

Syracuse University

SURFACE

Dissertations - ALL

SURFACE

January 2015

CHARACTERIZATION OF FUNDAMENTAL COMMUNICATION LIMITS OF STATE-DEPENDENT INTERFERENCE NETWORKS

Ruchen Duan
Syracuse University

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Engineering Commons](#)

Recommended Citation

Duan, Ruchen, "CHARACTERIZATION OF FUNDAMENTAL COMMUNICATION LIMITS OF STATE-DEPENDENT INTERFERENCE NETWORKS" (2015). *Dissertations - ALL*. 339.

<https://surface.syr.edu/etd/339>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

Interference management is one of the key techniques that drive evolution of wireless networks from one generation to another. Techniques in current cellular networks to deal with interference follow the basic principle of orthogonalizing transmissions in time, frequency, code, and space. My PhD work investigate information theoretic models that represent a new perspective/technique for interference management. The idea is to explore the fact that an interferer knows the interference that it causes to other users noncausally and can/should exploit such information for canceling the interference. In this way, users can transmit simultaneously and the throughput of wireless networks can be substantially improved. We refer to the interference treated in such a way as “dirty interference” or “noncausal state”.

In particular, my PhD thesis investigates two classes of information theoretic models and develops dirty interference cancelation schemes that achieve the fundamental communication limits. One class of models (referred to as state-dependent interference channels) capture the scenarios that users help each other to cancel dirty interference. The other class of models (referred to as state-dependent channels with helper) capture the scenarios that one dominate user interferes a number of other users and assists those users to cancel its dirty interference. For both classes of models, by comparing the corresponding achievable rate regions with the outer bounds on the capacity region. We characterize the channel parameters under which the developed inner bounds meet the outer bounds either partially or fully, and thus establish the capacity regions or partial boundaries of the capacity regions.

CHARACTERIZATION OF FUNDAMENTAL
COMMUNICATION LIMITS OF STATE-DEPENDENT
INTERFERENCE NETWORKS

by

Ruchen Duan

B.E.(EE), Beijing University of Posts and Telecommunications, 2010

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Syracuse University
August 2015

Copyright © 2015 Ruchen Duan

All rights reserved

ACKNOWLEDGMENTS

My deepest gratitude goes first and foremost to my advisor, Prof. Yingbin Liang. I sincerely thank her for her tutoring and support for the past five years. I learned a lot not only from her knowledge and wisdom, but also from her working style and passion. I still remember how she helped to formulate me from a immature student into a mature researcher. Every discussion with her was a training of critical thinking and logical analyzing, which makes my vague ideas into rigorous derivations. All my achievements are the results of her endeavor and patience.

I would also thank my collaborators, Prof. Shlomo Shamai and Prof. Ashish Khisti for their inspiring and stimulating suggestions and comments on our works. As learned scholars, their guidance enriched my ideology and broadened my horizon. Their insightful thoughts always contributed to a better understanding of the problems. Our joint works led to Chapter 4, 5, 6 and 7 of this thesis.

I thank my defense committee members, for carefully reading my thesis and giving me helpful suggestions. They are (in alphabetic order) Prof. Biao Chen, Prof. Mustafa C. Gursoy, Prof. Lixin Shen, Prof. Jian Tang and Prof. Pramod K. Varshney.

Sincere thanks also extend to my family for their selfless support. I am indebted to them for their belief in me. They are the people who make me complete, and who provide me the courage to face with anything confronted.

Thanks to my office mates, Jiayao Hu, Shaofeng Zou, Weiguang Wang, Huishuan Zhang, Anhong He, Yunhao Sun and Yi Zhou, for their help in study and life. Their

enthusiasm in research and professional attitude to work are great inspiration and encourage during my PhD study.

I acknowledge Syracuse University Graduate Fellowship program for providing funding for my Ph.D. study, and additional support from the National Science Foundation.

TABLE OF CONTENTS

Abstract	i
List of Figures	x
1 Introduction	1
1.1 Motivation	1
1.2 Channel Models	3
1.3 Related Work	9
1.4 Summary of Contributions and Thesis Organization	12
2 State-Dependent Interference Channel	15
2.1 Channel Model	15
2.2 Very Strong Interference Regime	17
2.2.1 State-Dependent Regular IC	17
2.2.2 State-Dependent Z-IC	22
2.2.3 Comparison of State-Dependent Regular IC and Z-IC	24
2.3 Strong Interference Regime	25
2.3.1 State-Dependent Regular IC	25
2.3.2 State-Dependent Z-IC	30
2.3.3 Comparison of State-Dependent Regular IC and Z-IC	34
2.4 Weak Interference Regime	35

3	State-Dependent Cognitive Interference Channel	37
3.1	Channel Model	38
3.2	The CIC-STR Model	40
3.2.1	Discrete Memoryless Channels	40
3.2.2	Gaussian Channels	43
3.3	The CIC-ST Model	46
3.3.1	Discrete Memoryless Channels	46
3.3.2	Gaussian Channels	52
4	State-Dependent Single-User Channel with a Helper	68
4.1	Channel Model	68
4.2	Achievable Scheme and Lower Bound	69
4.3	Capacity Results	72
5	State-Dependent Parallel Channel with a Common Helper	76
5.1	Channel Model	76
5.2	Model I: $K = 1$	78
5.3	Model II: $K = 2$	87
5.4	Model III: General K	93
6	State-Dependent Multiple Access Channel with a Helper	99
6.1	Channel Model	99
6.2	Outer and Inner Bounds on Capacity	100
6.3	Capacity Results	104
7	State-Dependent Broadcast Channel with a Helper	109
7.1	Channel Model	109
7.2	Scenario I: Common Message	111
7.3	Scenario II: Private Messages	115

8 Conclusion	120
A Proof for Chapter 2	122
A.1 Proof of Proposition 2.1	122
A.2 Proof of Proposition 2.2	125
A.3 Proof of Proposition 2.3	125
A.4 Proof of Proposition 2.4	130
A.5 Proof of Proposition 2.5	133
A.6 Proof of Corollary 2.2	134
A.7 Proof of Theorem 2.5	135
B Proof for Chapter 3	137
B.1 Proof of the Outer Bound (3.8a)-(3.8d)	137
B.2 Proof of the Converse for Theorem 3.2	142
B.3 Proof of Lemma 3.1	150
B.4 Proof of the Converse for Theorem 3.3	150
B.5 Proof of Lemma 3.2	154
B.6 Proof of Theorem 3.5	157
B.7 Proof of the Outer Bound for Theorem 3.6	160
B.8 Proof of the Converse for Theorem 3.7	161
C Proof for Chapter 5	163
C.1 Proof of Proposition 5.1	163
C.2 Proof of Proposition 5.2	165
C.3 Proof of Proposition 5.4	167
C.4 Proof of Proposition 5.5	168
C.5 Proof of Proposition 5.7	170
C.6 Proof of Theorem 5.5	173

D	Proof for Chapter 6	174
D.1	Proof of Proposition 6.1	174
D.2	Proof of Lemma 6.2	175
E	Proof for Chapter 7	178
E.1	Proof of Proposition 7.5	178
	References	181

LIST OF FIGURES

1.1	A practical example for the D2D communication in cellular system.	3
1.2	An illustration of the IC-ST models	4
1.3	An illustration of the CIC-STR (including the dashed line) and CIC-ST (without the dashed line) models	6
1.4	An illustration of the state-dependent single-user channel with a helper . . .	6
1.5	An illustration of the state-dependent parallel channel with a common helper	7
1.6	An illustration of the state-dependent MAC with a helper	8
1.7	An illustration of the state-dependent broadcast channel with a helper: Sce- nario with a common message	9
1.8	An illustration of the state-dependent broadcast channel with a helper: Sce- nario with private messages	9
2.1	The IC-ST model	16
2.2	Conditions on channel parameters (a, c) under which the state-dependent Gaussian regular IC and Z-IC achieve the capacity of the corresponding channel without state in the very strong regime.	24
2.3	Capacity region of the strong IC without state	26
2.4	Capacity region of the strong Z-IC without state	31
2.5	Ranges of c under which points on the sum capacity boundary of the strong regular/Z-IC without state can be achieved by the state-dependent channel .	35

3.1	The CIC-STR (including the dashed line) and the CIC-ST (without the dashed line) models	38
3.2	An illustration of the partial boundary of the capacity region for a Gaussian CIC-ST with $ a > 1$	56
3.3	An illustration of inner and outer bounds and the partial boundary of the capacity region for a Gaussian CIC-ST with $ a \leq 1$	64
4.1	The state-dependent single-user channel with a helper	68
4.2	Lower and upper bounds for the state-dependent single-user channel with a helper	74
4.3	Lower and upper bounds for the state-dependent single-user channel with a helper	75
4.4	Capacity achievable points	75
5.1	The state-dependent parallel channel with a common helper	76
5.2	The capacity region for case 1 with $P_0 = 1.5$ and $P_1 = 3$	83
5.3	Inner and outer bounds for case 2 with $P_0 < P_1$, which match partially on the boundaries	84
5.4	Inner and outer bounds for case 2 with $P_0 \geq P_1$, which match partially on the boundaries	84
5.5	Inner and outer bounds for case 3, which match partially on the boundaries.	88
5.6	Segments of the capacity boundary for the Gaussian channel of model II	90
5.7	Segment on the capacity region for the Gaussian channel of model III	97
5.8	Segment on the capacity region for the Gaussian channel of model III	97
5.9	Segment on the capacity region for the Gaussian channel of model III	98
6.1	The state-dependent MAC with a helper	99

6.2	An illustration of the capacity region for state-dependent Gaussian MAC with a helper for case 1 with $P_1 = P_2 = 3$, $P_0 = 7.5$ and arbitrary Q (characterized by Theorem 6.1) and case 2 with $P_1 = P_2 = 3$, $P_0 = 4.5$ and $Q = 8$ (characterized by Theorem 6.2).	105
6.3	A illustration of the segment of the capacity boundary for state-dependent Gaussian MAC with a helper: $P_0 < P_1 + P_2 - 1$	108
7.1	The state-dependent broadcast channel with a helper: Scenario with a common message	109
7.2	The state-dependent broadcast channel with a helper: Scenario with private messages	110

CHAPTER 1

INTRODUCTION

1.1 Motivation

New innovation of interference management is the major factor that drives evolution of cellular wireless networks from one generation to another. In second generation cellular systems, the frequency division multiplexing (FDM) and time division multiplexing (TDM) are adopted, in which multiple transmissions are orthogonalized in frequency or time to avoid interference among these transmissions. In third generation cellular system, orthogonal codes are used by simultaneous transmissions to avoid interference, which is widely known as code division multiple access (CDMA). In current fourth generation cellular networks, orthogonal frequency division multiplexing (OFDM) is implemented, which significantly improves data rate.

Despite the above new innovations, the demands for increasingly high data transmission rate continue to call for new interference management technologies for future cellular networks. Given that all up-to-date cellular wireless networks use the orthogonalization idea for handling interference, should/can new generation wireless networks employ non-orthogonalization idea so that all communication resources can be used simultaneously to improve throughputs?

A general non-orthogonal approach was proposed by Han and Kobayashi in [1] via information theoretic study of the interference channel. The idea is to split messages at transmitters so that receivers can decode part of messages intended for other receivers and remove these signals (i.e., interference) from their received outputs. Some special cases of such a scheme have been shown to be optimal (i.e., achieve the capacity region) in certain interference regime. (1) Fully decoding and canceling the interference has been shown in [2] to be optimal in the very strong interference regime. (2) Jointly decoding all messages has been shown in [3] to be optimal in the strong interference regime. (3) Treating interference as noise has been shown in [4–6] to be optimal (achieves the sum capacity) in the weak interference regime.

However, rate splitting requires user pairs to share codebooks, which substantially increases the complexity of design. In many cases, this is not possible in practice when transmissions are not within the same network domain. Furthermore, the interference can be superposition of signals to many receivers (in downlink), and it is difficult for a receiver to decode such interference. Although the special case of treating interference as noise does not require codebook sharing, it does not perform well in most scenarios.

In this thesis, we explore a new perspective/technique for interference management, which exploits the fact that an interferer knows the interference that it causes to other users noncausally and can/should exploit such information for canceling the interference. In this way, users can transmit simultaneously and the throughput of wireless networks can be substantially improved. Since the interference that is noncausally known at the transmitter is referred to as “dirty/state” corruption of the channel in information theory, we refer to the interference treated in such a way as “dirty interference” or “noncausal state”.

In the following, we use a practical example (see Fig. 1.1) to further illustrate our idea. Consider a cellular network that incorporates device-to-device (D2D) communications. It is typical that the cellular base station causes interference to D2D transmissions. In fact, the base station itself knows such interference noncausally, because the interference is the sig-

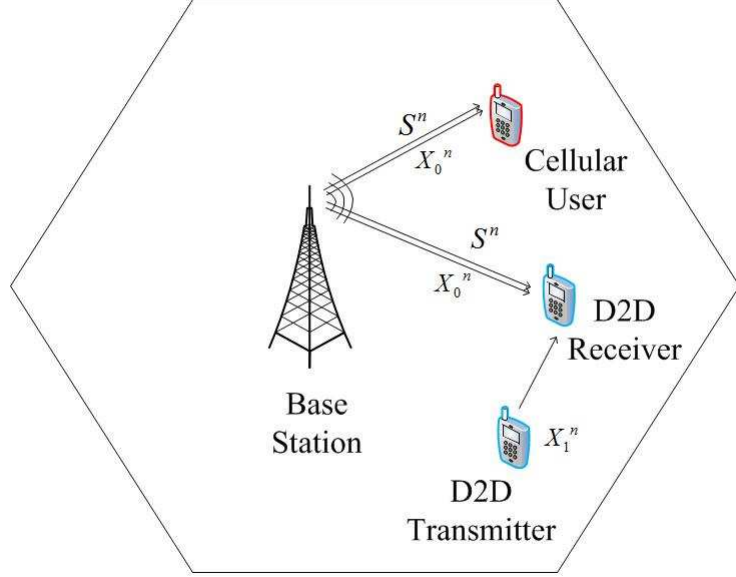


Fig. 1.1: A practical example for the D2D communication in cellular system.

nal that the base station sends to cellular receivers. Thus, the interference can be viewed as the noncausal state sequence (denoted as S^n in Fig. 1.1). The base station is then able to exploit such information about the interference (i.e., state) and send a help signal (denoted by X_0 in Fig. 1.1) to assist D2D users to cancel the interference. Since the help signal X_0 may also cause interference to the cellular receiver, simply reversing the state complete ruins the cellular communication. Therefore, a more sophisticated scheme should be designed to deal with the interference. More specifically, this thesis designs adapted dirty paper coding schemes for various of state-dependent models, in which the state information is precoded into help signal. The receiver then cancel the state interference with the assistance of the help signal. The throughput of the wireless networks can hence be significantly improved, compared with the orthogonalized transmission.

1.2 Channel Models

Towards designing a dirty interference cancelation framework for wireless networks, this thesis explores two classes of state-dependent interference networks and the goal is to de-

velop dirty interference cancellation schemes that achieve the fundamental communication limits. One class of models (referred to as state-dependent interference channels) capture the scenarios that users help each other to cancel dirty interference. The other class of models (referred to as state-dependent channels with helper) capture the scenarios that one dominate user interferes a number of other users and assists these users to cancel its dirty interference. We next introduce the models that we study in detail.

For the class of state-dependent interference channels, we study two models, i.e., the state-dependent interference channel with state known at both transmitters (IC-ST) and the cognitive interference channel with state known at the cognition transmitter (CIC-ST). In these models, transmitters interfere with each other, and the state models the additional interference due to the fact that the transmitters also send signals to other receivers (not included in the model) in broadcast scenarios.

For the IC-ST (see Fig. 1.2), two transmitters send two messages to two receivers, respectively, via an interference channel. The channel is corrupted by an independent and identically distributed (i.i.d.) state sequence, which is assumed to be known *noncausally* at both transmitters. One example scenario for this channel models is as follows. In cellular networks, two base stations communicate with two users which are near the edge between two adjacent cells. The state captures the signal that the base stations transmit to other users (not included in the model) in the network. For this model, we consider both the state-dependent regular IC and the state-dependent Z-IC.

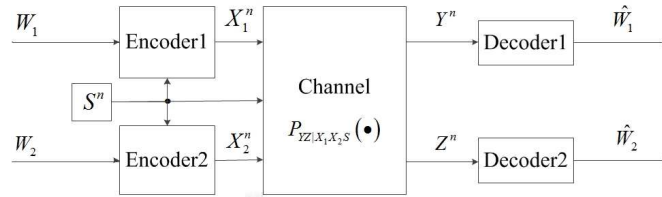


Fig. 1.2: An illustration of the IC-ST models

The second model we study is the CIC-ST (see Fig. 1.3), in which a primary trans-

mitter sends a message to two receivers (receivers 1 and 2) with assistance of a cognitive transmitter, and the cognitive transmitter also sends a separate message to receiver 2. The channel is corrupted by an i.i.d. state sequence. The state sequence is noncausally known at the cognitive transmitter. This model is well motivated in practical networks. For example, it is often the case in cognitive radio networks that a primary transmitter wishes to send a common message to a number of primary receivers, and a cognitive transmitter (which often knows the primary transmitter's message via its necessary coordination with the primary transmitter) can cooperatively send the common message to the primary receivers. This cognitive transmitter may also have its own message intended to one of the primary receivers. At the same time, the cognitive transmitter can communicate to some secondary receivers simultaneously, and its signals to these receivers then interfere with the primary receivers. Such a signal is clearly known by the cognitive transmitter noncausally, and is captured by the state in the model. A similar scenario can also occur in cellular networks. For example, two base stations may cooperatively send certain common information to many receivers which are near the edge between the two cells that the two base stations serve. In addition, one of the base stations may transmit additional information to receivers in its own cell.

For the CIC-ST model, we investigate two scenarios. The first scenario assumes that the state sequence is noncausally known at both the cognitive transmitter and receiver 2, and is referred to as the CIC-STR (which stands for the cognitive interference channel with state information noncausally known at both the cognitive transmitter and receiver 2). The second scenario assumes that the state sequence is noncausally known only at the cognitive transmitter, and is referred to as the CIC-ST (which stands for the cognitive interference channel with state information noncausally known at only the cognitive transmitter).

The second class of models we study are state-dependent channels with an additional helper, for which we study four models, i.e., the state-dependent single-user channel with a helper, the state-dependent parallel networks with a common state-cognitive helper, the

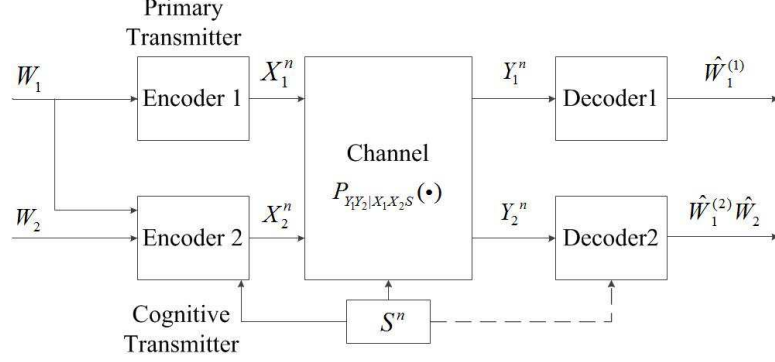


Fig. 1.3: An illustration of the CIC-STR (including the dashed line) and CIC-ST (without the dashed line) models

state-dependent multiple access channel (MAC) with a helper, and the state-dependent broadcast channel with a helper.

In the state-dependent single-user channel with a helper (see Fig. 1.4), the transmitter communicates with the receiver via a state-corrupted channel. The state is not known to the transmitter, but to a helper noncausally, which wishes to assist the receiver to cancel the state.

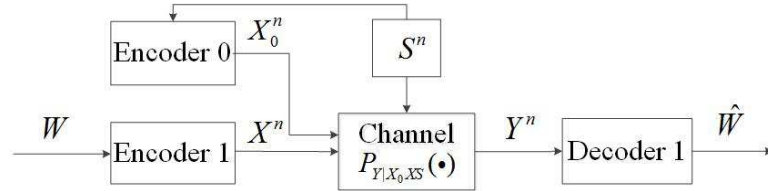


Fig. 1.4: An illustration of the state-dependent single-user channel with a helper

In the state-dependent parallel networks with a common state-cognitive helper (see Fig. 1.5), K transmitters wish to send K messages respectively to K receivers over K parallel channels, and the receivers are corrupted by states. The channel state is known to neither the transmitters nor the receivers, but to a helper noncausally. The helper hence assists these transmitter-receiver pairs to cancel state interference. Furthermore, the helper also has its own message to be sent simultaneously to its corresponding receiver. Since the state information is known only to the helper, but not to the corresponding transmitters, transmitter-side state cognition and receiver-side state interference are mismatched. The

practical motivation of such a channel can be referred to Fig. 1.1 with the understanding that the helper models the base station and the multiple parallel channels model multiple D2D transmissions.

More specifically, we study three (sub) models of the state-dependent parallel networks with a common helper. Model I serves as a basic model, which consists of only one state-corrupted receiver ($K = 1$) and a helper that assists this receiver to cancel state interference in addition to transmitting its own message. Our study of this model provides necessary techniques to deal with state in the mismatched context for studying more complicated models II and III. In fact, this model can be viewed as the state-dependent Z-interference channel, in which the interference is only at receiver 1 caused by the helper. In contrast to the state-dependent Z-interference channel studied previously in [7], which assumes that state interference at both receivers are known to both (corresponding) transmitters, our model assumes that state interference is known noncausally only to the helper, not to the corresponding transmitter 1. Model II consists of two transmitter-receiver pairs in addition to the helper, and only one receiver is interfered by a state sequence. Model III consists of a common helper assists multiple transmitter-receiver pairs with each receiver corrupted by an independently distributed state sequence.

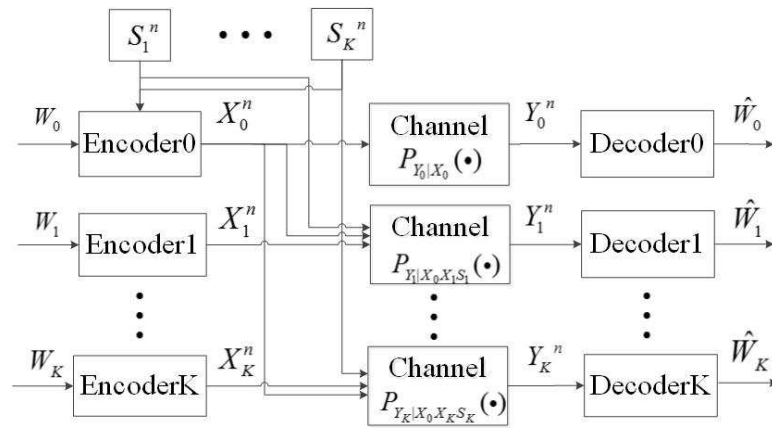


Fig. 1.5: An illustration of the state-dependent parallel channel with a common helper

In the state-dependent MAC with a helper (see Fig. 1.6), transmitter 1 and transmitter 2 send their own message to the receiver, respectively. The channel is corrupted by a state

sequence. The state sequence is known to neither the transmitters nor the receiver, but is known to a helper noncausally. Hence, the helper assists the receiver to cancel the state interference. A practical example for this model could be the multiple-access communications in a picocell located inside a macrocell of a cellular network. The macrocell user serves as a helper to assist the communications in the picocell to cancel the interference.

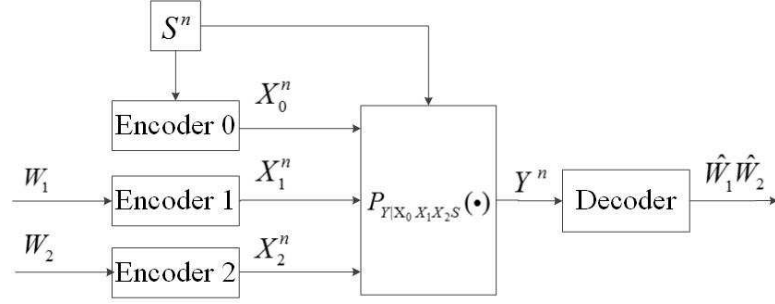


Fig. 1.6: An illustration of the state-dependent MAC with a helper

For the state-dependent broadcast channel with a helper, we studied two scenarios. In scenario I (see Fig. 1.7), a transmitter sends one common message to two receivers over the broadcast channel, which is interfered by a state sequence. The state sequence is known at neither the transmitter nor the receivers. A helper which knows the state sequence noncausally assists both receivers to deal with the channel state. Scenario II (see Fig. 1.8) is similar to scenario I with the difference being that the transmitter sends two independent messages to receivers 1 and 2, respectively. This model naturally arises in many practical scenarios, for example, downlink cellular communications. Consider two adjacent cells in a cellular network. It is likely that downlink transmission signals from one base station causes large interference to users in its adjacent cell. However, the base station can serve as a helper at the same time, and assists the downlink transmission in its adjacent cell to cancel its interference.

It is clear that in the first class of state-dependent interference channels, the state information is known at one or both transmitters and can hence be exploited for encoding messages. Hence, the focus is to design schemes that best exploit state information for encoding messages. However, for the second class of state-dependent channels with a helper,

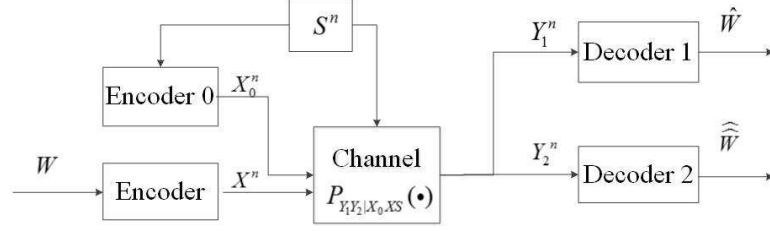


Fig. 1.7: An illustration of the state-dependent broadcast channel with a helper: Scenario with a common message

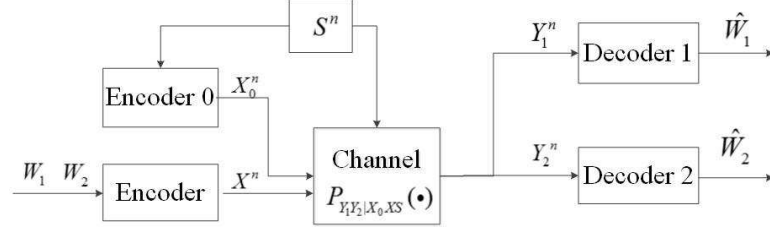


Fig. 1.8: An illustration of the state-dependent broadcast channel with a helper: Scenario with private messages

the state is known at only a helper which does not know messages. The key issue is to design encoding and state cancellation in a distributed manner to achieve the best overall performance. We are also interested in exploring whether distributed scheme can achieve the performance of the channel without the state corruption.

1.3 Related Work

Initiated by Shannon in [8], the channel with state corruption has been intensively studied for the past a few decades. Motivated by practical interests of modeling interference as state, our focus is on the cases in which the state is noncausally known at transmitters. In [9], the single-user channel with state known noncausally at the transmitter is studied, and the capacity is obtained for the discrete memoryless channel via Gel'fand-Pinsker binning. Based on this result, in [10], the capacity for the state-dependent single-user Gaussian channel is obtained, and it is shown that the state can be perfectly canceled as if there is no state interference. The achievable scheme is referred to as “dirty paper coding”.

Following similar schemes, various state-dependent network models are studied, and it

has been shown that the state interference can be perfectly or partially canceled at receivers. For example, the state-dependent broadcast channel has been studied in, e.g., [11–15], in which the transmitter knows the state noncausally and can exploit such information to select the codeword to be sent in the channel. In [16], the state-dependent relay channel is studied, in which the source node knows the state and can use such information for encoding. In [11, 17], the MAC with the receiver being corrupted by one state variable is studied. In such a model, both transmitters are assumed to know the state sequence noncausally, and can use the state information to independently encode their own messages. Similarly, in [18], the state-dependent cognitive MAC is studied, in which one transmitter knows both messages as well as the state, and can hence use state information to encode both messages.

More closely to our work, a few interference channel models with state noncausally known at transmitters have been studied. In [19] and [20], the interference channel with state known at both transmitters is studied. Various achievable schemes have been designed and the corresponding achievable regions are compared in [20]. The gap between inner and outer bounds on the capacity region has been characterized within certain finite bits in [19]. In [21], the interference channel is corrupted by two independent states, each interfering one receiver. The states are available at their corresponding transmitters. The capacity region is obtained for the strong interference regime with the state power going to infinity.

The Gaussian state-dependent IC model we study is the same as that studied in [19] and [20]. However, differently from [19, 20], our focus here is to characterize the exact capacity region, or points on the boundary of the capacity region. We note that the capacity region/the sum capacity has been characterized for the Gaussian interference channel *without* state in the following three regimes: (1) very strong interference channels [2]; (2) strong interference channels [3]; and (3) a certain weak interference channel [4–6] (based on the technique developed in [22]). And for the state-dependent Z-IC model, capacity/sum capacity has been characterized for the corresponding channel without state for the (1) very

strong Z-IC [3, 23]; strong Z-IC [3, 23]; and weak Z-IC [23]. We study whether or not the capacity region/the sum capacity in these regimes are achievable when the two receivers' outputs are also corrupted by differently scaled state, and if so, what transmission schemes are capacity achieving.

In [24] and [25], a model of the cognitive interference channel with state was studied, in which both transmitters (i.e., the primary and cognitive transmitters) jointly send one message to receiver 1, and the cognitive transmitter sends an additional message separately to receiver 2. The i.i.d. state sequence is noncausally known at the cognitive transmitter only. Inner and outer bounds on the capacity region were provided. The difference of our CIC-ST model from the model studied in [24] and [25] lies in that the common message jointly sent by both transmitters needs to be decoded at both receivers instead of only at receiver 1 as in [24] and [25]. Although the two models appear similar to each other, their capacity regions can have different forms, and the transmission schemes achieving these regions can also be different. This fact is already demonstrated by the two corresponding models without state studied respectively in [26–32] and [33]. The capacity bounds in [26–32] and the capacity region in [33] have different forms, and are achieved by different achievable schemes. Therefore, our study can lead to new information theoretic insights.

A common nature that the above models share is that the users are at the same level, thus, for each message to be transmitted, at least one transmitter in the system knows both the message and the state, and can incorporate the state information in encoding of the message so that state interference at the corresponding receiver can be cancelled. However, in practice, it is often the case that transmitters that have messages intended for receivers do not know the state, whereas some third-party nodes know the state, but do not know the message. In such a mismatched case, a dominant user will help all the interfered users to cancel state, though state information cannot be exploited in encoding of messages. A number of previously studied models capture such mismatched property. For example, in [34], a transmitter sends a message to a state-dependent receiver, and a helper knows

the state noncausally and can help the transmission. Lattice coding is designed in [34] for the helper to assist state cancelation at the receiver, and is shown to be optimal under certain channel conditions. In [35, 36], the state-dependent relay channel is studied, and the case with the state noncausally known only at the relay is the mismatched scenario. Furthermore, in [37], the state-dependent MAC channel is studied with the state known at only one transmitter. In such a case, the other transmitter's message cannot be encoded with the information of the state. In [38–40], the MAC is corrupted by two states that are respectively known at the two transmitters. In such a case, neither message can be encoded with the full information of the state. In our study, we also focus on the mismatched scenarios as discribed above. Howecer, we are interested in the following issues that are not captured in the previously studied models: (1) what is the optimal way for the helper to assist the interfered receiver (2) when there are multiple state-dependent transmitter-receiver links, how should the helper trade off among helping multiple state-interfered receivers; (3) when the helper has its own message intended for a separate receiver (not state-dependent), how should the helper trade off between sending its own message and assisting state-dependent receivers; and (4) under what channel conditions, the above two tradeoffs are optimal (i.e., achieve the boundary of the capacity region).

1.4 Summary of Contributions and Thesis Organization

As a summary, this thesis leads to one journal publication [41], two journal submissions [42] and [43], and eight conference publications [44–51]. In the following, I briefly summarize the contributions of my thesis.

In Chapter 2, we study the state-dependent IC/Z-IC. More specifically, in the very strong interference regime, we characterize the conditions on the channel parameters, under which the capacity region of the IC and Z-IC channels *without* state can be achieved by the

corresponding *state-dependent* IC/Z-IC. The capacity of the state-dependent IC/Z-IC are thus characterized under those cases. In the strong interference regime, we characterize the conditions on the channel parameters, under which points on the capacity region boundary of the channel *without* state can be achieved. Hence, these points also lie on the capacity region boundary of the state-dependent channel. For the weak interference regime, we obtain the sum capacity. We also compare the state-dependent regular IC with Z-IC, and provide a few insights.

In Chapter 3, we study the CIC-ST(R). We first study the CIC-STR. For this scenario, we obtain the capacity region for both the discrete memoryless and Gaussian channels. We further study the CIC-ST. For this scenario, we obtain the inner and outer bounds on the capacity region for the discrete memoryless channel and its degraded version. Then we characterize the capacity region for the degraded semideterministic channel and for channels that satisfy a less noisy condition. For the Gaussian channels, we partition the channel into two cases based on how the interference compares with the signal at receiver 1. For each case, we derive the inner and outer bounds on the capacity region, and characterize the partial boundaries of the capacity region. We also characterize the full capacity region for channels that satisfy certain conditions. We further show that certain Gaussian channels achieve the capacity of the same channels with state noncausally known at both the cognitive transmitter and receiver 2.

In Chapter 4, we study the state-dependent single-user channel with a helper. In the previous work [34], the capacity in the regime of infinite state power is characterized based on Lattice coding. In this thesis, we consider the general case with finite state power. We first derive the achievable scheme combining two methods to cancel state: 1. precoding the state with a single bin scheme; 2. directly reversing the state. By comparing the lower bound derived and the upper bound from the previous work, we characterize the capacity rate for channel under various channel parameters.

In Chapter 5, we study three models of parallel communication networks with a state-

cognitive helper. For each model, there is unique challenge to design capacity-achieving schemes for the helper to trade off among multiple functions. For model I, we design an adapted dirty paper coding together with superposition coding for the helper to trade off between assisting to cancel the state and transmitting its own message. We showed that such a scheme achieves the full capacity region or segments on the capacity region boundary for all channel parameters. For model II, we design a multi-layer scheme, such that the helper assists receiver 1 to cancel the infinite-power state while simultaneously eliminating its interference to receiver 2. Such a scheme achieves two segments on the capacity region. Over one segment, the helper is capable to fully cancel the interference that it causes to receiver 2, and simultaneously assists receiver 1 to achieve a certain positive rate. In the second segment, the sum capacity is obtained with the helper dedicated to help receiver 1. For model III, we employ a time-sharing scheme such that the helper alternatively assists each receiver, and we show that such a scheme achieves the sum capacity for certain channel parameters.

In Chapter 6, we study the state-dependent MAC with a helper. We first derive an outer bound on the capacity region, and then obtain an inner bound based on a dirty interference cancelation scheme. By comparing the inner and outer bounds, we characterize the full capacity region or segment on the boundary of the capacity region under various channel parameters.

In Chapter 7, we study the state-dependent broadcast channel with a helper. In scenario I, the transmitter sends one message to both receivers, and in scenario II, the transmitter sends two private messages respectively to two receivers. We derive inner and outer bounds for both scenarios. By comparing the inner and outer bounds, we characterize capacity/capacity region under various ranges of channel parameters.

In Chapter 8, we summarize the above results with some insights, and discuss about possible future works.

CHAPTER 2

STATE-DEPENDENT INTERFERENCE

CHANNEL

In this chapter, we study the state-dependent regular IC and the state-dependent Z-IC. We consider three regimes for each channel model, i.e. the very strong, strong and weak regime. For both very strong state-dependent regular IC and Z-IC, we characterize the capacity region and the conditions under which capacity region is obtained. For the strong (but not very strong) state-dependent regular IC and Z-IC, we characterize the points on the capacity boundary. For the weak state-dependent regular IC and Z-IC, we obtain the sum capacity. And for each regime, we make comparison between the result for regular IC and Z-IC, and reveal whether Z-IC has advantage over regular IC in cancelling state interference.

2.1 Channel Model

In the state-dependent IC and the state-dependent Z-IC (see Fig. 1.2 in Section 1.2. For convenience of reference, we include the figure again as Fig. 2.1 in this section), transmitter 1 sends a message W_1 to receiver 1, and transmitter 2 sends a message W_2 to receiver 2.

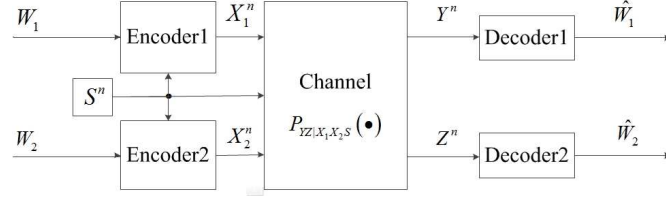


Fig. 2.1: The IC-ST model

The channel is corrupted by an i.i.d. state sequence S^n , which is assumed to be known *noncausally* at both transmitters. More specifically, encoder 1 $f_1 : \{\mathcal{W}_1, \mathcal{S}^n\} \rightarrow \mathcal{X}_1^n$ at transmitter 1 maps a message $w_1 \in \{1, \dots, 2^{nR_1}\}$ and a state sequence $s^n \in \mathcal{S}^n$ to an input x_1^n , and encoder 2 $f_2 : \{\mathcal{W}_2, \mathcal{S}^n\} \rightarrow \mathcal{X}_2^n$ at transmitter 2 maps a message $w_2 \in \{1, \dots, 2^{nR_2}\}$ and the state sequence s^n to an input x_2^n . For the state-dependent regular IC, these two inputs are sent over the memoryless interference channel characterized by $P_{YZ|X_1X_2S}$, and for the Z-IC, only receiver 1 is interfered by transmitter 2's signal, while receiver 2 is free from interference. Hence, the channel is characterized by $P_{Y_1|X_1X_2S}$ and $P_{Y_2|X_2S}$. Decoder 1 $g_1 : \mathcal{Y}_1^n \rightarrow \mathcal{W}_1$ at receiver 1 decodes W_1 and decoder 2 $g_2 : \mathcal{Y}_2^n \rightarrow \mathcal{W}_2$ at receiver 2 is required to decode W_2 , with the probability of error approaching zero as the codeword length n goes to infinity. The capacity region is defined to be the closure of the set of all achievable rate pairs (R_1, R_2) .

We study the Gaussian channel with the outputs at receivers 1 and 2 for one channel use given by

$$\begin{aligned} Y &= X_1 + aX_2 + S + N_1 \\ Z &= bX_1 + X_2 + cS + N_2 \end{aligned} \tag{2.1}$$

where a, b and c are constants, the noise variables $N_1, N_2 \sim \mathcal{N}(0, 1)$, and $S \sim \mathcal{N}(0, Q)$. Both the noise variables and the state variable are i.i.d. over channel uses. The channel inputs X_1 and X_2 are subject to the average power constraints P_1 and P_2 . For the Z-IC, the

channel parameter $b = 0$, and thus receiver 2 is not interfered by transmitter 1.

2.2 Very Strong Interference Regime

In this section, we study the state-dependent regular IC and Z-IC in the very strong regime, and characterize the conditions under which the capacity region can be obtained, i.e., the capacity region of the IC *without* state can be achieved. We also compare the results of the two channels.

2.2.1 State-Dependent Regular IC

In this subsection, we study the state-dependent regular IC in the very strong regime, in which the channel parameters satisfy

$$\begin{aligned} P_1 + a^2 P_2 + 1 &\geq (1 + P_1)(1 + P_2) \\ b^2 P_1 + P_2 + 1 &\geq (1 + P_1)(1 + P_2). \end{aligned} \quad (2.2)$$

In such a regime, the channel *without state* is the very strong IC, and its capacity region has been characterized in [2], which contains rate pairs (R_1, R_2) satisfying

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (2.3a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2). \quad (2.3b)$$

In this case, the two users achieve the single-user channel capacity even with interference. Our focus here is to study under what conditions on the channel parameters we can design schemes for the *state-dependent* IC to achieve the above capacity region, i.e., the state at receivers can be fully cancelled. Clearly, in this case, the above capacity region also serves as the capacity region for the state-dependent channel.

There are two challenges here. (1) Since the state are scaled differently at two receivers,

each transmitter needs to deal with the compound state corruption in two receivers. (2) The scheme to achieve the capacity region for the very strong IC without state suggests that the receivers decode the interference first, and then cancel it from the received output so that decoding of the intended input does not experience interference. For the state-dependent channel, if both transmitters employ dirty paper coding, receivers decode only auxiliary random variables, but not the exact input of the other transmitter. Hence, canceling the signal interference would cause certain left-over state interference.

In the following, we design a coding scheme to achieve the single-user channel capacity for each user based on *cooperative* dirty paper coding between the two transmitters such that (1) the two transmitters cooperatively cancel the compound states at the two receivers, and furthermore (2) each transmitter design its dirty paper input based on the original state plus the left-over interference by decoding the other transmitter's dirty paper coded interference. The cooperation between the transmitters is possible due to the state information known to both transmitters.

We first design an achievable scheme for the discrete memoryless channel, which is useful for the Gaussian channel. The two transmitters encode their messages W_1 and W_2 into two auxiliary random variables U and V , respectively, based on Gel'fand-Pinsker binning scheme [9]. Since the channel satisfies the very strong interference condition, each receiver first decodes the auxiliary random variable corresponding to the message intended for the other receiver, and then decodes its own message by decoding the auxiliary random variable for itself. For instance, receiver 1 first decodes V , then uses it to cancel the message interference and state interference, and finally decodes its message by decoding U . Such an achievable scheme yields the following achievable region.

Proposition 2.1. *For the state-dependent IC with state noncausally known at both trans-*

mitters, the achievable region consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \min\{I(U; Y_1V), I(U; Y_2)\} - I(U; S) \quad (2.4a)$$

$$R_2 \leq \min\{I(V; Y_2U), I(V; Y_1)\} - I(V; S) \quad (2.4b)$$

for some distribution $P_{SUVX_1X_2Y_1Y_2} = P_S P_{U|S} P_{X_1|US} P_{V|S} P_{X_2|VS} P_{Y_1Y_2|X_1X_2S}$, where U and V are auxiliary random variables.

Proof. See Appendix A.1. □

By choosing joint Gaussian distributions for the auxiliary random variables and the channel inputs in the achievable region given in Proposition 2.1, we can obtain an achievable region for the Gaussian channel. In particular, U is designed to deal with the state interference for Y_1 after cancelling V , and V is designed to deal with the state interference for Y_2 after cancelling U . Therefore, coefficients in dirty paper coding of U and V are jointly designed to cancel the states at the two receivers. Furthermore, by requiring $I(U; Y_1V) \geq I(U; Y_1)$ and $I(V; Y_1U) \geq I(V; Y_1)$ in (2.4a) and (2.4b), the resulting region is the same as the capacity region of the channel without state, and thus the capacity region of the state-dependent IC is established. We state this result in the following theorem.

Theorem 2.1. *Consider the state-dependent Gaussian IC with state noncausally known at both transmitters. If the channel parameters (a, b, c, P_1, P_2, Q) satisfy the following conditions:*

$$\frac{(b^2P_1 + P_2 + c^2Q + 1)}{(1 + P_2)\left(1 + \frac{(1+P_2)(c+cP_1-bP_1)^2Q+Q P_1(1+P_2-acP_2)^2}{((1+P_1)(1+P_2)-abP_1P_2)^2}\right)} \geq 1 + P_1 \quad (2.5a)$$

$$\frac{(P_1 + a^2P_2 + Q + 1)}{(1 + P_1)\left(1 + \frac{P_2(c+cP_1-bP_1)^2Q+Q(1+P_1)(1+P_2-acP_2)^2}{((1+P_1)(1+P_2)-abP_1P_2)^2}\right)} \geq 1 + P_2, \quad (2.5b)$$

then the capacity region consists of rate pairs (R_1, R_2) satisfying (2.3a) and (2.3b), i.e., is the same as the single-user capacity for both receivers.

Proof. In Proposition 2.1, we set U and V as $U = X_1 + \alpha S$, $V = X_2 + \beta S$, where X_1, X_2 and S are independent Gaussian variables with mean zero and variances P_1, P_2 and S , respectively. We then design α based on dirty paper coding for $Y'_1 = Y_1 - aV = X_1 + (1 - a\beta)S + N_1$, and design β based on dirty paper coding for $Y'_1 = Y_1 - bU = X_2 + (c - b\alpha)S + N_2$. We further require α and β to satisfy the following conditions:

$$\frac{\alpha}{1 - a\beta} = \frac{P_1}{P_1 + 1} \quad (2.6a)$$

$$\frac{\beta}{c - b\alpha} = \frac{P_2}{P_2 + 1}. \quad (2.6b)$$

By solving the equations (2.6a) and (2.6b), we have

$$\alpha = \frac{P_1(1 + P_2 - acP_2)}{(1 + P_1)(1 + P_2) - abP_1P_2}$$

$$\beta = \frac{cP_2(C(1 + P_1) - bP_1)}{(1 + P_1)(1 + P_2) - abP_1P_2}.$$

Then the bounds in equations (2.4a) and (2.4b) becomes

$$R_1 \leq \frac{1}{2} \log(1 + P_1)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2), \quad (2.7)$$

if

$$\frac{1}{2} \log(1 + P_1) \leq I(U; Y_2) - I(U; S)$$

$$\frac{1}{2} \log(1 + P_2) \leq I(V; Y_1) - I(V; S). \quad (2.8)$$

By computing the mutual information terms in the above equations based on the chosen distributions for U and V , we obtain the conditions given in the theorem. Such an achievable region is therefore the capacity region, because it is the same as the corresponding

channel without state. This can be formally shown by following steps similar to those in Appendix A.7. \square

Although the conditions (2.5a) and (2.5b) are expressed in complicated forms, they can be easily checked numerically. We provide numerical illustration in Section 2.2.3. Following Theorem 2.1, we also obtain the following result for the state-dependent symmetric Gaussian IC as a special case.

