CODING FOR RELAY NETWORKS WITH PARALLEL GAUSSIAN CHANNELS

A Dissertation

by

YU-CHIH HUANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Krishna R. Narayanan
Committee Members,	Tie Liu
	Srinivas Shakkottai
	Arun R. Srinivasa
Department Head,	Chanan Singh

May 2013

Major Subject: Electrical Engineering

ABSTRACT

A wireless relay network consists of multiple source nodes, multiple destination nodes, and possibly many relay nodes in between to facilitate its transmission. It is clear that the performance of such networks highly depends on information forwarding strategies adopted at the relay nodes. This dissertation studies a particular information forwarding strategy called compute-and-forward. Compute-and-forward is a novel paradigm that tries to incorporate the idea of network coding within the physical layer and hence is often referred to as physical layer network coding. The main idea is to exploit the superposition nature of the wireless medium to directly compute or decode functions of transmitted signals at intermediate relays in a network. Thus, the coding performed at the physical layer serves the purpose of error correction as well as permits recovery of functions of transmitted signals.

For the bidirectional relaying problem with Gaussian channels, it has been shown by Wilson *et al.* and Nam *et al.* that the compute-and-forward paradigm is asymptotically optimal and achieves the capacity region to within 1 bit; however, similar results beyond the memoryless case are still lacking. This is mainly because channels with memory would destroy the lattice structure that is most crucial for the compute-and-forward paradigm. Hence, how to extend compute-and-forward to such channels has been a challenging issue. This motivates this study of the extension of compute-and-forward to channels with memory, such as inter-symbol interference.

The bidirectional relaying problem with parallel Gaussian channels is also studied, which is a relevant model for the Gaussian bidirectional channel with inter-symbol interference and that with multiple-input multiple-output channels. Motivated by the recent success of linear finite-field deterministic model, we first investigate the

corresponding deterministic parallel bidirectional relay channel and fully characterize its capacity region. Two compute-and-forward schemes are then proposed for the Gaussian model and the capacity region is approximately characterized to within a constant gap.

The design of coding schemes for the compute-and-forward paradigm with low decoding complexity is then considered. Based on the separation-based framework proposed previously by Tunali *et al.*, this study proposes a family of constellations that are suitable for the compute-and-forward paradigm. Moreover, by using Chinese remainder theorem, it is shown that the proposed constellations are isomorphic to product fields and therefore can be put into a multilevel coding framework. This study then proposes multilevel coding for the proposed constellations and uses multistage decoding to further reduce decoding complexity.

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Professor Krishna Narayanan, who has been a mentor, teacher, colleague, and friend to me. I have always appreciated his intuition in research, his open-minded spirit, and his willingness to always go an extra mile to provide the help I need. I have benefitted tremendously from his continuous advice. Without his guidance and persistent help, this dissertation would not have been possible.

Lots of thanks to Professor Tie Liu, who has been like an older brother to me. His enthusiasm to learn new things, his diligence in research, and his patience while teaching me many things have made him a role model to me. I would also like to thank the other members of my dissertation committee. Professor Srinivas Shakkottai's Game Theory lectures were always enjoyable and interesting, and I gained a lot of information outside my immediate area of research. Professor Arun Srinivasa was also helpful, and I appreciated in particular that he taught me how to give a good engineering talk.

I would like to thank my officemates, Engin Tunali and Brett Hern. We joined the research group together, took classes together, learned new things together, attended conferences together, spent countless afternoons discussing our research together, and even published papers together. Without them, my graduate study would not have been so fruitful and enjoyable. I also want to thank my friends and colleagues in the TCSP group: Yung-Yih Jian for letting me share all the happiness and frustration along the journey, and Amir Salimi for introducing me to generalized cut-set bounds and for so many fruitful discussions over the caching problem. I would also like to thank Nariman Rahimian, Armin Banaei, Wei-Yu Chen, Jae Won Yoo, Jinjing Jiang,

Shou Li, Avinash Vem, Santhosh Kumar, and Fatemeh Hamidi Sepehr.

Special thanks to my family for their endless love and support. My parents have been there for me since day one. My father has been extremely hardworking for his entire life, which motivated me to pursue excellence. My mother never gave up on me during my rebellious teenage years. I especially want to thank my lovely wife, Hui-Chin, who never stopped loving and believing in me and gave birth to our precious daughter, Lauren. Without her love, all my achievement would be meaningless.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

After six decades of endeavor by many researchers, we can now safely say that we have a fair understanding of how to design point-to-point communication systems. In addition to computing the capacity of many channels, we also know how to build capacity-achieving codes to arbitrarily approach the capacity with a decoding complexity that is linear in block length. Since 1980s, a great deal of effort has been exerted on the design of wireless communication systems. There are two important characteristics that make wireless communications different from point-to-point communications, namely the *superposition* nature and the *broadcast* nature. The superposition nature basically captures the characteristic that in wireless communication, multiple nodes may send signals to a destination simultaneously, which results in multiple access channels (MAC). On the other hand, the broadcast nature captures the characteristic that if one sender transmits, all the nodes nearby would hear, which can be modeled as broadcast channels (BC). For MAC channels and two-user BC channels, the capacity regions have also been characterized in 1980s in [1] [2] for MAC channels and in [3] [4] [5] for BC channels.

Now, if we keep adding terminals on either the transmitter and/or the receiver ends, the problem becomes more and more complicated. Even for the two-user interference channel, which is the smallest example with multiple transmitters and receivers, the capacity region remains unknown and an approximate characterization of the capacity region has become available only very recently in [6]. However, modern wireless communication is all about networks in which there are multiple terminals. Unfortunately, our understanding of networks with multiple terminals such as multiple transmitters, multiple receivers, and/or multiple intermediate nodes

is *fairly limited*. The importance of wireless networks and the lack of a thorough understanding of such networks have urged researchers to explore the frontiers of network (or multi-terminal) information theory.

Although most problems in network information theory are fairly difficult to solve due to the distributed nature and the lack of structure in networks, there are two new exciting ideas that are driving progress in this area. The first is the deterministic model of Avestimehr, Diggavi, and Tse [7] which allows one to characterize the capacity region approximately by considering an appropriate finite-field model of the network. The second is a new line of research which tries to exploit interference in wireless networks either through aligning the interference or by exploiting it through structured coding at the physical layer [8] [9] [10]. In particular, the compute-and-forward paradigm [11] allows us to *harness* the interference instead of fighting it. This paradigm attempts to incorporate the idea of network coding within the physical layer and is hence called physical layer network coding as well. This strategy has been shown to be capable of providing substantial gains over conventional strategies in many networks. Recently the use of compute-and-forward paradigm and the construction of practical compute-and-forward schemes have become popular research areas.

Motivated by the success of the compute-and-forward strategy, the first part of this dissertation is devoted to study this strategy in more detail from an information-theoretic point of view and to extend this strategy to more general scenarios. Specifically, it has been shown in [12] and [13] that for bidirectional relaying problem with Gaussian channels, the compute-and-forward strategy is asymptotically optimal. However, in practice, channels are hardly memoryless. This urges us to extend the compute-and-forward paradigm to channels with memory such as inter-symbol interference (ISI). Moreover, in modern communication systems, devices are usually

equipped with multiple antennas for which the resulted channels are multiple-input multiple-output (MIMO) channels. This motivates the study of the MIMO bidirectional relaying problem in this dissertation.

One important step following the establishment of the information-theoretic results is to establish coding-theoretic results. In most of literatures on compute-and-forward, e.g., [12] [13] [11], (including the first part of this dissertation), infinite-dimensional lattices are used for the purpose of both shaping and channel coding; hence, the shaping and coding gains are inseparable. Although it has been shown that lattices that are simultaneously good for shaping and good for channel coding exist, the complexity of optimal algorithms for shaping and decoding is formidable. In one of our previous work [14], a separation-based framework has been proposed where channel coding and data modulation are separately designed. The constellation for data modulation has to possess some properties in order to benefit from compute-and-forward. In contrast to practically implementable lattice coding schemes that are available in the existing literature such as low density lattice codes [15] and signal codes [16], this framework enables one to improve the coding gain and shaping gain separately, thus resulting in increased computation rates. This separation also allows us to keep the constellation size q small so that optimal demodulation is feasible. Moreover, the separation allows the use of codes on graph as channel coding so that near optimal performance can be obtained with low-complexity message passing decoder. The second part of the dissertation focuses on designing practically implementable compute-and-forward schemes that have provable threshold with decoding complexity scaled linearly with code length. Based on the separation-based framework in [14], a low complexity compute-and-forward scheme is proposed which is particularly suitable for applications that operate in the moderate and high rate regimes.

1.1 Organization

The rest of the dissertation is organized as follows. In Chapter 2, we briefly introduce some background for lattices, nested lattice codes, and lattice partition chains which will be useful and crucial for the following chapters. We state basic definitions and some important existing results without proofs. After that, we introduce the problem of bidirectional relaying with Gaussian channels and explain the lattice-based compute-and-forward strategy over this problem. The reason that we particularly focus on Gaussian bidirectional relaying problem is that the idea of compute-and-forward can be well illustrated by and was first proposed for bidirectional relaying problem with Gaussian channels in [12] [13]. We restate the coding schemes as well as the main results in [12] [13] without proofs.

In Chapter 3, we extend compute-and-forward paradigm to channels with memory such as ISI. We provide inner and outer bounds on the capacity region for the Gaussian bidirectional relaying over ISI channels. The outer bound is obtained by the conventional cut-set argument. For the inner bound, we propose a compute-and-forward coding scheme based on lattice partition chains and study its achievable rate. The coding scheme is a time-domain coding scheme which uses a novel precoding scheme at the transmitter in combination with lattice precoding and a minimum mean squared error receiver to recover linear combinations of lattice codewords. The proposed compute-and-forward coding scheme substantially outperforms decode-and-forward schemes. We also compare the proposed time-domain scheme to the frequency-domain scheme that we proposed previously in [17]. The frequency domain scheme [17] is similar in spirit to coded orthogonal frequency division multiplexing (OFDM) with independent coding across sub-carriers and uses nested-lattice code with a power allocation strategy that can exploit the group property of lat-

tices. While it is well known that for the point-to-point communication case, both independent coding along sub-channels and time-domain coding can approach the capacity limit, we show that for the bidirectional relaying, the performance of the two coding schemes are different. Particularly, we show that independent coding across sub-channels is not optimal and coding across sub-channels can improve the capacity region for some channel realizations.

In Chapter 4, we consider the bidirectional relaying with parallel Gaussian channels. This is a relevant model for bidirectional relaying with ISI and MIMO bidirectional relaying as the problems can be converted into a set of parallel bidirectional relay channels by discrete Fourier transform and singular value decomposition, respectively. We first take a look at the corresponding linear deterministic model [7] and solve the corresponding network coding problem. Based on the insight obtained from the reasoning in the corresponding linear deterministic model, we propose two coding schemes that would substantially outperform the conventional strategies. We then show that both the proposed schemes achieve the capacity region to within a constant gap and one of the coding schemes is asymptotically optimal.

In Chapter 5, the problem of designing practically implementable compute-and-forward schemes is investigated. We first review the separation-based framework proposed in [14], which provides a design guideline for the design of compute-and-forward schemes. Motivated by the work of Feng *et al.* [18], we propose a family of constellations that satisfy the design guideline and thereby are suitable for compute-and-forward paradigm. This family of constellations is generated from quotient rings of Eisenstein integers and hence provides good shaping gains. Moreover, by Chinese Remainder Theorem, we show that the proposed family of constellations is isomorphic to the corresponding product fields. This allows us to incorporate the idea of multilevel coding so that for the proposed constellation with q^2 elements, the chan-

nel coding only has to work over $\mathbb{F}_q$, thereby resulting in a substantial complexity reduction.

In chapter 6, we summarize our prior work on the joint source-channel coding problem of transmitting a discrete-time analog source over an additive white Gaussian noise (AWGN) channel with interference known at transmitter. Particularly, we consider the case when the source and the interference are correlated. Although not directly relevant to the main theme of this dissertation, coding with side information is an important topic and a canonical problem in wireless networks and we feel appropriate to describe this work here. We first derive an outer bound on the achievable distortion and then, we propose two joint source-channel coding schemes to make use of the correlation between the source and the interference. The first scheme is the superposition of the uncoded signal and a digital part which is the concatenation of a Wyner-Ziv encoder and a dirty paper encoder. In the second scheme, the digital part is replaced by a hybrid digital and analog (HDA) scheme so that the proposed scheme can provide graceful degradation in the presence of signal-to-noise ratio (SNR) mismatch. Interestingly, unlike the independent interference setup, we show that neither of the two schemes outperform the other universally in the presence of SNR mismatch.

1.2 Notation

Throughout the dissertation, $\mathbb{C}$ represents the set of complex numbers, $\mathbb{E}$ denotes the expectation operation, and $\mathbb{P}(E)$ denotes the probability of the event E. Vectors are written in boldface. Also, for a given real number x, we use the convention $(x)^+ := \max\{0, x\}$ and $\log(x)^+ := \max\{0, \log(x)\}$ where all logarithms are to the base 2. $\times$ denotes the Cartesian product and $\oplus$ denotes the addition operation over some finite fields where the field size can be understood from the context if it is not

specified. Also, we do not distinguish the multiplication operation over the complex field and finite fields; which field the operation is taking over is understood from the context.

2. BACKGROUND

2.1 Lattices, Nested Lattice Codes, and Lattice Partition Chains

Since lattices and lattice codes have been well discussed in many recent works, we only summarize some required definitions and results without proofs. For more details about the lattices, nested lattice codes, and lattice partition chains, the reader is referred to [19, 20, 21, 13] and the reference therein.

2.1.1 Lattices

An n-dimensional lattice Λ^n is a discrete subgroup of $\mathbb{R}^n$ under ordinary vector addition operation. This implies that for all pairs $\lambda_1, \lambda_2 \in \Lambda^n$, we have $\lambda_1 + \lambda_2 \in \Lambda^n$. In Fig. 2.1 and 2.2, we provide two examples of two-dimensional lattices, namely Rectangular lattice (Z_2) and Hexagonal lattice (A_2), that corresponds to the ring of Gaussian integers and the ring of Eisenstein integers, respectively.

In the following, we define some important operations in lattices.

Definition 1 (Lattice Quantizer)*:* For a $\mathbf{x} \in \mathbb{R}^n$, the nearest neighbor quantizer associated with Λ is denoted as

$$Q(\mathbf{x}) = \lambda \in \Lambda^n;\ \|\mathbf{x} - \lambda\| \leq \|\mathbf{x} - \lambda'\|\ \forall \lambda' \in \Lambda^n, \tag{2.1}$$

where $\|.\|$ represents the L_2-norm operation.

Definition 2 (Fundamental Voronoi Region)*:* The fundamental Voronoi region $\mathcal{V}(\Lambda^n)$ is defined as

$$\mathcal{V}(\Lambda^n) = \{\mathbf{x} : Q_{\Lambda^n}(\mathbf{x}) = \mathbf{0}\}. \tag{2.2}$$

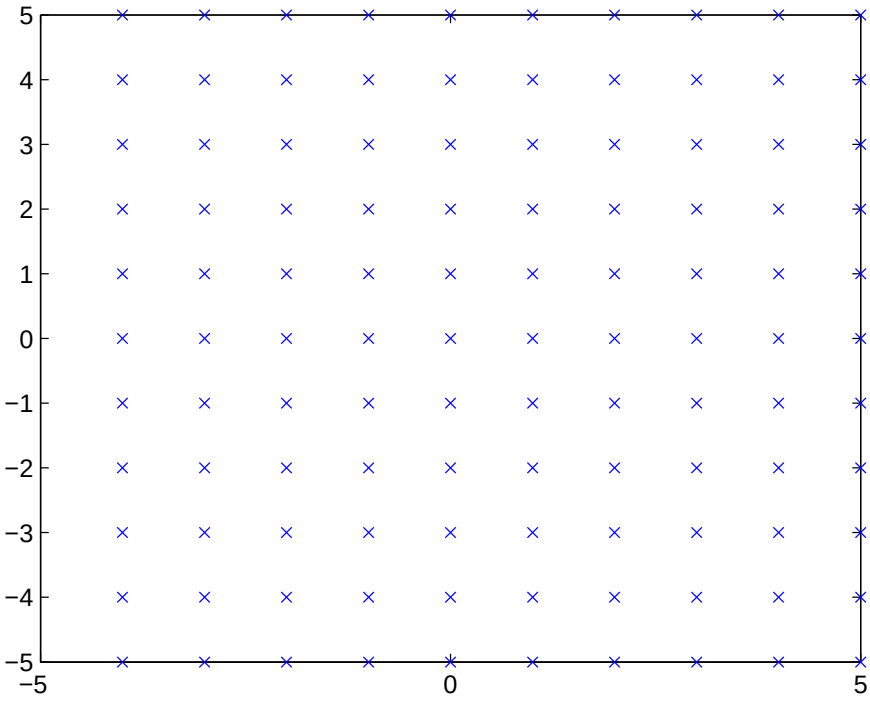

Figure 2.1: Rectangular (Z_2) lattice, Gaussian integers

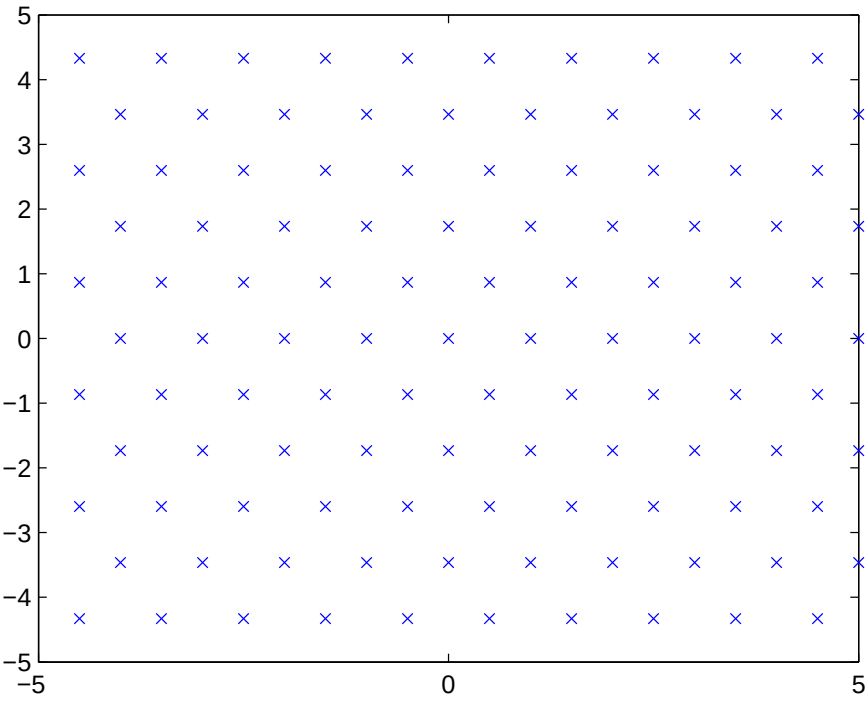

Figure 2.2: Hexagonal (A_2) lattice, Eisenstein integers

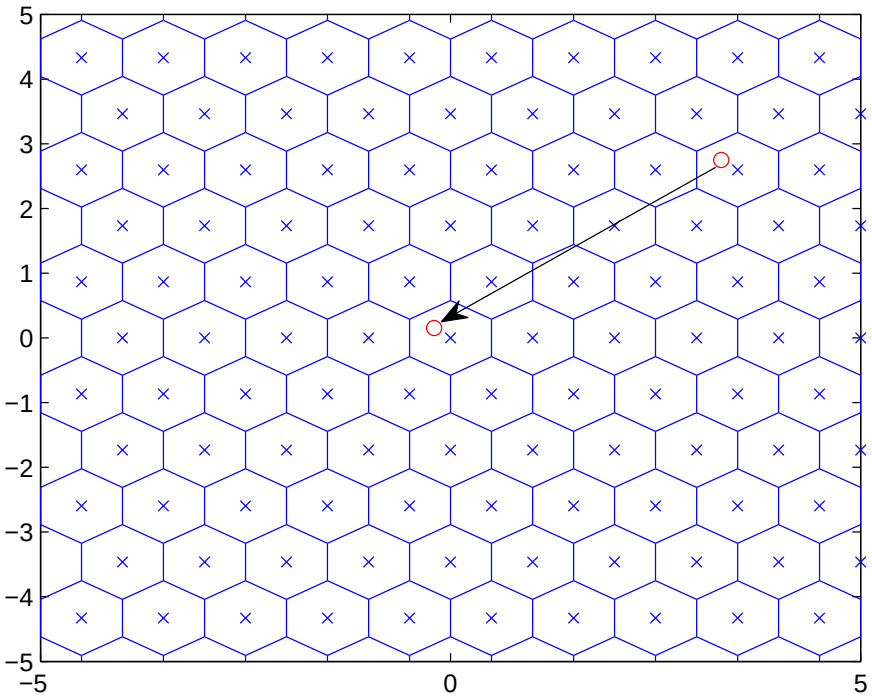

Figure 2.3: Example showing some lattice operations

Definition 3 (Modulo Operation)*:* The $\mod$ operation is represented as

$$\mathbf{x} \mod \Lambda^n = \mathbf{x} - Q_{\Lambda^n}(\mathbf{x}). \tag{2.3}$$

One example that explains the above operations can be found in Fig. 2.3 where A_2 lattice is considered and the circle in the upper right corner represents a vector $\mathbf{x} \in \mathbb{R}^2$. The nearest neighbor quantizer associated with A_2 will quantize $\mathbf{x}$ to the lattice point inside the same hexagon with $\mathbf{x}$. The hexagon circumventing the origin is the fundamental Voronoi region of A_2. Moreover, the modulo operation will map $\mathbf{x}$ to the corresponding position inside the fundamental Voronoi region as the circle shown in the middle of this figure.

In what follows, we define goodness of lattices from the perspective of packing, covering, quantization, and channel coding problems, respectively. Let us first denote by $\mathcal{B}$ the n-dimensional unit sphere centered at the origin and $r\mathcal{B}$ the n-dimensional sphere of radius r centered at the origin.

Packing Problem: For a lattice Λ^n, a radius $r > 0$ is said to be a packing radius if $\Lambda + r\mathcal{B}$ is a packing in Euclidean space. i.e., for $\mathbf{x} \neq \mathbf{y} \in \Lambda^n$, we have

$$(\mathbf{x} + r\mathcal{B}) \cap (\mathbf{y} + r\mathcal{B}) = \emptyset. \tag{2.4}$$

Definition 4 (Packing Radius)*:* Define the packing radius $r_{\Lambda^n}^{\text{pack}}$ of Λ^n by

$$r_{\Lambda^n}^{\text{pack}} = \sup\{r : \Lambda^n + r\mathcal{B} \text{ is a packing}\}. \tag{2.5}$$

Definition 5 (Effective Radius)*:* Define the effective radius $r_{\Lambda^n}^{\text{effec}}$ of Λ^n as a radius such that the volume of $r_{\Lambda^n}^{\text{effec}}\mathcal{B}$ is equal to the volume of the fundamental Voronoi region.

Definition 6 (Packing Efficiency)*:* Define the packing efficiency of a lattice Λ^n by

$$\rho_{\text{pack}}(\Lambda^n) = \frac{r_{\Lambda^n}^{\text{pack}}}{r_{\Lambda^n}^{\text{effec}}}. \tag{2.6}$$

Also, for all possible n-dimensional lattice Λ^n, we define the optimal asymptotic packing efficiency by

$$\rho_{\text{pack}}^* = \limsup_{n\to\infty} \sup_{\Lambda^n} \frac{r_{\Lambda^n}^{\text{pack}}}{r_{\Lambda^n}^{\text{effec}}}. \tag{2.7}$$

Definition 7 (Goodness for Packing)*:* We say that a sequence of lattices is asymptotically good for packing if it achieves the Minkowski lower bound defined as $\rho_{\text{pack}}^* = 1/2$.

Covering Problem: For a lattice Λ^n, $r > 0$ is said to be a covering radius if $\mathbb{R}^n \subseteq \Lambda^n + r\mathcal{B}$.

Definition 8 (Covering Radius)*:* Define the covering radius $r_{\Lambda^n}^{\text{cov}}$ of Λ^n by

$$r_{\Lambda^n}^{\text{cov}} = \min\{r : \Lambda^n + r\mathcal{B} \text{ is a covering}\}. \tag{2.8}$$

Definition 9 (Covering Efficiency)*:* Define the covering efficiency of a lattice Λ^n by

$$\rho_{\text{cov}}(\Lambda^n) = \frac{r_{\Lambda^n}^{\text{cov}}}{r_{\Lambda^n}^{\text{effec}}}. \tag{2.9}$$

Also, for all possible n-dimensional lattice Λ^n, we define the optimal asymptotic covering efficiency by

$$\rho_{\text{cov}}^* = \liminf_{n\to\infty} \inf_{\Lambda^n} \frac{r_{\Lambda^n}^{\text{cov}}}{r_{\Lambda^n}^{\text{effec}}}. \tag{2.10}$$

Definition 10 (Goodness for Covering)*:* We say that a sequence of lattices is asymptotically good for covering or Rogers-good if $\rho_{\text{cov}}^* = 1$.

MSE Quantization: In this problem, we consider the lattice quantizer for a lattice Λ^ndefined above. The second moment of a lattice is defined as the average energy per dimension of a uniform probability distribution over $\mathcal{V}(\Lambda^n)$ as

$$\sigma^2(\Lambda^n) = \frac{1}{V(\Lambda^n)} \frac{1}{n} \int_{\mathcal{V}(\Lambda^n)} \|\mathbf{x}\|^2 d\mathbf{x}, \tag{2.11}$$

where $V(\Lambda^n)$ is the volume of the $\mathcal{V}(\Lambda^n)$. The normalized second moment of the lattice is then defined as

$$G(\Lambda^n) = \frac{\sigma^2(\Lambda^n)}{V(\Lambda^n)^{2/n}}. \tag{2.12}$$

The minimum possible value $G(\Lambda^n)$ over all n-dimensional lattices is defined by G_n. Also, the normalized second moment of a sphere approaches $\frac{1}{2\pi e}$ as $n \to \infty$. The

isoperimetric inequality implies that $G_n > \frac{1}{2\pi e}$.

Definition 11 (Goodness for MSE Quantization)*:* We say that a sequence of lattices is asymptotically good for MSE quantization if

$$\lim_{n\to\infty} G_n = \frac{1}{2\pi e}. \tag{2.13}$$

Unconstraint AWGN Channel Coding Problem: Consider the AWGN channel $Y = X + N$ where X, Y, and $N \sim \mathcal{N}(0, P_N)$ represent the transmitted signal, the received signal, and the noise, respectively. Denote by $\mathbf{N}$ an i.i.d. vector of length n of noise random variables. We define the "effective radius" of the noise vector by $r_N = \sqrt{nP_N}$. Now, consider the case that we use an infinite lattice Λ^n as signal constellation, i.e., we don't have any power constraint.

Definition 12 (Voronoi-to-Noise Effective Radius Ratio)*:* Define the Voronoi-to-noise effective radius ratio of a lattice Λ^n by

$$\rho_{\text{AWGN}}(\Lambda^n, r_N) = \frac{r^{\text{effec}}_{\Lambda^n}}{r_N}. \tag{2.14}$$

Definition 13 (Goodness for AWGN Channel Coding)*:* We say that a sequence of lattices is asymptotically good for AWGN channel coding or Poltyrev-good if whenever $\rho_{\text{AWGN}}(\Lambda^n, r_N) > 1$, the error probability of decoding X from Y

$$P_e(\Lambda^n, r_N) = \mathbb{P}\{\mathbf{N} \notin \mathcal{V}(\Lambda^n)\} \tag{2.15}$$

tends to zero exponentially fast as n grows. i.e., achieves the Poltyrev exponent defined in [22].

We conclude this subsection by restating a very important theorem in [20] which

was proved by using the ensemble of lattices generated by Construction A [23].

Theorem 1 (Theorem 5 of [20])**.** *For asymptotically high dimension, there exist lattices that are simultaneously good for packing, covering, MSE quantization, and AWGN channel coding problems.*

2.1.2 Nested Lattice Codes

An n-dimensional nested lattice code consists of a fine lattice Λ_f^n and a coarse lattice Λ_c^n. The coarse lattice Λ_c^n is said to be nested in the fine lattice Λ_f^n if all the elements in Λ_c^n are also elements in Λ_f^n as shown in Fig. 2.4. i.e. $\Lambda_c^n \subseteq \Lambda_f^n$. A nested lattice code uses elements in the quotient group Λ_f^n/Λ_c^n as codewords. i.e., it takes all the fine lattice points in $\Lambda_f^n \cap \mathcal{V}(\Lambda_c^n)$ as codewords. The rate of a nested lattice code is given by

$$\frac{1}{n}\log|\Lambda_f^n \cap \mathcal{V}(\Lambda_c^n)| = \frac{1}{n}\log\frac{V(\Lambda_c^n)}{V(\Lambda_f^n)}. \tag{2.16}$$

To use a nested lattice code as transmission scheme, the coarse lattice has to be carefully chosen as it governs the shaping (closely related to the power constraint) as well as the rate of this nested lattice code.

In [19], Erez and Zamir consider the AWGN channel with the power constraint P and the noise variance N. They show the existence of an ensemble of good nested lattice codes and then use this ensemble to show that the AWGN capacity can be achieved by lattice encoding together with Euclidean lattice decoding. In contrast to maximum-likelihood decoding that finds the most likely lattice point inside the shaping region, Euclidean lattice decoding directly find the closet lattice point to the received signal, ignoring the boundary of the code. Lattice decoding has a much lower complexity than maximum likelihood decoding and has hence attracted a lot of interest. We restate these two Theorems in the following.

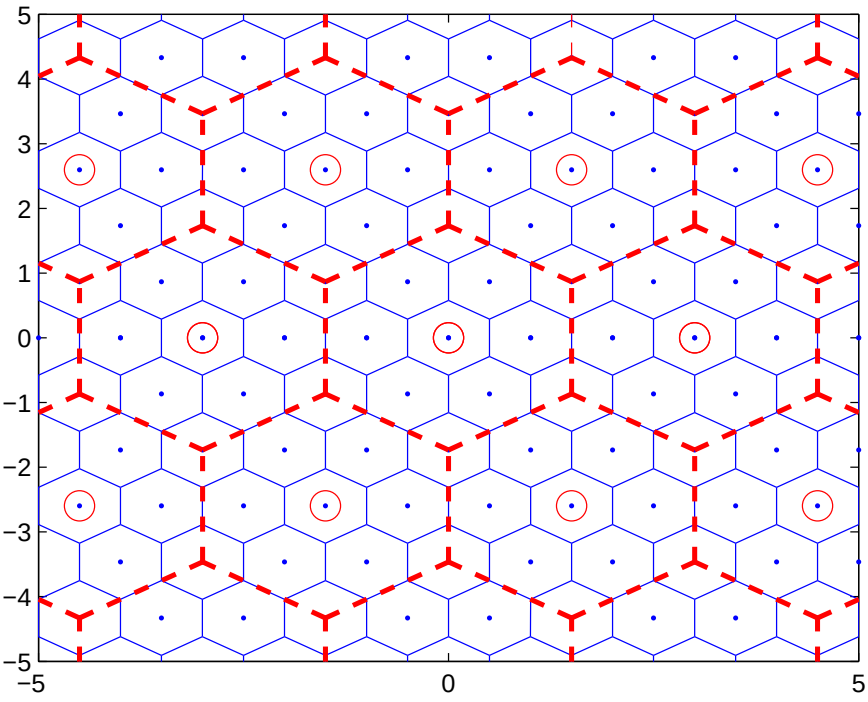

Figure 2.4: An example of a nested lattice code. •: points in Λ_f^n. ∘: points in Λ_c^n.

Theorem 2 (Section VII of [19]). *There exist a sequence of nested lattice codes in which the coarse lattices are simultaneously Rogers-good and Poltyrev-good and the fine lattices are Poltyrev good.*

Theorem 3 (Theorem 5 of [19]). *For any rate $R < C = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$, there exists a sequence of n-dimensional nested lattice codes Λ_f^n/Λ_c^n whose coding rate approaches R, and whose decoding error probability under Euclidean lattice decoding decays exponentially fast as $n \to \infty$.*

2.1.3 Lattice Partition Chains

An n-dimensional lattice partition chain $\Lambda_f^n/\Lambda_2^n/\Lambda_1^n$ consists three lattices Λ_f^n, Λ_1^n, and Λ_2^n, where $\Lambda_1^n \subseteq \Lambda_2^n \subseteq \Lambda_f^n$ as shown in Fig. 2.5. We restate an important theorem in [13].

Theorem 4 (Theorem 2 of [13]). *For any $P_1 \geq P_2 \geq 0$, a sequence of n-dimensional lattice partition chains $\Lambda_f^n/\Lambda_2^n/\Lambda_1^n$ exists that satisfies the following properties.*

1. *Λ_1^n and Λ_2^n are simultaneously Rogers-good and Poltyrev-good while Λ_c is Poltyrev-good.*

2. *For any $\epsilon > 0$, $P_i - \epsilon \leq \sigma^2(\Lambda_i^n) \leq P_i$, $i \in \{1,2\}$, for sufficiently large n.*

3. *The coding rate of the nested lattice code associated with the lattice partition Λ_f^n/Λ_2^n can approach any value $\gamma \geq 0$ as n tends to infinity, i.e.,*

$$R_2 = \frac{1}{n}\log\left(\frac{V(\Lambda_2^n)}{V(\Lambda_f^n)}\right) = \gamma + o_n(1), \tag{2.17}$$

where $o_n(1) \to 0$ as $n \to \infty$. Furthermore, the coding rate of the nested lattice

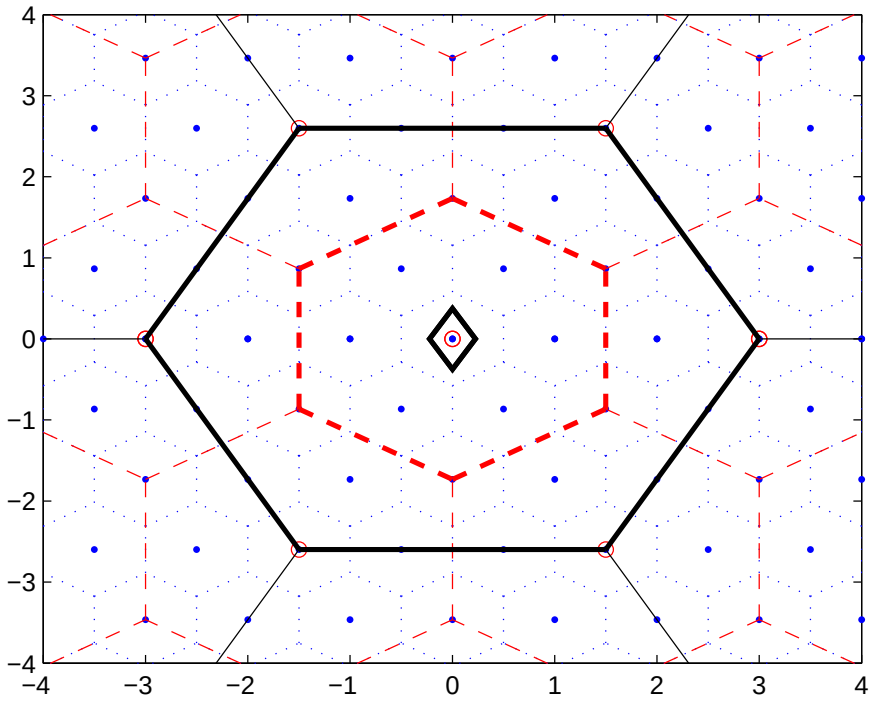

Figure 2.5: An example of a lattice partition chain. •: points in Λ_f^n. ∘: points in Λ_2^n. ◇: points in Λ_1^n

code associated with Λ_f^n/Λ_1^n *is given by*

$$R_1 = \frac{1}{n}\log\left(\frac{V(\Lambda_1^n)}{V(\Lambda_f^n)}\right) = R_2 + \frac{1}{2}\log\left(\frac{P_1}{P_2}\right) + o_n(1). \tag{2.18}$$

It is worth mentioning that lattice partition chains are not restricted to three lattices. A more general definition where one can have arbitrary number of coarse lattices has been introduced in [24, Theorem 2].

