Fundamental Limits of Gaussian Communication Networks in the Presence of Intelligent Jammers

by

Fatemeh Hosseinigoki

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2019 by the
Graduate Supervisory Committee:

Oliver Kosut, Chair
Junshan Zhang
Lalitha Sankar
Daniel Bliss

ARIZONA STATE UNIVERSITY

December 2019

ABSTRACT

The open nature of the wireless communication medium makes it inherently vulnerable to an active attack, wherein a malicious adversary (or jammer) transmits into the medium to disrupt the operation of the legitimate users. Therefore, developing techniques to manage the presence of a jammer and to characterize the effect of an attacker on the fundamental limits of wireless communication networks is important. This dissertation studies various Gaussian communication networks in the presence of such an adversarial jammer.

First of all, a standard Gaussian channel is considered in the presence of a jammer, known as a Gaussian arbitrarily-varying channel, but with list-decoding at the receiver. The receiver decodes a list of messages, instead of only one message, with the goal of the correct message being an element of the list. The capacity is characterized, and it is shown that under some transmitter's power constraints the adversary is able to suspend the communication between the legitimate users and make the capacity zero.

Next, generalized packing lemmas are introduced for Gaussian adversarial channels to achieve the capacity bounds for three Gaussian multi-user channels in the presence of adversarial jammers. Inner and outer bounds on the capacity regions of Gaussian multiple-access channels, Gaussian broadcast channels, and Gaussian interference channels are derived in the presence of malicious jammers. For the Gaussian multiple-access channels with jammer, the capacity bounds coincide. In this dissertation, the adversaries can send any arbitrary signals to the channel while none of the transmitter and the receiver knows the adversarial signals' distribution.

Finally, the capacity of the standard point-to-point Gaussian fading channel in the presence of one jammer is investigated under multiple scenarios of channel state information availability, which is the knowledge of exact fading coefficients. The

channel state information is always partially or fully known at the receiver to decode the message while the transmitter or the adversary may or may not have access to this information. Here, the adversary model is the same as the previous cases with no knowledge about the user's transmitted signal except possibly the knowledge of the fading path.

DEDICATION

To my beloved parents, Batoul and Abbas, my dear spouse, Mohammad,

and my dearest siblings, Narjes and Hossein,

for their love, unlimited support, encouragement and sacrifices.

ACKNOWLEDGMENTS

In the Name of Allah, the Most Merciful, the Most Compassionate. “All the praises and thanks be to Allah, who guided us to this, and in no way we could have been guided, unless Allah had guided us” (Quran surah Al A’raf 43).

I would like to thank my advisor, Professor Oliver Kosut for his advice and generous support through this challenging process. This work would have never become truth without his guidance. He was always so patient, helpful and encouraging. I owe him a debt of gratitude for teaching me how to think correctly, solve a problem and present my work.

I would like to thank Professor Junshan Zhang, Professor Lalitha Sankar, and Professor Daniel Bliss for accepting to be a member of my dissertation committee. I also thank Professor Antonia Papandreou-Suppappola, Professor Cihan Tepedelenlioglu and Professor Lei Ying for teaching me valuable courses in a way that I enjoyed learning.

Most importantly, I would like to thank my beloved parents, Batoul and Abbas, for their endless love and support which encourage me during this journey. I am grateful to my dear spouse, Mohammad, for his love and being a wonderful companion who helped me to accomplish my degree. I thank my dearest siblings Narjes and Hossein for always understanding me, giving me the best advice and being my best friends in my whole life.

I also thank my great friends at ASU, Tahora H. Nazer, Mansoure Moeini Rizi, Faezeh Kalantari, Mansooreh Karami, Jiachun Liao, and Allison Beemer.

This material is based upon work supported by the National Science Foundation under Grant No. CCF$-$1453718.

TABLE OF CONTENTS

Page

LIST OF TABLES ... viii

LIST OF FIGURES ... ix

CHAPTER

1 INTRODUCTION ... 1

1.1 Motivation and Overview ... 1

1.2 Gaussian Arbitrarily-Varying Channels ... 4

1.2.1 Prior Work ... 4

1.2.2 Contribution ... 8

1.3 Gaussian Arbitrarily-Varying Multiple-User Channels ... 11

1.3.1 Prior Work ... 11

1.3.2 Contribution ... 15

1.4 Gaussian Arbitrarily-Varying Fading Channels ... 20

1.4.1 Prior Work ... 20

1.4.2 Contribution ... 21

1.5 Organization ... 23

1.6 Notation ... 24

2 GAUSSIAN ARBITRARILY-VARYING CHANNEL WITH LIST-DECODING ... 25

2.1 Problem Statement ... 25

2.2 Main Results ... 27

2.3 Converse Proof ... 28

2.4 Achievability Proof ... 33

2.5 Proof of Lemma 4 ... 41

3 PACKING LEMMAS FOR GAUSSIAN ADVERSARIAL CHANNELS . 44
3.1 Packing Lemmas . . . 44
3.2 Proof of Lemma 5 . . . 50
3.3 Proof of Lemma 6 . . . 53
3.4 Proof of Lemma 7 . . . 56
3.5 Proof of Lemma 8 . . . 62
3.6 Proof of Lemma 9 . . . 67
4 GAUSSIAN ARBITRARILY-VARYING MULTIPLE ACCESS CHANNEL . . . 82
4.1 Problem Statement . . . 82
4.2 Main Results . . . 84
4.3 Converse Proof . . . 84
4.4 Achievability Proof . . . 85
5 GAUSSIAN ARBITRARILY-VARYING BROADCAST CHANNEL . . . 88
5.1 Problem Statement . . . 88
5.2 Main Results . . . 90
5.3 Proof of Outer Bound . . . 90
5.4 Proof of Inner Bound . . . 91
6 GAUSSIAN ARBITRARILY-VARYING INTERFERENCE CHANNEL 96
6.1 Problem Statement . . . 96
6.1.1 Generalized Jamming Model . . . 98
6.2 Main Results . . . 99
6.3 Discussion and Numerical Results . . . 101
6.4 Proof of Outer Bound . . . 106

6.5 Proof of Inner Bound 107
6.6 Proof of Proposition 14 112
7 GAUSSIAN ARBITRARILY-VARYING FADING CHANNELS 114
7.1 Problem Statement 115
7.2 Main Results 117
7.3 Auxiliary Results and Tools 120
7.4 Capacity Proof with Gains Available at Decoder 122
7.4.1 Converse Proof 122
7.4.2 Achievability Proof 123
7.5 Capacity Proof with Gains Available at Encoder and Decoder 123
7.5.1 Converse Proof 123
7.5.2 Achievability Proof 124
7.6 Capacity Proof with Gains Available at Decoder and Jammer 130
7.6.1 Converse Proof 130
7.6.2 Achievability Proof 132
7.7 Capacity Proof with Gains Available at Encoder, Decoder, and Jammer 133
7.7.1 Converse Proof 133
7.7.2 Achievability Proof (Gains Available Non-causally at Adversary and Causally at Encoder) 137
7.7.3 Achievability Proof (Gains Available Causally at Adversary and Non-causally at Encoder) 147
7.8 Proof of Lemmas 18 and 19 154
REFERENCES 156

LIST OF TABLES

Table Page

1. Summary of Literatures on the Point-To-Point AVC 9

LIST OF FIGURES

Figure Page

1. Gaussian Arbitrarily-Varying Channel 26

2. Two-User Gaussian Multiple Access Channel with a Jammer............... 83

3. Two-User Gaussian Broadcast Channel with Two Jammers. 89

4. Two-User Gaussian Interference Channel with Two Independent Jammers. . 97

5. Two-User Gaussian Interference Channel with G Independent Jammers..... 98

6. Bounds on the Symmetric Capacity $\mathscr{C}_{\text{sym}}(S, I, J)$ for $S = 4$, $I = 3$, and J between 0 and 5. 101

7. Bounds on the Symmetric Capacity $\mathscr{C}_{\text{sym}}(S, I, J)$ for Fixed $S_1 = S_2 = S$, $J_1 = J_2 = J$, and an Interval $I_1 = I_2 = I$ 102

8. Symmetric Degrees of Freedom for the GIC..............................105

9. Gaussian Arbitrarily-Varying Fading Channel............................115

10. Gaussian Arbitrarily-Varying Fading Channel Capacities for $P = 1, \sigma^2 = 0.25, 0 < \Lambda < 10$ with Rayleigh Fading. 120

Chapter 1

INTRODUCTION

1.1 Motivation and Overview

As wireless communications play an important role in upcoming technologies, the need for reliable and fast communications becomes more and more vital. Generally, in the near future, equipped machines with robust wireless communication technologies will control our cities and societies in a vast scale. There are many important applications using 5G cellular network technology such as self-driving cars, industry automation, mission critical applications, and smart cities that need invulnerable and robust data transfer in addition to fast communication. More network applications such as distributed machine learning also require high speed and reliable communications to reduce their intensive computations during the learning and training processes. However, the inherent feature of the wireless communication as an open environment to any unwelcome attacker could jeopardize robustness, reliability, speed and privacy of the communication network that may cause serious irreversible damages to the aforementioned applications.

These uninvited attackers in a wireless network can be passive or active. A passive attacker is an adversary who only listens and eavesdrops on the communications between the legitimate users. Channels with an eavesdropper are usually known as *wiretap* channels in which the main concern is providing privacy for the legitimate users. On the other hand, an active attacker is an adversary who maliciously transmits into the medium to disrupt the operation of the legitimate users or interrupt the ongoing communications between them, which seriously reduces reliability and speed of the communication channel. These active adversaries or jammers may have access to some information about the wireless network such as power constraints, channel fading information, transmitted code, and transmitted signals perfectly or through a noisy channel. Based on the jammers' knowledge, they are assigned to different categories.

We only provide four categories for jammers based on their knowledge about the legitimate users' transmitted code and the genuine transmitted signal. Omniscient adversaries are active attackers who have access to the genuine transmitted signal before selecting their own adversarial signal, so they benefit from this knowledge to disrupt the communication. Causal adversaries are active attackers with knowledge only of the past transmitted signal. Myopic adversaries refer to active attackers whose knowledge about the transmitted signal is through a noisy memoryless channel, so they have a noisy version of the genuine transmitter signal. Oblivious adversaries are those active attackers who only know the legitimate users' transmitted code but not the exact value of the transmitted signal. The adversary model that we assume throughout this dissertation is contained in the last category as *oblivious active adversaries*.

Finally, in order to benefit the upcoming technologies using fast and reliable wireless communications, we need to develop techniques to manage and control the presence of various active adversaries in the medium. Despite that, it is apparently very important to first characterize the effect of a jammer on the fundamental limits of wireless communication networks as a principal and prior task to achieve reliability, speed and robustness. Indeed, one initially needs to know how much the theoretical framework and boundaries allows them to construct and evolve new methods to combat the adversarial jammers. The hope is that fundamental research in this field will become a paradigm for secure and private wireless communication networks. One cannot solve all practical problems related to jammers in wireless networks at once; therefore, this dissertation focuses on a few well-defined problems on which we can make progress. In particular, we study the fundamental limits of Gaussian communication networks as one of the common wireless networks in the presence of intelligent active adversaries in this dissertation. The following sections introduce some prior work in the literature along with our contribution in deriving the fundamental limits of the intended Gaussian channels in the presence of oblivious jammers.

We investigate boundaries for the capacity region of five different Gaussian communication channels in the presence of oblivious adversaries. In this regard, we first consider a simple point-to-point standard Gaussian channel in the presence of a jammer. This channel is also known as a Gaussian arbitrarily-varying channel (AVC). Then, we introduce general adversarial packing lemmas in order to achieve the inner bounds for the capacities in three scenarios. We use our proposed adversarial packing lemmas to prove the capacity limits in Gaussian multi-user scenarios such as Gaussian multiple access channels, Gaussian broadcast channels and Gaussian interference channels all in the presence of jammers. Finally, we also explore the capacity of a point-to-point Gaussian fading channel in the presence of a jammer. This channel itself includes several cases based on the availability of channel fading information at the adversary or transmitter by which the capacity may vary.

1.2 Gaussian Arbitrarily-Varying Channels

1.2.1 Prior Work

An arbitrarily-varying channel (AVC) represents a memoryless channel including unknown parameters that are changing arbitrarily from channel use to channel use. Because these parameters (state) can change arbitrarily, we consider this to be a model for an active adversarial jammer. This adversary sends its signal to restrain the legitimate transmitter and receiver from maintaining reliable communication. In wireless channels, these unpleasant adversaries can easily enter channels, so it is of great importance to study the adversary's effect on the channel capacity. The capacity of the AVC depends on the coding method (such as randomized coding, stochastic encoder or deterministic coding), the performance criterion (such as average or maximum error probability) and the amount of adversary's knowledge about the transmitted signal (omniscient, myopic or oblivious adversary). Table 1 provides a summary of various models for point-to-point channels in the presence of adversaries appearing in the literature, including those considered in Chapter 2.

Blackwell et al. introduced AVC in Blackwell, Breiman, and Thomasian, (1960) and under the average error probability criterion they found the capacity of the discrete memoryless AVC to be given by a min-max expression over the mutual information of input and output. They employed *randomized coding* that is, common randomness between the encoder and the decoder and assumed the jammer to be *oblivious* that is, the jammer does not have any information about the transmitted signal except the code. In Stiglitz, (1966), it is shown that the min-max expression is equivalent to the corresponding max-min one. Further, in Ahlswede and Wolfowitz, (1969), the authors examined that this capacity remains the same even for the maximum error probability criterion, again provided access to common randomness. The case without common randomness between the transmitter and the receiver is referred to as the deterministic code setting. Ahlswede in a notable paper Ahlswede, (1978) characterized the deterministic capacity of a discrete AVC under the average probability of error through a dichotomy theorem. He proved that the capacity either corresponds to the AVC randomized code capacity or else it equals zero but he did not state any necessary or sufficient condition for which of the two cases prevails. Ericson, in Ericson, (1985), found the necessary condition for the positive alternative by defining *symmetrizability*. A symmetrizable AVC is an AVC in which the adversary can mimic the transmitted signal in order to prevent the decoder from distinguishing between the true message and an adversarial imitation. Thus, he showed that if the deterministic code capacity of an AVC is positive then the channel should be nonsymmetrizable. Later, in Csiszár and Narayan, (1988)(b), a sufficient condition was provided by Csiszár and Narayen stating that if the AVC is nonsymmetrizable then the deterministic code capacity is not zero. Therefore, considering both conditions, the deterministic code capacity of an AVC is positive if and only if the channel is nonsymmetrizable.

The capacity of discrete AVC is investigated also under input and state (or adversarial signal) constraints. Restricted by peak or average input and state cost constraints, the random code capacity of discrete AVC is studied in Csiszár and Narayan, (1988)(a) using the average probability of error as the performance criterion. Furthermore, the second part of Csiszár and Narayan work in Csiszár and Narayan, (1988)(b) focuses on the deterministic code capacity of AVC under input and state constraints for the same performance criterion. They proved that in this case if the capacity is positive then it is less than or equal to the corresponding random code capacity. In particular, with input and state constraints, the capacity can be positive but strictly less than the random code capacity. Note that this does not occur without cost constraints. Csiszár, in Csiszár, (1992), extended this result to general input and output alphabets and state sets rather than only finite alphabets and state sets.

There is a wide variety of research on different versions of AVCs under various adversaries model, including Sarwate, (2010); Dey et al., (2010), (2013). Sarwate, in Sarwate, (2010), considered a myopic adversary in which there is a discrete memoryless channel (DMC) between the legitimate user and the jammer and the jammer chooses its signal based on this noisy version of the user's codeword. He found the capacity by minimizing over all DMCs that the jammer can applied by its worst strategies. In Dey et al., (2010), single letter characterizations of capacity is obtained in the presence of a delayed adversary which can observe the transmitted signal after a delay. By assuming randomization at the encoder, the capacity is corresponding to the randomized code capacity. B. K. Dey et. al., in Dey et al., (2013), obtained upper bounds on the capacity of binary channel in the presence of a causal adversary for both maximal and average error probabilities.

This dissertation focuses on the Gaussian AVC, wherein all alphabets are continuous rather than discrete. Initially, Ahlswede, in Ahlswede, (1971), studied the capacity of a Gaussian AVC in which the adversary chooses the noise variance rather than an additive signal. Hughes and Narayan in Hughes and Narayan, (1987) determined the randomized code capacity of Gaussian AVC under the peak power input and state constraints. They further extended their result in Hughes and Narayan, (1988) for a vector Gaussian AVC. The deterministic code capacity of the Gaussian AVC, for the average probability of error, was found in Csiszár and Narayan, (1991). The authors showed that if the adversary's power is greater than the legitimate transmitter's power, then symmetrizability occurs causing the capacity to drop to zero. Note that for a discrete AVC with no cost constraint non-symmetrizability makes the deterministic capacity positive and equal to the randomized capacity Csiszár and Narayan, (1988)(b) (Theorem 1). It is further proved in Csiszár and Narayan, (1988)(b) (Theorem 3) that under input and state cost constraints, non-symmetrizability only results in positive deterministic capacity but it is sometimes strictly less than the randomized capacity. In the Gaussian case, even though there are input and state cost constraints, the behavior is like that of a discrete AVC with no cost constraint, in that if the channel is non-symmetrizable, then its deterministic capacity is positive and equal to the randomized capacity Csiszár and Narayan, (1991).

For the first time, in Hughes, (1997), Hughes showed that using list-decoding, in which the decoder can decode to a small (and bounded) list rather than a unique message estimate, causes positive capacity for most symmetrizable discrete-memoryless AVCs. Intuitively, list-decoding combats the symmetrizing attack by allowing the list to contain the true message as well as the counterfeit(s) generated by the adversary; thus, the receiver can successfully decode even if it cannot specify the correct message. Furthermore, the authors in Sarwate and Gastpar, (2012) extended the list-decoding result to the discrete-memoryless AVCs with state constraints. They determined upper and lower bounds on the capacity by introducing two notions of symmetrizability for this channel.

1.2.2 Contribution

In Chapter 2, we characterize the capacity of Gaussian AVC in Csiszár and Narayan, (1991), using list-decoding for any list size and almost all power values of the transmitter and adversary, a similar result to that of Hughes in Hughes, (1997) which obtained the list capacity for the discrete-memoryless AVC, for which the capacity was determined in Csiszár and Narayan, (1988)(b). We assume that the encoder may be stochastic— that is, the encoder has access to private randomness—and a deterministic list-decoder with constant list size L. Under the average probability of error criterion and without common randomness, we obtain the capacity of Gaussian AVC with list-decoding to be equal to the corresponding randomized code capacity if the list size is greater than the power ratio of the jammer to the legitimate user; otherwise, the capacity is zero. Generally, our problem is a generalized version of the multiple packing of spherical caps problem in Blachman and Few, (1963) with Gaussian noise; although, they assumed maximal probability of error as the performance criterion. Their upper bound, which is only calculated for the noiseless case, depends on the list size L even in the asymptotic case. However, we only have list size in our symmetrizability conditions rather than the capacity itself.

In our converse proof (in Section 2.3), the adversary focuses on two possible strategies, one of which is simply sending Gaussian noise which causes the channel to act as a standard Gaussian channel with increased noise variance. The second strategy for the adversary is to transmit the superposition of some random (counterfeit) codewords, which is shown to be possible with positive probability if its power is large enough. In our achievability proof (in Section 2.4), we employ Gaussian codewords with a particular version of minimum distance list-decoding based on typicality. We

Table 1: Summary of literatures on the point-to-point AVC

Reference	**Discrete Gaussian**	**Shared Randomness**	**Cost Constraints**	**Max Avg**	**List Decoding**	**Adversarial Model**
Blackwell, Breiman, and Thomasian, (1960)	Discrete	✓	✗	Avg	✗	Oblivious
Ahlswede and Wolfowitz, (1969)	Discrete	✓	✗	Max, Avg	✗	Oblivious
Ahlswede, (1978)	Discrete	✗	✗	Max, Avg	✗	Oblivious
Ericson, (1985)	Discrete	✓	✗	Avg	✗	Oblivious
Csiszár and Narayan, (1988)(b)	Discrete	✗	✓	Avg	✗	Oblivious
Csiszár and Narayan, (1988)(b)	Discrete	✗	✗	Avg	✗	Oblivious
Csiszár and Narayan, (1988)(a)	Discrete	✓	✓	Avg	✗	Oblivious
Csiszár, (1992)	Discrete	✗	✓	Avg	✗	Oblivious
Sarwate, (2010)	Discrete	✓	✗	Max	✗	Myopic
Dey et al., (2010)	Discrete	✓	✓	Max	✗	Delay
Dey et al., (2013)	Discrete	✗	✓	Max, Avg	✗	Causal
Hughes, (1997)	Discrete	✗	✗	Avg	✓	Oblivious
Sarwate and Gastpar, (2012)	Discrete	✓	✓	Max	✓	Oblivious
Sarwate and Gastpar, (2012)	Discrete	✗	✓	Avg	✓	Oblivious
Hughes and Narayan, (1987)	Gaussian	✓	✓	Max	✗	Oblivious
Csiszár and Narayan, (1991)	Gaussian	✗	✓	Avg	✗	Oblivious
This dissertation, Chapter 2	Gaussian	✗	✓	Avg	✓	Oblivious

Table 1 Continued: Summary of literatures on the point-to-point AVC

Reference	**Capacity** **or Notes**
Blackwell, Breiman, and Thomasian, (1960)	Blackwell et al. introduced AVC.
Ahlswede and Wolfowitz, (1969)	Capacity remains the same for Max. error Probability.
Ahlswede, (1978)	Studied deterministic capacity in a dichotomy theorem.
Ericson, (1985)	Found the necessary condition for the positive deterministic capacity by defining symmetrizability.
Csiszár and Narayan, (1988)(b)	$\begin{cases} \max\limits_{\substack{P:g(P)\leq\Gamma \\ \Lambda_0(P)\geq\Lambda}} \min\limits_{\substack{Y:P_{YXS}\in\mathscr{C}_0 \text{ for} \\ \text{some } S, \text{ with } P(X)=P}} I(X;Y), & \max\limits_{P:g(P)\leq\Gamma} \Lambda_0(P) > \Lambda \\ 0, & \max\limits_{P:g(P)\leq\Gamma} \Lambda_0(P) < \Lambda. \end{cases}$
Csiszár and Narayan, (1988)(b)	$\max\limits_{P} \min\limits_{\substack{Y:P_{YXS}\in\mathscr{C}_0 \text{ for} \\ \text{some } S, \text{ with } P(X)=P}} I(X;Y)$ if and only if $C > 0$
Csiszár and Narayan, (1988)(a)	$\max\limits_{X:\mathbb{E}g(X)\leq\Gamma} \min\limits_{S:\mathbb{E}S\leq\Lambda} I(X;Y_{X,S})$
Csiszár, (1992)	Capacity with general alphabets and states.
Sarwate, (2010)	$\max\limits_{P(x)} \min\limits_{V\in W} I(P,V)$ W is a set of transition matrices from X to Y
Dey et al., (2010)	$\max\limits_{P(x)} \min\limits_{Q\in\mathcal{Q}} I(P,W_Q)$ $\mathcal{Q}$ is a set of transition matrices from X to Y
Dey et al., (2013)	Upper bounds on the capacity
Hughes, (1997)	$\begin{cases} \max\limits_{P} \min\limits_{\substack{P_{YXS}:P_{YXS}\in\mathscr{C}_0 \text{ for} \\ \text{some } S, \text{ with } P(X)=P}} I(X;Y), & L > M \\ 0, & L \leq M. \end{cases}$
Sarwate and Gastpar, (2012)	$\max\limits_{P\in\mathcal{P}(\mathcal{X})} \min\limits_{U\in\mathcal{U}(P,\Lambda)} I(P,\sum W(y,x,s)U(s\mid x))$
Sarwate and Gastpar, (2012)	Upper and lower bounds on the capacity
Hughes and Narayan, (1987)	$C\left(\frac{P}{\Lambda+\sigma^2}\right)$
Csiszár and Narayan, (1991)	$\begin{cases} C\left(\frac{P}{\Lambda+\sigma^2}\right), & P > \Lambda \\ 0, & P \leq \Lambda. \end{cases}$
This dissertation, Chapter 2	$\begin{cases} C\left(\frac{P}{\Lambda+\sigma^2}\right), & L > \frac{\Lambda}{P} \\ 0, & L < \frac{\Lambda}{P}. \end{cases}$

extend the scheme of Csiszár and Narayan, (1991) to show that with high probability a random Gaussian codebook has desirable properties to make the probability of error zero. However, our achievability proof originates from the idea of Csiszár and Narayan, (1988)(b) based on typical sets, rather than the geometric approach of Csiszár and Narayan, (1991). This scheme allows us for simpler characterizations of codebook constraints. It is worth mentioning that we prove the achievability for the deterministic encoder since it suffices to achieve a rate even by a deterministic code, that is any deterministic code is a realistic value of a stochastic code. Our converse and achievability proof in Chapter 2 apply for any list size; our work Hosseinigoki and Kosut, (2018) provided proof only for list size $L = 2$. We published our results for the capacity of Gaussian AVC with list-decoding in 2018 IEEE International Symposium on Information Theory (ISIT) as Hosseinigoki and Kosut, (2018) and in Entropy Journal as Hosseinigoki and Kosut, (2019)(b).

1.3 Gaussian Arbitrarily-Varying Multiple-User Channels

1.3.1 Prior Work

Discrete AVCs are also studied in network settings throughout Jahn, (1981); Gubner, (1990), (1992); Ahlswede and Cai, (1999); Hof and Bross, (2006); Winshtok and Steinberg, (2006); Pereg and Steinberg, (2017), such as multiple-access and broadcast channels. Jahn, (1981) is the first study to determine the capacity region of arbitrarily-varying multiple-access channels (AVMACs) under the average probability of error criterion. Later, Gubner in Gubner, (1990) derives the sufficient condition of non-symmetrizablilty for the AVMAC to have non-empty deterministic-code capacity region, i.e. the two-user in AVMAC can have a reliable communication. He also provides three various symmetrizability conditions to get the aforementioned results. Two years later, in Gubner, (1992), he establishes the necessary and sufficient conditions for the deterministic-code capacity region of the AVMAC to be non-empty under a state constraint and average probability of error criterion. This capacity region is further specified one more time in Ahlswede and Cai, (1999) to eliminate some constraints on the former results by providing three non-symmetrizability conditions.

Moreover, a two-user discrete memoryless arbitrarily-varying broadcast channels (AVBCs) is addressed in Jahn, (1981), and an inner bound on the random-code capacity region is established. Jahn also prove that the deterministic-code capacity of the AVBC is equal to the random-code capacity of the AVBC if the interior region of the random-code capacity is not empty. However, he doe not specify when the random-code capacity is positive. In Hof and Bross, (2006), by defining different symmetrizable channels for two-user AVBC, a sufficient non-symmetrizable condition that makes the random-code capacity positive is attained under state and input constraints, i.e. if the AVBC is non-symmetrizable, then the random-code capacity is non-zero. Further, the capacity region for memoryless arbitrarily-varying degraded broadcast channels (AVDBCs) is derived in Winshtok and Steinberg, (2006) in which the transmitter knows channel state information (CSI) non-causally, and the stronger receiver has full access to the CSI. If the transmitter knows the channel state information causally, then inner and outer bounds are characterized for the deterministic-code capacity region and the random-code capacity region of AVDBCs in Pereg and Steinberg, (2017). The authors also find the conditions that lead the inner and outer bounds to a closed form capacity region formula.

The capacity of AVMAC was also studied with list-decoding in Nitinawarat, (2013); Boche and Schaefer, (2014). Sirin Nitinawarat in Nitinawarat, (2013) introduced symmetrizability of an AVMAC and showed that the capacity region for deterministic codes with fixed list-size is empty if the list size is less than the symmetrizability Ω. He obtained that the capacity corresponds to the random code capacity if the list size is greater than $(1+\Omega)^2$. H. Boche and R.F. Schaefer in Boche and Schaefer, (2014) obtained the list capacity region of AVMAC with conferencing encoders which is proved for large list size to be equal to the common randomness assisted capacity region. Moreover, in Schaefer and Boche, (2014), the authors found the deterministic code and random code capacity regions of AVBCs with receiver side information. By defining a concept of symmetrizability for the channel, they characterized deterministic list codes capacity region as either identical to the random code capacity region or empty. Note that these literatures study the discrete versions of AVC while in this dissertation our results are provided for Gaussian versions of AVC. Also, we consider list-decoding only for Gaussian AVC, but not for the network setting of AVC.

Considering the Gaussian networks without any adversary, the exact capacity regions of the Gaussian multiple-access channel (GMAC) and the Gaussian broadcast channel (GBC) are completely characterized. The GMAC consists of two transmitters and one receiver with additive Gaussian noise while the GBC includes one transmitter and two receivers with additive Gaussian noise at each receiver. The inner bound and the outer bound for these two channels are proved to be equal. The multiple-access channel was first introduced in Shannon, (1961). Further, the authors in T. M. Cover, (1975) and Wyner, (1974) characterize the capacity region of GMAC by showing the achievability and converse proof. The capacity region can be achieved by either simultaneous decoding or successive cancellation decoding with time sharing.

On the other hand, Cover in T. Cover, (1972) introduces GBC for the first time, and show the achievability proof for the capacity region of GBC. The main techniques that are used in the achievability proof of GBC are superposition coding and successive cancellation decoding. A converse proof for the capacity region is later investigated in Bergmans, (1974). It is worth mentioning that the GMAC and GBC has a duality in their structures and capacity regions formulations which is presented in Jindal, Vishwanath, and Goldsmith, (2004). They also determine this duality for fading GMAC and fading GBC under ergodic capacity. This result is based on the equivalence of channel gains and the equivalence of noise power at all receivers of dual channels. It is also proved in Gamal 1981 that the feedback can not increase the capacity region of physically degraded GBC.

However, the Gaussian interference channel (GIC) (without a jammer) is one of the fundamental problems in network information theory that the exact capacity region is still unknown in general. In GIC, there are two transmitters and two receivers who are interested in communicating with their corresponding transmitters, i.e. the is always an interference signal in each receiver. However, the Han-Kobayashi inner bound Han and Kobayashi, (1981) is optimal or near-optimal for many interference channels. The proof of this inner bound involves rate splitting, wherein each transmitter sends a common message, decoded by both receivers, as well as a private message, decoded by only the intended receiver. The authors in Etkin, Tse, and Wang, (2008) showed that for the GIC, the Han-Kobayashi comes within half a bit of the capacity region. Furthermore, Annapureddy and Veeravalli, (2009) obtains an outer bound on the capacity region of the GIC, and it is shown that for sufficiently weak interference signals, treating interference as noise achieves the sum capacity. The deterministic interference channel model is proposed by Bresler and Tse Bresler and Tse, (2008), and they show that the capacity of this channel is within a constant number of bits of the corresponding GIC.

1.3.2 Contribution

Our contribution begins with Chapter 3 where we introduce four versions of Gaussian AVC packing lemmas that are used to bound the error events in the achievability proof of Gaussian multiple-access channels (GMAC) in the presence of jammers, and to bound the error events in the proof of the inner bound of Gaussian broadcast channels (GBC) and Gaussian interference channels (GIC) in the presence of adversaries. The basic approach to these lemmas originates in Csiszár and Narayan, (1988)(b), Lemma 3 and Csiszár and Narayan, (1991), Lemma 1, and the proof is most similar to that of Csiszár and Narayan, (1988)(b), Lemma 3. The earlier lemma showed that a single random codebook satisfies several desirable properties with high probability. In multi-user cases, we need to show that multiple codebooks simultaneously satisfy desirable properties; thus we need a slightly more general approach.

Furthermore, we use Gaussian codewords instead of codewords uniformly distributed on the unit ball. The advantage of Gaussian codewords is that superposition of codewords are themselves Gaussian, and we are dealing with the summation of more than one codeword. Lemma 7 and 8 show that with high probability, two superposed Gaussian codebooks yield small probability of error. While the result is stated for two codebooks for simplicity, it applies for any number of codebooks. Note that Lemma 8 requires $\Lambda < 1$; i.e. the jammer's power must be less than the codeword power, which is necessary to avoid symmetrization. However, in Lemma 7 we do not introduce any power constraint since it benefits from common randomness to prevent symmetrization.

In both Csiszár and Narayan, (1988)(b) and Csiszár and Narayan, (1991), Csiszár and Narayan utilized lemmas (Lemma 3 in Csiszár and Narayan, (1988)(b) and Lemma 1 in Csiszár and Narayan, (1991)) which assert the existence of codebooks with desirable properties in order to prove achievability results. However, they only provided these lemmas for a single codebook, and for either discrete random vectors (in Csiszár and Narayan, (1988)(b)) or codewords uniformly distributed on the unit ball (in Csiszár and Narayan, (1991)). On the other hand, our proof requires a variation on the Gaussian AVC packing lemma that handles decoding of multiple superposed Gaussian codebooks. Our main technical tools for this goal are Lemmas 5, 6, 7, 8 and 9, proved in Sections 3.2, 3.3, 3.4, 3.5 and 3.6, respectively.

Chapter 4, 5 and 6 considers three Gaussian networks in the presence of adversarial jammers in which there are more than two legitimate users (the legitimate transmitters and receivers). We know these channels as Gaussian multiple-user channels. In order to achieve the inner bounds for the capacity regions of the multiple-user channels, we need the aforementioned adversarial packing lemmas that work for more than one codebook, which we propose in Chapter 3. Note that in Chapters 4, 5 and 6, we have some fixed gains for each path between each transmitter and each receiver, but these gains are fixed and not changing in time, so we do not have fading in the channel models.

In Chapters 4, we consider a Gaussian arbitrarily-varying multiple-access channels (GAVMAC). This channel consists of a standard GMAC and an intelligent jammer who sends its signal to the channel. We provide the exact capacity region of the GAVMAC with input power constraints and state power constraints under the average probability of error criterion. Since here we have two transmitters in GAVMAC, and the current literatures do not include a packing lemma for more than one codebook, we benefit from our proposed adversarial packing lemmas (in Chapter 3) to prove the achievability of the capacity region. Note that in this case, the outer bound and the inner bounds coincide and we find a concrete form of the capacity region. The achievability proof follows from simultaneous decoding technique and our proposed adversarial packing lemmas in Chapter 3. The converse proof is a straightforward proof following from the fact that the jammer can only send Gaussian noise if it does not have enough power to symmetrize the channel that is jammer's power is less than each of the transmitter's power; otherwise, the capacity is zero.

Further, in Chapters 5, we consider Gaussian Arbitrarily-varying broadcast channels (GAVBCs). This channel is equivalent to a standard two-user GBC in the presence of two jammers, one at each receiver. We determine an inner and outer bounds for the deterministic-code capacity region of GAVBC under the average probability error. It is also assumed that the channel has input and state power constraints. Again, since we have more than one receiver in the system model, we use our proposed adversarial packing lemmas to achieve the inner bound. In this case our proposed inner and outer bounds differs only in a power constraint between the transmitter and the jammer signals' power. In the inner bound proof, we utilize the superposition coding and successive cancellation decoding along with our proposed adversarial packing lemma. We show that the outer bound is zero if the jammers' power are both less than the corresponding transmitters' power. Otherwise, the outer bound is equal to the capacity of a standard two-user GBC with the noise variance that increases by the power of each jammer at each receiver. Note that even though the duality of GMACs and GBCs is provided in Jindal, Vishwanath, and Goldsmith, (2004), we do not have that sort of duality here in the arbitrarily-varying channels scenario.

In Chapter 6, we first provide a general model including an arbitrary number of jammers for two-user Gaussian interference channels with jammers. We show that its capacity region is equivalent to that of a simplified model in which the received jamming signal at each decoder is independent. Then, existing outer and inner bounds for the two-user Gaussian interference channel are generalized for this simplified jamming model using our proposed adversarial packing lemmas. We show that for certain problem parameters, precisely the same bounds hold, but with the noise variance increased by the received power of the jammer at each receiver. Thus, the jammers can do no better than to transmit Gaussian noise. For these problem parameters, this allows us to recover the half-bit theorem. In weak and strong interference regime, our inner bound coincides the corresponding Han-Kobayashi bound with increased noise variance by the received power of the jammer, and even in strong interference we achieve the exact capacity. Furthermore, we determine the symmetric degrees of freedom where the signal-to-noise, interference-to-noise and jammer-to-noise ratios are all tend to infinity. Moreover, we show that, if the jammer has greater received power than the legitimate user, symmetrizability makes the capacity zero. The proof of the outer bound is straightforward, while the inner bound generalizes the Han-Kobayashi rate splitting scheme. As a novel aspect, the inner bound takes advantage of the common message acting as common randomness for the private message; hence, the jammer cannot symmetrize only the private codeword without being detected. This complication requires an extra condition on the signal power, so that in general our inner bound is not identical to the Han-Kobayashi bound. However, we further are able to achieve the Han-Kobayashi bound if we apply Lemma 7.

Our main contribution in Chapter 6 is to generalize existing inner and outer bounds for the GIC in the presence of AVC-style jammers using our proposed adversarial packing lemmas. We provide a generalized GIC model with G jammers ($G \geq 1$), and show that the capacity region is equivalent to the capacity region of GIC with only two independent jammers. We show that the capacity region depends only on the received power of the jamming signal, not on the number of jammers. Moreover, we obtain the symmetric degrees of freedom (DoF) by taking the limit of the normalized symmetric capacity as signal-to-noise, interference-to-noise and jammer-to-noise ratios converge to infinity. This characterization generalized the so-called "W" DoF curve in ElGamal and Kim, (2011), p. 153. We also recover the optimal sum-rate for the weak interference regime, as well as the exact capacity region for the strong interference regime. We show that our bounds are within a half-bit.

We show that the outer bound in Etkin, Tse, and Wang, (2008) holds with the noise variance increased by the received power of the corresponding jammer at each receiver. The proof, given in Section 6.3, follows by applying the outer bound in Etkin, Tse, and Wang, (2008) with the jammers choosing to transmit Gaussian noise. Moreover, we show that if the jammer's received power at either receiver is larger than that of the intended transmitter, AVC symmetrizability prevents this message from being decoded, because the receiver cannot distinguish the legitimate codeword from the jammer's counterfeit; thus the capacity becomes zero.

We also provide a generalization of the Han-Kobayashi inner bound. For certain problem parameters—for example, in the symmetric case when the jammer's received power is less than that of the interfering user—this inner bound is precisely the Han-Kobayashi inner bound with the noise variance again increased by the received power of the jammer. Thus, for these problem parameters we recover the half-bit theorem of Etkin, Tse, and Wang, (2008), and we can prove that it holds in general by using Lemma 7. The proof of the inner bound, given in Section 6.4, is somewhat more involved, as the receivers must decode correctly no matter what the jammer transmits. One novel aspect of our inner bound proof is that we use the common message in the rate-splitting scheme as common randomness for the private message. Thus, if the jammer has more power than the private codeword but less than both together, it cannot use symmetrization without being detected, and thus the receiver can decode. The result of GAVIC capacity region is published in 2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton) as Hosseinigoki and Kosut, (2016), and the complete version is also available on arXiv as Hosseinigoki and Kosut, (2017).

1.4 Gaussian Arbitrarily-Varying Fading Channels

1.4.1 Prior Work

Wireless communications channels involve a large number of challenges, including noise and fading. Moreover, as it is stated before, wireless channels allow for a malicious intruder to disrupt the operation of the legitimate users. In Chapter 7, we explore how these various effects interact with one another. We adopt a fast fading model, wherein the channel gains form an i.i.d. sequence from a known distribution.

Goldsmith and Varaiya in Goldsmith and Varaiya, (1997) derived the capacity of a single-user fading channel (with no adversary) with the fading gains available at the decoder and possibly the encoder. They showed that if the channel gains are available at only the decoder then the capacity is equal to the expected value of the capacity of a standard Gaussian channel with the received signal to noise ratio multiplied by square of the channel gain. Moreover, if the channel gains are available at both decoder and encoder, then the encoder can choose its signal power as a function of the channel gains in order to maximize the capacity of the channel.

From the security point of view, the problem of slow fading channels, in which the fading gains are constant for all time, is considered in Barros and Rodrigues, (2006) in the presence of an eavesdropper who can only listen to the channel and does not send a signal. The outage capacity is obtained while the channel state information (CSI) is not known at the transmitter. Later in Wang, Yu, and Zhang, (2007), the outage capacity is generalized to the case of multiple eavesdroppers. Moreover, the authors in Li, Yates, and Trappe, (2010) explored the secrecy capacity for a Gaussian channel as the main channel and fast Rayleigh fading channel as the eavesdropper channel while the CSI are only known to the eavesdropper.

Note that in a standard AVC, there is not an eavesdropper or fading; instead, there is only an active attacker who sends its signal to the channel. It is worth mentioning that the "arbitrarily-varying" aspect of the AVC is the adversary's signal, not the channel gains, which we assume to be random from a known distribution. This adversary only knows the code of the legitimate users, but there is not any way for the jammer to access the user's message.

1.4.2 Contribution

In Chapter 7, we consider a Gaussian AVC with fast fading on the main path; we refer to this channel as the Gaussian arbitrarily-varying fading channel (GAVFC). We characterize the capacity of the GAVFC under the average probability of error criterion. Similar to the Gaussian fading channel, we also assume that everyone knows the fading gain distribution including the adversary, but they may or may not know the realization of the gain sequence. Note that the “arbitrarily-varying” aspect of the channel is the adversary’s signal, not the channel gains, which we assume to be random from a known distribution. The receiver always needs the exact fading gains to decode the message, while the adversary and the transmitter may or may not know the exact values of fading gains. Therefore, we derive the capacity of the GAVFC for four cases wherein the channel gains are available at the transmitter and/or adversary as follows:

- Neither the transmitter nor the adversary knows the channel gains.
- Only the transmitter knows the channel gains.
- Only the adversary knows the channel gains.
- Both the transmitter and the adversary knows the channel gains.

If the jammer does not know the channel gains, we show that the capacity is equal to the capacity of the corresponding fading channel with increased noise variance by the power of the jammer. If the jammer knows the exact fading gains, then it can choose its signal as a function of the gains, and under some power constraints it can symmetrize the channel and make the capacity zero. Note that if the channel gains are not available at the adversary, it does not have the required channel information to symmetrize the channel. Moreover, all the results still hold if the adversary and the encoder have the channel gains causally or non-causally, except one situation. If the adversary knows the channel gains causally while the encoder knows them non-causally, then the adversary cannot symmetrize the channel since the encoder possesses some extra information that the adversary does not. These results are published in 2019

53rd Annual Conference on Information Sciences and Systems (CISS) as Hosseinigoki and Kosut, (2019)(a).

1.5 Organization

This dissertation is organized as follows.

Chapter 2: We first describe the standard Gaussian AVC model in Chapter 2, and provide its capacity with list-decoding. We then present the converse and the achievability proof.

Chapter 3: Next, we introduce four adversarial packing lemmas in Chapter 3 for the multiple-user Gaussian AVCs. These Lemmas are used in the achievability proof of Chapters 4, 5 and 6. We proceed to prove each of these adversarial packing lemmas in Chapter 3.

Chapters 4 and 5: The system model of Gaussian arbitrarily-varying multiple-access channels and Gaussian arbitrarily-varying broadcast channels are first introduced in Chapters 4 and 5, respectively. Next, we go on to give the capacity regions limits of these two channels in each corresponding chapter. We then continue to provide the converse and the achievability proofs for the results in Chapter 4 and the inner bound and the outer bound proofs in Chapter 5.

Chapter 6: We describe the problem and the system model for Gaussian arbitrarily-varying interference channel in Chapter 6. Further, the main results including the outer and the inner bounds for the capacity region are given by two theorems. We also discuss implications of our bounds for different regimes, as well as illustrate numerical results in the chapter. The proof for the outer bound and the inner bounds are also provided.

Chapter 7: Finally, we describe the GAVFC model and define its capacities under different channel fading information in Chapter 7. We move on to state our main theorem including the capacity of the GAVFC for five cases. Before giving the proof of our main theorem, we need some auxiliary lemmas and tools. We also perform some numerical simulations for different GAVFCs' capacity region. We then show the converse and the achievability proof for each of the main results. We finally provide a brief proof for the auxiliary results.

1.6 Notation

Upper case letters denote random variables while lower case letters specify a deterministic value or the realization of a random variable. Bold letters denotes n-length vectors. We indicate Hadamard product (element-wise multiplication), inner product and 2-norm by $\circ$, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. We use $| \cdot |^+$ and $\mathbb{E}[\cdot]$ to denote the positive-part function and the expectation, respectively. Also, for an integer N, notation $[N]$ represents the set $\{1, 2, 3, \dots, N\}$. Notation $\mathbf{I}_n$ and $\mathbf{1}_n$ stand for the identity matrix of size n and a vector of size n with n ones elements, respectively. However, notation $\mathbb{1}(\cdot)$ refers to the indicator function. For a vector $\mathbf{v}$, we use superscript $\mathbf{v}^T$ to denote its transpose. Both $\log(\cdot)$ and $\exp(\cdot)$ functions has base 2, so we define the Gaussian channel capacity function $C(x) = \frac{1}{2}\log(1 + x)$, and $X \sim \mathcal{N}(\mu, \sigma^2)$ denotes Gaussian random variable X with mean μ and variance σ^2. We also have this definition $\bar{\alpha} = 1 - \alpha$.

Chapter 2

GAUSSIAN ARBITRARILY-VARYING CHANNEL WITH LIST-DECODING

In this chapter, we determine the capacity of the Gaussian arbitrarily-varying channel (Gaussian AVC) with a (possibly stochastic) encoder and a deterministic list-decoder under the average probability of error criterion. We assume that both the legitimate and the adversarial signals are restricted by their power constraints. We also assume that there is no path between the adversary and the legitimate user but the adversary knows the legitimate user's code. We show that for any list size L, the capacity is equivalent to the capacity of a point-to-point Gaussian channel with noise variance increased by the adversary power, if the adversary has less power than L times the transmitter power; otherwise, the capacity is zero. In the converse proof, we show that if the adversary has enough power, then the decoder can be confounded by the adversarial superposition of several codewords while satisfying its power constraint with positive probability. The achievability proof benefits from a novel variant of the Csiszár-Narayan method for the arbitrarily-varying channel.