Corollary 2.1. *For the state-dependent symmetric Gaussian IC with state noncausally known at the transmitters, i.e., $a = b$, $c = 1$, and $P_1 = P_2$, the capacity region contains rate pairs (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log(1 + P) \\ R_2 &\leq \frac{1}{2} \log(1 + P), \end{aligned} \tag{2.9}$$

if $a \geq a_{th}$, where a_{th} solves the following equation

$$\frac{(P + a^2P + Q + 1)(1 + P + aP)^2}{(1 + P)[(1 + P + aP)^2 + Q(1 + 2P)]} = 1 + P. \tag{2.10}$$

Proof. If $a = b$ and $c = 1$, then the conditions (2.5b) and (2.5a) reduce to the following single condition:

$$\frac{(P + a^2P + Q + 1)(1 + P + aP)^2}{(1 + P)[(1 + P + aP)^2 + Q(1 + 2P)]} \geq 1 + P. \tag{2.11}$$

Such a condition is equivalent to the one given in the corollary. \square

2.2.2 State-Dependent Z-IC

In this subsection, we study the state-dependent Z-IC, i.e., $b = 0$, in the very strong regime, in which the channel parameter satisfies

$$a^2 > 1 + P_1. \quad (2.12)$$

Under the above condition, the channel *without* state is very strong, and its capacity region contains rate pairs (R_1, R_2) satisfying (2.3a) and (2.3b), i.e., the two users achieve the single-user channel capacity.

Similarly to the regular IC, we also design cooperative dirty paper coding between the two transmitters, which encodes the messages W_1 and W_2 into two auxiliary random variables U and V , respectively. The difference from the scheme for the regular IC lies in the fact that since receiver 2 is interference free, V can be designed to fully cancel the state at receiver 2. Then receiver 1 first decodes the auxiliary random variable V to cancel the interference as well as partial state, and then decodes its own message and cancels the remaining state by decoding the auxiliary random variable U . Based on this achievable scheme, we have the following achievable region for the discrete memoryless channel.

Proposition 2.2. *For the state-dependent Z-IC with state noncausally known at both transmitters, the achievable region consists of rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned} R_1 &\leq I(U; VY_1) - I(U; S), \\ R_2 &\leq I(V; Y_2) - I(V; S) \end{aligned} \quad (2.13)$$

for some distribution $P_S P_{U|S} P_{V|S} P_{X_1|US} P_{X_2|VS} P_{Y_1|X_1X_2S} P_{Y_2|X_2S}$ that satisfies $I(V; Y_2) \leq I(V; Y_1)$.

Proof. See Appendix A.2. □

By choosing the joint Gaussian distribution for the auxiliary random variables and the channel inputs in the achievable region in Proposition 2.2, we obtain the achievable region for the state-dependent Gaussian Z-IC. In particular, the auxiliary random variable V is designed to deal with the state interference for Y_2 , but U is designed to deal with the state interference for Y_1 after cancelling V . By further comparing this achievable region with the capacity region of the Z-IC *without* state, we obtain the following capacity result.

Theorem 2.2. *For the state-dependent Gaussian Z-IC with state noncausally known at both transmitters, if its channel parameters (a, c, P_1, P_2, Q) satisfy the following conditions:*

$$\frac{P_2(a^2 P_2 + P_1 + 1)}{P_2 Q(1 - \alpha)^2 + (P_2 + \alpha^2 Q)(P_1 + 1)} \geq 1 + P_2,$$

where $\alpha = \frac{P_2}{P_2+1}c$, then the capacity region consists of rate pairs (R_1, R_2) satisfying (2.3a) and (2.3b).

Proof. We set U and V in Proposition 2.2 as $U = X_1 + \beta S$, $V = X_2 + \alpha S$, where X_1, X_2 and S are independent Gaussian variables with mean zero and variances P_1, P_2 and Q , respectively, and set α and β to be

$$\alpha = \frac{P_2}{(1 + P_2)}b, \quad \beta = \frac{P_1}{1 + P_1}(1 - \alpha).$$

Substituting the above choice of the Gaussian distribution into Proposition 2.2 yields the desired region and the condition in Theorem 2.2.

Since such an achievable region is the same as the capacity region of the corresponding channel without state, it can be shown to be the capacity region of the state-dependent channel. □

2.2.3 Comparison of State-Dependent Regular IC and Z-IC

In this subsection, we compare the result in Theorem 2.1 for the state-dependent *regular* IC and the result in Theorem 2.2 for the state-dependent *Z-IC*.

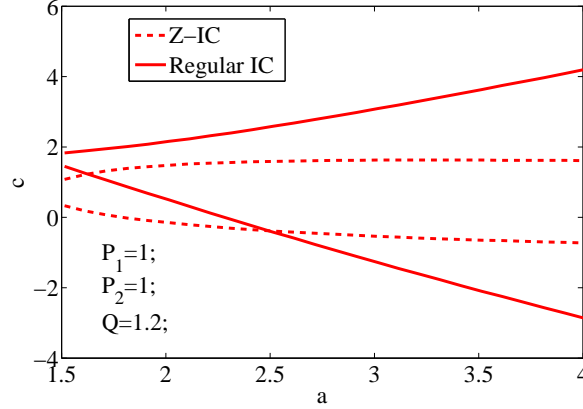


Fig. 2.2: Conditions on channel parameters (a, c) under which the state-dependent Gaussian regular IC and Z-IC achieve the capacity of the corresponding channel without state in the very strong regime.

We set $P_1 = 1$, $P_2 = 1$, and $Q = 1.2$ for both channels, and set the additional interference link in the regular IC to have the channel gain $b = 4$ such that it does not affect a fair comparison. In Fig. 2.2, we plot the range of parameter pairs (a, c) under which the two single-user channel capacities can be achieved for both state-dependent *regular IC* and *Z-IC*. The ranges between the two solid lines and between the two dashed lines respectively correspond to the regular IC and Z-IC. It is clear that the regular IC has a larger range than the Z-IC particularly for large a . Such observation suggests that it is easier to fully cancel the state for the regular IC than the Z-IC, which may appear counter intuitive, since the state-dependent Z-IC possesses an interference free link. In fact, it is reasonable, because receiver 2 in the regular IC can decode the dirty paper coded signal of transmitter 1 due to the very strong interference, via which it can cancel certain amount of state. In this way, the one more interference link to receiver 2 in the regular IC helps receiver 2 to cancel the state.

2.3 Strong Interference Regime

Since the very strong IC is studied separately in Section 2.2, in this section, we study the state-dependent regular IC and Z-IC in the strong, but not very strong regime, and characterize the conditions under which points on the capacity region boundary can be obtained. We then compare the results for the regular IC and Z-IC.

2.3.1 State-Dependent Regular IC

In this subsection, we study the state-dependent regular IC in the strong but not very strong regime, in which the channel parameters satisfy

$$\begin{aligned} a \geq 1, \quad b \geq 1, \\ \min\{P_1 + a^2 P_2 + 1, b^2 P_1 + P_2 + 1\} \leq (1 + P_1)(1 + P_2). \end{aligned} \quad (2.14)$$

Without loss of generality, we assume that $P_1 + a^2 P_2 + 1 \leq b^2 P_1 + P_2 + 1$. Under the above conditions, the IC *without* state is strong, and the capacity region was characterized in [3], which contains rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log(1 + P_1), \quad R_2 \leq \frac{1}{2} \log(1 + P_2), \\ R_1 + R_2 &\leq \frac{1}{2} \log(1 + P_1 + a^2 P_2). \end{aligned} \quad (2.15)$$

The above capacity is achieved by requiring both receivers to decode both messages, and hence the capacity region is the intersection of the capacity regions of two multiple-access channels. We illustrate such a capacity region in Fig. 2.3, as the pentagon O-A-B-E-F-O.

Our goal here is to study whether points on the boundary of such a pentagon (i.e., the capacity region boundary of the IC *without* state) can be achieved by the corresponding *state-dependent* IC. The main difference of the strong regime from the very strong regime

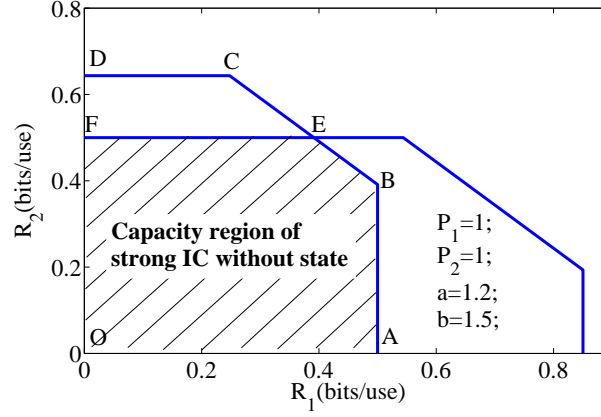


Fig. 2.3: Capacity region of the strong IC without state

studied in Section 2.2 is the additional sum rate constraint in the capacity region. Although the cooperative dirty paper coding scheme that we design for the very strong regime fully cancels the state in the single-user rate bounds, it does not fully cancel the state in the sum rate bound. Thus, new schemes need to be designed here in order for the state-dependent IC to achieve the sum rate boundary of the capacity region of the IC without state, i.e., the line B-E in Fig. 2.3. Then the points on the line A-B and the line F-E are achievable if the two corner points B and E on the sum rate boundary are achievable.

The idea of our achievable scheme is to exploit the fact that the sum rate boundary B-E is due to the decoding requirement at receiver 1 (as a receiver of the MAC), and hence every point on B-E can be achieved by message splitting and successive cancellation. For the state-dependent channel, in addition to rate splitting, we also utilize *layered dirty paper coding* and *successive state cancelation* to fully cancel the state at receiver 1. If such a coding scheme does not introduce extra bounds for receiver 2 to decode the two messages, then the sum rate boundary can be achieved.

Based on the above idea, we first design an achievable scheme for the corresponding discrete memoryless channel which is useful for studying the Gaussian channel. We split the message W_1 into two parts W_{11} and W_{12} , which are encoded into the auxiliary random variables U_1 and U_2 successively using Gel'fand-Pinsker binning. We also split the message

W_2 into two parts W_{21} and W_{22} , which are encoded into the auxiliary random variables V_1 and V_2 successively using Gel'fand-Pinsker binning scheme. Both receivers decode both messages with reasonable decoding orders, such that the decoding capability of the two receivers are accommodated. As an illustration, we next adopt the decoding order $W_{11}, W_{21}, W_{22}, W_{12}$ at receiver 1 and the decoding order $W_{21}, W_{11}, W_{12}, W_{22}$ at receiver 2. The resulting achievable rate region is given in the following Proposition.

Proposition 2.3. *For the state-dependent IC with state noncausally known at both transmitters, an achievable region consists of rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned}
R_1 &\leq \min\{I(U_1; Y_1), I(U_1; V_1 Y_2)\} \\
&\quad + \min\{I(U_2; V_1 V_2 Y_1 | U_1), I(U_2; V_1 Y_2 | U_1)\} - I(U_1 U_2; S) \\
R_2 &\leq \min\{I(V_1; Y_2), I(V_1; U_1 Y_1)\} \\
&\quad + \min\{I(V_2; U_1 U_2 Y_2 | V_1), I(V_2; U_1 Y_1 | V_1)\} - I(V_1 V_2; S)
\end{aligned} \tag{2.16}$$

for some distribution

$$P_{SU_1 U_2 V_1 V_2 X_1 X_2 Y_1 Y_2} = P_S P_{U_1|S} P_{U_2|SU_1} P_{X_1|U_1 U_2 S} P_{V_1|S} P_{V_2|SV_1} P_{X_2|V_1 V_2 S} P_{Y_1 Y_2|SX_1 X_2}$$

where U_1, U_2, V_1 , and V_2 are auxiliary random variables.

Proof. See Appendix A.3. □

Remark 2.1. *A more comprehensive achievable region can be obtained by taking the convex hull of the union over achievable regions resulting from all possible decoding orders of messages at the two receivers.*

Proposition 2.3 provides an example achievable region, based on which we next show that the designed scheme achieves the capacity region or partial boundary of the capacity region for the state-dependent Gaussian IC under certain conditions on channel parameters.

Namely, we characterize the conditions on the channel parameters under which points on the sum rate boundary of the IC without state (i.e., the line B-E in Fig. 2.3) can be achieved by the state-dependent Gaussian IC.

We note that any rate point on the line B-E can be characterized by

$$\begin{aligned} R_1 &= \frac{1}{2} \log \left(1 + \frac{P'_1}{a^2 P''_2 + P'_1 + 1} \right) + \frac{1}{2} \log (1 + P''_1) \\ R_2 &= \frac{1}{2} \log \left(1 + \frac{a^2 P'_2}{a^2 P''_2 + P'_1 + 1} \right) + \frac{1}{2} \log \left(1 + \frac{a^2 P''_2}{P''_1 + 1} \right) \end{aligned} \quad (2.17)$$

for some $P'_1, P''_1, P'_2, P''_2 \geq 0$, $P'_1 + P''_1 \leq P_1$, and $P'_2 + P''_2 \leq P_2$.

In order to achieve any rate point given in (2.17), we design layered dirty paper coding for the auxiliary random variables U_1, V_1, V_2 , and U_2 in order to successively decode messages and cancel the state at receiver 1. More specifically, dirty paper coding for U_1 is designed to cancel the state treating all other variables as noise, and then V_1, V_2 and U_2 are designed to successively cancel the residual state after subtracting the previously decoded auxiliary random variables from Y_1 . Furthermore, by requiring the rate bounds due to decoding at receiver 2 to be larger than those due to decoding at receiver 1, the rate point of interest is thus achievable for the state-dependent IC. We state this result in the following theorem.

Theorem 2.3. *Any rate point given in (2.17) with the parameters $(P'_1, P''_1, P'_2, P''_2)$ is on the capacity region boundary of the state-dependent IC if the channel parameters satisfy the following conditions*

$$\frac{1}{2} \log \left(1 + \frac{P'_1}{P''_1 + a^2 P_2 + 1} \right) \leq I(U_1; V_1 Y_2) \quad (2.18a)$$

$$\frac{1}{2} \log (1 + P''_1) \leq I(U_2; V_1 Y_2 | U_1) \quad (2.18b)$$

$$\frac{1}{2} \log \left(1 + \frac{a^2 P'_2}{P''_1 + a^2 P_2 + 1} \right) \leq I(V_1; Y_2) \quad (2.18c)$$

$$\frac{1}{2} \log \left(1 + \frac{a^2 P''_2}{P''_1 + 1} \right) \leq I(V_2; U_1 U_2 Y_2 | V_1) \quad (2.18d)$$

where the mutual information terms in the above conditions are computed based on the following auxiliary random variables

$$\begin{aligned} U_1 &= X'_1 + \alpha_1 S, & U_2 &= X''_1 + \alpha_2 S \\ V_1 &= aX'_2 + \beta_1 S, & V_2 &= aX''_2 + \beta_2 S \end{aligned} \quad (2.19)$$

where X'_1, X''_1, X'_2, X''_2 are independent Gaussian variables with mean zero and variances P'_1, P''_1, P'_2 and P''_2 , correspondingly, $X_1 = X'_1 + X''_1$, $X_2 = X'_2 + X''_2$, and $\alpha_1, \alpha_2, \beta_1$ and β_2 are given by

$$\begin{aligned} \alpha_1 &= \frac{P'_1}{P_1 + a^2 P_2 + 1}, & \alpha_2 &= \frac{P''_1}{P_1 + a^2 P_2 + 1} \\ \beta_1 &= \frac{a^2 P'_2}{P_1 + a^2 P_2 + 1}, & \beta_2 &= \frac{a^2 P''_2}{P_1 + a^2 P_2 + 1}. \end{aligned}$$

Proof. The achievability follows from Proposition 2.3 by choosing the auxiliary random variables U_1, U_2, V_1 , and V_2 as in (2.19) based on the successive dirty paper coding for removing the state from the received signal Y_1 so that the rate point given in (2.17) is achievable at receiver 1. For this rate point to be achievable also at receiver 2, following Proposition 2.3, the following conditions should be satisfied

$$I(U_1; Y_1) \leq I(U_1; V_1 Y_2) \quad (2.20a)$$

$$I(U_2; V_1 V_2 Y_1 | U_1) \leq I(U_2; V_1 Y_2 | U_1) \quad (2.20b)$$

$$I(V_1; U_1 Y_1) \leq I(V_1; Y_2) \quad (2.20c)$$

$$I(V_2; U_1 Y_1 | V_1) \leq I(V_2; U_1 U_2 Y_2 | V_1). \quad (2.20d)$$

By substituting the auxiliary random variables defined in (2.19) into (2.20a)-(2.20d), we obtain the conditions (2.18a)-(2.18d) on the channel parameters, under which the given boundary point is achievable by the state-dependent IC. Thus, such a point is on the capac-

ity region boundary, because it is on the capacity boundary of the channel without state, which serves as an outer bound. Formal justification can follow steps similar to those in Appendix A.7. \square

The mutual information terms in Theorem 2.4 can be explicitly computed in close forms. Thus, Theorem 2.4 provides a computable way for checking whether any point on the sum rate boundary of the capacity of the IC without state is also on the capacity boundary for the corresponding state-dependent channel under certain channel parameters. We provide an example range of parameters in Section 2.3.3.

2.3.2 State-Dependent Z-IC

In this subsection, we study the state-dependent Z-IC (i.e., $b = 0$) in the strong but not very strong regime, in which the channel parameters satisfy

$$1 \leq a^2 \leq (1 + P_1). \quad (2.21)$$

Under the above conditions, the Z-IC *without* state is strong (but not very strong Z-IC), and the capacity region is characterized in [3], which contains rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{2} \log(1 + P_1 + a^2 P_2) \\ R_1 &\leq \frac{1}{2} \log(1 + P_1), \quad R_2 \leq \frac{1}{2} \log(1 + P_2). \end{aligned} \quad (2.22)$$

The above capacity region is illustrated in Fig. 2.4 as the pentagon O-A-B-E-F-O, which is obtained by requiring receiver 1 to decode both messages and receiver 2 to decode the message W_2 .

Similarly to the regular IC, our goal here is also to study whether the points on the boundary of such a pentagon (i.e., the capacity region boundary of the Z-IC *without* state) can be achieved by the corresponding *state-dependent* Z-IC. We focus on the sum rate

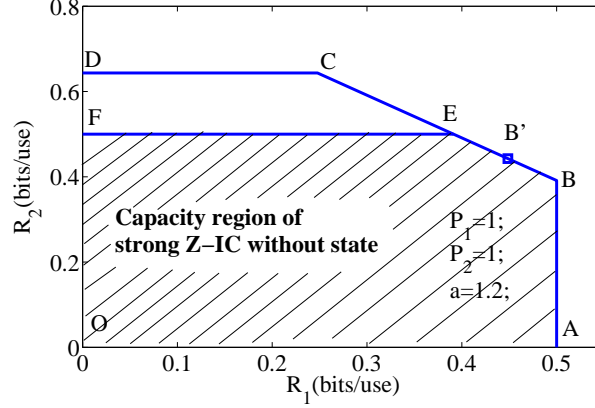


Fig. 2.4: Capacity region of the strong Z-IC without state

boundary of the pentagon (i.e., the line B-E in Fig. 2.4), and then the points on the line A-B and the line E-F are achievable if the two corner points B and E are achievable. We first design an achievable scheme for the state-dependent discrete memoryless Z-IC following the same idea as that for the regular IC based on rate splitting, layered dirty paper coding and successive state cancelation aiming at fully canceling the state at receiver 1. The only difference lies in that receiver 2 here decodes only W_{21} and W_{22} . Such a scheme then yields the following achievable rate region.

Proposition 2.4. *For the state-dependent Z-IC with state noncausally known at both transmitters, an achievable region consists of rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned}
 R_1 &\leq I(U_1; V_1 Y_1) + I(U_2; V_1 V_2 Y_1 | U_1) - I(U_1, U_2; S) \\
 R_2 &\leq \min\{I(V_1; Y_2), I(V_1; Y_1)\} \\
 &\quad + \min\{I(V_2; Y_2 | V_1), I(V_2; U_1 Y_1 | V_1)\} - I(V_1 V_2; S)
 \end{aligned} \tag{2.23}$$

for some distribution $P_{SU_1 U_2 V_1 V_2 X_2 X_1 Y_2 Y_1} = P_S P_{U_1 U_2 | S} P_{V_1 V_2 | S} P_{X_1 | U_1 U_2 S} P_{X_2 | V_1 V_2 S} P_{Y_1 | S X_1 X_2} P_{Y_2 | S X_2}$, where U_1, U_2, V_1 and V_2 are auxiliary random variables.

Proof. See Appendix A.4. □

Now specializing Proposition 2.4 to the Gaussian case yields an achievable region, based on which we can check if and under what conditions the points on the line B-E in

Fig. 2.4 are achievable. Since points on the line B-E can also be characterized in (2.17), we thus follow the same design of layered dirty paper coding for the auxiliary random variables U_1, V_1, V_2 , and U_2 as that for the regular IC in order to fully cancel the state at receiver 1 successively. Then by requiring the decoding bounds at receiver 2 to be larger than those of receiver 1, points on B-E can be shown to be achievable by the state-dependent Z-IC. We state this result in the following theorem.

Theorem 2.4. *Any rate point characterized in (2.17) with the parameters $(P'_1, P''_1, P'_2, P''_2)$ is on the capacity region boundary of the state-dependent Gaussian Z-IC with state non-causally known at the transmitters if the channel parameters satisfy the following conditions*

$$1 + \frac{a^2 P'_2}{a^2 P''_2 + P_1 + 1} \leq \frac{a^2 P'_2 (P_2 + b^2 Q + 1)}{P'_2 (ab - \alpha)^2 + (P''_2 + 1)(a^2 P'_2 + \alpha^2 Q)} \quad (2.24a)$$

$$1 + \frac{a^2 P''_2}{P''_1 + 1} \leq \frac{a^2 P''_2 [P'_2 (a^2 P''_2 + (ab - \alpha)^2 Q + a^2) + \alpha^2 Q (P''_2 + 1)]}{a^2 P''_2 (\alpha^2 Q + a^2 P'_2) + (a^2 b - a\alpha - a\gamma)^2 P'_2 P''_2 Q + a^2 \gamma^2 P'_2 Q} \quad (2.24b)$$

where $\alpha = \frac{a^2 P'_2}{a^2 P_2 + P_1 + 1}$, and $\gamma = \frac{a^2 P''_2}{a^2 P_2 + P_1 + 1}$.

Proof. In order to achieve a rate point given in (2.17) with the parameters $(P'_1, P''_1, P'_2, P''_2)$, we apply Proposition 2.4 and choose the auxiliary random variables U_1, U_2, V_1 , and V_2 based on the dirty paper coding as in (2.19) so that the state in the received signal Y_1 can be fully canceled.

In order for receiver 2 to decode at this rate point (without introducing more constraints on the rates), due to Proposition 2.4, the following conditions should be satisfied

$$I(V_1; Y_1) \leq I(V_1; Y_2), \quad I(V_2; U_1 Y_1 | V_1) \leq I(V_2; Y_2 | V_1). \quad (2.25)$$

By substituting the auxiliary random variables defined in (2.19) into (2.25), the conditions (2.24a) and (2.24b) on the channel parameters can be obtained, under which the rate point of interest is achievable over the state-dependent Z-IC. Thus, such a rate point is on

the capacity region boundary, because it is on the capacity region boundary of the channel without state, which serves as an outer bound. \square

Theorem 2.4 provides the conditions on the channel parameters under which a certain given point is on the capacity region boundary. In Proposition 2.5 and Corollary 2.2, we also characterize a line segment on the capacity region boundary for a given set of channel parameters.

Proposition 2.5. *For the state-dependent Gaussian Z-IC with state noncausally known at both transmitters, if a point (say B') on the line $B - E$ in Fig. 2.4 satisfies the conditions in Theorem 4, i.e., it is on the capacity region boundary, then the point B is also on the capacity region boundary, and thus the line segment $B' - B$ is on the capacity region boundary.*

Proof. See Appendix A.5. \square

Based on Proposition 2.5, we characterize a segment on the capacity region boundary in the following corollary.

Corollary 2.2. *For the state-dependent Gaussian Z-IC with state noncausally known at the transmitters, let $R_2^* = \frac{1}{2} \log(\frac{a^2 P_2 (P_2 + b^2 Q + 1)}{P_2 Q (ab - \beta)^2 + a^2 P_2 + \beta^2 Q})$, where $\beta = \frac{a^2 P_2}{a^2 P_2 + P_1 + 1}$. If $R_2^* > \frac{1}{2} \log(1 + \frac{a^2 P_2}{1 + P_1})$, then the line $B - B'$ are on the capacity region boundary with the rate coordinates of the points B and B' given by*

$$\begin{aligned} \text{Point } B &: \left(\frac{1}{2} \log(1 + P_1), \frac{1}{2} \log(1 + \frac{a^2 P_2}{1 + P_1}) \right) \\ \text{Point } B' &: \left(\frac{1}{2} \log(1 + a^2 P_2 + P_1) - R_2^*, R_2^* \right). \end{aligned} \quad (2.26)$$

Proof. See Appendix A.6. \square

2.3.3 Comparison of State-Dependent Regular IC and Z-IC

In this subsection, we compare the result in Theorem 2.3 for the state-dependent regular IC and the result in Theorem 2.4 for the state-dependent Z-IC in the strong interference regime.

In Fig. 2.5, we plot the parameter ranges characterized in Theorem 2.3 and in Theorem 2.4. For both the regular IC and the Z-IC, we set $P_1 = 1$, $P_2 = 1$, $Q = 2$ and $a = 1.2$. Moreover, for the regular IC, we set $b = 4$, which implies that the interference is strong enough such that its corresponding channel without state has the same capacity region as that of the Z-IC. Thus, the only flexible parameter left for both the regular IC and the Z-IC is the scaling coefficient c for the state. We study the range of c that guarantees the points on the line $B - E$ to be on the capacity region boundary of the state-dependent regular IC and Z-IC. We note that each point on the line $B - E$ can be parameterized as the rate pair $(R_1, R_2) = (R_1, \frac{1}{2} \log(1 + P_1 + a^2 P_2) - R_1)$, where R_1 changes from $R_1 = 0.5$ (corresponding to point B) to $R_1 = \frac{1}{2} \log 1.72$ (corresponding to point E). In Fig. 2.5, for each R_1 (and hence for each corresponding point on the $B - E$ line), we plot the range of c that guarantees the point (R_1, R_2) to be on the capacity region boundary of the state-dependent regular IC to be between the two solid lines, and plot the range of c that guarantees the point (R_1, R_2) to be on the capacity region boundary of the state-dependent Z-IC between the two dashed lines. Although the two ranges do not overlap, their structures are similar and the sizes of the ranges are comparable. This implies that both channels have the same flexibility to achieve the capacity region boundary point of the corresponding channel without state, and hence suggests that neither channel cancels the state more easily than the other. This is because for both the regular IC and the Z-IC, the layered dirty paper coding is designed in the same way to successively cancel the state for receiver 1. Hence, the advantage of the Z-IC at the other receiver is not significant due to the state interference that is not fully canceled. We further note that Fig. 2.5 also suggests

that it is easier to achieve a point on the $B - E$ line when the point is closer to the point B for both channels.

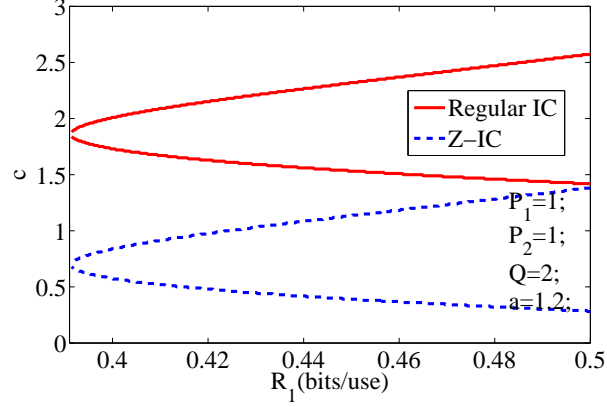


Fig. 2.5: Ranges of c under which points on the sum capacity boundary of the strong regular/Z-IC without state can be achieved by the state-dependent channel

2.4 Weak Interference Regime

In this section, we study both the state-dependent regular IC and Z-IC in the weak interference regime, in which the channel parameters satisfy $|a(1 + b^2 P_1)| + |b(1 + a^2 P_2)| \leq 1$ for the regular IC and satisfy $a^2 \leq 1$ for the Z-IC. Under such conditions, the sum capacity for the regular IC without state has been established in [4–6], and for the Z-IC without state has been established in [23]. In both cases, the sum capacity can be achieved by treating interference as noise at each receiver. Hence, for the corresponding state-dependent IC, independent dirty paper coding at two transmitters to cancel the state at their corresponding receivers (treating the interference as noise) can achieve the same sum capacity. Decoding at each receiver is not affected by how the interference signal is coded. Such an observation yields the following results.

Theorem 2.5. *For the state-dependent Gaussian IC with state noncausally known at both transmitters, if $|a(1 + b^2 P_1)| + |b(1 + a^2 P_2)| \leq 1$, then the sum capacity is given by*

$$C_{sum} = \frac{1}{2} \log \left(1 + \frac{P_1}{a^2 P_2 + 1} \right) + \frac{1}{2} \log \left(1 + \frac{P_2}{b^2 P_1 + 1} \right). \quad (2.27)$$

For the state-dependent Gaussian Z-IC with state noncausally known at both transmitters, if $a^2 \leq 1$, then the sum capacity is given by

$$C_{sum} = \frac{1}{2} \log \left(1 + \frac{P_1}{a^2 P_2 + 1} \right) + \frac{1}{2} \log (1 + P_2). \quad (2.28)$$

Proof. See Appendix A.7. □

CHAPTER 3

STATE-DEPENDENT COGNITIVE

INTERFERENCE CHANNEL

In this chapter, we study the cognitive interference channel with state. More specifically, we consider two sub models, i.e., the CIC-STR and CIC-ST. For the CIC-STR, we characterize the capacity region for both discrete memoryless channel and Gaussian channel. In particular, we partition the Gaussian CIC-STR into two sets based on the channel parameters, and derive the capacity region for the two sets, respectively. For the CIC-ST, we derive inner and outer bound for the discrete memoryless channel and its degraded version, and obtain the capacity region for channels that satisfy certain conditions. We then study the Gaussian CIC-ST. We also partition the channel into two sets, and derive inner and outer bounds for the two sets. By comparing the inner and outer bounds, we obtain the partial capacity boundary for the Gaussian CIC-ST, and full capacity region for channel with parameters satisfying certain conditions.

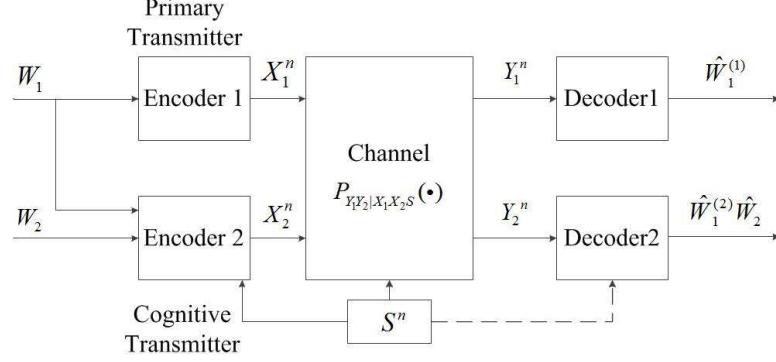


Fig. 3.1: The CIC-STR (including the dashed line) and the CIC-ST (without the dashed line) models

3.1 Channel Model

For the cognitive interference channel with state known at one transmitter (see Fig. 1.3 in Section 1.2. For convenience of reference, we include the figure again as Fig. 3.1 in this section), we investigate two scenarios, i.e., CIC-STR and the CIC-ST.

In the CIC-ST, two transmitters (referred to as the primary transmitter and the cognitive transmitter) jointly send a message W_1 to two receivers (say receivers 1 and 2), and the cognitive transmitter sends another message W_2 to receiver 2. The channel is also corrupted by an i.i.d. state sequence. The scenario, in which the state sequence is noncausally known at both the cognitive transmitter and receiver 2 (CIC-STR) and the scenario, in which the state sequence is noncausally known only at the cognitive transmitter (CIC-ST) are studied. More specifically, encoder 1 $f_1 : \mathcal{W}_1 \rightarrow \mathcal{X}_1^n$ at transmitter 1 maps a message $w_1 \in \mathcal{W}_1$ to a codeword $x_1^n \in \mathcal{X}_1^n$, and encoder 2 $f_2 : \mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n$ at transmitter 2 maps a message pair $(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_2$ and a state sequence $s^n \in \mathcal{S}^n$ to a codeword $x_2^n \in \mathcal{X}_2^n$. Decoder 1 $g_1 : \mathcal{Y}_1^n \rightarrow \mathcal{W}_1$ at receiver 1 maps a received sequence y_1^n into a message $\hat{w}_1^{(1)} \in \mathcal{W}_1$, and decoder 2 $g_2 : \mathcal{Y}_2^n \rightarrow \mathcal{W}_1 \times \mathcal{W}_2$ at receiver 2 maps a received sequence y_2^n into a message pair $(\hat{w}_1^{(2)}, \hat{w}_2) \in \mathcal{W}_1 \times \mathcal{W}_2$ with the probability of error approaching zero as the codeword length n goes to infinity. The capacity region is defined to be the closure of the set of all achievable rate pairs (R_1, R_2) .

We note that the above definition is also applicable to the CIC-STR, if the second decoder is changed to $g_2 : (\mathcal{Y}_2^n, S^n) \rightarrow \mathcal{W}_1 \times \mathcal{W}_2$.

In the following, we define a number of channel conditions for classifying the channels in our study:

$$\bullet P_{Y_1 Y_2 | X_1 X_2 S} = P_{Y_2 | X_1 X_2 S} P_{Y_1 | Y_2} \quad (3.1)$$

$$\bullet P_{Y_1 Y_2 | X_1 X_2 S} = P_{Y_2 | X_1 X_2 S} P_{Y_1 | Y_2 X_1 S} \quad (3.2)$$

$$\bullet P_{Y_1 Y_2 | X_1 X_2 S} = P_{Y_1 | X_1 X_2 S} P_{Y_2 | Y_1 X_1 S} \quad (3.3)$$

$$\bullet I(X_1; Y_1) \leq I(X_1; Y_2) \text{ and } I(U; Y_1 | X_1) \leq I(U; Y_2 | X_1) \\ \text{for all } P_{U X_1 X_2 S} \text{ s.t. } P_{X_1 S U X_2} = P_{X_1} P_S P_{U X_2 | S X_1} \quad (3.4)$$

$$\bullet I(X_1 U; Y_1) \geq I(X_1 U; Y_2) \\ \text{for all } P_{U X_1 X_2 S} \text{ s.t. } P_{X_1 S U X_2} = P_{X_1} P_S P_{U X_2 | S X_1} \quad (3.5)$$

We also study the Gaussian CIC-ST and CIC-STR models defined as follows. We note that the two models have the same input-output relationship. The Gaussian CIC-ST and CIC-STR have outputs at receivers 1 and 2 for one symbol time given by

$$Y_1 = X_1 + aX_2 + S + N_1 \quad (3.6a)$$

$$Y_2 = bX_1 + X_2 + cS + N_2 \quad (3.6b)$$

where the noise variables $N_1 \sim \mathcal{N}(0, 1)$ and $N_2 \sim \mathcal{N}(0, 1)$, and the state variable $S \sim \mathcal{N}(0, Q)$. Both the noise variables and the state variable are i.i.d. over channel uses. The channel inputs are subject to the average power constraints $\frac{1}{n} \sum_{i=1}^n X_{1i}^2 \leq P_1$, and $\frac{1}{n} \sum_{i=1}^n X_{2i}^2 \leq P_2$.

3.2 The CIC-STR Model

In this section, we present our results for the CIC-STR. We first provide the capacity region for the discrete memoryless channel, and then characterize the capacity region for the Gaussian model for two sets of channels: $|a| > 1$ and $|a| \leq 1$.

3.2.1 Discrete Memoryless Channels

We design an achievable scheme that employs rate-splitting, superposition coding and Gel'fand-Pinsker binning scheme. The primary transmitter first encodes W_1 . Then the cognitive transmitter cooperatively encodes and transmits W_1 using superposition. Moreover, the cognitive transmitter employs rate splitting for transmitting W_2 , i.e., splits W_2 into two components W_{21} and W_{22} with W_{21} intended for both receivers to decode and W_{22} intended only for receiver 2 to decode. The cognitive transmitter encodes W_{21} and W_{22} by superposing them on W_1 . Furthermore, since the cognitive transmitter knows the channel state information, it employs Gel'fand-Pinsker scheme via an auxiliary random variable U (in the following capacity region) to reduce state interference for receiver 1 to decode W_1 and W_{21} . Hence, U contains information of both W_1 and W_{21} , and plays dual roles: helping to cancel state interference and serving as a rate splitting random variable for carrying the message W_{21} . We also note that since receiver 2 has the knowledge of the state, no additional auxiliary random variable is needed for cancelling state interference for receiver 2.

The CIC-STR is easier to analyze than the CIC-ST, because receiver 2 knows the state and can hence remove the state interference from its output. In this way, the design of achievable schemes needs to deal with only the state interference at receiver 1. Whereas for the CIC-ST, in which the state information is known at neither receiver, the achievable scheme needs to deal with state interference at both receivers. This involves the design for compound states, and hence results in a more challenging problem.

We characterize the full capacity region for the CIC-STR in the following theorem.

Theorem 3.1. *The capacity region for the CIC-STR consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq I(X_1 U; Y_1) - I(U; S|X_1) \quad (3.7a)$$

$$R_2 \leq I(X_2; Y_2|S X_1) \quad (3.7b)$$

$$R_1 + R_2 \leq I(X_1 X_2; Y_2|S) \quad (3.7c)$$

$$R_1 + R_2 \leq I(X_1 U; Y_1) + I(X_2; Y_2|X_1 U S) - I(U; S|X_1) \quad (3.7d)$$

for some distribution $P_{X_1 S U X_2 Y_1 Y_2} = P_{X_1} P_S P_{U X_2|X_1 S} P_{Y_1 Y_2|S X_1 X_2}$, where U is an auxiliary random variable and its cardinality is bounded by $|\mathcal{U}| \leq |\mathcal{X}_1| |\mathcal{X}_2| |\mathcal{S}| + 1$.

Proof. Since the CIC-DM-STR can be viewed as a special case of the CIC-DM-ST with $Y_2 = (Y_2, S)$, the achievability proof follows directly from the achievable region for the CIC-DM-ST given in (3.18a)-(3.18e) by setting $T = X_1$, $V = X_2$ and $Y_2 = Y_2 S$.

For the converse, we first obtain the following outer bound consisting of rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(K X_1; Y_1) - I(K; S|X_1) \quad (3.8a)$$

$$R_2 \leq I(X_2; Y_2|S X_1) \quad (3.8b)$$

$$R_1 + R_2 \leq I(X_1 X_2; Y_2|S) \quad (3.8c)$$

$$R_1 + R_2 \leq I(T K X_1; Y_1) - I(T K; S|X_1) + I(X_2; Y_2|X_1 T K S) \quad (3.8d)$$

for some distribution $P_{X_1 S T K X_2 Y_1 Y_2} = P_{X_1} P_S P_{K T|X_1 S} P_{X_2|X_1 S K T} P_{Y_1 Y_2|S X_1 X_2}$, where K and T are auxiliary random variables. The proof is detailed in Appendix B.1.

In order to show that the region (3.7a)-(3.7d) is the capacity region, it is sufficient to show that the above outer bound (3.8a)-(3.8d) is a subset of the region (3.7a)-(3.7d). Towards this end, we apply the technique in [13] and analyze the outer bound (3.8a)-(3.8d)

by considering the following two cases.

If $I(T; Y_1 | KX_1) - I(T; S | KX_1) \leq 0$, the outer bound (3.8a)-(3.8d) can be further bounded as:

$$R_1 \leq I(KX_1; Y_1) - I(K; S | X_1) \quad (3.9a)$$

$$R_2 \leq I(X_2; Y_2 | SX_1) \quad (3.9b)$$

$$R_1 + R_2 \leq I(X_1X_2; Y_2 | S) \quad (3.9c)$$

$$\begin{aligned} R_1 + R_2 &\leq I(KX_1; Y_1) - I(K; S | X_1) + [I(T; Y_1 | KX_1) \\ &\quad - I(T; S | KX_1)] + I(X_2; Y_2 | X_1TKS) \\ &\leq I(KX_1; Y_1) - I(K; S | X_1) + I(X_2; Y_2 | X_1KS), \end{aligned} \quad (3.9d)$$

which implies that the outer bound (3.8a)-(3.8d) is contained in (3.7a)-(3.7d) by setting $U = K$ in (3.7a)-(3.7d).

If $I(T; Y_1 | KX_1) - I(T; S | KX_1) \geq 0$, the outer bound (3.8a)-(3.8d) can be further bounded as:

$$\begin{aligned} R_1 &\leq I(KX_1; Y_1) - I(K; S | X_1) \\ &= I(KTX_1; Y_1) - I(KT; S | X_1) - [I(T; Y_1 | KX_1) - I(T; S | KX_1)] \\ &\leq I(KTX_1; Y_1) - I(KT; S | X_1) \end{aligned} \quad (3.10a)$$

$$R_2 \leq I(X_2; Y_2 | SX_1) \quad (3.10b)$$

$$R_1 + R_2 \leq I(X_1X_2; Y_2 | S) \quad (3.10c)$$

$$R_1 + R_2 \leq I(TKX_1; Y_1) - I(TK; S | X_1) + I(X_2; Y_2 | X_1KTS) \quad (3.10d)$$

which also implies that the outer bound (3.8a)-(3.8d) is contained in (3.7a)-(3.7d) by setting $U = KT$ in (3.7a)-(3.7d). \square

3.2.2 Gaussian Channels

In this section, we characterize the capacity region for the Gaussian CIC-STR. We partition Gaussian channels into two classes based on the value of the channel parameter a , and characterize the capacity region for each class. We note that our results for Gaussian channels exploit the fact that for both $|a| \leq 1$ and $|a| > 1$, the Gaussian channel is stochastically degraded given X_1 and S , i.e., its marginal distributions at the two receivers are the same as a physically degraded Gaussian channel that satisfies the conditions (3.2) and (3.3), respectively. Because the capacities of the two Gaussian channels are the same, our results below are applicable to both stochastically degraded and physically degraded channels with the proofs exploiting the physical degradedness conditions (3.2) and (3.3).

We first provide the capacity region for the Gaussian channel with $|a| \leq 1$.

Theorem 3.2. *For the Gaussian CIC-STR, if $|a| \leq 1$, the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) + \frac{1}{2} \log \left(1 + \frac{a^2P_2'}{a^2P_2'' + 1} \right) \quad (3.11a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2'') \quad (3.11b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + b^2P_1 + 2b\rho_{21}\sqrt{P_1P_2} + (1 - \rho_{2s}^2)P_2 \right) \quad (3.11c)$$

where $P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $P_2' \geq 0$, $P_2'' \geq 0$, and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

We explain the achievable scheme used for obtaining the above capacity region as follows. Here, the cognitive transmitter's power P_2 is split into three parts: 1.cooperatively transmitting W_1 via beamforming, 2.transmitting additional W_1 via an auxiliary random variable U to deal with the state at receiver 1 using dirty paper coding, 3.transmitting W_2 . Here, rate splitting is not used, i.e., $W_{21} = \phi$, because for the case $|a| \leq 1$, forcing receiver 1 to decode certain W_{21} may reduce the achievable region.

Proof. Consider the following rate region, which consists of rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(X_1 U; Y_1) - I(U; S|X_1) \quad (3.12a)$$

$$R_2 \leq I(X_2; Y_2|U X_1 S) \quad (3.12b)$$

$$R_1 + R_2 \leq I(X_1 X_2; Y_2|S) \quad (3.12c)$$

for some distribution $P_{S X_1 U X_2 Y_1 Y_2} = P_{X_1} P_S P_{U X_2|X_1 S} P_{Y_2|X_1 X_2 S} P_{Y_1|Y_2 X_1 S}$. This region is contained in (3.7a)-(3.7d), and is hence achievable. This can be seen by observing that $I(X_2; Y_2|U X_1 S) \leq I(X_2 U; Y_2|X_1 S)$ and the sum rate bound (3.7d) is equal to the sum of the two bounds on the individual rates in (3.12a) and (3.12b).