2.2 Compute-and-Forward for Memoryless Bidirectional Relay channel

The main focus of this section is to explain the problem of memoryless bidirectional relay channel and how to use lattice codes for compute-and-forward as an information forwarding strategy. In a bidirectional relay channel shown in Fig. 2.6, two nodes A and B wish to exchange information with the help of an intermediate relay node between them. Also, we assume that the two nodes are remote from each other so that we can assume that there is no direct link between the two nodes. In the first phase, the MAC phase, both nodes send signals $\mathbf{x}_A$ and $\mathbf{x}_B$ with power constraints P_A and P_B, respectively, to the relay and the relay keeps silent. The received signal at the relay is given by

$$\mathbf{y}_R = h_A\mathbf{x}_A + h_B\mathbf{x}_B + \mathbf{z_R}, \tag{2.19}$$

where $\mathbf{z_R}$ is an i.i.d. noise random variables drawn from $\mathcal{N}(0, N_R)$. The relay then forms a signal $\mathbf{x}_R$ based on the observation $\mathbf{y}_R$. In the second pase, the BC phase, the relay broadcasts $\mathbf{x}_R$ with a power constrain P_R back to the nodes and the nodes keep silent. The received signals at the nodes are given by

$$\mathbf{y}_A = g_A\mathbf{x}_R + \mathbf{z_A}, \tag{2.20}$$

$$\mathbf{y}_B = g_B\mathbf{x}_R + \mathbf{z_B}, \tag{2.21}$$

respectively, where $\mathbf{z_A}$ and $\mathbf{z_B}$ are i.i.d. noise random variables drawn from $\mathcal{N}(0, N_A)$ and $\mathcal{N}(0, N_B)$, respectively. After this, the nodes try to decode the other's message according the received signals and their own messages as side information. It is worth mentioning that since h_A and h_B can be absorbed into the power constraints and g_A and g_B can be absorbed into the noise variances, one can equivalently consider the

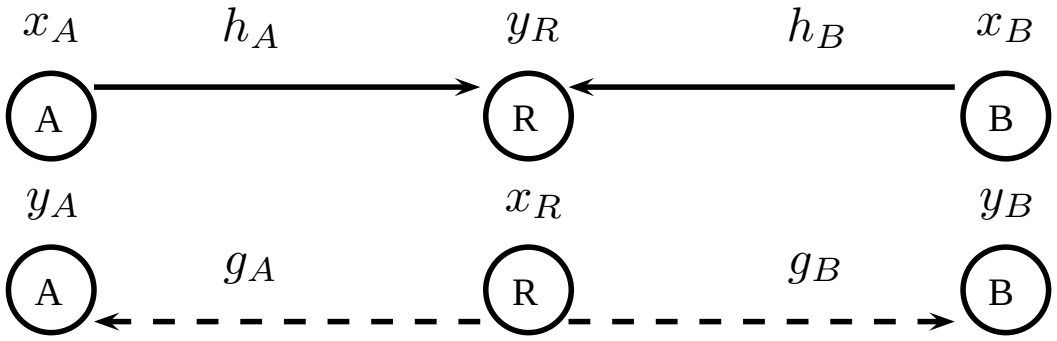

Figure 2.6: Bidirectional relay network.

case that all channel coefficients are 1, i.e., $h_A = h_B = g_A = g_B = 1$. Throughout the chapter, we consider this equivalent model.

For this problem, conventional strategies such as AF and DF either suffer from noise propagation or a loss in the multiplexing gain and hence are suboptimal. In the following subsections, we review the lattice-based compute-and-forward strategy which has been shown to be asymptotically optimal for memoryless bidirectional relay channel. The main idea of the compute-and-forward is based on the observation that the relay is not interested in the individual messages. The approach of compute-and-forward is to choose the codes used by the nodes very specifically such that it enables functional decoding or computing at the relay.

2.2.1 Symmetric Case

In this subsection, we consider the symmetric case where $P_A = P_B = P$ and $N_R = N_A = N_B = N$. For this case, Wilson *et al.* [12] exploit the linear structure of nested lattice codes and propose the following coding scheme. The nodes A and B adopt an identical sequence of good nested lattice codes (in the sense of Theorem 2) Λ_f^n/Λ_c^n where the coarse lattice are chosen to have a second moment (arbitrarily close to) P. They first map the messages to lattice points $\mathbf{t}_A, \mathbf{t}_B \in \Lambda_f^n/\Lambda_c^n$, respectively.

The dithered version

$$\mathbf{x}_A = [\mathbf{t}_A - \mathbf{d}_A] \mod \Lambda_c, \tag{2.22}$$

$$\mathbf{x}_B = [\mathbf{t}_B - \mathbf{d}_B] \mod \Lambda_c, \tag{2.23}$$

are then transmitted by the nodes A and B, respectively.

At the relay, due to the property that lattices are closed under integer combinations, the relay can decode the received signal to $\mathbf{t}_R = (\mathbf{t}_A + \mathbf{t}_B) \mod \Lambda_c^n$. The procedure is listed below.

1. Form MMSE estimate of $\mathbf{x}_A + \mathbf{x}_B$ from $\mathbf{y}_R$ as $\alpha \mathbf{y}_R$.
2. Subtract random dithers as $\alpha \mathbf{y}_R + \mathbf{d}_A + \mathbf{d}_B$.
3. Bring this estimate back to the fundamental Voronoi region of the coarse lattice by the modulo operation.
4. Quantize the result to the nearest fine lattice point.

The relay then re-encodes this estimate and broadcasts the codeword corresponding to $\mathbf{t}_R$.

Now suppose $\mathbf{t}_R$ can be reliably decoded at the two nodes, nodes A and B recover the other's message by the following operation

$$\mathbf{t}_B = [\mathbf{t}_R - \mathbf{t}_A] \mod \Lambda_c, \tag{2.24}$$

$$\mathbf{t}_A = [\mathbf{t}_R - \mathbf{t}_B] \mod \Lambda_c. \tag{2.25}$$

The performance of this coding scheme is described in the following Theorem which is the main result of [12].

Theorem 5 (Theorem 1 of [12])**.** *For Gaussian bidirectional relay channel, there exists at least a sequence of n-dimensional nested lattice codes Λ_f^n/Λ_c^n such that any exchange rate of $R_{ex} = R_A = R_B \leq \frac{1}{2}\log\left(\frac{1}{2}+\frac{P}{N}\right)^+$ is achievable using Euclidean lattice decoding.*

2.2.2 Asymmetric Case

In [13], Nam *et al.* consider the bidirectional relay channel with asymmetric power P_A and P_B and asymmetric noise variance N_A and N_B. They proposed a coding scheme as follows. Without loss of generality, one can assume that $P_A \geq P_B$. Otherwise, one can simply switch the role of node A and B. Consider a sequence of good lattice partition chains (in the sense of Theorem 4) $\Lambda_f^n/\Lambda_B^n/\Lambda_A^n$. Node A (the node having a larger power) encodes its message to $\mathbf{t}_A$ by the nested lattice code Λ_f^n/Λ_A^n and the weak node (the one having a weaker power) encodes its message to $\mathbf{t}_B$ by the nested lattice code Λ_f^n/Λ_B^n. Both nodes then transmit the corresponding dithered version of the lattice codewords as

$$\mathbf{x}_A = [\mathbf{t}_A - \mathbf{d}_A] \mod \Lambda_A, \tag{2.26}$$

$$\mathbf{x}_B = [\mathbf{t}_B - \mathbf{d}_B] \mod \Lambda_B, \tag{2.27}$$

respectively.

At the relay, due to the property that lattices are closed under integer combinations, the relay can decode the received signal to $\mathbf{t}_R = (\mathbf{t}_A + \mathbf{t}_B - Q_B(\mathbf{t}_B + \mathbf{u}_B)) \mod \Lambda_A^n$. The procedure is listed below.

1. Form MMSE estimate of $\mathbf{x}_A + \mathbf{x}_B$ from $\mathbf{y}_R$ as $\alpha\mathbf{y}_R$.

2. Subtract random dithers as $\alpha\mathbf{y}_R + \mathbf{d}_A + \mathbf{d}_B$.

3. Bring this estimate back to the fundamental Voronoi region of the coarse lattice Λ_A^n by the modulo operation.

4. Quantize the result to the nearest fine lattice point.

Note that here we have an extra term $-Q_B(\mathbf{t}_B + \mathbf{u}_B)$. This is a result of that now we adopt nested lattice codes with different coarse lattices and the relay performs the modulo operation with respect to Λ_A^n. The relay then re-encodes this estimate and broadcasts the codeword corresponding to $\mathbf{t}_R$.

Now suppose $\mathbf{t}_R$ can be reliably decoded at the two nodes, nodes A and B recover the other's message by the following operation

$$\mathbf{t}_B = [\mathbf{t}_R - \mathbf{t}_A] \mod \Lambda_B, \tag{2.28}$$

$$\mathbf{t}_A = [\mathbf{t}_R - \mathbf{t}_B + Q_B(\mathbf{t}_B + \mathbf{u}_B)] \mod \Lambda_A, \tag{2.29}$$

respectively. The performance of this coding scheme is described in the following Theorem which is the main result of [13].

Theorem 6 (Theorem 1 of [13])**.** *For Gaussian bidirectional relay channel, there exists a sequence of n-dimensional lattice partition chains $\Lambda_f^n/\Lambda_B^n/\Lambda_A^n$ such that any rate pair (R_A, R_B) satisfying*

$$R_A \leq \min\left\{\frac{1}{2}\log\left(\frac{P_A}{P_A+P_B}+\frac{P_A}{N_R}\right)^+, \frac{1}{2}\log\left(1+\frac{P}{N_B}\right)\right\}, \tag{2.30}$$

$$R_B \leq \min\left\{\frac{1}{2}\log\left(\frac{P_B}{P_A+P_B}+\frac{P_B}{N_R}\right)^+, \frac{1}{2}\log\left(1+\frac{P}{N_A}\right)\right\}. \tag{2.31}$$

is achievable using Euclidean lattice decoding.

We conclude this section by the following two remarks.

Remark: It has been shown in [13] that for Gaussian bidirectional relay channel, the above lattice-based compute-and-forward schemes is asymptotically optimal in the high SNR regime and achieves the capacity region to with 1/2 bit.

Remark: All the results and schemes in this section are over the real field. However, these can be easily extended to the complex field by regarding the complex filed as two orthogonal real fields and use the schemes above in each real field.

3. BIDIRECTIONAL RELAYING OVER ISI CHANNELS*

In this chapter, we provide inner and outer bounds on the capacity region for the Gaussian bidirectional relaying over inter-symbol interference channels. The outer bound is obtained by the conventional cut-set argument. For the inner bound, we propose a compute-and-forward coding scheme based on lattice partition chains and study its achievable rate. The coding scheme is a time-domain coding scheme which uses a novel precoding scheme at the transmitter in combination with lattice precoding and a minimum mean squared error receiver to recover linear combinations of lattice codewords. The proposed compute-and-forward coding scheme substantially outperforms decode-and-forward schemes. While it is well known that for the point-to-point communication case, both independent coding along sub-channels and time-domain coding can approach the capacity limit, as a byproduct of the proposed scheme, we show that for the bidirectional relay case, independent coding along sub-channels is not optimal in general and joint coding across sub-channels can improve the capacity for some channel realizations.

3.1 Introduction

The problem of communication over the bidirectional relay channel has attracted a great deal of attention [25] [26] [27] [28] [12] [13]. In a bidirectional relay channel, two nodes A and B wish to exchange information through a relay node R between them as shown in Fig. 3.1. In addition, there is no direct link from node A to node B and vice versa. For this setup with additive white Gaussian noise (AWGN) only, (i.e., channels do not have memory,) it has been shown that classical relaying

*

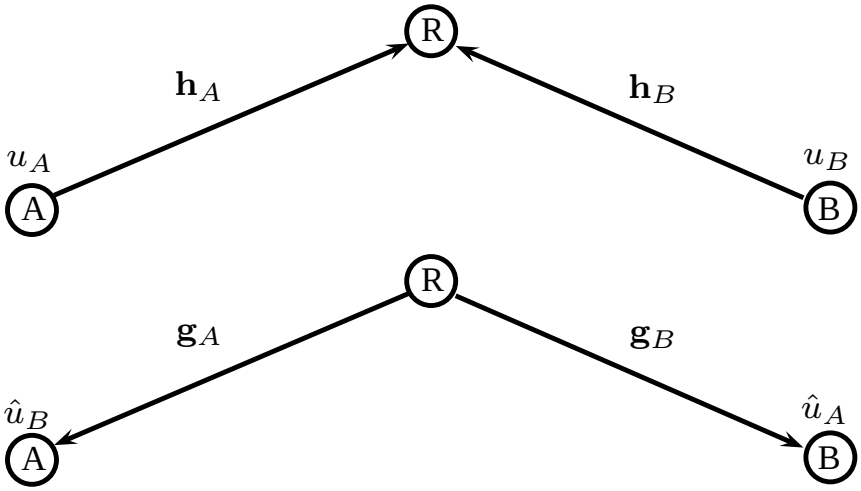

Figure 3.1: Bidirectional relay network with ISI.

strategies such as amplify-and-forward (AF) and decode-and-forward (DF) can be directly extended to this channel. However, the AF strategy which only scales and broadcasts the received signal at the relay is strictly suboptimal in the low signal-to-noise ratio (SNR) regime due to noise propagation. On the other hand, the DF strategy which requires the relay to decode the received signal to the individual messages may significantly reduce the multiplexing gain (pre-log factor) and is thus suboptimal in the high SNR regime. For more details about these classical strategies, the reader is referred to [25] and [26] and references therein.

Recently, a novel information-forwarding strategy called compute-and-forward (CF) (or physical layer network coding) for bidirectional relaying has been proposed and studied in [27] [28] [12] [13]. In the CF strategy, the relay only decodes a function of the transmitted signals instead of individual messages. The function is chosen specifically such that given this function and their own messages, both nodes can correctly recover each other's message. In [11], the CF strategy is further extended to a larger network.

In [12], the channels are assumed to be symmetric and memoryless and the CF strategy was realized using nested lattice codes (NLC). The main idea is to encode the messages by identical NLCs at the two nodes and then to exploit the group property of lattices at the relay. Since the channels are assumed to be symmetric, the two transmitted codewords lie in the same (fine) lattice at the relay so that the relay can directly decode the sum of two codewords modulo a coarse lattice from the received signal. The authors show that this CF scheme can achieve an exchange rate (i.e., symmetric rate) that is asymptotically optimal. For the case when channels are asymmetric, Wilson and Narayanan [29] proposed a scheme that uses a power allocation strategy for enforcing lattice alignment at the relay and then directly carries out the CF scheme in [12]. On the other hand, in [13], Nam *et al.* proposed another CF scheme in which different nesting ratios are used at the two encoders. Their coding scheme is based on lattice partition chains and has been shown to approach the capacity region to within one bit ($\frac{1}{2}$ bit per real dimension).

In this work, we study the Gaussian bidirectional relay channel with inter-symbol interference (ISI), i.e., with memory. We provide inner and outer bounds on the capacity region for this channel. The proposed outer bound is simply obtained by the conventional cut-set argument [30]. For the inner bound, it is unclear how to implement the above CF schemes for the channel with memory. To address this problem, we first propose a novel pre-filtering technique that again ensures that the two transmitted codewords lie in the same (fine) lattice at the relay. Then, we adopt lattice precoding [31] at the two nodes and an unbiased minimum mean squared error (MMSE) equalizer [32] at the relay. Our proposed scheme allows the use of NLC or lattice partition chains schemes [13] for the channels with ISI.

For the point-to-point communication over ISI channels, it has been known for quite a while that the capacity can be achieved by at least two different approaches.

The first approach proposed by Cioffi *et al.* [32] adopts single-carrier transmission with an infinite-length unbiased MMSE decision feedback equalizer (MMSE-DFE) at the receiver. Cioffi *et al.* showed that this unbiased MMSE-DFE scheme is capacity-achieving if the input power-spectral density is optimized. The second approach is by partitioning the whole spectrum (which is frequency-selective) into infinitely many infinitesimal sub-channels that are flat (frequency-nonselective) and then use a good AWGN code in each sub-channel separately. Moreover, the power allocated to each sub-channel can be obtained by the conventional water-filling power allocation strategy for maximizing the overall rate [33]. This leads to the family of multi-carrier transmission systems. This also implies that in point-to-point communication, joint coding across sub-channels is not necessary for achieving the capacity. Similar results for the Gaussian MAC channel and Gaussian BC channel with ISI are reported in [34] and [35], respectively. On the other hand, Cadambe and Jafar in [36] showed that for the parallel Gaussian interference channel, joint processing across sub-channels can potentially achieve a larger rate region than that achieved by separate coding. In this chapter, as a byproduct of our inner bound, we show that similar to the interference channel, joint processing across sub-channels is required in general to achieve the capacity region of the Gaussian bidirectional relay channel with ISI.

3.1.1 Organization

This chapter is organized as follows. In Section 3.2, we explain the problem considered in this chapter in detail and provide an outer bound on the capacity region in Section 3.3. In Section 3.4, we present the proposed scheme and analyze the corresponding achievable rate region. Some discussions and numerical results are given in Section 3.5 where we provide an example showing that joint coding across sub-channels is necessary in general to achieve the capacity region. After this, there

are some conclusions and potential future work in Section 4.5.

3.2 Problem Statement

In this chapter, we consider the bidirectional relay channel in which two nodes A and B wish to exchange information through a relay R between them as shown in Fig. 3.1. Different from other works on this problem, we consider the case that the communication takes place in the presence of inter-symbol interference (ISI). Specifically, the time-invariant vector channels from A to R and that from B to R are denoted as $\mathbf{h}_A \in \mathbb{C}^{L_{AR}}$ and $\mathbf{h}_B \in \mathbb{C}^{L_{BR}}$, respectively. Also, in the downlink, the channel from R to A and B are denoted as $\mathbf{g}_A \in \mathbb{C}^{L_{RA}}$ and $\mathbf{g}_B \in \mathbb{C}^{L_{RB}}$. Here, L_{AR}, L_{BR}, L_{RA}, and L_{RB} are finite integers and represent the channel impulse response length. These channel gains are assumed to be available at each node. Nodes A and B map their messages $u_A \in \{1, 2, \ldots, M_A\}$ and $u_B \in \{1, 2, \ldots, M_B\}$ to length-n channel input sequences $\mathbf{x}_A = \mathcal{E}_A^n(u_A)$ and $\mathbf{x}_B = \mathcal{E}_B^n(u_B)$, respectively, which are then transmitted to the relay. Each node is subject to a power constraint P_i, i.e., $\mathbb{E}[|x_i|^2] \leq P_i$, $i \in \{A, B, R\}$. Alternatively, the power constraint can be represented in frequency domain. Let us define $P_i(\theta)$ to be the corresponding power-spectral density of x_i, one can then write the feasible power constraint set as

$$\mathcal{P} = \left\{ P_A(\cdot), P_B(\cdot), P_R(\cdot) : \frac{1}{2\pi} \int_{-\pi}^{\pi} P_i(\theta) d\theta \leq P_i, i \in \{A, B, R\} \right\}. \tag{3.1}$$

Note that this individual power constraint is more practical than the total power constraint considered in [17] as nodes are assumed to be physically apart.

The transmission protocol we consider is the two-phase protocol consisting of a multiple access channel (MAC) phase and a broadcast channel (BC) phase. Each phase occupies a half of the channel uses and is assumed to be orthogonal to each other. This can be accomplished by assigning different frequency bands, time slots,

or spreading codes.

During the MAC phase, both nodes transmit their signals to the relay simultaneously and the relay remains silent. The received signal at the relay is given by

$$\mathbf{y}_R = \mathbf{x}_A * \mathbf{h}_A + \mathbf{x}_B * \mathbf{h}_B + \mathbf{z}_R, \tag{3.2}$$

where $*$ denotes the linear convolution operator and $\mathbf{z}_R \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ is i.i.d. Gaussian noise. Upon receiving $\mathbf{y}_R$, the relay maps it to the transmitted signal in the BC phase $\mathbf{x}_R = \mathcal{E}_R^n(\mathbf{y}_R)$. This mapping depends on the information forwarding strategy and will be discussed later.

During the BC phase, the relay broadcasts $\mathbf{x}_R$ back to nodes and both nodes remain silent. Then the received signal at both nodes are

$$\mathbf{y}_A = \mathbf{x}_R * \mathbf{g}_A + \mathbf{z}_A, \tag{3.3}$$

$$\mathbf{y}_B = \mathbf{x}_R * \mathbf{g}_B + \mathbf{z}_B, \tag{3.4}$$

where again $\mathbf{z}_A, \mathbf{z}_B \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$. Nodes A and B then form estimates of u_B and u_A as $\hat{u}_B = \mathcal{G}_A^n(\mathbf{y}_A, \mathbf{x}_A)$ and $\hat{u}_A = \mathcal{G}_B^n(\mathbf{y}_B, \mathbf{x}_B)$, respectively.

Definition 14*:* An (n, M_A, M_B) code consists of a set of encoding functions $(\mathcal{E}_A^n, \mathcal{E}_B^n, \mathcal{E}_R^n)$ with message cardinalities M_A and M_B at nodes A and B, respectively. The decoder uses a set of decoding functions $(\mathcal{G}_A^n, \mathcal{G}_B^n)$ and the error probability is then given by

$$\begin{aligned} P_e^{(n)} = &\sum_{u_A, u_B} \frac{1}{M_A M_B} \\ &\cdot \mathbb{P}\left(\{u_A \neq \hat{u}_A\} \cup \{u_B \neq \hat{u}_B\} | u_A,\ u_B \text{ are sent}\right). \end{aligned} \tag{3.5}$$

Definition 15*:* A rate pair (R_A, R_B) is achievable if, for any $\varepsilon > 0$, there exists

an n_o such that for $n \geq n_0$ there is an (n, M_A, M_B) code with

$$M_A \geq 2^{nR_A},\ M_B \geq 2^{nR_B},\ \text{and}\ P_e^{(n)} \leq \varepsilon. \tag{3.6}$$

The capacity region $\mathcal{C}$ of the Gaussian bidirectional relay channel with ISI is defined as the convex hull of the closure of the set of all achievable rate pairs (R_A, R_B). Note that since each phase occupies a half of the channel uses, the achievable rate and the capacity defined above are per two channel uses.

Also, as in [36], we define separate coding and separability of parallel channels (which include the ISI channel as a special case) as follows.

Definition 16*:* (Separate coding [36]) For a set of parallel channels, a coding scheme is said to be separate if no joint processing across sub-channels other than power allocation is permitted.

Definition 17*:* (Separability [36]) A set of parallel channels is said to be separable if its capacity can be achieved by a separate coding scheme (as defined above) and power allocation among sub-channels.

In this work, in addition to providing inner and outer bounds on the capacity region, we show the following theorem through providing a counterexample in Section 3.5.

Theorem 7. *Gaussian bidirectional relay channels with ISI are not always separable.*

3.3 Outer Bound

An outer bound on the capacity region can be obtained by the cut-set argument [30] as shown in the following lemma whose proof is straightforward and hence omitted. Before we start, let us define $C(x) = \log(1 + x)$.

Lemma 8. *$\mathcal{C} \subseteq \overline{\mathcal{C}}$, where*

$$\overline{\mathcal{C}} = \Big\{ (R_A, R_B) : 0 \le R_A \le \min\left\{\overline{C}_{AR}, \overline{C}_{RB}\right\}, \\ 0 \le R_B \le \min\left\{\overline{C}_{BR}, \overline{C}_{RA}\right\} \Big\}, \tag{3.7}$$

where for $i \in \{A, B\}$

$$\overline{C}_{iR} = \frac{1}{2\pi} \int_{-\pi}^{\pi} C\left(\frac{P_i^*(\theta)|h_i(\theta)|^2}{\sigma^2}\right) d\theta, \quad i \in \{A, B\} \\ \overline{C}_{Ri} = \frac{1}{2\pi} \int_{-\pi}^{\pi} C\left(\frac{P_{Ri}^*(\theta)|g_i(\theta)|^2}{\sigma^2}\right) d\theta \quad i \in \{A, B\},$$

with $P_A^*(\theta), P_B^*(\theta), P_{RA}^*(\theta)$, and $P_{RB}^*(\theta)$ be the conventional water-filling solutions [33] for a given $\mathbf{h}_A, \mathbf{h}_B, \mathbf{g}_A$, and $\mathbf{g}_B$, respectively.

Since we also wish to know whether Gaussian bidirectional channels with ISI are always separable or not, we propose the following a rate upper bound for all separate coding schemes..

Lemma 9. *For all separate coding schemes, the achievable rate region is contained in*

$$\overline{\mathcal{C}}^s = \left\{ (R_A, R_B) : 0 \le R_A \le \overline{C}_A^s, 0 \le R_B \le \overline{C}_B^s \right\}, \tag{3.8}$$

where

$$\overline{C}_i^s = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min\left\{ C\left(\frac{P_i^{**}(\theta)|h_i(\theta)|^2}{\sigma^2}\right), C\left(\frac{P_{Rj}^{**}(\theta)|g_j(\theta)|^2}{\sigma^2}\right) \right\} d\theta,$$

$i, j \in \{A, B\}$ $i \ne j$. The optimal power-spectral density $P_i^{**}(\theta)$, and $P_{Rj}^{**}(\theta)$ are given

by

$$P_i^{**}(\theta) = \left(\frac{|g_j(\theta)|^2}{v_1|g_j(\theta)|^2 + v_2|h_i(\theta)|^2} - \frac{\sigma^2}{|h_i(\theta)|^2} \right)^+, \tag{3.9}$$

$$P_{Rj}^{**}(\theta) = \left(\frac{|h_i(\theta)|^2}{v_1|g_j(\theta)|^2 + v_2|h_i(\theta)|^2} - \frac{\sigma^2}{|g_j(\theta)|^2} \right)^+, \tag{3.10}$$

where v_1 and v_2 are chosen such that the power constraints are satisfied.

Proof. The equation (3.8) can be derived in a similar way to the proof of the previous lemma, i.e., applying the cut-set bound to each sub-channel, and is hence omitted. The power-spectral density in (3.9) and (3.10) are derived in Appendix A.1. □

3.4 Proposed Scheme

In this section, we propose a time-domain coding scheme for the considered setting and then analyze the corresponding achievable rate region. This coding scheme can be regarded as a time-domain approach since we do not diagonalize the channel matrices via the discrete Fourier transform, i.e., we do not partition the spectrum into sub-channels. Instead, we code over the whole spectrum directly.

3.4.1 MAC Phase

We consider transmitting K messages together where nodes A and B map messages $u_{A,k}$ and $u_{B,k}$ with $k \in \{1, 2 \ldots, K\}$ to column vectors $\tilde{\mathbf{x}}_{A,k}$ and $\tilde{\mathbf{x}}_{B,k}$, respectively, and $\tilde{\mathbf{x}}_{A,k}$ and $\tilde{\mathbf{x}}_{B,k}$ will be determined later. The node $i \in \{A, B\}$ then passes those columns into the write column-wise transmit row-wise block interleaver [37] to form the sequence $\tilde{\mathbf{x}}_i$. Specifically, as shown in Fig. 3.2, the node i first aggregates the K column vectors and appends $L_g - 1$ columns of 0 to form a $n \times (K + L_g - 1)$ matrix where L_g is the length of $\mathbf{g}$, the equivalent ISI channel that will be defined shortly. It then forms the length-$n(K + L_g - 1)$ vector $\tilde{\mathbf{x}}_i$ by reading the matrix

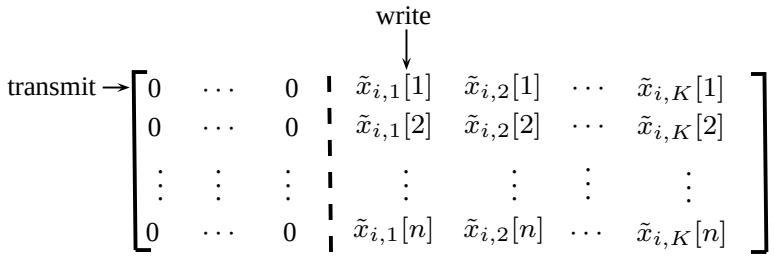

Figure 3.2: Write column-wise transmit row-wise interleaved signal for node $i \in \{A, B\}$

row-wisely.

Since we wish to make use of the group property of lattices, the key idea is to enforce the two ISI channels to be perfectly aligned (up to a scaling factor) at the relay. To this end, as shown in Fig. 3.3, we introduce linear filters $\mathbf{f}_A$ and $\mathbf{f}_B$, respectively, at two nodes so that the transmitted signals become

$$\mathbf{x}_A = \tilde{\mathbf{x}}_A * \mathbf{f}_A, \tag{3.11}$$

$$\mathbf{x}_B = \tilde{\mathbf{x}}_B * \mathbf{f}_B. \tag{3.12}$$

We enforce the power-spectral density of $\tilde{\mathbf{x}}_i$ to be white and $P_i(\theta) = P_i,\ \forall\theta$ and also enforce $\|\mathbf{f}_A\|^2 = \|\mathbf{f}_B\|^2 = 1$. Together they ensure that the power constraints are still satisfied. The key idea of our proposed coding scheme is to choose filters $\mathbf{f}_A$ and $\mathbf{f}_B$ such that the resultant ISI channels at the relay are collinear. i.e.,

$$\mathbf{f}_A * \mathbf{h}_A = a \cdot \mathbf{g}, \tag{3.13}$$

$$\mathbf{f}_B * \mathbf{h}_B = b \cdot \mathbf{g}, \tag{3.14}$$

where $a, b \in \mathbb{R}$ and $\mathbf{g}$ is a length-L_g vector representing the equivalent ISI channel seen at the relay. Also, we assume $a > b$ without loss of generality; otherwise, we can

switch the role of nodes A and B. By doing this, two lattices are perfectly aligned up to a scaling factor at the relay and the received signal at the relay (3.2) can be rewritten as

$$\begin{aligned}\mathbf{y}_R &= \tilde{\mathbf{x}}_A * \mathbf{f}_A * \mathbf{h}_A + \tilde{\mathbf{x}}_B * \mathbf{f}_B * \mathbf{h}_B + \mathbf{z}_R, \\ &= (a \cdot \tilde{\mathbf{x}}_A + b \cdot \tilde{\mathbf{x}}_B) * \mathbf{g} + \mathbf{z}_R \\ &= (\bar{\mathbf{x}}_A + \bar{\mathbf{x}}_B) * \mathbf{g} + \mathbf{z}_R, \end{aligned} \tag{3.15}$$

where $\bar{\mathbf{x}}_A = a \cdot \tilde{\mathbf{x}}_A$ and $\bar{\mathbf{x}}_B = b \cdot \tilde{\mathbf{x}}_B$ are the transmitted signals before pre-filtering. Also, we define $\bar{\mathbf{x}}_{A,k} = a \cdot \tilde{\mathbf{x}}_{A,k}$ and $\bar{\mathbf{x}}_{B,k} = b \cdot \tilde{\mathbf{x}}_{B,k}$ for all $k \in \{1, 2, \ldots, K\}$. It should be noted that in our proposed scheme, since we let the input power-spectral density of $\tilde{\mathbf{x}}_i$ to be $P_i(\theta) = P_i$ for all θ, $\bar{\mathbf{x}}_A$ and $\bar{\mathbf{x}}_B$ are subject to power constraints $\bar{P}_A = a^2 P_A$ and $\bar{P}_B = b^2 P_B$, respectively.

Now, two ISI channels have been perfectly aligned to an equivalent channel $\mathbf{g}$ so that the relay can perform unbiased MMSE equalization and the results in [32] can be directly applied with a slight modification. Let $c[m]$ be the tap-coefficients of the corresponding unbiased MMSE equalizer, the equalized signal at each time index m

is then given by

$$\begin{aligned}
v[m] &= \sum_{j=-\infty}^{\infty} c[j] y_R[m-j] \\
&= \bar{x}_A[m] + \bar{x}_B[m] \\
&\quad + \sum_{j=1}^{\infty} \bar{g}[j]\bar{x}_A[m-j] + \sum_{j=1}^{\infty} \bar{g}[-j]\bar{x}_A[m+j] \\
&\quad + \sum_{j=1}^{\infty} \bar{g}[j]\bar{x}_B[m-j] + \sum_{j=1}^{\infty} \bar{g}[-j]\bar{x}_B[m+j] + \bar{n}_R[m] \\
&\overset{(a)}{=} \bar{x}_A[m] + \bar{x}_B[m] + s_A[m] + s_B[m] \\
&\qquad\qquad + w_A[m] + w_B[m] + \bar{n}_R[m] \\
&= \bar{x}_A[m] + \bar{x}_B[m] + s_A[m] + s_B[m] + e[m], \qquad (3.16)
\end{aligned}$$

where $\bar{g}[j]$'s are the combined effect of actual channel gains and feed forward MMSE filter, and $\bar{n}_R[m]$ denotes the corresponding noise after the MMSE filter. Moreover, (a) follows from the definitions, for $i \in \{A, B\}$,

$$s_i[m] = \sum_{j=1}^{\infty} \bar{g}[j]\bar{x}_i[m-j], \qquad (3.17)$$

being the post-cursor ISI induced by the node $i \in \{A, B\}$, and

$$w_i[m] = \sum_{j=1}^{\infty} \bar{g}[-j]\bar{x}_i[m+j], \qquad (3.18)$$

representing the corresponding pre-cursor ISI. Furthermore, if we assume that previously decoded signals are correct, we can use a DFE to reconstruct $s_A[m]$ and $s_B[m]$ and then subtract them out. Therefore, we can write $e[m] = w_A[m] + w_B[m] + \bar{n}_R[m]$ as the error obtained by the unbiased MMSE-DFE equalizer. The variance of this

error is the mean squared error obtained by the unbiased MMSE-DFE as [32]

$$\sigma_e^2 = \frac{(\bar{P}_A + \bar{P}_B)N_0}{S_0 - N_0}, \tag{3.19}$$

with

$$\log \frac{S_0}{N_0} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left[\frac{\bar{P}_A + \bar{P}_B}{\sigma^2} |S_{\mathbf{g}}(\theta)|^2 + 1 \right] d\theta, \tag{3.20}$$

where $S_{\mathbf{g}}(\theta)$ is the power-spectral density of $\mathbf{g}$. In our proposed scheme, we do not literally perform the DFE at the relay; instead, we adopt lattice precoding at both nodes to get rid of the post-cursor ISI. The reason for this substitution will be explained in Remark 3.4.1.

The relay now pass the sequence $v[m]$ into the read row-wise output column-wise deinterleaver [37]. After removing the guard intervals, the k^{th} output column can be written as

$$\mathbf{v}_k = \bar{\mathbf{x}}_{A,k} + \bar{\mathbf{x}}_{B,k} + \mathbf{s}_{A,k} + \mathbf{s}_{B,k} + \mathbf{e}_k. \tag{3.21}$$

We pick an n-dimensional lattice partition chain $\Lambda_f^n/\Lambda_B^n/\Lambda_A^n$, where Λ_A^n and Λ_B^n are simultaneously Rogers-good and Poltyrev-good while Λ_f^n is Poltyrev-good (for the definition of goodness, the reader is referred to [20, 19]). The second moment of Λ_i^n is chosen to be (very close to) $\bar{P}_i$, $i \in \{A, B\}$. (Precisely, it should be $\bar{P}_i - \delta \leq \sigma^2(\Lambda_i^n) \leq \bar{P}_i$ for $\delta > 0$ and sufficiently large n. However, since eventually we let n tend to ∞ and we are interested in the asymptotic result, we will omit δ from here onwards for the sake of simplicity.) The existence of such lattices is provided in [13, Theorem 2].

Node i use NLC_i the NLC associated with Λ_f^n/Λ_i^n which is defined as $\{\Lambda_f^n \cap \mathcal{V}(\Lambda_i^n)\}$ with $\mathcal{V}(\Lambda_i^n)$ being the fundamental Voronoi region of Λ_i^n. By [13, Theorem 2], the coding rate of NLC_B can approach any value $\gamma \geq 0$ as $n \to \infty$ and that of NLC_A can

approach $\gamma + \frac{1}{2}\log\left(\frac{\bar{P}_A}{\bar{P}_B}\right)$ as $n \to \infty$. Node i then encodes its message to codewords $\mathbf{t}_{i,k} \in NLC_i$ in an injective manner. Observe that due to the interleaver/deinterleaver pair, while encoding the message k, $\bar{\mathbf{x}}_{A,k'}$ and $\bar{\mathbf{x}}_{B,k'}$ for $k' < k$ are encoded already. Now, since the post-cursor ISI are linear combinations of previous codewords which have been encoded, it is known at transmitter. We can then perform the lattice precoding [31] as a structured dirty paper coding to cancel the known interference. The lattice codewords are given by

$$\bar{\mathbf{x}}_{A,k} = [\mathbf{t}_{A,k} - \alpha \mathbf{s}_{A,k} - \mathbf{d}_{A,k}] \mod \Lambda_A^n, \tag{3.22}$$

$$\bar{\mathbf{x}}_{B,k} = [\mathbf{t}_{B,k} - \alpha \mathbf{s}_{B,k} - \mathbf{d}_{B,k}] \mod \Lambda_B^n, \tag{3.23}$$

where $\mathbf{d}_{A,k}$ and $\mathbf{d}_{B,k}$ are random dithers and

$$\alpha = \frac{\bar{P}_A + \bar{P}_B}{\bar{P}_A + \bar{P}_B + \sigma_e^2}, \tag{3.24}$$

is the MMSE coefficient. Note that each $\bar{\mathbf{x}}_{i,k}$ satisfies the power constraint $\bar{P}_i$ and hence $\bar{\mathbf{x}}_i$ satisfies the power constraint $\bar{P}_i$. The block diagram of the proposed transmitters at nodes A and B can be seen in Fig. 3.3 where we would like to recall that the actual transmitted signals are $\mathbf{x}_i$ instead of lattice codewords $\bar{\mathbf{x}}_{i,k}$. Note that $\mathbf{e}_k$ in (3.21) is in general not Gaussian since $\mathbf{w}_{A,k}$ and $\mathbf{w}_{B,k}$ are in general not Gaussian. However, from [38, Theorem 2], we know that if we choose the coarse lattices Λ_A^n and Λ_B^n to be good for quantization, then the distribution of $\bar{\mathbf{x}}_{A,k}$ and $\bar{\mathbf{x}}_{B,k}$ will tend to be a Gaussian distribution (so will $\mathbf{w}_{A,k}$ and $\mathbf{w}_{B,k}$) when $n \to \infty$. i.e., from now on, we can assume $\mathbf{e}_k$ to have a Gaussian distribution.

