Physical-Layer Security in Wireless Communication Systems

by

Ghadamali Bagheri-Karam

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2010

AUTHOR'S DECLARATION

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

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

Ghadamali Bagheri-Karam

Abstract

The use of wireless networks has grown significantly in contemporary times, and continues to develop further. The broadcast nature of wireless communications, however, makes them particularly vulnerable to eavesdropping. Unlike traditional solutions, which usually handle security at the application layer, the primary concern of this dissertation is to analyze and develop solutions based on coding techniques at the physical-layer.

First, in chapter 2, we consider a scenario where a source node wishes to broadcast two confidential messages to two receivers, while a wire-tapper also receives the transmitted signal. This model is motivated by wireless communications, where individual secure messages are broadcast over open media and can be received by any illegitimate receiver. The secrecy level is measured by the equivocation rate at the eavesdropper. We first study the general (non-degraded) broadcast channel with an eavesdropper, and present an inner bound on the secrecy capacity region for this model. This inner bound is based on a combination of random binning, and the Gelfand-Pinsker binning. We further study the situation in which the channels are degraded. For the degraded broadcast channel with an eavesdropper, we present the secrecy capacity region. Our achievable coding scheme is based on Cover's superposition scheme and random binning. We refer to this scheme as the Secret Superposition Scheme. Our converse proof is based on a combination of the converse proof of the conventional degraded broadcast channel and Csiszar Lemma. We then assume that the channels are Additive White Gaussian Noise and show that the Secret Superposition Scheme with Gaussian codebook is optimal. The converse proof is based on Costa's entropy power inequality. Finally, we use a broadcast strategy for the slowly fading wire-tap channel when only the eavesdropper's channel is fixed and known at the transmitter. We derive the optimum power allocation for the coding layers, which maximizes the total average rate.

Second, in chapter 3 , we consider the Multiple-Input-Multiple-Output (MIMO) scenario of a broadcast channel where a wiretapper also receives the transmitted signal via another MIMO channel. First, we assume that the channels are degraded and the wiretapper has the worst channel. We establish the capacity region of this scenario. Our achievability scheme is the Secret Superposition Coding. For the outerbound, we use notion of the enhanced

channels to show that the secret superposition of Gaussian codes is optimal. We show that we only need to enhance the channels of the legitimate receivers, and the channel of the eavesdropper remains unchanged. We then extend the result of the degraded case to a non-degraded case. We show that the secret superposition of Gaussian codes, along with successive decoding, cannot work when the channels are not degraded. We develop a Secret Dirty Paper Coding scheme and show that it is optimal for this channel. We then present a corollary generalizing the capacity region of the two receivers case to the case of multiple receivers. Finally, we investigate a scenario which frequently occurs in the practice of wireless networks. In this scenario, the transmitter and the eavesdropper have multiple antennae, while both intended receivers have a single antenna (representing resource limited mobile units). We characterize the secrecy capacity region in terms of generalized eigenvalues of the receivers' channels and the eavesdropper's channel. We refer to this configuration as the MISOME case. We then present a corollary generalizing the results of the two receivers case to multiple receivers. In the high SNR regime, we show that the capacity region is a convex closure of rectangular regions.

Finally, in chapter 4, we consider a K-user secure Gaussian Multiple-Access-Channel with an external eavesdropper. We establish an achievable rate region for the secure discrete memoryless MAC. Thereafter, we prove the secrecy sum capacity of the degraded Gaussian MIMO MAC using Gaussian codebooks. For the non-degraded Gaussian MIMO MAC, we propose an algorithm inspired by the interference alignment technique to achieve the largest possible total Secure-Degrees-of-Freedom . When all the terminals are equipped with a single antenna, Gaussian codebooks have shown to be inefficient in providing a positive S-DoF. Instead, we propose a novel secure coding scheme to achieve a positive S-DoF in the single antenna MAC. This scheme converts the single-antenna system into a multiple-dimension system with fractional dimensions. The achievability scheme is based on the alignment of signals into a small sub-space at the eavesdropper, and the simultaneous separation of the signals at the intended receiver. We use tools from the field of Diophantine Approximation in number theory to analyze the probability of error in the coding scheme. We prove that the total S-DoF of $\frac{K-1}{K}$ can be achieved for almost all channel gains. For the other channel gains, we propose a multi-layer coding scheme to achieve a positive S-DoF. As a function of channel gains, therefore, the achievable S-DoF is discontinued.

Acknowledgements

This dissertation is the result of years of research that has been done since I have come to the Coding and Signal Transmission (CST) Laboratory. I have had the rewardable opportunity to work with a great number of people whose contribution to the research and the making of the thesis deserves special mention. It is a pleasure to convey my gratitude to them in my humble acknowledgment.

First and foremost, I would like to heartily to thank my supervisor, Professor Amir K. Khandani, whose encouragement, guidance and support from the initial to the final stages has enabled me to develop a deeper understanding of the subject. It was a great opportunity for me to work with such a knowledgeable supervisor.

I am also thankful to the members of my dissertation committee, Professors Hesham El Gamal, Liang-Liang Xie, Murat Uysal, Kirsten Morris, and Oleg Michailovich for having accepted to take the time out of their busy schedules to read my thesis and to provide me with their comments and suggestions.

I offer my regards and blessings to all CST members, particularly Seyed Abolfazl Motahari, Hossein Bagheri, and Soroush Akhlaghi Esfahani, who supported me in all respects during the completion of the project.

Finally, I am grateful to my wife, Tahereh, for her patience and support. Without her this work would never have come into existence, literally.

To my Parents and Parents in-law,

To my beloved wife, Tahereh,

and

To my children.

Contents

List of Tables

List of Figures

List of Abbreviations

PHY	– – – – –	Physical-Layer
CSI	– – – – –	Channel-State-Information
SNR	– – – – –	Signal-to-Noise Ratio
AWGN	– – – – –	Additive-White-Gaussian-Noise
S-DoF	– – – – –	Secure Degrees-of-Freedom
DPC	– – – – –	Dirty-Paper-Coding
SDPC	– – – – –	Secure Dirty-Paper-Coding
BCE	– – – – –	Broadcast-Channel with an Eavesdropper
G-BCE	– – – – –	Gaussian Broadcast-Channel with an Eavesdropper
MIMO	– – – – –	Multiple-Input-Multiple-Output
MISOME	– – – – –	Multiple-Input-Single-Output-Multiple-Eavesdropper
SADBC	– – – – –	Secure-Aligned-Degraded MIMO Broadcast-Channel
SAMBC	– – – – –	Secure-Aligned MIMO Broadcast-Channel
SGMBC	– – – – –	Secure-Gaussian MIMO Broadcast-Channel
MAC	– – – – –	Multiple-Access-Channel

Notation

Boldface Upper-Case Letters	– – –	Matrices
Boldface Lower-Case Letters	– – –	Vectors
$\mathbf{A}^\dagger$	– – –	Hermitian transpose of $\mathbf{A}$
$\mathbf{A^{-1}}$	– – –	Inverse of $\mathbf{A}$
$\lvert A\rvert$	– – –	Determinant of the matrix $\mathbf{A}$
$Tr\{\mathbf{A}\}$	– – –	Trace of $\mathbf{A}$
$\mathbf{A} \succeq 0$	– – –	Matrix $\mathbf{A}$ is positive semi-definite
$\mathbf{U} \succeq \mathbf{V}$	– – –	$\mathbf{U} - \mathbf{V}$ is a positive semi-definite matrix
$\mathbf{I}$	– – –	The identity matrix
$\mathcal{K}$	– – –	The set $\{1, 2, ..., K\}$
x^n	– – –	The vector $(x_1, x_2, ..., x_n)$
x^{i-1}	– – –	The vector $(x_1, x_2, ..., x_{i-1})$
$\widetilde{x}^i$	– – –	The vector $(x_i, x_{i+1}, ..., x_n)$
$\mathcal{N}(m, \sigma^2)$	– – –	Gausian distribution with mean m and variance σ^2
$\mathbb{R}$	– – –	The set of real numbers
$\mathbb{R}^n$	– – –	The n-dimensional Euclidean space
$\mathbb{Q}$	– – –	The set of rational numbers
$\mathbb{N}$	– – –	The set of nonnegative integers
$E[X]$	– – –	The expectation of the random variable X
(m, n)	– – –	The greatest common divisor of integers m and n
$\lVert\mathcal{X}\rVert$	– – –	Cardinality of set $\mathcal{X}$
$\bigcup$	– – –	The union of sets

Chapter 1

Introduction

Security protocols are the most critical elements involved in enabling the growth of the wide range of wireless data networks and applications. The broadcast nature of wireless communications, however, makes them particularly vulnerable to eavesdropping. With the proliferation of more complex modern infrastructure systems, there is an increasing need for secure communication solutions. Cryptography is a traditional field that provides *computationally-secure* protocols at the application layer. The goal of cryptography has recently been diversified from providing the critical confidentiality service, to other issues including authentication, key exchange and management, digital signature, and more. Unlike the cryptographic approaches, the recently reintroduced *physical-layer* security aims to develop effective secure communication schemes exploiting the properties of the physical layer. This new paradigm can strength the security of existing systems by introducing a level of *information-theoretic* security which has provable security, as compared with computational security. Note that the physical-layer security has a complementary role and can integrated with existing security solutions to enhance the total level of security for communication systems.

1.1 Security Issues in Wireless Communications

Security issues arising in communication networks can be classified into four main areas including confidentiality, integrity, authentication, and non-repudiation. Confidentiality guarantees that legitimate recipients successfully obtain their intended information while the information are protected against eavesdropping. Integrity provides communicating parties with the assurance that a message is not modified during its transmission. Authentication ensures that a recipient of information be able to identify the sender of the information. Non-repudiation protects against denial by one of the entities involved in a communication

of having participated in all or part of the communication.

In addition to standard security issues, the inherent receptiveness of wireless signals has imposed extra security vulnerabilities on the wireless communication systems. Wireless channels are susceptible to channel jamming. An attacker can easily jam physical communication channels and prevent users from accessing the network. The goal of a jammer, here, is to interrupt the communication traffic instead of intercepting the transmitted information. Secondly, without a proper authentication mechanism, an attacker can have an unauthorized access to the network resources and bypass the security infrastructures. Finally, due to the openness of wireless media, eavesdropping can be performed easily. In particular, legitimate users in a network can be regarded as potential eavesdroppers.

To solve the aforementioned security issues, a layered protocol approach has been considered by many wireless service providers. Protocol layering is a common technique used to simplify networking designs, by dividing them into functional layers, and assigning protocols to perform each layer's task. Figure 1.1 illustrates the various layers considered in a typical wireless communication protocol, and indicates their specific purposes. For instance, channel coding is implemented at the Physical (PHY) layer to provide an error-free medium for the above layers, and admission control is handled at the Medium Access Control layer. Note that the design of modern communication protocols does not follow a strict layered approach. The protocol layering, however, is a convenient conceptual representation that we use in this dissertation.

Security solutions are handled in different layers; for examples, spread-spectrum modulation techniques are used at the PHY layer to mitigate channel jamming, authentication mechanisms are implemented at the link layer to prevent unauthorized access, and cryptographic message encryption is performed at the application layer to protect the messages against eavesdropping. Therefore, channel jamming and unauthorized access, which are vulnerabilities at the PHY layer and link layer, respectively, are performed by security solutions at their layers. Eavesdropping, however, which is also a PHY layer vulnerability is traditionally handled by a solution at the application layer. A natural question to ask is whether the physical phenomena occurring at the PHY layer can be exploited against eavesdropping.

1.2 Physical-Layer Security

To illustrate the general concept of physical layer security, consider the example of a three node wireless network in Figure 1.2. In this configuration, the communication between terminal A and B is being eavesdropped by terminal E. The communication channel between the legitimate users is called the main channel, whereas the communication channel between

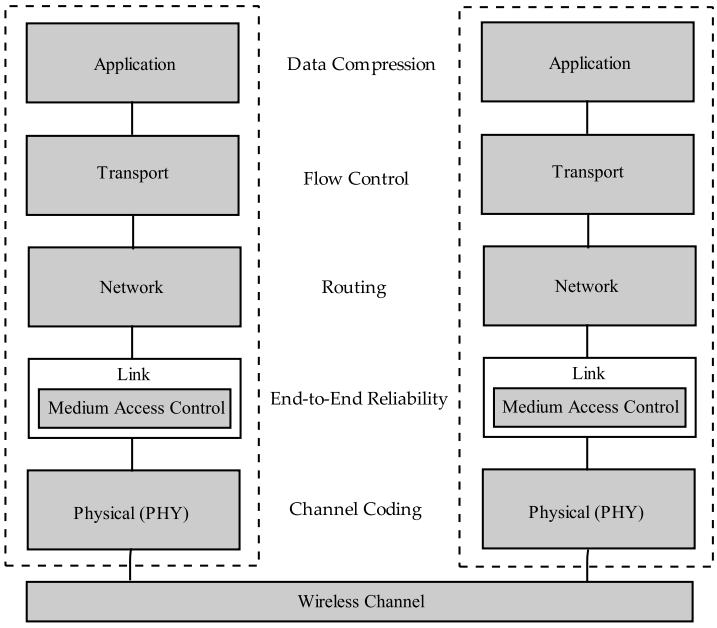

Figure 1.1: Layerd Protocol Architecture

terminals A and E is referred to as the eavesdroppers's channel. When the terminals B and C are not collocated, the observed signals by the legitimate receiver and the eavesdropper are usually different. The most notable effects for wireless communications are fading and path-loss. Fading is a self-interference phenomena that results from the multi-path propagation of the signals, while path-loss is simply the attenuation of signal amplitude with distance. If the transmission distance over the main channel is much smaller than the transmission distance over the eavesdropper's channel, the detected signal at terminal B is much stronger than at terminal E. For instance, if A broadcasts a video stream, the signal obtained by E is significantly degraded compared to the one received by B. This degradedness can be used at terminal A to prevent E from understanding the content of the video stream. Cryptographic security solutions implemented at application layer completely ignore these effects and operate as if the eavesdropper channel is an error-free channel. In contrast, the key idea of physical-layer security is to explicitly consider differences at the PHY layer to better protect the messages exchanged over the main channel.

In this dissertation, we study the physical-layer security from an information-theoretic and coding perspective, but we acknowledge that the scope of physical-layer security goes well beyond these considerations. In particular, we do not consider a large class of techniques that aim to modify the PHY layer to impair potential eavesdropping. Examples of such techniques are coded-division multiple-access signaling, which converts the signals into a

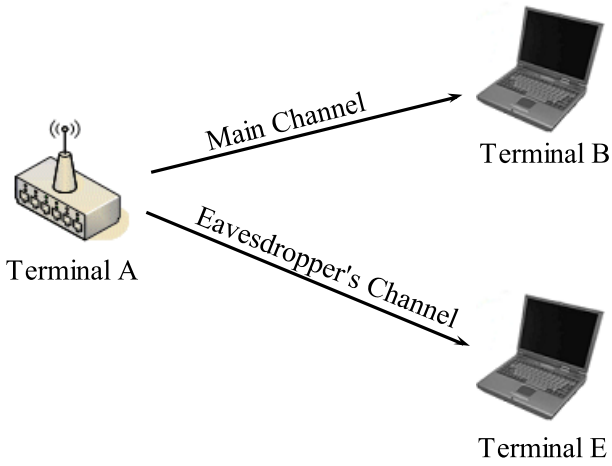

Figure 1.2: Illustration of Eavesdropping Scenario in Wireless Networks

noise level appearance, and beamforming with smart antenna, which essentially prevents the eavesdropper from detecting the transmitted signals.

Historically, the notion of information theoretic secrecy in communication systems was first introduced by Shannon in [1]. The information theoretic secrecy requires that the received signal of the eavesdropper not provide any information about the transmitted messages. Shannon considered a pessimistic situation where both the intended receiver and the eavesdropper have direct access to the transmitted signal (which is called ciphertext). Under these circumstances, he proved a negative result showing that perfect secrecy can be achieved only when the entropy of the secret key is greater than, or equal to, the entropy of the message. In modern cryptography, all practical cryptosystems are based on Shannnon's pessimistic assumption. Due to practical constraints, secret keys are much shorter than messages; therefore, these practical cryptosystems are theoretically susceptible to breaking by attackers. The goal of designing such practical ciphers, however, is to guarantee that no efficient algorithm exists for breaking them.

Wyner in [2] showed that the above negative result is a consequence of Shannon's restrictive assumption that the adversary has access to precisely the same information as the legitimate receiver. Wyner considered a scenario in which a wiretapper receives the transmitted signal over a degraded channel with respect to the legitimate receiver's channel. He further assumed that the wiretapper has no computational limitations and knows the codebook used by the transmitter. He measured the level of ignorance at the eavesdropper by its equivocation and characterized the capacity-equivocation region. Interestingly, a non-negative perfect secrecy capacity is always achievable for this scenario.

1.3 Wiretap Channel Model

In this section, we highlight the implicit assumptions inherent in the wiretap channel model and all the other generalized models.

1. **Knowledge of Channel-State-Information**: In the wiretap channel model, it is assumed that the Channel-State-Information (CSI) about the main channel and the eavesdropper's channel are perfectly available at the transmitter. While it is reasonable to assume that the main channel CSI is perfectly known, the availability of the eavesdropper's CSI is questionable. In the situations where transmitter is a wireless base station and the eavesdropper is a user in the network, however, the eavesdropper's CSI is in fact known at the transmitter. Moreover, the knowledge of the exact CSI can be replaced by a conservative assumption based on geographical information. As an example, the transmitter can certainly upper bound the signal-to-noise ratio at the eavesdropper, if it is known to be located outside a given region.

2. **Authentication**: The wiretap channel model implicitly assumes that the main channel is authenticated. As the authentication mechanisms can be implemented in the upper layers of the protocol stack, this particular assumption is not restrictive.

3. **Passive Eavesdropper**: In the wiretap channel model, the adversary is restricted to passive eavesdropping strategies. Therefore, additional techniques are required to cope with jamming.

4. **Availability of Random Generator**: Unlike traditional encoders which are deterministic functions, wiretap encoders are stochastic and rely on the availability of perfect random generators. In practice, strong pseudo-random generators could be used. The initialization mechanism of these generators need to be carefully considered.

1.4 Literature Review

Here, we summarize some of the most important definitions and results obtained on physical-layer security for wireless channels.

1.4.1 Gaussian and General Channel

The secrecy capacity for the Gaussian wiretap channel is characterized by Leung-Yan-Cheong in [3]. Wyner's work is then extended to the general (non-degraded) broadcast channel with

confidential messages by Csiszar and Korner [4]. They considered transmitting confidential information to the legitimate receiver while transmitting common information to both the legitimate receiver and the wiretapper. They established a capacity-equivocation region for this channel.

1.4.2 Fading Channel

Barros *et. al.* [5] provided a detailed characterization of the outage secrecy capacity of slow fading channels, and showed that fading alone guarantees that information-theoretic security is achieved, even when the eavesdropper has a better average signal-to-noise ratio than the legitimate receiver. The secrecy capacity of ergodic fading channels was derived independently by Liang *et. al.* [6, 7], Li *et. al.* [8], and Gopala *et. al.* [9], where the power and rate allocation schemes for secret communication over fading channels were presented.

1.4.3 Multiple-Input Multiple-Output Wiretap Channel

The extension of the wiretap channel when all the three nodes have multiple antennae were considered by different researchers. A closed form expression for the secrecy capacity of MIMO Gaussian wiretap channel was derived independently by Oggier *et. al.*, Khisti *et. al.*, and Liu *et. al.* in [10], [11], and [12], respectively. The special case where the intended receiver has a single antenna, is referred to as Multiple-Input Single-Output Multiple-Eavesdropper (MISOME) and is characterized by Khisti *et. al.* along with the optimum beam-forming in [13]. In [14, 15, 16] the MIMO wiretap channel is generalized to the case in which both receivers are to receive legitimate messages intended for each receiver, while being kept ignorant of each other's messages.

1.4.4 Relay Channel

The relay channel with confidential messages is studied in the works of [17], [18], [19], [20] and [21]. In this setup, one party communicates with another party directly, as well as through a relay node, which is used to increase the capacity between the two parties. The relay node, in this case, must be kept totally ignorant of the secret messages being transmitted. In a slightly different scenario of [22], the relay node is used to increase the capacity for the eavesdropper instead.

1.4.5 Interference Channel

The interference channel with confidential messages with and without an external eavesdropper is considered in [23, 24, 25, 26, 27, 28, 29]. In this scenario, the legitimate users try to communicate with each other through an interference channel when either the messages must be kept secret from other users or from an external eavesdropper.

1.4.6 Feedback Channel

The presence of feedback provides the wiretap channel with several advantages. First, when the legitimate channel is more noisy than the wiretap channel, feedback may permit unconditional secrecy, whereas without feedback this is not possible [30, 31]. Secondly, when both forward and feedback channels are noisy, it may be possible to increase the secrecy capacity to the usual capacity without secrecy constraint [32, 33]. Finally, the role of feedback in multiple user channels has been explored and found to aid secrecy in [34].

1.4.7 Wiretap Channel with Distortion Measure

In [35], rather than enforcing a minimum equivocation rate on the eavesdropper, a minimum distortion has been enforced. This alternative criterion could be useful for securing multimedia content such as video or voice.

1.4.8 Broadcast Channel

The broadcast channel with confidential messages has been considered in [36, 37, 38, 39]. This channel has recently been further studied in [23, 25, 40, 41], where the source node transmits a common message to both receivers, along with two additional confidential messages, each aimed at one of the two receivers. Here, the confidentiality of each message is measured with respect to the other user, and there is no external eavesdropper. In [42], the wiretap channel is extended to the parallel broadcast channels and also to the fading channels with multiple receivers. In [42], the secrecy constraint is a perfect equivocation for each of the messages, even if all the other messages are revealed to the eavesdropper. The secrecy sum capacity for a reverse broadcast channel is derived subject to this restrictive assumption.

1.4.9 Multiple-Access-Channel

The secure Gaussian MAC with/without an external eavesdropper is introduced in [43, 44, 45, 46]. The secure Gaussian MAC with an external eavesdropper consists of an ordinary

Gaussian MAC and an external eavesdropper. The capacity region of this channel is still an open problem in the information theory field. For this channel, an achievable rate scheme based on Gaussian codebooks is proposed in [46], and also the sum secrecy capacity of the degraded Gaussian channel is found in [44]. For some special cases, upper bounds, lower bounds, and some asymptotic results on the secrecy capacity exist, see for example [19, 47, 48, 49]. For the achievability part, Shannons random coding argument proves to be effective in these works.

1.5 Outline of the Dissertation

This dissertation is organized as follows: In chapter 2, we consider a scenario where a source node wishes to broadcast two confidential messages to two receivers, while a wiretapper also receives the transmitted signal. We study the general broadcast channel with an eavesdropper and the situation where the channels are degraded. We also characterize the secrecy capacity region when the channels are Additive White Gaussian Noise. Based on the rate characterization of the secure broadcast channel, we then use the broadcast strategy for the slow fading wiretap channel when only the eavesdropper's channel is fixed and known at the transmitter. In chapter 3, we establish the secrecy capacity region of the MIMO broadcast channel of chapter 2. Our achievability scheme is a combination of the dirty paper coding of Gaussian codes and randomization within the layers. To prove the converse, we use the notion of enhanced channel and show that the secret dirty paper coding of Gaussian codes is optimal. We investigate practical characterizations for the specific scenario in which the transmitter and the eavesdropper have multiple antennae, while both the intended receivers have a single antenna. We characterize the secrecy capacity region in terms of generalized eigenvalues of the receivers channels and the eavesdropper channel. In chapter 4 we establish the secrecy sum capacity of the degraded Gaussian MIMO MAC using random bininng of Gaussian codebooks. For the non-degraded channel, we present an algorithm inspired by the notion of signal alignment to achieve the largest Secure Degrees-of-Freedom (S-DoF) by using Gaussian codebooks. We then use the notion of *real alignment* to prove that for almost all channel gains in the secure K user single-antenna Gaussian MAC, we can achieve the S-DoF of $\frac{K-1}{K}$. Here, our scheme uses structure codes instead of Gaussian codebooks. In the case of the channel gains that the S-DoF of $\frac{K-1}{K}$ cannot be achieved, we propose a multi-layer coding scheme to achieve a positive S-DoF. Finally, chapter 5 concludes our works and presents an outline for the future researches. The results of this dissertation have been published/submitted in [50, 51, 52, 53, 54, 55, 56]. It should be noted that parallel to our works, references [57, 58, 59, 60], have independently derived some the results of this

dissertation.

Chapter 2

Secure Broadcasting

In this chapter, we consider a scenario in which a source node wishes to broadcast two confidential messages to two receivers, while a wiretapper also receives the transmitted signal. This model is motivated by wireless communications, where individual secure messages are broadcast over open media and can be received by any illegitimate receiver. The secrecy level is measured by the equivocation rate at the eavesdropper. We first study the general (non-degraded) broadcast channel with an eavesdropper. We present an inner bound on the secrecy capacity region for this model. This inner bound is based on a combination of random binning, and the Gelfand-Pinsker binning. We further study the situation in which the channels are degraded. For the degraded broadcast channel with an eavesdropper, we present the secrecy capacity region. Our achievable coding scheme is based on Cover's superposition scheme and random binning. We refer to this scheme as the Secret Superposition Scheme. Our converse proof is based on a combination of the converse proof of the conventional degraded broadcast channel and Csiszar Lemma. We then assume that the channels are Additive White Gaussian Noise (AWGN) and show that the Secret Superposition Scheme with Gaussian codebook is optimal. The converse proof is based on Costa's entropy power inequality. Finally, we use a broadcast strategy for the slowly fading wiretap channel when only the eavesdropper's channel is fixed and known at the transmitter. We derive the optimum power allocation for the coding layers, which maximizes the total average rate.

2.1 Preliminaries

Consider a Broadcast Channel with an Eavesdropper (BCE), as depicted in Figure 2.1.

In this confidential setting, the transmitter wishes to send two independent messages

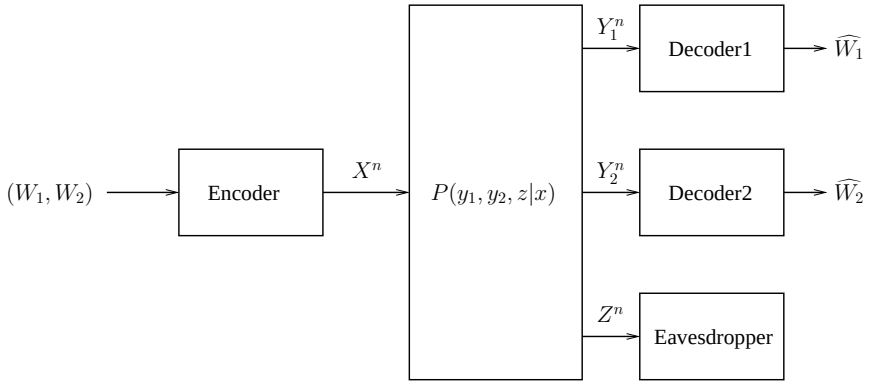

Figure 2.1: Broadcast Channel with an Eavesdropper (BCE)

(W_1, W_2) to the respective receivers in n uses of the channel and prevent the eavesdropper from having any information about the messages. A discrete memoryless broadcast channel with an eavesdropper is represented by $(\mathcal{X}, P, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z})$, where $\mathcal{X}$ is the finite input alphabet set, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ are three finite output alphabet sets, and P is the channel transition probability $P(y_1, y_2, z|x)$. The input of the channel is $x^n \in \mathcal{X}^n$ and the outputs are $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$, and $z^n \in \mathcal{Z}^n$ for Receiver 1, Receiver 2, and the eavesdropper, respectively. The channel is discrete memoryless in the sense that

$$P(y_1^n, y_2^n, z^n|x^n) = \prod_{i=1}^{n} P(y_{1,i}, y_{2,i}, z_i|x_i). \tag{2.1}$$

A $((2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel with an eavesdropper consists of a stochastic encoder

$$f : (\{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_2}\}) \to \mathcal{X}^n, \tag{2.2}$$

and two decoders,

$$g_1 : \mathcal{Y}_1^n \to \{1, 2, ..., 2^{nR_1}\} \tag{2.3}$$

and

$$g_2 : \mathcal{Y}_2^n \to \{1, 2, ..., 2^{nR_2}\}. \tag{2.4}$$

The average probability of error is defined as the probability that the decoded messages are not equal to the transmitted messages; that is,

$$P_e^{(n)} = P(g_1(Y_1^n) \neq W_1 \cup g_2(Y_2^n) \neq W_2). \tag{2.5}$$

The knowledge that the eavesdropper can extract about W_1 and W_2 from its received signal Z^n is measured by

$$I(Z^n, W_1) = H(W_1) - H(W_1|Z^n), \tag{2.6}$$

$$I(Z^n, W_2) = H(W_2) - H(W_2|Z^n), \tag{2.7}$$

and

$$I(Z^n,(W_1,W_2)) = H(W_1,W_2) - H(W_1,W_2|Z^n). \tag{2.8}$$

Perfect secrecy revolves around the idea that the eavesdropper should not obtain any information about the transmitted messages. Perfect secrecy thus requires that

$$\begin{aligned} I(Z^n,W_1) = 0 \Leftrightarrow H(W_1) = H(W_1|Z^n), \\ I(Z^n,W_2) = 0 \Leftrightarrow H(W_2) = H(W_2|Z^n), \end{aligned} \tag{2.9}$$

and

$$I(Z^n,(W_1,W_2)) = 0 \Leftrightarrow H(W_1,W_2) = H(W_1,W_2|Z^n). \tag{2.10}$$

where $n \to \infty$. The secrecy levels of confidential messages W_1 and W_2 are measured at the eavesdropper in terms of equivocation rates which are defined as follows:

Definition 1. *The equivocation rates R_{e1}, R_{e2} and R_{e12} for the broadcast channel with an eavesdropper are:*

$$\begin{aligned} R_{e1} &= \frac{1}{n} H(W_1|Z^n), \\ R_{e2} &= \frac{1}{n} H(W_2|Z^n), \\ R_{e12} &= \frac{1}{n} H(W_1,W_2|Z^n). \end{aligned} \tag{2.11}$$

The perfect secrecy rates R_1 and R_2 are the amount of information that can be sent to the legitimate receivers in a reliable and confidential manner.

Definition 2. *A secrecy rate pair (R_1,R_2) is said to be achievable if for any $\epsilon > 0, \epsilon_1 > 0, \epsilon_2 > 0, \epsilon_3 > 0$, there exists a sequence of $((2^{nR_1},2^{nR_2}),n)$ codes, such that for sufficiently large n, we have:*

$$P_e^{(n)} \leq \epsilon, \tag{2.12}$$

$$R_{e1} \geq R_1 - \epsilon_1, \tag{2.13}$$

$$R_{e2} \geq R_2 - \epsilon_2, \tag{2.14}$$

$$R_{e12} \geq R_1 + R_2 - \epsilon_3. \tag{2.15}$$

In the above definition, the first condition concerns the reliability, while the other conditions guarantee perfect secrecy for each individual message and the combination of the two messages, respectively. Since the messages are independent of each other, the conditions of (2.13) and (2.15) or (2.14) and (2.15) are sufficient to provide perfect secrecy.

The capacity region is defined as follows.

Definition 3. *The capacity region of the broadcast channel with an eavesdropper is the closure of the set of all achievable rate pairs (R_1,R_2).*

2.2 Achievable Rates for General BCE

In this section, we consider the general broadcast channel with an eavesdropper and present an achievable rate region. Our achievable coding scheme is based on a combination of the random binning, superposition coding, rate splitting, and Gelfand-Pinsker binning schemes [61]. Our binning approach is supplemented with superposition coding to accommodate the common message. We call this scheme the Secret Superposition Scheme. Additional binning is introduced for the confidentiality of private messages. We note that these double binning techniques have been used by various authors for secret communication (see e.g. [25, 41]). The following theorem illustrates the achievable rate region for this channel.

Theorem 1. *Let $\mathbb{R}_I$ denote the union of all non-negative rate pairs (R_0, R_1, R_2) satisfying*

$$\begin{aligned} R_0 &\leq \min\{I(U;Y_1), I(U;Y_2)\} - I(U;Z), \\ R_1 + R_0 &\leq I(V_1;Y_1|U) - I(V_1;Z|U) + \min\{I(U;Y_1), I(U;Y_2)\} - I(U;Z), \\ R_2 + R_0 &\leq I(V_2;Y_2|U) - I(V_2;Z|U) + \min\{I(U;Y_1), I(U;Y_2)\} - I(U;Z), \\ R_1 + R_2 + R_0 &\leq I(V_1;Y_1|U) + I(V_2;Y_2|U) - I(V_1,V_2;Z|U) - I(V_1;V_2|U) \\ &\quad + \min\{I(U;Y_1), I(U;Y_2)\} - I(U;Z), \end{aligned} \tag{2.16}$$

over all joint distributions $P(u)P(v_1,v_2|u)P(x|v_1,v_2)P(y_1,y_2,z|x)$. Then, any rate pair $(R_0,R_1,R_2) \in \mathbb{R}_I$ is achievable for the broadcast channel with an eavesdropper and with common information.

Please see section 2.6.1 for the proof.

Remark 1. *If we remove the secrecy constraints by removing the eavesdropper, the above rate region becomes Marton's achievable region with common information for the general broadcast channel.*

Remark 2. *If we remove one of the users, e.g. user 2 and the common message, then we get Csiszar and Korner's secrecy capacity for the other user.*

2.3 The Capacity Region of the Degraded BCE

In this section, we consider the degraded broadcast channel with an eavesdropper and establish its secrecy capacity region.

Definition 4. *A broadcast channel with an eavesdropper is said to be physically degraded, if $X \to Y_1 \to Y_2 \to Z$ forms a Markov chain. In other words, we have*

$$P(y_1,y_2,z|x) = P(y_1|x)P(y_2|y_1)P(z|y_2).$$

Definition 5. *A broadcast channel with an eavesdropper is said to be stochastically degraded if its conditional marginal distributions are the same as that of a physically degraded broadcast channel, i.e., if there exist two distributions $P^{'}(y_2|y_1)$ and $P^{'}(z|y_2)$, such that*

$$P(y_2|x) = \sum_{y_1} P(y_1|x)P^{'}(y_2|y_1),$$
$$P(z|x) = \sum_{y_2} P(y_2|x)P^{'}(z|y_2).$$

Lemma 1. *The secrecy capacity region of a broadcast channel with an eavesdropper depends only on the conditional marginal distributions $P(y_1|x)$, $P(y_2|x)$ and $P(z|x)$.*

Proof: It suffices to show that the error probability of $P_e^{(n)}$ and the equivocations of $H(W_1|Z^n)$, $H(W_2|Z^n)$ and $H(W_1, W_2|Z^n)$ are only functions of the marginal distributions when we use the same codebook and encoding schemes. Note that

$$\max\{P_{e,1}^{(n)}, P_{e,2}^{(n)}\} \leq P_e^{(n)} \leq P_{e,1}^{(n)} + P_{e,2}^{(n)}. \tag{2.17}$$

Therefore, $P_e^{(n)}$ is small if, and only if, both $P_{e,1}^{(n)}$ and $P_{e,2}^{(n)}$ are small. On the other hand, for a given codebook and encoding scheme, the decoding error probabilities $P_{e,1}^{(n)}$, $P_{e,2}^{(n)}$ and the equivocation rates depend only on the marginal channel probability densities of $P_{Y_1|X}$, $P_{Y_2|X}$ and $P_{Z|X}$. Thus, the same code and encoding scheme gives the same $P_e^{(n)}$ and equivocation rates. □

In the following theorem, we fully characterize the capacity region of the physically degraded broadcast channel with an eavesdropper.

Theorem 2. *The capacity region for transmitting independent secret information over the degraded broadcast channel is the convex hull of the closure of all (R_1, R_2) satisfying*

$$R_1 \leq I(X; Y_1|U) - I(X; Z|U), \tag{2.18}$$
$$R_2 \leq I(U; Y_2) - I(U; Z), \tag{2.19}$$

for some joint distribution $P(u)P(x|u)P(y_1, y_2, z|x)$.

Please refer to section 2.6.2 for the proof.

