of parallel AVCs.

Lemma 5.10. *[200] Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two AVCs and $\mathfrak{W} = \mathfrak{W}_1 \otimes \mathfrak{W}_2$ be their parallel combination. Then the following chain of inequalities holds:*

$$\max\{F_L(\mathfrak{W}_1); F_L(\mathfrak{W}_2)\} \leq F_L(\mathfrak{W}) \leq F_L(\mathfrak{W}_1) + F_L(\mathfrak{W}_2). \tag{5.102}$$

We are now ready to discuss the continuity behaviour of the public and secrecy capacity functions of AVCs and AVWCs under list decoding, Theorem 5.8 and Theorem 5.9 show that the public list capacity of an AVC $\mathfrak{W}$ and the list secrecy capacity of an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ depend on the L-symmetrizability property of the AVC $\mathfrak{W}$ between the transmitter and the legitimate receiver. One can notice from Definition 5.12 that the symmetrizability of any AVC depends on the channel state set $\mathcal{S}$. This implies that changing the channel state set $\mathcal{S}$ can turn a non-symmetrizable AVC into a symmetrizable one, which will consequently affects the list capacity. However, the crucial question is whether small variations in the channel state set $\mathcal{S}$ can cause such effect or only major changes can. The answer to this question determines the continuity behaviour as follows: If small variations in the channel state set $\mathcal{S}$ cannot turn a non-symmetrizable AVC into a symmetrizable one, then the list capacity is a continuous function in the AVC $\mathfrak{W}$, and vise verse.

From a mathematical point of view, Definition 5.8 implies that the public list capacity of an AVC $C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}, L)$ is discontinuous in $\mathfrak{W}$ if and only if there exists a sequence $\{\mathfrak{W}_n\}_{n=1}^{\infty}$ of finite AVCs satisfying (5.43), while simultaneously the following condition holds:

$$\lim_{n\to\infty} \sup C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}_n, L) > \lim_{n\to\infty} \inf C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}_n, L). \tag{5.103}$$

Using the same formulation, we can define the discontinuity condition for the list secrecy capacity of an AVWC $C^{\mathrm{s}}_{\mathrm{list}}(\mathfrak{Q}, L)$. In [199, 200], a complete characterization of the discontinuity points of the public list capacity of an AVC was established as follows:

Theorem 5.11. *The public list capacity of the AVC with respect to average decoding error probability constraint $C^{p}_{list}(\mathfrak{W}, L)$ is discontinuous in the finite AVC $\mathfrak{W}$ if and only if the following conditions hold:*

1. $C^{p}_{ran}(\mathfrak{W}) > 0$.
2. $F_L(\mathfrak{W}) = 0$ *and for every* $\epsilon > 0$ *there exists a finite AVC* $\tilde{\mathfrak{W}}$ *with* $D(\mathfrak{W}, \tilde{\mathfrak{W}}) < \epsilon$ *and* $F_L(\tilde{\mathfrak{W}}) > 0$.

The previous theorem interestingly characterizes the discontinuity behavior of the public list capacity in terms of two continuous functions: the correlated random public capacity $C^{\mathrm{p}}_{\mathrm{ran}}(\mathfrak{W})$, cf. Corollary 5.1 and the function $F_L(\mathfrak{W})$. A discontinuity point occurs if $\mathfrak{W}$ is L-symmetrizable, while $C^{\mathrm{p}}_{\mathrm{ran}}(\mathfrak{W})$ is greater than zero. In addition, there must be another AVC $\tilde{\mathfrak{W}}$ which is not L-symmetrizable, such that the distance between $\mathfrak{W}$ and $\tilde{\mathfrak{W}}$ is small. These two conditions implies that although $\mathfrak{W}$ and $\tilde{\mathfrak{W}}$ are close to each other, $C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}, L) = 0$ while $C^{\mathrm{p}}_{\mathrm{list}}(\tilde{\mathfrak{W}}, L) > 0$. That is why $\mathfrak{W}$ in that case is a discontinuity point. Next, we present a result that establishes a certain robustness property for the public list capacity as follows:

Theorem 5.12. *[200] Let $\mathfrak{W}$ be a finite AVC with $F_L(\mathfrak{W}) > 0$. Then there exists an $\epsilon > 0$ such that all finite AVCs $\tilde{\mathfrak{W}}$ with $D(\tilde{\mathfrak{W}}, \mathfrak{W}) < \epsilon$ are continuity points of $C^{p}_{list}(\mathfrak{W}, L)$.*

The previous result implies that as long as the AVC $\mathfrak{W}$ is not L-symmetrizable, then the public list capacity is continuous within a certain neighborhood. This result follows due to the continuity of correlated random public capacity $C^{\mathrm{p}}_{\mathrm{ran}}(\mathfrak{W})$ and the function $F_L(\mathfrak{W})$. Although Theorem 5.11 identifies the necessary conditions at which a discontinuity point occurs, Theorem 5.12 indicates the continuity of the public list capacity in a certain neighborhood. These two theorems raise an important question about the existence of an AVC that satisfies the discontinuity conditions in Theorem 5.11. In [200], the authors managed to show that for AVCs where the cardinality of the input and output alphabets is greater than or equal 4, i.e. $|\mathcal{X}|$, $|\mathcal{Y}| \geq 4$, there exists AVCs that satisfy the discontinuity conditions in Theorem 5.11. This implies that the set of discontinuity points of $C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}, L)$ is a non-empty set.

Corollary 5.3. *[166] For an AVC: $\mathfrak{W}$ the deterministic public capacity with respect to the average decoding error probability $C^{p}_{det}(\mathfrak{W})$ is discontinuous in the finite AVC $\mathfrak{W}$ if and only if the following conditions hold:*

1. *$C^{p}_{ran}(\mathfrak{W}) > 0$.*
2. *$F_1(\mathfrak{W}) = 0$ and for every $\epsilon > 0$ there exists a finite AVC $\tilde{\mathfrak{W}}$ with $D(\mathfrak{W}, \tilde{\mathfrak{W}}) < \epsilon$ and $F_1(\tilde{\mathfrak{W}}) > 0$.*

The previous result follows due to the relation between public list codes and public deterministic codes: A public deterministic code $\mathcal{C}^{\mathrm{p}}_{\mathrm{det}}$ is in fact a public list code $\mathcal{C}^{\mathrm{p}}_{\mathrm{list}}$ with list size $L = 1$. This implies that Theorem 5.11 generalizes the result established in Corollary 5.3. Theorem 5.11 also has a direct consequence on the continuity of the list secrecy capacity of AVWCs as follows:

Corollary 5.4. *For an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, the list secrecy capacity $C^{s}_{list}(\mathfrak{Q}, L)$ with respect to the average decoding error probability has the following characteristics:*

1. *The function $C^{s}_{list}(\mathfrak{Q}, L)$ is discontinuous at the point $(\mathfrak{W}, \mathfrak{V})$ if and only if the following holds: First $C^{s}_{ran}(\mathfrak{Q}) > 0$ and second $F_L(\mathfrak{W}) = 0$ such that for all $\epsilon > 0$ there is $\tilde{\mathfrak{W}}$ where $D(\mathfrak{W}, \tilde{\mathfrak{W}}) < \epsilon$ and $F_L(\tilde{\mathfrak{W}}) > 0$.*
2. *If $C^{s}_{list}(\mathfrak{Q}, L)$ is discontinuous at the point $(\mathfrak{W}, \mathfrak{V})$, then it is discontinuous at all points $(\mathfrak{W}, \tilde{\mathfrak{V}})$ for which $C^{s}_{ran}(\mathfrak{Q}) > 0$.*
3. *For every $\mathfrak{W}$, $C^{s}_{list}(\mathfrak{Q}, L)$ is continuous in the AVC $\mathfrak{V}$ between the transmitter and the eavesdropper.*

The previous result implies that the discontinuity in the list secrecy capacity only originates form the legitimate link $\mathfrak{W}$, while changing the eavesdropper link $\mathfrak{V}$ does not affect the discontinuity behavior. The only role played by the eavesdropper link $\mathfrak{V}$ in determining the discontinuity point is that the correlated random secrecy capacity at such $\mathfrak{V}$ has to be greater than zero. Corollary 5.4 indirectly indicates that the establishment of a stable reliable communication over an AVC is much more challenging than the assurance of the secrecy of such communication over an AVWC.

We now turn to the investigation of the additivity property of the public and secrecy capacity functions for AVCs and AVWCs under list decoding. The additivity property mainly allows us to decide whether to use a simple individual encoder-decoder pair for each parallel AVC or to use a joint encoder-decoder pair for the whole system. This is because if the capacity function under list decoding is super-additive, joint encoding and decoding will lead to a higher capacity results. We now present the main results that describes how the public list capacity of parallel AVCs behaves. We start by the following super-additivity result:

Theorem 5.13. *[199] Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two parallel AVCs, such that $\mathfrak{W} = \mathfrak{W}_1 \otimes \mathfrak{W}_2$. Then*

$$C^p_{list}(\mathfrak{W}, L) > C^p_{list}(\mathfrak{W}_1, L) + C^p_{list}(\mathfrak{W}_2, L)$$

if and only if

$$\min\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\} = 0, \tag{5.104a}$$

$$\max\{F_L(\mathfrak{W}_1), F_L(\mathfrak{W}_2)\} > 0, \tag{5.104b}$$

$$\min\{C^p_{list}(\mathfrak{W}_1), C^p_{list}(\mathfrak{W}_2)\} > 0. \tag{5.104c}$$

The previous theorem defines the necessary and sufficient conditions for the super-additivity of the public list capacity of parallel AVCs. The conditions in (5.104a) and (5.104b) are related to the characteristics of the public list capacity of the two parallel AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$ as follows: One of the two AVCs must be L-symmetrizable, while the other one is not. On the other hand, the condition in (5.104c) implies that correlated random public capacity of both channels is greater than zero. It is important to point out that the additivity result of the correlated random public capacity established in Theorem 5.7 plays an important role in the proof of Theorem 5.13.

Corollary 5.5. *[200] Let $\mathfrak{W}_1$ and $\mathfrak{W}_2$ be two parallel AVCs, such that $\mathfrak{W} = \mathfrak{W}_1 \otimes \mathfrak{W}_2$. Then*

$$C^p_{list}(\mathfrak{W}, L) = 0, \tag{5.105}$$

if and only if

$$C^p_{list}(\mathfrak{W}_1, L) = C^p_{list}(\mathfrak{W}_2, L) = 0. \tag{5.106}$$

The previous corollary shows that super-activation is not possible for parallel AVCs under list decoding. In fact, this result follows as a consequence of Theorem 5.13, because in order for super-activation to happen, the public list capacity needs to be super-additive and at the same time the public list capacity of the two AVCs $\mathfrak{W}_1$ and $\mathfrak{W}_2$ must be zero. However, if $C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}_1, L) = C^{\mathrm{p}}_{\mathrm{list}}(\mathfrak{W}_2, L) = 0$, then either the condition in (5.104b) or the one in (5.104c) is violated, which directly implies that the public list capacity of $\mathfrak{W}_1 \otimes \mathfrak{W}_2$ is not super-additive.

Theorem 5.13 implies that public list capacity of parallel AVCs is super-additive, yet it remains to investigate by how much is the additivity violated. This is similar to the question raised by Alon in [178] for the zero-error capacity. To answer this question, the normalized public list capacity of parallel AVCs that satisfies the supper-additivity constraints given by Theorem 5.13 was investigated. It was shown in [200], that the normalized public list capacity approaches zero as the list size $L \to \infty$. This implied that the additivity of the public list capacity is not violated in a strong form.

Next, we extend our investigation to the secrecy domain and aim to characterize the additivity behavior of the list secrecy capacity for parallel AVWCs. The problem with such investigation is that we do not know whether the correlated random secrecy capacity of parallel AVWCs is additive or super-additive. As highlighted in the proof of Theorem 5.13 that such result is usually crucial and without it, one can find it really difficult to characterize the behaviour under list decoding. Nevertheless, we were able to establish the following very interesting result:

Theorem 5.14. *Let $\mathfrak{Q}_1$ and $\mathfrak{Q}_2$ be two parallel AVWCs and let $\mathfrak{Q}$ be their parallel combination, then the following holds:*

1. *If $C^s_{list}(\mathfrak{Q}_1, L) = C^s_{list}(\mathfrak{Q}_2, L) = 0$, then $C^s_{list}(\mathfrak{Q}, L) > 0$ is true if and only if $\mathfrak{W}$ is not L-symmetrizable and $C^s_{ran}(\mathfrak{Q}) > 0$.*
2. *If $C^s_{ran}(\mathfrak{Q})$ shows super-activation, then $C^s_{list}(\mathfrak{Q}, L)$ shows super-activation as well, if and only if $\mathfrak{W}$ is not L-symmetrizable.*
3. *If $C^s_{ran}(\mathfrak{Q})$ shows no super-activation, then super-activation of $C^s_{list}(\mathfrak{Q}, L)$ can only happen if $\mathfrak{W}_1$ is not L-symmetrizable and $\mathfrak{W}_2$ is L-symmetrizable, in addition $C^s_{ran}(\mathfrak{Q}_1) = 0$, while $C^s_{ran}(\mathfrak{Q}_2) > 0$. The statement is independent of the specific labeling.*

The previous theorem is a generalization of the result established in [166, Theorem 5] for the deterministic secrecy capacity of parallel AVWCs $C^{\mathrm{s}}_{\mathrm{det}}(\mathfrak{Q})$. It was shown in [166] that there exists AVWCs that exhibit the behavior in the first statement of Theorem 5.14. In [200], an example for AVWCs that satisfy the third statement was constructed. This result is of a great impact as it shows that super-activation is not a unique feature for the field of quantum information theory. Moreover, Corollary 5.5 and Theorem 5.14 implies that super-activation is still a unique feature of secure communication over parallel AVWCs and is not possible for public communication. It is important to point out that the results established in this section are only valid for finite AVCs and finite AVWCs, i.e. $|\mathcal{S}| < \infty$. The extension of these results to the situation where $\mathcal{S}$ is infinite extremely difficult, specially because the characterization of the list capacity for infinite AVCs is still an open problem.

5.4 Public-Confidential Capacity Regions for AVWCs

In this section, we consider a communication scenario that involves the integration of public and confidential communication services simultaneously over an AVWC. In the beginning, we introduce our communication model along with an extension of the deterministic and correlated random coding techniques to the problem at hand. We then present a full characterization of the public-confidential capacity region for AVWCs under both coding techniques. The result established in this section were published in [201].

5.4.1 Motivation and Basic Definitions

Most of the literature that addressed the problem of secure communication over AVWCs considered the following communication scenario: A confidential message $m_c \in \mathcal{M}_c$ is sent over an AVWC, such that the legitimate receiver can reliably decode the message regardless of the jammer attack s^n while keeping the eavesdropper completely ignorant about the transmitted message. Real life communication scenarios usually involve more complex setups like a combination of public and confidential services or multi-user services as discussed in Chapter 3. To the best of our knowledge, the investigation of a public-confidential communication setup over a wiretap channel with imperfect CSI was only done in [202] for the compound wiretap channel and was never investigated for AVWCs.

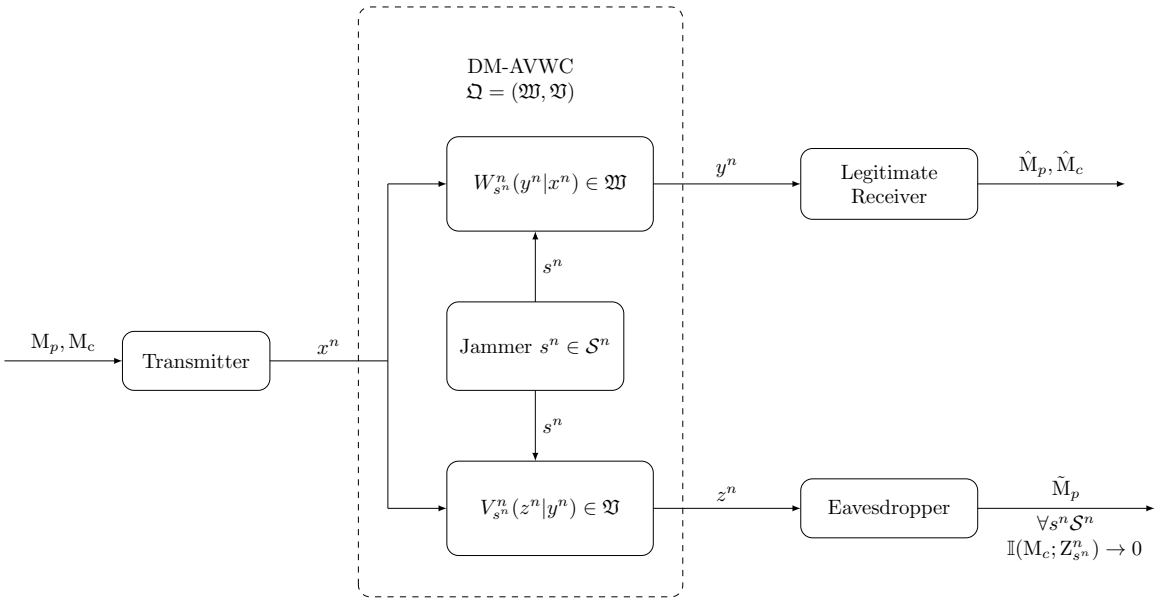


Figure 5.2: Public-Confidential communication scenario over an AVWC

Consider the communication scenario illustrated in Fig. 5.2, where a common public message $m_p \in \mathcal{M}_p$ and a confidential message $m_c \in \mathcal{M}_c$ are transmitted over an AVWC $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$. The public message needs to be reliably decoded by both the legitimate receiver and the eavesdropper. On the other hand, the confidential message needs to be only decoded by the legitimate receiver, while keeping the eavesdropper completely ignorant about it. These two requirements need to be fulfilled for all possible channel state sequences $s^n \in \mathcal{S}^n$ produced by the jammer. We will investigate this communication scenario using two coding schemes: deterministic codes and correlated random codes as follows:

Definition 5.17. A public-confidential deterministic code $\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}$ for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ consists of: a set of public messages $\mathcal{M}_p$, a set of confidential messages $\mathcal{M}_c$, a stochastic encoder:

$$E : \mathcal{M}_p \times \mathcal{M}_c \to \mathcal{P}(\mathcal{X}^n),$$

that maps a public-confidential message pair $(m_p, m_c) \in \mathcal{M}_p \times \mathcal{M}_c$ to a codeword $x^n(m_p, m_c) \in \mathcal{X}^n$ according to the conditional probability $E(x^n|m_p, m_c)$ and two deterministic decoders:

$$\varphi_1 : \mathcal{Y}^n \to \mathcal{M}_p \times \mathcal{M}_c \cup \{?\}$$
$$\varphi_2 : \mathcal{Z}^n \to \mathcal{M}_p \cup \{?\},$$

where each decoder maps the channel observation at its respective receiving node to the corresponding required messages or an error message.