The achievability of (3.11a)-(3.11c) is then obtained by choosing the following jointly Gaussian distribution for the random variables:

$$\begin{aligned} X_1 &\sim \mathcal{N}(0, P_1), \quad X'_2 \sim \mathcal{N}(0, P'_2), \quad X''_2 \sim \mathcal{N}(0, P''_2), \\ P'_2 + P''_2 &= (1 - \rho_{21}^2 - \rho_{2s}^2) P_2 \\ X_2 &= \rho_{21} \sqrt{\frac{P_2}{P_1}} X_1 + X'_2 + X''_2 + \rho_{2s} \sqrt{\frac{P_2}{Q}} S \\ U &= X'_2 + \alpha \left(1 + a \rho_{2s} \sqrt{\frac{P_2}{Q}} \right) S \end{aligned} \quad (3.13)$$

where X_1, X'_2, X''_2 and S are independent, and $\alpha = \frac{a^2 P'_2}{a^2 P'_2 + a^2 P''_2 + 1}$.

The converse proof is detailed in Appendix B.2. □

We next characterize the capacity region for the Gaussian channel with $|a| > 1$.

Theorem 3.3. *For the Gaussian CIC-STR, if $|a| > 1$, the capacity region consists of rate*

pairs (R_1, R_2) satisfying:

$$R_2 \leq \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \quad (3.14a)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + (1 - \rho_{2s}^2)P_2) \quad (3.14b)$$

$$\begin{aligned} R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1 P_2} + a^2 \rho_{21}^2 P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2 Q} + Q + 1} \right) \\ + \frac{1}{2} \log(1 + a^2(1 - \rho_{2s}^2 - \rho_{21}^2)P_2) \end{aligned} \quad (3.14c)$$

where $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

Differently from Theorem 3.2, due to the fact that $|\alpha| > 1$, receiver 1 is stronger in decoding W_2 . Hence, the achievable scheme sets $W_{21} = W_2$, i.e., requires receiver 1 to decode the full message W_2 . The cognitive transmitter's power P_2 is split into two parts: 1.cooperatively transmitting W_1 via beamforming, 2.transmitting additional W_1 and $W_{21} = W_2$ via an auxiliary random variable U to deal with the state at receiver 1 using dirty paper coding.

Proof. The achievability follows from (3.7a)-(3.7d) by choosing jointly Gaussian distribution for random variables as follows:

$$\begin{aligned} X_1 &\sim \mathcal{N}(0, P_1), \quad X'_2 \sim \mathcal{N}(0, (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \\ X_2 &= \rho_{21}\sqrt{\frac{P_2}{P_1}}X_1 + X'_2 + \rho_{2s}\sqrt{\frac{P_2}{Q}}S \\ U &= X'_2 + \alpha \left(1 + a\rho_{2s}\sqrt{\frac{P_2}{Q}} \right) S \end{aligned} \quad (3.15)$$

where X_1 , X'_2 and S are independent, and $\alpha = \frac{a^2(1-\rho_{21}^2-\rho_{2s}^2)P_2}{a^2(1-\rho_{21}^2-\rho_{2s}^2)P_2+1}$. We note that with this choice of the random variables, the first bound in (3.7a)-(3.7d) is redundant.

In order to prove the converse for Theorem 3.3, we first prove the following outer bound.

Lemma 3.1. *For the CIC-DM-STR, if it satisfies the condition (3.3), an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying*

$$R_2 \leq I(X_2; Y_2 | SX_1) \quad (3.16a)$$

$$R_1 + R_2 \leq I(X_1 X_2; Y_2 | S) \quad (3.16b)$$

$$R_1 + R_2 \leq I(X_1; Y_1) + I(X_2; Y_1 | SX_1) \quad (3.16c)$$

for some distribution $P_{SX_1 U X_2 Y_1 Y_2} = P_{X_1} P_S P_{U X_2 | X_1 S} P_{Y_1 | X_1 X_2 S} P_{Y_2 | Y_1 X_1 S}$.

The proof for the above lemma is detailed in Appendix B.3. For the Gaussian channel with $|a| > 1$, it satisfies the condition (3.3). We then use the above lemma for developing the converse proof, which is detailed in Appendix B.4. \square

3.3 The CIC-ST Model

In this section, we present our results for the CIC-ST. We first derive inner and outer bound for the discrete memoryless channel, and then characterize the capacity region for channel under certain conditions. For the Gaussian CIC-ST, we partition the channel into two classes based on the channel condition, and derive inner and outer bounds for both classes. By comparing the inner and outer bounds, we obtain partial boundary for the capacity region.

3.3.1 Discrete Memoryless Channels

In this section, we investigate the discrete memoryless CIC-ST model. We first provide inner and outer bounds on the capacity region, and then identify a few special cases, for which we establish the capacity region.

In order to derive an inner bound on the capacity region, we design an achievable scheme, which includes superposition coding, rate-splitting, and Gel'fand-Pinsker binning

scheme. The primary and cognitive transmitters cooperatively transmit W_1 . The cognitive transmitter splits W_2 into two components W_{21} and W_{22} with W_{21} intended for both receivers and W_{22} intended only for receiver 2. Differently from the scheme for the CIC-STR, here the cognitive transmitter employs Gel'fand-Pinsker scheme via three auxiliary random variables T , U and V (as in Lemma 3.2) to reduce state interference respectively for W_1 , W_{21} and W_{22} . In particular, T deals with state interference for either receiver 1 or receiver 2 to decode W_1 , U deals with state interference for either receiver 1 or receiver 2 to decode W_{21} , and V deals with state interference for receiver 2 to decode W_{22} . In particular, T and U cannot be combined because it is possible that U deals with the state at receiver 2 whereas T deals with the state at receiver 1. This also explains the reason that only one auxiliary random variable U is needed for obtaining the capacity region for the CIC-STR model, in which only state interference at receiver 1 needs to be handled, and hence a single auxiliary random variable U (combining T and U) is sufficient for receiver 1 to decode both W_1 and W_{21} . At the receiver end, since receiver 1 can decode W_{21} , it can eliminate the interference caused by this message when it decodes W_1 .

We now provide an achievable region based on the above achievable scheme, which is useful in establishing our main inner bound.

Lemma 3.2. *An achievable region for the CIC-ST consists of rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned}
R_2 &= R_{21} + R_{22}, \quad R_{21} \geq 0, \quad R_{22} \geq 0 \\
R_1 + R_{21} &\leq I(TUX_1; Y_1) - I(TU; S|X_1) \\
R_{22} &\leq I(V; Y_2|UTX_1) - I(V; S|UTX_1) \\
R_{21} + R_{22} &\leq I(UV; Y_2|X_1T) - I(UV; S|X_1T) \\
R_{21} + R_{22} &\leq I(TUV; Y_2|X_1) - I(TUV; S|X_1) \\
R_1 + R_{21} + R_{22} &\leq I(TUVX_1; Y_2) - I(TUV; S|X_1)
\end{aligned}$$

for some distribution $P_{X_1STUVX_2Y_1Y_2} = P_{X_1}P_S P_{TUVX_2|SX_1}P_{Y_1Y_2|SX_1X_2}$, where T , U and V

are auxiliary random variables.

Proof. The detailed proof is relegated to Appendix B.5 \square

Based on Lemma 3.2, our main inner bound on the capacity region is given in the following theorem.

Theorem 3.4. *For the CIC-ST, an achievable region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq I(X_1 T U; Y_1) - I(T U; S | X_1) \quad (3.18a)$$

$$R_2 \leq I(U V; Y_2 | X_1 T) - I(U V; S | X_1 T) \quad (3.18b)$$

$$R_2 \leq I(T U V; Y_2 | X_1) - I(T U V; S | X_1) \quad (3.18c)$$

$$R_1 + R_2 \leq I(X_1 T U V; Y_2) - I(T U V; S | X_1) \quad (3.18d)$$

$$R_1 + R_2 \leq I(X_1 T U; Y_1) + I(V; Y_2 | X_1 T U) - I(T U V; S | X_1) \quad (3.18e)$$

for some distribution $P_{X_1 S T U V X_2 Y_1 Y_2} = P_{X_1} P_S P_{T U V X_2 | S X_1} P_{Y_1 Y_2 | S X_1 X_2}$ that satisfies

$$I(V; Y_2 | U T X_1) - I(V; S | U T X_1) \geq 0.$$

Proof. By applying Fourier-Motzkin elimination [52], we eliminate R_{21} and R_{22} from the bounds in Lemma 3.2 and obtain the bounds in Theorem 3.4. \square

We next derive the following inner bound, which is achieved by a simpler scheme that combines T and U together as one auxiliary random variable. This inner bound is useful for studying Gaussian channels in Section 3.3.2.2.

Corollary 3.1. *For the CIC-ST, an achievable region consists of rate pairs (R_1, R_2) satis-*

fyng:

$$\begin{aligned}
R_1 &\leq I(X_1 T; Y_1) - I(T; S|X_1) \\
R_2 &\leq I(V; Y_2|X_1 T) - I(V; S|X_1 T) \\
R_2 &\leq I(TV; Y_2|X_1) - I(TV; S|X_1) \\
R_1 + R_2 &\leq I(X_1 TV; Y_2) - I(TV; S|X_1)
\end{aligned} \tag{3.19a}$$

for some distribution $P_{X_1 STV X_2 Y_1 Y_2} = P_{X_1} P_S P_{TV X_2|X_1 S} P_{Y_1 Y_2|S X_1 X_2}$ that satisfies

$$I(V; Y_2|T X_1) - I(V; S|T X_1) \geq 0. \tag{3.20}$$

Proof. The achievable region in Corollary 3.1 follows directly from Theorem 3.4 by setting $U = T$. \square

We next provide an outer bound on the capacity region for the CIC-ST.

Theorem 3.5. *An outer bound for the the CIC-ST consists of the rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned}
R_1 &\leq I(X_1 T U; Y_1) - I(T U; S|X_1) \\
R_2 &\leq I(TV; Y_2|X_1) - I(TV; S|X_1) \\
R_1 + R_2 &\leq I(X_1 TV; Y_2) - I(TV; S|X_1)
\end{aligned}$$

for some distribution $P_{X_1 STUV X_2 Y_1 Y_2} = P_{X_1} P_S P_{TUV X_2|X_1 S} P_{Y_1 Y_2|S X_1 X_2}$, which satisfies the Markov chain conditions $T \leftrightarrow UV \leftrightarrow X_1 X_2 S \leftrightarrow Y_1 Y_2$.

Proof. The proof employs the techniques in [9] for the Gel'fand-Pinsker model, and exploits independence properties among variables in our model. In particular, the auxiliary random variables are carefully constructed. The detailed proof is relegated to Appendix B.6. \square

We now provide inner and outer bounds for the degraded channel, which are useful for further identifying the cases for which we obtain the capacity region.

Theorem 3.6. *If the CIC-ST satisfies the degradedness condition (3.1) (i.e., receiver 1 is degraded with regard to receiver 2), then an achievable region consists of the rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq I(X_1 T; Y_1) - I(T; S|X_1) \quad (3.22a)$$

$$R_2 \leq I(V; Y_2|X_1 T) - I(V; S|X_1 T) \quad (3.22b)$$

$$R_2 \leq I(TV; Y_2|X_1) - I(TV; S|X_1) \quad (3.22c)$$

for some distribution $P_{X_1 S T V X_2 Y_1 Y_2} = P_{X_1} P_S P_{T V X_2 | X_1 S} P_{Y_1 Y_2 | S X_1 X_2}$ that satisfies

$$I(V; Y_2 | T X_1) - I(V; S | T X_1) \geq 0.$$

An outer bound on the capacity region for such a channel consists of the rate pairs (R_1, R_2) satisfying:

$$R_1 \leq I(X_1 T; Y_1) - I(T; S|X_1)$$

$$R_2 \leq I(TV; Y_2|X_1) - I(TV; S|X_1)$$

for some distribution $P_{X_1 S T V X_2 Y_1 Y_2} = P_{X_1} P_S P_{T V X_2 | X_1 S} P_{Y_1 Y_2 | S X_1 X_2}$, which satisfies the Markov chain conditions $T \leftrightarrow V \leftrightarrow X_1 X_2 S \leftrightarrow Y_1 Y_2$.

Proof. The achievability follows from the achievable region given in Corollary 3.1 by removing the bound (3.19a) due to the degradedness condition. The proof of the outer bound is detailed in Appendix B.7. \square

The inner and outer bounds given in Theorems 3.4 and 3.5 do not match in general. We next identify two classes of channels, for which we obtain the capacity region. We first

provide the capacity region for the degraded semideterministic channel in the following theorem.

Theorem 3.7. *If the CIC-ST model satisfies the degradedness condition (3.1) and the semideterministic condition such that Y_2 is a deterministic function of X_1 , X_2 and S , then the capacity region of the channel consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq I(X_1 T; Y_1) - I(T; S | X_1) \quad (3.24a)$$

$$R_2 \leq H(Y_2 | X_1 T S) \quad (3.24b)$$

$$R_2 \leq H(Y_2 | X_1) - I(T Y_2; S | X_1) \quad (3.24c)$$

for some distribution $P_{X_1 S T X_2 Y_1 Y_2} = P_{X_1} P_S P_{T X_2 | S X_1} P_{Y_2 | X_1 X_2 S} P_{Y_1 | Y_2}$, where T is an auxiliary random variable and its cardinality is bounded by $|\mathcal{T}| \leq |\mathcal{X}_1| |\mathcal{X}_2| |\mathcal{S}| + 1$.

Proof. The achievability follows from (3.22a)-(3.22c) by setting $V = Y_2$. The proof of the converse is detailed in Appendix B.8. \square

We next obtain the following capacity region when receiver 1 is less noisy than receiver 2, i.e, the channel satisfies the condition (3.5).

Theorem 3.8. *For the CIC-ST, if it satisfies the condition (3.5), the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_2 \leq I(U; Y_2 | X_1) - I(U; S | X_1)$$

$$R_1 + R_2 \leq I(X_1 U; Y_2) - I(U; S | X_1)$$

for some distribution $P_{X_1 S U X_2 Y_1 Y_2} = P_{X_1} P_S P_{U X_2 | X_1 S} P_{Y_1 Y_2 | S X_1 X_2}$, where U is an auxiliary random variable and its cardinality is bounded by $|\mathcal{U}| \leq |\mathcal{X}_1| |\mathcal{X}_2| |\mathcal{S}|$.

We note that if condition (3.5) is satisfied, receiver 1 is less noisy than receiver 2. Thus, bounds on achievable rates are dominated by receiver 2, and only one auxiliary random

variable U is needed for dealing with state interference for receiver 2 to decode all messages.

Proof. The achievability follows from Theorem 1 by setting $T = \phi$, $V = U$ and using (3.5) to remove the redundant bounds. The converse follows from the capacity region of the MAC (with its receiver being receiver 2 in our model) with state available at one transmitter given in [18], which clearly is an outer bound for our model. \square

3.3.2 Gaussian Channels

In this section, we consider the Gaussian CIC-ST model. Similarly to Section 3.2.2, we partition the Gaussian CIC-ST into two classes corresponding to $|a| > 1$ and $|a| \leq 1$, and study these two classes separately in this and next subsections. In each subsection, we first provide inner and outer bounds on the capacity region, and then characterize partial boundaries of the capacity region based on these bounds. We also obtain the full capacity region for channels that satisfy certain conditions.

3.3.2.1 Gaussian Channel: $|a| > 1$

If $|a| > 1$, the Gaussian channel satisfies the condition (3.3). We first provide an inner bound for this class of channels.

Proposition 3.1. *For the Gaussian CIC-ST, if $|a| > 1$, an inner bound consists of rate pairs (R_1, R_2) satisfying:*

$$R_2 \leq \frac{1}{2} \log(1 + P'_2) \quad (3.26a)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) + \frac{1}{2} \log(1 + P'_2) \quad (3.26b)$$

$$\begin{aligned} R_1 + R_2 \leq & \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1 P_2} + a^2 \rho_{21}^2 P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2 Q} + Q + 1} \right) \\ & + \frac{1}{2} \log \left(1 + \frac{a^2 P_2'^2 + 2a\rho_{2s1}\rho_{2s2}P_2' - a^2 \rho_{2s1}^2 P_2' - \rho_{2s1}^2}{a^2 \rho_{2s1}^2 P_2' + \rho_{2s2}^2 P_2' + P_2' + \rho_{2s1}^2 - 2a\rho_{2s1}\rho_{2s2}P_2'} \right) \end{aligned} \quad (3.26c)$$

where $P'_2 = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$ and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$, $\rho_{2s1} = \alpha(c\sqrt{Q} + \rho_{2s}\sqrt{P_2})$, $\rho_{2s2} = (\sqrt{Q} + a\rho_{2s}\sqrt{P_2})$, $\alpha = \frac{P'_2}{P'_2+1}$.

Similarly to the Gaussian CIC-STR, due to the fact that $|a| > 1$, i.e., receiver 1 is stronger in decoding W_2 , the achievable scheme sets $W_{21} = W_2$, i.e., requires receiver 1 to decode full message W_2 . The cognitive transmitter's power P_2 is split into two parts: 1.cooperatively transmitting W_1 via beamforming, 2.transmitting additional W_1 and $W_{21} = W_2$ via dirty paper coding. Differently from the CIC-STR, the auxiliary random variable U is used here to deal with the state interference at receiver 2 (instead of receiver 1 for the CIC-STR). This is also due to the fact that $|a| > 1$ so that receiver 2 is weaker in decoding information from the cognitive transmitter and hence needs additional help in state cancellation via dirty paper coding than receiver 1. Therefore, in the above achievable region, (3.26a) reflects the fact that receiver 2 decodes $W_{21} = W_2$, and (3.26b) and (3.26c) respectively reflect the facts that receiver 2 and receiver 1 decode both W_1 and $W_{21} = W_2$.

Proof. By setting $T = X_1$ and $U = V$ in the inner bound given in Theorem 3.4, we obtain an inner bound that includes the following bounds:

$$R_2 \leq I(U; Y_2 | X_1) - I(U; S | X_1) \quad (3.27a)$$

$$R_1 + R_2 \leq I(X_1 U; Y_2) - I(U; S | X_1) \quad (3.27b)$$

$$R_1 + R_2 \leq I(X_1 U; Y_1) - I(U; S | X_1) . \quad (3.27c)$$

Based on the above bounds, we choose the jointly Gaussian input distribution and employ dirty paper coding for U to deal with the state in Y_2 . More specifically, we set the

random variables as follows and obtain the desired inner bound:

$$\begin{aligned}
X_1 &\sim \mathcal{N}(0, P_1), \quad X'_2 \sim \mathcal{N}(0, P'_2) \\
X_2 &= \rho_{21} \sqrt{\frac{P_2}{P_1}} X_1 + X'_2 + \rho_{2s} \sqrt{\frac{P_2}{Q}} S \\
U &= X'_2 + \alpha \left(c + \rho_{2s} \sqrt{\frac{P_2}{Q}} \right) S
\end{aligned} \tag{3.28}$$

where X_1 , X'_2 and S are independent random variables, and $\alpha = \frac{P'_2}{P'_2+1}$. \square

We next provide an outer bound on the capacity region based on the following idea.

Since both W_1 and W_2 must be decoded at receiver 2, the two transmitters and receiver 2 form a cognitive MAC with state known at the cognitive transmitter. Hence, the capacity region for such a MAC serves as an outer bound for the Gaussian CIC-ST.

Proposition 3.2. *For the Gaussian CIC-ST, if $|a| > 1$, an outer bound consists of rate pairs (R_1, R_2) satisfying:*

$$R_2 \leq \frac{1}{2} \log(1 + P'_2) \tag{3.29a}$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) + \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \tag{3.29b}$$

where $P'_2 \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$ and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

Proof. It is clear that the outer bound in Proposition 3.2 is equivalent to the region that consists of rate pairs (R_1, R_2) satisfying:

$$R_2 \leq \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \tag{3.30a}$$

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) \\
&\quad + \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2)
\end{aligned} \tag{3.30b}$$

where $\rho_{21}^2 + \rho_{2s}^2 \leq 1$. This region is the capacity region of the MAC with state (with its receiver being receiver 2 in our model) given in [18], and hence serves as an outer bound for our model. \square

Although the inner bound (3.26a)-(3.26c) and the outer bound (3.29a)-(3.29b) do not match in general, we show that these bounds characterize some boundary points of the capacity region. In order to characterize boundary points of the capacity region, we first change the inner bound (3.26a)-(3.26c) into a more convenient form, which consists of rate pairs (R_1, R_2) satisfying

$$R_2 \leq \frac{1}{2} \log(1 + P'_2) \quad (3.31a)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) + \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \quad (3.31b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1 P_2} + a^2 \rho_{21}^2 P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2 Q} + Q + 1} \right) + \frac{1}{2} \log \left(1 + \frac{(a^2(1 - \rho_{21}^2 - \rho_{2s}^2)P_2 + 2a\rho_{2s1}\rho_{2s2} - a^2 \rho_{2s1}^2)(1 - \rho_{21}^2 - \rho_{2s}^2)P_2 - \rho_{2s1}^2}{(a^2 \rho_{2s1}^2 + \rho_{2s2}^2 + 1 - 2a\rho_{2s1}\rho_{2s2})(1 - \rho_{21}^2 - \rho_{2s}^2)P_2 + \rho_{2s1}^2} \right) \quad (3.31c)$$

where $P'_2 \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $\rho_{21}^2 + \rho_{2s}^2 \leq 1$, $\rho_{2s1} = \alpha(c\sqrt{Q} + \rho_{2s}\sqrt{P_2})$, $\rho_{2s2} = (\sqrt{Q} + a\rho_{2s}\sqrt{P_2})$, and $\alpha = \frac{(1 - \rho_{21}^2 - \rho_{2s}^2)P_2}{(1 - \rho_{21}^2 - \rho_{2s}^2)P_2 + 1}$. Such a region is equivalent to (3.26a)-(3.26c), because it is obtained by substituting the equality constraint $P'_2 = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$ into (3.26a) and (3.26b) (which does not change the bounds), and relaxing the constraint on P'_2 to be $P'_2 \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, which affects only (3.26a) and clearly does not enlarge the region. We now denote the bounds in (3.31a)-(3.31c) by $r_2(P'_2)$, $r_{12}(\rho_{21}, \rho_{2s})$, and $\tilde{r}_{12}(\rho_{21}, \rho_{2s})$. For $0 \leq P'_2 \leq P_2$, let

$$(\rho_{21}^*(P'_2), \rho_{2s}^*(P'_2)) = \underset{(\rho_{21}, \rho_{2s}): P'_2 \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2}{\operatorname{argmax}} r_{12}(\rho_{21}, \rho_{2s}). \quad (3.32)$$

Based on these notations, we characterize partial boundary of the capacity region for the Gaussian channel as follows.

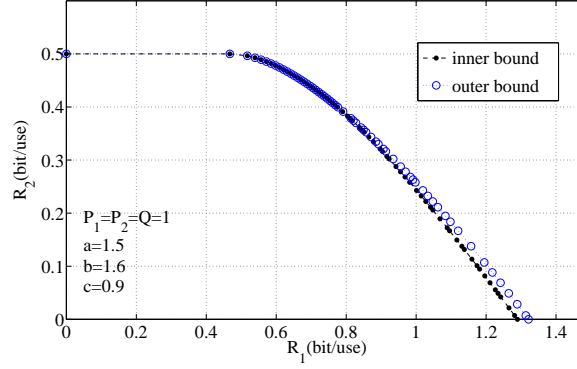


Fig. 3.2: An illustration of the partial boundary of the capacity region for a Gaussian CIC-ST with $|a| > 1$.

Theorem 3.9. *Consider the Gaussian CIC-ST with $|a| > 1$. For $0 \leq P'_2 \leq P_2$, the rate pairs $\left(r_{12}(\rho_{21}^*(P'_2), \rho_{2s}^*(P'_2)) - r_2(P'_2), r_2(P'_2) \right)$ are on the boundary of the capacity region if $r_{12}(\rho_{21}^*(P'_2), \rho_{2s}^*(P'_2)) \leq \tilde{r}_{12}(\rho_{21}^*(P'_2), \rho_{2s}^*(P'_2))$. The rate pairs $(R_1, r_2(P_2))$ are also on the boundary of the capacity region if $R_1 \leq \min\{r_{12}, \tilde{r}_{12}\}|_{\rho_{21}=0, \rho_{2s}=0} - r_2(P_2)$.*

Proof. The rate pairs given in the theorem are achievable due to the condition given in the theorem. They are also on the boundary of the outer bound given in Proposition 3.2, because r_2 and r_{12} are the same as the bounds on R_1 and on $R_1 + R_2$, respectively, and the chosen parameters $(\rho_{21}^*(P'_2), \rho_{2s}^*(P'_2))$ for each value of P'_2 guarantees that the rate pairs are on the boundary. The second statement is clear because when $P'_2 = P_2$, R_2 achieves the maximum value, and hence any such rate pair is on the boundary if it is achievable. \square

In Fig. 3.2, we demonstrate the partial boundary of the capacity region characterized in Theorem 3.9. We consider the channel defined by the parameters $P_1 = P_2 = Q = 1$, $a = 1.5$, $b = 1.6$ and $c = 0.9$. We plot the boundaries of the inner bound given in Proposition 3.1 and the outer bound given in Proposition 3.2, respectively. It is clear that the two boundaries match when R_2 is above a certain threshold, and this matching part thus characterizes some boundary points of the capacity region as studied in Theorem 3.9.

We next show that under certain channel conditions, the outer bound given in Proposition 3.2 fully characterizes the capacity region.

Theorem 3.10. *For the Gaussian CIC-ST, if $|a| > 1$ and the channel satisfies the condition (3.5), the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_2 \leq \frac{1}{2} \log(1 + P'_2) \quad (3.33a)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) + \frac{1}{2} \log(1 + P'_2) \quad (3.33b)$$

where $P'_2 = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$ and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

As explained after Theorem 3.8, if the less noisy condition (3.5) is satisfied, receiver 2 dominates the performance of the channel. Thus, the achievable scheme that uses the auxiliary random variable for dealing with the state at receiver 2 via dirty paper coding turns out to be optimal.

Proof. Following from the region in Theorem 3.8, we set the random variables as in (3.28) and obtain an achievable region as given in (3.33a)-(3.33b). Such an achievable region is equivalent to the outer bound given in Proposition 3.2 as we comment in the proof of Proposition 3.2. \square

3.3.2.2 Gaussian Channel: $|a| \leq 1$

We first note that the inner bound given in Proposition 3.1 for the case when $|a| > 1$ also serves as an inner bound for the case when $|a| \leq 1$. However, the choice of auxiliary random variables ($T = \phi$ and $U = V$) for obtaining this inner bound requires receiver 1 to decode all information for receiver 2. As such, this bound works well only when receiver 1 is stronger than receiver 2, and does not serve as a good inner bound for the case when $|a| \leq 1$. Thus, in this subsection, we develop two new inner bounds and one new outer

bound on the capacity region for the case when $|a| \leq 1$. We also note that the outer bound in Proposition 3.2 is applicable and useful here as demonstrated in the sequel.

We derive the two inner bounds based on the same achievable region for the discrete memoryless channel with different choices of the distributions for the auxiliary random variables.

Proposition 3.3. *For the Gaussian CIC-ST, if $|a| \leq 1$, then an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) + \frac{1}{2} \log \left(1 + \frac{a^2P_2'}{a^2P_2'' + 1} \right) \quad (3.34a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2'') \quad (3.34b)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{a^2P_2'^2 + 2a\rho_{2s1}\rho_{2s2}P_2' - \rho_{2s1}^2(P_2' + P_2'' + 1)}{a^2P_2'P_2'' + \rho_{2s1}^2(P_2' + P_2'' + 1) + a^2\rho_{2s2}P_2' + a^2P_2' - 2a\rho_{2s1}\rho_{2s2}P_2'} \right) + \frac{1}{2} \log(1 + P_2'') \quad (3.34c)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2P_1 + 2b\rho_{21}\sqrt{P_1P_2} + \rho_{21}^2P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2Q} + c^2Q + 1} \right) + \frac{1}{2} \log \left(1 + \frac{a^2P_2'^2 + 2a\rho_{2s1}\rho_{2s2}P_2' - \rho_{2s1}^2(P_2' + P_2'' + 1)}{a^2P_2'P_2'' + \rho_{2s1}^2(P_2' + P_2'' + 1) + a^2\rho_{2s2}P_2' + a^2P_2' - 2a\rho_{2s1}\rho_{2s2}P_2'} \right) + \frac{1}{2} \log(1 + P_2'') \quad (3.34d)$$

where $\rho_{2s1} = \alpha \left(1 + a\rho_{2s}\sqrt{\frac{P_2}{Q}} \right) \sqrt{Q}$, $\rho_{2s2} = \left(c + \rho_{2s}\sqrt{\frac{P_2}{Q}} \right) \sqrt{Q}$, $\alpha = \frac{a^2P_2'}{a^2P_2' + a^2P_2'' + 1}$, $|\rho_{21}| \leq 1$, $|\rho_{2s}| \leq 1$, $P_2' \geq 0$, $P_2'' \geq 0$, and $P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$.

Similarly to the CIC-STR, if $|a| \leq 1$, the cognitive transmitter's power P_2 is split into three parts: 1.cooperatively transmitting W_1 via beamforming, 2. $P_2' + \rho_{2s}^2P_2$ are for either transmitting additional W_1 or transmitting W_2 using dirty paper coding (via T) to deal with the state at receiver 1, 3.transmitting W_2 using dirty paper coding (via V) to deal with the state at receiver 2. Therefore, in the above achievable region, (3.34a) reflects the fact that receiver 1 decodes W_1 contained in both X_1 and T , (3.34b) reflects the fact that receiver 2

decodes W_2 contained in V , (3.34c) reflects the fact that receiver 2 decodes W_2 contained in both T and V , and (3.34d) reflects the fact that receiver 2 decodes W_1 contained in X_1 , and W_2 contained in both T and V . We note that T plays two roles: either transmitting W_1 or transmitting W_2 .

Proof. The above theorem is based on Corollary 3.1 by choosing (T, V, X_1, X_2) to be jointly Gaussian and employing dirty paper coding with T chosen for dealing with the state for Y_1 and V chosen for dealing with the state for Y_2 . More specifically, We set the random variables as follows:

$$\begin{aligned}
X_1 &\sim \mathcal{N}(0, P_1), \quad X'_2 \sim \mathcal{N}(0, P'_2), \quad X''_2 \sim \mathcal{N}(0, P''_2), \\
P'_2 + P''_2 &= (1 - \rho_{21}^2 - \rho_{2s}^2)P_2 \\
X_2 &= \rho_{21}\sqrt{\frac{P_2}{P_1}}X_1 + X'_2 + X''_2 + \rho_{2s}\sqrt{\frac{P_2}{Q}}S \\
T &= X'_2 + \alpha \left(1 + a\rho_{2s}\sqrt{\frac{P_2}{Q}} \right) S \\
V &= X''_2 + \beta \left(c - \alpha + (1 - a\alpha)\rho_{2s}\sqrt{\frac{P_2}{Q}} \right) S
\end{aligned} \tag{3.35}$$

where X_1, X'_2, X''_2 and S are independent random variables, $\alpha = \frac{a^2 P'_2}{a^2 P'_2 + a^2 P''_2 + 1}$, and $\beta = \frac{P''_2}{P''_2 + 1}$. \square

Proposition 3.4. *For the Gaussian CIC-ST, if $|a| \leq 1$, then an inner bound on the capacity*

region consists of rate pairs (R_1, R_2) satisfying

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) + \frac{1}{2} \log \left(1 + \frac{a^2P_2'^2 + 2a\rho_{2s1}\rho_{2s2}P_2' - a^2\rho_{2s1}^2(P_2' + P_2'') - \rho_{2s1}^2}{a^2\rho_{2s1}^2P_2' + \rho_{2s2}^2P_2' + a^2\rho_{2s1}^2P_2'' + a^2P_2'P_2'' + P_2' + \rho_{2s1}^2 - 2a\rho_{2s1}\rho_{2s2}P_2'} \right) \quad (3.36a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2'') \quad (3.36b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{b^2P_1 + 2b\rho_{21}\sqrt{P_1P_2} + \rho_{21}^2P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2Q} + c^2Q + 1} \right) + \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \quad (3.36c)$$

where $\rho_{2s1} = \alpha(c\sqrt{Q} + \rho_{2s}\sqrt{P_2})$, $\rho_{2s2} = (\sqrt{Q} + a\rho_{2s}\sqrt{P_2})$, $\alpha = \frac{P_2'}{P_2' + P_2'' + 1}$, $|\rho_{21}| \leq 1$, $|\rho_{2s}| \leq 1$, $P_2' \geq 0$, $P_2'' \geq 0$, and $P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$.

We note that the inner bounds in Proposition 3.3 and 3.4 are based on the same achievable region for the discrete memoryless channel, i.e., Corollary 3.1, except that the auxiliary random variable T is designed to deal with the state at receiver 2 in Proposition 3.4.

Proof. The above theorem is based on Corollary 3.1 by choosing (T, V, X_1, X_2) to be jointly Gaussian and employing dirty paper coding by choosing T and V as follows:

$$\begin{aligned} X_1 &\sim \mathcal{N}(0, P_1), \quad X_2' \sim \mathcal{N}(0, P_2'), \quad X_2'' \sim \mathcal{N}(0, P_2''), \\ P_2' + P_2'' &= (1 - \rho_{21}^2 - \rho_{2s}^2)P_2 \\ X_2 &= \rho_{21}\sqrt{\frac{P_2}{P_1}}X_1 + X_2' + X_2'' + \rho_{2s}\sqrt{\frac{P_2}{Q}}S \\ T &= X_2' + \alpha \left(c + \rho_{2s}\sqrt{\frac{P_2}{Q}} \right) S \\ V &= X_2'' + \beta(1 - \alpha) \left(c + \rho_{2s}\sqrt{\frac{P_2}{Q}} \right) S \end{aligned} \quad (3.37)$$

where X_1, X_2', X_2'' and S are independent random variables, $\alpha = \frac{P_2'}{P_2' + P_2'' + 1}$, and $\beta = \frac{P_2''}{P_2'' + 1}$.

Here, T is chosen for dealing with the state for Y_2 (differently from the proof for Proposition 3.3) based on dirty paper coding where X_2'' is taken as noise. We then subtract T from Y_2 and design V for dealing with the state for $Y_2' = Y_2 - T$ based on dirty paper coding. For this choice of the auxiliary random variables, the second bound on R_2 in Corollary 3.1 is redundant because $I(T; Y_2 | X_1) - I(T; S | X_1) > 0$. \square

We next provide two outer bounds, both of which are useful for characterizing the capacity results. The first outer bound is given by the capacity region of the Gaussian CIC-STR that we present in Theorem 3.2 in Section 3.2.2. For convenience, we rewrite this bound below.

Corollary 3.2. *For the Gaussian CIC-ST, if $|a| \leq 1$, then the capacity region of CIC-STR serves as an outer bound on the capacity region, which consists of rate pairs (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) \\ &\quad + \frac{1}{2} \log \left(1 + \frac{a^2P_2'}{a^2P_2'' + 1} \right) \\ R_2 &\leq \frac{1}{2} \log(1 + P_2'') \\ R_1 + R_2 &\leq \frac{1}{2} \log(1 + b^2P_1 + 2b\rho_{21}\sqrt{P_1P_2} + (1 - \rho_{2s}^2)P_2) \end{aligned}$$

where $P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $P_2' \geq 0$, $P_2'' \geq 0$, and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

As we comment at the beginning of this subsection, the outer bound in Proposition 3.2 is also applicable and useful for the case with $|a| \leq 1$. For convenience, we rewrite it below as a corollary.

Corollary 3.3. *For the Gaussian CIC-ST, if $|a| \leq 1$, an outer bound on the capacity region*

consists of rate pairs (R_1, R_2) satisfying:

$$\begin{aligned} R_2 &\leq \frac{1}{2} \log(1 + P_2'') \\ R_1 + R_2 &\leq \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right) \\ &\quad + \frac{1}{2} \log(1 + (1 - \rho_{21}^2 - \rho_{2s}^2)P_2) \end{aligned}$$

where $P_2'' \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $P_2'' \geq 0$, and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$.

For Gaussian channels with $|a| \leq 1$, we characterize partial boundaries of the capacity region based on the inner and outer bounds respectively given in Proposition 3.3 and 3.4, and Corollaries 3.2 and 3.3. Although the forms of inner bounds are complicated, we show that some boundary points on the capacity region are determined only by a subset of there bounds, and can hence be characterized via the given outer bounds.

We let $\Delta = (\rho_{21}, \rho_{2s}, P_2')$ and use $r_1'(\Delta, P_2'')$, $r_2'(P_2'')$, $\tilde{r}_2'(\Delta, P_2'')$, $r_{12}'(\Delta, P_2'')$ to denote the bounds (3.34a)-(3.34d) given in Proposition 3.3. For $0 \leq P_2'' \leq P_2$, let

$$\Delta^*(P_2'') = \underset{\Delta: P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2}{\operatorname{argmax}} r_1'(\Delta, P_2''). \quad (3.40)$$

Based on these notations, we characterize partial boundary of the capacity region for the Gaussian channel as follows.

Theorem 3.11. *Consider the Gaussian CIC-ST with $|a| \leq 1$. For $0 \leq P_2'' \leq P_2$, the rate pairs $(r_1'(\Delta^*(P_2''), P_2''), r_2'(P_2''))$ are on the boundary of the capacity region if $r_2'(P_2'') \leq \tilde{r}_2'(\Delta^*(P_2''), P_2'')$ and $r_1'(\Delta^*(P_2''), P_2'') + r_2'(P_2'') \leq r_{12}'(\Delta^*(P_2''), P_2'')$.*

Proof. The rate pairs given in the theorem are contained in inner bound 1 given in Proposition 3.3 due to the conditions given in the theorem. We next show that these rate pairs are also on the boundary of an outer bound. Following from outer bound 1 in Corollary 3.2, $R_1 \leq r_1'(\Delta, P_2'')$ and $R_2 \leq r_2'(P_2'')$ also determine an outer bound with (Δ, P_2'') taking the same values as in inner bound 1 given in Proposition 3.3. Then the chosen parameters

$\Delta^*(P_2'')$ for each value of P_2'' guarantee that the rate pairs are on the boundary of this outer bound. \square

We next characterize additional boundary points of the capacity region based on inner bound 2 given in Proposition 3.4 and outer bound 2 given in Corollary 3.3. We use $r_1''(\rho_{21}, \rho_{2s}, P_2', P_2'')$, $r_2''(P_2'')$, and $r_{12}''(\rho_{21}, \rho_{2s})$ to denote the bounds (3.36a)-(3.36c) given in Proposition 3.4. For $0 \leq P_2'' \leq P_2$, let

$$(\rho_{21}^*(P_2''), \rho_{2s}^*(P_2'')) = \underset{(\rho_{21}, \rho_{2s}): P_2'' \leq (1 - \rho_{21}^2 - \rho_{2s}^2)P_2}{\operatorname{argmax}} r_{12}''(\rho_{21}, \rho_{2s}), \quad (3.41)$$

and let $P_2'^*(P_2'') = (1 - \rho_{21}^*(P_2'')^2 - \rho_{2s}^*(P_2'')^2)P_2 - P_2''$. Based on these notations, we characterize partial boundary of the capacity region as follows.

Theorem 3.12. *Consider the Gaussian CIC-ST with $|a| \leq 1$. For $0 \leq P_2'' \leq P_2$, the rate pairs $(r_{12}''(\rho_{21}^*(P_2''), \rho_{2s}^*(P_2'')) - r_2''(P_2''), r_2''(P_2''))$ are on the boundary of the capacity region if $r_{12}''(\rho_{21}^*(P_2''), \rho_{2s}^*(P_2'')) - r_2''(P_2'') \leq r_1''(\rho_{21}^*(P_2''), \rho_{2s}^*(P_2''), P_2'^*(P_2''), P_2'')$. The rate pairs $(R_1, r_2''(P_2''))$ are also on the boundary of the capacity region if $R_1 \leq \min\{r_1'', r_{12}'' - r_2''(P_2'')\}|_{\rho_{21}=0, \rho_{2s}=0, P_2'=0}$.*

Proof. The rate pairs given in the theorem are clearly contained in inner bound 2 given in Proposition 3.4. These rate pairs are also on the boundary of outer bound 2 given in Corollary 3.3, because r_2'' and r_{12}'' are the same as the bounds on R_2 and on $R_1 + R_2$, respectively, and the chosen parameters $(\rho_{21}^*(P_2''), \rho_{2s}^*(P_2''))$ for each value of P_2'' guarantee that the rate pairs are on the boundary. The second statement is clear because when $P_2'' = P_2$, R_2 achieves the maximum value, and hence any rate pair with such R_2 is on the boundary if it is achievable. \square

Theorems 3.11 and 3.12 collectively characterize partial boundary of the capacity region for the Gaussian channel with $|a| \leq 1$. In Fig. 3.3, we demonstrate these boundary points of the capacity region for an example channel with the parameters $P_1 = P_2 = Q =$

1, $b = 0.85$, $c = 0.9$ and $a = 0.8$. We plot the boundaries of the two inner bounds given in Proposition 3.3 and Proposition 3.4, and the boundaries of the two outer bounds given in Corollary 3.2 and Corollary 3.3, respectively. It can be seen that the two inner bounds are very close to each other. The boundary of inner bound 1 matches the boundary of outer bound 1 when R_1 is above a certain value, and this part is thus on the boundary of the capacity region. We also note that this part of the boundary achieves the capacity region of the CIC-STR. It can further be seen that the boundary of inner bound 2 matches the boundary of outer bound 2 when R_2 is above a certain threshold, and this part is hence also on the boundary of the capacity region.

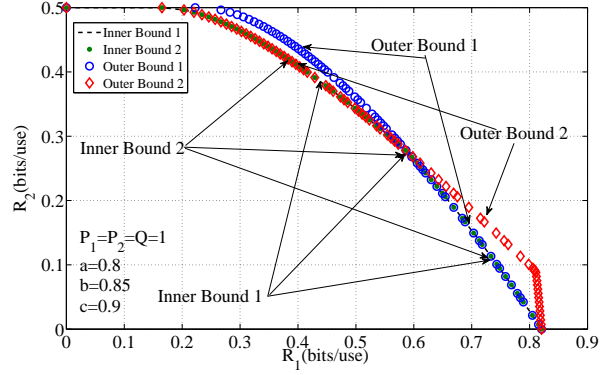


Fig. 3.3: An illustration of inner and outer bounds and the partial boundary of the capacity region for a Gaussian CIC-ST with $|a| \leq 1$

It can be seen that outer bounds 1 and 2 separately characterize certain parts of the boundary of the capacity region for Gaussian channels with $|a| \leq 1$. We further show that each of these two outer bounds can characterize the full capacity region for channels that satisfy certain conditions.

Theorem 3.13. *For the Gaussian CIC-ST, if $|a| \leq 1$ and the channel satisfies the condition*

(3.4), the capacity region consists of rate pairs (R_1, R_2) satisfying:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) \\ &\quad + \frac{1}{2} \log \left(1 + \frac{a^2P_2'}{a^2P_2'' + 1} \right) \\ R_2 &\leq \frac{1}{2} \log(P_2'' + 1) \end{aligned}$$

where $P_2' + P_2'' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $P_2' \geq 0$, $P_2'' \geq 0$ and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$, $|\rho_{21}| \leq 1$, $|\rho_{2s}| \leq 1$.

We note that the above capacity region matches the capacity region in [24] of a cognitive interference model with state, in which W_1 is intended only for receiver 1. This is reasonable because under the condition (3.4), receiver 1 is weaker in decoding W_1 than receiver 2, and receiver 2 can hence always decode W_1 , which satisfies the additional requirement in the channel model. Consequently, in the designation of auxiliary random variables, more resources are used to help receiver 1 to cancel signal and state interference. This is why only part of P_2 is used to transmit W_2 , and there is a tradeoff between the rates R_1 and R_2 .

Proof. Under the condition (3.4), the bounds in the achievable region in Corollary 3.1 reduce to:

$$R_1 \leq I(X_1T; Y_1) - I(T; S|X_1) \quad (3.43a)$$

$$R_2 \leq I(V; Y_2|X_1T) - I(V; S|X_1T) \quad (3.43b)$$

$$R_2 \leq I(TV; Y_2|X_1) - I(TV; S|X_1) \quad (3.43c)$$

Based on the above bounds, we choose the same jointly Gaussian input distribution as in (3.35). In particular, since the auxiliary random variable T is chosen to employ dirty paper coding to deal with the state in Y_1 , it guarantees that $I(T; Y_1|X_1) - I(T; S|X_1) \geq 0$, which implies that $I(T; Y_2|X_1) - I(T; S|X_1) \geq 0$ due to the condition (3.4). Hence, (3.43c) is

redundant. Thus, we obtain an achievable region that matches the first two bounds of outer bound 1 in Corollary 3.2 and is hence tight. \square

The following theorem identifies the channels for which outer bound 2 given in Corollary 3.3 characterizes the full capacity region.

Theorem 3.14. *For the Gaussian CIC-ST, if $|a| \leq 1$ and the channel satisfies the condition (3.5), the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P'_2}{P''_2 + 1} \right) \quad (3.44a)$$

$$+ \frac{1}{2} \log \left(1 + \frac{b^2 P_1 + 2b\rho_{21}\sqrt{P_1 P_2} + \rho_{21}^2 P_2}{(1 - \rho_{21}^2)P_2 + 2c\rho_{2s}\sqrt{P_2 Q} + c^2 Q + 1} \right)$$

$$R_2 \leq \frac{1}{2} \log(1 + P''_2) \quad (3.44b)$$

where $P'_2 + P''_2 = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2$, $P'_2 \geq 0$, $P''_2 \geq 0$ and $\rho_{21}^2 + \rho_{2s}^2 \leq 1$, $|\rho_{21}| \leq 1$, $|\rho_{2s}| \leq 1$.