Remark 3. *If we remove the secrecy constraints by removing the eavesdropper, then the above theorem becomes the capacity region of the degraded broadcast channel.*

The coding scheme is based on Cover's superposition coding [62] and random binning. We refer to this scheme as the Secure Superposition Coding scheme. The available resources at the encoder are used for two purposes: to confuse the eavesdropper so that perfect secrecy can be achieved for both layers, and to transmit the messages into the main channels. To satisfy confidentiality, the randomization used in the first layer is fully exploited in the second layer. This makes an increase of $I(U;Z)$ in the bound of R_1.

Remark 4. *As Lemma 2 bounds the secrecy rates for the general broadcast channel with an eavesdropper, then Theorem 2 is true when only the legitimate receivers are degraded.*

2.4 Capacity Region of Gaussian BCE

In this section, we consider the Gaussian Broadcast Channel with an Eavesdropper (G-BCE). Note that optimizing (2.18) and (2.19) for AWGN channels involves solving a non-convex functional. Usually nontrivial techniques and strong inequalities are used to solve the optimization problems of this type. In [3], Leung-Yan-Cheong successfully evaluated the capacity expression of the wiretap channel by using the entropy power inequality [63, 64]. Alternatively, it can also be evaluated using a classical result from the Estimation Theory and the relationship between mutual information and minimum mean-squared error estimation. On the other hand, the entropy power inequality is sufficient to establish the converse proof of a Gaussian broadcast channel without secrecy constraint. Unfortunately, the traditional entropy power inequality does not extend to the secure multi-user case. Here, by using Costa's version of the entropy power inequality, we show that secret superposition coding with Gaussian codebook is optimal.

Figure 2.2 shows the channel model. At time i the received signals are $Y_{1i} = X_i + N_{1i}$, $Y_{2i} = X_i + N_{2i}$ and $Z_i = X_i + N_{3i}$, where N_{ji} is a Gaussian random variable with zero mean and $Var(N_{ji}) = \sigma_j^2$ for $j = 1, 2, 3$. Here $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_3^2$. Assume that the transmitted power is limited to $E[X^2] \leq P$. Since the channels are degraded, the received signals can alternatively be written as $Y_{1i} = X_i + N_{1i}$, $Y_{2i} = Y_{1i} + N_{2i}^{'}$ and $Z_i = Y_{2i} + N_{3i}^{'}$, where N_{1i}'s are i.i.d $\mathcal{N}(0, \sigma_1^2)$, $N_{2i}^{'}$'s are i.i.d $\mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$, and $N_{3i}^{'}$'s are i.i.d $\mathcal{N}(0, \sigma_3^2 - \sigma_2^2)$. Figure 2.3 shows the equivalent channels for the G-BCE. The following theorem illustrates the secrecy capacity region of G-BCE.

Theorem 3. *The secrecy capacity region of the G-BCE is given by the set of rates pairs*

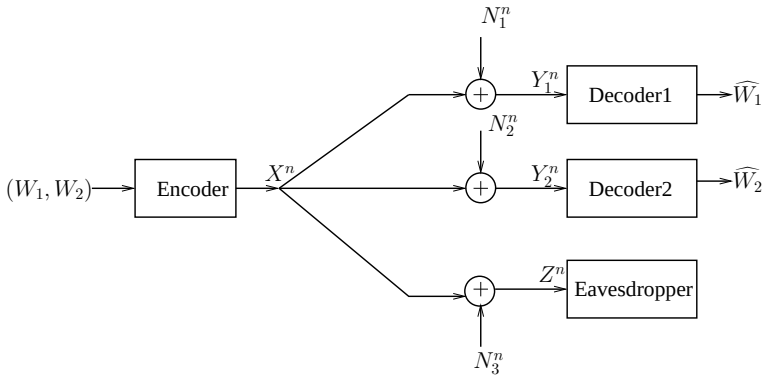

Figure 2.2: Gaussian Broadcast Channel with an Eavesdropper(G-BCE)

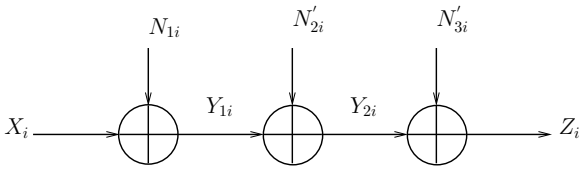

Figure 2.3: Equivalent Channels for the G-BCE

(R_1, R_2) *satisfying*

$$R_1 \leq C\left(\frac{\alpha P}{\sigma_1^2}\right) - C\left(\frac{\alpha P}{\sigma_3^2}\right), \tag{2.20}$$

$$R_2 \leq C\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) - C\left(\frac{(1-\alpha)P}{\alpha P + \sigma_3^2}\right). \tag{2.21}$$

for some $\alpha \in [0,1]$ *and* $C(x) = \frac{1}{2}\log(1+x)$.

Please see section 2.6.3 for the proof.

Figure 2.4 shows the capacity region of a degraded Gaussian broadcast channel with and without any secrecy constraint. In this figure $P = 20$, $N_1 = 0.9$, $N_2 = 1.5$ and $N_3 = 4$.

2.5 A Multilevel Coding Approach to the Slowly Fading Wiretap Channel

In this section, we use the secure degraded broadcast channel from the previous section to develop a new broadcast strategy for a slow fading wiretap channel. This strategy aims to maximize the average achievable rate where the main channel state information is not available at the transmitter. By assuming that there is an infinite number of ordered receivers

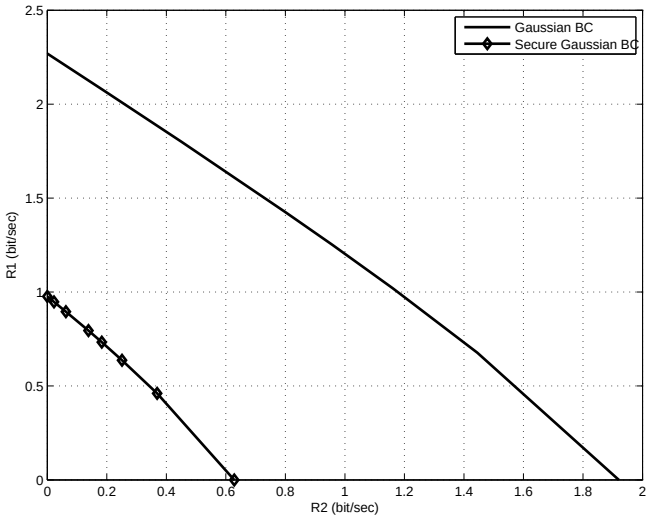

Figure 2.4: Secret versus Non-Secret Capacity Region of a Degraded Broadcast Channel. The gap between the regions corresponds to the cost of securing the system. The power used at the transmitter is beneficial for both the intended receiver and the eavesdropper such that the secure sum rate will be saturated in high SNR.

which correspond to different channel realizations, we propose a secret multilevel coding scheme that maximizes the underlying objective function. First, some preliminaries and definitions are given, and then the proposed multilevel coding scheme is described. Here, we follow the steps of the broadcast strategy for the slowly fading point-to-point channel of [65]. This method is used in several other papers; see, for example [66, 67, 68].

2.5.1 Channel Model

Consider a wiretap channel as depicted in Figure 2.5. The transmitter wishes to communicate with the destination in the presence of an eavesdropper. At time i, the signal received by the destination and the eavesdropper are given as follows

$$\begin{aligned} Y_i &= h_M X_i + N_{1i} \\ Z_i &= h_E X_i + N_{2i} \end{aligned} \tag{2.22}$$

where X_i is the transmitted symbol and h_M, h_E are the fading coefficients from the source to the legitimate receiver and to the eavesdropper, respectively. The fading power gains of the main and eavesdropper channels are given by $s = |h_M|^2$ and $\hat{s} = |h_E|^2$, respectively. N_{1i}, N_{2i} are the additive noise samples, which are Gaussian i.i.d with zero mean and unit variance. We assume that the channels are slowly fading, and also assume that the transmitter only

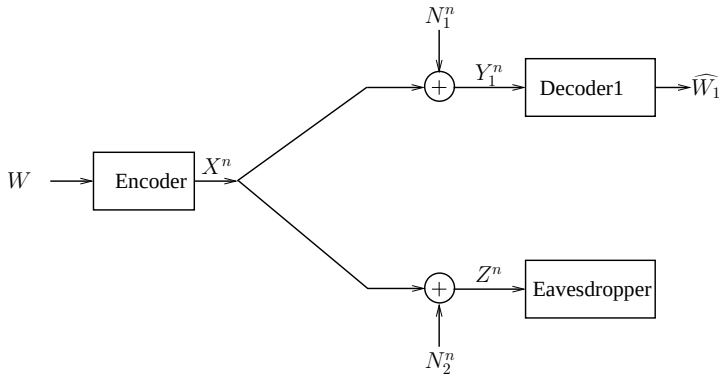

Figure 2.5: Gaussian Wiretap Channel

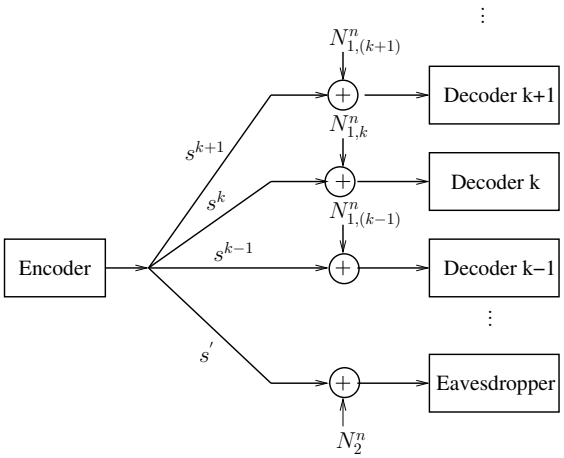

Figure 2.6: Equivalent Broadcast Channel Model.

knows the channel state information of the eavesdropper channel. A motivation for this assumption is that when both channels are unknown at the transmitter, we assume that $\hat{s} = |h_E|^2$ denotes the best-case eavesdropper channel gain. For each realization of h_M there is an achievable rate. Since the transmitter has no information about the main channel and the channels are slowly fading, then the system is non-ergodic. Here, we are interested in the average rate for various independent transmission blocks. The average shall be calculated over the distribution of h_M.

2.5.2 The Secret Multilevel Coding Approach

An equivalent broadcast channel for our channel is depicted in Figure 2.6. wherein the transmitter sends an infinite number of secure layers of coded information. The receiver is equivalent to a continuum of ordered users. For each channel realization h_M^k with the fading

power gain s^k, the information rate is $R(s^k, \hat{s})$. We drop the superscript k, and the realization of the fading power random variable S is denoted by s. Therefore, the transmitter views the main channel as a secure degraded Gaussian broadcast channel with an infinite number of receivers. The result of the previous section for the two receivers can easily be extended to an arbitrary number of users. According to theorem 3, the incremental differential secure rate is then given by

$$dR(s,\hat{s}) \;=\; \left[\frac{1}{2}\log\left(1+\frac{s\rho(s)ds}{1+sI(s)}\right)-\frac{1}{2}\log\left(1+\frac{\hat{s}\rho(s)ds}{1+\hat{s}I(s)}\right)\right]^{+}, \tag{2.23}$$

where $\rho(s)ds$ is the transmit power of a layer parameterized by s, intended for receiver s. As $\log(1+x)\approx x$ for $x\leq 1$ then the log function may be discarded. The function $I(s)$ represents the interference noise of the receivers indexed by $u>s$ which cannot be canceled at receiver s. The interference at receiver s is therefore given by

$$I(s)=\int_{s}^{\infty}\rho(u)d(u). \tag{2.24}$$

The total transmitted power is the summation of the power assigned to the layers

$$P=I(0)=\int_{0}^{\infty}\rho(u)d(u). \tag{2.25}$$

The total achievable rate for a fading realization s is an integration of the incremental rates over all receivers, which can successfully decode the respective layer

$$R(s,\hat{s})=\frac{1}{2}\int_{0}^{s}\left[\frac{u\rho(u)du}{1+uI(u)}-\frac{\hat{s}\rho(u)du}{1+\hat{s}I(u)}\right]^{+}. \tag{2.26}$$

Our goal is to maximize the total average rate over all fading realizations with respect to the power distribution $\rho(s)$ (or equivalently, with respect to $I(u)$, $u\geq 0$) under the power constraint of 2.25. The optimization problem may be written as

$$\begin{aligned} R_{\max} &= \max_{I(u)} \int_{0}^{\infty} R(u,\hat{s})f(u)du, \\ &s.t \\ P &= I(0)=\int_{0}^{\infty}\rho(u)d(u), \end{aligned} \tag{2.27}$$

where $f(u)$ is the probability distribution function (pdf) of the power gain S. Noting that the cumulative distribution function (cdf) is $F(u)=\int_{0}^{u}f(a)da$, the optimization problem may be written as

$$\begin{aligned} R_{\max} &= \frac{1}{2}\max_{I(u)} \int_{0}^{\infty} (1-F(u))G(u)du, \\ &s.t \\ P &= I(0)=\int_{0}^{\infty}\rho(u)d(u), \end{aligned} \tag{2.28}$$

where $G(u) = \left[\frac{u}{1+uI(u)} - \frac{\hat{s}}{1+\hat{s}I(u)}\right]^+ \rho(u)$. Note that $\rho(u) = -I^{'}(u)$. Therefore, the functional in (2.28) may be written as

$$J(x, I(x), I^{'}(x)) = \tag{2.29}$$

$$-(1 - F(x)) \left[\frac{x}{1 + xI(x)} - \frac{\hat{s}}{1 + \hat{s}I(x)}\right]^+ I^{'}(x). \tag{2.30}$$

The necessary condition for the maximization of an integral of J over x is

$$J_I - \frac{d}{dx} J_{I'} = 0, \tag{2.31}$$

where J_I means the derivation of function J with respect to I, and similarly $J_{I'}$ is the derivation of J with respect to $I^{'}$. After some manipulations, the optimum $I(x)$ is given by

$$I(x) = \begin{cases} \frac{1-F(x)-(x-\hat{s})f(x)}{\hat{s}(1-F(x))+x(x-\hat{s})f(x)}, & \max\{\hat{s}, x_0\} \leq x \leq x_1; \\ 0, & \text{otherwise}, \end{cases} \tag{2.32}$$

where x_0 is determined by $I(x_0) = P$, and x_1 by $I(x_1) = 0$.

As a special case, consider the Rayleigh flat fading channel. The random variable S is exponentially distributed with

$$f(s) = e^{-s}, \qquad F(s) = 1 - e^{-s}, \qquad s \geq 0. \tag{2.33}$$

Substituting $f(s)$ and $F(s)$ into the optimum $I(s)$ and taking the derivative with respect to the fading power s yields the following optimum transmitter power policy

$$\rho(s) = -\frac{d}{ds} I(s) = \begin{cases} \frac{-s^2+2(\hat{s}+1)s-s^{'2}}{(s^2-\hat{s}s+\hat{s})^2}, & \max\{\hat{s}, s_0\} \leq s \leq s_1; \\ 0, & \text{otherwise}, \end{cases} \tag{2.34}$$

where s_0 is the solution of the equation $I(s_0) = P$, which is

$$s_0 = \frac{-1 + Ps^{'} + \sqrt{P^2 s^{'2} + 2P(1-2P)\hat{s} + 4P + 1}}{2P}, \tag{2.35}$$

and s_1 is determined by $I(s_1) = 0$, which is

$$s_1 = 1 + \hat{s}. \tag{2.36}$$

2.6 Proofs for Chapter 2

2.6.1 Proof of Theorem 1

We split the private message $W_1 \in \{1, 2, ..., 2^{nR_1}\}$ into $W_{11} \in \{1, 2, ..., 2^{nR_{11}}\}$ and $W_{10} \in \{1, 2, ..., 2^{nR_{10}}\}$, and $W_2 \in \{1, 2, ..., 2^{nR_2}\}$ into $W_{22} \in \{1, 2, ..., 2^{nR_{22}}\}$ and $W_{20} \in \{1, 2, ..., 2^{nR_{20}}\}$,

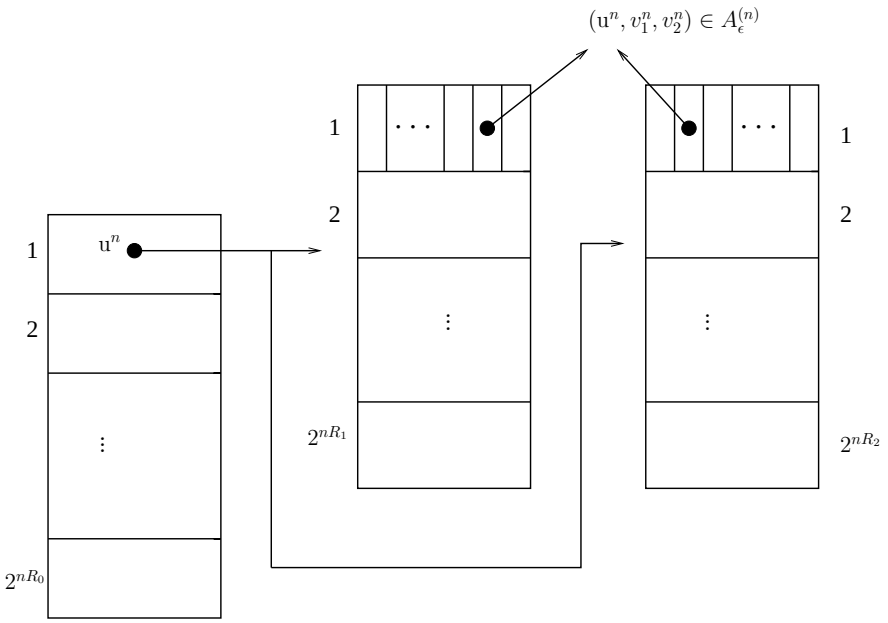

Figure 2.7: The Stochastic Encoder

respectively. W_{11} and W_{22} are only to be decoded by the intended receivers, while W_{10} and W_{20} are to be decoded by both receivers. Now, we combine (W_{10}, W_{20}, W_0) into a single auxiliary variable U. The messages W_{11} and W_{22} are represented by auxiliary variables V_1 and V_2, respectively. Here, $R_{10} + R_{11} = R_1$ and $R_{20} + R_{22} = R_2$.

1) *Codebook Generation*: The structure of the encoder is depicted in Figure 2.7. Fix $P(u)$, $P(v_1|u)$, $P(v_2|u)$ and $P(x|v_1, v_2)$. The stochastic encoding is as follows. Define

$$\begin{aligned} L_{11} &= I(V_1; Y_1|U) - I(V_1; Z, V_2|U), \\ L_{12} &= I(V_1; Z|V_2, U), \\ L_{21} &= I(V_2; Z|V_1, U) \\ L_{22} &= I(V_2; Y_2|U) - I(V_2; Z, V_1|U), \\ L_3 &= I(V_1; V_2|U) - \epsilon, \end{aligned} \tag{2.37}$$

Note that,

$$\begin{aligned} L_{11} + L_{12} + L_3 &= I(V_1; Y_1|U) - \epsilon, \\ L_{22} + L_{21} + L_3 &= I(V_2; Y_2|U) - \epsilon, \end{aligned} \tag{2.38}$$

We first prove the case where

$$R_{11} \geq L_{11} \geq 0, \tag{2.39}$$

$$R_{22} \geq L_{22} \geq 0. \tag{2.40}$$

Generate $2^{n(R_{10}+R_{20}+R_0)}$ independent and identically distributed (i.i.d) sequences $u^n(k)$ with $k \in \{1, 2, ..., 2^{R_{10}+R_{20}+R_0}\}$, according to the distribution $P(u^n) = \prod_{i=1}^{n} P(u_i)$. For each codeword $u^n(k)$, generate $2^{L_{11}+L_{12}+L_3}$ i.i.d codewords $v_1^n(i, i^{'}, i^{''})$, with $i \in \{1, 2, ..., 2^{nL_{11}}\}$, $i^{'} \in \{1, 2, ..., 2^{nL_{12}}\}$ and $i^{''} \in \{1, 2, ..., 2^{nL_3}\}$, according to $P(v_1^n|u^n) = \prod_{i=1}^{n} P(v_{1i}|u_i)$. The indexing presents an alternative interpretation of binning. Randomly distribute these sequences of v_1^ninto $2^{nL_{11}}$ bins indexed by i, for the codewords in each bin, randomly distribute them into $2^{nL_{12}}$ sub-bins indexed by $i^{'}$; thus $i^{''}$ is the index for the codeword in each sub-bin. Similarly, for each codeword u^n, generate $2^{L_{21}+L_{22}+L_3}$ i.i.d codewords $v_2^n(j, j^{'}, j^{''})$ according to $P(v_2^n|u^n) = \prod_{i=1}^{n} P(v_{2i}|u_i)$, where $j \in \{1, 2, ..., 2^{nL_{21}}\}$, $j^{'} \in \{1, 2, ..., 2^{nL_{22}}\}$ and $j^{''} \in \{1, 2, ..., 2^{nL_3}\}$.

2) *Encoding*: To send messages (w_{10}, w_{20}, w_0), we calculate the corresponding message index k and choose the corresponding codeword $u^n(k)$. Given this $u^n(k)$, there exists $2^{n(L_{11}+L_{12}+L_3)}$ codewords of $v_1^n(i, i^{'}, i^{''})$ to choose from for representing message w_{11}. Evenly map $2^{nR_{11}}$ messages w_{11} to $2^{nL_{11}}$ bins, then, given (2.39), each bin corresponds to at least one message w_{11}. Thus, given w_{11}, the bin index i can be decided.

1. If $R_{11} \leq L_{11} + L_{12}$, each bin corresponds to $2^{n(R_{11}-L_{11})}$ messages w_{11}. Evenly place the $2^{nL_{12}}$ sub-bins into $2^{n(R_{11}-L_{11})}$ cells. For each given w_{11}, we can find the corresponding cell. We, then, randomly choose a sub-bin from that cell, thus the sub-bin index $i^{'}$ can be decided. The codeword $v_1^n(i, i^{'}.i^{''})$ will be chosen properly from that sub-bin.

2. If $L_{11}+L_{12} \leq R_{11} \leq L_{11}+L_{12}+L_3$, then each sub-bin is mapped to at least one message w_{11}, therefore, given w_{11}; $i^{'}$ can be decided. In each sub-bin, there are $2^{n(R_{11}-L_{11}-L_{12})}$ messages. The codeword $v_1^n(i, i^{'}.i^{''})$ will be chosen randomly and properly from that sub-bin.

Given w_{22}, we select $v_2^n(j, j^{'}, j^{''})$ in the exact same manner. From the given sub-bins , the encoder chooses the codeword pair $(v_1^n(i, i^{'}, i^{''}), v_2^n(j, j^{'}, j^{''}))$ that satisfies the following property,

$$(v_1^n(i, i^{'}, i^{''}), v_2^n(j, j^{'}, j^{''})) \in A_{\epsilon}^{(n)}(V_1, V_2, U) \tag{2.41}$$

where $A_{\epsilon}^{(n)}(U, V_1, V_2)$ denotes the set of jointly typical sequences u^n, v_1^n, and v_2^n with respect to $P(u, v_1, v_2)$. If there is more than one such pair, the transmitter randomly chooses one; if there is no such pair, an error is declared.

Given v_1^n and v_2^n, the channel input x^n is generated i.i.d. according to the distribution $P(x^n|v_1^n, v_2^n) = \prod_{i=1}^{n} P(x_i|v_{1i}, v_{2i})$.

3) *Decoding*: The received signals at the legitimate receivers, y_1^n and y_2^n, are the outputs of the channels $P(y_1^n|x^n) = \prod_{i=1}^n P(y_{1,i}|x_i)$ and $P(y_2^n|x^n) = \prod_{i=1}^n P(y_{2,i}|x_i)$, respectively. The first receiver looks for the unique sequence $u^n(k)$ such that

$$(u^n(k), y_1^n) \in A_\epsilon^{(n)}(U, Y_1). \tag{2.42}$$

If such $u^n(k)$ exists and is unique, set $\hat{k} = k$; otherwise, declare an error. Upon decoding k, this receiver looks for sequences $v_1^n(i, i^{'}, i^{''})$ such that

$$(v_1^n(i, i^{'}, i^{''}), u^n(k), y_1^n) \in A_\epsilon^{(n)}(V_1, U, Y_1). \tag{2.43}$$

If such $v_1^n(i, i^{'}, i^{''})$ exists and is unique, set $\hat{i} = i$, $\hat{i}^{'} = i^{'}$, and $\hat{i}^{''} = i^{''}$; otherwise, declare an error. Using the values of $\hat{k}, \hat{i}, \hat{i}^{'}$ and $\hat{i}^{''}$, the decoder can calculate the message indices $\hat{w_0}, \hat{w}_{10}$ and $\hat{w}_{11}$. The decoding for the second decoder is similar.

4) *Error Probability Analysis*: Since the region of $\mathbb{R}_I$ is a subset of Marton's region, then the error probability analysis is the same as [69].

5) *Equivocation Calculation*: To meet the secrecy requirements, we need to prove that the common message W_0, the combination of (W_0, W_1), the combination of (W_0, W_2), and the combination of (W_0, W_1, W_2) are perfectly secured. The proof of secrecy requirement for the message W_0 is straightforward and is therefore omitted.

To prove the secrecy requirement for (W_0, W_1), we have

$$\begin{aligned}
nR_{e10} &= H(W_1, W_0|Z^n) \\
&= H(W_1, W_0, Z^n) - H(Z^n) \\
&= H(W_1, W_0, U^n, V_1^n, Z^n) - H(U^n, V_1^n|W_1, W_0, Z^n) - H(Z^n) \\
&= H(W_1, W_0, U^n, V_1^n) + H(Z^n|W_1, W_0, U^n, V_1^n) - H(U^n|W_1, W_0, Z^n) \\
&- H(V_1^n|W_1, W_0, Z^n, U^n) - H(Z^n) \\
&\overset{(a)}{\geq} H(W_1, W_0, U^n, V_1^n) + H(Z^n|W_1, W_0, U^n, V_1^n) - n\epsilon_n - H(Z^n) \\
&\overset{(b)}{=} H(W_1, W_0, U^n, V_1^n) + H(Z^n|U^n, V_1^n) - n\epsilon_n - H(Z^n) \\
&\overset{(c)}{\geq} H(U^n, V_1^n) + H(Z^n|U^n, V_1^n) - n\epsilon_n - H(Z^n) \\
&= H(U^n) + H(V_1^n|U^n) - I(U^n, V_1^n; Z^n) - n\epsilon_n \\
&\overset{(d)}{\geq} \min\{I(U^n; Y_1^n), I(U^n; Y_2^n)\} + I(V_1^n; Y_1^n|U^n) - I(V_1^n; Z^n|U^n) - I(U^n; Z^n) - n\epsilon_n \\
&\overset{(e)}{\geq} nR_1 + nR_0 - n\epsilon_n,
\end{aligned} \tag{2.44}$$

where (a) follows from Fano's inequality that bounds the term $H(U^n|W_1, W_0, Z^n) \leq h(P_{we0}^{(n)}) + nP_{we0}^n R_{w0} \leq n\epsilon_n/2$ and the term $H(V_1^n|W_1, W_0, Z^n, U^n) \leq h(P_{we1}^{(n)}) + nP_{we1}^n R_{w1} \leq n\epsilon_n/2$ for

sufficiently large n. Here P_{we0}^n and P_{we1}^n denote the wiretapper's error probability of decoding u^n and V_1^n in the case that the bin numbers w_0 and w_1 are known to the eavesdropper, respectively. The eavesdropper first looks for the unique u^n in bin w_0 of the first layer, such that it is jointly typical with z^n. As the number of candidate codewords is small enough, the probability of error is arbitrarily small for a sufficiently large n. Next, given u^n, the eavesdropper looks for the unique v_1^n in the bin w_1 which is jointly typical with z^n. Similarly, since the number of available candidates is small enough, then the probability of a decoding error is arbitrarily small. (b) follows from the fact that $(W_1, W_0) \rightarrow U^n \rightarrow V_1^n \rightarrow Z^n$ forms a Markov chain. Therefore, we have $I(W_1, W_0; Z^n|U^n, V_1^n) = 0$, where it is implied that $H(Z^n|W_1, W_0, U^n, V_1^n) = H(Z^n|U^n, V_1^n)$. (c) follows from the fact that $H(W_1, W_0, U^n, X^n) \geq H(U^n, X^n)$. (d) follows from that fact that $H(U^n) \geq \min\{I(U^n; Y_1^n), I(U^n; Y_2^n)\}$ and $H(V_1^n|U^n) \geq I(V_1^n; Y_1^n|U^n)$. (e) follows from Lemma 11 of the Appendix.

By using the same approach it is easy to show that,

$$\begin{aligned} nR_{e20} &= H(W_2, W_0|Z^n) \\ &\geq nR_2 + nR_0 - n\epsilon_n. \end{aligned} \tag{2.45}$$

Therefore, we only need to prove that (W_0, W_1, W_2) is perfectly secured; we have

$$
\begin{aligned}
nR_{e120} &= H(W_1, W_2, W_0|Z^n) \qquad (2.46)\\
&= H(W_1, W_2, W_0, Z^n) - H(Z^n)\\
&= H(W_1, W_2, W_0, U^n, V_1^n, V_2^n, Z^n) - H(U^n, V_1^n, V_2^n|W_1, W_2, W_0, , Z^n) - H(Z^n)\\
&= H(W_1, W_2, W_0, U^n, V_1^n, V_2^n) + H(Z^n|W_1, W_2, W_0, U^n, V_1^n, V_2^n)\\
&- H(U^n, V_1^n, V_2^n|W_1, W_2, W_0, Z^n) - H(Z^n)\\
&\overset{(a)}{\geq} H(W_1, W_2, W_0, U^n, V_1^n, V_2^n) + H(Z^n|W_1, W_2, W_0, U^n, V_1^n, V_2^n) - n\epsilon_n - H(Z^n)\\
&\overset{(b)}{=} H(W_1, W_2, W_0, U^n, V_1^n, V_2^n) + H(Z^n|U^n, V_1^n, V_2^n) - n\epsilon_n - H(Z^n)\\
&\overset{(c)}{\geq} H(U^n, V_1^n, V_2^n) + H(Z^n|U^n, V_1^n, V_2^n) - n\epsilon_n - H(Z^n)\\
&\overset{(d)}{=} H(U^n) + H(V_1^n|U^n) + H(V_2^n|U^n) - I(V_1^n; V_2^n|U^n) + H(Z^n|U^n, V_1^n, V_2^n)\\
&- n\epsilon_n - H(Z^n)\\
&\overset{(e)}{\geq} \min\{I(U^n; Y_1^n), I(U^n; Y_2^n)\} + I(V_1^n; Y_1^n|U^n) + I(V_2^n; Y_2^n|U^n) - I(V_1^n; V_2^n|U^n)\\
&- I(U^n, V_1^n, V_2^n; Z^n) - n\epsilon_n\\
&\overset{(f)}{\geq} \min\{I(U^n; Y_1^n), I(U^n; Y_2^n)\} + I(V_1^n; Y_1^n|U^n) + I(V_2^n; Y_2^n|U^n) - I(V_1^n; V_2^n|U^n)\\
&- I(V_1^n, V_2^n; Z^n|U^n) - I(U^n; Z^n) - n\epsilon_n\\
&\overset{(g)}{\geq} n\min\{I(U; Y_1), I(U; Y_2)\} + nI(V_1; Y_1|U) + nI(V_2; Y_2|U) - nI(V_1; V_2|U)\\
&- nI(V_1, V_2; Z|U) - nI(U; Z) - n\epsilon_n\\
&\geq nR_1 + nR_2 + nR_0 - n\epsilon_n,
\end{aligned}
$$

where (a) follows from Fano's inequality, which states that for sufficiently large n, we have $H(U^n, V_1^n, V_2^n|W_1, W_2, W_0, Z^n) \leq h(P_{we}^{(n)}) + nP_{we}^n R_w \leq n\epsilon_n$. Here P_{we}^n denotes the wiretapper's error probability of decoding (u^n, v_1^n, v_2^n) in the case that the bin numbers w_0, w_1, and w_2 are known to the eavesdropper. Since the sum rate is small enough, then $P_{we}^n \to 0$ for sufficiently large n. (b) follows from the following Markov chain: $(W_1, W_2, W_0) \to (U^n, V_1^n, V_2^n) \to Z^n$. Hence, we have $H(Z^n|W_1, W_2, W_0, U^n, V_1^n, V_2^n) = H(Z^n|U^n, V_1^n, V_2^n)$. (c) follows from the fact that $H(W_1, W_2, W_0, U^n, V_1^n, V_2^n) \geq H(U^n, V_1^n, V_2^n)$. (d) follows from that fact that $H(U^n, V_1^n, V_2^n) = H(U^n) + H(V_1^n|U^n) + H(V_2^n|U^n) - I(V_1^n; V_2^n|U^n)$. (e) follows from the fact that $H(U^n) \geq \min\{I(U^n; Y_1^n), I(U^n; Y_2^n)\}$ and $H(V_i^n|U^n) \geq I(V_i^n; Y_i^n|U^n)$ for $i = 1, 2$. (f) follows from the fact that $I(U^n, V_1^n, V_2^n; Z^n) = I(U^n; Z^n) + I(V_1^n, V_2^n; Z^n|U^n)$. (g) follows from Lemma 11 in the Appendix. This completes the achievability proof.

2.6.2 Proof of Theorem 2

Achievablity: We need to show that the region of (2.18) and (2.19) is a subset of the achievability region of Theorem 1. In the achievability scheme of Theorem 1, if we set $\mathcal{W}_2 = \emptyset$ and rename W_0 with W_2, then using the degradedness, we obtain the following region,

$$\begin{aligned} R_1 + R_2 &\leq I(V;Y_1|U) - I(V;Z|U) + I(U;Y_2) - I(U;Z), \\ R_2 &\leq I(U;Y_2) - I(U;Z). \end{aligned} \tag{2.47}$$

Note that since the first receiver decodes both messages, the total rate of this receiver is $R_1 \leftarrow R_1 + R_2$ and we have

$$\begin{aligned} R_1 &\leq I(UV;Y_1|U) + I(U;Y_2) - I(UV;Z), \\ R_2 &\leq I(U;Y_2) - I(U;Z). \end{aligned} \tag{2.48}$$

Now, since $U \rightarrow V \rightarrow X \rightarrow Y_2 \rightarrow Z$ is a markov chain, then the following region is a subset of the above region, and consequently, it is achievable,

$$\begin{aligned} R_1 &\leq I(X;Y_1|U) + I(U;Z) - I(X;Z), \\ R_2 &\leq I(U;Y_2) - I(U;Z). \end{aligned} \tag{2.49}$$

which is the same as that of region (2.18) and (2.19). This completes the achievability proof.