For each k, the relay first forms an estimate of $\mathbf{t}_{R,k} = (\mathbf{t}_{A,k} + \mathbf{t}_{B,k} - Q_{\Lambda_B^n}(\mathbf{t}_{B,k} -$

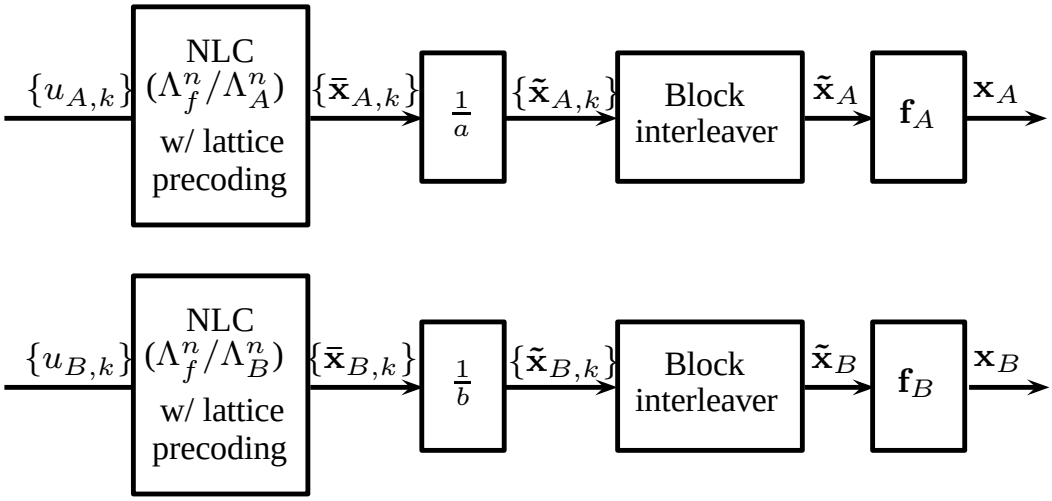

Figure 3.3: Block diagram of the proposed transmitters at nodes A and B

$\alpha\mathbf{s}_{B,k} - \mathbf{d}_{B,k}))) \mod \Lambda_A^n$ from (3.21) according to

$$\begin{aligned}
\hat{\mathbf{y}}_{R,k} &= [\alpha\mathbf{v}_k + \mathbf{d}_{A,k} + \mathbf{d}_{B,k}] \mod \Lambda_A^n \\
&= [\alpha\left(\bar{\mathbf{x}}_{A,k} + \bar{\mathbf{x}}_{B,k} + \mathbf{s}_{A,k} + \mathbf{s}_{B,k} + \mathbf{e}_k\right) \\
&\qquad + \mathbf{d}_{A,k} + \mathbf{d}_{B,k}] \mod \Lambda_A^n \\
&= [\mathbf{t}_{R,k} - (1-\alpha)(\bar{\mathbf{x}}_{A,k} + \bar{\mathbf{x}}_{B,k}) + \alpha\mathbf{e}_k] \mod \Lambda_A^n \\
&= [\mathbf{t}_{R,k} + \mathbf{n}_{eq,k}] \mod \Lambda_A^n, \qquad (3.25)
\end{aligned}$$

where

$$\mathbf{n}_{eq,k} = [\alpha\mathbf{e}_k - (1-\alpha)(\bar{\mathbf{x}}_{A,k} + \bar{\mathbf{x}}_{B,k})] \mod \Lambda_A^n. \qquad (3.26)$$

In deriving (3.25) we use the distributive property of modulo operation several times. Thanks to the interleaver/deinterleaver pair and random dithers, the pre-cursor ISI is linear combination of future codewords and hence independent of the current codewords. Thus, $\mathbf{e}_k$ is independent of $\bar{\mathbf{x}}_{A,k}$ and $\bar{\mathbf{x}}_{B,k}$. Now, with the choice of α

given in (3.24), the average second moment of $\mathbf{N}_{eq,k}$ (whose particular realizations are $\mathbf{n}_{eq,k}$) is given by

$$\begin{aligned}\frac{1}{n}\mathbb{E}\left[\|\mathbf{N}_{eq,k}\|^2\right] &\leq \frac{(\bar{P}_A+\bar{P}_B)\sigma_e^2}{\bar{P}_A+\bar{P}_B+\sigma_e^2}\\ &:= \sigma_{eq}^2. \end{aligned} \tag{3.27}$$

Note that according to the group property of lattices, $\mathbf{t}_{R,k}$ is a lattice point in NLC_A. Also, by applying the crypto-lemma in for example [21, Lemma 2], one has that $\mathbf{t}_{R,k}$ is uniformly distributed over $\{\Lambda_f \cap \mathcal{V}(\Lambda_A^n)\}$ and independent of $\mathbf{n}_{eq,k}$. The relay then attempts to decode to $\hat{\mathbf{t}}_{R,k}$ by finding the fine lattice point closest to $\hat{\mathbf{y}}_{R,k}$ (lattice decoding), i.e., $Q_{\Lambda_f^n}(\hat{\mathbf{y}}_{R,k})$. Thus, the decoding at the MAC phase is said to be successful if $\mathbf{t}_{R,k} = Q_{\Lambda_f^n}(\hat{\mathbf{y}}_{R,k})$. From (3.25), it is observed that to make the communication reliable during the MAC phase, we require

$$\mathbb{P}\left(\{\mathbf{N}_{eq,k} \notin \mathcal{V}(\Lambda_f^n)\}\right) \to 0, \text{ as } n \to \infty. \tag{3.28}$$

Note that after perfectly aligning two ISI channels, the derivations above are the same with those in [13] except that we incorporate the lattice precoding into the lattice partition chains scheme. Moreover, the loss of appending guard intervals can be made arbitrarily small as we let K tend to infinity and hence can be ignored. Thus, we can directly apply the Theorem 1 in [13] to obtain the achievable rate

region for the proposed scheme in the MAC (uplink) phase which is given by

$$\begin{aligned} R_{A,\mathrm{up}}(\mathbf{g}) &= \log\left(\frac{\bar{P}_A}{\sigma_{eq}^2}\right)^+ \\ &= \log\left(\frac{\bar{P}_A}{\bar{P}_A+\bar{P}_B}+\frac{\bar{P}_A}{\sigma_{eq}^2}\right)^+ \\ &= \log\left(\frac{a^2P_A}{a^2P_A+b^2P_B}+\frac{a^2P_A}{\sigma_{eq}^2}\right)^+, \end{aligned} \tag{3.29}$$

and

$$R_{B,\mathrm{up}}(\mathbf{g}) = \log\left(\frac{b^2P_B}{a^2P_A+b^2P_B}+\frac{b^2P_B}{\sigma_{eq}^2}\right)^+, \tag{3.30}$$

where the subscript "up" stands for the uplink.

Remark: In the MAC phase, the problem resembles the *dirty* MAC problem where two nodes wish to communicate with a destination node and the communication is corrupted by two pieces of interference. Each piece of interference is known to a different node. This problem has been studied in [39] where Philosof *et al.* showed that there is an inherent loss with respect to the *clean* MAC. However, here since the relay node is only interested in computing $\mathbf{t}_{R,k}$, we can achieve the computation rate as there is no interference at all as shown in (3.29) and (3.30).

Remark: One classical way to deal with the post-cursor ISI is to equip the receiver with a DFE so that one can subtract the post-cursor ISI once the previous codewords are decoded [32]. However, when we directly apply the lattice-based encoding and decoding techniques in [13] to here and use an unbiased MMSE-DFE at the relay, we still cannot get rid of the post-cursor ISI $\mathbf{s}_{A,k}+\mathbf{s}_{B,k}$. It is due to the fact that in lattice decoding, instead of decoding to $\mathbf{t}_{A,k}+\mathbf{t}_{B,k}$, the relay tries to decode to $\mathbf{t}_{R,k} = (\mathbf{t}_{A,k}+\mathbf{t}_{B,k}-Q_{\Lambda_B^n}(\mathbf{t}_{B,k}-\alpha\mathbf{s}_{B,k}-\mathbf{d}_{B,k}))\mod \Lambda_A^n$, making it very difficult for the relay to reconstruct $\mathbf{s}_{A,k}+\mathbf{s}_{B,k}$. Therefore, the lattice precoding is a reasonable

alternative of dealing with this issue.

It should be noted that very recently, Nazer proposed the successive compute-and-forward scheme in [40] that can effectively reconstruct the real sum from the modulo sum at the relay when two users adopt the same NLC. This approach can also be extended to the lattice partition chains scheme considered in this section.

Remark: Since all the pre-cursor ISI $\mathbf{w}_{A,k}$ and $\mathbf{w}_{B,k}$ are also known at its own transmitter, it is natural to argue that one can further improve the rate using the lattice precoding to get rid of them also. However, this is fallacious in the sense that one cannot process the future symbols which are yet to be determined. Specifically, after preprocessing the post-cursor ISI, every codeword is a function of the message at that time and some previous codewords (can be seen from (3.22) and (3.23)). Thus, any attempts of preprocessing the pre-cursor ISI may change the current symbol and hence may change the future symbols, which in turn changes the pre-cursor ISI itself. This leads to a logical fallacy. (The interested reader is referred to [31].)

3.4.2 BC Phase

In the BC (downlink) phase, we drop the subscript k as the coding scheme is the same for all k. The relay encodes the decoded $\mathbf{t}_R$ to $\mathbf{x}_R$ by a random Gaussian codebook. Nodes A and B first form estimates of $\mathbf{t}_R$, namely $\hat{\mathbf{t}}_{R,A}$ and $\hat{\mathbf{t}}_{R,B}$, respectively. The two nodes then form estimates of each other's message as

$$\hat{\mathbf{t}}_B = \left[\hat{\mathbf{t}}_{R,A} - \mathbf{t}_A\right] \mod \Lambda_B^n, \tag{3.31}$$

and

$$\hat{\mathbf{t}}_A = \left[\hat{\mathbf{t}}_{R,B} - (\mathbf{t}_B - Q_{\Lambda_B^n}(\mathbf{t}_B - \alpha\mathbf{s}_B - \mathbf{d}_B))\right] \mod \Lambda_A^n, \tag{3.32}$$

respectively.

Similar to [13] [41], we can now use a typical set decoder to show that the rate pair given below is achievable.

$$R_{A,\text{down}} \leq I(\mathbf{X}_R; \mathbf{Y}_A) = \log\left(1 + \frac{P_R}{\sigma_A^2}\right), \tag{3.33}$$

$$R_{B,\text{down}} \leq I(\mathbf{X}_R; \mathbf{Y}_B) = \log\left(1 + \frac{P_R}{\sigma_B^2}\right), \tag{3.34}$$

where the subscript "down" stands for the downlink and where σ_A^2 and σ_B^2 are derived by equipping each node with an optimal unbiased MMSE-DFE [32] and are given by

$$\sigma_A^2 = \frac{P_R N_0'}{S_0' - N_0'}, \tag{3.35}$$

$$\sigma_B^2 = \frac{P_R N_0''}{S_0'' - N_0''}, \tag{3.36}$$

with

$$\log\frac{S_0'}{N_0'} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[\frac{P_R}{\sigma^2}|S_{\mathbf{g}_A}(\theta)|^2 + 1\right] d\theta, \tag{3.37}$$

$$\log\frac{S_0''}{N_0''} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[\frac{P_R}{\sigma^2}|S_{\mathbf{g}_B}(\theta)|^2 + 1\right] d\theta. \tag{3.38}$$

3.4.3 Achievable Rate Region of the Proposed Scheme

Having derived the achievable rate of the proposed scheme in the MAC phase and the BC phase, we now summarize the achievable rate region of the proposed scheme in the following theorem.

Theorem 10. *For the Gaussian bidirectional relay channel with ISI* $\mathbf{h}_A$ *and* $\mathbf{h}_B$*, for every* $\mathbf{g}$ *such that there exist unit-norm filters* $\mathbf{f}_A$ *and* $\mathbf{f}_B$ *with* $\mathbf{f}_A * \mathbf{h}_A = a\mathbf{g}$ *and*

$\mathbf{f}_B * \mathbf{h}_B = b\mathbf{g}$, *respectively, the following rate pair is achievable.*

$$R_A(\mathbf{g}) \leq \min(R_{A,up}(\mathbf{g}), R_{B,down}), \tag{3.39}$$

$$R_B(\mathbf{g}) \leq \min(R_{B,up}(\mathbf{g}), R_{A,down}). \tag{3.40}$$

It should be noted that the rates described above are per two channel uses since we assume that each phase occupies a half of the channel uses. Notice that these rates are functions of the equivalent ISI channel $\mathbf{g}$ which is a result of pre-filters $\mathbf{f}_A$ and $\mathbf{f}_B$ that will be discussed in the following.

3.4.4 Filter Design

The problem now becomes how to choose a valid pair of unit-norm filters $\mathbf{f}_A$ and $\mathbf{f}_B$ such that (3.13) and (3.14) are satisfied. Here, only the case of $\mathbf{h}_A \neq \mathbf{h}_B$ is considered since if $\mathbf{h}_A = \mathbf{h}_B = \mathbf{h}$, there is no need to pre-filter the signals and choosing the pre-filters to be 1 automatically satisfies (3.13) and (3.14) with $a = b = 1$ and $\mathbf{g} = \mathbf{h}$.

For $\mathbf{h}_A \neq \mathbf{h}_B$, one valid choice is to make $\mathbf{f}_A = \mathbf{h}_B/\|\mathbf{h}_B\|$ and $\mathbf{f}_B = \mathbf{h}_A/\|\mathbf{h}_A\|$ so that

$$\mathbf{g} = \mathbf{h}_A * \mathbf{h}_B, \tag{3.41}$$

and $a = 1/\|\mathbf{h}_B\|$ and $b = 1/\|\mathbf{h}_A\|$. With this choice, the achievable rate in the MAC phase described in (3.29) and (3.30) becomes

$$R_{A,\text{up}} = \log\left(\frac{\|\mathbf{h}_A\|^2 P_A}{\|\mathbf{h}_A\|^2 P_A + \|\mathbf{h}_B\|^2 P_B} + \frac{P_A}{\|\mathbf{h}_B\|^2 \sigma_{eq}^2}\right)^+, \tag{3.42}$$

$$R_{B,\text{up}} = \log\left(\frac{\|\mathbf{h}_B\|^2 P_B}{\|\mathbf{h}_A\|^2 P_A + \|\mathbf{h}_B\|^2 P_B} + \frac{P_B}{\|\mathbf{h}_A\|^2 \sigma_{eq}^2}\right)^+. \tag{3.43}$$

Remark: This choice of $\mathbf{f}_A$ and $\mathbf{f}_B$ is by no means optimal. Of course one can try

to optimize the choice of transmitted filters $\mathbf{f}_A$ and $\mathbf{f}_B$ by solving the corresponding optimization problem. However, this problem seems to be non-convex. Thus, in the next section, we only give a numerical example of filter design problem for maximizing the exchange rate in the MAC phase given by

$$\begin{aligned} \text{maximize} \quad & \min\{R_{A,\text{up}}(\mathbf{g}), R_{B,\text{up}}(\mathbf{g})\} \\ \text{subject to} \quad & \mathbf{f}_A * \mathbf{h}_A = a \cdot \mathbf{g}, \\ & \mathbf{f}_B * \mathbf{h}_B = b \cdot \mathbf{g}, \end{aligned} \tag{3.44}$$

and leave the optimal filter design problem as potential future work.

Remark: For the case when $P_A = P_B$ and two channels have a same L_2-norm, the proposed coding scheme with filters chosen above reduces to the one proposed in [17] where, instead of lattice partition chains scheme, we use an identical NLC for both nodes as in [12].

Remark: It is worth mentioning that we will discuss another approach in Chapter for the considered problem where a frequency-domain approach is used in the MAC phase and joint coding across sub-channels is only adopted in the BC phase. This approach is shown to be asymptotically optimal in terms of the achievable rate and can approach the capacity region to within a constant bit. However, the time-domain approach proposed in this chapter is still interesting for several reasons. First, it is unclear which scheme would perform better in the finite SNR and finite block-length regime. Second, practical implementations of the time-domain approach may provide advantages over the frequency-domain approach in some situations of interest.

3.5 Numerical Results

In this section, we present numerical results for comparing the proposed scheme, the cut-set bound, and some existing schemes. As a byproduct, we also show that Gaussian bidirectional relay channels with ISI are not always separable. In Fig. 3.4, we plot the achievable rate region obtained by the scheme proposed in Section 3.4 and the outer bounds derived in Section 3.3 for P_A and P_B ranging from 0 to 20dB, and $P_R = \max\{P_A, P_B\}$. The channel parameters are set to be $\mathbf{h}_A = [1\ 1]^t$, $\mathbf{h}_B = [1\ -0.5]^t$, $\mathbf{g}_A = [1\ -0.5]^t$, and $\mathbf{g}_B = [1\ 1]^t$. Here, we choose the transmit filters as those in Section 3.4.4, i.e., $\mathbf{f}_A = \mathbf{h}_B/\|\mathbf{h}_B\|$ and $\mathbf{f}_B = \mathbf{h}_A/\|\mathbf{h}_A\|$. We observe that the proposed scheme offers a larger region than that provided by the cut-set bound for separate coding schemes. This implies that allowing joint coding across sub-channels enlarges the capacity region and that separate encoding schemes with optimal power allocation only is in general not sufficient to achieve the capacity. Therefore, although the choice of input linear filters is in general not optimal as mentioned in *Remark* 3.4.4, this at least provides an example showing that, similar to the parallel Gaussian interference channel [36], joint encoding across sub-channels is in general required for achieving the capacity.

Remark: It is worth mentioning that although the inner and outer bounds provided in Fig. 3.4 look like rectangles, the capacity region of the considered channel may not be a rectangle. Here, for the inner bound, we only plot the achievable rate for our proposed CF scheme, which appears to be a rectangle (but it is actually not a rectangle if one looks carefully). On the other hand, if we simply omit one node and implement point-to-point transmission for the other, a higher rate can be achieved for the operating node (see "Naïve scheme" in Fig. 3.4.) Now, by time-sharing between the proposed CF scheme and the naïve scheme, one obtains an inner bound

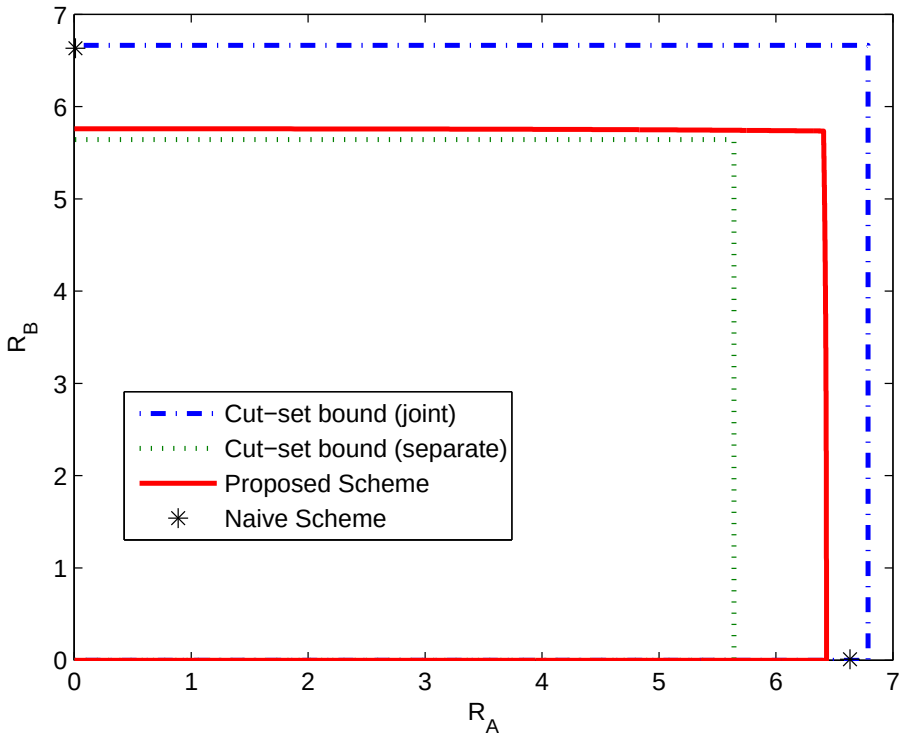

Figure 3.4: Inner and outer bounds on capacity region for SNR = 20dB, $\mathbf{h}_A = [1\ 1]^t$ and $\mathbf{h}_B = [1\ \ -0.5]^t$ case

on the capacity region which is not a rectangle.

3.5.1 Comparison

We now compare the performance of our proposed scheme with some existing schemes. To make the comparison easier, we set $P_A = P_B = P_R = P$ and only compare the maximum symmetric rate, i.e., the exchange rate as in [17], [12], and [42]. Moreover, we consider the reciprocal case where $\mathbf{h}_A = \mathbf{g}_A$ and $\mathbf{h}_B = \mathbf{g}_B$. For comparison, in addition to the proposed scheme, we also plot the exchange rate achieved by the frequency domain scheme in [17] where an orthogonal frequency division multiplexing-based scheme is adopted and a power allocation strategy is

performed so that each sub-channel can directly carry out the NLC scheme proposed in [12]. As a benchmark, the exchange rate provided by the two-phase DF scheme [41] is plotted where each sub-channel separately use the DF strategy and each phase uses a half of channel uses. Moreover, the outer bounds in Section 3.3 for both separate encoding and joint encoding cases are also provided.

In Fig. 3.5, we plot the exchange rate against the power available at each node P with ISI channels $\mathbf{h}_A = [1\ 1]^t$ and $\mathbf{h}_B = [1\ -1]^t$. It can be observed that both the proposed scheme and the frequency domain scheme in [17] outperform the DF substantially in the high SNR regime. This is a typical phenomenon when one compares the DF and CF strategies in the high SNR regime, since the DF strategy which requires the relay to decode the individual messages may result in a significant loss in multiplexing gain. Moreover, in this figure, one can observe that the proposed scheme outperforms the frequency domain scheme whose performance is upper bounded by the cut-set bound for separate coding schemes.

3.5.2 Filter Design Example

We give an example of designing the pair of filters despite the fact that the problem seems to be non-convex. The channel parameters are set to be $\mathbf{h}_A = [1\ 0.5]^t$ and $\mathbf{h_B} = [1\ 0.3\ 0.4]^t$. Let l be the maximum number of taps of filters $\mathbf{f}_A$ and $\mathbf{f}_B$. We numerically solve the optimization problem in (3.44) for $l = 3,\ 5$ and 20. In Table 3.1, we list the designed transmit filters before normalization $\tilde{\mathbf{f}}_\mathbf{i}$ (i.e., $\mathbf{f}_i = \tilde{\mathbf{f}}_i/\|\tilde{\mathbf{f}}_i\|$ for $i \in \{A, B\}$). The exchange rate of the proposed scheme with these transmit filters are presented in Fig. 3.6. One can observe in this example that there is a reasonable improvement on the achievable exchange rate when the filter length is increased from $l = 3$ to $l = 5$; however, the improvement from increasing l to 20 is marginal. Also, it should be noted that since the optimization problem may be non-convex, these

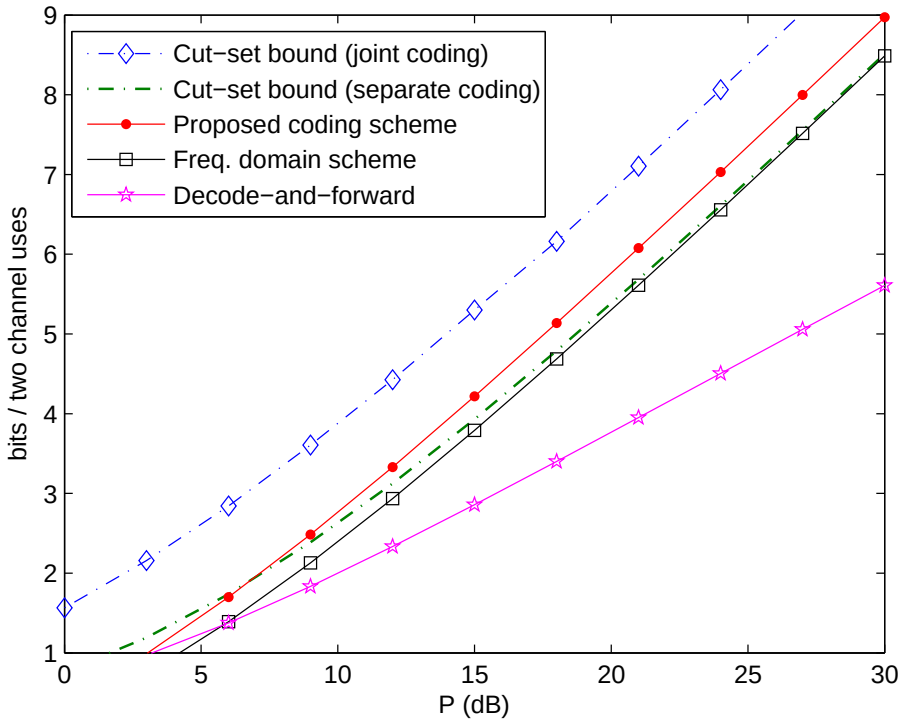

Figure 3.5: Exchange rate vs P for $\mathbf{h}_A = [1\ 1]^t$ and $\mathbf{h}_B = [1\ -1]^t$ case

	$\tilde{\mathbf{f}}_A$	$\tilde{\mathbf{f}}_B$
$l = 3$	[1, 0.3, 0.4]	[1, 0.5, 0]
$l = 5$	[1, -0.0299, 0.2391, -0.1505, -0.0248]	[1, 0.1701, -0.2269, -0.0309, 0]
$l = 20$	[1, -0.056, 0.251, -0.0656, -0.0231, 0.0371, -0.0153, 0.0167, -0.0083, 0.0176, -0.0095, 0.01734, -0.0088, 0.0167, -0.0084, 0.01686, -0.009, 0.0148, -0.0002, 0.0048]	[1, 0.144, -0.2202, 0.0684, 0.0117, -0.0053, 0.0001, 0.0111, -0.0033, 0.01, -0.0023, 0.0093, -0.002, 0.0091, -0.002, 0.0096, -0.0027, 0.0073, 0.006, 0]

Table 3.1: Transmit filters before normalization

solutions may be local.

3.6 Conclusions and Future Works

In this work, we provided inner and outer regions on the capacity region of the bidirectional relaying over ISI. The outer bound was obtained by the conventional cut-set argument while the inner bound was obtained by the proposed time-domain coding scheme. This scheme first pre-filters the transmitted signal at each node so that lattices used at each node are perfectly aligned up to a scaling factor at the relay. After that, with the help of adopting lattice precoding at each node, one can directly carry out the CF scheme in [13] for exploiting the group property in lattices. Although the results in this chapter are suited for all unit-norm pre-filters satisfying the lattice-alignment condition, designing optimal pre-filters seems to be quite challenging and is left as a potential future work.

Interestingly, in contrast to its point-to-point communication counterpart, for the Gaussian bidirectional relaying with ISI, there exist examples for which the proposed time-domain scheme beats the cut-set bound of separate encoding schemes despite the fact that the choice of linear filters may be suboptimal. This implies that Gaussian bidirectional relay channels with ISI are not always separable and to achieve the

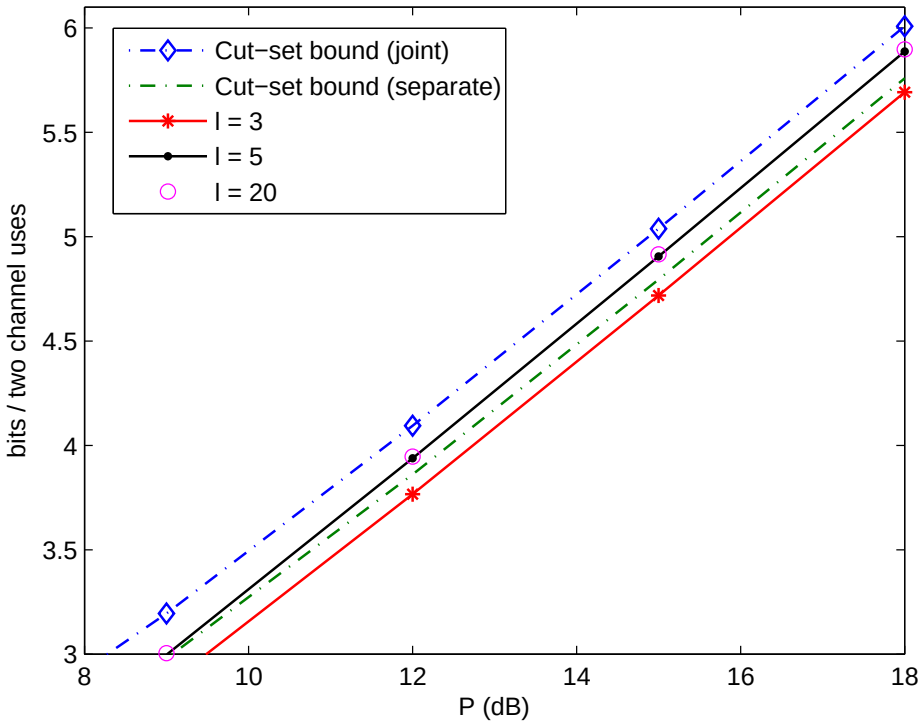

Figure 3.6: Exchange rate vs P for $\mathbf{h}_A = [1\ 0.5]^t$ and $\mathbf{h}_B = [1\ 0.3\ 0.4]^t$ case

capacity, joint processing across sub-channels is required in general.

4. CODING FOR PARALLEL GAUSSIAN BIDIRECTIONAL RELAY CHANNELS: A DETERMINISTIC APPROACH*

This chapter studies the capacity region and the efficient coding schemes for the parallel Gaussian bidirectional relay channel with L independent sub-channels. A two-step approach is considered. First, the corresponding finite-field linear deterministic model is studied, for which a compute-and-forward scheme is shown to achieve the cut-set outer bound of the channel. Next, based on the insight obtained, a lattice-based compute-and-forward scheme is proposed and is shown to achieve the capacity region of the Gaussian model to within L bits regardless of the channel parameters. Even though coding across different sub-channels is necessary for approaching the cut-set outer bound, it is shown that this can be realized via a simple bit reallocation (across different sub-channels) at the relay when the uplink and downlink channels are reciprocal. Numerical results show that the proposed scheme substantially outperforms the traditional decode-and-forward schemes and also provides nontrivial gains over the scheme recently proposed by Huang *et al.* when applied to the inter-symbol interference channel.

4.1 Introduction

In this chapter, we study the parallel Gaussian bidirectional relay network with half-duplex transceivers shown in Fig. 4.1. The channel state information at the transmitter (CSIT) is assumed at each node. The transmission protocol that we consider is a two-phase protocol which consists of a multiple-access (MAC) phase followed by a broadcast (BC) phase, where each phase occupies a half of the channel

uses. Many communication channels can be transformed into parallel Gaussian channels and, therefore, the parallel channel considered here represents a canonical model to study such communication scenarios. For example, inter-symbol interference (ISI) channels and multiple-input multiple-output (MIMO) channels can be converted into parallel Gaussian channels via multi-carrier systems such as orthogonal frequency division multiplex and via matrix decompositions, respectively.

The problem of communication over a single bidirectional relay channel has been intensively studied. Classical information forwarding strategies proposed for (unidirectional) relay channels such as amplify-and-forward [43] and decode-and-forward [43] [44] have been extended to the bidirectional relay problem, see for example [45], [26], and [41]. However, the amplify-and-forward strategy that directly forwards the received signals without cleaning up the noise suffers from noise amplification. The decode-and-forward strategy first decodes the individual messages sent in the MAC phase and then broadcasts them back to the end nodes in the BC phase. This strategy suffers from a loss of multiplexing gain and can be very suboptimal in the high signal-to-noise ratio (SNR) regime.

Recently, it has been shown that a coding scheme based on nested lattice codes at both nodes and compute-and-forward at the relay can substantially outperform classical forwarding strategies for the bidirectional relaying [12] [13]. The main idea comes from the observation that the relay does not need to decode the individual messages; instead, in compute-and-forward the relay is only required to decode the received signal to a function of the individual messages such that both end nodes can figure out the other's message from this function and from their own message as side information. In [12] and [42], the authors considered symmetric channel gains and adopted an identical nested lattice code at both nodes so that the lattice alignment is automatic at the relay. The relay then exploits the group property of

lattices and directly decodes the received signal to the modulo-sum of lattice points. When channel is asymmetric, in [13], Nam *et al.* proposed a scheme adopting lattice partition chains in the MAC phase and directly decoding the received signals to a version of modulo-sum at the relay. Both these two schemes are purely structured and turns out can approach the corresponding capacity region to within one bit per complex dimension (For more details about the schemes in [12] [13], the reader is referred to Chapter 2.) More recently, in [46], Lim *et al.* proposed a layer noisy network coding scheme and showed that this purely unstructured code can also approach the capacity region to within one bit.

Motivated by the success of nested lattice codes and compute-and-forward, in Chapter 3, we proposed a pre-filtering scheme for the bidirectional relay channel with ISI (a special case of the parallel Gaussian bidirectional relay channel) and provided an example showing that the Gaussian bidirectional relay channel is inseparable, i.e., to achieve its capacity, joint processing across sub-channels are required in general. The main idea of this coding scheme is to align two ISI channels by performing pre-filtering at both nodes so that the codewords sent by the two nodes would add up in a correct direction at the relay. Therefore, one can again apply the compute-and-forward strategy to the ISI setup. Notice that in Chapter 3, we only provide one feasible pair of filters and leave the optimal pair of filters design as an open problem since this problem seems to be non-convex and very difficult to solve. Coding schemes for the MIMO bidirectional relay channel (another special case of the parallel setting) have also been considered in [47] and [48]. However, the focus of [47] and [48] is mainly on how to create a parallel bidirectional relay channel from a MIMO problem while the focus of this chapter is on designing coding schemes when the parallel bidirectional relay channel has been created.

Different from [47] [48] and Chapter 3, in this chapter, our coding scheme is

motivated by the linear finite-filed deterministic model proposed by Avestimehr *et al.* [7] of the considered setup. This is a two-step approach that allows us to first ignore the background noise and then focus on the interaction between signals from different nodes. Coding schemes can then be designed according to the insight obtained from this relatively simple finite-field network coding problem. This approach has been very successful in capturing the features that a good coding scheme should possess and further characterizing the capacity region to within a constant bit for many networks. For instance, this model has been successfully used for the (arbitrary) Gaussian relay network [7], the two-user interference network [49], the bidirectional relay channel [50], the multi-pair bidirectional relay channel [51], the two user ZZ and ZS interference-relay channels [52], the parallel relay wire-tap network [53], and the many-to-one and one-to-many interference channel [54].

In this chapter, similar to [7], [49], [50], [51], [52], [53], and [54], we start by investigating the corresponding linear deterministic model and propose a coding scheme that achieves the cut-set bound for the deterministic model. This suggests an insight that for the original Gaussian problem, one should do independent coding for each sub-channel in the MAC phase and joint coding across sub-channels in the BC phase. A coding scheme is proposed for the original Gaussian model based on the above insight. In the MAC phase, the superposition of a nested lattice code and a random Gaussian code is adopted at each sub-channel independently. In the BC phase, the relay regards the problem as the bidirectional broadcast channel (BBC) with common message and directly apply the optimal BBC coding in [55] for joint coding across sub-channels. The proposed scheme is shown to approach the capacity region to within L bits, which provides a means to circumvent the optimal filters design problem posted in [17]. It is worth noting that when $L = 1$, although both the scheme in [50] and our proposed scheme achieve the cut-set bound for the

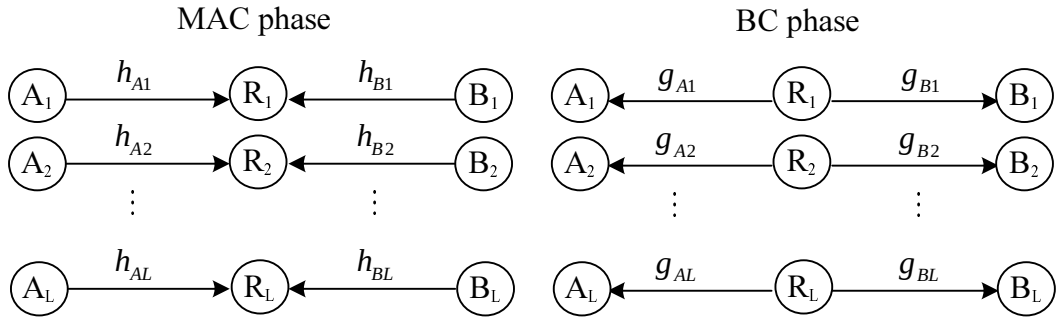

Figure 4.1: The parallel Gaussian bidirectional relay channel.

linear deterministic model, our schemes outperforms the one in [50] for the original Gaussian model and provides the same approximation with the state-of-the-art [13] [46]. Furthermore, although being a special case of general channels, channels with the reciprocity assumption will be studied and a very simple and structured coding scheme will be proposed.

The organization of this chapter is as follows. We introduce the channel model and the proposed scheme for the parallel deterministic model in Section 4.2. The channel model for the parallel Gaussian deterministic model is given in 4.3 where a coding scheme is proposed as well based on the insight obtained from Section 4.2. Some interesting discussions are provided in Section 4.4 where we extend the coding scheme in [13] to the parallel setting, consider the channel reciprocity assumption, compare the proposed scheme with some existing schemes, and show some numerical results on the exchange rate. In Section 4.5, there are some concluding remarks.

4.2 Parallel Deterministic Bidirectional Relay Channel

4.2.1 Channel Model

In this section, we study the parallel deterministic bidirectional relay channel where two nodes A and B wish to exchange information through a relay node R

between them. The channels between nodes can be described as collections of L parallel deterministic bidirectional relay channels where at the l^{th} sub-channel, there are n_l links connected from A_l to R_l and m_l links from B_l to R_l during the MAC phase. Also, in the BC phase, the number of links from R_l to A_l is denoted as r_l and that from R_l to B_l is denoted as k_l. Further, joint processing across sub-channels is permitted. Each node is assumed to have the global channel knowledge, i.e., global CSIT is assumed. One example of a parallel deterministic bidirectional relay channel with two sub-channels can be found in Fig. 4.2.