Proof. With the condition (3.5), it can be seen that an achievable region determined by the following bounds is contained in the inner bound given in Corollary 3.1, and is hence achievable.

$$R_1 \leq I(X_1 T; Y_2) - I(T; S|X_1) \quad (3.45a)$$

$$R_2 \leq I(V; Y_2|X_1 T) - I(V; S|X_1 T) \quad (3.45b)$$

$$R_2 \leq I(TV; Y_2|X_1) - I(TV; S|X_1) . \quad (3.45c)$$

The achievability follows from the above region by choosing the jointly Gaussian distribution and employing dirty paper coding for T to deal with the state for Y_2 and for V to deal with the remaining state for Y_2 after subtracting $\frac{1}{a}T$. More specifically, we set the

auxiliary random variable as follows:

$$\begin{aligned}
X_1 &\sim \mathcal{N}(0, P_1), \quad X'_2 \sim \mathcal{N}(0, P'_2), \quad X''_2 \sim \mathcal{N}(0, P''_2), \\
P'_2 + P''_2 &= (1 - \rho_{21}^2 - \rho_{2S}^2) \\
X_2 &= \rho_{21} \sqrt{\frac{P_2}{P_1}} X_1 + X'_2 + X''_2 + \rho_{2S} \sqrt{\frac{P_2}{Q}} S \\
T &= X'_2 + \alpha \left(c + \rho_{2S} \sqrt{\frac{P_2}{Q}} \right) S \\
V &= X''_2 + \beta (1 - \alpha) \left(c + \rho_{2S} \sqrt{\frac{P_2}{Q}} \right) S
\end{aligned} \tag{3.46}$$

where X_1 , X'_2 , X''_2 and S are independent random variables, $\alpha = \frac{P'_2}{P'_2 + P''_2 + 1}$, and $\beta = \frac{P''_2}{P''_2 + 1}$. Such a choice of the input distribution also implies that $I(T; Y_2 | X_1) - I(T; S | X_1) \geq 0$, and the bound (3.45c) is hence redundant. The proof for the converse follows by observing that the region (3.44a)-(3.44b) has the same boundary points as outer bound 2 given in Corollary 3.3, and hence the two regions are equivalent. \square

CHAPTER 4

STATE-DEPENDENT SINGLE-USER

CHANNEL WITH A HELPER

In this chapter, we study the state-dependent Gaussian single-user channel with a helper. In the previous work [34], the capacity in the regime of infinite state power is characterized based on lattice coding. In this thesis, we focus on the regime with general state-power. We design an achievable scheme combining two methods to cancel state: 1. precoding the state with a single bin scheme; 2. directly reversing the state. By comparing the lower bound derived based on the above scheme and the upper bound from the previous work, we characterize the capacity of the channel under various channel parameters.

4.1 Channel Model

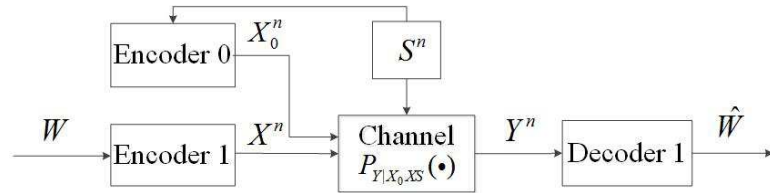


Fig. 4.1: The state-dependent single-user channel with a helper

In the state-dependent single-user channel with a helper (see Fig. 1.4 in Section 1.2. For convenience of reference, we include the figure again as Fig. 4.1 in this section), a transmitter wishes to send message W to a receiver over a state-corrupted channel, and a helper that knows the state information noncausally wishes to assist the receiver to cancel state interference.

More specifically, the transmitter has an encoder $f : \mathcal{W} \rightarrow \mathcal{X}^n$, which maps the message $w \in \mathcal{W}$ to a codeword $x^n \in \mathcal{X}^n$. The input x^n is transmitted over the channel. The receiver is interfered by an i.i.d. state sequence S^n , which is known at neither the transmitter nor the receiver, but at a helper noncausally. Thus, the encoder at the helper, $f_0 : \mathcal{S}^n \rightarrow \mathcal{X}_0^n$, maps the state sequences $s^n \in \mathcal{S}^n$ to a codeword $x_0^n \in \mathcal{X}_0^n$. The entire channel transition probability is given by $P_{Y|X_0, X, S}$. The decoder at the receiver, $g : \mathcal{Y}^n \rightarrow \mathcal{W}$, maps a received sequence y^n into a message $\hat{w} \in \mathcal{W}$.

We study the Gaussian channel model with the following output at the receiver for one symbol time:

$$Y = X_0 + X + S + N \quad (4.1)$$

where the noise variable N and the state variable S are Gaussian distributed with distributions $N \sim \mathcal{N}(0, 1)$ and $S \sim \mathcal{N}(0, Q)$, and all of the variables are independent and are i.i.d. over channel uses. The channel inputs X_0 , and X are subject to the average power constraints $\frac{1}{n} \sum_{i=1}^n X_{0i}^2 \leq P_0$ and $\frac{1}{n} \sum_{i=1}^n X_i^2 \leq P$.

4.2 Achievable Scheme and Lower Bound

In this section, we design an achievable scheme for the state-dependent Gaussian single-user channel with a helper. Two basic ideas to cancel the state are integrated together: 1. reversing the channel state directly; 2. precoding state into a help signal based on a single bin scheme. In [34], the focus of the design is on the regime of infinite state power. Hence,

only precoding the state is utilized, because it is impossible to reverse the infinite-power state directly. And the capacity result in [34] suggests that, precoding the state is capacity achieving for the infinite state power regime. However, for the regime with finite state power, it is useful to apply both methods as we demonstrate in our study. By integrating state reversion with single-bin scheme, we obtain the following achievable rate.

Proposition 4.1. *For the state-dependent Gaussian single-user channel with a helper, a rate R is achievable if it satisfies*

$$R \leq \min\{R_1(\alpha, \beta), R_2(\alpha, \beta)\}, \quad (4.2)$$

where

$$R_1(\alpha, \beta) = \frac{1}{2} \log\left(\frac{P'_0(P'_0 + (1 + \beta)^2 Q + P + 1)}{P'_0 Q(\alpha - 1 - \beta)^2 + P'_0 + \alpha^2 Q}\right) \quad (4.3a)$$

$$R_2(\alpha, \beta) = \frac{1}{2} \log\left(1 + \frac{P(P'_0 + \alpha^2 Q)}{P'_0 Q(\alpha - 1 - \beta)^2 + P'_0 + \alpha^2 Q}\right) \quad (4.3b)$$

for some (α, β, P'_0) such that $P'_0 + \beta^2 Q \leq P_0$.

Proof. We first derive an achievable rate in the following lemma for the discrete memoryless state-dependent single-user channel with a helper based on Proposition 5.2 by setting $X'_0 = \phi$.

Lemma 4.1. *For the discrete memoryless state-dependent single-user channel with a helper, a rate R is achievable if it satisfies*

$$R \leq I(UX; Y) - I(U; S) \quad (4.4a)$$

$$R \leq I(X; Y|U) \quad (4.4b)$$

for some distribution $P_S P_{X_0|S} P_X P_{Y|SX_0X}$.

Proposition 4.1 then follows from Lemma 4.1, by choosing the joint Gaussian distribution for the random variables as follows:

$$X_0 = X'_0 + \beta S$$

$$U = X'_0 + \alpha S$$

where $X'_0 \sim \mathcal{N}(0, P'_0)$, and $-\sqrt{\frac{P_0 - P'_0}{Q}} \leq \beta \leq \sqrt{\frac{P_0 - P'_0}{Q}}$. \square

We note that the achievable rate in Proposition 4.1 is optimized over α and β . It is clear that the optimization is a max-min problem, i.e., maximization of minimum of $R_1(\alpha, \beta)$ and $R_2(\alpha, \beta)$. In general, such optimization cannot be solved analytically with close form expressions. In order to obtain further insights of such a lower bound, we consider two special cases in which the optimization is solved analytically and the corresponding achievable rate turns out to achieve the capacity. The idea is to optimize $R_1(\alpha, \beta)$ and $R_2(\alpha, \beta)$ separately. For example, when $R_1(\alpha, \beta)$ is optimized, if $R_2(\alpha, \beta)$ at the optimizing values of α and β is greater than the optimal $R_1(\alpha, \beta)$, then the corresponding optimal $R_1(\alpha, \beta)$ is achievable. The same argument is applicable to optimize $R_2(\alpha, \beta)$ instead. Such an idea yield the following two corollaries on the achievable rate.

Corollary 4.1. *For the state-dependent Gaussian single-user channel with a helper, a rate R is achievable if it satisfies*

$$R \leq \frac{1}{2} \log\left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1}\right) + \frac{1}{2} \log(1 + P_0 - \rho_{0S}^2 P_0) \quad (4.5a)$$

$$R \leq \frac{1}{2} \log\left(1 + \frac{P((1 + P_0(1 - \rho_{0S}^2))^2 + (1 - \rho_{0S}^2)P_0(\sqrt{Q} + \rho_{0S}\sqrt{P_0})^2)}{(Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1)(1 + P_0 - \rho_{0S}^2 P_0)}\right), \quad (4.5b)$$

for some ρ_{0S} such that $-1 \leq \rho_{0S} \leq 1$.

Proof. It can be shown that $R_1(\alpha, \beta)$ is optimized by $\alpha = \frac{(1+\beta)P'_0}{P'_0+1}$. We then set $\beta = \rho_{0S}\sqrt{\frac{P_0}{Q}}$ to better illustrate the result, where $\rho_{0S} = \frac{E[X_0S]}{\sqrt{P_0Q}}$, and $-1 \leq \rho_{0S} \leq 1$. Corollary 4.1 then follows by substituting α and β into (4.3a) and (4.3b). \square

Corollary 4.2. *For the state-dependent Gaussian single-user channel with a helper, a rate R is achievable if it satisfies*

$$R \leq \frac{1}{2} \log \frac{P_0(P_0 + Q + P + 1)}{P_0 + Q}, \quad (4.6a)$$

$$R \leq \frac{1}{2} \log(1 + P). \quad (4.6b)$$

Proof. It can be shown that $R_2(\alpha, \beta)$ is optimized by setting $\alpha = 1$ and $\beta = 0$. Corollary 4.2 then follows by substituting α and β into (4.3a) and (4.3b). \square

4.3 Capacity Results

In order to characterize the capacity, we first present two upper bounds on the capacity in the following lemma. The first bound is characterized in [34], and the second bound is the capacity of the corresponding channel without state corruption.

Lemma 4.2. *For the state-dependent Gaussian single-user channel with a helper, the capacity is upper bounded by*

$$C \leq \max_{-1 \leq \rho_{0S} \leq 1} \frac{1}{2} \log\left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1}\right) + \frac{1}{2} \log(1 + P_0 - \rho_{0S}^2 P_0) \quad (4.7a)$$

$$C \leq \frac{1}{2} \log(1 + P) \quad (4.7b)$$

By comparing the achievable rate in Corollaries 4.1 and 4.2, we obtain the capacity results in the following two theorems.

Theorem 4.1. *For the state-dependent single-user channel with a helper, define*

$$\rho_{0S}^* = \underset{-1 \leq \rho_{0S} \leq 1}{\operatorname{argmax}} \left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1}\right)(1 + P_0 - \rho_{0S}^2 P_0) \quad (4.8a)$$

$$R_1(\rho_{0S}) = \frac{1}{2} \log\left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1}\right) + \frac{1}{2} \log(1 + P_0 - \rho_{0S}^2 P_0) \quad (4.8b)$$

$$R_2(\rho_{0S}) = \frac{1}{2} \log\left(1 + \frac{P[(P_0 - \rho_{0S}^2 P_0 + 1)^2 + (1 - \rho_{0S}^2)P_0(\sqrt{Q} + \rho_{0S}\sqrt{P_0})^2]}{(\sqrt{Q} + \rho_{0S}\sqrt{P_0})^2 + (P_0 - \rho_{0S}^2 P_0 + 1)^2 + (1 - \rho_{0S}^2)P_0(\sqrt{Q} + \rho_{0S}\sqrt{P_0})^2}\right). \quad (4.8c)$$

If the channel parameters satisfy the following condition:

$$R_1(\rho_{0S}^*) \leq R_2(\rho_{0S}^*), \quad (4.9)$$

then the channel capacity is given by $\mathcal{C} = R_1(\rho_{0S}^)$.*

Proof. Based on the achievable rate in Corollary 4.1, the bound in (4.5a) is optimized for $\rho_{0S}^* = \underset{\rho_{0S}}{\operatorname{argmax}} \left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0Q} + P_0 + 1}\right)(1 + P_0 - \rho_{0S}^2 P_0)$. If $R_1(\rho_{0S}^*) \leq R_2(\rho_{0S}^*)$, then $R_1(\rho_{0S}^*)$ is achievable, which matches the upper bound (4.7a) in Lemma 6.1, and is hence the capacity of the channel. \square

Theorem 4.2. *For the state-dependent single-user channel with a helper, channel state can be fully cancelled, if the channel parameters satisfy the following condition:*

$$P_0^2 + P_0Q - Q(P + 1) \geq 0 \quad (4.10)$$

then the channel capacity is $\mathcal{C} = \frac{1}{2} \log(1 + P)$.

Proof. Based on the achievable rate in Corollary 4.2, when (4.10), the bound in (4.6a) is larger than the RHS of (4.6b), then the capacity of the channel without state corruption $\frac{1}{2} \log(1 + P)$ is achieved, which is thus the capacity of the state-dependent channel. \square

In Fig. 4.2, we plot the lower bounds in Corollary 4.1 and 4.2 and the upper bounds in Lemma 4.2 as a function of the helper's power P_0 , for the channel with $P = 5$, and

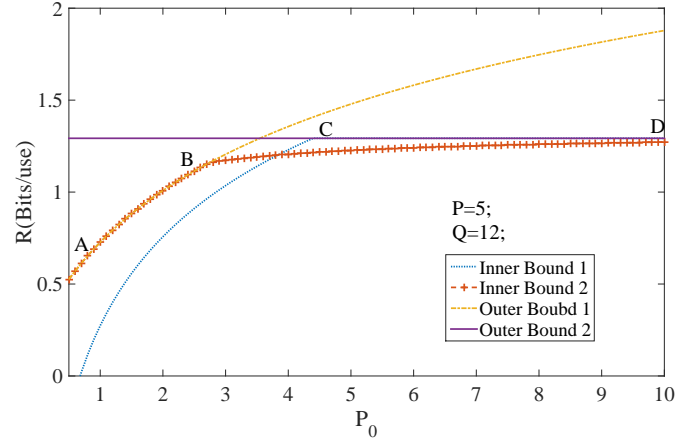


Fig. 4.2: Lower and upper bounds for the state-dependent single-user channel with a helper

$Q = 12$. The solid line and the dashed line are the two bounds in (4.7a) and (4.7b), and the dot line and the cross line are the two lower bounds in Corollary 4.1 and 4.2. Therefore, the points on the line A - B correspond to the capacity result in Theorem 4.1, and the points on the line C - D correspond to the capacity result in Theorem 4.2. The result suggests that when P_0 is small, the channel capacity is determined by a function of the helper's power P_0 and the state power Q . As P_0 becomes large enough, the channel capacity is determined only by the transmitter's power P , i.e., the state is perfectly canceled. We further note that the channel capacity without state can even be achieved by points with $P_0 < Q$ (i.e., some points on the line C - D). This indicates that for these points, the state are fully cancelled not only by reversing the state, but also by precoding the state.

In Fig. 4.3, we plot the lower bound in Proposition 4.1 (dashed line), the lower bound achieved by single bin scheme only (dashed-dot line), and the lower bound achieved by direct reversion only (solid line) as a function of the helper's power P_0 . It can be seen that the combination of the two methods provides larger achievable rate.

In Fig. 4.4, we plot the set of channel parameters (Q, P_0) for which our scheme achieves the capacity. Each point in the figure corresponds to the channel with certain P_0 and Q with fixed $P = 5$. The points in the upper part correspond to channel parameters that satisfy

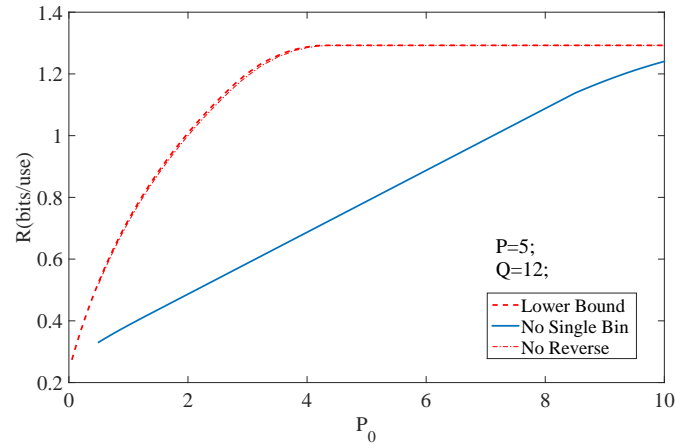


Fig. 4.3: Lower and upper bounds for the state-dependent single-user channel with a helper

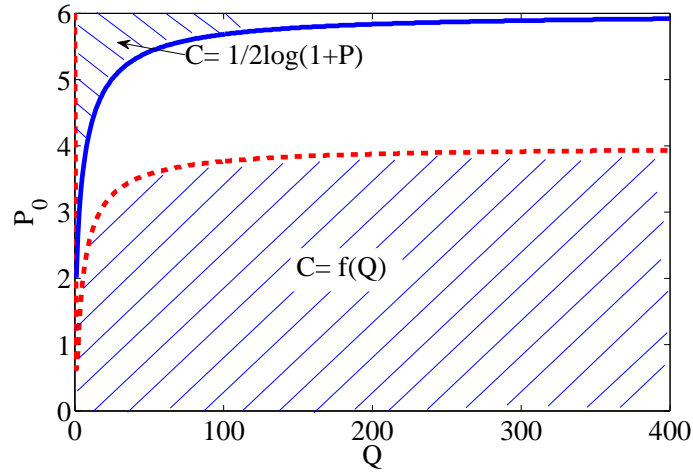


Fig. 4.4: Capacity achievable points

(4.10), and hence the capacity for single-user channel without state is obtained. The points in the lower part corresponds to channel parameters that satisfy (4.8a)-(4.8c), and hence the capacity is characterized by a function of not only P , but also P_0 and Q . As the state power Q goes to infinity, the result matches the result in [34].

CHAPTER 5

STATE-DEPENDENT PARALLEL

CHANNEL WITH A COMMON HELPER

In this chapter, we study the state-dependent parallel channel with a common helper. We consider three submodels for the channel. For each model, we derive inner and outer bounds. By comparing inner and outer bounds, we characterize the segments on the capacity boundary for the Gaussian channel with the state power goes to infinity.

5.1 Channel Model

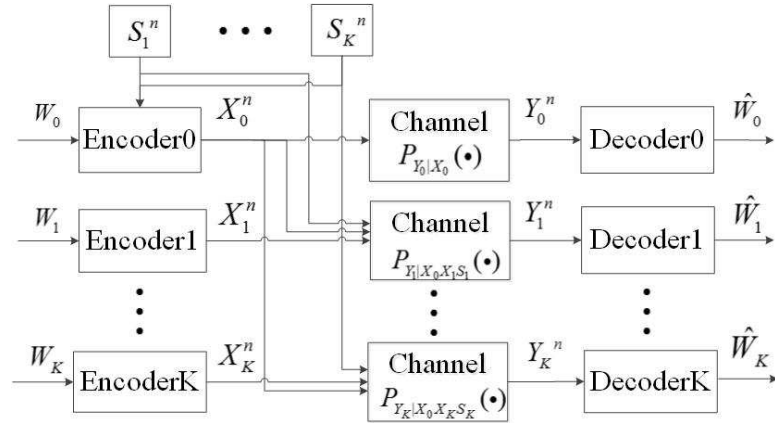


Fig. 5.1: The state-dependent parallel channel with a common helper

In the state-dependent parallel channel with a common helper (see Fig. 1.5 in Section 1.2. For convenience of reference, we include the figure again as Fig. 5.1 in this section), each transmitter (say transmitter k) has an encoder $f_k : \mathcal{W}_k \rightarrow \mathcal{X}_k^n$, which maps a message $w_k \in \mathcal{W}_k$ to a codeword $x_k^n \in \mathcal{X}_k^n$ for $k = 1, \dots, K$. The K inputs x_1^n, \dots, x_K^n are transmitted over K parallel channels, respectively. Each receiver (say receiver k) is interfered by an i.i.d. state sequence S_k^n for $k = 1, \dots, K$, which is known at none of transmitters $1, \dots, K$ and receivers $1, \dots, K$. A common helper is assumed to know all state sequences S_k^n for $k = 1, \dots, K$ noncausally. Thus, the encoder at the helper, $f_0 : \mathcal{W}_0 \times \{\mathcal{S}_1^n, \dots, \mathcal{S}_K^n\} \rightarrow \mathcal{X}_0^n$, maps a message $w_0 \in \mathcal{W}_0$ and the state sequences $(s_1^n, \dots, s_K^n) \in \mathcal{S}_1^n \times \dots \times \mathcal{S}_K^n$ to a codeword $x_0^n \in \mathcal{X}_0^n$. The entire channel transition probability is given by $P_{Y_0|X_0} \prod_{k=1}^K P_{Y_k|X_0, X_k, S_k}$. There are $K + 1$ decoders with each at one receiver, $g_k : \mathcal{Y}_k^n \rightarrow \mathcal{W}_k$, maps a received sequence y_k^n into a message $\hat{w}_k \in \mathcal{W}_k$ for $k = 0, 1, \dots, K$.

We study the following three Gaussian channel models.

In model I, $K = 1$, i.e., the helper assists one transmitter-receiver pair. The channel outputs at receiver 0 and 1 for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (5.1a)$$

$$Y_1 = X_0 + X_1 + S_1 + N_1. \quad (5.1b)$$

In model II, $K = 2$, in which one helper assists two transmitter-receiver pairs, and only one receiver is interfered by a state sequence. The channel outputs at receivers 0, 1 and 2 for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (5.2a)$$

$$Y_1 = X_0 + X_1 + S_1 + N_1, \quad (5.2b)$$

$$Y_2 = X_0 + X_2 + N_2. \quad (5.2c)$$

In model III, K is general, in which a common helper assists multiple transmitter-receiver pairs with each receiver corrupted by an independently distributed state sequence. This model is more general than model I, but does not include model II as a special case (due to infinite state power). The channel outputs at receivers 0 and receivers $1, \dots, K$ for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (5.3a)$$

$$Y_k = X_0 + X_k + S_k + N_k, \quad \text{for } k = 1, \dots, K \quad (5.3b)$$

In the above three models, the noise variables N_0, N_1, \dots, N_K and the state variables S_1, \dots, S_K are Gaussian distributed with distributions $N_0, \dots, N_K \sim \mathcal{N}(0, 1)$ and $S_k \sim \mathcal{N}(0, Q_k)$ for $k = 1, \dots, K$, and all of the variables are independent and are i.i.d. over channel uses. The channel inputs X_0, X_1, \dots, X_K are subject to the average power constraints $\frac{1}{n} \sum_{i=1}^n X_{ki}^2 \leq P_k$ for $k = 0, 1, \dots, K$.

We are interested in the regime of high state power, i.e., as $Q_k \rightarrow \infty$ for $k = 1, \dots, K$. Our goal is to design helper strategies in order to cancel the high power state interference and to further characterize the capacity region in this regime.

5.2 Model I: $K = 1$

In this section, we study the model I with $K = 1$. It is a basic model, in which the helper assists one transmitter-receiver pair. Understanding this model will help the study of the general parallel network. In this section, we first develop outer and inner bounds on the capacity region, and then characterize the boundary of the capacity region based on these bounds.

We first provide an outer bound on the capacity region in high state power regime.

Proposition 5.1. *For the Gaussian channel of model I, an outer bound on the capacity*

region for the regime when $Q_1 \rightarrow \infty$ consists of rate pairs (R_0, R_1) satisfying:

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (5.4a)$$

$$R_0 + R_1 \leq \frac{1}{2} \log(1 + P_0). \quad (5.4b)$$

The bound (5.4a) on R_1 follows simply from the capacity of the single-user channel between transmitter 1 and receiver 1 without signal and state interference. The bound (5.4b) on the sum rate is limited only by the power P_0 of the helper, and does not depend on the power P_1 of transmitter 1. Intuitively, this is because P_0 is split for transmission of W_0 and for helping transmission of W_1 by removing state interference, and hence P_0 determines a trade-off between R_0 and R_1 . On the other hand, improving the power P_1 , although may improve R_1 , can also cause more interference for receiver 1 to decode the auxiliary variable for canceling state and interference. Thus, the balance of the two effects determines that P_1 does not affect the sum rate.

Proof. The proof is detailed in Appendix C.1. □

We further note that although the sum-rate upper bound (5.4b) can be achieved easily by keeping transmitter 1 silent (i.e., R_0 achieves the sum rate bound with $R_1 = 0$), we are interested in characterizing the capacity region (i.e., the trade-off between R_0 and R_1) rather than a single point that achieves the sum-rate capacity. In the next section, we characterize such optimal trade-off based on the sum-rate bound.

We then design a coding scheme and derive the achievable region accordingly. The major challenge in designing an achievable scheme arises from the mismatched property due to transmitter-side state cognition and receiver-side state interference, i.e., state interference to receiver 1 is known noncausally only to the helper, not to the corresponding transmitter 1. Since we study the regime with large state power, transmitter 1 can send information to receiver 1 only if the helper assists to cancel the state. Thus, the helper needs to resolve the tension between transmitting its own message to receiver 0 and helping receiver 1 to cancel

its interference. A simple scheme of time-sharing between the two transmitters in general is not optimal.

We design a layered coding scheme as follows. The helper splits its signal into two parts in a layered fashion: one (represented by X'_0 in Proposition 5.2) for transmitting its own message and the other (represented by U in Proposition 5.2) for helping receiver 1 to remove both state and signal interference. In particular, the second part of the scheme applies a single-bin dirty paper coding scheme, in which transmission of W_1 and treatment of state interference for decoding W_1 are performed separately by transmitter 1 and the helper. This is because the helper knows the state but does not know the message (of transmitter 1) that the state interferes, and hence cannot encode this message via the regular multi-bin dirty paper coding as in [10]. Based on such a scheme, we obtain the following achievable rate region for the discrete memoryless channel, which is useful for deriving an inner bound for the Gaussian channel.

Proposition 5.2. *For the discrete memoryless channel of model I, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_0 \leq I(X'_0; Y_0) \tag{5.5a}$$

$$R_1 \leq I(X_1; Y_1 | U) \tag{5.5b}$$

$$R_1 \leq I(X_1 U; Y_1) - I(U; S_1 X'_0) \tag{5.5c}$$

for some distribution $P_{S_1} P_{X'_0} P_{U|S_1 X'_0} P_{X_0|U S_1 X'_0} P_{X_1} P_{Y_0|X_0} P_{Y_1|S_1 X_0 X_1}$.

Proof. The proof is detailed in Appendix C.2. □

Based on Proposition 5.2, we have the following simpler inner bound.

Corollary 5.1. *For the discrete memoryless channel of model I, an inner bound on the*

capacity region consists of rate pairs (R_0, R_1) satisfying:

$$R_0 \leq I(X'_0; Y_0) \quad (5.6a)$$

$$R_1 \leq I(X_1; Y_1|U) \quad (5.6b)$$

for some distribution $P_{S_1}P_{X'_0}P_{U|S_1X'_0}P_{X_0|US_1X'_0}P_{X_1}P_{Y_0|X_0}P_{Y_1|S_1X_0X_1}$ that satisfies

$$I(U; Y_1) \geq I(U; S_1X'_0). \quad (5.7)$$

Proof. The region follows from Proposition 5.2 because (5.5c) is redundant due to the condition (5.7). \square

The inner bound in Corollary 5.1 corresponds to an intuitive achievable scheme based on successive cancelation. Namely, the condition guarantees that receiver 1 decodes the auxiliary random variable U first, and then removes it from its output and decodes the message, which results in the bound (5.6b). In particular, cancelation of U leads to cancelation of signal and state interference at receiver 1.

We next derive an inner bound for the Gaussian channel of model I based on Corollary 5.1.

Proposition 5.3. *For the Gaussian channel of model I, in the regime when $Q_1 \rightarrow \infty$, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_0 \leq \frac{1}{2} \log \left(1 + \frac{\bar{\beta}P_0}{\beta P_0 + 1} \right) \quad (5.8a)$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{1 + (1 - \frac{1}{\alpha})^2 \beta P_0} \right) \quad (5.8b)$$

for some real constants $\alpha \geq 0$ and $0 \leq \beta \leq 1$ that satisfy $\alpha \leq \frac{2\beta P_0}{\beta P_0 + P_1 + 1}$.

Proof. Proposition 5.3 follows from Corollary 5.1 by choosing the joint Gaussian distribu-

tion for random variables as follows:

$$U = X_0'' + \alpha(S_1 + X_0'), \quad X_0 = X_0' + X_0''$$

$$X_0' \sim \mathcal{N}(0, \bar{\beta}P_0), \quad X_0'' \sim \mathcal{N}(0, \beta P_0)$$

$$X_1 \sim \mathcal{N}(0, P_1)$$

where X_0' , X_0'' , X_1 and S_1 are independent, $\alpha > 0$, $0 \leq \beta \leq 1$, and $\bar{\beta} = 1 - \beta$. \square

We note that in Proposition 5.3, the parameter α captures correlation between the state variable S_1 and the auxiliary variable U for dealing with the state, and can be chosen to optimize the rate region. This is in contrast to the classical dirty paper coding [10], in which such correlation parameter is fixed for state cancelation. Therefore, although Corollary 5.1 may provide a smaller inner bound than that given in Proposition 5.2, it can be shown that two inner bounds are equivalent for our chosen auxiliary random variables and input distribution after optimizing over α .

By comparing the inner and outer bounds, we characterize the boundary points of the capacity region for the Gaussian channel of model I based on the inner and outer bounds given in Propositions 5.3 and 5.1, respectively. We divide the Gaussian channel into three cases based on the conditions on the power constraints: (1) $P_1 \geq P_0 + 1$; (2) $P_0 - 1 \leq P_1 < P_0 + 1$ and (3) $0 \leq P_1 < P_0 - 1$. For each case, we optimize the dirty paper coding parameter α in the inner bound in Proposition 5.3 to find achievable rate points that lie on the sum-rate upper bound (5.4b) in order to characterize the boundary points of the capacity region.

Case 1: $P_1 \geq P_0 + 1$. The capacity region is fully characterized in the following theorem.

Theorem 5.1. *For the Gaussian channel of model I, in the regime when $Q_1 \rightarrow \infty$, if*

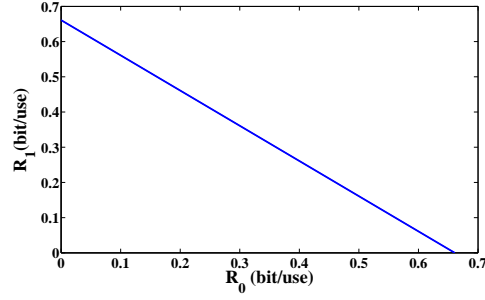


Fig. 5.2: The capacity region for case 1 with $P_0 = 1.5$ and $P_1 = 3$.

$P_1 \geq P_0 + 1$, the capacity region consists of the rate pairs (R_0, R_1) satisfying

$$R_0 + R_1 \leq \frac{1}{2} \log(1 + P_0). \quad (5.9)$$

Proof. Let \tilde{P}_1 be the actual power for transmitting W_1 . Then the inner bound (5.8b) on R_1 is optimized when $\alpha = \frac{2\beta P_0}{\beta P_0 + \tilde{P}_1 + 1}$. By setting $\tilde{P}_1 = \beta P_0 + 1$, the inner bound given in Proposition 5.3 matches the outer bound given in Proposition 5.1, and hence is the capacity region. \square

The capacity region of case 1 is illustrated in Fig. 5.2.

Theorem 5.1 implies that when P_1 is large enough, the power of the helper limits the system performance. Furthermore, since P_1 for transmission of W_1 causes interference to receiver 1 to decode the auxiliary variable for canceling state and interference, beyond a certain value, increasing P_1 does not improve the rate region any more. Theorem 5.1 also suggests that in order to achieve different points on the boundary of the capacity region (captured by the parameter β), different amounts of power \tilde{P}_1 should be applied.

Case 2: $P_0 - 1 \leq P_1 < P_0 + 1$. We summarize the capacity result in the following theorem.

Theorem 5.2. *Consider the Gaussian channel of model I in the regime when $Q_1 \rightarrow \infty$, and $P_0 - 1 \leq P_1 < P_0 + 1$. If $P_1 \geq 1$, the rate points (R_0, R_1) on the line A-B (see Fig. 5.3 (a) and Fig. 5.4 (a)) are on the capacity region boundary. More specifically, the points A*

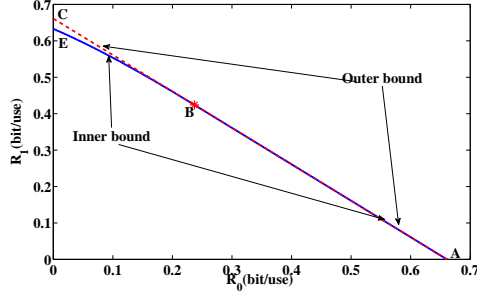
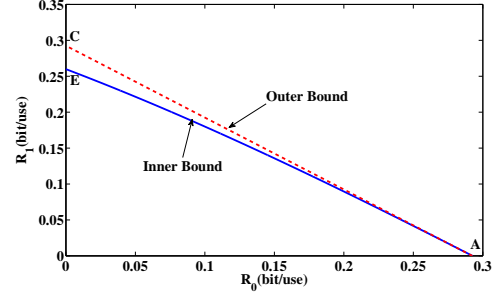
(a) $P_1 \geq 1$ with $P_0 = 1.5$ and $P_1 = 1.8$ (b) $P_1 < 1$ with $P_0 = 0.5$ and $P_1 = 0.8$

Fig. 5.3: Inner and outer bounds for case 2 with $P_0 < P_1$, which match partially on the boundaries

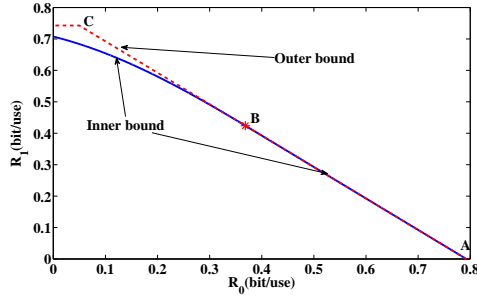
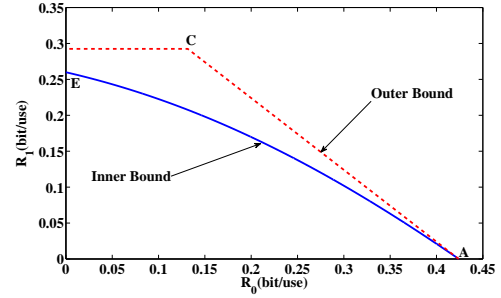
(a) $P_1 \geq 1$ with $P_0 = 2$ and $P_1 = 1.8$ (b) $P_1 < 1$ with $P_0 = 0.8$ and $P_1 = 0.5$

Fig. 5.4: Inner and outer bounds for case 2 with $P_0 \geq P_1$, which match partially on the boundaries

and B are characterized as:

$$\begin{aligned} \text{Point } A : & \left(\frac{1}{2} \log(1 + P_0), 0 \right) \\ \text{Point } B : & \left(\frac{1}{2} \log\left(1 + \frac{P_0 - P_1 + 1}{P_1}\right), \frac{1}{2} \log P_1 \right) \end{aligned}$$

If $P_1 < 1$ the rate point A (see Fig. 5.3 (b) and Fig. 5.4 (b)) is on the capacity region boundary, and is characterized as:

$$\text{Point } A : \left(\frac{1}{2} \log(1 + P_0), 0 \right)$$

Proof. We first set $\alpha = \frac{2\beta P_0}{\beta P_0 + P_1 + 1}$, and then substitute α into (5.8b) and obtain the following inner bound:

$$R_0 \leq \frac{1}{2} \log \left(1 + \frac{\bar{\beta} P_0}{\beta P_0 + 1} \right) \quad (5.12a)$$

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{4\beta P_0 P_1}{4\beta P_0 + (P_1 + 1 - \beta P_0)^2} \right). \quad (5.12b)$$

When $P_1 \geq 1$, by setting $\beta = \frac{P_1 - 1}{P_0}$, we obtain an achievable rate point B given by $\left(\frac{1}{2} \log(1 + \frac{P_0 - P_1 + 1}{P_1}), \frac{1}{2} \log P_1 \right)$, which is also on the outer bound. It is also clear that the point A given by $\left(\frac{1}{2} \log(1 + P_0), 0 \right)$ is achievable by setting $\beta = 0$, which is also on the outer bound. Thus, the line $A - B$ is on the boundary of the capacity region due to time sharing. \square

For this case, if $P_1 \geq 1$, i.e., P_1 is larger than the noise power, inner and outer bounds match over the line A-B as illustrated in Fig. 5.3 (a) and Fig. 5.4 (a), and thus optimal trade-off between R_0 and R_1 is achieved over the points on the line A-B. If $P_1 < 1$, the inner and outer bounds match only at the rate point A as illustrated in Fig. 5.3 (b) and Fig. 5.4 (b), which achieves the sum-rate capacity. We further note that Fig. 5.3 is different from Fig. 5.4 in the outer bound. Fig. 5.4 corresponds to the case with $P_0 \geq P_1$, and hence the capacity region is also upper bounded by the single-user capacity of R_1 . Such a bound is redundant in Fig. 5.3 which corresponds to the case with $P_0 < P_1$, because P_0 is not large enough to perfectly cancel state and signal interference at receiver 1. However, in case 3, we show that this single-user capacity of R_1 is achievable simultaneously with a certain positive R_0 .

Case 3: $P_1 < P_0 - 1$. We first summarize the capacity results in the following theorem.

Theorem 5.3. *Consider the Gaussian channel of model I in the regime when $Q_1 \rightarrow \infty$, and $P_1 < P_0 - 1$. If $P_1 \geq 1$, the rate points (R_0, R_1) on the line A-B (see Fig. 5.5 (a)) are on the boundary of the capacity region. More specifically, the points A and B are characterized*

as:

$$\begin{aligned} \text{Point } A &: \left(\frac{1}{2} \log(1 + P_0), 0 \right) \\ \text{Point } B &: \left(\frac{1}{2} \log\left(1 + \frac{P_0 - P_1 + 1}{P_1}\right), \frac{1}{2} \log P_1 \right) \end{aligned}$$

And the rate points (R_0, R_1) on the line D-E (see Fig. 5.5 (a)) are on the boundary of the capacity region. The points D and E are characterized as:

$$\begin{aligned} \text{Point } D &: \left(\frac{1}{2} \log\left(\frac{P_0 + 1}{P_1 + 2}\right), \frac{1}{2} \log(1 + P_1) \right) \\ \text{Point } E &: \left(0, \frac{1}{2} \log(1 + P_1) \right) \end{aligned}$$

If $P_1 < 1$, then point A (see Fig. 5.5 (b)) is on the capacity region boundary. The point A is characterized as:

$$\text{Point } A : \left(\frac{1}{2} \log(1 + P_0), 0 \right)$$

And the rate points (R_0, R_1) on the line D-E (see Fig. 5.5 (b)) are on the boundary of the capacity region. The points D and E are characterized as:

$$\begin{aligned} \text{Point } D &: \left(\frac{1}{2} \log\left(\frac{P_0 + 1}{P_1 + 2}\right), \frac{1}{2} \log(1 + P_1) \right) \\ \text{Point } E &: \left(0, \frac{1}{2} \log(1 + P_1) \right) \end{aligned}$$

Proof. For case 3, the inner bound boundary given in Proposition 5.3 is characterized by segment I consisting of rate points satisfying:

$$R_0 \leq \frac{1}{2} \log \left(1 + \frac{\bar{\beta} P_0}{1 + \beta P_0} \right) \quad (5.17a)$$

$$R_1 \leq \frac{1}{2} \log(1 + \beta P_0) \quad (5.17b)$$

for $0 \leq \beta \leq \frac{P_1+1}{P_0}$; and segment II consisting of rate points satisfying

$$R_0 \leq \frac{1}{2} \log \left(1 + \frac{\bar{\beta} P_0}{1 + \beta P_0} \right) \quad (5.18a)$$

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (5.18b)$$

for $\frac{P_1+1}{P_0} \leq \beta \leq 1$. Segment I is obtained by setting $\alpha = \frac{2\beta P_0}{\beta P_0 + P_1 + 1}$, and segment II is obtained by setting $\alpha = 1$.

For segment I, if $P_1 \geq 1$, the line A-B is on the boundary of the capacity region as shown in Fig. 5.5 (a). If $P_1 < 1$, only point A is on the capacity boundary as shown in Fig. 5.5 (b). For segment II, it is clear that the single-user channel capacity for R_1 is achievable. Furthermore, by setting $\beta = \frac{P_1+1}{P_0}$, the point D is achievable. Thus, the line D – E as shown in Fig. 5.5 (a) and (b) is on the boundary of the capacity region. \square

Similarly to cases 2, the inner and outer bounds match partially over the sum rate bound, i.e., the two bounds match over the line A-B (see Fig. 5.5 (a)) if $P_1 \geq 1$ and match at only the point A (see Fig. 5.5 (b)) if $P_1 < 1$. However, differently from case 2, the inner and outer bounds also match when $R_1 = \frac{1}{2} \log(1 + P_1)$ over the line D-E (see Fig. 5.5 (a) and (b)). This is because the power P_0 of the helper in this case is large enough to fully cancel state and signal interference so that transmitter 1 is able to reach its maximum single-user rate to receiver 1 without interference. Furthermore, the helper is also able to simultaneously transmit its own message at a certain positive rate.

5.3 Model II: $K = 2$

In this section, we study the model II with $K = 2$, and only receiver 1 corrupted by the channel state. In this model, the challenge lies in the fact that the helper needs to assist receiver 1 to remove the state interference, but such signal inevitably causes interference to receiver 2. To better understand the function of the helper, we study the case with $W_0 = \phi$,

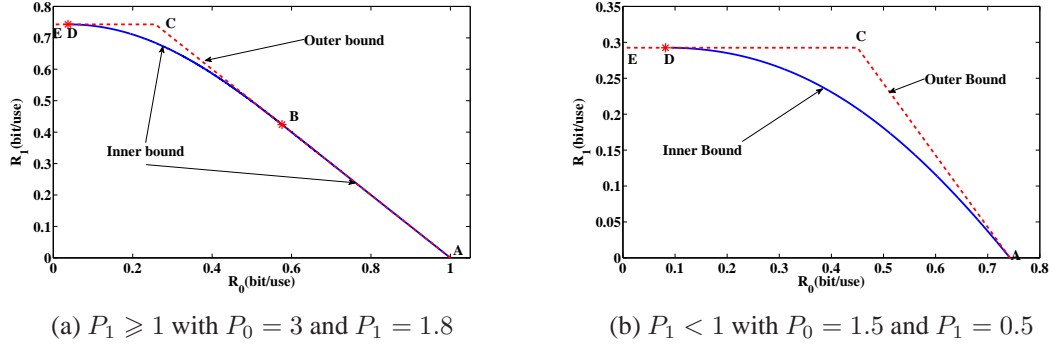


Fig. 5.5: Inner and outer bounds for case 3, which match partially on the boundaries.

and hence $Y_0 = \phi$. It is straight-forward to generalize these results to the model with $W_0 \neq \phi$.

We first provide a useful outer bound for Model II.

Proposition 5.4. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \rightarrow \infty$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \min \left\{ \frac{1}{2} \log(1 + P_0), \frac{1}{2} \log(1 + P_1) \right\} \quad (5.19a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (5.19b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + P_0 + P_2). \quad (5.19c)$$

Proof. The proof is detailed in Appendix C.3. □

We note that (5.19a) represents the best single-user rate of receiver 1 with the helper dedicated to help it as shown in Equation (5.4a) and (5.4a), (5.19b) is the single-user capacity for receiver 2, and (5.19c) implies that although the two transmitters communicate over parallel channels to their corresponding receivers, due to the shared common helper, the sum rate is still subject to a certain rate limit.

We next describe our idea to design an achievable scheme. We first note that although receiver 2 is not interfered by the state, the signal that the helper sends to assist receiver 1

to deal with the state still causes unavoidable interference to receiver 2. A natural idea to optimize the transmission rate to receiver 2 is simply to keep the helper silent. In this case, without the helper's assistance, receiver 1 gets zero rate due to infinite state power. Here, we design a novel scheme, which enables the single-user channel capacity for receiver 2 and a certain positive rate for receiver 1 simultaneously. Consequently, the helper is able to assist receiver 1 without causing interference to receiver 2. In our achievable scheme, the signal of the helper is split into two parts, represented by U and V as in Proposition 5.5. Here, U is designed to help receiver 1 to cancel the state while treating V as noise, and V is designed to help receiver 2 to cancel the interference caused by U . Since there is no state interference at receiver 2, U is decoded only at receiver 1. Based on such an achievable scheme, we obtain the following achievable region.