Converse: The transmitter sends two independent secret messages W_1 and W_2 to Receiver 1 and Receiver 2, respectively. Let us define $U_i = (W_2, Y_1^{i-1})$. The following Lemma bounds the secrecy rates for a general case of $(W_1, W_2) \rightarrow X^n \rightarrow Y_1^n Y_2^n Z^n$:

Lemma 2. *For the broadcast channel with an eavesdropper, the perfect secrecy rates are bounded as follows,*

$$\begin{aligned} nR_1 &\leq \sum_{i=1}^{n} I(W_1;Y_{1i}|W_2, Z_i, Y_1^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3, \\ nR_2 &\leq \sum_{i=1}^{n} I(W_2;Y_{2i}|Z_i, Y_2^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2. \end{aligned} \tag{2.50}$$

Proof: We need to prove the second bound. The first bound can similarly be proven. nR_2 is bounded as follows:

$$\begin{aligned} nR_2 &\overset{(a)}{\leq} H(W_2|Z^n) + n\epsilon_2 \\ &\overset{(b)}{\leq} H(W_2|Z^n) - H(W_2|Y_2^n) + n\delta_1 + n\epsilon_2 \\ &= I(W_2;Y_2^n) - I(W_2;Z^n) + n\delta_1 + n\epsilon_2 \end{aligned} \tag{2.51}$$

where (a) follows from the secrecy constraint that $H(W_2|Z^n) \geq H(W_2) - n\epsilon_2$. (b) follows from Fano's inequality that $H(W_2|Y_2^n) \leq n\delta_1$. Next, we expand $I(W_2; Y_2^n)$ and $I(W_2; Z^n)$ as follows.

$$\begin{aligned} I(W_2; Y_2^n) &= \sum_{i=1}^{n} I(W_2; Y_{2i}|Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2, \widetilde{Z}^{i+1}; Y_{2i}|Y_2^{i-1}) - I(\widetilde{Z}^{i+1}; Y_{2i}|W_2, Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2; Y_{2i}|Y_2^{i-1}, \widetilde{Z}^{i+1}) + I(\widetilde{Z}^{i+1}; Y_{2i}|Y_2^{i-1}) - I(\widetilde{Z}^{i+1}; Y_{2i}|W_2, Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2; Y_{2i}|Y_2^{i-1}, \widetilde{Z}^{i+1}) + \Delta_1 - \Delta_2, \end{aligned} \tag{2.52}$$

where, $\Delta_1 = \sum_{i=1}^{n} I(\widetilde{Z}^{i+1}; Y_{2i}|Y_2^{i-1})$ and $\Delta_2 = \sum_{i=1}^{n} I(\widetilde{Z}^{i+1}; Y_{2i}|W_2, Y_2^{i-1})$. Similarly, we have,

$$\begin{aligned} I(W_2; Z^n) &= \sum_{i=1}^{n} I(W_2; Z_i|\widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2, Y_2^{i-1}; Z_i|\widetilde{Z}^{i+1}) - I(Y_2^{i-1}; Z_i|W_2, \widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2; Z_i|Y_2^{i-1}, \widetilde{Z}^{i+1}) + I(Y_2^{i-1}; Z_i|\widetilde{Z}^{i+1}) - I(Y_2^{i-1}; Z_i|W_2, \widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2; Z_i|Y_2^{i-1}, \widetilde{Z}^{i+1}) + \Delta_1^* - \Delta_2^*, \end{aligned} \tag{2.53}$$

where, $\Delta_1^* = \sum_{i=1}^{n} I(Y_2^{i-1}; Z_i|\widetilde{Z}^{i+1})$ and $\Delta_2^* = \sum_{i=1}^{n} I(Y_2^{i-1}; Z_i|W_2, \widetilde{Z}^{i+1})$. According to Lemma 7 of [4], $\Delta_1 = \Delta_1^*$ and $\Delta_2 = \Delta_2^*$. Thus, we have,

$$\begin{aligned} nR_2 &\leq \sum_{i=1}^{n} I(W_2; Y_{2i}|Y_2^{i-1}, \widetilde{Z}^{i+1}) - I(W_2; Z_i|Y_2^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &= \sum_{i=1}^{n} H(W_2|Z_i, Y_2^{i-1}, \widetilde{Z}^{i+1}) - H(W_2|Y_{2i}, Y_2^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &\overset{(a)}{\leq} \sum_{i=1}^{n} H(W_2|Z_i, Y_2^{i-1}, \widetilde{Z}^{i+1}) - H(W_2|Y_{2i}, Z_i, Y_2^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &= \sum_{i=1}^{n} I(W_2; Y_{2i}|Z_i, Y_2^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2, \end{aligned} \tag{2.54}$$

where (a) follows from the fact that conditioning always decreases the entropy. $\square$

Now according to the above Lemma, the secrecy rates are bounded as follows:

$$
\begin{aligned}
nR_1 &\overset{(a)}{\leq} \sum_{i=1}^{n} I(W_1; Y_{1,i}|W_2, Z_i, Y_1^{i-1}, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3 \qquad (2.55)\\
&= \sum_{i=1}^{n} I(W_1; Y_{1,i}|U_i, Z_i, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3\\
&\overset{(b)}{\leq} \sum_{i=1}^{n} I(X_i; Y_{1,i}|U_i, Z_i, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3\\
&\overset{(c)}{=} \sum_{i=1}^{n} I(X_i; Y_{1,i}, U_i, Z_i|\widetilde{Z}^{i+1}) - I(X_i; Z_i|\widetilde{Z}^{i+1}) - I(X_i; U_i|Z_i, \widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3\\
&\overset{(d)}{=} \sum_{i=1}^{n} I(X_i; Y_{1,i}|U_i, \widetilde{Z}^{i+1}) + I(X_i; U_i|\widetilde{Z}^{i+1}) - I(X_i; Z_i|\widetilde{Z}^{i+1}) - I(X_i; U_i|Z_i, \widetilde{Z}^{i+1})\\
&+ n\delta_1 + n\epsilon_3\\
&\overset{(e)}{=} \sum_{i=1}^{n} I(X_i; Y_{1,i}|U_i, \widetilde{Z}^{i+1}) - I(X_i; Z_i|\widetilde{Z}^{i+1}) + I(Z_i; U_i|\widetilde{Z}^{i+1}) - I(Z_i; U_i|X_i, \widetilde{Z}^{i+1})\\
&+ n\delta_1 + n\epsilon_3\\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_{1,i}|U_i, \widetilde{Z}^{i+1}) - I(X_i; Z_i|\widetilde{Z}^{i+1}) + I(Z_i; U_i|\widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_3,
\end{aligned}
$$

where (a) follows from the Lemma (2). (b) follows from the data processing theorem. (c) follows from the chain rule. (d) follows from the fact that $I(X_i; Y_{1,i}, U_i, Z_i|\widetilde{Z}^{i+1}) = I(X_i; U_i|\widetilde{Z}^{i+1}) + I(X_i; Y_{1,i}|U_i, \widetilde{Z}^{i+1}) + I(X_i; Z_i|Y_{1,i}, U_i, \widetilde{Z}^{i+1})$ and from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow X_i \rightarrow Y_{1,i} \rightarrow Y_{2,i} \rightarrow Z_i$ forms a Markov chain, which means that $I(X_i; Z_i|Y_{1,i}, U_i, \widetilde{Z}^{i+1}) = 0$. (e) follows from the fact that $I(X_i; U_i|\widetilde{Z}^{i+1}) - I(X_i; U_i|Z_i, \widetilde{Z}^{i+1}) = I(Z_i; U_i|\widetilde{Z}^{i+1}) - I(Z_i; U_i|X_i, \widetilde{Z}^{i+1})$. (f) follows from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow X_i \rightarrow Z_i$ forms a Markov chain. Thus, $I(Z_i; U_i\widetilde{Z}^{i+1}|X_i) = 0$ which implies that $I(Z_i; U_i|X_i, \widetilde{Z}^{i+1}) = 0$.

For the second receiver, we have

$$\begin{aligned}
nR_2 &\overset{(a)}{\leq} \sum_{i=1}^{n} I(W_2; Y_{2,i}|Y_2^{i-1}, Z_i, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1 \qquad (2.56)\\
&= \sum_{i=1}^{n} H(Y_{2,i}|Y_2^{i-1}, Z_i, \widetilde{Z}^{i+1}) - H(Y_{2,i}|W_2, Y_2^{i-1}, Z_i, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&\overset{(b)}{\leq} \sum_{i=1}^{n} H(Y_{2,i}|Z_i, \widetilde{Z}^{i+1}) - H(Y_{2,i}|W_2, Y_1^{i-1}, Y_2^{i-1}, Z_i, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&\overset{(c)}{=} \sum_{i=1}^{n} H(Y_{2,i}|Z_i, \widetilde{Z}^{i+1}) - H(Y_{2,i}|U_i, Z_i, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&= \sum_{i=1}^{n} I(Y_{2,i}; U_i|Z_i, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&= \sum_{i=1}^{n} I(Y_{2,i}; U_i|\widetilde{Z}^{i+1}) + I(Y_{2,i}; Z_i|U_i, \widetilde{Z}^{i+1}) - I(Y_{2,i}; Z_i|\widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&= \sum_{i=1}^{n} I(Y_{2,i}; U_i|\widetilde{Z}^{i+1}) - I(Z_i; U_i|\widetilde{Z}^{i+1}) + I(Z_i; U_i|Y_{2,i}, \widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1\\
&\overset{(d)}{=} \sum_{i=1}^{n} I(Y_{2,i}; U_i|\widetilde{Z}^{i+1}) - I(Z_i; U_i|\widetilde{Z}^{i+1}) + n\delta_2 + n\epsilon_1,
\end{aligned}$$

where (a) follows from the lemma (2). (b) follows from the fact that conditioning always decreases the entropy. (c) follows from the fact that $Y_2^{i-1} \rightarrow W_2\widetilde{Z}^{i+1}Y_1^{i-1} \rightarrow Y_{2i} \rightarrow Z_i$ forms a Markov chain. (d) follows from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow Y_{2,i} \rightarrow Z_i$ forms a Markov chain. Thus $I(Z_i; U_i\widetilde{Z}^{i+1}|Y_{2i}) = 0$ which implies that $I(Z_i; U_i|Y_{2i}, \widetilde{Z}^{i+1}) = 0$. Now, following [62], let us define the time sharing random variable Q which is uniformly distributed over $\{1, 2, ..., n\}$ and independent of $(W_1, W_2, X^n, Y_1^n, Y_2^n)$. Let us define $U = U_Q,\ V = (\widetilde{Z}^{Q+1}, Q),\ X = X_Q,\ Y_1 = Y_{1,Q},\ Y_2 = Y_{2,Q},\ Z = Z_Q$, then R_1 and R_2 can be written as

$$R_1 \leq I(X; Y_1|U, V) + I(U; Z|V) - I(X; Z|V), \qquad (2.57)$$

$$R_2 \leq I(U; Y_2|V) - I(U; Z|V). \qquad (2.58)$$

Note that the boundary of this region is characterized by the maximization of $R_1 + \mu R_2$ over this region for $\mu \geq 1$. On the other hand we have,

$$R_1 + \mu R_2 \leq I(X; Y_1|U, V) + I(U; Z|V) - I(X; Z|V) + \mu\left(I(U; Y_2|V) - I(U; Z|V)\right) \qquad (2.59)$$

Since conditional mutual information is the average of the unconditional ones, the largest region is achieved when V is a constant. This proves the converse part.

2.6.3 Proof of Theorem 3

Achievability: Let $U \sim \mathcal{N}(0,(1-\alpha)P)$ and $X^{'} \sim \mathcal{N}(0,\alpha P)$ be independent and $X = U + X^{'} \sim \mathcal{N}(0,P)$. Now consider the following secure superposition coding scheme:

1) *Codebook Generation*: Generate $2^{nI(U;Y_2)}$ i.i.d Gaussian codewords u^n with average power $(1-\alpha)P$ and randomly distribute these codewords into 2^{nR_2} bins. Then index each bin by $w_2 \in \{1,2,...,2^{nR_2}\}$. Generate an independent set of $2^{nI(X^{'};Y_1)}$ i.i.d Gaussian codewords $x^{'n}$ with average power αP. Then, randomly distribute them into 2^{nR_1} bins. Index each bin by $w_1 \in \{1,2,...,2^{nR_1}\}$.

2) *Encoding*: To send messages w_1 and w_2, the transmitter randomly chooses one of the codewords in bin w_2, (say u^n) and one of the codewords in bin w_1 (say $x^{'n}$). The transmitter then simply transmits $x^n = u^n + x^{'n}$.

3) *Decoding*: The received signal at the legitimate receivers are y_1^n and y_2^n, respectively. Receiver 2 determines the unique u^n such that (u^n, y_2^n) are jointly typical and declares the index of the bin containing u^n as the message received. If there is none of such or more than one of such, an error is declared. Receiver 1 uses the successive cancelation method; it first decodes u^n and subtracts it from y_1^n and then looks for the unique $x^{'n}$ such that $(x^{'n}, y_1^n - u^n)$ are jointly typical and declares the index of the bin containing $x^{'n}$ as the message received.

It can be shown that if R_1 and R_2 satisfy (2.20) and (2.21), the error probability analysis and equivocation calculation is straightforward and may therefore be omitted.

Converse: According to the previous section, R_2 is bounded as follows:

$$nR_2 \leq I(Y_2^n;U^n|Z^n) = h(Y_2^n|Z^n) - h(Y_2^n|U^n,Z^n), \tag{2.60}$$

where h is the differential entropy. The classical entropy power inequality states that:

$$2^{\frac{2}{n}h(Y_2^n+N_3^{'n})} \geq 2^{\frac{2}{n}h(Y_2^n)} + 2^{\frac{2}{n}h(N_3^{'n})} \tag{2.61}$$

Therefore, $h(Y_2^n|Z^n)$ may be written as follows:

$$\begin{aligned} h(Y_2^n|Z^n) &= h(Z^n|Y_2^n) + h(Y_2^n) - h(Z^n) \\ &= \frac{n}{2}\log 2\pi e(\sigma_3^2-\sigma_2^2) + h(Y_2^n) - h(Y_2^n + N_3^{'n}) \\ &\leq \frac{n}{2}\log 2\pi e(\sigma_3^2-\sigma_2^2) + h(Y_2^n) - \frac{n}{2}\log(2^{\frac{2}{n}h(Y_2^n)} + 2\pi e(\sigma_3^2-\sigma_2^2)). \end{aligned} \tag{2.62}$$

On the other hand, for any fixed $a \in \mathcal{R}$, the function

$$f(t,a) = t - \frac{n}{2}\log(2^{\frac{2}{n}t} + a) \tag{2.63}$$

is concave in t and has a global maximum at the maximum value of t. Thus, $h(Y_2^n|Z^n)$ is maximized when Y_2^n (or equivalently X^n) has a Gaussian distribution. Hence,

$$\begin{aligned} h(Y_2^n|Z^n) &\leq \frac{n}{2}\log 2\pi e(\sigma_3^2-\sigma_2^2)+\frac{n}{2}\log 2\pi e(P+\sigma_2^2)-\frac{n}{2}\log 2\pi e(P+\sigma_3^2) \\ &= \frac{n}{2}\log\left(\frac{2\pi e(\sigma_3^2-\sigma_2^2)(P+\sigma_2^2)}{P+\sigma_3^2}\right). \end{aligned} \tag{2.64}$$

Note that another method to obtain (2.64) is using the worst additive noise lemma (see [70, 71] for details). Now consider the term $h(Y_2^n|U^n,Z^n)$. This term is lower bounded with $h(Y_2^n|U^n,X^n,Z^n)=\frac{n}{2}\log 2\pi e(\sigma_2^2)$ which is greater than $\frac{n}{2}\log 2\pi e(\frac{\sigma_2^2(\sigma_3^2-\sigma_2^2)}{\sigma_3^2})$. Hence,

$$\frac{n}{2}\log 2\pi e(\frac{\sigma_2^2(\sigma_3^2-\sigma_2^2)}{\sigma_3^2}) \leq h(Y_2^n|U^n,Z^n) \leq h(Y_2^n|Z^n). \tag{2.65}$$

Inequalities (2.64) and (2.65) imply that there exists an $\alpha\in[0,1]$ such that

$$h(Y_2^n|U^n,Z^n)=\frac{n}{2}\log\left(\frac{2\pi e(\sigma_3^2-\sigma_2^2)(\alpha P+\sigma_2^2)}{\alpha P+\sigma_3^2}\right). \tag{2.66}$$

Substituting (2.66) and (2.64) into (2.60) yields the desired bound

$$\begin{aligned} nR_2 &\leq h(Y_2^n|Z^n)-h(Y_2^n|U^n,Z^n) && (2.67)\\ &\leq \frac{n}{2}\log\left(\frac{(P+\sigma_2^2)(\alpha P+\sigma_3^2)}{(P+\sigma_3^2)(\alpha P+\sigma_2^2)}\right) \\ &= nC\left(\frac{(1-\alpha)P}{\alpha P+\sigma_2^2}\right)-nC\left(\frac{(1-\alpha)P}{\alpha P+\sigma_3^2}\right). && (2.68) \end{aligned}$$

Note that the left hand side of (2.66) can be written as $h(Y_2^n,Z^n|U^n)-h(Z^n|U^n)$ which implies that

$$h(Y_2^n|U^n)-h(Z^n|U^n)=\frac{n}{2}\log\left(\frac{\alpha P+\sigma_2^2}{\alpha P+\sigma_3^2}\right). \tag{2.69}$$

Since $\sigma_1^2\leq\sigma_2^2\leq\sigma_3^2$, there exists a $0\leq\beta\leq 1$ such that $\sigma_2^2=(1-\beta)\sigma_1^2+\beta\sigma_3^2$, or equivalently, $\sigma_2^2=\sigma_1^2+\beta(\sigma_3^2-\sigma_1^2)$. Therefore, since $Y_1^n\to Y_2^n\to Z^n$ forms a Markov chain, the received signals Z^n and Y_2^n can be written as $Z^n=Y_1^n+\widetilde{N}^n$ and $Y_2^n=Y_1^n+\sqrt{\beta}\widetilde{N}^n$ where $\widetilde{N}$ is an independent Gaussian noise with variance $\widetilde{\sigma}^2=\sigma_3^2-\sigma_1^2$. All noises are Gaussian random n-vector with a positive definite covariance matrix. Costa's entropy power inequality [72] states that (see also [73] for its linear version),

$$2^{\frac{2}{n}h(Y_1^n+\sqrt{\beta}\widetilde{N}^n|U^n)} \geq (1-\beta)2^{\frac{2}{n}h(Y_1^n|U^n)}+\beta 2^{\frac{2}{n}h(Y_1^n+\widetilde{N}^n|U^n)} \tag{2.70}$$

for any random n-vector Y_1^n and Gaussian random n-vector of $\widetilde{N}^n$. Equivalently we have,

$$2^{\frac{2}{n}h(Y_2^n|U^n)} \geq (1-\beta)2^{\frac{2}{n}h(Y_1^n|U^n)} + \beta 2^{\frac{2}{n}h(Z^n|U^n)} \tag{2.71}$$

After some manipulations of (2.71), we obtain

$$\begin{aligned} & h(Y_1^n|U^n) - h(Z^n|U^n) \\ & \leq \frac{n}{2}\log\left(\frac{\alpha P + \sigma_2^2 - \beta(\alpha P + \sigma_3^2)}{(1-\beta)(\alpha P + \sigma_3^2)}\right) \\ & = \frac{n}{2}\log\left(\frac{\alpha P + \sigma_1^2}{\alpha P + \sigma_3^2}\right). \end{aligned} \tag{2.72}$$

The rate R_1 is bounded as follows

$$\begin{aligned} nR_1 &\leq I(X^n; Y_1^n|U^n) - I(X^n; Z^n) + I(U^n; Z^n) \\ &= h(Y_1^n|U^n) - h(Y_1^n|X^n, U^n) + h(Z^n|X^n) - h(Z^n|U^n) \\ &= h(Y_1^n|U^n) - h(Z^n|U^n) + \frac{n}{2}\log(\frac{\sigma_3^2}{\sigma_1^2}) \\ &\overset{(a)}{\leq} \frac{n}{2}\log\left(\frac{\alpha P + \sigma_1^2}{\alpha P + \sigma_3^2}\frac{\sigma_3^2}{\sigma_1^2}\right) \\ &= nC\left(\frac{\alpha P}{\sigma_1^2}\right) - nC\left(\frac{\alpha P}{\sigma_3^2}\right), \end{aligned} \tag{2.73}$$

where (a) follows from (2.72).

Chapter 3

Secure Gaussian MIMO Broadcast Channel

In this chapter, we consider a scenario where a source node wishes to broadcast two confidential messages for two respective receivers via a Gaussian MIMO broadcast channel. A wiretapper also receives the transmitted signal via another MIMO channel. First, we assume that the channels are degraded and the wiretapper has the worst channel. We establish the capacity region of this scenario. Our achievability scheme is a combination of the superposition of Gaussian codes and randomization within the layers, which we will refer to as the Secret Superposition Coding. For the outerbound, we use notion of the enhanced channels to show that the secret superposition of Gaussian codes is optimal. We show that we only need to enhance the channels of the legitimate receivers, and the channel of the eavesdropper remains unchanged. We then extend the result of the degraded case to a non-degraded case. We show that the secret superposition of Gaussian codes, along with successive decoding, cannot work when the channels are not degraded. We develop a Secret Dirty Paper Coding (SDPC) scheme and show that SDPC is optimal for this channel. We then present a corollary generalizing the capacity region of the two receivers' case to the case of multiple receivers. Finally, we investigate a scenario which frequently occurs in the practice of wireless networks. In this scenario, the transmitter and the eavesdropper have multiple antennae, while both intended receivers have a single antenna (representing resource limited mobile units). We characterize the secrecy capacity region in terms of generalized eigenvalues of the receivers' channels and the eavesdropper channel. We refer to this configuration as the MISOME (Multiple-Input-Single-Output-Multiple-Eavesdropper) case. We then present a corollary generalizing the results of the two receivers' case to multiple receivers. In the high SNR regime, we show that the capacity region is a convex closure of rectangular regions.

3.1 Introduction

Recently, significant research has been conducted in both theoretical and practical aspects of wireless communication systems with Multiple-Input Multiple-Output (MIMO) antennae. Most works have focused on the role of MIMO in enhancing the throughput and robustness of such systems. In this work, however, we focus on the role of such multiple antennae in enhancing wireless security.

The capacity region of the conventional Gaussian MIMO broadcast channel is studied in [74] by Weingarten *et. al.* The notion of an enhanced broadcast channel is introduced in [74] and is used jointly with entropy power inequality to characterize the capacity region of the degraded vector Gaussian broadcast channel. They have shown that the superposition of Gaussian codes is optimal for the degraded vector Gaussian broadcast channel, and that dirty-paper coding is optimal for the nondegraded case.

We have published/submitted the results of this chapter in [53] and [54]. Parallel to our work, references [59] and [60] independently considered the secure MIMO broadcast channel and established its capacity region. Reference [59] used the relationships between the minimum-mean-square-error and the mutual information, and equivalently, the relationships between the Fisher information and the differential entropy to provide the converse proof. Reference [60] considered the vector Gaussian MIMO broadcast channel with and without an external eavesdropper. They presented a vector generalization of Costa's Entropy Power Inequality to provide their converse proof. In our proof, however, we enhance the channels properly and show that the enhanced channels are proportional. We then use the proportionality characteristic to provide the converse proof.

3.2 Preliminaries

Consider a Secure Gaussian Multiple-Input Multiple-Output Broadcast Channel (SGMBC), as depicted in Figure 3.1. In this setting, the transmitter wishes to send two independent messages (W_1, W_2) to the respective receivers in n uses of the channel while preventing the eavesdropper from having any information about the messages. At a specific time, the signals received by the destinations and the eavesdropper are given by

$$\begin{aligned} \mathbf{y_1} &= \mathbf{H}_1\mathbf{x} + \mathbf{n_1} \\ \mathbf{y_2} &= \mathbf{H}_2\mathbf{x} + \mathbf{n_2} \\ \mathbf{z} &= \mathbf{H}_3\mathbf{x} + \mathbf{n_3}, \end{aligned} \tag{3.1}$$

where

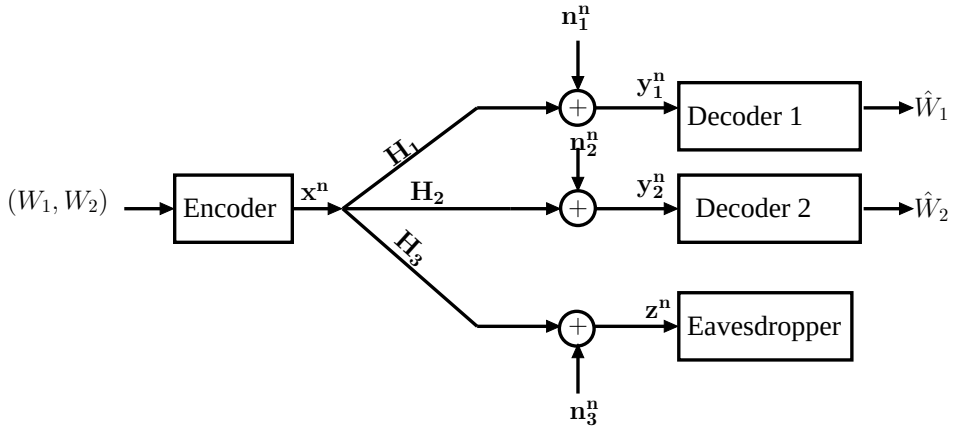

Figure 3.1: Secure Gaussian MIMO Broadcast Channel

- $\mathbf{x}$ is a real input vector of size $t \times 1$ under an input covariance constraint. We require that $E[\mathbf{x}\mathbf{x}^T] \preceq \mathbf{S}$ for a positive semi-definite matrix $\mathbf{S} \succeq 0$. Here, $\prec, \preceq, \succ$, and $\succeq$ represent partial ordering between symmetric matrices where $\mathbf{B} \succeq \mathbf{A}$ means that $(\mathbf{B} - \mathbf{A})$ is a positive semi-definite matrix.

- $\mathbf{y_1}$, $\mathbf{y_2}$, and $\mathbf{z}$ are real output vectors which are received by the destinations and the eavesdropper, respectively. These are vectors of size $r_1 \times 1$, $r_2 \times 1$, and $r_3 \times 1$, respectively.

- $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are fixed, real gain matrices which model the channel gains between the transmitter and the receivers. These are matrices of size $r_1 \times t$, $r_2 \times t$, and $r_3 \times t$, respectively. The channel state information is assumed to be known perfectly at the transmitter and at all the receivers.

- $\mathbf{n_1}$, $\mathbf{n_2}$ and $\mathbf{n_3}$ are real Gaussian random vectors with zero means and covariance matrices $\mathbf{N_1} = E[\mathbf{n_1}\mathbf{n_1}^T] \succ 0$, $\mathbf{N_2} = E[\mathbf{n_2}\mathbf{n_2}^T] \succ 0$, and $\mathbf{N_3} = E[\mathbf{n_3}\mathbf{n_3}^T] \succ 0$, respectively.

Let W_1 and W_2 denote the the message indices of user 1 and user 2, respectively. Furthermore, let $\mathbf{x^n}$, $\mathbf{y_1^n}$, $\mathbf{y_2^n}$, and $\mathbf{z^n}$ denote the channel input and channel output matrices over a block of n samples. Let $\mathbf{n_1^n}$, $\mathbf{n_2^n}$, and $\mathbf{n_3^n}$ denote the additive noise components for these channels. Thus,

$$\begin{aligned} \mathbf{y_1^n} &= \mathbf{H}_1\mathbf{x^n} + \mathbf{n_1^n} \\ \mathbf{y_2^n} &= \mathbf{H}_2\mathbf{x^n} + \mathbf{n_2^n} \\ \mathbf{z^n} &= \mathbf{H}_3\mathbf{x^n} + \mathbf{n_3^n}. \end{aligned} \tag{3.2}$$

Note that $\mathbf{n_i^n}$ is an $r_i \times n$ random matrix and $\mathbf{H}_i$ is an $r_i \times t$ deterministic matrix where $i = 1, 2, 3$. The columns of $\mathbf{n_i^n}$ are independent Gaussian random vectors with covariance matrices $\mathbf{N_i}$ for $i = 1, 2, 3$. In addition $\mathbf{n_i^n}$ is independent of $\mathbf{x^n}$, W_1 and W_2. A $((2^{nR_1}, 2^{nR_2}), n)$ code for the above channel consists of a stochastic encoder

$$f : (\{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_2}\}) \rightarrow \mathcal{X}^{\mathbf{n}}, \tag{3.3}$$

and two decoders,

$$g_1 : \mathcal{Y}_{\mathbf{1}}^{\mathbf{n}} \rightarrow \{1, 2, ..., 2^{nR_1}\}, \tag{3.4}$$

and

$$g_2 : \mathcal{Y}_{\mathbf{2}}^{\mathbf{n}} \rightarrow \{1, 2, ..., 2^{nR_2}\}. \tag{3.5}$$

The average probability of error is defined as the probability that the decoded messages are not equal to the transmitted messages; that is,

$$P_e^{(n)} = P(g_1(\mathbf{y_1^n}) \neq W_1 \cup g_2(\mathbf{y_2^n}) \neq W_2). \tag{3.6}$$

The secrecy levels of confidential messages W_1 and W_2 are measured at the eavesdropper in terms of equivocation rates, which are defined as follows.

Definition 6. *The equivocation rates R_{e1}, R_{e2} and R_{e12} for the secure broadcast channel are:*

$$\begin{aligned} R_{e1} &= \frac{1}{n} H(W_1|\mathbf{z^n}) \\ R_{e2} &= \frac{1}{n} H(W_2|\mathbf{z^n}) \\ R_{e12} &= \frac{1}{n} H(W_1, W_2|\mathbf{z^n}). \end{aligned} \tag{3.7}$$

The perfect secrecy rates R_1 and R_2 are the amount of information that can be sent to the legitimate receivers reliably *and* confidentially.

The model presented in (3.1) is a SGMBC. However, we will initially consider two subclasses of this channel and then generalize our results for the SGMBC.

The first subclass that we will consider is the Secure Aligned Degraded MIMO Broadcast Channel (SADBC). The MIMO broadcast channel of (3.1) is said to be aligned if the number of transmit antennae is equal to the number of receive antennae at each of the users and the eavesdropper ($t = r_1 = r_2 = r_3$) and the gain matrices are all identity matrices ($\mathbf{H}_1 = \mathbf{H}_2 = \mathbf{H}_3 = \mathbf{I}$). Furthermore, if the additive noise vectors' covariance matrices are ordered such that $0 \prec \mathbf{N_1} \preceq \mathbf{N_2} \preceq \mathbf{N_3}$, then the channel is SADBC.

The second subclass we consider is a generalization of the SADBC. The MIMO broadcast channel of (3.1) is said to be Secure Aligned MIMO Broadcast Channel (SAMBC) if it is aligned and not necessarily degraded. In other words, the additive noise vector covariance matrices are not necessarily ordered. A time sample of an SAMBC is given by the following expressions,

$$\begin{aligned} \mathbf{y_1} &= \mathbf{x} + \mathbf{n_1} \\ \mathbf{y_2} &= \mathbf{x} + \mathbf{n_2} \\ \mathbf{z} &= \mathbf{x} + \mathbf{n_3}, \end{aligned} \tag{3.8}$$

where, $\mathbf{y_1}$, $\mathbf{y_2}$, $\mathbf{z}$, $\mathbf{x}$ are real vectors of size $t \times 1$ and $\mathbf{n_1}$, $\mathbf{n_2}$, and $\mathbf{n_3}$ are independent and real Gaussian noise vectors such that $\mathbf{N_i} = \mathbf{E}[\mathbf{n_i}\mathbf{n_i}^{\mathbf{T}}] \succ 0_{t\times t}$ for $i = 1, 2, 3$.

3.3 The Capacity Region of The SADBC

In this section, we characterize the capacity region of the SADBC. In Chapter 2, we considered the degraded broadcast channel with confidential messages and establish its secrecy capacity region. The following remark recalls this result from Chapter 2.

Remark 5. *The capacity region for transmitting independent secret messages over the degraded broadcast channel is the convex hull of the closure of all* (R_1, R_2) *satisfying*

$$R_1 \leq I(X; Y_1|U) - I(X; Z|U) \tag{3.9}$$

$$R_2 \leq I(U; Y_2) - I(U; Z), \tag{3.10}$$

for some joint distribution $P(u)P(x|u)P(y_1, y_2, z|x)$.

Note that evaluating (3.9) and (3.10) involves solving a functional, nonconvex optimization problem. Usually nontrivial techniques and strong inequalities are used to solve optimization problems of this type. Indeed, for the single antenna case, we successfully evaluated the capacity expression of (3.9) and (3.10) in [52]. Liu *et. al.* in [12] evaluated the capacity expression of the MIMO wiretap channel by using the channel enhancement method. In the following section, we state and prove our result for the capacity region of SADBC.

First, we define the achievable rate region due to Gaussian codebook under a covariance matrix constraint $\mathbf{S} \succeq 0$. The achievability scheme of Remark 5 is the secret superposition of Gaussian codes and successive decoding at the first receiver. According to the above Remark, for any covariance matrix input constraint $\mathbf{S}$ and two semi-definite matrices $\mathbf{B_1} \succeq 0$

and $\mathbf{B_2} \succeq 0$ such that $\mathbf{B_1} + \mathbf{B_2} \preceq \mathbf{S}$, it is possible to achieve the following rates:

$$R_1^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) = \frac{1}{2}\left[\log\left|\mathbf{N_1^{-1}}(\mathbf{B_1} + \mathbf{N_1})\right| - \frac{1}{2}\log\left|\mathbf{N_3^{-1}}(\mathbf{B_1} + \mathbf{N_3})\right|\right]^+ \tag{3.11}$$

$$R_2^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) = \frac{1}{2}\left[\log\frac{|\mathbf{B_1} + \mathbf{B_2} + \mathbf{N_2}|}{|\mathbf{B_1} + \mathbf{N_2}|} - \frac{1}{2}\log\frac{|\mathbf{B_1} + \mathbf{B_2} + \mathbf{N_3}|}{|\mathbf{B_1} + \mathbf{N_3}|}\right]^+.$$

The Gaussian rate region of SADBC is defined as follows.

Definition 7. *Let* $\mathbf{S}$ *be a positive semi-definite matrix. Then, the Gaussian rate region of SADBC under a covariance matrix constraint* $\mathbf{S}$ *is given by*

$$\mathcal{R}^G(\mathbf{S}, \mathbf{N_{1,2,3}}) = \left\{ \begin{array}{l} \left(R_1^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}), R_2^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}})\right) \mid \\ s.t \;\; \mathbf{S} - (\mathbf{B_1} + \mathbf{B_2}) \succeq 0, \;\; \mathbf{B_k} \succeq 0, \;\; k = 1, 2 \end{array} \right\}. \tag{3.12}$$

We will show that $\mathcal{R}^G(\mathbf{S}, \mathbf{N_{1,2,3}})$ is the capacity region of the SADBC; however, certain preliminaries need to be addressed first. We begin by characterizing the boundary of the Gaussian rate region.

Remark 6. *Note that in characterizing the capacity region of the conventional Gaussian MIMO broadcast channel Weingarten* et. al. *[74] have proven that on the boundary of the above region we have* $\mathbf{B_1} + \mathbf{B_2} = \mathbf{S}$ *which maximizes the rate* R_2*. In our argument, however, the boundary is not characterized with this equality as rate* R_2 *may decrease by increasing* $\mathbf{B_1} + \mathbf{B_2}$*.*

Definition 8. *The rate vector* $R^* = (R_1, R_2)$ *is said to be an optimal Gaussian rate vector under the covariance matrix* $\mathbf{S}$*, if* $R^* \in \mathcal{R}^G(\mathbf{S}, \mathbf{N_{1,2,3}})$ *and if there is no other rate vector* $R'^* = (R_1', R_2') \in \mathcal{R}^G(\mathbf{S}, \mathbf{N_{1,2,3}})$ *such that* $R_1' \geq R_1$ *and* $R_2' \geq R_2$ *where at least one of the inequalities is strict. The set of positive semi-definite matrices* $(\mathbf{B_1^*}, \mathbf{B_2^*})$ *such that* $\mathbf{B_1^*} + \mathbf{B_2^*} \preceq \mathbf{S}$ *is said to be realizing matrices of an optimal Gaussian rate vector if the rate vector* $\left(R_1^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}), R_2^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}})\right)$ *is an optimal Gaussian rate vector.*

In general, there is no known closed form solution for the realizing matrices of an optimal Gaussian rate vector. Note that finding an optimal Gaussian rate vector again involves solving a nonconvex optimization problem. The realizing matrices of an optimal Gaussian rate vector, $\mathbf{B_1^*}, \mathbf{B_2^*}$ are the solution of the following optimization problem:

$$\begin{array}{l} \max\limits_{(\mathbf{B_1}, \mathbf{B_2})} R_1^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) + \mu R_2^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) \\ \text{s.t} \;\; \mathbf{B_1} \succeq 0, \quad \mathbf{B_2} \succeq 0, \quad \mathbf{B_1} + \mathbf{B_2} \preceq \mathbf{S}, \end{array} \tag{3.13}$$

where $\mu \geq 1$. Next, we define a class of enhanced channels. The enhanced channel has some fundamental properties which help us to characterize the secrecy capacity region. We will discuss its properties further on.

Definition 9. *A SADBC with noise covariance matrices* $(\mathbf{N}_1^{'}, \mathbf{N}_2^{'}, \mathbf{N}_3^{'})$ *is an enhanced version of another SADBC with noise covariance matrices* $(\mathbf{N_1}, \mathbf{N_2}, \mathbf{N_3})$ *if*

$$\mathbf{N}_1^{'} \preceq \mathbf{N_1}, \quad \mathbf{N}_2^{'} \preceq \mathbf{N_2}, \quad \mathbf{N}_3^{'} = \mathbf{N_3}, \quad \mathbf{N}_1^{'} \preceq \mathbf{N}_2^{'}. \tag{3.14}$$

It is apparent that the capacity region of the enhanced version contains the capacity region of the original channel. Note that in characterizing the capacity region of the conventional Gaussian MIMO broadcast channel, all channels must be enhanced by reducing the noise covariance matrices. In our scheme, however, we only enhance the channels for the legitimate receivers and the channel of the eavesdropper remains unchanged. This is due to the fact that the capacity region of the enhanced channel must contain the original capacity region. Reducing the noise covariance matrix of the eavesdropper's channel, however, may reduce the secrecy capacity region. The following theorem connects the definitions of the optimal Gaussian rate vector and the enhanced channel.