In order to measure the reliability performance of $\mathcal{C}_{\text{det}}^{\text{pc}}$, we define the following average decoding error probabilities for $s^n \in \mathcal{S}^n$:

$$\bar{\mathbf{P}}_{e1}(s^n|\mathcal{C}_{\text{det}}^{pc}) = \frac{1}{|\mathcal{M}_p||\mathcal{M}_c|} \sum_{m_p,m_c} \sum_{x^n} \sum_{y^n:\varphi_1(y^n)\neq(m_p,m_c)} W_{s^n}^n(y^n|x^n)E(x^n|m_p,m_c) \quad (5.107)$$

$$\bar{\mathbf{P}}_{e2}(s^n|\mathcal{C}_{\text{det}}^{pc}) = \frac{1}{|\mathcal{M}_p|} \sum_{m_p} \sum_{x^n} \sum_{z^n:\varphi_2(z^n)\neq m_p} V_{s^n}^n(z^n|x^n)E(x^n|m_p,\cdot). \quad (5.108)$$

On the other hand, the secrecy performance of $\mathcal{C}_{\text{det}}^{\text{pc}}$ is evaluated with respect to the strong secrecy criterion which considers the information leakage of the confidential message to the eavesdropper for every channel state s^n as follows:

$$\mathbb{L}(\mathcal{C}_{\text{det}}^{\text{pc}}) = \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}_{s^n}^n|\mathcal{C}_{\text{det}}^{\text{pc}}), \quad (5.109)$$

where M_c represents a uniformly distributed random variable over the message set $\mathcal{M}_c$, while $\mathrm{Z}_{s^n}^n$ is the channel output at the eavesdropper for a state sequence s^n.

Definition 5.18. A rate pair $(R_p, R_c) \in \mathbb{R}_+^2$ is an achievable deterministic public-confidential rate pair for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\delta, \lambda, \tau)$, there exists a sequence of deterministic codes $(\mathcal{C}_{\text{det}}^{\text{pc}})_n$ that satisfies the following constraints:

$$\frac{1}{n}\log|\mathcal{M}_p| \geq R_p - \delta, \qquad \frac{1}{n}\log|\mathcal{M}_c| \geq R_c - \delta, \quad (5.110)$$

$$\max_{s^n\in\mathcal{S}^n} \bar{\mathbf{P}}_{e1}(s^n|\mathcal{C}_{\text{det}}^{\text{pc}}) + \max_{s^n\in\mathcal{S}^n} \bar{\mathbf{P}}_{e2}(s^n|\mathcal{C}_{\text{det}}^{\text{pc}}) \leq \lambda, \quad (5.111)$$

$$\mathbb{L}(\mathcal{C}_{\text{det}}^{\text{pc}}) \leq \tau. \quad (5.112)$$

The public-confidential deterministic strong secrecy capacity region $C_{\text{det}}^{\text{pc}}(\mathfrak{Q})$ is given by the set of all achievable rate pairs (R_p, R_c).

Definition 5.19. A public-confidential correlated random code $\mathcal{C}_{\text{ran}}^{\text{pc}}$ for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ is given by a family of public-confidential deterministic codes of Definition 5.17 as follows:

$$\mathcal{C}_{\text{ran}}^{\text{pc}} = \{\mathcal{C}_{\text{det}}^{\text{pc}}(\gamma) : \gamma \in \mathcal{G}\}$$

together with a random variable Γ that takes values $\gamma \in \mathcal{G}$ according to $P_\Gamma(\gamma)$.

In order to measure the reliability performance of $\mathcal{C}_{\text{ran}}^{\text{pc}}$, we build on the average decoding errors of a deterministic code in (5.107) and (5.108). Thus, we have

$$\bar{\mathbf{P}}_{ei}(s^n|\mathcal{C}_{\text{ran}}^{\text{pc}}) = \sum_{\gamma} \bar{\mathbf{P}}_{ei}\big(s^n|\mathcal{C}_{\text{det}}^{\text{pc}}(\gamma)\big)P_\Gamma(\gamma), \qquad i \in \{1,2\}. \quad (5.113)$$

On the other hand, the secrecy performance of $\mathcal{C}_{\text{ran}}^{\text{pc}}$ is evaluated with respect to the maximum strong secrecy criterion which considers the maximum information leakage

of the confidential message to the eavesdropper for every channel state s^n and every realization of the common randomness γ as follows:

$$\mathbb{L}^{\max}(\mathcal{C}_{\text{ran}}^{\text{pc}}) = \max_{s^n \in \mathcal{S}^n} \max_{\gamma} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}_{s^n}^n(\gamma) | \mathcal{C}_{\text{ran}}^{\text{pc}}), \tag{5.114}$$

where $\mathrm{Z}_{s^n}^n(\gamma)$ is the channel output at the eavesdropper for a state sequence s^n, when the realization of the common randomness is $\Gamma = \gamma$.

Definition 5.20. A rate pair $(R_p, R_c) \in \mathbb{R}_+^2$ is an achievable public-confidential correlated random rate pair for the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\delta, \lambda, \tau)$, there exists a sequence of correlated random codes $(\mathcal{C}_{\text{ran}}^{\text{pc}})_n$ that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_p| \geq R_p - \delta, \qquad \frac{1}{n} \log |\mathcal{M}_c| \geq R_c - \delta, \tag{5.115}$$

$$\max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_{e1}(s^n | \mathcal{C}_{\text{ran}}^{\text{pc}}) + \max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_{e2}(s^n | \mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \lambda, \tag{5.116}$$

$$\mathbb{L}^{\max}(\mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \tau. \tag{5.117}$$

The public-confidential correlated random maximum strong secrecy capacity region $C_{\text{ran}}^{\text{pc}}(\mathfrak{Q})$ is given by the set of all achievable rate pairs (R_p, R_c).

5.4.2 An Auxiliary Channel Model

In order to investigate the capacity regions acknowledged by Definitions 5.18 and 5.20, we need to introduce an auxiliary channel model that will facilitate our investigation. Consider a channel model that consists of one transmitting node T_X and three receiving nodes $R_{X_1}, R_{X_2}, R_{X_3}$ such that: the channels from T_X to R_{X_1} and from T_X to R_{X_2} are compound channels, while the channel from T_X to R_{X_3} is an AVC. For n channel uses, two finite channel state sets $\mathcal{R}$ and $\mathcal{S}$, finite input alphabet $\mathcal{X}$ and two finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, our discrete memoryless channel model is given by the following transition matrices:

$$W_r^n(y^n|x^n) = \prod_{i=1}^n W_r(y_i|x_i), \qquad V_r^n(z^n|x^n) = \prod_{i=1}^n V_r(z_i|x_i), \tag{5.118}$$

$$V_{s^n}^n(z^n|x^n) = \prod_{i=1}^n V_{s_i}(z_i|x_i). \tag{5.119}$$

The previous equation indicates that the transition matrix $V_r^n(z^n|x^n)$ describes a compound channel from the input alphabet $\mathcal{X}$ to the output alphabet $\mathcal{Z}$. This is because the channel state r remains constant for all the n uses of the channel. On the other hand, the transition matrix $\mathrm{V}_{s^n}^n(z^n|x^n)$ describe an AVC where the channel state s varies for every channel use. We will call this model the compound broadcast channel with an arbitrarily varying wiretapper (CBAVWC) and denote it as follows:

$$(\mathcal{W}, \mathcal{V}, \mathfrak{V}) = \big\{(W_r^n, V_r^n, V_{s^n}^n) : r \in \mathcal{R}, s^n \in \mathcal{S}^n, n = 1, 2, \dots\big\}, \tag{5.120}$$

Now consider a communication scenario in which a common public message $m_p \in \mathcal{M}_p$ and a confidential message $m_c \in \mathcal{M}_c$ are transmitted over the CBAVWC $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ such that: The public message needs to be reliably decoded by nodes R_{X_1} and R_{X_2},

i.e. the compound components (5.118) of the CBAVWC. On the other hand, the confidential message only needs to be decoded by node R_{X_1}, while keeping node R_{X_3} - AV component (5.119) - completely ignorant about it. Of course, these two requirements need to be fulfilled for all possible compound states $r \in \mathcal{R}$ and AV state sequences $s^n \in \mathcal{S}^n$.

In order to achieve this task, we extend Definition 5.17 to define a public-confidential deterministic code $\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}$ for our CBAVWC. In order to measure the reliability performance of $\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}$, we modify the average decoding error probabilities in (5.107) and (5.108) such that $W_{s^n}^n$ and $V_{s^n}^n$ are replaced by their corresponding compound channels W_r^n and V_r^n. Moreover, the decoding error probabilities are defined with respect to the states $r \in \mathcal{R}$ instead of $s^n \in \mathcal{S}^n$. On the other hand, the secrecy performance of $\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}$ is evaluated with respect to the strong secrecy criterion as in (5.109).

Definition 5.21. A rate pair $(R_p, R_c) \in \mathbb{R}_+^2$ is an achievable public-confidential rate pair for the CBAVWC: $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\delta, \lambda, \tau)$, there exists a sequence of deterministic codes $(\overline{\mathcal{C}_{\text{det}}^{\text{pc}}})_n$ that satisfies the following constraints:

$$\frac{1}{n}\log|\mathcal{M}_p| \geq R_p - \delta, \qquad \frac{1}{n}\log|\mathcal{M}_c| \geq R_c - \delta, \tag{5.121}$$

$$\max_{r\in\mathcal{R}} \bar{\mathbf{P}}_{e1}(r|\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}) + \max_{r\in\mathcal{R}} \bar{\mathbf{P}}_{e2}(r|\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}) \leq \lambda, \tag{5.122}$$

$$\mathbb{L}(\overline{\mathcal{C}_{\text{det}}^{\text{pc}}}) \leq \tau. \tag{5.123}$$

The public-confidential deterministic strong secrecy capacity region $C_{\text{det}}^{\text{pc}}(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ is given by the set of all achievable rate pairs (R_p, R_c).

We introduced the model of CBAVWC because it serves as an intermediate step in the establishment of the public-confidential correlated random secrecy capacity of AVWCs. This technique was used in [159], where an auxiliary channel called the compound arbitrary varying wiretap channel was introduced to facilitate the establishment of the correlated random secrecy capacity of an AVWC.

Before, we present our coding theorem, we need to define the following multi-letter expressions. For an arbitrary but fixed $n \in \mathbb{N}$, a CBAVWC $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ and two random variables (U, V) that can take values in the finite alphabets $(\mathcal{U}, \mathcal{V})$ such that $\mathrm{P}_{\mathrm{UVX}^n}$ denotes a joint probability distribution over the random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X}^n)$, we let $\mathcal{R}_n^{\mathrm{P}}(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ be a public-confidential rate region that contains the set of rate pairs $(R_p, R_c) \in \mathbb{R}_+^2$ that satisfy the following constraints:

$$R_p \leq \frac{1}{n}\min_{r\in\mathcal{R}}\min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_r^n), \mathbb{I}(\mathrm{U};\mathrm{Z}_r^n)\Big] \tag{5.124a}$$

$$R_c \leq \frac{1}{n}\Big[\min_{r\in\mathcal{R}}\mathbb{I}(\mathrm{V};\mathrm{Y}_r^n|\mathrm{U}) - \max_{s^n\in\mathcal{S}^n}\mathbb{I}(\mathrm{V};\mathrm{Z}_{s^n}^n|\mathrm{U})\Big]. \tag{5.124b}$$

Further, we use $\mathcal{R}_n^{\mathrm{P}}(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ to define two additional rate regions for the CBAVWC $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ as follows:

$$\mathcal{R}_n^*(\mathcal{W}, \mathcal{V}, \mathfrak{V}) = \bigcup_{\mathrm{P}_{\mathrm{UVX}^n}} \mathcal{R}_n^{\mathrm{P}}(\mathcal{W}, \mathcal{V}, \mathfrak{V}), \tag{5.125}$$

$$\mathcal{R}^*(\mathcal{W},\mathcal{V},\mathfrak{V}) = \bigcup_{n=1}^{\infty} \mathcal{R}_n^*(\mathcal{W},\mathcal{V},\mathfrak{V}). \tag{5.126}$$

We are now ready to present our coding theorem for the deterministic public-confidential capacity region of our auxiliary channel $(\mathcal{W},\mathcal{V},\mathfrak{V})$ as follows:

Theorem 5.15. *For the CBAVWC $(\mathcal{W},\mathcal{V},\mathfrak{V})$, we have*

$$C_{det}^{pc}(\mathcal{W},\mathcal{V},\mathfrak{V}) = \mathcal{R}^*(\mathcal{W},\mathcal{V},\mathfrak{V}).$$

The achievability of the previous theorem is based on the results established in [202] for the compound broadcast channel with common and confidential message along with the strong secrecy results established for AVWC in [159, 166]. In order to prove that $\mathcal{R}^*(\mathcal{W},\mathcal{V},\mathfrak{V})$ is achievable, we need to show that the rate region $\mathcal{R}_{\tilde{n}}^{\tilde{\mathrm{P}}}(\mathcal{W},\mathcal{V},\mathfrak{V})$ is achievable for any $\tilde{n} \in \mathbb{N}$ and any distribution $\tilde{\mathrm{P}}_{\mathrm{UVX}^{\tilde{n}}}$. Based on standard channel prefixing arguments, cf. [159, Appendix D] along with the concept of block coding cf. [166, Section IV-F], it is enough to show that:

$$R_0 \leq \min_{r\in\mathcal{R}} \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_r), \mathbb{I}(\mathrm{U};\mathrm{Z}_r)\Big] \tag{5.127a}$$

$$R_1 \leq \min_{r\in\mathcal{R}} \mathbb{I}(\mathrm{X};\mathrm{Y}_r|\mathrm{U}) - \max_{q\in\mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X};\mathrm{Z}_q|\mathrm{U}) \tag{5.127b}$$

is achievable. In order to do so, we present the following coding scheme:

1) Codebook Generation: For a fixed block length t and a given probability distribution $\mathrm{P}_{\mathrm{UX}} \in \mathcal{P}(\mathcal{U}\times\mathcal{X})$, we define a probability measure $\hat{\mathrm{P}}_{\mathrm{U}^t\mathrm{X}^t} \in \mathcal{P}(\mathcal{U}^t\times\mathcal{X}^t)$ as follows:

$$\hat{\mathrm{P}}_{\mathrm{U}^t\mathrm{X}^t}(u^t,x^t) \triangleq \frac{\mathrm{P}_{\mathrm{UX}}^t(u^t,x^t)}{\mathrm{P}_{\mathrm{UX}}^t\big(\mathcal{T}_\epsilon^t(\mathrm{P}_{\mathrm{UX}})\big)} \tag{5.128}$$

if $(u^t,x^t) \in \mathcal{T}_\epsilon^t(\mathrm{P}_{\mathrm{UX}})$ and $\hat{\mathrm{P}}_{\mathrm{U}^t\mathrm{X}^t}(u^t,x^t) = 0$ otherwise, for some typicality constant $\epsilon > 0$. Next, we generate the codewords $u^t(m_p)$ for $m_p \in \mathcal{M}_p$ by using $\hat{\mathrm{P}}_{\mathrm{U}^t}$. Then, for every $u^t(m_p)$, we generate the codewords $x^t(m_p,m_c,m_k)$ for $m_c \in \mathcal{M}_p$ and $m_k \in \mathcal{M}_k$ according to $\hat{\mathrm{P}}_{\mathrm{X}^t|\mathrm{U}^t}$, where $\mathcal{M}_k$ is a randomization set used for the realization of our stochastic encoder.

2) Encoding and Decoding: Given a message pair (m_p,m_c), the encoder E choose a message m_k uniformly at random from the set $\mathcal{M}_k$ and transmit the corresponding codeword $x^t(m_p,m_c,m_k)$. At the legitimate receiver, the decoder φ_1 searches for the unique message triple $(\ddot{m}_p,\ddot{m}_c,\ddot{m}_k)$ such that the received sequence y_r^n is jointly typical with the codewords $u^t(\ddot{m}_p)$ and $x^t(\ddot{m}_p,\ddot{m}_c,\ddot{m}_k)$. Similarly, at the eavesdropper, the decoder φ_2 searches for the unique message $\check{m}_p$ such that the received sequence z_r^t is jointly typical with the codeword $u^t(\check{m}_p)$.

3) Reliability and Secrecy Analysis: Based on the results established in [202, Lemma 1] along with the standard joint-typicality decoding arguments in [34,35], we can show that the average decoding error probability of our coding scheme decays exponentially in t as long as:

$$R_p \leq \min_{r\in\mathcal{R}} \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_r), \mathbb{I}(\mathrm{U};\mathrm{Z}_r)\Big] \tag{5.129a}$$

$$R_c + R_k \leq \min_{r \in \mathcal{R}} \mathbb{I}(\mathrm{X}; \mathrm{Y}_r | \mathrm{U}). \tag{5.129b}$$

On the other hand, by adapting the results established in [166, Lemma 2] and [159, Corollary 2] along with the strong secrecy techniques discussed in Section 2.2.3, keeping in mind that the codewords X^t were generated conditioned on the random variable U^t, we can show that $\lim_{t\to\infty} \max_{s^t \in \mathcal{S}^t} \mathbb{I}(\mathrm{M}_1; \mathrm{Z}^t_{s^t}) = 0$ as long as:

$$R_k > \max_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{X}; \mathrm{Z}_q | \mathrm{U}). \tag{5.130}$$

By combining the rate constraints in (5.129) and (5.130), it directly follows that the rate region in (5.127) is achievable for the CBAVWC $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$. This concludes our achievability proof for Theorem 5.15. On the other hand, the converse proof follows easily by combining the techniques used in [202] along with the concepts used in Section 5.3.2.

The establishment of the capacity region in Theorem 5.15 is only the first step in our analysis. Next, we need to consider a stronger notion for the secrecy capacity of the CBAVWC known as the permutation invariant secrecy capacity defined as follows: Given a public-confidential deterministic code $\overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}$ and a permutation $\pi \in \Pi_n$ where Π_n is the set of all permutations over $\{1, 2, \ldots, n\}$, we define the set of permuted codes $\overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)$, where each code in this set has a stochastic encoder E^π and two deterministic decoders $(\varphi^\pi_1, \varphi^\pi_2)$ given by:

$$\begin{aligned} E^\pi\big(x^n | m_p, m_c\big) &\triangleq E\big(\pi^{-1}(x^n) | m_p, m_c\big), \\ \varphi^\pi_1(y^n) &\triangleq \varphi_1\big(\pi^{-1}(y^n)\big), \\ \varphi^\pi_2(z^n) &\triangleq \varphi_2\big(\pi^{-1}(z^n)\big), \end{aligned}$$

where for a sequence a^n, $\pi(a^n) = \big(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(n)}\big)$. Accordingly, we present the following definition:

Definition 5.22. A rate pair $(R_p, R_c) \in \mathbb{R}^2_+$ is an achievable public-confidential permutation invariant rate pair for the CBAVWC: $(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ if for all $\delta, \lambda, \tau > 0$ there is an $n(\delta, \lambda, \tau) \in \mathbb{N}$, such that for all $n > n(\delta, \lambda, \tau)$, there exists a sequence of deterministic codes $(\overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}})_n$ that satisfies the following constraints:

$$\frac{1}{n} \log |\mathcal{M}_p| \geq R_p - \delta, \qquad \frac{1}{n} \log |\mathcal{M}_c| \geq R_c - \delta, \tag{5.131}$$

$$\max_{\pi \in \Pi_n} \Big[\max_{r \in \mathcal{R}} \bar{\mathbf{P}}_{e1}\big(r | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)\big) + \max_{r \in \mathcal{R}} \bar{\mathbf{P}}_{e2}\big(r | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)\big) \Big] \leq \lambda, \tag{5.132}$$

$$\max_{s^n \in \mathcal{S}^n} \max_{\pi \in \Pi_n} \mathbb{I}\big(\mathrm{M}_c; \mathrm{Z}^n_{s^n}(\pi) | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)\big) \leq \tau. \tag{5.133}$$

The public-confidential permutation invariant strong secrecy capacity region $C^{\mathrm{pc}}_{\mathrm{per}}(\mathcal{W}, \mathcal{V}, \mathfrak{V})$ is given by the set of all achievable permutation invariant rate pairs (R_p, R_c).