Proposition 5.5. *For the Gaussian channel of model II with $W_0 = \phi$, an achievable region consists of the rate pair (R_1, R_2) satisfying*

$$R_1 \leq I(X_1; Y_1 | U), \quad (5.20a)$$

$$R_2 \leq I(X_2; Y_2 | V), \quad (5.20b)$$

for some distribution $P_{S_1 UV X_0 X_1 X_2} = P_{S_1} P_{UV X_0 | S_1} P_{X_1} P_{X_2}$, where U and V are auxiliary random variables satisfying that

$$I(U; Y_1) \geq I(U; S_1), \quad (5.21a)$$

$$I(V; Y_2) \geq I(V; U S_1). \quad (5.21b)$$

Proof. The proof is detailed in Appendix C.4. □

Following from the above achievable region, we obtain an achievable region for the Gaussian channel by setting an appropriate joint input distribution.

Proposition 5.6. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when*

$Q_1 \rightarrow \infty$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{(1 - \frac{1}{\alpha})^2 P_{01} + P_{02} + 1} \right) \quad (5.22a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2}{1 + \frac{(\beta-1)^2 P_{02} P_{01}}{P_{02} + \beta^2 P_{01}}} \right) \quad (5.22b)$$

where $P_{01}, P_{02} \geq 0$, $P_{01} + P_{02} \leq P_0$, $0 \leq \alpha \leq \frac{2P_{01}}{1+P_0+P_1}$, and $P_{02}^2 + 2\beta P_{01} P_{02} \geq \beta^2 P_{01} (P_{02} + P_2 + 1)$.

Proof. The region follows from Proposition 5.5 by choosing jointly Gaussian distribution for random variables as follows:

$$U = X_{01} + \alpha S_1, \quad V = X_{02} + \beta X_{01}$$

$$X_0 = X_{01} + X_{02}$$

$$X_{01} \sim \mathcal{N}(0, P_{01}), \quad X_{02} \sim \mathcal{N}(0, P_{02})$$

$$X_1 \sim \mathcal{N}(0, P_1), \quad X_2 \sim \mathcal{N}(0, P_2)$$

where X_{01}, X_{02}, X_1, X_2 and S_1 are independent. □

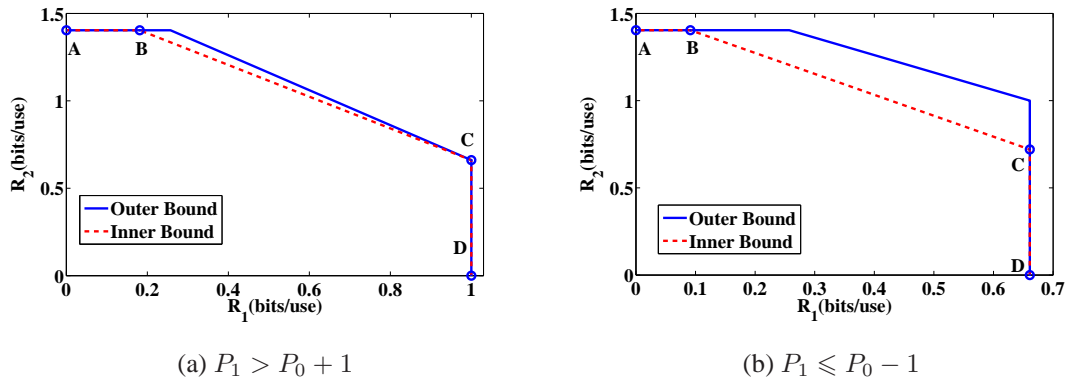


Fig. 5.6: Segments of the capacity boundary for the Gaussian channel of model II

Comparing the inner and outer bounds given in Propositions 5.6 and 5.4, respectively, we characterize two segments of the boundary of the capacity region, over which the two

bounds meet.

Theorem 5.4. *Consider the Gaussian model II with $W_0 = \phi$, in the regime when $Q_1 \rightarrow \infty$, the rate points on the line A-B (see Fig. 5.6-(a)) are on the capacity region boundary. More specifically, if $\frac{1}{2}(1 + P_0 + P_1) \geq \frac{P_0^2}{P_0 + P_2 + 1}$, points A and B are characterized as*

$$\begin{aligned} \text{Point A} &: \left(0, \frac{1}{2} \log(1 + P_2)\right) \\ \text{Point B} &: \left(\frac{1}{2} \log \left(1 + \frac{4P_1 P_0^2}{(1 + P_0 + P_1)^2(1 + P_0 + P_2) - 4P_1 P_0^2}\right), \frac{1}{2} \log(1 + P_2)\right). \end{aligned}$$

If $\frac{1}{2}(1 + P_0 + P_1) < \frac{P_0^2}{P_0 + P_2 + 1}$, points A and B are characterized as

$$\begin{aligned} \text{Point A} &: \left(0, \frac{1}{2} \log(1 + P_2)\right) \\ \text{Point B} &: \left(\frac{1}{2} \log \left(1 + \frac{P_1(P_0 + P_2 + 1)}{P_0 + (P_0 + 1)(P_2 + 1)}\right), \frac{1}{2} \log(1 + P_2)\right). \end{aligned}$$

Furthermore, the rate points on the line C-D (see Fig. 5.6-(a)) are also on the capacity region boundary. If $P_1 \geq P_0 + 1$, the points C and D are characterized as

$$\begin{aligned} \text{Point C} &: \left(\frac{1}{2} \log(1 + P_0), \frac{1}{2} \log \left(1 + \frac{P_2}{P_0 + 1}\right)\right) \\ \text{Point D} &: \left(\frac{1}{2} \log(1 + P_0), 0\right), \end{aligned}$$

as illustrated in Fig. 5.6-(a).

If $P_1 \leq P_0 - 1$, the points C and D is characterized as

$$\begin{aligned} \text{Point C} &: \left(\frac{1}{2} \log(1 + P_1), \frac{1}{2} \log \left(1 + \frac{P_2}{P_1 + 2}\right)\right) \\ \text{Point D} &: \left(\frac{1}{2} \log(1 + P_1), 0\right), \end{aligned}$$

as illustrated in Fig. 5.6-(b).

Proof. We first show that the line A-B is achievable. The point A is achievable by keeping

the helper silent. To show that the point B is achievable, we set $\beta = 1$ in Proposition 5.6, and hence the achievable rate R_2 in (5.22b) reaches the single-user channel capacity, and the condition $P_{02}^2 + 2\beta P_{01}P_{02} \geq \beta^2 P_{01}(P_{02} + P_2 + 1)$ becomes $P_{01} \leq \frac{P_0^2}{P_0 + P_2 + 1}$. We set $P_{01} = \frac{P_0^2}{P_0 + P_2 + 1}$.

If $\frac{1}{2}(1 + P_0 + P_1) \geq \frac{P_0^2}{P_0 + P_2 + 1}$, we have $\frac{2P_{01}}{1 + P_0 + P_1} \leq 1$. Thus, setting $\alpha = \frac{2P_{01}}{1 + P_0 + P_1}$, (5.22a) and (5.22b) imply that the point B is achievable.

If $\frac{1}{2}(1 + P_0 + P_1) \leq \frac{P_0^2}{P_0 + P_2 + 1}$, we have $\frac{2P_{01}}{1 + P_0 + P_1} \geq 1$. By setting $\alpha = 1$, (5.22a) and (5.22b) imply that the point B is achievable.

We next show that the line C - D is on the capacity boundary.

As implied by Theorems 5.1 and 5.3, the only possible cases that the outer bound (5.19a) (i.e. the maximum rate R_1 with the helper fully assisting receiver 1) can be achieved are when $P_1 \leq P_0 - 1$ and $P_1 \geq P_0 + 1$.

If $P_1 \geq P_0 + 1$, setting the actual transmission power of transmitter 1 as $\tilde{P}_1 = P_0 + 1$, $P_{01} = P_0$, $\alpha = \frac{P_0}{1 + P_0}$ and $\beta = 0$, then (5.22a) and (5.22b) imply that the rate point C is achievable. This point also achieves the sum capacity. It is obvious that point D is achievable, and hence the points on the line C - D are on the capacity boundary.

If $P_1 \leq P_0 - 1$, by setting $\beta = 0$, $\alpha = 1$ and $P_{01} = \tilde{P}_0 = P_1 + 1$ (where \tilde{P}_0 is the actual transmission power of the helper), then (5.22a) and (5.22b) imply that the rate point C is achievable. In particular, the actual power the helper uses is $P_1 + 1$ rather than P_0 , because larger P_0 does not help receiver 1 to decode more, but increases interference to receiver 2. It is clear that the point D is achievable. Hence, the points on the line C - D are on the capacity boundary. \square

The capacity result for the line A - B in Theorem 5.4 indicates that our coding scheme effectively enables the helper to assist receiver 1 without causing interference to receiver 2. Hence, R_2 achieves the corresponding single-user channel capacity, while transmitter 1 and receiver 1 communicate at a certain positive rate R_1 with the assistance of the helper.

The capacity result for the line C - D in Theorem 5.4 can be achieved based on a scheme,

in which the helper assists receiver 1 to deal with the state and receiver 2 treats the helper's signal as noise. Such a scheme is guaranteed to be the best by the outer bound if receiver 1's rate is maximized.

Corollary 5.2. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \rightarrow \infty$, if $P_1 \geq P_0 + 1$, the sum capacity is given by $\frac{1}{2} \log(1 + P_0 + P_2)$.*

5.4 Model III: General K

In this section, we consider the Gaussian channel of model III with $K \geq 2$, in which there are multiple receivers with each interfered by an independent state. In this section, we present the results for the scenario, in which the helper dedicates to help two users without transmitting its own message, i.e., $K = 2$ and $W_0 = \phi$. It is straight-forward to extend the result to the more general scenario, in which the helper assists more than two users and transmits its own message at the same time, i.e., $K > 2$ and $W_0 \neq \phi$.

We note that model III is more general than model I, but does not include model II as a special case, because model II has one receiver that is not corrupted by state, but each receiver (excluding the helper's targeted receiver) in model III is corrupted by an infinitely powered state sequence. Hence for model III, the challenge lies in the fact that the helper needs to assist multiple receivers to cancel interference caused by independent states. In this subsection, we first derive an outer bound on the capacity region, and then derive an inner bound based on a time-sharing scheme for the helper. Somewhat interestingly, comparing the inner and outer bounds concludes that the time-sharing scheme achieves the sum capacity under certain channel parameters, and we hence characterize segments of the capacity region boundary corresponding to the sum capacity under these channel parameters.

We first derive an outer bound on the capacity region.

Proposition 5.7. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime when $Q_1, Q_2 \rightarrow \infty$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (5.27)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (5.28)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + P_0). \quad (5.29)$$

Proof. The proof is detailed in Appendix C.5. □

Although the two transmitters transmit over parallel channels, the above outer bound suggests that their sum rate is still subject to a certain constraint determined by the helper's power. This implies that it is not possible for one common helper to cancel the two independent high-power states simultaneously (i.e., using the common resource). This fact also suggests that a time-sharing scheme, in which the helper alternatively assists each receiver, can be desirable to achieve the sum rate upper bound (i.e., to achieve the sum capacity).

We hence design a time-sharing achievable scheme. The helper splits its transmission duration into two time slots with the fraction γ of the total time duration for assisting receiver 1 and the fraction $1 - \gamma$ for assisting receiver 2. Each transmitter transmits only during the time slot that it is assisted by the helper, and keeps silent while the helper assisting the other transmitter. We note that the power constraints for transmitters 1 and 2 in their corresponding transmission time slots are $\frac{P_1}{\gamma}$ and $\frac{P_2}{1-\gamma}$, respectively.

Now at each transmission slot, the channel consists of one transmitter-receiver pair with the receiver corrupted by a infinite-power state, and one helper that assists the receiver to cancel the state interference. Such a model is equivalent to the state-dependent single-user

channel with a helper studied in [34]. We rewrite the achievable rate as follows:

$$R(P, P_0) := \begin{cases} \frac{1}{2} \log(1 + P_0), & P \geq P_0 + 1 \\ \frac{1}{2} \log\left(1 + \frac{4P_0P}{4P_0 + (P_0 - P - 1)^2}\right), & P_0 - 1 \leq P \leq P_0 + 1 \\ \frac{1}{2} \log(1 + P), & P \leq P_0 - 1. \end{cases} \quad (5.30)$$

By employing the time-sharing scheme between the helper assisting one receiver and the other alternatively, we obtain the following achievable region.

Proposition 5.8. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \rightarrow \infty$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \gamma R\left(\frac{P_1}{\gamma}, P_0\right) \quad (5.31a)$$

$$R_2 \leq (1 - \gamma) R\left(\frac{P_2}{1 - \gamma}, P_0\right) \quad (5.31b)$$

where $0 \leq \gamma \leq 1$ is the time-sharing coefficient, and the function $R(\cdot, \cdot)$ is defined in (5.30).

We note that following from (5.30), the best possible single-user rate is $\frac{1}{2} \log(1 + P_0)$, which can be achieved if $P \geq P_0 + 1$. This best rate may not be possible if P is not large enough. Interestingly, in a time-sharing scheme, both transmitters can simultaneously achieve the best single user rate $\frac{1}{2} \log(1 + P_0)$ over their transmission fraction of time, because both of their powers get boosted over a certain fraction of time, although neither power is larger than $P_0 + 1$. In this way, the sum rate upper bound (5.29) can be achieved. The following theorem characterizes the sum capacity of the channel for the scenario described above.

Theorem 5.5. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \rightarrow \infty$, if $P_1 + P_2 \geq P_0 + 1$, then the sum capacity equals*

$\frac{1}{2} \log(1 + P_0)$. The rate points that achieve the sum capacity (i.e. on the capacity region boundary) are characterized as $(R_1, R_2) = \left(\gamma R(\frac{P_1}{\gamma}, P_0), (1 - \gamma) R(\frac{P_2}{1 - \gamma}, P_0) \right)$ for $\gamma \in \left(\max(1 - \frac{P_2}{P_0 + 1}, 0), \min(\frac{P_1}{P_0 + 1}, 1) \right)$.

Proof. The proof is detailed in Appendix C.6. \square

The above theorem implies the following characterization of the full capacity region under certain parameters.

Corollary 5.3. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \rightarrow \infty$, if $P_1, P_2 \geq P_0 + 1$, then the capacity region consists of the rate pair (R_1, R_2) satisfying $R_1 + R_2 \leq \frac{1}{2} \log(1 + P_0)$.*

We next provide channel examples to understand the outer and inner bounds respectively in Proposition 5.7 and 5.8, and in the sum capacity in Theorem 5.5. It can be seen that the power constraints fall into four cases, among which we consider the following three cases: case 1. $P_1 \geq P_0, P_2 \geq P_0$; case 2. $P_1 \geq P_0, P_2 < P_0$; and case 3. $P_1 < P_0, P_2 < P_0$ by noting that case 4 is opposite to case 2 and is omitted due to symmetry of the two transmitters.

- **Case 1:** $P_1 \geq P_0, P_2 \geq P_0$. We consider an example channel with $P_0 = 1$, $P_1 = 1.8$ and $P_2 = 1.5$. Fig. 5.7 plots the inner and outer bounds on the capacity region. In particular, the two bounds meet over the line segment B-C, which corresponds to the rate points $(R_1, R_2) = \left(\gamma R(\frac{P_1}{\gamma}, P_0), (1 - \gamma) R(\frac{P_2}{1 - \gamma}, P_0) \right)$ for $\gamma \in \left(\max(1 - \frac{P_2}{P_0 + 1}, 0), \min(\frac{P_1}{P_0 + 1}, 1) \right)$ as characterized in Theorem 5.5. All these rate points achieve the sum capacity. It can also be seen that although neither transmitter achieves the best possible single-user rate, the sum capacity can be achieved due to the time-sharing scheme. We also note that, in this case, if the conditions in Corollary 5.3 are satisfied, the full capacity region is characterized.

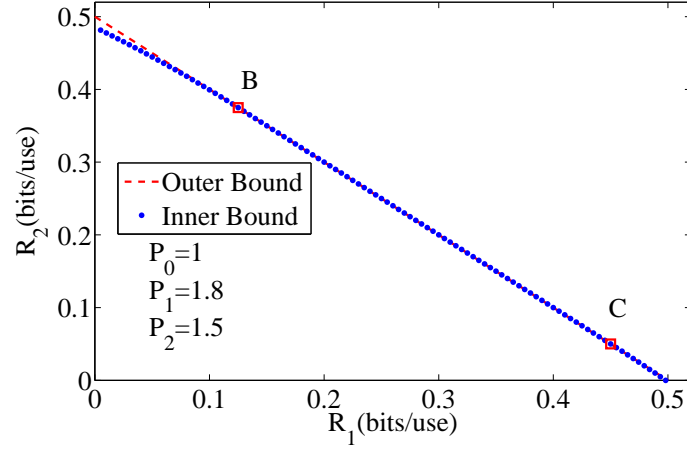


Fig. 5.7: Segment on the capacity region for the Gaussian channel of model III

- Case 2: $P_1 > P_0, P_2 \leq P_0$. We consider an example channel with $P_0 = 2, P_1 = 2.5$ and $P_2 = 0.8$. Fig. 5.8 plots the inner and outer bounds on the capacity region. Similarly to case 1, the two bounds meet over the line segment B-C as characterized in Theorem 5.5, and the points on such a line segment achieve the sum capacity. Differently from case 1, transmitter 2 achieves its single-user channel capacity indicated by the point A in Figure 5.8. This is consistent with the single user rate provided in (5.30) for the case with $P_2 \leq P_0 - 1$.

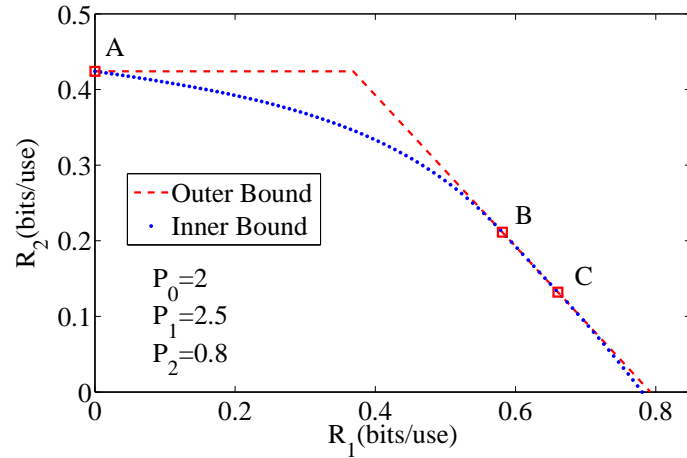


Fig. 5.8: Segment on the capacity region for the Gaussian channel of model III

- Case 3: $P_1 < P_0, P_2 < P_0$. We consider an example channel with $P_0 = 4, P_1 = 3$ and $P_2 = 3$. Figure 5.9 plots the inner and outer bounds on the capacity region. The points on the line segment B-C achieve the sum capacity as characterized in Theorem 5.5, and the points A and D respectively achieve the single-user capacity for two transceiver pairs. This is consistent with the single-user rate provided in (5.30) for the case with $P_1, P_2 \leq P_0 - 1$.

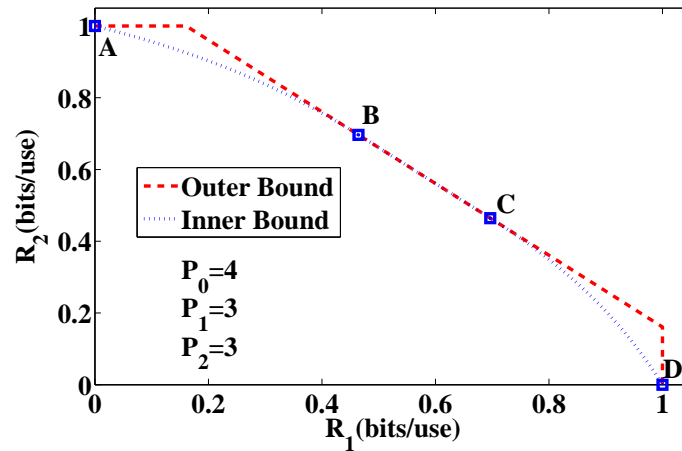


Fig. 5.9: Segment on the capacity region for the Gaussian channel of model III

CHAPTER 6

STATE-DEPENDENT MULTIPLE ACCESS CHANNEL WITH A HELPER

In this chapter, we study the state-dependent MAC with a helper. Our focus is on the Gaussian channel with additive state. We derive an outer bound on the capacity region, and obtain an inner bound based on a dirty interference cancelation scheme. By comparing the inner and outer bounds, we characterize the full capacity region or segment on the boundary of the capacity region under various channel parameters.

6.1 Channel Model

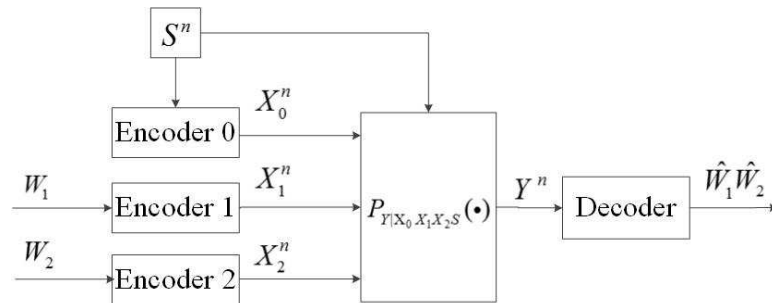


Fig. 6.1: The state-dependent MAC with a helper

In the state-dependent MAC with a common helper (see Fig. 1.6 in Section 1.2. For

convenience of reference, we include the figure again as Fig. 6.1 in this section), transmitter 1 and transmitter 2 send their own messages to the receiver, respectively. The channel is corrupted by a state sequence. The state sequence is known to neither the transmitters nor the receiver, but is known to a helper noncausally. Hence, the helper assists the receiver to cancel the state interference. More specifically, two encoders, each at one transmitter, $f_k : \mathcal{W}_k \rightarrow \mathcal{X}_k^n$ map the message $w_k \in \mathcal{W}_k$ to a codeword $x_k^n \in \mathcal{X}_k^n$ for $k = 1, 2$. The encoder at the helper, $f_0 : \mathcal{S}^n \rightarrow \mathcal{X}_0^n$ maps the state sequence $s^n \in \mathcal{S}^n$ into a codeword $x_0^n \in \mathcal{X}_0^n$. The help signal x_0^n and the inputs x_1^n, x_2^n are transmitted over the MAC to the receiver. The channel transition probability is given by $P_{Y|X_0X_1X_2S}$. The decoder at the receiver, $g : \mathcal{Y}^n \rightarrow (\mathcal{W}_1, \mathcal{W}_2)$ maps the received sequence y^n into two messages $\hat{w}_k \in \mathcal{W}_k$ for $k = 1, 2$.

We focus on the Gaussian channel with the output at receiver for one channel use given by

$$Y = X_0 + X_1 + X_2 + S + N \quad (6.1)$$

where the noise variable $N \sim \mathcal{N}(0, 1)$ and the state variable $S \sim \mathcal{N}(0, Q)$. Both the noise and state variables are i.i.d. over channel uses. The channel inputs X_0 , X_1 and X_2 are subject to the average power constraints P_0 , P_1 and P_2 .

6.2 Outer and Inner Bounds on Capacity

In this section, we provide outer and inner bounds on the capacity region for the state-dependent Gaussian MAC with a helper. We start with an outer bound as follows.

Proposition 6.1. *For the state-dependent Gaussian MAC with a helper, an outer bound on*

the capacity region consists of the rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \min \left\{ \frac{1}{2} \log(1 + P_1), \frac{1}{2} \log(1 + P_0) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\rho_{0S}\sqrt{P_0 Q} + P_1 + 1}{Q} \right) \right\} \quad (6.2a)$$

$$R_2 \leq \min \left\{ \frac{1}{2} \log(1 + P_2), \frac{1}{2} \log(1 + P_0) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\rho_{0S}\sqrt{P_0 Q} + P_2 + 1}{Q} \right) \right\} \quad (6.2b)$$

$$R_1 + R_2 \leq \min \left\{ \frac{1}{2} \log(1 + P_1 + P_2), \frac{1}{2} \log(1 + P_0) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\rho_{0S}\sqrt{P_0 Q} + P_1 + P_2 + 1}{Q} \right) \right\} \quad (6.2c)$$

for some ρ_{0S} that satisfies $-1 \leq \rho_{0S} \leq 1$.

Proof. The first bounds in (6.2a)-(6.2c) follow from the capacity of the Gaussian MAC without state. The remaining bounds arise due to capability of the helper for assisting state cancelation. Detailed proof is relegated to D.1. \square

In particular, we are interested in the large state power regime, i.e., $Q \rightarrow \infty$. The following outer bound for such a regime follows readily from Proposition 6.1.

Corollary 6.1. *For the state-dependent Gaussian MAC with a helper, in the regime that $Q \rightarrow \infty$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (6.3a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (6.3b)$$

$$R_1 + R_2 \leq \min \left\{ \frac{1}{2} \log(1 + P_1 + P_2), \frac{1}{2} \log(1 + P_0) \right\} \quad (6.3c)$$

We note that as $Q \rightarrow \infty$, the communication rates are not only bounded by the power

constraints of transmitters 1 and 2, but also by the power of the helper. This is because as the state power becomes asymptotically large, the receiver must remove the state interference first in order to decode useful information. In this case, increasing the powers P_1 and P_2 causes more interference for the receiver to remove the state, and hence may reduce the sum rate. Thus, when $P_1 + P_2$ is large enough, the sum rate depends only on the power of the helper that affects how the state can be removed.

We next derive an achievable region for the channel. The basic idea of the achievable scheme is to employ a dirty interference cancelation scheme, i.e., the helper incorporates two schemes for canceling state interference: scheme 1 cancels some state power by signals that exactly reverses the state realization; and scheme 2 uses dirty paper coding via generation of an auxiliary variable (represented by U in Proposition 6.2) to incorporate the state information so that the receiver decodes such variable first to cancel the state and then decode the users' information. Based on such an achievable scheme, we derive the following inner bound on the capacity region. The detailed proof is omitted due to the space limitations.

Proposition 6.2. *For the discrete memoryless state-dependent MAC with a helper, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_1 \leq I(X_1; Y | U X_2) \quad (6.4a)$$

$$R_2 \leq I(X_2; Y | U X_1) \quad (6.4b)$$

$$R_1 + R_2 \leq I(X_1 X_2; Y | U) \quad (6.4c)$$

for some distribution $P_S P_{U|S} P_{X_0|US} P_{X_1} P_{X_2} P_{Y|SX_0X_1X_2}$ such that

$$I(U; Y) \geq I(U; S). \quad (6.5)$$

Proof. The proof is detailed in Appendix D.2. □

We note that the constraint (6.5) is imposed because the receiver needs to decode the auxiliary codeword (with single letter representation U) that the helper generates to cancel the state. Based on the above inner bound, we derive the following inner bound for the Gaussian channel.

Proposition 6.3. *For the state-dependent Gaussian MAC with a helper, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{\frac{(\alpha-1-\beta)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + 1} \right) \quad (6.6a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2}{\frac{(\alpha-1-\beta)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + 1} \right) \quad (6.6b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + P_2}{\frac{(\alpha-1-\beta)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + 1} \right) \quad (6.6c)$$

for some real constants α, β , and $0 \leq P_{00} \leq P_0$ that satisfy

$$-\sqrt{\frac{P_0 - P_{00}}{Q}} \leq \beta \leq \sqrt{\frac{P_0 - P_{00}}{Q}}, \quad \text{and} \quad \alpha^2 Q(P_1 + P_2 + 1 + P_{00}) - 2\alpha P_{00} Q(1 + \beta) - P_{00}^2 \leq 0. \quad (6.7)$$

Proof. The region follows from Proposition 6.2 by choosing the joint Gaussian distribution for random variables as follows:

$$U = X_{00} + \alpha S, \quad X_0 = X_{00} + \beta S, \quad X_{00} \sim \mathcal{N}(0, P_{00}),$$

where X_{00} and S are independent. The constraints on β follows due to the power constraints on X_0 . \square

We note that the above construction of the input X_0 of the helper reflects two state cancelation schemes: the term βS represents directly cancelation of some state power via reverse of the state realization; and the variable X_{00} is used for dirty paper coding via

generation of the state-correlated auxiliary variable U . Hence, the parameter β controls the balance of two schemes in the integrated scheme, and can be optimized to achieve the best performance. This scheme is also equivalent to the one with $U = X_0 + \alpha S$, where X_0 and S are correlated. While such approaches have been considered in the literature (see e.g., [37]), we believe that selecting U and X_0 successively provides a more operational meaning to the correlation structure.

In the high state power regime, i.e., $Q \rightarrow \infty$, it is necessary that $\beta = 0$, because the helper's input X_0 has only limited power. Hence, in this case, the achievable scheme completely uses dirty paper coding for state cancelation. Such a scheme yields the following inner bound.

Corollary 6.2. *For the state-dependent Gaussian MAC with a helper, in the regime with $Q \rightarrow \infty$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{(1 - \frac{1}{\alpha})^2 P_0 + 1} \right) \quad (6.8a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2}{(1 - \frac{1}{\alpha})^2 P_0 + 1} \right) \quad (6.8b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1 + P_2}{(1 - \frac{1}{\alpha})^2 P_0 + 1} \right) \quad (6.8c)$$

for some constant α that satisfies $0 \leq \alpha \leq \frac{2P_0}{P_0 + P_1 + P_2 + 1}$.

6.3 Capacity Results

In this section, by comparing the inner and outer bounds, we characterize the capacity region or segment on the capacity boundary under various channel parameters. We first characterize the capacity region for case 1 when the helper's power is relatively large.

Theorem 6.1. *For the state-dependent Gaussian MAC with a helper, if $P_0 \geq P_1 + P_2 + 1$,*

the capacity region consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (6.9a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (6.9b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + P_1 + P_2) \quad (6.9c)$$

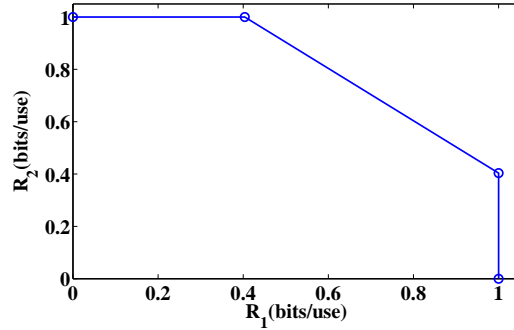


Fig. 6.2: An illustration of the capacity region for state-dependent Gaussian MAC with a helper for case 1 with $P_1 = P_2 = 3$, $P_0 = 7.5$ and arbitrary Q (characterized by Theorem 6.1) and case 2 with $P_1 = P_2 = 3$, $P_0 = 4.5$ and $Q = 8$ (characterized by Theorem 6.2).

Proof. The achievability follows by setting $\alpha = 1$ and $\beta = 0$ in Proposition 6.3. It is easy to check that the condition (6.7) is satisfied given $P_0 \geq P_1 + P_2 + 1$. It is clear that such an inner bound matches the outer bound in Proposition 6.1. \square

Theorem 6.1 implies that if the helper's power is above a certain threshold, the capacity region of the state-dependent MAC with a helper is the same as the capacity region of the MAC without state. Thus, the helper is capable to fully cancel the state interference at the receiver. In particular, the statement holds for any state interference power, which can be as large as infinite. In Fig. 6.2, we illustrate the capacity region of case 1 under an example set of channel parameters, i.e., $P_1 = P_2 = 3$, $P_0 = 7.5$, and arbitrary Q .

We next characterize the capacity region for case 2 when the helper's power is not large enough.

Theorem 6.2. *For the state-dependent Gaussian MAC with a helper, if $P_0 < P_1 + P_2 + 1$, and if*

$$\sqrt{Q} \leq \max_{0 \leq P_{00} \leq P_0} \sqrt{P_0 - P_{00}} + \frac{P_{00}}{\sqrt{P_1 + P_2 + 1 - P_{00}}}, \quad (6.10)$$

then the capacity region consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \frac{1}{2} \log(1 + P_1) \quad (6.11a)$$

$$R_2 \leq \frac{1}{2} \log(1 + P_2) \quad (6.11b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + P_1 + P_2). \quad (6.11c)$$

Proof. When (6.10) is satisfied, by setting $\alpha = 1 + \beta$ in Proposition 6.3, the region in (6.11a)-(6.11c) is achieved, which matches with the outer bound, and hence is the capacity region. \square

Theorem 6.2 implies that if the helper's power is below a certain threshold, then the capacity region of the state-dependent MAC with a helper is the same as the capacity region of the MAC without state when the state power is lower than a certain value. Thus, the helper can fully cancel the state interference at the receiver only for a certain range of state power. It can also be checked that the threshold on Q given in (6.10) can be larger than P_0 , which implies that dirty paper coding is necessary in the achievable scheme to fully cancel state interference.

The capacity region in Fig. 6.2 is also applicable to case 2 characterized in Theorem 6.2 under certain channel parameters, for example, when $P_1 = P_2 = 3$, $P_0 = 4.5$ and $Q = 8$. Compared to case 1, the helper's power is smaller, but achieves the same capacity region. This is reasonable, because the state power Q in case 2 is limited by a certain threshold, but the power Q in case 1 can be arbitrary.

We finally study case 3 with the helper's power being small. For this case, the capacity

region is limited by the helper's power constraint. The following theorem characterizes the sum capacity and the segment of the boundary of the capacity region in the large state power regime.

Theorem 6.3. *For the state-dependent Gaussian MAC with a helper, if $P_0 < P_1 + P_2 - 1$, in the regime of $Q \rightarrow \infty$, the sum capacity equals to $\frac{1}{2} \log(1 + P_0)$. Furthermore, the points on the line B-C (see Fig. 6.3 for an illustration) are on the boundary of the capacity region, where the points B and C are characterized as*

$$\begin{aligned} B : & \left(\frac{1}{2} \log(1 + P_0) - R_B, R_B \right) \\ \text{where } R_B = & \frac{1}{2} \log\left(1 + \frac{P_0 \min\{P_2, P_0 + 1\}}{1 + P_0}\right) \\ C : & \left(R_C, \frac{1}{2} \log(1 + P_0) - R_C \right) \\ \text{where } R_C = & \frac{1}{2} \log\left(1 + \frac{P_0 \min\{P_1, P_0 + 1\}}{1 + P_0}\right). \end{aligned}$$

Proof. Achievability of the sum capacity follows from Corollary 6.2 by setting the actual transmission powers of the two transmitters to be $0 \leq \tilde{P}_1 \leq P_1$ and $0 \leq \tilde{P}_2 \leq P_2$ such that $\tilde{P}_1 + \tilde{P}_2 = P_0 + 1$. The upper bound on the sum capacity follows from Corollary 6.1. The points B and C are characterized by setting $\tilde{P}_2 = \min\{P_2, P_0 + 1\}$ and $\tilde{P}_1 = \min\{P_1, P_0 + 1\}$, respectively. \square

We illustrate an example of case 3 in Fig. 6.3, in which the inner and outer bounds match over the line B-C.

Theorem 6.3 implies the characterization of the full capacity under further conditions, as given in the following corollary.

Corollary 6.3. *For the state-dependent Gaussian MAC with a helper, if $P_0 < \min\{P_1, P_2\} - 1$, in the regime of $Q \rightarrow \infty$, the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + P_0). \quad (6.12)$$

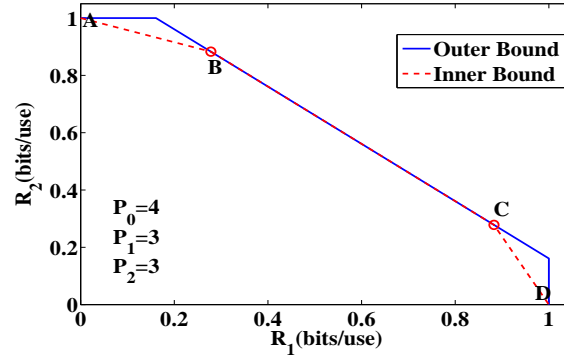


Fig. 6.3: A illustration of the segment of the capacity boundary for state-dependent Gaussian MAC with a helper: $P_0 < P_1 + P_2 - 1$.

Theorem 6.3 and its Corollary 6.3 imply that if the helper's power is below a certain threshold, then the capacity region of the state-dependent MAC with a helper is strictly smaller than the capacity region of the MAC without state. Thus, the helper is not able to fully cancel the state interference at the receiver. This is particularly reflected in the asymptotical regime as the state power $Q \rightarrow \infty$.

CHAPTER 7

STATE-DEPENDENT BROADCAST

CHANNEL WITH A HELPER

In this chapter, we study the state-dependent broadcast channel with a helper for two scenarios. In scenario 1, the transmitter sends one message to both receivers, and in scenario II, the transmitter sends two private messages respectively to two receivers. Our focus is on the Gaussian channel with additive state. We derive inner and outer bounds for both scenarios. By comparing the inner and outer bounds, we characterize the capacity/capacity region under various ranges of channel parameters.

7.1 Channel Model

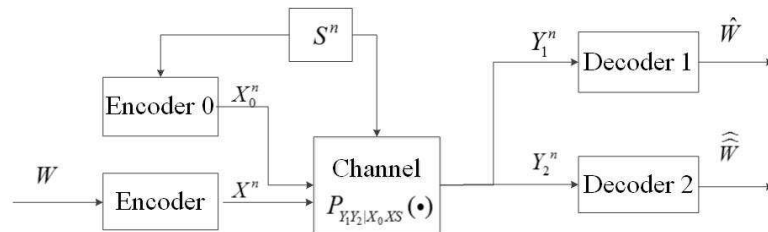


Fig. 7.1: The state-dependent broadcast channel with a helper: Scenario with a common message

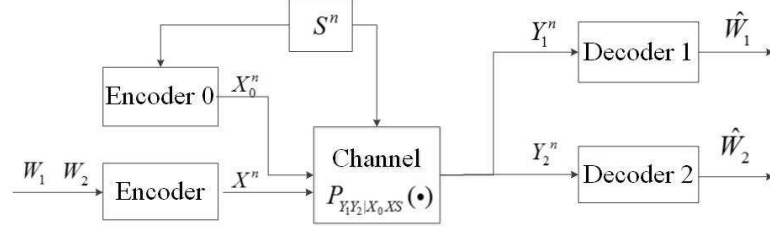


Fig. 7.2: The state-dependent broadcast channel with a helper: Scenario with private messages

We study two scenarios for the state-dependent broadcast channel with a helper. In scenario I (see Fig. 1.7 in Section 1.2. For convenience of reference, we include the figure again as Fig. 7.1 in this section), the transmitter wishes to transmit one common message $W \in \mathcal{W}$ to two receivers. The encoder $f : \mathcal{W} \rightarrow \mathcal{X}^n$, maps a message $w \in \mathcal{W}$ to a codeword $x^n \in \mathcal{X}^n$. The input x^n is transmitted over the broadcast channel, which is interfered by an i.i.d. state sequence S^n . The state sequence is known at neither the transmitter nor the receivers. A helper which knows the state sequence noncausally assists both receivers to deal with the channel state. Thus, the encoder at the helper, $f_0 : \mathcal{S}^n \rightarrow \mathcal{X}_0^n$, maps the state sequences $s^n \in \mathcal{S}^n$ to a codeword $x_0^n \in \mathcal{X}_0^n$. The channel transition probability is given by $P_{Y_1 Y_2 | X_0 X S}$. Two decoders with each at one receiver, $g_k : \mathcal{Y}_k^n \rightarrow \mathcal{W}$, maps a received sequence y_k^n into the message $\hat{w} \in \mathcal{W}$ for $k = 1, 2$.

Scenario II (see Fig. 1.8 in Section 1.2. For convenience of reference, we include the figure again as Fig. 7.2 in this section) is similar to scenario I with the difference being that the transmitter sends two independent messages $W_1 \in \mathcal{W}_1$ and $W_2 \in \mathcal{W}_2$ to receivers 1 and 2, respectively. Hence, the encoder, $f : (\mathcal{W}_1, \mathcal{W}_2) \rightarrow \mathcal{X}^n$, maps two messages $w_1 \in \mathcal{W}_1$ and $w_2 \in \mathcal{W}_2$ to a codeword $x^n \in \mathcal{X}^n$. The helper now assists both receivers to deal with the channel state. The two decoders with each at one receiver, $g_k : \mathcal{Y}_k^n \rightarrow \mathcal{W}_k$, maps a received sequence y_k^n into a message $\hat{w}_k \in \mathcal{W}_k$ for $k = 1, 2$.

We study the Gaussian state-dependent broadcast channel, in which the outputs at the

two receivers for one channel use are given by

$$Y_1 = X_0 + S + \frac{1}{a}(X + Z_1), \quad (7.1a)$$

$$Y_2 = X_0 + S + X + Z_2. \quad (7.1b)$$

where the noise variables Z_1 and Z_2 and the state variables S are Gaussian distributed with distributions $Z_1 \sim \mathcal{N}(0, N_1)$, $Z_2 \sim \mathcal{N}(0, N_2)$ and $S \sim \mathcal{N}(0, Q)$, and all of these variables are independent and are i.i.d. over channel uses. The channel inputs X_0 and X are subject to the average power constraints $\frac{1}{n} \sum_{i=1}^n X_{0i}^2 \leq P_0$ and $\frac{1}{n} \sum_{i=1}^n X_i^2 \leq P$.

7.2 Scenario I: Common Message

In this section, we study scenario I, in which only one common message is transmitted from the transmitter to both receivers. We first derive a useful upper bound.

Proposition 7.1. *For the state-dependent Gaussian broadcast channel in scenario I, an upper bound on the capacity is given by*

$$\begin{aligned} R \leq & \min \left\{ \frac{1}{2} \log \left(1 + \frac{P}{N_1} \right), \frac{1}{2} \log \left(1 + \frac{P}{N_2} \right), \right. \\ & \frac{1}{2} \log \left(1 + \frac{a^2 P_0}{N_1} \right) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\sqrt{P_0 Q} + \frac{1}{a^2}(P + N_1)}{Q} \right), \\ & \left. \frac{1}{2} \log \left(1 + \frac{P_0}{N_2} \right) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\sqrt{P_0 Q} + P + N_2}{Q} \right) \right\}. \end{aligned} \quad (7.2)$$

We note that, in (7.2), the first two terms represent the capacity for the compound channel without state. The third and fourth terms equal to the best single-user rates of receivers 1 and 2, respectively, with the helper dedicated to help each receiver, which can be reduced from the result in Proposition 5.1.

We next derive an achievable rate based on the dirty interference cancelation scheme, in which the helper incorporates two schemes for canceling state interference: scheme 1 can-

cels some state power by a signal that exactly reverses the state realization; and scheme 2 uses dirty paper coding via generation of an auxiliary variable (represented by U in Proposition 7.2) to incorporate the state information so that the receiver decodes such variable first to cancel the state and then decode the users' information. We first provide an achievable region for the discrete memoryless channel in the following proposition.

Proposition 7.2. *For the state-dependent broadcast channel with a helper, a rate R is achievable if it satisfies*

$$R \leq I(X; Y_k | U) \text{ for } k=1, 2, \quad (7.3)$$

for some distribution $P_{SUX_0X} = P_S P_{U|X_0} P_X$, where U is an auxiliary random variable such that

$$I(U; Y_k) \geq I(U; S) \text{ for } k=1, 2. \quad (7.4)$$

Proof. The achievable region follows from a coding scheme in which the state is encoded using a single-bin coding at the helper, and a successive cancellation at each receivers. This is similar to the coding scheme in Proposition 5.2 for each receiver. \square

Following from Proposition 7.2, we obtain an achievable rate for the Gaussian channel.

Proposition 7.3. *For the state-dependent Gaussian broadcast channel in scenario I, a rate R is achievable if it satisfies*

$$R \leq \min \left\{ \frac{1}{2} \log \left(1 + \frac{P}{\frac{(1+\beta-\alpha)^2 a^2 P_{00} Q}{P_{00} + \alpha^2 Q} + N_1} \right), \quad (7.5a)$$

$$\frac{1}{2} \log \left(1 + \frac{P}{\frac{(1+\beta-\alpha)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + N_2} \right) \right\}, \quad (7.5b)$$

where $P_{00} + \beta^2 Q \leq P_0$, $P_{00} \geq 0$,

$$\alpha^2 Q \frac{P + N_1}{a^2} + \alpha^2 P_{00} Q - 2\alpha(1 + \beta)P_{00} Q \leq P_{00}^2, \quad \text{and}$$

$$\alpha^2 Q(P + N_2 + P_{00}) - 2\alpha(1 + \beta)P_{00} Q \leq P_{00}^2.$$

Proof. The achievability follows from Proposition 7.2 by choosing jointly Gaussian distribution as follows:

$$U = X_{00} + \alpha S, \quad X_0 = X_{00} + \beta S$$

$$X_{00} \sim \mathcal{N}(0, P_{00}), \quad X \sim \mathcal{N}(0, P)$$

where X_{00} , X and S are independent. □

Comparing the lower and upper bounds given in Propositions 7.3 and 7.1, respectively, we characterize the capacity for three ranges of channel parameters, respectively, in the following three theorems.