Theorem 4. *Consider a SADBC with positive definite noise covariance matrices* $(\mathbf{N_1}, \mathbf{N_2}, \mathbf{N_3})$. *Let* $\mathbf{B}_1^*$ *and* $\mathbf{B}_2^*$ *be realizing matrices of an optimal Gaussian rate vector under a transmit covariance matrix constraint* $\mathbf{S} \succ 0$. *There then exists an enhanced SADBC with noise covariance matrices* $(\mathbf{N}_1^{'}, \mathbf{N}_2^{'}, \mathbf{N}_3^{'})$ *for which the following properties hold.*

1. *Enhancement:*
   $\mathbf{N}_1^{'} \preceq \mathbf{N_1}, \quad \mathbf{N}_2^{'} \preceq \mathbf{N_2}, \quad \mathbf{N}_3^{'} = \mathbf{N_3}, \quad \mathbf{N}_1^{'} \preceq \mathbf{N}_2^{'}$,

2. *Proportionality: There exists an* $\alpha \geq 0$ *and a matrix* $\mathbf{A}$ *such that,*
   $(\mathbf{I} - \mathbf{A})(\mathbf{B}_1^* + \mathbf{N}_1^{'}) = \alpha \mathbf{A}(\mathbf{B}_1^* + \mathbf{N}_3^{'})$,

3. *Rate and optimality preservation:*
   $R_k^G(\mathbf{B}_{1,2}^*, \mathbf{N}_{1,2,3}) = R_k^G(\mathbf{B}_{1,2}^*, \mathbf{N}_{1,2,3}^{'}) \quad \forall k = 1, 2$, *furthermore,* $\mathbf{B}_1^*$ *and* $\mathbf{B}_2^*$ *are realizing matrices of an optimal Gaussian rate vector in the enhanced channel.*

Proof: The realizing matrices $\mathbf{B}_1^*$ and $\mathbf{B}_2^*$ are the solution of the optimization problem of (3.13). Using the Lagrange Multiplier method, this constraint optimization problem is equivalent to the following unconstrained optimization problem:

$$\begin{aligned} \max_{(\mathbf{B_1}, \mathbf{B_2})} & R_1^G(\mathbf{B}_{1,2}, \mathbf{N}_{1,2,3}) + \mu R_2^G(\mathbf{B}_{1,2}, \mathbf{N}_{1,2,3}) + Tr\{\mathbf{B_1 O_1}\} \\ & + Tr\{\mathbf{B_2 O_2}\} + Tr\{(\mathbf{S} - \mathbf{B_1} - \mathbf{B_2})\mathbf{O_3}\}, \end{aligned} \tag{3.15}$$

where $\mathbf{O_1}$, $\mathbf{O_2}$, and $\mathbf{O_3}$ are positive semi-definite $t \times t$ matrices such that $Tr\{\mathbf{B}_1^*\mathbf{O_1}\} = 0$, $Tr\{\mathbf{B}_2^*\mathbf{O_2}\} = 0$, and $Tr\{(\mathbf{S} - \mathbf{B}_1^* - \mathbf{B}_2^*)\mathbf{O_3}\} = 0$. As all $\mathbf{B_k^*}$, $k = 1, 2$, $\mathbf{O_i}$, $i = 1, 2, 3$, and

$\mathbf{S}-\mathbf{B}_1^*-\mathbf{B}_2^*$ are positive semi-definite matrices, then we must have $\mathbf{B_k^*O_k}=0,\quad k=1,2$ and $(\mathbf{S}-\mathbf{B}_1^*-\mathbf{B}_2^*)\mathbf{O_3}=0$. According to the necessary KKT conditions, and after some manipulations, we have:

$$(\mathbf{B}_1^*+\mathbf{N_1})^{-1}+(\mu-1)(\mathbf{B}_1^*+\mathbf{N_3})^{-1}+\mathbf{O_1}=\mu(\mathbf{B}_1^*+\mathbf{N_2})^{-1}+\mathbf{O_2} \tag{3.16}$$

$$\mu(\mathbf{B}_1^*+\mathbf{B}_2^*+\mathbf{N_2})^{-1}+\mathbf{O_2}=\mu(\mathbf{B}_1^*+\mathbf{B}_2^*+\mathbf{N_3})^{-1}+\mathbf{O_3}. \tag{3.17}$$

We choose the noise covariance matrices of the enhanced SADBC as follows:

$$\begin{aligned}\mathbf{N_1^{'}}&=\left(\mathbf{N_1}^{-1}+\mathbf{O_1}\right)^{-1}\\ \mathbf{N_2^{'}}&=\left((\mathbf{B}_1^*+\mathbf{N_2})^{-1}+\frac{1}{\mu}\mathbf{O_2}\right)^{-1}-\mathbf{B}_1^*\\ \mathbf{N_3^{'}}&=\mathbf{N_3}.\end{aligned} \tag{3.18}$$

As $\mathbf{O_1}\succeq 0$ and $\mathbf{O_2}\succeq 0$, then the above choice has the enhancement property. Note that

$$\begin{aligned}\left((\mathbf{B}_1^*+\mathbf{N_1})^{-1}+\mathbf{O_1}\right)^{-1}&=\left((\mathbf{B}_1^*+\mathbf{N_1})^{-1}\left(\mathbf{I}+(\mathbf{B}_1^*+\mathbf{N_1})\,\mathbf{O_1}\right)\right)^{-1}\\ &\overset{(a)}{=}(\mathbf{I}+\mathbf{N_1O_1})^{-1}(\mathbf{B}_1^*+\mathbf{N_1})-\mathbf{B}_1^*+\mathbf{B}_1^*\\ &=(\mathbf{I}+\mathbf{N_1O_1})^{-1}\left((\mathbf{B}_1^*+\mathbf{N_1})-(\mathbf{I}+\mathbf{N_1O_1})\,\mathbf{B}_1^*\right)+\mathbf{B}_1^*\\ &\overset{(b)}{=}(\mathbf{I}+\mathbf{N_1O_1})^{-1}\mathbf{N_1}+\mathbf{B}_1^*\\ &=\left(\mathbf{N_1}\left(\mathbf{N_1^{-1}}+\mathbf{O_1}\right)\right)^{-1}\mathbf{N_1}+\mathbf{B}_1^*\\ &=\left(\mathbf{N_1^{-1}}+\mathbf{O_1}\right)^{-1}+\mathbf{B}_1^*\\ &=\mathbf{B}_1^*+\mathbf{N_1^{'}},\end{aligned} \tag{3.19}$$

where (a) and (b) follows from the fact that $\mathbf{B_1^*O_1}=0$. Therefore, according to (3.16), the following property holds for the enhanced channel,

$$(\mathbf{B}_1^*+\mathbf{N_1^{'}})^{-1}+(\mu-1)(\mathbf{B}_1^*+\mathbf{N_3^{'}})^{-1}=\mu(\mathbf{B}_1^*+\mathbf{N_2^{'}})^{-1}.$$

Since $\mathbf{N_1^{'}}\preceq\mathbf{N_2^{'}}\preceq\mathbf{N_3^{'}}$ then there exists a matrix $\mathbf{A}$ such that $\mathbf{N_2^{'}}=(\mathbf{I}-\mathbf{A})\mathbf{N_1^{'}}+\mathbf{AN_3^{'}}$ where $\mathbf{A}=(\mathbf{N_2^{'}}-\mathbf{N_1^{'}})(\mathbf{N_3^{'}}-\mathbf{N_1^{'}})^{-1}$. Therefore, the above equation can be written as:

$$\begin{aligned}&(\mathbf{B}_1^*+\mathbf{N_1^{'}})^{-1}+(\mu-1)(\mathbf{B}_1^*+\mathbf{N_3^{'}})^{-1}=\\ &\mu\left[(\mathbf{I}-\mathbf{A})(\mathbf{B}_1^*+\mathbf{N_1^{'}})+\mathbf{A}(\mathbf{B}_1^*+\mathbf{N_3^{'}})\right]^{-1}.\end{aligned} \tag{3.20}$$

Let $(\mathbf{I}-\mathbf{A})(\mathbf{B}_1^*+\mathbf{N_1^{'}})=\alpha\mathbf{A}(\mathbf{B}_1^*+\mathbf{N_3^{'}})$, then after some manipulations, the above equation becomes

$$\frac{1}{\alpha}\mathbf{I}+(\mu-1-\frac{1}{\alpha})\mathbf{A}=\frac{\mu}{\alpha+1}\mathbf{I}. \tag{3.21}$$

The above equation is satisfied by $\alpha = \frac{1}{\mu-1}$ which completes the proportionality property. We can now prove the rate conservation property. The expression $\frac{\left|\mathbf{B_1^*}+\mathbf{N_1'}\right|}{\left|\mathbf{N_1'}\right|}$ can be written as follows:

$$\begin{aligned}
\frac{\left|\mathbf{B_1^*}+\mathbf{N_1'}\right|}{\left|\mathbf{N_1'}\right|} &= \frac{|\mathbf{I}|}{\left|\mathbf{N_1'}\left(\mathbf{B_1^*}+\mathbf{N_1'}\right)^{-1}\right|} \\
&= \frac{|\mathbf{I}|}{\left|\left(\mathbf{B_1^*}+\mathbf{N_1'}-\mathbf{B_1^*}\right)\left(\mathbf{B_1^*}+\mathbf{N_1'}\right)^{-1}\right|} \\
&= \frac{|\mathbf{I}|}{\left|\mathbf{I}-\mathbf{B_1^*}\left(\mathbf{B_1^*}+\mathbf{N_1'}\right)^{-1}\right|} \\
&= \frac{|\mathbf{I}|}{|\mathbf{I}-\mathbf{B_1^*}\left((\mathbf{B_1^*}+\mathbf{N_1})^{-1}+\mathbf{O_1}\right)|} \\
&\overset{(a)}{=} \frac{|\mathbf{I}|}{\left|\mathbf{I}-\mathbf{B_1^*}\left(\mathbf{B_1^*}+\mathbf{N_1}\right)^{-1}\right|} \\
&= \frac{|\mathbf{B_1^*}+\mathbf{N_1}|}{|\mathbf{N_1}|},
\end{aligned} \tag{3.22}$$

where (a) once again follows from the fact that $\mathbf{B_1^*O_1} = 0$. To complete the proof of rate conservation, consider the following equalities:

$$\begin{aligned}
\frac{\left|\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2'}\right|}{\left|\mathbf{B_1^*}+\mathbf{N_2'}\right|} &= \frac{\left|\mathbf{B_2^*}\left(\mathbf{B_1^*}+\mathbf{N_2'}\right)^{-1}+\mathbf{I}\right|}{|\mathbf{I}|} \\
&= \frac{\left|\mathbf{B_2^*}\left(\left(\mathbf{B_1^*}+\mathbf{N_2}\right)^{-1}+\frac{1}{\mu}\mathbf{O_2}\right)+\mathbf{I}\right|}{|\mathbf{I}|} \\
&\overset{(a)}{=} \frac{|\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2}|}{|\mathbf{B_1^*}+\mathbf{N_2}|},
\end{aligned} \tag{3.23}$$

where (a) follows from the fact $\mathbf{B_2^*O_2} = 0$. Therefore, according to (3.22), (3.23), and the fact that $\mathbf{N_3'} = \mathbf{N_3}$, the rate preservation property holds for the enhanced channel. To prove the optimality preservation, we need to show that $(\mathbf{B_1^*}, \mathbf{B_2^*})$ are also realizing matrices of an optimal Gaussian rate vector in the enhanced channel. To establish this point, we show that the necessary KKT conditions for the enhanced channel coincides with the KKT conditions

for the original channel. The expression $\mu(\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2^{'}})^{-1}$ can be written as follows:

$$\begin{aligned}
\mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2^{'}}\right)^{-1} &\overset{(a)}{=} \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}+\left(\mathbf{N_2}^{-1}+\frac{1}{\mu}\mathbf{O_2}\right)^{-1}\right)^{-1} \\
&= \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}\left(\mathbf{I}+\mathbf{B_2^*}^{-1}\left(\mathbf{N_2}^{-1}+\frac{1}{\mu}\mathbf{O_2}\right)^{-1}\right)\right)^{-1} \\
&= \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}\left(\mathbf{I}+\left(\left(\mathbf{N_2}^{-1}+\frac{1}{\mu}\mathbf{O_2}\right)\mathbf{B_2^*}\right)^{-1}\right)\right)^{-1} \\
&\overset{(b)}{=} \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}\left(\mathbf{I}+\left(\mathbf{N_2}^{-1}\mathbf{B_2^*}\right)^{-1}\right)\right)^{-1} \\
&= \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}\left(\mathbf{I}+\mathbf{B_2^*}^{-1}\mathbf{N_2}\right)\right)^{-1} \\
&= \mu\left(\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2}\right)^{-1},
\end{aligned} \tag{3.24}$$

where (a) follows from the definition of $\mathbf{N_2^{'}}$ and (b) follows from the fact that $\mathbf{B_2^*O_2}=0$. Therefore, according to (3.19) and the above equation, the KKT conditions of (3.16) and (3.17) for the original channel can be written as follows for the enhanced channel:

$$(\mathbf{B_1^*}+\mathbf{N_1^{'}})^{-1}+(\mu-1)(\mathbf{B_1^*}+\mathbf{N_3^{'}})^{-1}=\mu(\mathbf{B_1^*}+\mathbf{N_2^{'}})^{-1} \tag{3.25}$$

$$\mu(\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_2^{'}})^{-1}=\mu(\mathbf{B_1^*}+\mathbf{B_2^*}+\mathbf{N_3^{'}})^{-1}+\mathbf{O_3}-\mathbf{O_2}, \tag{3.26}$$

where $\mathbf{O_3}-\mathbf{O_2}\succeq 0$. Therefore, $R_1^G(\mathbf{B_{1,2}},\mathbf{N_{1,2,3}^{'}})+\mu R_2^G(\mathbf{B_{1,2}},\mathbf{N_{1,2,3}^{'}})$ is maximized when $\mathbf{B_k}=\mathbf{B_k^*}$ for $k=1,2$. □

We can now use Theorem 4 to prove that $\mathcal{R}^G(\mathbf{S},\mathbf{N_{1,2,3}})$ is the capacity region of the SADBC. We follow Bergman's approach [75] to prove a contradiction. Note that since the original channel is not proportional, we cannot apply Bergman's proof to the original channel directly. Here we apply his proof to the enhanced channel instead.

Theorem 5. *Consider a SADBC with positive definite noise covariance matrices* $(\mathbf{N_1},\mathbf{N_2},\mathbf{N_3})$. *Let* $\mathcal{C}(\mathbf{S},\mathbf{N_{1,2,3}})$ *denote the capacity region of the SADBC under a covariance matrix constraint* $\mathbf{S}\succ 0$ *. Then,* $\mathcal{C}(\mathbf{S},\mathbf{N_{1,2,3}})=\mathcal{R}^G(\mathbf{S},\mathbf{N_{1,2,3}})$.

Proof: The achievability scheme is secret superposition coding with Gaussian codebook. For the converse proof, we use a contradiction argument and assume that there exists an achievable rate vector $\bar{R}=(R_1,R_2)$ which is not in the Gaussian region. We can apply the steps of Bergman's proof of [75] to the enhanced channel and show that this assumption is impossible. Since $\bar{R}\notin\mathcal{R}^G(\mathbf{S},\mathbf{N_{1,2,3}})$, there exist realizing matrices of an optimal Gaussian

rate vector $\mathbf{B_1^*}, \mathbf{B_2^*}$ such that

$$\begin{aligned} R_1 &\geq R_1^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}) \\ R_2 &\geq R_2^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}) + b, \end{aligned} \tag{3.27}$$

for some $b > 0$. We know by Theorem 4 that for every set of realizing matrices of an optimal Gaussian rate vector $\mathbf{B_1^*}, \mathbf{B_2^*}$, there exists an enhanced SADBC with noise covariance matrices $\mathbf{N_1^{'}}, \mathbf{N_2^{'}}$, such that the proportionality and rate preservation properties hold. According to the rate preservation property, we have $R_k^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2}}) = R_k^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2}^{'}}), \quad k = 1, 2$. Therefore, the preceding expression can be rewritten as follows:

$$\begin{aligned} R_1 &\geq R_1^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}) = R_1^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}^{'}}) \\ R_2 &\geq R_2^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}) + b = R_2^G(\mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}^{'}}) + b. \end{aligned} \tag{3.28}$$

According to the Theorem 5, R_1 and R_2 are bounded as follows:

$$\begin{aligned} R_1 &\leq h(\mathbf{y_1}|\mathbf{u}) - h(\mathbf{z}|\mathbf{u}) - (h(\mathbf{y_1}|\mathbf{x}, \mathbf{u}) - h(\mathbf{z}|\mathbf{x}, \mathbf{u})) \\ R_2 &\leq h(\mathbf{y_2}) - h(\mathbf{z}) - (h(\mathbf{y_2}|\mathbf{u}) - h(\mathbf{z}|\mathbf{u})). \end{aligned} \tag{3.29}$$

Let $\mathbf{y_1^{'}}$ and $\mathbf{y_2^{'}}$ denote the enhanced channel outputs of each of the receiving users. As $\mathbf{u} \rightarrow \mathbf{y_k^{'}} \rightarrow \mathbf{y_k}$ forms a Markov chain for $k = 1, 2$ and $\mathbf{z}^{'} = \mathbf{z}$, then we can use the data processing inequality to rewrite the above region as follows:

$$\begin{aligned} R_1 &\leq h(\mathbf{y_1^{'}}|\mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{u}) - \Big(h(\mathbf{y_1^{'}}|\mathbf{x}, \mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{x}, \mathbf{u})\Big) \\ &= h(\mathbf{y_1^{'}}|\mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{u}) - \frac{1}{2}\Big(\log|\mathbf{N_1^{'}}| - \log|\mathbf{N_3^{'}}|)\Big) \end{aligned} \tag{3.30}$$

$$R_2 \leq h(\mathbf{y_2^{'}}) - h(\mathbf{z}^{'}) - \Big(h(\mathbf{y_2^{'}}|\mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{u})\Big). \tag{3.31}$$

Now, the inequalities of (3.28) and (3.30) have shifted to the enhanced channel.

Since $R_1 > R_1^G(\mathbf{B_{1,2}}, \mathbf{N_{1,2,3}^{'}})$, the inequality (3.30) means that

$$h(\mathbf{y_1^{'}}|\mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{u}) > \frac{1}{2}\Big(\log|\mathbf{B_1^*} + \mathbf{N_1^{'}}| - \log|\mathbf{B_1^*} + \mathbf{N_3^{'}}|)\Big). \tag{3.32}$$

By the definition of matrix $\mathbf{A}$ and since $\mathbf{y_1^{'}} \rightarrow \mathbf{y_2^{'}} \rightarrow \mathbf{z}^{'}$ forms a Morkov chain, the received signals $\mathbf{z}^{'}$ and $\mathbf{y_2^{'}}$ can be written as $\mathbf{z}^{'} = \mathbf{y_1^{'}} + \widetilde{\mathbf{n}}$ and $\mathbf{y_2^{'}} = \mathbf{y_1^{'}} + \mathbf{A}^{\frac{1}{2}}\widetilde{\mathbf{n}}$ where $\widetilde{\mathbf{n}}$ is an independent Gaussian noise with covariance matrix $\widetilde{\mathbf{N}} = \mathbf{N_3^{'}} - \mathbf{N_1^{'}}$. According to Costa's Entropy Power Inequality and the previous inequality, we have

$$\begin{aligned} h(\mathbf{y_2^{'}}|\mathbf{u}) - h(\mathbf{z}^{'}|\mathbf{u}) &\geq \frac{t}{2}\log\left(|\mathbf{I} - \mathbf{A}|^{\frac{1}{t}}2^{\frac{2}{t}\left(h(\mathbf{y_1^{'}}|\mathbf{u}) - h(\mathbf{z}|^{'}\mathbf{u})\right)} + |\mathbf{A}|^{\frac{1}{t}})\right) \\ &> \frac{t}{2}\log\left(\frac{|\mathbf{I} - \mathbf{A}|^{\frac{1}{t}}|\mathbf{B_1^*} + \mathbf{N_1^{'}}|^{\frac{1}{t}}}{|\mathbf{B_1^*} + \mathbf{N_3^{'}}|^{\frac{1}{t}}} + |\mathbf{A}|^{\frac{1}{t}})\right) \\ &\overset{(a)}{=} \frac{1}{2}\log(\mathbf{B_1^*} + \mathbf{N_2^{'}}) - \frac{1}{2}\log(\mathbf{B_1^*} + \mathbf{N_3^{'}}), \end{aligned} \tag{3.33}$$

where (a) is due to the proportionality property. Using (3.31) and the fact that $R_2 > R_2^G(\mathbf{B_{1,2}}, \mathbf{N'_{1,2,3}})$, the inequality (3.31) means that

$$\begin{aligned} &h(\mathbf{y_2'}) - h(\mathbf{z'}) \geq R_2 + h(\mathbf{y_2'}|\mathbf{u}) - h(\mathbf{z'}|\mathbf{u}) > \\ &\frac{1}{2}\log(\mathbf{B_1^*} + \mathbf{B_2^*} + \mathbf{N_2'}) - \frac{1}{2}\log(\mathbf{B_1^*} + \mathbf{B_2^*} + \mathbf{N_3'}). \end{aligned} \tag{3.34}$$

On the other hand, Gaussian distribution maximizes $h(\mathbf{x} + \mathbf{n_2}) - h(\mathbf{x} + \mathbf{n_3})$ (See [71]) and $(\mathbf{B_1^*}, \mathbf{B_2^*})$ satisfying the KKT conditions of (3.26). Therefore, the above inequality is a contradiction. □

3.4 The Capacity Region of the SAMBC

In this section, we characterize the secrecy capacity region of the aligned (but not necessarily degraded) MIMO broadcast channel. Note that since the SAMBC is not degraded, there is no single optimozation formula for its capacity region. In addition, the secret superposition of Gaussian codes along with successive decoding cannot work when the channel is not degraded. In [50], we presented an achievable rate region for the General Secure Broadcast Channel. Our achievable coding scheme is based on a combination of the random binning and the Gelfand-Pinsker binning schemes. We first review this scheme and based on this result, we develop an achievable secret coding scheme for the SAMBC. Based on Theorem 4, we then provide a full characterization of the capacity region of SAMBC.

3.4.1 Secret Dirty-Paper Coding Scheme and Achievability Proof

In [50], we established an achievable rate region for the general secure broadcast channel. This scheme enables both joint encoding at the transmitter by using Gelfand-Pinsker binning and preserving confidentiality by using random binning. The following theorem summarizes the encoding strategy.

Theorem 6. *: Let V_1 and V_2 be auxiliary random variables and Ω be the class of joint probability densities $P(v_1, v_2 x, y_1, y_2, z)$ that factors as $P(v_1, v_2)P(x|v_1, v_2)P(y_1, y_2, z|x)$. Let $\mathcal{R}_I(\pi)$ denote the union of all non-negative rate pairs (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &\leq I(V_1; Y_1) - I(V_1; Z) \\ R_2 &\leq I(V_2; Y_2) - I(V_2; Z) \\ R_1 + R_2 &\leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1, V_2; Z) - I(V_1; V_2), \end{aligned} \tag{3.35}$$

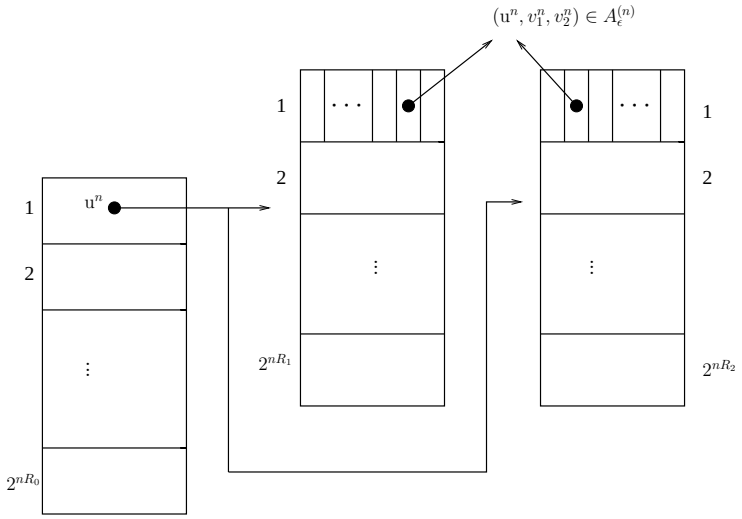

Figure 3.2: The Stochastic Encoder

for a given joint probability density $\pi \in \Omega$. For the general broadcast channel with confidential messages, the following region is achievable

$$\mathcal{R}_I = conv\left\{\bigcup_{\pi\in\Omega} \mathcal{R}_I(\pi)\right\}, \tag{3.36}$$

where conv is the convex closure operator.

Remark 7. *If we remove the secrecy constraints by removing the eavesdropper, then the above rate region becomes Marton's achievable region for the general broadcast channel.*

Proof: 1) *Codebook Generation*: The structure of the encoder is depicted in Figure 3.2. Fix $P(v_1)$, $P(v_2)$ and $P(x|v_1, v_2)$. The stochastic encoder generates $2^{n(I(V_1;Y_1)-\epsilon)}$ independent and identically distributed sequences v_1^n according to the distribution $P(v_1^n) = \prod_{i=1}^n P(v_{1,i})$. Next, randomly distribute these sequences into 2^{nR_1} bins such that each bin contains $2^{n(I(V_1;Z)-\epsilon)}$ codewords. Similarly, it generates $2^{n(I(V_2;Y_2)-\epsilon)}$ independent and identically distributed sequences v_2^n according to the distribution $P(v_2^n) = \prod_{i=1}^n P(v_{2,i})$. Randomly distribute these sequences into 2^{nR_2} bins such that each bin contains $2^{n(I(V_2;Z)-\epsilon)}$ codewords. Index each of the above bins by $w_1 \in \{1, 2, ..., 2^{nR_1}\}$ and $w_2 \in \{1, 2, ..., 2^{nR_2}\}$, respectively.

2) *Encoding*: To send messages w_1 and w_2, the transmitter looks for v_1^n in bin w_1 of the first bin set and looks for v_2^n in bin w_2 of the second bin set, such that $(v_1^n, v_2^n) \in A_\epsilon^{(n)}(P_{V_1,V_2})$ where $A_\epsilon^{(n)}(P_{V_1,V_2})$ denotes the set of jointly typical sequences v_1^n and v_2^n with respect to $P(v_1, v_2)$. The rates are such that there exist more than one joint typical pair. The transmitter randomly chooses one of them and then generates x^n according to $P(x^n|v_1^n, v_2^n) =$

$\prod_{i=1}^{n} P(x_i|v_{1,i}, v_{2,i})$. This scheme is equivalent to the scenario in which each bin is divided into subbins and the transmitter randomly chooses one of the subbins of bin w_1 and one of the subbins of bin w_2. It then looks for a joint typical sequence (v_1^n, v_2^n) in the corresponding subbins and generates x^n.

3) *Decoding*: The received signals at the legitimate receivers, y_1^n and y_2^n, are the outputs of the channels $P(y_1^n|x^n) = \prod_{i=1}^{n} P(y_{1,i}|x_i)$ and $P(y_2^n|x^n) = \prod_{i=1}^{n} P(y_{2,i}|x_i)$, respectively. The first receiver looks for the unique sequence v_1^n such that (v_1^n, y_1^n) is jointly typical and declares the index of the bin containing v_1^n as the message received. The second receiver uses the same method to extract the message w_2.

4) *Error Probability Analysis*: Since the region of (3.35) is a subset of Marton region, then the error probability analysis is the same as [69].

5) *Equivocation Calculation*: Please see the proof of Theorem 1. □

The achievability scheme in Theorem 6 introduces random binning. When we want to construct the rate region of (3.36), however, it is not clear how to choose the auxiliary random variables V_1 and V_2. Here, we employ the Dirty-Paper Coding (DPC) technique to develop the secret DPC (SDPC) achievable rate region for the SAMBC. We consider a secret dirty-paper encoder with Gaussian codebooks as follows.

First, we separate the channel input $\mathbf{x}$ into two random vectors $\mathbf{b_1}$ and $\mathbf{b_2}$ such that

$$\mathbf{b_1} + \mathbf{b_2} = \mathbf{x}. \tag{3.37}$$

Here, $\mathbf{b_1}$ and $\mathbf{b_2}$ and $\mathbf{v_1}$ and $\mathbf{v_2}$ are chosen as follows:

$$\begin{aligned} \mathbf{b_1} &\sim \mathcal{N}(0, \mathbf{B_1}) \\ \mathbf{b_2} &\sim \mathcal{N}(0, \mathbf{B_2}) \\ \mathbf{v_2} &= \mathbf{b_2} \\ \mathbf{v_1} &= \mathbf{b_1} + \mathbf{C}\mathbf{b_2}, \end{aligned} \tag{3.38}$$

where $\mathbf{B_1} = E[\mathbf{b_1}\mathbf{b_1^T}] \succeq 0$ and $\mathbf{B_2} = E[\mathbf{b_2}\mathbf{b_2^T}] \succeq 0$ are covariance matrices such that $\mathbf{B_1} + \mathbf{B_2} \preceq \mathbf{S}$, and the matrix $\mathbf{C}$ is given as follows:

$$\mathbf{C} = \mathbf{B_1}\left(\mathbf{N_1} + \mathbf{B_1}\right)^{-1}. \tag{3.39}$$

By substituting (3.38) into the Theorem 6, we obtain the following SDPC rate region for the SAMBC.

Lemma 3. *(SDPC Rate Region): Let* $\mathbf{S}$ *be a positive semi-definite matrix then the following SDPC rate region of an SAMBC with a covariance matrix constraint* $\mathbf{S}$ *is achievable*

$$\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}}) = conv\left\{\bigcup_{\pi\in\prod} \mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,2,3}})\right\}, \tag{3.40}$$

where $\prod$ *is the collection of all possible permutations of the ordered set* $\{1, 2\}$*, conv is the convex closure operator and* $\mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,2,3}})$ *is given as follows:*

$$\mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,2,3}}) = \left\{ \begin{array}{l} (R_1, R_2) \,\big| R_k = R^{SDPC}_{\pi^{-1}(k)}(\pi, \mathbf{B_{1,2}}, \mathbf{N_{1,2,3}})\ k = 1, 2 \\ s.t\ \ \mathbf{S} - (\mathbf{B_1} + \mathbf{B_2}) \succeq 0,\ \mathbf{B_1} \succeq 0,\ \mathbf{B_2} \succeq 0 \end{array} \right\}, \tag{3.41}$$

where

$$R^{SDPC}_{\pi^{-1}(k)}(\pi, \mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) = \frac{1}{2}\left[\log\frac{\left|\sum_{i=1}^{\pi^{-1}(k)} \mathbf{B}_{\pi(\mathbf{i})} + \mathbf{N_k}\right|}{\left|\sum_{i=1}^{\pi^{-1}(k)-1} \mathbf{B}_{\pi(\mathbf{i})} + \mathbf{N_k}\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{\pi^{-1}(k)} \mathbf{B}_{\pi(\mathbf{i})} + \mathbf{N_3}\right|}{\left|\sum_{i=1}^{\pi^{-1}(k)-1} \mathbf{B}_{\pi(\mathbf{i})} + \mathbf{N_3}\right|}\right]^+.$$

Note that for the identity permutation, π_I, where $\pi_I(k) = k$, we have,

$$\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}}) = \mathcal{R}^{G}(\mathbf{S}, \mathbf{N_{1,2,3}}). \tag{3.42}$$

Proof: We prove the lemma for the case of identity permutation $\pi_I = \{1, 2\}$. This proof can similarly be used for the case that $\pi = \{2, 1\}$. According to the Theorem 6, we have,

$$\begin{aligned} R_1 &\leq \min\{I(V_1; Y_1) - I(V_1; Z), I(V_1; Y_1) + I(V_2; Z) - I(V_1, V_2; Z) - I(V_1; V_2)\} \\ &\overset{(a)}{\leq} \min\{I(V_1; Y_1) - I(V_1; Z), I(V_1; Y_1) - I(V_1; Z|V_2) - I(V_1; V_2)\} \\ &\overset{(b)}{\leq} I(V_1; Y_1) - I(V_1; Z|V_2) - I(V_1; V_2) \\ R_2 &\leq I(V_2; Y_2) - I(V_2; Z), \end{aligned} \tag{3.43}$$

where (a) follows from the fact that $I(V_1, V_2; Z) = I(V_2; Z) + I(V_1; Z|V_2)$ and (b) follows from the fact that $I(V_1; Z|V_2) + I(V_1; V_2) = I(Z, V_2; V_1) \geq I(Z; V_1)$. To calculate the upper-bound of R_1, we need to review the following lemma which has been noted by several authors [76].

Lemma 4. *Let* $\mathbf{y_1} = \mathbf{b_1} + \mathbf{b_2} + \mathbf{n_1}$*, where* $\mathbf{b_1}$*,* $\mathbf{b_2}$ *and* $\mathbf{n_1}$ *are Gaussian random vectors with covariance matrices* $\mathbf{B_1}$*,* $\mathbf{B_2}$ *and* $\mathbf{N_1}$*, respectively. Let* $\mathbf{b_1}$*,* $\mathbf{b_2}$ *and* $\mathbf{n_1}$ *be independent, and let* $\mathbf{v_1} = \mathbf{b_1} + \mathbf{C}\mathbf{b_2}$*, where* $\mathbf{C}$ *is an* $t \times t$ *matrix. Then an optimal matrix* $\mathbf{C}$ *which maximizes* $I(\mathbf{v_1}; \mathbf{y_1}) - I(\mathbf{v_1}; \mathbf{b_2})$ *is* $\mathbf{C} = \mathbf{B_1}(\mathbf{N_1} + \mathbf{B_1})^{-1}$*. Further more, the maximum value of* $I(\mathbf{v_1}; \mathbf{y_1} - I(\mathbf{v_1}; \mathbf{b_2})$ *is* $I(\mathbf{v_1}; \mathbf{y_1}|\mathbf{b_2})$*.*

Now, using the above Lemma and substituting (3.38) into (3.43), we obtain the following achievable rate region when $\pi = \pi_I$.

$$R_1 \leq \frac{1}{2}\left[\log\left|\mathbf{N_1^{-1}}(\mathbf{B_1}+\mathbf{N_1})\right| - \frac{1}{2}\log\left|\mathbf{N_3^{-1}}(\mathbf{B_1}+\mathbf{N_3})\right|\right]^+ \tag{3.44}$$
$$R_2 \leq \frac{1}{2}\left[\log\frac{|\mathbf{B_1}+\mathbf{B_2}+\mathbf{N_2}|}{|\mathbf{B_1}+\mathbf{N_2}|} - \frac{1}{2}\log\frac{|\mathbf{B_1}+\mathbf{B_2}+\mathbf{N_3}|}{|\mathbf{B_1}+\mathbf{N_3}|}\right]^+.$$

□

3.4.2 SAMBC- Converse Proof

For the converse part, note that not all points on the boundary of $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ can be directly obtained using a single SDPC scheme. Instead, we must use time-sharing between points corresponding to different permutations. Therefore, unlike the SADBC case, we cannot use a similar notion to the optimal Gaussian rate vectors, as not all the boundary points can immediately be characterized as a solution of an optimization problem. Instead, as the SDPC region is convex by definition, we use the notion of supporting hyperplanes of [74] to define this region.

In this section, we first define the supporting hyperplane of a closed and bounded set. We then present the relation between the ideas of a supporting hyperplane and the enhanced channel in Theorem 7. This theorem is an extension of Theorem 4 to the SAMBC case. Finally, we use Theorem 7 to prove that $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ is indeed the capacity region of the SAMBC.