In the MAC phase, at the l^{th} sub-channel, A_l and B_l map the messages $w_A \in \{1, 2, \ldots, M_A\}$ and $w_B \in \{1, 2, \ldots, M_B\}$ to $q \times N$ codeword matrices $\mathbf{X}_{Al} = \mathcal{E}_{Al}^N(w_A) \in (\mathbb{F}_2^q)^N$ and $\mathbf{X}_{Bl} = \mathcal{E}_{Bl}^N(w_B) \in (\mathbb{F}_2^q)^N$, respectively, independent of each other and of all other sub-channels' messages. Here $q = \max(n_1 \ldots, n_L, m_1 \ldots, m_L, r_1 \ldots, r_L, k_1 \ldots, k_L)$. The received signal at R_l is given by

$$\mathbf{Y}_{Rl} = \mathbf{S}^{q-n_l}\mathbf{X}_{Al} \oplus \mathbf{S}^{q-m_l}\mathbf{X}_{Bl}. \tag{4.1}$$

where $\oplus$ is the XOR operation and

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \tag{4.2}$$

is a $q \times q$ downshift matrix. Let $\mathbf{Y}_R = [\mathbf{Y}_{R1}^t, \mathbf{Y}_{R2}^t, \ldots, \mathbf{Y}_{RL}^t]^t$ be the collection of received signals at the relay.

The relay first maps the received signal $\mathbf{Y}_R$ to the transmitted signal in the BC

phase at the R_l as

$$\mathbf{X}_{Rl} = \mathcal{E}_{Rl}^{N}(\mathbf{Y}_R), \tag{4.3}$$

where $\mathcal{E}_{Rl}^{N}(\cdot)$ represents an information forwarding strategy for generating transmitted signals at R_l. Note that since we allow joint processing across sub-channels, $\mathbf{X}_{Rl}$ depends on the whole $\mathbf{Y}_R$ instead of just $\mathbf{Y}_{Rl}$.

In the BC phase, at the l^{th} sub-channel, R_l broadcasts $\mathbf{X}_{Rl}$ to both nodes and the received signals are given by

$$\mathbf{Y}_{Al} = \mathbf{S}^{q-r_l}\mathbf{X}_{Rl}, \tag{4.4}$$

$$\mathbf{Y}_{Bl} = \mathbf{S}^{q-k_l}\mathbf{X}_{Rl}. \tag{4.5}$$

Nodes A and B then collect all the received signals to form $\mathbf{Y}_A = [\mathbf{Y}_{A1}^t, \mathbf{Y}_{A2}^t, \ldots, \mathbf{Y}_{AL}^t]^t$ and $\mathbf{Y}_B = [\mathbf{Y}_{B1}^t, \mathbf{Y}_{B2}^t, \ldots, \mathbf{Y}_{BL}^t]^t$, respectively, and then try to figure out the other's message from the collected signals together with its own message as side information.

Definition 18*:* An (N, M_A, M_B) code consists of a sequence of encoding functions $\left(\mathcal{E}_{Al}^{N}, \mathcal{E}_{Bl}^{N}, \mathcal{E}_{Rl}^{N}\right)$, decoding functions $\left(\mathcal{G}_{Al}^{N}, \mathcal{G}_{Bl}^{N}\right)$, $l \in \{1, 2, \ldots, L\}$, and an error probability

$$\begin{aligned} P_e^{(N)} = & \sum_{w_A, w_B} \frac{1}{M_A M_B} \\ & \cdot \mathbb{P}\left(\{w_A \neq \hat{w}_A\} \cup \{w_B \neq \hat{w}_B\} | w_A w_B \text{ are sent}\right). \end{aligned} \tag{4.6}$$

Definition 19*:* A rate pair (R_{AB}^d, R_{BA}^d) is said to be achievable if, for any $\varepsilon > 0$, there is an (N, M_A, M_B) code such that

$$M_A \geq 2^{NR_{AB}^d} \text{ and } M_B \geq 2^{NR_{BA}^d} \tag{4.7}$$

and $P_e^{(N)} \leq \varepsilon$. The capacity region $\mathcal{C}^d$ of the parallel deterministic bidirectional relay channel is then defined as the closure of the set of all achievable rate pairs.

An upper bound on the capacity region of the parallel deterministic bidirectional relay channel can be easily derived from the cut-set argument [30] as follows,

Lemma 11. *$\mathcal{C}^d \subseteq \overline{\mathcal{C}}^d$, where*

$$\begin{aligned}\overline{\mathcal{C}}^d = \Bigg\{ & (R_{AB}^d, R_{BA}^d) : 0 \leq R_{AB}^d \leq \overline{C}_{AB}^d, 0 \leq R_{BA}^d \leq \overline{C}_{BA}^d, \\ & \overline{C}_{AB}^d = \min\left\{ \sum_{l=1}^{L} n_l, \sum_{l=1}^{L} k_l \right\}, \\ & \overline{C}_{BA}^d = \min\left\{ \sum_{l=1}^{L} m_l, \sum_{l=1}^{L} r_l \right\} \Bigg\}, \end{aligned} \tag{4.8}$$

per two channel uses (one for each phase.)

The main result of this section stated in below completely characterizes the capacity region of the parallel deterministic relay channel.

Theorem 12. *The capacity region of the parallel deterministic bidirectional relay channel is equal to the cut-set upper bound. i.e., $\mathcal{C}^d = \overline{\mathcal{C}}^d$.*

Sometimes, it is also interesting to study the exchange rate and exchange capacity defined as follows.

Definition 20*:* An exchange rate R_{ex}^d is said to be achievable if (R_{ex}^d, R_{ex}^d) lies inside the capacity region $\mathcal{C}^d$. The exchange capacity C_{ex} is then defined as the supremum of all achievable exchange rates. Therefore, it is the maximum rate that the two nodes can reliably transmit at a same rate simultaneously. i.e., the maximum symmetric rate.

4.2.2 Toy Example

For the deterministic model, before discussing the coding scheme, we emphasize via Example 1 that different from the conventional bidirectional relay channel (with only one sub-channel) [50], routing is in general not optimal and network coding is required for the parallel setting.

Example 1. Let us consider the example with $L = 2$ given in Fig. 4.2. The cut-set bound is $(C_{AB}^d, C_{BA}^d) = (5, 5)$ bits per two channel uses. It is obvious that for this example, any routing scheme is not able to achieve the cut-set bound. We now provide a network coding scheme that can achieve this bound. As shown in the part (a) of Fig. 4.2, in the MAC phase, A_l and B_l send n_l and m_l data streams, respectively. Each sub-channel first shifts bits belonging to the aligned part to the top. After this, the relay still wants to broadcast B_{11}, B_{12}, B_{13} to the node A and A_{21}, A_{22}, A_{23} to the node B. Moreover, the relay still has 2 bidirectional links, 1 unidirectional link to the node A, and 1 unidirectional link to the node B. Since this is the case, the relay shifts B_{13} to the unidirectional link to the node A and also A_{23} to that to the node B. Further, the relay performs linear network coding and send $A_{21} \oplus B_{11}$ and $A_{22} \oplus B_{12}$ through two bidirectional links, respectively. This coding and forwarding scheme is illustrated in Fig. 4.2-(b) where one can verify that indeed we can achieve 5 bits. $\Diamond$

From the above example, one can see that blindly extending the coding scheme in [50] to here results in a suboptimal scheme. In what follows, we will provide a coding scheme that achieves the cut-set upper bound given in Lemma 11. This will complete the proof of Theorem 12.

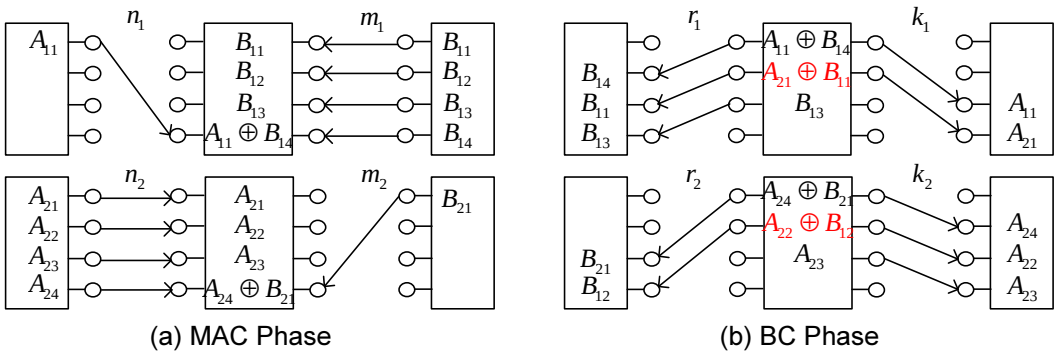

Figure 4.2: Example 1.

4.2.3 Proposed Scheme

We outline the proposed coding scheme for the deterministic parallel bidirectional relay channel as follows:

step 1: At l^{th} sub-channel, in the MAC phase, A_l and B_l send data streams to R_l through some of its n_l and m_l most significant bit (MSB), respectively.

step 2: The relay collects all the received signals on all sub-channels and then jointly encode them across sub-channels for broadcasting. One should notice that there are three different types of messages waiting for broadcasting, namely the common message intended for both end nodes, the private message from node A to node B, and that from node B to node A, respectively.

step 3: Upon receiving, each end node recovers the other's message according the received signal and its own message as side-information.

Note that, in the MAC phase, the rate constraints for the nodes A and B are

given by

$$R_{AB}^{d} \leq \sum_{l=1}^{L} n_l, \tag{4.9}$$

$$R_{BA}^{d} \leq \sum_{l=1}^{L} m_l, \tag{4.10}$$

respectively.

After collecting all the received signals and the received functions, the problem in the BC phase can be regarded as a special broadcast channel where we have one common message and two private messages, one for each direction. Moreover, each private message is known perfectly by the corresponding non-intended receiver as side information. This problem is usually referred to as bidirectional broadcast channel (BBC) with common message whose capacity region has been characterized in [55]. By directly applying the result in [55], one obtains an upper bound on the rate region in the BC phase given by

$$R_{BA}^{d} = R_{C}^{d} + R_{A}^{d} \leq I(\mathbf{X}_R; \mathbf{Y}_B) = \sum_{l=1}^{L} k_l, \tag{4.11}$$

$$R_{AB}^{d} = R_{C}^{d} + R_{B}^{d} \leq I(\mathbf{X}_R; \mathbf{Y}_A) = \sum_{l=1}^{L} r_l, \tag{4.12}$$

where R_C^d, R_A^d, and R_B^d are the rates for the common message, the private message from the node A to B, and that from the node B to A, respectively. In the following lemma, we show that all the rate pairs satisfying (4.9)-(4.12) are indeed achievable. This will complete the proof of Theorem 12

Lemma 13. *For the parallel deterministic bidirectional relay channel, all the rate*

pairs (R_{AB}^d, R_{BA}^d) *satisfying*

$$R_{AB}^d \leq \min\left\{\sum_{l=1}^{L} n_l, \sum_{l=1}^{L} k_l\right\}, \tag{4.13}$$

$$R_{BA}^d \leq \min\left\{\sum_{l=1}^{L} m_l, \sum_{l=1}^{L} r_l\right\}, \tag{4.14}$$

are achievable.

Proof. We start by denoting $R_{MAC,C}^d$, $R_{MAC,A}^d$, and $R_{MAC,B}^d$ as the maximum possible rates for the message sent through aligned bit pipes, message sent through non-aligned bit pipes from the node A, and that from the node B, respectively, in the MAC phase without any coding and any rate allocation. It is easy to see that

$$R_{MAC,C}^d \triangleq \sum_{l=1}^{L} \min(n_l, m_l), \tag{4.15}$$

$$R_{MAC,A}^d \triangleq \sum_{l=1}^{L} (n_l - m_l)^+, \tag{4.16}$$

$$R_{MAC,B}^d \triangleq \sum_{l=1}^{L} (m_l - n_l)^+. \tag{4.17}$$

In what follows, we separately discuss two different cases that contain all the possibilities.

case 1: $(R_{MAC,C}^d \leq \min\left\{\sum_{l=1}^{L} r_l, \sum_{l=1}^{L} k_l\right\})$

In this case, since $R_{MAC,C}^d \leq \min\left\{\sum_{l=1}^{L} r_l, \sum_{l=1}^{L} k_l\right\}$, the BC phase is able to support the transmission of a common message of a rate described in (4.15). Thus, in the MAC phase, the two nodes send bits over all the aligned bit pipes, i.e., $\min(n_l, m_l)$ bits in sub-channel l. This accounts for

$$R_C^d = \sum_{l=1}^{L} \min(n_l, m_l), \tag{4.18}$$

bits of common message (sum of two messages) at the relay. Besides, the relay can still send a private message of a rate up to $\sum_{l=1}^{L} r_l - R_C^d$ bits to the node A and another private message of a rate up to $\sum_{l=1}^{L} k_l - R_C^d$ bits to the node B. According to these constraints in the BC phase together with the rate upper bounds in (4.16) and (4.17), nodes A and B send messages through those non-aligned bit pipes of rates

$$R_A^d = \min\left\{R_{MAC,A}^d, \sum_{l=1}^{L} k_l - R_C^d\right\}, \tag{4.19}$$

$$R_B^d = \min\left\{R_{MAC,B}^d, \sum_{l=1}^{L} r_l - R_C^d\right\}, \tag{4.20}$$

respectively. Therefore, the achievable rate from the node A to B and that from the node B to A are given by

$$R_{AB}^d = R_C^d + \min\left\{R_{MAC,A}^d, \sum_{l=1}^{L} k_l - R_C^d\right\} = \min\left\{\sum_{l=1}^{L} n_l, \sum_{l=1}^{L} k_l\right\}, \tag{4.21}$$

$$R_{BA}^d = R_C^d + \min\left\{R_{MAC,B}^d, \sum_{l=1}^{L} r_l - R_C^d\right\} = \min\left\{\sum_{l=1}^{L} m_l, \sum_{l=1}^{L} r_l\right\}, \tag{4.22}$$

respectively.

case 2: $(R_{MAC,C}^d > \min\left\{\sum_{l=1}^{L} r_l, \sum_{l=1}^{L} k_l\right\})$

In this case, we assume that $\sum_{l=1}^{L} r_l \geq \sum_{l=1}^{L} k_l$ without loss of generality. The other case follows immediately if we switch the role of nodes A and B as we don't specify the relationship between n_l and m_l in the proof. Now, in the BC phase, the relay is unable to support the transmission of a common message of a rate in (4.15) to the node B. Thus, in this case, both nodes only use the same $\sum_{l=1}^{L} k_l$ bit pipes

belonging to the aligned part. This accounts for

$$R_C^d = \sum_{l=1}^{L} k_l, \tag{4.23}$$

bits of common message at the relay. After this, since $\sum_{l=1}^{L} r_l \geq R_C^d$ the relay can still send a private message of a rate up to $\sum_{l=1}^{L} r_l - R_C^d$ bits to the node A. Also, in the MAC phase, since both nodes only uses R_C^d bits in the aligned part, there are still $R_{MAC,C}^d - R_C^d$ bits that can be used by node B. Therefore, node B sends a message through those non-aligned bit pipes and the remaining aligned bit pipes of a rate

$$R_B^d = \min\left\{R_{MAC,B}^d + R_{MAC,C}^d - R_C^d, \sum_{l=1}^{L} r_l - R_C^d\right\}. \tag{4.24}$$

Therefore, the achievable rate from node A to B and that from node B to A are given by

$$R_{AB}^d = R_C^d = \sum_{l=1}^{L} k_l = \min\left\{\sum_{l=1}^{L} n_l, \sum_{l=1}^{L} k_l\right\}, \tag{4.25}$$

$$R_{BA}^d = R_C^d + \min\left\{R_{MAC,B}^d + R_{MAC,C}^d - R_C^d, \sum_{l=1}^{L} r_l - R_C^d\right\} = \min\left\{\sum_{l=1}^{L} m_l, \sum_{l=1}^{L} r_l\right\}, \tag{4.26}$$

respectively. □

4.3 Parallel Gaussian Bidirectional Relay Channel

In this section, we study the parallel Gaussian bidirectional relay channel based on the insight learned from the corresponding linear deterministic model. The scheme in Section 4.2.3 suggests independently encoding among sub-channels in the MAC phase and jointly encoding across sub-channels in the BC phase. According to this

insight, a coding scheme is proposed which can be regarded as an extension of the lattice alignment scheme proposed in [12]. We then show that the proposed scheme achieves the capacity region of the parallel Gaussian bidirectional relay channel to within L bits.

4.3.1 Channel Model

The parallel Gaussian bidirectional relay channel shown in Fig. 4.1 consists of three nodes A, B, and R where A and B wish to exchange information through a relay node R between them via a set of L parallel bidirectional relay channels. The channel coefficients are assumed to be fixed in each sub-channel but may vary from one sub-channel to another. As shown in Fig. 4.1, in the parallel Gaussian bidirectional relay channel, each sub-channel $l \in \{1, 2, \ldots, L\}$ can be modeled as a bidirectional relay channel with sub-nodes A_l, B_l, and R_l. In addition, since we wish to model the ISI and the MIMO channels, joint processing (such as joint coding) across sub-channels is permitted. During the l^{th} sub-channel, the channel gains from nodes A_l and B_l to R_l are denoted as $h_{Al} \in \mathbb{C}$ and $h_{Bl} \in \mathbb{C}$, respectively, and that from R_l to the nodes are denoted as $g_{Al} \in \mathbb{C}$ and $g_{Bl} \in \mathbb{C}$, respectively. Each node is assumed to have the global channel knowledge, i.e., global CSIT is assumed.

During the l^{th} sub-channel, nodes A and B respectively map the messages $w_A \in \{1, 2, \ldots, M_A\}$ and $w_B \in \{1, 2, \ldots, M_B\}$, independent of each other and of all other sub-channels' messages, to length-N codewords $\mathbf{x_{Al}} = \mathcal{E}_{Al}^N(w_A)$ and $\mathbf{x_{Bl}} = \mathcal{E}_{Bl}^N(w_B)$. Each sub-node is subject to an individual power constraint P. The transmission protocol we consider is a two phase protocol consisting of a MAC phase and a BC phase. Each of phases occupies a half of channel uses and is assumed to be orthogonal to each other.

In the MAC phase, both nodes transmit their signals to the relay simultaneously

and the relay keeps silent. The received signal at R_l is then given by

$$\mathbf{y}_{Rl} = h_{Al}\mathbf{x}_{Al} + h_{Bl}\mathbf{x}_{Bl} + \mathbf{z}_{Rl}, \tag{4.27}$$

$\mathbf{z}_{Rl} \sim \mathcal{CN}(0, \mathbf{I})$ be i.i.d. Gaussian noise. The relay first collects all $\mathbf{y}_{Rl}$ for $l \in \{1, 2, \ldots, L\}$ and each R_l generates the transmitted signal $\mathbf{x}_{Rl} = \mathcal{E}_{Rl}^N(\mathbf{y}_R)$.

In the BC phase, each R_l broadcasts $\mathbf{x}_{Rl}$ back to A_l and B_l and both nodes keep silent. The received signal at end nodes are given by

$$\mathbf{y}_{Al} = g_{Al}\mathbf{x}_{Rl} + \mathbf{z}_{Al}, \tag{4.28}$$

$$\mathbf{y}_{Bl} = g_{Bl}\mathbf{x}_{Rl} + \mathbf{z}_{Bl}, \tag{4.29}$$

respectively, where again $\mathbf{z}_{Al}, \mathbf{z}_{Bl} \sim \mathcal{CN}(0, \mathbf{I})$. Nodes A and B then collect its received signals to form $\mathbf{y}_A$ and $\mathbf{y}_B$. After that, estimates $\hat{w}_B = \mathcal{G}_A^N(\mathbf{y}_A, \mathbf{x}_A)$ and $\hat{w}_A = \mathcal{G}_B^N(\mathbf{y}_B, \mathbf{x}_B)$ are formed at nodes A and B, respectively.

Here, the codes, the achievable rate region, and the capacity region can be defined in a similar way as those in Section 4.2.1. Clearly, one can again have an upper bound on the capacity region by using the cut-set argument [30] as

Lemma 14. *$\mathcal{C} \subseteq \overline{\mathcal{C}}$, where*

$$\begin{aligned}
\overline{\mathcal{C}} = \Bigg\{ & (R_{AB}, R_{BA}) : 0 \le R_{AB} \le \overline{C}_{AB}, 0 \le R_{BA} \le \overline{C}_{BA}, \\
& \overline{C}_{AB} = \min \left\{ \sum_{l=1}^{L} \log\left(1 + P|h_{Al}|^2\right), \sum_{l=1}^{L} \log\left(1 + P|g_{Bl}|^2\right) \right\}, \\
& \overline{C}_{BA} = \min \left\{ \sum_{l=1}^{L} \log\left(1 + P|h_{Bl}|^2\right), \sum_{l=1}^{L} \log\left(1 + P|g_{Al}|^2\right) \right\} \Bigg\}.
\end{aligned} \tag{4.30}$$

The main result in this section is stated in the following theorem which approxi-

mately characterizes the capacity region of the parallel Gaussian bidirectional relay channel.

Theorem 15. *The capacity region of the parallel Gaussian bidirectional relay channel satisfies* $\overline{\mathcal{C}} - L \subseteq \mathcal{C} \subseteq \overline{\mathcal{C}}$.

It is worth noting that the approximation only depends on the number of sub-channels and is independent of specific channel realizations.

Again, one can define the exchange rate and exchange capacity for the Gaussian bidirectional relay channel as follows.

Definition 21*:* An exchange rate R_{ex} is achievable if (R_{ex}, R_{ex}) lies inside the capacity region $\mathcal{C}$. The exchange capacity C_{ex} is then defined as the supremum of all achievable exchange rate.

4.3.2 Proposed Scheme - Lattice Alignment Scheme + Coding Across Sub-channels

As suggested by the linear deterministic model, in the MAC phase, we independently encode signals for each sub-channel in the sense that no coding across sub-channels is used. For the l^{th} sub-channel, the weak node transmits a lattice codeword only and the strong node transmits the superposition of a random codeword and a lattice codeword. In addition, the lattice codewords are chosen from an identical nested lattice code and are perfect aligned at the relay. Upon receiving, the relay decodes the non-aligned codeword (corresponding to the non-aligned bits in the linear deterministic model) and decodes the lattice function (corresponds to the XOR bits in the linear deterministic model). In the BC phase, after collecting all the decoded codewords and functions, the relay performs coding across sub-channels as suggested by the linear deterministic model.

In general, in the MAC phase, the transmitted signals of the proposed scheme at

the l^{th} sub-channel from nodes A and B are respectively given by

$$x_{Al} = \sqrt{\alpha_{Al}} x_{Al}^{(1)} + \sqrt{1-\alpha_{Al}} x_{Al}^{(2)}, \tag{4.31}$$

$$x_{Bl} = \sqrt{\alpha_{Bl}} x_{Bl}^{(1)} + \sqrt{1-\alpha_{Bl}} x_{Bl}^{(2)}, \tag{4.32}$$

where $x_{Al}^{(1)}$ and $x_{Bl}^{(1)}$ are codewords chosen from an identical nested lattice code with codebook size $2^{nR_{Al}^{(1)}} = 2^{nR_{Bl}^{(1)}}$. Moreover, $x_{Al}^{(2)}$ and $x_{Bl}^{(2)}$ are codewords chosen from random codes with codebook size $2^{nR_{Al}^{(2)}}$ and $2^{nR_{Bl}^{(2)}}$, respectively.

At the relay, each sub-channel always decodes random Gaussian codewords first, subtracts the decoded codewords, and then computes lattice functions. Note that the order of decoding at the relay matters. It is because that lattice functions contain information for both directions; hence, one would decode it last as this would make the equivalent channel for the aligned part to be as clean as possible.

As in Section 4.2, the resulted problem in the BC phase can be thought of as a set of parallel broadcast channels with a common message (of a rate R_C) and two private messages (of rates R_A and R_B) where each private message is exactly known at the corresponding non-intended receiver; i.e., a BBC with a common message. We then directly apply the result in [55] to the parallel Gaussian bidirectional relay channel and get the following rate constraints in the BC phase as

$$R_{AB} = R_C + R_A \leq \sum_{l=1}^{L} \log\left(1 + P|g_{Al}|^2\right), \tag{4.33}$$

$$R_{BA} = R_C + R_B \leq \sum_{l=1}^{L} \log\left(1 + P|g_{Bl}|^2\right). \tag{4.34}$$

In the following lemma, we show that there exists a power allocation strategy so that the achievable rate region of the proposed scheme approaches the capacity

region to within L bits. This will complete the proof of Theorem 15.

Lemma 16. *For a parallel Gaussian bidirectional relay channel, there always exist α_{Al} and α_{Bl} for the proposed scheme such that the resulted rate region achieves the capacity region to within L bits.*

Proof. We start by defining

$$R_{MAC,C} \triangleq \sum_{l=1}^{L} \log\left(\frac{1}{2} + P\min(|h_{Al}|^2, |h_{Bl}|^2)\right)^+, \tag{4.35}$$

In what follows, we separately discuss two different cases that contain all the possibilities.

case 1: $(R_{MAC,C} \leq \min\left\{\sum_{l=1}^{L}\log(1+P|g_{Al}|^2), \sum_{l=1}^{L}\log(1+P|g_{Bl}|^2)\right\})$ In this case, the two nodes use the power allocation in the MAC phase described as follows.

For the sub-channels having a equal gain, i.e., $|h_{Al}|^2 = |h_{Bl}|^2$, we set $\alpha_{Al} = 1$, $\alpha_{Bl} = 1$, $R_{Al}^{(2)} = 0$, and $R_{Bl}^{(2)} = 0$. The relay directly decodes the received signal to the modulo-sum of two lattice codewords (see Chapter 2 for details.) This results in

$$\begin{aligned} R_{Al}^{(1)} = R_{Bl}^{(1)} &\leq \log\left(\frac{1}{2} + P|h_{Al}|^2\right)^+ \\ &= \left(\log\left(1+2P|h_{Al}|^2\right) - 1\right)^+ \\ &\geq \log\left(1+P|h_{Al}|^2\right) - 1. \end{aligned} \tag{4.36}$$

For the sub-channels with $|h_{Al}|^2 > |h_{Bl}|^2$, node A can have one extra codeword in the MAC phase. We therefore set $\alpha_{Bl} = 1$ and $R_{Bl}^{(2)} = 0$ and choose the power allocation at node A such that the two lattices are perfectly aligned at the relay as

$$\alpha_{Al} = \frac{|h_{Bl}|^2}{|h_{Al}|^2}. \tag{4.37}$$

The relay first decodes the extra codeword $\mathbf{x}_{Al}^{(2)}$ by treating the lattice part as noise. This results in

$$R_{Al}^{(2)} \leq \log\left(1 + \frac{P|h_{Al}|^2(1-\alpha_{Al})}{1+2P|h_{Bl}|^2}\right). \tag{4.38}$$

It then subtracts the decoded codeword and tries to decode the lattice function. This gives us

$$R_{Al}^{(1)} = R_{Bl}^{(1)} \leq \log\left(\frac{1}{2} + P|h_{Bl}|^2\right)^+. \tag{4.39}$$

Notice that one can bound the sum rates at the each sub-channel as

$$R_{Bl}^{(1)} + R_{Bl}^{(2)} \geq \log\left(1 + P|h_{Bl}|^2\right) - 1, \tag{4.40}$$

and

$$\begin{aligned} R_{Al}^{(1)} + R_{Al}^{(2)} &= \log\left(\frac{1}{2} + P|h_{Bl}|^2\right)^+ + \log\left(1 + \frac{P|h_{Al}|^2(1-\alpha_{Al})}{1+2P|h_{Bl}|^2}\right) \\ &\geq \log\left(1 + 2P|h_{Bl}|^2 + P|h_{Al}|^2(1-\alpha_{Al})\right) - 1 \\ &= \log\left(1 + P|h_{Al}|^2 + P|h_{Bl}|^2\right) - 1 \\ &\geq \log\left(1 + P|h_{Al}|^2\right) - 1. \end{aligned} \tag{4.41}$$

Similarly, for the sub-channels with $|h_{Al}|^2 < |h_{Bl}|^2$, we switch the role of nodes A and B and hence we have $\alpha_{Al} = 1$, $R_{Al}^{(2)} = 0$, and

$$\alpha_{Bl} = \frac{|h_{Al}|^2}{|h_{Bl}|^2}. \tag{4.42}$$

For this case, the relay again first decodes the extra codeword $\mathbf{x}_{Bl}^{(2)}$ by treating the lattice part as noise, subtracts the decoded codeword out, and then computes the

lattice function. Thus, one has

$$R_{Bl}^{(2)} \leq \log\left(1 + \frac{P|h_{Bl}|^2(1-\alpha_{Bl})}{1+2P|h_{Al}|^2}\right), \tag{4.43}$$

and

$$R_{Al}^{(1)} = R_{Bl}^{(1)} \leq \log\left(\frac{1}{2} + P|h_{Al}|^2\right)^+. \tag{4.44}$$

Again, notice that one can bound the sum rates as

$$R_{Al}^{(1)} + R_{Al}^{(2)} \geq \log\left(1 + P|h_{Al}|^2\right) - 1, \tag{4.45}$$

and

$$\begin{aligned} R_{Bl}^{(1)} + R_{Bl}^{(2)} &= \log\left(\frac{1}{2} + P|h_{Al}|^2\right)^+ + \log\left(1 + \frac{P|h_{Bl}|^2(1-\alpha_{Bl})}{1+2P|h_{Al}|^2}\right) \\ &\geq \log\left(1 + 2P|h_{Al}|^2 + P|h_{Bl}|^2(1-\alpha_{Bl})\right) - 1 \\ &= \log\left(1 + P|h_{Al}|^2 + P|h_{Bl}|^2\right) - 1 \\ &\geq \log\left(1 + P|h_{Bl}|^2\right) - 1. \end{aligned} \tag{4.46}$$

In the BC phase, we have a common message (lattice functions) at the relay of

$$R_C = \sum_{l=1}^{L} R_{Al}^{(1)} = \sum_{l=1}^{L} R_{Bl}^{(1)} = R_{MAC,C}, \tag{4.47}$$

bits. After this, the relay can still send a private message of a rate up to $\sum_{l=1}^{L}\log(1+P|g_{Al}|^2) - R_C$ bits to node A and another private message of a rate up to $\sum_{l=1}^{L}\log(1+P|g_{Bl}|^2) - R_C$ bits to node B. According to these constraints in the BC phase together with the rate constraints in the MAC phase, the private messages (Gaussian

codewords) sent from nodes A and B can have rates up to

$$R_A = \min\left\{\sum_{l=1}^{L} R_{Al}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Bl}|^2) - R_C\right\}, \tag{4.48}$$

$$R_B = \min\left\{\sum_{l=1}^{L} R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2) - R_C\right\}, \tag{4.49}$$

respectively. Therefore, the achievable rates from node A to B is given by

$$\begin{aligned}
R_{AB} &= R_C + R_A \\
&= R_C + \min\left\{\sum_{l=1}^{L} R_{Al}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Bl}|^2) - R_C\right\} \\
&= \min\left\{\sum_{l=1}^{L} R_{Al}^{(1)} + R_{Al}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Bl}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Al}|^2) - 1, \sum_{l=1}^{L} \log(1+P|g_{Bl}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Al}|^2), \sum_{l=1}^{L} \log(1+P|g_{Bl}|^2)\right\} - L,
\end{aligned} \tag{4.50}$$

and that from node B to A is given by

$$\begin{aligned}
R_{BA} &= R_C + R_B \\
&= R_C + \min\left\{\sum_{l=1}^{L} R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2) - R_C\right\} \\
&= \min\left\{\sum_{l=1}^{L} R_{Bl}^{(1)} + R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Bl}|^2) - 1, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Bl}|^2), \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} - L.
\end{aligned} \tag{4.51}$$

case 2: $(R_{MAC,C} > \min\left\{\sum_{l=1}^{L}\log(1+P|g_{Al}|^2), \sum_{l=1}^{L}\log(1+P|g_{Bl}|^2)\right\})$

Without loss of generality, we assume that $\sum_{l=1}^{L}\log(1+P|g_{Al}|^2) \geq \sum_{l=1}^{L}\log(1+P|g_{Bl}|^2)$. Otherwise, one can switch the role of nodes A and B. In this case, we observe that the downlink channel is not able to support the whole transmission of the lattice function if we still use the same power allocation strategy as that in the previous case. Thus, one has to modify the power allocation strategy as suggested by the deterministic model. In what follows, we propose a power allocation strategy that can again achieve the cut-set bound to within L bits. In our strategy, in some sub-channels, the nodes may have to reduce the power assigned to the lattice codewords in order to make the resultant $\sum_{l=1}^{L}R_{Al}^{(1)} = \sum_{l=1}^{L}R_{Bl}^{(1)} = \sum_{l=1}^{L}\log(1+P|g_{Bl}|^2)$. Moreover, for those sub-channels which reduce the power for the lattice codewords, node B may increase the power assigned to the Gaussian codewords to somehow compensate the rate loss. On the other hand, node A cannot do that simply because the downlink will not be able to support any rate more than $\sum_{l=1}^{L}\log(1+P|g_{Bl}|^2)$ to node B.

Let $0 \leq \delta_l \leq P$. For sub-channels with $|h_{Al}|^2 \geq |h_{Bl}|^2$, we set $\alpha_{Al} = \frac{P-\delta_l}{P}\frac{|h_{Bl}|^2}{|h_{Al}|^2}$, $\alpha_{Bl} = \frac{P-\delta_l}{P}$, and $R_{Al}^{(2)} = 0$ (i.e., we don't send $\mathbf{x}_{Al}^{(2)}$). The relay first decode the $\mathbf{x}_{Bl}^{(2)}$, subtracts it out, and then decodes the lattice function. This leads to

$$R_{Bl}^{(2)} = \log\left(1+\frac{\delta_l|h_{Bl}|^2}{1+2(P-\delta_l)|h_{Bl}|^2}\right), \tag{4.52}$$

and

$$R_{Al}^{(1)} = R_{Bl}^{(1)} = \log\left(\frac{1}{2}+(P-\delta_l)|h_{Bl}|^2\right)^{+}. \tag{4.53}$$

Notice that one can bound the sum rate of node B as

$$\begin{aligned} R_{Bl}^{(1)} + R_{Bl}^{(2)} &= \log\left(\frac{1}{2} + (P-\delta_l)|h_{Bl}|^2\right)^+ + \log\left(1 + \frac{\delta_l|h_{Bl}|^2}{1+2(P-\delta_l)|h_{Bl}|^2}\right) \\ &\geq \log\left(1 + 2(P-\delta_l)|h_{Bl}|^2 + \delta_l|h_{Bl}|^2\right) - 1 \\ &= \log\left(1 + P|h_{Bl}|^2 + (P-\delta_l)|h_{Bl}|^2\right) - 1 \\ &\geq \log\left(1 + P|h_{Bl}|^2\right) - 1. \end{aligned} \tag{4.54}$$

For sub-channels with $|h_{Al}|^2 < |h_{Bl}|^2$, we set $\alpha_{Al} = \frac{P-\delta_l}{P}$, $\alpha_{Bl} = \frac{P-\delta_l}{P}\frac{|h_{Al}|^2}{|h_{Bl}|^2}$, and $R_{Al}^{(2)} = 0$ (i.e., we again don't send $\mathbf{x}_{Al}^{(2)}$). The relay first decode the $\mathbf{x}_{Bl}^{(2)}$, subtracts it out, and then decodes the lattice function. This leads to

$$R_{Bl}^{(2)} = \log\left(1 + \frac{P|h_{Bl}|^2 - P|h_{Al}|^2 + \delta_l|h_{Al}|^2}{1+2(P-\delta_l)|h_{Al}|^2}\right), \tag{4.55}$$

and

$$R_{Al}^{(1)} = R_{Bl}^{(1)} = \log\left(\frac{1}{2} + (P-\delta_l)|h_{Al}|^2\right)^+. \tag{4.56}$$

Notice that again one can bound the sum rate of node B as

$$\begin{aligned} R_{Bl}^{(1)} + R_{Bl}^{(2)} &= \log\left(\frac{1}{2} + (P-\delta_l)|h_{Al}|^2\right)^+ + \log\left(1 + \frac{P|h_{Bl}|^2 - P|h_{Al}|^2 + \delta_l|h_{Al}|^2}{1+2(P-\delta_l)|h_{Al}|^2}\right) \\ &\geq \log\left(1 + (P-\delta_l)|h_{Al}|^2 + P|h_{Bl}|^2\right) - 1 \\ &\geq \log\left(1 + P|h_{Bl}|^2\right) - 1. \end{aligned} \tag{4.57}$$

Now we choose δ_ls such that

$$\sum_{l=1}^{L} R_{Al}^{(1)} = \sum_{l=1}^{L} R_{Bl}^{(1)} = \sum_{l=1}^{L} \log(1 + P|g_{Bl}|^2). \tag{4.58}$$

In the BC phase, one has

$$R_C = \sum_{l=1}^{L} R_{Al}^{(1)} = \sum_{l=1}^{L} R_{Bl}^{(1)} = \sum_{l=1}^{L} \log(1 + P|g_{Bl}|^2), \tag{4.59}$$

bits of common message (lattice functions) at the relay. After this, the relay can still send a private message of a rate up to $\sum_{l=1}^{L} \log(1 + P|g_{Al}|^2) - R_C$ bits to node A. According to this constraint in the BC phase together with the constraints in the MAC phase, the private message sent from node B that is correctly decodable at node A can have a rate up to

$$R_B = \min \left\{ \sum_{l=1}^{L} R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1 + P|g_{Al}|^2) - R_C \right\}. \tag{4.60}$$

Therefore, the achievable rates from node A to B is given by

$$\begin{aligned} R_{AB} = R_C &= \sum_{l=1}^{L} \log(1 + P|g_{Bl}|^2) \\ &\geq \min \left\{ \sum_{l=1}^{L} \log(1 + P|h_{Al}|^2), \sum_{l=1}^{L} \log(1 + P|g_{Bl}|^2) \right\} - L, \end{aligned} \tag{4.61}$$

and that from node B to A is given by

$$\begin{aligned}
R_{BA} &= R_C + R_B \\
&= R_C + \min\left\{\sum_{l=1}^{L} R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2) - R_C\right\} \\
&= \min\left\{\sum_{l=1}^{L} R_{Bl}^{(1)} + R_{Bl}^{(2)}, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Bl}|^2) - 1, \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} \\
&\geq \min\left\{\sum_{l=1}^{L} \log(1+P|h_{Bl}|^2), \sum_{l=1}^{L} \log(1+P|g_{Al}|^2)\right\} - L. \qquad (4.62)
\end{aligned}$$

□

4.4 Discussions

In this section, we first propose another coding scheme that can be regarded as a natural extension of the scheme in [13] to the parallel setting. We also study a special case of the parallel Gaussian bidirectional relay channel, namely the reciprocal case. i.e., the channel reciprocity is assumed. For this case, we show that routing across sub-channels is optimal for the corresponding linear deterministic model and propose a (conceptually) easier coding scheme for the Gaussian model. In addition, we also compare the proposed schemes with some existing schemes.