Theorem 7.1. *For the state-dependent Gaussian broadcast channel in scenario I, if $P_0 \geq \max\{P + N_2, \frac{P+N_1}{a^2}\}$, the capacity is given by*

$$C = \min \left\{ \frac{1}{2} \log \left(1 + \frac{P}{N_1} \right), \frac{1}{2} \log \left(1 + \frac{P}{N_2} \right) \right\}. \quad (7.6)$$

Proof. When $P_0 \geq \max\{P + N_2, \frac{P+N_1}{a^2}\}$, by setting $\alpha = 1 + \beta$ for (7.5a) and (7.5b), the rate in (7.6) is achievable which matches the outer bound in Proposition 7.1, and hence is the capacity rate. □

Theorem 7.1 indicates that when the helper's power is large enough, it can help the receivers to fully cancel the state interference. In particular, this holds even when the state power is arbitrarily large. This is very useful as finite amount of helper's power can help to

cancel infinite amount of interference power. We next consider the case when the helper's power is below a certain threshold.

Theorem 7.2. *For the state-dependent Gaussian broadcast channel in scenario I, if $P_0 < \max\{P + N_2, \frac{P+N_1}{a^2}\}$, and*

$$\sqrt{Q} \leq \max_{0 \leq P_{00} \leq P_0} \sqrt{P_0 - P_{00}} + \frac{P_{00}}{\sqrt{\max\{\frac{P+N_1}{a^2}, P + N_2\} - P_{00}}}, \quad (7.7)$$

the channel capacity is given by

$$C = \min\left\{\frac{1}{2} \log\left(1 + \frac{P}{N_1}\right), \frac{1}{2} \log\left(1 + \frac{P}{N_2}\right)\right\}. \quad (7.8)$$

Proof. When (7.7) is satisfied, by setting $\alpha = 1 + \beta$ in Proposition 7.3, the rate in (7.8) is achieved, which matches the upper bound in Proposition 7.1. Hence, the capacity rate is obtained. \square

Theorem 7.2 implies that when the helper's power is below a certain threshold, only a limited power of state interference can be fully canceled with the assistance of the helper. We note that such power Q of the state can still be larger than the helper's power P_0 , which implies that the combined scheme in Proposition 7.2 is necessary to fully cancel the state interference. One example of such channel parameters can be given by $\max\{P + N_2, \frac{P+N_1}{a^2}\} = 7$, $P_0 = 4.5$, and $Q = 8$.

We note that if the state power is asymptotically large, the upper bound in Proposition 7.1 (and hence the capacity) can be determined only by the helper's power, as summarized in the following theorem. It is also clear that when $Q \rightarrow \infty$, with limited helper's power, direct cancellation does not lead to any positive transmission rate, and dirty paper coding is necessary for state cancelation.

Theorem 7.3. *For the state-dependent Gaussian broadcast channel in scenario I, suppose*

$Q \rightarrow \infty$, and $P \geq \max\{a^2 P_0 + N_1, P_0 + N_2\}$. If the channel parameters satisfy

$$\begin{aligned} a^2 P_0 + N_1 + N_2 &\leq P_0 + \frac{2}{a^2} N_1, \quad \text{and} \\ N_1^2(1 - a^2) &\leq a^4 P_0(N_1 - N_2), \end{aligned} \quad (7.9)$$

then, the channel capacity is given by $C = \frac{1}{2} \log(1 + \frac{a^2 P_0}{N_1})$. Furthermore, if the channel parameters satisfy

$$\begin{aligned} \frac{1}{a^2}(P_0 + N_1 + N_2) &\leq P_0 + 2N_2, \quad \text{and} \\ P_0(N_2 - N_1) &\geq (a^2 - 1)N_2^2, \end{aligned} \quad (7.10)$$

then, the channel capacity is given by $C = \frac{1}{2} \log(1 + \frac{P_0}{N_2})$.

Proof. When (7.9) is satisfied, by setting $\tilde{P} = a^2 P_0 + N_1$ and $\alpha = \frac{a^2 P_0}{a^2 P_0 + N_1}$, the rate $C = \frac{1}{2} \log(1 + \frac{a^2 P_0}{N_1})$ is achieved, which matches with the outer bound. Hence, the capacity rate is obtained.

Similarly, when (7.10) is satisfied, by setting $\tilde{P} = P_0 + N_2$ and $\alpha = \frac{P_0}{P_0 + N_2}$, the rate $C = \frac{1}{2} \log(1 + \frac{P_0}{N_2})$ is achieved, which matches the outer bound. Hence, the capacity rate is obtained. \square

We note that the two ranges of channel parameters in Theorem 7.3 respectively correspond to the cases with the channel performance bounded by receivers 1 and 2.

7.3 Scenario II: Private Messages

In this section, we study scenario II, in which the transmitter sends two independent messages to the two receivers, respectively. Without loss of generality, we assume that $N_1 \geq N_2$, which implies that in the original broadcast channel without state, receiver 1's channel quality is worse than receiver 2. We first derive an outer bound on the capacity region.

Proposition 7.4. *For the state-dependent Gaussian broadcast channel in scenario II with $N_1 \geq N_2$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \min \left\{ \frac{1}{2} \log \left(1 + \frac{P_1}{P - P_1 + N_1} \right), \right. \\ \left. \frac{1}{2} \log \left(1 + \frac{a^2 P_0}{N_1} \right) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\sqrt{P_0 Q} + \frac{1}{a^2}(P + N_1)}{Q} \right) \right\}, \quad (7.11a)$$

$$R_2 \leq \min \left\{ \frac{1}{2} \log \left(1 + \frac{P - P_1}{N_2} \right), \right. \\ \left. \frac{1}{2} \log \left(1 + \frac{P_0}{N_2} \right) + \frac{1}{2} \log \left(1 + \frac{P_0 + 2\sqrt{P_0 Q} + P + N_2}{Q} \right) \right\}, \quad (7.11b)$$

where $0 \leq P_1 \leq P$.

The outer bound for each rate consists of two bounds. The first one is based on the capacity region of the Gaussian broadcast channel without state. The second one is the best single-user rate with the helper dedicated to help each receiver, which can be reduced from the result in [47].

We then derive the following achievable region based on the helper employing the dirty interference cancellation scheme as for scenario I. Furthermore, superposition coding is used for broadcasting two messages.

Proposition 7.5. *For the state-dependent broadcast channel in scenario II, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq I(V; Y_1|U), \quad (7.12a)$$

$$R_2 \leq I(X; Y_2|UV), \quad (7.12b)$$

$$R_1 + R_2 \leq I(X; Y_2|U), \quad (7.12c)$$

for some distribution $P_S P_{U|S} P_{X_0|SU} P_V P_{X|V}$, where $I(U; Y_k) \geq I(U; S)$ for $k = 1, 2$.

Following from Proposition 7.5, we obtain the following achievable rate region for the Gaussian channel.

Proposition 7.6. *For the state-dependent Gaussian broadcast channel in scenario II with $N_1 \geq N_2$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{\frac{(1+\beta-\alpha)^2 a^2 P_{00} Q}{P_{00} + \alpha^2 Q} + P - P_1 + N_1} \right), \quad (7.13a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P - P_1}{\frac{(1+\beta-\alpha)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + N_2} \right), \quad (7.13b)$$

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P}{\frac{(1+\beta-\alpha)^2 P_{00} Q}{P_{00} + \alpha^2 Q} + N_2} \right), \quad (7.13c)$$

where $P_{00} + \beta^2 Q \leq P_0$, $P_{00} \geq 0$, $0 \leq P_1 \leq P$,

$$\alpha^2 Q \frac{P + N_1}{a^2} + \alpha^2 P_{00} Q - 2\alpha(1 + \beta)P_{00} Q \leq P_{00}^2, \quad \text{and}$$

$$\alpha^2 Q(P + N_2 + P_{00}) - 2\alpha(1 + \beta)P_{00} Q \leq P_{00}^2.$$

Proof. The proof follows from Proposition 7.5 by choosing jointly Gaussian distribution as follows:

$$U = X_{00} + \alpha S, \quad X_0 = X_{00} + \beta S$$

$$X = V + X', \quad X_{00} \sim \mathcal{N}(0, P_{00})$$

$$V \sim \mathcal{N}(0, P_1), \quad X' \sim \mathcal{N}(0, P - P_1)$$

where X_{00} , V , X' and S are independent. □

By comparing the outer and inner bounds, we characterize the capacity region for two ranges of channel parameters.

Theorem 7.4. *For the state-dependent Gaussian broadcast channel in scenario II with $N_1 \geq N_2$, if $P_0 \geq \max\{P + N_2, \frac{P+N_1}{a^2}\}$, the capacity region consists of rate pairs (R_1, R_2)*

satisfying:

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{P - P_1 + N_1} \right), \quad (7.14a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P - P_1}{N_2} \right). \quad (7.14b)$$

Proof. When $P_0 \geq \max\{P + N_2, \frac{P+N_1}{a^2}\}$, by setting $\alpha = 1 + \beta$ for (7.13a)- (7.13c), the region in (7.14a) and (7.14b) is achievable which matches the outer bound in Proposition 7.4, and hence is the capacity region. \square

Similarly to Theorem 7.1, Theorem 7.4 implies that when the helper's power is larger than a certain threshold, the state is fully canceled with the assistance of the helper, and the state power can be arbitrarily large. Thus, the capacity region of the corresponding channel without state is achieved.

We next study the case with the helper's power being smaller than a threshold following similar step in Theorem 7.2.

Theorem 7.5. *For the state-dependent Gaussian broadcast channel in scenario II with $N_1 \geq N_2$, if $P_0 < \max\{P + N_2, \frac{P+N_1}{a^2}\}$, and*

$$\sqrt{Q} \leq \max_{0 \leq P_{00} \leq P_0} \sqrt{P_0 - P_{00}} + \frac{P_{00}}{\sqrt{\max\{\frac{P+N_1}{a^2}, P + N_2\} - P_{00}}}, \quad (7.15)$$

the capacity region consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{P - P_1 + N_1} \right), \quad (7.16a)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P - P_1}{N_2} \right). \quad (7.16b)$$

Proof. When (7.15) is satisfied, by setting $\alpha = 1 + \beta$ for (7.13a)- (7.13c), the region in (7.16a) and (7.16b) is achieved, which matches the outer bound in Proposition 7.4. Hence, the capacity region is obtained. \square

Theorem 7.5 implies that if the helper's power is not large enough, only the state with

limited power can be fully cancelled to result in the capacity region of the corresponding broadcast channel without state. Nevertheless, such state power can still be larger than the helper's power demonstrating necessity of using dirty paper coding. One example of such channel parameters is given by $\max\{P + N_2, \frac{P+N_1}{a^2}\} = 7.5$, $P_0 = 5$, and $Q = 9$.

CHAPTER 8

CONCLUSION

In this thesis, we studied the state-dependent interference channels in two classes. One class of models, including the state-dependent interference channel and the state-dependent cognitive interference channel, capture the scenarios that the state cancellation and the message transmission are performed by the same node. The other class of models, including the state-dependent single-user channel with a helper, the state-dependent parallel channel with a common helper, the state-dependent MAC with a helper, and the state-dependent broadcast channel with a helper, capture the scenarios that the state cancellation is performed by a separate helper. For each channel model, we derived the inner and outer bounds on the capacity region, and characterized the capacity partially/fully for various channel parameters. In particular, for the second class of models, our results demonstrate that the capacity region is not only bounded by the transmitter's power, but also by the helper's power. This suggests that the state cannot always be perfectly cancelled.

This thesis demonstrates that interference in wireless networks can be effectively canceled by its source node via dirty interference cancelation. Thus, users can transmit simultaneously as well as enjoy low or no interference transmission environments. In this way, dirty interference cancelation is very promising to substantially improve the performance of wireless networks. Future work can be focused on three aspects:

1. For the MAC and broadcast channel, we derived inner and outer bounds, and characterized capacity region for channel with various channel parameters. In particular, the outer bound is only tight for the regimes that the state can be perfectly cancelled, and fail to characterize the relationship among the capacity region, the helper's power and the state power. In the future, we will develop more sophisticated outer bounds, and study how does the helper and the state power influence the channel capacity.
2. Since dirty interference cancelation provide a new technique that enables simultaneous transmission, it is also interesting to compare our dirty interference cancellation scheme with the conventional interference management techniques based on the orthogonality idea. We will compare the achievable region in the model consisting of a base station and a D2D transmitter, sending two messages to their corresponding receivers. The base station has high transmission power and interferes the D2D receiver. We assume that the cellular and D2D transmissions do not share codebooks and hence Han-Kobayashi rate splitting cannot be applied. For such a model, we compare the performance of two schemes: dirty interference cancellation and the orthogonalized transmission via time sharing.
3. Since D2D communications can be diversified, and can include multi-access transmissions and broadcast transmissions, we will also extend the model to include multiple D2D user pairs with more complex structures. For such a scenario, we will compare the sum rate over the cellular receiver and the D2D receivers for the two schemes: dirty interference cancellation, and time sharing scheme. In particular, the sum rate depends on the locations of the D2D receivers from the base station. We will assume that the D2D receivers are located uniformly and independently over a certain range, and derive average sum rates for the two schemes. Comparison of the two schemes will provide us the gain that dirty interference cancelation yields on average.

APPENDIX A

PROOF FOR CHAPTER 2

A.1 Proof of Proposition 2.1

We use random codes and fix the following joint distribution:

$$P_{SUX_1VX_2Y_1Y_2} = P_S P_{U|S} P_{X_1|US} P_{V|S} P_{X_2|VS} P_{Y_1Y_2|X_1X_2S}.$$

Let $T_\epsilon^n(P_{SUX_1VX_2Y_1Y_2})$ denote the strongly joint ϵ -typical set based on the above distribution.

Code Construction:

1. Generate $2^{n(R_1+R'_1)}$ codewords $U^n(w_1, l_1)$ with i.i.d. components based on P_U . Index these codewords by $w_1 = 1, \dots, 2^{nR_1}$. $l_1 = 1, 2, \dots, 2^{nR'_1}$.
2. Generate $2^{n(R_2+R'_2)}$ codewords $V^n(w_2, l_2)$ with i.i.d. components based on P_V . Index these codewords by $w_2 = 1, \dots, 2^{nR_2}$. $l_2 = 1, 2, \dots, 2^{nR'_2}$.

Encoding:

1. Encoder 1: Given w_1 , and s^n , select $u^n(w_1, \tilde{l}_1)$ such that

$$(u^n(w_1, \tilde{l}_1), s^n) \in T_\epsilon^n(P_{US}).$$

Otherwise, set $\tilde{l}_1 = 1$. It can be shown that for large n , such u^n exists with high probability if

$$R'_1 > I(U; S). \quad (\text{A.1})$$

Given selected $u^n(w_1, \tilde{l}_1)$ and s^n , generate x_1^n with i.i.d. components based on $P_{X_1|US}$ for transmission.

2. Encoder 2: Given w_2 , and s^n , select $v^n(w_2, \tilde{l}_2)$ such that

$$(v^n(w_2, \tilde{l}_2), s^n) \in T_\epsilon^n(P_{VS}).$$

Otherwise, set $\tilde{l}_2 = 1$. It can be shown that for large n , such v^n exists with high probability if

$$R'_2 > I(V; S). \quad (\text{A.2})$$

Given selected $v^n(w_2, \tilde{l}_2)$ and s^n , generate x_2^n with i.i.d. components based on $P_{X_2|VS}$ for transmission.

Decoding:

1. Decoder 1: Given y_1^n , find the unique pair (\hat{w}_2, \hat{l}_2) such that

$$(v^n(\hat{w}_2, \hat{l}_2), y_1^n) \in T_\epsilon^n(P_{VY_1}).$$

If no or more than one such pairs (\hat{w}_2, \hat{l}_2) can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_2 + R'_2 \leq I(V; Y_1). \quad (\text{A.3})$$

After successfully decoding v^n , find the unique pair (\hat{w}_1, \hat{l}_1) such that

$$(v^n(\hat{w}_2, \hat{l}_2), u^n(\hat{w}_1, \hat{l}_1), y_1^n) \in T_\epsilon^n(P_{VUY_1}).$$

If no or more than one such pairs with different w_1 can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_1 + R'_1 \leq I(U; VY_1). \quad (\text{A.4})$$

2. Decoder 2: Given y_2^n , find the unique pair (\hat{w}_1, \hat{l}_1) such that

$$(u^n(\hat{w}_1, \hat{l}_1), y_2^n) \in T_\epsilon^n(P_{UY_2}).$$

If no or more than one such pairs (\hat{w}_1, \hat{l}_1) can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_1 + R'_1 \leq I(U; Y_2). \quad (\text{A.5})$$

After successfully decoding u^n , find the unique pair (\hat{w}_2, \hat{l}_2) such that

$$(u^n(\hat{w}_1, \hat{l}_1), v^n(\hat{w}_2, \hat{l}_2), y_2^n) \in T_\epsilon^n(P_{UVY_2}).$$

If no or more than one such pairs with different w_2 can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_2 + R'_2 \leq I(V; UY_2). \quad (\text{A.6})$$

Proposition 2.1 is thus proved by combining (A.1)-(A.6).

A.2 Proof of Proposition 2.2

The coding scheme for the very strong Z-IC is similar to that for the regular IC. More specifically, codebook generation, encoding and decoding for decoder 1 are the same as those in Appendix A.1. We next describe decoding for decoder 2 as follows.

Decoding for decoder 2:

Given y_2^n , find the unique pair (\hat{w}_2, \hat{l}_2) such that

$$(v^n(\hat{w}_2, \hat{l}_2), y_2^n) \in T_\epsilon^n(P_{VY_2}).$$

If no or more than one such pairs with different w_2 can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability if

$$R_2 + R'_2 \leq I(V; Y_2).$$

If $I(V; Y_2) \leq I(V; Y_1)$, then the bound $R_2 + R'_2 \leq I(V; Y_1)$ obtained in decoding for decoder 1 (see (A.3)) is redundant. Hence, the corresponding achievable region is as given in Proposition 2.2.

A.3 Proof of Proposition 2.3

The achievable scheme applies rate splitting, superposition and Gel'fand-Pinsker binning. In particular, we split the message W_1 into two components W_{11} and W_{12} , and split W_2 into two components W_{21} and W_{22} . We use random codes and fix the following joint

distribution:

$$P_{SU_1U_2X_1V_1V_2X_2Y_1Y_2} = P_S P_{U_1U_2|S} P_{X_1|U_1U_2S} P_{V_1V_2|S} P_{X_2|V_1V_2S} P_{Y_1Y_2|X_1X_2S}.$$

Code Construction:

1. Generate $2^{n(R_{11}+R'_{11})}$ codewords $U_1^n(w_{11}, l_{11})$ with i.i.d. components based on P_{U_1} . Index these codewords by $w_{11} = 1, \dots, 2^{nR_{11}}, l_{11} = 1, 2, \dots, 2^{nR'_{11}}$.
2. For each $u_1^n(w_{11}, l_{11})$, generate $2^{n(R_{12}+R'_{12})}$ codewords $U_2^n(w_{11}, l_{11}, w_{12}, l_{12})$ with i.i.d. components based on $P_{U_2|U_1}$. Index these codewords by $w_{12} = 1, \dots, 2^{nR_{12}}, l_{12} = 1, 2, \dots, 2^{nR'_{12}}$.
3. Generate $2^{n(R_{21}+R'_{21})}$ codewords $V_1^n(w_{21}, l_{21})$ with i.i.d. components based on P_{V_1} . Index these codewords by $w_{21} = 1, \dots, 2^{nR_{21}}, l_{21} = 1, 2, \dots, 2^{nR'_{21}}$.
4. For each $v_1^n(w_{21}, l_{21})$, generate $2^{n(R_{22}+R'_{22})}$ codewords $V_2^n(w_{21}, l_{21}, w_{22}, l_{22})$ with i.i.d. components based on $P_{V_2|V_1}$. Index these codewords by $w_{22} = 1, \dots, 2^{nR_{22}}, l_{22} = 1, 2, \dots, 2^{nR'_{22}}$.

Encoding:

1. Encoder 1: Given w_{11} , and s^n , select $u_1^n(w_{11}, \tilde{l}_{11})$ such that

$$(u_1^n(w_{11}, \tilde{l}_{11}), s^n) \in T_\epsilon^n(P_{U_1S}).$$

Otherwise, set $\tilde{l}_{11} = 1$. It can be shown that for large n , such u_1^n exists with high probability if

$$R'_{11} > I(U_1; S). \quad (\text{A.7})$$

Given w_{12} , w_{11} , \tilde{l}_{11} , and s^n , select $u_2^n(w_{11}, \tilde{l}_{11}, w_{12}, \tilde{l}_{12})$ such that

$$(u_2^n(w_{11}, \tilde{l}_{11}, w_{12}, \tilde{l}_{12}), u_1^n(w_{11}, \tilde{l}_{11}), s^n) \in T_\epsilon^n(P_{U_2SU_1}).$$

Otherwise, set $\tilde{l}_{12} = 1$. It can be shown that for large n , such u_2^n exists with high probability if

$$R'_{12} > I(U_2; S|U_1). \quad (\text{A.8})$$

Given $u_1^n(w_{11}, \tilde{l}_{11})$, $u_2^n(w_{11}, \tilde{l}_{11}, w_{12}, \tilde{l}_{12})$, and s^n , generate x_1^n with i.i.d. components based on $P_{X_1|U_1U_2S}$ for transmission.

2. Encoder 2: Given w_{21} , and s^n , select $v_1^n(w_{21}, \tilde{l}_{21})$ such that

$$(v_1^n(w_{21}, \tilde{l}_{21}), s^n) \in T_\epsilon^n(P_{V_1S}).$$

Otherwise, set $\tilde{l}_{21} = 1$. It can be shown that for large n , such v_1^n exists with high probability if

$$R'_{21} > I(V_1; S). \quad (\text{A.9})$$

Given w_{22} , w_{21} , \tilde{l}_{21} and s^n , select $v_2^n(w_{21}, \tilde{l}_{21}, w_{22}, \tilde{l}_{22})$ such that

$$(v_2^n(w_{21}, \tilde{l}_{21}, w_{22}, \tilde{l}_{22}), v_1^n(w_{21}, \tilde{l}_{21}), s^n) \in T_\epsilon^n(P_{V_2SV_1}).$$

Otherwise, set $\tilde{l}_{22} = 1$. It can be shown that for large n , such v_2^n exists with high probability if

$$R'_{22} > I(V_2; S|V_1). \quad (\text{A.10})$$

Given $v_1^n(w_{21}, \tilde{l}_{21})$, $v_2^n(w_{21}, \tilde{l}_{21}, w_{22}, \tilde{l}_{22})$ and s^n , generate x_2^n with i.i.d. components based on $P_{X_2|V_1V_2S}$ for transmission.

Decoding:

1. Decoder 1: Given y_1^n , find the unique pair $(\hat{w}_{11}, \hat{l}_{11})$ such that

$$(u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \in T_\epsilon^n(P_{U_1Y_1}).$$

If no or more than one such pairs with different w_{11} can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{11} + R'_{11} \leq I(U_1; Y_1).$$

After successfully decoding u_1^n , find the unique tuple $(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22})$ such that

$$\begin{aligned} & (v_1^n(\hat{w}_{21}, \hat{l}_{21}), v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \\ & \in T_\epsilon^n(P_{V_1 V_2 U_1 Y_1}). \end{aligned}$$

If no or more than one such tuples with different rate pairs (w_{21}, w_{22}) can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$\begin{aligned} R_{21} + R'_{21} & \leq I(V_1; U_1 Y_1) \\ R_{22} + R'_{22} & \leq I(V_2; U_1 Y_1 | V_1) \end{aligned}$$

After successfully decoding u_1^n , v_1^n and v_2^n , we find the unique pair $(\hat{w}_{12}, \hat{l}_{12})$ such that

$$\begin{aligned} & (u_2^n(\hat{w}_{11}, \hat{l}_{11}, \hat{w}_{12}, \hat{l}_{12}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \\ & \in T_\epsilon^n(P_{U_2 U_1 V_1 V_2 Y_1}). \end{aligned}$$

If no or more than one such pair with different w_{12} can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability

if

$$R_{12} + R'_{12} \leq I(U_2; V_1 V_2 Y_1 | U_1) \quad (\text{A.11})$$

2. Decoder 2: Given y_2^n , find the unique pair $(\hat{w}_{21}, \hat{l}_{21})$ such that

$$(v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_2^n) \in T_\epsilon^n(P_{V_1 Y_2}).$$

If no or more than one such pairs with different w_{21} can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability

if

$$R_{21} + R'_{21} \leq I(V_1; Y_2).$$

After successfully decoding v_1^n , find the unique tuple $(\hat{w}_{11}, \hat{l}_{11}, \hat{w}_{12}, \hat{l}_{12})$ such that

$$(u_1^n(\hat{w}_{11}, \hat{l}_{11}), u_2^n(\hat{w}_{11}, \hat{l}_{11}, \hat{w}_{12}, \hat{l}_{12}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_2^n) \in T_\epsilon^n(P_{U_1 U_2 V_1 Y_2}).$$

If no or more than one such tuples with different rate pair (w_{11}, w_{12}) can be found,

then declare error. One can show that for sufficiently large n , decoding is correct

with high probability if

$$R_{11} + R'_{11} \leq I(U_1; V_1 Y_2)$$

$$R_{12} + R'_{12} \leq I(U_2; V_1 Y_2 | U_1)$$

After successfully decoding v_1^n , u_1^n and u_2^n , we find the unique pair $(\hat{w}_{22}, \hat{l}_{22})$ such

that

$$\begin{aligned} & (v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), u_2^n(\hat{w}_{11}, \hat{l}_{11}, \hat{w}_{12}, \hat{l}_{12}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_2^n) \\ & \in T_\epsilon^n(P_{V_2 V_1 U_1 U_2 Y_2}). \end{aligned}$$

If no or more than one such pair with different w_{22} can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{22} + R'_{22} \leq I(V_2; U_1 U_2 Y_2 | V_1) \quad (\text{A.12})$$

The corresponding achievable region is thus characterized by

$$\begin{aligned} R_{11} & \leq \min\{I(U_1; Y_1), I(U_1; V_1 Y_2)\} - I(U_1; S) \\ R_{12} & \leq \min\{I(U_2; V_1 V_2 Y_1 | U_1), I(U_2; V_1 Y_2 | U_1)\} - I(U_2; S | U_1) \\ R_{21} & \leq \min\{I(V_1; Y_2), I(V_1; U_1 Y_1)\} - I(V_1; S) \\ R_{22} & \leq \min\{I(V_2; U_1 U_2 Y_2 | V_1), I(V_2; U_1 Y_1 | V_1)\} - I(V_2; S | V_1) \end{aligned}$$

Proposition 6.2 follows by setting $R_1 = R_{11} + R_{12}$ and $R_2 = R_{21} + R_{22}$, and applying Fourier-Motzkin elimination to the above region.

A.4 Proof of Proposition 2.4

The coding scheme for the strong Z-IC is similar to that for the regular IC. More specifically, codebook generation and encoding for the strong Z-IC are the same as those for the regular IC provided in Appendix A.3. We next describe decoding as follows.

Decoding:

1. Decoder 1: Given y_1^n , find the unique pair $(\hat{w}_{21}, \hat{l}_{21})$ such that

$$(v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_1^n) \in T_\epsilon^n(P_{V_1 Y_1}).$$

If no or more than one such pairs with different w_{21} can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{21} + R'_{21} \leq I(V_1; Y_1).$$

After successfully decoding v_1^n , find the unique pair $(\hat{w}_{11}, \hat{l}_{11})$ such that

$$(v_1^n(\hat{w}_{21}, \hat{l}_{21}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \in T_\epsilon^n(P_{V_1 U_1 Y_1}).$$

If no or more than one such rate pairs with different w_{11} can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{11} + R'_{11} \leq I(U_1; V_1 Y_1).$$

After successfully decoding u_1^n and v_1^n we find the unique pair $(\hat{w}_{22}, \hat{l}_{22})$ such that

$$\begin{aligned} (v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \\ \in T_\epsilon^n(P_{U_1 V_1 V_2 Y_1}). \end{aligned}$$

If no or more than one such pairs with different w_{22} can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability

if

$$R_{22} + R'_{22} \leq I(V_2; U_1 Y_1 | V_1) \quad (\text{A.13})$$

After successfully decoding u_1^n , v_1^n and v_2^n , we find the unique pair $(\hat{w}_{12}, \hat{l}_{12})$ such that

$$\begin{aligned} & (u_2^n(\hat{w}_{11}, \hat{l}_{11}, \hat{w}_{12}, \hat{l}_{12}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), u_1^n(\hat{w}_{11}, \hat{l}_{11}), y_1^n) \\ & \in T_\epsilon^n(P_{U_2 U_1 V_1 V_2 Y_1}). \end{aligned}$$

If no or more than one such pairs with different w_{12} can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{12} + R'_{12} \leq I(U_2; V_1 V_2 Y_1 | U_1). \quad (\text{A.14})$$

2. Decoder 2: Given y_2^n , find the unique pair $(\hat{w}_{21}, \hat{l}_{21})$ such that

$$(v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_2^n) \in T_\epsilon^n(P_{V_1 Y_2}).$$

If no or more than one such pairs with different w_{21} can be found, then declare error. One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{21} + R'_{21} \leq I(V_1; Y_2).$$

After successfully decoding v_1^n , find the unique pair $(\hat{w}_{22}, \hat{l}_{22})$ such that

$$(v_2^n(\hat{w}_{21}, \hat{l}_{21}, \hat{w}_{22}, \hat{l}_{22}), v_1^n(\hat{w}_{21}, \hat{l}_{21}), y_2^n) \in T_\epsilon^n(P_{V_2 V_1 Y_2}).$$

If no or more than one such pairs with different w_{22} can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability if

$$R_{22} + R'_{22} \leq I(V_2; Y_2 | V_1). \quad (\text{A.15})$$

The corresponding achievable region is thus characterized by

$$R_{11} \leq I(U_1; Y_1 V_1) - I(U_1; S)$$

$$R_{12} \leq I(U_2; V_1 V_2 Y_1 | U_1) - I(U_2; S | U_1)$$

$$R_{21} \leq \min\{I(V_1; Y_2), I(V_1; Y_1)\} - I(V_1; S)$$

$$R_{22} \leq \min\{I(V_2; Y_2 | V_1), I(V_2; U_1 Y_1 | V_1)\} - I(V_2; S | V_1).$$

Proposition 2.4 then follows by setting $R_1 = R_{11} + R_{12}$ and $R_2 = R_{21} + R_{22}$, and applying Fourier-Motzkin elimination to the above region.

A.5 Proof of Proposition 2.5

Assume $P'_{1B'}$, $P''_{1B'}$, $P'_{2B'}$ and $P''_{2B'}$ are power allocation parameters corresponding to the given point B' under which the conditions in (2.24a) and (2.24b) are satisfied. In order to prove that the point B is also achievable, we design the following coding scheme. We split W_1 into W_{11} and W_{12} , and split W_2 into W_{21} and W_{22} . We then encode the messages W_{11} , W_{12} , W_{21} and W_{22} into auxiliary random variables U_1 , U_2 , V_1 , and V_2 , respectively. Then receiver 1 decodes in the order of V_1 , V_2 , U_1 and U_2 , and receiver 2 decodes in the order of

V_1 and V_2 . It can be shown that (R_1, R_2) is achievable if it satisfies

$$\begin{aligned} R_1 &\leq I(U_1; Y_1 V_1 V_2) + I(U_2; V_1 V_2 Y_1 | U_1) - I(U_1, U_2; S) \\ R_2 &\leq \min\{I(V_1; Y_2), I(V_1; Y_1)\} \\ &\quad + \min\{I(V_2; Y_2 | V_1), I(V_2; Y_1 | V_1)\} - I(V_1 V_2; S) \end{aligned} \quad (\text{A.16})$$

for some distribution $P_{SU_1 U_2 V_1 V_2 X_2 X_1 Y_2 Y_1} = P_S P_{U_1 U_2 | S} P_{V_1 V_2 | S} P_{X_1 | U_1 U_2 S} P_{X_2 | V_1 V_2 S} P_{Y_1 | S X_1 X_2} P_{Y_2 | S X_2}$. We now compute (A.16) by setting the auxiliary random variables as in (2.19), with the power allocations $P'_{1B'}$, $P''_{1B'}$, $P'_{2B'}$ and $P''_{2B'}$ for X'_1 , X''_1 , X'_2 and X''_2 in U_1 , U_2 , V_1 and V_2 , respectively. It can be verified that due to (2.24a) and (2.24b) that the power allocation parameters satisfy, the two mutual information terms $I(V_1; Y_2)$ and $I(V_2; Y_2 | V_1)$ in R_2 become redundant. It can then be verified that the rate pair corresponding to the point B satisfies the resulting (A.16), and is hence achievable. Thus, the line $B - B'$ is achievable by time sharing.

A.6 Proof of Corollary 2.2

It is sufficient to show that the point B' satisfies Theorem 2.4, i.e., it is on the capacity region boundary. Then following Proposition 2.5, the line $B - B'$ is on the capacity region boundary. It can be verified that the point B' is characterized by (2.17) by setting $P'_2 = 0$, $P''_2 = P_2$, P''_1 to satisfy

$$1 + \frac{a^2 P_2}{P''_1 + 1} \leq \frac{a^2 P_2 (P_2 + b^2 Q + 1)}{P_2 Q (ab - \beta)^2 + a^2 P_2 + \beta^2 Q}, \quad (\text{A.17})$$

and $P'_1 = P_1 - P''_1$. Then it can be verified that the condition (2.24a) and (2.24b) in Theorem 2.4 are satisfied by the point B' .

A.7 Proof of Theorem 2.5

Similarly to [4], [5] and [6], to achieve the sum capacity for the state-dependent Gaussian IC, we apply dirty paper coding for X_1 treating $aX_2 + N_1$ as noise and apply dirty paper coding for X_2 treating $bX_1 + N_1$ as noise. Thus, the point $(R_1, R_2) = (\frac{1}{2} \log(1 + \frac{P_1}{a^2 P_2 + 1}), \frac{1}{2} \log(1 + \frac{P_2}{b^2 P_1 + 1}))$ can be achieved.

For the outer bound, applying Fano's inequality, we have

$$\begin{aligned}
 nR_1 &\leq I(W_1; Y_1^n) + n\epsilon_n \\
 &\leq I(W_1; Y_1^n S^n) + n\epsilon_n \\
 &= I(W_1; Y_1^n | S^n) + n\epsilon_n \\
 &\leq I(W_1 X_1^n; Y_1^n | S^n) + n\epsilon_n \\
 &= I(X_1^n; Y_1^n | S^n) + I(W_1; Y_1^n | S^n X_1^n) + n\epsilon_n \\
 &= I(X_1^n; Y_1^n | S^n) + n\epsilon_n \\
 &= I(X_1^n; X_1^n + aX_2^n + S^n + N_1^n | S^n) + n\epsilon_n \\
 &= I(X_1^n; X_1^n + aX_2^n + N_1^n | S^n) + n\epsilon_n \\
 &= \sum_{s^n} p(S^n = s^n) I(X_1^n; X_1^n + aX_2^n + N_1^n | S^n = s^n) + n\epsilon_n. \tag{A.18}
 \end{aligned}$$

Similarly, we have

$$nR_2 \leq \sum_{s^n} p(S^n = s^n) I(X_2^n; bX_1^n + X_2^n + N_2^n | S^n = s^n) + n\epsilon_n. \tag{A.19}$$

Combining (A.18) and (A.19), we obtain

$$\begin{aligned}
& n(R_1 + R_2) \\
& \leq \sum_{s^n} p(S^n = s^n) \max_{P_{X_1^n|S^n} P_{X_2^n|S^n}} [I(X_1^n; X_1^n + aX_2^n + N_1^n | S^n = s^n) \\
& \quad + I(X_2^n; bX_1^n + X_2^n + N_2^n | S^n = s^n)] + 2n\epsilon_n \\
& = \sum_{s^n} p(S^n = s^n) \max_{P_{X_1^n} P_{X_2^n}} [I(X_1^n; X_1^n + aX_2^n + N_1^n) + I(X_2^n; bX_1^n + X_2^n + N_2^n)] + 2n\epsilon_n \\
& = \max_{P_{X_1^n} P_{X_2^n}} [I(X_1^n; X_1^n + aX_2^n + N_1^n) + I(X_2^n; bX_1^n + X_2^n + N_2^n)] + 2n\epsilon_n.
\end{aligned}$$

If $|a(1 + b^2 P_1)| + |b(1 + a^2 P_2)| \leq 1$, following the results in [5, Section IV.C], we further obtain

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{a^2 P_2 + 1} \right) + \frac{1}{2} \log \left(1 + \frac{P_2}{b^2 P_1 + 1} \right) + 2\epsilon_n.$$

Hence, the rate point $(R_1, R_2) = (\frac{1}{2} \log(1 + \frac{P_1}{a^2 P_2 + 1}), \frac{1}{2} \log(1 + \frac{P_2}{b^2 P_1 + 1}))$ is sum-rate optimal. Thus, the sum capacity is obtained.

APPENDIX B

PROOF FOR CHAPTER 3

B.1 Proof of the Outer Bound (3.8a)-(3.8d)

Consider a $(2^{nR_1}, 2^{nR_2}, n)$ code with an average error probability $P_e^{(n)}$. The probability distribution on $\mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{S}^n \times \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n$ is given by

$$P_{W_1 W_2 S^n X_1^n X_2^n Y_1^n Y_2^n} = P_{W_1} P_{W_2} \left[\prod_{i=1}^n P_{S_i} \right] \\ P_{X_1^n | W_1} P_{X_2^n | W_1 W_2 S^n} \prod_{i=1}^n P_{Y_{1i} Y_{2i} | X_{1i} X_{2i} S_i}.$$

By Fano's inequality, we have

$$H(W_1 | Y_1^n) \leq nR_1 P_e^{(n)} + 1 = n\delta_{1n} \\ H(W_1 W_2 | S^n Y_2^n) \leq n(R_1 + R_2) P_e^{(n)} + 1 = n\delta_{2n} \quad (\text{B.1})$$

where $\delta_{1n}, \delta_{2n} \rightarrow 0$ as $n \rightarrow +\infty$. Let $\delta_n = \delta_{1n} + \delta_{2n}$, which also satisfies that $\delta_n \rightarrow 0$ as $n \rightarrow +\infty$.

We define the following auxiliary random variables:

$$\begin{aligned} K_i &= (W_1, S_{i+1}^n, X_1^n, Y_1^{i-1}) \\ T_i &= Y_{2(i+1)}^n \end{aligned} \tag{B.2}$$

which satisfies the Markov chain condition:

$$K_i T_i \leftrightarrow X_{1i} X_{2i} S_i \leftrightarrow Y_{1i} Y_{2i} \tag{B.3}$$

for $i = 1, \dots, n$.

We first bound R_1 based on the Fano's inequality as follows:

$$\begin{aligned} nR_1 &\leq I(W_1; Y_1^n) + n\delta_n \\ &\stackrel{(a)}{=} \sum_{i=1}^n [I(W_1 S_{i+1}^n; Y_1^i) - I(W_1 S_i^n; Y_1^{i-1})] + n\delta_n \\ &\stackrel{(b)}{=} \sum_{i=1}^n [I(W_1 S_{i+1}^n; Y_1^{i-1}) + I(W_1 S_{i+1}^n; Y_{1i} | Y_1^{i-1}) \\ &\quad - I(W_1 S_{i+1}^n; Y_1^{i-1}) - I(S_i; Y_1^{i-1} | W_1 S_{i+1}^n)] + n\delta_n \\ &= \sum_{i=1}^n [I(W_1 S_{i+1}^n; Y_{1i} | Y_1^{i-1}) - I(S_i; Y_1^{i-1} | W_1 S_{i+1}^n)] + n\delta_n \\ &= \sum_{i=1}^n [H(Y_{1i} | Y_1^{i-1}) - H(Y_{1i} | W_1 S_{i+1}^n Y_1^{i-1}) \\ &\quad - H(S_i | W_1 S_{i+1}^n) + H(S_i | W_1 S_{i+1}^n Y_1^{i-1})] + n\delta_n \\ &\stackrel{(c)}{\leq} \sum_{i=1}^n [H(Y_{1i}) - H(Y_{1i} | W_1 S_{i+1}^n Y_1^{i-1} X_1^n) - (H(S_i | X_{1i}) \\ &\quad + H(S_i | W_1 S_{i+1}^n Y_1^{i-1} X_1^n))] + n\delta_n \\ &\stackrel{(d)}{\leq} \sum_{i=1}^n [I(K_i X_{1i}; Y_{1i}) - I(K_i; S_i | X_{1i})] + n\delta_n \end{aligned} \tag{B.4}$$

where (a) follows due to cancellation of the terms in the sum, and the fact that $Y_1^0 = \phi$, (b)

follows from the chain rule of mutual information, (c) follows because X_1^n is a function of W_1 , and (d) follows from the definition of K_i . The single letter characterization follows standard steps and is hence omitted.

We next bound R_2 as follows:

$$\begin{aligned}
nR_2 &\stackrel{(a)}{\leq} I(W_2; Y_2^n S^n) + n\delta_n \\
&\leq I(W_2; Y_2^n S^n W_1) + n\delta_n \\
&\stackrel{(b)}{\leq} I(W_2; Y_2^n | W_1 S^n) + n\delta_n \\
&= \sum_{i=1}^n I(W_2; Y_{2i} | Y_{2(i+1)}^n S^n W_1 X_1^n) + n\delta_n \\
&\stackrel{(c)}{\leq} \sum_{i=1}^n [H(Y_{2i} | S_i X_{1i}) - H(Y_{2i} | W_2 Y_{2(i+1)}^n S^n W_1 X_1^n X_{2i})] + n\delta_n \\
&\stackrel{(d)}{\leq} \sum_{i=1}^n [H(Y_{2i} | S_i X_{1i}) - H(Y_{2i} | S_i X_{1i} X_{2i})] + n\delta_n \\
&= \sum_{i=1}^n I(X_{2i}; Y_{2i} | S_i X_{1i}) + n\delta_n. \tag{B.5}
\end{aligned}$$

where (a) follows from Fano's inequality (B.1), (b) follows from chain rule and the fact that W_2 and (W_1, S^n) are independent, (c) follows because conditioning does not increase entropy, and (d) follows from the Markov chain $K_i T_i \leftrightarrow X_{1i} X_{2i} S_i \leftrightarrow Y_{1i} Y_{2i}$.

We further bound $R_1 + R_2$ based on Fano's inequality as follows:

$$\begin{aligned}
n(R_1 + R_2) &\leq I(W_1 W_2; Y_2^n S^n) + n\delta_n \\
&= \sum_{i=1}^n I(W_2 W_1; Y_{2i} | Y_{2(i+1)}^n S^n) + n\delta_n \\
&\stackrel{(a)}{\leq} \sum_{i=1}^n [H(Y_{2i} | S_i) - H(Y_{2i} | W_2 Y_{2(i+1)}^n S^n W_1 X_{1i} X_{2i})] + n\delta_n \\
&\stackrel{(b)}{\leq} \sum_{i=1}^n [H(Y_{2i} | S_i) - H(Y_{2i} | S_i X_{1i} X_{2i})] + n\delta_n
\end{aligned}$$

$$= \sum_{i=1}^n I(X_{1i}X_{2i}; Y_{2i}|S_i) + n\delta_n. \quad (\text{B.6})$$

where (a) follows because conditioning does not increase entropy, and (b) follows because Y_{2i} is independent of other variables given X_{1i} , X_{2i} and S_i .

We introduce a lemma which is useful in the proof.

Lemma B.1. : [53, Lemma 7] For a set of random variables $(T, Y_1, \dots, Y_n, Z_1, \dots, Z_n)$,

$$\sum_{i=1}^n I(Y_i; Z^{i-1}|TY_{i+1}^n) = \sum_{i=1}^n I(Y_{i+1}^n; Z_i|TZ^{i-1}). \quad (\text{B.7})$$

We proceed to derive an alternative bound on $R_1 + R_2$ as follows:

$$\begin{aligned} n(R_1 + R_2) &\leq I(W_1; Y_1^n) + I(W_2; Y_2^n S^n) + n\delta_n \\ &\stackrel{(a)}{\leq} I(W_1; Y_1^n) + I(W_2; Y_2^n S^n|W_1) + n\delta_n \end{aligned} \quad (\text{B.8})$$

where (a) follows because W_1 and W_2 are independent.

The first term in (B.8) can be bounded as follows:

$$\begin{aligned} I(W_1; Y_1^n) &= \sum_{i=1}^n I(W_1; Y_{1i}|Y_1^{i-1}) \\ &\stackrel{(a)}{\leq} \sum_{i=1}^n I(W_1 Y_1^{i-1}; Y_{1i}) \\ &\stackrel{(b)}{=} \sum_{i=1}^n [I(W_1 Y_1^{i-1} S_{i+1}^n Y_{2(i+1)}^n; Y_{1i}) - I(S_{i+1}^n Y_{2(i+1)}^n; Y_{1i}|W_1 Y_1^{i-1})] \\ &\stackrel{(c)}{=} \sum_{i=1}^n [I(W_1 Y_1^{i-1} S_{i+1}^n Y_{2(i+1)}^n; Y_{1i}) - I(S_i Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n)] \\ &\stackrel{(d)}{=} \sum_{i=1}^n [I(W_1 Y_1^{i-1} S_{i+1}^n Y_{2(i+1)}^n; Y_{1i}) - I(S_i Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n) \\ &\quad + I(W_1 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i})] \end{aligned}$$

$$\begin{aligned}
& \stackrel{(e)}{=} \sum_{i=1}^n [I(W_1 Y_1^{i-1} S_{i+1}^n Y_{2(i+1)}^n; Y_{1i}) - I(S_i; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n) \\
& \quad + I(W_1 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i}) - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i)] \\
& \stackrel{(f)}{=} \sum_{i=1}^n [I(T_i K_i X_{1i}; Y_{1i}) - I(T_i K_i X_{1i}; S_i) \\
& \quad + I(W_1 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i}) - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i)] \tag{B.9}
\end{aligned}$$

where (a) follows from chain rule and the fact that mutual information is nonnegative, (b) follows from chain rule, (c) follows from Lemma B.1, (d) and (e) follows from chain rule, and (f) follows from the definition for T_i and K_i .