Definition 10. *The set $\{\overline{R} = (R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$, for fixed and given scalars γ_1, γ_2 and, b, is a supporting hyperplane of a closed and bounded set $\mathcal{X} \subset \mathbb{R}^m$, if $\gamma_1 R_1 + \gamma_2 R_2 \leq b$ $\forall (R_1, R_2) \in \mathcal{X}$, with equality for at least one rate vector $(R_1, R_2) \in \mathcal{X}$.*

Note that as $\mathcal{X}$ is closed and bounded, $\max_{(R_1,R_2)\in\mathcal{X}} \gamma_1 R_1 + \gamma_2 R_2$, exists for any γ_1, γ_2. Thus, we always can find a supporting hyperplane for the set $\mathcal{X}$. As $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ is a closed and convex set, for each rate pair of $\overline{R}^o = (R_1^o, R_2^o) \notin \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ which lies outside the set, there exists a separating hyperplane $\{(R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$ where $\gamma_1 \geq 0, \gamma_1 \geq 0, b \geq 0$ and,

$$\gamma_1 R_1 + \gamma_2 R_2 \leq b \quad \forall (R_1, R_2) \in \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}}) \tag{3.45}$$
$$\gamma_1 R_1^o + \gamma_2 R_2^o > b.$$

The following theorem illustrates the relation between the ideas of enhanced channel and a supporting hyperplane.

Theorem 7. *Consider a SAMBC with noise covariance matrices* $(\mathbf{N_1}, \mathbf{N_2}, \mathbf{N_3})$ *and an average transmit covariance matrix constraint* $\mathbf{S} \succ 0$. *Assume that* $\{(R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$ *is a supporting hyperplane of the rate region* $\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}})$ *such that* $0 \leq \gamma_1 \leq \gamma_2$, $\gamma_2 > 0$ *and* $b \geq 0$. *Then, there exists an enhanced SADBC with noise covariance matrices* $(\mathbf{N_1'}, \mathbf{N_2'}, \mathbf{N_3'})$ *such that the following properties hold.*

1. *Enhancement:*
   $\mathbf{N_1'} \preceq \mathbf{N_1}, \quad \mathbf{N_2'} \preceq \mathbf{N_2}, \quad \mathbf{N_3'} = \mathbf{N_3}, \quad \mathbf{N_1'} \preceq \mathbf{N_2'},$

2. *Supporting hyperplane preservation:*
   $\{(R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$ *is also a supporting hyperplane of the rate region* $\mathcal{R}^G(\mathbf{S}, \mathbf{N'_{1,2,3}})$

Proof: To prove this theorem, we can follow the steps of the proof of Theorem 4. Assume that the hyperplane $\{(R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$ touches the region $\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}})$ at the point (R_1^*, R_2^*). Let $\mathbf{B_1^*}, \mathbf{B_2^*}$ be two positive semi-definite matrices such that $\mathbf{B_1^*} + \mathbf{B_2^*} \preceq \mathbf{S}$ and such that

$$R_k^{SDPC}(\pi_I, \mathbf{B_{1,2}^*}, \mathbf{N_{1,2,3}}) = R_k^*, \qquad k = 1, 2. \tag{3.46}$$

By definition of the supporting hyperplane, the scalar b and the matrices $(\mathbf{B_1^*}, \mathbf{B_2^*})$ are the solution of the following optimization problem:

$$\begin{aligned} &\max_{\mathbf{B_1}, \mathbf{B_2}} \gamma_1 R_1^{SDPC}(\pi_I, \mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) + \gamma_2 R_2^{SDPC}(\pi_I, \mathbf{B_{1,2}}, \mathbf{N_{1,2,3}}) \\ &\text{s.t} \quad \mathbf{B_1} + \mathbf{B_2} \preceq \mathbf{S} \quad \mathbf{B_k} \succeq 0 \qquad k = 1, 2. \end{aligned} \tag{3.47}$$

We define the noise covariance matrices of the enhanced SADBC as (3.18). Since for the permutation $\pi = \pi_I$ we have $\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}}) = \mathcal{R}^G(\mathbf{S}, \mathbf{N_{1,2,3}})$, the supporting hyperplane $\{(R_1, R_2) | \gamma_1 R_1 + \gamma_2 R_2 = b\}$ is also a supporting hyperplane of the rate region $\mathcal{R}^G(\mathbf{S}, \mathbf{N'_{1,2,3}})$. □

We can now use Theorem 7 and the capacity result of the SADBC to prove that $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ is indeed the capacity region of the SAMBC. The following theorem formally states the main result of this section.

Theorem 8. *Consider a SAMBC with positive definite noise covariance matrices* $(\mathbf{N_1}, \mathbf{N_2}, \mathbf{N_3})$. *Let* $\mathcal{C}(\mathbf{S}, \mathbf{N_{1,2,3}})$ *denote the capacity region of the SAMBC under a covariance matrix constraint* $\mathbf{S} \succ 0$. *Then,* $\mathcal{C}(\mathbf{S}, \mathbf{N_{1,2,3}}) = \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$.

Proof: To prove this theorem, we use Theorem 7 to show that for every rate vector $\overline{R}^o$, which lies outside the region $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$, we can find an enhanced SADBC, whose capacity region does not contain $\overline{R}^o$. As the capacity region of the enhanced channel outer bounds that of the original channel, $\overline{R}^o$ cannot be an achievable rate vector.

Let $\overline{R}^o = (R_1^o, R_2^o)$ be a rate vector which lies outside the region $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$. There exists a supporting and separating hyperplane $\{(R_1, R_2)|\gamma_1 R_1 + \gamma_2 R_2 = b\}$ where $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, and at least one of the γ_k's is positive. Without loss of generality, we assume that $\gamma_2 \geq \gamma_1$. If this is not the case, we can always reorder the indices of the users, such that this assumption will hold. By definition of the region $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$, we have,

$$\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}}) \subseteq \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}}). \tag{3.48}$$

Note that, as $\{(R_1, R_2)|\gamma_1 R_1 + \gamma_2 R_2 = b\}$ is a supporting hyperplane of $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$, we can write,

$$\begin{aligned} b' &= \max_{(R_1,R_2)\in\mathcal{R}^{SDPC}(\pi_I,\mathbf{S},\mathbf{N_{1,2,3}})} \gamma_1 R_1 + \gamma_2 R_2 \\ &\leq \max_{(R_1,R_2)\in\mathcal{R}^{SDPC}(\mathbf{S},\mathbf{N_{1,2,3}})} \gamma_1 R_1 + \gamma_2 R_2 = b. \end{aligned} \tag{3.49}$$

Furthermore, we can also write,

$$\gamma_1 R_1^o + \gamma_2 R_2^o > b \geq b'. \tag{3.50}$$

Therefore, the hyperplane of $\{(R_1, R_2)|\gamma_1 R_1 + \gamma_2 R_2 = b'\}$ is a supporting and separating hyperplane for the rate region $\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}})$. By Theorem 7, we know that there exists an enhanced SADBC whose Gaussian rate region $\mathcal{R}^G(\mathbf{S}, \mathbf{N'_{1,2,3}})$ lies under the supporting hyperplane and hence $(R_1^o, R_2^o) \notin \mathcal{R}^G(\mathbf{S}, \mathbf{N'_{1,2,3}})$. Therefore, (R_1^o, R_2^o) must lie outside the capacity region of the enhanced SADBC. To complete the proof, note that the capacity region of the enhanced SADBC contains that of the original channel and therefore, (R_1^o, R_2^o) must lie outside the capacity region of the original SAMBC. As this statement is true for all rate vectors which lie outside $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$, we have $\mathcal{C}(\mathbf{S}, \mathbf{N_{1,2,3}}) \subseteq \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$. However, $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$ is the set of achievable rates and therefore, $\mathcal{C}(\mathbf{S}, \mathbf{N_{1,2,3}}) = \mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,2,3}})$. □

Using the same discussion as [74], the result of SAMBC can extend to the SGMBC which is omitted here. The results of the secrecy capacity region for two receivers can be extended for m receivers as follows.

Corollary 1. *Consider a SGMBC with m receivers and one external eavesdropper. Let $\mathbf{S}$ be a positive semi-definite matrix. The SDPC rate region of $\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,\ldots,m}}, \mathbf{H_{1,\ldots,m}})$, which*

is defined by the following convex closure is indeed the secrecy capacity region of the SGMBC under a covariance constraint $\mathbf{S}$.

$$\mathcal{R}^{SDPC}(\mathbf{S}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}}) = conv\left\{ \bigcup_{\pi \in \prod} \mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}}) \right\}, \tag{3.51}$$

where $\prod$ *is the collection of all possible permutations of the ordered set* $\{1, ..., m\}$*, conv is the convex closure operator and* $\mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}})$ *is given as follows:*

$$\mathcal{R}^{SDPC}(\pi, \mathbf{S}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}}) = \left\{ \begin{array}{l} (R_1, R_2) \,\big| R_k = R^{SDPC}_{\pi^{-1}(k)}(\pi, \mathbf{B_{1,...,m}}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}}) \\ s.t \;\; \mathbf{S} - \sum_{i=1}^{m} \mathbf{B}_i \succeq 0, \;\; \mathbf{B}_i \succeq 0, \;\; k, i = 1, ..., m. \end{array} \right\},$$

where

$$\begin{aligned} R^{SDPC}_{\pi^{-1}(k)}(\pi, \mathbf{B_{1,...,m}}, \mathbf{N_{1,...,m}}, \mathbf{H_{1,...,m}}) = \frac{1}{2}\Big[\log \frac{\left| \mathbf{H_k} \left(\sum_{i=1}^{\pi^{-1}(k)} \mathbf{B}_{\pi(\mathbf{i})} \right) \mathbf{H_k^\dagger} + \mathbf{N_k} \right|}{\left| \mathbf{H_k} \left(\sum_{i=1}^{\pi^{-1}(k)-1} \mathbf{B}_{\pi(\mathbf{i})} \right) \mathbf{H_k^\dagger} + \mathbf{N_k} \right|} \\ - \frac{1}{2} \log \frac{\left| \mathbf{H_3} \left(\sum_{i=1}^{\pi^{-1}(k)} \mathbf{B}_{\pi(\mathbf{i})} \right) \mathbf{H_3^\dagger} + \mathbf{N_3} \right|}{\left| \mathbf{H_3} \left(\sum_{i=1}^{\pi^{-1}(k)-1} \mathbf{B}_{\pi(\mathbf{i})} \right) \mathbf{H_3^\dagger} + \mathbf{N_3} \right|} \Big]^+. \end{aligned} \tag{3.52}$$

3.5 Multiple-Input Single-Outputs Multiple Eavesdropper (MISOME) Channel

In this section we investigate a scenario in which the transmitter and the eavesdropper have multiple antennae, while both intended receivers have a single antenna. We refer to this configuration as the MISOME case. The significance of this model is apparent when a base station wishes to broadcast secure information to small mobile units. In this scenario small mobile units have single antenna while the base station and the eavesdropper can afford multiple antennae. We characterize the secrecy capacity region in terms of generalized eigenvalues of the receivers' channels and the eavesdropper channel. Note that in this case our analysis is valid for a general number of single antenna receivers. We can rewrite the signals received by the destination and the eavesdropper for the MISOME channel as follows:

$$\begin{aligned} y_1 &= \mathbf{h}_1^\dagger \mathbf{x} + n_1 \\ y_2 &= \mathbf{h}_2^\dagger \mathbf{x} + n_2 \\ \mathbf{z} &= \mathbf{H}_3 \mathbf{x} + \mathbf{n_3}, \end{aligned} \tag{3.53}$$

where $\mathbf{h}_1$ and $\mathbf{h}_2$ are fixed, real matrices which model the channel gains between the transmitter and the legitimate receivers. These are matrices of size $t \times 1$. The channel state

information again is assumed to be known perfectly at the transmitter and at all the receivers. Here, the superscript † denotes the Hermitian transpose of a vector. Without loss of generality, we assume that n_1 and n_2 are i.i.d real Gaussian random variables with zero means and unit covariances, i.e., $n_1, n_2 \sim \mathcal{N}(0,1)$. Furthermore, we assume that $\mathbf{n_3}$ is a Gaussian random vector with zero mean and covariance matrix $\mathbf{I}$. In this section, we assume that the input $\mathbf{x}$ satisfies a total power constraint of P, i.e.,

$$Tr\{E(\mathbf{x}\mathbf{x}^T)\} \leq P.$$

Before we state our results for the MISOME channel, we need to review some properties of generalized eigenvalues and eigenvectors. For more details of this topic, see, e.g.,[77].

Definition 11. *(Generalized eigenvalue-eigenvector) Let* $\mathbf{A}$ *be a Hermitian matrix and* $\mathbf{B}$ *be a positive definite matrix.* $(\lambda, \boldsymbol{\psi})$ *is a generalized eigenvalue-eigenvector pair if it satisfies the following equation.*

$$\mathbf{A}\boldsymbol{\psi} = \lambda \mathbf{B}\boldsymbol{\psi}. \tag{3.54}$$

Note that as $\mathbf{B}$ is invertible, the generalized eigenvalues and eigenvectors of the pair $(\mathbf{A}, \mathbf{B})$ are the regular eigenvalues and eigenvectors of the matrix $\mathbf{B}^{-1}\mathbf{A}$. The following Lemma describes the variational characterization of the generalized eigenvalue-eigenvector pair.

Lemma 5. *(Variational Characterization) Let* $r(\boldsymbol{\psi})$ *be the Rayleigh quotient defined as follows:*

$$r(\boldsymbol{\psi}) = \frac{\boldsymbol{\psi}^\dagger \mathbf{A}\boldsymbol{\psi}}{\boldsymbol{\psi}^\dagger \mathbf{B}\boldsymbol{\psi}}. \tag{3.55}$$

The generalized eigenvectors of $(\mathbf{A}, \mathbf{B})$ *are the stationary point solution of the Rayleigh quotient* $r(\boldsymbol{\psi})$*. Specifically, the largest generalized eigenvalue* $\lambda_{\max}$ *is the maximum of the Rayleigh quotient* $r(\boldsymbol{\psi})$ *and the optimum is attained by the eigenvector* $\boldsymbol{\psi}_{\max}$ *which corresponds to* $\lambda_{\max}$*, i.e.,*

$$\max_{\boldsymbol{\psi}} r(\boldsymbol{\psi}) = \frac{\boldsymbol{\psi}^\dagger{}_{\max} \mathbf{A}\boldsymbol{\psi}_{\max}}{\boldsymbol{\psi}^\dagger{}_{\max} \mathbf{B}\boldsymbol{\psi}_{\max}} = \lambda_{\max}. \tag{3.56}$$

Now consider the MISOME channel of (3.53). Assume that $0 \leq \alpha \leq 1$ and P are fixed. Let us define the following matrices for this channel

$$\begin{aligned} \mathbf{A_{1,1}} &= \mathbf{I} + \alpha P \mathbf{h_1} \mathbf{h_1^\dagger} \\ \mathbf{B_{1,1}} &= \mathbf{I} + \alpha P \mathbf{H_3^\dagger} \mathbf{H_3}. \end{aligned} \tag{3.57}$$

Suppose that $(\lambda_{(1,1)\max}, \boldsymbol{\psi}_{1\max})$ is the largest generalized eigenvalue and the corresponding eigenvector pair of the pencil $(\mathbf{A_{1,1}}, \mathbf{B_{1,1}})$. Furthermore, we define the following matrices for the MISOME channel.

$$\begin{aligned}\mathbf{A_{2,2}} &= \mathbf{I} + \frac{(1-\alpha)P}{1+\alpha P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}|^2}\mathbf{h_2}\mathbf{h_2^\dagger} \\ \mathbf{B_{2,2}} &= \mathbf{I} + (1-\alpha)P\mathbf{H_3^\dagger}\left(\mathbf{I}+\alpha P\mathbf{H_3}\boldsymbol{\psi}_{1\max}\boldsymbol{\psi}_{1\max}^\dagger\mathbf{H_3}\right)^{-1}\mathbf{H_3}.\end{aligned} \tag{3.58}$$

Assume that $(\lambda_{(2,2)\max}, \boldsymbol{\psi}_{2\max})$ is the largest generalized eigenvalue and the corresponding eigenvector pair of the pencil $(\mathbf{A_{2,2}}, \mathbf{B_{2,2}})$. Moreover, consider the following matrices for this channel.

$$\begin{aligned}\mathbf{A_{2,1}} &= \mathbf{I} + (1-\alpha)P\mathbf{h_2}\mathbf{h_2^\dagger} \\ \mathbf{B_{2,1}} &= \mathbf{I} + (1-\alpha)P\mathbf{H_3^\dagger}\mathbf{H_3} \\ \mathbf{A_{1,2}} &= \mathbf{I} + \frac{\alpha P}{1+(1-\alpha)P|\mathbf{h_1^\dagger}\boldsymbol{\psi}_{3\max}|^2}\mathbf{h_1}\mathbf{h_1^\dagger} \\ \mathbf{B_{1,2}} &= \mathbf{I} + \alpha P\mathbf{H_3^\dagger}\left(\mathbf{I}+(1-\alpha)P\mathbf{H_3}\boldsymbol{\psi}_{3\max}\boldsymbol{\psi}_{3\max}^\dagger\mathbf{H_3}\right)^{-1}\mathbf{H_3},\end{aligned} \tag{3.59}$$

where we assume that $(\lambda_{(2,1)\max}, \boldsymbol{\psi}_{3\max})$, and $(\lambda_{(1,2)\max}, \boldsymbol{\psi}_{4\max})$ are the largest generalized eigenvalue and the corresponding eigenvector pair of the pencils $(\mathbf{A_{2,1}}, \mathbf{B_{2,1}})$, and $(\mathbf{A_{1,2}}, \mathbf{B_{1,2}})$, respectively. The following theorem then characterizes the capacity region of the MISOME channel under a total power constraint P based on the above parameters.

Theorem 9. *Let $\mathcal{C}^{MISOME}$ denote the secrecy capacity region of the the MISOME channel under an average total power constraint P. Let $\prod$ be the collection of all possible permutations of the ordered set $\{1,2\}$ and conv be the convex closure operator, then $\mathcal{C}^{MISOME}$ is given by*

$$\mathcal{C}^{MISOME} = conv\left\{\bigcup_{\pi\in\prod}\mathcal{R}^{MISOME}(\pi)\right\},$$

where $\mathcal{R}^{MISOME}(\pi)$ is given by

$$\mathcal{R}^{MISOME}(\pi) = \bigcup_{0\le\alpha\le1}\mathcal{R}^{MISOME}(\pi,\alpha),$$

where $\mathcal{R}^{MISOME}(\pi,\alpha)$ is the set of all (R_1,R_2) satisfying the following condition.

$$R_k \le \frac{1}{2}\left[\log\lambda_{(k,\pi^{-1}(k))\max}\right]^+, \qquad k=1,2.$$

Proof: This theorem is a special case of Theorem 8 and corollary 1. First assume that the permutation $\pi = \pi_I = \{1, 2\}$. In the SDPC achievable rate region of (3.52), we choose the covariance matrices $\mathbf{B_1}$ and $\mathbf{B_2}$ as follows.

$$\begin{aligned} \mathbf{B_1} &= \alpha P \boldsymbol{\psi}_{1\max} \boldsymbol{\psi}_{1\max}^{\dagger} \\ \mathbf{B_2} &= (1-\alpha) P \boldsymbol{\psi}_{2\max} \boldsymbol{\psi}_{2\max}^{\dagger}. \end{aligned} \tag{3.60}$$

In other words, the channel input $\mathbf{x}$ is separated into two vectors $\mathbf{b_1}$ and $\mathbf{b_2}$ such that

$$\begin{aligned} \mathbf{x} &= \mathbf{b_1} + \mathbf{b_2} \\ \mathbf{b_1} &= u_1 \boldsymbol{\psi}_{1\max} \\ \mathbf{b_2} &= u_2 \boldsymbol{\psi}_{2\max}, \end{aligned} \tag{3.61}$$

where $u_1 \sim \mathcal{N}(0, \alpha P)$, $u_2 \sim \mathcal{N}(0, (1-\alpha)P)$, and $0 \leq \alpha \leq 1$. Using these parameters, the region of $\mathcal{R}^{SDPC}(\pi_I, \mathbf{S}, \mathbf{N_{1,2,3}})$ becomes as follows:

$$\begin{aligned} R_1 &\leq \frac{1}{2}\left[\log\left|1 + \mathbf{h_1^{\dagger}} \mathbf{B_1} \mathbf{h_1}\right| - \frac{1}{2}\log\left|\mathbf{I} + \mathbf{H_3} \mathbf{B_1} \mathbf{H_3^{\dagger}}\right|\right]^{+} \\ &= \frac{1}{2}\left[\log\left|(1 + \alpha P \mathbf{h_1^{\dagger}} \boldsymbol{\psi}_{1\max} \boldsymbol{\psi}_{1\max}^{\dagger} \mathbf{h_1})\right| - \frac{1}{2}\log\left|(\mathbf{I} + \alpha P \mathbf{H_3} \boldsymbol{\psi}_{1\max} \boldsymbol{\psi}_{1\max}^{\dagger} \mathbf{H_3^{\dagger}})\right|\right]^{+} \\ &= \frac{1}{2}\left[\log \frac{\boldsymbol{\psi}_{1\max}^{\dagger}\left(\mathbf{I} + \alpha P \mathbf{h_1} \mathbf{h_1^{\dagger}}\right)\boldsymbol{\psi}_{1\max}}{\boldsymbol{\psi}_{1\max}^{\dagger}\left(\mathbf{I} + \alpha P \mathbf{H_3^{\dagger}} \mathbf{H_3}\right)\boldsymbol{\psi}_{1\max}}\right]^{+} \\ &\overset{(a)}{=} \frac{1}{2}\left[\log \lambda_{(1,1)\max}\right]^{+}, \end{aligned} \tag{3.62}$$

where (a) is due to the fact that $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$ and the fact that $\boldsymbol{\psi}_{1\max}^{\dagger} \boldsymbol{\psi}_{1\max} = 1$ Similarly, for the R_2 we have,

$$\begin{aligned} R_2 &\leq \frac{1}{2}\left[\log \frac{\left|1 + \mathbf{h_2^{\dagger}}\left(\mathbf{B_1} + \mathbf{B_2}\right)\mathbf{h_2}\right|}{\left|1 + \mathbf{h_2^{\dagger}} \mathbf{B_1} \mathbf{h_2}\right|} - \frac{1}{2}\log \frac{\left|\mathbf{I} + \mathbf{H_3}\left(\mathbf{B_1} + \mathbf{B_2}\right)\mathbf{H_3^{\dagger}}\right|}{\left|\mathbf{I} + \mathbf{H_3} \mathbf{B_1} \mathbf{H_3^{\dagger}}\right|}\right]^{+} \\ &= \frac{1}{2}\left[\log\left|1 + \frac{\mathbf{h_2^{\dagger}} \mathbf{B_2} \mathbf{h_2}}{1 + \mathbf{h_2^{\dagger}} \mathbf{B_1} \mathbf{h_2}}\right| - \frac{1}{2}\log\left|\mathbf{I} + \frac{\mathbf{H_3} \mathbf{B_2} \mathbf{H_3^{\dagger}}}{\mathbf{I} + \mathbf{H_3} \mathbf{B_1} \mathbf{H_3^{\dagger}}}\right|\right]^{+} \\ &= \frac{1}{2}\left[\log \frac{\boldsymbol{\psi}_{2\max}^{\dagger}\left(\mathbf{I} + \frac{(1-\alpha)P \mathbf{h_2} \mathbf{h_2^{\dagger}}}{1 + \alpha P |\mathbf{h_2^{\dagger}} \boldsymbol{\psi}_{1\max}|^2}\right)\boldsymbol{\psi}_{2\max}}{\boldsymbol{\psi}_{2\max}^{\dagger}\left(\mathbf{I} + (1-\alpha) P \mathbf{H_3^{\dagger}}\left(\mathbf{I} + \alpha P \mathbf{H_3} \boldsymbol{\psi}_{1\max} \boldsymbol{\psi}_{1\max}^{\dagger} \mathbf{H_3^{\dagger}}\right)^{-1} \mathbf{H_3}\right)\boldsymbol{\psi}_{2\max}}\right]^{+} \\ &= \frac{1}{2}\left[\log \lambda_{(2,2)\max}\right]^{+}. \end{aligned} \tag{3.63}$$

Similarly, when $\pi = \{2, 1\}$, in the SDPC region, we choose $\mathbf{b_1} = u_1\boldsymbol{\psi}_{4\max}$ and $\mathbf{b_2} = u_2\boldsymbol{\psi}_{3\max}$. Then the SDPC region is given by

$$\begin{aligned} R_1 \leq & \frac{1}{2}\left[\log\frac{\left|1+\mathbf{h_1^\dagger}\left(\mathbf{B_1}+\mathbf{B_2}\right)\mathbf{h_1}\right|}{\left|1+\mathbf{h_1^\dagger}\mathbf{B_2}\mathbf{h_1}\right|} - \frac{1}{2}\log\frac{\left|\mathbf{I}+\mathbf{H_3}\left(\mathbf{B_1}+\mathbf{B_2}\right)\mathbf{H_3^\dagger}\right|}{\left|\mathbf{I}+\mathbf{H_3}\mathbf{B_2}\mathbf{H_3^\dagger}\right|}\right]^+ \\ = & \frac{1}{2}\left[\log\left|1+\frac{\mathbf{h_1^\dagger}\mathbf{B_1}\mathbf{h_1}}{1+\mathbf{h_1^\dagger}\mathbf{B_2}\mathbf{h_1}}\right| - \frac{1}{2}\log\left|\mathbf{I}+\frac{\mathbf{H_3}\mathbf{B_1}\mathbf{H_3^\dagger}}{\mathbf{I}+\mathbf{H_3}\mathbf{B_2}\mathbf{H_3^\dagger}}\right|\right]^+ \\ = & \frac{1}{2}\left[\log\frac{\boldsymbol{\psi}_{4\max}^\dagger\left(\mathbf{I}+\frac{\alpha P\mathbf{h_1}\mathbf{h_1^\dagger}}{1+(1-\alpha)P|\mathbf{h_1^\dagger}\boldsymbol{\psi}_{3\max}|^2}\right)\boldsymbol{\psi}_{4\max}}{\boldsymbol{\psi}_{4\max}^\dagger\left(\mathbf{I}+\alpha P\mathbf{H_3^\dagger}\left(\mathbf{I}+(1-\alpha)P\mathbf{H_3}\boldsymbol{\psi}_{3\max}\boldsymbol{\psi}_{3\max}^\dagger\mathbf{H_3^\dagger}\right)^{-1}\mathbf{H_3}\right)\boldsymbol{\psi}_{4\max}}\right]^+ \\ = & \frac{1}{2}\left[\log\lambda_{(1,2)\max}\right]^+, \end{aligned} \tag{3.64}$$

and R_2 is bounded as follows:

$$\begin{aligned} R_2 \leq & \frac{1}{2}\left[\log\left|1+\mathbf{h_2^\dagger}\mathbf{B_2}\mathbf{h_2}\right| - \frac{1}{2}\log\left|\mathbf{I}+\mathbf{H_3}\mathbf{B_2}\mathbf{H_3^\dagger}\right|\right]^+ \\ = & \frac{1}{2}\left[\log\left|(1+(1-\alpha)P\mathbf{h_2^\dagger}\boldsymbol{\psi}_{3\max}\boldsymbol{\psi}_{3\max}^\dagger\mathbf{h_2})\right| - \frac{1}{2}\log\left|(\mathbf{I}+(1-\alpha)P\mathbf{H_3}\boldsymbol{\psi}_{3\max}\boldsymbol{\psi}_{3\max}^\dagger\mathbf{H_3^\dagger})\right|\right]^+ \\ = & \frac{1}{2}\left[\log\frac{\boldsymbol{\psi}_{3\max}^\dagger\left(\mathbf{I}+(1-\alpha)P\mathbf{h_2}\mathbf{h_2^\dagger}\right)\boldsymbol{\psi}_{3\max}}{\boldsymbol{\psi}_{3\max}^\dagger\left(\mathbf{I}+(1-\alpha)P\mathbf{H_3^\dagger}\mathbf{H_3}\right)\boldsymbol{\psi}_{3\max}}\right]^+ \\ = & \frac{1}{2}\left[\log\lambda_{(2,1)\max}\right]^+. \end{aligned} \tag{3.65}$$

□

Note that the eigenvalues $\lambda_{(l,k)\max} = \lambda_{(l,k)\max}(\alpha, P)$ and the eigenvector $\boldsymbol{\psi}_{k\max} = \boldsymbol{\psi}_{k\max}(\alpha, P)$ for $l, k = 1, 2$ are the functions of α and P. The following corollary characterizes the secrecy capacity region of the MISOME channel in high SNR regime.

Corollary 2. *In the high SNR regime, the secrecy capacity region of the MISOME channel is given as follows:*

$$\lim_{P\to\infty}\mathcal{C}^{MISOME} = conv\left\{\bigcup_{\pi\in\prod}\mathcal{R}_\infty^{MISOME}(\pi)\right\}, \tag{3.66}$$

where

$$\mathcal{R}_{\infty}^{MISOME}(\pi=\{1,2\}) = \tag{3.67}$$
$$\left\{(R_1,R_2), R_1 \leq \frac{1}{2}\left[\log \lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right)\right]^{+}, R_2 \leq \frac{1}{2}\left[\log \frac{\lambda_{\max}\left(\mathbf{h_2}\mathbf{h_2}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right)}{b}\right]^{+}\right\}$$
$$\mathcal{R}_{\infty}^{MISOME}(\pi=\{2,1\}) =$$
$$\left\{(R_1,R_2), R_1 \leq \frac{1}{2}\left[\log \frac{\lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right)}{a}\right]^{+}, R_2 \leq \frac{1}{2}\left[\log \lambda_{\max}\left(\mathbf{h_2}\mathbf{h_2}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right)\right]^{+}\right\},$$

where $(\lambda_{\max}(\mathbf{A_i},\mathbf{B}), \boldsymbol{\psi}_{i\,\max})$ *denotes the largest eigenvalue and corresponding eigenvector of the pencil* $(\mathbf{A_i},\mathbf{B})$ *and* $b = \frac{|\mathbf{h_2}^{\dagger}\boldsymbol{\psi}_{1\,\max}|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{1\,\max}\|^2}$, $a = \frac{|\mathbf{h_1}^{\dagger}\boldsymbol{\psi}_{2\,\max}|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{2\,\max}\|^2}$.

Note that the above secrecy rate region is independent of α and therefore is a convex closure of two rectangular regions.

Proof: We restrict our attention to the case that $\lambda_{(l,k)\,\max}(\alpha,P) > 1$ for $l,k=1,2$ where the rates R_1 and R_2 are nonzero. First suppose that $\pi = \pi_I = \{1,2\}$. We show that

$$\lim_{P\to\infty} \lambda_{(1,1)\,\max}(\alpha,P) = \lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right) \tag{3.68}$$
$$\lim_{P\to\infty} \lambda_{(2,2)\,\max}(\alpha,P) = \frac{\lambda_{\max}\left(\mathbf{h_2}\mathbf{h_2}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right)}{b}. \tag{3.69}$$

Note that since

$$\lambda_{(1,1)\,\max}(\alpha,P) = \frac{1+\alpha P|\mathbf{h_1}^{\dagger}\boldsymbol{\psi}_{1\,\max}(\alpha,P)|^2}{1+\alpha P\|\mathbf{H_3}\boldsymbol{\psi}_{1\,\max}(\alpha,P)\|^2} > 1, \tag{3.70}$$

where

$$\boldsymbol{\psi}_{1\,\max}(\alpha,P) = \arg \max_{\{\boldsymbol{\psi}_1:\|\boldsymbol{\psi}_1\|^2=1\}} \frac{1+\alpha P|\mathbf{h_1}^{\dagger}\boldsymbol{\psi}_1(\alpha,P)|^2}{1+\alpha P\|\mathbf{H_3}\boldsymbol{\psi}_1(\alpha,P)\|^2}, \tag{3.71}$$

for all $P>0$ we have,

$$|\mathbf{h_1}^{\dagger}\boldsymbol{\psi}_{1\,\max}(\alpha,P)|^2 > |\mathbf{H_3}\boldsymbol{\psi}_{1\,\max}(\alpha,P)\|^2. \tag{3.72}$$

Therefore, $\lambda_{(1,1)\,\max}$ is an increasing function of P. Thus,

$$\begin{aligned}\lambda_{(1,1)\,\max}(\alpha,P) &\leq \frac{|\mathbf{h_1}^{\dagger}\boldsymbol{\psi}_{1\,\max}(\alpha,P)|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{1\,\max}(\alpha,P)\|^2} \\ &\leq \lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1}^{\dagger}, \mathbf{H_3}^{\dagger}\mathbf{H_3}\right).\end{aligned} \tag{3.73}$$

Since $\lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1^\dagger}, \mathbf{H_3^\dagger}\mathbf{H_3}\right)$ is independent of P, we have

$$\lim_{P\to\infty} \lambda_{(1,1)\max} \leq \lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1^\dagger}, \mathbf{H_3^\dagger}\mathbf{H_3}\right). \tag{3.74}$$

Next, defining

$$\boldsymbol{\psi}_1(\infty) = \arg \max_{\{\boldsymbol{\psi}_1 : \|\boldsymbol{\psi}_1\|^2=1\}} \frac{|\mathbf{h_1^\dagger}\boldsymbol{\psi}_1|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_1\|^2}, \tag{3.75}$$

we have the following lower bound

$$\begin{aligned}\lim_{P\to\infty} \lambda_{(1,1)\max}(\alpha, P) &\geq \lim_{P\to\infty} \frac{\frac{1}{P} + \alpha|\mathbf{h_1^\dagger}\boldsymbol{\psi}_{1\max}(\infty)|^2}{\frac{1}{P} + \alpha\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\infty)\|^2} \\ &= \lambda_{\max}\left(\mathbf{h_1}\mathbf{h_1^\dagger}, \mathbf{H_3^\dagger}\mathbf{H_3}\right).\end{aligned} \tag{3.76}$$

As the lower bound and upper bound coincide, we obtain (3.68). Similarly, to obtain (3.69), we note that since

$$\lambda_{(2,2)\max}(\alpha, P) = \frac{1 + \frac{(1-\alpha)P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}(\alpha,P)|^2}{1+\alpha P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\alpha,P)|^2}}{1 + \frac{(1-\alpha)P\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}(\alpha,P)\|^2}{1+\alpha P\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\alpha,P)\|^2}} > 1, \tag{3.77}$$

where

$$\boldsymbol{\psi}_{2\max}(\alpha, P) = \arg \max_{\{\boldsymbol{\psi}_2 : \|\boldsymbol{\psi}_2\|^2=1\}} \frac{1 + \frac{(1-\alpha)P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}(\alpha,P)|^2}{1+\alpha P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\alpha,P)|^2}}{1 + \frac{(1-\alpha)P\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}(\alpha,P)\|^2}{1+\alpha P\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\alpha,P)\|^2}}, \tag{3.78}$$

for all $P > 0$, we have,

$$\frac{(1-\alpha)P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}(\alpha, P)|^2}{1 + \alpha P|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\alpha, P)|^2} > \frac{(1-\alpha)P\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}(\alpha, P)\|^2}{1 + \alpha P\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\alpha, P)\|^2}. \tag{3.79}$$

Therefore, we have

$$\begin{aligned}\lim_{P\to\infty} \lambda_{(2,2)\max}(\alpha, P) &\leq \frac{\frac{|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}(\infty)|^2}{|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\infty)|^2}}{\frac{\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}(\infty)\|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\infty)\|^2}} \\ &\leq \frac{\lambda_{\max}\left(\mathbf{h_2}\mathbf{h_2^\dagger}, \mathbf{H_3^\dagger}\mathbf{H_3}\right)}{b},\end{aligned} \tag{3.80}$$

where

$$\begin{aligned}b &= \frac{|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\infty)|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\infty)\|^2} \\ \boldsymbol{\psi}_2(\infty) &= \arg \max_{\{\boldsymbol{\psi}_2 : \|\boldsymbol{\psi}_2\|^2=1\}} \frac{|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}|^2}{\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}\|^2}.\end{aligned} \tag{3.81}$$

On the other hand, we have the following lower bound

$$\lim_{P\to\infty} \lambda_{(2,2)\max}(\alpha, P) \geq \frac{1+\frac{(1-\alpha)|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{2\max}(\infty)|^2}{\alpha|\mathbf{h_2^\dagger}\boldsymbol{\psi}_{1\max}(\infty)|^2}}{1+\frac{(1-\alpha)\|\mathbf{H_3}\boldsymbol{\psi}_{2\max}(\infty)\|^2}{\alpha\|\mathbf{H_3}\boldsymbol{\psi}_{1\max}(\infty)\|^2}}. \tag{3.82}$$

Note that $0 \leq \alpha \leq 1$. It is easy to show that the right side of the equation of (3.82) is a decreasing function of α and therefore the maximum value of this function is when $\alpha = 0$. Thus we have,

$$\lim_{P\to\infty} \lambda_{(2,2)\max}(\alpha, P) \geq \frac{\lambda_{\max}\left(\mathbf{h_2}\mathbf{h_2^\dagger}, \mathbf{H_3^\dagger}\mathbf{H_3}\right)}{b}. \tag{3.83}$$

As the lower bound and upper bound coincide, we obtain (3.69). When $\pi = 2, 1$, the proof is similar and may be omitted here. □

Now consider the MISOME channel with m single antenna receivers and an external eavesdropper. Let $\mathbf{x} = \sum_{k=1}^{m} \mathbf{b_k}$, where $\mathbf{b_k} = u_k\psi_{k\max}$, $u_k \sim \mathcal{N}(0, \alpha_k P)$, and $\sum_{k=1}^{m}\alpha_k = 1$. Assume that $(\lambda_{k\max}, \boldsymbol{\psi}_{k\max})$ is the largest generalized eigenvalue and the corresponding eigenvector pair of the pencil

$$\left(\mathbf{I} + \frac{\alpha_k P\mathbf{h_k}\mathbf{h_k^\dagger}}{1+\mathbf{h_k^\dagger}\mathbf{A}\mathbf{h_k}}, \mathbf{I} + \alpha_k P\mathbf{H_3^\dagger}\left(\mathbf{I}+\mathbf{H_3}\mathbf{A}\mathbf{H_3^\dagger}\right)^{-1}\mathbf{H_3}\right), \tag{3.84}$$

where $\mathbf{A} = (\sum_{i=1}^{\pi^{-1}(k)-1}\alpha_{\pi(i)}P\psi_{\pi(i)\max}\psi_{\pi(i)\max}^\dagger)$. The following corollary then characterizes the capacity region of the MISOME channel with m receivers under a total power constraint P.