In [171], it was shown that the average decoding error probability is independent of the value of the permutation $\pi \in \Pi_n$. In particular it was shown that for any $\pi \in \Pi_n$:

$$\bar{\mathbf{P}}_{ei}\big(r | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)\big) = \bar{\mathbf{P}}_{ei}\big(r | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\mathrm{id})\big) = \bar{\mathbf{P}}_{ei}\big(r | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}\big), \tag{5.134}$$

for $i = 1, 2$ and (id) denotes the identity permutation. Moreover, it was shown in [159, Lemma 13] that if there exists an $\epsilon > 0$ such that $\max_{s^n} \mathbb{I}\big(\mathrm{M}_c; \mathrm{Z}^n_{s^n}(\mathrm{id}) | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\mathrm{id})\big) \leq \epsilon$, then it directly follows that:

$$\max_{\pi \in \Pi_n} \max_{s^n \in \mathcal{S}^n} \mathbb{I}\big(\mathrm{M}_c; \mathrm{Z}^n_{s^n}(\pi) | \overline{\mathcal{C}^{\mathrm{pc}}_{\mathrm{det}}}(\pi)\big) \leq \epsilon. \tag{5.135}$$

The two results in (5.134) and (5.135) directly imply that Theorem 5.15 does not only establish the public-confidential deterministic capacity region of our CBAVWC, but it also establishes the permutation invariant capacity region. In particular, we have

$$C^{\mathrm{pc}}_{\mathrm{per}}(\mathcal{W}, \mathcal{V}, \mathfrak{V}) = C^{\mathrm{pc}}_{\mathrm{det}}(\mathcal{W}, \mathcal{V}, \mathfrak{V}). \tag{5.136}$$

5.4.3 Coding Theorems: Achievability and Converse

In this section, we will use the result established in Theorem 5.15 to derive the public-confidential correlated random capacity regions of AVWCs by using Ahlswede's robustification technique [171]. We will then use the elimination of correlation approach [163] to derive the corresponding public-confidential deterministic capacity region. Before, we present our coding theorem, we need to define some multi-letter expressions similar to the ones in (5.124).

For an arbitrary but fixed $n \in \mathbb{N}$, an AVWC: $\mathfrak{Q} = (\mathcal{W}, \mathfrak{V})$ and two random variables (U, V) that can take values in the finite alphabets $(\mathcal{U}, \mathcal{V})$ such that $\mathrm{P}_{\mathrm{UVX}^n}$ denotes a joint probability distribution over the random variables $(\mathrm{U}, \mathrm{V}, \mathrm{X}^n)$, we let $\mathcal{R}^{\mathrm{P}}_n(\mathfrak{Q})$ be a public-confidential rate region that contains the set of rate pairs $(R_p, R_c) \in \mathbb{R}^2_+$ which satisfy the following constraints:

$$R_p \leq \frac{1}{n} \min_{q \in \mathcal{P}(\mathcal{S})} \min \Big[\mathbb{I}(\mathrm{U}; \mathrm{Y}^n_q), \mathbb{I}(\mathrm{U}; \mathrm{Z}^n_q)\Big] \tag{5.137a}$$

$$R_c \leq \frac{1}{n} \Big[\min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{V}; \mathrm{Y}^n_q | \mathrm{U}) - \max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{V}; \mathrm{Z}^n_{s^n} | \mathrm{U})\Big]. \tag{5.137b}$$

We then use $\mathcal{R}^{\mathrm{P}}_n(\mathfrak{Q})$ to define the two rate regions: $\mathcal{R}^*_n(\mathfrak{Q})$ and $\mathcal{R}^*(\mathfrak{Q})$ accordingly as in (5.125) and (5.126). Moreover, for an AVWC: $\mathfrak{Q}$ and the joint distribution $\mathrm{P}_{\mathrm{VX}^n}$, we let $\mathcal{R}^{\mathrm{P}}_{c,n}(\mathfrak{Q})$ be the set of confidential rates pairs $R_c \in \mathbb{R}_+$ that satisfy:

$$R_c \leq \frac{1}{n} \Big[\min_{q \in \mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{V}; \mathrm{Y}^n_q) - \max_{s^n \in \mathcal{S}^n} \mathbb{I}(\mathrm{V}; \mathrm{Z}^n_{s^n})\Big]. \tag{5.138}$$

We then use it to define $\mathcal{R}^*_{c,n}(\mathfrak{Q})$ and $\mathcal{R}^*_c(\mathfrak{Q})$ in the same manner as in (5.125) and (5.126). We now present the main result of this section as follows:

Theorem 5.16. *For the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, it follows that:*

$$C^{pc}_{cr}(\mathfrak{Q}) = \mathcal{R}^*(\mathfrak{Q}). \tag{5.139}$$

$$C^{pc}_{det}(\mathfrak{Q}) = \begin{cases} \mathcal{R}^*(\mathfrak{Q}) & \textit{if } \mathcal{A} \\ \big(0, \mathcal{R}^*_c(\mathfrak{Q})\big) & \textit{if } \mathcal{B} \\ (0, 0) & \textit{otherwise} \end{cases} \tag{5.140}$$

where the event $\mathcal{A}$ occurs when both $\mathfrak{W}$ and $\mathfrak{V}$ are not symmetrizable, while the event $\mathcal{B}$ occurs when $\mathfrak{W}$ is not symmetrizable but $\mathfrak{V}$ is symmetrizable.

For the achievability proof of the previous theorem, we need to recall the notation of average channel for an AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ given in (5.10) and use it to define the following families: $\overline{\mathfrak{W}} \triangleq \{\mathrm{W}_q^n : q \in \mathcal{P}(\mathcal{S}^n), n = 1, 2, \dots\}$ and $\overline{\mathfrak{V}} \triangleq \{\mathrm{V}_q^n : q \in \mathcal{P}(\mathcal{S}^n), n = 1, 2, \dots\}$. Next, we observe that for an arbitrary but fixed number of channel uses n, the AVWCs $\overline{\mathfrak{W}}$ and $\overline{\mathfrak{V}}$ simplify to their corresponding compound channels $\mathcal{W}$ and $\mathcal{V}$ as highlighted in Section 5.2.2. This implies that for a fixed n, $\overline{\mathfrak{W}}$, $\overline{\mathfrak{V}}$ and $\mathfrak{V}$ define a CBAVWC $(\overline{\mathfrak{W}}, \overline{\mathfrak{V}}, \mathfrak{V})$.

Based on Theorem 5.15, it follows that for a fixed $n \in \mathbb{N}$, $\lambda, \tau > 0$, we can construct a deterministic code $\overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}$ such that for every permutation $\pi \in \Pi_n$:

$$\max_{\pi \in \Pi_n} \Big[\max_{q \in \mathcal{P}(\mathcal{S}^n)} \bar{\mathbf{P}}_{e1}\big(q | \overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}(\pi)\big) + \max_{q \in \mathcal{P}(\mathcal{S}^n)} \bar{\mathbf{P}}_{e2}\big(q | \overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}(\pi)\big) \Big] \leq 2^{-n\lambda} \tag{5.141}$$

$$\max_{s^n \in \mathcal{S}^n} \max_{\pi \in \Pi_n} \mathbb{I}\big(\mathrm{M}_c; \mathrm{Z}_{s^n}^n(\pi) | \overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}(\pi)\big) \leq 2^{-n\tau}, \tag{5.142}$$

while the rate pair (R_p, R_c) satisfies the constraints of the region $\mathcal{R}_n^*(\overline{\mathfrak{W}}, \overline{\mathfrak{V}}, \mathfrak{V})$ in (5.125) after replacing r with q and $\mathcal{R}$ with $\mathcal{P}(\mathcal{S}^n)$ in (5.127). Our next step is to transform the different permutations of the deterministic code $\overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}$ constructed for the CBAVWC $(\overline{\mathfrak{W}}, \overline{\mathfrak{V}}, \mathfrak{V})$ into a correlated random code $\mathcal{C}_{\mathrm{ran}}^{\mathrm{pc}}$ for our AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$. This is done by letting Γ –the random variable that represents the correlated randomness of a correlated random code– to be uniformly distributed over the permutations set Π_n. Based on this strategy, the average decoding error probability of $\mathcal{C}_{\mathrm{ran}}^{\mathrm{pc}}$ is given by:

$$\sum_{i=1}^{2} \bar{\mathbf{P}}_{ei}(\mathcal{C}_{\mathrm{ran}}^{\mathrm{pc}}) = \sum_{i=1}^{2} \max_{s^n \in \mathcal{S}^n} \frac{1}{n!} \sum_{\pi \in \Pi_n} \bar{\mathbf{P}}_{ei}(s^n | \overline{\mathcal{C}_{\mathrm{det}}^{\mathrm{pc}}}(\pi)), \tag{5.143}$$

where $n!$ is the total number of permutations in the set $\Pi_n = \mathcal{G}$. In order to derive an upper-bound for this decoding error probability, we use the robustification lemma, cf. Lemma 5.1 which implies that if (5.141) holds then there exists $\tilde{\lambda} > 0$ such that the term in (5.143) is less than or equal $2^{-n\tilde{\lambda}}$. On the other hand, the bound in (5.142) directly implies that:

$$\max_{s^n \in \mathcal{S}^n} \max_{\gamma \in \Pi_n} \mathbb{I}(\mathrm{M}; \mathrm{Z}_{s^n}^n(\gamma) | \mathcal{C}_{\mathrm{ran}}^{\mathrm{pc}}) \leq 2^{-n\tau}. \tag{5.144}$$

The previous two steps show that for an arbitrary but fixed number of channel uses n, there exists a public-confidential correlated random code $\mathcal{C}_{\mathrm{ran}}^{\mathrm{pc}}$ for the AVWC $(\mathfrak{Q})$ that satisfies the reliability and secrecy constraints in Definition 5.20 as long as the rate pair (R_p, R_c) are bounded as in $\mathcal{R}_n^*(\mathfrak{Q})$. Finally, if we generalized our coding scheme for all $n \in \mathbb{N}$, it follows that the rate region $\mathcal{R}^*(\mathfrak{Q})$ is an achievable public-confidential correlated rate region for the AVWC: $\mathfrak{Q}$ and this completes the achievability proof of the first statement in Theorem 5.16.

Now, for the second statement in Theorem 5.16, we start by recalling the result established in [165, Theorem 1] and generalized in Lemma 5.2 which states that: if the AVC between the transmitter and the respective receiver is symmetrizable, then no reliable communication can be established over this AVC using deterministic codes. This implies that if $\mathfrak{W}$ is symmetrizable, then neither the public message $m_p \in \mathcal{M}_p$ nor the confidential message $m_c \in \mathcal{M}_c$ can be reliably transmitted to the legitimate receiver.

This directly implies that for this class of channels it follows that: $C_{\text{det}}^{\text{pc}}(\mathfrak{Q}) = (0,0)$. Moreover, the same argument implies that if only $\mathfrak{V}$ is symmetrizable, then the public message $m_p \in \mathcal{M}_p$ can not be reliably transmitted to the eavesdropper. This means that our problem simplifies to the characterization of the deterministic secrecy capacity of the AVWC $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$ when the legitimate AVC $\mathfrak{W}$ is not symmetrizable. Based on Theorem 5.5, it follows that $C_{\text{det}}^{\text{pc}}(\mathfrak{Q}) = (0, \mathcal{R}_c^*(\mathfrak{Q}))$. Thus, it only remains to prove that if both $\mathfrak{W}$ and $\mathfrak{V}$ are not symmetrizable, then $\mathcal{R}^*(\mathfrak{Q})$ can be achieved using a public-confidential deterministic code.

In order to prove this last case, we need to recall an important result regarding the amount of correlated randomness needed to construct the correlated random code $\mathcal{C}_{\text{ran}}^{\text{pc}}$ that establish the achievability of the first statement of Theorem 5.16. According to our proof, it is sufficient to have Γ be uniformly distributed over a set $\mathcal{G}$ such that $|\mathcal{G}| = n!$. Nevertheless, it was shown in [14, 159] as highlighted in Section 5.2.3 that this amount can be considerably reduced and in fact $|\mathcal{G}| = \mathcal{O}(n^2)$ is sufficient. Based on this result, we can use the elimination of correlation principle to construct a public-confidential deterministic code $\mathcal{C}_{\text{det}}^{\text{pc}}$ by concatenating a public deterministic code $\mathcal{C}_{\text{det}}^{\text{p}}$ and the public-confidential correlated random code $\mathcal{C}_{\text{ran}}^{pc}$. In this concatenation, $\mathcal{C}_{\text{det}}^{\text{p}}$ acts as a prefix code and is used to transmit a public message $m \in \mathcal{M} = \tilde{\mathcal{G}}$.

Now, it only remains to show two small points: The first point is that if both $\mathfrak{W}$ and $\mathfrak{V}$ are not symmetrizable, then both the legitimate receiver and the eavesdropper can decode the prefix message $m \in \mathcal{M}$. The second point is to show that adding the prefix code $\mathcal{C}_{\text{det}}^{\text{p}}$ on top of the correlated random code $\mathcal{C}_{\text{ran}}^{\text{pc}}$ has no effect on the total rate of the system. Based on the results established in [163, 165], we know that both the legitimate receiver and the eavesdropper are able to decode a message $m \in \mathcal{M}$ as long as R is small enough. Since $\mathcal{M} = \tilde{\mathcal{G}} = \mathcal{O}(n^2)$, we have

$$R = \lim_{n\to\infty} \frac{1}{n} \log |\tilde{\mathcal{G}}| \to 0. \tag{5.145}$$

The previous relation directly implies that the two previous requirements are satisfied, i.e. decoding error of m at both the legitimate receiver and the eavesdropper vanishes, the rates (R_p, R_c) of $\mathcal{C}_{\text{det}}^{\text{pc}}$ are identical to the ones of $\mathcal{C}_{\text{ran}}^{\text{pc}}$. This concludes the achievability proof of the first condition in the second statement of Theorem 5.16. It is important to highlight that the procedure used to prove this condition is related to the one used in Section 5.3.4 to provide an achievability proof for Theorem 5.9.

We now turn to the last part in the proof of Theorem 5.16 which is the converse proof. We will only focus on the converse of the public-confidential correlated random secrecy capacity $C_{\text{ran}}^{\text{pc}}(\mathfrak{Q})$ as the converse of the deterministic capacity $C_{\text{det}}^{\text{pc}}(\mathfrak{Q})$ follow accordingly using the same concepts. For the AVWC: $\mathfrak{Q} = (\mathfrak{W}, \mathfrak{V})$, assume that there exists a public-confidential correlated random code $\mathcal{C}_{\text{ran}}^{\text{pc}}$ of Definition 5.19, such that for $\epsilon > 0$ and a sufficiently large code block length n the following holds:

$$nR_j = \mathbb{H}(\mathrm{M}_j|\mathcal{C}_{\text{ran}}^{\text{pc}}) \qquad j \in \{p, c\} \tag{5.146a}$$

$$\max_{s^n \in \mathcal{S}^n} \bar{\mathbf{P}}_{ei}(s^n|\mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \epsilon \qquad i \in \{1, 2\} \tag{5.146b}$$

$$\max_{s^n \in \mathcal{S}^n} \max_{\gamma} \mathbb{I}(\mathrm{M}_c; \mathrm{Z}_{s^n}^n(\gamma)|\mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \epsilon. \tag{5.146c}$$

Since the average decoding error probability of the correlated random code is affine in the channel, it follows that $\bar{e}_i(\mathcal{C}_{\text{ran}}^{\text{pc}})$ does not change if one passes to the generalized channel state space $\mathcal{P}(\mathcal{S}^n)$. Thus, we have:

$$\max_{q\in\mathcal{P}(\mathcal{S})} \bar{\mathbf{P}}_{ei}(q|\mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \max_{\tilde{q}\in\mathcal{P}(\mathcal{S}^n)} \bar{\mathbf{P}}_{ei}(\tilde{q}|\mathcal{C}_{\text{ran}}^{\text{pc}}) = \max_{s^n\in\mathcal{S}^n} \bar{\mathbf{P}}_{ei}(s^n|\mathcal{C}_{\text{ran}}^{\text{pc}}) \leq \epsilon, \tag{5.147}$$

for $i \in \{1, 2\}$. Now, if we apply Fano's inequality cf. Lemma A.2 to the previous bounds, we can show that:

$$\max_{q\in\mathcal{P}(\mathcal{S})} \mathbb{H}(\mathrm{M}_p|\mathrm{Z}_q^n) \leq \max_{s^n\in\mathcal{S}^n} \mathbb{H}(\mathrm{M}_p|\mathrm{Z}_{s^n}^n) \leq 1 + n\epsilon R_p \tag{5.148}$$

$$\max_{q\in\mathcal{P}(\mathcal{S})} \mathbb{H}(\mathrm{M}_p\mathrm{M}_c|\mathrm{Y}_q^n) \leq \max_{s^n\in\mathcal{S}^n} \mathbb{H}(\mathrm{M}_p\mathrm{M}_c|\mathrm{Y}_{s^n}^n) \leq 1 + n\epsilon(R_p + R_c). \tag{5.149}$$

Based on these relations together with Eq. (5.146a), we can derive an upper-bound on the common rate R_p as follows:

$$\begin{aligned} R_p &\leq \frac{1}{n(1-\epsilon)}\Big(\min_{q\in\mathcal{P}(\mathcal{S})} \min\left[\mathbb{I}(\mathrm{M}_0;\mathrm{Y}_q^n), \mathbb{I}(\mathrm{M}_0;\mathrm{Z}_q^n)\right] + 1\Big) \\ &\overset{(a)}{\leq} \frac{1}{n(1-\epsilon)} \min_{q\in\mathcal{P}(\mathcal{S})} \min\Big[\mathbb{I}(\mathrm{U};\mathrm{Y}_q^n), \mathbb{I}(\mathrm{U};\mathrm{Z}_q^n)\Big] + \epsilon, \end{aligned} \tag{5.150}$$

where (a) follows by defining the mapping function $\mathcal{F}_p : \mathcal{M}_p \to \mathcal{U}$ and the fact that for sufficiently large n, $\frac{1}{n(1-\epsilon)} \leq \epsilon$. Next, we use [122, Lemma 2] and the bound in (5.146c) to show that:

$$\begin{aligned} \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_c;\mathrm{Z}_{s^n}^n|\mathrm{M}_p) &= \max_{s^n\in\mathcal{S}^n} \sum_{\gamma\in\mathcal{G}} \mathbb{I}(\mathrm{M}_c;\mathrm{Z}_{s^n}^n(\gamma)|\mathrm{M}_p)P_\Gamma(\gamma) \\ &\leq \max_{s^n\in\mathcal{S}^n} \max_{\gamma} \mathbb{I}(\mathrm{M}_c;\mathrm{Z}_{s^n}^n(\gamma)|\mathrm{M}_p) \leq 2\epsilon. \end{aligned} \tag{5.151}$$

Now, we can use the previous bound along with the bounds in (5.146a) and (5.149) to bound the confidential rate R_c as follows:

$$\begin{aligned} R_c &\overset{(a)}{\leq} \frac{1}{n(1-\epsilon)}\Big(\min_{q\in\mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{M}_1;\mathrm{Y}_q^n|\mathrm{U}) - \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{M}_1;\mathrm{Z}_{s^n}^n|\mathrm{U}) + 1 + 2\epsilon\Big) \\ &\overset{(b)}{\leq} \frac{1}{n(1-\epsilon)}\Big[\min_{q\in\mathcal{P}(\mathcal{S})} \mathbb{I}(\mathrm{V};\mathrm{Y}_q^n|\mathrm{U}) - \max_{s^n\in\mathcal{S}^n} \mathbb{I}(\mathrm{V};\mathrm{Z}_{s^n}^n|\mathrm{U})\Big] + 2\epsilon \end{aligned} \tag{5.152}$$

where (a) follows by using the same argument used in (5.150) along with the bound in (5.151), while (b) follows by defining the auxiliary function $\mathcal{F}_c : \mathcal{M}_c \times \mathcal{U} \to \mathcal{V}$ and the fact that for sufficiently large n, $\frac{1+2\epsilon}{n(1-\epsilon)} \leq 2\epsilon$. Finally, since we are only interested in the rates as $\epsilon \to 0$, it follows from Eqs. (5.150) and (5.152) that for any achievable rate pair (R_p, R_c), there exists an $\tilde{n} \in \mathbb{N}$ and a distribution $\tilde{\mathrm{P}}_{\mathrm{UVX}^{\tilde{n}}}$ such that $(R_p, R_c) \in \mathcal{R}_{\tilde{n}}^{\tilde{\mathrm{P}}}(\mathfrak{Q})$. This consequently implies that $C_{\text{ran}}^{\text{pc}}(\mathfrak{Q}) \subseteq \mathcal{R}^*(\mathfrak{Q})$ which completes our converse.