We next consider the last two terms in (B.9) together with the second term in (B.8) as follows:

$$\begin{aligned}
& I(W_2; Y_2^n S^n | W_1) + \sum_{i=1}^n \left[I(W_1 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i}) - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i) \right] \\
& \stackrel{(a)}{=} \sum_{i=1}^n \left[I(W_2; Y_{2i} S_i | W_1 S_{i+1}^n Y_{2(i+1)}^n) + I(W_1 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i}) \right. \\
& \quad \left. - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i) \right] \\
& \stackrel{(b)}{=} \sum_{i=1}^n [I(W_1 W_2 S_{i+1}^n Y_{2(i+1)}^n; S_i Y_{2i}) + I(S^{i-1}; S_i Y_{2i} | W_1 W_2 S_{i+1}^n Y_{2(i+1)}^n) \\
& \quad - I(S_{i+1}^n Y_{2(i+1)}^n; S_i | W_1 W_2 S^{i-1}) - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i)] \\
& \stackrel{(c)}{=} \sum_{i=1}^n [I(W_1 W_2 S_{i+1}^n S^{i-1} Y_{2(i+1)}^n; S_i Y_{2i}) - I(S_{i+1}^n Y_{2(i+1)}^n; S_i | W_1 W_2 S^{i-1}) \\
& \quad - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i)] \\
& \stackrel{(d)}{=} \sum_{i=1}^n [I(W_1 W_2 S_{i+1}^n S^{i-1} Y_{2(i+1)}^n; S_i Y_{2i}) - I(S_{i+1}^n Y_{2(i+1)}^n W_1 W_2 S^{i-1}; S_i) \\
& \quad - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i)] \\
& \stackrel{(e)}{=} \sum_{i=1}^n \left[I(W_1 W_2 S_{i+1}^n S^{i-1} Y_{2(i+1)}^n; Y_{2i} | S_i) - I(Y_{2i}; Y_1^{i-1} W_1 S_{i+1}^n Y_{2(i+1)}^n | S_i) \right]
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(f)}{\leq} \sum_{i=1}^n \left[H(Y_{2i}|S_i Y_1^{i-1} W_1 X_1^n S_{i+1}^n Y_{2(i+1)}^n) - H(Y_{2i}|S_i Y_1^{i-1} W_1 X_1^n W_2 S_{i+1}^n S^{i-1} Y_{2(i+1)}^n X_{2i}) \right] \\
&\stackrel{(g)}{\leq} \sum_{i=1}^n \left[H(Y_{2i}|S_i Y_1^{i-1} W_1 X_1^n S_{i+1}^n Y_{2(i+1)}^n) - H(Y_{2i}|S_i Y_1^{i-1} W_1 X_1^n S_{i+1}^n Y_{2(i+1)}^n X_{2i}) \right] \\
&\stackrel{(h)}{=} \sum_{i=1}^n I(X_{2i}; Y_{2i}|X_{1i} T_i K_i S_i)
\end{aligned} \tag{B.10}$$

where (a) follows from chain rule, (b) follows from chain rule to combine the first two terms in the previous step and Lemma B.1, (c) follows from chain rule, (d) follows from chain rule and because W_1 , W_2 and S^{i-1} are independent from S_i , (e) follows from chain rule, (f) follows because X_1^n is a function of W_1 and conditioning does not increase entropy, (g) follows because Y_{2i} is independent of other variables given X_{1i} , X_{2i} and S_i , and (h) follows from the definition of T_i and K_i .

Therefore, substituting (B.9) and (B.10) into (B.8), we obtain

$$\begin{aligned}
n(R_1 + R_2) &\leq \sum_{i=1}^n [I(T_i K_i X_{1i}; Y_{1i}) - I(T_i K_i; S_i|X_{1i}) \\
&\quad + I(X_{2i}; Y_{2i}|X_{1i} T_i K_i S_i)] + n\delta_n.
\end{aligned} \tag{B.11}$$

B.2 Proof of the Converse for Theorem 3.2

For the Gaussian channel, if $|a| \leq 1$, it satisfies the condition (3.2). For these channels, we first prove the following bounds.

$$nR_1 \leq \sum_{i=1}^n [I(U_i X_{1i}; Y_{1i}) - I(U_i; S_i|X_{1i})] + n\delta_n \tag{B.12a}$$

$$nR_2 \leq \sum_{i=1}^n I(X_{2i}; Y_{2i}|U_i X_{1i} S_i) + n\delta_n \tag{B.12b}$$

$$n(R_1 + R_2) \leq \sum_{i=1}^n I(X_{1i} X_{2i}; Y_{2i}|S_i) + n\delta_n \tag{B.12c}$$

The bound (B.12a) follows from (B.4) by setting $U_i = K_i = (W_1 S_{i+1}^n X_1^n Y_1^{i-1})$ for

$i = 1, \dots, n$. The bound (B.12c) follows from (B.6).

We then bound R_2 as follows and obtain (B.12b):

$$\begin{aligned}
nR_2 &= I(W_2; Y_2^n S^n) + n\delta_n \\
&\stackrel{(a)}{\leq} I(W_2; Y_2^n S^n | W_1) + n\delta_n \\
&\stackrel{(b)}{=} I(W_2; Y_2^n | W_1 S^n) + n\delta_n \\
&= \sum_{i=1}^n I(W_2; Y_{2i} | W_1 S^n Y_2^{i-1}) + n\delta_n \\
&= \sum_{i=1}^n [H(Y_{2i} | W_1 S^n Y_2^{i-1}) - H(Y_{2i} | W_1 W_2 S^n Y_2^{i-1})] + n\delta_n \\
&\stackrel{(c)}{=} \sum_{i=1}^n [H(Y_{2i} | W_1 S^n X_1^n Y_1^{i-1} Y_2^{i-1}) - H(Y_{2i} | W_1 W_2 S^n Y_2^{i-1} X_1^n Y_1^{i-1})] + n\delta_n \\
&\stackrel{(d)}{\leq} \sum_{i=1}^n [H(Y_{2i} | W_1 S_{i+1}^n X_1^n Y_1^{i-1} S_i) - H(Y_{2i} | W_1 S_{i+1}^n S_i X_1^n Y_1^{i-1} X_{2i})] + n\delta_n \\
&\stackrel{(e)}{\leq} \sum_{i=1}^n [H(Y_{2i} | S_i X_{1i} U_i) - H(Y_{2i} | S_i X_{1i} U_i X_{2i})] + n\delta_n \\
&\leq \sum_{i=1}^n I(X_{2i}; Y_{2i} | U_i X_{1i} S_i) + n\delta_n \tag{B.13}
\end{aligned}$$

where (a) follows because W_1 and W_2 are independent, (b) follows because W_2 and S are independent, (c) follows from the degradedness condition (3.2) so that X_1^n and Y_1^{i-1} can be added into the conditioning, (d) follows from the fact that given X_{1i} , X_{2i} , and S_i , Y_{2i} is independent of all other variables, and (e) follows from the definition of U_i .

We further derive the bounds (B.12a)-(B.12c) for Gaussian channels. We first consider

the bound on R_1 as follows:

$$\begin{aligned}
R_1 &\leq \frac{1}{n} \sum_{i=1}^n [I(X_{1i}U_i; Y_{1i}) - I(U_i; S_i|X_{1i})] \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(Y_{1i}|X_{1i}U_i) - h(S_i|X_{1i}) + h(S_i|X_{1i}U_i)] \\
&\stackrel{(a)}{=} \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(Y_{1i}|X_{1i}U_iS_i) - I(S_i; Y_{1i}|X_{1i}U_i) - h(S_i|X_{1i}) + h(S_i|X_{1i}U_i)] \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(Y_{1i}|X_{1i}U_iS_i) - h(S_i|X_{1i}) + h(S_i|X_{1i}U_iY_{1i})] \\
&\stackrel{(b)}{\leq} \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(Y_{1i}|X_{1i}U_iS_i) - h(S_i) + h(S_i|X_{1i}Y_{1i})] \tag{B.14}
\end{aligned}$$

where (a) follows because addition of the second and third terms equals the second term in the previous step, and (b) follows because S_i and X_{1i} are independent and conditioning does not increase entropy.

We then derive bound for each term in (B.14) respectively as follows. The first term in (B.14) can be derived as:

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^n h(Y_{1i}) \\
&\stackrel{(a)}{\leq} \frac{1}{2n} \sum_{i=1}^n \log 2\pi e (E(X_{1i} + aX_{2i} + S_i + N_i)^2) \\
&\leq \frac{1}{2n} \sum_{i=1}^n \log 2\pi e \left(E[X_{1i}^2] + 2aE(X_{1i}X_{2i}) + a^2E[X_{2i}^2] \right. \\
&\quad \left. + 2aE(X_{2i}S_i) + E[S_i^2] + E[N_i^2] \right) \\
&\stackrel{(b)}{\leq} \frac{1}{2} \log 2\pi e \left(\frac{1}{n} \sum_{i=1}^n E[X_{1i}^2] + \frac{2a}{n} \sum_{i=1}^n E(X_{1i}X_{2i}) + \frac{a^2}{n} \sum_{i=1}^n E[X_{2i}^2] \right.
\end{aligned}$$

$$\begin{aligned}
& + \frac{2a}{n} \sum_{i=1}^n E(X_{2i}S_i) + \frac{1}{n} \sum_{i=1}^n E[S_i^2] + \frac{1}{n} \sum_{i=1}^n E[N_i^2]) \Bigg) \\
& \stackrel{(c)}{\leq} \frac{1}{2} \log 2\pi e \left(P_1 + a^2 P_2 + Q + 1 + \frac{2a}{n} \sum_{i=1}^n E(X_{1i}X_{2i}) + \frac{2a}{n} \sum_{i=1}^n E(X_{2i}S_i) \right) \\
& \leq \frac{1}{2} \log 2\pi e \left(P_1 + a^2 P_2 + Q + 1 + 2a\rho_{21}\sqrt{P_1P_2} + 2a\rho_{2s}\sqrt{P_2Q} \right) \quad (\text{B.15})
\end{aligned}$$

where $\rho_{21} = \frac{\frac{1}{n} \sum_{i=1}^n E(X_{1i}X_{2i})}{\sqrt{P_1P_2}}$ and $\rho_{2s} = \frac{\frac{1}{n} \sum_{i=1}^n E(X_{2i}S_i)}{\sqrt{P_2Q}}$. In the above derivation, (a) follows from the fact that the Gaussian distribution maximizes the entropy given the variance of the random variable, (b) follows from the concavity of the logarithm function and Jensen's inequality, and (c) follows from the power constraints.

We next quantify the term $\frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}U_iS_i)$ via its upper and lower bounds. We first have

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}X_{2i}S_i) \\
& \stackrel{(a)}{\leq} \frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}U_iS_i) \leq \frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}S_i) \quad (\text{B.16})
\end{aligned}$$

where (a) follows because conditioning does not increase entropy and given X_{1i} , X_{2i} , and S_i , Y_{1i} is independent of all other variables.

For the left-hand side, we have

$$\frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}X_{2i}S_i) = \frac{1}{2} \log 2\pi e. \quad (\text{B.17})$$

For the right-hand side, by setting $\alpha = a\rho_{21}\sqrt{\frac{P_2}{P_1}}$ and $\beta = a\rho_{2S}\sqrt{\frac{P_2}{Q}}$, we have

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}S_i) \\
&= \frac{1}{n} \sum_{i=1}^n h(X_{1i} + aX_{2i} + S_i + N_{1i}|S_iX_{1i}) \\
&= \frac{1}{n} \sum_{i=1}^n h(aX_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i|S_iX_{1i}) \\
&\stackrel{(a)}{\leq} \frac{1}{n} \sum_{i=1}^n h(aX_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i) \\
&\stackrel{(b)}{\leq} \frac{1}{2n} \sum_{i=1}^n \log(2\pi e E[(aX_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i)^2]) \\
&\stackrel{(c)}{\leq} \frac{1}{2} \log 2\pi e \left(a^2 P_2 + 1 + \alpha^2 P_1 + \beta^2 Q \right. \\
&\quad \left. - 2a\alpha \frac{1}{n} \sum_{i=1}^n E[X_{1i}X_{2i}] - 2a\beta \frac{1}{n} \sum_{i=1}^n E[X_{2i}S_i] \right) \\
&= \frac{1}{2} \log 2\pi e (1 + a^2(1 - \rho_{2S}^2 - \rho_{21}^2)P_2). \tag{B.18}
\end{aligned}$$

where (a) follows because conditioning does not increase entropy, (b) follows because the Gaussian distribution maximizes the entropy for variables with certain variance, and (c) follows from the concavity of the log function and Jensen's inequality.

Therefore, combining (B.17) and (B.18), we conclude that there exists $0 \leq P_2'' \leq (1 - \rho_{2S}^2 - \rho_{21}^2)P_2$ such that

$$\frac{1}{n} \sum_{i=1}^n h(Y_{1i}|X_{1i}U_iS_i) = \frac{1}{2} \log 2\pi e (1 + a^2 P_2''). \tag{B.19}$$

The third term in (B.14) is given by

$$\frac{1}{n} \sum_{i=1}^n h(S_i) = \frac{1}{2} \log 2\pi e Q. \quad (\text{B.20})$$

Finally, for the fourth term in (B.14), we first define $\alpha' = \frac{-a\rho_{21}\sqrt{P_2P_1}(a\rho_{2s}\sqrt{P_2Q}+Q)}{(a^2(1-\rho_{21}^2)P_2+Q+2a\rho_{2s}\sqrt{P_2Q}+1)P_1}$ and $\beta' = -\frac{P_1}{a\rho_{21}\sqrt{P_1P_2}}\alpha'$, and then have

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n h(S_i|X_{1i}Y_{1i}) \\ &= \frac{1}{n} \sum_{i=1}^n h(S_i|X_{1i}, X_{1i} + aX_{2i} + S_i + N_{1i}) \\ &= \frac{1}{n} \sum_{i=1}^n h(S_i - \alpha'X_{1i} - \beta'(aX_{2i} + S_i + N_{1i})|X_{1i}, X_{1i} + aX_{2i} + S_i + N_{1i}) \\ &\stackrel{(a)}{\leq} \frac{1}{n} \sum_{i=1}^n h(S_i - \alpha'X_{1i} - \beta'(aX_{2i} + S_i + N_{1i})) \\ &\stackrel{(b)}{\leq} \frac{1}{n} \sum_{i=1}^n \log \left(2\pi e E(S_i - \alpha'X_{1i} - \beta'(aX_{2i} + S_i + N_{1i}))^2 \right) \\ &\stackrel{(c)}{\leq} \frac{1}{2} \log 2\pi e \left(Q + \alpha'^2 P_1 + a^2 \beta'^2 P_2 + \beta'^2 Q + 2a\beta'^2 \frac{1}{n} \sum_{i=1}^n E(X_{2i}S_i) + \beta'^2 \right. \\ &\quad \left. + 2\alpha'\beta'a \frac{1}{n} \sum_{i=1}^n E(X_{1i}X_{2i}) - 2\beta'a \frac{1}{n} \sum_{i=1}^n E(X_{2i}S_i) - 2\beta'Q \right) \\ &= \frac{1}{2} \log 2\pi e \frac{(a^2(1-\rho_{21}^2-\rho_{2s}^2)P_2+1)Q}{a^2(1-\rho_{21}^2)P_2+2a\rho_{2s}\sqrt{P_2Q}+Q+1} \end{aligned} \quad (\text{B.21})$$

where (a) follows because conditioning does not increase entropy, (b) follows because the Gaussian distribution maximizes the entropy for variables with certain variance, and (c) follows from the concavity of the log function and Jensen's inequality.

Substituting (B.15), (B.17), (B.20) and (B.21) into (B.14), we obtain

$$\begin{aligned}
 R_1 \leq & \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1P_2} + a^2\rho_{21}^2P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2Q} + Q + 1} \right) \\
 & + \frac{1}{2} \log \left(1 + \frac{a^2P_2'}{a^2P_2'' + 1} \right)
 \end{aligned} \tag{B.22}$$

where $P_2' = (1 - \rho_{21}^2 - \rho_{2s}^2)P_2 - P_2''$.

We then bound R_2 by further deriving (B.12b). When $a \leq 1$, we have $Y_{1i} = aY_{2i} + (1 - ab)X_{1i} + (1 - ac)S_i + N_i'$, where $N_i' \sim \mathcal{N}(0, 1 - a^2)$ and is independent from Y_2^n , X_1^n and S^n . By applying the conditional entropy power inequality [54], we have

$$\begin{aligned}
 2^{2h(Y_{1i}|U_iS_iX_{1i})} &= 2^{2h(aY_{2i} + (1-ab)X_{1i} + (1-ac)S_i + N_i'|U_iS_iX_{1i})} \\
 &= 2^{2h(aY_{2i} + N_i'|U_iS_iX_{1i})} \\
 &\geq 2^{2h(aY_{2i}|U_iS_iX_{1i})} + 2^{2h(N_i'|U_iS_iX_{1i})} \\
 &= 2^{2h(Y_{2i}|U_iS_iX_{1i}) + \log(a^2)} + 2\pi e(1 - a^2).
 \end{aligned} \tag{B.23}$$

Thus,

$$\begin{aligned}
 & \frac{1}{n} \sum_{i=1}^n h(Y_{2i}|U_iS_iX_{1i}) \\
 & \leq \frac{1}{n} \sum_{i=1}^n \frac{1}{2} \log \left(\frac{2^{2h(Y_{1i}|U_iS_iX_{1i})} - 2\pi e(1 - a^2)}{a^2} \right) \\
 & \stackrel{(a)}{\leq} \frac{1}{2} \log \left(\frac{2^{\frac{1}{n} \sum_{i=1}^n h(Y_{1i}|U_iS_iX_{1i})} - 2\pi e(1 - a^2)}{a^2} \right) \\
 & \stackrel{(b)}{=} \frac{1}{2} \log(2\pi e(1 + P_2''))
 \end{aligned} \tag{B.24}$$

where (a) follows from the concavity of the function $\log(2^x - b)$ for $b \geq 0$, and (b) follows

from (B.19).

Therefore, we have

$$\begin{aligned}
R_2 &\leq \frac{1}{n} \sum_{i=1}^n I(X_{2i}; Y_{2i} | X_{1i} S_i U_i) \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{2i} | X_{1i} S_i U_i) - h(Y_{2i} | X_{1i} S_i X_{2i})] \\
&\stackrel{(a)}{=} \frac{1}{2} \log(2\pi e(1 + P_2'')) - \frac{1}{2} \log(2\pi e) \\
&= \frac{1}{2} \log(1 + P_2'') .
\end{aligned} \tag{B.25}$$

where (a) follows from (B.19).

We finally bound $R_1 + R_2$ by further deriving (B.12c). We set $\alpha'' = \rho_{2s} \sqrt{\frac{P_2}{Q}}$, and have

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{n} \sum_{i=1}^n I(X_{1i} X_{2i}; Y_{2i} | S_i) \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{2i} | S_i) - h(Y_{2i} | X_{1i} S_i X_{2i})] \\
&= \frac{1}{n} \sum_{i=1}^n h(bX_{1i} + X_{2i} + cS_i + N_{1i} | S_i) - \frac{1}{2} \log 2\pi e \\
&= \frac{1}{n} \sum_{i=1}^n h(bX_{1i} + X_{2i} + N_{1i} - \alpha'' S_i | S_i) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(a)}{\leq} \frac{1}{n} \sum_{i=1}^n h(bX_{1i} + X_{2i} + N_{1i} - \alpha'' S_i) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(b)}{\leq} \frac{1}{n} \sum_{i=1}^n \log(2\pi e E(bX_{1i} + X_{2i} + N_{1i} - \alpha'' S_i)^2) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(c)}{\leq} \frac{1}{2} \log 2\pi e \left(b^2 P_1 + P_2 + 1 + \alpha''^2 Q + 2b \frac{1}{n} \sum_{i=1}^n E[X_{1i} X_{2i}] - 2\alpha'' \frac{1}{n} \sum_{i=1}^n E[X_{2i} S_i] \right) \\
&\quad - \frac{1}{2} \log 2\pi e
\end{aligned}$$

$$= \frac{1}{2} \log(b^2 P_1 + P_2 + 1 + 2b\rho_{21}\sqrt{P_1 P_2} - \rho_{2s}^2 P_2). \quad (\text{B.26})$$

where (a) follows because conditioning does not increase entropy, (b) follows because the Gaussian distribution maximizes the entropy for variables with certain variance, and (c) follows from the concavity of the log function and Jensen's inequality.

B.3 Proof of Lemma 3.1

Following (B.5) and (B.6), we obtain

$$nR_2 \leq \sum_{i=1}^n I(X_{2i}; Y_{2i} | S_i X_i) + n\delta_n \quad (\text{B.27a})$$

$$n(R_1 + R_2) \leq \sum_{i=1}^n I(X_{1i} X_{2i}; Y_{2i} | S_i) + n\delta_n. \quad (\text{B.27b})$$

We then prove an alternative bound on $R_1 + R_2$ as in (B.28) on the top of next page, where (a) follows due to the chain rule and the fact that W_1 and W_2 are independent, (b) follows because conditioning does not increase entropy, (c) follows from degradedness condition (3.3), (d) follows because the term $H(Y_{1i} | X_{1i})$ is added and subtracted, (e) follows because conditioning does not increase entropy, (f) follows because given X_{1i} , X_{2i} , and S_i , Y_{2i} is independent of all other variables, and (g) follows because X_1^n is a function of W_1 and conditioning does not increase entropy.

B.4 Proof of the Converse for Theorem 3.3

Based on the outer bound derived in Appendix B.3, we further derive an outer bound for the Gaussian channel. We first derive a bound on R_2 based on (B.27a). We set $\alpha = \rho_{21} \sqrt{\frac{P_2}{P_1}}$

$$\begin{aligned}
& n(R_1 + R_2) \\
& \leq I(W_1; Y_1^n) + I(W_2; Y_2^n | S^n) + n\delta_n \\
& \stackrel{(a)}{\leq} I(W_1; Y_1^n) + I(W_2; Y_2^n | S^n W_1) + n\delta_n \\
& = I(W_1; Y_1^n) + H(W_2 | S^n W_1) - H(W_2 | S^n W_1 Y_2^n) + n\delta_n \\
& \stackrel{(b)}{\leq} I(W_1; Y_1^n) + H(W_2 | S^n W_1) - H(W_2 | S^n W_1 Y_2^n Y_1^n X_1^n) + n\delta_n \\
& \stackrel{(c)}{=} I(W_1; Y_1^n) + H(W_2 | S^n W_1) - H(W_2 | S^n W_1 Y_1^n) + n\delta_n \\
& = I(W_1; Y_1^n) + I(W_2; Y_1^n | S^n W_1) + n\delta_n \\
& = \sum_{i=1}^n [H(Y_{1i} | Y_1^{i-1}) - H(Y_{1i} | W_1 Y_1^{i-1}) + H(Y_{1i} | S^n W_1 Y_1^{i-1}) - H(Y_{1i} | S^n W_1 W_2 Y_1^{i-1})] + n\delta_n \\
& \stackrel{(d)}{=} \sum_{i=1}^n [H(Y_{1i} | Y_1^{i-1}) - H(Y_{1i} | X_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | W_1 Y_1^{i-1}) \\
& \quad + H(Y_{1i} | S^n W_1 Y_1^{i-1}) - H(Y_{1i} | S^n W_1 W_2 Y_1^{i-1})] + n\delta_n \\
& \stackrel{(e)}{\leq} \sum_{i=1}^n [H(Y_{1i}) - H(Y_{1i} | X_{1i}) + H(Y_{1i} | X_{1i}) - I(S^n; Y_{1i} | W_1 Y_1^{i-1}) \\
& \quad - H(Y_{1i} | S^n X_{1i} X_{2i} W_1 W_2 Y_1^{i-1})] + n\delta_n \\
& \stackrel{(f)}{=} \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i})] - I(S^n; Y_1^n | W_1) + n\delta_n \\
& = \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i})] - H(S^n) + H(S^n | Y_1^n W_1) + n\delta_n \\
& = \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i}) - H(S_i) + H(S_i | Y_1^n W_1 S_{i+1}^n)] + n\delta_n \\
& \stackrel{(g)}{\leq} \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i}) - H(S_i) + H(S_i | Y_{1i} X_{1i})] + n\delta_n \\
& = \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + H(Y_{1i} | X_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i}) - I(S_i; Y_{1i} | X_{1i})] + n\delta_n \\
& = \sum_{i=1}^n [I(X_{1i}; Y_{1i}) - H(Y_{1i} | S_i X_{1i} X_{2i}) + H(Y_{1i} | S_i X_{1i})] + n\delta_n \\
& = \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + I(X_{2i}; Y_{1i} | S_i X_{1i})] + n\delta_n \tag{B.28}
\end{aligned}$$

and $\beta = \rho_{2s} \sqrt{\frac{P_2}{Q}}$, where $\rho_{21} = \frac{\frac{1}{n} \sum_{i=1}^n E(X_{1i}X_{2i})}{\sqrt{P_1 P_2}}$ and $\rho_{2s} = \frac{\frac{1}{n} \sum_{i=1}^n E(X_{2i}S_i)}{\sqrt{P_2 Q}}$. We then obtain:

$$\begin{aligned}
R_2 &\leq \frac{1}{n} \sum_{i=1}^n h(Y_{2i}|X_{1i}S_i) - h(Y_{2i}|X_{1i}X_{2i}S_i) \\
&= \frac{1}{n} \sum_{i=1}^n h(bX_{1i} + X_{2i} + cS_i + N_{1i}|S_iX_{1i}) - \frac{1}{2} \log 2\pi e \\
&= \frac{1}{n} \sum_{i=1}^n h(X_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i|S_iX_{1i}) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(a)}{\leq} \frac{1}{n} \sum_{i=1}^n h(X_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(b)}{\leq} \frac{1}{2n} \sum_{i=1}^n \log (2\pi e E(X_{2i} + N_{1i} - \alpha X_{1i} - \beta S_i)^2) - \frac{1}{2} \log 2\pi e \\
&\stackrel{(c)}{\leq} \frac{1}{2} \log \left(P_2 + 1 + \alpha^2 P_1 + \beta^2 Q - 2\alpha \frac{1}{n} \sum_{i=1}^n E[X_{1i}X_{2i}] - 2\beta \frac{1}{n} \sum_{i=1}^n E[X_{2i}S_i] \right) \\
&= \frac{1}{2} \log(1 + (1 - \rho_{2s}^2 - \rho_{21}^2)P_2) \tag{B.29}
\end{aligned}$$

where (a) follows because conditioning does not increase entropy, (b) follows because the Gaussian distribution maximizes the entropy for variables with certain variance, and (c) follows from the concavity of the log function and Jensen's inequality.

Following (B.26), we obtain the following bound on $R_1 + R_2$ based on (B.27b)

$$R_1 + R_2 \leq \frac{1}{2} \log \left(b^2 P_1 + P_2 + 1 + 2b\rho_{21}\sqrt{P_1 P_2} - \rho_{2s}^2 P_2 \right). \tag{B.30}$$

We further derive (B.28) for the Gaussian channel as follows:

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{n} \sum_{i=1}^n [I(X_{1i}; Y_{1i}) + I(X_{2i}; Y_{1i} | X_{1i} S_i)] \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(Y_{1i} | X_{1i}) + h(Y_{1i} | X_{1i} S_i) - h(Y_{1i} | S_i X_{1i} X_{2i})] \\
&= \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - I(S_i; Y_{1i} | X_{1i}) - h(Y_{1i} | S_i X_{1i} X_{2i})] \\
&\stackrel{(a)}{=} \frac{1}{n} \sum_{i=1}^n [h(Y_{1i}) - h(S_i) + h(S_i | X_{1i} Y_{1i}) - h(Y_{1i} | S_i X_{1i} X_{2i})]
\end{aligned} \tag{B.31}$$

where (a) follows because S_i and X_{1i} are independent.

Following (B.15), (B.17), (B.20), and (B.21) in Appendix B.2, we obtain

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n h(Y_{1i}) &\leq \frac{1}{2} \log 2\pi e (P_1 + a^2 P_2 + Q + 1 + 2a\rho_{21}\sqrt{P_1 P_2} + 2a\rho_{2s}\sqrt{P_2 Q}) \\
\frac{1}{n} \sum_{i=1}^n h(Y_{1i} | X_{1i} X_{2i} S_i) &= \frac{1}{2} \log 2\pi e \\
\frac{1}{n} \sum_{i=1}^n h(S_i) &= \frac{1}{2} \log 2\pi e Q \\
\frac{1}{n} \sum_{i=1}^n h(S_i | X_{1i} Y_{1i}) &\leq \frac{1}{2} \log 2\pi e \frac{(a^2(1 - \rho_{21}^2 - \rho_{2s}^2)P_2 + 1)Q}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2 Q} + Q + 1}
\end{aligned}$$

Substituting the above bounds into (B.31), we obtain

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{2} \log \left(1 + \frac{P_1 + 2a\rho_{21}\sqrt{P_1 P_2} + a^2 \rho_{21}^2 P_2}{a^2(1 - \rho_{21}^2)P_2 + 2a\rho_{2s}\sqrt{P_2 Q} + Q + 1} \right) \\
&\quad + \frac{1}{2} \log (1 + a^2(1 - \rho_{2s}^2 - \rho_{21}^2)P_2)
\end{aligned} \tag{B.32}$$

which concludes the proof.

B.5 Proof of Lemma 3.2

The achievable scheme applies rate splitting, superposition coding and Gel'fand-Pinsker binning scheme. We use random codes and fix the following joint distribution:

$$P_{SX_1TUVX_2Y_1Y_2} = P_{X_1}P_S P_{T|X_1S} P_{U|X_1TS} P_{V|TUX_1S} P_{X_2|TUVX_1S} P_{Y_1Y_2|X_1X_2S}.$$

Let $T_\epsilon^n(P_{SX_1TUVX_2Y_1Y_2})$ denote the strongly joint ϵ -typical set based on the above distribution. For a given sequence x^n , let $T_\epsilon^n(P_{U|X}|x^n)$ denote the set of sequences u^n such that (u^n, x^n) is jointly typical based on the distribution P_{XU} .

Code Construction:

1. Generate 2^{nR_1} codewords $x_1^n(w_1)$ with i.i.d. components based on P_{X_1} . Index these codewords by $w_1 = 1, \dots, 2^{nR_1}$.
2. For each $x_1^n(w_1)$, generate $t^n(w_1, v_1)$ with i.i.d. components based on $P_{T|X_1}$. Index these codewords by $v_1 = 1, \dots, 2^{n\tilde{R}_1}$.
3. For each $x_1^n(w_1)$ and $t^n(w_1, v_1)$, generate $u^n(w_1, v_1, w_{21}, v_{21})$ with i.i.d. components based on $P_{U|X_1T}$. Index these codewords by $w_{21} = 1, \dots, 2^{nR_{21}}$ and $v_{21} = 1, \dots, 2^{n\tilde{R}_{21}}$.
4. For each $x_1^n(w_1)$, $t^n(w_1, v_1)$, and $u^n(w_1, v_1, w_{21}, v_{21})$, generate $v^n(w_1, v_1, w_{21}, v_{21}, w_{22}, v_{22})$ with i.i.d. components based on $P_{V|X_1TU}$. Index these codewords by $w_{22} = 1, \dots, 2^{nR_{22}}$ and $v_{22} = 1, \dots, 2^{n\tilde{R}_{22}}$.

Encoding:

1. Encoder 1: Given w_1 , map w_1 into $x_1^n(w_1)$ for transmission.
2. Encoder 2:

- Given $w_1, x_1^n(w_1)$ and s^n , select $t^n(w_1, \tilde{v}_1)$ such that

$$(t^n(w_1, \tilde{v}_1), s^n, x_1^n(w_1)) \in T_\epsilon^n(P_{X_1}P_SP_{T|X_1S}).$$

Otherwise, set $\tilde{v}_1 = 1$. It can be shown that for large n , such t^n exists with high probability if

$$\tilde{R}_1 > I(T; S|X_1). \quad (\text{B.33})$$

- Given w_{21} and selected $t^n(w_1, \tilde{v}_1)$, select $u^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21})$ such that

$$\begin{aligned} & (u^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}), t^n(w_1, \tilde{v}_1), s^n, x_1^n(w_1)) \\ & \in T_\epsilon^n(P_{X_1}P_SP_{T|X_1S}P_{U|X_1ST}). \end{aligned}$$

Otherwise, set $\tilde{v}_{21} = 1$. It can be shown that for large n , such u^n exists with high probability if

$$\tilde{R}_{21} > I(U; S|X_1T). \quad (\text{B.34})$$

- Given w_{22} and selected $u^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21})$, select $v^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}, w_{22}, \tilde{v}_{22})$ such that

$$\begin{aligned} & (v^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}, w_{22}, \tilde{v}_{22}), \\ & u^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}), t^n(w_1, \tilde{v}_1), s^n, x_1^n(w_1)) \\ & \in T_\epsilon^n(P_{X_1}P_SP_{T|X_1S}P_{U|X_1ST}P_{V|UX_1ST}). \end{aligned} \quad (\text{B.35})$$

Otherwise, set $\tilde{v}_{22} = 1$. It can be shown that for large n , such v^n exists with high probability if

$$\tilde{R}_{22} > I(V; S|UX_1T). \quad (\text{B.36})$$

- Given selected $x_1^n(w_1), t^n(w_1, \tilde{v}_1), u^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}), v^n(w_1, \tilde{v}_1, w_{21}, \tilde{v}_{21}, w_{22}, \tilde{v}_{22})$ and s^n , generate x_2^n with i.i.d. components based on $P_{X_2|TUVX_1S}$ for transmission.

Decoding:

1. Decoder 1: Given y_1^n , find the unique tuple $(\hat{w}_1, \hat{v}_1, \hat{w}_{21}, \hat{v}_{21})$ such that

$$\begin{aligned} & (x_1^n(\hat{w}_1), t^n(\hat{w}_1, \hat{v}_1), u^n(\hat{w}_1, \hat{v}_1, \hat{w}_{21}, \hat{v}_{21}), y_1^n) \\ & \in T_\epsilon^n(P_{X_1TUY_1}). \end{aligned}$$

If no or more than one such tuples with different w_1 can be found, then declare error.

One can show that for sufficiently large n , decoding is correct with high probability if

$$R_1 + \tilde{R}_1 + R_{21} + \tilde{R}_{21} \leq I(TUX_1; Y_1) \quad (\text{B.37})$$

We note that since receiver 1 is not required to decode W_{21} correctly by the channel model, the corresponding error events do not need to be analyzed.

2. Decoder 2: Given y_2^n , find a tuple $(\hat{w}_1, \hat{v}_1, \hat{w}_{21}, \hat{v}_{21}, \hat{w}_{22}, \hat{v}_{22})$ such that

$$\begin{aligned} & (x_1^n(\hat{w}_1), t^n(\hat{w}_1, \hat{v}_1), u^n(\hat{w}_1, \hat{v}_1, \hat{w}_{21}, \hat{v}_{21}), \\ & v^n(\hat{w}_1, \hat{v}_1, \hat{w}_{21}, \hat{v}_{21}, \hat{w}_{22}, \hat{v}_{22}), y_2^n) \in T_\epsilon^n(P_{X_1TUVY_2}). \end{aligned}$$

If no or more than one such tuples can be found, then declare error. It can be shown that

for sufficiently large n , decoding is correct with high probability if

$$R_{22} + \tilde{R}_{22} \leq I(V; Y_2 | U X_1 T) \quad (\text{B.38a})$$

$$R_{21} + \tilde{R}_{21} + R_{22} + \tilde{R}_{22} \leq I(UV; Y_2 | X_1 T) \quad (\text{B.38b})$$

$$\tilde{R}_1 + R_{21} + \tilde{R}_{21} + R_{22} + \tilde{R}_{22} \leq I(TUV; Y_2 | X_1) \quad (\text{B.38c})$$

$$R_1 + \tilde{R}_1 + R_{21} + \tilde{R}_{21} + R_{22} + \tilde{R}_{22} \leq I(TUV X_1; Y_2) \quad (\text{B.38d})$$

Lemma 3.2 is thus proved by combining (B.33)-(B.38d).

B.6 Proof of Theorem 3.5

Consider a $(2^{nR_1}, 2^{nR_2}, n)$ code with an average error probability $P_e^{(n)}$. The probability distribution on $\mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{S}^n \times \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n$ is given by

$$P_{W_1 W_2 S^n X_1^n X_2^n Y_1^n Y_2^n} = P_{W_1} P_{W_2} \left[\prod_{i=1}^n P_{S_i} \right] P_{X_1^n | W_1} P_{X_2^n | W_1 W_2 S^n} \prod_{i=1}^n P_{Y_{1i} Y_{2i} | X_{1i} X_{2i} S_i}. \quad (\text{B.39})$$

By Fano's inequality, we have

$$H(W_1 | Y_1^n) \leq n R_1 P_e^{(n)} + 1 = n \delta_{1n} \quad (\text{B.40a})$$

$$H(W_1 W_2 | Y_2^n) \leq n (R_1 + R_2) P_e^{(n)} + 1 = n \delta_{2n} \quad (\text{B.40b})$$

where $\delta_{1n}, \delta_{2n} \rightarrow 0$ as $n \rightarrow +\infty$. Let $\delta_n = \delta_{1n} + \delta_{2n}$, which also satisfies that $\delta_n \rightarrow 0$ as $n \rightarrow +\infty$.

We define the following auxiliary random variables:

$$\begin{aligned} T_i &= (W_1, S_{i+1}^n, X_1^n) \\ U_i &= (T_i, Y_1^{i-1}) \\ V_i &= (T_i, W_2, Y_2^{i-1}) \end{aligned} \tag{B.41}$$

which satisfy the Markov chain conditions:

$$T_i \longleftrightarrow U_i V_i \longleftrightarrow X_{1i} X_{2i} S_i \longleftrightarrow Y_{1i} Y_{2i} \tag{B.42}$$

for $i = 1, \dots, n$.

The following bound on R_1 follows the same steps as in (B.4) in Appendix B.1, and we have

$$nR_1 \leq \sum_{i=1}^n [I(T_i U_i X_{1i}; Y_{1i}) - I(T_i U_i; S_i | X_{1i})] + n\delta_n. \tag{B.43}$$

I.e. we define $T_i = (W_1, S_{i+1}^n, X_1^n)$ and $U_i = (T_i, Y_1^{i-1})$

We next bound R_2 and obtain

$$\begin{aligned} nR_2 &= I(W_2; Y_2^n) + n\delta_n \leq I(W_2; Y_2^n | W_1) + n\delta_n \\ &\stackrel{(a)}{=} \sum_{i=1}^n [I(W_2 S_{i+1}^n; Y_2^i | W_1) - I(W_2 S_i^n; Y_2^{i-1} | W_1)] + n\delta_n \\ &\stackrel{(b)}{=} \sum_{i=1}^n [I(W_2 S_{i+1}^n; Y_2^{i-1} | W_1) + I(W_2 S_{i+1}^n; Y_{2i} | W_1 Y_2^{i-1}) \\ &\quad - I(W_2 S_{i+1}^n; Y_2^{i-1} | W_1) - I(S_i; Y_2^{i-1} | W_1 W_2 S_{i+1}^n)] + n\delta_n \\ &= \sum_{i=1}^n [I(W_2 S_{i+1}^n; Y_{2i} | W_1 Y_2^{i-1}) - I(S_i; Y_2^{i-1} | W_1 W_2 S_{i+1}^n)] + n\delta_n \\ &= \sum_{i=1}^n [H(Y_{2i} | W_1 Y_2^{i-1}) - H(Y_{2i} | W_1 W_2 S_{i+1}^n Y_2^{i-1})] \end{aligned}$$

$$\begin{aligned}
& -H(S_i|W_1W_2S_{i+1}^n) + H(S_i|W_1W_2S_{i+1}^nY_2^{i-1})] + n\delta_n \\
& \stackrel{(c)}{=} \sum_{i=1}^n [H(Y_{2i}|W_1Y_2^{i-1}X_{1i}) - H(Y_{2i}|W_1W_2S_{i+1}^nX_1^nY_2^{i-1}) \\
& \quad - H(S_i|W_1W_2S_{i+1}^nX_{1i}) + H(S_i|W_1W_2S_{i+1}^nX_1^nY_2^{i-1})] + n\delta_n \\
& \stackrel{(d)}{\leq} \sum_{i=1}^n [H(Y_{2i}|X_{1i}) - H(Y_{2i}|X_{1i}T_iV_i) - H(S_i|X_{1i}) + H(S_i|X_{1i}T_iV_i)] + n\delta_n \\
& = \sum_{i=1}^n [I(T_iV_i; Y_{2i}|X_{1i}) - I(T_iV_i; S_i|X_{1i})] + n\delta_n.
\end{aligned} \tag{B.45}$$

where (a) follows due to cancellation of the terms in the sum and because $Y_1^0 = \phi$, (b) follows from chain rule, (c) follows because X_1^n is a function of W_1 , and (d) follows because conditioning does not increase entropy, and from the definition of T_i and V_i .

We then bound the sum rate $R_1 + R_2$ as follows.

$$\begin{aligned}
n(R_1 + R_2) &= I(W_1W_2; Y_2^n) + n\delta_n \\
&\stackrel{(a)}{=} \sum_{i=1}^n [I(W_1W_2S_{i+1}^n; Y_2^i) - I(W_1W_2S_i^n; Y_2^{i-1})] + n\delta_n \\
&\stackrel{(b)}{=} \sum_{i=1}^n [I(W_1W_2S_{i+1}^n; Y_2^{i-1}) + I(W_1W_2S_{i+1}^n; Y_{2i}|Y_2^{i-1}) \\
&\quad - I(W_1W_2S_{i+1}^n; Y_2^{i-1}) - I(S_i; Y_2^{i-1}|W_1W_2S_{i+1}^n)] + n\delta_n \\
&= \sum_{i=1}^n [I(W_1W_2S_{i+1}^n; Y_{2i}|Y_2^{i-1}) - I(S_i; Y_2^{i-1}|S_{i+1}^nW_1W_2)] + n\delta_n \\
&= \sum_{i=1}^n [H(Y_{2i}|Y_2^{i-1}) - H(Y_{2i}|W_1W_2S_{i+1}^nY_2^{i-1}) \\
&\quad - H(S_i|S_{i+1}^nW_1W_2) + H(S_i|S_{i+1}^nW_1W_2Y_2^{i-1})] + n\delta_n \\
&\stackrel{(c)}{\leq} \sum_{i=1}^n [H(Y_{2i}) - H(Y_{2i}|W_1W_2S_{i+1}^nX_1^nY_2^{i-1}) \\
&\quad - H(S_i|X_{1i}) + H(S_i|W_1W_2S_{i+1}^nX_1^nY_2^{i-1})] + n\delta_n \\
&\stackrel{(d)}{=} \sum_{i=1}^n [I(X_{1i}T_iV_i; Y_{2i}) - I(T_iV_i; S_i|X_{1i})] + n\delta_n
\end{aligned} \tag{B.46}$$

where (a) follows due to cancellation of the terms in the sum and because $Y_1^0 = \phi$, (b) follows due to chain rule, (c) follows because S^n is independent of (X_1^n, W_1, W_2) , S^n is i.i.d. and because X_1^n is a function of W_1 , and (d) follows from the definition of T_i and V_i .

B.7 Proof of the Outer Bound for Theorem 3.6

We define the following auxiliary random variables:

$$\begin{aligned} T_i &= (W_1, S_{i+1}^n, X_1^n, Y_1^{i-1}) \\ V_i &= (T_i, W_2, Y_2^{i-1}) \end{aligned} \tag{B.47}$$

which satisfy the Markov chain conditions:

$$T_i \longleftrightarrow V_i \longleftrightarrow X_{1i} X_{2i} S_i \longleftrightarrow Y_{1i} \longleftrightarrow Y_{2i} \tag{B.48}$$

for $i = 1, \dots, n$.

By following the step similar to those in (B.4), we obtain the following bound on R_1 :

$$nR_1 \leq \sum_{i=1}^n [I(T_i X_{1i}; Y_{1i}) - I(T_i; S_i | X_{1i})] + n\delta_n. \tag{B.49}$$

We next derive a bound on R_2 by continuing to derive the bound (B.44) as follows:

$$\begin{aligned} nR_2 &\leq \sum_{i=1}^n [H(Y_{2i} | W_1 Y_2^{i-1}) - H(Y_{2i} | W_1 W_2 S_{i+1}^n Y_2^{i-1}) \\ &\quad - H(S_i | W_1 W_2 S_{i+1}^n) + H(S_i | W_1 W_2 S_{i+1}^n Y_2^{i-1})] + n\delta_n \\ &\stackrel{(a)}{=} \sum_{i=1}^n [H(Y_{2i} | W_1 Y_2^{i-1} X_{1i}) - H(Y_{2i} | W_1 W_2 S_{i+1}^n X_1^n Y_1^{i-1} Y_2^{i-1}) \\ &\quad - H(S_i | W_1 W_2 S_{i+1}^n X_{1i}) + H(S_i | W_1 W_2 S_{i+1}^n X_1^n Y_1^{i-1} Y_2^{i-1})] + n\delta_n \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{i=1}^n [H(Y_{2i}|X_{1i}) - H(Y_{2i}|X_{1i}T_iV_i) \\
&\quad - H(S_i|X_{1i}) + H(S_i|X_{1i}T_iV_i)] + n\delta_n \\
&= \sum_{i=1}^n [I(T_iV_i; Y_{2i}|X_i) - I(T_iV_i; S_i|X_{1i})] + n\delta_n.
\end{aligned} \tag{B.50}$$

where (a) follows due to the degradedness condition (3.2), and because X_{1i} is a function of W_1 .