4.4.1 Extension of Lattice Partition Chain Scheme

For the parallel Gaussian bidirectional relay channel, once we realize that separate coding in the MAC phase and joint coding across sub-channel in the BC phase may be a right approach, it is natural to extend the lattice partition chain scheme proposed by Nam *et al.* [13] to the parallel setting. This has been done independently in [48] and in the conference version of this work [56]. We summarize the coding scheme

and result in the following and refer this coding scheme to as the proposed scheme 2.

In the MAC phase, again following the linear deterministic model, for each sub-channel, we independently encode the message by a sequence of lattice partition chains proposed by Nam *et al.* [13]. The relay then decodes and re-encodes the computed functions by a random codebook that performs joint coding across sub-channels.

Specifically, in the MAC phase, at the l^{th} sub-channel, any computation rate satisfies the following is achievable

$$R_{Al,Nam} \leq \log\left(\frac{|h_{Al}|^2}{|h_{Al}|^2+|h_{Bl}|^2}+P|h_{Al}|^2\right)^+, \tag{4.63}$$

$$R_{Bl,Nam} \leq \log\left(\frac{|h_{Bl}|^2}{|h_{Al}|^2+|h_{Bl}|^2}+P|h_{Bl}|^2\right)^+. \tag{4.64}$$

The total achievable rates in the MAC phase from the nodes A and B to the relay are given by

$$R_{AB,Nam} \leq \sum_{l=1}^{L} R_{Al,Nam}, \tag{4.65}$$

$$R_{BA,Nam} \leq \sum_{l=1}^{L} R_{Bl,Nam}, \tag{4.66}$$

respectively.

The relay collects all the decoded functions and performs joint coding across

sub-channels in the BC phase. This results in

$$R_{BA,Nam} \leq \sum_{l=1}^{L} \log\left(1 + P|g_{Al}|^2\right), \tag{4.67}$$

$$R_{AB,Nam} \leq \sum_{l=1}^{L} \log\left(1 + P|g_{Bl}|^2\right). \tag{4.68}$$

Both nodes then identify the other's message from the received signals and their own message as side information. The achievable rate region of this coding scheme is then given by

$$R_{AB,Nam} \leq \min\left\{\sum_{l=1}^{L} \log\left(\frac{|h_{Al}|^2}{|h_{Al}|^2 + |h_{Bl}|^2} + P|h_{Al}|^2\right)^+, \sum_{l=1}^{L} \log\left(1 + P|g_{Bl}|^2\right)\right\}, \tag{4.69}$$

$$R_{BA,Nam} \leq \min\left\{\sum_{l=1}^{L} \log\left(\frac{|h_{Bl}|^2}{|h_{Al}|^2 + |h_{Bl}|^2} + P|h_{Bl}|^2\right)^+, \sum_{l=1}^{L} \log\left(1 + P|g_{Al}|^2\right)\right\}. \tag{4.70}$$

Following the proof in [13], it is easy to verify that this scheme also achieves the capacity region to within L bits. This provides an alternative proof of the achievability part of Theorem 15.

4.4.2 Reciprocal Case

Under the channel reciprocity assumption, the capacity region is fully characterized by the exchange capacity, C_{ex}^{d} and C_{ex} for deterministic and Gaussian models, respectively.

For the linear deterministic model with channel reciprocity, we note that an upper

bound on C_{ex}^{d} can be obtained from Lemma 11 as

$$\overline{C}_{ex}^{d}, = \min\left(\sum_{l=1}^{L} n_l, \sum_{l=1}^{L} m_l\right), \tag{4.71}$$

per two channel uses (one for each phase.) The main result on the reciprocal case is stated in the following theorem.

Theorem 17. *(*Reciprocal case*) For reciprocal channels, the exchange capacity (or capacity region) of the parallel deterministic bidirectional relay channel is equal to the cut-set upper bound given in* (4.71)*. Furthermore, the exchange capacity can be achieved by a simple routing scheme.*

We postpone the proof to Appendix B.1 but only outline the coding scheme. A small example is also provided to demonstrate the main idea behind the coding scheme.

In the following, we outline the proposed coding scheme, which is nothing but routing.

step 1: In the MAC phase, A_l and B_l send n_l and m_l independent data streams to R_l through its n_l and m_l MSB, respectively.

step 2: Each R_l shifts the aligned part to the MSB because it contains information intended for both directions. Note that this step is always possible because of the channel reciprocity assumption.

step 3: For the non-aligned part, the relay *reorders* the signals to other sub-channels that can support the transmission to the desired destination.

Example 2. Let us consider the example with two sub-channels (i.e., $L = 2$) given in Fig. 4.3 where n_1, n_2, m_1, and m_2 are equal to $4, 1, 2$, and 3, respectively. Therefore, the cut-set upper bound for this example is 5 bits per two channel uses (one for

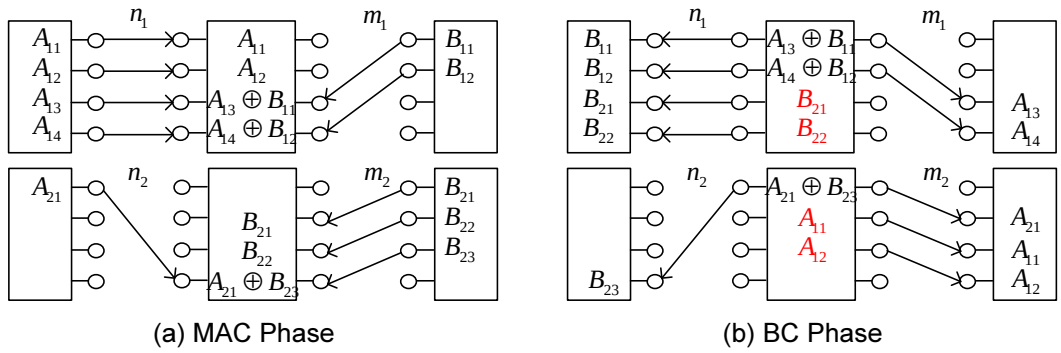

Figure 4.3: An example of a linear deterministic model of the considered setup.

each phase). To achieve this upper bound, in the MAC phase, A_l sends $A_{l1}, \ldots, A_{ln_l}$ independent bits to R_l through its n_l MSB. Similarly, B_l sends $[B_{l1}, \ldots, B_{lm_l}]$ independent bits to R_l through its m_l MSB bits.

The received signals at R_1 contains an aligned part $A_{13} \oplus B_{11}, A_{14} \oplus B_{12}$ and an non-aligned part A_{11}, A_{12}. Likewise, we have an aligned part $A_{21} \oplus B_{23}$ and a non-aligned part B_{21}, B_{22} at R_2. At each sub-channel, the relay first shifts the aligned part to its MSB. After this, the relay still needs to arrange broadcasting A_{11}, A_{12} to the node B and B_{21}, B_{22} to the node A. It then *reorders* those remaining bits across sub-channels such that it will broadcast A_{11}, A_{12} at the second sub-channel and broadcast B_{21}, B_{22} at the first sub-channel. Each node then cancels out their own messages for the aligned part and recovers 5 bits from the other node. $\Diamond$

For the Gaussian model with channel reciprocity, we again note that for the exchange capacity, the upper bound in Lemma 14 reduces to

$$\overline{C}_{ex} = \min \left\{ \sum_{l=1}^{L} \log \left(1 + P|h_{Al}|^2\right), \sum_{l=1}^{L} \log \left(1 + P|h_{Bl}|^2\right) \right\}. \tag{4.72}$$

According to the coding scheme for the corresponding deterministic model, we propose a coding scheme (referred to as the proposed scheme 3) that is conceptually

easier than the coding scheme in 4.3.2 in the sense that no joint coding across sub-channels are required. We then show that this coding scheme is again able to approach the upper bound to within L bit as well. Since this coding is motivated by the insight obtained from the reordering scheme described above, it works only under the channel reciprocity assumption.

In the MAC phase, we use the same lattice alignment scheme as that in the proposed scheme 1. After correctly decoding signals and functions at the relay, due to the channel reciprocity, the lattice functions are guaranteed to be broadcasted within the same sub-channel. For the codewords that correspond to the non-aligned part, as suggested by the linear deterministic model, the relay collects all the non-aligned codewords and then reallocates bits to each sub-channel. This bit reallocation scheme mimics the reordering scheme in the deterministic model to the Gaussian scenario.

In the BC phase, each sub-channel now independently considers the problem as a corresponding BBC problem, i.e., separate coding, with a common message corresponding to the lattice function received in this sub-channel and a private message corresponding to the bits assigned to this sub-channel.

The performance of this scheme for the reciprocal case is given as follows.

Theorem 18. *The proposed bit reallocation scheme achieves the exchange capacity to within L bits.*

Proof. See Appendix B.2. □

Remark: From the above result, one can see that under the assumption of channel reciprocity, simple joint operations across sub-channels such as bit reallocation together with an optimal separate coding in the BC phase is able to achieve the exchange capacity to within a constant gap. This allows us to conceptually simplify the proposed scheme as joint coding across sub-channels is not performed. However,

for the extension of lattice partition chain scheme, it is not clear to us how to exploit the channel reciprocity to further simplify the scheme.

Remark: It is worth noting that the proposed bit reallocation scheme here is different from the codeword reallocation scheme in the conference version of this chapter [56]. In [56], for the reciprocal case, the relay reorders the whole codeword to another sub-channel such that the induced rate mismatch is minimized. The codeword allocation problem is then formulated as a weighted bipartite matching problem and the solution is given by the Hungarian method. This coding scheme is in fact not able to approximate the exchange capacity to within a constant bit. It is due to the fact that when adopting the codeword allocation scheme, we are not allowed to split the private messages and this results in a gap that is dependent of channel parameters.

4.4.3 Comparisons

Remark: (*Single sub-channel case*) When $L = 1$, the problem reduces to the conventional bidirectional relay channel. We first compare the proposed scheme with the one proposed by Avestimehr *et al.* since their coding scheme was inspired by the linear deterministic model also. In [50], they showed that their coding scheme that adopts superposition of unstructured codes approaches the capacity region to within 3 bits. When $L = 1$, our proposed scheme 1 can be regarded as a modification of their scheme that replaces the coding scheme for the aligned part by a structured code. As shown in Theorem 15, the proposed scheme which is a mixture of structured and random schemes guarantees to approach the capacity region to within 1 bit, which is the same as the completely structured one in [13] and the completely random one in [46].

Remark: (*Comparison to the scheme in [47]*) In [47], Khina *et al.* study the

achievable exchange rate of the MIMO bidirectional relay channel and propose a coding scheme based on the joint matrix decomposition under the assumption that $\det \mathbf{H}_A\mathbf{H}_A^* = \det \mathbf{H}_B\mathbf{H}_B^* = 1$ and $\mathbf{H}_A, \mathbf{H}_B$ are full rank. This joint matrix decomposition allows us converting the MIMO bidirectional relay channel into a parallel bidirectional relay channel with *equal channel gains* at each sub-channel, i.e., $|h_{Al}|^2 = |h_{Bl}|^2$. Thus, one can directly carry out the nested lattice coding scheme in [12] for compute-and-forward at each sub-channel. They also show that this scheme achieves an asymptotically optimal exchange rate equal to $N_r \log\left(\frac{P_T}{N_r}\right)$ where N_r is the number of receive antenna and P_T is the total power constraint at each node.

Now, let's compare our proposed schemes to the joint matrix decomposition scheme. We first notice that a parallel bidirectional relay channel with channel gains $h_{A1}, \ldots, h_{AL}$ and $h_{B1}, \ldots, h_{BL}$ can be thought of as a MIMO bidirectional relay channel with L transmit antenna and L receive antenna and whose channel matrices are $\mathbf{H}_A = \mathrm{diag}(h_{A1}, \ldots, h_{AL})$ and $\mathbf{H}_B = \mathrm{diag}(h_{B1}, \ldots, h_{BL})$. The assumption that two MIMO channels are full rank is equivalent to assuming that channel gains of both directions are non-zero for all sub-channels, i.e., $h_{Al} \neq 0$ and $h_{Bl} \neq 0$ for all l. Moreover, the assumption of $\det \mathbf{H}_A\mathbf{H}_A^* = \det \mathbf{H}_B\mathbf{H}_B^* = 1$ is equivalent to that

$$\prod_{l=1}^{L} |h_{Al}|^2 = \prod_{l=1}^{L} |h_{Bl}|^2 = 1. \tag{4.73}$$

With these two assumptions, we now take a closer look at the performance achieved in the MAC phase of the proposed scheme 2. Note that since we consider the asymptotic (in P) behavior, we can assume all the terms inside the logarithm is greater than 1;

i.e., we can ignore the $(\cdot)^+$ sign. One has that

$$\begin{aligned} R_{AB} &\geq \sum_{l=1}^{L} \log\left(1 + P|h_{Al}|^2\right) - 1 \\ &\geq \log\left(\prod_{l=1}^{L} P|h_{Al}|^2\right) - L = L\log\left(\frac{LP}{L}\right) - L \\ &\stackrel{P\to\infty}{=} L\log\left(\frac{LP}{L}\right), \end{aligned} \tag{4.74}$$

and similarly for R_{BA}. This implies that the proposed scheme is asymptotically optimal when P approach ∞ and is as good as the one in [47] in the asymptotic regime when the corresponding $\mathbf{H}_A$ and $\mathbf{H}_B$ are full-rank.

More importantly, the joint matrix decomposition scheme cannot handle the channels which are not full rank. However, the proposed scheme is able to work under non-full rank channel matrices (i.e., channel gains for some sub-channels may be 0.) One may think that it is unfair to compare the joint matrix decomposition scheme with ours since they start by a MIMO problem and convert it into a parallel setup, while we start by assuming that a parallel setup has been created. However, in the following extreme example, it can easily be verified that our scheme is optimal in the MAC phase while the joint matrix decomposition fails to work.

Example 3. (*Example in [47]*) Consider a pair of channel matrices as

$$\mathbf{H}_A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.75}$$

and

$$\mathbf{H}_B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.76}$$

For this example, the proposed scheme reduces to the decode-and-forward and is optimal in the MAC phase while the joint matrix decomposition fails to work. $\Diamond$

This points out that when one of or both channel matrices are not full rank, the joint matrix decomposition scheme does not work while our proposed scheme can still work if the corresponding parallel setup can somehow be created.

4.4.4 Numerical Results

In this section, we provide a numerical example to compare the performance (in terms of the exchange rate) of the proposed scheme, the extension of lattice partition chain scheme, the codeword allocation scheme proposed in [56], and the pre-filtering scheme proposed in [17]. We also plot the cut-set bound in (4.72) as an upper bound and the decode-and-forward scheme with an optimal power allocation as a baseline scheme. The channel parameters are set to be $\mathbf{h_A} = [9,\ 1,\ 5,\ 4]$, $\mathbf{h_B} = [8,\ 4,\ 2,\ 2]$, and $L = 4$.

In Fig. 4.4, one observes that the decode-and-forward scheme trying to decode individual messages at the relay is very sub-optimal in the high SNR regime. For the pre-filtering scheme and the codeword allocation scheme, although adopting the idea of compute-and-forward strategy makes the exchange rate scale correctly, there is no guarantee of a constant gap to the exchange capacity. Therefore, the gap between the cut-set bound and these two schemes can be very large. However, as promised by the Theorem 15 and Theorem 18, the proposed scheme is able to provide an exchange rate that is within a constant gap to the cut-set bound.

4.5 Conclusions

In this chapter, we studied the achievable rate region for the parallel Gaussian bidirectional relay channel. We first looked into the corresponding linear deterministic model and proposed a coding scheme that can achieve the capacity region for this

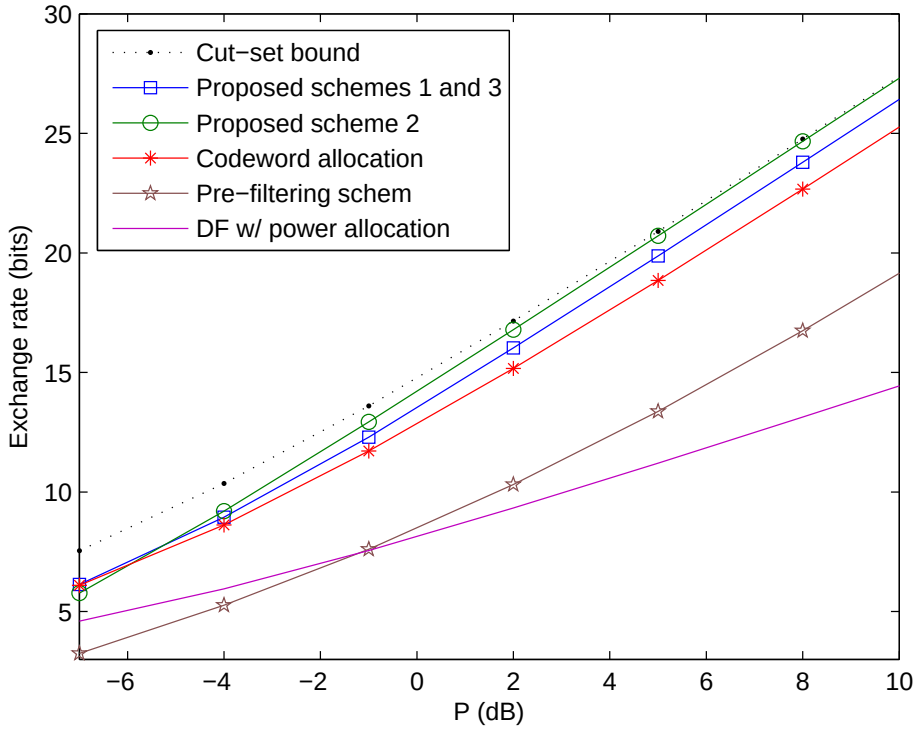

Figure 4.4: Performance comparison when $\mathbf{h_A} = [9,\ 1,\ 5,\ 4]$ and $\mathbf{h_b} = [8,\ 4,\ 2,\ 2]$.

deterministic model. Based on the insight obtained from the deterministic model, we proposed a coding scheme that can approach the capacity region to within L bits. This provided a detour to avoid the optimal linear filter designs for the scheme in [17]. For the reciprocal channels, we also showed that a very simple routing scheme is in fact optimal for the deterministic model. This idea was then leveraged to design a conceptually easier and more structured coding scheme for the Gaussian model with channel reciprocity assumption. Numerical results showed that the proposed scheme substantially outperform the decode-and-forward scheme and provide non-trivial gains over the scheme recently proposed in [17].

5. DESIGN OF PRACTICALLY IMPLEMENTABLE COMPUTE-AND-FORWARD SCHEMES

A novel construction of lattices is proposed. This construction can be thought of as Construction A with linear codes over $\mathbb{F}_{q^2}$ chosen from a restricted subset of codes where each code in this subset can be represented as the Cartesian product of two linear codes over $\mathbb{F}_q$; hence, is referred to as the product construction. The existence of a sequence of Poltyrev-good lattices generated by the product construction is shown. This family of lattices is then used to generate signal constellations with q^2 elements for the separation-based compute-and-forward framework proposed by Tunali *et al.*. Moreover, a multilevel coding/multistage decoding scheme tailored for these constellations is proposed so that the channel coding has to work only over $\mathbb{F}_q$ instead of $\mathbb{F}_{q^2}$.

5.1 Introduction

Compute-and-forward is a novel information forwarding paradigm in wireless communications in which relays in a network decode functions of signals transmitted from multiple transmitters and forward them to a central destination. If these functions are chosen as linear integer combinations, lattice codes are one of the most effective ways to implement a compute-and-forward scheme since a lattice is closed under addition. If the channel state information is not available at the transmitters, compute-and-forward can be implemented effectively by allowing the relay to choose integer coefficients depending on the channel coefficients and SNR. Such a scheme which uses lattices over integers has been analyzed by Nazer and Gastpar for AWGN networks in [11] where achievable information rates were derived. Feng, Silva, and Kschischang [57] have extended the framework in [11] towards the design of efficient

and practical schemes via an algebraic approach. Based on this approach, in [14], Tunali *et al.* considered the use of lattices over Eisenstein integers for the compute-and-forward paradigm and successfully extended the result on achievable rates in [11] to lattices over Eisenstein integers.

The schemes in [11] and [14] are based on Construction A over $\mathbb{F}_q$, which uses a linear code over $\mathbb{F}_q$ in conjunction with a constellation cropped from the corresponding integers (integers in [11] and Eisenstein integers in [14]) with q elements. One of the main drawbacks of this scheme is that the decoding complexity increases with q and, hence, the computational complexity of this scheme is quite high for high rates. In this chapter, we aim to construct lattice coding schemes with lower decoding complexity while still maintaining desirable properties such as good shaping gain and the ability to perform compute-and-forward. The main contributions of this work are the following. We first propose a novel construction of lattices inspired by Theorem 2 in [57] that can be thought of as construction A with linear codes over $\mathbb{F}_{q^2}$ chosen from a particular subset of codes. Specifically, we restrict the linear codes over $\mathbb{F}_{q^2}$ to those that can be represented as the Cartesian product of two linear codes over $\mathbb{F}_q$. This construction is then shown to be able to generate sequences of Poltyrev-good lattices. Moreover, based on the proposed lattices, we propose two families of signal constellations of q^2 elements that are suitable for compute-and-forward. A multilevel coding/multistage decoding scheme is then proposed for these constellations (with q^2 elements) so that the channel coding only has to work over $\mathbb{F}_q$ instead of $\mathbb{F}_{q^2}$. This substantially reduces the decoding complexity for a given size of the constellation (or, equivalently, rate) and hence makes the proposed scheme more practically implementable than the existing ones [11] [57] [14].

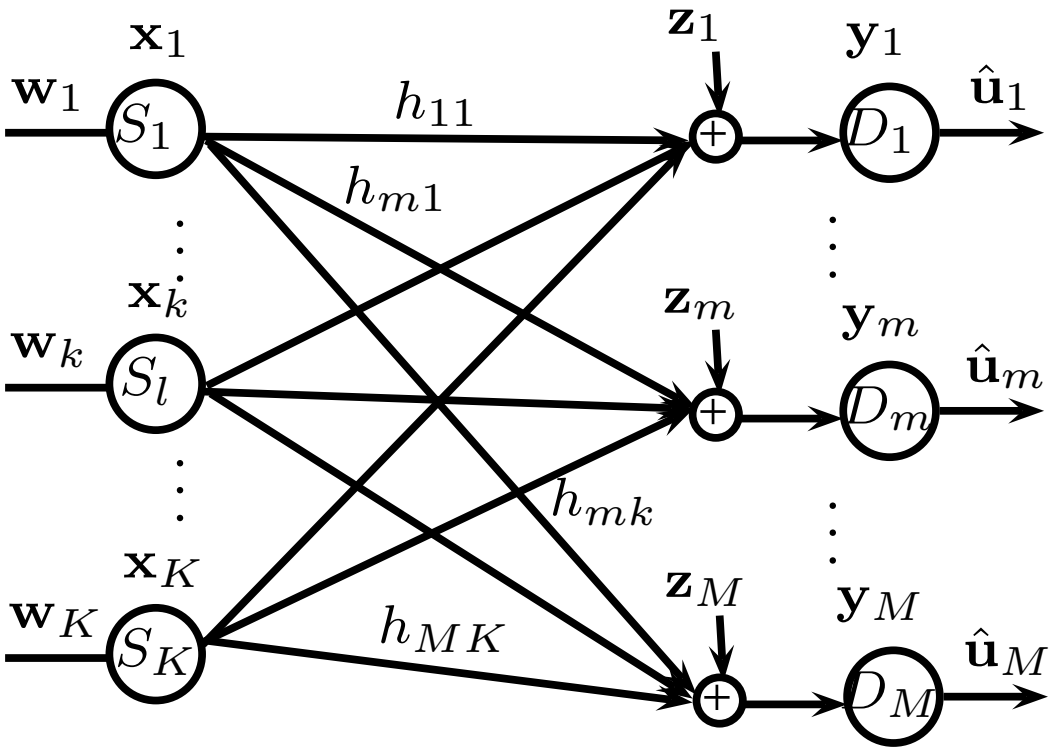

Figure 5.1: A compute-and-forward relay network where $S_1, \ldots, S_K$ are source nodes and $D_1, \ldots, D_M$ are destination nodes.

5.2 Problem Statement and Background

5.2.1 Problem Statement

The network considered in this chapter is the compute-and-forward relay network first studied by Nazer and Gastpar in [11]. Consider a K source nodes K destinations AWGN network as shown in Fig 5.1. Each source node has a message $w_k \in \{1, 2, \ldots, N'\}$, $k \in \{1, \ldots, K\}$ which is fed into an encoder $\mathcal{E}_k^N$ whose output is a length-N codeword $\mathbf{x}_k \in \mathbb{C}^N$. Each codeword is subject to a power constraint given by

$$\frac{1}{N}\|\mathbf{x}_k\|^2 = \frac{1}{N}\sum_{n=1}^{N} |x_k[n]|^2 \leq P. \tag{5.1}$$

The signal observed at destination m is given by

$$y_m[n] = \sum_{k=1}^{K} h_{mk} x_k[n] + z_m[n], \tag{5.2}$$

where $h_{mk} \in \mathbb{C}$ is the channel coefficient between the source node k and destination m, and $z_m[n] \sim \mathcal{CN}(0,1)$. One can also define the channel model for using the channel N times as

$$\mathbf{y}_m = \sum_{k=1}^{K} h_{mk} \mathbf{x}_k + \mathbf{z}_m. \tag{5.3}$$

In Nazer and Gastpar's setting, instead of individual messages, each destination node is only interested in recovering a linear combination of messages given by

$$\mathbf{u}_m = b_{m1}\mathbf{w}_1 \oplus \ldots \oplus b_{mK}\mathbf{w}_K, \tag{5.4}$$

where $b_{m1}, \ldots, b_{mK}$ are elements in the same field with $\mathbf{w}_k$ and the operations are elementwise. Upon observing $\mathbf{y}_m$, the destination node m forms $\hat{\mathbf{u}}_m = \mathcal{G}_m^N(\mathbf{y}_m)$ an estimate of $\mathbf{u}_m$.

Definition (Computation codes)**.** For a given set of $(b_{m1}, \ldots, b_{mK})$, a (N, N') computation code consists of a sequence of encoding/decoding functions $(\mathcal{E}_1^N, \ldots, \mathcal{E}_K^N)/(\mathcal{G}_1^N, \ldots, \mathcal{G}_M^N)$ described above and an error probability given by

$$P_e^{(N)} \triangleq \mathbb{P}\left(\bigcup_{m=1}^{M} \{\hat{\mathbf{u}}_m \neq \mathbf{u}_m\}\right). \tag{5.5}$$

Definition (Computation rate of relay m)**.** For a given channel vector $\mathbf{h}_m \triangleq [h_{m1}, \ldots, h_{mK}]^T$ and equation coefficient vector $\mathbf{b}_m \triangleq [b_{m1}, \ldots, b_{mK}]^T$, a computation rate $R(\mathbf{h}_m, \mathbf{b}_m)$ is achievable at relay m if for any $\varepsilon > 0$ there is an (N, N') computation code such

that

$$N' \geq 2^{NR(\mathbf{h}_m, \mathbf{b}_m)} \text{ and } P_e^{(N)} \leq \varepsilon. \tag{5.6}$$

In the following, we will consider the case that we have two source nodes and only one destination node for the sake of simplicity. However, all the results can be extend to the general setting straightforwardly. With this simplification, we will restrict $k \in \{1, 2\}$ and drop the subscript m from now on. The received signal can then be compressed to

$$\mathbf{y} = h_1 \mathbf{x}_1 + h_2 \mathbf{x}_2 + \mathbf{z}. \tag{5.7}$$

5.2.2 Background

In this subsection, we provide some preliminaries that will be useful in explaining our results in the following sections. All the Lemmas are provided without proofs for the sake of brevity; however, their proofs can be found in standard textbooks of abstract algebra, see for example [58].

Lemma 19. *If $\mathcal{R}$ is a principal ideal domain (PID), then every non-zero prime ideal is maximal.*

Lemma 20. *Let $\mathcal{I}$ be an ideal in a ring $\mathcal{R}$ with identity $1_{\mathcal{R}} \neq 0$. If $\mathcal{I}$ is maximal and $\mathcal{R}$ is commutative, then the quotient ring $\mathcal{R}/\mathcal{I}$ is isomorphic to a field.*

Lemma 21 (Chinese Remainder Theorem)**.** *Let $\mathcal{R}$ be a commutative ring, and $\mathcal{I}_1, \ldots, \mathcal{I}_n$ be ideals in $\mathcal{R}$, such that they are relatively prime. Then,*

$$\mathcal{R}/\cap_{i=1}^{n} \mathcal{I}_i \cong \mathcal{R}/\mathcal{I}_1 \times \ldots \times \mathcal{R}/\mathcal{I}_n. \tag{5.8}$$

The ring of Eisenstein integers $\mathbb{Z}[\omega]$ is the collection of complex numbers of the form $a + b\omega$ where $a, b \in \mathbb{Z}$ and $\omega = -\frac{1}{2} + j\frac{\sqrt{3}}{2}$. The ring of Gaussian integers $\mathbb{Z}[i]$

is the collection of complex numbers of the form $a + bj$ where again $a, b \in \mathbb{Z}$. Both $\mathbb{Z}[\omega]$ and $\mathbb{Z}[i]$ are PIDs. An Eisenstein integer (Gaussian integer) π is an Eisenstein prime (Gaussian prime) if either one of the following mutually exclusive conditions hold:

1. π is equal to the product of a unit and any natural prime congruent to 2 mod 3 (3 mod 4).

2. $|\pi|^2$ is any natural prime congruent to 0 mod 3 or 1 mod 3 (2 or 1 mod 4).

For those Eisenstein primes (Gaussian primes) π with $|\pi|^2 = \pi \cdot \bar{\pi}$ being natural primes congruent to 1 mod 3 (1 mod 4), one can verify that π and $\bar{\pi}$ are both Eisenstein primes (Gaussian primes) but they are not associates. Moreover, it has been shown in [59] that for every $x \geq 7$, there exists a natural prime of this form between x and $2x$. Thus, the choices of π satisfying the above property are abundant.

5.2.3 Proposed Product Construction of Lattices

Motivated by Theorem 2 in [57], we propose the product construction of lattices shown in Fig. 5.2. Let $\sigma : \mathbb{Z}[\omega] \to \mathbb{Z}[\omega]/q\mathbb{Z}[\omega] \leftrightarrow \mathbb{F}_q^2$ be a $\mathbb{Z}$-module homomorphism and let $\mathcal{M} \triangleq \sigma^{-1}$ be a mapping from $\mathbb{F}_q^2$ to $\mathbb{Z}[\omega]/q\mathbb{Z}[\omega]$. Throughout the chapter, we will refer to the set of coset representatives of $\mathbb{Z}[\omega]/q\mathbb{Z}[\omega]$ as *signal constellation* or *constellation* in short. Also, let $\mathcal{B}^l$, $l \in \{1, 2\}$, be the set of *all* linear (N, m_l) codes over $\mathbb{F}_q$ and $\mathcal{B} \triangleq \mathcal{B}^1 \times \mathcal{B}^2$. i.e., $\mathcal{B}$ is the collection of all linear codes over $\mathbb{F}_{q^2}$ that can be represented as the Cartesian product of two linear codes whose input lengths are m_1 and m_2, respectively, over $\mathbb{F}_q$. The construction consists of the following steps.

1. Let $\mathcal{C} = \mathcal{C}^1 \times \mathcal{C}^2 \in \mathcal{B}$ where $\mathcal{C}^l \in \mathcal{B}^l$, $l \in \{1, 2\}$.

2. Define $\Lambda^* \triangleq \mathcal{M}(\mathcal{C}^1, \mathcal{C}^2)$ where for two vectors $\mathbf{c}^1$ and $\mathbf{c}^2$ with equal length, $\mathcal{M}(\mathbf{c}^1, \mathbf{c}^2)$ is defined as the elementwise mapping.

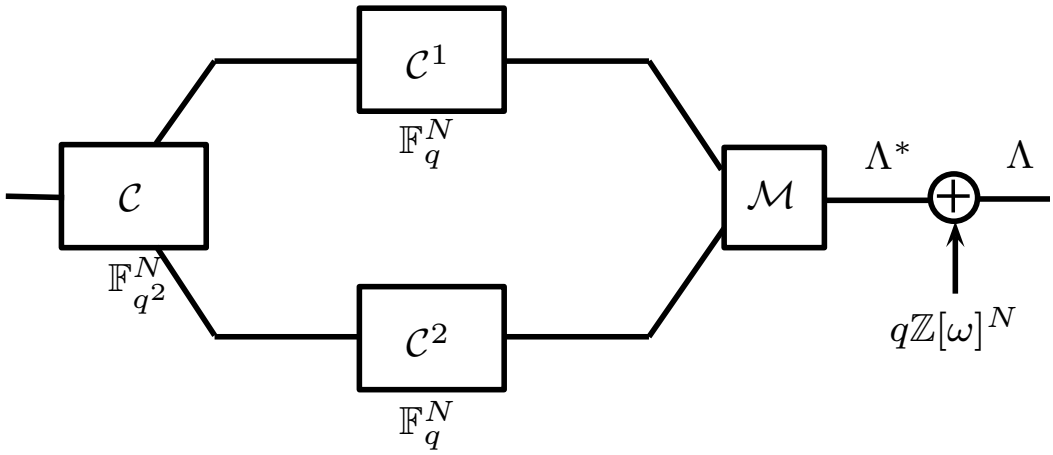

Figure 5.2: The proposed production construction of lattices.

3. Replicate Λ^* over the entire $\mathbb{C}^N$ to form $\Lambda \triangleq \Lambda^* + q\mathbb{Z}[\omega]^N$.

Theorem 22. *Λ is a lattice and there exists a sequence of lattices generated by this construction that is Poltyrev-good.*

Proof. (Sketch) We first verify that $\mathcal{B}$ is a balanced set. i.e., each element in $\mathbb{F}_{q^2}^N$ is contained in the same number of codes in this restricted set. This in turn shows the Minkowski-Hlawka Theorem and the remaining steps of the proof in [60] can be repeated. □

Here, by Poltyrev-good, we mean that the lattices achieve the Poltyrev limit; instead of the Poltyrev exponent whose proof is more involved. This theorem establishes the fact that the proposed lattice is good for channel coding if shaping is ignored.

5.2.4 Separation-Based Compute-and-Forward by Tunali et al.

Typically, in addition to the Poltyrev-goodness, shaping has to be taken into account in order to use a lattice for communication. However, due to the lack of

efficient shaping techniques in practice, in this chapter, we consider the separation-based framework discussed in Section V of [14]. This framework attempts to separate the design of channel coding and data modulation so that N the dimension of the channel coding can still grow to infinity and the shaping is designed to be optimal in a small dimensional space (although suboptimal in the N dimensional space). The channel coding employed by all the source nodes is restricted to be the same linear code over $\mathbb{F}_q$ in order to make linear combinations (over $\mathbb{F}_q$) of codewords valid codewords. On the other hand, the constellation has to be carefully chosen so that one can still benefit from the structural gain offered by compute-and-forward strategy. It turns out that the key condition for this is a ring homomorphism between the extended version (to infinite constellation) of the signal constellation and $\mathbb{F}_q$ the field that the channel coding is working over.

In Section V of [14], a scaled version of the quotient ring of Eisenstein integers shown in the following is chosen as signal constellation at each source node,

$$\mathcal{A} \triangleq \gamma \left(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \right), \tag{5.9}$$

where $|\pi|^2 \triangleq q$ is a natural prime congruent to 1 mod 3 and γ is for satisfying the power constraint. Since $\mathbb{Z}[\omega]$ is a PID, from Lemma 19, one has that $\pi\mathbb{Z}[\omega]$ is a maximal ideal. Hence, from Lemma 20, $\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \cong \mathbb{F}_q$. i.e., the following ring homomorphism σ exist,

$$\sigma : \mathbb{Z}[\omega] \to \mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \leftrightarrow \mathbb{F}_q. \tag{5.10}$$

In addition to the ring homomorphisms, this choice of constellation provides other properties such as optimal shaping gain (in two-dimensional space) and good quan-

tization of channel coefficients as $\mathbb{Z}[\omega]$ corresponds to hexagonal lattices.

Unlike the framework proposed in [11] and Section III of [14] in which infinitely-dimensional lattices are employed as channel coding and data modulation jointly, the separation approach allows one to let the dimension of channel coding grow while keeps the constellation size small so that optimal decoding is feasible. This also allows the use of well-developed codes on graph (e.g., non-binary LDPC) as channel coding for which one can carry out iterative decoding to further reduce the decoding complexity.

5.3 Proposed Constellations

In this section, we propose two families of constellations for the separation-based compute-and-forward. The constellations in the first family are isomorphic to the corresponding extension fields so that ring homomorphisms exist. Therefore, these constellations can be directly used for compute-and-forward. For the constellations in the second family, although the ring homomorphisms are not guaranteed, the existence of $\mathbb{Z}$-module homomorphisms can be shown by using the Chinese remainder theorem. The $\mathbb{Z}$-module homomorphisms will be exploited for a novel multilevel compute-and-forward strategy in the next section. It is worth emphasizing that the theory required for showing the existence of the homomorphisms has been developed in [57, Theorem 1 and Theorem 2]. But these specific constructions are not explicitly proposed in [57] and the multilevel coding/multistage decoding in Section 5.4.1 is not given in [57].