5.5 Discussion

The model of AVCs was first introduced to capture the uncertainty of the channel state during transmission and to model channels with imperfect CSI. Nevertheless, it

has become a very valuable and resourceful model specially in the secrecy domain. This is because the model of an AVWC can be interpreted as channel that undergoes two classes of attacks: A passive eavesdropper that only listen to the transmitted signal and tries to infer any information about the transmission along with an active jammer that threatens the communication by maliciously manipulating the channel state. Moreover, the model of AVWCs can be considered as a generalization to other important channel models like the multi-user wiretap channel and the compound wiretap channel. In this chapter, we discussed some of the major results established in previous literature for AVCs and AVWCs. We also presented some very interesting new results. These results allowed us to gain a better understanding to the problem of secure communication over AVWCs as highlighted by the following points:

1. One of the major issues of communication over AVCs is the symmetrizability phenomena. This phenomena captures the ability of AVCs to emulate a valid channel input making the design of an encoding-decoding rule for such channels a very challenging task. In particular, deterministic codes with a fixed encoder-decoder pair through the whole transmission fails to establish reliable communication over symmetrizable AVWCs, despite being useful for wiretap channels where perfect CSI is available. This observation implies that the development of coding schemes that works against passive eavesdropping is much easier than those who can overcome active jamming strategies.

2. Although correlated random codes manage to overcome the problem of symmetrizable AVWCs, the usage of correlated random codes in most practical scenarios is still questionable. This because as highlighted by Definition 5.5, correlated random codes require some sort of a pre-established coordination between the transmitter and the receiving nodes which might not always be feasible. Moreover, the reliability analysis of correlated random codes is derived under the assumption that the jammer possesses no information about this coordination. This assumption directly enforces some restrictions on the communication link between the eavesdropper and the jammer. In particular, the jammer is allowed to inform the eavesdropper about the jamming strategy used or even the exact channel state selected, but the eavesdropper is not allowed to inform the jammer about the exact realization of the common randomness.

3. The concept of list decoding appeared to be a promising solution for secure communication over AVWCs for two reasons: It provide an alternative coding solution if no common randomness is shared between the transmitting and receiving nodes. It can overcome the symmetrizability issue by using a list size L which is greater than the order of symmetrizability of the AVC between the transmitter and the legitimate receiver. Moreover, there exists some list encoding schemes that does not enforce any restrictions on the communication link between the jammer and the eavesdropper.

4. In Section 5.3, we introduced two coding scheme to derive the achievability of Theorem 5.9: A direct coding approach whose reliability analysis works for any AVWC and an indirect approach that utilizes the principle of elimination of correlation whose reliability analysis is only valid under the assumption that the jammer possesses no information about the messages transmitted over the channel. Although this argument might implies that the direct coding approach is better than the indirect one, there are some situations where the indirect approach is preferable. This is because the

indirect approach allows the integration of some public services like synchronization and channel estimation that do not need to be protected against eavesdropping and at the same time only utilizes negligible transmission rate. Thus, for communication scenario where a combination of these public services along with the usual confidential services is required, it is more advantageous to use the indirect approach.

5. One of the main limitations of the secrecy capacity results established in this chapter and for AVWCs in general is that no single-letter capacity description exists for the general AVWC and only multi-letter descriptions have been established. The dispute between single-letter and multi-letter capacity descriptions is a major issue that does not only occur in the classical information theory domain but its shadow has been casted to the quantum domain as well. In particular, for the class of classical-quantum channels (CQC), it was shown that although a single-letter characterization of the capacity in terms of mutual information is unknown, a multi-letter description is possible. This observation motivated Holevo to tackle the capacity characterization problem using a different information quantity other than the mutual information. He introduced the Holevo quantity in [203] and used it to establish a single-letter description of the CQC capacity in [204].

6. Despite the previous argument, we have shown in this chapter that multi-letter capacity descriptions can still be useful even in the absence of single-letter characterization. We used them to establish some properties regarding the continuity and the additivity behavior of the list capacity function for AVCs and AVWCs as well. We even showed that the list secrecy capacity of parallel AVWCs can endorse an extreme super-additivity feature known as super-activation, where joint decoding of two parallel AVWCs with a vanishing list secrecy capacity can achieve a non-zero secrecy rate.

Chapter 6

Conclusion and Future Work

With the recent growth in the field of communication systems and data transmission, it has become more crucial to guarantee a certain level of secrecy. However, the task of securing communication networks is becoming more challenging due to the rapid increase and the variability of the demands required to be achieved by these networks. In general, the term secure communication implies a list of various requirements. In this thesis, we limited our investigation to the confidentiality of the information transmitted over discrete memoryless channels from an information theoretic perspective.

We started our investigation in Chapter 2 by presenting the problem of secure communication over the wiretap channel introduced by Wyner in [2]. We then highlighted the different concepts used to define an information theoretic secrecy measure. These measures started with the weak secrecy notation introduced by Wyner, followed by the notation of strong secrecy [38] and more recently the concept of effective secrecy [31] and semantic secrecy [205]. We then discussed the different wiretap random coding techniques and the corresponding secrecy analysis used to establish secure communication over wiretap channels with respect to the required secrecy measure.

The topics discussed in the beginning of Chapter 2 were necessary to explore the main target of this thesis. Our aim was to investigate the problem of secure communication over two setups of wiretap channels with respect to the strong secrecy measure. The first setup considered wiretap channels with multiple legitimate receivers. During investigating this setup, we focused on the case where perfect CSI is available non-causally at the transmitting node. The second setup considered wiretap channels that suffers from another class of attacks beside passive eavesdropping known as active jamming. As an example of this setup we investigated secure communication over AVWCs which was originally introduced to model wiretap channels with imperfect CSI.

In previous literature, different channel models have been used to address the problem of secure communication over multi-receiver wiretap channels. However, all of these works only considered a conservative secrecy criterion known as **joint secrecy** with respect to the **weak secrecy** measure. In Section 2.4, we gave a brief summary for the achievable secrecy rate regions established for the two-receiver WBC. The presented results indicated that the development of coding schemes for the multi-receiver wiretap channel is not an easy task. This is because, despite the tremendous efforts conducted on this problem, the secrecy capacity of the simplest communication scenario, where a

common confidential message is transmitted over a two-receiver WBC is still unknown. Nevertheless, the integration of secure transmission over multi-user networks seems to be an inevitable requirement in all future communication system.

Bearing the previous argument in mind, we decided to uncover the potential advantages of introducing an additional legitimate receiver to the wiretap channel model. In [119], secure communication over a two-receiver WBC with receiver side information was investigated. We noticed that one of the coding schemes proposed therein did not fulfill the conservative joint secrecy criterion. Instead, it satisfied a weaker notation of secrecy that was only considered for MAWC known as the *individual secrecy* criterion. This criterion is based on the concept of mutual trust between the legitimate receivers. This implies that it is a unique feature for secure communication over wiretap channels with more than one legitimate receiver. This observation motivated us to examine how the individual secrecy criterion can affect the previous results established for the two-receiver DM-WBC.

Beside the concept of individual secrecy, there was another motive for us to consider secure communication over multi-receiver wiretap channels. Most of the coding schemes established for this class of channels were derived under the weak secrecy measure. This is because multi-user communication scenarios over BCs require complex codebook structures such as Marton coding. Differently from superposition encoding, where secrecy analysis techniques are available for both weak and **strong secrecy** measures, the secrecy analysis developed for Marton wiretap random codes in [12] is only valid for the weak secrecy notation. The only work that considered Marton wiretap codes under the strong secrecy notation was [127], yet it did not establish a full Marton coding region in the conventional sense.

The previous two reasons motivated us to study the communication scenario introduced in Chapter 3. In this scenario, a common public message, a common confidential message and two individual confidential messages are transmitted over a two-receiver DM-WBC. The model has a very interesting feature; as each legitimate receiver is only interested in one of the individual confidential messages, while being fully cognizant of the other one. At first, we investigated this scenario under the **joint-strong secrecy** criterion. We derived one of the main results of this thesis which is a full Matron coding achievable region that adhere to the strong secrecy measure. In particular, we showed that the strong secrecy analysis introduced in [127] can be adapted to a full Marton coding region using the concepts of time sharing and rate splitting. Beside Marton coding, our coding scheme utilized the principles of superposition encoding and indirect decoding. The combination of these various techniques appeared to be optimal in the sense that our achievable region is tight for all classes of the two-receiver DM-WBC where the joint secrecy capacity region is known. To the best of our knowledge, this is the first result to demonstrate the optimality of indirect decoding for a secure communication scenario.

Next, we extended our investigation to the **individual-strong secrecy** criterion. We derived a general achievable rate region that utilized the principle of secret key encoding in addition to the coding techniques used for the joint secrecy criterion (superposition encoding, Marton coding and indirect decoding). We showed that the

usage of secret key encoding is possible because each legitimate receiver knows the individual confidential message of the other one. However, this usage is only valid under the individual secrecy criterion, where a mutual trust exists between the legitimate receivers. Moreover, we established the individual secrecy capacity of some classes of the less noisy two-receiver DM-WBCs by showing the optimality of using superposition encoding and secret key encoding only. Our results showed that the individual secrecy criterion allows a larger capacity region compared to the joint one.

The results established in Chapter 3 raised a very interesting question about the usability of the individual secrecy criterion for general multi-receiver WBCs, where no receiver side information is available. This question was motivated by the fact that individual secrecy managed to provide a larger capacity region compared to joint secrecy by applying the concept of secret key encoding. This utilization was based on the fact that in the communication scenario investigated in Chapter 3, each legitimate receiver is cognizant of the individual confidential message of the other one. Thus, it is not clear how individual secrecy coding schemes can be constructed, if such information is not available. In order to answer this question, we investigated another communication scenario in Chapter 4, where two individual confidential messages are transmitted over a two-receiver DM-WBC without any receiver side information. This scenario have been previously investigated in previous literature but only under the **joint-weak** secrecy criterion, cf. [73, 84].

In Chapter 4, we showed that the results established for this communication scenario under the joint-weak secrecy criterion are also valid for the joint-strong secrecy criterion. In doing so, we derived a general achievable rate region using a coding scheme that combined the principle of superposition encoding and Marton encoding, where the strong secrecy analysis is based on the techniques developed in Chapter 3. We noticed that in our coding scheme the information encoded in the superposition variable (cloud center of our total coding scheme) can be decoded by both legitimate receivers. Thus, by encoding part of the confidential message of the first legitimate receiver in the superposition variable, we can allow the second legitimate receiver to create some sort of receiver side information and vice versa. This implies that we can use secret key encoding to encode confidential messages on higher layers.

The previous observation was the main key in investigating the broadcasting of individual confidential messages over the two-receiver DM-WBC without message cognition under the individual-strong secrecy criterion. With that in mind, we derived a general achievable rate region that combined the techniques used for joint secrecy along with secret key encoding. We then extended our investigation to the multi-receiver degraded DM-WBC. For this class of channels, we established the joint and individual-strong secrecy capacity regions. Our results indicated that even in the absence of receiver message cognition, the individual secrecy criterion can provide a larger capacity region compared to the joint one. Finally, we showed that our results can be used to derive the joint and individual secrecy capacity regions of the multi-receiver Gaussian SISO WBCs and the degraded multi-receiver Gaussian MIMO WBCs.

In Chapter 5, we investigated our second setup which considers secure communication over a channel that undergoes two different classes of attacks at the same time: A

passive eavesdropper that threatens the secrecy of the communication by eavesdropping upon the transmitted signal and an **active jammer** that threatens the reliability of the communication by maliciously manipulating the channel state. This setup has been addressed in previous literature from a different perspective by investigating secure communication over wiretap channels with imperfect CSI. Nevertheless, the two problems are equivalent. In previous literature, two main classes of channels have been used to model the behaviour of a wiretap channel with imperfect CSI: Compound wiretap channels and arbitrary varying wiretap channels (AVWCs). The main difference between these two classes lies in the rate at which the channel state varies as follows: An AVWC models a scenario where the channel state can vary from one time to another during a single transmission; while in a compound wiretap channel, the channel state remains constant for the transmission of each codeword. We limited our investigation to the class of AVWCs because they simulate a more general scenario and they capture the behaviour of various real life communication channels.

We started our investigation by introducing the two main coding schemes that have been used to establish a reliable and secure communication over AVWCs: Deterministic codes, in which a fixed encoder-decoder pair is used through out the whole transmission and correlated random codes, in which an encoder-decoder pair is selected based on some sort of common randomness shared between the transmitter and the receiving nodes. Unfortunately, both schemes suffer from various drawbacks as there exist some scenarios where neither of the two schemes can provide a secure communication over an AVWC at a non-vanishing rate. In particular, for an AVWC where the channel between the transmitter and the legitimate receiver is symmetrizable and the amount of common randomness is less than $\mathcal{O}(n^2)$, both coding schemes fails to establish a reliable transmission link. The previous result motivated us to use the principle of **list decoding** to develop an alternative coding scheme that can overcome the limitations of both deterministic and correlated random codes.

In Section 5.3, we derived a full characterization of the secrecy capacity of the AVWC under list decoding, showing that the list secrecy capacity is equivalent to the correlated random secrecy capacity, as long as the list size L is greater than the order of symmetrizability of the AVC between the transmitter and the legitimate receiver. Our result implies that even in the absence of a correlated randomness, we can still achieve the maximum secrecy rate for an AVWC by using an appropriate list size L. Moreover, we showed that even if we used a smaller list size L, we can still achieve rate closer to the capacity at the cost of a non-vanishing error probability. Surprisingly, such modification did not have any consequences on the secrecy constraints and a vanishing information leakage is still possible under a smaller list size. Furthermore, we used the established capacity characterization to investigate the analytical properties of the capacity function such as: continuity and additivity of parallel AVWCs. In particular, we derived the conditions that define a discontinuity point for the list secrecy capacity of an AVWC and showed that such points do exist. We also showed that there exists some scenarios for which the list secrecy capacity of an AVWC exhibits super-activation.

Although the investigation of secure communication over AVWC is not a new topic, most of the work presented in previous literature only considered one communication

scenario in which only one confidential message is transmitted over the channel. In Section 5.4, we considered a communication scenario in which we integrated both public and confidential services over an AVWC. We established both the deterministic and correlated random capacities for transmitting a common public message to both the legitimate receiver and the eavesdropper and a confidential message to the legitimate receiver. Our result aimed to be the first step in combining the two communication scenarios considered in this thesis by investigating the problem of secure communication over an AVWC with multiple legitimate receivers.

Future Research Directions

Through out this thesis, we extended the investigation of secure communication over wiretap channels to a more realistic and general scenarios. In the first half of the thesis, we considered secure communication over wiretap channels with multiple legitimate receivers, while in the second half, we considered the problem of active jamming and wiretap channels with imperfect CSI. A common observation for the two scenarios is that we only managed to establish a single-letter characterization for the secrecy capacity of some special cases, while for the general setup a single-letter expression for the secrecy capacity has remained unknown. However, in both setups, we were able to derive a multi-letter description for the secrecy capacity. We also showed that although multi-letter capacity descriptions are not efficiently computable, they can still play an important role in deriving some of the analytical properties of the capacity function and provide further useful insights. Nevertheless, we believe that there might be an alternative solution to tackle this problem.

Our belief originates from the work of Holevo in [206, 207] on the classical-quantum channel (CQC), where the channel input is a classical random variable, while the channel output is a quantum state. For such channel, a single-letter characterization of the capacity in terms of the mutual information is unknown, while a multi-letter description is possible. Holevo suggested to tackle the capacity characterization problem using a different information quantity instead of the mutual information. He introduced the Holevo quantity in [206] and used it to establish a single-letter description for the capacity of the CQC in [207]. We believe that this result might motivate researchers to investigate whether other information quantities can be used to establish a single-letter description for the secrecy capacity, where the mutual information can only provide a multi-letter description for the capacity. This argument was discussed recently in [208] and it seems to be a very interesting extension for our work.

In Chapter 5, we investigate the problem of secure communication over AVWC, where an active jammer manipulates the channel state s of the channel based on a certain distribution $q(s) \in \mathcal{P}(\mathcal{S})$. This setup is known as the AVWC with uncorrelated jamming because the jammer selects a certain channel state sequence s^n independently from both the transmitted message m and the channel input sequence x^n. In [209–211], an alternative setup for AVC was investigated. In this setup, the jammer chooses a certain channel state sequence s^n based on the transmitted message m or the channel input sequence x^n as shown in Fig. 6.1. This class of AVCs is very interesting because it captures a more sophisticated and challenging class of attacks.