2.1 Problem Statement

A Gaussian AVC is a modified standard point-to-point Gaussian channel in the presence of an additive arbitrarily chosen adversary signal as it is shown in Figure 1. Both transmitter and receiver do not know anything about the adversary signal and the adversary do not have any information about the transmitted signal except the

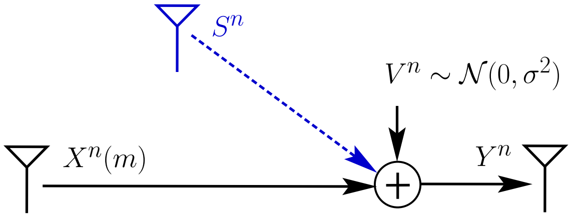


Figure 1: Gaussian Arbitrarily-Varying Channel

codebooks. The received signal is given by

$$\mathbf{Y} = \mathbf{x} + \mathbf{s} + \mathbf{V} \tag{2.1}$$

where the n-length vector $\mathbf{x}$ is the legitimate transmitter's signal, $\mathbf{s}$ represents the independent adversary signal and noise $\mathbf{V}$ is a sequence of n-length i.i.d. zero mean Gaussian random variables with variance σ^2, independent of $\mathbf{x}$ and $\mathbf{s}$.

We have the assumption of peak power constraints for the transmitter and adversary signals respectively as $\|\mathbf{x}\|^2 \leq \ nP$ and $\|\mathbf{s}\|^2 \leq n\Lambda$. In addition, the transmitter and receiver are assumed to know the three parameters P, Λ and σ^2.

An (n, N, N_r, L) stochastic list code for the Gaussian AVC is given by:

- Message set $\mathcal{M} = [N]$ and encoder private randomness set $\mathcal{K} = [N_r]$,
- Stochastic encoding function $\mathbf{x}(M, K) : \mathcal{M} \times \mathcal{K} \to \mathbb{R}^n$,
- List decoding function $\phi(\mathbf{y}) : \mathbb{R}^n \to \mathcal{D}_L = \{\mathcal{L} : \mathcal{L} \subset [N], |\mathcal{L}| \leq L\}$,

where the rate of the code is $R = \frac{1}{n}\log(N/L)$. The transmitter encodes the message M and its private randomness K to $\mathbf{x}(M, K)$ where M and K are chosen uniformly respectively from sets $\mathcal{M}$ and $\mathcal{K}$. At the receiver, signal $\mathbf{Y}$ is decoded by a deterministic function ϕ to the set $\mathcal{D}_L$ which is the set of all subsets of $[N]$ with cardinality at most L. In other words, L is the size of the list decoder.

First, define the probability of error $e(\mathbf{s}, i)$ for a specific message $i \in [N]$ in the presence of a specific adversary signal $\mathbf{s} \in \mathbb{R}^n$ as the probability that $i \notin \phi(\mathbf{y})$. Therefore, the average probability of error for $\mathbf{s}$ is given by

$$\bar{e}(\mathbf{s}) = \frac{1}{N} \sum_{i=1}^{N} e(\mathbf{s}, i). \tag{2.2}$$

Finally, the overall probability of error $P_e^{(n)}$ is obtained by maximizing over all possible choices of jammers' sequences $\mathbf{s}$ satisfying peak power constraint $\|\mathbf{s}\|^2 \leq n\Lambda$. Suppose r is rate of private randomness. Given L and r, rate R is *achievable* if there exists a sequence of $\left(n, L2^{nR}, 2^{nr}, L\right)$ codes such that $\lim_{n \to \infty} P_e^{(n)} = 0$. The list-code capacity $\mathscr{C}(L, r)$ is the supremum of all achievable rates given L and r.

2.2 Main Results

Theorem 1 *The list-code capacity of Gaussian AVC is given by*

$$\mathscr{C}(L, r) = \begin{cases} C\left(\frac{P}{\Lambda + \sigma^2}\right), & L > \frac{\Lambda}{P} \\ 0, & L < \frac{\Lambda}{P}. \end{cases} \tag{2.3}$$

Note that the capacity for $\Lambda = LP$ is unsolved.

Remark 1 *Note that this result holds for all r, including $r = 0$ which corresponds to a deterministic encoder. That is, the capacity does not depend on the amount of private randomness.*

Remark 2 *The condition on the ratio $\frac{\Lambda}{P}$ determines whether it is possible for the adversary to launch a symmetrizing attack, wherein it transmits a superposition of codewords. Since each codeword has power P, the most codewords that the adversary can superpose while obeying its power constraint of Λ is the $\lfloor \frac{\Lambda}{P} \rfloor$. Thus, if the allowable* *list size is greater than $\frac{\Lambda}{P}$, then even under this attack the decoder can output a list made up of the true message and the superposed codewords selected by the adversary. Of course, the decoder does not know which is which but it can still guarantee that the true message is in the list. Thus, the worst the adversary can do is to act as an extra additive Gaussian noise with variance Λ, so the capacity is equal to the capacity of a standard Gaussian channel with increased noise variance as in $C(\frac{P}{\Lambda+\sigma^2})$. However, if the allowable list size is less than $\frac{\Lambda}{P}$, then there are too many possibilities for the decoder to decode correctly, so the capacity is zero. Note that none of this depends on the channel noise, so σ^2 does not come into play in the condition on L.*

Remark 3 *For the achievability proof, we make no assumptions about what the adversary does. However, for the converse proof, it is allowable to weaken the adversary by making certain assumptions about its behavior, because doing so can only increase the achievable rates. Since the converse is an upper bound on achievable rates, weakening the adversary in this manner still yields a valid upper bound.*

In our proofs in Sections 2.3 and 2.4, without loss of generality we restrict ourselves to the transmitter's power $P = 1$ which can be done by scaling the output signal.

2.3 Converse Proof

Without loss of generality, suppose $P = 1$. For the first case where $\Lambda < L$, we assume that we have a code $(n, L2^{nR}, 2^{nr}, L)$ with vanishing probability of error. Since these codes must function for arbitrary jamming signals, we may derive an outer bound by assuming the adversary transmits Gaussian noise with variance $\Lambda - \gamma$ for any $\gamma > 0$ or $\mathbf{0}$ if Gaussian realization has power greater than Λ. By the law of large numbers, with high probability the resulting channel is equivalent to a standard Gaussian channel with noise variance $\sigma^2 + \Lambda - \gamma$. Thus, since γ can be chosen arbitrarily small, from the capacity of a non-adversarial Gaussian channel,

$$\mathscr{C}(L, r) \le C\left(\frac{1}{\sigma^2 + \Lambda}\right). \tag{2.4}$$

Now, assume the symmetrizable case where $\Lambda > L$. In order to show $\mathscr{C}(L, r) = 0$, first consider a sequence of stochastic codebooks and probability of error $P_e^{(n)}$. We claim that if $R > 0$ and the jammer has the following strategy, then $P_e^{(n)}$ is bounded away from zero for sufficiently large n: The jammer randomly and uniformly chooses L messages $M_1, \ldots, M_L$ from $[L2^{nR}]$ and also L private keys $K_1, \ldots, K_L$ from $[2^{nr}]$ where M_i and K_i are independent. Note that the jammer knows the transmitted codebook. The jammer then constructs

$$\mathbf{Z} = \mathbf{x}(M_1, K_1) + \ldots + \mathbf{x}(M_L, K_L) - L\boldsymbol{\mu} \tag{2.5}$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is a constant vector that we will choose later. The jammer transmits $\mathbf{S} = \mathbf{Z}$ if $\|\mathbf{Z}\|^2 \le n\Lambda$ or else transmits $\mathbf{S} = \mathbf{0}$. In the former case, the received signal is

$$\begin{aligned} \mathbf{Y} &= \mathbf{x}(M_0, K_0) + \mathbf{x}(M_1, K_1) + \ldots + \mathbf{x}(M_L, K_L) - L\boldsymbol{\mu} + \mathbf{V} && (2.6)\\ &= \boldsymbol{\mu} + \sum_{i=0}^{L} (\mathbf{x}(M_i, K_i) - \boldsymbol{\mu}) + \mathbf{V} && (2.7) \end{aligned}$$

where M_0 is the true message. If $M_0, M_1, \ldots, M_L$ are all different and for all sets $D \subset \{0, 1, \ldots, L\}$ with $|D| = L$,

$$\left\| \sum_{i \in D} \mathbf{x}(M_i, K_i) - L\boldsymbol{\mu} \right\|^2 \le n\Lambda, \tag{2.8}$$

then from the decoder's perspective, any L of the $L + 1$ messages might have been forged by the adversary. Therefore, the list decoder with list size at most L has a probability of error at least $\frac{1}{L+1}$; that is, the probability that the decoder chooses L

from the $L+1$ messages that does not contain the true message M_0. That is,

$$P_e^{(n)} \geq \frac{1}{L+1}\mathbb{P}\Bigg(\Bigg\|\sum_{i\in D}\mathbf{x}(M_i,K_i) - L\boldsymbol{\mu}\Bigg\|^2 \leq n\Lambda \quad \text{for all } D\subset\{0,1,\dots,L\} \text{ with } |D|=L, \text{ and } M_0,M_1,\dots,M_L \text{ are distinct}\Bigg) \tag{2.9}$$

$$\geq \frac{1}{L+1}\Bigg[\mathbb{P}\Bigg(\Bigg\|\sum_{i\in D}\mathbf{X}_i - L\boldsymbol{\mu}\Bigg\|^2 \leq n\Lambda \text{ for all } D\subset\{0,1,\dots,L\} \text{ with } |D|=L\Bigg) - \Bigg(1 - \frac{L2^{nR}-1}{L2^{nR}}\cdot\frac{L2^{nR}-2}{L2^{nR}}\cdots\frac{L2^{nR}-L}{L2^{nR}}\Bigg)\Bigg] \tag{2.10}$$

where $\mathbf{X}_i = \mathbf{x}(M_i,K_i)$ and the second term in (2.10) shows the probability of messages $M_0,\dots,M_L$ not being distinct which tends to zero as $n\to\infty$. Note that $\mathbf{X}_0,\mathbf{X}_1,\dots,\mathbf{X}_L$ are independent and each distributed as a transmitted sequence from the code. We proceed to show that there exists a choice of $\boldsymbol{\mu}$ based only on the codebook such that (2.10) is bounded away from zero for sufficiently large n if $R>0$.

Let

$$\alpha^\star = \inf\left\{\alpha : \liminf_{n\to\infty}\max_{\boldsymbol{\mu}\in\mathbb{R}^n}\mathbb{P}(\|\mathbf{X}_0-\boldsymbol{\mu}\|^2 \leq n\alpha) > 0\right\}. \tag{2.11}$$

Note that $\alpha^\star \leq 1$, since by the power constraint we always have $\mathbb{P}(\|\mathbf{X}_0\|^2 \leq n) = 1$. Fix any $\delta > 0$ and let $\alpha = \alpha^\star + \delta/2$. Let

$$\gamma = \liminf_{n\to\infty}\max_{\boldsymbol{\mu}\in\mathbb{R}^n}\mathbb{P}(\|\mathbf{X}_0-\boldsymbol{\mu}\|^2 \leq n\alpha). \tag{2.12}$$

Since $\alpha > \alpha^\star$ we have $\gamma > 0$. Thus for n sufficiently large, there exists $\boldsymbol{\mu}\in\mathbb{R}^n$ such that

$$\mathbb{P}(\|\mathbf{X}_0-\boldsymbol{\mu}\|^2 \leq n\alpha) \geq \gamma - \delta. \tag{2.13}$$

This $\boldsymbol{\mu}$ is the one to be used by the jammer in (2.5). Let $\mathcal{B}_n$ be the set of all $\mathbf{x}$ that satisfy $\|\mathbf{x}-\boldsymbol{\mu}\|^2 \le n\alpha$. Note that $\mathbb{P}(\mathbf{X}_0 \in \mathcal{B}_n) \ge \gamma - \delta$.

Since $\alpha - \delta < \alpha^\star$, by the definition of $\alpha^\star$,

$$\liminf_{n\to\infty} \max_{\boldsymbol{\mu}'\in\mathbb{R}} \mathbb{P}(\|\mathbf{X}_0 - \boldsymbol{\mu}'\|^2 \le n(\alpha-\delta)) = 0. \tag{2.14}$$

Specifically, there exists n sufficiently large so that for all $\boldsymbol{\mu}' \in \mathbb{R}^n$,

$$\mathbb{P}(\|\mathbf{X}_0 - \boldsymbol{\mu}'\|^2 \le n(\alpha-\delta)) \le \delta. \tag{2.15}$$

Fix any $\mathbf{x}_1 \in \mathcal{B}_n$ and consider those $\mathbf{x} \in \mathcal{B}_n$ such that

$$\langle \mathbf{x}-\boldsymbol{\mu}, \mathbf{x}_1 - \boldsymbol{\mu}\rangle > n\sqrt{\delta\alpha} \tag{2.16}$$

which implies

$$\left\|\boldsymbol{\mu} + \sqrt{\delta\alpha^{-1}}(\mathbf{x}_1-\boldsymbol{\mu}) - \mathbf{x}\right\|^2 = \|\mathbf{x}-\boldsymbol{\mu}\|^2 + \delta\alpha^{-1}\|\mathbf{x}_1-\boldsymbol{\mu}\|^2 - 2\sqrt{\delta\alpha^{-1}}\langle\mathbf{x}-\boldsymbol{\mu},\mathbf{x}_1-\boldsymbol{\mu}\rangle \tag{2.17}$$

$$< n\alpha + n\delta - n2\delta \tag{2.18}$$

$$= n(\alpha-\delta). \tag{2.19}$$

Thus, we obtain the following by applying (2.15) with $\boldsymbol{\mu}' = \boldsymbol{\mu} + \sqrt{\delta\alpha^{-1}}(\mathbf{x}_1 - \boldsymbol{\mu})$ as

$$\mathbb{P}(\langle \mathbf{X}_0 - \boldsymbol{\mu}, \mathbf{X}_1 - \boldsymbol{\mu}\rangle > n\sqrt{\delta\alpha}, \mathbf{X}_0, \mathbf{X}_1 \in \mathcal{B}_n)$$

$$\le \max_{\mathbf{x}_1\in\mathcal{B}_n} \mathbb{P}(\langle \mathbf{X}_0 - \boldsymbol{\mu}, \mathbf{x}_1 - \boldsymbol{\mu}\rangle > n\sqrt{\delta\alpha},\ \mathbf{X}_0 \in \mathcal{B}_n) \tag{2.20}$$

$$\le \delta. \tag{2.21}$$

Moreover, if $\mathbf{x}_1, \ldots, \mathbf{x}_L \in \mathcal{B}_n$ satisfy $\langle \mathbf{x}_i - \boldsymbol{\mu}, \mathbf{x}_j - \boldsymbol{\mu} \rangle \le n\sqrt{\delta\alpha}$ for all $i \neq j \in \{1, \ldots, L\ \}$, then

$$\begin{aligned}
\|\mathbf{x}_1 + \ldots + \mathbf{x}_L - L\boldsymbol{\mu}\|^2 &= \sum_{i=0}^{L-1} \|\mathbf{x}_i - \boldsymbol{\mu}\|^2 + 2\sum_{i=0}^{L-2} \sum_{j=i+1}^{L-1} \langle \mathbf{x}_i - \boldsymbol{\mu}, \mathbf{x}_j - \boldsymbol{\mu} \rangle && (2.22)\\
&\le n\big(L\alpha + L(L-1)\sqrt{\delta\alpha}\big) && (2.23)\\
&\le n\big(L + L\delta + L(L-1)\sqrt{\delta\alpha}\big) && (2.24)\\
&< n\Lambda && (2.25)
\end{aligned}$$

where (2.24) holds since $\alpha < \alpha^\star + \delta \le 1 + \delta$ and (2.25) holds for sufficiently small δ and by assumption $\Lambda > L$. Now we have

$$\begin{aligned}
&\mathbb{P}\left(\left\| \sum_{i \in D} \mathbf{X}_i - L\boldsymbol{\mu} \right\|^2 \le n\Lambda \text{ for all distinct set } D \subset \{0, 1, \ldots, L\} \text{ with } |D| = L \right)\\
&\quad \ge \mathbb{P}\Big(\langle \mathbf{X}_i - \boldsymbol{\mu}, \mathbf{X}_j - \boldsymbol{\mu} \rangle \le n\sqrt{\delta\alpha} \text{ for all } i, j \in \{0, 1, \ldots, L\}, i \neq j, \mathbf{X}_0, \ldots, \mathbf{X}_L \in \mathcal{B}_n \Big) && (2.26)\\
&\quad \ge \mathbb{P}(\mathbf{X}_0 \in \mathcal{B}_n)^{(L+1)} - \frac{L(L+1)}{2} \mathbb{P}(\langle \mathbf{X}_0 - \boldsymbol{\mu}, \mathbf{X}_1 - \boldsymbol{\mu} \rangle > n\sqrt{\delta\alpha}, \mathbf{X}_0, \mathbf{X}_1 \in \mathcal{B}_n) && (2.27)\\
&\quad \ge (\gamma - \delta)^{(L+1)} - \frac{L(L+1)\delta}{2} && (2.28)
\end{aligned}$$

where (2.26) follows from the analysis leading to (2.25), (2.27) follows from the union bound and the fact that $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_L$ are independent and (2.28) follows from the lower bound on the probability of $\mathcal{B}_n$ and (2.21). For sufficiently small δ, (2.28) is bounded away from zero, so by (2.10), $P_e^{(n)}$ is also bounded away from zero for sufficiently large n if $R > 0$.

Before proceeding to the proof, let define the typical set for Gaussian random variables $X_1, \ldots, X_k$ as:

$$\mathcal{T}_\epsilon^{(n)}(X_1, \ldots, X_k) = \left\{(\mathbf{x}_1, \ldots, \mathbf{x}_k) : \mathbb{E}(X_i X_j) - \epsilon \leq \frac{1}{n}\langle \mathbf{x}_i, \mathbf{x}_j \rangle \leq \mathbb{E}(X_i X_j) + \epsilon \text{ for all } i, j \in [1:k]\right\}. \quad (2.29)$$

Also, we use the following lemmas several times throughout the thesis, so we provide them here. These lemmas can be easily generalized for Gaussian random variables by following the corresponding lemmas in ElGamal and Kim, (2011) for discrete memoryless random variables.

Lemma 2 (Conditional Typicality Lemma) *: Let $(X, Y) \sim f(x, y)$ be jointly Gaussian random variables. Suppose that $\mathbf{x} \in \mathcal{T}_\epsilon^{(n)}(X)$ and $\mathbf{Y} \sim f(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n f_{Y|X}(y_i|x_i)$. Then, for every $\epsilon > \epsilon'$,*

$$\lim_{n\to\infty} \mathbb{P}\{(\mathbf{x}, \mathbf{Y}) \in \mathcal{T}_\epsilon^{(n)}(X, Y)\} = 1. \quad (2.30)$$

Lemma 3 (Joint Typicality Lemma) *: Let $(X, Y, Z) \sim f(x, y, z)$ be jointly Gaussian random variables. If $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is a pair of arbitrary sequences and $\tilde{\mathbf{Z}} \sim \prod_{i=1}^n f_{Z|X}(\tilde{z}_i|\tilde{x}_i)$ then there exists $\delta(\epsilon) > 0$ that tends to zero as $\epsilon \to 0$ such that*

$$\mathbb{P}\{(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{Z}}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, Z)\} \leq \exp(-n(I(Y; Z|X) - \delta(\epsilon))). \quad (2.31)$$

Now, without loss of generality, assume $P = 1$, $r = 0$ and

$$\Lambda < L, \quad (2.32)$$

$$R < C\left(\frac{1}{\Lambda + \sigma^2}\right). \quad (2.33)$$

Note that assuming $r = 0$ makes the code deterministic. Thus, it suffices to achieve the list-code capacity by a $(n, L2^{nR}, L)$ deterministic code construction as follows:

Codebook generation: Fix $\epsilon > \epsilon' > \gamma > 0$. We generate $L2^{nR}$ i.i.d zero mean Gaussian sequences $\mathbf{X}(m)$ with variance $(1 - \gamma)$ for each $m \in [L2^{nR}]$.

Encoding: The transmitter sends $\mathbf{X} = \mathbf{X}(m)$ if its power is less than 1, otherwise it sends zero.

Decoding: First, given $\mathbf{y}$, for $1 \leq \ell \leq L$ let set $\mathscr{S}_\ell$ be the set of ℓ-tuple messages $(m_1, \ldots, m_\ell)$ that $(\mathbf{x}(m_1), \ldots, \mathbf{x}(m_\ell), \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(X_1, \ldots, X_\ell, Y)$ for any set of zero-mean Gaussian variables $(X_1, \ldots, X_\ell, Y)$ such that

$$\text{Cov}(X_1, \ldots, X_\ell, Y) = \begin{bmatrix} \mathbf{I}_\ell & \mathbf{1}_\ell^T \\ \mathbf{1}_\ell & A \end{bmatrix}_{(\ell+1)\times(\ell+1)} \tag{2.34}$$

and $1 \leq A \leq 1 + \sigma^2 + \Lambda$.

Now we define the decoding function as

$$\phi(\mathbf{y}) = \operatorname*{arg\,min}_{\mathcal{L} \in \bigcup_{\ell=1}^{L} \mathscr{S}_\ell} \left\| \mathbf{y} - \sum_{m \in \mathcal{L}} \mathbf{x}(m) \right\|. \tag{2.35}$$

where we choose between multiple minimizing $\mathcal{L}$ arbitrarily.

Analysis of the probability of error: To prove that the constructed code is achievable meaning that the probability of error tends to zero as $n \to \infty$, we utilize several lemmas including the following. We provide some necessary codebook properties that hold with high probability in Lemma 4, the proof for which is in the Section 2.5.

Lemma 4 *Let $\mathbf{X}(m)$ for $m \in [N]$, $N = L2^{nR}$, be the Gaussian codebook described above. With probability approaching 1 as $n \to \infty$, the codebook satisfies the following, for any $\mathbf{x}, \mathbf{s}$ where $\|\mathbf{s}\|^2 \leq n\Lambda$ and any zero-mean jointly Gaussian random vector $(X, X_1, \ldots, X_\ell, S)$ with positive definite covariance matrices with diagonals at most*

$(1, 1, \ldots, 1, \Lambda)$ *for all* $1 \le \ell \le L$:

$$\frac{1}{N}\left|\left\{m\colon(\mathbf{x}(m),\mathbf{s}) \notin \bigcup_{\substack{(X,S)\ \textit{independent:}\\ EX^2=1, ES^2\le\Lambda}} \mathcal{T}_{\epsilon'}^{(n)}(X,S)\right\}\right| \le \exp(-n\delta(\epsilon)), \tag{2.36}$$

$$\left|\left\{m_1 : (\mathbf{x},\mathbf{x}(m_1),\mathbf{s}) \in \mathcal{T}_{\epsilon}^{(n)}(X,X_1,S)\right\}\right| \le \exp\left\{n\left[|R - I(X_1;XS)|^+ + \delta(\epsilon)\right]\right\}, \tag{2.37}$$

$$\frac{1}{N}\left|\left\{m\colon(\mathbf{x}(m),\mathbf{x}(m_1),\mathbf{s}) \in \mathcal{T}_{\epsilon}^{(n)}(X,X_1,S)\ \textit{for some}\ m_1 \ne m\right\}\right| \le 2\exp\{-n\delta(\epsilon)/2\},$$

$$\textit{if}\ I(X;X_1S) \ge |R - I(X_1;S)|^+ + \delta(\epsilon), \tag{2.38}$$

$$\left|\left\{(m_1,\ldots,m_\ell)\colon(\mathbf{x},\mathbf{x}(m_1),\ldots,\mathbf{x}(m_\ell),\mathbf{s}) \in \mathcal{T}_{\epsilon}^{(n)}(X,X_1,\ldots,X_\ell,S)\right\}\right| \le \exp\left\{n\delta(\epsilon)\right\}$$

$$\textit{if}\ R < \min_{k\in\{1,\ldots,\ell\}} I(X_k;S), \tag{2.39}$$

$$\frac{1}{N}\left|\left\{m\colon(\mathbf{x}(m),\mathbf{x}(m_1),\ldots,\mathbf{x}(m_\ell),\mathbf{s}) \in \mathcal{T}_{\epsilon}^{(n)}(X,X_1,\ldots,X_\ell,S)\ \textit{for some}\ m_1,\ldots,m_\ell \ne m\right\}\right|$$

$$\le \exp\{-n\delta(\epsilon)/2\}, \quad \textit{if}\ I(X;X_1\ldots X_\ell S) \ge \delta(\epsilon)\ \textit{and}\ R < \min_{k\in\{1,\ldots,\ell\}} I(X_k;S). \tag{2.40}$$

Assume that the legitimate user transmits message M. Then, the probability of error is upper bounded by the sum of the following L error events probabilities:

$$P_\ell = \mathbb{P}\left\{\|\mathbf{Y}-\mathbf{x}(m_1)-\ldots-\mathbf{x}(m_\ell)\|^2 \le \min_{\substack{\hat{m}_1,\ldots,\hat{m}_\ell:\\ (M,\hat{m}_1,\ldots,\hat{m}_\ell)\in\mathscr{S}_{\ell+1}}} \|\mathbf{Y}-\mathbf{x}(M)-\mathbf{x}(\hat{m}_1)-\ldots-\mathbf{x}(\hat{m}_\ell)\|^2\right.$$

$$\left.\text{for some } (m_1,\ldots,m_\ell) \in \mathscr{S}_\ell, m_i \ne M \text{ for all } i \in [\ell]\right\} \text{ for } 1 \le \ell < L, \tag{2.41}$$

$$P_L = \mathbb{P}\left\{\exists(m_1,\ldots,m_L) \in \mathscr{S}_L : m_\ell \ne M,\ \text{for all } \ell \in [L]\right\}. \tag{2.42}$$

By Lemma 4, we may assume we have a deterministic codebook that satisfies (2.36)–(2.40). Consider any state sequence $\mathbf{s}$. By (2.36), with high probability

$(\mathbf{x}(M), \mathbf{s}) \in \mathcal{T}_{\epsilon'}^{(n)}(X, S)$ where (X, S) are independent and $\mathbb{E}X^2 = 1, \mathbb{E}S^2 \leq \Lambda$ (2.36). Thus, by the conditional typicality lemma 2, for every $\epsilon > \epsilon'$ with high probability $(\mathbf{x}_i, \mathbf{s}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(X, S, V)$ where (X, S, V) are mutually independent and $\mathbb{E}V^2 = \sigma^2$.

We first bound probability event P_ℓ for $1 \leq \ell < L$. Define the shorthand $\vec{X}_\ell = (XX_1 \dots X_\ell SV)$. Let $\mathcal{V}_\ell$ denote a finite set of Gaussian distributions of $\vec{X}_\ell$ that is ϵ-dense in the set of all Gaussian distributions of $\vec{X}_\ell$ with variances at most $(1, 1, \dots, 1, \Lambda, \sigma^2)$. Note that the cardinality of $\mathcal{V}_\ell$ does not depend on n. We may upper bound P_ℓ by

$$\sum_{\vec{X}_\ell \in \mathcal{V}_\ell} \frac{1}{N} \sum_{m=1}^{N} e_{\vec{X}_\ell}(m, \mathbf{s}) \tag{2.43}$$

where

$$\begin{aligned} e_{\vec{X}_\ell}(m, \mathbf{s}) = \mathbb{P}\Big\{ &(\mathbf{x}(m), \mathbf{x}(m_1), \dots, \mathbf{x}(m_\ell), \mathbf{s}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(\vec{X}_\ell), \\ &\|\mathbf{x}(m) + \mathbf{s} + \mathbf{V} - \mathbf{x}(m_1) - \dots - \mathbf{x}(m_\ell)\|^2 \leq \min_{\substack{\hat{m}_1, \dots, \hat{m}_\ell: \\ (m, \hat{m}_1, \dots, \hat{m}_\ell) \in \mathscr{S}_{\ell+1}}} \|\mathbf{s} + \mathbf{V} - \mathbf{x}(\hat{m}_1) - \dots - \mathbf{x}(\hat{m}_\ell)\|^2 \\ &\text{for some } (m_1, \dots, m_\ell) \in \mathscr{S}_\ell \text{ and } m_i \neq m \text{ for all } i \in [\ell] \Big\}. \end{aligned} \tag{2.44}$$

We will show that $\frac{1}{N} \sum_{m=1}^{N} e_{\vec{X}_\ell}(m, \mathbf{s}) \to 0$ for all Gaussian vectors $\vec{X}_\ell$ (whether or not they are in $\mathcal{V}_\ell$). We may restrict ourselves to $\vec{X}_\ell$ where

$$(X, S, V) \text{ are mutually independent,} \tag{2.45}$$

$$\mathbb{E}X^2 = \mathbb{E}X_1^2 = \dots = \mathbb{E}X_\ell^2 = 1, \quad \mathbb{E}V^2 = \sigma^2, \quad \mathbb{E}S^2 \leq \Lambda \tag{2.46}$$

$$(X_i, X + S + V - X_i) \text{ are independent for all } i \in [\ell], \tag{2.47}$$

$$\mathbb{E}(X + S + V - X_i)^2 \leq \Lambda + \sigma^2 \text{ for all } i \in [\ell], \tag{2.48}$$

where (2.45) holds since the legitimate transmitter, the jammer and the noise are independently generated, (2.46) follows from the assumptions for $\vec{X}_\ell$, (2.47) corresponds to $\mathbb{E}X_i(Y - X_i) = 0$ following from (2.34) and the assumption that $(m_1, \dots, m_\ell) \in \mathscr{S}_\ell$

and (2.48) is obtained by (2.34) as follows:

$$\begin{aligned}\mathbb{E}(X+S+V-X_i)^2 &= \mathbb{E}(Y-X_i)^2 && (2.49)\\ &= \mathbb{E}Y^2 + \mathbb{E}X_i^2 - 2\mathbb{E}YX_i && (2.50)\\ &\le 1+\sigma^2+\Lambda+1-2 && (2.51)\\ &= \Lambda+\sigma^2. && (2.52)\end{aligned}$$

Now, suppose

$$I(XV; SX_1 \dots X_\ell) = 0. \tag{2.53}$$

Then we would have $\mathbb{E}XX_i = 0$ for all $i \in [\ell]$. Thus, $(X, X_1, \dots, X_\ell, S+V-X_1-\dots-X_\ell)$ are mutually independent since

$$\mathbb{E}X(S+V-X_1-\dots-X_\ell) = \mathbb{E}X(S+V) = 0, \tag{2.54}$$

and

$$\begin{aligned}\mathbb{E}X_i(S+V-X_1-\dots-X_\ell) &= \mathbb{E}X_i(Y-X-X_1-\dots-X_\ell) && (2.55)\\ &= \mathbb{E}X_i(Y-X_i) && (2.56)\\ &= 0, \text{ for all } i \in [\ell], && (2.57)\end{aligned}$$

where (2.54) follows from (2.45) and (2.56)–(2.57) follow from (2.34) and $\mathbb{E}XX_i = 0$. Hence, if the message $(m_1, \dots, m_\ell)$ satisfies the conditions in the probability in (2.44), then $(m, m_1, \dots, m_\ell) \in \mathscr{S}_{\ell+1}$. This implies that $(\hat{m}_1, \dots, \hat{m}_\ell)$ takes on the value $(m_1, \dots, m_\ell)$ in the minimum, so $\|\mathbf{Y} - \mathbf{x}(m_1) - \dots - \mathbf{x}(m_\ell)\|^2 \le \|\mathbf{Y} - \mathbf{x}(m_1) - \dots - \mathbf{x}(m_\ell) - \mathbf{x}(M)\|^2$ and so we must have

$$\mathbb{E}(X+S+V-X_1-\dots-X_\ell)^2 \le \mathbb{E}(S+V-X_1-\dots-X_\ell)^2. \tag{2.58}$$

However, this contradicts the assumptions that X is independent from $S, X_1, \ldots, X_\ell, V$, since

$$\mathbb{E}(X+S+V-X_1-\ldots,X_\ell)^2 = \mathbb{E}X^2 + \mathbb{E}(S+V-X_1-\ldots,X_\ell)^2 \tag{2.59}$$

$$= 1 + \mathbb{E}(S+V-X_1-\ldots,X_\ell)^2. \tag{2.60}$$

Therefore, the assumption in (2.53) is false and there exists $\eta > 0$ such that

$$\eta \le I(XV; SX_1\ldots X_\ell) = I(XV; X_1\ldots X_\ell|S). \tag{2.61}$$

Now, consider the following two cases.

Case (a): $R < \min\{I(X_1;S),\ldots,I(X_\ell;S)\}$. By (2.40), we only need to consider distributions where

$$I(X; X_1\ldots X_\ell S) < \delta(\epsilon). \tag{2.62}$$

Then for any $m, \mathbf{s}$

$$e_{\vec{X}_\ell}(m,\mathbf{s}) \le \sum_{m_1,\ldots,m_\ell} \mathbb{P}\left\{(\mathbf{x}(m),\mathbf{x}(m_1),\ldots,\mathbf{x}(m_\ell),\mathbf{s},\mathbf{V})\in\mathcal{T}_\epsilon^{(n)}(X,X_1,\ldots,X_\ell,S,V)\right\} \tag{2.63}$$

$$\le \exp\left\{-n\left(I(V;X_1\ldots X_\ell|XS)-\delta(\epsilon)\right)\right\} \tag{2.64}$$

$$= \exp\{-n(I(XV;X_1\ldots X_\ell|S) - I(X;X_1\ldots X_\ell|S) - \delta(\epsilon))\} \tag{2.65}$$

$$\le \exp\{-n(\eta - 2\delta(\epsilon))\}. \tag{2.66}$$

where in (2.63) the sum is over all $m_1,\ldots,m_\ell : (\mathbf{x}(m),\mathbf{x}(m_1),\ldots,\mathbf{x}(m_\ell),\mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X,X_1,\ldots,X_\ell,S)$, in (2.64) we use (2.39), the joint typicality lemma 3 and $I(V;XS)=0$ and (2.66) follows from (2.61) and (2.62). Thus, (2.66) tends to zero exponentially fast for sufficiently small $\delta(\epsilon)$.

Case (b): $R \geq \min\{I(X_1;S),\ldots,I(X_\ell;S)\}$. We may assume without loss of generality that $R \geq I(X_1;S)$. Now, we upper bound (2.44) by

$$e_{\vec{X}_\ell}(m,\mathbf{s}) \;\leq \sum_{\hat{m}:(\mathbf{x}(m),\mathbf{x}(\hat{m}),\mathbf{s})\in\mathcal{T}_\epsilon^{(n)}(X,X_1,S)} \mathbb{P}\left\{(\mathbf{x}(m),\mathbf{x}(\hat{m}),\mathbf{s},\mathbf{V})\in\mathcal{T}_\epsilon^{(n)}(X,X_1,S,V)\right\}. \tag{2.67}$$

Note that by (2.38), we may narrow the distributions to those with

$$I(X;X_1S) < R - I(X_1;S) + \delta(\epsilon). \tag{2.68}$$

Therefore,

$$\begin{aligned} e_{\vec{X}_\ell}(m,\mathbf{s}) &\leq \exp\left\{n\left[|R-I(X_1;XS)|^+ - I(V;X_1|XS)+2\delta(\epsilon)\right]\right. && (2.69)\\ &\leq \exp\left(n\left[R-I(X_1;XS)-I(V;X_1|XS)+2\delta(\epsilon)\right]\right) && (2.70)\\ &= \exp(n[R-I(X_1;XSV)+2\delta(\epsilon)]) && (2.71)\end{aligned}$$

where (2.69) follows from (2.37) and the joint typicality lemma 3, (2.70) follows since by $R \geq I(X_1;S)$ and (2.68) we have

$$\begin{aligned} R &> I(X;X_1S) + I(X_1;S) - \delta(\epsilon) && (2.72)\\ &= I(X;X_1|S) + I(X_1;S) - \delta(\epsilon) && (2.73)\\ &= I(X_1;XS) - \delta(\epsilon). && (2.74)\end{aligned}$$

Let $Z = X + S + V - X_1$. From (2.47)–(2.48), we get

$$\begin{aligned} I(X_1;XSV) &\geq I(X_1;X_1+Z) && (2.75)\\ &= C\left(\frac{1}{\mathbb{E}Z^2}\right) && (2.76)\\ &\geq C\left(\frac{1}{\Lambda+\sigma^2}\right) && (2.77)\end{aligned}$$

Using this result in (2.71), we obtain

$$e_{\vec{X}_\ell}(m, \mathbf{s}) \leq \exp\left\{n\left[R - C\left(\frac{1}{\Lambda + \sigma^2}\right) + 2\delta(\epsilon)\right]\right\} \tag{2.78}$$

meaning that $e_{\vec{X}_\ell}(m, \mathbf{s})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small and the rate condition in (2.33) holds.

Now, we consider error probability P_L. Define $\vec{X}_L = (XX_1X_2 \dots X_LSV)$. Let $\mathcal{V}_L$ denote a finite ϵ-dense subset of Gaussian vectors $\vec{X}_L$ with variances at most $1, 1, 1, \dots, 1, \Lambda, \sigma^2$. Thus, P_L can be upper bounded by

$$\sum_{\vec{X}_L \in \mathcal{V}_L} \frac{1}{N} \sum_{m=1}^{N} e_{\vec{X}_L}(m, \mathbf{s}) \tag{2.79}$$

where

$$\begin{aligned} e_{\vec{X}_L}(m, \mathbf{s}) = \mathbb{P}\Big\{&(\mathbf{x}(m), \mathbf{x}(m_1), \dots, \mathbf{x}(m_L), \mathbf{s}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(\vec{X}_L), \\ &\text{for some } (m_1, \dots, m_L) \in \mathscr{S}_L \text{ and } m_\ell \neq m \text{ for all } \ell \in [L]\Big\}. \end{aligned} \tag{2.80}$$

Thus, we need to show that $\frac{1}{N}\sum_{m=1}^{N} e_{\vec{X}_L}(m, \mathbf{s})$ vanishes as $n \to \infty$ for all Gaussian vectors $\vec{X}_L$ that satisfy (2.45)–(2.48) for all $\ell \in [L]$ and

$$\mathbb{E}X_L^2 = 1, \quad (X_1, \dots, X_L) \text{ are independent,} \tag{2.81}$$

where (2.81) follows from (2.34) in which $\mathrm{Cov}(X_iX_j) = 0$ for all $i \neq j \in [L]$.

Observe that if $I(XV; SX_1 \dots X_L) = 0$, then we would have for all $\ell \in [L]$

$$0 = \mathbb{E}X_\ell(X + S + V - X_\ell) \tag{2.82}$$

$$= \mathbb{E}X_\ell(S - X_\ell) \tag{2.83}$$

$$= \mathbb{E}X_\ell S - 1, \tag{2.84}$$

where (2.84) follows from (2.34).

Since $\mathbb{E}X_iX_j = 0$ for all $i, j \in [L]$ and $i \neq j$, the covariance matrix of $(S, X_1, \ldots, X_L)$ is equal to

$$\begin{bmatrix} \mathbb{E}S^2 & \mathbf{1}_L^T \\ \mathbf{1}_L & \mathbf{I}_L \end{bmatrix} \tag{2.85}$$

which has the determinant of $\mathbb{E}S^2 - L$. This determinant should be positive since the covariance matrix $\text{Cov}(S, X_1, \ldots, X_L)$ is positive definite. However, since $\mathbb{E}S^2 \leq \Lambda$, this assumption contradicts the assumption that $\Lambda < L$ in (2.32). Thus, there exists $\eta > 0$ such that

$$\eta \leq I(XV; SX_1 \ldots X_L) = I(XV; X_1 \ldots X_L|S) \tag{2.86}$$

where we have used the fact that $I(XV; S) = 0$.

Now, we may consider two cases $R < \min\{I(X_1; S), \ldots, I(X_L; S)\}$ and $R \geq \min\{I(X_1; S), \ldots, I(X_L; S)\}$. Therefore, using an identical argument as in the cases (a) and (b) for P_ℓ, it follows that $e_{\vec{X}_L}$ is also exponentially vanishing.

2.5 Proof of Lemma 4

In order to prove (2.36), we use our proof in Hosseinigoki and Kosut, (2017) (Lemma 6) for one codebook. Moreover, to obtain (2.37)–(2.40), we apply the corresponding proof of these four equations in Hughes, (1997) (Lemma 1) for Gaussian distributions. Note that Hughes, (1997) focuses on discrete alphabets but the same proofs can be extended to Gaussian distributions by quantization of the set of continuous random variables in the following way.

Let $\mathbf{X}_i$ be Gaussian i.i.d. n-length random vector (codebook) independent from each other with $\text{Var}(X) = 1$. Fix $\mathbf{x} \in \mathcal{T}_\epsilon^{(n)}(X), \mathbf{s} \in \mathscr{S}^n$ and a covariance matrix

$\mathrm{Cov}(X, X_1, \ldots, X_\ell, S) \in \mathcal{V}^{(\ell+2)\times(\ell+2)}$ such that $\mathscr{S}^n$ is a ν-dense subset of $\mathbb{R}^n$ for $\mathbf{s}$ such that $||\mathbf{s}||^2 \le n\Lambda$ and $\mathcal{V}^{(\ell+2)\times(\ell+2)}$ is a ν-dense subset of $\mathbb{R}^{(\ell+2)\times(\ell+2)}$ for positive definite covariance matrices with diagonals at most $(1, 1, \ldots, 1, \Lambda)$.

Using the similar proof in Hughes, (1997) (Lemma 1), we obtain for given $\mathbf{x}, \mathbf{s}$ and covariance matrix $\mathrm{Cov}(X, X_1, \ldots, X_\ell, S)$ that the complement of each event in (2.37)–(2.40) happens with decreasingly doubly exponential probability for sufficiently large n meaning that

$$\begin{aligned}
&\mathbb{P}\big\{\big|\big\{m_1 : (\mathbf{x}, \mathbf{x}(m_1), \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, S)\big\}\big| \le \exp\big\{n\big[|R - I(X_1; XS)|^+ + \delta(\epsilon)\big]\big\}\big\} \\
&\quad < \exp(-\exp(n\sigma(\epsilon))),
\end{aligned} \tag{2.87}$$

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{N}\big|\big\{m : (\mathbf{x}(m), \mathbf{x}(m_1), \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, S) \text{ for some } m_1 \neq m\big\}\big| \le 2\exp\{-n\delta(\epsilon)/2\}\right\} \\
&\quad < \exp[-\exp(n\sigma(\epsilon))], \text{ if } I(X; X_1S) \ge |R - I(X_1; S)|^+ + \delta(\epsilon),
\end{aligned} \tag{2.88}$$

$$\begin{aligned}
&\mathbb{P}\left\{\big|\big\{(m_1, \ldots, m_\ell) : (\mathbf{x}, \mathbf{x}(m_1), \ldots, \mathbf{x}(m_\ell), \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, \ldots, X_\ell, S)\big\}\big| \le \exp\left[n\delta(\epsilon)\right]\right\} \\
&\quad < \exp(-\exp(n\sigma(\epsilon))) \text{ if } R < \min_{k \in \{1, \ldots, \ell\}} I(X_k; S),
\end{aligned} \tag{2.89}$$

$$\begin{aligned}
&\mathbb{P}\Bigg\{\frac{1}{N}\big|\big\{m : (\mathbf{x}(m), \mathbf{x}(m_1), \ldots, \mathbf{x}(m_\ell), \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, \ldots, X_\ell, S) \\
&\qquad\qquad \text{for some } m_1, \ldots, m_\ell \neq m\big\}\big| \le \exp\{-n\delta(\epsilon)/2\}\Bigg\} < \exp(-\exp(n\sigma(\epsilon))) \\
&\text{if } I(X; X_1 \ldots X_\ell S) \ge \delta(\epsilon) \text{ and } R < \min_{k \in \{1, \ldots, \ell\}} I(X_k; S).
\end{aligned} \tag{2.90}$$

Then, in order to complete the proof, since for any fixed ν the cardinality of finite set $\mathscr{S}^n$ is only increasingly exponentially in n and the set $\mathcal{V}^{(\ell+2)\times(\ell+2)}$ is finite along with the doubly decreasing exponential probabilities in (2.87)–(2.90), we derive that with probability approaching to 1, all inequalities in (2.37)–(2.40) hold simultaneously for sufficiently large n. Since these inequalities hold for every element in the finite sets $\mathscr{S}^n$ and $\mathcal{V}^{(\ell+2)\times(\ell+2)}$, then for any vector $\mathbf{s}, \mathbf{x}$ and any given covariance matrix

$\mathrm{Cov}(X, X_1, \ldots, X_\ell, S)$ (with $\|\mathbf{x}\|^2 = n, \|\mathbf{s}\|^2 \leq n\Lambda$) which is not in its corresponding ν-dense subset, there exists a point in the corresponding ν-dense subset that is close enough to it (in its ν distance neighborhood). Now, by using the continuity properties of all sets, we may conclude that (2.37)–(2.40) hold also for any point which is not in the ν-dense subset.

Chapter 3

PACKING LEMMAS FOR GAUSSIAN ADVERSARIAL CHANNELS

In this chapter, we introduce four adversarial packing lemmas as Lemmas 5, 6, 7 and 8, and prove them in separate sections. We propose these lemmas to use them to achieve the lower bounds for the capacity regions of Gaussian arbitrarily-varying multiple-access channels, Gaussian arbitrarily-varying broadcast channels and Gaussian arbitrarily-varying interference channels in Chapters 4, 5 and 6. We also provide Lemma 9 and its proof in Sections 3.6 which is required in the proofs of proposed adversarial packing lemmas. Note that Lemma 7 and 8 are introduced for two codebooks while Lemma 5 and 6 focus on just one codebook. Moreover, Lemma 6 and 7 differ from Lemma 5 and 8 in those they take into account common randomness between the encoder and the decoder for one codebook and two codebooks scenarios, respectively. Lemma 6 and 7 are used to bound the error events with the common message using as common randomness. The advantage of arbitrarily-varying channel coding with common randomness is that it is not susceptible to symmetrization. Thus, in Lemma 6 and 7 there is no requirement that $\Lambda < 1$.