B.8 Proof of the Converse for Theorem 3.7

We define the auxiliary random variable $T_i = (W_1 S_{i+1}^n X_1^n Y_1^{i-1})$, which satisfies the Markov chain:

$$T_i \leftrightarrow X_{1i} X_{2i} S_i \leftrightarrow Y_{1i} Y_{2i}, \quad \text{for } i = 1, \dots, n. \tag{B.51}$$

Following (B.49), we obtain

$$nR_1 \leq \sum_{i=1}^n [I(T_i X_{1i}; Y_{1i}) - I(T_i; S_i|X_{1i})] + n\delta_n.$$

We next bound R_2 as follows.

$$\begin{aligned}
nR_2 &= I(W_2; Y_2^n) + n\delta_n \\
&\stackrel{(a)}{\leq} I(W_2; Y_2^n | W_1 S^n X_1^n) + n\delta_n \\
&= \sum_{i=1}^n [I(W_2; Y_{2i} | W_1 S^n X_1^n Y_2^{i-1})] + n\delta_n \\
&\leq \sum_{i=1}^n H(Y_{2i} | W_1 S^n X_1^n Y_2^{i-1}) + n\delta_n \\
&\stackrel{(b)}{=} \sum_{i=1}^n H(Y_{2i} | W_1 S^n X_1^n Y_1^{i-1} Y_2^{i-1}) + n\delta_n
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(c)}{\leq} \sum_{i=1}^n H(Y_{2i}|W_1 S_{i+1}^n X_1^n Y_1^{i-1} S_i) + n\delta_n \\
&= \sum_{i=1}^n H(Y_{2i}|X_{1i} T_i S_i) + n\delta_n
\end{aligned} \tag{B.52}$$

where (a) follows because W_2 is independent of (W_1, S^n, X_1^n) , (b) follows due to the degradedness condition (3.1), and (c) follows because conditioning does not increase entropy.

We then derive another bound on R_2 by continuing to derive the bound (B.44) as follows:

$$\begin{aligned}
&nR_2 \\
&\leq \sum_{i=1}^n [H(Y_{2i}|W_1 Y_2^{i-1}) - H(Y_{2i}|W_1 W_2 S_{i+1}^n Y_2^{i-1}) \\
&\quad - H(S_i|W_1 W_2 S_{i+1}^n) + H(S_i|W_1 W_2 S_{i+1}^n Y_2^{i-1})] + n\delta_n \\
&\stackrel{(a)}{=} \sum_{i=1}^n [H(Y_{2i}|W_1 X_1^n Y_2^{i-1}) - H(Y_{2i}|W_1 W_2 S_{i+1}^n Y_2^{i-1}) - H(S_i|W_1 W_2 X_1^n S_{i+1}^n) \\
&\quad + H(S_i|W_1 W_2 X_1^n S_{i+1}^n Y_1^{i-1} Y_2^{i-1} Y_{2i}) + I(Y_{2i}; S_i|W_1 W_2 S_{i+1}^n Y_2^{i-1})] + n\delta_n \\
&\stackrel{(b)}{\leq} \sum_{i=1}^n [H(Y_{2i}|X_{1i}) - H(S_i|X_{1i}) + H(S_i|X_{1i} T_i Y_{2i})] + n\delta_n \\
&= \sum_{i=1}^n [H(Y_{2i}|X_{1i}) - I(T_i Y_{2i}; S_i|X_{1i})] + n\delta_n.
\end{aligned} \tag{B.53}$$

where (a) follows because X_1^n is a function of W_1 and from the degradedness condition (3.1), and (b) follows because S_i is independent of (W_1, W_2, X_1^n) , and conditioning does not increase entropy, and follows from the definition of T_i .

APPENDIX C

PROOF FOR CHAPTER 5

C.1 Proof of Proposition 5.1

The first bound follows easily from the single-user rate bound of receiver 1 as follows.

$$\begin{aligned}
 nR_1 &\leq I(W_1; Y_1^n) + n\epsilon_n \\
 &\leq I(W_1; Y_1^n S_1^n X_0^n) + n\epsilon_n \\
 &= I(W_1; Y_1^n | S_1^n X_0^n) + n\epsilon_n \\
 &\leq h(Y_1^n | S_1^n X_0^n) - h(Y_1^n | W_1 S_1^n X_1^n X_0^n) + n\epsilon_n \\
 &= h(X_1^n + N_1^n) - h(N_1^n) + n\epsilon_n \\
 &\leq \frac{n}{2} \log(1 + P_1)
 \end{aligned} \tag{C.1}$$

We then bound the sum rate as follows. For the message W_0 , based on Fano's inequality, we have

$$\begin{aligned}
 nR_0 &\leq I(W_0; Y_0^n) + n\epsilon_n \\
 &= h(Y_0^n) - h(Y_0^n | W_0) + n\epsilon_n,
 \end{aligned} \tag{C.2}$$

where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$.

For the message W_1 , based on Fano's inequality, we have

$$\begin{aligned}
 nR_1 &\leq I(W_1; Y_1^n) + n\epsilon_n \\
 &= h(Y_1^n) - h(Y_1^n | W_1) + n\epsilon_n \\
 &\leq h(Y_1^n) - h(Y_1^n | W_1 X_1^n) + n\epsilon_n \\
 &= h(Y_1^n) - h(X_0^n + S_1^n + N_1^n) + n\epsilon_n \\
 &\leq h(Y_1^n) - h(X_0^n + S_1^n + N_1^n | W_0 Y_0^n) + n\epsilon_n
 \end{aligned} \tag{C.3}$$

Summation of (C.2) and (C.3) yields

$$\begin{aligned}
 n(R_0 + R_1) &\leq h(Y_0^n) + h(Y_1^n) - h(Y_0^n, X_0^n + S_1^n + N_1^n | W_0) + n\epsilon_n \\
 &= h(Y_0^n) + h(Y_1^n) - h(X_0^n + N_0^n, X_0^n + S_1^n + N_1^n | W_0) + n\epsilon_n
 \end{aligned} \tag{C.4}$$

Since the two receivers perform decoding independently, the capacity region of the channel depends on only the marginal distributions of (X_0, Y_0) and (X_0, X_1, S, Y_1) . It is clear that setting $N_1 = N_0$ does not change the two marginal distributions respectively involving Y_0 and Y_1 , and hence does not affect the capacity region. Thus,

$$\begin{aligned}
 n(R_0 + R_1) &\leq h(Y_0^n) + h(Y_1^n) - h(X_0^n + N_1^n, X_0^n + S_1^n + N_1^n | W_0) + n\epsilon_n \\
 &\leq h(Y_0^n) + h(Y_1^n) - h(S_1^n, X_0^n + N_1^n | W_0) + n\epsilon_n \\
 &\leq h(Y_0^n) + h(Y_1^n) - h(S_1^n) - h(N_1^n) + n\epsilon_n \\
 &\leq \frac{n}{2} \log(1 + P_0) + \frac{n}{2} \log \left(1 + \frac{P_0 + 2\sqrt{P_0 Q_1} + P_1 + 1}{Q_1} \right) + n\epsilon_n
 \end{aligned} \tag{C.5}$$

As $Q_1 \rightarrow \infty$, the second term of the above bound goes to 0, and we have

$$R_0 + R_1 \leq \frac{1}{2} \log(1 + P_0). \tag{C.6}$$

C.2 Proof of Proposition 5.2

We use random codes and fix the following joint distribution:

$$P_{S_1 X'_0 U X_0 X_1 Y_0 Y_1} = P_{S_1} P_{X'_0} P_{U|S_1 X'_0} P_{X_0|U S_1 X'_0} P_{X_1} P_{Y_0|X_0} P_{Y_1|X_0 X_1 S_1}.$$

Let $T_\epsilon^n(P_{S_1 X'_0 U X_0 X_1 Y_0 Y_1})$ denote the strongly joint ϵ -typical set (see, e.g., [55, Sec. 10.6] and [56, Sec. 1.3] for definition) based on the above distribution. For a given sequence x^n , let $T_\epsilon^n(P_{U|X}|x^n)$ denote the set of sequences u^n such that (u^n, x^n) is jointly typical based on the distribution P_{XU} .

1. Codebook Generation

- Generate $2^{n\tilde{R}}$ i.i.d. codewords $u^n(t)$ according to $P(u^n) = \prod_{i=1}^n P_U(u_i)$ for the fixed marginal probability P_U as defined, in which $t \in [1, 2^{n\tilde{R}}]$.
- Generate 2^{nR_0} i.i.d codewords $x_0'^n(w_0)$ according to $P(x_0'^n) = \prod_{i=1}^n P_{X'_0}(x_{0i}')$ for the fixed marginal probability $P_{X'_0}$ as defined, in which $w_0 \in [1, 2^{nR_0}]$.
- Generate 2^{nR_1} i.i.d. codewords $x_1^n(w_1)$ according to $P(x_1^n) = \prod_{i=1}^n P_{X_1}(x_{1i})$ for the fixed marginal probability P_{X_1} as defined, in which $w_1 \in [1, 2^{nR_1}]$.

2. Encoding

- Encoder at the helper: Given w_0 , map w_0 into $x_0'^n(w_0)$. For each $x_0'^n(w_0)$, select \tilde{t} such that $(u^n(\tilde{t}), s_1^n, x_0'^n(w_0)) \in T_\epsilon^n(P_{S_1} P_{X'_0} P_{U|S_1 X'_0})$. If $u^n(\tilde{t})$ cannot be found, set $\tilde{t} = 1$. Then map $(s_1^n, u^n(\tilde{t}), x_0'^n(w_0))$ into $x_0^n = f_0^{(n)}(x_0'^n(w_0), s_1^n, u^n(\tilde{t}))$. Based on the rate distortion type of argument [55, Sec. 10.5] or the Covering Lemma [57, Sec. 3.7], it can be shown that such $u^n(\tilde{t})$ exists with high probability for large n if

$$\tilde{R} > I(U; S_1 X'_0). \quad (\text{C.7})$$

- Encoder 1: Given w_1 , map w_1 into $x_1^n(w_1)$.

3. Decoding

- Decoder 0: Given y_0^n , find \hat{w}_0 such that $(x_0'^n(\hat{w}_0), y_0^n) \in T_\epsilon^n(P_{X_0'Y_0})$. If no or more than one \hat{w}_0 can be found, declare an error. It can be shown that the decoding error is small for sufficient large n if

$$R_0 \leq I(X_0'; Y_0). \quad (\text{C.8})$$

The proof for the above bound (and the similar bounds in the sequel) follows the standard techniques as given in [55, Sec. 7.7], and hence is omitted.

- Decoder 1: Given y_1^n , find a pair (\hat{t}, \hat{w}_1) such that $(u^n(\hat{t}), x_1^n(\hat{w}_1), y_1^n) \in T_\epsilon^n(P_{UX_1Y_1})$. If no or more than one such pair can be found, then declare an error. It can be shown that decoding is successful with small probability of error for sufficiently large n if the following conditions are satisfied

$$R_1 \leq I(X_1; Y_1|U), \quad (\text{C.9})$$

$$\tilde{R} \leq I(U; Y_1|X_1), \quad (\text{C.10})$$

$$R_1 + \tilde{R} \leq I(UX_1; Y_1). \quad (\text{C.11})$$

We note that (C.10) corresponds to the decoding error for the index t , which is not the message of interest. Hence, the bound (C.10) can be removed. Hence, combining (C.7), (C.8), (C.9), and (C.11), and eliminating \tilde{R} , we obtain the desired achievable region.

C.3 Proof of Proposition 5.4

The single rate bounds follow from Proposition 5.1 and the single-user channel capacity.

For the sum rate bound, based on Fano's inequality, we have

$$\begin{aligned}
n(R_1 + R_2) &\leq I(W_1; Y_1^n) + I(W_2; Y_2^n) + n\epsilon_n \\
&= h(Y_1^n) - h(Y_1^n | W_1) + h(Y_2^n) - h(Y_2^n | W_2) + n\epsilon_n \\
&\stackrel{(a)}{=} h(Y_1^n) - h(Y_1^n | W_1 X_1^n) + h(Y_2^n) - h(Y_2^n | W_2 X_2^n) + n\epsilon_n \\
&= h(Y_1^n) - h(X_0^n + S_1^n + N_1^n) + h(Y_2^n) - h(X_0^n + N_2^n) + n\epsilon_n \\
&\leq h(Y_1^n) - h(X_0^n + S_1^n + N_1^n | X_0^n + N_1^n) \\
&\quad + h(Y_2^n) - h(X_0^n + N_2^n) + n\epsilon_n
\end{aligned}$$

where (a) follows from the fact that X_1^n is a function of W_1 , and X_2^n is a function of W_2 , and they are independent from X_0^n , state and noise. As argued in Appendix C.1, setting $N_1^n = N_2^n$ does not change the capacity region. Thus,

$$\begin{aligned}
n(R_1 + R_2) &\leq h(Y_1^n) - h(X_0^n + S_1^n + N_1^n, X_0^n + N_1^n) + h(Y_2^n) + n\epsilon_n \\
&= h(Y_1^n) - h(S_1^n, X_0^n + N_1^n) + h(Y_2^n) + n\epsilon_n \\
&= h(Y_1^n) - h(S_1^n) - h(X_0^n + N_1^n | S_1^n) + h(Y_2^n) + n\epsilon_n \\
&\leq h(Y_1^n) - h(S_1^n) - h(X_0^n + N_1^n | S_1^n, X_0^n) + h(Y_2^n) + n\epsilon_n \\
&\stackrel{(b)}{=} h(X_0^n + X_1^n + S_1^n + N_1^n) - h(S_1^n) - h(N_1^n) + h(X_0^n + X_2^n + N_1^n) + n\epsilon_n \\
&\leq \frac{n}{2} \log 2\pi e(P_1 + P_0 + 2\sqrt{P_0 Q_1} + Q_1 + 1) - \frac{n}{2} \log(2\pi e Q_1) \\
&\quad + \frac{n}{2} \log 2\pi e(P_0 + P_2 + 1) - \frac{n}{2} \log(2\pi e) + n\epsilon_n \\
&= \frac{n}{2} \log \left(\frac{P_1 + P_0 + 2\sqrt{P_0 Q_1} + Q_1 + 1}{Q_1} \right) + \frac{n}{2} \log(P_0 + P_2 + 1) + n\epsilon_n \\
&\rightarrow \frac{n}{2} \log(P_0 + P_2 + 1) \text{ as } Q_1 \rightarrow \infty
\end{aligned}$$

where (b) follows from the fact that X_0^n and S_1^n are independent from N_1^n .

C.4 Proof of Proposition 5.5

We use random codes and fix the following joint distribution:

$$P_{S_1UVX_0X_1X_2Y_1Y_2} = P_{VUS_1}P_{X_0|VUS_1}P_{X_1}P_{X_2}P_{Y_1|X_0X_1S_1}P_{Y_2|X_0X_2}.$$

Let $T_\epsilon^n(P_{S_1UVX_0X_1X_2Y_1Y_2})$ denote the strongly joint ϵ -typical set based on the above distribution.

1. Codebook Generation

- Generate $2^{n\tilde{R}_1}$ i.i.d. codewords $u^n(t)$ according to $P(u^n) = \prod_{i=1}^n P_U(u_i)$ for the fixed marginal probability P_U as defined, in which $t \in [1, 2^{n\tilde{R}_1}]$.
- Generate $2^{n\tilde{R}_2}$ i.i.d. codewords $v^n(k)$ according to $P(v^n) = \prod_{i=1}^n P_V(v_i)$ for the fixed marginal probability P_V as defined, in which $k \in [1, 2^{n\tilde{R}_2}]$.
- Generate 2^{nR_1} i.i.d. codewords $x_1^n(w_1)$ according to $P(x_1^n) = \prod_{i=1}^n P_{X_1}(x_{1i})$ for the fixed marginal probability P_{X_1} as defined, in which $w_1 \in [1, 2^{nR_1}]$.
- Generate 2^{nR_2} i.i.d. codewords $x_2^n(w_2)$ according to $P(x_2^n) = \prod_{i=1}^n P_{X_2}(x_{2i})$ for the fixed marginal probability P_{X_2} as defined, in which $w_2 \in [1, 2^{nR_2}]$.

2. Encoding

- Encoder at the helper: Given s_1^n , find \tilde{t} , such that $(u^n(\tilde{t}), s_1^n) \in T_\epsilon^n(P_{S_1U})$. Such $u^n(\tilde{t})$ exists with high probability for large n if

$$\tilde{R}_1 \geq I(U; S_1). \quad (\text{C.12})$$

- For each \tilde{t} selected, select \tilde{k} , such that $(v^n(\tilde{k}), u^n(\tilde{t}), s_1^n) \in T_\epsilon^n(P_{VUS_1})$. Such $v^n(\tilde{k})$ exists with high probability for large n if

$$\tilde{R}_2 \geq I(V; S_1 U). \quad (\text{C.13})$$

- Map (s_1^n, u^n, v^n) into $x_0^n = f_0^{(n)}(u^n(\tilde{t}), v^n(\tilde{k}), s_1^n)$.
- Encoder 1: Given w_1 , map w_1 into $x_1^n(w_1)$.
- Encoder 2: Given w_2 , map w_2 into $x_2^n(w_2)$.

3. Decoding

- Decoder 1: Given y_1^n , find (\hat{w}_1, \hat{t}) such that $(x_1^n(\hat{w}_1), u^n(\hat{t}), y_1^n) \in T_\epsilon^n(P_{X_1UY_1})$.
If no or more than one \hat{w}_1 can be found, declare an error. One can show that the decoding error is small for sufficient large n if

$$R_1 \leq I(X_1; Y_1 U) \quad (\text{C.14})$$

$$R_1 + \tilde{R}_1 \leq I(X_1 U; Y_1). \quad (\text{C.15})$$

- Decoder 2: Given y_2^n , find (\hat{w}_2, \hat{k}) such that $(x_2^n(\hat{w}_2), v^n(\hat{k}), y_2^n) \in T_\epsilon^n(P_{X_2VY_2})$.
If no or more than one \hat{w}_2 can be found, declare an error. One can show that the decoding error is small for sufficient large n if

$$R_2 \leq I(X_2; Y_2 V), \quad (\text{C.16})$$

$$R_2 + \tilde{R}_2 \leq I(X_2 V; Y_2). \quad (\text{C.17})$$

Combining (C.12)-(C.17), and eliminating \tilde{R}_1 and \tilde{R}_2 , we have

$$R_1 \leq I(X_1; Y_1 U) \quad (\text{C.18a})$$

$$R_1 \leq I(X_1 U; Y_1) - I(U; S_1) \quad (\text{C.18b})$$

$$R_2 \leq I(X_2; Y_2 V) \quad (\text{C.18c})$$

$$R_2 \leq I(X_2 V; Y_2) - I(V; U S_1) \quad (\text{C.18d})$$

When conditions (5.21a) and (5.21b) are satisfied, (C.18b) and (C.18d) are redundant, and hence, we have the desired achievable region.

C.5 Proof of Proposition 5.7

The bounds on R_1 and R_2 follow from the single-user channel capacity. For the sum rate bound, based on the Fano's inequality, we have

$$\begin{aligned} n(R_1 + R_2) &\leq I(W_1; Y_1^n) + I(W_2; Y_2^n) + n\epsilon_n \\ &= h(Y_1^n) - h(Y_1^n | W_1) + h(Y_2^n) - h(Y_2^n | W_2) + n\epsilon_n \\ &\stackrel{(a)}{=} h(Y_1^n) - h(Y_1^n | W_1 X_1^n) + h(Y_2^n) - h(Y_2^n | W_2 X_2^n) + n\epsilon_n \\ &= h(Y_1^n) - h(X_0^n + S_1^n + N_1^n) + h(Y_2^n) - h(X_0^n + S_2^n + N_2^n) + n\epsilon_n \\ &\leq h(Y_1^n) - h(X_0^n + S_1^n + N_1^n | X_0^n + N_1^n) \\ &\quad + h(Y_2^n) - h(X_0^n + S_2^n + N_2^n | X_0^n + N_2^n, X_0^n + S_1^n + N_1^n) \\ &\quad + h(X_0^n + N_1^n) - h(X_0^n + N_1^n) + n\epsilon_n \end{aligned}$$

where (a) follows from the fact that X_1^n is a function of W_1 , X_2^n is a function of W_2 , and they are independent from X_0^n , S_1^n , S_2^n , N_1^n and N_2^n . Since receivers 1 and 2 decode based on the marginal distributions only, setting $N_1^n = N_2^n$ does not affect the channel capacity.

Therefore,

$$\begin{aligned}
& n(R_1 + R_2) \\
& \leq h(Y_1^n) - h(X_0^n + S_1^n + N_1^n, X_0^n + S_2^n + N_1^n, X_0^n + N_1^n) \\
& \quad + h(Y_2^n) + h(X_0^n + N_1^n) + n\epsilon_n \\
& = h(Y_1^n) - h(S_1^n, S_2^n, X_0^n + N_1^n) + h(Y_2^n) + h(X_0^n + N_1^n) + n\epsilon_n \\
& = h(Y_1^n) - h(S_1^n) - h(S_2^n) - h(X_0^n + N_1^n | S_1^n, S_2^n) \\
& \quad + h(Y_2^n) + h(X_0^n + N_1^n) + n\epsilon_n \\
& \leq h(Y_1^n) - h(S_1^n) - h(S_2^n) - h(X_0^n + N_1^n | S_1^n, S_2^n, X_0^n) \\
& \quad + h(Y_2^n) + h(X_0^n + N_1^n) + n\epsilon_n \\
& \stackrel{(b)}{=} h(Y_1^n) - h(S_1^n) - h(S_2^n) - h(N_1^n) + h(Y_2^n) + h(X_0^n + N_1^n) + n\epsilon_n \\
& \stackrel{(c)}{=} \sum_{i=1}^n h(Y_{1i} | Y_1^{i-1}) - h(S_{1i}) - h(S_{2i}) - h(N_{1i}) \\
& \quad + h(Y_{2i} | Y_2^{i-1}) + h(X_{0i} + N_{1i} | X_0^{i-1} + N_1^{i-1}) + n\epsilon_n \\
& \leq \sum_{i=1}^n h(Y_{1i}) - h(S_{1i}) - h(S_{2i}) - h(N_{1i}) + h(Y_{2i}) + h(X_{0i} + N_{1i}) + n\epsilon_n \\
& = \sum_{i=1}^n [h(X_{0i} + X_{1i} + S_{1i} + N_{1i}) - h(S_{1i}) - h(S_{2i}) - h(N_{1i}) \\
& \quad + h(X_{0i} + X_{2i} + S_{2i} + N_{1i}) + h(X_{0i} + N_{1i})] + n\epsilon_n \tag{C.19}
\end{aligned}$$

We then derive the items respectively. The first term in (C.19) can be derived as

$$\begin{aligned}
& \sum_{i=1}^n h(X_{0i} + X_{1i} + S_{1i} + N_{1i}) \\
& \stackrel{(d)}{\leq} \frac{1}{2} \sum_{i=1}^n \log 2\pi e (E(X_{0i} + X_{1i} + S_i + N_i)^2) \\
& \leq \frac{1}{2} \sum_{i=1}^n \log 2\pi e \left(E[X_{0i}^2] + E(X_{0i}S_i) + E[S_i^2] + E[X_{1i}^2] + E[N_i^2] \right)
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(e)}{\leq} \frac{n}{2} \log 2\pi e \left(\frac{1}{n} \sum_{i=1}^n E[X_{0i}^2] + \frac{2}{n} \sum_{i=1}^n E(X_{0i}S_i) \right. \\
& \quad \left. + \frac{1}{n} \sum_{i=1}^n E[S_i^2] + \frac{1}{n} \sum_{i=1}^n E[X_{1i}^2] + \frac{1}{n} \sum_{i=1}^n E[N_i^2] \right) \\
& \stackrel{(f)}{\leq} \frac{n}{2} \log 2\pi e \left(P_0 + Q + P_1 + 1 + \frac{2}{n} \sum_{i=1}^n E(X_{1i}S_i) \right) \\
& \leq \frac{n}{2} \log 2\pi e \left(P_0 + P_1 + Q + 1 + 2\sqrt{P_0Q} \right) \tag{C.20}
\end{aligned}$$

where (d) follows from the fact that the Gaussian distribution maximizes the entropy given the variance of the random variable, (e) follows from the concavity of the logarithm function and Jensen's inequality, and (f) follows from the power constraints. Similarly, we have

$$\begin{aligned}
\sum_{i=1}^n h(X_{0i} + X_{2i} + S_{1i} + N_{1i}) & \leq \frac{n}{2} \log 2\pi e (P_2 + P_0 + 2\sqrt{P_0Q_2} + Q_2 + 1) \\
\sum_{i=1}^n h(X_{0i} + N_{1i}) & \leq \frac{n}{2} \log(P_0 + 1)
\end{aligned}$$

And hence, we have

$$\begin{aligned}
& n(R_1 + R_2) \\
& \leq \frac{n}{2} \log 2\pi e (P_1 + P_0 + 2\sqrt{P_0Q_1} + Q_1 + 1) - \frac{n}{2} \log(2\pi e Q_1) - \frac{n}{2} \log(2\pi e Q_2) \\
& \quad - \frac{n}{2} \log(2\pi e) + \frac{n}{2} \log 2\pi e (P_2 + P_0 + 2\sqrt{P_0Q_2} + Q_2 + 1) + \frac{n}{2} \log 2\pi e (P_0 + 1) + n\epsilon_n \\
& \leq \frac{n}{2} \log \left(\frac{P_1 + P_0 + 2\sqrt{P_0Q_1} + Q_1 + 1}{Q_1} \right) + \frac{n}{2} \log \left(\frac{P_2 + P_0 + 2\sqrt{P_0Q_2} + Q_2 + 1}{Q_2} \right) \\
& \quad + \frac{n}{2} \log(P_0 + 1) + n\epsilon_n \\
& \rightarrow \frac{n}{2} \log(P_0 + 1) \quad \text{as } Q_1 \rightarrow \infty, Q_2 \rightarrow \infty
\end{aligned}$$

where (b) follows from the fact that X_0^n , S_1^n and S_2^n are independent from N_1^n .

C.6 Proof of Theorem 5.5

The proof contains two parts: 1. we first show that if $P_1 + P_2 \geq P_0 + 1$, then the sum capacity can be obtained; 2. we further characterize the time allocation parameters γ that achieves the sum capacity.

1. For a given P_0 , we consider the following two cases.

a). If the power constraint satisfies $P_1 + P_2 = P_0 + 1$, by applying Proposition 5.8, and by setting $\gamma = \frac{P_1}{P_1 + P_2}$, the point $(R_1, R_2) = (\frac{P_1}{2(P_1 + P_2)} \log(1 + P_0), \frac{P_2}{2(P_1 + P_2)} \log(1 + P_0))$ is achievable, which achieves the sum rate outer bound in Proposition 5.7.

b). If $P_1 + P_2 \geq P_0 + 1$, we set the actual transmission power \tilde{P}_1 and \tilde{P}_2 of transmitters 1 and 2 to satisfy $\tilde{P}_1 + \tilde{P}_2 = P_0 + 1$, $\tilde{P}_1 \leq P_1$ and $\tilde{P}_2 \leq P_2$. Then following a), the sum capacity is obtained.

2. In order for each transmitter to achieve the sum capacity during its own transmission slot, (5.30) together with (5.31a) and (5.31b) imply that

$$\frac{P_1}{\gamma} \geq P_0 + 1 \quad (\text{C.21})$$

$$\frac{P_2}{1 - \gamma} \geq P_0 + 1. \quad (\text{C.22})$$

It is clear that (C.21) implies

$$\gamma \leq \frac{P_1}{P_0 + 1},$$

and (C.22) implies

$$\gamma \geq 1 - \frac{P_2}{P_0 + 1}.$$

Considering $0 \leq \gamma \leq 1$, we obtain the desired bounds on γ .

APPENDIX D

PROOF FOR CHAPTER 6

D.1 Proof of Proposition 6.1

We show the outer bound that involves the impact of the helper. In particular, we focus on the sum rate bound. The single rate bounds follow from the similar steps.

For the sum rate bound, based on Fano's inequality, we have

$$\begin{aligned}
& n(R_1 + R_2) \\
& \leq I(W_1 W_2; Y^n) + n\epsilon_n \\
& = h(Y^n) - h(Y^n | W_1 W_2) + n\epsilon_n \\
& \leq h(Y^n) - h(X_0^n + X_1^n + X_2^n + S^n + N^n | W_1 X_1^n W_2 X_2^n) + n\epsilon_n \\
& \stackrel{(a)}{=} h(Y^n) - h(X_0^n + S^n + N^n) + n\epsilon_n \\
& \leq h(Y^n) - h(X_0^n + S^n + N^n | X_0^n + N^n) + n\epsilon_n \\
& = h(Y^n) - h(X_0^n + S^n + N^n | X_0^n + N^n) + h(X_0^n + N^n) - h(X_0^n + N^n) + n\epsilon_n \\
& = h(Y^n) - h(X_0^n + S^n + N^n, X_0^n + N^n) + h(X_0^n + N^n) + n\epsilon_n \\
& = h(Y^n) - h(S^n, X_0^n + N^n) + h(X_0^n + N^n) + n\epsilon_n
\end{aligned}$$

$$\begin{aligned}
&= h(Y^n) - h(S^n) - h(X_0^n + N^n | S^n) + h(X_0^n + N^n) + n\epsilon_n \\
&\leq h(Y^n) - h(S^n) - h(X_0^n + N^n | S^n X_0^n) + h(X_0^n + N^n) + n\epsilon_n \\
&\stackrel{(b)}{=} h(Y^n) - h(S^n) - h(N^n) + h(X_0^n + N^n) + n\epsilon_n \\
&\leq \frac{n}{2} \log 2\pi e (P_0 + P_1 + P_2 + 2\rho_{0S} \sqrt{P_0 Q} + Q + 1) \\
&\quad - \frac{n}{2} \log(2\pi e Q) + \frac{n}{2} \log 2\pi e (P_0 + 1) - \frac{n}{2} \log(2\pi e) \\
&\leq \frac{n}{2} \log(P_0 + 1) \\
&\quad + \frac{n}{2} \log\left(\frac{P_0 + P_1 + P_2 + 2\rho_{0S} \sqrt{P_0 Q} + Q + 1}{Q}\right)
\end{aligned}$$

where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$, $\rho_{0S} = \frac{\sum_{i=1}^n E(X_{0i} S_i)}{\sqrt{P_0 Q}}$, (a) follows because (X_0^n, S^n, N^n) are independent from (W_1, X_1^n, W_2, X_2^n) , and (b) follows because X_0^n and S^n are independent from N^n .

D.2 Proof of Lemma 6.2

We design the following scheme for the discrete memoryless with state noncausally known at the helper.

We use random codes and fix the following joint distribution:

$$P_{SUX_0X_1X_2Y} = P_S P_{U|S} P_{X_0|US} P_{X_1} P_{X_2} P_{Y|X_0X_1X_2S}$$

Let $T_\epsilon^n(P_{SUX_0X_1X_2Y})$ denote the strongly joint ϵ -typical set based on the above distribution.

For a given sequence x^n , let $T_\epsilon^n(P_{U|X}|x^n)$ denote the set of sequences u^n such that (u^n, x^n) is jointly typical based on the distribution P_{XU} .

1. Codebook Generation

- Generate $2^{n\tilde{R}}$ codewords $U^n(t)$ with the probability of P_U , in which $t \in [1, 2^{n\tilde{R}}]$.

- Generate 2^{nR_1} codewords $X_1^n(w_1)$ with the probability of P_{X_1} , in which $w_1 \in [1, 2^{nR_1}]$.
- Generate 2^{nR_2} codewords $X_2^n(w_2)$ with the probability of P_{X_2} , in which $w_2 \in [1, 2^{nR_2}]$.

2. Encoding

- Encoder 0: For given s^n , select \tilde{t} such that $(u^n(\tilde{t}), s^n) \in T_\epsilon^n(P_{SU})$. If $u^n(\tilde{t})$ can be found, map $(s^n, u^n(\tilde{t}))$ into x_0^n , else, $x_0^n = f^n(s^n, u^n(1))$.

It is easy to show that such $u^n(\tilde{t})$ exists with high probability for large n if

$$\tilde{R} > I(U; S). \quad (\text{D.1})$$

- Encoder 1: Given w_1 , map w_1 into $x_1^n(w_1)$.
- Encoder 2: Given w_2 , map w_2 into $x_2^n(w_2)$.

3. Decoding: Given y^n ,

- Find \hat{t} such that $(u^n(\hat{t}), y^n) \in T_\epsilon^n(P_{UY})$. One can show that the decoding error is small for sufficient large n if

$$\tilde{R} \leq I(U; Y). \quad (\text{D.2})$$

- For selected u^n , find \hat{w}_1 and \hat{w}_2 such that $(x_1(\hat{w}_1), (x_2(\hat{w}_2), u^n(\hat{t}), y^n) \in T_\epsilon^n(P_{X_1'UY})$. One can show that the decoding error is small for sufficient large n if

$$R_1 \leq I(X_1; Y | U X_2) \quad (\text{D.3})$$

$$R_2 \leq I(X_2; Y | U X_1) \quad (\text{D.4})$$

$$R_1 + R_2 \leq I(X_1 X_2; Y | U) \quad (\text{D.5})$$

According to (D.1)- (D.5), exploit the Foriour-Mozkin elimination to eliminate \tilde{R} , we have the achievable region as in Lemma 6.2.

APPENDIX E

PROOF FOR CHAPTER 7

E.1 Proof of Proposition 7.5

We use random codes and fix the following joint distribution:

$$P_{SUX_0VXY_1Y_2} = P_S P_{U|S} P_{X_0|US} P_V P_{X|V} P_{Y_1|X_0XS} P_{Y_2|X_0XS}.$$

Let $T_\epsilon^n(P_{SUX_0VXY_1Y_2})$ denote the strongly joint ϵ -typical set based on the above distribution.

1. Code Construction:

- (a) Generate 2^{nR_1} codewords $v^n(w_1)$ with i.i.d. components based on P_V . Index these codewords by $w_1 = 1, \dots, 2^{nR_1}$.
- (b) For each $v^n(w_1)$, generate 2^{nR_2} codewords $x^n(w_1, w_2)$ with i.i.d. components based on $P_{X|V}$. Index these codewords by $w_2 = 1, \dots, 2^{nR_2}$.
- (c) Generate $2^{n\tilde{R}}$ codewords $U^n(l)$ with i.i.d. components based on P_U . Index these codewords by $l = 1, 2, \dots, 2^{n\tilde{R}}$.

2. Encoding:

- (a) Encoder: Given w_1 , map it into $v^n(w_1)$. Given v^n and w_2 , map it into $x^n(w_1, w_2)$.
- (b) Encoder at the helper: Given s^n , select \tilde{l} such that

$$(u^n(\tilde{l}), s^n) \in T_\epsilon^n(P_{US}).$$

Otherwise, set $\tilde{l} = 1$. It can be shown that for large n , such u^n exists with high probability if

$$\tilde{R} > I(U; S). \quad (\text{E.1})$$

- (c) Given selected $u^n(\tilde{l})$ and s^n , generate x_0^n with i.i.d. components based on $P_{X_0|US}$ for transmission.

3. Decoding: Given y^n ,

(a) Decoder 1:

- i. Find \hat{l} such that $(u^n(\hat{l}), y_1^n) \in T_\epsilon^n(P_{UY_1})$. One can show that the decoding error is small for sufficient large n if

$$\tilde{R} \leq I(U; Y_1). \quad (\text{E.2})$$

- ii. For selected u^n , find \hat{w}_1 such that $(v(\hat{w}_1), u^n(\hat{l}), y_1^n) \in T_\epsilon^n(P_{VUY_1})$. One can show that the decoding error is small for sufficient large n if

$$R_1 \leq I(V; Y_1|U). \quad (\text{E.3})$$

(b) Decoder 2:

- i. Find \hat{l} such that $(u^n(\hat{l}), y_2^n) \in T_\epsilon^n(P_{UY_2})$. One can show that the decoding error is small for sufficient large n if

$$\tilde{R} \leq I(U; Y_2). \quad (\text{E.4})$$

- ii. For selected u^n , find \hat{w}_2 such that $(v(\hat{w}_1), x(\hat{w}_1, \hat{w}_2), u^n(\hat{l}), y_2^n) \in T_\epsilon^n(P_{VXY_2})$. One can show that the decoding error is small for sufficient large n if

$$R_2 \leq I(X; Y_2 | VU) \quad (\text{E.5})$$

$$R_1 + R_2 \leq I(X; Y_2 | U) \quad (\text{E.6})$$

According to (E.1)- (E.6), exploit the Foriour-Mozkin elimination to eliminate \tilde{R} , we have the achievable region as in Proposition 7.5.

REFERENCES

- [1] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
- [2] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inform. Theory*, vol. 21, no. 5, Sep. 1975.
- [3] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.
- [4] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
- [5] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.
- [6] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.
- [7] S. Hajizadeh, M. Monemizadeh, E. Bahmani, G. A. Hodtani, and M. Joneidi, "State-dependent Z channel," Available at <http://arxiv.org/abs/1301.6272>.
- [8] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, Tech. Rep., 1958.

- [9] S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, Jan. 1980.
- [10] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [11] S. Gel'fand and M. Pinsker, "On Gaussian channels with random parameters," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Russian, Jun. 1984.
- [12] R. Khosravi-Farsani and F. Marvasti, "Capacity bounds for multiuser channels with non-causal channel state information at the transmitters," in *Proc. IEEE Information Theory Workshop (ITW)*, Paraty, Brazil, Oct. 2011.
- [13] A. Lapidoth and L. Wang, "The state-dependent semideterministic broadcast channel," *IEEE Trans. Inform. Theory*, vol. 59, no. 4, pp. 2242–2251, Apr. 2013.
- [14] Y. Steinberg, "Correspondence coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, Aug. 2005.
- [15] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Jul. 2005.
- [16] A. Zaidi, S. Shamai (Shitz), P. Piantanida, and L. Vandendorpe, "Bounds on the capacity of the relay channel with noncausal state at source," *IEEE Trans. Inform. Theory*, vol. 59, no. 5, pp. 2639–2672, May 2013.
- [17] Y.-H. Kim, A. Sutivong, and S. Sigurjónsson, "Multiple user writing on dirty paper," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, Illinois, Jul. 2004.

- [18] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdú, “Cooperative multiple-access encoding with states available at one transmitter,” *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4448–4469, Oct. 2008.
- [19] L. Zhang, T. Liu, and S. Cui, “Symmetric Gaussian interference channel with state information,” in *Proc. 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2011.
- [20] L. Zhang, J. Jiang, and S. Cui, “Gaussian interference channel with state information,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4058–4071, Aug. 2013.
- [21] S. Ghasemi-Goojani and H. Behroozi, “On the achievable rate-regions for state-dependent Gaussian interference channel,” Available at <http://arxiv.org/abs/1301.5535>, submitted in January 2013.
- [22] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. Inform. Theory*, vol. 54, no. 12, Dec. 2008.
- [23] I. Sason, “On achievable rate regions for the Gaussian interference channel,” *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.
- [24] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdú, “Cognitive interference channels with state information,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Toronto, Canada, Jul. 2008.
- [25] M. Kazemi and A. Vosoughi, “On the capacity of the state-dependent cognitive interference channel,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Istanbul, Turkey, Jul. 2013.
- [26] W. Wu, S. Vishwanath, and A. Arapostathis, “Capacity of a class of cognitive radio channels: Interference channels with degraded message sets,” *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

- [27] A. Jovicic and P. Viswanath, “Cognitive radio: An information-theoretic perspective,” *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.
- [28] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai (Shitz), “On the capacity of interference channels with one cooperating transmitter,” *European Transactions on Telecommunications*, vol. 19, pp. 405–420, Apr. 2008.
- [29] S. Rini, D. Tuninetti, and N. Devroye, “New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results,” *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4087–4109, Jul. 2011.
- [30] —, “Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results,” *IEEE Trans. Inform. Theory*, vol. 58, no. 2, pp. 820 – 848, Jul. 2012.
- [31] N. Devroye, P. Mitran, and V. Tarokh, “Achievable rates in cognitive radio channels,” *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
- [32] J. Jiang and Y. Xin, “On the achievable rate regions for interference channels with degraded message sets,” *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4707–4712, Oct. 2008.
- [33] Y. Liang, A. Bomekh-baruch, S. S. H. V. Poor, and S. Verdú, “Capacity of cognitive interference channels with and without secrecy,” *IEEE Trans. Inform. Theory*, vol. 55, no. 2, Feb. 2009.
- [34] S. Mallik and R. Koetter, “Helpers for cleaning dirty papers,” in *Proc. IEEE International ITG Conference on Source and Channel Coding (SCC)*, Ulm, Germany, Jan. 2008.

- [35] B. Akhbari, M. Mirmohseni, and M. R. Aref, “Compress-and-forward strategy for the relay channel with non-causal state information,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jul. 2009.
- [36] A. Zaidi, S. P. Kotagiri, J. N. Laneman, and L. Vandendorpe, “Cooperative relaying with state available noncausally at the relay,” *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2272–2298, May 2011.
- [37] S. P. Kotagiri and J. N. Laneman, “Multiaccess channels with state known to some encoders and independent messages,” *EURASIP Journal on Wireless Communications and Networking*, 2008.
- [38] I.-H. Wang, “Distributed interference cancellation in multiple access channels,” *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp. 2781–2787, May 2012.
- [39] T. Philosof and R. Zamir, “The rate loss of single letter characterization for the “dirty” multiple access channel,” in *Proc. IEEE Information Theory Workshop (ITW)*, May 2008, pp. 31–35.
- [40] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, “Lattice strategies for the dirty multiple access channel,” *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5006–5035, Aug. 2011.
- [41] R. Duan and Y. Liang, “Bounds and capacity theorems for cognitive interference channels with state,” *IEEE Trans. Inform. Theory*, vol. 61, no. 1, pp. 280–304, Jan. 2015.
- [42] R. Duan, Y. Liang, A. Khisti, and S. Shamai (Shitz), “Parallel Gaussian networks with a common state-cognitive helper,” Submitted to *IEEE Trans. Inform. Theory*, 2014.

- [43] R. Duan, Y. Liang, and S. Shamai (Shitz), “State-dependent Gaussian interference channels: Can state be fully cancelled?” Submitted to *IEEE Trans. Inform. Theory*, 2015.
- [44] R. Duan and Y. Liang, “Capacity bounds for a class of cognitive interference channels with state,” in *Proc. Annu. Allerton Conf. Communication, Control and Computing*, Allerton, IL, Sep. 2011, pp. 603–608.
- [45] —, “Gaussian cognitive interference channels with state,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Boston, MA, Jul. 2012.
- [46] R. Duan, Y. Liang, and S. Shamai (Shitz), “On the capacity region of Gaussian interference channels with state,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Istanbul, Turkey, Jul. 2013.
- [47] R. Duan, Y. Liang, A. Khisti, and S. Shamai (Shitz), “State-dependent gaussian Z-channel with mismatched side-information and interference,” in *Proc. IEEE Information Theory Workshop (ITW)*, Sevilla, Spain, Sep. 2013.
- [48] —, “State-dependent parallel Gaussian channels with a common helper in high power regime,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Honolulu, HI, Jul. 2014.
- [49] R. Duan, Y. Liang, and S. Shamai (Shitz), “Dirty interference cancellation for Gaussian broadcast channels,” in *Proc. IEEE Information Theory Workshop (ITW)*, Hobart, Tasmania, Australia, Nov. 2014.
- [50] —, “Dirty interference cancelation for multiple access channels,” in *Proc. IEEE Int. Symp. Information Theory and Its Applications (ISITA)*, Melbourne, Australia, Oct. 2014.

- [51] ———, “State-dependent Gaussian Z-interference channel: Capacity results,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Hongkong, China, Jun. 2015.
- [52] A. Schrijver, *Theory of Linear and Integer Programming*. New York: Wiley, 1998.
- [53] I. Csiszár and J. Körner, *Information Theory*. Budapest: Akademiai Kiado, 1981.
- [54] N. M. Blachman, “The convolution inequality for entropy powers,” *IEEE Trans. Inform. Theory*, vol. 11, no. 2, pp. 267–271, Apr. 1965.
- [55] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition*. New York: Wiley, 2006.
- [56] G. Kramer, *Topics in Multi-User Information Theory*. Hanover, MA: Now Publishers Inc, 2008.
- [57] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York: Cambridge University Press, 2011.

VITA

RUCHEN DUAN: She received the B.E. degree from Beijing University of Posts and Telecommunications, Beijing, China in 2010. Since September 2010, she has been a PhD student at Syracuse University. During this period, she received the Syracuse University Fellowship Award for years 2010-2012. Ruchen's research interests focus on information theory and wireless communications.