Corollary 3. *Let $\prod$ be the collection of all possible permutations of the ordered set $\{1, ..., m\}$ and conv be the convex closure operator, then $\mathcal{C}^{MISOME}$ is given as follows:*

$$\mathcal{C}^{MISOME} = conv\left\{\bigcup_{\pi\in\prod}\mathcal{R}^{MISOME}(\pi)\right\},$$

where $\mathcal{R}^{MISOME}(\pi)$ is given by

$$\mathcal{R}^{MISOME}(\pi) = \bigcup_{0\leq\alpha_k\leq 1, \sum_{k=1}^{m}\alpha_k=1} \mathcal{R}^{MISOME}(\pi, \alpha_1, ..., \alpha_m),$$

where $\mathcal{R}^{MISOME}(\pi, \alpha_1, ..., \alpha_m)$ is the set of all $(R_1, ..., R_m)$ satisfying the following condition.

$$R_k \leq \frac{1}{2}\left[\log\lambda_{k\max}\right]^+, \qquad k = 1, ..., m.$$

Chapter 4

Secure Gaussian Multiple-Access-Channel

In this chapter, we consider a K-user secure Gaussian Multiple-Access-Channel (MAC) with an external eavesdropper. We establish an achievable rate region for the secure discrete memoryless MAC. We prove the secrecy sum capacity of the degraded Gaussian MIMO MAC using Gaussian codebooks. For the non-degraded Gaussian MIMO MAC, we propose an algorithm inspired by the interference alignment technique to achieve the largest possible total Secure-Degrees-of-Freedom (S-DoF). When all the terminals are equipped with a single antenna, we show that Gaussian codebooks are inefficient in providing a positive S-DoF. Instead, we propose a novel secure coding scheme to achieve a positive S-DoF in the single antenna MAC. This scheme converts the single-antenna system into a multiple-dimension system with fractional dimensions. The achievability scheme is based on the alignment of signals into a small sub-space at the eavesdropper, and the simultaneous separation of the signals at the intended receiver. We use tools from the field of Diophantine Approximation in the number theory to analyze the probability of error in the coding scheme. We prove that the total S-DoF of $\frac{K-1}{K}$ can be achieved for almost all channel gains. For the other channel gains, we propose a multi-layer coding scheme to achieve a positive S-DoF. As a function of channel gains, therefore, the achievable S-DoF is discontinued.

4.1 Introduction

The secure MAC generalizes the wiretap channel. In the wiretap channel, the direct coding scheme uses the framework of random coding, which is widely used in the analysis of multi-terminal source and channel coding problems. One approach to finding achievable sum rates

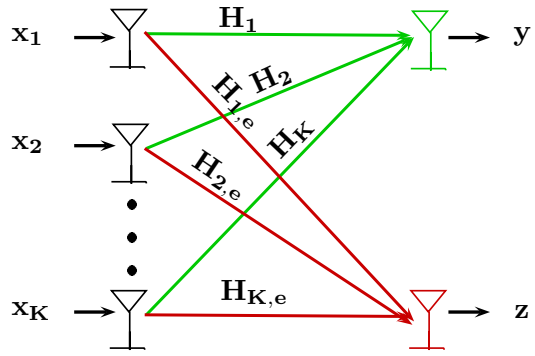

Figure 4.1: Secure K-user Gaussian MIMO Multiple-Access-Channel

for the secure MAC is to extend the random coding solution to the secure multi-user. As we will show in this chapter, this extension leads to a single-letter characterization for the secure rate region of the MAC. Our achievability, as usual, is based on the i.i.d random binning scheme.

On the other hand, it is shown that the random coding argument may be insufficient to prove capacity theorems for certain channels; instead, structure codes can be used to construct efficient channel codes for Gaussian channels. In reference [78], nested lattice codes are used to provide secrecy in two-user Gaussian channels. In [78] it is shown that structure codes can achieve a positive S-DoF in a two-user MAC. In particular, the achievability scheme of [78] provides an S-DoF of $\frac{1}{2}$ for a small category of channel gains and for the other categories, it provides a S-DoF of strictly less than $\frac{1}{2}$.

In reference [79], the concept of interference alignment is introduced and has illustrated its capability in achieving the full DoF of a class of two-user X channels. In reference [80], and [81], a novel coding scheme applicable in networks with single antenna nodes is proposed. This scheme converts a single antenna system into an equivalent Multiple Input Multiple Output (MIMO) system with fractional dimensions.

4.2 Preliminaries

Consider a secure K-user Gaussian MIMO Multiple-Access-Channel (MAC) as depicted in Figure 4.1. In this confidential setting, each user k ($k \in \mathcal{K} \stackrel{\triangle}{=} \{1, 2, ..., K\}$) wishes to send a message W_k to the intended receiver in n uses of the channel simultaneously, and prevent the eavesdropper from having any information about the messages. At a specific time, the

signals received by the intended receiver and the eavesdropper is given by

$$\begin{aligned} \mathbf{y} &= \sum_{k=1}^{K} \mathbf{H_k x_k} + \mathbf{n_1} \\ \mathbf{z} &= \sum_{k=1}^{K} \mathbf{H_{k,e} x_k} + \mathbf{n_2} \end{aligned} \tag{4.1}$$

where

- $\mathbf{x_k}$ for $k = 1, 2, ..., K$ is a real input vector of size $M_k \times 1$ under an input average power constraint. We require that $Tr(\mathbf{Q_k}) \leq P$, where $\mathbf{Q_k} = E[\mathbf{x_k x_k^\dagger}]$. Here, the superscript $\dagger$ denotes the Hermitian transpose of a vector and $Tr(.)$ denotes the Trace operator on the matrices.

- $\mathbf{y}$ and $\mathbf{z}$ are real output vectors which are received by the destination and the eavesdropper, respectively. These are vectors of size $N \times 1$ and $N_e \times 1$, respectively.

- $\mathbf{H_k}$ and $\mathbf{H_{k,e}}$ for $k = 1, 2, ..., K$ are fixed, real gain matrices which model the channel gains between the transmitters and the intended receiver, and the eavesdropper, respectively. These are matrices of size $N \times M_k$ and $N_e \times M_k$, respectively. The channel state information is assumed to be known perfectly at all the transmitters and at all receivers.

- $\mathbf{n_1}$ and $\mathbf{n_2}$ are real Gaussian random vectors with zero means and covariance matrices $\mathbf{N_1} = E[\mathbf{n_1 n_1}^T] = \mathbf{I_N}$ and $\mathbf{N_2} = E[\mathbf{n_2 n_2}^T] = \mathbf{I_{N_e}}$, respectively. Here, $\mathbf{I_M}$ represents the identity matrix of size $M \times M$.

Let $\mathbf{x_k^n}$, $\mathbf{y^n}$ and $\mathbf{z^n}$ denote the random channel inputs and random channel outputs matrices over a block of n samples. Furthermore, let $\mathbf{n_1^n}$, and $\mathbf{n_2^n}$ denote the additive noises of the channels. Therefore, we have

$$\begin{aligned} \mathbf{y^n} &= \sum_{k=1}^{K} \mathbf{H_k x_k^n} + \mathbf{n_1^n} \\ \mathbf{z^n} &= \sum_{k=1}^{K} \mathbf{H_{k,e} x_k^n} + \mathbf{n_2^n}. \end{aligned} \tag{4.2}$$

Note that bold vectors are random while the matrices $\mathbf{H_k}$ and $\mathbf{H_{k,e}}$ are deterministic matrices for all $k \in \mathcal{K}$. The columns of $\mathbf{n_1^n}$ and $\mathbf{n_2^n}$ are independent Gaussian random vectors with covariance matrices $\mathbf{I_N}$ and $\mathbf{I_{N_e}}$, respectively. In addition $\mathbf{n_1^n}$ and $\mathbf{n_2^n}$ are independent of

$\mathbf{x_k^n}$'s and W_k's. A $((2^{nR_1}, 2^{nR_2}, ..., 2^{nR_k}), n)$ secret code for the above channel consists of the following components:

1) K secret message sets $\mathcal{W}_k = \{1, 2, ..., 2^{nR_k}\}$.

2) K stochastic encoding functions $f_k(.)$ which map the secret messages to the transmitted symbols, i.e., $f_k : w_k \rightarrow \mathbf{x_k^n}$ for each $w_k \in \mathcal{W}_k$. At encoder k, each codeword is designed according to the transmitter's average power constraint P.

3) A decoding function $\phi(.)$ which maps the received symbols to estimate the messages: $\phi(\mathbf{y^n}) \rightarrow (\hat{W}_1, ..., \hat{W}_K)$.

The reliability of the transmission is measured by the average probability of error, which is defined as the probability that the decoded messages are not equal to the transmitted messages; that is,

$$P_e^{(n)} = \frac{1}{\prod_{k=1}^{K} 2^{nR_k}} \sum_{(w_1,...,w_K)\in\mathcal{W}_1\times....\times\mathcal{W}_\mathcal{K}} P(\phi(\mathbf{y^n}) \neq (w_1, ..., w_K)|(w_1, ..., w_k) \text{ is sent}). \quad (4.3)$$

The secrecy level is measured by the normalized equivocation defined as follows: The normalized equivocation for each subset of messages $W_\mathcal{S}$ for $\mathcal{S} \subseteq \mathcal{K}$ is

$$\Delta_\mathcal{S} \triangleq \frac{H(W_\mathcal{S}|\mathbf{z^n})}{H(W_\mathcal{S})}. \quad (4.4)$$

The rate-equivocation tuple $(R_1, ..., R_K, d)$ is said to be achievable for the Gaussian MIMO Multiple-Access-Channel with confidential messages, if for any $\epsilon > 0$, there exists a sequence of $((2^{nR_1}, ..., 2^{nR_K}), n)$ secret codes, such that for sufficiently large n,

$$P_e^{(n)} \leq \epsilon, \quad (4.5)$$

and

$$\Delta_\mathcal{S} \geq d - \epsilon, \qquad \forall \mathcal{S} \subseteq \mathcal{K}. \quad (4.6)$$

The perfect secrecy rate tuple $(R_1, ..., R_K)$ is said to be achievable when $d = 1$. When all the transmitted messages are perfectly secure, we have

$$\Delta_\mathcal{K} \geq 1 - \epsilon, \quad (4.7)$$

or equivalently

$$H(W_\mathcal{K}|\mathbf{z^n}) \geq H(W_\mathcal{K}) - \epsilon H(W_\mathcal{K}). \quad (4.8)$$

The normalized equivocation of each subset of messages can then be written as follows:

$$\begin{aligned} H(W_{\mathcal{S}}|\mathbf{z^n}) &\overset{(a)}{=} H(W_{\mathcal{S}}, W_{\mathcal{S}^c}|\mathbf{z^n}) - H(W_{\mathcal{S}^c}|W_{\mathcal{S}}, \mathbf{z^n}) \\ &= H(W_{\mathcal{K}}|\mathbf{z^n}) - H(W_{\mathcal{S}^c}|W_{\mathcal{S}}, \mathbf{z^n}) \\ &\overset{(b)}{\geq} H(W_{\mathcal{K}}) - \epsilon H(W_{\mathcal{K}}) - H(W_{\mathcal{S}^c}|W_{\mathcal{S}}, \mathbf{z^n}) \\ &\overset{(c)}{=} H(W_{\mathcal{S}}) + H(W_{\mathcal{S}^c}|W_{\mathcal{S}}) - \epsilon H(W_{\mathcal{K}}) - H(W_{\mathcal{S}^c}|W_{\mathcal{S}}, \mathbf{z^n}) \\ &\overset{(d)}{\geq} H(W_{\mathcal{S}}) - \epsilon H(W_{\mathcal{K}}), \end{aligned} \tag{4.9}$$

where (a) and (c) follow from the chain rule, (b) follows from (4.7) and (d) follows from the fact that conditioning always decreases the amount of entropy. Therefore, the normalized equivocation of each subset of messages is

$$\Delta_{\mathcal{S}} \geq 1 - \epsilon^{'}, \tag{4.10}$$

where $\epsilon^{'} = \frac{H(W_{\mathcal{K}})}{H(W_{\mathcal{S}})}\epsilon$. Thus, when all of the K messages are perfectly secure then it is guaranteed that any subset of the messages becomes perfectly secure.

The total Secure Degrees-of Freedom (S-DoF) of η is said to be achievable, if the rate-equivocation tuple $(R_1, ..., R_K, d = 1)$ is achievable, and

$$\eta = \lim_{P\to\infty} \frac{\sum_{k=1}^{K} R_k}{\frac{1}{2}\log P} \tag{4.11}$$

4.3 Secure DoF of the Multiple-Antenna Multiple-Access-Channel

In this section, we first present an achievability rate region for the secure discrete memoryless MAC. We then characterize the sum capacity of the degraded secure discrete memoryless and degraded Gaussian MIMO MAC. We present an achievable S-DoF of the non-degraded Gaussian MIMO MAC under the perfect secrecy constraint using Gaussian codebooks. In order to satisfy the perfect secrecy constraint, we use the random binning coding scheme to generate the codebooks. To maximize the achievable degrees of freedom, we adopt the signal alignment scheme to separate the signals at the intended receiver and simultaneously align the signals into a small subspace at the eavesdropper.

4.3.1 Discrete Memoryless MAC

In this subsection, we study the secure discrete MAC of $P(y, z|x_1, ..., x_K)$ with K users and an external eavesdropper. The following theorems illustrate our results:

Theorem 10. *For the perfectly secure discrete memoryless MAC of $P(y,z|x_1,...,x_K)$, the region of*

$$\left\{(R_1,...,R_K)|\sum_{i\in\mathcal{S}}R_i\leq I(U_{\mathcal{S}};Y|U_{\mathcal{S}^c}),\sum_{k\in\mathcal{K}}R_k\leq\left[I(U_{\mathcal{K}};Y)-I(U_{\mathcal{K}};Z)\right]^+,\forall\mathcal{S}\subset\mathcal{K}\right\},\quad(4.12)$$

for any distribution of $P(u_1)P(u_2)....P(u_K)P(x_1|u_1)P(x_2|u_2)...P(x_K|u_K)P(y,z|x_1,...,x_K)$, is achievable.

Please see section 4.5.1 for the proof.

In this Theorem $[x]^+$ denotes the positivity operator, i.e., $[x]^+=\max(x,0)$. Reference [46] derived an achievable rate region with Gaussian codebooks and power control for the Gaussian secure MAC when all the transmitters and receivers are equipped with a single antenna. Theorem 10, however, gives an achievability secrecy rate region for the general discrete memoryless MAC. Our achievability rate region is also larger than the region of [46] in the special Gaussian channel case. Therefore, we have the following achievable sum rate for the secure discrete memoryless MAC:

Corollary 4. *For the secure discrete memoryless MAC of $P(y,z|x_1,...,x_K)$, the following sum rate is achievable:*

$$R_{sum}=\max\left[I(U_{\mathcal{K}};Y)-I(U_{\mathcal{K}};Z)\right]^+,\quad(4.13)$$

where the maximization is over all distributions of

$$P(u_1)...P(u_K)P(x_1|u_1)...P(x_K|u_K)P(y,z|x_1,...,x_K)$$

that satisfy the markov chain $W_{\mathcal{K}}\rightarrow U_{\mathcal{K}}\rightarrow X_{\mathcal{K}}\rightarrow YZ$.

4.3.2 Gaussian MIMO MAC

Consider the secure Gaussian MIMO MAC of (4.1) which can be re-written as follows:

$$\begin{aligned}\mathbf{y}&=\mathbf{Hx}+\mathbf{n_1}\\ \mathbf{z}&=\mathbf{H_e x}+\mathbf{n_2}\end{aligned}\quad(4.14)$$

where, $\mathbf{H}=[\mathbf{H_1},\mathbf{H_2},...,\mathbf{H_K}]$, $\mathbf{H_e}=[\mathbf{H_{1,e}},\mathbf{H_{2,e}},...,\mathbf{H_{K,e}}]$, and $\mathbf{x}=\left[\mathbf{x_1^\dagger},\mathbf{x_2^\dagger},...,\mathbf{x_K^\dagger}\right]^\dagger$. Without loss of generality, assume that all nodes are equipped with the same number of antennas, i.e., $M_k=N=N_e$ for all $k\in\mathcal{K}$. Note that when the channel gain matrices $\mathbf{H_k}$ and $\mathbf{H_{k,e}}$ are identity matrices we can determine that one channel output is degraded w.r.t. another

by examining whether their noise covariances can be ordered correctly. In (4.1), however, all noise covariances are identity matrices and the receive vectors differ only in their channel gain matrices. Therefore, similar to [82], we use the following definition to determine a degradedness order:

Definition 12. *A receive vector* $\mathbf{z} = \mathbf{H_e x} + \mathbf{n_2}$ *is said to be degraded w.r.t.* $\mathbf{y} = \mathbf{Hx} + \mathbf{n_1}$ *if there exists a matrix* $\mathbf{D}$ *such that* $\mathbf{DH} = \mathbf{H_e}$ *and such that* $\mathbf{DD}^\dagger \preceq \mathbf{I}$. *Alternatively, we say that* $\mathbf{H_e}$ *is degraded w.r.t.* $\mathbf{H}$.

According to this definition, it is easy to see that $\mathbf{y}$ can be approximated by multiplying $\mathbf{Dy}$. The approximated channel has a different additive noise which is now given by $\mathbf{Dn_1} \sim \mathcal{N}(0, \mathbf{DD}^\dagger)$ compared to the original channel. As this approximated channel has less noise ($\mathbf{DD}^\dagger \preceq \mathbf{I}$), however, it is clear that any message that can be decoded by the eavesdropper, can also be decoded by the intended receiver. In the other words $W_\mathcal{K} \rightarrow \mathbf{x} \rightarrow \mathbf{y} \rightarrow \mathbf{z}$ forms a Markov chain.

Theorem 11. *The secrecy sum capacity of the degraded Gaussian MIMO MAC is given by*

$$C_{sum} = \max_{\mathbf{Q_k}:\mathbf{Q_k}\succeq 0, Tr(\mathbf{Q_k})\leq P} \frac{1}{2}\log\left|I + \sum_{k\in\mathcal{K}} \mathbf{H_k Q_k H_k^\dagger}\right| - \frac{1}{2}\log\left|I + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e} Q_k H_{k,e}^\dagger}\right| . \quad (4.15)$$

Proof: We need to show that the secrecy sum capacity is as follows:

$$C_{\text{sum}} = \frac{1}{2}\log\left|I + \sum_{k\in\mathcal{K}} \mathbf{H_k Q_k H_k^\dagger}\right| - \frac{1}{2}\log\left|I + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e} Q_k H_{k,e}^\dagger}\right|, \quad (4.16)$$

if the inputs are subject to the following covariance matrices constraints:

$$\mathbf{K_{x_k}} \leq \mathbf{Q_k}, \qquad \forall k \in \mathcal{K}, \quad (4.17)$$

where $\mathbf{K_{x_k}}$ denotes the covariance matrix of $\mathbf{x_k}$. Theorem 11 then follows by maximization over all $\mathbf{Q_1}$, $\mathbf{Q_2}$,..., and $\mathbf{Q_K}$ that satisfy the power constraint, i.e., $Tr(\mathbf{Q_k}) \leq P$, for all $k \in \mathcal{K}$.

The achievability of this theorem follows from Theorem 4 by choosing $U_k = \mathbf{x_k} \sim \mathbf{N}(0, \mathbf{Q_k})$. The converse proof is presented in section 4.5.2. □

According to (4.55), it is easy to show that:

Corollary 5. *The total S-DoF for the degraded Gaussian MIMO MAC is* $\eta = 0$.

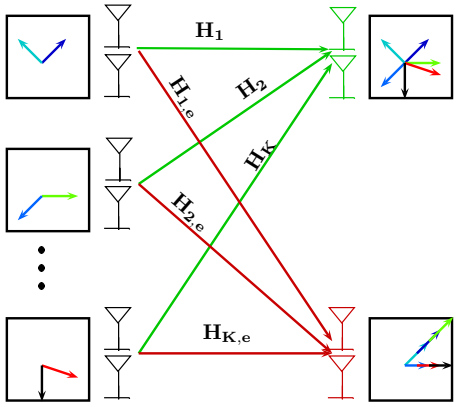

Figure 4.2: Separation/Alignment of Signals at the Intended Receiver/Eavesdropper

Now consider the general Gaussian MIMO MAC where it's not necessarily degraded. According to Theorem 10, by choosing $U_k = \mathbf{x_k} \sim \mathcal{N}(0, \mathbf{Q_k})$, it is easy to see that the following secrecy sum rate is achievable.

Corollary 6. *For the Gaussian MIMO MAC, an achievable secrecy sum rate is given by*

$$R_{sum} = \sum_{k\in\mathcal{K}} R_k = \max_{\mathbf{Q}:\mathbf{Q}\succeq 0, Tr(\mathbf{Q})\leq P} \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k Q_k H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e} Q_k H_{k,e}^\dagger}\right|. \tag{4.18}$$

Note that in the above achievable scheme, we chose $\mathbf{K_{x_k}} = \mathbf{Q_k}$ which is generally not optimal but indeed is achievable. In general, solving the maximization problem of (4.18) is difficult. Reference [46], however, has solved this problem for a single antenna case and derived the optimal power control policy. As shown in [46] even for a single antenna case some users need to be silent and therefore those users can cooperatively help to jam the eavesdropper.

We study the S-DoF defined in (4.11) to analyze the behavior of the sum rate in high SNR. We design the following strategy scheme at the transmitters.

To achieve the largest value for S-DoF we need to separate the received signals at the legitimate receiver, such that each received signal has a different dimension in the signal space of the legitimate receiver. At the same time all the received signals at the eavesdropper need to be aligned in a minimal subspace of the signal space of the eavesdropper (see Figure 4.2).

Let $\mathbf{x_k} = \mathbf{F_k v_k}$ where, $\mathbf{F_k}$ is a pre-coding matrix such that $\mathbf{F_k F_k^\dagger} = \mathbf{Q_k}$ and $\mathbf{v_k}$ is a vector with i.i.d Gaussian components with zero mean and unit variance, i.e, $\mathbf{v_k} = [v_k^1, ..., v_k^N]^\dagger$, such that $v_k^j \sim \mathcal{N}(0,1)$ for $j = 1, 2, ..., N$. Let $\boldsymbol{\psi}_i = [0, 0, ..., 0, 1, 0, ..., 0]^\dagger$ be a $N \times 1$ vector that all of its elements are zero except the ith element which is 1. Let $\mathbf{F_k} = [\mathbf{f_k^1}, \mathbf{f_k^2}, ..., \mathbf{f_k^N}]$,

where $\mathbf{f_k^i}$'s for $i = 1, 2, ..., N$ are $N \times 1$ vectors that represent the columns of $\mathbf{F_k}$. We use the following algorithm to choose $\mathbf{f_k^i}$'s:

- Assume that for users $k = 1, 2, ..., J$, the matrix $\mathbf{H_{k,e}}$ has non-empty null space and the null space of $\mathbf{H_{k,e}}$ for users $k = J+1, J+2, ..., K$ is empty. The users $k = 1, 2, ..., \min\{J, N\}$ choose $\mathbf{f_k^1}$ such that $\mathbf{H_{k,e}f_k^1} = 0$. Almost surely, we can assume that these vectors occupy separate dimensions at the legitimate receiver.

- If $J \geq N$, then all the N dimensions at the legitimate receiver are full and $\eta = N$.

- If $J < N$, users $k = J+1, J+2, ..., \min\{K, N\}$ then create a vector $\mathbf{f_k^1}$ such that $\mathbf{H_{k,e}f_k^1} = \boldsymbol{\psi}_1$. Theses vectors, therefore, are aligned at the eavesdropper in one dimension and almost surely occupy separate dimensions in the remaining subspace of the legitimate receiver.

- If $N \leq K$, at this step all the dimensions at the legitimate receiver are full and one dimension at the eavesdropper has a non-zero signal. Thus, $\eta = N - 1$.

- If $N > K$, users $k = 1, 2, ..., J$ then also creates a vector $\mathbf{f_k^2}$ such that $\mathbf{H_{k,e}f_k^2} = \boldsymbol{\psi}_1$. These vectors are also aligned at the dimension $\boldsymbol{\psi}_1$ and almost surely $\eta = N - 1$

- We repeat the above steps such that all the dimensions at the legitimate receiver become full.

The above algorithm can be followed when the users and the receivers are equipped with different numbers of antennae. The following theorem characterizes the maximum amount of the total S-DoF that can be achieved by Gaussian codebooks.

Theorem 12. *For the Gaussian MIMO MAC, the following η for S-DoF can almost surely be achieved for almost all channel gains by using Gaussian codebooks and under perfect secrecy constraint:*

$$\eta = \left[\min \left\{ \sum_{k \in \mathcal{K}} M_k, N \right\} - r \right]^+, \tag{4.19}$$

where $0 \leq r \leq \min\{\sum_{k \in \mathcal{K}} M_k, N_e\}$ depends on the channel gain matrix $\mathbf{H_e}$.

Note that in Theorem 12, it is emphasized that total $\left[\min\left\{\sum_{k \in \mathcal{K}} M_k, N\right\} - r\right]^+$ S-DoF is achievable for almost all channel gains. It means the set of all possible gains that the total amount of $\left[\min\left\{\sum_{k \in \mathcal{K}} M_k, N\right\} - r\right]^+$ S-DoF may not be achieved has the Lebesgue measure zero. In other words, if all the channel gains are drawn independently from a random distribution, then almost all of them satisfy properties required to achieve the total S-DoF, almost surely. The term "almost surely" means with a probability arbitrary close to 1.

Remark 8. *In the achievability scheme of Theorem 12, all the transmitted signals are aligned into an r-dimensional subspace at the eavesdropper, and hence, impair the ability of the eavesdropper to distinguish any of the secure messages efficiently.*

Now assume that the transmitters can cooperate with each other: we have a MIMO wiretap channel where the transmitter has $\sum_{k\in\mathcal{K}} M_k$ antennas and the legitimate and eavesdropper have N and N_e antennas, respectively. The secrecy capacity of this channel is indeed an upper-bound for the secrecy sum capacity of the Gaussian MIMO MAC. As the capacity of a non-secret MIMO channel is an upper-bound for the secrecy capacity of the MIMO wiretap channel, we have the following upper-bound for S-DoF for the secure Gaussian MIMO MAC.

Lemma 6. *For the Gaussian MIMO MAC, the maximum total achievable S-DoF under perfect secrecy constraint is given by*

$$\eta_{\max} = \min\left\{\sum_{k\in\mathcal{K}} M_k, N\right\}. \tag{4.20}$$

The maximum penalty for the achievable S-DoF in Theorem 12 is therefore r, where, $0 \leq r \leq \min\{\sum_{k\in\mathcal{K}} M_k, N_e\}$. We note that an achievable S-DoF in the MIMO wiretap channel using zero-forcing beamforming is given as follows.

Lemma 7. *In the MIMO wiretap channel, the following S-DoF is achievable, almost surely.*

$$\eta = \min\left\{\left[\sum_{k\in\mathcal{K}} M_k - r^{'}\right]^{+}, N\right\}, \tag{4.21}$$

where $1 \leq r^{'} \leq N_e$ depends on the channel gain matrix $\mathbf{H_e}$.

Remark 9. *When the transmitters and the intended receiver are equipped with a sufficiently large number of antennae while the eavesdropper is equipped with a limited number of antennae, then Gaussian codebooks provide a near optimum total of S-DoF for the Gaussian MIMO multiple-access-channel under perfect secrecy constraint.*

Note that when the transmitters and the receivers are equipped with a single antenna, i.e., $M_k = N_e = N = 1$, then the total achieved S-DoF is 0. It should be noted that this result comes from the lack of sufficient dimensions for signal management at the receivers by using Gaussian codebooks. In our achievability scheme, nodes $\mathcal{K}-1$ send sequences from a codebook randomly generated in an i.i.d. fashion according to a Gaussian distribution. These are the worst noises from the eavesdroppers perspective if Gaussian i.i.d. signaling is used in X_1, see [70, 71]. However, since the channel is fully connected, $\mathcal{K}-1$ are also the worst noises for the intended receiver. This effect causes the secrecy rate to saturate, leading to zero S-DoF.

4.4 Secure DoF of the Single-Antenna Multiple-Access-Channel

In this section we consider the secure multiple-access-channel of (4.1) when all the transmitters and the receivers have a single antenna, i.e., $M_k = N = N_e = 1$ for all $k \in \mathcal{K}$. We have shown in the previous section that Gaussian codebooks lead to zero total S-DoF. Here, we will provide a coding scheme based on integer codebooks and show that for almost all channel gains a positive total S-DoF is achievable, almost surely. The following theorem illustrates our results.

Theorem 13. *For the Gaussian single antenna multiple-access-channel of (4.1) with $M_k = N = N_e = 1$, a total $\frac{K-1}{K}$ secure degrees-of-freedom can be achieved for almost all channel gains, almost surely.*

Proof: When the transmitters and the receivers are equipped with a single antenna, then the channel model of (4.1) is equivalent as follows:

$$\begin{aligned} Y &= \sum_{k=1}^{K} h_k X_k + \widetilde{W}_1 \\ Z &= \sum_{k=1}^{K} h_{k,e} X_k + \widetilde{W}_2, \end{aligned} \tag{4.22}$$

where $\widetilde{W}_1 \sim \mathcal{N}(0,1)$, $\widetilde{W}_2 \sim \mathcal{N}(0,1)$, and $E[X_k^2] \leq P$ for all $k \in \mathcal{K}$. Let us define $\widetilde{X}_k \stackrel{\triangle}{=} \frac{h_{k,e}}{A} X_k$ and $\widetilde{h}_k \stackrel{\triangle}{=} \frac{h_k}{h_{k,e}}$ and without loss of generality assume that $\widetilde{h}_K = 1$, then the channel model is equivalent as follows:

$$\begin{aligned} Y &= A \left[\sum_{k=1}^{K-1} \widetilde{h}_k \widetilde{X}_k + \widetilde{X}_K \right] + \widetilde{W}_1 \\ Z &= A \sum_{k=1}^{K} \widetilde{X}_k + \widetilde{W}_2, \end{aligned} \tag{4.23}$$

where, $A^2 E[\widetilde{X}_k^2] \leq \widetilde{P} \stackrel{\triangle}{=} h_{k,e}^2 P$. In this model we say that the signals are aligned at the eavesdropper according to the following definition:

Definition 13. *The signals $\widetilde{X}_1$, $\widetilde{X}_2$,...,$\widetilde{X}_K$ are said to be aligned at a receiver if its received signal is a rational combination of them.*

Note that, in n-dimensional Euclidean spaces ($n \geq 2$), two signals are aligned when they are received in the same direction at the receiver. In general, m signals are aligned

at a receiver if they span a subspace with dimension less than m. The above definition, however, generalizes the concept of alignment for the one-dimensional real numbers. Our coding scheme is based on integer codebooks, which means that $\widetilde{X}_k \in \mathbb{Z}$ for all $k \in \mathcal{K}$. If some integer signals are aligned at a receiver, then their effect is similar to a single signal at high SNR regimes. This is due to the fact that rational numbers form a field and therefore the sum of constellations from $\mathbb{Q}$ form a constellation in $\mathbb{Q}$ with an enlarged cardinality.

Before we present our achievability scheme, we need to define the rational dimension of a set of real numbers.

Definition 14. *(*Rational Dimension*) The rational dimension of a set of real numbers* $\{\widetilde{h}_1, \widetilde{h}_2, ..., \widetilde{h}_{K-1}, \widetilde{h}_K = 1\}$ *is* M *if there exists a set of real numbers* $\{g_1, g_2, ..., g_M\}$ *such that each* $\widetilde{h}_k$ *can be represented as a rational combination of* g_i*s, i.e.,* $\widetilde{h}_k = a_{k,1}g_1 + a_{k,2}g_2 + ... + a_{k,M}g_M$*, where* $a_{k,i} \in \mathbb{Q}$ *for all* $k \in \mathcal{K}$ *and* $i \in \mathcal{M}$*.*

In fact, the rational dimension of a set of channel gains is the effective dimension seen at the corresponding receiver. In particular, $\{\widetilde{h}_1, \widetilde{h}_2, ..., \widetilde{h}_K\}$ are *rationally independent* if the rational dimension is K, i.e., none of the $\widetilde{h}_k$ can be represented as the rational combination of other numbers.

Note that all of the channel gains $\widetilde{h}_k$ are generated independently with a distribution. From the number theory, it is known that the set of all possible channel gains that are rationally independent have a Lebesgue measure 1. Therefore, we can assume that $\{\widetilde{h}_1, \widetilde{h}_2, ..., \widetilde{h}_K\}$ are rationally independent, almost surely. Our achievability coding scheme is as follows:

Encoding

Each transmitter limits its input symbols to a finite set which is called the transmit constellation. Even though it has access to the continuum of real numbers, restriction to a finite set has the benefit of easy and feasible decoding at the intended receiver. The transmitter k selects a constellation $\mathcal{V}_k$ to send message W_k. The constellation points are chosen from integer points, i.e., $\mathcal{V}_k \subset \mathbb{Z}$. We assume that $\mathcal{V}_k$ is a bounded set. Hence, there is a constant Q_k such that $\mathcal{V}_k \subset [-Q_k, Q_k]$. The cardinality of $\mathcal{V}_k$ which limits the rate of message W_k is denoted by $\|\mathcal{V}_k\|$.

Having formed the constellation, the transmitter k constructs a random codebook for message W_k with rate R_k. This can be accomplished by choosing a probability distribution on the input alphabets. The uniform distribution is the first candidate and it is selected for the sake of simplicity. Therefore, the stochastic encoder k generates $2^{n(I(\widetilde{X}_k;Y|\widetilde{X}_{(\mathcal{K}-k)^c})+\epsilon_k)}$ independent and identically distributed sequences $\widetilde{x}_k^n$ according to the distribution $P(\widetilde{x}_k^n) = \prod_{i=1}^n P(\widetilde{x}_{k,i})$, where $P(\widetilde{x}_{k,i})$ denotes the probability distribution function of the uniformly

distributed random variable $\widetilde{x}_{k,i}$ over $\mathcal{V}_k$. Next, randomly distribute these sequences into 2^{nR_k} bins. Index each of the bins by $w_k \in \{1, 2, ..., 2^{nR_k}\}$.