3.1 Packing Lemmas

Lemmas 5, 6, 7 and 8 are proved in Sections 3.2, 3.3, 3.4, 3.5 respectively. Before proceeding to the packing lemmas, we first define the following typical set for Gaussian random variables $X_1, \ldots, X_k$ as:

$$\mathcal{T}_\epsilon^{(n)}(X_1, \ldots, X_k) = \\ \left\{(\mathbf{x}_1, \ldots, \mathbf{x}_k) : \mathbb{E}(X_i X_j) - \epsilon \leq \frac{1}{n}\langle \mathbf{x}_i, \mathbf{x}_j \rangle \leq \mathbb{E}(X_i X_j) + \epsilon \text{ for all } i, j \in [1:k]\right\}. \quad (3.1)$$

Lemma 5 (One-Codebook Adversarial Packing Lemma) *Fix σ^2, $\Lambda \geq 0$, $N = 2^{nR}$. Let $\mathbf{X}, \ldots, \mathbf{X}_N \in \mathbb{R}^n$ be independent zero mean Gaussian random vectors (codebooks) with variance matrices $\mathbf{I}_n$. Let Λ, R satisfy*

$$\Lambda < 1, \quad (3.2)$$

$$R < C\left(\frac{1}{\Lambda + \sigma^2}\right), \quad (3.3)$$

Let $\mathcal{U}$ be the set of pairs $(\mathbf{x}, \mathbf{z}) \in \mathcal{T}_\epsilon^{(n)}(X, Z)$ for some Gaussian pairs (X, Z) where

$$\mathbb{E}X^2 = 1, \quad (X, Z) \text{ are independent.} \quad (3.4)$$

Define

$$p(\mathbf{x}, \ldots, \mathbf{x}_N | \mathbf{w}) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{P}\big\{\exists j \neq i : \|\mathbf{x}_i + \mathbf{w} + \mathbf{V} - \mathbf{x}_j\|^2 \leq \|\mathbf{w} + \mathbf{V}\|^2, \\ (\mathbf{x}_i, \mathbf{w} + \mathbf{V}) \in \mathcal{U}, (\mathbf{x}_j, \mathbf{x}_i + \mathbf{w} + \mathbf{V} - \mathbf{x}_j) \in \mathcal{U}\big\} \quad (3.5)$$

where $\mathbf{V}$ is Gaussian noise distributed as $\mathbf{V} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. There exists $\rho > 0$ such that

$$\lim_{n \to \infty} \mathbb{P}\left[\sup_{\mathbf{w}:||\mathbf{w}||^2 \leq n\Lambda} p(\mathbf{X}, \ldots, \mathbf{X}_N | \mathbf{w}) \geq \exp(-n\rho)\right] = 0. \quad (3.6)$$

Lemma 6 (One-Codebook Adversarial Packing Lemma with Common Randomness) *Fix σ^2, $\Lambda \geq 0$, $N = 2^{nR}$ and $K \geq n^2$. Let $\mathbf{X}_i(k)$ for $i = 1, \ldots, N$, $k = 1, \ldots, K$ be independent zero mean Gaussian random vectors with covariance* *matrix* $\mathbf{I}_n$. *Let* R *satisfy* $R < C(1/(\Lambda + \sigma^2))$. *Define*

$$p_2\left(\mathbf{x}_1(1), \ldots, \mathbf{x}_1(K), \mathbf{x}_2(1), \ldots, \mathbf{x}_2(K), \ldots, \mathbf{x}_N(1), \ldots, \mathbf{x}_N(K)|\mathbf{w}\right) = \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} \mathbb{P}\Big\{\exists j \neq i : \|\mathbf{x}_i(k) + \mathbf{w} + \mathbf{V} - \mathbf{x}_j(k)\|^2 \leq \|\mathbf{w} + \mathbf{V}\|^2, \mathbf{x}_i(k) \in \mathcal{T}_\epsilon^{(n)}(X), \mathbf{x}_j(k) \in \mathcal{T}_\epsilon^{(n)}(X)\Big\} \quad (3.7)$$

where $\mathbf{V}$ *is Gaussian noise distributed as* $\mathbf{V} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$. *There exists* $\rho > 0$ *such that*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{\mathbf{w}:||\mathbf{w}||^2 \leq n\Lambda} p_2(\mathbf{X}_1(1), \ldots, \mathbf{X}_1(K), \ldots, \mathbf{X}_N(1), \ldots, \mathbf{X}_N(K)|\mathbf{w}) \geq \rho\right] = 0. \quad (3.8)$$

Lemma 7 (Two-Codebook Adversarial Packing Lemma with Common Randomness) *Fix* $\theta \in [0, 1]$, σ^2, $\Lambda \geq 0$, $N_1 = 2^{nR_1}$, $N_2 = 2^{nR_2}$ *and* $K \geq n^2$. *Let* $\mathbf{X}_1(k), \ldots, \mathbf{X}_{N_1}(k) \in \mathbb{R}^n$ *and* $\mathbf{Y}_1, \ldots, \mathbf{Y}_{N_2} \in \mathbb{R}^n$ *for* $k = 1, \ldots, K$ *be independent zero mean Gaussian random vectors (codebooks) with covariance matrices* $\theta\,\mathbf{I}_n$ *and* $\bar{\theta}\,\mathbf{I}_n$, *respectively. Let* R_1, R_2 *satisfy*

$$R_1 < C\left(\frac{\theta}{\Lambda + \sigma^2}\right), \quad (3.9)$$

$$R_2 < C\left(\frac{\bar{\theta}}{\Lambda + \sigma^2}\right), \quad (3.10)$$

$$R_1 + R_2 < C\left(\frac{1}{\Lambda + \sigma^2}\right). \quad (3.11)$$

Let $\mathcal{U}$ *be the set of triples* $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, Z)$ *for some Gaussian triple* (X, Y, Z) *where*

$$\mathbb{E}X^2 = \theta, \quad \mathbb{E}Y^2 = \bar{\theta}, \quad (X, Y, Z) \text{ are mutually independent.} \quad (3.12)$$

Define

$$\begin{aligned}&p_1(\mathbf{x}_1(1),\dots,\mathbf{x}_1(K),\mathbf{x}_2(1),\dots,\mathbf{x}_2(K),\dots,\mathbf{x}_{N_1}(1),\dots,\mathbf{x}_{N_1}(K),\mathbf{y}_1,\dots,\mathbf{y}_{N_2}|\mathbf{w}) =\\ &\frac{1}{N_1N_2K}\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{k=1}^{K}\mathbb{P}\Big\{\exists j_1\neq i_1, j_2\neq i_2 : \|\mathbf{x}_{i_1}(k)+\mathbf{y}_{i_2}+\mathbf{w}+\mathbf{V}-\mathbf{x}_{j_1}(k)-\mathbf{y}_{j_2}\|^2\leq\|\mathbf{w}+\mathbf{V}\|^2,\\ &(\mathbf{x}_{i_1}(k),\mathbf{y}_{i_2},\mathbf{w}+\mathbf{V})\in\mathcal{U},(\mathbf{x}_{j_1}(k),\mathbf{y}_{j_2},\mathbf{x}_{i_1}(k)+\mathbf{y}_{i_2}+\mathbf{w}+\mathbf{V}-\mathbf{x}_{j_1}(k)-\mathbf{y}_{j_2})\in\mathcal{U}\Big\}\end{aligned} \tag{3.13}$$

where $\mathbf{V}$ *is Gaussian noise distributed as* $\mathbf{V}\sim\mathcal{N}(\mathbf{0},\sigma^2\mathbf{I}_n)$*. There exists* $\rho>0$ *such that*

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}\Bigg[&\sup_{\mathbf{w}:||\mathbf{w}||^2\leq n\Lambda} p_1\left(\mathbf{X}_1(1),\dots,\mathbf{X}_1(K),\dots,\mathbf{X}_{N_1}(1),\dots,\mathbf{X}_{N_1}(K),\mathbf{Y}_1,\dots,\mathbf{Y}_{N_2}|\mathbf{w}\right)\\ &\geq\exp(-n\rho)\Bigg]=0.\end{aligned} \tag{3.14}$$

Lemma 8 (Two-Codebook Adversarial Packing Lemma) *Fix* $\theta\in[0,1]$*,* σ^2*,* $\Lambda\geq 0$*,* $N_1=2^{nR_1}$ *and* $N_2=2^{nR_2}$*. Let* $\mathbf{X}_1,\dots,\mathbf{X}_{N_1}\in\mathbb{R}^n$ *and* $\mathbf{Y}_1,\dots,\mathbf{Y}_{N_2}\in\mathbb{R}^n$ *be independent zero mean Gaussian random vectors (codebooks) with covariance matrices* $\theta\,\mathbf{I}_n$ *and* $\bar{\theta}\,\mathbf{I}_n$*, respectively. Let* Λ, R_1, R_2 *satisfy*

$$\Lambda<1, \tag{3.15}$$

$$R_1<C\left(\frac{\theta}{\Lambda+\sigma^2}\right), \tag{3.16}$$

$$R_2<C\left(\frac{\bar{\theta}}{\Lambda+\sigma^2}\right), \tag{3.17}$$

$$R_1+R_2<C\left(\frac{1}{\Lambda+\sigma^2}\right). \tag{3.18}$$

Let $\mathcal{U}$ *be the set of triples* $(\mathbf{x},\mathbf{y},\mathbf{z})\in\mathcal{T}_\epsilon^{(n)}(X,Y,Z)$ *for some Gaussian triple* (X,Y,Z) *where*

$$\mathbb{E}X^2=\theta,\quad \mathbb{E}Y^2=\bar{\theta},\quad (X,Y,Z)\ \textit{are mutually independent.} \tag{3.19}$$

Define

$$p_1(\mathbf{x}_1,\ldots,\mathbf{x}_{N_1},\mathbf{y}_1,\ldots,\mathbf{y}_{N_2}|\mathbf{w}) =$$
$$\frac{1}{N_1N_2}\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\mathbb{P}\big\{\exists j_1\neq i_1, j_2\neq i_2 : \|\mathbf{x}_{i_1}+\mathbf{y}_{i_2}+\mathbf{w}+\mathbf{V}-\mathbf{x}_{j_1}-\mathbf{y}_{j_2}\|^2 \leq \|\mathbf{w}+\mathbf{V}\|^2,$$
$$(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w}+\mathbf{V})\in\mathcal{U}, (\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{x}_{i_1}+\mathbf{y}_{i_2}+\mathbf{w}+\mathbf{V}-\mathbf{x}_{j_1}-\mathbf{y}_{j_2})\in\mathcal{U}\big\} \quad (3.20)$$

where $\mathbf{V}$ *is Gaussian noise distributed as* $\mathbf{V}\sim\mathcal{N}(\mathbf{0},\sigma^2\mathbf{I}_n)$*. There exists* $\rho > 0$ *such that*

$$\lim_{n\to\infty}\mathbb{P}\left[\sup_{\mathbf{w}:||\mathbf{w}||^2\leq n\Lambda} p_1(\mathbf{X}_1,\ldots,\mathbf{X}_{N_1},\mathbf{Y}_1,\ldots,\mathbf{Y}_{N_2}|\mathbf{w}) \geq \exp(-n\rho)\right] = 0. \quad (3.21)$$

Since we are dealing with more than one Gaussian codeword in this thesis, we need a new version of Csiszár and Narayan, (1988)(b), Lemma 3 and Csiszár and Narayan, (1991), Lemma 1 not only for Gaussian vectors, but also for multiple codebooks to prove the packing lemmas. It did not appear possible to use the properties derived from these lemmas on each codebook individually; instead, we must prove a new lemma establishing joint properties among more than one codebook. This new lemma, Lemma 9, provides the main properties that the Gaussian codebooks need as part of the proof of Lemma 8 by (3.27)-(3.29). The proof of Lemma 9 is provided in Section 3.6.

Lemma 9 *Fix* $\theta\in[0,1]$*,* $N_1 = 2^{nR_1}$ *and* $N_2 = 2^{nR_2}$*. Given any random variables* X, Y, W*, define the quantity*

$$J_{X;Y;W}(R_1,R_2) =$$
$$\max\big\{0, R_1 - I(X;WY), R_2 - I(Y;WX), R_1+R_2 - I(XY;W) - I(X;Y)\big\}. \quad (3.22)$$

Let $\mathbf{X}_{i_1}$ and $\mathbf{Y}_{i_2}$ be Gaussian i.i.d. n-length random vectors (codebooks) independent from each other with zero mean and $\mathrm{Cov}(\mathbf{X}_{i_1}) = \theta\,\mathbf{I}_n$, $\mathrm{Cov}(\mathbf{Y}_{i_2}) = \bar{\theta}\,\mathbf{I}_n$ where $i_1 \in \{1, 2, \ldots, N_1\}$ and $i_2 \in \{1, 2, \ldots, N_2\}$. With probability approaching 1 as $n \to \infty$, they satisfy the following, for any $\mathbf{x}, \mathbf{y}, \mathbf{w}$ where $\|\mathbf{w}\|^2 \leq n\Lambda$ and any zero mean jointly Gaussian random vector (X, Y, X', Y', W) with positive definite covariance matrices with diagonals at most $(\theta, \bar{\theta}, \theta, \bar{\theta}, \Lambda)$. (3.82)

$$\frac{1}{N_1N_2}\left|\left\{(i_1,i_2) : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \notin \bigcup_{\substack{(X,Y,W) \textit{ mutually independent:}\\ EX^2=\theta, EY^2=\bar{\theta}, EW^2 \leq \Lambda}} \mathcal{T}_\epsilon^{(n)}(X,Y,W)\right\}\right| \leq \exp(-n\delta(\epsilon)). \tag{3.23}$$

$$\left|\left\{j : (\mathbf{x}, \mathbf{x}_j, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, X', W)\right\}\right| \leq \exp\left\{n\left[|R - I(X'; XW)|^+ + \delta(\epsilon)\right]\right\} \tag{3.24}$$

$$\frac{1}{N}\left|\left\{i : (\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, X', W) \textit{ for some } j \neq i\right\}\right|$$
$$\leq \exp\{n(|R - I(X'; W)|^+ - I(X; X'W) + \delta(\epsilon))\} \tag{3.25}$$

$$\frac{1}{N}\left|\left\{i : (\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, X', W) \textit{ for some } j \neq i\right\}\right| \leq 2\exp(-n\delta(\epsilon)/2)$$
$$\textit{if } |R - I(X'; W)|^+ \leq I(X; X'W) - 2\delta(\epsilon) \tag{3.26}$$

$$\left|\left\{(i_1, i_2) : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W)\right\}\right| \leq \exp\left\{n\left[J_{X';Y';W}(R_1, R_2) + \delta(\epsilon)\right]\right\} \tag{3.27}$$

$$\left|\left\{(i_1, i_2) : (\mathbf{x}, \mathbf{y}, \mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W)\right\}\right|$$
$$\leq \exp\left\{n\left[J_{X';Y';XYW}(R_1, R_2) + \delta(\epsilon)\right]\right\} \tag{3.28}$$

$$\frac{1}{N_1N_2}\left|\left\{(i_1, i_2) : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W) \textit{ for some } j_1 \neq i_1, j_2 \neq i_2\right\}\right|$$
$$\leq 8\exp\{-n\delta(\epsilon)/4\} \textit{ if } J_{X';Y';W}(R_1, R_2) \leq I(XY; X'Y'W) - 2\delta(\epsilon). \tag{3.29}$$

3.2 Proof of Lemma 5

Since Lemma 5 is a special case of Lemma 8 for one codebook. Note that Csiszár and Narayan, (1991) find the capacity of Gaussian AVC channels by other tools rather than Lemma 5. We are the first one to introduce the adversarial packing lemma to prove the Gaussian AVC and its network setting.

We apply the single-codebook results of Lemma 9 to assume the codebook satisfies the single-codebook version of (3.23)–(3.29). To prove (3.6), first note that by the single-codebook version of (3.23), with high probability $(\mathbf{x}_i, \mathbf{w}) \in \mathcal{T}_{\epsilon'}^{(n)}(X, W)$ where (X, W) are independent, and

$$\mathbb{E}X^2 = 1, \quad \mathbb{E}W^2 \leq \Lambda. \tag{3.30}$$

Thus, by the conditional typicality lemma 2, for every $\epsilon > \epsilon'$ with high probability $(\mathbf{x}_i, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(X, W, V)$ where (X, W, V) are mutually independent, and $\mathbb{E}V^2 = \sigma^2$. This implies that $(\mathbf{x}_i, \mathbf{w} + \mathbf{V}) \in \mathcal{U}$, and also that

$$\|\mathbf{w} + \mathbf{V}\|^2 \leq n(\Lambda + \sigma^2 + \epsilon). \tag{3.31}$$

We use shorthand $\vec{X} = (XX'WV)$. For $i, \mathbf{w}$ and any Gaussian distribution on $\vec{X}$, define

$$e_{\vec{X}}(i, \mathbf{w}) = \mathbb{P}\Big\{(\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(\vec{X}) \text{ for some } j \neq i\Big\}. \tag{3.32}$$

We need to show that for some $\rho > 0$,

$$\frac{1}{N} \sum_i e_{\vec{X}}(i, \mathbf{w}) \leq \exp(-n\rho) \tag{3.33}$$

for all $\vec{X}$ where

$$(X,W,V) \text{ are mutually independent,} \tag{3.34}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = 1, \quad \mathbb{E}V^2 = \sigma^2 \tag{3.35}$$

$$\mathbb{E}W^2 \leq \Lambda, \quad \mathbb{E}(X+W+V-X')^2 \leq \Lambda + \sigma^2 \tag{3.36}$$

$$(X', X+W+V-X') \text{ are independent.} \tag{3.37}$$

Observe that if $I(XV;WX')=0$, then we would have

$$\Lambda + \sigma^2 \geq \mathbb{E}(X+W+V-X')^2 \tag{3.38}$$

$$= \mathbb{E}(X+V)^2 + \mathbb{E}(W-X')^2 \tag{3.39}$$

$$\geq 1+\sigma^2. \tag{3.40}$$

But this cannot happen since $\Lambda < 1$ by (3.2). Thus, there exists $\eta > 0$ where

$$I(XV;WX') \geq \eta. \tag{3.41}$$

Recalling that $I(XV;W)=0$, this implies

$$I(XV;X'|W) \geq \eta. \tag{3.42}$$

Also, by (3.26), we may restrict ourselves to distributions where

$$R - I(X';W) \geq I(X;X'W) - 2\delta(\epsilon). \tag{3.43}$$

We may now write, for any i and any $\mathbf{w}$

$$e_{\vec{X}}(i,\mathbf{w}) \leq \sum_{j:(\mathbf{x}_i,\mathbf{x}_j,\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}} \mathbb{P}\Big\{(\mathbf{x}_i,\mathbf{x}_j,\mathbf{w},\mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}\Big\} \tag{3.44}$$

$$\leq \exp\Big\{n\Big[|R - I(X';XW)|^+ - I(V;X'|XW) + \delta(\epsilon)\Big]\Big\}, \tag{3.45}$$

where in (3.45) we have applied (3.24), the joint typicality lemma 3, and the fact that $I(V;XW) = 0$.

We consider two cases.

Case (a): $R - I(X';W) \leq 0$. Note that $R - I(X';XW) \leq R - I(X';W)$ so in this case we also have $R - I(X';XW) \leq 0$. By (3.43), $I(X;X'W) \leq 2\delta(\epsilon)$. Thus, from (3.42)

$$\eta \leq I(XV;X'|W) \tag{3.46}$$

$$= I(X;X'|W) + I(V;X'|XW) \tag{3.47}$$

$$\leq 2\delta(\epsilon) + I(V;X'|XW). \tag{3.48}$$

From (3.45), we have

$$e_{\vec{X}}(i,\mathbf{w}) \leq \exp\{n[-I(V;X'|XW) + \delta(\epsilon)]\} \tag{3.49}$$

$$\leq \exp\{n[-\eta + 3\delta(\epsilon)]\}. \tag{3.50}$$

This vanishes exponentially fast if $\delta(\epsilon)$ is sufficiently small.

Case (b): $R - I(X';W) > 0$. By (3.43), we have

$$-2\delta(\epsilon) \leq R - I(X';W) - I(X;X'W). \tag{3.51}$$

Note that

$$I(X';W) + I(X;X'W) \geq I(X';W) + I(X;X'|W) \tag{3.52}$$

$$= I(X';XW). \tag{3.53}$$

Thus

$$-2\delta(\epsilon) \leq R - I(X';W) - I(X;X'W) \leq R - I(X';XW). \tag{3.54}$$

By (3.45), we have

$$\frac{1}{n}\log e_{\vec{X}}(i,\mathbf{w}) \leq |R - I(X';XW)|^+ - I(V;X'|XW) + \delta(\epsilon) \tag{3.55}$$

$$\leq R - I(X';XW) - I(V;X'|XW) + 3\delta(\epsilon) \tag{3.56}$$

$$\leq R - I(X';XWV) + 3\delta(\epsilon). \tag{3.57}$$

Let $Z = X + W + V - X'$. Recalling that (X', Z) are mutually independent and $\mathbb{E}Z^2 \leq \Lambda + \sigma^2$, we have

$$I(X';XWV) \geq I(X';X+W+V) \tag{3.58}$$

$$= I(X';X'+Z) \tag{3.59}$$

$$= h(X'+Z) - h(X'+Z|X') \tag{3.60}$$

$$= \frac{1}{2}\log 2\pi e(1+\mathbb{E}Z^2) - h(Z|X') \tag{3.61}$$

$$= \frac{1}{2}\log 2\pi e(1+\mathbb{E}Z^2) - \frac{1}{2}\log 2\pi e\mathbb{E}Z^2 \tag{3.62}$$

$$= C\left(\frac{1}{\mathbb{E}Z^2}\right) \tag{3.63}$$

$$\geq C\left(\frac{1}{\Lambda+\sigma^2}\right). \tag{3.64}$$

Thus

$$e_{\vec{X}}(i,\mathbf{w}) \leq \exp\left\{n\left[R - C\left(\frac{1}{\Lambda+\sigma^2}\right) + 3\delta(\epsilon)\right]\right\}. \tag{3.65}$$

Therefore, $e_{\vec{X}}(i,\mathbf{w})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small and (3.3) holds.

3.3 Proof of Lemma 6

We prove this lemma using a random code reduction, as in Csiszár and Körner, (2011), Lemma 12.8. We first show that a Gaussian codebook independent of the jammer's signal achieves small probability of error, and then we show that a finite number of deterministic codebooks achieve essentially the same probability.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_N$ be Gaussian random vectors with zero mean and covariance $\mathbf{I}_n$. We will prove that, for any $i \in [N]$ and any $\mathbf{w}$ such that $\|\mathbf{w}\|^2 \leq n\Lambda$

$$\mathbb{P}\big\{\exists j \neq i : \|\mathbf{X}_i + \mathbf{w} + \mathbf{V} - \mathbf{X}_j\|^2 \leq \|\mathbf{w} + \mathbf{V}\|^2\big\} \to 0 \tag{3.66}$$

as $n \to \infty$, where $\mathbf{V} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$. To prove this, we adopt the basic approach of Lapidoth, (1996). In particular, let $\mathbf{Z} = \mathbf{w} + \mathbf{V}$, and let U be a unitary matrix that maps $\mathbf{Z}$ to $(\|\mathbf{Z}\|, 0, \ldots, 0)$. Then we may write

$$\mathbb{P}\big\{\exists j \neq i : \|\mathbf{X}_i + \mathbf{Z} - \mathbf{X}_j\|^2 \leq \|\mathbf{Z}\|^2\big\} = \mathbb{P}\big\{\exists j \neq i : \|U\mathbf{X}_i + U\mathbf{Z} - U\mathbf{X}_j\|^2 \leq \|U\mathbf{Z}\|^2\big\} \tag{3.67}$$

$$= \mathbb{P}\big\{\exists j \neq i : \|\mathbf{X}_i + U\mathbf{Z} - \mathbf{X}_j\|^2 \leq \|U\mathbf{Z}\|^2\big\} \tag{3.68}$$

where (3.68) follows from the spherical symmetry of the codebook distribution. Now if we define, for any $\Sigma > 0$,

$$e(\Sigma) = \mathbb{P}\{\exists j \neq i : \|\mathbf{X}_i + (\sqrt{n\Sigma}, 0, \ldots, 0) - \mathbf{X}_j\|^2 \leq n\Sigma\} \tag{3.69}$$

then the probability in (3.68) may be written as $\mathbb{E}e(\|\mathbf{Z}\|^2)$. Note that for any δ,

$$\lim_{n\to\infty} \mathbb{P}\big\{\big|\|\mathbf{Z}\|^2 - n(\sigma^2 + \Lambda)\big| > n\delta\big\} \to 0. \tag{3.70}$$

Moreover, $e(\Sigma)$ is non-decreasing in Σ. Thus, for any $\delta > 0$, if we let $\tilde{\mathbf{V}} \sim \mathcal{N}(0, \sigma^2 + \Lambda + \delta)$, for sufficiently large n we have $\mathbb{E}e(\|\mathbf{Z}\|^2) \leq \mathbb{E}e(\|\tilde{\mathbf{V}}\|^2)$. Now, $\mathbb{E}e(\|\tilde{\mathbf{V}}\|^2)$ is simply the probability of error for a Gaussian channel with noise variance $\sigma^2 + \Lambda + \delta$. Since Gaussian codebooks achieve capacity for Gaussian channels with minimum distance decoding, this quantity vanishes with n as long as

$$R < C\left(\frac{1}{\sigma^2 + \Lambda + \delta}\right) \tag{3.71}$$

which holds for small enough δ by the assumption that $R < C(\frac{1}{\sigma^2+\Lambda})$. This proves (3.66).

Now let $\mathbf{X}_i(k)$ for $i = 1, \ldots, N$ and $k = 1, \ldots, K$ be independent Gaussian vectors with zero mean and covariance $\mathbf{I}_n$. For any $i \in [N]$, $k \in [K]$, and $\mathbf{w}$ such that $\|\mathbf{w}\|^2 \le n\Lambda$, let

$$\begin{aligned}\mathcal{E}(k, i, \mathbf{w}) = \mathbb{P}\Big\{\exists j \neq i : \|\mathbf{X}_i(k) + \mathbf{w} + \mathbf{V} - \mathbf{X}_j(k)\|^2 \le \|\mathbf{w} + \mathbf{V}\|^2,\\ \mathbf{X}_i(k) \in \mathcal{T}_{\epsilon'}^{(n)}, \mathbf{X}_j(k) \in \mathcal{T}_{\epsilon'}^{(n)} \Big| \mathbf{X}_1(k), \ldots, \mathbf{X}_N(k)\Big\}. \quad (3.72)\end{aligned}$$

To prove the lemma, we need to show that, for any $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left\{ \bigcup_{\mathbf{w}:\|\mathbf{w}\|^2 \le n\Lambda} \left\{ \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} \mathcal{E}(k, i, \mathbf{w}) > \epsilon \right\} \right\} \to 0. \quad (3.73)$$

From (3.66), we know that for any $\delta > 0$ and sufficiently large n for any $k, i, \mathbf{w}$, we have $\mathbb{E}\mathcal{E}(k, i, \mathbf{w}) \le \delta$ (The conditions on $\mathbf{X}_i$ and $\mathbf{X}_j$ only decrease the probability.) Thus, for fixed i and $\mathbf{w}$ we have

$$\begin{aligned}\mathbb{P}\left\{ \frac{1}{K} \sum_{k=1}^{K} \mathcal{E}(k, i, \mathbf{w}) > \epsilon/2 \right\} &= \mathbb{P}\left\{ 2^{\sum_{k=1}^{K} \mathcal{E}(k,i,\mathbf{w})} > 2^{K\epsilon/2} \right\} && (3.74)\\ &\le 2^{-K\epsilon/2} \prod_{k=1}^{K} \mathbb{E}2^{\mathcal{E}(k,i,\mathbf{w})} && (3.75)\\ &\le 2^{-K\epsilon/2}(1 + \mathbb{E}\mathcal{E}(1, i, \mathbf{w}))^K && (3.76)\\ &= 2^{-K(\epsilon/2 - \log(1+\delta))} && (3.77)\end{aligned}$$

where (3.75) holds by Markov's inequality, and (3.76) holds since $2^t \le 1 + t$ if $t \in [0, 1]$. Thus, if we let $\mathbf{w}_1, \ldots, \mathbf{w}_L$ be any finite set of vectors with norm at most $\sqrt{n\Lambda}$, we may apply the union bound to find

$$\mathbb{P}\left\{ \bigcup_{\substack{l\in[L]\\ i\in[N]}} \left\{ \frac{1}{K} \sum_{k=1}^{K} \mathcal{E}(k, i, \mathbf{w}_l) > \frac{\epsilon}{2} \right\} \right\} \le LN2^{-K(\frac{\epsilon}{2} - \log(1+\delta))}. \quad (3.78)$$

In particular, let $\mathbf{w}_1, \ldots, \mathbf{w}_L$ be a ν-dense subset of points in the sphere of radius $\sqrt{n\Lambda}$. There exists such a set with $L = 2^{n\rho}$ for some ρ. Since $\mathcal{E}(k, i, \mathbf{w})$ is continuous in $\mathbf{w}$, for sufficiently small ν, if the probability of error for all $\mathbf{w}_l$ is at most $\epsilon/2$, then the probability of error for all $\mathbf{w}$ is at most ϵ. Thus we may bound the probability in (3.73) by

$$2^{n\rho} 2^{nR} 2^{-K(\epsilon/2 - \log(1+\delta))}.$$

As long as δ is small enough so that $\epsilon/2 - \log(1+\delta) > 0$ and $K/n \to \infty$, this probability vanishes in n.

3.4 Proof of Lemma 7

We prove this lemma using a random code reduction, as in Csiszár and Körner, (2011), Lemma 12.8. We first show that a Gaussian codebook independent of the jammer's signal achieves small probability of error, and then we show that a finite number of deterministic codebooks achieve essentially the same probability.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_{N_1}$ be Gaussian random vectors with zero mean and covariance $\theta \mathbf{I}_n$ for all $\mathbf{y}_{i_2}$ with $\|\mathbf{y}_{i_2}\|^2 = \bar{\theta}$. We will prove that, for any $i_1 \in [N_1]$ and any $\mathbf{w}$ such that $\|\mathbf{w}\|^2 \leq n\Lambda$

$$\frac{1}{N_2} \sum_{i_2=1}^{N_2} \mathbb{P}_{XV} \left\{ \exists j_1 \neq i_1, j_2 \neq i_2 : \|\mathbf{X}_{i_1} + \mathbf{y}_{i_2} + \mathbf{w} + \mathbf{V} - \mathbf{X}_{j_1} - \mathbf{y}_{j_2}\|^2 \leq \|\mathbf{w} + \mathbf{V}\|^2 \right\} \to 0 \tag{3.79}$$

as $n \to \infty$, where $\mathbf{V} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$.

To prove this, we apply Lemma 9 to assume the two codebooks satisfy (3.23)–(3.29). To prove (3.21), first note that by (3.23), with high probability $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_{\epsilon'}^{(n)}(X, Y, W)$ where (X, Y, W) are mutually independent, and

$$\mathbb{E}X^2 = \theta, \quad \mathbb{E}Y^2 = \bar{\theta}, \quad \mathbb{E}W^2 \leq \Lambda. \tag{3.80}$$

Thus, by the conditional typicality lemma 2, for every $\epsilon > \epsilon'$ with high probability $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, W, V)$ where (X, Y, W, V) are mutually independent, and $\mathbb{E}V^2 = \sigma^2$. This implies that $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w} + \mathbf{V}) \in \mathcal{U}$, and also that

$$\|\mathbf{w} + \mathbf{V}\|^2 \leq n(\Lambda + \sigma^2 + \epsilon). \tag{3.81}$$

We use shorthand $\vec{X} = (XYX'Y'WV)$. For $i_1, i_2, \mathbf{w}$ and any Gaussian distribution on $\vec{X}$, define

$$\begin{aligned}
&e_{\vec{X}}(i_1, i_2, \mathbf{w}) \\
&= \frac{1}{N_2} \sum_{i_2=1}^{N_2} \mathbb{P}_{XV}\Big\{\exists j_1 \neq i_1, j_2 \neq i_2 : \|\mathbf{X}_{i_1} + \mathbf{y}_{i_2} + \mathbf{w} + \mathbf{V} - \mathbf{X}_{j_1} + \mathbf{y}_{j_2}\|^2 \leq \|\mathbf{w} + \mathbf{V}\|^2, \\
&\quad (\mathbf{X}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w} + \mathbf{V}) \in \mathcal{U}, (\mathbf{X}_{j_1}, \mathbf{y}_{j_2}, \mathbf{X}_{i_1} + \mathbf{y}_{i_2} + \mathbf{w} + \mathbf{V} - \mathbf{X}_{j_1} - \mathbf{y}_{j_2}) \in \mathcal{U}\Big\} &(3.82)\\
&\leq \frac{1}{N_2} \sum_{i_2=1}^{N_2} \mathbb{P}_{XV}\Big\{\exists j_1 \neq i_1, j_2 \neq i_2 : (\mathbf{X}_{i_1}, \mathbf{y}_{i_2}, \mathbf{X}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W, V)\Big\} \\
& &(3.83)\\
&= \frac{1}{N_2} \sum_{i_2=1}^{N_2} \sum_{j_1 \neq i_1} \mathbb{P}_{XV}\Big\{\exists j_2 \neq i_2 : (\mathbf{X}_{i_1}, \mathbf{y}_{i_2}, \mathbf{X}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W, V)\Big\} \\
& &(3.84)\\
&= \sum_{j_1 \neq i_1} \frac{1}{N_2} \sum_{\substack{i_2:\ \exists j_2 \neq i_2 \\ (\mathbf{y}_{i_2}, \mathbf{y}_{j_2}, \mathbf{w}) \mathcal{T}_\epsilon^{(n)}}} \mathbb{P}_{XV}\Big\{(\mathbf{X}_{i_1}, \mathbf{y}_{i_2}, \mathbf{X}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W, V)\Big\} \\
& &(3.85)\\
&\leq \sum_{j_1 \neq i_1} \exp\{n(|R_2 - I(Y'; W)|^+ - I(Y; Y'W) + 2\delta(\epsilon) - I(XX'V; YY'W))\} &(3.86)\\
&\leq \exp\{n(R_1 + |R_2 - I(Y'; W)|^+ - I(Y; Y'W) - I(XX'V; YY'W) + 2\delta(\epsilon))\} &(3.87)
\end{aligned}$$

By Lemma (9) and its proof in (3.194).

We need to show that for some $\rho > 0$,

$$e_{\vec{X}}(i_1, i_2, \mathbf{w}) \leq \exp(-n\rho) \tag{3.88}$$

for all $\vec{X}$ where

$$(X, X', Y, W, V) \text{ are mutually independent,} \tag{3.89}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = \theta, \quad \mathbb{E}Y^2 = \mathbb{E}Y'^2 = \bar{\theta}, \quad \mathbb{E}V^2 = \sigma^2 \tag{3.90}$$

$$\mathbb{E}W^2 \leq \Lambda, \quad \mathbb{E}(X + Y + W + V - X' - Y')^2 \leq \Lambda + \sigma^2 \tag{3.91}$$

$$(X', Y', X + Y + W + V - X' - Y') \text{ are mutually independent.} \tag{3.92}$$

Case(a): If $R_2 \leq I(Y'; W)$ meaning that $|R_2 - I(Y'; W)|^+ = 0$ then by assuming $I(X'; XV) = 0$ we have

$$
\begin{aligned}
&e_{\vec{X}}(i_1, i_2, \mathbf{w}) \\
&\leq \exp\{n(R_1 - I(Y; Y'W) - I(XX'V; YY'W) + 2\delta(\epsilon))\} && (3.93) \\
&= \exp\{n(R_1 - I(Y; Y'W) - I(XV; YY'W) - I(X'; YY'W|XV) + 2\delta(\epsilon))\} && (3.94) \\
&= \exp\{n(R_1 - I(Y; Y'W) - I(XV; YY'W) - I(X'; YY'W|XV) - I(X'; XV) + 2\delta(\epsilon))\} \\
& && (3.95) \\
&= \exp\{n(R_1 - I(Y; Y'W) - I(XV; YY'W) - I(X'; YY'WXV) + 2\delta(\epsilon))\} && (3.96) \\
&\leq \exp\{n(R_1 - I(X'; YY'WXV) + 2\delta(\epsilon))\} && (3.97)
\end{aligned}
$$

Let $Z = X+Y+W+V-X'-Y'$. Recalling that (X', Y', Z) are mutually independent and $\mathbb{E}Z^2 \leq \Lambda + \sigma^2$, we have

$$\begin{aligned} I(X'; XYWY'V) &\geq I(X'; X+Y+W+V-Y') && (3.98)\\ &= I(X'; X'+Z) && (3.99)\\ &= h(X'+Z) - h(X'+Z|X') && (3.100)\\ &= \frac{1}{2}\log 2\pi e(\theta + \mathbb{E}Z^2) - h(Z|X') && (3.101)\\ &= \frac{1}{2}\log 2\pi e(\theta + \mathbb{E}Z^2) - \frac{1}{2}\log 2\pi e\mathbb{E}Z^2 && (3.102)\\ &= C\left(\frac{\theta}{\mathbb{E}Z^2}\right) && (3.103)\\ &\geq C\left(\frac{\theta}{\Lambda+\sigma^2}\right). && (3.104) \end{aligned}$$

Therefore,

$$e_{\vec{X}}(i_1, i_2, \mathbf{w}) \leq \exp\left\{n\left(R_1 - C\left(\frac{\theta}{\Lambda+\sigma^2}\right) + 2\delta(\epsilon)\right)\right\}, \quad (3.105)$$

and $e_{\vec{X}}(i_1, i_2, \mathbf{w})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small and $R_1 < C\left(\frac{\theta}{\Lambda+\sigma^2}\right)$.

Case(b): If $R_2 > I(Y';W)$ meaning that $|R_2 - I(Y';W)|^+ = R_2 - I(Y';W)$ then by assuming $I(X';XV) = 0$ we have

$$\begin{aligned}
&e_{\vec{X}}(i_1, i_2, \mathbf{w}) \\
&\leq \exp\{n(R_1 + R_2 - I(Y';W) - I(Y;Y'W) - I(XX'V;YY'W) + 2\delta(\epsilon))\} && (3.106) \\
&= \exp\{n(R_1 + R_2 - I(Y';W) - I(Y;Y'|W) - I(Y;W) - I(XX'V;YY'W) + 2\delta(\epsilon))\} && (3.107) \\
&= \exp\{n(R_1 + R_2 - I(Y';YW) - I(XX'V;YY'W) + 2\delta(\epsilon))\} && (3.108) \\
&= \exp\{n(R_1 + R_2 - I(Y';YW) - I(X';YY'W) - I(XV;YY'W|X') + 2\delta(\epsilon))\} && (3.109) \\
&= \exp\{n(R_1 + R_2 - I(Y';YW) - I(X';Y') - I(X';YW|Y') - I(XV;YY'WX') \\
&\qquad + I(XV;X') + 2\delta(\epsilon))\} && (3.110) \\
&= \exp\{n(R_1 + R_2 - I(Y'X';YW) - I(XV;YW) - I(XV;X'Y'|YW) + 2\delta(\epsilon))\} && (3.111) \\
&= \exp\{n(R_1 + R_2 - I(Y'X';YW) - I(X'Y';XVYW) + I(X'Y';YW) + 2\delta(\epsilon))\} && (3.112) \\
&= \exp\{n(R_1 + R_2 - I(X'Y';XVYW) + 2\delta(\epsilon))\} && (3.113)
\end{aligned}$$

Moreover,

$$\begin{aligned}
I(X'Y';XYWV) &\geq I(X'Y';Z + X' + Y') && (3.114) \\
&= h(Z + X' + Y') - h(Z) && (3.115) \\
&= C\left(\frac{1}{\mathbb{E}Z^2}\right) && (3.116) \\
&\geq C\left(\frac{1}{\Lambda + \sigma^2}\right). && (3.117)
\end{aligned}$$

Thus,

$$e_{\vec{X}}(i_1, i_2, \mathbf{w}) \leq \exp\left\{n\left(R_1 + R_2 - C\left(\frac{1}{\Lambda + \sigma^2}\right)\right) + 2\delta(\epsilon)\right\}. \tag{3.118}$$

Therefore, $e_{\vec{X}}(i_1, i_2, \mathbf{w})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small and (3.16)–(3.18) hold.

Note that in the following if $I(X'; XYWV) = 0$ then $I(Y'; X'XYVW) = I(X'Y'; XYVW)$.

$$\begin{aligned} I(Y'; X'XYVW) &= I(Y'; X') + I(Y'; XYVW|X') && (3.119) \\ &= I(X'Y'; XYVW) - I(X'; XYVW). && (3.120) \end{aligned}$$

This proves (3.79).

Now, let $\mathbf{X}_{i_1}(k)$ for $i_1 = 1, \ldots, N_1$ and $k = 1, \ldots, K$ be independent Gaussian vectors with zero mean and covariance $\theta \mathbf{I}_n$. For any $i_1 \in [N_1]$, $k \in [K]$, and $\mathbf{w}$ such that $\|\mathbf{w}\|^2 \leq n\Lambda$, let

$$\begin{aligned} \mathcal{E}(k, i_1, i_2, \mathbf{w}) = \frac{1}{N_2} \sum_{i_2=1}^{N_2} \mathbb{P}_V \Big\{ &\exists j_1 \neq i_1, j_2 \neq i_2 : \\ \|\mathbf{X}_{i_1}(k) + \mathbf{y}_{i_2} + \mathbf{w} + \mathbf{V} - \mathbf{X}_{j_1}(k) + \mathbf{y}_{j_2}\|^2 &\leq \|\mathbf{w} + \mathbf{V}\|^2, (\mathbf{X}_{i_1}(k), \mathbf{y}_{i_2}, \mathbf{w} + \mathbf{V}) \in \mathcal{U}, \\ (\mathbf{X}_{j_1}(k), \mathbf{y}_{j_2}, \mathbf{X}_{i_1}(k) + \mathbf{y}_{i_2} + \mathbf{w} + \mathbf{V} - \mathbf{X}_{j_1}(k) - \mathbf{y}_{j_2}) &\in \mathcal{U} \Big| \mathbf{X}_1(k), \ldots, \mathbf{X}_{N_1}(k) \Big\}. \end{aligned} \tag{3.121}$$

To prove the lemma, we need to show that, for any $\epsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}\left\{ \bigcup_{\mathbf{w}: \|\mathbf{w}\|^2 \leq n\Lambda} \left\{ \frac{1}{N_1 K} \sum_{i_1=1}^{N_1} \sum_{k=1}^{K} \mathcal{E}(k, i_1, i_2, \mathbf{w}) > \epsilon \right\} \right\} \to 0. \tag{3.122}$$

From (3.66), we know that for any $\delta > 0$ and sufficiently large n for any $k, i_1, i_2, \mathbf{w}$, we have $\mathbb{E}_X \mathcal{E}(k, i_1, i_2, \mathbf{w}) \leq \delta$ (The conditions on $\mathbf{X}_{i_1}$ and $\mathbf{X}_{j_1}$ only decrease the probability.) Thus, for fixed i_1 and $\mathbf{w}$ we have

$$\mathbb{P}\left\{\frac{1}{K}\sum_{k=1}^{K}\mathcal{E}(k,i_1,i_2,\mathbf{w})>\epsilon/2\right\}=\mathbb{P}\left\{2^{\sum_{k=1}^{K}\mathcal{E}(k,i_1,i_2,\mathbf{w})}>2^{K\epsilon/2}\right\} \tag{3.123}$$

$$\leq 2^{-K\epsilon/2}\prod_{k=1}^{K}\mathbb{E}2^{\mathcal{E}(k,i_1,i_2,\mathbf{w})} \tag{3.124}$$

$$\leq 2^{-K\epsilon/2}(1+\mathbb{E}\mathcal{E}(1,i_1,i_2,\mathbf{w}))^K \tag{3.125}$$

$$= 2^{-K(\epsilon/2-\log(1+\delta))} \tag{3.126}$$

where (3.124) holds by Markov's inequality, and (3.125) holds since $2^t \leq 1+t$ if $t\in[0,1]$. Thus, if we let $\mathbf{w}_1,\dots,\mathbf{w}_L$ be any finite set of vectors with norm at most $\sqrt{n\Lambda}$, we may apply the union bound to find

$$\mathbb{P}\left\{\bigcup_{\substack{l\in[L]\\ i_1\in[N_1]}}\left\{\frac{1}{K}\sum_{k=1}^{K}\mathcal{E}(k,i_1,i_2,\mathbf{w}_l)>\frac{\epsilon}{2}\right\}\right\}\leq LN_1 2^{-K(\frac{\epsilon}{2}-\log(1+\delta))}. \tag{3.127}$$

In particular, let $\mathbf{w}_1,\dots,\mathbf{w}_L$ be a ν-dense subset of points in the sphere of radius $\sqrt{n\Lambda}$. There exists such a set with $L=2^{n\rho}$ for some ρ. Since $\mathcal{E}(k,i_1,i_2,\mathbf{w})$ is continuous in $\mathbf{w}$, for sufficiently small ν, if the probability of error for all $\mathbf{w}_l$ is at most $\epsilon/2$, then the probability of error for all $\mathbf{w}$ is at most ϵ. Thus we may bound the probability in (3.122) by

$$2^{n\rho}2^{nR_1}2^{-K(\epsilon/2-\log(1+\delta))}.$$

As long as δ is small enough so that $\epsilon/2-\log(1+\delta)>0$ and $K/n\to\infty$, this probability vanishes in n.

3.5 Proof of Lemma 8

We apply Lemma 9 to assume the two codebooks satisfy (3.23)–(3.29). To prove (3.21), first note that by (3.23), with high probability $(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\in\mathcal{T}_{\epsilon'}^{(n)}(X,Y,W)$

where (X, Y, W) are mutually independent, and

$$\mathbb{E}X^2 = \theta, \quad \mathbb{E}Y^2 = \bar{\theta}, \quad \mathbb{E}W^2 \le \Lambda. \tag{3.128}$$

Thus, by the conditional typicality lemma 2, for every $\epsilon > \epsilon'$ with high probability $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, W, V)$ where (X, Y, W, V) are mutually independent, and $\mathbb{E}V^2 = \sigma^2$. This implies that $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w} + \mathbf{V}) \in \mathcal{U}$, and also that

$$\|\mathbf{w} + \mathbf{V}\|^2 \le n(\Lambda + \sigma^2 + \epsilon). \tag{3.129}$$

We use shorthand $\vec{X} = (XYX'Y'WV)$. For $i_1, i_2, \mathbf{w}$ and any Gaussian distribution on $\vec{X}$, define

$$e_{\vec{X}}(i_1, i_2, \mathbf{w}) = \mathbb{P}\Big\{(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(\vec{X}) \text{ for some } j_1 \neq i_1, j_2 \neq i_2\Big\}. \tag{3.130}$$

We need to show that for some $\delta > 0$,

$$\frac{1}{N_1 N_2} \sum_{i_1, i_2} e_{\vec{X}}(i_1, i_2, \mathbf{w}) \le \exp(-n\rho) \tag{3.131}$$

for all $\vec{X}$ where

$$(X, Y, W, V) \text{ are mutually independent,} \tag{3.132}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = \theta, \quad \mathbb{E}Y^2 = \mathbb{E}Y'^2 = \bar{\theta}, \quad \mathbb{E}V^2 = \sigma^2 \tag{3.133}$$

$$\mathbb{E}W^2 \le \Lambda, \quad \mathbb{E}(X + Y + W + V - X' - Y')^2 \le \Lambda + \sigma^2 \tag{3.134}$$

$$(X', Y', X + Y + W + V - X' - Y') \text{ are mutually independent.} \tag{3.135}$$

Observe that if $I(XYV; WX'Y') = 0$, then we would have

$$\Lambda + \sigma^2 \ge \mathbb{E}(X + Y + W + V - X' - Y')^2 \tag{3.136}$$

$$= \mathbb{E}(X + Y + V)^2 + \mathbb{E}(W - X' - Y')^2 \tag{3.137}$$

$$\ge 1 + \sigma^2. \tag{3.138}$$

But this cannot happen since $\Lambda < 1$ by (3.15). Thus, there exists $\eta > 0$ where

$$I(XYV; WX'Y') \geq \eta. \tag{3.139}$$

Recalling that $I(XYV;W) = 0$, this implies

$$I(XYV; X'Y'|W) \geq \eta. \tag{3.140}$$

Also, by (3.29), we may restrict ourselves to distributions where

$$J_{X';Y';W}(R_1, R_2) \geq I(XY; X'Y'W) - 2\delta(\epsilon). \tag{3.141}$$

We may now write, for any (i_1, i_2) and any $\mathbf{w}$

$$\begin{aligned} e_{\vec{X}}(i_1, i_2, \mathbf{w}) &\leq \sum_{(j_1,j_2):(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}} \mathbb{P}\Big\{(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{y}_{j_2}, \mathbf{w}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}\Big\} & (3.142) \\ &\leq \exp\Big\{n\Big[J_{X';Y';XYW}(R_1, R_2) - I(V; X'Y'|XYW) + \delta(\epsilon)\Big]\Big\}, & (3.143) \end{aligned}$$

where in (3.143) we have applied (3.28), the joint typicality lemma 3, and the fact that $I(V; XYW) = 0$.