5.3.1 The First Proposed Family of Constellations

Let π be an Eisenstein prime with π being the product of a unit and a natural prime q congruent to 2 mod 3. Since $\mathbb{Z}[\omega]$ is a PID, from Lemma 19, $\pi\mathbb{Z}[\omega]$ is a

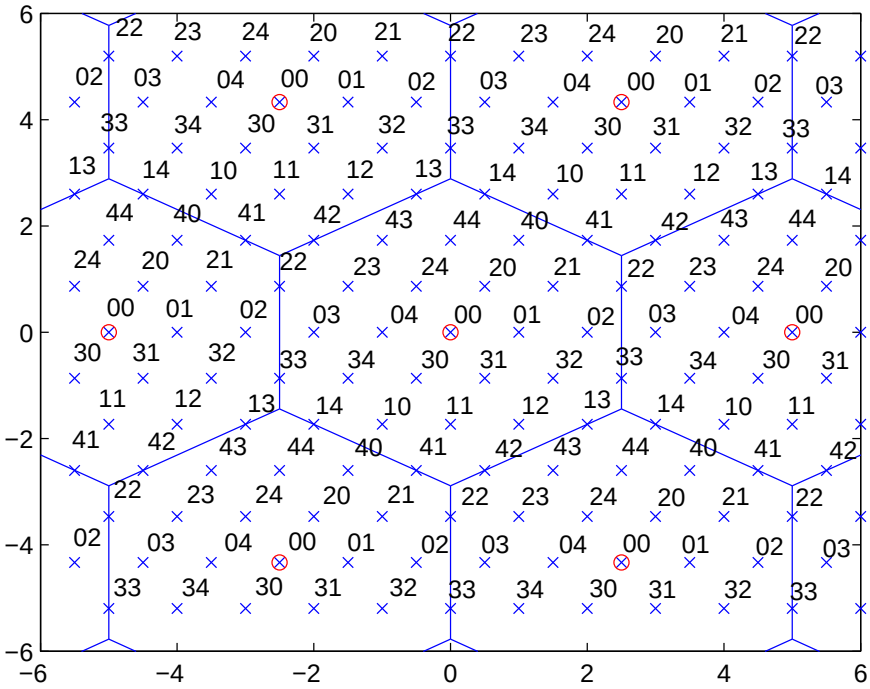

Figure 5.3: The proposed constellation with $q = 5$ and a ring homomorphism shown as the labeling with the irreducible polynomial $x^2 + 2x + 4$.

prime ideal and hence a maximal ideal. From Lemma 20, one has that

$$\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \cong \mathbb{F}_{q^2}. \tag{5.11}$$

Thus, the following ring homomorphism exists

$$\sigma : \mathbb{Z}[\omega] \to \mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \leftrightarrow \mathbb{F}_{q^2}. \tag{5.12}$$

Example 4. One example of this construction with $q = 5$ is given in Fig. 5.3 where the labeling is the ring homomorphism and the multiplication in $\mathbb{F}_{25}$ is defined by the irreducible polynomial $x^2 + 2x + 4$ over $\mathbb{F}_5$. $\Diamond$

Similarly, let π be a Gaussian prime with π being the product of a unit and a natural prime q congruent to 3 mod 4. Again, one has that

$$\mathbb{Z}[i]/\pi\mathbb{Z}[i] \cong \mathbb{F}_{q^2}. \tag{5.13}$$

Thus, the following ring homomorphism exists

$$\sigma : \mathbb{Z}[i] \to \mathbb{Z}[i]/\pi\mathbb{Z}[i] \leftrightarrow \mathbb{F}_{q^2}. \tag{5.14}$$

Example 5. One example of this construction with $q = 3$ is given in Fig. 5.4 where the labeling is the ring homomorphism and the multiplication in $\mathbb{F}_9$ is defined by the irreducible polynomial $x^2 + 1$ over $\mathbb{F}_3$. $\Diamond$

One straightforward way to exploit this property is to use

$$\mathcal{A} \triangleq \gamma\left(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega]\right) \text{ or } \gamma\left(\mathbb{Z}[i]/\pi\mathbb{Z}[i]\right), \tag{5.15}$$

where γ is for the power constraint, as signal constellation for the separation-based compute-and-forward framework in [14] and to directly apply the above ring homomorphism as data modulation. However, this implies that one has to work with a very large field $\mathbb{F}_{q^2}$ which results in a formidable decoding complexity. In the next section, we propose a novel encoding/decoding pair that incorporates the idea of multilevel coding and multistage decoding [61]. The proposed encoding/decoding allows us to work over a potentially much smaller field $\mathbb{F}_q$.

5.3.2 The Second Proposed Family of Constellations

Let π be an Eisenstein prime with $|\pi|^2 = q$ congruent to 1 mod 3. As mentioned before, one has that $\bar{\pi}$ is also an Eisenstein prime and π and $\bar{\pi}$ are not associates,

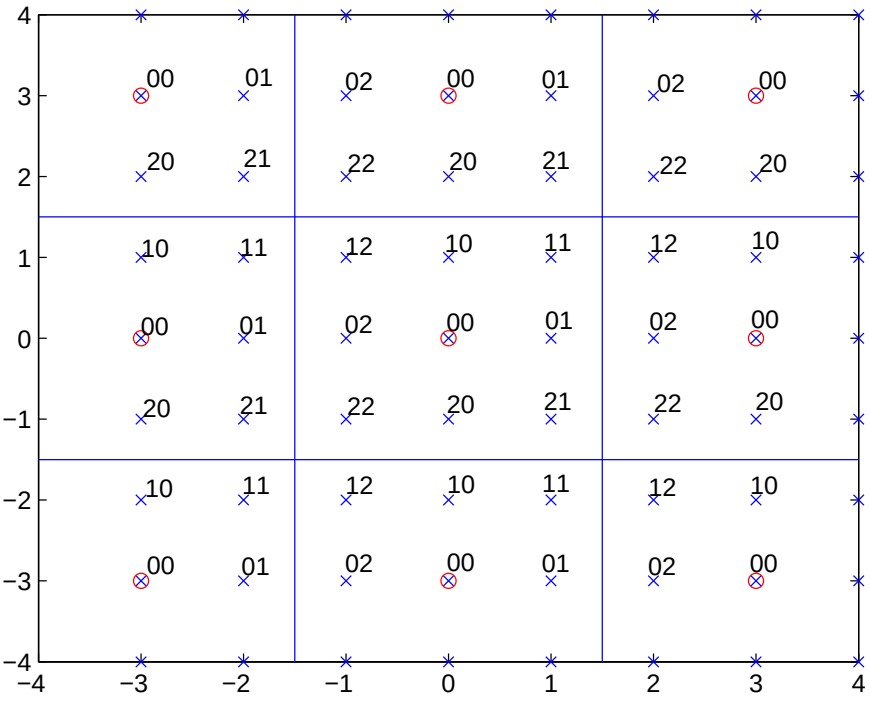

Figure 5.4: The proposed constellation with $q = 3$ and a ring homomorphism shown as the labeling.

i.e., they are relatively prime. Now, one has that

$$\begin{aligned}\mathbb{Z}[\omega]/\pi\bar{\pi}\mathbb{Z}[\omega] &\overset{(a)}{\cong} \mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \times \mathbb{Z}[\omega]/\bar{\pi}\mathbb{Z}[\omega] \\ &\overset{(b)}{\cong} \mathbb{F}_q \times \mathbb{F}_q = \mathbb{F}_q^2,\end{aligned} \tag{5.16}$$

where (a) follows from the Chinese Remainder Theorem in Lemma 21 and (b) is from Lemma 20 and the fact that $\mathbb{Z}[\omega]$ is a PID. Therefore, this implies that the following $\mathbb{Z}$-module homomorphism exists,

$$\sigma : \mathbb{Z}[\omega] \to \mathbb{Z}[\omega]/\pi\bar{\pi}\mathbb{Z}[\omega] \leftrightarrow \mathbb{F}_q^2. \tag{5.17}$$

One way to generate $\mathbb{Z}$-module homomorphism is to choose

$$\sigma^{-1}(c^1, c^2) \triangleq c^1 + c^2\omega \mod \pi\bar{\pi}\mathbb{Z}[\omega]. \tag{5.18}$$

Similarly, let π be a Gaussian prime with $|\pi|^2 = q$ congruent to 1 mod 4. Then, $\bar{\pi}$ and π are both Gaussian primes and are relatively prime. Hence, one again has

$$\mathbb{Z}[i]/\pi\bar{\pi}\mathbb{Z}[i] \cong \mathbb{F}_q^2, \tag{5.19}$$

$$\begin{aligned}\mathbb{Z}[i]/\pi\bar{\pi}\mathbb{Z}[i] &\cong \mathbb{Z}[i]/\pi\mathbb{Z}[i] \times \mathbb{Z}[i]/\bar{\pi}\mathbb{Z}[i] \\ &\cong \mathbb{F}_q \times \mathbb{F}_q = \mathbb{F}_q^2,\end{aligned} \tag{5.20}$$

and

$$\sigma : \mathbb{Z}[i] \to \mathbb{Z}[i]/\pi\bar{\pi}\mathbb{Z}[i] \leftrightarrow \mathbb{F}_q^2. \tag{5.21}$$

A corresponding $\mathbb{Z}$-module homomorphism can be obtained by replacing ω by j

and $\mathbb{Z}[\omega]$ by $\mathbb{Z}[i]$ in (5.18). For this family of constellations, although the ring homomorphisms are not guaranteed, one can still exploit the existence of $\mathbb{Z}$-module homomorphisms for compute-and-forward as described in the next section.

5.4 Proposed Multilevel Coding Scheme

In this section, we propose a multilevel encoding/multi-stage decoding scheme where only $\mathbb{Z}$-module homomorphisms are required. For the first proposed family of constellations (whose $\mathbb{Z}$-module homomorphisms can be trivially obtained from the ring homomorphisms), the proposed multilevel coding scheme allows one to significantly reduce the decoding complexity at a price of slight rate reduction. On the other hand, for the second proposed family of constellations, this multilevel coding scheme enables the compute-and-forward strategy. It is worth emphasizing that for both the proposed families of constellations with q^2 elements, the employment of the proposed multilevel coding/multistage decoding allows the channel coding to work only over $\mathbb{F}_q$.

5.4.1 Encoding/Decoding

Without loss of generality, here, we equivalently consider the messages to be length L vectors over the extension field $\mathbb{F}_{q^2}$ such that

$$L \approx \log_{q^2}(N'), \tag{5.22}$$

where the approximation can be made precise when N' is large. Each source node S_k first splits its input stream $\mathbf{w}_k \in (\mathbb{F}_{q^2})^L$ into two streams, namely $\mathbf{w}_k^1 \in \mathbb{F}_q^{m_1}$ and $\mathbf{w}_k^2 \in \mathbb{F}_q^{m_2}$ with $m_1+m_2 = 2L$. Let $\mathcal{C}^1$ and $\mathcal{C}^2$ be linear codes over $\mathbb{F}_q$ adopted in levels 1 and 2 and G^1 and G^2 be the generator matrices, respectively. The rate of these linear codes are chosen to be $R_1 = m_1/N$ and $R_2 = m_2/N$ such that the outputs

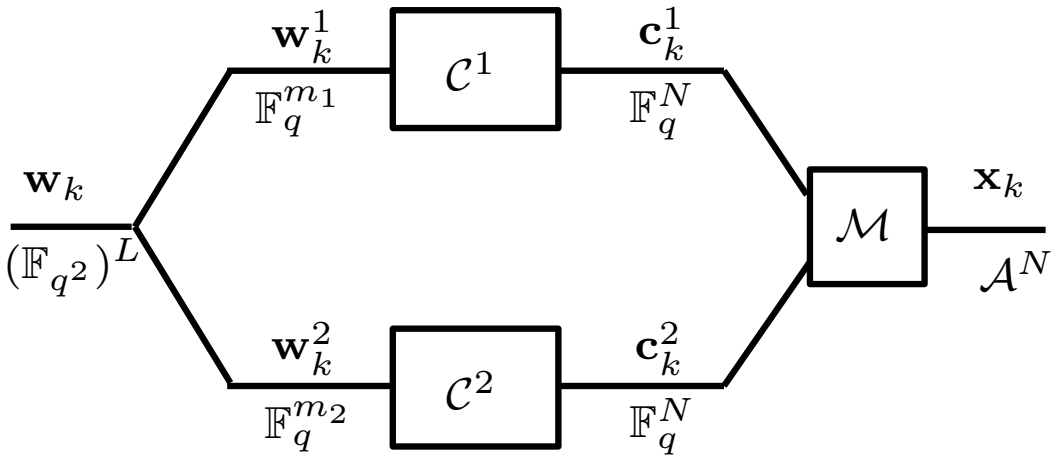

Figure 5.5: The encoder of the proposed multilevel coding scheme.

have a same length N. We individually encode each stream with the corresponding linear code as $\mathbf{c}_k^1 = \mathbf{w}_k^1 G^1$ and $\mathbf{c}_k^2 = \mathbf{w}_k^2 G^2$. The encoder k then takes two of these symbols and maps them to a symbol from $\mathcal{A}$ via $\mathcal{M} \triangleq \gamma \cdot \sigma^{-1}$. The overall encoding process is summarized in Fig. 5.5.

The decoder at the destination is a multistage decoder in which we first decode the first stream by treating the other stream as unknown and then decode the second stream by assuming the previous decoding is correct. It can be seen from the examples of the proposed constellations in Section 5.3 that given one level (or ignoring one level by treating it as noise), the resulted constellation for the other level is a q-ary pulse amplitude modulation (PAM). Therefore, using the multilevel coding with multistage decoding for compute-and-forward only requires $\mathbb{Z}$-module homomorphisms instead of ring homomorphisms since we only consider one level at a time.

Given (b_1^1, b_2^1), in order to decode the first data stream, the decoder first computes

the *a posteriori* probabilities given by

$$\mathbb{P}\left(\tilde{c}_R^1[n] = l | y[n]\right) \propto \sum_{\substack{c_1^1, c_2^1 \in \mathbb{F}_q: \\ b_1^1 c_1^1 \oplus b_2^1 c_2^1 = l}} \sum_{c_1^2, c_2^2 \in \mathbb{F}_q} \exp\left[- \left\| h_1 \mathcal{M}(c_1^1, c_1^2) + h_2 \mathcal{M}(c_2^1, c_2^2) - y[n] \right\|^2 \right], \tag{5.23}$$

for all $l \in \mathbb{F}_q$ and for each codeword dimension n. According to these *a posteriori* probabilities, the decoder forms the first level's estimate given by

$$\hat{\mathbf{c}}_R^1 = \arg\max_{\mathbf{c} \in \mathcal{C}^1} \prod_{n=1}^{N} \mathbb{P}\left(\tilde{c}_R^1[n] = l | y[n]\right). \tag{5.24}$$

Given (b_1^2, b_2^2), in the decoding of the second data stream, the destination assumes the decoding in the first level is correct and again forms the corresponding *a posteriori* probabilities given by

$$\mathbb{P}\left(\tilde{c}_R^2[n] = l | y[n], \hat{c}_R^1[n]\right) \propto \sum_{\substack{c_1^1, c_2^1 \in \mathbb{F}_q: \\ b_1^1 c_1^1 \oplus b_2^1 c_2^1 = \hat{c}_R^1[n]}} \sum_{\substack{c_1^2, c_2^2 \in \mathbb{F}_q: \\ b_1^2 c_1^2 \oplus b_2^2 c_2^2 = l}} \exp\left[- \left\| h_1 \mathcal{M}(c_1^1, c_1^2) + h_2 \mathcal{M}(c_2^1, c_2^2) - y[n] \right\|^2 \right], \tag{5.25}$$

for all $l \in \mathbb{F}_q$ and for each codeword dimension n. Similar to the first level, the decoder then forms the second level's estimate as

$$\hat{\mathbf{c}}_R^2 = \arg\max_{\mathbf{c} \in \mathcal{C}^2} \prod_{n=1}^{N} \mathbb{P}\left(\tilde{c}_R^2[n] = l | y[n], \hat{c}_R^1[n]\right). \tag{5.26}$$

5.4.2 Achievable Computation Rate

In [61], using the chain rule of mutual information, Wachsmann *et al.* show that the multilevel coding incurs no loss in terms of the achievable information rate for point to point communication. The same proof works for our construction as well, which we summarize here. Let $\mathcal{C}^1$ and $\mathcal{C}^2$ be the codebooks used for level 1 and 2, respectively, and let C^1 and C^2 be the corresponding random variables. Also, notice that the mapping between (C^1, C^2) and $\mathcal{A}$ is bijective. One has that

$$\begin{aligned} R_{\text{AWGN}} &= I(Y;A) = I(Y;\mathcal{M}(C^1,C^2)) \\ &\overset{(a)}{=} I(Y;C^1,C^2) = I(Y;C^1) + I(Y;C^2|C^1), \end{aligned} \tag{5.27}$$

where (a) is due to the fact that $\mathcal{M}$ is bijective.

Now, we provide the achievable information rates of the proposed schemes for compute-and-forward. For the first proposed family of constellations, one can choose to directly work over $\mathbb{F}_{q^2}$. Let $\mathcal{C}$ be a codebook over $\mathbb{F}_{q^2}$ and let C_1 and C_2 be the corresponding random variables at source nodes 1 and 2, respectively. The achievable computation rate of directly working over $\mathbb{F}_{q^2}$ can be written as

$$R_{\text{direct}} = \max_{b_1,b_2\in\mathbb{F}_{q^2}} I(Y; b_1C_1 \oplus b_2C_2), \tag{5.28}$$

where the subscript "direct" stands for that we directly work over the extension field. For the multilevel coding/multistage decoding scheme (which works for both constructions), let $\mathcal{C}^1$ and $\mathcal{C}^2$ be the linear codebooks adopted for level 1 and 2, respectively, and let C_k^1 and C_k^2 be the corresponding random variables at source

node k. One has the achievable computation rate given by

$$\begin{aligned} R_{\text{MLC}} &= \max_{b_1^1,b_1^2,b_2^1,b_2^2 \in \mathbb{F}_q} I(Y; b_1^1 C_1^1 \oplus b_2^1 C_2^1, b_1^2 C_1^2 \oplus b_2^2 C_2^2) \\ &= \max_{b_1^1,b_1^2,b_2^1,b_2^2 \in \mathbb{F}_q} I(Y; b_1^1 C_1^1 \oplus b_2^1 C_2^1) \\ &\qquad + I(Y; b_1^2 C_1^2 \oplus b_2^2 C_2^2 | b_1^1 C_1^1 \oplus b_2^1 C_2^1). \end{aligned} \tag{5.29}$$

It should be noted that R_{direct} and R_{MLC} are in general not the same. It is because for $C_k = (\tilde{C}_k^1, \tilde{C}_k^2)$, given $b_1, b_2 \in \mathbb{F}_{q^2}$, there may not exist $\tilde{b}_1^1, \tilde{b}_1^2, \tilde{b}_2^1, \tilde{b}_2^2 \in \mathbb{F}_q$ such that

$$b_1 C_1 \oplus b_2 C_2 = (\tilde{b}_1^1 \tilde{C}_1^1 \oplus \tilde{b}_2^1 \tilde{C}_2^1, \tilde{b}_1^2 \tilde{C}_1^2 \oplus \tilde{b}_2^2 \tilde{C}_2^2), \tag{5.30}$$

and $\tilde{C}_k^1, \tilde{C}_k^2$ are valid codewords over $\mathbb{F}_q$.

5.5 Simulation Results

In this section, we use the Monte-Carlo method to simulate the achievable computation rates. We will focus on the ring of Eisenstein integers for the sake of brevity but similar results can be observed for the Gaussian integers. The Eisenstein prime that we use is $\pi = 5$, i.e., $|\mathcal{A}| = 25$. Since 5 is congruent to 2 mod 3, this belongs to the first proposed family of constellations and ring homomorphisms exist. Hence, one can choose either to directly carry out the separation-based scheme with the ring homomorphism given in Example 4 or to implement the proposed multilevel coding and multistage decoding scheme. In what follows, we do both and compare the achievable rates of these approaches given in (5.28) and (5.29).

In Fig. 5.6, we show the achievable rates of the proposed constellations with multilevel coding where each level employs a linear code over $\mathbb{F}_5$ and the proposed construction with a linear code over $\mathbb{F}_{25}$. The ring homomorphism adopted is as

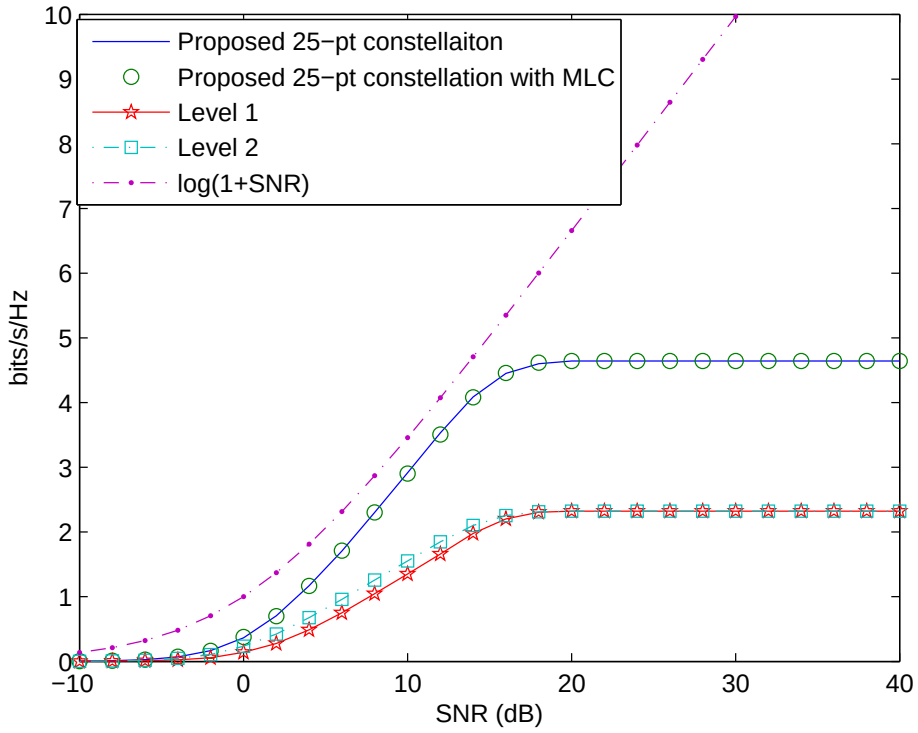

Figure 5.6: Achievable rates of the proposed construction with and without multilevel coding. For the one with multilevel coding, the achievable rate achieved by each level is also plotted. The channel coefficients are set to be $h_1 = h_2 = 1$.

shown in Example 4. For the multilevel coding scheme, the rate achieved by each level is also shown. The transmitted SNR is ranging from -10 dB to 40 dB. The channel coefficients are set to be $h_1 = h_2 = 1$ in order to simulate the scenario when there is no self-interference. In this case, one can see from the figure that using the proposed constellations with multilevel coding incur no rate loss compared to the scheme directly working over $\mathbb{F}_{25}$. It is because for this case, both schemes would choose to decode the received signal to the sum of the messages (over the corresponding fields) and digit-wise addition in the base field is equivalent to addition in the extension field.

We then compare the average achievable rates of the proposed constellation in Fig. 5.3 with and without multilevel coding, the constellations in [14] with 7 elements, and that with 19 elements. The channel coefficients are drawn from $\mathcal{CN}(0,1)$ (i.e., its norm has a Rayleigh distribution). We average over 100 pairs of channel coefficients and the results are shown in Fig. 5.7. One can see that the scheme directly working over $\mathbb{F}_{25}$ provides a higher rate than that provided by the multilevel coding which works over $\mathbb{F}_5$. However, the gap becomes smaller and smaller as the SNR increases. One also observes that after roughly 26 dB, the gap becomes negligible and the proposed multilevel coding scheme over $\mathbb{F}_5$ outperforms the scheme working over $\mathbb{F}_{19}$. This shows that using the proposed scheme over $\mathbb{F}_5$, one can perform very close to the scheme over $\mathbb{F}_{25}$ and outperform the scheme over $\mathbb{F}_{19}$ in the high SNR regime with a substantially lower computational complexity.

5.6 Extensions

One interesting extension is to restrict the codes used in the two levels to be the same, i.e., $\mathcal{C}^1 = \mathcal{C}^2 = \mathcal{C}$. Then, there always exist parameters such that (5.30) holds. However, this still doesn't mean that the multilevel coding scheme would achieve the same rate with that provided by the code over $\mathbb{F}_{q^2}$. It is because the symmetric capacity may not touch the boundary of the sum rate limit for the underlying MAC channel. It is interesting to see when this is true and currently we have been able to identify some special cases. Moreover, by setting $\mathcal{C}^1 = \mathcal{C}^2 = \mathcal{C}$, one can further incorporate the idea of flexible decoding [62] into our framework. Specifically, in addition to the original choice we have had

$$b_1^1 c_1^1 \oplus b_2^1 c_2^1 \text{ and } b_1^2 c_1^2 \oplus b_2^2 c_2^2, \tag{5.31}$$

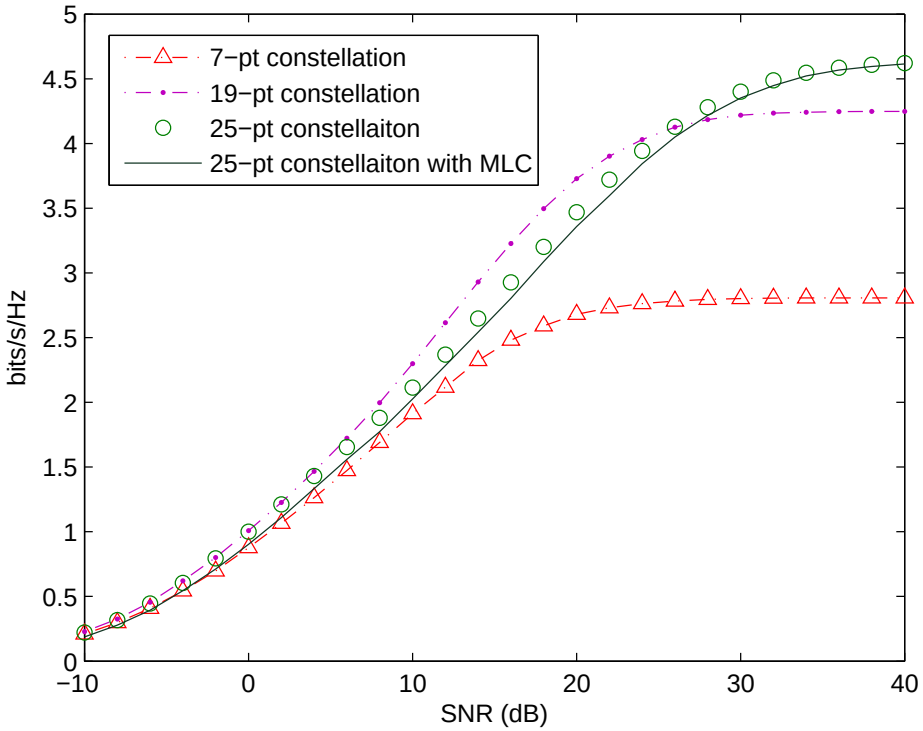

Figure 5.7: Average information rates of the proposed constellation with multilevel coding over $\mathbb{F}_5$, that with a linear code over $\mathbb{F}_{25}$, and the construction by Tunali *et al.* with a linear code over $\mathbb{F}_7$, and that over $\mathbb{F}_{19}$.

where $b_1^1, b_1^2, b_2^1, b_2^2 \in \mathbb{F}_q$, one can decode to something else. For example,

$$\tilde{b}_1^1 c_1^1 \oplus \tilde{b}_2^2 c_2^2 \text{ and } \tilde{b}_1^2 c_1^2 \oplus \tilde{b}_2^1 c_2^1. \tag{5.32}$$

where $\tilde{b}_1^1, \tilde{b}_1^2, \tilde{b}_2^1, \tilde{b}_2^2 \in \mathbb{F}_q$. More precisely, one can decode the received signal to

$$[c_R^1, c_R^2]^T = [\mathbf{B}_1 \mathbf{B}_2] \begin{bmatrix} c_1^1 \\ c_1^2 \\ c_1^1 \\ c_1^2 \end{bmatrix}, \tag{5.33}$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are 2 by 2 full-rank matrices with elements in $\mathbb{F}_q$. This approach allows rich choices of functions that one can decode to and hence may result in a higher rate in general.

5.7 Conclusions

In this chapter, we studied the design of practically implementable compute-and-forward schemes. We focused on the separation-based framework that allow separately designing channel coding and data modulation. We first proposed two families of constellations that are suitable for the separation-based framework while still provide good shaping gains. Using the idea of multilevel coding and multistage decoding, we then proposed a low complexity scheme that would substantially reduce the decoding complexity in the high rate regime. We also showed that the use of multilevel coding and multistage decoding incurs no essential rate loss in the regions that one would operate on. Moreover, the proposed scheme can be further extended by incorporating the idea of flexible decoding in [62].

6. JOINT SOURCE-CHANNEL CODING WITH CORRELATED INTERFERENCE*

We study the joint source-channel coding problem of transmitting a discrete-time analog source over an additive white Gaussian noise (AWGN) channel with interference known at transmitter. We consider the case when the source and the interference are correlated. We first derive an outer bound on the achievable distortion and then, we propose two joint source-channel coding schemes. The first scheme is the superposition of the uncoded signal and a digital part which is the concatenation of a Wyner-Ziv encoder and a dirty paper encoder. In the second scheme, the digital part is replaced by the hybrid digital and analog scheme proposed by Wilson *et al.* When the channel signal-to-noise ratio (SNR) is perfectly known at the transmitter, both proposed schemes are shown to provide identical performance which is substantially better than that of existing schemes. In the presence of an SNR mismatch, both proposed schemes are shown to be capable of graceful enhancement and graceful degradation. Interestingly, unlike the case when the source and interference are independent, neither of the two schemes outperforms the other universally. As an application of the proposed schemes, we provide both inner and outer bounds on the distortion region for the generalized cognitive radio channel.

6.1 Introduction and Problem Statement

In this chapter, we consider transmitting a length-n i.i.d. zero-mean Gaussian source $\mathbf{V} = (V(1), V(2), \ldots, V(n))$ over n uses of an additive white Gaussian noise (AWGN) channel with noise $\mathbf{Z} \sim \mathcal{N}(0, N \cdot I)$ in the presence of Gaussian interfer-

*

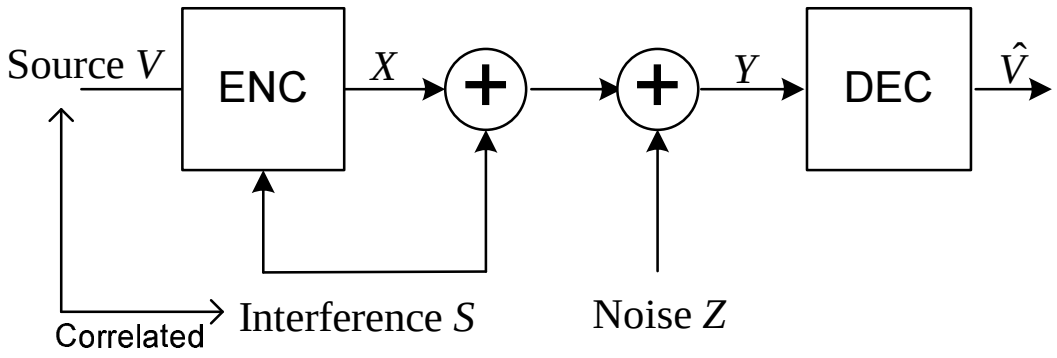

Figure 6.1: Joint source-channel coding with interference known at transmitter.

ence $\mathbf{S}$ which is known at the transmitter as shown in Fig. 6.1. Throughout the chapter, we only focus on the bandwidth-matched case, i.e., the signalling rate over the channel is equal to the sampling rate of the source. The transmitted signal $\mathbf{X} = (X(1), X(2), \ldots, X(n))$ is subject to a power constraint

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X(i)^2] \leq P, \tag{6.1}$$

where $\mathbb{E}[\cdot]$ represents the expectation operation. The received signal $\mathbf{Y}$ is given by

$$\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{Z}. \tag{6.2}$$

We are interested in the expected distortion between the source and the estimate $\widehat{\mathbf{V}}$ at the output of the decoder given by

$$d = \mathbb{E}[d(\mathbf{V}, g(f(\mathbf{V}, \mathbf{S}) + \mathbf{S} + \mathbf{Z}))], \tag{6.3}$$

where f and g are a pair of source-channel coding encoder and decoder, respectively,

and $d(.,.)$ is the mean squared error (MSE) distortion measure given by

$$d(\mathbf{v}, \hat{\mathbf{v}}) = \frac{1}{n}\sum_{i=1}^{n}(v(i) - \hat{v}(i))^2. \tag{6.4}$$

Here the lower case letters represent realizations of random variables denoted by upper case letters. As in [30], a distortion D is achievable under power constraint P if for any $\varepsilon > 0$, there exists a source-channel code and a sufficiently large n such that $d \leq D + \varepsilon$.

When V and S are uncorrelated, it is known that an optimal quantizer followed by a Costa's dirty paper coding (DPC) [63] is optimal and the corresponding joint source-channel coding problem is fully discussed in [64]. However, different from the typical writing on dirty paper problem, in this chapter, we consider the case where the source and the interference are correlated with a covariance matrix given by

$$\Lambda_{VS} = \begin{pmatrix} \sigma_V^2 & \rho\sigma_V\sigma_S \\ \rho\sigma_V\sigma_S & \sigma_S^2 \end{pmatrix}. \tag{6.5}$$

Under this assumption, separate source and channel coding using DPC naively may not be a good candidate for encoding $\mathbf{V}$ in general. It is due to the fact that in Costa's DPC scheme, the transmitted signal is designed to be orthogonal to the interference and, hence, the DPC scheme cannot exploit the correlation between the source and the interference. Also, the purely uncoded scheme fails to avoid the interference and is suboptimal in general. In this chapter, we first derive an outer bound on the achievable distortion region and then, we propose two joint source-channel coding schemes which exploit the correlation between $\mathbf{V}$ and $\mathbf{S}$, thereby outperforming the naive DPC scheme. The first scheme is a superposition of the uncoded scheme and a digital part formed by a Wyner-Ziv coding [65] followed by

a DPC, which we refer to as a digital DPC based scheme (or just the digital DPC scheme). The second scheme is obtained by replacing the digital part by a hybrid digital and analog (HDA) scheme given in [64] that has been shown to provide graceful improvement when the actual SNR (SNR_a) is better than the design SNR (SNR_d). We then analyze the performance of these two proposed schemes when there is an SNR mismatch. It is shown that both the HDA scheme and the digital DPC scheme benefit from a higher channel SNR and provide graceful enhancement; however, interestingly, for this case neither of schemes dominate the other universally and which one performs better depends on the designed SNR. When ρ is small, the HDA scheme outperforms the digital DPC scheme and when ρ is large, the digital DPC scheme outperforms the HDA scheme. When the channel deteriorates, both the proposed schemes perform identically and are able to provide graceful degradation.

One interesting application of this problem is to derive an achievable distortion region for the generalized cognitive radio channel with correlated sources. This channel can be modeled as a typical two-user interference channel except that one of them knows exactly what the other plans to transmit. Moreover, two users' sources are assumed to be correlated. One can regard the informed user's channel as the setup we consider here and then directly apply the schemes we propose as the coding scheme for the informed user. For the generalized cognitive radio channel with correlated sources, we provide inner and outer bounds on the distortion region where the inner bound largely relies on the coding schemes proposed in this chapter.

The rest of the chapter is organized as follows. In Section 6.2, we present some prior work which is closely related to ours. The outer bound is given in Section 6.3 and two proposed schemes are given in Section 6.4. In Section 6.5, we analyze the performance of the proposed schemes under SNR mismatch. These proposed schemes are then extended to the generalized cognitive radio channel in Section 6.6.

Some conclusions are given in Section 6.7.