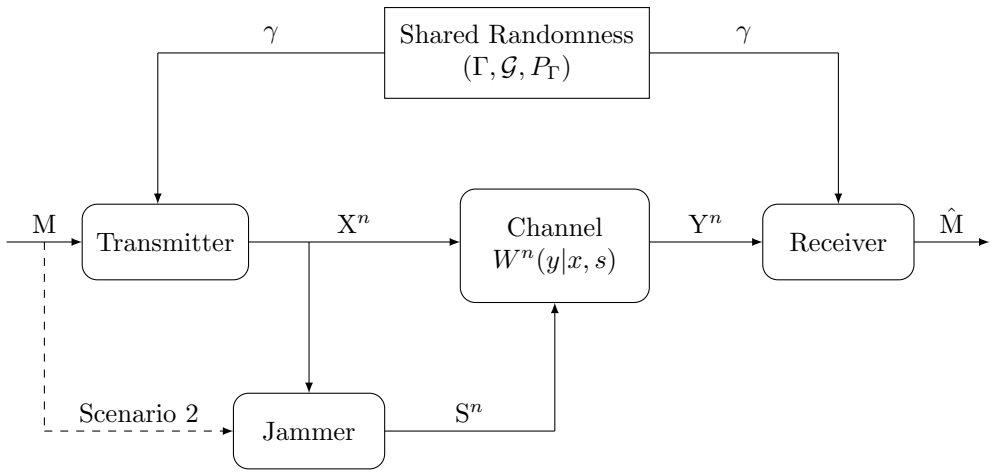


Figure 6.1: Arbitrarily varying channel with correlated jamming

In [209], Sarwate established the correlated random capacity of this channel using some of the results established by of Ahlswede in [212] along with some list decoding techniques [213]. The concept of correlated jamming was further investigated in [214] but this time over a classical-quantum AVC. The authors established the correlated random capacity for this channel using a different coding scheme from the one introduced by Sarwate in [210]. Despite these various investigations, the problem of secure communication over AVWC with correlated jamming is still an open problem and it also might be an interesting extension for our work.

Appendix A

Fundamental Tools in Information Theory

In this appendix, we discuss some of the basic and fundamental concepts of information theory. We also introduce various mathematical tools that play an important role in establishing the results presented in this thesis.

A.1 Entropy, Mutual Information and Divergence

Consider a discrete random variable X that can take values form the alphabet $\mathcal{X}$ according to the probability distribution $\mathrm{P_X}$, such that $\mathrm{P_X}(x)$ is the probability that the random variable X takes the value x. In information theory, one of the most important properties of a discrete random variable is entropy because it measures the uncertainty of the random variable [215]. In general, entropy plays a huge role in various applications such as astronomy, cryptography, signal processing, statistics, physics, image analysis in neuron-science, network theory, and bio-informatics [216]. In information theory, the term entropy usually refers to the Shannon entropy introduced by Claude E. Shannon in [217].

Definition A.1. For a discrete a random variable X, the entropy is denoted by $\mathbb{H}(\mathrm{X})$ and is defined as follows:

$$\mathbb{H}(\mathrm{X}) = -\sum_{x\in\mathcal{X}} \mathrm{P_X}(x) \cdot \log_b \mathrm{P_X}(x), \tag{A.1}$$

with the convention that $0 \log_b 0 = 0$.

Throughout this thesis, base b of the logarithm is taken to be 2, unless specified otherwise. This implies that entropy is measured in bits. Now, consider another discrete random variable Y with probability distribution $\mathrm{P_Y}$ and let $\mathrm{P_{XY}}$ be the joint probability distribution for two discrete random variables X and Y. The conditional entropy of Y given X is defined as follows:

$$\mathbb{H}(\mathrm{Y}|\mathrm{X}) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathrm{P_{XY}}(x,y) \log_b \mathrm{P_{Y|X}}(y|x). \tag{A.2}$$

The next theorem highlights some of the main properties of the entropy of discrete random variables.

Theorem A.1. *[215, 218] For a discrete random variable* X, $\mathbb{H}(\mathrm{X})$ *has the following properties:*

1. $0 \leq \mathbb{H}(\mathrm{X}) \leq \log_b |\mathcal{X}|$, *where the lower-bound is achieved if* X *is deterministic, while the upper-bound is achieved if* X *is uniformly distributed over the whole alphabet* $\mathcal{X}$.
2. *Conditioning reduces entropy: for a discrete random variable* Y, *it follows that* $\mathbb{H}(\mathrm{X}|\mathrm{Y}) \leq \mathbb{H}(\mathrm{X})$, *where the equality occurs if and only if* X *and* Y *are independent.*
3. *For a group of random variables, the following chain rule holds:*

$$\mathbb{H}(\mathrm{X}_1, \ldots \mathrm{X}_n) = \sum_{i=1}^{n} \mathbb{H}(\mathrm{X}_i|\mathrm{X}_1 \ldots \mathrm{X}_{i-1}). \tag{A.3}$$

Entropy is also used to calculate the mutual information between two random variables X and Y denoted by $\mathbb{I}(\mathrm{X};\mathrm{Y})$. The mutual information is a measure of the statistical dependence between those two random variables [215,219] and is given by the following relation:

$$\mathbb{I}(\mathrm{X};\mathrm{Y}) = \mathbb{H}(\mathrm{X}) - \mathbb{H}(\mathrm{X}|\mathrm{Y}). \tag{A.4}$$

The mutual information is also related to the Kullback-Leibler divergence [215], because $\mathbb{I}(\mathrm{X};\mathrm{Y})$ is equivalent to the divergence between the joint distribution P_{XY} and the product of the marginal distributions P_{X} and P_{Y} as follows:

$$\begin{aligned}\mathbb{I}(\mathrm{X};\mathrm{Y}) &= D_{\mathrm{KL}}(\mathrm{P}_{\mathrm{XY}}\|\mathrm{P}_{\mathrm{X}}\mathrm{P}_{\mathrm{Y}}) \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathrm{P}_{\mathrm{XY}}(x,y) \log \frac{\mathrm{P}_{\mathrm{XY}}(x,y)}{\mathrm{P}_{\mathrm{X}}(x)\mathrm{P}_{\mathrm{Y}}(y)}.\end{aligned} \tag{A.5}$$

The next theorem highlights some of the main properties of the mutual information.

Theorem A.2. *[215, 218] For the discrete random variables* X, Y *and* Z, $\mathbb{I}(\mathrm{X};\mathrm{Y})$ *has the following properties:*

1. $0 \leq \mathbb{I}(\mathrm{X};\mathrm{Y}) \leq \min[\mathbb{H}(\mathrm{X}), \mathbb{H}(\mathrm{Y})]$, *where the lower-bound is achieved if and only if* X *and* Y *are independent.*
2. *Unlike the entropy conditioning does not necessarily decrease the mutual information. However, if* $\mathrm{Z} - \mathrm{X} - \mathrm{Y}$ *forms a Markov chain, then conditioning can decrease the mutual information as follows:* $\mathbb{I}(\mathrm{X};\mathrm{Y}|\mathrm{Z}) \leq \mathbb{I}(\mathrm{X};\mathrm{Y})$.
3. *For a group of random variables, the following chain rule holds:*

$$\mathbb{I}(\mathrm{X}_1, \ldots \mathrm{X}_n;\mathrm{Y}) = \sum_{i=1}^{n} \mathbb{I}(\mathrm{X}_i;\mathrm{Y}|\mathrm{X}_1 \ldots \mathrm{X}_{i-1}). \tag{A.6}$$

A.2 Discrete Memoryless Channels

In his seminal paper "A mathematical theory of communication" [217] Claude E. Shannon stated that the problem of communication is to be able to reproduce at one point either exactly or approximately a message selected at another point. In order to do this reproduction, some amount of information must be transmitted over the physical channel between these two points. Based on this principle, Shannon proposed a general model for the point-to-point communication as shown in Fig.A.1, which is referred to as the discrete memoryless channel. The model of a discrete memoryless channel consists of:

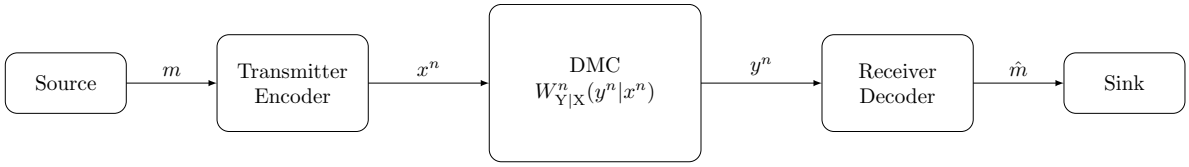


Figure A.1: Discrete memoryless channel model

- An information source that produces a message $m \in \mathcal{M}$ according to a certain probability distribution $\mathrm{P_M}$. All of the information sources considered in this thesis are assumed to have a uniform output distribution unless stated otherwise.
- A transmitter that transforms the produced message into a signal to be transmitted over the channel. In order to do so, an encoder $E : \mathcal{M} \to \mathcal{X}^n$ is used to map each message m into a codeword x^n, where n is the length of the codeword.
- A discrete memoryless channel (DMC) $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ that maps an input symbol x into an output symbol y according to the conditional probability distribution $W_{\mathrm{Y|X}}$. Since the channel input is a codeword of length n, the channel is used n times producing the output sequence y^n. The memoryless property indicates that the output of the channel at a certain instant only depends on the input at that instant and is conditionally independent of the previous channel inputs or outputs [220].
- A receiver that aims to reconstruct the transmitted message based on his observation. This is done by using a decoder $\varphi : \mathcal{Y}^n \to \mathcal{M} \cup \{?\}$ that maps maps the received sequence y^n into a message $\hat{m} \in \mathcal{M}$ or an error message.

In order to establish a reliable communication over a given DMC, one need to define an encoding-decoding strategy such that the reconstructed message $\hat{m}$ is most probably equivalent to the transmitted message m. In order to make things more abstract, the following definition is introduced:

Definition A.2. [221] A $(2^{nR}, n)$ code $\mathcal{C}$ for the DMC (W) consists of a message set $\mathcal{M} = [\![1, 2^{nR}]\!]$, an encoder $E : \mathcal{M} \to \mathcal{X}^n$ and a decoder $\varphi : \mathcal{Y}^n \to \mathcal{M} \cup \{?\}$.

R is defined as the rate of the code, the set of codewords $\{x^n(m) \in \mathcal{X}^n : m \in \mathcal{M}\}$ is called the codebook of the code. Now, it remains to define a measure for the reliability performance of a given code $\mathcal{C}$. In general, two reliability measures have been used [222, 223]. The first is the average probability of error given by:

$$\begin{aligned}\bar{\mathbf{P}}_e(\mathcal{C}) &= \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \sum_{y^n : \varphi(y^n) \neq m} W(y^n|x^n(m)) \\ &= \mathbb{P}(\hat{\mathrm{M}} \neq \mathrm{M}|\mathcal{C}),\end{aligned} \tag{A.7}$$

where $\hat{\mathrm{M}}$ is a random variable that represents the estimated message at the receiver. The second reliability measure used is the maximum error probability and is given by:

$$\tilde{\mathbf{P}}_e(\mathcal{C}) = \max_{m \in \mathcal{M}} \sum_{y^n : \varphi(y^n) \neq m} W(y^n|x^n(m)). \tag{A.8}$$

Definition A.3. A rate R is an achievable rate for the DMC (W) with respect to the average error probability measure, if for all $\eta > 0$ and $\lambda > 0$, there is an $n(\eta, \lambda) \in \mathbb{N}$

such that for all $n > n(\eta, \lambda)$ there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}\}_{n\geq 1}$ according to Definition A.2 such that:

$$\frac{1}{n}\log|\mathcal{M}| \geq R - \eta, \tag{A.9}$$

$$\bar{\mathbf{P}}_e(\mathcal{C}) \leq \lambda. \tag{A.10}$$

The channel capacity of the DMC (W) is given by the supremum of all achievable rates R and is denoted by $C(W)$.

Similarly, One can use the maximum error probability constraint to define the corresponding channel capacity and achievable rate. Nevertheless, most of the results derived in this thesis only consider the average error probability constraint.

Information theory aims to characterize the capacity of different DMC in terms of information-theoretic quantities that mainly depend only on the channel transmission matrix W. This characterization is called a coding theorem. In order to establish a coding theorem, one needs to proof two results. The first is called the direct result or the achievability proof which confirms the existence of a code whose rate is equivalent to the capacity expression. The second result is called the converse proof and it assures that no code can have an achievable rate greater than the capacity expression. In [217], Shannon presented the following coding theorem for DMCs,

Theorem A.3. *[217] The capacity of the DMC (W) is given by the following expression:*

$$C(W) = \max_{\mathrm{P_X}} \mathbb{I}(\mathrm{X};\mathrm{Y}). \tag{A.11}$$

where the maximum is taken over all possible input distributions $\mathrm{P_X}$.

The previous theorem is called the channel coding theorem and it is considered as one of the fundamental results in information theory [136]. In the following sections, some of the mathematical tools needed to establish this coding theorem are introduced. These tools are very important as they play an important role in confirming various results in the field of information theory.

A.3 Channel Comparison and Basic Lemmas

In investigating the channel capacity of DMCs, the notation of channel comparison always comes into play [224]. The notation of channel comparison identifies and compare the decoding capabilities of each channel by measuring the statistical dependence between a random variable that serves as a common input for all channels and the random variables that represent the outputs of each channel. The concept of channel comparison is extremely helpful for *multi-user* broadcasting scenarios because it is one of the main factors that defines the optimum coding scheme for such scenarios [225].

Now, consider a DMC $Q : \mathcal{X} \to \mathcal{P}(\mathcal{Y}, \mathcal{Z})$ given by the probability transition matrix $Q_{\mathrm{YZ|X}}$ and is characterized by the following marginal transition probabilities $W_{\mathrm{Y|X}}$ and $V_{\mathrm{Z|X}}$. This implies that Q can be interpreted as two DMCs $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$. The next definition gathers the most common notations used to compare DMCs.

Definition A.4. [226] A DMC (V) is said to be physically degraded with respect to the DMC (W) if

$$\forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \qquad Q_{\mathrm{YZ|X}}(y, z|x) = \sigma_{\mathrm{Z|Y}}(z|y) \cdot W_{\mathrm{Y|X}}(y|x), \tag{A.12}$$

for some transition probability $\sigma_{\mathrm{Z|Y}}$. In other words, physically degraded follows if $\mathrm{X} - \mathrm{Y} - \mathrm{Z}$ forms a Markov chain. On the other hand, V is said to be stochastically degraded with respect to W if there exists a DMC $\sigma : \mathcal{Y} \to \mathcal{P}(\mathcal{Z})$ such that

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z} \qquad V_{\mathrm{Z|X}}(z|x) = \sum_{y \in \mathcal{Y}} \sigma_{\mathrm{Z|Y}}(z|y) \cdot W_{\mathrm{Y|X}}(y|x). \tag{A.13}$$

The DMC (W) is said to be less noisy than the DMC (V), if for every random variable U such that $\mathrm{U} - \mathrm{X} - \mathrm{YZ}$ forms a Markov chain, the following holds:

$$\mathbb{I}(\mathrm{U}; \mathrm{Y}) \geq \mathbb{I}(\mathrm{U}; \mathrm{Z}). \tag{A.14}$$

The previous notation is sometimes expressed as Z is noisier than Y and is formulated mathematically as $\mathrm{Y} \succeq \mathrm{Z}$. Finally, the DMC (W) is said to be more capable than the DMC (V), if for every distribution on X, the following holds:

$$\mathbb{I}(\mathrm{X}; \mathrm{Y}) \geq \mathbb{I}(\mathrm{X}; \mathrm{Z}). \tag{A.15}$$

Theorem A.4. *[16, 226] Let $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ and $V : \mathcal{X} \to \mathcal{P}(\mathcal{Z})$ be two DMCs and consider the following statements:*

1. *(V) is physically degraded with respect to (W).*
2. *(V) is stochastically degraded with respect to (W).*
3. *(W) is less noisy than (V).*
4. *(W) is more capable than (V).*

Then, the following relation holds:

$$1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4. \tag{A.16}$$

The previous theorem establishes a relation between the different classes of channel introduced in Definition A.4. It clearly indicates that the class of more capable channels is strictly wider than the less noisy one. It also implies that the class of less noisy channels contains the physically and stochastically degraded channels.

In addition to the concept of channel comparison, this section highlights some important tools that are crucial to the characterization of the channel capacity for various communication scenarios.

Lemma A.1. *[227] Let* X, Y *and* Z *be three discrete random variables such that* $\mathrm{X} - \mathrm{Y} - \mathrm{Z}$ *forms a Markov chain. Then, the following holds*

$$\mathbb{I}(\mathrm{X}; \mathrm{Z}) \leq \mathbb{I}(\mathrm{X}; \mathrm{Y}). \tag{A.17}$$

The previous result is knows as the data-processing inequality and it implies that on average taking a processed version of the random variable Y which in our case is Z can only lead to an increase in our uncertainty about the original random variable X.

Lemma A.2. *[22] Let* X *be a discrete random variable on the alphabet* $\mathcal{X}$ *and* $\hat{\mathrm{X}}$ *be an estimate of* X *that takes values in the same alphabet. Then for* $\boldsymbol{P}_e = \mathbb{P}[\mathrm{X} \neq \hat{\mathrm{X}}]$*, we have*

$$\mathbb{H}(\mathrm{X}|\hat{\mathrm{X}}) \leq \mathbb{H}_2(\boldsymbol{P}_e) + \boldsymbol{P}_e \log_2(|\mathcal{X}| - 1), \tag{A.18}$$

where $\mathbb{H}_2(\boldsymbol{P}_e)$ *is the binary entropy function and is defined as* $\mathbb{H}_2(\boldsymbol{P}_e) = -\boldsymbol{P}_e \log_2 \boldsymbol{P}_e$ $-(1 - \boldsymbol{P}_e)\ \log_2(1 - \boldsymbol{P}_e)$.

The previous relation is known as Fano's inequality and it is one of the most crucial results for communication because it establishes a relation between an operational quantity (decoding error probability) and an information-theoretic quantity (the conditional entropy) [215]. One can easily show that if $\mathbf{P}_e$ approaches zero, then $\mathbb{H}(\mathrm{X}|\hat{\mathrm{X}})$ also approaches zero.

Another important result is a bound that exploits the concentration of independent and identically distributed (i.i.d.) random variables around their expectation as follows:

Lemma A.3. *[228] Let* $b > 0$ *and let* $\mathrm{Z}_1, \ldots, \mathrm{Z}_L$ *be i.i.d. random variables with values in* $[0, b]$*. Let* $\mu = \mathbb{E}[\mathrm{Z}_1]$ *be the expectation of* Z_1*. Then*

$$\mathbb{P}\left[\frac{1}{L}\sum_{i=1}^{L} \mathrm{Z}_i \notin [(1 \pm \epsilon)\mu]\right] \leq 2\exp\left(-L \cdot \frac{\epsilon^2 \mu}{2b \ln 2}\right), \tag{A.19}$$

where $[(1 \pm \epsilon)\mu]$ *denotes the interval* $[(1 - \epsilon)\mu, (1 + \epsilon)\mu]$.