We consider two cases.

Case (a): $J_{X';Y';W}(R_1, R_2) = 0$. Note that $J_{X';Y';XYW}(R_1, R_2) \leq J_{X';Y';W}(R_1, R_2)$ so in this case we also have $J_{X';Y';XYW}(R_1, R_2) = 0$. By (3.141), $I(XY; X'Y'W) \leq 2\delta(\epsilon)$. Thus, from (3.140)

$$\begin{aligned} \eta &\leq I(XYV; X'Y'|W) & (3.144) \\ &= I(XY; X'Y'|W) + I(V; X'Y'|XYW) & (3.145) \\ &\leq 2\delta(\epsilon) + I(V; X'Y'|XYW). & (3.146) \end{aligned}$$

From (3.143), we have

$$\begin{aligned} e_{\vec{X}}(i_1, i_2, \mathbf{w}) &\leq \exp\{n[-I(V; X'Y'|XYW) + \delta(\epsilon)]\} & (3.147) \\ &\leq \exp\{n[-\eta + 3\delta(\epsilon)]\}. & (3.148) \end{aligned}$$

This vanishes exponentially fast if $\delta(\epsilon)$ is sufficiently small.

Case (b): $J_{X';Y';W}(R_1, R_2) > 0$. This implies that (recalling that $I(X';Y') = 0$)

$$J_{X';Y';W}(R_1, R_2) = \max\{R_1 - I(X';WY'), R_2 - I(Y';WX'), R_1 + R_2 - I(X'Y';W)\}. \tag{3.149}$$

By (3.141), we have

$$\begin{aligned}&-2\delta(\epsilon) \leq \\ &\max\{R_1 - I(X';WY'), R_2 - I(Y';WX'), R_1 + R_2 - I(X'Y';W)\} - I(XY;X'Y'W).\end{aligned} \tag{3.150}$$

Note that

$$I(X';WY') + I(XY;X'Y'W) \geq I(X';WY') + I(XY;X'|WY') \tag{3.151}$$

$$= I(X';XYWY'). \tag{3.152}$$

Similarly

$$I(Y';WX') + I(XY;X'Y'W) \geq I(Y';XYWX'), \tag{3.153}$$

$$I(X'Y';W) + I(XY;X'Y'W) \geq I(X'Y';XYW). \tag{3.154}$$

Thus

$$-2\delta(\epsilon) \leq \max\{R_1 - I(X';XYWY'), R_2 - I(Y';XYWX'), R_1 + R_2 - I(X'Y';XYW)\}. \tag{3.155}$$

Hence

$$\begin{aligned}&J_{X';Y';XYW}(R_1, R_2) \\ &\leq \max\{R_1 - I(X';XYWY'), R_2 - I(Y';XYWX'), R_1 + R_2 - I(X'Y';XYW)\} + 2\delta(\epsilon).\end{aligned} \tag{3.156}$$

By (3.143), we have

$$\begin{aligned}
&\frac{1}{n}\log e_{\vec{X}}(i_1, i_2, \mathbf{w}) \\
&\leq J_{X';Y';XYW}(R_1, R_2) - I(V; X'Y'|XYW) + \delta(\epsilon) && (3.157) \\
&\leq \max\{R_1 - I(X'; XYWY'), R_2 - I(Y'; XYWX'), R_1 + R_2 - I(X'Y'; XYW)\} \\
&\quad - I(V; X'Y'|XYW) + 3\delta(\epsilon) && (3.158) \\
&\leq \max\{R_1 - I(X'; XYWY'V), R_2 - I(Y'; XYWX'V), R_1 + R_2 - I(X'Y'; XYWV)\} \\
&\quad + 3\delta(\epsilon). && (3.159)
\end{aligned}$$

Let $Z = X + Y + W + V - X' - Y'$. Recalling that (X', Y', Z) are mutually independent and $\mathbb{E}Z^2 \leq \Lambda + \sigma^2$, we have

$$\begin{aligned}
I(X'; XYWY'V) &\geq I(X'; X + Y + W + V - Y') && (3.160) \\
&= I(X'; X' + Z) && (3.161) \\
&= h(X' + Z) - h(X' + Z|X') && (3.162) \\
&= \frac{1}{2}\log 2\pi e(\theta + \mathbb{E}Z^2) - h(Z|X') && (3.163) \\
&= \frac{1}{2}\log 2\pi e(\theta + \mathbb{E}Z^2) - \frac{1}{2}\log 2\pi e\mathbb{E}Z^2 && (3.164) \\
&= C\left(\frac{\theta}{\mathbb{E}Z^2}\right) && (3.165) \\
&\geq C\left(\frac{\theta}{\Lambda + \sigma^2}\right). && (3.166)
\end{aligned}$$

Similarly

$$I(Y'; XYWX'V) \geq C\left(\frac{\bar{\theta}}{\Lambda + \sigma^2}\right). \qquad (3.167)$$

Moreover,

$$
\begin{aligned}
I(X'Y';XYWV) &\geq I(X'Y';Z+X'+Y') && (3.168)\\
&= h(Z+X'+Y')-h(Z) && (3.169)\\
&= C\left(\frac{1}{\mathbb{E}Z^2}\right) && (3.170)\\
&\geq C\left(\frac{1}{\Lambda+\sigma^2}\right). && (3.171)
\end{aligned}
$$

Thus

$$
e_{\vec{X}}(i_1,i_2,\mathbf{w}) \leq \\
\exp\Big\{n\Big[\max\big\{R_1 - C\big(\tfrac{\theta}{\Lambda+\sigma^2}\big), R_2 - C\big(\tfrac{\bar{\theta}}{\Lambda+\sigma^2}\big), R_1+R_2-C\big(\tfrac{1}{\Lambda+\sigma^2}\big)\big\} + 3\delta(\epsilon)\Big]\Big\}. \tag{3.172}
$$

Therefore, $e_{\vec{X}}(i_1,i_2,\mathbf{w})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small and (3.16)–(3.18) hold.

3.6 Proof of Lemma 9

In this section, we provide the proofs for (3.23)–(3.29). Since (3.24) is a special case of (3.28) for single codebook, and (3.25)–(3.26) are special cases of (3.29) for single codebook, we refer the proofs for single codebook to the proofs with two codebooks. Moreover, since we frequently use Csiszár and Narayan, (1988)(b), Lemma A1 throughout this section, we provide the statement of this lemma here as Lemma 10.

Lemma 10 *Let $\mathbf{Z}_1,\ldots,\mathbf{Z}_N$ be arbitrary random variables, and let $f_i(\mathbf{Z}_1,\ldots,\mathbf{Z}_i)$ be arbitrary with $0 \leq f_i \leq 1$, $i=1,\ldots,N$. Then the condition*

$$
\mathbb{E}\left[f_i(\mathbf{Z}_1,\ldots,\mathbf{Z}_i)|\mathbf{Z}_1,\ldots,\mathbf{Z}_{i-1}\right] \leq a \text{ a.s.}, \quad i=1,\ldots,N, \tag{3.173}
$$

implies that

$$\mathbb{P}\left\{\frac{1}{N}\sum_{i=1}^{N} f_i(\mathbf{Z}_1,\ldots,\mathbf{Z}_i) > t\right\} \leq \exp\{-N(t - a\log e)\}. \tag{3.174}$$

Let $\mathbf{X}_{i_1}$ and $\mathbf{Y}_{i_2}$ be Gaussian i.i.d. n-length random vectors (codebooks) independent from each other with $\mathrm{Var}(X) = \theta$ and $\mathrm{Var}(Y) = \bar{\theta}$. Fix $\mathbf{x} \in \mathcal{T}_{\epsilon}^{(n)}(X), \mathbf{y} \in \mathcal{T}_{\epsilon}^{(n)}(Y), \mathbf{w} \in \mathscr{S}^n$ and a covariance matrix $\mathrm{Cov}(X, Y, X', Y', W) \in \mathcal{V}^{5\times 5}$ such that $\mathscr{S}^n$ is a ν-dense subset of $\mathbb{R}^n$ for $\mathbf{w}$ such that $||\mathbf{w}||^2 \leq n\Lambda$, and $\mathcal{V}^{5\times 5}$ is a ν-dense subset of $\mathbb{R}^{5x5}$ for positive definite covariance matrices with diagonals at most $(\theta, \bar{\theta}, \theta, \bar{\theta}, \Lambda)$.

Let

$$A_{\epsilon}^{n}(X, W) = \bigcup_{\substack{X,W \text{ independent}\\ \mathbb{E}X^2=\theta, \mathbb{E}W^2\leq\Lambda}} \mathcal{T}_{\epsilon}^{(n)}(X, W) \tag{3.175}$$

and

$$A_{\epsilon}^{n}(X, Y, W) = \bigcup_{\substack{(X,Y,W) \text{ mutually independent}\\ EX^2=\theta, EY^2=\bar{\theta}, EW^2\leq\Lambda}} \mathcal{T}_{\epsilon}^{(n)}(X, Y, W). \tag{3.176}$$

To prove (3.23), first define h_{i_1} as a function of $\mathbf{X}_1, \ldots, \mathbf{X}_{i_1}$ as follows:

$$h_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1}) = \begin{cases} 1, & \text{if } (\mathbf{X}_{i_1}, \mathbf{w}) \notin A_{\epsilon}^{n}(X, W) \\ 0, & \text{otherwise .} \end{cases} \tag{3.177}$$

Then the expected value of h_{i_1} is given as

$$\begin{aligned}
\mathbb{E}[h_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1})|\mathbf{X}_1, \ldots, \mathbf{X}_{i_1-1}] &= \mathbb{E}\left[\mathbb{1}\left((\mathbf{X}_{i_1}, \mathbf{w}) \notin A_{\epsilon}^{n}(X, W)\right)\right] && (3.178)\\
&= \mathbb{P}\left\{(\mathbf{X}_{i_1}, \mathbf{w}) \notin A_{\epsilon}^{n}(X, W)\right\} && (3.179)\\
&\leq \mathbb{P}\left\{\frac{1}{n}\|\mathbf{w}\|^2 \geq \Lambda + \epsilon\right\} + \mathbb{P}\left\{\frac{1}{n}|\langle \mathbf{X}_{i_1}, \mathbf{w}\rangle| \geq \epsilon\right\} \\
&\quad + \mathbb{P}\left\{|\frac{1}{n}\|\mathbf{X}_{i_1}\|^2 - \theta| \geq \epsilon\right\} && (3.180)\\
&\leq \exp(-nr_1(\epsilon)) && (3.181)
\end{aligned}$$

where (3.180) follows since the union in $A_\epsilon^n(X, W)$ is over independent X, W such that $\mathbb{E}X^2 = \theta, \mathbb{E}W^2 \leq \Lambda$, and directly from the definition of typical set in (3.1) we obtain that it only suffices to find the probability of those $(\mathbf{X}_{i_1}, \mathbf{w})$ that simultaneously do not satisfy the conditions of the union and the typical set definition's inequalities. The upper bound in (3.181) follows since the first term in (3.180) is equal to zero (assumption $\|\mathbf{w}\|^2 \leq n\Lambda$), and the other terms are exponentially vanishing by using the large deviation theory for Gaussian distributions X with positive function $r_1(\epsilon)$.

Now, using Lemma 10, we have

$$\mathbb{P}\left\{\frac{1}{N_1}|\{i_1 : (\mathbf{X}_{i_1}, \mathbf{w}) \notin A_\epsilon^n(X, W)\}| > \exp(-n\delta_1(\epsilon))\right\}$$

$$\leq \exp\{-N_1[\exp(-n\delta_1(\epsilon)) - \exp(-nr_1(\epsilon)) \log e]\} \quad (3.182)$$

$$\leq \exp(-\exp(n\rho_1(\epsilon))). \quad (3.183)$$

where the last inequality follows as long as $\delta_1(\epsilon) < r_1(\epsilon)$ for some $\rho_1(\epsilon) > 0$. Thus, the probability vanishes doubly exponentially as $n \to \infty$, and with high probability we have

$$\frac{1}{N_1}|\{i_1 : (\mathbf{x}_{i_1}, \mathbf{w}) \notin A_\epsilon^n(X, W)\}| \leq \exp(-n\delta_1(\epsilon)). \quad (3.184)$$

Fix $\mathbf{x}_{i_1}$, and for any i_2 define $\tilde{h}_{i_2}$ as

$$\tilde{h}_{i_2}(\mathbf{Y}_1, \ldots, \mathbf{Y}_{i_2}) = \frac{1}{N_1} \sum_{i_1 : (\mathbf{x}_{i_1}, \mathbf{w}) \in A_\epsilon^n(X, W)} \mathbb{1}\left((\mathbf{x}_{i_1}, \mathbf{Y}_{i_2}, \mathbf{w}) \notin A_\epsilon^n(X, Y, W)\right). \quad (3.185)$$

The expected value of $\tilde{h}_{i_2}$ can be written as

$$\mathbb{E}\left[\tilde{h}_{i_2}(\mathbf{Y}_1, \ldots, \mathbf{Y}_{i_2}) | \mathbf{Y}_1, \ldots, \mathbf{Y}_{i_2 - 1}\right] = \frac{1}{N_1} \sum_{i_1 : (\mathbf{x}_{i_1}, \mathbf{w}) \in A_\epsilon^n(X, W)} \mathbb{P}\left\{(\mathbf{x}_{i_1}, \mathbf{Y}_{i_2}, \mathbf{w}) \notin A_\epsilon^n(X, Y, W)\right\}$$

$$\leq \mathbb{P}\left\{\frac{1}{n}\|\mathbf{w}\|^2 \geq \Lambda + \epsilon\right\} + \mathbb{P}\left\{\frac{1}{n}|\langle \mathbf{x}_{i_1}, \mathbf{w}\rangle| \geq \epsilon\right\} + \mathbb{P}\left\{\left|\frac{1}{n}\|\mathbf{x}_{i_1}\|^2 - \theta\right| \geq \epsilon\right\}$$

$$+ \mathbb{P}\left\{\frac{1}{n}|\langle \mathbf{Y}_{i_2}, \mathbf{w}\rangle| \geq \epsilon\right\} + \mathbb{P}\left\{\frac{1}{n}|\langle \mathbf{Y}_{i_2}, \mathbf{x}_{i_1}\rangle| \geq \epsilon\right\} + \mathbb{P}\left\{\left|\frac{1}{n}\|\mathbf{Y}_{i_2}\|^2 - \bar{\theta}\right| \geq \epsilon\right\} \quad (3.186)$$

$$\leq \exp(-nr_2(\epsilon)) \quad (3.187)$$

where (3.186) follows directly from the definition of typical set and the union's conditions, and (3.187) follows since the first three terms in (3.186) are equal to 0 due to the assumptions and the other terms in (3.186) vanish exponentially by the large deviation theory for Gaussian distributions Y with positive function $r_2(\epsilon)$

Using Lemma 10, we have

$$\mathbb{P}\left\{\frac{1}{N_1N_2}\sum_{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in A_\epsilon^n(X,W)}|\{i_2:(\mathbf{x}_{i_1},\mathbf{Y}_{i_2},\mathbf{w})\notin A_\epsilon^n(X,Y,W)\}|>\exp(-n\delta_2(\epsilon))\right\}$$

$$\leq \exp\{-N_2(\exp(-n\delta_2(\epsilon)-\exp(-nr_2(\epsilon))\log e))\} \tag{3.188}$$

$$\leq \exp(-\exp(n\rho_2(\epsilon))) \tag{3.189}$$

where (3.189) follows if $\delta_2(\epsilon)<r_2(\epsilon)$ for some $\rho_2(\epsilon)>0$. Therefore, with probability approaching 1, as $n\to\infty$ we have

$$\frac{1}{N_1N_2}\sum_{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in A_\epsilon^n(X,W)}|\{i_2:(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\notin A_\epsilon^n(X,Y,W)\}|\leq\exp(-n\delta_2(\epsilon)). \tag{3.190}$$

Eventually, we easily use (3.184) and (3.190) to bound the fraction in (3.23) as follows:

$$\begin{aligned}\frac{1}{N_1N_2}|\{(i_1,i_2):&(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\notin A_\epsilon^n(X,Y,W)\}|\\ &\leq\frac{1}{N_1N_2}\sum_{i_1:(\mathbf{x}_{i_1},\mathbf{w})\notin A_\epsilon^n(X,W)}|\{i_2:(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\notin A_\epsilon^n(X,Y,W)\}|\\ &\quad+\frac{1}{N_1N_2}\sum_{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in A_\epsilon^n(X,W)}|\{i_2:(\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\notin A_\epsilon^n(X,Y,W)\}|\end{aligned} \tag{3.191}$$

$$\leq\exp(-n\delta_1(\epsilon))+\exp(-n\delta_2(\epsilon)) \tag{3.192}$$

$$\leq\exp(-n\delta(\epsilon)). \tag{3.193}$$

This proves (3.23).

Now, in order to prove (3.26) in Lemma (9), first let

$$A_i = \big\{j : j < i, (\mathbf{x}_j, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X', S)\big\}, \tag{3.194}$$

$$\tilde{A}_i = \begin{cases} A_i, & \text{if } |A_i| \leq \exp\left\{n\left(|R - I(X'; S)|^+ + \delta(\epsilon)\right)\right\} \\ \emptyset, & \text{otherwise.} \end{cases} \tag{3.195}$$

Define

$$g_i(\mathbf{x}_1, \ldots, \mathbf{x}_i) = \begin{cases} 1, & \text{if } (\mathbf{x}_i, \mathbf{x}_j, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S) \text{ for some } j \in \tilde{A}_i \\ \emptyset, & \text{otherwise.} \end{cases} \tag{3.196}$$

It is notable that since by (3.24)

$$\mathbb{P}\left\{\left|\{j : (\mathbf{x}_j, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X', S)\}\right| > \exp\{n(|R - I(X'; S)|^+ + \delta(\epsilon))\}\right\} \tag{3.197}$$

tends to zero as n grows, the probability that $\tilde{A}_i \neq A_i$ for some i vanishes as $n \to \infty$.

Finding the expected values of g_i, we have

$$\begin{aligned} &\mathbb{E}[g_i(\mathbf{X}_1, \ldots, \mathbf{X}_i) | \mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_{i-1} = \mathbf{x}_{i-1}] \\ &\quad = \mathbb{P}\Big\{(\mathbf{X}_i, \mathbf{x}_j, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S) \text{ for some } j \in \tilde{A}_i\Big\} && (3.198) \\ &\quad \leq \sum_{j \in \tilde{A}_i} \mathbb{P}\left\{(\mathbf{X}_i, \mathbf{x}_j, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S)\right\} && (3.199) \\ &\quad \leq \exp(n(|R - I(X'; S)|^+ + \delta(\epsilon))) \max_{\hat{\mathbf{x}}} \mathbb{P}\left\{(\mathbf{X}_i, \hat{\mathbf{x}}, \mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S)\right\} && (3.200) \\ &\quad \leq \exp\{-n(-|R - I(X'; S)|^+ + I(X; X'S) - \delta(\epsilon)/2)\} && (3.201) \end{aligned}$$

where (3.200) follows since by (3.195) the size of $\tilde{A}_i$ is almost surely less than $\exp(n(|R - I(X'; S)|^+ + \delta(\epsilon)))$, (3.201) follows by joint typicality lemma 3, and (3.241) follows by the condition in (3.29).

Therefore, using Lemma 10, we have

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{N}\sum_{i=1}^{N} g_i(\mathbf{x}_1,\ldots,\mathbf{x}_i) > \exp\{-n(-|R-I(X';S)|^+ + I(X;X'S)-\delta(\epsilon))\})\right\}\\
&\quad = \mathbb{P}\Bigg\{\frac{1}{N}\left|\Big\{i:(\mathbf{x}_i,\mathbf{x}_j,\mathbf{s})\in\mathcal{T}_\epsilon^{(n)}(X,X',S) \text{ for some } j\in\tilde{A}_i\Big\}\right|\\
&\qquad > \exp\{-n(-|R-I(X';S)|^+ + I(X;X'S)-\delta(\epsilon)/2)\}\Bigg\} \qquad (3.202)\\
&\quad \le \exp[-\exp(nR)(\exp\{-n(-|R-I(X';S)|^+ + I(X;X'S)-\delta(\epsilon))\}\\
&\qquad -\exp\{-n(-|R-I(X';S)|^+ + I(X;X'S)-\delta(\epsilon)/2)\}\log e)] \qquad (3.203)\\
&\quad \le \exp(-\exp(n\sigma(\epsilon))) \qquad (3.204)
\end{aligned}$$

for $\sigma(\epsilon) > 0$ i.e. with high probability

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N} g_i(\mathbf{x}_1,\ldots,\mathbf{x}_i) &= \frac{1}{N}|\{i:(\mathbf{x}_i,\mathbf{x}_j,\mathbf{s})\in\mathcal{T}_\epsilon^{(n)}(X,X',S) \text{ for some } j\in\tilde{A}_i\}| \qquad (3.205)\\
&\le \exp\{-n(-|R-I(X';S)|^+ + I(X;X'S)-\delta(\epsilon))\}. \qquad (3.206)
\end{aligned}$$

We may use this argument as we used here for the case $j > i$ by only reversing the order of the codewords for the same A_i. Finally, by the same decreasing exponential function for each case, we obtain (3.26) as

$$\frac{1}{N}\Big|\Big\{i:(\mathbf{x}_i,\mathbf{x}_j,\mathbf{s})\in\mathcal{T}_\epsilon^{(n)}(X,X',S) \text{ for some } j\neq i\Big\}\Big| \le 2\exp\{-n\delta(\epsilon)/2\} \tag{3.207}$$

if we have $|R-I(X';S)|^+ \le I(X;X'S)-2\delta(\epsilon)$.

Next, define function f_{i_1} as follows:

$$f_{i_1}(\mathbf{X}_1,\ldots,\mathbf{X}_{i_1}) = \begin{cases} 1, & \text{if } (\mathbf{X}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\\ 0, & \text{otherwise .}\end{cases} \tag{3.208}$$

Then, by joint typicality lemma 3 we get

$$\begin{aligned}\mathbb{E}[f_{i_1}(\mathbf{X}_1,\dots,\mathbf{X}_{i_1})|\mathbf{X}_1,\dots,\mathbf{X}_{i_1-1}] &= \mathbb{E}\left[\mathbb{1}\left((\mathbf{X}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\right)\right] && (3.209)\\ &= \mathbb{P}\left((\mathbf{X}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\right) && (3.210)\\ &\le \exp(-nI(X',W)+n\delta(\epsilon)). && (3.211)\end{aligned}$$

Thus, using Lemma 10, we have

$$\begin{aligned}&\mathbb{P}\left\{\left|\left\{i_1:(\mathbf{X}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\right\}\right|>\exp\left(n|R_1-I(X',W)|^+ + n2\delta(\epsilon)\right)\right\}\\ &\le \exp\left\{-\exp\left(n|R_1-I(X',W)|^+ + n2\delta(\epsilon)\right)+\exp(-nI(X',W)+n\delta(\epsilon)+nR_1)\log e\right\}.\end{aligned} \tag{3.212}$$

If $R_1 > I(X',W)$, (3.212) becomes less than doubly exponentially function $\exp(-\exp(n\sigma(\epsilon)))$ where $\sigma(\epsilon)>0$ since for large enough n we obtain $\exp(n\delta(\epsilon)) > \log e$. Now, if $R_1 \le I(X',W)$ then (3.212) is less than

$$\exp\{-\exp(n2\delta(\epsilon))+\exp(nR_1+n\delta(\epsilon)-nI(X',W))\log e\}\le\exp(-\exp(n\sigma(\epsilon)))$$

where $\sigma(\epsilon)>0$. In both cases, this doubly decreasing exponential function vanishes as $n\to\infty$. Hence, with high probability, we have

$$\left|\left\{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\right\}\right|\le\exp\left\{n\left[|R_1-I(X';W)|^+ +\delta(\epsilon)\right]\right\}. \tag{3.213}$$

For any i_2,

$$\begin{aligned}&\mathbb{P}\left\{(\mathbf{x}_{i_1},\mathbf{Y}_{i_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',Y',W)\text{ for some }i_1\right\}\\ &\quad\le \sum_{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)}\mathbb{P}\left\{(\mathbf{x}_{i_1},\mathbf{Y}_{i_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\} && (3.214)\\ &\quad\le \left|\left\{i_1:(\mathbf{x}_{i_1},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',W)\right\}\right|\max_{\hat{\mathbf{x}}}\mathbb{P}\left\{(\hat{\mathbf{x}},\mathbf{Y}_{i_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\} && (3.215)\\ &\quad\le \exp\left\{n\left(|R_1-I(X';W)|^+ - I(Y';X'W)+\delta(\epsilon)\right)\right\}, && (3.216)\end{aligned}$$

where (3.214) follows since if $(\mathbf{x}_{i_1}, \mathbf{Y}_{i_2}, \mathbf{w})$ is typical then $(\mathbf{x}_{i_1}, \mathbf{w})$ is always typical and (3.216) follows from (3.213) and joint typicality lemma 3. Thus, applying Lemma 10 in the same way that we used it to get (3.213), with high probability we attain

$$\begin{aligned}&\left|\left\{i_2 : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W) \text{ for some } i_1\right\}\right| \\ &\qquad\le \exp\Big\{n\Big[\big|R_2 + |R_1 - I(X';W)|^+ - I(Y';X'W)\big|^+ + \delta(\epsilon)\Big]\Big\}. \qquad (3.217)\end{aligned}$$

Since $(\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W)$ implies $(\mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(Y', W)$, we have the simpler bound

$$\begin{aligned}\left|\left\{i_2 : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W) \text{ for some } i_1\right\}\right| &\le \left|\left\{i_2 : (\mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(Y', W)\right\}\right| \qquad (3.218) \\ &\le \exp\big\{n\big[|R_2 - I(Y';W)|^+ + \delta(\epsilon)\big]\big\}, \qquad (3.219)\end{aligned}$$

where (3.219) is similar to (3.213). Moreover, if we replace vector $\mathbf{w}$ by $(\mathbf{y}_{i_2}, \mathbf{w})$ in (3.213), then for any i_2 we get

$$\left|\left\{i_1 : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W)\right\}\right| \le \exp\big\{n\big[|R_1 - I(X';Y'W)|^+ + \delta(\epsilon)\big]\big\}. \qquad (3.220)$$

Therefore,

$$\begin{aligned}&\left|\left\{(i_1, i_2) : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W)\right\}\right| \\ &\le \sum_{i_2 : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W) \text{ for some } i_1} \left|\left\{i_1 : (\mathbf{x}_{i_1}, \mathbf{y}_{i_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X', Y', W)\right\}\right| \qquad (3.221) \\ &\le \exp\Big\{n\Big[\big|R_2 + |R_1 - I(X';W)|^+ - I(Y';X'W)\big|^+ + |R_1 - I(X';Y'W)|^+ + \delta(\epsilon)\Big]\Big\} \qquad (3.222)\end{aligned}$$

where (3.222) follows from (3.217) and (3.220). If $R_1 \leq I(X';Y'W)$, then we have

$$\begin{aligned}
&\left|\left\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\}\right| \\
&\leq \exp\left\{n\Big[\big|R_2 + |R_1 - I(X';W)|^+ - I(Y';X'W)\big|^+ + \delta(\epsilon)\Big]\right\} && (3.223) \\
&= \exp\left\{n\Big[\max\big\{0, R_2 - I(Y';X'W), R_1 + R_2 - I(X'Y';W) - I(X';Y')\big\} + \delta(\epsilon)\Big]\right\} && (3.224) \\
&= \exp\big\{n\big[J_{X';Y';W}(R_1,R_2) + \delta(\epsilon)\big]\big\}. && (3.225)
\end{aligned}$$

Using (3.219) and (3.220), we alternatively bound

$$\begin{aligned}
&\left|\left\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\}\right| \\
&\qquad\leq \exp\big\{n\big[|R_2 - I(Y';W)|^+ + |R_1 - I(X';Y'W)|^+ + \delta(\epsilon)\big]\big\}. && (3.226)
\end{aligned}$$

In particular, if $R_1 > I(X';Y'W)$ then we have

$$\begin{aligned}
&\left|\left\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\}\right| \\
&\leq \exp\big\{n\big[|R_2 - I(Y';W)|^+ + R_1 - I(X';W|Y') + \delta(\epsilon)\big]\big\} && (3.227) \\
&= \exp\left\{n\Big[\max\big\{R_1 - I(X';W|Y'), R_1 + R_2 - I(X'Y';W)\big\} + \delta(\epsilon)\Big]\right\} && (3.228) \\
&= \exp\big\{n\big[J_{X';Y';W}(R_1,R_2) + \delta(\epsilon)\big]\big\}. && (3.229)
\end{aligned}$$

This proves (3.27). An identical calculation with (X,Y,W) in place of W gives (3.28). We indeed use Lemma 10 to prove that the complement of two events (3.27) and (3.28) happen with decreasingly doubly exponential probability for sufficiently large n as follows:

$$\begin{aligned}
&\mathbb{P}\Big\{\left|\left\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X',Y',W)\right\}\right| > \exp\big\{n\big[J_{X';Y';W}(R_1,R_2) + \delta(\epsilon)\big]\big\}\Big\} \\
&\qquad < \exp(-\exp(n\sigma(\epsilon))), && (3.230)
\end{aligned}$$

$$\mathbb{P}\big\{\big|\big\{(i_1,i_2):(\mathbf{x},\mathbf{y},\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W)\big\}\big|$$

$$>\exp\big\{n\big[J_{X';Y';XYW}(R_1,R_2)+\delta(\epsilon)\big]\big\}\big\}<\exp(-\exp(n\sigma(\epsilon))). \quad (3.231)$$

Now, in order to prove (3.29), first let

$$A_{(i_1,i_2)}=\big\{(j_1,j_2):j_1<i_1,j_2\neq i_2,(\mathbf{x}_{j_1},\mathbf{Y}_{j_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',Y',W)\big\}, \quad (3.232)$$

$$\tilde{A}_{(i_1,i_2)}=\begin{cases}A_{(i_1,i_2)}, & \text{if } |A_{(i_1,i_2)}|\leq\exp\left\{n\left(J_{X';Y';W}(R_1,R_2)+\delta(\epsilon)\right)\right\}\\ \emptyset, & \text{otherwise,}\end{cases} \quad (3.233)$$

where $\tilde{A}_{(i_1,i_2)}$ is defined for fixed value of $\mathbf{x}_1,\ldots,\mathbf{x}_{i_1-1}$ and random $\mathbf{Y}_1,\ldots,\mathbf{Y}_{j_2}$, i.e. $\tilde{A}_{(i_1,i_2)}$ is a random set. Define

$$g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1})=$$

$$\mathbb{P}\left\{(\mathbf{x}_{i_1},\mathbf{Y}_{i_2},\mathbf{x}_{j_1},\mathbf{Y}_{j_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W)\text{ for some }(j_1,j_2)\in\tilde{A}_{(i_1,i_2)}\right\} \quad (3.234)$$

and

$$\tilde{g}_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1})=\begin{cases}1, & \text{if } g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1})>\exp(-n\delta(\epsilon)/2)\\ \emptyset, & \text{otherwise.}\end{cases} \quad (3.235)$$

It is notable that since by (3.27)

$$\mathbb{P}\left\{\big|\big\{(j_1,j_2):(\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w})\in\mathcal{T}_\epsilon^{(n)}(X',Y',W)\big\}\big|>\exp\{n(J_{X';Y';W}(R_1,R_2)+\delta(\epsilon))\}\right\} \quad (3.236)$$

tends to zero as n grows, the probability that $\tilde{A}_{(i_1,i_2)}\neq A_{(i_1,i_2)}$ for some i_1,i_2 vanishes as $n\to\infty$.

Finding the expected values of g_{i_1} and $\tilde{g}_{i_1}$, we have

$$\mathbb{E}[g_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1})|\mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_{i_1-1} = \mathbf{x}_{i_1-1}]$$

$$= \mathbb{P}\Big\{(\mathbf{X}_{i_1}, \mathbf{Y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{Y}_{j_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W) \text{ for some } (j_1, j_2) \in \tilde{A}_{(i_1,i_2)}\Big\} \tag{3.237}$$

$$\leq \sum_{(j_1,j_2)\in\tilde{A}_{(i_1,i_2)}} \mathbb{P}\left\{(\mathbf{X}_{i_1}, \mathbf{Y}_{i_2}, \mathbf{x}_{j_1}, \mathbf{Y}_{j_2}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W)\right\} \tag{3.238}$$

$$\leq \exp(nJ_{X';Y';W}(R_1, R_2) + \delta(\epsilon)) \max_{\hat{\mathbf{x}},\hat{\mathbf{y}}} \mathbb{P}\left\{(\mathbf{X}_{i_1}, \mathbf{Y}_{i_2}, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X, Y, X', Y', W)\right\} \tag{3.239}$$

$$\leq \exp\{-n(-J_{X';Y';W}(R_1, R_2) + I(XY; X'Y'W) - \delta(\epsilon))\} \tag{3.240}$$

$$\leq \exp(-n\delta(\epsilon)) \tag{3.241}$$

where (3.239) follows since by (3.233) the size of $\tilde{A}_{(i_1,i_2)}$ is almost surely less than $\exp(nJ_{X';Y';W}(R_1, R_2) + \delta(\epsilon))$, (3.240) follows by joint typicality lemma 3, and (3.241) follows by the condition in (3.29). Moreover, by Markov's inequality we have

$$\mathbb{E}[\tilde{g}_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1})|\mathbf{X}_1, \ldots, \mathbf{X}_{i_1-1}] = \mathbb{P}\{g_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1}) > \exp(-n\delta/2)|\mathbf{X}_1, \ldots, \mathbf{X}_{i_1-1}\} \tag{3.242}$$

$$\leq \frac{\mathbb{E}[g_{i_1}(\mathbf{X}_1, \ldots, \mathbf{X}_{i_1})|\mathbf{X}_1, \ldots, \mathbf{X}_{i_1-1}]}{\exp(-n\delta(\epsilon)/2)} \tag{3.243}$$

$$\leq \exp(-n\delta(\epsilon) + n\delta(\epsilon)/2) \tag{3.244}$$

$$= \exp(-n\delta(\epsilon)/2). \tag{3.245}$$

Therefore, using Lemma 10, we have

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{N_1}\sum_{i_1=1}^{N_1}\tilde{g}_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) > \exp(-n\delta(\epsilon)/4)\right\} \\
&\quad = \mathbb{P}\left\{\frac{1}{N_1}|\{i_1 : g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) > \exp(-n\delta(\epsilon)/2)\}| > \exp(-n\delta(\epsilon)/4)\right\} && (3.246)\\
&\quad \le \exp\{-\exp(nR_1)(\exp(-n\delta(\epsilon)/4) - \exp(-n\delta(\epsilon)/2)\log e)\} && (3.247)\\
&\quad \le \exp(-\exp(n\sigma(\epsilon))). && (3.248)
\end{aligned}$$

for $\sigma(\epsilon) > 0$ i.e. with high probability

$$\begin{aligned}
\frac{1}{N_1}\sum_{i_1}\tilde{g}_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) &= \frac{1}{N_1}|\{i_1 : g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) > \exp(-n\delta(\epsilon)/2)\}| && (3.249)\\
&\le \exp(-n\delta(\epsilon)/4). && (3.250)
\end{aligned}$$

Let

$$f_{(i_1,i_2)}(\mathbf{y}_1,\ldots,\mathbf{y}_{i_2}) = \begin{cases} 1, & \text{if } (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W), \\ & \text{for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2 \\ 0, & \text{otherwise.} \end{cases} \qquad (3.251)$$

Now, fix an i_1 such that $g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) \le \exp(-n\delta(\epsilon)/2)$. Therefore, we have

$$\begin{aligned}
&\mathbb{E}\left[f_{(i_1,i_2)}(\mathbf{Y}_1,\ldots,\mathbf{Y}_{i_2})|\mathbf{Y}_1,\ldots,\mathbf{Y}_{i_2-1}\right] \\
&\quad = \mathbb{P}\Big\{(\mathbf{x}_{i_1},\mathbf{Y}_{i_2},\mathbf{x}_{j_1},\mathbf{Y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \\
&\qquad \text{for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2 \Big| \mathbf{Y}_1,\ldots,\mathbf{Y}_{i_2-1}\Big\} && (3.252)\\
&\quad \le g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) && (3.253)\\
&\quad \le \exp(-n\delta(\epsilon)/2) && (3.254)
\end{aligned}$$

where (3.253) and (3.254) follow directly from g_{i_1} definition and our assumption for the g_{i_1}. Thus, using Lemma 10 we get

$$\mathbb{P}\left(\frac{1}{N_2}\sum_{i_2=1}^{N_2} f_{(i_1,i_2)}(\mathbf{Y}_1,\ldots,\mathbf{Y}_{i_2}) > \exp(-n\delta(\epsilon)/4)\right) \leq \exp(-\exp(n\sigma(\epsilon))) \tag{3.255}$$

where $\sigma(\epsilon) > 0$. If we sum over all i_1's, we obtain

$$\sum_{i_1=1}^{N_1}\mathbb{P}\left(\frac{1}{N_2}\sum_{i_2=1}^{N_2} f_{(i_1,i_2)}(\mathbf{Y}_1,\ldots,\mathbf{Y}_{i_2}) > \exp(-n\delta(\epsilon)/4)\right) \leq \exp(nR_1 - \exp(n\sigma(\epsilon))), \tag{3.256}$$

that is this doubly exponential function still tends to zero as $n \to \infty$. Therefore, with probability approaching 1, for every i_1 that $g_{i_1}(\mathbf{x}_1,\ldots,\mathbf{x}_{i_1}) \leq \exp(-n\delta(\epsilon)/2)$ we have

$$\frac{1}{N_2}\Big|\Big\{i_2 : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \\ \text{for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2\Big\}\Big| \leq \exp(-n\delta(\epsilon)/4). \tag{3.257}$$

In general, for all i_1 we have

$$\frac{1}{N_1 N_2}\Big|\Big\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \\ \text{for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2\Big\}\Big| \tag{3.258}$$

$$\leq \frac{1}{N_1}\sum_{i_1=1}^{N_1}\frac{1}{N_2}\Big|\Big\{i_2 : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \\ \text{for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2\Big\}\Big| \tag{3.259}$$

$$\leq \exp(-n\delta(\epsilon)/4) + \frac{1}{N_1}\sum_{i_1 : g_{i_1} \leq \exp(-n\delta(\epsilon)/2)}\frac{1}{N_2}\Big|\Big\{i_2 : \\ (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \text{ for some } (j_1,j_2) \in \tilde{A}_{(i_1,i_2)} \text{ and } j_2 < i_2\Big\}\Big| \tag{3.260}$$

$$\leq 2\exp(-n\delta(\epsilon)/4) \tag{3.261}$$

where (3.260) and (3.261) follow from (3.250) and (3.257), respectively.

We may use this argument to upper bound the probability in (3.258) for the remain three cases $(j_1 < i_1, j_2 > i_2)$, $(j_1 > i_1, j_2 < i_2)$ and $(j_1 > i_1, j_2 > i_2)$ by defining different $A_{(i_1,i_2)}$'s, and conclude the same decreasing exponential function. Finally, we obtain

$$\frac{1}{N_1 N_2}\big|\{(i_1,i_2) : (\mathbf{x}_{i_1},\mathbf{y}_{i_2},\mathbf{x}_{j_1},\mathbf{y}_{j_2},\mathbf{w}) \in \mathcal{T}_\epsilon^{(n)}(X,Y,X',Y',W) \\ \text{for some } j_1 \neq i_1, j_2 \neq i_2\}\big| \leq 8\exp\{-n\delta(\epsilon)/4\} \tag{3.262}$$

if we have $J_{X';Y';W}(R_1,R_2) \leq I(XY;X'Y'W) - 2\delta(\epsilon)$.

In order to complete the proof, since for any fixed ν the cardinality of finite set $\mathscr{S}^n$ is only increasingly exponentially in n, and the set $\mathcal{V}^{5\times 5}$ is finite along with the doubly decreasing exponential probabilities in (3.230) and (3.231), we derive that with probability approaching to 1, all inequalities in (3.23), (3.27), (3.28) and (3.29) hold simultaneously for sufficiently large n. Since these inequalities hold for every element in the finite sets $\mathscr{S}^n$ and $\mathcal{V}^{5\times 5}$, then for any vector $\mathbf{w}, \mathbf{x}, \mathbf{y}$ and any given covariance matrix $\mathrm{Cov}(X, Y, X', Y', W)$ (with $\|\mathbf{x}\|^2 = n\theta$, $\|\mathbf{y}\|^2 = n\bar{\theta}$, $\|\mathbf{w}\|^2 \leq n\Lambda$) which is not in its corresponding ν-dense subset, there exists a point in the corresponding ν-dense subset that is close enough to it (in its ν distance neighborhood). Now, by using the continuity properties of all sets, we may conclude that (3.23), (3.27), (3.28) and (3.29) hold also for any point which is not in the ν-dense subset.

Chapter 4

GAUSSIAN ARBITRARILY-VARYING MULTIPLE ACCESS CHANNEL

4.1 Problem Statement

The Gaussian multiple access channel with a jammer is shown in Fig. 2, in which two users send their messages to one receiver in the presence of one jammer. This channel is also known as Gaussian arbitrarily-varying multiple-access channel (Gaussian AVMAC). The jammer is assumed not to have any information about the user's signals (but know the code). In particular, the received signal is given by

$$\mathbf{Y} = g_1\mathbf{X}_1 + g_2\mathbf{X}_2 + \mathbf{S} + \mathbf{V} \tag{4.1}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are n-length vectors representing the user's signals, $\mathbf{S}$ is the adversarial jammer signal, g_1 and g_2 are the channel gains, and $\mathbf{V}$ is the n-length noise vector distributed as a sequence of i.i.d. zero mean Gaussian random variables with variance σ^2 which is independent of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{S}$.

The transmitters and jammer signals are constrained to satisfy power constraints $\|\mathbf{X}_i\|^2 \leq \ nP_i$, for $i = 1, 2$ and $\|\mathbf{S}\|^2 \leq n\Lambda$. We define the received signal-to-noise ratios as $S_1 = g_1^2 P_1/\sigma^2$, $S_2 = g_2^2 P_2/\sigma^2$. We also denote the jammer-to-noise ratio as $J = \Lambda/\sigma^2$. Note that the vector $\mathbf{S}$ refers to the jammer signal while the scaler values S_1 and S_2 denotes the received signal-to-noise ratios. We assume that the transmitters and receiver know the signal-to-noise ratios, but they need not know the jammer-to-noise ratio. However, we require small probability of error only when the

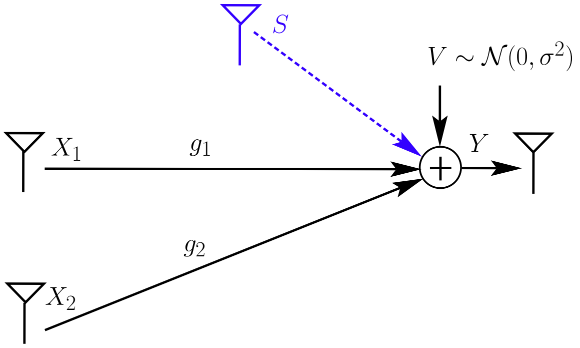


Figure 2: Two-user Gaussian Multiple Access Channel with a Jammer.

jammer-to-noise ratio does not exceed J; thus the code is independent of the jammer's power up to a point, and beyond that it may fail to decode correctly.

A $\left(2^{nR_1}, 2^{nR_2}, n\right)$ deterministic code is given by:

- Message sets $\mathcal{M}_1 = [2^{nR_1}]$ and $\mathcal{M}_2 = [2^{nR_2}]$,
- Encoding functions $\mathbf{x}_i : \mathcal{M}_i \to \mathbb{R}^n$ for $i = 1, 2$, and
- Decoding function $\phi : \mathbb{R}^n \to (\mathcal{M}_1, \mathcal{M}_2)$.

For $i = 1, 2$, the message M_i is chosen uniformly from the set $\mathcal{M}_i$, and each transmitter encodes its own message to $\mathbf{X}_i$. At the receiver, the received signal $\mathbf{Y}$ is decoded by function ϕ to $(\hat{M}_1, \hat{M}_2) = \phi(\mathbf{Y})$. The average probability of error $P_e^{(n)}$ is now given by the probability that $(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)$, maximized over all possible choices of jammer's sequence $\mathbf{S}$. A rate pair (R_1, R_2) is *achievable* if there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}, n\right)$ codes where $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region $\mathscr{C}$ is the closure of the set of all achievable rate pairs (R_1, R_2).

4.2 Main Results

The following theorem provides the capacity region of two-user Gaussian MAC with a jammer.