6.2 Related Work on JSCC with Interference Known at Transmitter

In [66], Sutivong *et al.* consider the problem of sending a digital source in the presence of interference (or, channel state) which is known at the transmitter and is assumed to be independent of the source. The optimal tradeoff between the achievable rate for transmitting the digital source and the distortion in estimating the interference is then studied. A coding scheme that is able to achieve the optimal tradeoff is also provided in [66]. This coding scheme uses a portion of the power to amplify the interference and uses the remaining power to transmit the digital source via DPC. This coding scheme can be extended to the problem we consider as follows. Since the source and the interference are jointly Gaussian, we can first rewrite the source as $V = \rho\frac{\sigma_V}{\sigma_S}S + N'_\rho$ with S and the innovation N'_ρ being independent of each other. Now if one quantizes N'_ρ into digital data, the setup becomes the one considered by Sutivong *et al.* and their proposed scheme can be applied directly. For any power allocation between the analog part and digital part, using this scheme to operate on the boundary of the optimal tradeoff, the optimal distortion in estimating $\rho\frac{\sigma_V}{\sigma_S}S$ and that in estimating N'_ρ is achieved. The distortion in estimating V for this power allocation strategy is the sum of the above two distortions. One can then optimize the power allocation strategy to get the minimum distortion for this coding scheme. It is worth pointing out that this coding scheme is in general suboptimal for our problem although it achieves the optimal tradeoff between estimating S and N'_ρ individually. This is because, our interest is in estimating V directly and it is importantly to carefully take advantage of the correlation in the estimation error in estimating S and N'_ρ. The coding scheme in [66] is not naturally suited to take advantage of this correlation. One numerical example is shown in Fig. 6.2 where we can

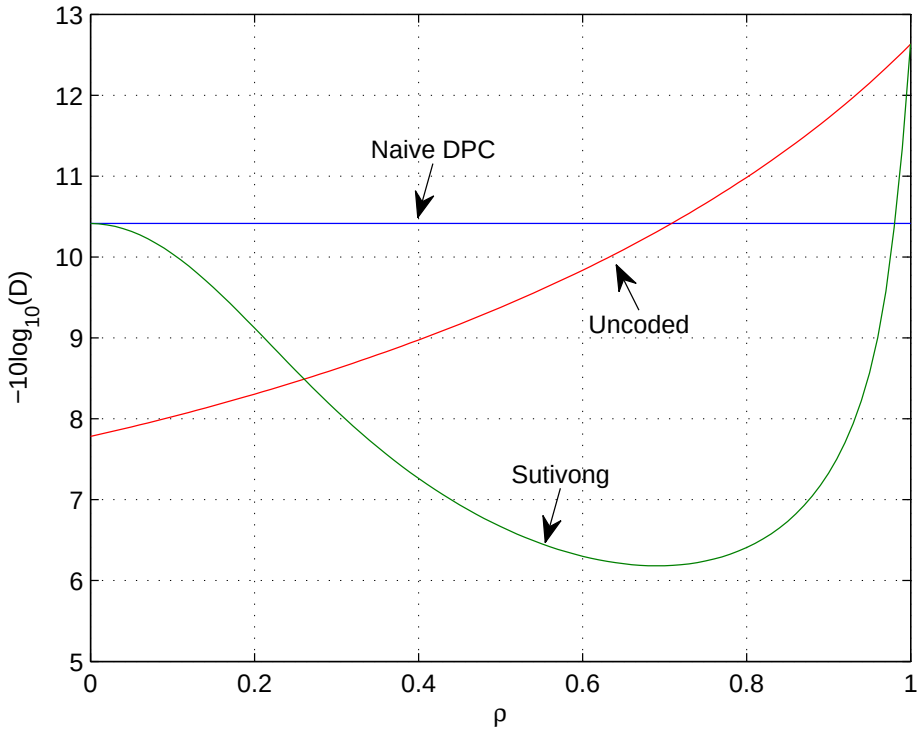

Figure 6.2: Distortions for the naive DPC, the uncoded scheme, and the extension of of Sutivong *et al.*'s scheme.

see that the union of the uncoded scheme and the naive DPC scheme outperforms the extension of Sutivong *et al.*'s scheme.

In [67], Lapidoth *et al.* consider the 2×1 multiple access channel in which two transmitters wish to communicate their sources, which are drawn from a bivariate Gaussian distribution, to a receiver which is interested in reconstructing both sources. There are some similarities between the proposed work and the work in [67] if we regard one of the users', say the user 2's, signal as interference. However, an important difference is that in [67], the transmitters are not allowed to cooperate with each other, i.e., for the transmitter 1, the interference (user 2's signal) is not known. Moreover, this interference now depends on the signalling scheme adopted

at user 2 and may not be correlated to the source anymore.

In [68] [69] [70] [71], transmitting a bi-variate Gaussian source over a 1×2 Gaussian Broadcast Channel is considered. In their setup, the source consists of two components $\mathbf{V}_1$ and $\mathbf{V}_2$ which are memoryless and stationary bi-variate Gaussian random variables and each receiver is only interested in one part of the sources. In [71], Tian *et al.* propose a HDA scheme that achieves the outer bound given in [68] and therefore leads to a complete characterization of the distortion region. This problem is similar to ours if we only focus on one receiver, say the first receiver. However, a crucial difference is that the interference now is a function of $\mathbf{V}_2$ which depends on the broadcast encoding scheme and may not be correlated to $\mathbf{V}_1$. The joint source-channel coding problem for broadcasting a single memoryless Gaussian source under bandwidth mismatch is considered in [72] [73] [74]. However, different from its bandwidth matched counterpart [75], only approximation characterizations of the achievable distortion region are available for this problem. Broadcasting a colored Gaussian source over a colored Gaussian broadcast channel to a digital receiver and a analogy receiver is considered in [76] where Prabhakaran *et al.* propose a HDA scheme that achieves the entire distortion region for the problem they consider.

Joint source-channel coding for point to point communications over Gaussian channels has also been widely discussed. See e.g. [64],[77] [78] [79] [80]. However, they either don't consider interference ([77] [78] [79] [80]) or assume independence of source and interference ([64]). In [64], Wilson *et al.* proposed a HDA coding scheme for the typical writing on dirty paper problem in which the source is independent of the interference. This HDA scheme was originally proposed to perform well in the case of a SNR mismatch. In [64], the authors showed that their HDA scheme not only achieves the optimal distortion in the absence of SNR mismatch but also provides gracefully degradation in the presence of SNR mismatch. In the following

sections, we will discuss this scheme in detail and then propose a coding scheme based on this one.

6.3 Outer Bounds

6.3.1 Outer Bound 1

For comparison, we first present a genie-aided outer bound. This outer bound is derived in a similar way to the one in [70] in which we assume that $\mathbf{S}$ is revealed to the decoder by a genie. Thus, we have

$$\begin{aligned}
\frac{n}{2}\log\frac{\sigma_V^2(1-\rho^2)}{D_{ob}} &\overset{(a)}{\leq} I(\mathbf{V};\widehat{\mathbf{V}}|\mathbf{S}) \\
&\overset{(b)}{\leq} I(\mathbf{V};\mathbf{Y}|\mathbf{S}) \\
&= h(\mathbf{Y}|\mathbf{S}) - h(\mathbf{Y}|\mathbf{S},\mathbf{V}) \\
&= h(\mathbf{X}+\mathbf{Z}|\mathbf{S}) - h(\mathbf{Z}) \\
&\overset{(c)}{\leq} h(\mathbf{X}+\mathbf{Z}) - h(\mathbf{Z}) \\
&\overset{(d)}{\leq} \frac{n}{2}\log\left(1+\frac{P}{N}\right),
\end{aligned} \tag{6.6}$$

where (a) follows from the rate-distortion theorem [30], (b) is from the data processing inequality, (c) is due from that conditioning reduces differential entropy and (d) comes from the fact that Gaussian density maximizes the differential entropy and all random variables involved are i.i.d. Therefore, we have the outer bound as

$$D_{ob,1} = \frac{\sigma_V^2(1-\rho^2)}{1+P/N}. \tag{6.7}$$

Note that this outer bound in general may not be tight for our setup since in the presence of correlation, giving $\mathbf{S}$ to the decoder also offers a correlated version of the source that we wish to estimate. For example, in the case of $\rho = 1$, giving $\mathbf{S}$ to the

decoder implies that the outer bound is $D_{ob} = 0$ no matter what the received signal $\mathbf{Y}$ was. On the other hand, if $\rho = 0$, the setup reduces to the one with uncorrelated interference and we know that this outer bound is tight. Now, we present another outer bound that improves this outer bound for some values of ρ.

6.3.2 Outer Bound 2

Since $S(i)$ and $V(i)$ are jointly Gaussian distributed with covariance matrix given in (6.5), we can write

$$S(i) = \rho\frac{\sigma_S}{\sigma_V}V(i) + N_\rho(i), \tag{6.8}$$

where $N_\rho(i) \sim \mathcal{N}\left(0, (1-\rho^2)\sigma_S^2\right)$ representing the innovation and is independent to $V(i)$. Now, suppose a genie reveals only the n-letter collection of innovation $\mathbf{N}_\rho$ to the decoder, we have

$$\begin{aligned}
\frac{n}{2}\log\frac{\sigma_V^2}{D_{ob,2}} &= \frac{n}{2}\log\frac{\mathbf{var}(V|N_\rho)}{D_{ob}} \\
&\overset{(a)}{\leq} I(\mathbf{V};\widehat{\mathbf{V}}|\mathbf{N}_\rho) \\
&\overset{(b)}{\leq} I(\mathbf{V};\mathbf{Y}|\mathbf{N}_\rho) \\
&= h(\mathbf{Y}|\mathbf{N}_\rho) - h(\mathbf{Y}|\mathbf{N}_\rho,\mathbf{V}) \\
&= h(\mathbf{X} + \rho\frac{\sigma_S}{\sigma_V}\mathbf{V} + \mathbf{Z}|\mathbf{N}_\rho) - h(\mathbf{Z}) \\
&\overset{(c)}{\leq} h(\mathbf{X} + \rho\frac{\sigma_S}{\sigma_V}\mathbf{V} + \mathbf{Z}) - h(\mathbf{Z}) \\
&\overset{(d)}{\leq} \frac{n}{2}\log\left(\frac{\mathbf{var}\left(X + \rho\frac{\sigma_S}{\sigma_V}V + Z\right)}{N}\right) \\
&\overset{(e)}{\leq} \frac{n}{2}\log\left(1 + \frac{(\sqrt{P} + \rho\sqrt{\sigma_S^2})^2}{N}\right),
\end{aligned} \tag{6.9}$$

where (a)-(d) follow from the same reasons with those in the previous outer bound and (e) is due from the Cauchy-Schwartz inequality that states that the maximum occurs when X and V are collinear. Thus, we have

$$D_{ob,2} = \frac{\sigma_V^2}{1 + (\sqrt{P} + \rho\sqrt{\sigma_S^2})^2/N}. \tag{6.10}$$

Note that although the encoder knows the interference $\mathbf{S}$ exactly instead of just $\mathbf{N}_\rho$, the inequality in step (a) does not decrease the knowledge about $\mathbf{S}$ at the transmitter since $\mathbf{S}$ is a deterministic function of $\mathbf{V}$ and $\mathbf{N}_\rho$.

Remark: If $\rho = 0$, this outer bound reduces to the previous one and is tight. If $\rho = 1$, the genie actually reveals nothing to the decoder and the setup reduces to the one considered in [66], i.e., the encoder is interested in revealing the interference to the decoder. For this case, we know that this outer bound is tight. However, this outer bound is in general optimistic except for two extremes. It is due to the fact that in derivations, we assume that we can simultaneously ignore the $\mathbf{N}_\rho$ and use all the power to take advantage of the coherent part. Despite this, the outer bound still provides an insight that in order to build a good coding scheme that one should try to use a portion of power to make use of the correlation and then use the remaining power to avoid $\mathbf{N}_\rho$. Further, it is natural to combine these two outer bounds as $D_{ob} = \max\{D_{ob,1}, D_{ob,2}\}$.

From now on, since the channel we consider is memoryless and all the random variables we consider are i.i.d. in time, i.e. $V(i)$ is independent of $V(j)$ for $i \neq j$, we will drop the index i for the sake of convenience.

6.4 Proposed Schemes

6.4.1 Digital DPC Based Scheme

We now propose a digital DPC scheme which retains the advantages of the above two schemes. This scheme can be regarded as an extended version of the coding scheme in [78] to the setup we consider. As shown in Fig. 6.3, the transmitted signal of this scheme is the superposition of the analog part X_a with power P_a and the digital part X_d with power $P - P_a$. The motivation here is to allocate some power for the analog part to make use of the interference which is somewhat coherent to the source for large ρ's and to assign more power to the digital part to avoid the interference when ρ is small. The analog part is the scaled version of linear combination of source and interference as

$$X_a = \sqrt{a}\left(\gamma V + (1-\gamma)S\right), \tag{6.11}$$

where $P_a \in [0, P]$, $a = P_a/\sigma_a^2$, $\gamma \in [0, 1]$ and

$$\sigma_a^2 = \gamma^2\sigma_V^2 + (1-\gamma)^2\sigma_S^2 + 2\gamma(1-\gamma)\rho\sigma_V\sigma_S. \tag{6.12}$$

The received signal is given by

$$\begin{aligned} Y &= X_d + X_a + S + Z \\ &= X_d + \sqrt{a}\gamma V + \left(1 + \sqrt{a}(1-\gamma)\right) S + Z \\ &= X_d + S' + Z, \end{aligned} \tag{6.13}$$

where X_d is chosen to be orthogonal to S and V and $S' = \sqrt{a}\gamma V + (1 + \sqrt{a}(1-\gamma))\, S$ is the effective interference. The receiver first makes an estimate from Y only as

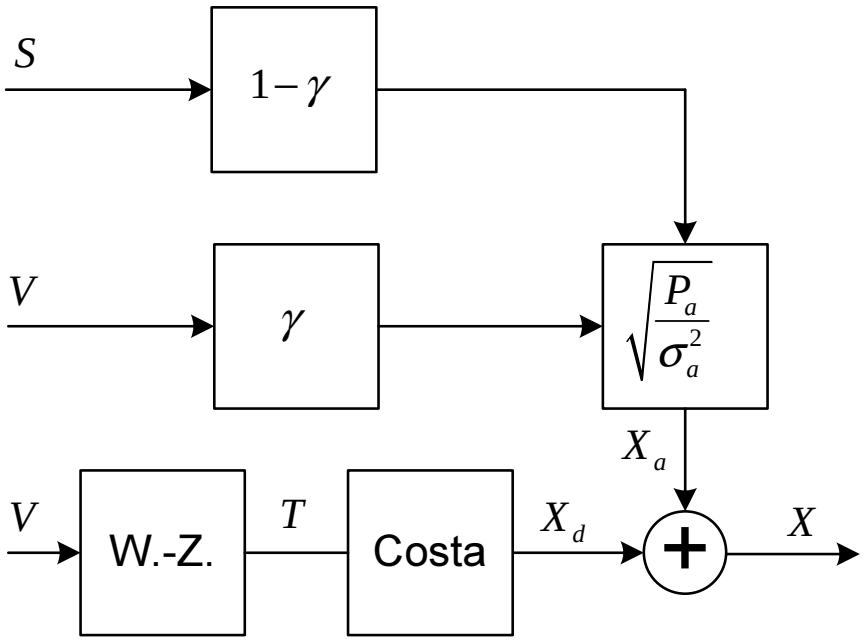

Figure 6.3: Digital DPC scheme.

$V' = \beta Y$ with

$$\beta = \frac{\mathbb{E}[VY]}{\mathbb{E}[Y^2]} = \frac{\sqrt{a}(\gamma\sigma_V^2 + (1-\gamma)\rho\sigma_V\sigma_S) + \rho\sigma_V\sigma_S}{P + N + \sigma_S^2 + +2\sqrt{a}\left((1-\gamma)\sigma_S^2 + \gamma\rho\sigma_V\sigma_S\right)}. \tag{6.14}$$

The corresponding MSE is

$$\begin{aligned} D^* &= \sigma_V^2 - \beta\mathbb{E}[VY] \\ &= \sigma_V^2 \left[1 - \beta\left(\sqrt{a}(\gamma + (1-\gamma)\rho\frac{\sigma_S}{\sigma_V}) + \rho\frac{\sigma_S}{\sigma_V}\right)\right]. \end{aligned} \tag{6.15}$$

Thus, we can write $V = V' + W$ with $W \sim \mathcal{N}(0, D^*)$.

We now refine the estimate through the digital part, which is the concatenation of a Wyner-Ziv coding and a DPC. Since the DPC achieves the rate equal to that when there is no interference at all, the encoder can use the remaining power $P - P_a$

to reliably transmit the refining bits T with a rate arbitrarily close to

$$R = \frac{1}{2}\log\left(1 + \frac{P - P_a}{N}\right). \tag{6.16}$$

The resulting distortion after refinement is then given as

$$D_{sep} = \inf_{\gamma,\ P_a} \frac{D^*}{1 + \frac{P-P_a}{N}}. \tag{6.17}$$

In Appendix C.2, for self-containedness, we briefly summarize the digital Wyner-Ziv scheme to illustrate how to achieve the above distortion.

Remark: The performance of our proposed scheme can be improved by allowing the coefficient in front of S to be negative. i.e., $\gamma > 1$ is allowed. When this coefficient is allowed to be negative, performance similar to that reported in [79] can be achieved.

Remark: Different from the setup considered in [78] that the optimal distortion can be achieved by any power allocation between coded and uncoded transmissions, in our setup the optimal distortion is in general achieved by a particular power allocation which is a function of ρ.

6.4.2 HDA Scheme

Now, let us focus on the HDA scheme obtained by replacing the digital part in Fig. 6.3 by the HDA scheme given in [64]. The analog signal remains the same as in (6.11) and the HDA output is referred to as X_h. Therefore, we have $Y = X_h + S' + Z$. Again, the HDA scheme regards S' as interference and V' described previously as side-information. The encoding and decoding procedures are similar to that in [64] but the coefficients need to be re-derived to fit our setup (the reader is referred to [64] for details).

Let the auxiliary random variable U be

$$U = X_h + \alpha S' + \kappa V, \tag{6.18}$$

where $X_h \sim \mathcal{N}(0, P_h)$ independent to S' and V and $P_h = P - P_a$. The covariance matrix of S' and V can be computed by (6.5).

Codebook Generation: Generate a random i.i.d. codebook $\mathcal{U}$ with 2^{nR_1} codewords, reveal the codebook to both transmitter and receiver.

Encoding: Given realizations $\mathbf{s}'$ and $\mathbf{v}$, find a $\mathbf{u} \in \mathcal{U}$ such that $(\mathbf{s}', \mathbf{v}, \mathbf{u})$ is jointly typical. If such a $\mathbf{u}$ can be found, transmit $\mathbf{x}_h = \mathbf{u} - \alpha\mathbf{s}' - \kappa\mathbf{v}$. Otherwise, an encoding failure is declared.

Decoding: The decoder looks for a $\hat{\mathbf{u}}$ such that $(\mathbf{y}, \mathbf{v}', \hat{\mathbf{u}})$ is jointly typical. A decoding failure is declared if none or more than one such $\hat{\mathbf{u}}$ are found. It is shown in [64] that if $n \to \infty$ and the condition given in (6.21) is satisfied, the probability of $\hat{\mathbf{u}} \neq \mathbf{u} \to 0$.

Estimation: After decoding $\mathbf{u}$, the receiver forms a linear MMSE estimate of $\mathbf{v}$ from $\mathbf{y}$ and $\mathbf{u}$. The distortion is then obtained as

$$D_{hda} = \inf_{\gamma,\ P_a} \left[\sigma_V^2 - \Gamma^T \Lambda_{UY}^{-1} \Gamma\right], \tag{6.19}$$

where Λ_{UY} is the covariance matrix of U and Y, and $\Gamma = [\mathbb{E}[VU], \mathbb{E}[VY]]^T$.

In the encoding step, to make sure the probability of encoding failure vanishes

with increasing n, we require

$$\begin{aligned} R_1 &> I(U; S', V) \\ &= h(U) - h(X_h + \alpha S' + \kappa V | S', V) \\ &\overset{(a)}{=} h(U) - h(X_h) \\ &= \frac{1}{2} \log \frac{\mathbb{E}[U^2]}{P_h}, \end{aligned} \tag{6.20}$$

where (a) follows because X_h is independent of S' and V.

Further, to guarantee the decodability of U in the decoding step, one requires

$$\begin{aligned} R_1 &\overset{(a)}{<} I(U; Y, V') \\ &= h(U) - h(U | Y, V') \\ &= h(U) - h(U - \alpha Y - \kappa V' | Y, V') \\ &\overset{(b)}{=} h(U) - h(\kappa W + (1-\alpha) X_h - \alpha Z | Y), \end{aligned} \tag{6.21}$$

where (a) follows from the error analysis of $\mathcal{E}_3$ in Section III of [81] and (b) is due to the fact that $V' = \beta Y$. By choosing

$$\alpha = \frac{P_h}{P_h + N}, \quad \kappa^2 = \frac{P_h^2}{(P_h + N) D^*}, \tag{6.22}$$

one can verify that (6.20) and (6.21) are satisfied. Note that in (6.20) what we really need is $R_1 \geq I(U; S', V) + \varepsilon$ and in (6.21) it is $R_1 \leq I(U; Y, V') - \delta$. However, since ε and δ can be made arbitrarily small, these are omitted for the sake of convenience and to maintain clarity.

Remark: It is shown in Appendix C.1 that the distortions in (6.17) and (6.19) are exactly the same. However, as we will see in the next section, two schemes perform

differently when $\text{SNR}_a > \text{SNR}_d$.

6.4.3 Numerical Results

In Fig. 6.4, we plot the distortion (in $-10\log_{10}(D)$) for coding schemes and outer bounds described above as a function of SNR. In this figure, we set $\sigma_V^2 = \sigma_S^2 = 1$ and $\rho = 0.3$. Note that for this choice of σ_V^2, what we plot is actually the signal-to-distortion ratio. As expected, the two proposed schemes have exactly the same performance. Moreover, for this case, these two schemes not only outperform others but also approach the outer bound (maximum of two) very well.

We then fix the SNR and plot the distortion as a function of ρ in Fig. 6.5. The parameters are set to be $\sigma_V^2 = \sigma_S^2 = 1$, $P = 10$, and $N = 1$. It can be seen that both the proposed schemes perform exactly the same and that the achievable distortion region with the proposed scheme is larger than what is achievable with a separation based scheme using DPC and a uncoded scheme. Further, although the proposed schemes perform close to the outer bound over a wide range of ρs, the outer bound and the inner bound do not coincide however, leaving room for improvement either of the outer bound or the schemes.

6.5 Performance Analysis in the Presence of SNR Mismatch

In this section, we study the distortions for the proposed schemes in the presence of SNR mismatch i.e., we consider the scenario where instead of knowing the exact channel SNR, the transmitter only knows a lower bound on the channel SNR. Specifically, we assume that the actual channel noise to be $Z_a \sim \mathcal{N}(0, N_a)$ but the transmitter only knows that $N_a \leq N$ so that it designs the coefficients assuming the noise variance is N. In what follows, we analyze the performance for both proposed schemes under the above assumption.

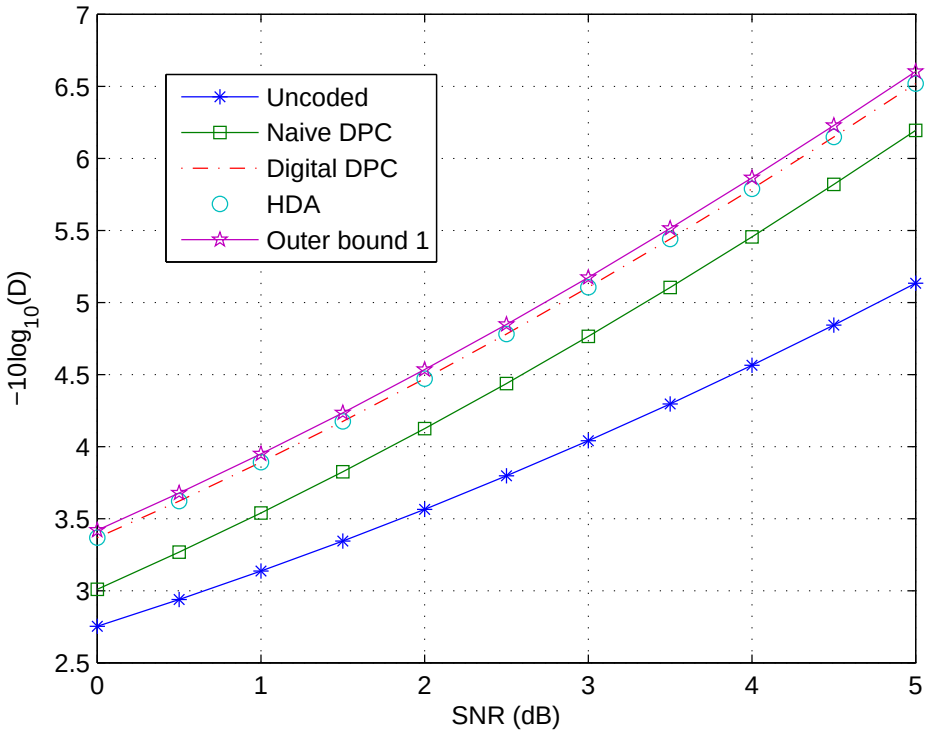

Figure 6.4: $\frac{P}{N}$ vs D with $\sigma_V^2 = \sigma_S^2 = 1$ and $\rho = 0.3$.

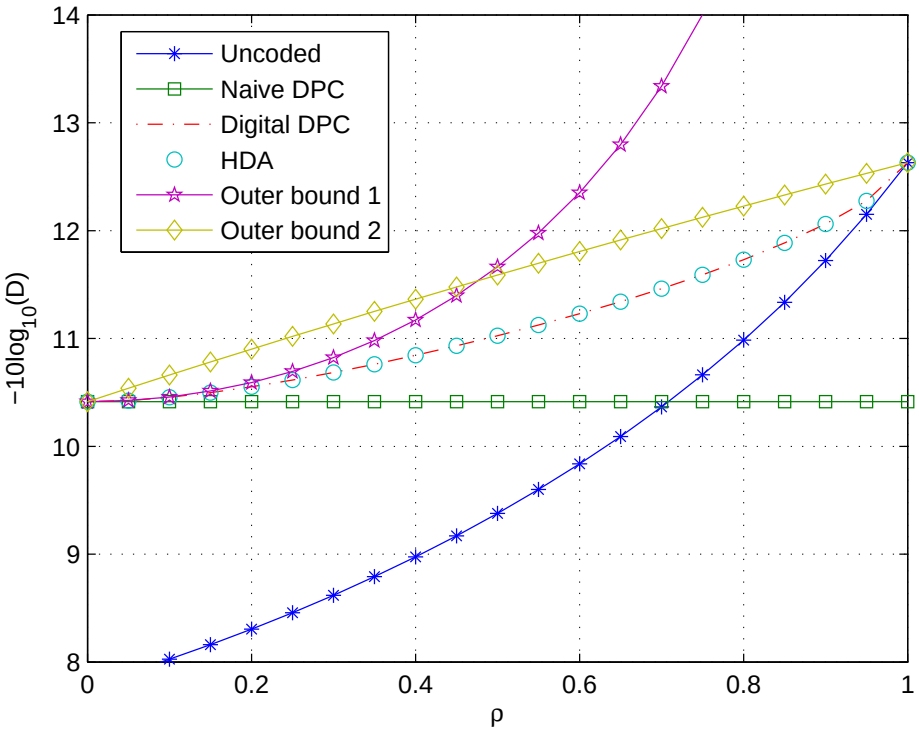

Figure 6.5: ρ vs D with $\sigma_V^2 = \sigma_S^2 = 1$ and $\frac{P}{N} = 10$.

6.5.1 Digital DPC Based Scheme

Since the transmitter designs its coefficients for N, it aims to achieve the distortion D_{sep} given in (6.17). It first quantizes the source to T by a Wyner-Ziv coding with side-information D^* given in (6.15) and then encodes the quantization output by a DPC with a rate

$$R = \frac{1}{2}\log\left(1 + \frac{P - \tilde{P}_a}{N}\right), \tag{6.23}$$

where $\tilde{P}_a$ is the power allotted to X_a such that the distortion in the absence of SNR mismatch is minimized. i.e.,

$$\tilde{P}_a = \arg\inf_{P_a} \frac{D^*}{1 + \frac{P - P_a}{N}}. \tag{6.24}$$

At receiver, since $N_a \leq N$, the DPC decoder can correctly decode T with high probability. Moreover, the receiver forms the MMSE estimate of V from Y as $V_a' = \beta_a Y$ with β_a and the corresponding MSE D_a^* derived by substituting N_a for N in (6.14) and (6.15), respectively. After that, the problem reduces to the Wyner-Ziv problem with mismatched side-information. In Appendix C.3, we show that for this problem, one can achieve

$$D_{sep,mis} = \frac{D^* D_a^*}{D^* D_a^* + (D^* - D_a^*) D_{sep}} D_{sep}. \tag{6.25}$$

Unlike the typical separation-based scheme that we have seen in [64], the proposed digital DPC scheme (whose digital part can be regarded as a separation-based scheme) can still take advantage of better channels through mismatched side-information.

6.5.2 HDA Scheme

Different from the digital DPC scheme, in the presence of SNR mismatch, the performance analysis of the HDA scheme cannot be converted to the Wyner-Ziv problem with mismatched side-information. It is because that in the HDA scheme, we jointly form an estimate of V from U and Y. Fortunately, as shown in [64], the HDA scheme is capable of making use of an SNR mismatch.

Similar to the digital DPC scheme, we design the coefficients for noise variance N. The HDA scheme regards D^* as side-information and S' as interference. It generates the auxiliary random variable U given by (6.18) with coefficients described by (6.22). Since $N_a \leq N$, the receiver can correctly decode U with high probability. The receiver then forms the MMSE as described in (6.19). Note that $\mathbb{E}[Y^2]$ in Λ_{UY} should be modified appropriately to address the fact that the actual noise variance is N_a in this case.

Remark: In [64], the authors compare the distortions of the digital scheme and the HDA scheme in estimating the source V and the interference S as we move away from the designed SNR. One important observation is that the HDA scheme outperforms the separation-based scheme in estimating the source; however, the separation-based scheme is better than the HDA scheme if one is interested in estimating the interference. Here, since the *effective* interference S' includes the uncoded signal $\sqrt{a}V$ in part and the source is correlated to the interference, estimating the source V is equivalent to estimating a part of S'. Thus, one can expect that if P_a and ρ are large enough, the digital DPC scheme may outperform the HDA scheme in the presence of SNR mismatch. One the other hand, if P_a and ρ are relatively small, one can expect the reverse.

Remark: Note that we have only discussed the case when the actual channel turns

out to be better than that expected by the transmitter. On the other hand, when the channel deteriorates, the digital DPC scheme and the HDA scheme are not able to decode the digital part and the HDA part, respectively. For the digital DPC scheme, this is due to the fact that a capacity-approaching code is used so that the decoding will fail if the channel is no longer being able to support this rate. For the HDA scheme, this inability to decode U is because the constraint (6.21) is no longer satisfied if the channel is worse than that expected. However, both schemes can still form the MMSE estimate of the source from the received signal Y. Therefore, for a same choice of P_a, the resulting distortion of two proposed schemes would be the same and is equal to D_a^*. This implies that both the proposed schemes are able to provide graceful degradation when channel deteriorates.

6.5.3 Numerical Results

Now, we compare the performance of the above two schemes and the scheme that knows the actual SNR. The parameters are set to be $\sigma_V^2 = \sigma_S^2 = 1$. We plot $-10\log_{10}(D)$ as we move away from the designed SNR for both $\rho = 0.1$ and $\rho = 0.5$ cases. Two examples for designed SNR = 0 dB and 10 dB are given in Fig. 6.6 and Fig. 6.7, respectively.

In Fig. 6.6, we consider the case that the designed SNR is 0 dB which is relatively small compared to the variance of interference. We plot both the cases when channel improves and deteriorates in this figure. When channel deteriorates, as discussed in Remark 6.5.2, both the proposed scheme can provide a distortion equal to D_a^*. For the case when channel improves, we can see that which scheme performs better in the presence of SNR mismatch really depends on ρ. It can be explained by the observations made in *Remark* 6.5.2 and the power allocation strategy. For this case the optimal power allocation $\tilde{P}_a$ is proportional to ρ. For $\rho = 0.1$ case, since the

correlation is small and the assigned $\tilde{P}_a$ is also small, the HDA scheme is better than the digital DPC scheme. On the other hand, for $\rho = 0.5$ case, we allot a relatively large power to $\tilde{P}_a$ so that one may get a better estimate if we try to use the digital DPC scheme to estimate a part of S'. This property is further discussed in the Appendix C.4.

In Fig. 6.7, we design the coefficients for SNR = 10 dB which can be regarded as relatively large SNR compared to the variance of interference. Here, we only plot the case when channel improves for the sake of brevity. For this case, the optimal power allocation $\tilde{P}_a$ for both $\rho = 0.1$ and $\rho = 0.5$ are relatively small. Therefore, the performance improvement provided by the HDA scheme is larger than that provided by the digital DPC scheme for both cases. It is worth noting that in Fig. 6.6 and Fig. 6.7, both two proposed schemes can provide a performance improvement although some of them may be hard to observe due to that the scale of the y-axis is in dB.

In Fig. 6.8, we plot the performance of the proposed schemes with different choices of P_a for the same channel parameters with those in the previous figure for $\rho = 0.1$. We observe that for both schemes, if we compromise the optimality at the designed SNR, it is possible to get better slopes of distortion than that obtained by setting $P_a = \tilde{P}_a$. In other words, we can obtain a family of achievable distortion under SNR mismatch by choosing $P_a \in [0, P]$.

6.6 JSCC for the Generalized Cognitive Radio Channel

An interesting application of the joint source-channel coding problem considered in this chapter is in the transmission of analog sources over a cognitive radio channel. In this section, we will first formally state the problem, derive an outer bound on the achievable distortion region, and then propose a coding scheme based on the schemes

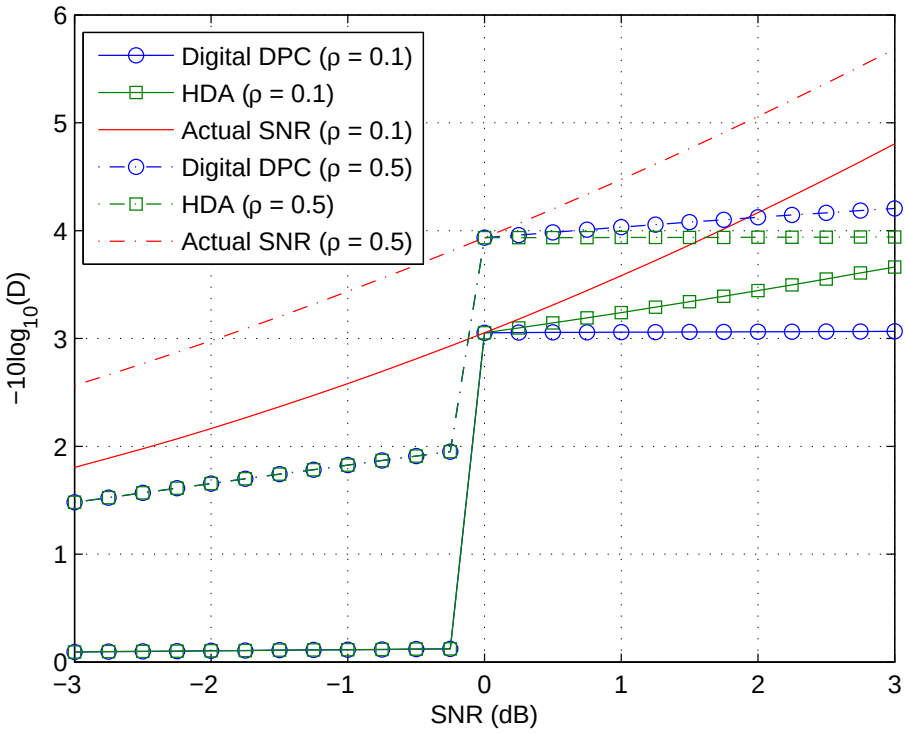

Figure 6.6: SNR mismatch case for SNR = 0dB.

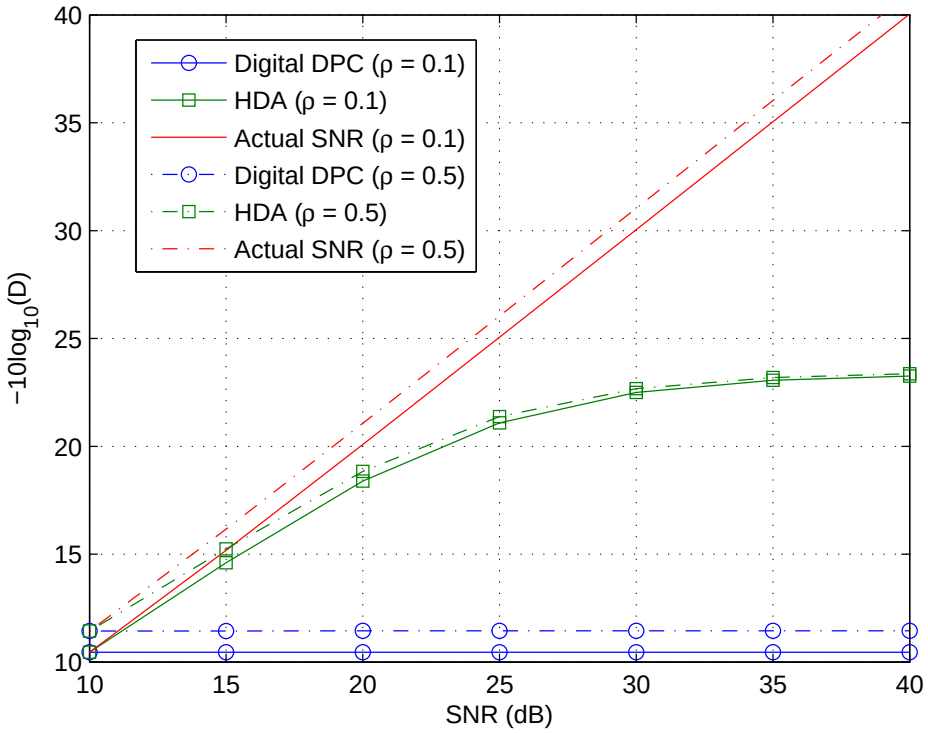

Figure 6.7: SNR mismatch case for SNR = 10dB.