The previous bound is known as the Chernoff-Hoeffding bound because it was introduced by Hoeffding in [228] based on an adaptation to the Chernoff bound introduced in [229]. Another, important result is the following lemma:

Lemma A.4. *[16, Lemma 2.1] Let* X_n *be a random variable on the alphabet* $\mathcal{X}_n$ *and let* $\mathcal{F}$ *be a finite set of functions* $f : \mathrm{X}_n \to \mathbb{R}^+$ *such that* $|\mathcal{F}|$ *does not depend on* n*. Now, if*

$$\forall f \in \mathcal{F} \quad \mathbb{E}_{\mathrm{X}_n}[f(\mathrm{X}_n)] \leq \delta(n), \tag{A.20}$$

then there exists a specific realization x_n *of* X_n *such that*

$$\forall f \in \mathcal{F} \quad f(x_n) \leq \delta(n). \tag{A.21}$$

The previous lemma is known as the "selection lemma" because it implies that if a random variable satisfies some constraints on average, then there must exist a realization of this random variable that satisfies these constraints. This lemma directly follows from the Markov's inequality [230].

Lemma A.5. *[231] Let* P_X *and* Q_X *be two probability distribution for the discrete random variable* X *that takes values in the finite alphabet* $\mathcal{X}$*, then the following holds*

$$D_{\mathrm{KL}}(\mathrm{P}_\mathrm{X} \| \mathrm{Q}_\mathrm{X}) \geq \frac{1}{2\ln 2}\|\mathrm{P}_\mathrm{X} - \mathrm{Q}_\mathrm{X}\|^2, \tag{A.22}$$

where $D_{\mathrm{KL}}(\mathrm{P_X}\|\mathrm{Q_X})$ *is the Kullback-Leibler divergence defined in the previous section, while* $\|\mathrm{P_X} - \mathrm{Q_X}\|$ *is the total variation distance between* $\mathrm{P_X}$ *and* $\mathrm{Q_X}$ *is defined as follows:*

$$\|\mathrm{P_X} - \mathrm{Q_X}\| = \sum_{a \in \mathcal{X}} |\mathrm{P_X}(a) - \mathrm{Q_X}(a)|. \tag{A.23}$$

The previous result is known as Pinsker's inequality and it is of high importance because it establishes a relation between two fundamental measures for discrete probability distributions. A refinement of the constant in Pinsker's inequality has been addressed in [232]. It was shown in [233] that a general inverse for Pinsker's inequality cannot be established. This is because one can show that for a random variable X, there exist distributions $\mathrm{P_X}$ and $\mathrm{Q_X}$ such that for $0 < \epsilon \leq 2$, the following holds:

$$\|\mathrm{P_X} - \mathrm{Q_X}\| \leq \epsilon, \quad \text{while} \quad D_{\mathrm{KL}}(\mathrm{P_X}\|\mathrm{Q_X}) = \infty \tag{A.24}$$

Nevertheless, it was shown that for some special cases an inverse inequality does exist cf. [234–236]. In particular, the bound established in the following lemma is og high importance:

Lemma A.6. *[237, Lemma 2.7] Let* X *and* Y *be two discrete random variables that take values in the finite alphabets* $\mathcal{X}$ *and* $\mathcal{Y}$ *respectively. Next, let* $\mathrm{P_{XY}}$ *be a joint distribution on* X *and* Y *with marginal distributions* $\mathrm{P_X}$ *and* $\mathrm{P_Y}$*. Now, if* $\|\mathrm{P_{XY}} - \mathrm{P_X P_y}\| \leq \epsilon \leq \frac{1}{2}$*, then*

$$\mathbb{I}(\mathrm{X};\mathrm{Y}) = D_{\mathrm{KL}}(\mathrm{P_{XY}}\|\mathrm{P_X P_y}) \leq -\epsilon \log \frac{\epsilon}{|\mathcal{X}||\mathcal{Y}|}. \tag{A.25}$$

The proof of the previous lemma is based on the definition of the mutual information along with the continuity of the entropy function.

A.4 Types and Typical Sequences

The principle of types and typical sequences is one of the main pillars of information theory due to the asymptotic equipartition property. This property plays an important role in establishing the direct part of a coding theorem [238]. There exists two notation for typicality: weak typicality and strong typicality. In this thesis, only the strong typicality notation is utilized.

Consider a discrete memoryless source (DMS) that outputs n symbols from a discrete and finite alphabet $\mathcal{X}$ producing the sequence x^n. The symbols are produced in an i.d.d. manner according to the probability distribution $\mathrm{P_X}$. This implies that the probability of a sequence x^n to be produced by this DMS is given by:

$$\mathrm{P}_{\mathrm{X}}^n(x^n) = \prod_{i=1}^{n} \mathrm{P_X}(x_i). \tag{A.26}$$

Definition A.5. The type of a sequence $x^n \in \mathcal{X}^n$ is the empirical distribution induced by the sequence x^n over the alphabet $\mathcal{X}$ and is given by:

$$\mathrm{Q_X}(a) = \frac{1}{n} N(a|x^n) \quad \text{for every} \quad a \in \mathcal{X}, \tag{A.27}$$

$N(a|x^n)$ denotes the number of occurrences of the element a in x^n.

Although x^n was generated according to the distribution $\mathrm{P_X}$, there is no guarantee that the type of x^n given by $\mathrm{Q_X}$ is identical to $\mathrm{P_X}$. That is why we present the following definition:

Definition A.6. Consider a probability distribution $\mathrm{P_X}$ on the finite alphabet $\mathcal{X}$ and let $\epsilon > 0$. A sequence $x^n \in \mathcal{X}^n$ is said to be strongly ϵ-typical with respect to $\mathrm{P_X}$, if the type of x^n is close to $\mathrm{P_X}$ as follows:

$$\left|\frac{N(a|x^n)}{n} - \mathrm{P_X}(a)\right| \leq \epsilon \cdot \mathrm{P_X}(a) \quad \text{for every} \quad a \in \mathcal{X}. \tag{A.28}$$

The set of sequences x^n that satisfies the constraint in (A.28) is called the ϵ-strongly typical set with respect to $\mathrm{P_X}$ and is denoted by $\mathcal{T}_\epsilon^n(\mathrm{P_X})$.

The notation of strong typicality is also known in literature as letter typicality and can only be applied to discrete random variables. Unlike the strong typicality, the weak typicality notation can be used for both discrete and continuous random variables. This is because the weak notation requires the empirical entropy of a sequence x^n to be close to the true entropy of the corresponding random variable. That is why the weak typicality notation is also known as the entropy typicality [215].

Theorem A.5. *[237, 239] Consider a probability distribution* $\mathrm{P_X}$ *on a finite alphabet* $\mathcal{X}$ *and let* X^n *be a random variable that represents the produced sequences by the DMS* $\mathrm{P_X}$. *Then for* $\epsilon > 0$, *we have*

$$\forall x^n \in \mathcal{T}_\epsilon^n(\mathrm{P_X}) \quad 2^{-n(1+\epsilon)\mathbb{H}(\mathrm{X})} \leq \mathrm{P_X^n}(x^n) \leq 2^{-n(1-\epsilon)\mathbb{H}(\mathrm{X})} \tag{A.29}$$

$$1 - \delta_n(\epsilon) \leq \mathbb{P}[\mathrm{X}^n \in \mathcal{T}_\epsilon^n(\mathrm{P_X})] \leq 1 \tag{A.30}$$

$$(1 - \delta_n(\epsilon))2^{n(1-\epsilon)\mathbb{H}(\mathrm{X})} \leq |\mathcal{T}_\epsilon^n(\mathrm{P_X})| \leq 2^{n(1+\epsilon)\mathbb{H}(\mathrm{X})}, \tag{A.31}$$

where $\delta_n(\epsilon) = 2|\mathcal{X}| \cdot e^{-2n\epsilon^2\mu_{\mathrm{X}}^2}$ *and* $\mu_{\mathrm{X}} = \min_{a \in supp(\mathrm{P_X})} \mathrm{P_X}(a)$.

One can notice that $\lim_{n\to\infty} \delta_n(\epsilon) = 0$. The previous theorem is known as the the asymptotic equipartition property (AEP) and it implies that if n is sufficiently large, then with a probability close to one the produced sequence x^n belongs to the strongly typical set. The proof of the inequality in (A.29) follows directly from the constraint of strong typical sequences in (A.28). On the other hand, the inequality in (A.30) follows by using the Chernoff bound and the constraint in (A.29) along with Pinsker's inequality. The last inequality in (A.31) follows directly from the first two inequalities. The the notion of strong typicality can be easily generalized to jointly distributed random variables as follows:

Definition A.7. Consider the joint probability distribution $\mathrm{P_{XY}}$ on the finite alphabet $\mathcal{X} \times \mathcal{Y}$ and let $\epsilon > 0$. The sequence $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$ are said to be ϵ-strongly jointly typical with respect to $\mathrm{P_{XY}}$, if

$$\left|\frac{N(a,b|x^n,y^n)}{n} - \mathrm{P_{XY}}(a,b)\right| \leq \epsilon \cdot \mathrm{P_{XY}}(a,b) \quad \text{for every} \quad (a,b) \in \mathcal{X} \times \mathcal{Y}. \tag{A.32}$$

The ϵ-strongly jointly typical set with respect to $\mathrm{P_{XY}}$ is denoted by $\mathcal{T}_\epsilon^n(\mathrm{P_{XY}})$.

One can easily show that if the sequences $(x^n, y^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_{XY}})$ then $x^n \in \mathcal{T}_\epsilon^n(\mathrm{P_X})$ and $y^n \in \mathcal{T}_\epsilon^n(\mathrm{P_Y})$. This property is known as the consistency of joint typicality [16]. Another important generalization is the conditional typicality defined as follows:

Definition A.8. Consider the joint probability distribution $\mathrm{P_{XY}}$ on the finite alphabet $\mathcal{X} \times \mathcal{Y}$ and let $\epsilon > 0$. For a sequence $x^n \in \mathcal{T}_\epsilon^n(\mathrm{P_X})$, the set

$$\mathcal{T}_\epsilon^n(\mathrm{P_{XY}}|x^n) = \{y^n \in \mathcal{Y}^n : (x^n, y^n) \in \mathcal{T}_\epsilon^n(\mathrm{P_{XY}})\} \tag{A.33}$$

is called the ϵ-strongly conditional typical set with respect to x^n.

Theorem A.6. *[237] Consider a joint probability distribution* $\mathrm{P_{XY}}$ *on the finite alphabet* $\mathcal{X} \times \mathcal{Y}$ *and let* $(\mathrm{X}^n, \mathrm{Y}^n)$ *be the random variables that represent the sequences produced by the DMS* $\mathrm{P_{XY}}$*. Then for* $0 < \epsilon_1 < \epsilon_2$ *and a sequence* $x^n \in \mathcal{T}_{\epsilon_1}^n(\mathrm{P_X})$*, we have*

$$\forall y^n \in \mathcal{T}_{\epsilon_1}^n(\mathrm{P_{XY}}|x^n) \quad 2^{-n(1+\epsilon_1)\mathbb{H}(\mathrm{Y|X})} \leq \mathrm{P}_{\mathrm{Y|X}}^n(y^n|x^n) \leq 2^{-n(1-\epsilon_1)\mathbb{H}(\mathrm{Y|X})} \tag{A.34}$$

$$1 - \delta_n(\epsilon_1, \epsilon_2) \leq \mathbb{P}[\mathrm{Y}^n \in \mathcal{T}_{\epsilon_2}^n(\mathrm{P_{XY}}|x^n)|\mathrm{X}^n = x^n] \leq 1 \tag{A.35}$$

$$(1 - \delta_n(\epsilon_1, \epsilon_2))2^{n(1-\epsilon_2)\mathbb{H}(\mathrm{Y|X})} \leq |\mathcal{T}_{\epsilon_2}^n(\mathrm{P_{XY}}|x^n)| \leq 2^{n(1+\epsilon_2)\mathbb{H}(\mathrm{Y|X})}, \tag{A.36}$$

where $\delta_n(\epsilon_1, \epsilon_2) = 2|\mathcal{X}||\mathcal{Y}| \exp\left(-2n\left(\frac{\epsilon_2-\epsilon_1}{1+\epsilon_1}\right)^2 \mu_{\mathrm{XY}}^2\right)$ *and* $\mu_{\mathrm{XY}} = \min_{(a,b)\in supp(\mathrm{P_{XY}})} \mathrm{P_{XY}}(a, b)$.

The previous theorem implies that for a typical sequence x^n and a random sequence Y^n generated by the following distribution $\mathrm{P}_{\mathrm{Y}}^n(y^n) = \prod_{i=1}^n \mathrm{P_{Y|X}}(y_i|x_i)$, then Y^n is jointly typical with x^n with probability that approaches one as n approaches infinity. Next, we consider scenarios where X^n and Y^n are generated in a different way as follows:

Theorem A.7. *[16, 237] Consider the joint probability distribution* $\mathrm{P_{XY}}$ *on the finite alphabet* $\mathcal{X} \times \mathcal{Y}$ *and let* X^n *and* Y^n *be two random variables that represent the sequences generated by two independent DMS* $\mathrm{P_X}$ *and* $\mathrm{P_Y}$ *respectively. Then for* $0 < \epsilon_1 < \epsilon_2$ *and a sequence* $x^n \in \mathcal{T}_{\epsilon_1}^n(\mathrm{P_X})$*, we have*

$$(1 - \delta_n(\epsilon_1))2^{-n[\mathbb{I}(\mathrm{X;Y})+2\epsilon_1\gamma_{xy}]} \leq \mathbb{P}\left[(\mathrm{X}^n, \mathrm{Y}^n) \in \mathcal{T}_{\epsilon_1}^n(\mathrm{P_{XY}})\right] \leq 2^{-n[\mathbb{I}(\mathrm{X;Y})-2\epsilon_1\gamma_{xy}]} \tag{A.37}$$

$$(1 - \delta_n(\epsilon_1, \epsilon_2))2^{-n[\mathbb{I}(\mathrm{X;Y})+2\epsilon_2\gamma_y]} \leq \mathbb{P}\left[\mathrm{Y}^n \in \mathcal{T}_{\epsilon_2}^n(\mathrm{P_{XY}}|x^n)\right] \leq 2^{-n[\mathbb{I}(\mathrm{X;Y})-2\epsilon_2\gamma_y]}, \tag{A.38}$$

where $\gamma_{xy} = \log|\mathcal{X}||\mathcal{Y}|$ *and* $\gamma_y = \log|\mathcal{Y}|$*, while* $\delta_n(\epsilon_1)$ *and* $\delta_n(\epsilon_1, \epsilon_2)$ *are as defined before.*

The previous theorem can be easily generalized to joint distributions with more than two random variables. This generalization is used a lot throughout this thesis.

A.5 Useful Bounding Lemmas

In this section, we discuss some useful lemmas that play an important role in establishing some of the results presented in this thesis. Most of these lemmas are novel results established by us to serve as a supporting arguments for our main results. We now present these lemmas and prove their validity.

Lemma A.7. *[121, Lemma 2] Consider two independent random variables* M *and* W*, such that* $\mathbb{H}(\mathrm{W}|\mathrm{Z}^n) \leq \alpha$ *and* $\mathbb{I}(\mathrm{M};\mathrm{Z}^n) \leq \beta$*, where* $\alpha, \beta > 0$*. Then, the following relation holds:* $\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{W}) \leq \alpha + \beta$*.*

Proof.

$$\begin{aligned}
\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{W}) &= \mathbb{H}(\mathrm{M}|\mathrm{W}) - \mathbb{H}(\mathrm{M}|\mathrm{Z}^n\mathrm{W}) \\
&\leq \mathbb{H}(\mathrm{M}|\mathrm{W}) - \mathbb{H}(\mathrm{M}|\mathrm{Z}^n\mathrm{W}) - \mathbb{H}(\mathrm{W}|\mathrm{Z}^n) + \alpha \\
&\overset{(a)}{=} \mathbb{H}(\mathrm{M}) - \mathbb{H}(\mathrm{W}\mathrm{M}|\mathrm{Z}^n) + \alpha \\
&= \mathbb{H}(\mathrm{M}) - \mathbb{H}(\mathrm{M}|\mathrm{Z}^n) - \mathbb{H}(\mathrm{W}|\mathrm{Z}^n\mathrm{M}) + \alpha \\
&\overset{(b)}{\leq} \mathbb{I}(\mathrm{M};\mathrm{Z}^n) + \alpha \leq \alpha + \beta,
\end{aligned}$$

where (a) follows due to the chain rule of entropy along with the fact that M and W are independent; while (b) follows from the properties of entropy, i.e. $\mathbb{H}(\mathrm{W}|\mathrm{Z}^n\mathrm{M}) \geq 0$. ■

Lemma A.8. *Let* $Q_{\mathrm{Y},\mathrm{Z}|\mathrm{X}}$ *be a discrete memoryless BC and assume that* $\mathrm{Y} \succeq \mathrm{Z}$*. Consider two independent random variables* M *and* W*, such that* $(\mathrm{M},\mathrm{W}) - \mathrm{X}^n - (\mathrm{Y}^n,\mathrm{Z}^n)$ *forms a Markov chain. Then the following holds:*

$$\mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{W}) \geq \mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{W}).$$

Proof. We define $\Delta = \frac{1}{n}[\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{W}) - \mathbb{I}(\mathrm{M};\mathrm{Y}^n|\mathrm{W})]$ and prove that if $\mathrm{Y} \succeq \mathrm{Z}$, $\Delta \leq 0$ and this directly implies our proposition. Let $\mathrm{U}_i \triangleq (\mathrm{W}, \tilde{\mathrm{Z}}^{i+1}, \mathrm{Y}^{i-1})$ and $\mathrm{V}_i \triangleq (\mathrm{M}, \mathrm{U}_i)$, we have

$$\begin{aligned}
\Delta &= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{W}\tilde{\mathrm{Z}}^{i+1}) - \mathbb{I}(\mathrm{M};\mathrm{Y}_i|\mathrm{W}\mathrm{Y}^{i-1})\Big] \\
&\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{M};\mathrm{Z}_i|\mathrm{W}\tilde{\mathrm{Z}}^{i+1}\mathrm{Y}^{i-1}) - \mathbb{I}(\mathrm{M};\mathrm{Y}_i|\mathrm{W}\tilde{\mathrm{Z}}^{i+1}\mathrm{Y}^{i-1})\Big] \\
&= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathbb{I}(\mathrm{V}_i;\mathrm{Z}_i|\mathrm{U}_i) - \mathbb{I}(\mathrm{V}_i;\mathrm{Y}_i|\mathrm{U}_i)\Big] \\
&\overset{(b)}{=} \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}) \\
&= \mathbb{E}_{\mathrm{U}}\Big[\mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}) - \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U})\Big] \\
&\overset{(c)}{\leq} \mathbb{I}(\mathrm{V};\mathrm{Z}|\mathrm{U}=\bar{u}) - \mathbb{I}(\mathrm{V};\mathrm{Y}|\mathrm{U}=\bar{u}) \\
&\overset{(d)}{\leq} \mathbb{I}(\bar{\mathrm{V}};\mathrm{Z}) - \mathbb{I}(\bar{\mathrm{V}};\mathrm{Y}) \overset{(e)}{\leq} 0
\end{aligned} \tag{A.39}$$

where (a) follows from the Csiszár sum identity [10, Lemma 7]; (b) follows by introducing a random variable T independent of all others and uniformly distributed over $[\![1,n]\!]$, then letting $\mathrm{U} = (\mathrm{U}_{\mathrm{T}}, \mathrm{T})$, $\mathrm{V} = \mathrm{V}_{\mathrm{T}}$, $\mathrm{Y} = \mathrm{Y}_{\mathrm{T}}$ and $\mathrm{Z} = \mathrm{Z}_{\mathrm{T}}$; (c) follows as $\bar{u}$ is the realization of U that maximizes the difference; (d) follows as $\bar{\mathrm{V}}$ is distributed as $\mathrm{Q}(v|u=\bar{u})$ cf. [5, Corollary 2.3] and (e) follows since $\bar{\mathrm{V}} - \mathrm{X} - (\mathrm{Y},\mathrm{Z})$ forms a Markov chain and $\mathrm{Y} \succeq \mathrm{Z}$. ■