Theorem 11 *Assume $S_1 > J$ and $S_2 > J$. The capacity region of Gaussian multiple access channel is the set of rate pairs (R_1, R_2) such that*

$$\begin{aligned} R_1 &< C\left(\frac{S_1}{J+1}\right) = C\left(\frac{g_1^2 P_1}{\Lambda + \sigma^2}\right) \\ R_2 &< C\left(\frac{S_2}{J+1}\right) = C\left(\frac{g_2^2 P_2}{\Lambda + \sigma^2}\right) \\ R_1 + R_2 &< C\left(\frac{S_1 + S_2}{J+1}\right) = C\left(\frac{g_1^2 P_1 + g_2^2 P_2}{\Lambda + \sigma^2}\right) \end{aligned} \tag{4.2}$$

4.3 Converse Proof

Consider a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with vanishing probability of error. Since these codes must function for arbitrary jamming signals, we may assume that the jammer transmits Gaussian noise with variance Λ. Thus, we follow the capacity for the Gaussian MAC with no jammer ElGamal and Kim, (2011), Chapter 4.6.1, p. 94 and the noise power $\sigma^2 + \Lambda$.

Moreover, if $J \geq S_1$, based on the assumption that the jammer knows the code, the jammer can choose an arbitrary message $\tilde{m}_1$ and transmit a scaled form of the corresponding codeword $\mathbf{s} = \mathbf{x}_1(\tilde{m}_1)g_1$. Given $\mathbf{Y} = g_1\mathbf{x}_1(m_1) + g_2\mathbf{x}_2(m_2) + g_1\mathbf{x}_1(\tilde{m}_1) + \mathbf{V}$, the decoder cannot decode the message of transmitter 1 since it does not know whether the true message is m_1 or $\tilde{m}_1$. The same scenario can happen if $J \geq S_2$. This attack constitutes AVC symmetrization.

4.4 Achievability Proof

Our achievability proof is a generalization of the achievability proof in Chapter 4.5.1, p. 87 of ElGamal and Kim, (2011). Before proceeding to the proof, we first define the following typical set for Gaussian random variables $X_1, \ldots, X_k$ as:

$$\begin{aligned}&\mathcal{T}_\epsilon^{(n)}(X_1,\ldots,X_k)\\&=\left\{(\mathbf{x}_1,\ldots,\mathbf{x}_k) : \mathbb{E}(X_iX_j)-\epsilon \le \frac{1}{n}\langle \mathbf{x}_i,\mathbf{x}_j\rangle \le \mathbb{E}(X_iX_j)+\epsilon \text{ for all } i,j\in[k]\right\}. \quad (4.3)\end{aligned}$$

Codebook generation: Fix $\gamma > 0$. For $i = 1, 2$, we generate 2^{nR_i} i.i.d zero mean Gaussian sequences $\mathbf{X}_i(m_i)$ with variance $(1-\gamma)P_i$ for each $m_i \in [2^{nR_i}]$.

Encoding: For $i = 1, 2$, transmitter i sends $\mathbf{X}_i = \mathbf{X}_i(m_i)$ if its power is less than P_i, otherwise it sends zero.

Decoding: First, let

$$\mathscr{S} = \left\{(m_1,m_2) : (\mathbf{x}_1(m_1),\mathbf{x}_2(m_2),\mathbf{y}) \in \bigcup \mathcal{T}_\epsilon^{(n)}(X_1,X_2,Y)\right\} \quad (4.4)$$

where the union is over all joint Gaussian distributions X_1, X_2, Y such that $(X_1, X_2, Y - g_1X_1 - g_2X_2)$ are mutually independent.

Given $\mathbf{y}$, the decoder finds

$$(\hat{m}_1,\hat{m}_2) = \underset{(m_1,m_2)\in\mathscr{S}}{\arg\min} \left\|\mathbf{y} - g_1\mathbf{x}_1(m_1) - g_2\mathbf{x}_2(m_2)\right\|. \quad (4.5)$$

If there is more than one minimum, choose between them arbitrarily. The decoder then outputs the message estimate $(\hat{m}_1, \hat{m}_2)$.

Analysis of the probability of error: Assume the two users send messages (M_1, M_2). Define the error event

$$\mathcal{E}_0 = \{(M_1,M_2) \notin \mathscr{S}\}. \quad (4.6)$$

To consider error events in which a false message set appears correct, we define the set

$$\mathscr{T} = \big\{(m_1, m_2) \in \mathscr{S} : \\ \|\mathbf{Y} - g_1\mathbf{x}_1(m_1) - g_2\mathbf{x}_2(m_2)\|^2 \leq \|\mathbf{Y} - g_1\mathbf{x}_1(M_1) - g_2\mathbf{x}_2(M_2)\|^2\big\}. \quad (4.7)$$

An error can only occur if there exists some $(m_1, m_2) \in \mathscr{T}$ where $(m_1, m_2) \neq (M_{1c}, M_{1p})$. We divide this event into the following three error events:

$$\mathcal{E}_1 = \{(\tilde{m}_1, M_2) \in \mathscr{T} \text{ for some } \tilde{m}_1 \neq M_1\} \quad (4.8)$$

$$\mathcal{E}_2 = \{(M_1, \tilde{m}_2) \in \mathscr{T} \text{ for some } \tilde{m}_2 \neq M_2\} \quad (4.9)$$

$$\mathcal{E}_3 = \{(\tilde{m}_1, \tilde{m}_2) \in \mathscr{T} \text{ for some } \tilde{m}_1 \neq M_1, \tilde{m}_2 \neq M_2\}\,. \quad (4.10)$$

We will prove that the probability of each one of the error events converges to zero as long as the conditions in (4.2) are satisfied.

We now consider each of the four error events, beginning with $\mathcal{E}_0$. By the law of large numbers, $\mathbb{P}(\mathcal{E}_0)$ tends to zero as $n \to \infty$.

To bound the probability of event $\mathcal{E}_1$, we apply Lemma 5 with the following:

- $i = M_1$, $j = \tilde{m}_1$,
- $\mathbf{x}_i = g_1\mathbf{x}_1(M_1)$,
- $\mathbf{x}_j = g_1\mathbf{x}_1(\tilde{m}_1)$.

Note that event $\mathcal{E}_1$ occurs if

$$\|g_1\mathbf{X}_1(M_1) + \mathbf{s} + \mathbf{V} - g_1\mathbf{X}_1(\tilde{m}_1)\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2. \quad (4.11)$$

Thus, by Lemma 5, if $R_1 < C\left(\frac{(1-\gamma)S_1}{J+1}\right)$ and $S_1 > J$ then with high probability the codebook $\mathbf{X}_1$ will be such that $\mathbb{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$. Error event $\mathcal{E}_2$ can be bounded by the same argument but for $\mathbf{x}_i = g_2\mathbf{x}_2(M_2)$ and $\mathbf{x}_j = g_2\mathbf{x}_2(\tilde{m}_2)$ if $R_2 < C\left(\frac{(1-\gamma)S_2}{J+1}\right)$ and $S_2 > J$.

We now bound event $\mathcal{E}_3$ by applying Lemma 8 with the following particularizations:

- $i_1 = M_1$, $i_2 = M_2$, $j_1 = \tilde{m}_1$, $j_2 \; = \; (\tilde{m}_2)$,
- $\mathbf{x}_{i_1} = g_1\mathbf{x}_1(M_1)$, $\mathbf{y}_{i_2} = g_2\mathbf{x}_2(M_2)$,
- $\mathbf{x}_{j_1} = g_1\mathbf{x}_1(\tilde{m}_1)$, $\mathbf{y}_{j_2} = g_2\mathbf{x}_2(\tilde{m}_2)$.

Note that event $\mathcal{E}_3$ occurs if

$$\|g_1\mathbf{X}_1(M_1) + g_2\mathbf{X}_2(M_2) + \mathbf{s} + \mathbf{V} - g_1\mathbf{X}_1(\tilde{m}_1) - g_2\mathbf{X}_2(\tilde{m}_2)\|^2 \leq \|\mathbf{s}{+}\mathbf{V}\|^2. \quad (4.12)$$

Therefore, we can conclude by Lemma 8 that with high probability as $n \to \infty$, $\mathbb{P}(\mathcal{E}_3) \to 0$ if $J < S_1 + S_2$,

$$R_1 + R_2 < C\left(\frac{(1-\gamma)(S_1+S_2)}{J+1}\right). \quad (4.13)$$

We finally get all the equations in (4.2) as $\gamma \to 0$.

Chapter 5

GAUSSIAN ARBITRARILY-VARYING BROADCAST CHANNEL

5.1 Problem Statement

The Gaussian broadcast channel with two jammers is shown in Fig. 3, in which one transmitter sends its messages to two receivers in the presence of two independent jammers. This channel is also known as Gaussian arbitrarily-varying broadcast channel (Gaussian AVBC). The jammers are assumed not to have any information about the user's signal (but know the code). In particular, the received signals are given by

$$\begin{aligned}\mathbf{Y}_1 &= g_1\mathbf{X} + \mathbf{S}_1 + \mathbf{V}_1 \\ \mathbf{Y}_2 &= g_2\mathbf{X} + \mathbf{S}_2 + \mathbf{V}_2\end{aligned} \tag{5.1}$$

where $\mathbf{X}$ is an n-length vector representing the user's signal, $\mathbf{S}_1$ and $\mathbf{S}_2$ are the adversarial jammers' signals, g_1 and g_2 are the channel gains, and $\mathbf{V}_1$ and $\mathbf{V}_2$ are two independent n-length noise vectors distributed as two sequences of i.i.d. zero mean Gaussian random variables with variances σ_1^2 and σ_2^2 respectively. These two noise are assumed to be independent of $\mathbf{X}$, $\mathbf{S}_1$ and $\mathbf{S}_2$.

The transmitter and jammers signals are constrained to satisfy power constraints $\|\mathbf{X}\|^2 \leq nP$ and $\|\mathbf{S}_i\|^2 \leq n\Lambda$, for $i=1,2$. Without loss of generality, we assume $\frac{g_1^2}{\sigma_1^2} > \frac{g_2^2}{\sigma_2^2}$, i.e. receiver 1 is the stronger receiver from the signal-to-noise ratio perspective. Note that the vectors $\mathbf{S}_1$ and $\mathbf{S}_2$ refer to the jammers signals while the scaler values S_1 and S_2 denotes the signal-to-noise ratios. We assume that the transmitter and receivers know the signal-to-noise ratios, but they need not know the jammer-to-noise ratio. However, we require small probability of error only when the jammer-to-noise ratios

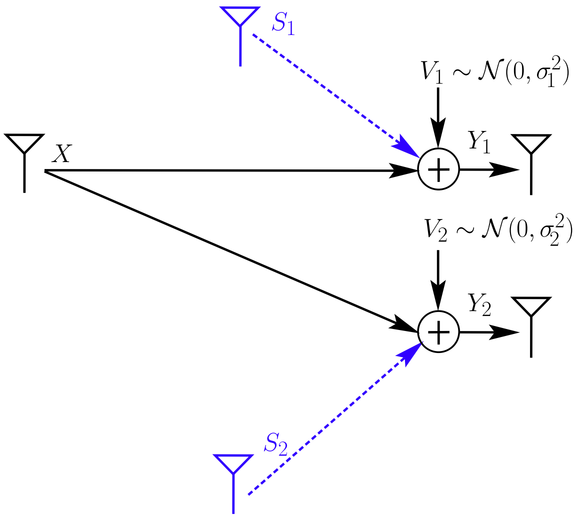

S_1
$V_1 \sim \mathcal{N}(0,\sigma_1^2)$
X
Y_1
$V_2 \sim \mathcal{N}(0,\sigma_2^2)$
Y_2
S_2

Figure 3: Two-user Gaussian Broadcast Channel with Two Jammers.

do not exceed J_1 and J_2; thus the code is independent of the jammer's power up to a point, and beyond that it may fail to decode correctly.

A $\left(2^{nR_1}, 2^{nR_2}, n\right)$ deterministic code is given by:

- Message sets $\mathcal{M}_1 = [2^{nR_1}]$ and $\mathcal{M}_2 = [2^{nR_2}]$,
- Encoding function $\mathbf{x} : (\mathcal{M}_1, \mathcal{M}_2) \to \mathbb{R}^n$, and
- Decoding functions $\phi_1 : \mathbb{R}^n \to \mathcal{M}_1$ and $\phi_2 : \mathbb{R}^n \to \mathcal{M}_2$.

For $i = 1, 2$, the message M_i is chosen uniformly from the set $\mathcal{M}_i$, and the transmitter encodes two messages to $\mathbf{X}$. At the receiver, the received signal $\mathbf{Y}_i$ is decoded by function ϕ_i to $\hat{M}_i = \phi_i(\mathbf{Y}_i)$ for $i = 1, 2$. The average probability of error $P_e^{(n)}$ is now given by the probability that $(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)$, maximized over all possible choices of jammer's sequence $\mathbf{S}$. A rate pair (R_1, R_2) is *achievable* if there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}, n\right)$ codes where $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region $\mathscr{C}$ is the closure of the set of all achievable rate pairs (R_1, R_2).

5.2 Main Results

The following theorems provides the inner and outer bounds for the capacity region of two-user Gaussian broadcast channel with two independent jammer.

Theorem 12 (outer bound) *Assume $g_1^2P > \Lambda$ and $g_2^2P > \Lambda$. If the rate pair (R_1, R_2) is achievable then*

$$\begin{aligned} R_1 &< C\left(\frac{\alpha g_1^2 P}{\Lambda + \sigma_1^2}\right) \\ R_2 &< C\left(\frac{\bar{\alpha} g_2^2 P}{\alpha g_2^2 P + \Lambda + \sigma_2^2}\right) \end{aligned} \tag{5.2}$$

for some $\alpha \in [0, 1]$.

Theorem 13 (inner bound) *Assume $g_1^2P > \Lambda$ and $g_2^2P > \Lambda$. The rate pair (R_1, R_2) is achievable if*

$$\begin{aligned} R_1 &< C\left(\frac{\alpha g_1^2 P}{\Lambda + \sigma_1^2}\right) \\ R_2 &< C\left(\frac{\bar{\alpha} g_2^2 P}{\alpha g_2^2 P + \Lambda + \sigma_2^2}\right) \end{aligned} \tag{5.3}$$

for some $\alpha \in [0, 1]$ where $\bar{\alpha} g_2^2 P > \Lambda$.

5.3 Proof of Outer Bound

Consider a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with vanishing probability of error. Since these codes must function for arbitrary jamming signals, we may assume that the jammer transmits Gaussian noise with variance Λ. Thus, we follow the capacity for the Gaussian broadcast channel with no jammer ElGamal and Kim, (2011), Chapter 5.5.1, p. 118 and the noise power $\sigma_1^2 + \Lambda$ and $\sigma_2^2 + \Lambda$ at the receiver 1 and receiver 2, respectively.

Moreover, if $\Lambda \geq g_1^2 P$, based on the assumption that the jammer knows the code, the jammer can choose two arbitrary messages $\tilde{m}_1$ and $\tilde{m}_2$ and transmit a scaled form of the corresponding codeword $\mathbf{s} = g_1 \mathbf{x}(\tilde{m}_1, \tilde{m}_2)$. Given $\mathbf{Y}_1 = g_1 \mathbf{x}_1(m_1, m_2) + g_1 \mathbf{x}_1(\tilde{m}_1, \tilde{m}_2) + \mathbf{V}_1$, the decoder cannot decode any message since it does not know whether the true message is m_1 or $\tilde{m}_1$ and the same for m_2 or $\tilde{m}_2$. The same scenario can happen if $\Lambda \geq g_2^2 P$ at receiver 2 for a given $\mathbf{Y}_2$. This attack constitutes AVC symmetrization.

5.4 Proof of Inner Bound

Before proceeding to the proof, we first define the following typical set for Gaussian random variables $X_1, \ldots, X_k$ as:

$$\begin{aligned} &\mathcal{T}_\epsilon^{(n)}(X_1, \ldots, X_k) \\ &= \left\{ (\mathbf{x}_1, \ldots, \mathbf{x}_k) : \mathbb{E}(X_i X_j) - \epsilon \leq \frac{1}{n} \langle \mathbf{x}_i, \mathbf{x}_j \rangle \leq \mathbb{E}(X_i X_j) + \epsilon \text{ for all } i, j \in [k] \right\}. \end{aligned} \tag{5.4}$$

Codebook generation: Fix $\alpha \in [0, 1]$ and $\gamma > 0$. We generate 2^{nR_2} i.i.d zero mean Gaussian sequences $\mathbf{X}_2(m_2)$ with variance $(1-\gamma)\bar{\alpha}P$ for each $m_2 \in [2^{nR_2}]$. We generate 2^{nR_1} i.i.d. zero mean Gaussian sequences $\mathbf{X}_1(m_1, m_2)$ with variance $(1-\gamma)\alpha P$ for each $m_1 \in [2^{nR_1}]$ and $m_2 \in [2^{nR_2}]$.

Encoding: The transmitter sends $\mathbf{X} = \mathbf{X}_1(m_1, m_2) + \mathbf{X}_2(m_2)$ if its power is less than P, otherwise it sends zero.

Decoding: We first describe the decoding procedure for receiver 2. First, let

$$\mathscr{S} = \left\{ m_2 : (\mathbf{x}_2(m_2), \mathbf{y}_2) \in \bigcup \mathcal{T}_\epsilon^{(n)}(X_2, Y_2) \right\} \tag{5.5}$$

where the union is over all joint Gaussian distributions X_2, Y_2 such that $(X_2, Y_2 - g_2X_2)$ are mutually independent. Given $\mathbf{y}_2$, decoder 2 finds

$$\hat{m}_2 = \underset{m_2 \in \mathscr{S}}{\arg\min} \left\| \mathbf{y}_2 - g_2\mathbf{x}_2(m_2) \right\| \tag{5.6}$$

If there is more than one minimum, choose between them arbitrarily.

Now, decoder 1's structure to find message $\hat{m}_1$ is as follows. Let

$$\mathscr{S}_1 = \left\{ (m_1, m_2) : (\mathbf{x}_1(m_1, m_2), \mathbf{x}_2(m_2), \mathbf{y}_1) \in \bigcup \mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y_1), \right\} \tag{5.7}$$

where the union is over all joint Gaussian distributions X_1, X_2, Y_1 such that $(X_1, X_2, Y_1 - g_1X_1 - g_2X_2)$ are mutually independent. Also, let

$$\mathscr{S}_2 = \left\{ m_2 : (\mathbf{x}_2(m_2), \mathbf{y}_1) \in \bigcup \mathcal{T}_\epsilon^{(n)}(X_2, Y_1), \right\} \tag{5.8}$$

where the union is over all joint Gaussian distributions X_2, Y_1 such that $(X_2, Y_1 - g_2X_2)$ are mutually independent. Given $\mathbf{y}_1$, decoder 1 first finds $\hat{m}_2$ if it is a unique message such that

$$\hat{m}_2 = \underset{m_2 \in \mathscr{S}_2}{\arg\min} \left\| \mathbf{y}_1 - g_2\mathbf{x}_2(m_2) \right\| \tag{5.9}$$

Then, if such an $\hat{m}_2$ exists, then the decoder 1 declares the unique $\hat{m}_1$ such that

$$\hat{m}_1 = \underset{(m_1, \hat{m}_2) \in \mathscr{S}_1}{\arg\min} \left\| \mathbf{y}_1 - g_1\mathbf{x}_1(m_1, \hat{m}_2) - g_2\mathbf{x}_2(\hat{m}_2) \right\| . \tag{5.10}$$

If there is more than one minimum, choose between them arbitrarily.

Analysis of the probability of error: Assume the user sends messages (M_1, M_2). We first analyze the average probability of error for decoder 2, and then for decoder 1. For decoder 2, define the following error event

$$\mathcal{E}_{20} = \{M_2 \notin \mathscr{S}\} . \tag{5.11}$$

Moreover, to consider error event in which a false message set appears correct, we define the set

$$\mathscr{T} = \left\{ m_2 \in \mathscr{S} : \|\mathbf{Y}_2 - g_2\mathbf{x}_2(m_2)\|^2 \le \|\mathbf{Y}_2 - g_2\mathbf{x}_2(M_2)\|^2 \right\}. \tag{5.12}$$

An error can only occur if there exists some $m_2 \in \mathscr{T}$ where $m_2 \neq M_2$. We only have then one event of

$$\mathcal{E}_{21} = \{\exists\, \tilde{m}_2 \neq M_2 : \tilde{m}_2 \in \mathscr{T}\}. \tag{5.13}$$

The decoder 1 makes error if at least one of the following events $\mathcal{E}_{10}$, $\mathcal{E}_{11}$ and $\mathcal{E}_{12}$ happens.

$$\mathcal{E}_{10} = \{(M_1, M_2) \notin \mathscr{S}_1\}. \tag{5.14}$$

Moreover, to consider error events in which a false message appears correct, we define the following sets

$$\mathscr{T}_1 = \left\{ (m_1, m_2) \in \mathscr{S}_1 : \|\mathbf{Y}_1 - g_1\mathbf{x}_1(m_1, m_2) - g_2\mathbf{x}_2(m_2)\|^2 \le \|\mathbf{s}_1 + \mathbf{V}_1\|^2 \right\} \tag{5.15}$$

$$\mathscr{T}_2 = \left\{ m_2 \in \mathscr{S}_2 : \|\mathbf{Y}_1 - g_2\mathbf{x}_2(m_2)\|^2 \le \|\mathbf{Y}_1 - g_2\mathbf{x}_2(M_2)\|^2 \right\}. \tag{5.16}$$

An error can occur if we have one of the following:

$$\mathcal{E}_{11} = \{\exists\, \tilde{m}_2 \neq M_2 : \tilde{m}_2 \in \mathscr{T}_2\} \tag{5.17}$$

$$\mathcal{E}_{12} = \{\exists\, \tilde{m}_1 \neq M_1 : (\tilde{m}_1, M_2) \in \mathscr{T}_1\}. \tag{5.18}$$

We now consider each of the five error events, beginning with $\mathcal{E}_{20}$ and $\mathcal{E}_{10}$. Using the law of large numbers, we conclude that with high probability both $\mathcal{E}_{10}$ and $\mathcal{E}_{20}$ tend to zero as $n \to \infty$.

To bound the probability of event $\mathcal{E}_{21}$, we apply Lemma 5 with the following:

- $i = M_2$, $j = \tilde{m}_2$,

- $\mathbf{x}_i = g_2\mathbf{x}_2(M_2)$,
- $\mathbf{x}_j = g_2\mathbf{x}_2(\tilde{m}_2)$.

Note that event $\mathcal{E}_{21}$ occurs if

$$\|g_1\mathbf{X}_1(M_1, M_2)+g_2\mathbf{X}_2(M_2)+\mathbf{s}_2+\mathbf{V}_2-g_2\mathbf{X}_2(\tilde{m}_2)\|^2 \leq \|\mathbf{s}_2+\mathbf{V}_2+g_1\mathbf{X}_1(M_1, M_2)\|^2. \tag{5.19}$$

Thus, by Lemma 5, if $R_2 < C\left(\frac{(1-\gamma)\bar{\alpha}g_2^2P}{\sigma_2^2+\Lambda+(1-\gamma)\alpha g_1^2P}\right)$ and $\Lambda < \bar{\alpha}g_2^2P$ then with high probability $\mathbb{P}(\mathcal{E}_{21}) \to 0$ as $n \to \infty$.

To bound the probability of event $\mathcal{E}_{11}$, we apply Lemma 5 with the following:

- $i = M_2$, $j = \tilde{m}_2$,
- $\mathbf{x}_i = g_2\mathbf{x}_2(M_2)$,
- $\mathbf{x}_j = g_2\mathbf{x}_2(\tilde{m}_2)$

Note that event $\mathcal{E}_{11}$ occurs if

$$\|g_1\mathbf{X}_1(M_1, M_2)+g_2\mathbf{X}_2(M_2) + \mathbf{s}_1+\mathbf{V}_1-g_2\mathbf{X}_2(\tilde{m}_2)\|^2 \leq \|g_1\mathbf{X}_1(M_1, M_2) + \mathbf{s}_1 + \mathbf{V}_1\|^2. \tag{5.20}$$

Thus, by Lemma 5, if $R_2 < C\left(\frac{(1-\gamma)\bar{\alpha}g_2^2P}{\sigma_1^2+\Lambda+(1-\gamma)\alpha g_1^2P}\right)$ and $\Lambda < \bar{\alpha}g_2^2P$, then with high probability $\mathbb{P}(\mathcal{E}_{11}) \to 0$ as $n \to \infty$.

To bound the probability of event $\mathcal{E}_{12}$, we apply Lemma 6 with the following:

- $i = M_1$, $j = \tilde{m}_1$, $k = M_2$,
- $\mathbf{x}_i(k) = g_1\mathbf{x}_1(M_1, M_2) + g_2\mathbf{x}(M_2)$,
- $\mathbf{x}_j(k) = g_1\mathbf{x}_1(\tilde{m}_1, M_2) + g_2\mathbf{x}(M_2)$

In this case, $K = 2^{nR_2} \geq n^2$ for sufficiently large n as long as $R_2 > 0$. Note that event $\mathcal{E}_{12}$ occurs if

$$\|g_1\mathbf{X}_1(M_1, M_2) + \mathbf{s}_1 + \mathbf{V}_1 - g_1\mathbf{X}_1(\tilde{m}_1, M_2)\|^2 \leq \|\mathbf{s}_1 + \mathbf{V}_1\|^2. \tag{5.21}$$

Thus, by Lemma 6, if $R_1 < C\left(\frac{(1-\gamma)\alpha g_1^2 P}{\sigma_1^2+\Lambda}\right)$, then with high probability $\mathbb{P}(\mathcal{E}_{12}) \to 0$ as $n \to \infty$. Thus, we get all equations in (5.3) as $\gamma \to 0$.

Chapter 6

GAUSSIAN ARBITRARILY-VARYING INTERFERENCE CHANNEL

6.1 Problem Statement

The Gaussian interference channel with two independent jammers is shown in Fig. 4, in which two users send their messages to their own receivers in the presence of one or two jammers. The jammers are assumed not to have any information about the user's signals (but know the code). This channel is also known as Gaussian arbitrarily-varying interference channel (Gaussian AVIC). In particular, the received signals are given by

$$\begin{aligned} \mathbf{Y}_1 &= h_{11}\mathbf{X}_1 + h_{12}\mathbf{X}_2 + g_1\mathbf{W}_1 + \mathbf{V}_1 \\ \mathbf{Y}_2 &= h_{21}\mathbf{X}_1 + h_{22}\mathbf{X}_2 + g_2\mathbf{W}_2 + \mathbf{V}_2 \end{aligned} \tag{6.1}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are n-length vectors representing the user's signals, $\mathbf{W}_1$ and $\mathbf{W}_2$ are the independent adversarial jammer signals, h_{ij} and g_i for $i, j \in \{1, 2\}$ are the channel gains, and $\mathbf{V}_i$ is the n-length noise vector distributed as a sequence of i.i.d. zero mean Gaussian random variables with variance σ^2 which is independent of $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{W}_1$ and $\mathbf{W}_2$.

The transmitter and jammer signals are constrained to satisfy power constraints $\|\mathbf{X}_i\|^2 \leq \ nP_i$ and $\|\mathbf{W}_i\|^2 \leq n\Lambda$, for $i = 1, 2$, respectively. We define the received signal-to-noise and interference-to-noise ratios as $S_1 = h_{11}^2 P_1/\sigma^2$, $S_2 = h_{22}^2 P_2/\sigma^2$, $I_1 = h_{12}^2 P_2/\sigma^2$ and $I_2 = h_{21}^2 P_1/\sigma^2$. We also denote the jammer-to-noise ratios as $J_1 = g_1^2\Lambda/\sigma^2$ and $J_2 = g_2^2\Lambda/\sigma^2$. We assume that the transmitters and receivers know the signal-to-noise and interference-to-noise ratios, but they need not know the

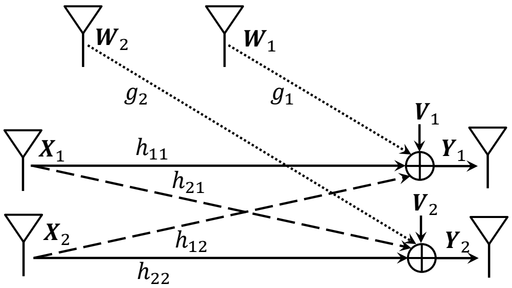


Figure 4: Two-user Gaussian Interference Channel with Two Independent Jammers.

jammer-to-noise ratios. However, we require small probability of error only when the jammer-to-noise ratios do not exceed J_1,J_2; thus the code is independent of the jammer's power up to a point, and beyond that it may fail to decode correctly.

A $\left(2^{nR_1}, 2^{nR_2}, n\right)$ deterministic code is given by:

- Message sets $\mathcal{M}_1 = [2^{nR_1}]$ and $\mathcal{M}_2 = [2^{nR_2}]$,
- Encoding functions $\mathbf{x}_i : \mathcal{M}_i \rightarrow \mathbb{R}^n$ for $i = 1, 2$, and
- Decoding functions $\phi_i : \mathbb{R}^n \rightarrow \mathcal{M}_i$ for $i = 1, 2$.

For $i = 1, 2$, the message M_i is chosen uniformly from the set $\mathcal{M}_i$, and each transmitter encodes its own message to $\mathbf{X}_i$. At each receiver, the received signal $\mathbf{Y}_i$ is decoded by function ϕ_i to $\hat{M}_i = \phi_i(\mathbf{Y}_i)$. The average probability of error $P_e^{(n)}$ is now given by the probability that $(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)$, maximized over all possible choices of jammers' sequences $\mathbf{W}_1$ and $\mathbf{W}_2$. A rate pair (R_1, R_2) is *achievable* if there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}, n\right)$ codes where $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region $\mathscr{C}$ is the closure of the set of all achievable rate pairs (R_1, R_2).

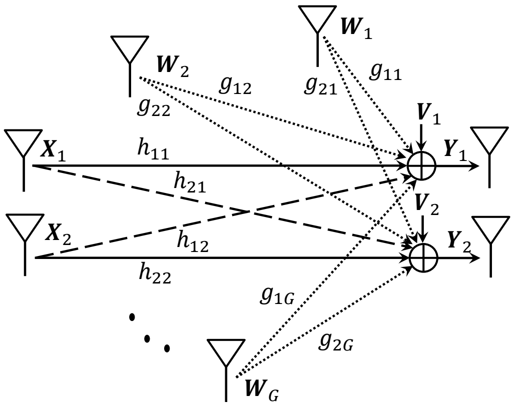


Figure 5: Two-user Gaussian Interference Channel with G Independent Jammers.

6.1.1 Generalized Jamming Model

Generally speaking, if there are G jammers ($G \geq 1$) with the cross matrix $\mathscr{G}$

$$\mathscr{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1G} \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2G} \end{bmatrix} \tag{6.2}$$

as shown in Fig. 5, then the received signals at each decoder are given by

$$\begin{aligned} \mathbf{Y}_1 &= h_{11}\mathbf{X}_1 + h_{12}\mathbf{X}_2 + g_{11}\mathbf{W}_1 + g_{12}\mathbf{W}_2 + \ldots + g_{1G}\mathbf{W}_G + \mathbf{V}_1 \\ \mathbf{Y}_2 &= h_{21}\mathbf{X}_1 + h_{22}\mathbf{X}_2 + g_{21}\mathbf{W}_1 + g_{22}\mathbf{W}_2 + \ldots + g_{2G}\mathbf{W}_G + \mathbf{V}_2 \end{aligned} \tag{6.3}$$

where $\|\mathbf{W}_i\|^2 \leq n\Lambda$ for $i = 1, 2, \ldots, G$. This includes the case where there is only one jammer ($G = 1$). We refer the capacity region of this channel as $\mathscr{C}_G$, and state the following proposition for the relation between $\mathscr{C}_G$ and $\mathscr{C}$. Indeed, the capacity region depends only on the received signal at each decoder and not the number of jammers.

Proposition 14 *We have $\mathscr{C}_G = \mathscr{C}$ as long as*

$$\begin{aligned} |g_{11}| + |g_{12}| + \ldots + |g_{1G}| &= |g_1| \\ |g_{21}| + |g_{22}| + \ldots + |g_{2G}| &= |g_2| \end{aligned} \tag{6.4}$$

where the jammer-to-noise ratios are then given by $J_1 = (g_{11} + g_{12} + \ldots + g_{1G})^2 \Lambda/\sigma^2$ *and* $J_2 = (g_{21} + g_{22} + \ldots + g_{2G})^2 \Lambda/\sigma^2$.

The proof is provided in Appendix 6.6.

6.2 Main Results

In this section, we present inner and outer bounds on the capacity region $\mathscr{C}$ (two-user Gaussian interference channel with two independent jammers). Before stating the main results, we define regions $\mathscr{R}_o(S_1, S_2, I_1, I_2)$ and $\mathscr{R}_i(S_1, S_2, I_1, I_2)$ as the previously-derived outer and inner bounds respectively for the Gaussian interference channel with no jammer; namely $\mathscr{R}_o$ is the outer bound of Etkin, Tse, and Wang, (2008), and $\mathscr{R}_i$ is the Han-Kobayashi inner bound Han and Kobayashi, (1981). When we write an expression with i and j, we mean for it to hold for both $(i, j) = (1, 2)$ and $(i, j) = (2, 1)$.

Define $\mathscr{R}_o(S_1, S_2, I_1, I_2)$ as the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_i &\leq C\left(S_i\right) \\ R_i + R_j &\leq C\left(\tfrac{S_i}{1+I_j}\right) + C\left(I_j + S_j\right) \\ R_1 + R_2 &\leq C\left(\tfrac{S_1+I_1+I_1 I_2}{1+I_2}\right) + C\left(\tfrac{S_2+I_2+I_1 I_2}{1+I_1}\right) \\ 2R_i + R_j &\leq C\left(\tfrac{S_i}{1+I_j}\right) + C\left(S_i + I_i\right) + C\left(\tfrac{S_j+I_j+I_i I_j}{1+I_i}\right). \end{aligned}$$

Define $\mathscr{R}_i(S_1, S_2, I_1, I_2)$ as the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_i &< C\left(\tfrac{S_i}{1+\alpha_j I_i}\right) \\ R_i + R_j &< C\left(\tfrac{S_i+\bar{\alpha}_j I_i}{1+\alpha_j I_i}\right) + C\left(\tfrac{\alpha_j S_j}{1+\alpha_i I_j}\right) \\ R_1 + R_2 &< C\left(\tfrac{\alpha_1 S_1+\bar{\alpha}_2 I_1}{1+\alpha_2 I_1}\right) + C\left(\tfrac{\alpha_2 S_2+\bar{\alpha}_1 I_2}{1+\alpha_1 I_2}\right) \\ 2R_i + R_j &< C\left(\tfrac{S_i+\bar{\alpha}_j I_i}{1+\alpha_j I_i}\right)+C\left(\tfrac{\alpha_i S_i}{1+\alpha_j I_i}\right)+C\left(\tfrac{\alpha_j S_j+\bar{\alpha}_i I_j}{1+\alpha_i I_j}\right) \end{aligned}$$

for some α_i in $[0,1]$ where α_i implies the portion of the private message power in the Han-Kobayashi inner bound proof at user i. Note that in the Han-Kobayashi inner bound proof encoder i divides the message m_i into private message m_{ip} and common message m_{ic} with power $\alpha_i P_i$ and $\bar{\alpha}_i P_i$ respectively.

Define $S_i' = \frac{S_i}{1+J_i}$ and $I_i' = \frac{I_i}{1+J_i}$. We now state our main outer and inner bounds.

Theorem 15 (Outer Bound) *$\mathscr{C} \subseteq \mathscr{R}_o(S_1', S_2', I_1', I_2')$. Moreover, if $S_1 \leq J_1$ or $S_2 \leq J_2$, then $\mathscr{C} = \emptyset$.*

Theorem 16 (Inner Bound) *Assume $S_i > J_i$ for $i = 1, 2$. Let $\tilde{\mathscr{R}}_i(S_1', S_2', I_1', I_2')$ be the subset of rate pairs in $\mathscr{R}_i(S_1', S_2', I_1', I_2')$ achieved by $\alpha_i \in [0,1]$ satisfying*

$$\alpha_i S_i + \bar{\alpha}_j I_i > J_i \text{ for } (i,j) = (1,2),(2,1). \tag{6.5}$$

Then $\tilde{\mathscr{R}}_i(S_1', S_2', I_1', I_2') \subseteq \mathscr{C}$.

Note that we are also able to remove power constraint 6.5 in the inner bound by using Lemma 7 in the proof.

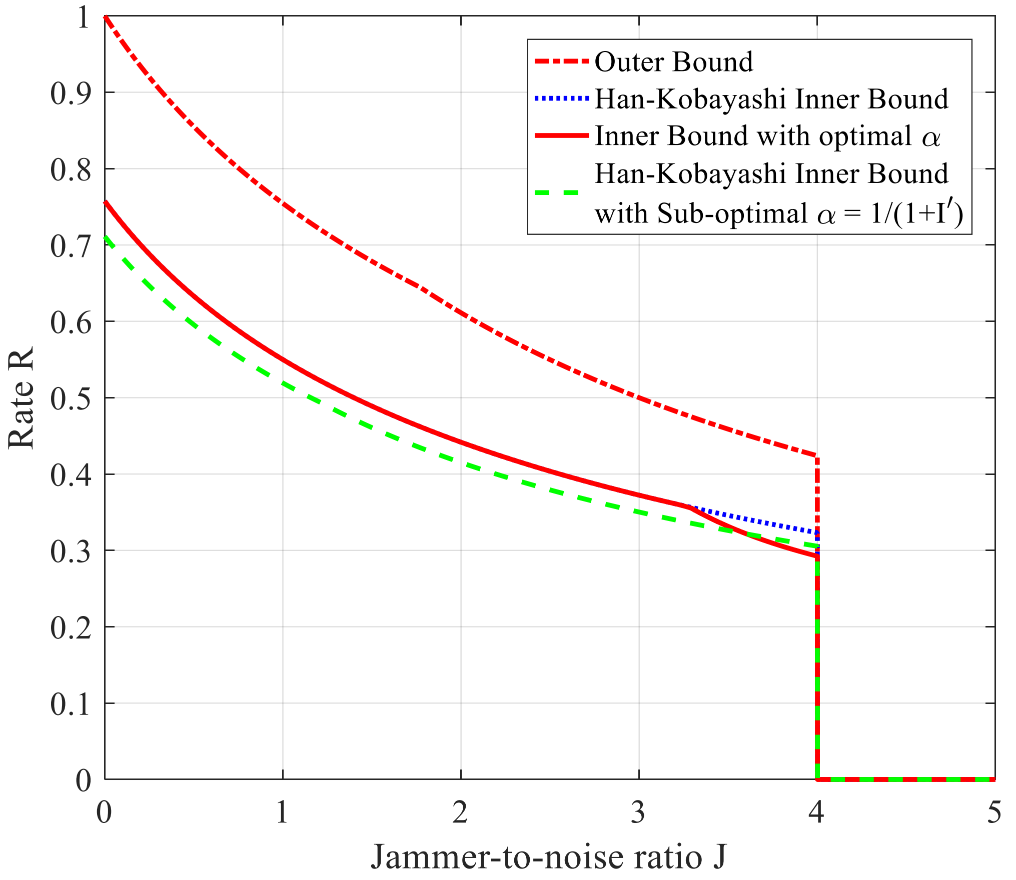
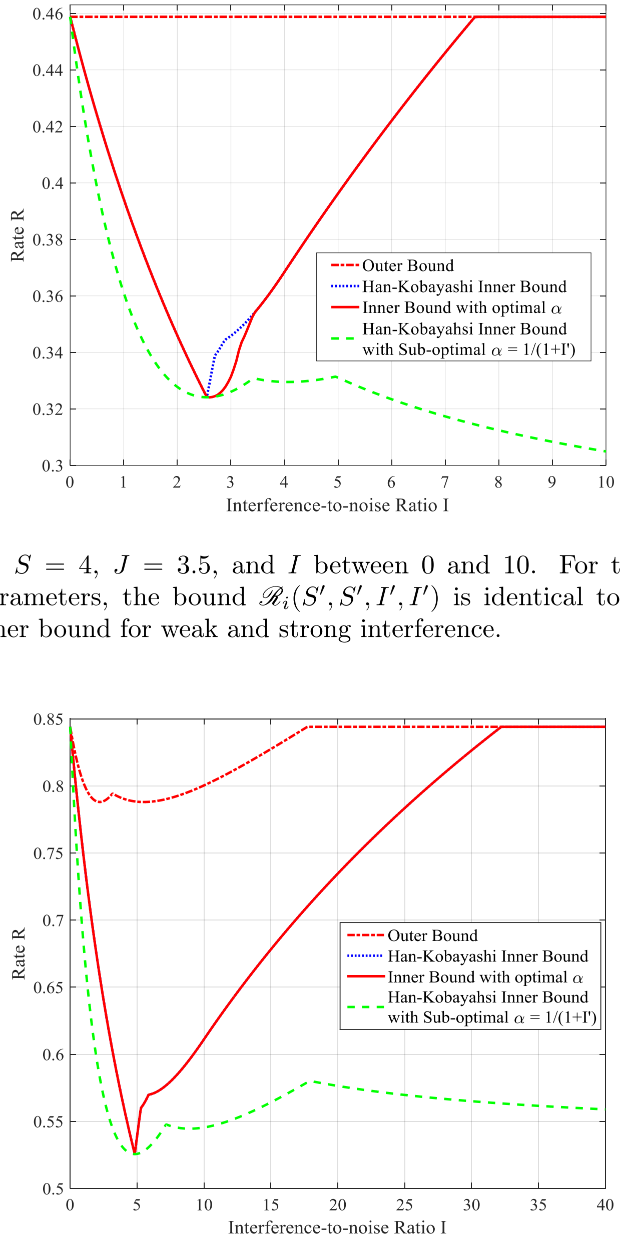


Figure 6: Bounds on the symmetric capacity $\mathscr{C}_{\text{sym}}(S, I, J)$ for $S = 4$, $I = 3$, and J between 0 and 5.

Note: Bounds on the symmetric capacity $\mathscr{C}_{\text{sym}}(S, I, J)$ for $S_1 = S_2 = S = 4$, $I_1 = I_2 = I = 3$, and $J_1 = J_2 = J$ between 0 and 5. In addition to our inner $\tilde{\mathscr{R}}_i(S', S', I', I')$ and outer bounds $\mathscr{R}_o(S', S', I', I')$, also shown the bound $\mathscr{R}_i(S', S', I', I')$ and shown $\mathscr{R}_i(S', S', I', I')$ with sub-optimal $\alpha = \frac{1}{1+I'}$. For these parameters, the bound $\mathscr{R}_i(S', S', I', I')$ is identical to our inner bound if the jammer-to-noise ratio is less than 3.2.

6.3 Discussion and Numerical Results

Note that the inner bound differs from $\mathscr{R}_i(S_1', S_2', I_1', I_2')$ only when the optimal α_i parameters do not satisfy (6.5). However, in several regimes of interest, this constraint



(a) $S = 4$, $J = 3.5$, and I between 0 and 10. For these parameters, the bound $\mathscr{R}_i(S', S', I', I')$ is identical to our inner bound for weak and strong interference.

(b) $S = 10$, $J = 3.5$, and I between 0 and 40. For these parameters, the bound $\mathscr{R}_i(S', S', I', I')$ is identical to our inner bound for high signal-to-noise ratio $S = 10$.

Figure 7: Bounds on the symmetric capacity $\mathscr{C}_{\text{sym}}(S, I, J)$ for fixed $S_1 = S_2 = S$, $J_1 = J_2 = J$, and an interval $I_1 = I_2 = I$

Note: In addition to our inner $\tilde{\mathscr{R}}_i(S', S', I', I')$ and outer bounds $\mathscr{R}_o(S', S', I', I')$, also shown the bound $\mathscr{R}_i(S', S', I', I')$ and shown $\mathscr{R}_i(S', S', I', I')$ with sub-optimal $\alpha = \frac{1}{1+I'}$.

is not active. For example, if the channel has weak interference, in the sense of ElGamal and Kim, (2011), eq (6.8) $\sqrt{\frac{I'_j}{S'_i}}(1+I'_i) \leq \rho_i(1-\rho_j)$ for some $\rho_1, \rho_2 \in [0,1]$ and $(i,j) = (1,2),(2,1)$ then treating interference as noise is optimal for the sum-rate Theorem 6.3 of ElGamal and Kim, (2011). Treating interference as noise corresponds to $\alpha_1 = \alpha_2 = 1$, under which (6.5) holds. Therefore, in the weak interference regime, our inner bound matches $\mathscr{R}_i(S'_1, S'_2, I'_1, I'_2)$, and it also achieves the exact sum-rate capacity. On the other hand, when the channel has strong interference in both users $I'_2 \geq S'_1$ and $I'_1 \geq S'_2$, by choosing $\alpha_1 = \alpha_2 = 0$ each transmitter only sends its own common message, and both messages can be decoded at each receiver. Therefore, (6.5) holds if we have $I_1 > J_1$ and $I_2 > J_2$. Thus, we obtain the exact capacity region for the strong interference regime Theorem 6.2 of ElGamal and Kim, (2011).

In the Theorem 6.6 of ElGamal and Kim, (2011), it is shown that using Han-Kobayashi inner bound with sub-optimal choices $\alpha_1 = \frac{1}{1+I'_2}$ and $\alpha_2 = \frac{1}{1+I'_1}$ yields an inner bound that is always within half a bit of the outer bound. Therefore, if $\alpha_1 = \frac{1}{1+I'_2}$ and $\alpha_2 = \frac{1}{1+I'_1}$ satisfy our conditions in (6.5) then our inner bound is guaranteed to be within half a bit of our outer bound; that is, if $J_1 < \frac{S_1}{1+I'_2} + \frac{I_1^2}{1+I'_1}$ and $J_2 < \frac{S_2}{1+I'_1} + \frac{I_2^2}{1+I'_2}$, our inner and outer bounds are within half a bit. However, we are now able to prove that our inner and outer bounds are always within half a bit by using Lemma 7 in the proof of the inner bound.

Now, consider the symmetric case; i.e. $S_1 = S_2 = S$, $I_1 = I_2 = I$, $J_1 = J_2 = J$, and $R_1 = R_2 = R$. Clearly in this case it is optimal to choose $\alpha_1 = \alpha_2 = \alpha$ for inner bound. Define the symmetric capacity of the channel as $C_{\text{sym}}(S,I,J) = \max\{R : (R,R) \in \mathscr{C}\}$. We illustrate the bounds for $C_{\text{sym}}(S,I,J)$ in Fig. 6 and Fig. 7 including our outer bound $\mathscr{R}_o(S',S',I',I')$, our inner bound $\tilde{\mathscr{R}}_i(S',S',I',I')$ with optimal α, the Han-Kobayashi inner bound with the noise variance increased by the received power of the jammer $\mathscr{R}_i(S', S', I', I')$ and the latter with sub-optimal $\alpha = \frac{1}{1+I'}$. Note that we are in the strong interference regime only if $I' \geq S'$.