For each user $k \in \mathcal{K}$, to send message w_k, the transmitter looks for a $\widetilde{x}_k^n$ in bin w_k. The rates are such that there exist more than one $\widetilde{x}_k^n$. The transmitter randomly chooses one of them and sends $x_k^n = A\frac{\widetilde{x}_k^n}{h_{k,e}}$. The parameter A controls the input power.

Decoding

At a specific time, the received signal at the legitimate receiver is as follows:

$$Y = A\left[\widetilde{h}_1\widetilde{X}_1 + \widetilde{h}_2\widetilde{X}_2 + ... + \widetilde{h}_{K-1}\widetilde{X}_{K-1} + \widetilde{X}_K\right] + \widetilde{W}_1 \tag{4.24}$$

The legitimate receiver passes the received signal Y through a hard decoder. The hard decoder looks for a point $\widetilde{Y}$ in the received constellation

$$\mathcal{V}_r = A\left[\widetilde{h}_1\mathcal{V}_1 + \widetilde{h}_2\mathcal{V}_2 + ... + \widetilde{h}_{K-1}\mathcal{V}_{K-1} + \mathcal{V}_K\right]$$

which is the nearest point to the received signal Y. Therefore, the continuous channel changes to a discrete one in which the input symbols are taken from the transmit constellations $\mathcal{V}_k$ and the output symbols belongs to the received constellation $\mathcal{V}_r$. $\widetilde{h}_k$'s are rationally independent which means that the equation $A\left[\widetilde{h}_1X_1 + \widetilde{h}_2X_2 + ... + \widetilde{h}_{K-1}X_{K-1} + X_K\right] = 0$ has no rational solution. This property implies that any real number v_r belonging to the constellation $\mathcal{V}_r$ is uniquely decomposable as $v_r = A\sum_{k=1}^{K}\widetilde{h}_k\widehat{\widetilde{X}}_k$. Note that if there exists another possible decomposition $\widetilde{v}_r = A\sum_{k=1}^{K}\widetilde{h}_k\widehat{\widetilde{X}}_k^{'}$, then $\widetilde{h}_k$'s have to be rationally-dependent, which is a contradiction. We refer to this property as property Γ. This property in fact implies that if there is no additive noise in the channel, then the receiver can decode all the transmitted signals with zero error probability.

Remark 10. *In a random environment it is easy to show that the set of channel gains which are rationally-dependent have a measure of zero with respect to the Lebesque measure. Therefore, Property* Γ *is almost surely satisfied.*

Error Probability Analysis

Let $d_{\min}$ denote the minimum distance in the received constellation $\mathcal{V}_r$. Having property Γ, the receiver can decode the transmitted signals. Let V_r and $\hat{V}_r$ be the transmitted and decoded symbols, respectively. The probability of error, i.e., $P_e = P(\hat{V}_r \neq V_r)$, is bounded as follows:

$$P_e \leq Q(\frac{d_{\min}}{2}) \leq \exp(-\frac{d_{\min}^2}{8}) \tag{4.25}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{t^2}{2})dt$. Note that finding $d_{\min}$ is not easy in general. Using Khintchine and Groshev theorems, however, it is possible to lower bound the minimum distance. Here we explain some background information for using the theorems of Khintchine and Groshev.

The field of Diophantine approximation in number theory deals with the approximation of real numbers with rational numbers. The reader is referred to [83, 84] and the references therein. The Khintchine theorem is one of the cornerstones in this field. This theorem provides criteria for a given function $\psi : \mathbb{N} \to \mathbb{R}_+$ and real number $\widetilde{h}$, such that $|p + \widetilde{h}q| < \psi(|q|)$ has either infinitely many solutions or at most, finitely many solutions for $(p, q) \in \mathbb{Z}^2$. Let $\mathcal{A}(\psi)$ denote the set of real numbers such that $|p + \widetilde{h}q| < \psi(|q|)$ has infinitely many solutions in integers. The theorem has two parts. The first part is the convergent part and states that if $\psi(|q|)$ is convergent, i.e.,

$$\sum_{q=1}^{\infty} \psi(q) < \infty \tag{4.26}$$

then $\mathcal{A}(\psi)$ has a measure of zero with respect to the Lebesgue measure. This part can be rephrased in a more convenient way, as follows. For almost all real numbers, $|p+\widetilde{h}q| > \psi(|q|)$ holds for all $(p, q) \in \mathbb{Z}^2$ except for finitely many of them. Since the number of integers violating the inequality is finite, one can find a constant c such that

$$|p + \widetilde{h}q| > c\psi(|q|) \tag{4.27}$$

holds for all integers p and q, almost surely. The divergent part of the theorem states that $\mathcal{A}(\psi)$ has the full measure, i.e. the set $\mathbb{R} - \mathcal{A}(\psi)$ has measure zero, provided that ψ is decreasing and $\psi(|q|)$ is divergent, i.e.,

$$\sum_{q=1}^{\infty} \psi(q) = \infty. \tag{4.28}$$

There is an extension to Khintchines theorem which regards the approximation of linear forms. Let $\widetilde{\mathbf{h}} = (\widetilde{h}_1, \widetilde{h}_2, ..., \widetilde{h}_{K-1})$ and $\mathbf{q} = (q_1, q_2, ..., q_{K-1})$ denote $(K-1)$-tuples in $\mathbb{R}^{K-1}$ and $\mathbb{Z}^{K-1}$, respectively. Let $\mathcal{A}_{K-1}(\psi)$ denote the set of $(K-1)$-tuple real numbers $\widetilde{\mathbf{h}}$ such that

$$|p + q_1\widetilde{h}_1 + q_2\widetilde{h}_2 + ... + q_{K-1}\widetilde{h}_{K-1}| < \psi(|\mathbf{q}|_\infty) \tag{4.29}$$

has infinitely many solutions for $p \in \mathbb{Z}$ and $\mathbf{q} \in \mathbb{Z}^{K-1}$. Here, $|\mathbf{q}|_\infty$ is the supremum norm of $\mathbf{q}$ which is defined as $\max_k |q_k|$. The following theorem illustrates the Lebesque measure of the set $\mathcal{A}_{K-1}(\psi)$.

Theorem 14. *(Khintchine-Groshev) Let $\psi : \mathbb{N} \to \mathbb{R}_+$. Then, the set $\mathcal{A}_{K-1}(\psi)$ has measure zero provided that*

$$\sum_{q=1}^{\infty} q^{K-2}\psi(q) < \infty \tag{4.30}$$

and has the full measure if

$$\sum_{q=1}^{\infty} q^{K-2}\psi(q) = \infty \quad \text{and } \psi \text{ is monotonic} \tag{4.31}$$

In this chapter, we are interested in the convergent part of the theorem. Moreover, given an arbitrary $\epsilon > 0$ the function $\psi(q) = \frac{1}{q^{K-1+\epsilon}}$ satisfies the condition of (4.30). In fact, the convergent part of the above theorem can be stated as follows. For almost all $K-1$-tuple real numbers $\widetilde{\mathbf{h}}$ there exists a constant c such that

$$|p + q_1\widetilde{h}_1 + q_2\widetilde{h}_2 + ... + q_{K-1}\widetilde{h}_{K-1}| > \frac{c}{(\max_k |q_k|)^{K-1+\epsilon}} \tag{4.32}$$

holds for all $p \in \mathbb{Z}$ and $\mathbf{q} \in \mathbb{Z}^{K-1}$. The Khintchine-Groshev theorem can be used to bound the minimum distance of points in the received constellation $\mathcal{V}_r$. In fact, a point in the received constellation has a linear form of

$$v_r = A\left[\widetilde{h}_1 v_1 + \widetilde{h}_2 v_2 + ... + \widetilde{h}_{K-1}v_{K-1} + v_K\right], \tag{4.33}$$

Therefore, we can conclude that

$$d_{\min} > \frac{Ac}{(\max_{k\in\{1,2,...,K-1\}} Q_k)^{K-1+\epsilon}}. \tag{4.34}$$

The probability of error in hard decoding, see (4.25), can be bounded as:

$$P_e < \exp\left(-\frac{(Ac)^2}{8(\max_{k\in\{1,2,...,K-1\}} Q_k)^{2K-2+2\epsilon}}\right) \tag{4.35}$$

Let us assume that Q_k for all $k \in \{1, 2, ..., K-1\}$ is $Q = \lfloor \widetilde{P}^{\frac{1-\epsilon}{2(K+\epsilon)}} \rfloor$. Moreover, since $E[\widetilde{X}_k^2] \leq A^2 Q_k^2 \leq \widetilde{P}$, we can choose $A = \widetilde{P}^{\frac{K-1+2\epsilon}{2(K+\epsilon)}}$. Substituting in (4.35) yields

$$P_e < \exp(-\frac{c^2}{8}\widetilde{P}^{\epsilon}). \tag{4.36}$$

Thus, $P_e \to 0$ when $\widetilde{P} \to \infty$ or equivalently $P \to \infty$.

Equivocation Calculation

Since the equivocation analysis of Theorem 10 is valid for any input distribution, therefore integer inputs satisfy the perfect secrecy constraint.

S-DoF Calculation

The maximum achievable sum rate is as follows:

$$\begin{aligned}\sum_{k\in\mathcal{K}} R_k &= I(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K;Y) - I(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K;Z) \\ &= H(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K|Z) - H(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K|Y) \\ &\overset{(a)}{\geq} H(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K|Z) - 1 - P_e\log\|\widetilde{\mathcal{X}}\| \\ &\overset{(b)}{\geq} H(\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_K|\sum_{k\in\mathcal{K}}\widetilde{X}_k) - 1 - P_e\log\|\widetilde{\mathcal{X}}\| \\ &\overset{(c)}{=} \sum_{k\in\mathcal{K}} H(\widetilde{X}_k) - H(\sum_{k\in\mathcal{K}}\widetilde{X}_k) - 1 - P_e\log\|\widetilde{\mathcal{X}}\| \\ &\overset{(d)}{=} K\log(2Q+1) - \log(2KQ+1) - 1 - P_e\log\|\widetilde{\mathcal{X}}\|,\end{aligned} \tag{4.37}$$

where (a) follows from Fano's inequality, (b) follows from the fact that conditioning always decreases entropy, (c) follows from chain rule, and (d) follows from the fact that $\widetilde{X}_k$ has uniform distribution over $\mathcal{V}_k = [-Q, Q]$. The S-DoF can therefore be computed as follows:

$$\begin{aligned}\eta &= \lim_{P\to\infty}\frac{\sum_{k\in\mathcal{K}} R_k}{\frac{1}{2}\log P} \\ &= \frac{(K-1)(1-\epsilon)}{K+\epsilon}\end{aligned} \tag{4.38}$$

Since ϵ can be arbitrarily small, then $\eta = \frac{K-1}{K}$ is indeed achievable. □

As we saw in the previous section, multiple-antennas (or equivalently, time-varying and/or frequency-selective channels) provide enough freedom, which allows us to choose appropriate signaling directions to separate between the messages at the intended receiver and at the same time pack the signals into a low dimensionality subspace at the eavesdropper. In contrast, it was commonly believed that time-invariant frequency flat single-antenna channels cannot provide any degrees-of-freedom. In Theorem 13, however, we developed a machinery that transforms the single-antenna systems to a pseudo multiple-antenna system with some antennas. The number of available dimensions in the equivalent pseudo multiple-antenna systems is K when all of K channel gains between the transmitters and the intended receiver are rationally-independent (this condition is satisfied almost surely). The equivalent pseudo multiple-antenna system can simulate the behaviour of a multiple-dimensional system (in time/frequency/space) and allows us to simultaneously separate the signals at the intended receiver and align them to the eavesdropper.

Note that in the MISOSE wiretap channel (Multiple-Input Single-Output Single-Eavesdropper), when the channel realization is unique, we can achieve the optimum S-DoF of 1 through cooperation among transmitters. This theorem clarifies the fact that we lose the amount of $\frac{1}{K}$ in S-DoF due to the lack of cooperation between the transmitters. We still gain through the possibility of signal alignment at the eavesdropper, however.

4.4.1 Rationally-Dependent Channel Gains: Multiple-layer coding

When the channel gains are rationally dependent, then a more sophisticated multiple-layer constellation design is required to achieve higher S-DoF. The reason is that some messages share the same dimension at the intended receiver and as a result, splitting them requires more structure in constellations. We propose a multiple-layer constellation that can not only be distinguished at the intended receiver but are also packed efficiently at the eavesdropper. This is accomplished by allowing a carry-over from different levels. In this subsection, for the sake of simplicity, we consider a two user-secure MAC. This channel is modeled as follows:

$$\begin{aligned} Y&= A\left[\widetilde{h}_1\widetilde{X}_1+\widetilde{X}_2\right]+\widetilde{W}_1 \\ Z&= A\left[\widetilde{X}_1+\widetilde{X}_2\right]+\widetilde{W}_2, \end{aligned} \tag{4.39}$$

where A controls the output power. When the channel gain $\widetilde{h}_1$ is irrational then according to Theorem 13 the total S-DoF of $\frac{1}{2}$ is indeed achievable. For the rational channel gain $\widetilde{h}_1=\frac{n}{m},\ n,m\in\mathbb{N}, m\neq 0$, however, the coding scheme of Theorem 13 fails and we need to use a multiple-layer coding scheme.

In multiple-layer coding scheme, we select the constellation points in the base $W\in\mathbb{N}$ as follows:

$$v(\mathbf{b})=\sum_{l=0}^{L-1} b_l W^l, \tag{4.40}$$

where, b_l for all $l\in\{0,1,...,L-1\}$ are independent random variables which take value from $\{0,1,2,...,,a-1\}$ with uniform probability distribution. $\mathbf{b}$ represents the vector $\mathbf{b}=\{b_0,b_1,...,b_{L-1}\}$ and the parameter a controls the number of constellation points. We assume that $a<W$ and therefore, all constellation points are distinct and the size of the constellations are $|\mathcal{V}_1|=|\mathcal{V}_2|=a^L$. The maximum possible rate for each user is therefore bounded by $L\log a$.

At each transmitter a random codebook is generated by randomly choosing b_l according to a uniform distribution. The signal transmitted by users 1 and 2 are $\widetilde{X}_1=v(\mathbf{b})$ and $\widetilde{X}_2=$

$v(\mathbf{b}^{'})$, respectively. Note that the above multiple-layer constellation had a DC component and this component needs to be removed at the transmitters. The DC component, however, duplicated the achievable rates and has no effect on the S-DoF.

To calculate A, since b_l and b_j are independent for $l \neq j$, we have the following chain of inequalities:

$$\begin{aligned} A^2 E[\widetilde{X}_1^2] &= A^2 W^{2(L-1)} \sum_{l=0}^{L-1} E[b_l^2] W^{-2l} \\ &\leq A^2 W^{2(L-1)} \frac{(a-1)(2a-1)}{6} \sum_{l=0}^{\infty} W^{-2l} \\ &\leq A^2 W^{2(L-1)} \frac{a^2}{3} \frac{1}{1-W^{-2}} \\ &\leq \frac{A^2 a^2 W^{2L}}{W^2-1}. \end{aligned} \tag{4.41}$$

Therefore, by choosing $A = \frac{\sqrt{(W^2-1)\widetilde{P}}}{aW^L}$ the power constraint $A^2 E\widetilde{X_1^2} \leq \widetilde{P}$ is satisfied. The received constellation at the intended receiver and the eavesdropper can be written as follows, respectively:

$$\begin{aligned} Y &= \frac{A}{m} \sum_{l=0}^{L-1} \left(n b_l + m b_l^{'} \right) W^l + \widetilde{W}_1 \\ Z &= A \sum_{l=0}^{L-1} \left(b_l + b_l^{'} \right) W^l + \widetilde{W}_2. \end{aligned} \tag{4.42}$$

A point in the received constellation $\mathcal{V}_r$ of the intended receiver can be represented as follows:

$$v_r(\mathbf{b}, \mathbf{b}^{'}) = \frac{A}{m} \sum_{l=0}^{L-1} \left(n b_l + m b_l^{'} \right) W^l. \tag{4.43}$$

Note that the received constellation needs to satisfy the property Γ, as the intended receiver needs to uniquely decode the transmitted signals. The following theorem characterizes the total achievable S-DoF.

Theorem 15. *The following S-DoF is achievable for the two-user single antenna MAC with rational channel gain $\widetilde{h}_1 = \frac{n}{m}$:*

$$\eta = \begin{cases} \frac{\log(n)}{\log(n(2n-1))}, & \text{if } 2n \geq m \\ \frac{\log(s+1)}{\log((s+1)(2s+1))}, & \text{if } 2n < m \text{ and } m = 2s+1 \\ \frac{\log(s)}{\log(2s^2-n)}, & \text{if } 2n < m \text{ and } m = 2s \end{cases} \tag{4.44}$$

Proof: Let us first assume that the property Γ is satisfied for given W and a. It is easy to show that the minimum distance in the received constellation $\mathcal{V}_r$ is $d_{\min} = \frac{A}{m}$. The probability of error is therefore bounded as follows:

$$\begin{aligned} P_e &\leq \exp(-\frac{d_{\min}^2}{8}) \\ &= \exp\left(-\frac{(W^2-1)\widetilde{P}}{8a^2m^2W^{2L}}\right). \end{aligned} \tag{4.45}$$

Let us choose L as

$$L = \left\lfloor \frac{\log(\widetilde{P}^{0.5-\epsilon})}{\log(W)} \right\rfloor, \tag{4.46}$$

where $\epsilon > 0$ is an arbitrary small constant. Clearly, with this choice of L, $P_e \leq \exp(-\gamma \widetilde{P}^{2\epsilon})$ where γ is a constant. Thus, when $SNR \to \infty$, then $P_e \to 0$. The S-DoF of the system can therefore be derived as follows:

$$\begin{aligned} \eta &= \lim_{\widetilde{P}\to\infty} \frac{L\log(a)}{\frac{1}{2}\log\widetilde{P}} \\ &= \lim_{\widetilde{P}\to\infty} \frac{\left\lfloor \frac{\log(\widetilde{P}^{0.5-\epsilon})}{\log(W)} \right\rfloor \log(a)}{\frac{1}{2}\log\widetilde{P}} \\ &= (1-2\epsilon)\frac{\log(a)}{\log(W)}. \end{aligned} \tag{4.47}$$

Since ϵ is an arbitrary small constant, the total S-DoF of the system is

$$\eta = \frac{\log(a)}{\log(W)}. \tag{4.48}$$

This equation implies that to achieve the maximum possible η, we need to maximize a and minimize W with the constraint that the property Γ is satisfied. Table I shows the choices of a and W for Theorem 15. To complete the proof we need to show that with the choices of Table I, the property Γ is satisfied.

Lemma 8. *The property Γ holds for all the choices of Table I.*

Please see section 4.5.3. □

Note that this result implies that the total achievable S-DoF by using integer lattice codes is discontinuous with respect to the channel coefficients.

Table 4.1: Chosen a and W to satisfy property Γ

	$\widetilde{h}_1 = \frac{n}{m}$	a	W
Case 1	$2n \geq m$	n	$n(2n-1)$
Case 2	$2n < m$ and $m = 2s+1$	$s+1$	$(s+1)(2s+1)$
Case 3	$2n < m$ and $m = 2s$	s	$2s^2 - n$

4.5 Proofs for Chapter 4

4.5.1 Proof of Theorem 10

1) *Codebook Generation*: The structure of the encoder for user $k \in \mathcal{K}$ is depicted in Figure Fix $P(u_k)$ and $P(x_k|u_k)$. The stochastic encoder k generates $2^{n(I(U_k;Y|U_{(\mathcal{K}-k)^c})+\epsilon_k)}$ independent and identically distributed sequences u_k^n according to the distribution $P(u_k^n) = \prod_{i=1}^n P(u_{k,i})$. Next, randomly distribute these sequences into 2^{nR_k} bins. Index each of the bins by $w_k \in \{1, 2, ..., 2^{nR_k}\}$.

2) *Encoding*: For each user $k \in \mathcal{K}$, to send message w_k, the transmitter looks for a u_k^n in bin w_k. The rates are such that there exist more than one u_k^n. The transmitter randomly chooses one of them and then generates x_k^n according to $P(x_k^n|u_k^n) = \prod_{i=1}^n P(x_{k,i}|u_{k,i})$ and sends it.

3) *Decoding*: The received signals at the legitimate receiver, y^n, is the output of the channel $P(y^n|x_{\mathcal{K}}^n) = \prod_{i=1}^n P(y_i|x_{\mathcal{K},i})$. The legitimate receiver looks for the unique sequence $u_{\mathcal{K}}^n$ such that $(u_{\mathcal{K}}^n, y^n)$ is jointly typical and declares the indices of the bins containing u_k^n as the messages received.

4) *Error Probability Analysis*: Since the region of (4.12) is a subset of the capacity region of the multiple-access-channel without secrecy constraint, then the error probability analysis is the same as [62] and omitted here.

5) *Equivocation Calculation*: To satisfy the perfect secrecy constraint, we need to prove

the requirement of (4.7). From $H(W_{\mathcal{K}}|Z^n)$ we have

$$\begin{aligned} H(W_{\mathcal{K}}|Z^n) &= H(W_{\mathcal{K}}, Z^n) - H(Z^n) \\ &= H(W_{\mathcal{K}}, U_{\mathcal{K}}^n, Z^n) - H(U_{\mathcal{K}}^n|W_{\mathcal{K}}, Z^n) - H(Z^n) \\ &= H(W_{\mathcal{K}}, U_{\mathcal{K}}^n) + H(Z^n|W_{\mathcal{K}}, U_{\mathcal{K}}^n) - H(U_{\mathcal{K}}^n|W_{\mathcal{K}}, Z^n) - H(Z^n) \\ &\overset{(a)}{\geq} H(W_{\mathcal{K}}, U_{\mathcal{K}}^n) + H(Z^n|W_{\mathcal{K}}, U_{\mathcal{K}}^n) - n\epsilon_n - H(Z^n) \\ &\overset{(b)}{=} H(W_{\mathcal{K}}, U_{\mathcal{K}}^n) + H(Z^n|U_{\mathcal{K}}^n) - n\epsilon_n - H(Z^n) \\ &\overset{(c)}{\geq} H(U_{\mathcal{K}}^n) + H(Z^n|U_{\mathcal{K}}^n) - n\epsilon_n - H(Z^n) \\ &= H(U_{\mathcal{K}}^n) - I(U_{\mathcal{K}}^n; Z^n) - n\epsilon_n \\ &\overset{(d)}{\geq} I(U_{\mathcal{K}}^n; Y^n) - I(U_{\mathcal{K}}^n; Z^n) - n\epsilon_n \\ &\overset{(e)}{\geq} n\sum_{k\in\mathcal{K}} R_k - n\epsilon_n - n\delta_{1n} - n\delta_{4n} \\ &= H(W_{\mathcal{K}}) - n\epsilon_n - n\delta_{1n} - n\delta_{4n}, \end{aligned} \tag{4.49}$$

where (a) follows from Fano's inequality, which states that for sufficiently large n, $H(U_{\mathcal{K}}^n|W_{\mathcal{K}}, Z^n) \leq h(P_{we}^{(n)}) + nP_{we}^n R_w \leq n\epsilon_n$. Here P_{we}^n denotes the wiretapper's error probability of decoding $u_{\mathcal{K}}^n$ in the case that the bin numbers $w_{\mathcal{K}}$ are known to the eavesdropper and $R_w = I(U_{\mathcal{K}}; Z)$. Since the sum rate is small enough, then $P_{we}^n \to 0$ for sufficiently large n. (b) follows from the following Markov chain: $W_{\mathcal{K}} \to U_{\mathcal{K}}^n \to Z^n$. Hence, we have $H(Z^n|W_{\mathcal{K}}, U_{\mathcal{K}}^n) = H(Z^n|U_{\mathcal{K}}^n)$. (c) follows from the fact that $H(W_{\mathcal{K}}, U_{\mathcal{K}}^n) \geq H(U_{\mathcal{K}}^n)$. (d) follows from that fact that $H(U_{\mathcal{K}}^n) \geq I(U_{\mathcal{K}}^n; Y^n)$. (e) follows from the lemma 12 in the Appendix.

4.5.2 Proof of the Converse for Theorem 11

Before starting the proof, we first present some useful lemmas.

Lemma 9. *The secrecy sum capacity of the Gaussian MIMO MAC is upper-bounded by*

$$C_{sum} \leq \max_{P(\mathbf{x_1})P(\mathbf{x_2})...P(\mathbf{x_K})} I(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_K}; \mathbf{y}|\mathbf{z}), \tag{4.50}$$

where maximization is over all distributions $P(\mathbf{x_1})P(\mathbf{x_2})...P(\mathbf{x_K})$ that satisfy the power constraint, i.e., $Tr(\mathbf{x}^\dagger\mathbf{x}) \leq P$.

Proof: According to Fano's inequality and the perfect secrecy constraint, we have

$$\begin{aligned} n\sum_{k\in\mathcal{K}} R_k \leq & I(W_{\mathcal{K}};\mathbf{y^n}) - I(W_{\mathcal{K}};\mathbf{z^n}) \\ & \overset{(a)}{\leq} I(W_{\mathcal{K}};\mathbf{y^n},\mathbf{z^n}) - I(W_{\mathcal{K}};\mathbf{z^n}) \\ & \overset{(b)}{=} I(W_{\mathcal{K}};\mathbf{y^n}|\mathbf{z^n}) \\ & = h(\mathbf{y^n}|\mathbf{z^n}) - h(\mathbf{y^n}|W_{\mathcal{K}},\mathbf{z^n}) \\ & \overset{(c)}{\leq} h(\mathbf{y^n}|\mathbf{z^n}) - h(\mathbf{y^n}|W_{\mathcal{K}},\mathbf{x_{\mathcal{K}}^n},\mathbf{z^n}) \\ & \overset{(d)}{\leq} h(\mathbf{y^n}|\mathbf{z^n}) - h(\mathbf{y^n}|\mathbf{x_{\mathcal{K}}^n},\mathbf{z^n}) \\ & \overset{(e)}{\leq} h(\mathbf{y^n}|\mathbf{z^n}) - \sum_{i=1}^{n} h(\mathbf{y(i)}|\mathbf{x}_{\mathcal{K}}(\mathbf{i}),\mathbf{z(i)}) \\ & \overset{(f)}{\leq} \sum_{i=1}^{n} h(\mathbf{y(i)}|\mathbf{z(i)}) - h(\mathbf{y(i)}|\mathbf{x}_{\mathcal{K}}(\mathbf{i}),\mathbf{z(i)}) \\ & \leq nI(\mathbf{x}_{\mathcal{K}};\mathbf{y}|\mathbf{z},q) \\ & \overset{(g)}{\leq} nI(\mathbf{x}_{\mathcal{K}};\mathbf{y}|\mathbf{z}), \end{aligned} \tag{4.51}$$

where (a) and (b) follows from chain rule, (c) follows from the fact that conditioning decreases the differential entropy, (d) follows from the Markov chain $W_{\mathcal{K}} \rightarrow (\mathbf{x_{\mathcal{K}}^n},\mathbf{z^n}) \rightarrow \mathbf{y^n}$, (e) follows from the fact that the channel is memoryless, (f) is obtained by defining a time-sharing random variable q that takes values uniformly over the index set $\{1,2,...,n\}$ and defining $(\mathbf{x}_{\mathcal{K}},\mathbf{y},\mathbf{z})$ to be the tuple of random variables that conditioned on $q=i$, have the same joint distribution as $(\mathbf{x}_{\mathcal{K}}(\mathbf{i}),\mathbf{y(i)},\mathbf{z(i)})$. Finally, (g) follows from the fact that $I(\mathbf{x}_{\mathcal{K}};\mathbf{y}|\mathbf{z})$ is concave in $P(\mathbf{x_1})...P(\mathbf{x_K})$ (see, e.g., [42], Appendix I for a proof), so that Jensens Inequality can be applied. □

Lemma 10. *If* $\mathbf{DH_k} = \mathbf{H_{k,e}}$ *for all* $k \in \mathcal{K}$ *and* $\mathbf{DD}^\dagger \preceq \mathbf{I}$*, then the function*

$$f(\mathbf{X_1},\mathbf{X_2},...,\mathbf{X_K}) = \frac{1}{2}\log\left|\mathbf{I}+\sum_{k\in\mathcal{K}}\mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I}+\sum_{k\in\mathcal{K}}\mathbf{H_{k,e}}\mathbf{X_k}\mathbf{H_{k,e}^\dagger}\right| \tag{4.52}$$

is a concave function of $(\mathbf{X_1},...,\mathbf{X_K})$ *for* $\mathbf{X_k} \succeq 0$ *for all* $k \in \mathcal{K}$*. Moreover, for* $(\mathbf{X_1},...,\mathbf{X_K})$ *such that* $\mathbf{X_k} \succeq 0$ *and* $(\mathbf{\Delta}_1,...,\mathbf{\Delta}_K)$ *such that* $\mathbf{\Delta}_k \succeq 0$*, we have*

$$f(\mathbf{X_1},...,\mathbf{X_K}) \leq f(\mathbf{X_1}+\mathbf{\Delta}_1,...,\mathbf{X_K}+\mathbf{\Delta}_K). \tag{4.53}$$

Proof: Using the degradedness property of $\mathbf{DH_k} = \mathbf{H_{k,e}}$, the function $f(.)$ can be re-written as follows:

$$f(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_K}) = \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e}}\mathbf{X_k}\mathbf{H_{k,e}^\dagger}\right| \tag{4.54}$$

$$= \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{D}\mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\mathbf{D}^\dagger\right|$$

$$\overset{(a)}{=} \frac{1}{2}\log\frac{\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right|}{\left|\left[(\mathbf{D}^\dagger\mathbf{D})^{-1} - \mathbf{I}\right] + \left[\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right]\right| |\mathbf{D}^\dagger\mathbf{D}|}, \tag{4.55}$$

where (a) follows from the fact that $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$. According to [70], Lemma II.3, this function is concave with regard to $\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}$, and is therefore concave with regard to $(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_K})$.

To prove the property of (4.53), note that for any $\mathbf{A} \succeq 0$, $\mathbf{\Delta} \succeq 0$ and $\mathbf{B} \succ 0$, we have the following property [82]:

$$\frac{|\mathbf{B}|}{|\mathbf{A} + \mathbf{B}|} \leq \frac{|\mathbf{B} + \mathbf{\Delta}|}{|\mathbf{A} + \mathbf{B} + \mathbf{\Delta}|}. \tag{4.56}$$

We choose $\mathbf{\Delta} = \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{\Delta}_k\mathbf{H_k^\dagger}$ and apply the above property to (4.55). Thus, we obtain,

$$f(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_K}) = \frac{1}{2}\log\frac{\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right|}{\left|\left[(\mathbf{D}^\dagger\mathbf{D})^{-1} - \mathbf{I}\right] + \left[\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{X_k}\mathbf{H_k^\dagger}\right]\right| |\mathbf{D}^\dagger\mathbf{D}|} \tag{4.57}$$

$$\leq \frac{1}{2}\log\frac{\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\left(\mathbf{X_k} + \mathbf{\Delta_k}\right)\mathbf{H_k^\dagger}\right|}{\left|\left[(\mathbf{D}^\dagger\mathbf{D})^{-1} - \mathbf{I}\right] + \left[\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\left(\mathbf{X_k} + \mathbf{\Delta_k}\right)\mathbf{H_k^\dagger}\right]\right| |\mathbf{D}^\dagger\mathbf{D}|}$$

$$= f(\mathbf{X_1} + \mathbf{\Delta}_1, ..., \mathbf{X_K} + \mathbf{\Delta}_K).$$

□

To prove the converse part, we first start with Lemma 9 to bound the sum rate as follows:

$$\sum_{k\in\mathcal{K}} R_k \leq I(\mathbf{x}_\mathcal{K}; \mathbf{y}|\mathbf{z}) \tag{4.58}$$

$$= h(\mathbf{y}|\mathbf{z}) - h(\mathbf{y}|\mathbf{x}_\mathcal{K}, \mathbf{z})$$

$$= h(\mathbf{y}|\mathbf{z}) - h(\mathbf{n_1})$$

$$\overset{(a)}{=} \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{K_{x_k}}\mathbf{H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e}}\mathbf{K_{x_k}}\mathbf{H_{k,e}^\dagger}\right|$$

$$\overset{(b)}{\leq} \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_k}\mathbf{Q_k}\mathbf{H_k^\dagger}\right| - \frac{1}{2}\log\left|\mathbf{I} + \sum_{k\in\mathcal{K}} \mathbf{H_{k,e}}\mathbf{Q_k}\mathbf{H_{k,e}^\dagger}\right|,$$

where (a) follows from the fact that $h(\mathbf{y}|\mathbf{z})$ is maximized by jointly Gaussian $\mathbf{y}$ and $\mathbf{z}$ for fixed covariance matrix $\mathbf{Q}_{\mathbf{y},\mathbf{z}}$ and (b) follows from the degradedness assumption and therefore concavity and monotonicity properties given in Lemma 10 and the fact that $\mathbf{K}_{\mathbf{x_k}} \preceq \mathbf{Q_k}$.

4.5.3 Proof of Lemma 8

We prove this lemma by induction on L. To show the lemma for $L = 0$, it is sufficient to prove that the equation

$$n(b_0 - \widetilde{b}_0) + m(b_0^{'} - \widetilde{b}_0^{'}) = 0, \tag{4.59}$$

has no nontrivial solution when $b_0, b_0^{'}, \widetilde{b}_0, \widetilde{b}_0^{'} \in \{0, 1, 2, ..., a-1\}$. The necessary and sufficient conditions for the equation (4.59) are that $b_0^{'} - \widetilde{b}_0^{'}$ is dividable by n **and** $b_0 - \widetilde{b}_0$ is dividable by n. We show that one of these conditions does not hold for all choices of Table I.

Case 1: In this case $a = n$. Following the fact that $-(n-1) \leq b_0^{'} - \widetilde{b}_0^{'} \leq n-1$, it is easy to deduce that $n \nmid (b_0^{'} - \widetilde{b}_0^{'})$.,

Case 2: In this case $a = s+1$ where $m = 2s+1$. Following the fact that $-(2s+1) \leq b_0 - \widetilde{b}_0 \leq 2s+1$, it is easy to deduce that $m \nmid b_0 - \widetilde{b}_0$.

Case 3: In this case $a = s$ where $m = 2s$. Following the fact that $-2s \leq b_0 - \widetilde{b}_0 \leq 2s$, it is easy to show that $m \nmid b_0 - \widetilde{b}_0$.

Now assume that property Γ holds for $L-1$. We need to show that the equation

$$\frac{A}{m} \sum_{l=0}^{L-1} \left(n(b_l - \widetilde{b}_l) + m(b_l^{'} - \widetilde{b}_l^{'}) \right) W^l = 0, \tag{4.60}$$

has no nontrivial solution. Equivalently, this equation can be written as follows:

$$n(b_0 - \widetilde{b}_0) + m(b_0^{'} - \widetilde{b}_0^{'}) = W \sum_{l=0}^{L-2} \left(n(b_{l+1} - \widetilde{b}_{l+1}) + m(b_{l+1}^{'} - \widetilde{b}_{l+1}^{'}) \right) W^l. \tag{4.61}$$

We prove that the above equation has no nontrivial solution in two steps. First, assume that the right hand side of (4.61) is zero. This equation therefore reduces to

$$n(b_0 - \widetilde{b}_0) + m(b_0^{'} - \widetilde{b}_0^{'}) = 0, \tag{4.62}$$

which we have already shown has no nontrivial solution for all three cases.

Secondly, assume that the right side of (4.61) is non-zero. The equation (4.61) can therefore be written as follows:

$$n(b_0 - \widetilde{b}_0) + m(b_0^{'} - \widetilde{b}_0^{'}) = cW, \tag{4.63}$$

where $c \in \mathbb{Z}$ and $c \neq 0$. We need to prove that equation (4.63) has no nontrivial solution for all three cases.

Case 1: In this case $W = n(2n-1)$ and n divides $n(b_0 - \widetilde{b}_0)$ and cW as well. However, as $(m,n) = 1$ and $-(n-1) \leq b_0 - \widetilde{b}_0 \leq n-1$, the equation (4.63) has a solution iff $b_0' = \widetilde{b}_0'$ which is a contradiction with the fact that $n|b_0 - \widetilde{b}_0| < n(n-1) < |c|n(2n-1) = |c|W$.