Lemma A.9. *Consider a multi-receiver Gaussian SISO WBC as defined in* (4.35)*, where* $\mathbb{E}[\mathrm{X}^2] \leq P$ *and the variances of the Gaussian noises satisfy the order in* (4.36)*. Moreover, assume that the following Markov chain* $\mathrm{U}_k - \cdots - \mathrm{U}_2 - \mathrm{X} - \mathrm{Y}_1 - \cdots - \mathrm{Y}_k - \mathrm{Z}$ *holds. Now, if the following inequality is true:*

$$\mathbb{I}(\mathrm{U}_j;\mathrm{Y}_j|\mathrm{U}_{j+1}) - \mathbb{I}(\mathrm{U}_j;\mathrm{Z}|\mathrm{U}_{j+1}) \leq f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P + \sigma_j^2}\right) - f\left(\frac{\alpha_j P}{\sum_{i=1}^{j-1}\alpha_i P + \sigma_Z^2}\right), \tag{A.40}$$

for every $j \in [\![1,k]\!]$ *and* $\alpha_j \in [0,1]$*, where* $\mathrm{U}_{k+1} = \emptyset$*,* $\mathrm{U}_1 = \mathrm{X}$*,* $\sum_{i=1}^{k}\alpha_i = 1$ *and* $f(a) \triangleq \frac{1}{2}\log(1+a)$*, then it follows directly that:*

$$\mathbb{E}[\mathrm{U}_j^2] \leq \sum_{i=j}^{k}\alpha_i P. \tag{A.41}$$

Proof. We start by considering the random variable U_k, i.e. $j = k$. Based on the results established in [94], we know that the $\mathbb{LHS}$ of Eq. (A.40) is maximized by Gaussian signaling. Thus, we let U_k be a Gaussian random variable and assume that $\mathbb{E}[\mathrm{U}_k^2] = (\alpha_k + \gamma)P$, where $\gamma \geq 0$. Further, we let $\mathrm{X} = \mathrm{U}_k + \bar{\mathrm{V}}_k$, where $\bar{\mathrm{V}}_k$ is a Gaussian random variable independent from U_k. This implies that $\mathbb{E}[\bar{\mathrm{V}}_k^2] = \big(1 - (\alpha_k + \gamma)\big)P = \Big(\sum_{i=1}^{k-1}\alpha_i - \gamma\Big)P$. Under the previous power assumptions, we have:

$$\mathbb{I}(\mathrm{U}_k;\mathrm{Y}_k) - \mathbb{I}(\mathrm{U}_k;\mathrm{Z}) = f\left(\frac{(\alpha_k+\gamma)P}{\Big(\sum_{i=1}^{k-1}\alpha_i - \gamma\Big)P + \sigma_k^2}\right) f\left(\frac{(\alpha_k+\gamma)P}{\Big(\sum_{i=1}^{k-1}\alpha_i - \gamma\Big)P + \sigma_Z^2}\right). \tag{A.42}$$

However, this condition contradicts the given constraint given in (A.40) at $j = k$, unless $\gamma \leq 0$ which consequently implies that $\mathbb{E}[\mathrm{U}_k^2] \leq \alpha_k P$. Now, by repeating these steps recursively for every random variable U_j, we can show that $\mathbb{E}[\mathrm{U}_j^2] \leq \sum_{i=j}^{k}\alpha_i P$ and this completes our proof. ■

Lemma A.10. *For the degraded Gaussian MIMO multi-receiver wiretap channel defined in* (4.44)*, where the Gaussian noise vectors' covariance matrices satisfy the semi-definite order in* (4.45)*, if the bound in* (4.48) *holds with all its constraints, then the vector realizations* $\mathbf{U}_j$ *of the auxiliary random variables* U_j *in* (4.48) *must satisfy the following covariance constraint:* $\mathbb{E}[\mathbf{U}_j\mathbf{U}_j^\top] \preceq \sum_{i=j}^{k}\mathbf{K}_i$.

Proof. The proof follows by adapting the same techniques used to prove Lemma A.9 to the vector nature of the degraded Gaussian MIMO multi-receiver wiretap channel. ■

List of Figures

2.1 Shannon's ciphering system 8
2.2 Discrete memoryless degraded wiretap channel 9
2.3 Wiretap channel with public and confidential messages 11
2.4 Wiretap channel with shared secret key 23
2.5 Two-receiver DM-WBC with public and confidential messages 32
2.6 Two-receiver DM-WBC with two individual confidential messages . . . 35

3.1 Secure bidirectional relaying in three-nodes network 48
3.2 Two-receiver DM-WBC with degraded message sets and message cognition 50
3.3 Individual secrecy for Shannon's ciphering system 54
3.4 Rate Splitting and Time Sharing for the Two-receiver DM-WBC 76

4.1 Multi-receiver DM-WBC 115
4.2 Joint and individual secrecy capacity regions of a two-receiver Gaussian SISO-WBC. 120
4.3 Degraded multi-receiver DM-WBC 121

5.1 Discrete memoryless arbitrary varying wiretap channel 159
5.2 Public-Confidential communication scenario over an AVWC 196

6.1 Arbitrarily varying channel with correlated jamming 214

A.1 Discrete memoryless channel model 217

Bibliography

[1] C. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech.*, vol. 28, pp. 656–715, October 1949.

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

[3] Deutsche Telekom AG Laboratories, "Next Generation Mobile Networks – (R)evolution in Mobile Communications," Technology Radar Edition III, Tech. Rep., 2010.

[4] G. Fettweis, H. Boche, T. Wiegand, and E. Z. et al., "The tactile internet," ITU-T Technology Watch, Tech. Rep., August 2014, available at:http://www.itu.int/oth/T2301000023/en.

[5] Y. Liang, H. V. Poor, and S. S. (Shitz), "Information theoretic security." *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2008.

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

[7] R. E. Blahut, *Cryptography and Secure Communication.* Cambridge: Cambridge University Press, 2014.

[8] R. F. Schaefer, H. Boche, A. Khisti, and H. V. Poor, *Information Theoretic Security and Privacy of Information Systems.* Cambridge University Press, 2017.

[9] R. F. Schaefer and H. Boche, "Physical layer service integration in wireless networks : Signal processing challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 147–156, May 2014.

[10] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[11] S. Leung-Yan-Cheong and M. Hellman, "The gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, July 1978.

[12] Y. K. Chia and A. E. Gamal, "Three-receiver broadcast channels with common and confidential messages," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2748–2765, May 2012.

[13] M. Bloch and J. N. Laneman, "On the secrecy capacity of arbitrary wiretap channels," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, September 2008, pp. 818–825.

[14] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity results for arbitrarily varying wiretap channels," *Information Theory, Combinatorics, and Search Theory*, pp. 123–144, 2013.

[15] E. A. Jorswieck, A. Wolf, and S. Gerbracht, "Secrecy on the physical layer in wireless networks," *Trends in Telecommunications Technologies*, pp. 413–435, March 2010.

[16] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering.* Cambridge University Press, 2011.

[17] J. G. Proakis and M. Salehi, *Digital Communications.* McGraw-Hill, 2008.

[18] A. R. Barron and A. Joseph, "Toward fast reliable communication at rates near capacity with gaussian noise," in *2010 IEEE International Symposium on Information Theory*, Austin, Texas, USA, June 2010, pp. 315–319.

[19] X. Zhou, L. Song, and Y. Zhang, *Physical Layer Security in Wireless Communications.* CRC Press, November 2013.

[20] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, January 1972.

[21] J. Körner and K. Marton, "A source network problem involving the comparison of two channels ii," *Trans. Colloquim Inform. Theory,*, August 1975.

[22] R. Fano, *Transmission of Information: A Statistical Theory of Communications.* M.I.T. Press, 1961.

[23] R. F. Wyrembelski and H. Boche, "Physical layer integration of private, common, and confidential messages in bidirectional relay networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3170–3179, September 2012.

[24] B. He and X. Zhou, "New physical layer security measures for wireless transmissions over fading channels," in *2014 IEEE Global Communications Conference*, Austin, Texas, USA, December 2014, pp. 722–727.

[25] R. Rajesh, S. M. Shah, and V. Sharma, "On secrecy above secrecy capacity," in *2012 IEEE International Conference on Communication Systems (ICCS)*, November 2012, pp. 70–74.

[26] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[27] U. Maurer, "Information-theoretic cryptography," in *Advances in Cryptology — CRYPTO' 99*, M. Wiener, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 47–65.

[28] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8077–8098, December 2013.

[29] R. F. Wyrembelski, M. Wiese, and H. Boche, "Strong secrecy in bidirectional broadcast channels with confidential messages," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 2, pp. 324–334, February 2013.

[30] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems of Information Transmission*, vol. 49, no. 1, pp. 73–98, 2013.

[31] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *IEEE International Symposium on Information Theory (ISIT)*, Honolulu, HI, USA, June 2014, pp. 601–605.

[32] R. Liu and W. Trappe, *Securing Wireless Communications at the Physical Layer*, 1st ed. Springer Publishing Company, Incorporated, 2009.

[33] J. L. Massey, "A simplified treatment of wyner's wiretap channels," in *Proceedings of* 21st *Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, October 1983, pp. 268–276.

[34] G. Kramer, "Lecture notes for multi-user information theory," Munich, Germany, April 2013.

[35] A. El Gamal and Y.-H. Kim, *Network Information Theory.* New York, NY, USA: Cambridge University Press, 2012.

[36] J. Barros and M. Bloch, "Strong secrecy for wireless channels," in *Proceedings of Third International Conference on Information Theoretic Security (ICITS)*, Calgary, Alberta, Canada, August 2008, pp. 40–53.

[37] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Lecture Notes in Computer Science*, vol. 1807. Springer-Verlag, 2000, pp. 351–368.

[38] I. Csiszár, "Almost independence and secrecy capacity," *Probl. Peredachi Inf.*, vol. 32, no. 1, pp. 48–57, 1996.

[39] M. Bloch and J. N. Laneman, "On the secrecy capacity of arbitrary wiretap channels," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA,, September 2008, pp. 818–825.

[40] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.

[41] A. J. Pierrot and M. R. Bloch, "Strongly secure communications over the two-way wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 595–605, September 2011.

[42] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, January 2005.

[43] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 163–179, March 1975.

[44] T. S. Han and S. Verdu, "Approximation theory of output statistics," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[45] A. Winter, "Secret, public and quantum correlation cost of triples of random variables," in *Proceeding of IEEE International Symposium on Information Theory, 2005 (ISIT)*, Adelaide, Australia, September 2005, pp. 2270–2274.

[46] J. Hou and G. Kramer, "Informational divergence approximations to product distributions," in *13th Canadian Workshop on Information Theory (CWIT)*, Toronto, Canada, June 2013, pp. 76–81.

[47] J. L. W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, vol. 30, pp. 175–193, 1906.

[48] H. Yamamoto, "Rate-distortion theory for the shannon cipher system," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 827–835, May 1997.

[49] N. Merhav, "Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2723–2734, June 2008.

[50] 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, December 2009.

[51] R. Ahlswede and N. Cai, *Transmission, Identification and Common Randomness Capacities for Wire-Tape Channels with Secure Feedback from the Decoder.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, vol. 4123, pp. 258–275.

[52] D. Gunduz, D. R. Brown, and H. V. Poor, "Secret communication with feedback," in *2008 International Symposium on Information Theory and Its Applications*, Auckland, New Zealand, December 2008, pp. 1–6.

[53] W. Kang and N. Liu, "Wiretap channel with shared key," in *IEEE Information Theory Workshop 2010 (ITW)*, Dublin, Ireland, September 2010, pp. 1–5.

[54] N. Cai, "The maximum error probability criterion, random encoder, and feedback, in multiple input channels," *Entropy*, vol. 16, no. 3, pp. 1211–1242, 2014.

[55] D. J. C. MacKay, *Information Theory, Inference & Learning Algorithms.* New York, NY, USA: Cambridge University Press, 2002.

[56] T. M. Cover, "Comments on broadcast channels," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2524–2530, October 1998.

[57] I. B. Gattegno, Z. Goldfeld, and H. H. Permuter, "Fourier-motzkin elimination for information theory (fme-it) package for matlab," http://www.ee.bgu.ac.il/~fmeit/index.html, September 2015, version R0.6.6.

[58] J. Wolfowitz, "The coding of messages subject to chance errors," *Illinois Journal of Mathematics*, vol. 1, no. 4, pp. 591–606, December 1957.

[59] M. Fekete, "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten," *Math Zeitung*, vol. 17, no. 1, pp. 228–249, 1923.

[60] H. Boche, R. F. Schaefer, and H. V. Poor, "On the continuity of the secrecy capacity of compound and arbitrarily varying wiretap channels," *Information Forensics and Security, IEEE Transactions on*, vol. 10, no. 12, pp. 2531–2546, Dec 2015.

[61] R. Liu, I. Maric, P. Spasojevic, 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.

[62] R. F. Wyrembelski and H. Boche, "How to achieve privacy in bidirectional relay networks," in *Proceeding of IEEE International Symposium on Information Theory, 2011 (ISIT)*, Saint-Petersburg, Russia, July 2011, pp. 1891–1895.

[63] 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, October 2009.

[64] Y. K. Chia and A. E. Gamal, "3-receiver broadcast channels with common and confidential messages," in *Proceeding of IEEE International Symposium on Information Theory, 2009 (ISIT)*, Seoul, Korea, June 2009, pp. 1849–1853.

[65] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 60–64, January 1977.

[66] S. Diggavi and D. Tse, "On opportunistic codes and broadcast codes with degraded message sets," in *IEEE Information Theory Workshop 2006 (ITW)*, Punta del Este, Uruguay, March 2006, pp. 227–231.

[67] V. Prabhakaran, S. Diggavi, and D. Tse, "Broadcasting with degraded message sets: A deterministic approach," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing*, Allerton Park, IL, USA, September 2007.

[68] C. Nair and Z. Wang, "On 3-receiver broadcast channels with 2-degraded message sets," in *Proceedings of IEEE International Symposium on Information Theory 2009 (ISIT)*, Seoul, Korea, June 2009, pp. 1844–1848.

[69] 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*, Toronto, Canada, July 2008, pp. 1706–1710.

[70] 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.

[71] C. Nair and A. El Gamal, "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, October 2009.

[72] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *The Annals of Mathematical Statistics*, vol. 30, no. 4, pp. 1229–1241, 1959.

[73] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "Secrecy rate region of the broadcast channel with an eavesdropper," in *Proceedings of the 47th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, September 2009, pp. 834–841.

[74] L.-C. Choo and K.-K. Wong, "Three-receiver broadcast channel with confidential messages," in *Proceedings of the International Symposium on Communication Theory and Applications*, Ambleside, Lake District, UK, July 2009.

[75] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP Journal of Wireless Communication Networks*, pp. 1–29, March 2009.

[76] L. C. Choo and K. K. Wong, "Another proof for the secrecy capacity of the k-receiver broadcast channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2142–2152, April 2013.

[77] Y. Liang, H. V. Poor, and S. S. (Shitz), "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.

[78] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, June 2008.

[79] T. Cover, "An achievable rate region for the broadcast channel," *IEEE Transactions on Information Theory*, vol. 21, no. 4, pp. 399–404, July 1975.

[80] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Transactions on Information Theory*, vol. 19, no. 2, pp. 197–207, March 1973.

[81] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredachi Inf.*, vol. 10, no. 3, pp. 3–14, 1974.

[82] J. Körner and K. Marton, "Comparison of two noisy channels," *Topics in Information Theory, Coll. Math. Soc. J. Bolyai*, no. 16, pp. 411–423, 1977.

[83] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 166–169, March 1979.

[84] E. Ekrem and S. Ulukus, "Multi-receiver wiretap channel with public and confidential messages," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2165–2177, April 2013.

[85] L. Wang, E. Sasoglu, B. Bandemer, and Y. H. Kim, "A comparison of superposition coding schemes," in *2013 IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013, pp. 2970–2974.

[86] E. van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 180–190, Mar 1975.

[87] T. Berger, *Rate Distortion Theory and Data Compression.* Vienna: Springer Vienna, 1975, pp. 1–39.

[88] A. E. Gamal and E. van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 120–122, January 1981.

[89] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[90] M. Fogiel, *The Handbook of Electrical Engineering*, ser. Handbook Series. Research & Education Association, 1996.

[91] A. J. Stam, "Some inequalities satisfied by the quantities of information of fisher and shannon," *Information and Control*, vol. 2, no. 2, pp. 101 – 112, June 1959.

[92] N. Blachman, "The convolution inequality for entropy powers," *IEEE Transactions on Information Theory*, vol. 11, no. 2, pp. 267–271, April 1965.

[93] R. Liu, T. Liu, H. V. Poor, and S. S. (Shitz), "A vector generalization of costa's entropy-power inequality with applications," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1865–1879, April 2010.

[94] E. Ekrem and S. Ulukus, "The secrecy capacity region of the gaussian mimo multi-receiver wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2083–2114, April 2011.

[95] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 279–280, March 1974.

[96] A. E. Gamal, "Lecture notes on multiple user information theory," California, USA, 2010.

[97] D. Guo, S. S. (Shitz), and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[98] D. Guo, Y. Wu, S. S. Shitz, and S. Verdu, "Estimation in gaussian noise: Properties of the minimum mean-square error," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2371–2385, April 2011.

[99] O. Rioul, "Information theoretic proofs of entropy power inequalities," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 33–55, January 2011.

[100] M. Costa, "A new entropy power inequality," *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 751–760, November 1985.

[101] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy capacity region of the gaussian mimo broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2673–2682, May 2013.

[102] H. Weingarten, Y. Steinberg, and S. S. (Shitz), "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, September 2006.

[103] T. Liu and S. S. (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, June 2009.

[104] E. Ekrem and S. Ulukus, "Gaussian mimo broadcast channels with common and confidential messages," in *Proceedings of IEEE International Symposium on Information Theory 2010 (ISIT)*, Austin, Texas, USA, June 2010, pp. 2583–2587.

[105] ——, "Capacity region of gaussian mimo broadcast channels with common and confidential messages," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5669–5680, September 2012.

[106] R. F. Wyrembelski and H. Boche, "Physical layer integration of private, common, and confidential messages in bidirectional relay networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3170–3179, September 2012.

[107] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Multiple access channels with generalized feedback and confidential messages," in *IEEE Information Theory Workshop 2007 (ITW)*, Lake Tahoe, California, USA, September 2007, pp. 608–613.