Define the normalized symmetric capacity as $d_{\text{sym}} = \frac{C_{\text{sym}}(S,I,J)}{C(S)}$. Then the symmetric degrees of freedom (DoF) d^*_{sym} is given by

$$d^*_{\text{sym}}(\beta, \delta) = \lim_{S \to \infty} \frac{C_{\text{sym}}(S, S^\beta, S^\delta)}{C(S)}. \tag{6.6}$$

By substituting $I = S^\beta$ and $J = S^\delta$ in our outer and inner bounds $\mathscr{R}_o(S', S', I', I')$ and $\tilde{\mathscr{R}}_i(S', S', I', I')$, we find the upper bound for $C_{\text{sym}}(S, S^\beta, S^\delta)$ given by

$$\begin{aligned} C_{\text{sym}}(S, S^\beta, S^\delta) &\leq \max\{R : (R, R) \in \mathscr{R}_o(S', S', I', I')\} = \\ &\min\Bigg\{ C\left(\tfrac{S}{1+S^\delta}\right), \frac{1}{2}C\left(\tfrac{S}{1+S^\delta+S^\beta}\right) + \frac{1}{2}C\left(\tfrac{S+S^\beta}{1+S^\delta}\right), C\left(\tfrac{S+S^\beta+\frac{S^{2\beta}}{1+S^\delta}}{1+S^\delta+S^\beta}\right), \\ &\qquad \frac{1}{3}C\left(\tfrac{S}{1+S^\delta+S^\beta}\right) + \frac{1}{3}C\left(\tfrac{S+S^\beta}{1+S^\delta}\right) + \frac{1}{3}C\left(\tfrac{S+S^\beta+\frac{S^{2\beta}}{1+S^\delta}}{1+S^\delta+S^\beta}\right) \Bigg\}, \end{aligned} \tag{6.7}$$

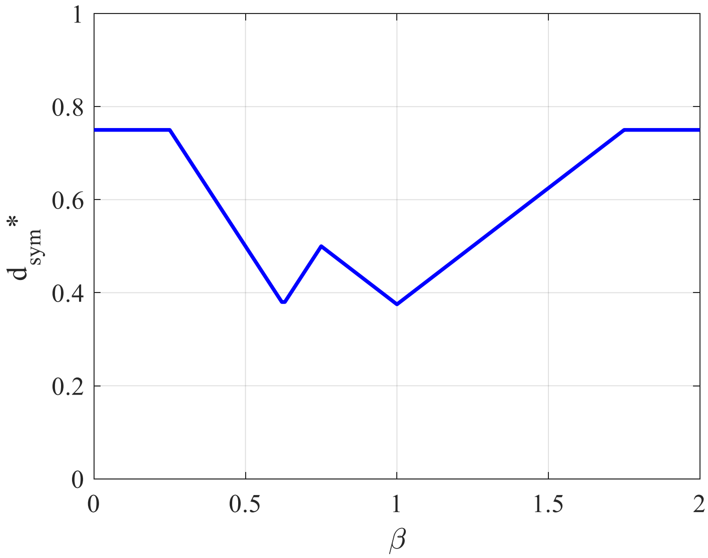
and the lower bound for $C_{\text{sym}}(S, S^\beta, S^\delta)$ given by

$$\begin{aligned} C_{\text{sym}}(S, S^\beta, S^\delta) &\geq \max\{R : (R, R) \in \tilde{\mathscr{R}}_i(S', S', I', I')\} = \\ &\max_{\alpha : \alpha S + \bar{\alpha} S^\beta > S^\delta} \Bigg\{ \min\Bigg\{ C\left(\tfrac{S}{1+S^\delta+\alpha S^\beta}\right), \frac{1}{2}C\left(\tfrac{S+\bar{\alpha}S^\beta}{1+S^\delta+\alpha S^\beta}\right) + \frac{1}{2}C\left(\tfrac{\alpha S}{1+S^\delta+\alpha S^\beta}\right), \\ &C\left(\tfrac{\alpha S+\bar{\alpha}S^\beta}{1+S^\delta+\alpha S^\beta}\right), \frac{1}{3}C\left(\tfrac{S+\bar{\alpha}S^\beta}{1+S^\delta+\alpha S^\beta}\right) + \frac{1}{3}C\left(\tfrac{\alpha S}{1+S^\delta+\alpha S^\beta}\right) + \frac{1}{3}C\left(\tfrac{\alpha S+\bar{\alpha}S^\beta}{1+S^\delta+\alpha S^\beta}\right) \Bigg\}\Bigg\}. \end{aligned} \tag{6.8}$$

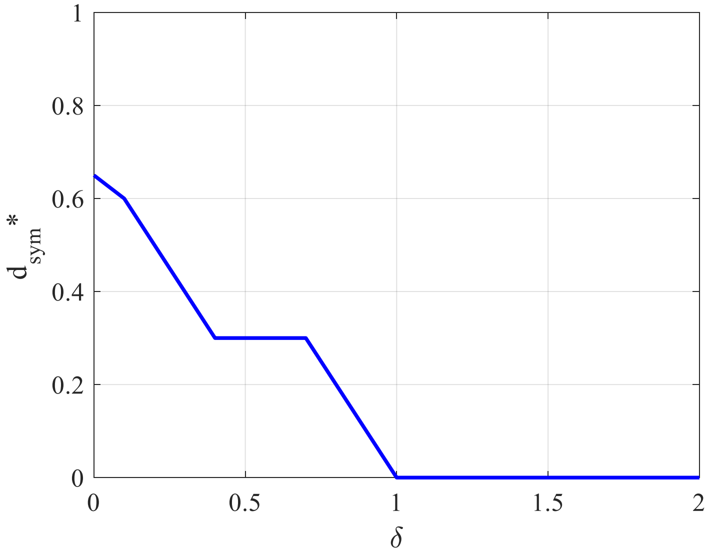
We may further lower bound the symmetric capacity by choosing $\alpha = \frac{1}{1+I'} = \frac{1+S^\delta}{1+S^\delta+S^\beta}$ as long as this choice satisfies (6.5). In particular, we claim that this value of α always satisfies (6.5) for sufficiently large S. We show this by substituting this value of α to find

$$\alpha S + \bar{\alpha} S^\beta = \frac{1+S^\delta}{1+S^\delta+S^\beta} S + \frac{S^\beta}{1+S^\delta+S^\beta} S^\beta \tag{6.9}$$

$$= \frac{S + S^{1+\delta} + S^{2\beta}}{1+S^\delta+S^\beta}. \tag{6.10}$$



(a) DoF for $\delta = \frac{1}{4}$ and β between $0, 2$.



(b) DoF for $\beta = 0.7$ and δ between $0, 2$.

Figure 8: Symmetric degrees of freedom for the GIC.

Since the capacity region is empty when $S \leq J = S^{\delta}$, it suffices to consider (6.10) only for $\delta < 1$. The dominant power of S in (6.10) is given by

$$\max\{1+\delta, 2\beta\} - \max\{\delta, \beta\} > \max\{2\delta, 2\beta\} - \max\{\delta, \beta\} = \max\{\delta, \beta\} \geq \delta. \quad (6.11)$$

Therefore, (6.10) is larger than $J = S^{\delta}$ for sufficiently large S, thus $\alpha = \frac{1}{1+I'}$ satisfies (6.5). Now, we may substitute this choice of α into (6.8), and take the limits of (6.8) and (6.7) as $S \to \infty$. Therefore, we find that the symmetric DoF is given by

$$\begin{aligned} &d^*_{\text{sym}}(\beta, \delta) = \\ &\quad \min\left\{\max\{0, 1-\delta\}, \max\left\{0, 1-\beta, \beta-\delta\right\}, \max\left\{0, 1-\frac{\beta}{2}-\frac{\delta}{2}, \frac{\beta}{2}-\frac{\delta}{2}\right\}\right\} \quad (6.12) \end{aligned}$$

which is illustrated for fixed $\delta = 1/4$ in Fig. 8a and fixed $\beta = 0.7$ in Fig. 8b. Note that for the interference channel with no jammer ElGamal and Kim, (2011), p. 153, the DoF exhibits a "W" shape for a fixed $\delta = \frac{\log J}{\log S}$.

6.4 Proof of Outer Bound

Consider a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with vanishing probability of error. Since these codes must function for arbitrary jamming signals, we may derive an outer bound by assuming the jammers transmit Gaussian noise with variance Λ. Thus, we follow the outer bound for the Gaussian interference channel with no jammer Chapter 6.7.2, p. 151 and the noise power $\sigma^2 + g_i^2\Lambda$. This yields the outer bound $\mathscr{R}_o(S_1', S_2', I_1', I_2')$.

Moreover, if $J_1 \geq S_1$, based on the assumption that the jammer knows the code, the jammer can choose an arbitrary message $\tilde{m}_1$ and transmit a scaled form of the corresponding codeword $\mathbf{w}_1 = \mathbf{x}_1(\tilde{m}_1)h_{11}/g_1$. Given $\mathbf{Y}_1 = h_{11}\mathbf{x}_1(m_1) + h_{12}\mathbf{x}_2(m_2) + h_{11}\mathbf{x}_1(\tilde{m}_1) + \mathbf{V}_1$, decoder 1 cannot decode the message since it does not know whether the true message is m_1 or $\tilde{m}_1$. The same scenario can happen for decoder 2 if $J_2 \geq S_2$. This attack constitutes AVC symmetrization.

6.5 Proof of Inner Bound

Our inner bound proof is a generalization of the Han-Kobayashi bound Chapter 6.5.1, p. 144 inElGamal and Kim, (2011). Using rate splitting, we represent message m_i from user i for $i = 1, 2$, by independent common message m_{ic} at rate R_{ic} and private message m_{ip} at rate R_{ip} such that $R_i = R_{ic} + R_{ip}$. Thus, each receiver will decode its own common and private messages and the common message of the interfering user. Assuming $S_i > J_i$ for $i = 1, 2$, we show that $(R_{1c}, R_{1p}, R_{2c}, R_{2p})$ is achievable if

$$\begin{aligned} R_{ip} &< C\left(\tfrac{\alpha_i S_i}{1+J_i+\alpha_j I_i}\right) \\ R_{ip} + R_{ic} &< C\left(\tfrac{S_i}{1+J_i+\alpha_j I_i}\right) \\ R_{ip} + R_{jc} &< C\left(\tfrac{\alpha_i S_i+\bar{\alpha}_j I_i}{1+J_i+\alpha_j I_i}\right) \\ R_{ip} + R_{ic} + R_{jc} &< C\left(\tfrac{S_i+\bar{\alpha}_j I_i}{1+J_i+\alpha_j I_i}\right) \end{aligned} \tag{6.13}$$

for some $\alpha_i \in [0, 1]$ satisfying $\alpha_i S_i + \bar{\alpha}_j I_i > J_i$, and again the above holds for $(i, j) = (1, 2)$ and $(i, j) = (2, 1)$. This achieves the region $\tilde{\mathscr{R}}_i$ by substituting $R_1 = R_{1c} + R_{1p}$ and $R_2 = R_{2c} + R_{2p}$, and applying the Fourier-Motzkin procedure to eliminate R_{ic} and R_{ip}.

Before proceeding to the proof, we first define the following typical set for Gaussian random variables $X_1, \dots, X_k$ as:

$$\begin{aligned} &\mathcal{T}_\epsilon^{(n)}(X_1, \dots, X_k) \\ &\quad = \left\{(\mathbf{x}_1, \dots, \mathbf{x}_k) : \mathbb{E}(X_i X_j) - \epsilon \leq \frac{1}{n}\langle \mathbf{x}_i, \mathbf{x}_j\rangle \leq \mathbb{E}(X_i X_j) + \epsilon \text{ for all } i, j \in [k]\right\}. \end{aligned} \tag{6.14}$$

Codebook generation: Fix $\alpha_1, \alpha_2 \in [0,1]$ and $\gamma > 0$. For $i = 1, 2$, we generate $2^{nR_{ic}}$ i.i.d zero mean Gaussian sequences $\mathbf{X}_{ic}(m_{ic})$ with variance $(1-\gamma)\bar{\alpha}_i P_i$ for each $m_{ic} \in [2^{nR_{ic}}]$. Also, for each $m_{ic} \in [2^{nR_{ic}}]$, generate $2^{nR_{ip}}$ i.i.d. zero mean Gaussian sequences $\mathbf{X}_{ip}(m_{ic}, m_{ip})$ with variance $(1-\gamma)\alpha_i P_i$ for each $m_{ip} \in [2^{nR_{ip}}]$ for $i = 1, 2$.

Encoding: For $i = 1, 2$, write message m_i as (m_{ic}, m_{ip}) where $m_{ic} \in [2^{nR_{ic}}]$ and $m_{ip} \in [2^{nR_{ip}}]$. Transmitter i sends $\mathbf{X}_i = \mathbf{X}_{ic}(m_{ic}) + \mathbf{X}_{ip}(m_{ic}, m_{ip})$ if its power is less than P_i, otherwise it sends zero.

Decoding: We describe the decoding procedure for receiver 1; that of receiver 2 is similar. First, let

$$\mathscr{S} = \\ \Big\{(m_{1c}, m_{1p}, m_{2c}) : (\mathbf{x}_{1c}(m_{1c}), \mathbf{x}_{1p}(m_{1c}, m_{1p}), \mathbf{x}_{2c}(m_{2c}), \mathbf{y}_1) \in \bigcup \mathcal{T}_\epsilon^{(n)}(X_{1c}, X_{1p}, X_{2c}, Y_1)\Big\} \tag{6.15}$$

where the union is over all joint Gaussian distributions $X_{1c}, X_{1p}, X_{2c}, Y_1$ such that $(X_{1c}, X_{1p}, X_{2c}, Y_1 - h_{11}X_{1c} - h_{11}X_{1p} - h_{12}X_{2c})$ are mutually independent.

Given $\mathbf{y}_1$, decoder 1 finds

$$(\hat{m}_{1c}, \hat{m}_{1p}, \hat{m}_{2c}) = \operatorname*{arg\,min}_{(m_{1c}, m_{1p}, m_{2c}) \in \mathscr{S}} \left\| \mathbf{y}_1 - h_{11}\mathbf{x}_{1c}(m_{1c}) - h_{11}\mathbf{x}_{1p}(m_{1c}, m_{1p}) - h_{12}\mathbf{x}_{2c}(m_{2c}) \right\|. \tag{6.16}$$

If there is more than one minimum, choose between them arbitrarily. The decoder then outputs the message estimate $\hat{m}_1 = (\hat{m}_{1c}, \hat{m}_{1p})$.

Analysis of the probability of error: Assume the two users send messages $\big((M_{1c}, M_{1p}), (M_{2c}, M_{2p})\big)$. We will obtain the average probability of error for decoder 1 and similarly generalize the results for decoder 2. Define the error event

$$\mathcal{E}_0 = \{(M_{1c}, M_{1p}, M_{2c}) \notin \mathscr{S}\}. \tag{6.17}$$

To consider error events in which a false message set appears correct, we define the set

$$\begin{aligned}\mathscr{T} = \big\{(m_{1c}, m_{1p}, m_{2c}) \in \mathscr{S} : \|\mathbf{Y}_1 - h_{11}\mathbf{x}_{1c}(m_{1c}) - h_{11}\mathbf{x}_{1p}(m_{1c}, m_{1p}) - h_{12}\mathbf{x}_{2c}(m_{2c})\|^2 \\ \leq \|\mathbf{Y}_1 - h_{11}\mathbf{x}_{1c}(M_{1c}) - h_{11}\mathbf{x}_{1p}(M_{1c}, M_{1p}) - h_{12}\mathbf{x}_{2c}(M_{2c})\|^2\big\}. \quad (6.18)\end{aligned}$$

An error can only occur if there exists some $(m_{1c}, m_{1p}, m_{2c}) \in \mathscr{T}$ where $(m_{1c}, m_{1p}) \neq (M_{1c}, M_{1p})$. We divide this event into the following 4 error events:

$$\mathcal{E}_1 = \{\exists\, \tilde{m}_{1p} \neq M_{1p} : (M_{1c}, \tilde{m}_{1p}, M_{2c}) \in \mathscr{T}\} \quad (6.19)$$

$$\mathcal{E}_2 = \{\exists\, \tilde{m}_{1c} \neq M_{1c}, \tilde{m}_{1p} : (\tilde{m}_{1c}, \tilde{m}_{1p}, M_{2c}) \in \mathscr{T}\} \quad (6.20)$$

$$\mathcal{E}_3 = \{\exists\, \tilde{m}_{1p} \neq M_{1p}, \tilde{m}_{2c} \neq M_{2c} : (M_{1c}, \tilde{m}_{1p}, \tilde{m}_{2c}) \in \mathscr{T}\} \quad (6.21)$$

$$\mathcal{E}_4 = \{\exists\, \tilde{m}_{1c} \neq M_{1c}, \tilde{m}_{1p}, \tilde{m}_{2c} \neq M_{2c} : (\tilde{m}_{1c}, \tilde{m}_{1p}, \tilde{m}_{2c}) \in \mathscr{T}\}. \quad (6.22)$$

We will prove that the probability of each one of the error events converges to zero as long as the conditions in (6.13) are satisfied.

We now consider each of the five error events, beginning with $\mathcal{E}_0$. Define set $\mathcal{T}_\epsilon^{(n,k)}$ as $\bigcup \mathcal{T}_\epsilon^{(n)}(X_1, \ldots, X_k)$ over all joint Gaussian distributions $X_1, \ldots, X_k$ such that $(X_1, \ldots, X_k)$ are mutually independent. For every $\epsilon > \epsilon'$, we have

$$\begin{aligned}\mathbb{P}(\mathcal{E}_0) &= \mathbb{P}\{(M_{1c}, M_{1p}, M_{2c}) \notin \mathscr{S}\} \quad (6.23)\\ &= \mathbb{P}\left\{(\mathbf{x}_{1c}(M_{1c}), \mathbf{x}_{1p}(M_{1c}, M_{1p}), \mathbf{x}_{2c}(M_{2c}), \mathbf{Y}_1) \notin \bigcup \mathcal{T}_\epsilon^{(n)}(X_{1c}, X_{1p}, X_{2c}, Y_1)\right\} \\ &\quad (6.24)\\ &= \mathbb{P}\left\{(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, h_{12}\mathbf{x}_{2p} + g_1\mathbf{w}_1 + \mathbf{V}_1) \notin \mathcal{T}_\epsilon^{(n,4)}\right\} \quad (6.25)\\ &\leq \mathbb{P}\left\{(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1, \mathbf{V}_1) \notin \mathcal{T}_\epsilon^{(n,6)}\right\} \quad (6.26)\\ &\leq \mathbb{P}\left\{(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1) \notin \mathcal{T}_{\epsilon'}^{(n,5)}\right\} \\ &\quad + \mathbb{P}\left\{(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1, \mathbf{V}_1) \notin \mathcal{T}_\epsilon^{(n,6)} \,\middle|\, (\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1) \in \mathcal{T}_{\epsilon'}^{(n,5)}\right\} \\ &\quad (6.27)\end{aligned}$$

where the union in (6.24) is exactly the one in set $\mathscr{S}$ definition (6.15). (6.26) follows because if $(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1, \mathbf{V}_1)$ is typical for independent distributions then $(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, h_{12}\mathbf{x}_{2p} + g_1\mathbf{w}_1 + \mathbf{V}_1)$ would be typical. The probabilities in (6.27) follows from the fact that if $(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1) \notin \mathcal{T}_{\epsilon'}^{(n,5)}$ then $(\mathbf{x}_{1c}, \mathbf{x}_{1p}, \mathbf{x}_{2c}, \mathbf{x}_{2p}, \mathbf{w}_1, \mathbf{V}_1) \notin \mathcal{T}_{\epsilon}^{(n,6)}$. Finally, as $n \to \infty$ the first term in (6.27) vanishes exponentially by using the general version of Lemma 9-(3.23), and the second term in (6.27) tends to zero by using conditional typicality lemma (see (2) in Section 2.4). Then, $\mathbb{P}(\mathcal{E}_0)$ tends to zero as $n \to \infty$.

To bound the probability of event $\mathcal{E}_1$, we apply Lemma 6 with the following:

- $i = m_{1p}$, $j = \tilde{m}_{1p}$, $k = m_{1c}$,
- $\mathbf{x}_i(k) = \frac{h_{11}\mathbf{x}_{1p}(m_{1c},m_{1p})}{\sqrt{(1-\gamma)\alpha_1\sigma^2 S_1}}$,
- $\mathbf{x}_j(k) = \frac{h_{11}\mathbf{x}_{1p}(m_{1c},\tilde{m}_{1p})}{\sqrt{(1-\gamma)\alpha_1\sigma^2 S_1}}$,
- $\mathbf{V} = \frac{h_{12}\mathbf{x}_{2p}(M_{2c},M_{2p})+\mathbf{V}_1}{\sqrt{(1-\gamma)\alpha_1\sigma^2 S_1}}$,
- $\mathbf{w} = \frac{g_1\mathbf{w}_1}{\sqrt{(1-\gamma)\alpha_1\sigma^2 S_1}}$.

In this case, $K = 2^{nR_{1c}} \geq n^2$ for sufficiently large n as long as $R_{1c} > 0$. Note that event $\mathcal{E}_1$ occurs if

$$\begin{aligned}\|h_{11}\mathbf{X}_{1p}(m_{1c}, m_{1p}) + g_1\mathbf{w}_1 + h_{12}\mathbf{X}_{2p}(M_{2c}, M_{2p}) + \mathbf{V}_1 - h_{11}\mathbf{X}_{1p}(m_{1c}, \tilde{m}_{1p})\|^2 \\ \leq \|g_1\mathbf{w}_1 + h_{12}\mathbf{X}_{2p}(M_{2c}, M_{2p}) + \mathbf{V}_1\|^2. \quad (6.28)\end{aligned}$$

Thus, by Lemma 6, if $R_{1p} < C\left(\frac{(1-\gamma)\alpha_1 S_1}{1+J_1+(1-\gamma)\alpha_2 I_1}\right) = C\left(\frac{(1-\gamma)\alpha_1 S_1'}{1+(1-\gamma)\alpha_2 I_1'}\right)$, then with high probability the codebook $\mathbf{X}_{1p}$ will be such that $\mathbb{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$.

We now bound event $\mathcal{E}_2$ by applying Lemma 8 with the following particularizations:

- $i_1 = m_{1c}$, $i_2 = (m_{1c}, m_{1p})$, $j_1 = \tilde{m}_{1c}$, $j_2 = (\tilde{m}_{1c}, \tilde{m}_{1p})$,
- $\mathbf{x}_{i_1} = \frac{h_{11}\mathbf{x}_{1c}(m_{1c})}{\sqrt{(1-\gamma)\sigma^2 S_1}}$, $\mathbf{y}_{i_2} = \frac{h_{11}\mathbf{x}_{1p}(m_{1c},m_{1p})}{\sqrt{(1-\gamma)\sigma^2 S_1}}$,

- $\mathbf{x}_{j_1} = \frac{h_{11}\mathbf{x}_{1c}(\tilde{m}_{1c})}{\sqrt{(1-\gamma)\sigma^2 S_1}}$, $\mathbf{y}_{j_2} = \frac{h_{11}\mathbf{x}_{1p}(\tilde{m}_{1c},\tilde{m}_{1p})}{\sqrt{(1-\gamma)\sigma^2 S_1}}$,
- $\mathbf{V} = \frac{h_{12}\mathbf{x}_{2p}(M_{2c},M_{2p})+\mathbf{V}_1}{\sqrt{(1-\gamma)\sigma^2 S_1}}$,
- $\mathbf{w} = \frac{g_1\mathbf{w}_1}{\sqrt{(1-\gamma)\sigma^2 S_1}}$.

Note that event $\mathcal{E}_2$ occurs if

$$\|h_{11}\mathbf{X}_{1c}(m_{1c}) + h_{11}\mathbf{X}_{1p}(m_{1c}, m_{1p}) + \mathbf{w} + \mathbf{V} - h_{11}\mathbf{X}_{1c}(\tilde{m}_{1c}) - h_{11}\mathbf{X}_{1p}(\tilde{m}_{1c}, \tilde{m}_{1p})\|^2 \leq \|\mathbf{w}+\mathbf{V}\|^2.$$

Therefore, we can conclude by Lemma 8 that with high probability as $n \to \infty$, $\mathbb{P}(\mathcal{E}_2) \to 0$ if $\frac{J_1}{S_1} < 1$,

$$R_{1c} < C\left(\frac{(1-\gamma)\bar{\alpha}_1 S_1}{1+J_1+(1-\gamma)\alpha_2 I_1}\right) \tag{6.29}$$

$$R_{1p} < C\left(\frac{(1-\gamma)\alpha_1 S_1}{1+J_1+(1-\gamma)\alpha_2 I_1}\right) \tag{6.30}$$

$$R_{1c} + R_{1p} < C\left(\frac{(1-\gamma) S_1}{1+J_1+(1-\gamma)\alpha_2 I_1}\right). \tag{6.31}$$

Similarly, the probability of event $\mathcal{E}_3$ can be bounded using Lemma 8 as long as we have $\alpha_1 S_1 + \bar{\alpha}_2 I_1 > J_1$, (6.30),

$$R_{2c} < C\left(\frac{(1-\gamma)\bar{\alpha}_2 I_1}{1+J_1+(1-\gamma)\alpha_2 I_1}\right) \tag{6.32}$$

$$R_{1p} + R_{2c} < C\left(\frac{(1-\gamma)(\alpha_1 S_1+\bar{\alpha}_2 I_1)}{1+J_1+(1-\gamma)\alpha_2 I_1}\right). \tag{6.33}$$

Note that if we apply Lemma 7, which benefits from the common randomness M_{1c} between the encoder and the decoder, instead of Lemma 8, we can delete the power constraint $\alpha_1 S_1 + \bar{\alpha}_2 I_1 > J_1$. Finally, the probability of event $\mathcal{E}_4$ may be bounded using Lemma 7 and Lemma 8 under the conditions $S_1 + \bar{\alpha}_2 I_1 > J_1$, (6.29), (6.30), (6.31), (6.32), (6.33), $R_{1c} + R_{2c} < C\left(\frac{(1-\gamma)(\bar{\alpha}_1 S_1+\bar{\alpha}_2 I_1)}{1+J_1+(1-\gamma)\alpha_2 I_1}\right)$ and $R_{1c} + R_{1p} + R_{2c} < C\left(\frac{(1-\gamma)(S_1+\bar{\alpha}_2 I_1)}{1+J_1+(1-\gamma)\alpha_2 I_1}\right)$. Note that in this case we use a version of Lemma 7 and Lemma 8 for three independent codebooks rather than two. We finally get all equations in (6.13) as $\gamma \to 0$.

6.6 Proof of Proposition 14

First, we prove that $\mathscr{C} \subseteq \mathscr{C}_G$. Consider rate pair $(R_1', R_2') \in \mathscr{C}$ meaning that there exists a $\left(2^{nR_1'}, 2^{nR_2'}, n\right)$ code that yields an arbitrary small probability of error for any possible adversary action of two independent jammers. Now, we must show that this same code will also work for G jammers. In the G jammer model, let $\mathbf{w}_1, \ldots, \mathbf{w}_G$ be any jamming signals. We may define equivalent jamming signals for the model with two independent jammers as

$$\mathbf{w}_1' = \frac{g_{11}\mathbf{w}_1 + g_{12}\mathbf{w}_2 + \ldots + g_{1G}\mathbf{w}_G}{g_1}$$

and

$$\mathbf{w}_2' = \frac{g_{21}\mathbf{w}_1 + g_{22}\mathbf{w}_2 + \ldots + g_{2G}\mathbf{w}_G}{g_2}.$$

Note that the received signal in the G jammer model is identical to that in the 2 jammer model with jamming signals $\mathbf{w}_1', \mathbf{w}_2'$. Moreover, in order to show that $\mathbf{w}_1'$ and $\mathbf{w}_2'$ satisfy power constraints, we have

$$\begin{aligned}
\|\mathbf{w}_i'\|^2 &= \frac{\|g_{i1}\mathbf{w}_1 + g_{i2}\mathbf{w}_2 + \ldots + g_{iG}\mathbf{w}_G\|^2}{|g_i|^2} && (6.34)\\
&\leq \frac{\left(|g_{i1}|\|\mathbf{w}_1\| + |g_{i2}|\|\mathbf{w}_2\| + \ldots + |g_{iG}|\|\mathbf{w}_G\|\right)^2}{|g_i|^2} && (6.35)\\
&\leq \frac{\left(|g_{i1}|\sqrt{n\Lambda} + |g_{i2}|\sqrt{n\Lambda} + \ldots + |g_{iG}|\sqrt{n\Lambda}\right)^2}{|g_i|^2} && (6.36)\\
&\leq n\Lambda\frac{\left(|g_{i1}| + |g_{i2}| + \ldots + |g_{iG}|\right)^2}{|g_i|^2} && (6.37)\\
&= n\Lambda && (6.38)
\end{aligned}$$

where $i = 1, 2$, and 6.38 follows from the assumption in the proposition. Therefore, the probability of error under the G jammer model is at most that of the 2 jammer model.

Now, we prove $\mathscr{C}_G \subseteq \mathscr{C}$. Let rate $(R_1, R_2) \in \mathscr{C}_G$. Therefore, there exists a sequence of $\left(2^{nR_1}, 2^{nR_2}, n\right)$ code that has arbitrary small probability of error P_e^G for any possible adversary actions with G jammers. Now, if we use this code for two-jammer scenario Fig. 4, the probability of error at decoder $i = 1, 2$ is given as

$$P_{ei} = \max_{\mathbf{w}_i'} \mathbb{P}(\hat{M}_i \neq M_i) \tag{6.39}$$

$$= \max_{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_G} \mathbb{P}(\hat{M}_i \neq M_i) \tag{6.40}$$

where the last equality follows because of the same power constraints $\|\mathbf{w}_j\|^2 \leq n\Lambda$ and $\|\mathbf{w}_i'\|^2 \leq n\Lambda$ (6.38) for $j = 1, 2, \ldots, G$ and $i = 1, 2$ in both models meaning that the set of received jammer signals $g_{i1}\mathbf{w}_1 + g_{i2}\mathbf{w}_2 + \ldots + g_{iG}\mathbf{w}_G$ or $g_i\mathbf{w}_i'$ are identical at each decoder. Note that we have the assumption $|g_{i1}| + |g_{i2}| + \ldots + |g_{iG}| = |g_i|$ for $i = 1, 2$. By the equivalent expression for P_{ei} in (6.40), the probability of error in the G jammer model can be lower bounded by

$$P_e^G \geq \max\{P_{e1}, P_{e2}\}. \tag{6.41}$$

In addition, the overall probability of error is upper bounded by

$$P_e^{(n)} \leq P_{e1} + P_{e2}. \tag{6.42}$$

Since $P_e^G \to 0$, from (6.41) both P_{e1} and P_{e2} must tend to zero, too. Therefore, the sum in (6.42) also tends to zero, so the overall probability of error with two jammers $P_e^{(n)}$ vanishes as $n \to \infty$.

Chapter 7

GAUSSIAN ARBITRARILY-VARYING FADING CHANNELS

In this chapter, we consider an arbitrarily-varying fading channel consisting of one transmitter, one receiver and an arbitrarily varying adversary. The channel is assumed to have additive Gaussian noise and fast fading of the gain from the legitimate user to the receiver. We study four variants of the problem depending on whether the transmitter and/or adversary have access to the fading gains; we assume the receiver always knows the fading gains. In two variants the adversary does not have access to the gains, so the capacity corresponds to the capacity of a standard point-to-point fading channel with increased noise variance. The capacity of the other two cases, in which the adversary has knowledge of the channel gains, are determined by the worst-case noise variance as a function of the channel gain subject to the jammer's power constraint; if the jammer has enough power, then it can imitate the legitimate user's channel, causing the capacity to drop to zero. We also show that having the channel gains causally or non-causally at the encoder and/or the adversary does not change the capacity, except for the case where all parties know the channel gains. In this case, if the transmitter knows the gains non-causally, while the adversary knows the gains causally, then it is possible for the legitimate users to keep a secret from the adversary. We show that in this case the capacity is always positive

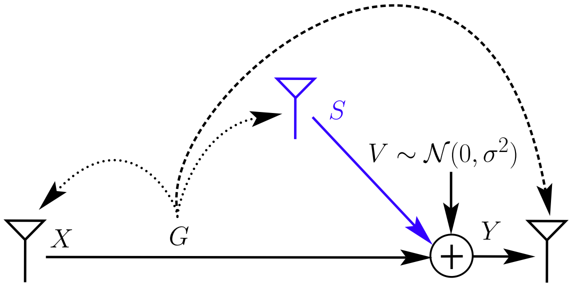


Figure 9: Gaussian Arbitrarily-Varying Fading Channel.

7.1 Problem Statement

The Gaussian arbitrarily-varying fading channel (GAVFC) in Fig. 9 is a point-to-point fading channel with additive Gaussian noise and an intelligent adversary who does not have any information about the transmitted signal except the code. The received signal is given by

$$\mathbf{Y} = \mathbf{G} \circ \mathbf{x} + \mathbf{s} + \mathbf{V} \tag{7.1}$$

where $\mathbf{G}$ is a random sequence of identical and independently distributed (i.i.d.) fast fading channel gains from the legitimate transmitter to the receiver drawn from continuous distribution $f_G(g)$ assumed to have positive and finite variance, $\mathbf{x}$ is the n-length deterministic vector representing the user's signal, $\mathbf{s}$ is the adversary signal chosen arbitrarily, and $\mathbf{V}$ is a random n-length noise vector distributed as a sequence of i.i.d. zero mean Gaussian random variables with variance σ^2, independent of $\mathbf{x}$, $\mathbf{G}$ and $\mathbf{s}$. Note that the receiver always knows the exact fading coefficients $\mathbf{g}$ while the transmitter and the adversary either not know the gains, know them causally, or know them non-causally.

Define an (N, n) code for the GAVFC by a message set, an encoding function and a decoding function as follows:

- Message set $\mathcal{M} = [N]$,
- Encoding function (one of the following)

No knowledge $\mathbf{x}(m) : \mathcal{M} \to \mathbb{R}^n$ where $\mathbf{x} = (x_1, \dots, x_n)$

Causal $x_i(m, \mathbf{g}^i) : \mathcal{M} \times \mathbb{R}^i \to \mathbb{R}$ where $\mathbf{g}^i = (g_1, \dots, g_i)$ and $\mathbf{x} = (x_1, \dots, x_n)$ for $i \in [n]$

Non-Causal $x_i(m, \mathbf{g}) : \mathcal{M} \times \mathbb{R}^n \to \mathbb{R}$, where $\mathbf{g} = (g_1, \dots, g_n)$ and $\mathbf{x} = (x_1, \dots, x_n)$ for $i \in [n]$

- Decoding function $\Theta(\mathbf{y}, \mathbf{g}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{M}$,

where the rate of the code is $R = \frac{1}{n}\log(N)$. The message m is drawn uniformly from the set $\mathcal{M}$. If the encoder does not know the channel gains, it maps the message to $\mathbf{x}(m) \in \mathbb{R}^n$. If the encoder knows the channel gains causally, then it maps the message to $x_i(m, \mathbf{g}^i) \in \mathbb{R}$, and if the encoder knows the channel gains non-causally, then it maps the message to $x_i(m, \mathbf{g}) \in \mathbb{R}$ where $\mathbf{x} = (x_1, \dots, x_n)$. Given channel gains $\mathbf{g}$ at the receiver, the signal $\mathbf{y}$ is decoded by function $\Theta(\mathbf{y}, \mathbf{g})$ to the message $\hat{m}$. Moreover, we assume that if the channel gains are available at the transmitter then the transmitter's signal satisfies the expected power constraints $\mathbb{E}\left[\|\mathbf{X}(m, \mathbf{G})\|^2\right] \leq nP$ for any message $m \in \mathcal{M}$. Otherwise, the power constraint is $\|\mathbf{x}(m)\|^2 \leq nP$. The same definition applies to the adversary's signal power constraint, i.e. if the adversary knows the channel gains, the constraint is $\mathbb{E}\left[\|\mathbf{S}(\mathbf{G})\|^2\right] \leq n\Lambda$; otherwise, it is $\|\mathbf{s}\|^2 \leq n\Lambda$. The three parameters P, Λ, and σ^2 as well as the distribution of fading gains $f_G(g)$ are known to all parties.

The probability of error $e(\mathbf{s}, m)$ for the message $m \in \mathcal{M}$ in the presence of adversary signal $\mathbf{s} \in \mathbb{R}^n$ is now given by the probability that $\hat{m} \neq m$. Thus, the average probability of error for a specific $\mathbf{s} \in \mathbb{R}^n$ is

$$\bar{e}(\mathbf{s}) = \frac{1}{N}\sum_{m=1}^{N} e(\mathbf{s}, m). \tag{7.2}$$

If the adversary knows the channel gains non-causally then his signal is denoted by $s_i(\mathbf{g})$ for $i \in [n]$. Alternatively, if the adversary knows the gains causally, then the adversary's action is given by functions $s_i(\mathbf{g}^i)$ for $i \in [n]$ where $\mathbf{s} = (s_1, \cdots, s_n)$ and $\mathbf{g}^i = (g_1, \cdots, g_i)$. Finally, the overall probability of error $P_e^{(n)}$ is maximized over all possible choices of jammers' sequences $\mathbf{s}$ which satisfy either $\mathbb{E}\left[\|\mathbf{S}\|^2\right] \leq n\Lambda$ or $\|\mathbf{s}\|^2 \leq n\Lambda$. Rate R is *achievable* if there exists a sequence of $\left(2^{nR}, n\right)$ codes where $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity is the supremum of all achievable rates. We denote the capacity of the GAVFC as $C_{\alpha,\beta}$ where α denotes the transmitter's knowledge, and β denotes the adversary's knowledge; α and β can be U, C, or N depending on whether the transmitter or adversary does not know the gains (U = unknown), knows the gains causally (C), or knows the gains non-causally (N). For example $C_{\mathrm{U,N}}$ is the capacity where the transmitter does not know the gains and the adversary knows the gains non-causally.

7.2 Main Results

We present our results for the capacity of GAVFC whether the fading channel gains $\mathbf{G}$ are available causally or non-causally at the encoder and/or the adversary (the decoder always knows the gains) in the following theorems.

Theorem 17 *The capacities of the GAVFC are given by*

$$C_{U,U} = \mathbb{E}_G\left[C\left(\frac{G^2P}{\Lambda+\sigma^2}\right)\right], \tag{7.3}$$

$$C_{N,U} = C_{C,U} = \max_{\varphi(g):\mathbb{E}\varphi(G)\leq P} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\Lambda+\sigma^2}\right)\right], \tag{7.4}$$

$$C_{U,N} = C_{U,C} = \begin{cases} \min\limits_{\psi(g):\mathbb{E}\psi(G)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2P}{\psi(G)+\sigma^2}\right)\right], & \mathbb{E}G^2P > \Lambda \\ 0, & \mathbb{E}G^2P \leq \Lambda \end{cases} \tag{7.5}$$

$$C_{N,N} = C_{C,C} = C_{C,N} =$$

$$\begin{cases} \max\limits_{\substack{\varphi(g):\mathbb{E}\varphi(G)\leq P,\\ \mathbb{E}G^2\varphi(G)\geq\Lambda}} \ \min\limits_{\psi(g):\mathbb{E}\psi(G)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\psi(G)+\sigma^2}\right)\right] & \textit{if} \max\limits_{\varphi(g):\mathbb{E}\varphi(G)\leq P} \mathbb{E}G^2\varphi(G) > \Lambda \\ 0, & \textit{if} \max\limits_{\varphi(g):\mathbb{E}\varphi(G)\leq P} \mathbb{E}G^2\varphi(G) \leq \Lambda. \end{cases} \tag{7.6}$$

$$C_{N,C} = \max_{\varphi(g):\mathbb{E}\varphi(G)\leq P} \ \min_{\psi(g):\mathbb{E}\psi(G)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\psi(G)+\sigma^2}\right)\right] \tag{7.7}$$

Note that when the encoder knows the gains in (7.4), (7.6) and (7.7), the capacity expression includes a maximization of the input power as a function $\varphi(\cdot)$ of the gain, similar to the result in Goldsmith and Varaiya, (1997). Similarly, when the jammer knows the gains in (7.5), (7.6) and (7.7), the capacity expression includes a minimization that represents the jammer's choice of noise power as a function $\psi(\cdot)$ of the gain. Moreover, when the jammer knows the gains, with enough power it can symmetrize the channel by mimicking the legitimate signal, thus reducing the capacity to zero. However, in (7.7) we have assumed that the adversary knows the gains causally and the encoder and the decoder know the gains non-causally. Thus, the encoder and decoder effectively share a secret (the channel gains at the end of the block) unknown to the adversary, so the adversary cannot symmetrize the channel. It is also worth mentioning that for the other cases (except (7.7)) our proof works exactly the same whether the transmitter and/or the adversary know the gain sequence causally, non-causally, or even memorylessly (i.e., at time i, you only know the gain value at time i).

In the Fig. 10, the capacity of GAVFC with Rayleigh fading is shown for $P = 1, \sigma^2 = 0.25, 0 < \Lambda < 10$ whether the channel gains are available at the encoder and/or adversary. However, C is the capacity of standard Gaussian arbitrary-varying channel without any fading. It is notable that if the encoder knows the channel gains, then it can choose its signal as a function of gains to increase the capacity of the channel. On the other hand, the knowledge of adversary about the channel gains may decrease the capacity, and in this case if the adversary's power exceeds 2, the capacity will be zero by the symmetrizability.

It is worth mentioning that all of the achievability proofs follow a very similar structure, and the main differences origin from the knowledge of the adversary about the channel gains. Codebook generation, encoding, decoding and the error events are mostly the same. However, each proof changes in the way that we analyze the probability of error and show that it vanishes subject to some rate and power constraints. We provide the achievability proof for the three cases of $C_{\mathrm{N,U}}$, $C_{\mathrm{N,N}}$ and $C_{\mathrm{N,C}}$ in which the encoder knows the channel gains, but the adversary knowledge about the channel gains changes from unknown, non-causally known and causally known. Since the proof does not change too much whether the encoder knows the channel gains or not, the achievability proof for $C_{\mathrm{U,U}}$ and $C_{\mathrm{U,N}}$ become special cases of $C_{\mathrm{N,U}}$ and $C_{\mathrm{N,N}}$, respectively. Note that the achievability proves for the other cases as $C_{\mathrm{C,U}}$, $C_{\mathrm{U,C}}$, $C_{\mathrm{C,C}}$ and $C_{\mathrm{C,N}}$ are equal to one of those proves that we have already covered.

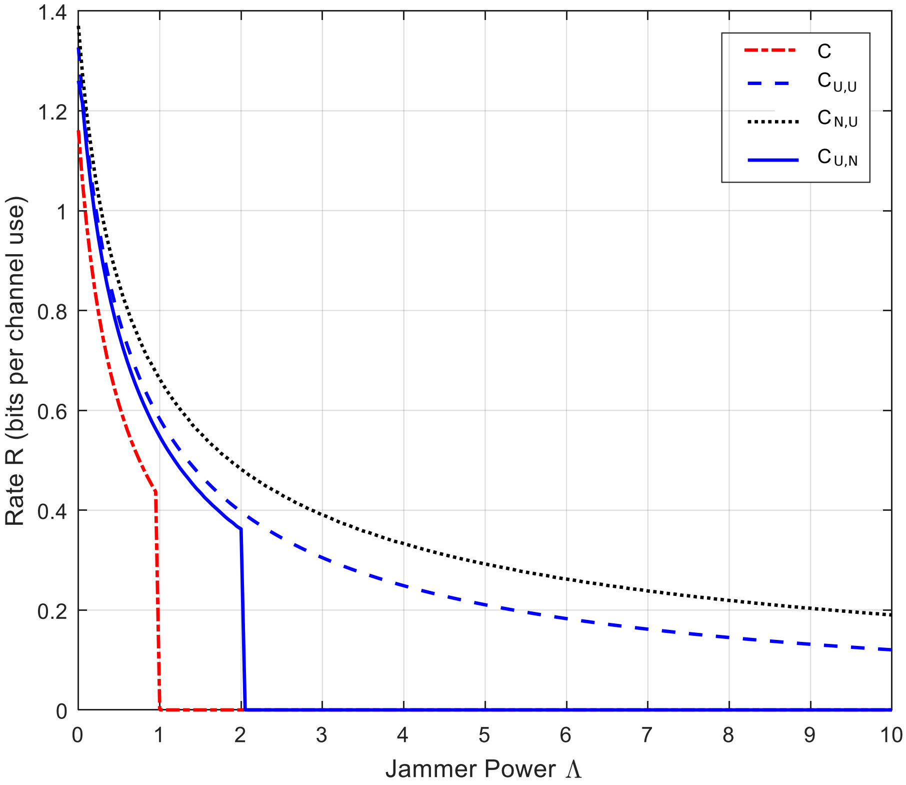


Figure 10: Gaussian arbitrarily-varying fading channel capacities for $P = 1, \sigma^2 = 0.25, 0 < \Lambda < 10$ with Rayleigh fading.

Note: C is the capacity of the standard Gaussian channel without fading.

7.3 Auxiliary Results and Tools

Before proceeding to the proofs, we first define the typical set for continuous random variables $X_1, \dots, X_k$ with probability density function $f_{X_1,\dots,X_k}(x_1, \dots, x_k)$ as follows:

$$\mathcal{T}_\epsilon^{(n)}(X_1, \dots, X_k) = \left\{ (x_1, \dots, x_k) \colon \left| -\frac{1}{n} \log f_{X_A}(x_A) - h(X_A) \right| \le \epsilon \text{ for all } A \subset [k] \right\} \tag{7.8}$$

where $h(X_A)$ is the differential entropy of $(X_i : i \in A)$. Next, we define the typical set for continuous random variables $X_1, \ldots, X_k$ with probability density function $f_{X_1,\ldots,X_k}(x_1, \ldots, x_k)$ and a discrete random variable $\tilde{G}$ with probability mass function $P_{\tilde{G}(\tilde{g})}$ as follows:

$$\mathcal{T}_{\epsilon}^{(n)}(X_1, \ldots, X_k, \tilde{G}) = \Bigg\{(x_1, \ldots, x_k, \hat{g})\colon \left|-\frac{1}{n}\log P_{\tilde{G}}(\tilde{g}) - H(\tilde{G})\right| \leq \epsilon,$$
$$\left|-\frac{1}{n}\log f_{X_A}(x_A) - h(X_A)\right| \leq \epsilon, \left|-\frac{1}{n}\log f_{X_A|\tilde{G}}(x_A|\tilde{g}) - h(X_A|\tilde{G})\right| \leq \epsilon, \text{for all } A \subset [k]\Bigg\}. \tag{7.9}$$

where $H(\tilde{G})$ and $h(X_A|\tilde{G})$ denote the entropy of G and the conditional differential entropy of X_A given $\tilde{G}$.