Case 2: In this case $W = (s+1)(2s+1)$, $n = s+1$ and $m = 2s+1$. Thus, $2s+1$ divides both $m(b_0' - \widetilde{b}_0')$ and cW. Since $(2n, m = 2s+1) = 1$ and $-2s \leq b_0 - \widetilde{b}_0 \leq 2s$, therefore, $2s+1$ cannot divide $n(b_0 - \widetilde{b}_0)$. Hence, equation (4.63) has a solution iff $b_0 = \widetilde{b}_0$ which contradicts the fact that $m|b_0' - \widetilde{b}_0'| < |c|W$.

Case 3: In this case $W = 2s^2 - n$, $a = s$, $2n < m$ and $m = 2s$. We have

$$\begin{aligned} |n(b_0 - \widetilde{b}_0) + m(b_0' - \widetilde{b}_0')| &< m|b_0 - \widetilde{b}_0 + b_0' - \widetilde{b}_0'| \\ &\leq 2m(a-1) \\ &= 4s(s-1) \\ &< 2W \end{aligned} \tag{4.64}$$

and therefore, it suffices to assume $c = 1$. Substituting $W = 2s^2 - n$ in (4.63), we have the following equation:

$$n(b_0 - \widetilde{b}_0 + 1) + 2s(b_0' - \widetilde{b}_0') = 2s^2. \tag{4.65}$$

Obviously, $2s$ divides $2s(b_0' - \widetilde{b}_0')$ and $= 2s^2$. However, since $(2s, n) = 1$ and $-(2s-1) \leq b_0 - \widetilde{b}_0 \leq 2s-1$, equation (4.65) has a solution iff $b_0 = \widetilde{b}_0 - 1$ which is impossible due to the fact $2s|b_0' - \widetilde{b}_0'| < 2s^2$. This completes the proof.

Chapter 5

Conclusion

In this final chapter, we summarize our contributions and point out several directions for future research.

5.1 Contributions

The notion of physical-layer security is based on the idea that noise and losses are resources for information-theoretic security. Physical-layer security has the potential of significantly strengthening the security level of current systems, by introducing a level of information-theoretic security instead of computational security. In this dissertation, we investigated the design of secure channel coding schemes for multi-user wireless communication systems. In particular, the core of this dissertation focuses on multi-user information-theoretic security analysis.

In chapter 2, a generalization of the wiretap channel in the case of two receivers and one eavesdropper was considered. We established an inner bound for the general (non-degraded) case. This bound matches Marton's bound on broadcast channels without security constraint. Furthermore, we considered the scenario in which the channels are degraded. We established the perfect secrecy capacity region for this case. The achievability coding scheme is a secret superposition scheme where randomization in the first layer helps the secrecy of the second layer. The converse proof combines the converse proof for the degraded broadcast channel without security constraint, and the perfect secrecy constraint. We proved that the secret superposition scheme with the Gaussian codebook is optimal in AWGN-BCE. The converse proof is based on Costa's entropy power inequality and Csiszar lemma. Based on the rate characterization of the AWGN-BCE, the broadcast strategy for the slowly fading wiretap channel were used. In this strategy, the transmitter only knows the eavesdropper's channel

and the source node sends secure layered coding. The receiver is viewed as a continuum of ordered users. We derived the optimum power allocation for the layers, which maximizes the total average rate.

In chapter 3, a scenario in which a source node wishes to broadcast two confidential messages for two respective receivers via a Gaussian MIMO broadcast channel, while a wiretapper also receives the transmitted signal via another MIMO channel is considered. We considered the secure vector Gaussian degraded broadcast channel and established its capacity region. Our achievability scheme was the secret superposition of Gaussian codes. Instead of solving a nonconvex problem, we used the notion of an enhanced channel to show that secret superposition of Gaussian codes is optimal. To characterize the secrecy capacity region of the vector Gaussian degraded broadcast channel, we only enhanced the channels for the legitimate receivers, and the channel of the eavesdropper remained unchanged. We then extended the result of the degraded case to a non-degraded case. We showed that the secret superposition of Gaussian codes, along with successive decoding, cannot work when the channels are not degraded. We developed a Secret Dirty Paper Coding (SDPC) scheme and showed that SDPC is optimal for this channel. We then extended the results of the two user case for the general multiple receivers case. Finally, we investigated a scenario which frequently occurs in the practice of wireless networks. In this scenario, the transmitter and the eavesdropper have multiple antennae, while the intended receivers have a single antenna (representing resource limited mobile units). We characterized the secrecy capacity region in terms of generalized eigenvalues of the receivers' channels and the eavesdropper channel. We showed that our analysis is valid for a general number of single antenna receivers. In high SNR, we showed that the capacity region is a convex closure of rectangular regions.

Finally, in chapter 4, a K-user secure Gaussian MAC with an external eavesdropper was considered. We proved an achievable rate region for the secure discrete memoryless MAC and thereafter we established the secrecy sum capacity of the degraded Gaussian MIMO MAC using Gaussian codebooks. For the non-degraded Gaussian MIMO MAC, we proposed an algorithm inspired by the interference alignment technique to achieve the largest possible total S-DoF. When all the terminals are equipped with single antenna, the Gaussian codebooks lead to zero S-DoF. Therefore, we proposed a novel secure coding scheme to achieve positive S-DoF in the single antenna MAC. This scheme converts the single-antenna system into a multiple-dimension system with fractional dimensions. The achievability scheme is based on the alignment of signals into a small sub-space at the eavesdropper, and the simultaneous separation of the signals at the intended receiver. We proved that total S-DoF of $\frac{K-1}{K}$ can be achieved for almost all channel gains which are rationally independent. For the rationally dependent channel gains, we illustrated the power of the multi-layer coding

scheme, through an example channel, to achieve a positive S-DoF. As a function of channel gains, therefore, we showed that the achievable S-DoF is discontinues.

5.2 Future Research

Although information-theoretic secret communication for wireless channels was first studied by Wyner in 1975, the research on this topic has been disregarded for decades since then. Recently, this topic has attracted more attention from the information theory society. This thesis considered the basic extensions of the original wiretap channel and characterized essential limitations in down-link and up-link wireless channels to design perfect secure codes. The works presented in this dissertation can be extended in the following interesting directions:

- **Practical Code Design**: The research on the physical-layer security topic has mainly focused on the information-theoretic limit of the secrecy rate over various wireless channels. To implement the secure channel codes in a real system, however, we need to study the practical coding schemes with finite code-lengths and analyze their performance. LDPC codes and turbo codes are the most promising candidates as they achieve communication rates close to the shannon capacity without any secrecy requirement. Studying how these codes can be modified for secret communication, and evaluating their performance and sensitivity to coding parameters is an interesting direction for future research. References [85, 86, 87, 88, 89] are valuable resources to start this topic.

- **Cross-Layer Security Protocol Design**: In the physical-layer security solutions, the computational power of the eavesdropper is assumed to be infinite which is seen to be possible in the future quantum systems; however, these solutions rely on other assumptions about the communication channels which may not be entirely accurate in practice. The eavesdropper may, for instance, feedback a smaller CSI than the real one to improve her capabilities in the next transmissions. In light of these considerations, it is likely that the implementation of physical-layer security in a real system will be part of a layered-approach, and the design of protocols that combine traditional cryptographic techniques with physical-layer techniques is an interesting research direction. A key part of this research is the definition of relevant metrics that would make it possible to analyze the performance of these hybrid schemes.

- **Availability of the Intended Receiver's CSI**: In this thesis, we have almost always assumed that the transmitters know the channel state information associated with the

intended receiver. The channel state information can be obtained either by a feedback channel from the receiver, or by utilizing the reciprocity and the duplex properties of the link, i.e. by the intended receiver transmitting a training sequence so that the transmitter can estimate the channel. However, sometimes getting the channel state information at the transmitters is so difficult or inconvenient that it is desirable to work without the channel state information at the transmitters. The secrecy capacity for this scenario alongside the feasible secret communication scheme under this scenario would be very interesting to study.

- **Secure Space-Time Codes**: Multiple antennae have been shown to be beneficial for both the conventional capacity and the secrecy capacity. Optimal space-time codings on multiple-antenna systems for conventional capacity have been extensively studied since the pioneering work in [90]. When it comes to secrecy capacity, the same set of questions arise: what are the tradeoffs among the secrecy, rate and diversity? The problem becomes even more interesting when one considers multi-user communication systems. In [91, 92], this topic is addressed.

- **Secure Gaussian Interference Channel**: Another direction that is worth exploring is the secrecy capacity in a network scenario with multiple transmitter-receiver pairs. Recently, tight outer bounds for the capacity without the secrecy requirement of the two-user Gaussian interference channel and X-channel were derived in [93, 94, 95]. A similar approach might be useful to study the secrecy capacity of the two-user Gaussian interference channel and X-channel. Extending the results to multiple users could be even more exciting, because the mixture of the interference from the multiple users might provide a natural means to protect each individual message against a non-intended receiver or the out-of-network eavesdropper. The interference alignment approach proposed in [79, 96, 80, 81] could be an excellent option because the interference from multiple transmitters are aligned perfectly. Due to the careful alignment, the intended receiver can extract the signal sent to her cleanly, while the non-intended receiver would only get a mixture consisting of signals from all transmitters, and would not have enough SNR to extract any particular message from the mixture. Some results have been developed in [29, 97].

- **Compound Secure Gaussian MAC**: The compound channel models the transmissions over a channel that takes a finite number of states. We studied the Gaussian MAC with single antenna in chapter 4. The idea was to align the secret messages at the eavesdropper and to separate them at the intended receiver. The compound se-

cure Gaussian MAC can be considered as a secure Gaussian MAC with finite number of intended receivers and eavesdroppers. Similarly, by aligning the secure signals at all the eavesdroppers, one can achieve a positive S-DoF in this channel. The results would be more interesting by providing a tight upper-bound on the S-DoF. The works of [98, 99, 100] are useful to start this topic.

Appendix A

Lemma 11. *Assume U^n, V_1^n, V_2^n, Y_1^n, Y_2^n and Z^n are generated according to the achievability scheme of Theorem 1, then we have,*

$$nI(V_1;Y_1|U) - n\delta_{1n} \leq I(V_1^n;Y_1^n|U^n) \leq nI(V_1;Y_1|U) + n\delta_{2n} \tag{A.1}$$
$$nI(V_2;Y_2|U) - n\delta_{3n} \leq I(V_2^n;Y_2^n|U^n) \leq nI(V_2;Y_2|U) + n\delta_{4n} \tag{A.2}$$
$$nI(V_1;Z|U) - n\delta_{5n} \leq I(V_1^n;Z^n|U^n) \leq nI(V_1;Z|U) + n\delta_{6n} \tag{A.3}$$
$$nI(V_1;V_2|U) - n\delta_{7n} \leq I(V_1^n;V_2^n|U^n) \leq nI(V_1;V_2|U) + n\delta_{8n} \tag{A.4}$$
$$nI(V_1,V_2;Z|U) - n\delta_{9n} \leq nI(V_1^n,V_2^n;Z^n|U^n) \leq nI(V_1,V_2;Z|U) + n\delta_{10n} \tag{A.5}$$
$$nI(U;Y_1) - n\delta_{11n} \leq I(U^n;Y_1^n) \leq nI(U;Y_1) + n\delta_{12n} \tag{A.6}$$
$$nI(U;Y_2) - n\delta_{13n} \leq I(U^n;Y_2^n) \leq nI(U;Y_2) + n\delta_{14n} \tag{A.7}$$
$$nI(U;Z) - n\delta_{15n} \leq I(U^n;Z^n) \leq nI(U;Z) + n\delta_{16n}, \tag{A.8}$$

where, $\delta_{in} \to 0$ when $n \to \infty$ for all $i = 1, 2, ..., 16$.

Proof: Here, we only prove (A.1) and using the same method the other inequalities can be proven. Let $A_\epsilon^n(P_{U,V_1,Y_1})$ denote the set of typical sequences (U^n, V_1^n, Y_1^n) with respect to P_{U,V_1,Y_1}, and

$$\zeta = \begin{cases} 1, & (U^n,V_1^n,Y_1^n) \notin A_\epsilon^n(P_{U,V_1,Y_1}); \\ 0, & \text{otherwise,} \end{cases}$$

be the corresponding indicator function. We expand $I(V_1^n;Y_1^n,\zeta|U^n)$ and $I(V_1^n,\zeta;Y_1^n|U^n)$ as follows:

$$\begin{aligned} I(V_1^n;Y_1^n,\zeta|U^n) &= I(V_1^n;Y_1^n|U^n,\zeta) + I(V_1^n;\zeta|U^n) \\ &= I(V_1^n;\zeta|U^n,Y_1^n) + I(V_1^n;Y_1^n|U^n), \end{aligned} \tag{A.9}$$

$$\begin{aligned} I(V_1^n,\zeta;Y_1^n|U^n) &= I(V_1^n;Y_1^n|U^n,\zeta) + I(\zeta;Y_1^n|U^n) \\ &= I(\zeta;Y_1^n|U^n,V_1^n) + I(V_1^n;Y_1^n|U^n), \end{aligned} \tag{A.10}$$

Therefore, we have,

$$I(V_1^n;Y_1^n|U^n,\zeta)-I(\zeta;Y_1^n|U^n,V_1^n)\leq I(V_1^n;Y_1^n|U^n)\leq I(V_1^n;Y_1^n|U^n,\zeta)+I(V_1^n;\zeta|U^n). \quad \text{(A.11)}$$

Note that $I(V_1^n;\zeta|U^n)\leq H(\zeta)\leq 1$ and $I(\zeta;Y_1^n|U^n,V_1^n)\leq H(\zeta)\leq 1$. Thus, the above inequality implies that,

$$I(V_1^n;Y_1^n|U^n,\zeta)-1\leq I(V_1^n;Y_1^n|U^n)\leq I(V_1^n;Y_1^n|U^n,\zeta)+1. \quad \text{(A.12)}$$

Or equivalently,

$$\sum_{j=0}^{1}P(\zeta=j)I(V_1^n;Y_1^n|U^n,\zeta=j)-1\leq I(V_1^n;Y_1^n|U^n)\leq\sum_{j=0}^{1}P(\zeta=j)I(V_1^n;Y_1^n|U^n,\zeta=j)+1. \quad \text{(A.13)}$$

According to the joint typicality property, we know that,

$$0\leq P(\zeta=1)I(V_1^n;Y_1^n|U^n,\zeta=1)\leq nP((u^n,v_1^n,y_1^n)\notin A_\epsilon^{(n)}(P_{U,V_1,Y_1}))\log\|\mathcal{Y}_1\|\leq n\epsilon_{1n}\log\|\mathcal{Y}_1\|. \quad \text{(A.14)}$$

Now consider the term $P(\zeta=0)I(V_1^n;Y_1^n|U^n,\zeta=0)$. Following the sequence joint typicality properties, we have

$$(1-\epsilon_n)I(V_1^n;Y_1^n|U^n,\zeta=0)\leq P(\zeta=0)I(V_1^n;Y_1^n|U^n,\zeta=0)\leq I(V_1^n;Y_1^n|U^n,\zeta=0), \quad \text{(A.15)}$$

where

$$\begin{aligned}I(V_1^n;Y_1^n|U^n,\zeta=0)=&\sum_{(U^n,V_1^n,Y_1^n)\in A_\epsilon^n}P(U^n,V_1^n,Y_1^n)\big(\log P(V_1^n,Y_1^n|U^n)\\&-\log P(V_1^n|U^n)-\log P(Y_1^n|U^n)\big).\end{aligned} \quad \text{(A.16)}$$

Since,

$$\begin{aligned}H(V_1,Y_1|U)-\epsilon_n\leq-\frac{1}{n}\log P(V_1^n,Y_1^n|U^n)\leq H(V_1,Y_1|U)+\epsilon_n \quad \text{(A.17)}\\H(V_1|U)-\epsilon_n\leq-\frac{1}{n}\log P(V_1^n|U^n)\leq H(V_1|U)+\epsilon_n\\H(Y_1|U)-\epsilon_n\leq-\frac{1}{n}\log P(Y_1^n|U^n)\leq H(Y_1|U)+\epsilon_n\end{aligned}$$

then, we have

$$n\left[I(V_1;Y_1|U)-3\epsilon_n\right]\leq I(V_1^n;Y_1^n|U^n,\zeta=0)\leq n\left[I(V_1;Y_1|U)+3\epsilon_n\right]. \quad \text{(A.18)}$$

By substituting (A.18) into (A.15) and then substituting (A.15) and (A.14) into (A.13), we get the desired result:

$$nI(V_1; Y_1|U) - n\delta_{1n} \leq I(V_1^n; Y_1^n|U^n) \leq nI(V_1; Y_1|U) + n\delta_{2n}, \tag{A.19}$$

where,

$$\begin{aligned} \delta 1n &= \epsilon_n I(V_1; Y_1|U) + 3(1-\epsilon_n)\epsilon_n + \frac{1}{n} \\ \delta_{2n} &= 3\epsilon_n + \epsilon_n \log \|\mathcal{Y}_1\| + \frac{1}{n}. \end{aligned} \tag{A.20}$$

□

Similarly, we have the following lemma:

Lemma 12. *Assume $U_{\mathcal{K}}^n$, Y^n and Z^n are generated according to the achievability scheme of Theorem 10. Then, we have*

$$nI(U_{\mathcal{K}}; Y) - n\delta_{1n} \leq I(U_{\mathcal{K}}^n; Y^n) \leq nI(U_{\mathcal{K}}; Y) + n\delta_{2n} \tag{A.21}$$

$$nI(U_{\mathcal{K}}; Z) - n\delta_{3n} \leq I(U_{\mathcal{K}}^n; Z^n) \leq nI(U_{\mathcal{K}}; Z) + n\delta_{4n}, \tag{A.22}$$

where, $\delta_{1n}, \delta_{2n}, \delta_{3n}, \delta_{4n} \to 0$ when $n \to \infty$.

References

[1] C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, vol. 28, pp. 656-715, Oct. 1949.

[2] A. Wyner, "The Wire-tap Channel", *Bell System Technical Journal*, vol. 54, pp. 1355-1387, 1975.

[3] S. K. Leung-Yan-Cheong and M. E. Hellman, "Gaussian Wiretap Channel", *IEEE Trans. Inform. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.

[4] I. Csiszar and J. Korner, "Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339-348, May 1978.

[5] J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels", *in Proc. of ISIT 2006*, pp. 356-360, July 2006.

[6] Y. Liang, H. V. Poor, and S. Shamai, "Secrecy capacity region of fading broadcast channels," *in Proc. ISIT,* (Nice, France), June 2007.

[7] Y. Liang; H.V. Poor and S. Shamai, "Secure Communication over Fading Channels", *IEEE Trans. Inf. Theory*, Volume 54 , Issue 6 pp: 2470 - 2492, 2008.

[8] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," *in Proc. of 44th Annual Allerton Conference,* (Monticello, IL, USA), September 2006.

[9] P. K. Gopala, L. Lai and H. El-Gamal, " On the Secrecy Capacity of Fading Channels", *in IEEE Trans. on Info. Theory*, Volume 54, Issue 10, pp. 4687-4698, Oct. 2008.

[10] F. Oggier, B. Hassibi, " The MIMO Wiretap Channel", *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on.*, pp. 213-218, Mar. 2008.

[11] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO Wiretap Channel", *in Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.

[12] T. Liu and S. Shamai (Shitz), "A Note on the Secrecy Capacity of the Multi-antenna Wiretap Channel", *IEEE Trans. Inf. Theory*, Volume 55, Issue 6, pp.2547 - 2553,June 2009.

[13] A. Khisti and G. Wornell, "Secure Transmission with Multiple Antennas: The MISOME Wiretap Channel", submitted to *IEEE Trans. Inf. Theory.*

[14] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz),"Multiple Input Multiple Output Gaussian Broadcast Channels with Confidential Messages", Submitted to *IEEE Trans. Inform. Theory*, March 2009.

[15] R. Liu, H. V. Poor,"Multi-Antenna Gaussian Broadcast Channels with Confidential Messages", *in Proc IEEE International Symposium on Information Theory*, Toronto, pp.2202 - 2206, July 2008.

[16] R. Liu, H. V. Poor,"Secrecy Capacity Region of a Multiple-Antenna Gaussian Broadcast Channel With Confidential Messages ", *IEEE Trans. Inform. Theory*, Volume 55, Issue 3, pp.1235 - 1249, March 2009.

[17] Y. Oohama, "Coding for Relay Channels with Confidential messages", *in Proc. Of IEEE Information Theory Workshop*, pp. 87-89, Sep. 2001.

[18] Y. Oohama, "Capacity Theorems for Relay Channels with Confidential Messages ", *in Proc. of ISIT 2007*, pp. 926-930, Jun. 2007.

[19] L. Lai and H. El Gamal, "The Relay-Eavesdropper Channel: Cooperation for Secrecy", *IEEE Trans. Inf. Theory*, Volume 54, Issue 9, pp. 4005-4019, Sept. 2008.

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

[21] V. Aggarwal, L. Sankar, A. R. Calderbank, and H. V. Poor, "Information secrecy from multiple eavesdroppers in orthogonal relay channels," *in Proc. IEEE International Symposium on Information Theory,* Seoul, Korea, Jun.-Jul. 2009.

[22] M. Yuksel and E. Erkip, "Secure communication with a relay helping the wiretapper," *in Proc. of IEEE Information Theory Workshop,* Lake Tahoe, CA, September 2007, pp. 595 - 600.

[23] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages," *in Proc. Allerton 124 Conference on Communication, Control, and Computing,* Urbana, IL, Sept. 2629 2006, pp. 305313.

[24] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Multi-terminal communications with confidential messages,", *in Proc. Workshop on Information Theory and Applications,* La Jolla, CA, Jan. 29 Feb. 2 2007, pp. 370 377.

[25] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. on Information Theory,* vol. 54, no. 6, pp. 2493 2507, June 2008.

[26] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Interference-assisted secret communication," *in Proc. of IEEE Information Theory Workshop,* Porto, Portugal, May 59 2008, pp. 405-409.

[27] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "The Gaussian wiretap channel with a helping interferer," *in Proc. of IEEE International Symposium on Information Theory,* Toronto, Ontario, Canada, July 611 2008,

[28] X. He and A. Yener, "Secure Degrees of Freedom for Gaussian Channels with Interference: Structured Codes Outperform Gaussian Signalling," *In IEEE Global Telecommunication Conference,* November 2009.

[29] O.Ozan Koyluoglu, H.El-Gamal, "On the Secure Degrees of Freedom in the K-User Gaussian Interference Channel", *in Proc. of ISIT 2008*, pp. 384-388, Jul. 2008.

[30] S. K. Leung-Yan-Cheong, "Multi-user and wire-tap channels including feedback," Ph.D. dissertation, Stanford University, Stanford, CA, 1976.

[31] R. M. Kahn, "Privacy in multi-user information theory," Ph.D. dissertation, Stanford University, Stanford, CA, 1979.

[32] E. Tekin and A. Yener, "Achievable rates for two-way wire-tap channels," *in IEEE International Symposium on Information Theory,* Nice, France, June 24 29 2007, pp. 1150-1154.

[33] L. Lai, H. E. Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. on Information Theory,* 2007, submitted.

[34] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Multiple access channels with generalized feedback and confidential messages," *in Proc. of IEEE Information Theory Workshop,* Lake Tahoe, CA, September 2007, pp. 608 613.

[35] H. Yamamoto, "Rate-distortion theory for the shannon cipher system," *IEEE Trans. Info. Theory,* vol. 43, no. 3, pp. 827-835, 1997.

[36] A. B. Carleial and M. E. Hellman, "A note on Wyners wiretap channel," *IEEE Trans. on Information Theory,* vol. 23, no. 3, pp. 387390, May 1977.

[37] S. Leung-Yan-Cheong, "On a special class of wiretap channels," *IEEE Trans. on Information Theory,* vol. 23, no. 5, pp. 625627, September 1977.

[38] J. L. Massey, "A simplified treatment of Wyners wire-tap channel," *in Proc. 21st Allerton Conference on Communication, Control and Computing,* Monticello, IL, October 57 1983, pp. 268276.

[39] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Trans. on Information Theory,* vol. 43, no. 2, pp. 712714, March 1997.

[40] J. Xu and B. Chen, "Broadcast Confidential and Public Messages", *in Proc. 42nd Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, pp. 630-635 Mar. 2008.

[41] J. Xu, Y. Cao, and B. Chen ,"Capacity Bounds for Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*,Volume 55, Issue 10, pp.4529 - 4542, Oct. 2009.

[42] A. Khisti, A. Tchamkerten and G. W. Wornell, "Secure Broadcasting", submitted to *IEEE Trans. Inf. Theory.*

[43] E. Tekin, S. Serbetli, and A. Yener, "On secure Signaling for the Gaussian Multiple Access Wire-tap Channel", *in Proc. 2005 Asilomar Conf. On Signals, Systems, and Computers*, Asilomar, CA, pp. 1747-1751, November 2005.

[44] E. Tekin and A. Yener, "The Gaussian Multiple Access Wire-tap Channel", *IEEE Trans. Informa. Theory*, Volume 54, Issue 12, pp: 5747-5755, December 2008.

[45] Y. Liang and H.V. Poor, "Multiple-Access Channels with Confidential Messages", *IEEE Trans.Inform. Theory.*, Volume 54, Issue 3, pp: 976-1002, March 2008.

[46] E. Tekin and A. Yener, "The General Gaussian Multiple Access and Two-Way Wire-Tap Channels: Achievable Rates and Cooperative Jamming", *IEEE Trans. Inform. Theory.*, Volume 54, Issue 6, pp: 2735-2751, June 2008.

[47] X. He and A. Yener, "Cooperation with an Untrusted Relay: A Secrecy Perspective", Submitted to *IEEE Trans. Inform. Theory.*, October 2008.

[48] X. He and A. Yener, "Two-hop Secure Communication Using an Untrusted Relay", Accepted for publication in *EURASIP Journal on Communications and Networks, Special Issue on Wireless Physical Layer Security.*

[49] O. Koyluoglu and H. El-Gamal, "Cooperative binning and channel prefixing for secrecy in interference channels", Submitted to *IEEE Trans. Inform. Theory.*, May 2009.

[50] G. Bagherikaram,A. S. Motahari, A. K. Khandani "Secure broadcasting : The Secrecy Rate Region", *in Proc. of 46th Annual Allerton Conference on Communication, Control, and Computing* , pp.834 - 841, 23-26 Sept. 2008

[51] G. Bagherikaram, A. S. Motahari and A. K. Khandani, "Secrecy Rate Region of the Broadcast Channel with an Eavesdropper", Revised for Publication in *IEEE Trans. Inf. Theory*, September 2009.

[52] G. Bagherikaram, A. S. Motahari and A. K. Khandani, "Secrecy capacity region of Gaussian broadcast channel," *in Proc. 43rd Annual Conference on Information Sciences and Systems,CISS 2009,* 2009 , pp. 152 - 157.

[53] G. Bagherikaram, A. S. Motahari and A. K. Khandani, "The Secrecy Capacity Region of the Degraded Vector Gaussian Broadcast Channel", *in Proc IEEE International Symposium on Information Theory*, South Korea, pp.2772 - 2776, July 2009.

[54] G. Bagherikaram, A. S. Motahari and A. K. Khandani, "The Secrecy Capacity Region of the Gaussian MIMO Broadcast Channel," *IEEE Trans. Inf. Theory*, submitted, October, 2009.

[55] G. Bagherikaram, A. S. Motahari, A. K. Khandani, "On the Secure DoF of the Single-Antenna MAC", Accepted for presentation in *IEEE International Symposium on Information Theory (ISIT)*, ISIT 2010, Austin, Texas, June 2010.

[56] G. Bagherikaram, A. S. Motahari, A. K. Khandani, "On the Secure Degrees-of-Freedom of the Multiple-Access-Channel ", to be submitted to *IEEE Trans. Inf. Theory.*

[57] E. Ekrem, S. Ulukus "Secrecy Capacity of a Class of Broadcast Channels with an Eavesdropper", submitted to *EURASIP Journal on Wireless Communications and Networking* , Dec. 2008.

[58] E. Ekrem, S. Ulukus "Secrecy Capacity Region of the Gaussian Multi-Receiver Wiretap Channel", *in Proc. of IEEE International Symposium on Information Theory (ISIT)*, pp.2612 - 2616, June 28 2009-July 3 2009

[59] E. Ekrem and S.Ulukus, "The Secrecy Capacity Region of the Gaussian MIMO Multi-Receiver Wiretap Channel", Submitted to *IEEE Trans. Inform. Theory*, March 2009.

[60] R. Liu, T. Liu, H. V. Poor, S. Shamai(Shitz), "A Vector Generalization of Costa Entropy Power Inequality and Applications", Submitted to *IEEE Trans. on Inf. Theory*, Mar. 2009.

[61] S. I. Gelfand and M. S. Pinsker, "Coding for Channel with Random Parameters", *Problemy Peredachi Informatsii*, vol. 9, no. 1, pp. 19-31, 1980.

[62] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley Sons, Inc., 1991.

[63] C. E. Shannon, "A mathematical Theory of Communication", *Bell Syst. Tech. J.*, vol. 27, pp. 379423 and 623656, Jul. and Oct. 1948

[64] A. J. Stam, " Some Inequalities Satisfied by the Quantities of Information of Fisher and Shannon " *Information and Control*, vol.2 pp. 101-112, Jun. 1959

[65] S. Shamai, A. Steiner, "A Broadcast Approach for a Single-User Slowly Fading MIMO Channel", *in IEEE Trans. on Info. Theory*, Volume 49, Issue 10, pp. 2617-2635, Oct. 2003.

[66] S. Shamai, "A Broadcast Strategy for the Gaussian Slowly Fading Channel", *in IEEE International Symposium on Info. Theory*, pp. 150, July 1997.

[67] A. Steiner, S. Shamai, "Single-User Broadcasting Protocols Over a Two-Hop Relay Fading Channel", *in IEEE Trans. on Info. Theory*, Volume 52, Issue 11, pp. 4821-4838, Nov. 2006.

[68] A. Steiner, S. Shamai, "Multi-Layer Broadcasting Hybrid-ARQ Strategies for Block Fading Channels", *in IEEE Trans. on Wireless Communications*, Volume 7, Issue 7, pp. 2640-2650, July 2008.

[69] K. Marton, "A Coding Theorem for the Discrete Memoryless Broadcast Channel", *IEEE Trans. on Inf. Theory*, vol. 25, no. 1, pp. 306-311, May 1979.

[70] S. N. Digagvi, T. M. Cover, " The Worst Additive Noise Under a Covariance Constraint", *IEEE Trans. on Info. Theory*, Vol. 47, No. 7, pp. 3072-3081, Nov. 2001.

[71] T. Liu, P. Viswanath, "An Extremal Inequality Motivated by Multiterminal Information Theoretic Problems", *IEEE Trans. on Inf. Theory*, vol. 53, no. 5, pp. 1839-1851, May 2007.

[72] M. H. M. Costa, " A New Entropy Power Inequality",*IEEE Trans. on Inf. Theory*, vol. 31, o. 6 pp. 751-760, Nov. 1985.

[73] R. Liu, T. Liu, H. V. Poor, S. Shamai(Shitz), "A Vector Generalization of Costa Entropy Power Inequality and Applications", Submitted to *IEEE Trans. on Inf. Theory*, Mar. 2009.

[74] H. Weingarten, Y. Steinberg, S. Shamai(Shitz), "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel", *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936-3964, September 2006.

[75] P. P. Bergmans, "A Simple Converse for Broadcast Channels with Additive White Gaussian Noise", *IEEE Trans. Inform. Theory*, vol. IT-20, no. 2, pp. 279-280, March 1974.

[76] W. Yu, and J.M. Ciofi, "Sum Capacity of Gaussian Vector Broadcast Cahnnels", *IEEE Trans. on Inf. Theory*, vol. 50, pp. 1875-1893, September 2004.

[77] G. Strang, *Linear Algebra and Its Applications.* Wellesley, MA: Wellesley-Cambridge Press, 1998.

[78] X. He and A. Yener, "Providing Secrecy With Structured Codes: Tools and Applications to Two-User Gaussian Channels", submitted to *IEEE Trans. Inform. Theory*, Jully 2009.

[79] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis, *IEEE Trans. Inf. Theory,* vol. 54, no. 8, pp. 3457-3470, August 2008 (also see earlier technical reports by the same authors referenced therein).

[80] A. S. Motahari,S. O. Gharan, and A. K. Khandani, "Real Interference Alignment with Real Numbers", Submitted to *IEEE Trans. Inf. Theory.*

[81] A. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real Interference Alignment: Exploiting the Potential of Single Antenna Systems", Submitted to *IEEE Trans. Inf. Theory*, November 2009.

[82] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, P. Viswanath, "The Capacity Region of the Degraded MIMO Compound Broadcast Channel", in Proc. *IEEE International Symposium on Information Theory (ISIT)*, pp: 766 - 770, June 2007.

[83] W. M. Schmidt, *Diophantine approximation*. Berlin, Springer-Verlag, 1980.

[84] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers,* fifth edition, Oxford science publications, 2003.

[85] L. H. Ozarow and A. D. Wyner, "Wire-Tap Channel II," B.S.T.J., vol. 63, no. 10, pp. 2135-2157, Dec. 1984.

[86] A. Thangaraj, S. Dihidar,A.R. Calderbank, S.W. McLaughlin, J-M. Merolla, "Applications of LDPC Codes to the Wiretap Channel," *IEEE Trans. on Inform. Theory,*Volume: 53 , Issue: 8 pp. 2933 - 2945, 2007.

[87] G. Cohen, G. Zemor, "Generalized coset schemes for the wire-tap channel: application to biometrics," in Proc *International Symposium on Information Theory,* pp.45, 2004

[88] D. Klinc, Ha. Jeongseok, S.W. McLaughlin,J. Barros, Byung-Jae Kwak, "LDPC codes for the Gaussian wiretap channel,", in Proc *Information Theory Workshop (ITW)*, PP. 95 - 99, 2009.

[89] M. Bloch, A. Thangaraj, S.W. McLaughlin,J.-M. Merolla,"LDPC-based secret key agreement over the Gaussian wiretap channel," in Proc *IEEE International Symposium Information Theory,* pp. 1179- 1183, 2006.

[90] L. Zheng and D. Tse," Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. on Inform. Theory,* Vol. 49, Issue 5, pp.1073-1096, 2003.

[91] Hero, A.O., III; "Secure space-time communication," *IEEE Trans. on Inform. Theory,* Volume: 49 , Issue: 12 pp. 3235 - 3249, 2003.

[92] Yuu-Seng Lau; Lin, K.H.; Hussain, Z.M.;" Space-Time Encoded Secure Chaos Communications with Transmit Beamforming," in *IEEE Region 10 TENCON 2005,* pp. 1 - 5, 2005.

[93] R. Etkin, D. Tse, and H. Wang. "Gaussian interference channel capacity to within one bit," *IEEE Trans. on Info. Theory,* Vol. 54, Issue 12, pp. 5534 -5562, 2008.

[94] A.S. Motahari, A.K. Khandani, "Capacity Bounds for the Gaussian Interference Channel," *IEEE Transactions on Information Theory,* Volume: 55 , Issue: 2, pp. 620 - 643, 2009.

[95] M.A. Maddah-Ali, A.S. Motahari, A.K. Khandani, "Decomposition of the MIMO X Channels," *Canadian Workshop on Information Theory, CWIT,* pp. 110 - 113, 2007.

[96] Cadambe, V.R.; Jafar, S.A.; "Interference alignment via random codes and he capacity of a class of deterministic interference channels,", in Proc *47th Annual Allerton Conference on Communication, Control, and Computing,* pp. 67 - 74, 2009.

[97] O.O. Koyluoglu, H. El Gamal, "On the secrecy rate region for the interference channel," in proc *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications,* pp. 1- 5,PIMRC 2008.

[98] Tie Liu, V. Prabhakaran, S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in Proc *IEEE International Symposium on Information Theory,* pp. 116-120, ISIT 2008.

[99] Yingbin Liang, G. Kramer, H.V. Poor, S. Shamai, "Recent results on compound wiretap channels," in Proc *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications,* pp.1-5 ,PIMRC 2008.

[100] M. Kobayashi, Yingbin Liang, S. Shamai, M. Debbah, "On the compound MIMO broadcast channels with confidential messages," in Proc *IEEE International Symposium on Information Theory,* pp.1283 - 1287, ISIT 2009.