[108] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, March 2008.

[109] M. H. Yassaee and M. R. Aref, "Multiple access wiretap channels with strong secrecy," in *IEEE Information Theory Workshop 2010 (ITW)*, Cairo,Egypt, August 2010, pp. 1–5.

[110] E. Tekin and A. Yener, "The gaussian multiple access wire-tap channel," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5747–5755, December 2008.

[111] X. He and A. Yener, "A new outer bound for the secrecy capacity region of the gaussian two-way wiretap channel," in *IEEE International Conference on Communications 2010 (ICC)*, Cape Town, South Africa, May 2010, pp. 1–5.

[112] A. E. Gamal, O. O. Koyluoglu, M. Youssef, and H. E. Gamal, "Achievable secrecy rate regions for the two-way wiretap channel," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8099–8114, December 2013.

[113] A. J. Pierrot and M. R. Bloch, "Strongly secure communications over the two-way wiretap channel," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 595–605, September 2011.

[114] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 454–458, January 2008.

[115] Y. Wu, "Broadcasting when receivers know some messages a priori," in *Proceedings of IEEE International Symposium on Information Theory 2007 (ISIT)*, Nice, France, June 2007, pp. 1141–1145.

[116] G. Kramer and S. S. (Shitz), "Capacity for classes of broadcast channels with receiver side information," in *IEEE Information Theory Workshop 2007 (ITW)*, Lake Tahoe, California, USA, September 2007, pp. 313–318.

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

[118] T. J. Oechtering, M. Wigger, and R. Timo, "Broadcast capacity regions with three receivers and message cognition," in *Proceedings of IEEE International Symposium on Information Theory 2012 (ISIT)*, MIT, Cambridge, MA, USA, July 2012, pp. 388–392.

[119] R. F. Wyrembelski, A. Sezgin, and H. Boche, "Secrecy in broadcast channels with receiver side information," in 45th *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, November 2011, pp. 290–294.

[120] A. S. Mansour, R. F. Schaefer, and H. Boche, "Joint and individual secrecy in broadcast channels with receiver side information," in *IEEE* 15th *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Toronto, Canada, June 2014, pp. 369–373.

[121] ——, "Secrecy measures for broadcast channels with receiver side information: Joint vs individual," in *Information Theory Workshop (ITW)*, Hobart, Tasmania, Australia, November 2014, pp. 426–430.

[122] ——, "Capacity regions for broadcast channels with degraded message sets and message cognition under different secrecy constraints," November 2014, submitted to IEEE Transactions of Information Theory.

[123] O. O. Koyluoglu, Y. Chen, and A. Sezgin, "Broadcast channel with receiver side information: Achieving individual secrecy," in *Proc. of International Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, February 2014, pp. 67–70.

[124] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "On the achievable individual-secrecy rate region for broadcast channels with receiver side information," in *IEEE International Symposium on Information Theory (ISIT)*, Honolulu, HI, USA, June 2014, pp. 26–30.

[125] ——, "On the individual secrecy for gaussian broadcast channels with receiver side information," in *The IEEE International Conference on Communications (ICC 2015)*, London, UK, June 2015, pp. 503–508.

[126] ——, "Individual secrecy for broadcast channels with receiver side information," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4687–4708, July 2017.

[127] Z. Goldfeld, G. Kramer, H. H. Permuter, and P. Cuff, "Strong secrecy for cooperative broadcast channels," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 469–495, January 2017.

[128] N. A. Weiss, P. T. Holmes, and M. Hardy, *A Course in Probability.* Pearson Addison Wesley, 2005.

[129] G. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., ser. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[130] I. Schneider, *Statisticians of the Centuries: Jakob Bernoulli*, C. C. Heyde, E. Seneta, P. Crépel, S. E. Fienberg, and J. Gani, Eds. New York, NY: Springer New York, 2001.

[131] E. C. Song, P. W. Cuff, and H. V. Poor, "A rate-distortion based secrecy system with side information at the decoders," in 52nd *Annual Allerton Conference on Communication, Control, and Computing*, Monticell, IL, USA, September 2014, pp. 755–762.

[132] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7071–7096, November 2013.

[133] I. Rozanov and R. Silverman, *Probability Theory: A Concise Course*, ser. Dover Books on Mathematics. Dover Publications, 1977.

[134] G. Kramer, "Topics in multi-user information theory," *Found. Trends Commun. Inf. Theory*, vol. 4, no. 4-5, pp. 265–444, April 2008.

[135] I. Bárány and R. Karasev, "Notes about the carathéodory number," *Discrete & Computational Geometry*, vol. 48, no. 3, pp. 783–792, October 2012.

[136] A. Feinstein, "A new basic theorem of information theory," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 2–22, September 1954.

[137] C. Nair and Z. V. Wang, "The capacity region of the three receiver less noisy broadcast channel," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4058–4062, July 2011.

[138] A. S. Mansour, R. F. Schaefer, and H. Boche, "The individual secrecy capacity of degraded multi-receiver wiretap broadcast channels," in *IEEE International Conference on Communications (ICC)*, London, United Kingdom, June 2015, pp. 4181–4186.

[139] ——, "The individual secrecy capacity of the Gaussian SISO and degraded Gaussian MIMO multi-receiver wiretap channel," in *IEEE* 16th *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, June 2015, pp. 365–369.

[140] ——, "On the individual secrecy capacity regions of the general, degraded, and gaussian multi-receiver wiretap broadcast channel," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2107–2122, September 2016.

[141] M. Benammar and P. Piantanida, "On the secrecy capacity region of the wiretap broadcast channel," in *IEEE Information Theory Workshop (ITW)*, Hobart, Tasmania, Australia, November 2014, pp. 421–425.

[142] ——, "Secrecy capacity region of some classes of wiretap broadcast channels," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5564–5582, October 2015.

[143] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "Individual secrecy for the broadcast channel," in *International Symposium on Information Theory and Its Applications (ISITA)*, Monterey, California, USA, October 2016, pp. 611–615.

[144] ——, "Individual secrecy for the broadcast channel," *IEEE Transactions on Information Theory*, vol. 63, no. 9, pp. 5981–5999, September 2017.

[145] A. Memon, *Advances in Computers, Volume 101*, 1st ed. Orlando, FL, USA: Academic Press, Inc., 2016.

[146] C. Nair and V. W. Zizhou, "On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels," in *Information Theory and Applications Workshop*, San Diego, California, USA, January 2008, pp. 226–229.

[147] Y. Liang, G. Kramer, and H. V. Poor, "On the equivalence of two achievable regions for the broadcast channel," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 95–100, January 2011.

[148] A. A. Gohari, A. E. Gamal, and V. Anantharam, "On an outer bound and an inner bound for the general broadcast channel," in *IEEE International Symposium on Information Theory*, Austin, Texas, USA, June 2010, pp. 540–544.

[149] O. Oyerinde, *Channel Estimation for Wireless Communication Systems.* Lambert Academic Publishing, 2011.

[150] M. R. Bloch and J. N. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1840–1849, September 2013.

[151] X. He and A. Yener, "Providing secrecy when the eavesdropper channel is arbitrarily varying: A case for multiple antennas," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, September 2010, pp. 1228–1235.

[152] Y. Liang, G. Kramer, H. V. Poor, and S. S. (Shitz), "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking*, no. 1, p. 142374, October 2009.

[153] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, USA, September 2009, pp. 1069–1075.

[154] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," *The Annals of Mathematical Statistics*, vol. 31, no. 3, pp. 558–567, 09 1960.

[155] J. Wolfowitz, "Simultaneous channels," *Archive for Rational Mechanics and Analysis*, vol. 4, pp. 371–386, January 1959.

[156] H. Weingarten, T. Liu, S. S. (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5011–5023, November 2009.

[157] E. Ekrem and S. Ulukus, "Degraded compound multi-receiver wiretap channels," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5681–5698, September 2012.

[158] J. Kiefer and J. Wolfowitz, "Channels with arbitrarily varying channel probability functions," *Information and Control*, vol. 5, no. 1, pp. 44 – 54, 1962.

[159] M. Wiese, J. Nötzel, and H. Boche, "A channel under simultaneous jamming and eavesdropping attack – correlated random coding capacities under strong secrecy criteria," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3844–3862, 2016.

[160] R. Ahlswede and J. Wolfowitz, "Correlated decoding for channels with arbitrarily varying channel probability functions," *Information and Control*, vol. 14, no. 5, pp. 457 – 473, 1969.

[161] ——, "The capacity of a channel with arbitrarily varying channel probability functions and binary output alphabet," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 15, no. 3, pp. 186–194, September 1970.

[162] N. S. Kambo and S. Singh, "The capacities of certain special channels with arbitrarily varying channel probability functions," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 29, no. 4, pp. 331–344, December 1974.

[163] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 44, no. 2, pp. 159–175, 1978.

[164] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 31, no. 1, pp. 42–48, January 1985.

[165] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: positivity, constraints," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, March 1988.

[166] J. Nötzel, M. Wiese, and H. Boche, "The arbitrarily varying wiretap channel – secret randomness, stability, and super-activation," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3504–3531, June 2016.

[167] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Arbitrarily varying wiretap channels with type constrained states," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7216–7244, December 2016.

[168] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, October 1998.

[169] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[170] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[171] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," *IEEE Transactions on Information Theory*, vol. 32, no. 5, pp. 621–629, September 1986.

[172] H. Boche, R. F. Schaefer, and H. V. Poor, "On the continuity of the secrecy capacity of compound and arbitrarily varying wiretap channels," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2531–2546, December 2015.

[173] M. Hayashi, S. Ishizaka, A. Kawachi, G. Kimura, and T. Ogawa, *Introduction to Quantum Information Science*, 1st ed. Springer Publishing Company, Incorporated, 2016.

[174] D. Leung and G. Smith, "Continuity of quantum channel capacities," *Communications in Mathematical Physics*, vol. 292, no. 1, pp. 201–215, November 2009.

[175] K. Matsumoto, T. Shimono, and A. Winter, "Remarks on additivity of the holevo channel capacity and of the entanglement of formation," *Communications in Mathematical Physics*, vol. 246, no. 3, pp. 427–442, April 2004.

[176] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, September 1956.

[177] W. Haemers, "On some problems of lovasz concerning the shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 231–232, March 1979.

[178] N. Alon, "The shannon capacity of a union," *Combinatorica*, vol. 18, no. 3, pp. 301–310, 1998.

[179] P. Keevash and E. Long, "On the normalized shannon capacity of a union," *Combinatorics, Probability and Computing*, vol. 25, no. 5, pp. 766–767, 2016.

[180] H. Boche, R. F. Schaefer, and H. V. Poor, "Identification over channels with feedback: Discontinuity behavior and super-activation," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 256–260.

[181] ——, "Identification capacity of channels with feedback: Discontinuity behavior, super-activation, and turing computability," January 2018, submitted to IEEE Transactions of Information Theory.

[182] R. Ahlswede, "A note on the existence of the weak capacity for channels with arbitrarily varying channel probability functions and its relation to shannon's zero error capacity," *Ann. Math. Statist.*, vol. 41, no. 3, pp. 1027–1033, June 1970.

[183] L. Lovasz, "On the shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, January 1979.

[184] R. F. Schaefer, H. Boche, and H. V. Poor, "Super-activation as a unique feature of arbitrarily varying wiretap channels," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 3077–3081.

[185] H. Boche and R. F. Schaefer, "Capacity results and super-activation for wiretap channels with active wiretappers," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 9, pp. 1482–1496, September 2013.

[186] G. Giedke and M. M. Wolf, "Quantum communication: Super-activated channels," *Nature Photonics*, vol. 5, no. 10, pp. 578–580, October 2011.

[187] R. F. Schaefer, H. Boche, and H. V. Poor, "Super-activation as a unique feature of secure communication in malicious environments," *Information*, vol. 7, no. 24, 2016.

[188] A. S. Mansour, H. Boche, and R. F. Schaefer, "List decoding for arbitrarily varying wiretap channels," in *IEEE Conference on Communications and Network Security (CNS)*, Philadelphia, PA, USA, October 2016, pp. 611–615.

[189] A. S. Mansour and H. Boche, "Strong secrecy capactiy of arbitrarily varying wiretap channels with finite coordination resources," in *IEEE Global Conference on Signal and Information Processing*, Washington, D.C., USA, December 2016, pp. 1002–1006.

[190] A. S. Mansour, H. Boche, and R. F. Schaefer, "Stabilizing the secrecy capacity of the arbitrarily varying wiretap channel and transceiver synchronization using list decoding," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sapporo, Japan, July 2017, pp. 1–5.

[191] ——, "The secrecy capacity of the arbitrarily varying wiretap channel under list decoding," June 2017, submitted for publication.

[192] A. S. Mansour, R. F. Schaefer, and H. Boche, "Secrecy capacity under list decoding for a channel with a passive eavesdropper and an active jammer," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Alberta, Canada, April 2018.

[193] P. Elias, "List decoding for noisy channels," *1957 IRE Wescon convention record*, vol. 2, no. 335, pp. 94–104, September 1957.

[194] V. M. Blinovskii, P. Narajan, and M. S. Pinsker, "Capacity of the arbitrarily varying channel under list decoding," *Problems Inform. Transmission*, vol. 31, no. 2, pp. 99–113, 1995.

[195] B. L. Hughes, "The smallest list for the arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 803–815, May 1997.

[196] S. Nitinawarat, "On the deterministic code capacity region of an arbitrarily varying multiple-access channel under list decoding," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2683–2693, May 2013.

[197] H. Boche and R. F. Schaefer, "Arbitrarily varying multiple access channels with conferencing encoders: List decoding and finite coordination resources," *Adv. in Math. of Comm.*, vol. 10, no. 2, pp. 333–354, 2016.

[198] ——, "Arbitrarily varying wiretap channels with finite coordination resources," in *2014 IEEE International Conference on Communications Workshops (ICC)*, Sydney, Australia, June 2014, pp. 746–751.

[199] H. Boche, R. F. Schaefer, and H. V. Poor, "Characterization of super-additivity and discontinuity behavior of the capacity of arbitrarily varying channels under list decoding," in *2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, June 2017, pp. 2820–2824.

[200] ——, "Analytical properties of shannon's capacity of arbitrarily varying channels under list decoding: Super-additivity and discontinuity behavior," January 2018, submitted for publication.

[201] A. S. Mansour, R. F. Schaefer, and H. Boche, "The deterministic and correlated random public-confidential capacity regions of the arbitrarily varying wiretap channel," May 2018, submitted for publication.

[202] R. F. Schaefer and H. Boche, "Robust broadcasting of common and confidential messages over compound channels: Strong secrecy and decoding performance," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 10, pp. 1720–1732, October 2014.

[203] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, vol. 9, no. 3, pp. 177–183, 1973.

[204] ——, "The capacity of the quantum channel with general signal states," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 269–273, January 1998.

[205] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Semantic-security capacity for wiretap channels of type ii," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3863–3879, July 2016.

[206] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, vol. 9, no. 3, pp. 177–183, 1973.

[207] ——, "The capacity of the quantum channel with general signal states," *Information Theory, IEEE Transactions on*, vol. 44, no. 1, pp. 269–273, Jan 1998.

[208] H. Boche, N. Cai, and J. Nötzel, "The classical-quantum channel with random state parameters known to the sender," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 19, p. 195302, 2016.

[209] A. D. Sarwate and M. Gastpar, "Channels with nosy "noise"," in *IEEE International Symposium on Information Theory*, Nice, France, June 2007, pp. 996–1000.

[210] A. D. Sarwate, "Robust and adaptive communication under uncertain interference," Ph.D. dissertation, EECS Department, University of California, Berkeley, July 2008.

[211] ——, "An avc perspective on correlated jamming," in *2012 International Conference on Signal Processing and Communications (SPCOM)*, July 2012, pp. 1–5.

[212] R. Ahlswede, "The maximal error capacity of arbitrarily varying channels for constant list sizes," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1416–1417, Jul 1993.

[213] A. D. Sarwate and M. Gastpar, "List-decoding for the arbitrarily varying channel under state constraints," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1372–1384, March 2012.

[214] H. Boche, M. Cai, and N. Cai, "Message transmission over classical quantum channels with a jammer with side information: Message transmission capacity and resources," *CoRR*, vol. abs/1801.10550, January 2018.

[215] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 99th ed. Wiley-Interscience, August 1991.

[216] R. Wieczorkowski and P. Grzegorzewski, "Entropy estimators; improvements and comparisons," *Communications in Statistics - Simulation and Computation*, vol. 28, no. 2, pp. 541–567, 1999.

[217] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, pp. pp. 379–423, July 1948.

[218] B. Behmardi, R. Raich, and A. O. Hero, "Entropy estimation using the principle of maximum entropy," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 2008–2011.

[219] C. Schmitz, A. Schmeink, and R. Mathar, "Estimating mutual information in genetics," in *8th International Symposium on Wireless Communication Systems (ISWCS)*, November 2011, pp. 246–250.

[220] R. W. Yeung, *Discrete Memoryless Channels.* Boston, MA: Springer US, 2008, pp. 137–182.

[221] V. B. Balakirsky, *Algebraic Coding.* Springer Berlin Heidelberg, 1992, ch. Coding theorem for discrete memoryless channels with given decision rule, pp. 142–150.

[222] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65–103, 1967.

[223] ——, "Lower bounds to error probability for coding on discrete memoryless channels. II," *Information and Control*, vol. 10, no. 5, pp. 522–552, 1967.

[224] E. Torgersen, *Comparison of Statistical Experiments*, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1991.

[225] J. COHEN, J. Kempermann, and G. Zbaganu, *Comparisons of Stochastic Matrices with Applications in Information Theory, Statistics, Economics and Population.* Birkhäuser Boston, 1998.

[226] A. E. Gamal, "Broadcast channels with and without feedback," in 11th *Asilomar Conference on Circuits, Systems and Computers*, vol. 1807, 1977, pp. 180–183.

[227] N. J. Beaudry and R. Renner, "An intuitive proof of the data processing inequality," *Quantum Information and Computation*, vol. 12, July 2011.

[228] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.

[229] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, 12 1952.

[230] A. Markov, "On certain applications of algebraic continued fractions," Ph.D. dissertation, St. Petersburg, 1884.

[231] M. Pinsker, *Information and information stability of random variables and processes.* Holden-Day, 1964, originally published in Russian in 1960.

[232] A. A. Fedotov, P. Harremoes, and F. Topsoe, "Refinements of pinsker's inequality," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1491–1498, June 2003.

[233] M. Basu and T. K. Ho, *Data Complexity in Pattern Recognition (Advanced Information and Knowledge Processing)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[234] I. Csiszar and Z. Talata, "Context tree estimation for not necessarily finite memory processes, via bic and mdl," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1007–1016, March 2006.

[235] S. Verdú, "Total variation distance and the distribution of relative information," in *Information Theory and Applications Workshop (ITA)*, February 2014, pp. 1–3.

[236] I. Sason and S. Verdú, "f -divergence inequalities," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 5973–6006, November 2016.

[237] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[238] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1978.

[239] I. Csiszar, "The method of types," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2505–2523, October 1998.