Throughout the achievability proofs, we will utilize several lemmas including the joint typicality lemma as Lemma 3 and conditional typicality lemma as Lemma 2 for Gaussian random variables. The main two lemmas in this proof are described as follows and they show that with high probability a Gaussian codebook satisfies several desirable properties. The proofs are given in Section 7.8.

Lemma 18 *Fix $\epsilon' > 0$. There exists $\gamma > 0$ such that the following holds. Let $\mathbf{X}(m)$ for $m \in [N]$, $N = 2^{nR}$ be a zero mean Gaussian codebook with variance $1-\gamma$. Consider a random variable G drawn from probability density function $f_G(g)$. With probability approaching 1 as $n \to \infty$, for any $\mathbf{s}, \mathbf{g}$ where $\|\mathbf{s}\|^2 \leq n\Lambda$, there exists a function $\delta(\epsilon') > 0$ such that*

$$\frac{1}{N}\left|\left\{m : (\mathbf{x}(m), \mathbf{s}, \mathbf{g}) \notin \bigcup_{\substack{X \text{ independent of } (S,G): \\ EX^2=1, ES^2 \leq \Lambda}} \mathcal{T}_{\epsilon'}^{(n)}(X, S, G)\right\}\right| \leq \exp(-n\delta(\epsilon')), \tag{7.10}$$

where the union is over zero mean conditionally Gaussian random vectors (X, S) given G.

Lemma 19 *Fix $\epsilon > 0$. There exists $\gamma > 0$ such that the following holds. Let $\mathbf{X}(m)$ for $m \in [N]$, $N = 2^{nR}$ be a zero mean Gaussian codebook with variance $1-\gamma$. Let G be drawn from probability density function $f_G(g)$. With probability approaching 1 as $n \to \infty$, for any*

- *zero-mean conditionally Gaussian random vector (X, X', S) given G where $\mathbb{E}X^2 = \mathbb{E}X'^2 = 1$ and $\mathbb{E}S^2 \leq \Lambda$,*
- *$\mathbf{x}, \mathbf{s}, \mathbf{g}$ where $\|\mathbf{s}\|^2 \leq n\Lambda$,*

there exists a function $\delta(\epsilon) > 0$ such that

$$\mathbb{P}\left\{\left|\left\{(\mathbf{x}(m'),\mathbf{s},\mathbf{G})\in\mathcal{T}_\epsilon^{(n)}(X',S,G) \textit{ for some } m'\right\}\right|\right\} \leq 2\exp\{-n\delta(\epsilon)/2\}, \qquad \textit{if } I(G;X'S) \geq |R - I(X';S)|^+ + \delta(\epsilon), \tag{7.11}$$

$$\left|\left\{m' : (\mathbf{x}(m'),\mathbf{s}) \in \mathcal{T}_\epsilon^{(n)}(X',S)\right\}\right| \leq \exp\left\{n\left[|R - I(X';S)|^+ + \delta(\epsilon)\right]\right\}, \tag{7.12}$$

$$\left|\left\{m' : (\mathbf{x},\mathbf{x}(m'),\mathbf{s},\mathbf{g}) \in \mathcal{T}_\epsilon^{(n)}(X,X',S,G)\right\}\right| \leq \exp\left\{n\left[|R - I(X';XSG)|^+ + \delta(\epsilon)\right]\right\}, \tag{7.13}$$

$$\frac{1}{N}\left|\left\{m : (\mathbf{x}(m),\mathbf{x}(m'),\mathbf{s},\mathbf{g}) \in \mathcal{T}_\epsilon^{(n)} \textit{ for some } m' \neq m\right\}\right| \leq 2\exp\{-n\delta(\epsilon)/2\}, \qquad \textit{if } I(X;X'SG) \geq |R - I(X';SG)|^+ + \delta(\epsilon). \tag{7.14}$$

7.4 Capacity Proof with Gains Available at Decoder

7.4.1 Converse Proof

We initially assume that for any arbitrary adversary strategy there is a sequence of $(2^{nR}, n)$ codes with vanishing probability of error. The adversary can generate a

Gaussian sequence with variance $\Lambda - \gamma$ for any $\gamma > 0$; if this sequence has power less than Λ, it is transmitted, otherwise, the adversary sends the all-zero sequence. Note that the power of this Gaussian sequence exceeds Λ only with small probability by the law of large numbers. With this choice of adversary, the channel corresponds to a standard Gaussian fading channel with the noise variance $\Lambda + \sigma^2 - \gamma$ where the channel gains are available only at the decoder. Therefore, using capacity of a non-adversarial Gaussian fading channel ElGamal and Kim, (2011) for arbitrarily small γ, we may upper bound the capacity by

$$C \leq \mathbb{E}_G \left[C \left(\frac{G^2 P}{\Lambda + \sigma^2} \right) \right]. \tag{7.15}$$

7.4.2 Achievability Proof

The achievability proof of this case can be counted as a special case of $C_{N,U}$ in Sec. 7.5.2 where both encoder and decoder know the channel gains. However, in this case since the encoder does not know the channel gains, we do not have any $\varphi(g)$ function at the encoder. In other words, the achievability proof for this case is identical to that in Sec. 7.5.2 with $\varphi(g) = P$.

7.5 Capacity Proof with Gains Available at Encoder and Decoder

7.5.1 Converse Proof

As in the previous case, the adversary can simply send Gaussian noise with variance $\Lambda - \gamma$. By the law of large numbers, the resulting channel is equivalent to a standard Gaussian fading channel with the knowledge of gains at both encoder and decoder and noise variance $\Lambda + \sigma^2 - \gamma$ with high probability. Thus, since γ can be chosen arbitrarily small, from the capacity of a non-adversarial Gaussian fading channel Goldsmith and Varaiya, (1997), we have

$$C \leq \max_{\varphi(g):\mathbb{E}\varphi(G)\leq P} \mathbb{E}_G \left[C\left(\frac{G^2\varphi(G)}{\Lambda+\sigma^2} \right) \right]. \tag{7.16}$$

The optimum value of $\varphi^*(g) = \left|\lambda - \frac{\Lambda+\sigma^2}{g^2}\right|^+$ where λ is obtained by $\mathbb{E}[\varphi^*(G)] = P$.

7.5.2 Achievability Proof

For simplicity we assume $P = 1$. Suppose any arbitrary function $\varphi(G)$ that satisfies $\mathbb{E}\varphi(G) \leq 1$ and $\mathrm{Var}(G\sqrt{\varphi(G)}) > 0$. We further assume that $G^2\varphi(G)$ has a positive variance. Note that this is only a concern if the optimum $\varphi^*(G) = \frac{c}{G^2}$; in this case, we can instead take $\varphi(G) = \frac{c}{(G-d)^2}$ where c, d are two positive constants and d can be chosen arbitrarily small. Let

$$R < \mathbb{E}_G \left[C\left(\frac{G^2\varphi(G)}{\Lambda+\sigma^2} \right) \right]. \tag{7.17}$$

We now propose a $(2^{nR}, n)$ code sequence, and prove that using this code the probability of error tends to zero as $n \to \infty$.

Codebook generation: Fix $\epsilon > \epsilon' > \gamma > 0$. We generate 2^{nR} i.i.d zero mean Gaussian sequences $\mathbf{X}(m)$ with variance $(1-\gamma)$ for each $m \in [2^{nR}]$. By Lemma 18 and Lemma 19, we assume that the deterministic codebook satisfies (7.10)–(7.14).

Encoding: Since the transmitter knows the channel gains, it sends $\sqrt{\varphi(\mathbf{g})} \circ \mathbf{x}(m)$ (at time i signal $\sqrt{\varphi(g_i)}x_i(m)$ is sent) if its power is less than 1, otherwise it sends zero.

Decoding: Given $\mathbf{y}$, let $\mathscr{S}$ be the set of messages $\hat{m}$ such that $(\mathbf{x}(\hat{m}), \mathbf{g}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(X', G, Y)$ for some random variables $X' \sim \mathcal{N}(0,1)$, $G \sim f_G(g)$ and zero mean Gaussian $Y - G\sqrt{\varphi(G)}X'$ where $(X', G, Y - G\sqrt{\varphi(G)}X')$ are mutually independent.

Now, we define the decoding function as

$$\Theta(\mathbf{y}, \mathbf{g}) = \underset{\hat{m} \in \mathscr{S}}{\arg\min} \left\| \mathbf{y} - \mathbf{g} \circ \sqrt{\varphi(\mathbf{g})} \circ \mathbf{x}(\hat{m}) \right\|^2. \tag{7.18}$$

Analysis of the probability of error: Suppose the true message sent by the legitimate user is message M with the power constraint $\|\mathbf{x}(M)\|^2 \leq n(1-\gamma)$. Then, the overall probability of error is upper bounded by $P_e^{(n)} \leq P_0 + P_1$ where

$$P_0 = \mathbb{P}\left\{M \notin \mathscr{S}\right\}, \tag{7.19}$$

$$P_1 = \mathbb{P}\left\{ \left\| \mathbf{Y} - \mathbf{G} \circ \sqrt{\varphi(\mathbf{G})} \circ \mathbf{x}(\hat{m}) \right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 \text{ for some } \hat{m} \in \mathscr{S} \setminus \{M\} \right\}. \tag{7.20}$$

Consider any state sequence $\mathbf{s}$. By (7.10), with high probability $(\mathbf{x}(M), \mathbf{s}, \mathbf{G}) \in \mathcal{T}_{\epsilon'}^{(n)}(X, S, G)$ where (X, S, G) are independent, and $\mathbb{E}X^2 = 1, \mathbb{E}S^2 \leq \Lambda$. By the conditional typicality lemma 2, for every $\epsilon > \epsilon'$ with high probability $(\mathbf{x}(M), \mathbf{s}, \mathbf{G}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, S, G, V)$ where (X, S, G, V) are mutually independent, and $\mathbb{E}V^2 = \sigma^2$. Thus, according to the definition of $\mathscr{S}$, with high probability $M \in \mathscr{S}$ and P_0 tends to zero as $n \to \infty$.

Define the shorthand $\vec{X} = (XX'SGV)$. Let $\mathcal{V}$ be a finite ϵ-dense subset in the set of all distributions of random vectors $\vec{X}$ that are determined by $f_G(g)$ and jointly zero mean Gaussian vector $(XX'SV)$ independent of G with bounded covariances at most $(1, 1, \Lambda, \sigma^2)$. Note that because the distribution of $f_G(g)$ is fixed, the overall distribution of $\vec{X}$ can be determined by the covariance matrix of $(XX'SV)$, so $\mathcal{V}$ only needs to cover a compact set. Now, we may upper bound P_1 by

$$\sum_{\vec{X} \in \mathcal{V}} \frac{1}{N} \sum_{m=1}^{N} \mathbb{E}_G[e_{\vec{X}}(m, \mathbf{s}, \mathbf{G})] \tag{7.21}$$

where

$$e_{\vec{X}}(m,\mathbf{s},\mathbf{g}) = \mathbb{P}\Big\{(\mathbf{x}(m),\mathbf{x}(\hat{m}),\mathbf{s},\mathbf{g},\mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(\vec{X}),$$
$$\|\mathbf{g}\circ\sqrt{\varphi(\mathbf{g})}\circ\mathbf{x}(m)+\mathbf{s}+\mathbf{V}-\mathbf{g}\circ\sqrt{\varphi(\mathbf{g})}\circ\mathbf{x}(\hat{m})\|^2 \le \|\mathbf{s}+\mathbf{V}\|^2 \text{ for some } \hat{m} \in \mathscr{S}\backslash\{m\}\Big\}. \tag{7.22}$$

We will show that $\frac{1}{N}\sum_{m=1}^{N} e_{\vec{X}}(m,\mathbf{s},\mathbf{g}) \to 0$ for all vectors $\mathbf{g}$ and all vectors $(XX'SV)$ which are Gaussian given G (whether or not they are in $\mathcal{V}$). Let $Z = G\sqrt{\varphi(G)}X + S + V - G\sqrt{\varphi(G)}X'$. We may restrict ourselves to $\vec{X}$ where

$$(X,S,G,V) \text{ are mutually independent,} \tag{7.23}$$

$$(X,X',S,V) \text{ are zero mean Gaussian} \tag{7.24}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = 1, \quad \mathbb{E}V^2 = \sigma^2, \quad \mathbb{E}S^2 \le \Lambda \tag{7.25}$$

$$(X',G,Z) \text{ are independent,} \tag{7.26}$$

$$\mathbb{E}\left[Z^2\right] \le \Lambda + \sigma^2. \tag{7.27}$$

where (7.23) holds since the input X, adversary S, fading gains G and noise V are all generated independently, (7.24)–(7.25) follows from $m,\hat{m} \in \mathscr{S}$, and $\vec{X} \in \mathcal{V}$, (7.26) holds since we have $(X',G,Y-GX')$ are mutually independent using $\mathbf{x}(\hat{m}) \in \mathscr{S}$, and (7.27) corresponds to $\mathbb{E}\left[\left(Y - G\sqrt{\varphi(G)}X'\right)^2\right]$ which is less than $\Lambda + \sigma^2$ from (7.22).

Observe that if $I(X,V,G;X',S) = 0$, then we would have

$$\begin{aligned}
0 &= \mathbb{E}[X'Z] && (7.28)\\
&= \mathbb{E}[X'(G\sqrt{\varphi(G)}X + S + V - G\sqrt{\varphi(G)}X')] && (7.29)\\
&= \mathbb{E}[X'(S - G\sqrt{\varphi(G)}X')] && (7.30)\\
&= \mathbb{E}[X'S] - \mathbb{E}G\sqrt{\varphi(G)}. && (7.31)
\end{aligned}$$

where (7.28) follows from (7.26), (7.30) holds because (X', G, X, V) are all mutually independent by the assumption $I(X, V, G; X', S) = 0$ and (7.23), and the last equality holds since X' is independent of G and because $\mathbb{E}[X'^2] = 1$. Therefore, $\mathbb{E}[X'S] = \mathbb{E}G\sqrt{\varphi(G)}$.

Moreover, from (7.22) we have

$$\mathbb{E}(S+V)^2 \geq \mathbb{E}(G\sqrt{\varphi(G)}X+S+V-G\sqrt{\varphi(G)}X')^2 \tag{7.32}$$

$$= \mathbb{E}G^2\varphi(G)(X-X')^2 + 2\mathbb{E}G\sqrt{\varphi(G)}(X-X')(S+V) + \mathbb{E}(S+V)^2 \tag{7.33}$$

$$= \mathbb{E}G^2\varphi(G)\mathbb{E}X^2 + \mathbb{E}G^2\varphi(G)\mathbb{E}X'^2 - 2\mathbb{E}G\sqrt{\varphi(G)}X'S + \mathbb{E}(S+V)^2 \tag{7.34}$$

$$= 2\mathbb{E}G^2\varphi(G) - 2\mathbb{E}G\sqrt{\varphi(G)}\mathbb{E}X'S + \mathbb{E}(S+V)^2 \tag{7.35}$$

where (7.34) holds because $\mathbb{E}X = \mathbb{E}X' = \mathbb{E}V = 0$, (X, X', G) are mutually independent, (X, S, V) are mutually independent, and (X', V) are independent by (7.23), (7.24) and the assumption $I(X, V, G; X', S) = 0$. Canceling $\mathbb{E}(S+V)^2$ from both sides of (7.35) gives us

$$\mathbb{E}G^2\varphi(G) - \mathbb{E}G\sqrt{\varphi(G)}\mathbb{E}X'S \leq 0. \tag{7.36}$$

Now, if we apply the result from (7.31) to (7.36), we get

$$\mathbb{E}G^2\varphi(G) - \mathbb{E}G\sqrt{\varphi(G)}\mathbb{E}X'S = \mathbb{E}G^2\varphi(G) - \mathbb{E}G\sqrt{\varphi(G)}\mathbb{E}G\sqrt{\varphi(G)} \tag{7.37}$$

$$= \mathbb{E}G^2\varphi(G) - \mathbb{E}^2G\sqrt{\varphi(G)} \tag{7.38}$$

$$= \operatorname{Var} G\sqrt{\varphi(G)} \tag{7.39}$$

$$\leq 0. \tag{7.40}$$

which is a contradiction since we assume $\operatorname{Var}(G\sqrt{\varphi(G)})$ is always positive. Thus, there exists an $\eta > 0$ such that

$$\eta \leq I(XVG; X'S). \tag{7.41}$$

Also, by (7.14), we may restrict ourselves to distributions where

$$I(X; X'SG) < |R - I(X'; SG)|^+ + \delta(\epsilon) \tag{7.42}$$

and

$$I(G; X'S) < |R - I(X'; S)|^+ + \delta(\epsilon). \tag{7.43}$$

Note that $I(X; X'SG) = I(X; X'|SG)$. We also have the upper bound

$$e_{\vec{X}}(m, \mathbf{s}, \mathbf{g}) \leq \sum_{\hat{m}: (\mathbf{x}(m), \mathbf{x}(\hat{m}), \mathbf{s}, \mathbf{g}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S, G)} \mathbb{P}\left\{(\mathbf{x}(m), \mathbf{x}(\hat{m}), \mathbf{s}, \mathbf{g}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S, G, V)\right\} \tag{7.44}$$

$$\leq \exp\left\{n\left[|R - I(X'; XSG)|^+ - I(V; X'|XSG) + \delta(\epsilon)\right]\right. \tag{7.45}$$

where (7.45) follows from $I(V; XSG) = 0$, (7.13) and the joint typicality lemma 3.

Now, let us consider three cases as follows:

Case (a): $R < I(X'; S)$ that implies $R < I(X'; XSG)$. From (7.45), for any $m, \mathbf{s}, \mathbf{g}$

$$e_{\vec{X}}(m, \mathbf{s}, \mathbf{g}) \leq \exp\{-n(I(V; X'|XSG) - \delta(\epsilon))\} \tag{7.46}$$

$$= \exp\{-n(I(XV; X'|SG) - I(X; X'|SG) - I(XV; S|G) - \delta(\epsilon))\} \tag{7.47}$$

$$= \exp\{-n(I(XV; X'S|G) - I(X; X'|SG) - \delta(\epsilon))\} \tag{7.48}$$

$$= \exp\{-n(I(XVG; X'S) - I(G; X'S) - I(X; X'|SG) - \delta(\epsilon))\} \tag{7.49}$$

$$\leq \exp\{-n(\eta - \delta(\epsilon) - \delta'(\epsilon))\} \tag{7.50}$$

where (7.50) follows from (7.41), (7.42) and (7.43). Therefore, $e_{\vec{X}}(m, \mathbf{s}, \mathbf{g})$ vanishes exponentially fast if $\delta(\epsilon)$ is sufficiently small.

Case (b): $I(X'; S) \leq R$. Since $R \geq I(X'; S)$ and $I(G; S) = 0$, from (7.43) we have

$$R > I(G; X'S) + I(X'; S) - \delta(\epsilon) \tag{7.51}$$

$$= I(G; S) + I(G; X'|S) + I(X'; S) - \delta(\epsilon) \tag{7.52}$$

$$= I(X'; SG) - \delta(\epsilon). \tag{7.53}$$

Using this result in (7.42), we have

$$I(X; X'SG) < R - I(X'; SG) + \delta(\epsilon) + \delta(\epsilon). \tag{7.54}$$

Therefore,

$$R > I(X; X'SG) + I(X'; SG) - 2\delta(\epsilon) \tag{7.55}$$

$$\geq I(X'; XSG) - 2\delta(\epsilon). \tag{7.56}$$

Now, from (7.45), we have for any $m, \mathbf{s}, \mathbf{g}$

$$e_{\vec{X}}(m, \mathbf{s}, \mathbf{g}) \leq \exp\left\{n\big[|R - I(X'; XSG)|^{+} - I(V; X'|XSG) + \delta(\epsilon)\big]\right. \tag{7.57}$$

$$\leq \exp\left\{n\big[R - I(X'; XSG) + 2\delta(\epsilon) - I(V; X'|XSG) + \delta(\epsilon)\big]\right. \tag{7.58}$$

$$= \exp(n[R - I(X'; XSGV) + 3\delta(\epsilon)]) \tag{7.59}$$

where (7.58) follows from (7.56). We now lower bound $I(X';XSVG)$ as follows:

$$\begin{aligned}
I(X';XSVG) &= I(X';XSV|G) + I(X';G) && (7.60)\\
&\geq I(X';G\sqrt{\varphi(G)}X+S+V|G) && (7.61)\\
&= I(X';Z+G\sqrt{\varphi(G)}X'|G) && (7.62)\\
&= h(Z+G\sqrt{\varphi(G)}X'|G) - h(Z+G\sqrt{\varphi(G)}X'|G,X') && (7.63)\\
&= \mathbb{E}\left[\frac{1}{2}\log 2\pi e\left(G^2\varphi(G)+\mathbb{E}[Z^2|G]\right) - \frac{1}{2}\log 2\pi e\mathbb{E}[Z^2|G]\right] && (7.64)\\
&= \mathbb{E}\left[C\left(\frac{G^2\varphi(G)}{\mathbb{E}[Z^2|G]}\right)\right] && (7.65)\\
&\geq \mathbb{E}\left[C\left(\frac{G^2\varphi(G)}{\Lambda+\sigma^2}\right)\right] && (7.66)
\end{aligned}$$

where (7.66) follows from (7.26) and (7.27). Replacing this result in (7.59), we obtain

$$e_{\vec{X}}(m,\mathbf{s},\mathbf{g}) \leq \exp\left\{n\left[R-\mathbb{E}\left[C\left(\frac{G^2\varphi(G)}{\Lambda+\sigma^2}\right)\right]+3\delta(\epsilon)\right]\right\} \tag{7.67}$$

meaning that $e_{\vec{X}}(m,\mathbf{s},\mathbf{g})$ is exponentially vanishing if $\delta(\epsilon)$ is sufficiently small, and (7.17) holds.

7.6 Capacity Proof with Gains Available at Decoder and Jammer

7.6.1 Converse Proof

Consider a sequence of $(2^{nR},n)$ codes with vanishing probability of error that must function for arbitrary jamming signals. Because we are proving the converse, we may assume the best case scenario from the legitimate user's perspective; in particular, that the adversary only knows the channel gains causally.

We begin with the case that $\Lambda \leq \mathbb{E}G^2P$. Given any function $\psi(g)$ satisfying $\mathbb{E}\psi(G) \leq \Lambda$, we may obtain an upper bound by assuming that the jammer transmits a random sequence $\mathbf{S} = (S_1, \cdots, S_n)$ where S_i is Gaussian with mean zero and variance $\psi(G_i)$ for $i = 1, \cdots, n$. Note that

$$\begin{aligned}\mathbb{E}[\|\mathbf{S}\|^2] &= \mathbb{E}\sum_{i=1}^{n} S_i^2 && (7.68)\\ &= \sum_{i=1}^{n} \mathbb{E}S_i^2 && (7.69)\\ &= \sum_{i=1}^{n} \psi(G_i) && (7.70)\\ &\leq n\Lambda. && (7.71)\end{aligned}$$

The resulting channel is equivalent to a standard Gaussian fading channel with the knowledge of gains only at the decoder and noise variance $\psi(g)+\sigma^2$. From the capacity of a non-adversarial Gaussian fading channel

$$C \leq \mathbb{E}_G\left[C\left(\frac{G^2P}{\psi(G)+\sigma^2}\right)\right]. \tag{7.72}$$

Therefore, the capacity is also less than the minimum over all $\psi(G)$ that satisfies $\mathbb{E}\psi(G) \leq \Lambda$.

$$C \leq \min_{\psi(G):\mathbb{E}\psi(G)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2P}{\psi(G)+\sigma^2}\right)\right]. \tag{7.73}$$

For the case $\Lambda > \mathbb{E}G^2P$, we first show that the adversary has enough power to choose a codeword and send it to the channel. Let $\tilde{M}$ be a uniformly chosen message by the adversary and M be the true message send by the legitimate transmitter. Suppose the adversary chooses $\mathbf{S} = \mathbf{G} \circ \mathbf{x}(\tilde{M})$ then the adversary power constraint is satisfied as follows:

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{S}\|^2\right] &= \mathbb{E}\left[\|\mathbf{G}\circ\mathbf{x}(\tilde{M})\|^2\right] && (7.74)\\
&= \mathbb{E}\left[\sum_{i=1}^{n} G_i^2 x_i^2(\tilde{M})\right] && (7.75)\\
&< \sum_{i=1}^{n} x_i^2(\tilde{M})\frac{\Lambda}{P} && (7.76)\\
&\le n\Lambda && (7.77)
\end{aligned}$$

where (7.76) follows from the assumption $\Lambda > \mathbb{E}G^2P$, and (7.77) follows from the codebook power constraint $\|\mathbf{x}^2\| \le nP$. Given this choice of $\mathbf{S}$, $\mathbf{Y} = \mathbf{G}\circ\mathbf{x}(M) + \mathbf{G}\circ\mathbf{x}(\tilde{M}) + \mathbf{V}$. Thus, with high probability the decoder cannot decode the message since it does not know whether the true message is M or $\tilde{M}$. In other words, the adversary symmetrizes the channel and makes the capacity zero if $\Lambda > \mathbb{E}G^2P$.

7.6.2 Achievability Proof

The achievability proof of this case is very similar to the achievability proof of Sec. 7.7.2 where the encoder, the decoder and the adversary all know the channel gains. Here, the transmitter does not know the channel gain so it cannot leverage this knowledge to choose its transmit power. However, the achievability proof for this case is identical to that in Sec. 7.7.2 except that the transmitter's power function is constant; i.e., $\varphi(g) = 1$.

7.7 Capacity Proof with Gains Available at Encoder, Decoder, and Jammer

In this section, we first provide the converse proof for the case that the channel gains are available at the encoder, the decoder and the adversary in Sec. 7.7.1. The converse proof includes all the four cases in which each of the adversary and the encoder knows the fading gains causally or non-causally. In Sec. 7.7.2, we show the achievability proof of the case that the channel gains are available non-causally at the adversary and causally at the encoder. This proof also works for the two cases of channel gains being available causally at both the adversary and the encoder or non-causally at both ends. Finally, we provide the achievability proof for the last case when the channel gains are causally available at the adversary and non-causally available at the encoder in Sec. 7.7.3.

7.7.1 Converse Proof

Consider a sequence of $(2^{nR}, n)$ code with vanishing probability of error. Since in this case both the encoder and the adversary know the channel gains, we consider four cases to prove the converse whether each of them knows the fading gains causally or non-causally.

First assume that both the encoder and adversary know the exact channel gains causally. Let $\varphi_i(g) = \frac{1}{N}\sum_{m=1}^{N}\mathbb{E}[X_i^2(m, G^i)|G_i = g]$ and $\varphi(g) = \frac{1}{n}\sum_{i=1}^{n}\varphi_i(g)$ where

$G^i = (G_1, ...G_i)$, for $i \in [n]$. Thus, $\varphi(g)$ satisfies $\mathbb{E}\varphi(G) \leq P$ as follows:

$$\begin{aligned}\mathbb{E}\varphi(G) &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\varphi_i(G)\right] && (7.78)\\ &= \frac{1}{N}\sum_{m=1}^{N}\frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n}X_i^2(m, G^i)\right] && (7.79)\\ &\leq P && (7.80)\end{aligned}$$

where (7.80) follows by the power constraint for the input signal.

Now, similar to the previous case, where the adversary and decoder know the channel gains, we also have symmetrizability and non-symmetrizability cases, but with different conditions. We first show the symmetrizability case, that is if $\Lambda \geq \mathbb{E}G^2\varphi(G)$, then the jammer can symmetrize the channel. Suppose the adversary chooses a message $\tilde{M}$ uniformly at random and sends $S_i = G_i X_i(\tilde{M}, G^i)$ where $G^i = (G_1, \cdots, G_i)$ for $i \in [n]$. Note that this selection of jamming signal is a causal function of the channel gains. Then we have

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{S}\|^2\right] &= \mathbb{E}\left[\sum_{i=1}^{n}S_i^2\right] && (7.81)\\ &= \frac{1}{N}\sum_{\tilde{m}=1}^{N}\mathbb{E}\left[\sum_{i=1}^{n}G_i^2 X_i^2(\tilde{m}, G^i)\right] && (7.82)\\ &= \sum_{i=1}^{n}\mathbb{E}_G\left[G^2\frac{1}{N}\sum_{\tilde{m}=1}^{N}\mathbb{E}\left[X_i^2(\tilde{m}, G^i)|G_i = G\right]\right] && (7.83)\\ &= \sum_{i=1}^{n}\mathbb{E}_G\left[G^2\varphi_i(G)\right] && (7.84)\\ &= \mathbb{E}_G\left[G^2\sum_{i=1}^{n}\varphi_i(G)\right] && (7.85)\\ &= n\mathbb{E}_G\left[G^2\varphi(G)\right] && (7.86)\\ &\leq n\Lambda && (7.87)\end{aligned}$$

Therefore, this choice of jammer satisfies the adversary power constraint. Given $\mathbf{Y} = \mathbf{g} \circ \mathbf{x}(M, \mathbf{g}) + \mathbf{g} \circ \mathbf{x}(\tilde{M}, \mathbf{g}) + \mathbf{V}$, the decoder cannot determine the correct message between true message M or the adversary message $\tilde{M}$ with high probability. Thus, the probability of error is bounded away from zero. By the above argument, if $\mathbb{E}G^2\varphi(G) \le \Lambda$ for all $\varphi(g)$ where $\mathbb{E}\varphi(G) \le P$, then the capacity cannot be positive; the adversary can always symmetrize the channel, so the capacity is 0.

On the other hand, consider the case where there exists some function $\varphi(g)$ where $\mathbb{E}G^2\varphi(G) > \Lambda$ and $\mathbb{E}\varphi(G) \le P$. Let $\psi_i(g)$ be given by

$$\psi_i(g) = \operatorname*{arg\,min}_{\psi(g):\mathbb{E}\psi(G)\le\Lambda} \mathbb{E}\left[C\left(\frac{G^2\varphi_i(G)}{\sigma^2 + \psi(G)}\right)\right]. \tag{7.88}$$

Since the transmitted codes should work for arbitrary jamming signals, an outer bound may be obtained by assuming the adversary sends $S_i \sim \mathcal{N}(0, \psi_i(G))$. By the assumption that $\mathbb{E}\psi_i(G) \le \ \Lambda$, the jammer's expected power constraint is satisfied. Therefore, the rate is upper bounded by

$$\begin{aligned}
nR &\le \sum_{i=1}^{n} I(X_i; Y_i|G_i) && (7.89)\\
&= \sum_{i=1}^{n} I(X_i; G_iX_i + S_i + V_i|G_i) && (7.90)\\
&\le \sum_{i=1}^{n} \mathbb{E}_{G_i}\left[C\left(\frac{G_i^2\varphi_i(G_i)}{\psi_i(G_i) + \sigma^2}\right)\right] && (7.91)\\
&= \sum_{i=1}^{n} \min_{\psi(g):\mathbb{E}\psi(g)\le\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi_i(G)}{\psi(G) + \sigma^2}\right)\right] && (7.92)\\
&\le n \min_{\psi(g):\mathbb{E}\psi(g)\le\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\frac{1}{n}\sum_{i=1}^{n}\varphi_i(G)}{\psi(G) + \sigma^2}\right)\right] && (7.93)\\
&\le n \max_{\substack{\varphi(g):\mathbb{E}\varphi(G)\le P\\ \mathbb{E}G^2\varphi(G)\ge\Lambda}} \min_{\psi(g):\mathbb{E}\psi(g)\le\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\psi(G) + \sigma^2}\right)\right] && (7.94)
\end{aligned}$$

where (7.91) follows since the mutual information is less than the capacity of equivalent standard fading channel with noise variance $\psi_i(g_i) + \sigma^2$, and the gains being available at both encoder and decoder, (7.92) follows by the definition of $\psi_i(g)$, (7.93) follows by the concavity of $C(\cdot)$ with respect to $\varphi_i(g)$ and Jensen's inequality, and (7.94) follows since we have established that $\varphi(g) = \frac{1}{n}\sum_{i=1}^{n}\varphi_i(g)$ satisfies $\mathbb{E}\varphi(G) \leq P$ and $\mathbb{E}G^2\varphi(G) \geq \Lambda$.

Moreover, if the encoder knows the channel gains causally, and the adversary knows them non-causally, then the adversary is stronger than in the previous case, so exactly the same bound holds. If both encoder and adversary know the channel gains non-causally, then we instead assume

$$\varphi_i(g) = \frac{1}{N}\sum_{m=1}^{N}\mathbb{E}[X_i^2(m, \mathbf{G})|G_i = g] \tag{7.95}$$

where $\mathbf{G} = (G_1, \ldots, G_n)$ and $S_i = G_i X_i(\tilde{m}, \mathbf{G})$, so we get the same upper bound.

However, the challenging case happens if the encoder knows the channel gains non-causally, and the adversary knows them causally. In this case, the encoder may send $X_i^2(m, \mathbf{G})$ while the adversary does not have any access to $(G_{i+1}, \ldots, G_n)$ to construct $S_i = G_i X_i(\tilde{m}, \mathbf{G})$. Thus, it cannot do better than sending Gaussian noise. In this case, the jammer can not use its knowledge of channel gains, and it can not symmetrize the channel. In fact, the Gaussian noise bound works essentially the same, even though the symmetrizability bound does not, and we do not have symmetrizability case for this scenario. Hence, we obtain the following bound for this case:

$$R \leq \max_{\varphi(g):\mathbb{E}\varphi(G)\leq P}\ \min_{\psi(g):\mathbb{E}\psi(g)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\psi(G)+\sigma^2}\right)\right] \tag{7.96}$$

Note that here we do not have the constraint $\mathbb{E}G^2\varphi(G) \geq \Lambda$ on the adversary signal since we do not have the symmetrizability case.

7.7.2 Achievability Proof (Gains Available Non-causally at Adversary and Causally at Encoder)

We first quantize G in the following way. Fix $\nu > 0$. Given the assumption that G has finite variance, there exists a real-valued random variable $\tilde{G}$ with a finite support such that $\tilde{G}$ is a deterministic function of G and $\mathbb{E}[(G - \tilde{G})^2] \leq \nu$. We further assume that $\tilde{G}$ is the expected value of G within each quantization set; that is, $\mathbb{E}[G|\tilde{G}] = \tilde{G}$.

Without loss of generality, assume $P = 1$. Let $\varphi(\tilde{g})$ be any concave function satisfying

$$\mathbb{E}\varphi(\tilde{G}) \leq 1 \tag{7.97}$$

$$\Lambda < \mathbb{E}\tilde{G}^2\varphi(\tilde{G}) \tag{7.98}$$

$$R < \min_{\psi(\tilde{g}):\mathbb{E}\psi(\tilde{G})\leq\Lambda} \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\psi(\tilde{G})+\sigma^2}\right)\right]. \tag{7.99}$$

We construct a $(2^{nR}, n)$ code as follows:

Codebook generation: Fix $\epsilon > \ \epsilon'' > \epsilon' > \lambda > 0$. Generate 2^{nR} i.i.d. zero mean Gaussian sequences $\mathbf{X}(m)$ with variance $(1-\gamma)$ for each $m \in [2^{nR}]$. By Lemmas 19 and Lemma 18, we may assume that the deterministic codebook satisfies (7.10)–(7.14).

Encoding: Given message m and gain sequence $\mathbf{g}$, the transmitter computes $\tilde{g}$ from the quantization function, and then sends $\sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{x}(m)$ (at time i signal $\sqrt{\varphi(\tilde{g}_i)}x_i(m)$ is sent) if $\|\mathbf{x}(m)\|^2 \leq n$; otherwise, it sends zero. Note that here we assume that the encoder knows the channel gains causally.

Decoding: Given $\mathbf{y}$ and $\mathbf{g}$, let $\nu < \epsilon$ and $\mathscr{S}$ be the set of messages $\hat{m}$ such that $(\mathbf{x}(\hat{m}), \tilde{\mathbf{g}}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(X', \tilde{G}, Y)$ where $\tilde{G}$ is the quantized random variable from G and some random variables (X', Y) that are conditionally Gaussian given $\tilde{G} = \tilde{g}$ with zero mean and covariance

$$\mathrm{Cov}\left(X', Y \middle| \tilde{G} = \tilde{g}\right) = \begin{bmatrix} 1 & \tilde{g}\sqrt{\varphi(\tilde{g})} \\ \tilde{g}\sqrt{\varphi(\tilde{g})} & a_{\tilde{g}} \end{bmatrix} \tag{7.100}$$

where $a_{\tilde{g}} \geq \tilde{g}^2\varphi(\tilde{g}) + \sigma^2$. Note that the following can be shown from (7.100).

$$X' \text{ is independent of } \tilde{G} \tag{7.101}$$

$$\mathbb{E}X'^2 = 1 \tag{7.102}$$

$$Y - \tilde{G}\sqrt{\varphi(\tilde{G})}X' \text{ is independent of } X' \text{ given } \tilde{G} \tag{7.103}$$

$$\mathrm{Var}\left(Y - \tilde{G}\sqrt{\varphi(\tilde{G})}X' \middle| \tilde{G}\right) \geq \sigma^2 \tag{7.104}$$

Now, we define the decoding function as

$$\Theta(\mathbf{y}, \tilde{\mathbf{g}}) = \underset{\hat{m} \in \mathscr{S}}{\arg\min} \left\| \mathbf{y} - \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{x}(\hat{m}) \right\|^2 \tag{7.105}$$

Analysis of the probability of error: Assume the legitimate transmitter sends message M. Then, we can upper bound the probability of error by the summation of the following error probabilities:

$$P_0 = \mathbb{P}\left\{M \notin \mathscr{S}\right\}, \tag{7.106}$$

$$P_1 = \mathbb{P}\left\{ \left\| \mathbf{Y} - \tilde{\mathbf{G}} \circ \sqrt{\varphi(\tilde{\mathbf{G}})} \circ \mathbf{x}(\hat{m}) \right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 \text{ for some } \hat{m} \in \mathscr{S} \setminus \{M\} \right\}. \tag{7.107}$$

We can prove with high probability

$$\begin{aligned} \frac{1}{n}\left\| \mathbf{x} \circ \left(\mathbf{G} - \tilde{\mathbf{G}}\right) \right\|^2 &= \frac{1}{n}\sum_{i=1}^{n} \left(x_i \left(G_i - \tilde{G}_i \right) \right)^2 & (7.108) \\ &\leq \frac{1}{n}\sum_{i=1}^{n} x_i^2 \mathbb{E}\left(G_i - \tilde{G}_i \right)^2 + \nu & (7.109) \\ &\leq 2\nu & (7.110) \end{aligned}$$

where (7.109) follows from the law of large numbers for non-identical independent random variables $x_i^2\left(G_i-\tilde{G}_i\right)^2$ and (7.110) follows from the facts that $\mathbb{E}\left[\left(G-\tilde{G}\right)^2\right]\leq\nu$, $\frac{1}{n}\sum_{i=1}^n x_i^2\leq 1$ and ν is sufficiently smaller than ϵ.

Consider any jammer sequence $\mathbf{s}$. We may assume sequence $\mathbf{G}$ is typical since it is drawn i.i.d. from the distribution $f_G(g)$. Similarly, $\tilde{\mathbf{G}}$ is also typical because it is from the corresponding discrete distribution $P_{\tilde{G}}(\tilde{g})$. Thus, $(\mathbf{s},\tilde{\mathbf{G}})$ is also typical with respect to some distribution $P_{\tilde{G}}(\tilde{g})f_{S|\tilde{G}}(s|\tilde{g})$ where $f_{S|\tilde{G}}(s|\tilde{g})$ is conditionally Gaussian. Note that we can make no assumptions about the conditional variances defining $f_{S|\tilde{G}}$, because the adversary is assumed to know G in its choice of s. By (7.10), with high probability $(\mathbf{x}(M),\mathbf{s},\tilde{\mathbf{G}})\in\mathcal{T}_{\epsilon'}^{(n)}(X,S,\tilde{G})$ where X is independent of $(S,\tilde{G})$, and $\mathbb{E}X^2=1,\mathbb{E}S^2\leq\Lambda$. Thus, by the conditional typicality lemma 2, with high probability $(\mathbf{x},\mathbf{s},\tilde{\mathbf{G}},\mathbf{V})\in\mathcal{T}_{\epsilon''}^{(n)}(X,S,\tilde{G},V)$ where $X,S,\tilde{G}$ are independent of V, and $\mathbb{E}V^2=\sigma^2$. Hence, using 7.110, we have $\left(\mathbf{x},\mathbf{s},\tilde{\mathbf{G}},\mathbf{V}+\mathbf{x}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}\circ\left(\mathbf{G}-\tilde{\mathbf{G}}\right)\right)\in\mathcal{T}_{\epsilon}^{(n)}(X,S,\tilde{G},V)$. Note that $\mathbf{Y}-\mathbf{x}\circ\tilde{\mathbf{G}}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}-\mathbf{s}=\mathbf{V}+\mathbf{x}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}\circ\left(\mathbf{G}-\tilde{\mathbf{G}}\right)$ and ν is sufficiently small compared to ϵ. In order to show that with high probability $M\in\mathscr{S}$, we need to compute the covariance matrix of (X,Y), where $Y=G\sqrt{\phi(\tilde{G})}X+S+V$, and show that it is in the form of (7.100). First, $\mathbb{E}\left(X^2|\tilde{G}=\tilde{g}\right)=\mathbb{E}X^2=1$ since X

is independent of $\tilde{G}$,

$$\begin{aligned}
&\mathbb{E}\left(X\left(G\sqrt{\varphi(\tilde{G})}X+S+V\right)\middle|\tilde{G}=\tilde{g}\right)= \\
&\qquad\qquad\mathbb{E}\left(X\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V+X\sqrt{\varphi(\tilde{G})}\left(G-\tilde{G}\right)\right)\middle|\tilde{G}=\tilde{g}\right)
\end{aligned}\tag{7.111}$$

$$=\mathbb{E}\left(X\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)\middle|\tilde{G}=\tilde{g}\right)+\mathbb{E}\left(X^2\sqrt{\varphi(\tilde{G})}\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right)\tag{7.112}$$

$$=\tilde{g}\sqrt{\varphi(\tilde{g})}\mathbb{E}X^2+\mathbb{E}\left(XS|\tilde{G}=\tilde{g}\right)+\mathbb{E}\left(XV|\tilde{G}=\tilde{g}\right)+\sqrt{\varphi(\tilde{g})}\mathbb{E}X^2\mathbb{E}\left(G-\tilde{G}\middle|\tilde{G}=\tilde{g}\right)\tag{7.113}$$

$$=\tilde{g}\sqrt{\varphi(\tilde{g})}+\sqrt{\varphi(\tilde{g})}\mathbb{E}X^2\left(\mathbb{E}\left(G\middle|\tilde{G}=\tilde{g}\right)-\tilde{g}\right)\tag{7.114}$$

$$=\tilde{g}\sqrt{\varphi(\tilde{g})},\tag{7.115}$$

where $\mathbb{E}\left(XS\middle|\tilde{G}=\tilde{g}\right)=0$ follows from the weak union rule since X is independent of (S,G), and $\mathbb{E}\left(G-\tilde{G}\middle|\tilde{G}=\tilde{g}\right)=0$ follows from the the definition of $\tilde{G}$.

$$
\begin{aligned}
&\mathbb{E}\left(\left(G\sqrt{\varphi(\tilde{G})}X+S+V\right)^2\middle|\tilde{G}=\tilde{g}\right)= \\
&\qquad\qquad \mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V+X\sqrt{\varphi(\tilde{G})}\left(G-\tilde{G}\right)\right)^2\middle|\tilde{G}=\tilde{g}\right) && (7.116)\\
&\quad=\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)^2\middle|\tilde{G}=\tilde{g}\right)+\mathbb{E}\left(X^2\varphi(\tilde{G})\left(G-\tilde{G}\right)^2\middle|\tilde{G}=\tilde{g}\right)\\
&\quad+2\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)X\sqrt{\varphi(\tilde{G})}\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right) && (7.117)\\
&\quad=\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)^2\middle|\tilde{G}=\tilde{g}\right)+\varphi(\tilde{g})\mathbb{E}X^2\mathbb{E}\left(\left(G-\tilde{G}\right)^2\middle|\tilde{G}=\tilde{g}\right)\\
&\qquad+2\tilde{g}\varphi(\tilde{g})\mathbb{E}\left(X^2\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right)+2\sqrt{\varphi(\tilde{g})}\mathbb{E}\left(XS\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right)\\
&\qquad+2\sqrt{\varphi(\tilde{g})}\mathbb{E}\left(XV\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right) && (7.118)\\
&\quad\geq\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)^2\middle|\tilde{G}=\tilde{g}\right)+2\tilde{g}\varphi(g)\mathbb{E}X^2\mathbb{E}\left(G-\tilde{G}\middle|\tilde{G}=\tilde{g}\right)\\
&\qquad+2\sqrt{\varphi(\tilde{g})}\mathbb{E}X\mathbb{E}\left(S\left(G-\tilde{G}\right)\middle|\tilde{G}=\tilde{g}\right)+2\sqrt{\varphi(\tilde{g})}\mathbb{E}X\mathbb{E}V\mathbb{E}\left(G-\tilde{G}\middle|\tilde{G}=\tilde{g}\right)\\
& && (7.119)\\
&=\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V\right)^2\middle|\tilde{G}=\tilde{g}\right) && (7.120)
\end{aligned}
$$

where (7.119) holds because $\varphi(\tilde{g})\mathbb{E}X^2\mathbb{E}\left(\left(G-\tilde{G}\right)^2\middle|\tilde{G}=\tilde{g}\right)>0$ and (7.120) follows from the fact that $(\tilde{G},G)$ are independent of (X,V), $\mathbb{E}V=\mathbb{E}X=0$ and the definition

$\tilde{g} = \mathbb{E}(G|\tilde{G} = \tilde{g})$.

$$\begin{aligned}
\mathbb{E}\left(\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X + S + V\right)^2 \middle| \tilde{G} = \tilde{g}\right) &= \mathbb{E}\left(\tilde{G}^2\varphi(\tilde{G})X^2 \middle| \tilde{G} = \tilde{g}\right) + \mathbb{E}\left(S^2 \middle| \tilde{G} = \tilde{g}\right) \\
&\quad + \mathbb{E}\left(V^2 \middle| \tilde{G} = \tilde{g}\right) + 2\mathbb{E}\left(\tilde{G}\sqrt{\varphi(\tilde{G})}XS \middle| \tilde{G} = \tilde{g}\right) + 2\mathbb{E}\left(\tilde{G}X\sqrt{\varphi(\tilde{G})}V \middle| \tilde{G} = \tilde{g}\right) \\
&\quad + 2\mathbb{E}\left(SV \middle| \tilde{G} = \tilde{g}\right) \qquad (7.121) \\
&= \tilde{g}^2\varphi(\tilde{g}) + \mathbb{E}\left(S^2 \middle| \tilde{G}=\tilde{g}\right) + \sigma^2 + 2\tilde{g}\sqrt{\varphi(\tilde{g})}\mathbb{E}\left(XS \middle| \tilde{G}=\tilde{g}\right) + 2\tilde{g}\sqrt{\varphi(\tilde{g})}\mathbb{E}\left(XV \middle| \tilde{G}=\tilde{g}\right) \\
&\quad + 2\mathbb{E}\left(SV \middle| \tilde{G}=\tilde{g}\right) \qquad (7.122) \\
&= \tilde{g}^2\varphi(\tilde{g}) + \mathbb{E}\left(S^2 \middle| \tilde{G} = \tilde{g}\right) + \sigma^2 \qquad (7.123) \\
&\geq \tilde{g}^2\varphi(\tilde{g}) + \sigma^2 \qquad (7.124)
\end{aligned}$$

where (7.123) follows from the weak union rule for X independent of $(S, \tilde{G})$ and V independent of $(S, \tilde{G})$. Therefore, the conditional covariance matrix of (X, Y) can be obtain from $\mathbb{E}X^2 = 1$, (7.115) and (7.124), and is the same as (7.100). Now, since $(\mathbf{x}(\hat{M}), \tilde{\mathbf{g}}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(X, \tilde{G}, Y)$ and the conditional covariance matrix of $(X(M), Y)$ satisfies (7.100), with high probability $M \in \mathscr{S}$, and P_0 vanishes as $n \to \infty$.

Using (7.110) and triangle inequality, we may upper bound P_1 by the following:

$$P_1 \leq \mathbb{P}\left\{ \left\|\mathbf{x}(m) \circ \tilde{\mathbf{G}}\sqrt{\varphi(\tilde{\mathbf{G}})} + \mathbf{s} + \mathbf{V} - \mathbf{x}(\hat{m}) \circ \tilde{\mathbf{G}}\sqrt{\varphi(\tilde{\mathbf{G}})}\right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 + 2n\nu \right. \\ \left. \text{for some } \hat{m} \in \mathscr{S} \setminus \{m\} \right\} \qquad (7.125)$$

Define the shorthand $\vec{X} = (XX'S\tilde{G}V)$. Let $\mathcal{V}$ denote a finite ϵ-dense subset in the set of all distributions of random vectors $\vec{X}$ that are determined by $P_{\tilde{G}}(\tilde{g})$ and a random vector $(XX'SV)$ distributed conditionally zero mean Gaussian given $\tilde{G}$ with bounded covariances at most $(1, 1, \Lambda, \sigma^2)$. Note that because the distribution of $P_{\tilde{G}}(\tilde{g})$ is completely known, the overall distribution of $\vec{X}$ can be determined by the conditional covariance matrix of $(XX'SV)$ given $\tilde{G} = \tilde{g}$ for each of the finitely many $\tilde{g}$ realizations, so $\mathcal{V}$ only needs to cover a compact set. Now, we may upper bound P_1 by

$$\sum_{\vec{X} \in \mathcal{V}} \frac{1}{N} \sum_{m=1}^{N} \mathbb{E}_{\tilde{G}} \left[e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{G}}) \right] \tag{7.126}$$

where

$$\begin{aligned} e_{\vec{X}}\left(m, \mathbf{s}, \tilde{\mathbf{g}}\right) = \mathbb{P}\Bigg\{ & \left(\mathbf{x}(m), \mathbf{x}(\hat{m}), \mathbf{s}, \tilde{\mathbf{g}}, \mathbf{V}\right) \in \mathcal{T}_\epsilon^{(n)}\left(\vec{X}\right), \\ & \left\| \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{x}(m) + \mathbf{s} + \mathbf{V} - \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{x}(\hat{m}) \right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 + 2n\nu \\ & \text{for some } \ \hat{m} \in \mathscr{S} \setminus \{m\} \Bigg\}. \end{aligned} \tag{7.127}$$

We will show that $\frac{1}{N}\sum_{m=1}^{N} e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) \to 0$ for all vectors $\tilde{\mathbf{g}}$ and all vectors $(XX'SV)$ which are Gaussian given $\tilde{G}$ (whether or not they are in $\mathcal{V}$). Let $Z = \tilde{G}\sqrt{\varphi(\tilde{G})}X + S + V - \tilde{G}\sqrt{\varphi(\tilde{G})}X'$. We may restrict ourselves to $\vec{X}$ where

$$\tilde{G} \sim P_{\tilde{G}}(\tilde{g}) \tag{7.128}$$

$$(X, X', S, V) \text{ are zero mean Gaussian given } \tilde{G} \tag{7.129}$$

$$X, (S, \tilde{G}), V \text{ are mutually independent,} \tag{7.130}$$

$$(X', \tilde{G}) \text{ are independent,} \tag{7.131}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = 1, \mathbb{E}S^2 \leq \Lambda, \mathbb{E}V^2 = \sigma^2 \tag{7.132}$$

$$(X', Z) \text{ are independent given } \tilde{G}, \tag{7.133}$$

$$\mathbb{E}\left[Z^2 \middle| \tilde{G}\right] \geq \sigma^2 \tag{7.134}$$

$$\mathrm{Var}(Z) \leq \sigma^2 + \Lambda + 2\nu \tag{7.135}$$

$$I(X; X'S\tilde{G}) < |R - I(X'; S\tilde{G})|^+ + \delta(\epsilon) \tag{7.136}$$

where (7.128)–(7.129) are obtained by the definition of $\mathscr{S}$, (7.130) holds since the codebook X, Gaussian noise V and fading gains $\tilde{G}$ are generated independently, and the adversary signal S may depend on $\tilde{G}$ but not the others, (7.131) follows from (7.101), (7.132) follows from the power constraints of the codebook, the adversary and the distribution of noise, (7.133)-(7.134) follows from (7.103)-(7.104), and (7.135) follows from (7.127). Let $\psi(\tilde{g}) = \mathbb{E}\left[Z^2 \middle| \tilde{G} = \tilde{g}\right] - \sigma^2$. Therefore, using (7.134) we have $\psi(\tilde{g}) \geq 0$, and by (7.135) we get $\mathbb{E}\psi(\tilde{G}) = \mathrm{Var}(Z) - \sigma^2 \leq \Lambda + 2\nu$. Note that using (7.14), we only need to consider the distributions that satisfies (7.136).

Observe that if $I(XV; X'S|\tilde{G}) = 0$, then we would have

$$\begin{aligned}
0 &= \mathbb{E}\left[X'Z|\tilde{G}\right] && (7.137)\\
&= \mathbb{E}\left[X'\left(\tilde{G}\sqrt{\varphi(\tilde{G})}X + S + V - \tilde{G}\sqrt{\varphi(\tilde{G})}X'\right)\middle|\tilde{G}\right] && (7.138)\\
&= \mathbb{E}\left[X'S\middle|\tilde{G}\right] - \mathbb{E}\left[\tilde{G}\sqrt{\varphi(\tilde{G})}X'^2\middle|\tilde{G}\right] && (7.139)\\
&= \mathbb{E}\left[X'S\middle|\tilde{G}\right] - \tilde{G}\sqrt{\varphi(\tilde{G})} && (7.140)
\end{aligned}$$

where (7.138) follows from (7.133), (7.139) follows from the assumption $I(XV; X'S|\tilde{G}) = 0$ in which X' is independent of (X, V), and (7.140) holds since X' is independent of $\tilde{G}$. Therefore, $\mathbb{E}\left[X'S\middle|\tilde{G}\right] = \tilde{G}\sqrt{\varphi(\tilde{G})}$ and the covariance matrix of S, X' given $\tilde{G}$ is equal to

$$\mathrm{Cov}\left(S, X'\middle|\tilde{G}\right) = \begin{bmatrix} \mathbb{E}\left[S^2\middle|\tilde{G}\right] & \tilde{G}\sqrt{\varphi(\tilde{G})} \\ \tilde{G}\sqrt{\varphi(\tilde{G})} & 1 \end{bmatrix}. \qquad (7.141)$$

The determinant of $\mathrm{Cov}\left(S, X'\middle|\tilde{G}\right)$ is $\mathbb{E}\left[S^2\middle|\tilde{G}\right] - \tilde{G}^2\varphi(\tilde{G})$ that should be non-negative since the covariance matrix must be positive semi-definite. Thus, its expectation is also non-negative:

$$0 \leq \mathbb{E}S^2 - \mathbb{E}\tilde{G}^2\varphi(\tilde{G}). \qquad (7.142)$$

However, since $\mathbb{E}S^2 \leq \Lambda$, (7.142) contradicts the initial assumption on φ in (7.98). Thus, there exists $\eta > 0$ such that

$$\eta \leq I(XV; X'S|\tilde{G}) = I(XV; X'|S\tilde{G}) \tag{7.143}$$

where we have used the fact that $I(XV; S) = 0$.

Probability $e_{\vec{X}}$ may be upper bounded by

$$e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) \leq \sum_{\hat{m}:(\mathbf{x}(m),\mathbf{x}(\hat{m}),\mathbf{s},\tilde{\mathbf{g}})\in\mathcal{T}_\epsilon^{(n)}(X,X',S,\tilde{G})} \mathbb{P}\left\{(\mathbf{x}(m), \mathbf{x}(\hat{m}), \mathbf{s}, \tilde{\mathbf{g}}, \mathbf{V}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S, \tilde{G}, V)\right\} \tag{7.144}$$

$$\leq \exp\left\{n\left[|R - I(X'; XS\tilde{G})|^+ - I(V; X'|XS\tilde{G}) + \delta(\epsilon)\right]\right. \tag{7.145}$$

where (7.145) follows from (7.13) and the joint typicality lemma 3.

We consider the following two cases.

Case (a): $R < I(X'; S\tilde{G})$. Applying this condition to (7.136), we get

$$\delta(\epsilon) > I(X; X'S\tilde{G}) \tag{7.146}$$

$$= I(X; X'|S\tilde{G}). \tag{7.147}$$

Since $I(X'; S\tilde{G}) \leq I(X'; XS\tilde{G})$ then $R - I(X'; XS\tilde{G}) < 0$. Considering (7.145), for any $m, \mathbf{s}, \tilde{\mathbf{g}}$ we have

$$e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) \leq \exp\left\{-n\left(I(V; X'|XS\tilde{G}) - \delta(\epsilon)\right)\right\} \tag{7.148}$$

$$= \exp\{-n(I(XV; X'|S\tilde{G}) - I(X; X'|S\tilde{G}) - \delta(\epsilon))\} \tag{7.149}$$

$$\leq \exp\{-n(\eta - 2\delta(\epsilon))\} \tag{7.150}$$

where (7.150) follows from (7.143) and (7.147). Therefore, $e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}})$ vanishes exponentially fast if $\delta(\epsilon)$ is sufficiently small.

Case (b): $R \geq I(X'; S\tilde{G})$. Then we may apply this condition to (7.136) as

$$\begin{aligned}
R &> I(X; X'S\tilde{G}) + I(X'; S\tilde{G}) - \delta(\epsilon) && (7.151)\\
&\geq I(X; X'|S\tilde{G}) + I(X'; S\tilde{G}) - \delta(\epsilon) && (7.152)\\
&= I(X'; XS\tilde{G}) - \delta(\epsilon). && (7.153)
\end{aligned}$$

Since $R - I(X'; XS\tilde{G}) + \delta(\epsilon) > 0$, we may upper bound (7.145) by

$$\begin{aligned}
e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) &\leq \exp\left(n\left[R - I(X'; XS\tilde{G}) - I(V; X'|XS\tilde{G}) + 2\delta(\epsilon)\right]\right) && (7.154)\\
&= \exp(n[R - I(X'; XS\tilde{G}V) + 2\delta(\epsilon)]) && (7.155)\\
&\leq \exp(n[R - I(X'; XSV|\tilde{G}) + 2\delta(\epsilon)]) && (7.156)
\end{aligned}$$

From (7.133)–(7.134), we obtain

$$\begin{aligned}
I\left(X'; XSV\middle|\tilde{G}\right) &\geq I\left(X'; \tilde{G}\sqrt{\varphi(\tilde{G})}X + S + V\middle|\tilde{G}\right) && (7.157)\\
&= I\left(X'; Z + \tilde{G}\sqrt{\varphi(\tilde{G})}X'\middle|\tilde{G}\right) && (7.158)\\
&= h\left(Z + \tilde{G}\sqrt{\varphi(\tilde{G})}X'\middle|\tilde{G}\right) - h\left(Z + \tilde{G}\sqrt{\varphi(\tilde{G})}X'\middle|\tilde{G}, X'\right) && (7.159)\\
&= \mathbb{E}_{\tilde{G}}\left[\frac{1}{2}\log 2\pi e\left(\tilde{G}^2\varphi(\tilde{G}) + \mathbb{E}\left[Z^2\middle|\tilde{G}\right]\right) - \frac{1}{2}\log 2\pi e \mathbb{E}\left[Z^2\middle|\tilde{G}\right]\right] && (7.160)\\
&= \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\mathbb{E}\left[Z^2|\tilde{G}\right]}\right)\right] && (7.161)\\
&= \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\psi(\tilde{G}) + \sigma^2 + 2\nu}\right)\right] && (7.162)
\end{aligned}$$

where (7.157) follows from data processing inequality, (7.161) follows from standard argument for the capacity of Gaussian channel, and (7.162) follows from the definition of ψ. Therefore, by the assumptions about R and Λ in (7.98)–(7.99), $R < I(X'; XSV|\tilde{G})$, so by (7.156) $e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}})$ is exponentially vanishing if $\delta(\epsilon)$ and ν are sufficiently small.

It is worth mentioning that this achievability proof also works for the case where both the adversary and encoder know the channel gains causally, or both know the gains non-causally. Since in all three cases the knowledge of the encoder is not more than the knowledge of the adversary, the jammer is able to impersonate the legitimate transmitter, and thereby symmetrize the channel, depending on the power allocation.

7.7.3 Achievability Proof (Gains Available Causally at Adversary and Non-causally at Encoder)

In this case, both the encoder and the decoder know the channel gains non-causally meaning that they know the whole $\mathbf{g}$ string including $(g_1, g_2, \cdots, g_n)$. However, the adversary only knows the gains causally, so at time i it only has access to $(g_1, g_2, \cdots, g_i)$. Therefore, both the encoder and the decoder have some extra common information $(g_{i+1}, g_{i+2}, \cdots, g_n)$ that the adversary does not know. In particular, the encoder and the decoder have always g_n which the adversary never knows except at time n. Hence, we can leverage this common knowledge between the encoder and the decoder as common randomness that is unknown to the jammer. Moreover, by the assumption that G is a continuous random variable with positive variance, in fact just G_n has infinite entropy, and thus can be considered a source of an infinite number of bits of common randomness. Therefore, we proceed to provide an achievability proof where the encoder and decoder are assumed to share an infinite source of common randomness. However, note that implementing this approach would require measuring G_n to an arbitrarily level of precision, which is not practical. Even so, the random code reduction technique of, for example, Csiszár and Körner, (2011), Lemma 12.8, can be used to show that only $O(\log n)$ bits of common randomness need to be extracted from G_n (or perhaps $G_{n-k}, \dots, G_n$ for some k) in order to achieve the same rate.

We first quantize G similar to the previous quantization in the achievability proof in Sec. 7.7.2. For a fix $\nu > 0$, $\tilde{G}$ is a deterministic function of G and $\mathbb{E}[(G-\tilde{G})^2] \leq \nu$. We also define $\tilde{G}$ as the expected value of G within each quantization set; that is, $\mathbb{E}[G|\tilde{G}] = \tilde{G}$.

Assume we have infinite amount of common randomness between the encoder and the decoder. Without loss of generality, assume $P = 1$. Let $\varphi(\tilde{g})$ be any function satisfying

$$\mathbb{E}\varphi(\tilde{G}) \leq 1 \tag{7.163}$$

$$R < \min_{\psi(\tilde{g}):\mathbb{E}\psi(\tilde{G})\leq\Lambda} \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\psi(\tilde{G})+\sigma^2}\right)\right]. \tag{7.164}$$

We construct a $(2^{nR}, n)$ code as follows:

Codebook generation: Let $\mathbf{X}(m)$ be a Gaussian codebook with variance $1-\gamma$ satisfying (7.10). This random codebook is generated from the infinite source of common randomness, so it is unknown to the adversary.

Encoding: Given message m and gain sequence $\mathbf{g}$, the transmitter first computes $\tilde{g}$ from the quantization function, and then sends $\sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{X}(m)$ (at time i signal $\sqrt{\varphi(\tilde{g}_i)}X_i(m)$ is sent) if $\mathbb{E}X^2 \leq 1$; otherwise, it sends zero. Note that here we assume that the encoder knows the channel gains non-causally.

Decoding: Given $\mathbf{y}$ and $\mathbf{g}$, let $\nu < \epsilon$ and let $\mathscr{S}$ be the set of messages $\hat{m}$ such that $(\mathbf{X}(\hat{m}), \tilde{\mathbf{g}}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(X', \tilde{G}, Y)$ where $\tilde{G}$ is the quantized random variable from G and (X', Y) are conditionally Gaussian given $\tilde{G} = \tilde{g}$ with zero mean and covariance matrix

$\Sigma_{\tilde{g}}$ as follows:

$$\Sigma_{\tilde{g}} = \text{Cov}\left(X', Y \middle| \tilde{G} = \tilde{g}\right) = \begin{bmatrix} 1 & \tilde{g}\sqrt{\varphi(\tilde{g})} \\ \tilde{g}\sqrt{\varphi(\tilde{g})} & a_{\tilde{g}} \end{bmatrix} \tag{7.165}$$

where $a_{\tilde{g}} \geq \tilde{g}^2\varphi(\tilde{g}) + \sigma^2$. Note that the following can be shown from (7.165):

$$X' \text{ is independent of } \tilde{G} \tag{7.166}$$

$$\mathbb{E}X'^2 = 1 \tag{7.167}$$

$$Y - \tilde{G}\sqrt{\varphi(\tilde{G})}X' \text{ is independent of } X' \text{ given } \tilde{G} \tag{7.168}$$

$$\text{Var}\left(Y - \tilde{G}\sqrt{\varphi(\tilde{G})}X' \middle| \tilde{G}\right) \geq \sigma^2 \tag{7.169}$$

Now, we define the decoding function as

$$\Theta(\mathbf{y}, \tilde{\mathbf{g}}) = \underset{\hat{m} \in \mathscr{S}}{\arg\min} \left\| \mathbf{y} - \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{X}(\hat{m}) \right\|^2 \tag{7.170}$$

Analysis of the probability of error: Assume the legitimate transmitter sends message M. Then, we can upper bound the probability of error by the summation of the following error probabilities:

$$P_0 = \mathbb{P}\left\{M \notin \mathscr{S}\right\}, \tag{7.171}$$

$$P_1 = \mathbb{P}\left\{ \left\| \mathbf{Y} - \tilde{\mathbf{G}} \circ \sqrt{\varphi(\tilde{\mathbf{G}})} \circ \mathbf{X}(\hat{m}) \right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 \text{ for some } \hat{m} \in \mathscr{S} \setminus \{M\} \right\} \tag{7.172}$$

We can prove with high probability

$$\begin{aligned} \frac{1}{n}\left\| \mathbf{X} \circ \left(\mathbf{G} - \tilde{\mathbf{G}}\right) \right\|^2 &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i\left(G_i - \tilde{G}_i\right)\right)^2 && (7.173) \\ &\leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}X_i^2 \mathbb{E}\left(G_i - \tilde{G}_i\right)^2 + \nu && (7.174) \\ &\leq 2\nu && (7.175) \end{aligned}$$

where (7.174) follows from the law of large numbers for non-identical independent random variables $X_i^2\left(G_i-\tilde{G}_i\right)^2$ and (7.175) follows from the facts that $\mathbb{E}\left[\left(G-\tilde{G}\right)^2\right]\leq\nu$, $\mathbb{E}X_i^2\leq 1$ and ν is sufficiently smaller than ϵ.

Consider any jammer sequence $\mathbf{s}$. We may assume sequence $\mathbf{G}$ is typical since it is drawn i.i.d. from the distribution $f_G(g)$. The quantized version $\tilde{\mathbf{G}}$ is also typical since it is a discrete function of $\mathbf{G}$ with distribution $P_{\tilde{G}}(\tilde{g})$. Thus, $(\mathbf{s},\tilde{\mathbf{G}})$ is also typical with respect to some distribution $P_{\tilde{G}}(\tilde{g})f_{S|\tilde{G}}(s|\tilde{g})$ where $f_{S|\tilde{G}}(s|\tilde{g})$ is conditionally Gaussian. Note that we can make no assumptions about the conditional variances defining $f_{S|\tilde{G}}$, because the adversary is assumed to know $\tilde{G}$ in its choice of s. By (7.10), with high probability $(\mathbf{X}(M),\mathbf{s},\tilde{\mathbf{G}})\in\mathcal{T}_{\epsilon'}^{(n)}(X,S,\tilde{G})$ where X is independent of $(S,\tilde{G})$, and $\mathbb{E}X^2=1,\mathbb{E}S^2\leq\Lambda$. Thus, by the conditional typicality lemma 2, with high probability $(\mathbf{X},\mathbf{s},\tilde{\mathbf{G}},\mathbf{V})\in\mathcal{T}_{\epsilon}^{(n)}(X,S,\tilde{G},V)$ where $X,S,\tilde{G}$ are independent of V, and $\mathbb{E}V^2=\sigma^2$. Hence, using 7.175, we have $\left(\mathbf{X},\mathbf{s},\tilde{\mathbf{G}},\mathbf{V}+\mathbf{X}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}\circ\left(\mathbf{G}-\tilde{\mathbf{G}}\right)\right)\in\mathcal{T}_{\epsilon}^{(n)}(X,S,\tilde{G},V)$. Note that $\mathbf{Y}-\mathbf{X}\circ\tilde{\mathbf{G}}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}-\mathbf{s}=\mathbf{V}+\mathbf{X}\circ\sqrt{\varphi(\tilde{\mathbf{G}})}\circ\left(\mathbf{G}-\tilde{\mathbf{G}}\right)$ and ν is sufficiently small compared to ϵ. Referring to the previous achievability proof in Sec. 7.7.2, the conditional covariance matrix of (X,Y) can be similarly obtained from $\mathbb{E}X^2=1$, (7.115) and (7.124), so the conditional covariance matrix is the same as the one in (7.165). Now, since $(\mathbf{X}(\hat{M}),\tilde{\mathbf{g}},\mathbf{y})\in\mathcal{T}_{\epsilon}^{(n)}(X,\tilde{G},Y)$ and the conditional covariance matrix of $(X(M),Y)$ satisfies (7.165), with high probability $M\in\mathscr{S}$, so P_0 vanishes as $n\to\infty$.

Using (7.175) and triangle inequality, we may upper bound P_1 by the following:

$$P_1\leq\mathbb{P}\Bigg\{\left\|\mathbf{X}(m)\circ\tilde{\mathbf{G}}\sqrt{\varphi(\tilde{\mathbf{G}})}+\mathbf{s}+\mathbf{V}-\mathbf{X}(\hat{m})\circ\tilde{\mathbf{G}}\sqrt{\varphi(\tilde{\mathbf{G}})}\right\|^2\leq\|\mathbf{s}+\mathbf{V}\|^2+2n\nu \\ \text{for some } \hat{m}\in\mathscr{S}\setminus\{m\}\Bigg\} \quad (7.176)$$

Define the shorthand $\vec{X} = (XX'S\tilde{G}V)$. Let $\mathcal{V}$ denote a finite ϵ-dense subset in the set of all distributions of random vectors $\vec{X}$ that are determined by $P_{\tilde{G}}(\tilde{g})$ and a random vector $(XX'SV)$ distributed conditionally zero mean Gaussian given $\tilde{G}$ with bounded covariances at most $(1, 1, \Lambda, \sigma^2)$. Note that because the distribution of $P_{\tilde{G}}(\tilde{g})$ is completely known, the overall distribution of $\vec{X}$ can be determined by the conditional covariance matrix of $(XX'SV)$ given $\tilde{G} = \tilde{g}$ for each of the finitely many $\tilde{g}$ realizations, so $\mathcal{V}$ only needs to cover a compact set. We may now upper bound P_1 by

$$\sum_{\vec{X} \in \mathcal{V}} \frac{1}{N} \sum_{m=1}^{N} \mathbb{E}_{\tilde{G}} \left[e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{G}}) \right] \tag{7.177}$$

where

$$\begin{aligned} e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) = \Big\{ &(\mathbf{X}(m), \mathbf{X}(\hat{m}), \mathbf{s}, \tilde{\mathbf{g}}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(\vec{X}), \\ &\left\| \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{X}(m) + \mathbf{s} + \mathbf{V} - \tilde{\mathbf{g}} \circ \sqrt{\varphi(\tilde{\mathbf{g}})} \circ \mathbf{X}(\hat{m}) \right\|^2 \leq \|\mathbf{s} + \mathbf{V}\|^2 + 2n\nu \\ &\text{for some } \hat{m} \in \mathscr{S} \setminus \{M\} \Big\} \end{aligned} \tag{7.178}$$

Now, it suffices to show that $\frac{1}{N} \sum_{m=1}^{N} e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}})$ vanishes for all typical vectors $\mathbf{g}$ and all vectors $(XX'SV)$ which are Gaussian given $\tilde{G}$ (whether or not they are in $\mathcal{V}$). Let $Z = \tilde{G}\sqrt{\varphi(\tilde{G})}X + S + V - \tilde{G}\sqrt{\varphi(\tilde{G})}X'$. We have established that $\vec{X}$ satisfies the following:

$$\tilde{G} \sim f_G(\tilde{g}) \tag{7.179}$$

$$(X, X', S, V) \text{ are zero mean Gaussian given } \tilde{G} \tag{7.180}$$

$$X, (S, \tilde{G}), V \text{ are mutually independent,} \tag{7.181}$$

$$X' \text{ is independent of } (\tilde{G}, S), \tag{7.182}$$

$$\mathbb{E}X^2 = \mathbb{E}X'^2 = 1, \mathbb{E}S^2 \leq \Lambda, \mathbb{E}V^2 = \sigma^2 \tag{7.183}$$

$$(X', Z) \text{ are independent given } \tilde{G}, \tag{7.184}$$

$$\mathbb{E}\left[Z^2 \middle| \tilde{G}\right] \geq \sigma^2 \tag{7.185}$$

$$\mathrm{Var}(Z) \leq \sigma^2 + \Lambda. \tag{7.186}$$

where (7.179)–(7.180) are obtained by the definition of $\mathscr{S}$, (7.181) follows since the codebook X, Gaussian noise V, fading gains $\tilde{G}$ are generated independently while $\tilde{G}$ may depend on S but not the others, (7.182) follows from (7.166), (7.183) follows from the power constraints for the codebook, the adversary and the Gaussian noise, (7.184)-(7.185) follows by (7.168)-(7.169), and (7.186) follows by (7.178). Let $\psi(\tilde{g}) = \mathbb{E}[Z^2|\tilde{G} = \tilde{g}] - \sigma^2$. By (7.185) $\psi(\tilde{g}) \geq 0$, and by (7.186) $\mathbb{E}\psi(\tilde{G}) = \mathrm{Var}(Z) - \sigma^2 \leq \Lambda + 2\nu$.

Now, using jointly typicality lemma in ElGamal and Kim, (2011), Remark 2.2 we may upper bound $e_{\vec{X}}$ as follows:

$$e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}}) \leq \sum_{\hat{m} \in \mathscr{S} \setminus \{m\}} \mathbb{P}\left\{(\mathbf{X}(m), \mathbf{X}(\hat{m}), \mathbf{s}, \tilde{\mathbf{g}}, \mathbf{V}) \in \mathcal{T}_{\epsilon}^{(n)}(X, X', S, \tilde{G}, V)\right\} \tag{7.187}$$

$$\leq \exp\{n(R - I(X'; X, S, V, \tilde{G}) + \epsilon)\} \tag{7.188}$$

where $\mathbf{X}(\hat{m})$ is independent of $(\mathbf{X}(m), \mathbf{s}, \tilde{\mathbf{g}}, \mathbf{V})$. From (7.184)–(7.185), we obtain

$$\begin{aligned}
I(X';XS\tilde{G}V) &= I(X';XSV|\tilde{G}) + I(X';\tilde{G}) && (7.189)\\
&\geq I(X';\tilde{G}\sqrt{\varphi(\tilde{G})}X+S+V|\tilde{G}) && (7.190)\\
&= I(X';Z+\tilde{G}\sqrt{\varphi(\tilde{G})}X'|\tilde{G}) && (7.191)\\
&= h(Z+\tilde{G}\sqrt{\varphi(\tilde{G})}X'|\tilde{G})\\
&\quad - h(Z+\tilde{G}\sqrt{\varphi(\tilde{G})}X'|\tilde{G},X') && (7.192)\\
&= \mathbb{E}_{\tilde{G}}\left[\frac{1}{2}\log 2\pi e(\tilde{G}^2\varphi(\tilde{G})\mathbb{E}[X'^2|\tilde{G}]+\mathbb{E}[Z^2|\tilde{G}])\right.\\
&\qquad\qquad \left. -\frac{1}{2}\log 2\pi e\mathbb{E}[Z^2|\tilde{G}]\right] && (7.193)\\
&= \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\mathbb{E}[Z^2|\tilde{G}]}\right)\right] && (7.194)\\
&= \mathbb{E}_{\tilde{G}}\left[C\left(\frac{\tilde{G}^2\varphi(\tilde{G})}{\psi(\tilde{G})+\sigma^2+2\nu}\right)\right] && (7.195)
\end{aligned}$$

where (7.190) follows from data processing inequality, (7.194) follows from standard argument for the capacity of Gaussian channel, and (7.195) follows from the definition of ψ. Therefore, by the assumptions about R and Λ in (7.163)–(7.164), $R < I(X';XSV|\tilde{G})$, so by (7.188) $e_{\vec{X}}(m, \mathbf{s}, \tilde{\mathbf{g}})$ is exponentially vanishing if $\delta(\epsilon)$ and ν are sufficiently small.

Therefore, if we have infinite amount of common randomness between the encoder and the decoder which the adversary does not know it but knows the distribution of $\mathbf{X}$, then the adversary can choose its signal as a function of both the channel gains and a random codeword. Note that the adversary knows the codebook but not the common randomness since it contains infinite amount of numbers. However, since it does not know the common randomness, with high probability (as we have proven above) it can not symmetrize the channel any more. Thus, according to the random code reduction in Csiszár and Körner, (2011), Lemma 12.8 we only faced with a standard fading channel without any adversary with the channel gains available at the encoder and the decoder with the increased noise variance by the power of the adversary as in $\psi(G)+\sigma^2$, and the rate should be less than $\min_{\psi(g):\mathbb{E}\psi(G)\leq\Lambda} \mathbb{E}_G\left[C\left(\frac{G^2\varphi(G)}{\psi(G)+\sigma^2}\right)\right]$.

7.8 Proof of Lemmas 18 and 19

In order to prove (7.10), we use our proof in Hosseinigoki and Kosut, (2017), Lemma 6 for one codebook. Moreover, to obtain (7.13)–(7.14), we apply the corresponding proof of the equations in Hughes, (1997), Lemma 1 for Gaussian distributions. Note that Hughes, (1997) focuses on discrete alphabets, but the same proofs can be extended to Gaussian distributions by quantization of the set of continuous random variables in the following way.

Let $\mathbf{X}_i$ be Gaussian i.i.d. n-length random vectors (codebook) independent from each other with $\text{Var}(X)=1$. First let $\mathbf{g}\in\mathbb{R}^n$ be a typical realization of n i.i.d. continuous random variable G with probability density function $f_G(g)$. Next, we quantize the set of all $\mathbf{g}\in\mathbb{R}^n$, into a ν-dense subset $\mathcal{G}^n$. For a fixed $\mathbf{g}\in\mathcal{G}^n$, fix $\mathbf{x}\in\mathcal{T}_\epsilon^{(n)}(X), \mathbf{s}\in\mathscr{U}^n$ and a covariance matrix $\text{Cov}(X,X',S|G=g)\in\mathcal{V}^{3\times3}$ such that $\mathscr{U}^n$ is a ν-dense subset of $\mathbb{R}^n$ for $\mathbf{s}$ such that $||\mathbf{s}||^2\leq n\Lambda$, and $\mathcal{V}^{3\times3}$ is a ν-dense subset of $\mathbb{R}^{3\times3}$ for positive definite covariance matrices with diagonals at most $(1,1,\Lambda)$.

Using the similar proof Lemma 1 inHughes, (1997), we obtain for given $(\mathbf{x},\mathbf{s},\mathbf{g})$ and covariance matrix $\text{Cov}(X,X',S|G=g)$ that the complement of each event in (7.13)–(7.14) happens with decreasingly doubly exponential probability for sufficiently large n meaning that

$$\mathbb{P}\Big\{\big|\{m' : (\mathbf{x}, \mathbf{x}(m'), \mathbf{s}, \mathbf{g}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S, G)\}\big| \le \exp\big\{n\big[|R - I(X'; XSG)|^+ + \delta(\epsilon)\big]\big\}\Big\}$$
$$< \exp(-\exp(n\sigma(\epsilon))), \tag{7.196}$$

$$\mathbb{P}\Big\{\frac{1}{N}\big|\{m : (\mathbf{x}(m), \mathbf{x}(m'), \mathbf{s}, \mathbf{g}) \in \mathcal{T}_\epsilon^{(n)}(X, X', S, G) \text{ for some } m' \ne m\}\big| \le 2\exp\{-n\delta(\epsilon)/2\}\Big\}$$
$$< \exp(-\exp(n\sigma(\epsilon))), \text{ if } I(X; X'SG) \ge |R - I(X'; SG)|^+ + \delta(\epsilon), \tag{7.197}$$

Then, in order to complete the proof, since for any fixed ν the cardinality of finite set $\mathscr{U}^n$ is only increasingly exponentially in n, and the set $\mathcal{V}^{3\times3}$ is finite along with the doubly decreasing exponential probabilities in (7.196)–(7.197), we derive that with probability approaching to 1, all inequalities in (7.13)–(7.14) hold simultaneously for sufficiently large n. Since these inequalities hold for every element in the finite sets $\mathscr{U}^n$ and $\mathcal{V}^{3\times3}$, then for any vector $\mathbf{s}, \mathbf{x}$ and any given covariance matrix $\mathrm{Cov}(X, X', S|G = g)$ (with $\|\mathbf{x}\|^2 = n, \|\mathbf{s}\|^2 \le n\Lambda$) which is not in its corresponding ν-dense subset, there exists a point in the corresponding ν-dense subset that is close enough to it (in its ν distance neighborhood). Now, by using the continuity properties of all sets, we may conclude that (7.13)–(7.14) hold also for any point which is not in the ν-dense subset.

REFERENCES

Ahlswede, R. (1971). “The capacity of a channel with arbitrarily varying additive Gaussian channel probability functions.” In *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes,* 13–21. September.

———. (1978). “Elimination of correlation in random codes for arbitrarily varying channels.” *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 44 (2): 159–175. doi:10.1007/BF00533053.

Ahlswede, R., and N. Cai. (1999). “Arbitrarily varying multiple-access channels. I. Ericson’s symmetrizability is adequate, Gubner’s conjecture is true.” *IEEE Transactions on Information Theory* 45, no. 2 (March): 742–749. doi:10.1109/18.749024.

Ahlswede, R., and J. Wolfowitz. (1969). “Correlated decoding for channels with arbitrarily varying channel probability functions.” *Information and Control* 14 (5): 457–473. doi:https://doi.org/10.1016/S0019-9958(69)90157-0.

Annapureddy, V. S., and V. V. Veeravalli. (2009). “Gaussian Interference Networks: Sum Capacity in the Low-Interference Regime and New Outer Bounds on the Capacity Region.” *IEEE Transactions on Information Theory* 55, no. 7 (July): 3032–3050. doi:10.1109/TIT.2009.2021380.

Barros, J., and M. R. D. Rodrigues. (2006). “Secrecy Capacity of Wireless Channels.” In *2006 IEEE International Symposium on Information Theory,* 356–360. July. doi:10.1109/ISIT.2006.261613.

Bergmans, P. (1974). “A simple converse for broadcast channels with additive white Gaussian noise (Corresp.)” *IEEE Transactions on Information Theory* 20, no. 2 (March): 279–280. doi:10.1109/TIT.1974.1055184.

Blachman, N. M., and L. Few. (1963). “Multiple packing of spherical caps.” *Mathematika* 10 (1): 84–88. doi:10.1112/S0025579300003405.

Blackwell, D., L. Breiman, and A. J. Thomasian. (1960). “The Capacities of Certain Channel Classes Under Random Coding.” *Ann. Math. Statist.* 31, no. 3 (September): 558–567. doi:10.1214/aoms/1177705783.

Boche, H., and R. F. Schaefer. (2014). "List decoding for arbitrarily varying multiple access channels with conferencing encoders." In *2014 IEEE International Conference on Communications (ICC),* 1934–1940. June. doi:10.1109/ICC.2014.6883606.

Bresler, G., and D. Tse. (2008). "The two-user Gaussian interference channel: a deterministic view." *European Transactions on Telecommunications* 19 (4): 333–354. doi:10.1002/ett.1287.

Cover, T. (1972). "Broadcast channels." *IEEE Transactions on Information Theory* 18, no. 1 (January): 2–14. doi:10.1109/TIT.1972.1054727.

Cover, Thomas M. (1975). "Some Advances in Broadcast Channels," edited by A.J. Viterbi, 4:229–260. Advances in Communication Systems. Elsevier.

Csiszár, I. (1992). "Arbitrarily varying channels with general alphabets and states." *IEEE Transactions on Information Theory* 38, no. 6 (November): 1725–1742. doi:10.1109/18.165446.

Csiszár, I., and J. Körner. (2011). *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge.

Csiszár, I., and P. Narayan. (1991). "Capacity of the Gaussian arbitrarily varying channel." *IEEE Transactions on Information Theory* 37, no. 1 (January): 18–26. doi:10.1109/18.61125.

———. (1988)(a). "Arbitrarily varying channels with constrained inputs and states." *IEEE Transactions on Information Theory* 34, no. 1 (January): 27–34. doi:10.1109/18.2598.

———. (1988)(b). "The capacity of the arbitrarily varying channel revisited: positivity, constraints." *IEEE Transactions on Information Theory* 34, no. 2 (March): 181–193. doi:10.1109/18.2627.

Dey, B. K., S. Jaggi, M. Langberg, and A. D. Sarwate. (2010). "Coding against delayed adversaries." In *2010 IEEE International Symposium on Information Theory,* 285–289. June. doi:10.1109/ISIT.2010.5513325.

———. (2013). "Upper Bounds on the Capacity of Binary Channels With Causal Adversaries." *IEEE Transactions on Information Theory* 59, no. 6 (June): 3753–3763. doi:10.1109/TIT.2013.2245721.

ElGamal, A., and Y. H. Kim. (2011). *Network information theory.* Cambridge University Press.

Ericson, T. (1985). “Exponential error bounds for random codes in the arbitrarily varying channel.” *IEEE Transactions on Information Theory* 31, no. 1 (January): 42–48. doi:10.1109/TIT.1985.1056995.

Etkin, R. H., D. N. C. Tse, and Hua Wang. (2008). “Gaussian Interference Channel Capacity to Within One Bit.” *Information Theory, IEEE Transactions on* 54, no. 12 (December): 5534–5562. doi:10.1109/TIT.2008.2006447.

Gamal, A. E. 1981. “The capacity of the physically degraded Gaussian broadcast channel with feedback (Corresp.)” *IEEE Transactions on Information Theory* 27, no. 4 (July): 508–511. doi:10.1109/TIT.1981.1056372.

Goldsmith, A. J., and P. P. Varaiya. (1997). “Capacity of fading channels with channel side information.” *IEEE Transactions on Information Theory* 43, no. 6 (November): 1986–1992. doi:10.1109/18.641562.

Gubner, J. A. (1990). “On the deterministic-code capacity of the multiple-access arbitrarily varying channel.” *IEEE Transactions on Information Theory* 36, no. 2 (March): 262–275. doi:10.1109/18.52472.

———. (1992). “On the capacity region of the discrete additive multiple-access arbitrarily varying channel.” *IEEE Transactions on Information Theory* 38, no. 4 (July): 1344–1347. doi:10.1109/18.144713.

Han, Te, and K. Kobayashi. (1981). “A new achievable rate region for the interference channel.” *Information Theory, IEEE Transactions on* 27, no. 1 (January): 49–60. doi:10.1109/TIT.1981.1056307.

Hof, E., and S. I. Bross. (2006). “On the Deterministic-Code Capacity of the Two-User Discrete Memoryless Arbitrarily Varying General Broadcast Channel With Degraded Message Sets.” *IEEE Transactions on Information Theory* 52, no. 11 (November): 5023–5044. doi:10.1109/TIT.2006.883543.

Hosseinigoki, F., and O. Kosut. (2016). “The Gaussian interference channel in the presence of a malicious jammer.” In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton),* 679–686. September. doi:10.1109/ALLERTON.2016.7852297.

———. (2017). *The Gaussian Interference Channel in the Presence of Malicious Jammers.* [Online] arXiv:1712.04133. Submitted to IEEE Trans. on Information Theory. December.

Hosseinigoki, F., and O. Kosut. (2018). “Capacity of the Gaussian Arbitrarily-Varying Channel with List Decoding.” In *2018 IEEE International Symposium on Information Theory (ISIT),* 471–475. June. doi:10.1109/ISIT.2018.8437866.

———. (2019)(a). “Capacity of Gaussian Arbitrarily-Varying Fading Channels.” In *2019 53rd Annual Conference on Information Sciences and Systems (CISS),* 1–6. March. doi:10.1109/CISS.2019.8693030.

———. (2019)(b). “List-Decoding Capacity of the Gaussian Arbitrarily-Varying Channel.” *Entropy* 21 (6). doi:10.3390/e21060575.

Hughes, B. L. (1997). “The smallest list for the arbitrarily varying channel.” *IEEE Transactions on Information Theory* 43, no. 3 (May): 803–815. doi:10.1109/18.568692.

Hughes, B., and P. Narayan. (1987). “Gaussian arbitrarily varying channels.” *IEEE Transactions on Information Theory* 33, no. 2 (March): 267–284. doi:10.1109/TIT.1987.1057288.

———. (1988). “The capacity of a vector Gaussian arbitrarily varying channel.” *IEEE Transactions on Information Theory* 34, no. 5 (September): 995–1003. doi:10.1109/18.21222.

Jahn, J. (1981). “Coding of arbitrarily varying multiuser channels.” *IEEE Transactions on Information Theory* 27, no. 2 (March): 212–226. doi:10.1109/TIT.1981.1056320.

Jindal, N., S. Vishwanath, and A. Goldsmith. (2004). “On the duality of Gaussian multiple-access and broadcast channels.” *IEEE Transactions on Information Theory* 50, no. 5 (May): 768–783. doi:10.1109/TIT.2004.826646.

Lapidoth, A. (1996). “Nearest neighbor decoding for additive non-Gaussian noise channels.” *IEEE Transactions on Information Theory* 42, no. 5 (September): 1520–1529. doi:10.1109/18.532892.

Li, Z., R. Yates, and W. Trappe. (2010). “Achieving Secret Communication for Fast Rayleigh Fading Channels.” *IEEE Transactions on Wireless Communications* 9, no. 9 (September): 2792–2799. doi:10.1109/TWC.2010.080210.090948.

Nitinawarat, S. (2013). “On the Deterministic Code Capacity Region of an Arbitrarily Varying Multiple-Access Channel Under List Decoding.” *IEEE Transactions on Information Theory* 59, no. 5 (May): 2683–2693. doi:10.1109/TIT.2013.2242514.

Pereg, U., and Y. Steinberg. (2017). "The arbitrarily varying degraded broadcast channel with causal side information at the encoder." In *2017 IEEE International Symposium on Information Theory (ISIT),* 1033–1037. June. doi:10.1109/ISIT.2017.8006685.

Sarwate, A. D. (2010). "Coding against myopic adversaries." In *2010 IEEE Information Theory Workshop,* 1–5. August. doi:10.1109/CIG.2010.5592896.

Sarwate, A. D., and M. Gastpar. (2012). "List-Decoding for the Arbitrarily Varying Channel Under State Constraints." *IEEE Transactions on Information Theory* 58, no. 3 (March): 1372–1384. doi:10.1109/TIT.2011.2178153.

Schaefer, R. F., and H. Boche. (2014). "List Decoding for Arbitrarily Varying Broadcast Channels With Receiver Side Information." *IEEE Transactions on Information Theory* 60, no. 8 (August): 4472–4487. doi:10.1109/TIT.2014.2326412.

Shannon, Claude E. (1961). "Two-way Communication Channels." In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics,* 611–644. University of California Press. https://projecteuclid.org/euclid.bsmsp/1200512185.

Stiglitz, I. (1966). "Coding for a class of unknown channels." *IEEE Transactions on Information Theory* 12, no. 2 (April): 189–195. doi:10.1109/TIT.1966.1053861.

Wang, P., G. Yu, and Z. Zhang. (2007). "On the Secrecy Capacity of Fading Wireless Channel with Multiple Eavesdroppers." In *2007 IEEE International Symposium on Information Theory,* 1301–1305. June. doi:10.1109/ISIT.2007.4557128.

Winshtok, A., and Y. Steinberg. (2006). "The Arbitrarily Varying Degraded Broadcast Channel with States Known at the Encoder." In *2006 IEEE International Symposium on Information Theory,* 2156–2160. July. doi:10.1109/ISIT.2006.261932.

Wyner, A. (1974). "Recent results in the Shannon theory." *IEEE Transactions on Information Theory* 20, no. 1 (January): 2–10. doi:10.1109/TIT.1974.1055171.