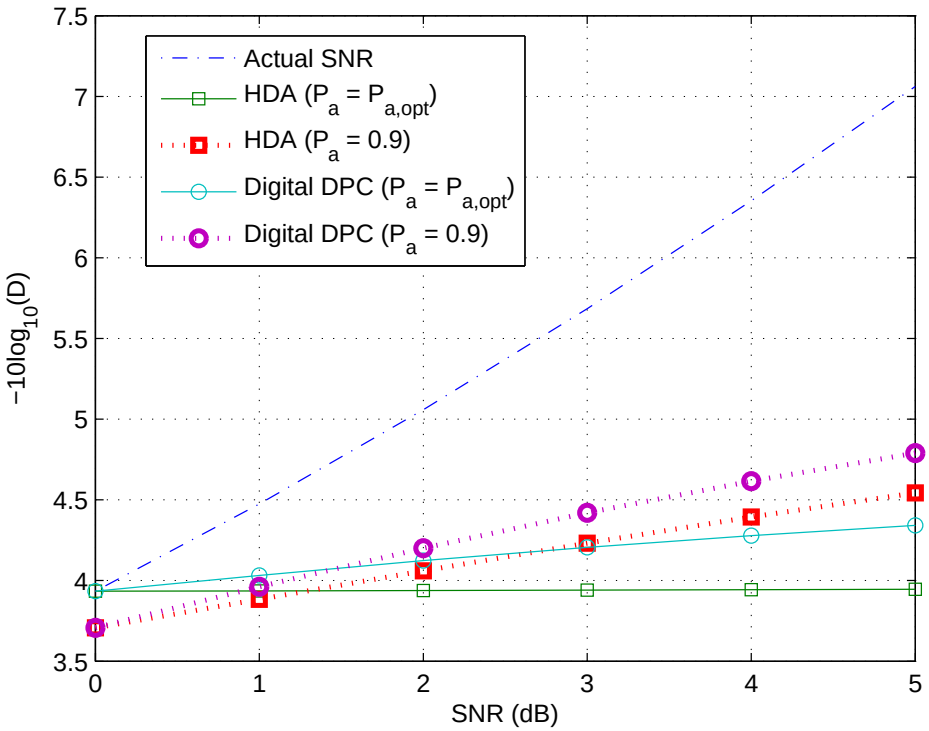

Figure 6.8: Proposed schemes with different choices of P_a.

given in Section 6.4.

6.6.1 Problem Statement

Recently, there has been a lot of interest in cognitive radio since it was proposed in [82] for flexible communication devices and higher spectral efficiency. In a conventional cognitive radio setup, the lower priority user (usually referred to as the secondary user) listens to the wireless channel and transmits the signal only through the spectrum not used by the higher priority user (referred to as the primary user). In a generalized cognitive radio channel, simultaneous transmission over the same time and frequency is allowed. As shown in Fig. 6.9, the problem can be modeled as an interference channel with direct channel gain 1 and cross channels h_1 and h_2 representing the real-valued channel gains from user 1 to user 2 and vice versa, respectively. The average power constraints imposed on the outputs of user 1 and 2 are P_1 and P_2, respectively. Different from interference channels, in cognitive radio channels, we further assume that the secondary user knows V_1 non-causally. Here, we also assume that the channel coefficient h_1 is known by the secondary user. The received signals are given by

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & h_1 \\ h_2 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}. \tag{6.26}$$

where $Z_i \sim \mathcal{N}(0,1)$ for $i \in \{1,2\}$. The capacity region of this channel has been studied and is known for some special cases, e.g., the weak interference case [83] [84], the very-strong interference case [85], and the primary-decode-cognitive case [86].

In this section, we consider the same generalized cognitive radio channel but our focus is on the case when both users have *analog* information V_1 and V_2, respectively. We are interested in the distortion region which describes how much distortion two

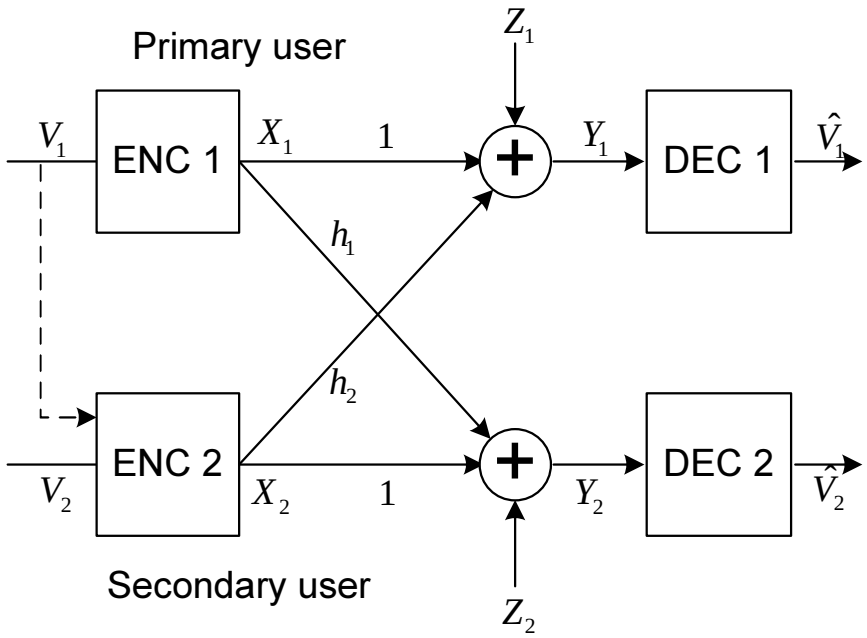

Figure 6.9: System model for a cognitive radio channel.

users can achieve simultaneously. In particular, we consider the case when the two sources are correlated with a covariance matrix given by

$$\Lambda_{V_1V_2} = \begin{pmatrix} \sigma_{V_1}^2 & \rho\sigma_{V_1}\sigma_{V_2} \\ \rho\sigma_{V_1}\sigma_{V_2} & \sigma_{V_2}^2 \end{pmatrix}. \tag{6.27}$$

The distortion measure is the MSE distortion measure defined in (6.4). An achievable distortion region can be obtained by first enforcing the primary user to use the uncoded scheme and using the proposed schemes given in section 6.4 for the secondary user. In fact, since the primary user does not have any side-information, analog transmission is an optimal choice [75] [87] in terms of the distortion achieved at the primary receiver. Further notice that since we do not consider SNR mismatch here, it makes no difference which proposed scheme we use.

6.6.2 Outer Bound

In this subsection, we derive an outer bound on the distortion region for the generalized cognitive radio channel with $nR_1 = I(\mathbf{X}_1; \mathbf{Y}_1)$ and $nR_2 = I(\mathbf{X}_2; \mathbf{Y}_2|\mathbf{X}_1)$. Then, for the primary user, we have

$$\begin{aligned} \frac{n}{2}\log\frac{\sigma_{V_1}^2}{D_1} &\overset{(a)}{\leq} I(\mathbf{V}_1; \widehat{\mathbf{V}}_1) \\ &\overset{(b)}{\leq} I(\mathbf{X}_1; \mathbf{Y}_1) \\ &= nR_1, \end{aligned} \tag{6.28}$$

where (a) follows from rate distortion theory and (b) follows from the data processing inequality. Also, for the secondary user, we have

$$\begin{aligned} \frac{n}{2}\log\frac{\sigma_{V_2}(1-\rho^2)}{D_2} &\leq I(\mathbf{V}_2; \widehat{\mathbf{V}}_2|\mathbf{V}_1) \\ &\overset{(a)}{=} I(\mathbf{V}_2; \widehat{V}_2|\mathbf{V}_1, \mathbf{X}_1) \\ &\overset{(b)}{=} I(\mathbf{V}_2; \widehat{\mathbf{V}}_2|\mathbf{X}_1) \\ &\overset{(c)}{\leq} I(\mathbf{X}_2; \mathbf{Y}_2|\mathbf{X}_1) \\ &= nR_2, \end{aligned} \tag{6.29}$$

where (a) is due to the fact that $\mathbf{X}_1$ is a deterministic function of $\mathbf{V}_1$, (b) follows from the Markov chain $\mathbf{V}_1 \leftrightarrow (\mathbf{X}_1, \mathbf{X}_2) \leftrightarrow (\mathbf{Y}_1, \mathbf{Y}_2)$, and (c) follows from the data processing inequality. Thus, we have

$$D_{ob1} = \frac{\sigma_{V_1}^2}{R_1}, \tag{6.30}$$

$$D_{ob2} = \frac{\sigma_{V_2}^2(1-\rho^2)}{R_2}, \tag{6.31}$$

where (R_1, R_2) must lie inside the capacity region of the generalized cognitive radio channel.

As mentioned earlier, the capacity region of this channel setup is only known for some special cases. Fortunately, for those cases whose capacity regions remain unknown, outer bounds on (R_1, R_2) are available (see e.g. [86] wherein the authors give an unified view of outer bounds for different cases) and therefore we can still obtain the outer bound given in (6.30) and (6.31).

6.6.3 Proposed Coding Scheme

Let the primary user simply transmit the scaled version of the uncoded source $X_1 = \sqrt{P_1/\sigma_{V_1}^2}V_1$. Therefore, the bottom channel in Fig. 6.9 reduces to the situation we considered in the previous section with source $V = V_2$ and interference $S = h_1 X_1$. The covariance matrix becomes (6.5) with

$$\sigma_V^2 = \sigma_{V_2}^2, \tag{6.32}$$

$$\sigma_S^2 = h_1^2 P_1. \tag{6.33}$$

The secondary user then encodes its source to X_2 by the HDA scheme described previously in section 6.4.2 with power $P_2 = P_h + P_a$ and coefficients according to (6.22). With these coefficients, the corresponding distortion D_2 is computed by (6.19). At the receiver 1, the received signal is

$$\begin{aligned} Y_1 &= X_1 + h_2 X_2 + Z_1 \\ &= \left(1 + (1-\gamma)\sqrt{a}h_1 h_2\right) X_1 + h_2 X_h + h_2\sqrt{a}\gamma V_2 + Z_1. \end{aligned} \tag{6.34}$$

Decoder 1 then forms a linear MMSE estimate from Y_1 given by $\widehat{V}_1 = \beta_1 Y_1$, where $\beta_1 = \mathbb{E}[V_1 Y_1]/\mathbb{E}[Y_1^2]$ and

$$\mathbb{E}[V_1 Y_1] = \left(1 + (1-\gamma)\sqrt{a}h_1 h_2\right)\sqrt{P_1 \sigma_{V_1}^2} + h_2\sqrt{a}\gamma\rho\sigma_{V_1}\sigma_{V_2} \tag{6.35}$$

$$\begin{aligned}\mathbb{E}[Y_1^2] = &\left(1 + (1-\gamma)\sqrt{a}h_1 h_2\right)^2 P_1 + a h_2^2 \gamma^2 \sigma_{V_2}^2 + \\ &h_2^2 P_h + 2\sqrt{a}h_2\gamma\rho\sqrt{P_1\sigma_{V_2}^2}\left(1 + (1-\gamma)\sqrt{a}h_1 h_2\right) + N_1.\end{aligned} \tag{6.36}$$

Therefore, the corresponding distortion is $D_1 = \sigma_{V_1}^2 - \beta_1 \mathbb{E}[V_1 Y_1]$.

It can be verified that assigning $\gamma = 1$ may lead to a suboptimal D_1 in general. Thus, as we mentioned in Section 6.4.1, one may want to assign a non-zero power to transmit S in order to achieve a larger distortion region. We can then optimize the power allocation for particular performance criteria. For instance, if one desires achieving the minimum distortion for the secondary user, γ should be set to be 1. However, if the aim is to obtain the largest achievable distortion region, one should optimize over $P_a \in [0, P_1]$ and $\gamma \in [0, 1]$.

6.6.4 Discussions and Numerical Results

Here, we give examples to compare the performance of the outer bound and the proposed coding scheme for two cases whose capacity region is known, namely the weak interference case and very-strong interference case. Also, similar to [84], we present the distortion for the secondary user under the coexistence conditions.

1. Weak interference case: When the interference is weak, i.e., $|h_2| \leq 1$, the capacity region is given by [83] [84]

$$\begin{aligned}R_1 &\leq \frac{1}{2}\log\left(1 + \frac{P_1(1 + h_2\rho_x\sqrt{P_2/P_1})^2}{1 + (1-\rho_x^2)h_2^2 P_2}\right) \\ R_2 &\leq \frac{1}{2}\log\left(1 + (1-\rho_x^2)P_2\right),\end{aligned} \tag{6.37}$$

where $\rho_x \in [0,1]$. One can see that for a fixed ρ_x, increasing R_2 will not affect R_1. Thus, for this case, the outer bounds in (6.30) and (6.31) become

$$D_{ob1} = \frac{\sigma_{V_1}^2}{1 + \frac{P_1(1+h_2\rho_x\sqrt{P_2/P_1})^2}{1+(1-\rho_x^2)h_2^2P_2}}, \tag{6.38}$$

$$D_{ob2} = \frac{\sigma_{V_2}^2(1-\rho^2)}{1+(1-\rho_x^2)P_2}. \tag{6.39}$$

One example of the distortion region for this case is shown in Fig. 6.10 in which we plot the outer bound and the boundary of the distortion region achieved by the proposed coding scheme. The parameters are set to be $\sigma_{V_1}^2 = \sigma_{V_2}^2 = 1$, $h_1 = h_2 = 0.5$, and the power constraints are $P_1 = P_2 = 1$. In this figure, One can observe that when $\rho = 0$, the outer bound is tight and the proposed coding scheme is optimal. However, the inner and outer bound do not coincide for other ρs and one can see that the gap increases as ρ increases.

2. Very-strong interference case: The channel is said to be in the very-strong interference regime if the following conditions are satisfied,

$$|h_2| \geq 1, \tag{6.40}$$

$$|h_1\sqrt{P_1/P_2} + 1| \geq |\sqrt{P_1/P_2} + h_2|, \tag{6.41}$$

$$|h_1\sqrt{P_1/P_2} - 1| \geq |\sqrt{P_1/P_2} - h_2|. \tag{6.42}$$

The capacity region of this case is the union of (R_1, R_2) satisfying [85]

$$\begin{aligned} R_2 &\leq \frac{1}{2}\log\left(1+(1-\rho_x^2)P_2\right) \\ R_1 + R_2 &\leq \frac{1}{2}\log\left(1+P_1+h_2^2P_2+2\rho_x h_2\sqrt{P_1P_2})^2\right), \end{aligned} \tag{6.43}$$

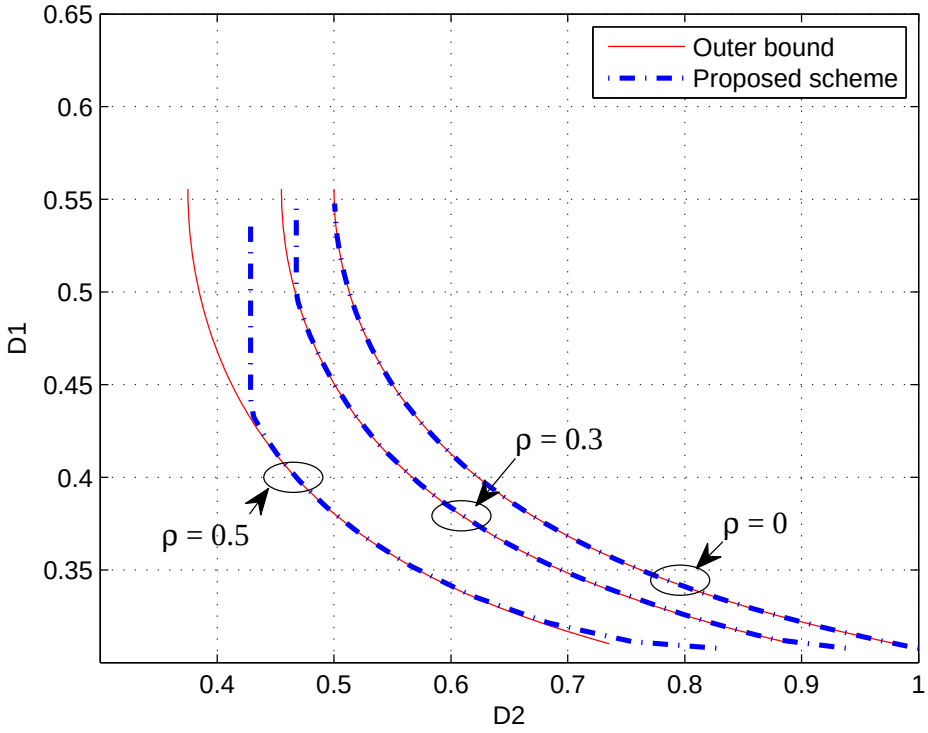

Figure 6.10: Distortion region for the weak interference case with $P_1 = P_2 = 1, \sigma^2_{V_1} = \sigma^2_{V_2} = 1$, and $h_1 = h_2 = 0.5$.

where $\rho_x \in [0, 1]$. For this case, different choices of R_2 may lead to different upper bounds for R_1. Thus, the outer bound can be obtained by collecting all the Pareto minimal points of (D_1, D_2) among all choices of (R_1, R_2) and ρ_x.

In Fig. 6.11, the outer bound and the boundary of the distortion region achieved by the proposed scheme are plotted. All the parameters are set to be the same as those in the previous figure except for $h_1 = h_2 = 1.5$ now. It is easy to see that (6.40)-(6.42) are satisfied under these parameters. One can see that for this case the inner and outer bound do not coincide even for $\rho = 0$ case. This may be due to the fact that in the proposed coding scheme, the primary decoder treats the signal from the secondary user as extra noise. This violates the insight of the very-strong interference regime that one should first decode interfering signal and then cancel it out since the interference is "very-strong" and is regarded as easier to decode. However, for the proposed scheme, the primary decoder is not able to obtain an improvement from this decoding strategy. This is because the digital part (or the HDA part, depends on which scheme is used) of the interfering signal is a function of V_1 and the bin index (or U). Therefore, decoding the bin index (or U) only is not enough to reconstruct X_d (or X_h).

On the other hand, if one simply ignores the correlation and uses an optimal separate source-channel code at the secondary user, this coding scheme is guaranteed to achieve the outer bound for $\rho = 0$ but this scheme is unable to adapt with ρ, i.e., the performance is fixed for all ρs. Therefore, when ρ is large, one may obtain a lower distortion by using the proposed scheme although it fails to achieve the outer bound for any ρ. One example is given in Fig. 6.11 that when $\rho = 0.5$, the distortion region achieved by the proposed scheme is larger than that achieved by an optimal separate coding scheme (whose performance is the same as the outer bound for $\rho = 0$). It is interesting to build a coding scheme that achieves the outer bound for $\rho = 0$ and

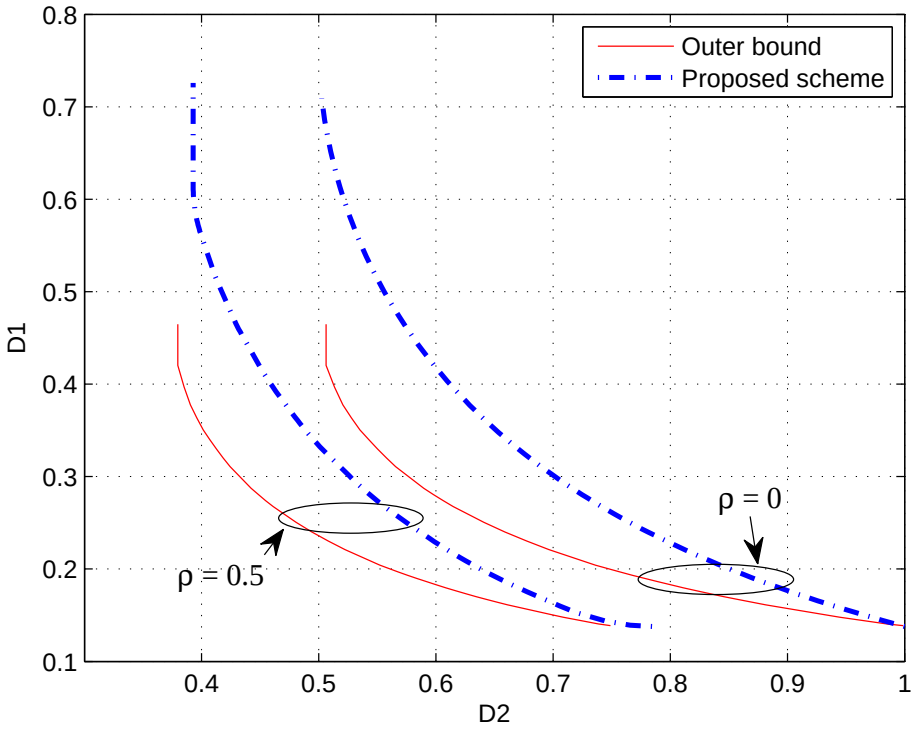

Figure 6.11: Distortion region for the very-strong interference case with $P_1 = P_2 = 1, \sigma^2_{V_1} = \sigma^2_{V_2} = 1$, and $h_1 = h_2 = 1.5$.

is capable of adapting with ρ for the very-strong interference case; however, this is beyond the scope of this chapter.

3. Coexistence conditions: In [84], the coexistence conditions are introduced to understand the system-wise benefits of cognitive radio. The authors study the largest rate that the cognitive radio can achieve under these coexistence conditions described as follows.

$\diamond$ the presence of cognitive radio should not create rate degradation for the primary user, and

$\diamond$ the primary user does not need to use a more sophisticated decoder than it would use in the absence of the cognitive radio. i.e, a single-user decoder is enough.

Similar to this idea, we study the distortion of the secondary user under the following conditions

♣ the presence of cognitive radio should not create distortion increment for the primary user, and

♣ the primary user uses a single-user decoder.

We present the outer bound and the signal-to-distortion ratio for the secondary user obtained by the proposed scheme under coexistence conditions. Here the outer bound is given by

$$D_{ob2,coexist} = \inf_{D_{ob1} \leq \frac{\sigma_{V_1}^2}{1+P_1}} D_{ob2}, \tag{6.44}$$

where D_{ob1} and D_{ob2} are given in (6.30) and (6.31), respectively, and R_1 and R_2 therein can be further bounded by the capacity region or upper bounds on the capacity region as mentioned. Note that when taking the infimum, we simply constrain the distortion of the primary user to be at most the one achieved when there is no interference at all and ignore the second coexistence condition. i.e., this outer bound allows the primary decoder to be any possible decoder, not necessary a single-user decoder.

In Fig. 6.12 and Fig. 6.13, the achievable distortion for the secondary user is plotted for the same set of parameters as in Fig. 6.10 and Fig. 6.11, respectively. As shown in these figures, the proposed scheme is able to increase the secondary user's signal-to-distortion ratio without degrading the performance of the primary user. Moreover, one can observe that at $\rho = 0$ the proposed coding is optimal for the weak interference case but not for the very-strong interference case. This may be due to the fact that in the proposed coding scheme the interfering signal is not fully decoded. This may also be the consequence of ignoring the second condition when deriving the outer bound. Another interesting observation is that in Fig. 6.13, the signal-to-

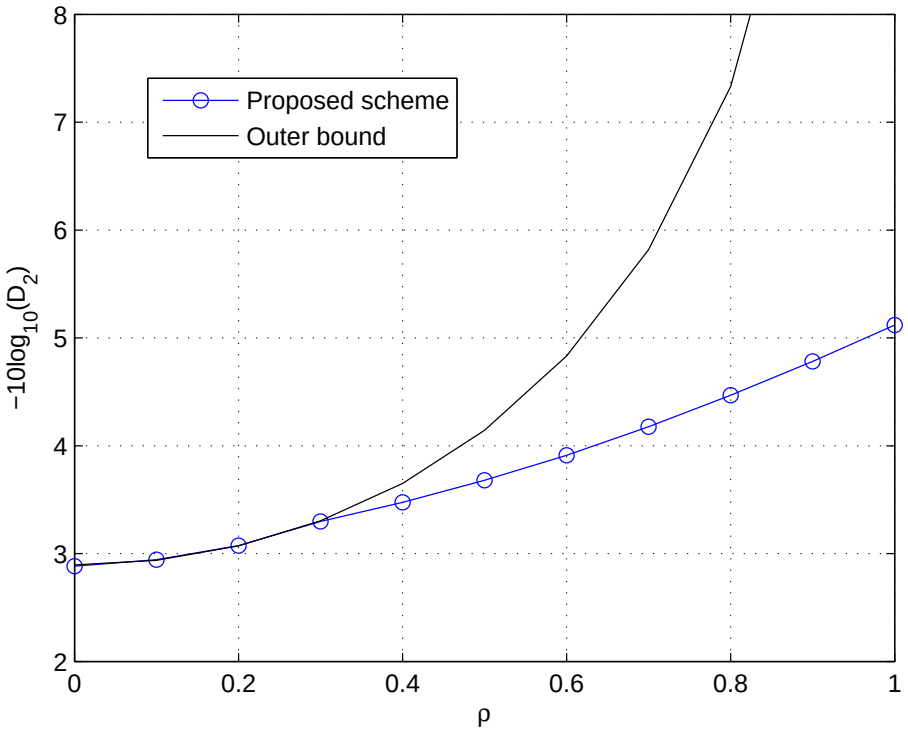

Figure 6.12: Distortion region for the weak interference case under coexistence conditions, $h_1 = h_2 = 0.5$.

distortion ratio increases more rapidly than that in Fig. 6.12. This is because in the very-strong interference case, the channel would amplify the secondary user's signal much more than that in the weak interference case. So the secondary user could use less power to boost the primary signal such that the coexistence conditions are satisfied and then use the remaining power to decrease its own distortion.

6.7 Conclusions

In this chapter, we have discussed the joint source-channel coding problem with interference known at the transmitter. In particular, we considered the case that the source and the interference are correlated with each other. We proposed a digital

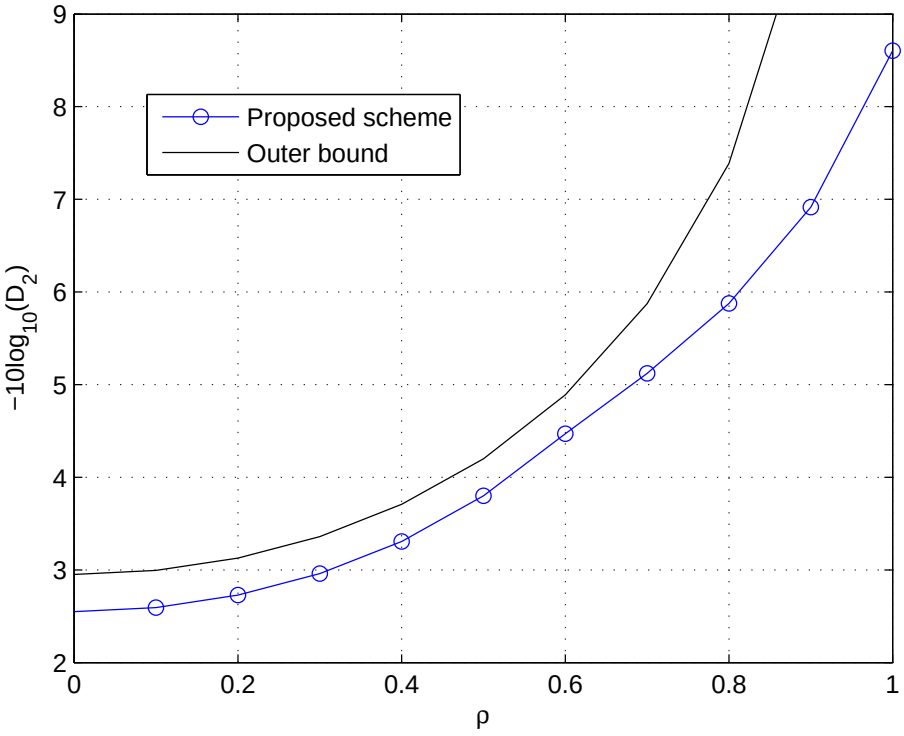

Figure 6.13: Distortion region for the very-strong interference case under coexistence conditions, $h_1 = h_2 = 1.5$.

DPC scheme and a HDA scheme and showed that both two schemes can adapt with ρ. The performance of these two schemes under SNR mismatch are also discussed. Different from typical separation-based schemes which are not able to take advantage of a better channel SNR and suffer from abrupt degradation when the channel deteriorates, both the proposed schemes can benefit from a better side-information acquired at the decoder and also provide a graceful degradation and improvement under SNR mismatch. However, there is a difference between the performance of the two proposed schemes when $\text{SNR}_a > \text{SNR}_d$ and which scheme is better depends on the designed SNR and ρ.

These two schemes are then applied to the generalized cognitive radio channel for deriving an achievable distortion region. Outer bounds on distortion region for this channel are also provided. To the best of our knowledge, this is the first joint source-channel coding scheme that has been proposed for the generalized cognitive radio channel. Numerical results suggest that, in the weak interference regime, the gap between the inner and outer bound is reasonably small for small and medium ρ and increases as ρ increases. Moreover, in the very-strong interference regime, there exist ρs such that the proposed joint source-channel coding scheme outperforms optimal separate coding scheme. The system-wise benefits of cognitive radio in terms of distortion are also studied via imposing the coexistence conditions.

7. CONCLUSIONS

The bidirectional relay channel with ISI and MIMO bidirectional relay channel have been studied in the first part of this dissertation. We have extended the compute-and-forward strategy to these scenarios which has been very successful in memoryless bidirectional relay channel. For the bidirectional relay channel with ISI, we proposed a time-domain approach that prefilters the input codewords so that the equivalent ISI channel seen by the relay is perfectly aligned. After this, the lattice precoding and an unbiased MMSE equalizer was adopted so that the problem can be converted into a memoryless bidirectional relay channel; hence, the compute-and-forward strategy can be directly carried out. Our results have shown that the bidirectional relay channel with memory is inseparable, i.e., in order to achieve the capacity, joint processing across sub-channels other than power allocation is required.

We then studied a relevant model for the considered scenarios, namely the parallel bidirectional relay channels. The corresponding linear deterministic model has been considered and its capacity region has been fully characterized. Based on the insight obtained from the study of the deterministic model, we proposed two coding schemes that are able to approach the capacity region to within a constant gap. This led us to an approximate characterization of the capacity region for the considered setting.

After establishing the information-theoretic results, we focused on designing practically implementable compute-and-forward schemes. The separation-based compute-and-forward framework that we previously proposed was considered. We proposed two families of constellations that are suitable for compute-and-forward. The constellations in the first family are isomorphic to the corresponding extension fields and hence can be directly used for the separation-based compute-and-forward. On

the other hand, using the Chinese Remainder Theorem, we showed that the constellations in the second family are isomorphic to the corresponding product fields. We then used the idea of multilevel coding/multistage decoding to enable the compute-and-forward for the constellations in the second family. The introduction of this multilevel coding/multistage decoding also resulted in a huge reduction in computational complexity for the constellations in the first family. For example, with the proposed multilevel coding/multistage decoding, for a constellation with q^2 elements, the channel coding has to work only over $\mathbb{F}_q$ instead of $\mathbb{F}_{q^2}$.

REFERENCES

[1] R. Ahlswede, "Multiway communication channels," in *Proc. IEEE ISIT*, 1971.

[2] H. Liao, "Multiple-access channels," *Ph.D. dissertation*. Dept. Elec. Eng., Univ. Hawaii, Honolulu, 1972.

[3] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, pp. 2–14, Jan. 1972.

[4] P. P. Bergmans, "Random coding theorems for broadcast channels with degrarded components," *IEEE Trans. Inf. Theory*, vol. 19, pp. 197–207, Mar. 1973.

[5] R. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Inf. Transm.*, pp. 3–14, July-Sept. 1974.

[6] R. Etkin and D. N. C. Tse, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534–5562, Dec. 2008.

[7] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1872–1905, Apr. 2011.

[8] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[9] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[10] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," in *Proc. IEEE ISIT*, pp. 1769–1773, June 2009.

[11] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.

[12] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5641–5654, Nov. 2010.

[13] W. Nam, S.-Y. Chung, and Y. Lee, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5488–5494, Nov. 2010.

[14] N. E. Tunali, K. Narayanan, J. Boutros, and Y.-C. Huang, "Lattices over Eisenstein integers for compute-and-forward," in *Proc. Allerton Conf.*, Oct. 2012.

[15] N. Sommer, M. Feder, and O. Shalvi, "Low density lattice codes," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1561–1585, Apr. 2008.

[16] N. Sommer, M. Feder, and O. Shalvi, "Signal codes," in *Proc. IEEE ITW*, pp. 332–336, 2003.

[17] Y.-C. Huang, N. Tunali, and K. Narayanan, "A compute-and-forward scheme for Gaussian bi-directional relaying over inter-symbol interference," *IEEE Trans. Commun.*, vol. 61, pp. 1011–1019, Mar. 2013.

[18] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *arXiv:1108.1695v2 [cs.IT]*, Oct. 2012.

[19] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2293–2314, Oct. 2004.

[20] S. L. U. Erez and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3401–3416, Oct. 2005.

[21] G. D. Forney, "On the role of mmse estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. Allerton Conf.*, pp. 430–439, Oct. 2003.

[22] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 40, pp. 409–417, Mar. 1994.

[23] J. Conway and N. Sloane, *Sphere Packings, Lattices, and Groups*. New York: Springer Verlag, 1999.

[24] W. Nam, S.-Y. Chung, and Y. Lee, "Nested lattice codes for Gaussian relay networks with interference," *IEEE Trans. Inf. Theory*, vol. 57, pp. 7733–7745, Dec. 2011.

[25] P. Larsson, N. Johansson, and K. E. Sunell, "Coded bi-directional relaying," in *Proc. IEEE VTC*.

[26] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions for bi-directional relaying," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534–5562, Dec. 2008.

[27] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *Proc. ACM MobiCom*, Sept. 2006.

[28] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proc. IEEE ICC*, pp. 707–712, Oct. 2007.

[29] M. Wilson and K. Narayanan, "Power allocation strategies and lattice based coding schemes for bi-directional relaying," in *Proc. IEEE ISIT*, pp. 344–348, June 2009.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[31] S. S. R. Zamir and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1250–1276, June 2002.

[32] M. E. J. Cioffi, G. Dudevoir and D. Forney, "Mmse decision-feedback equalizers and coding - part i: Equalization results, part ii: Coding results," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2604, Oct. 1995.

[33] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 34, pp. 380–388, May 1988.

[34] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE ISIT*, June 1997.

[35] D. Tse and S. Hanly, "Multiaccess fading channels. Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[36] V. Cadambe and S. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3983–3990, Sept. 2009.

[37] T. Guess and M. Varanasi, "A new successively decodable coding technique for intersymbol-interference channels," in *Proc. IEEE ISIT*, p. 102, June 2000.

[38] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1152–1159, July 1996.

[39] U. E. T. Philosof, R. Zamir and A. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Trans. Inf. Theory*, vol. 57, pp. 5006–5035, Aug. 2011.

[40] B. Nazer, "Successive compute-and-forward," in *International Zurich Seminar*, Mar. 2012.

[41] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454–458, 2008.

[42] B. Nazer and M. Gastpar, "Lattice coding increases multicast rates for Gaussian multiple-access networks," in *Proc. Allerton Conf.*, Sept. 2007.

[43] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behaviour," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[44] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sept. 1972.

[45] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proc. IEEE VTC*, May 2006.

[46] S.-H. Lim, Y.-H. Kim, A. E. Gamal, and S.-Y. Chung, "Layered noisy network coding," in *Proc. 3rd IEEE Int. Workshop on Wireless Network Coding*, pp. 1–6, June 2010.

[47] A. Khina, Y. Kochman, and U. Erez, "Physical-layer MIMO relaying," in *Proc. IEEE ISIT*, pp. 2437–2441, July 2011.

[48] T. Yang, X. Yuan, L. Ping, I. Collings, and J. Yuan, "A new eigen-direction alignment algorithm for physical-layer network coding in MIMO two-way relay channels," in *Proc. IEEE ISIT*, pp. 2253–2257, July 2011.

[49] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," *Europ. Trans. Telecommun.*, vol. 19, pp. 333–354, Apr. 2008.

[50] S. Avestimehr, A. Sezgin, and D. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *Proc. Allerton Conf.*, Nov. 2008.

[51] A. Sezgin, A. Avestimehr, M. Khajehnejad, and B. Hassibi, "Divide-and-conquer: Approaching the capacity of the two-pair bidirectional gaussian relay network," *IEEE Trans. Inf. Theory*, vol. 58, pp. 2434–2454, Apr. 2012.

[52] S. Mohajer, S. Diggavi, C. Fragouli, and D. Tse, "Approximate capacity of a class of Gaussian interference-relay networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 2837–2864, May 2011.

[53] S. Mohajer, S. Diggave, H. V. Poor, and S. Shamai, "On the parallel relay wire-tap network," in *Proc. Allerton Conf.*, Sept. 2011.

[54] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4566–4592, Sept. 2010.

[55] R. F. Wyrembelski, T. J. Oechtering, and H. Boche, "MIMO Gaussian bidirectional broadcast channels with common messages," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2950–2959, Sept. 2011.

[56] K. N. Y.-C. Huang and T. Liu, "Coding for parallel Gaussian bi-directional relay channels: A deterministic approach," in *Proc. Allerton Conf.*, Sept. 2011.

[57] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," in *Proc. IEEE ISIT*, pp. 1017–1021, June 2010.

[58] T. W. Hungerford, *Algebra (Graduate Texts in Mathematics)*. New York: Springer-Verlag, 1974.

[59] R. Breusch, "Zur verallgemeinerung des bertrandsehen postulates, dab zwischen x und 2x stets primzalflen liegen," *Mathematische Zeitschrift*, vol. 34, no. 1, pp. 505–526, 1932.

[60] H.-A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.

[61] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1361–1391, July 1999.

[62] B. Hern and K. Narayanan, "Multilevel coding schemes for compute-and-forward," in *Proc. IEEE ISIT*, July 2011.

[63] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, pp. 439–441, May 1983.

[64] M. P. Wilson, K. Narayanan, and G. Caire, "Joint source channel coding with side information using hybrid digital analog codes," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4922–4940, Oct. 2010.

[65] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, pp. 1–10, Jan. 1976.

[66] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1486–1495, Apr. 2005.

[67] A. Lapidoth and S. Tinguely, "Sending a bi-varite gaussian source over a gaussian mac," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2714–2752, June 2010.

[68] S. Bross, A. Lapidoth, and S. Tinguely, "Broadcasting correlated Gaussians," in *Proc. IEEE ISIT*, July 2008.

[69] H. Behroozi, F. Alajaji, and T. Linder, "Hybrid digital-analog joint source-channel coding for broadcasting correlated Gaussian sources," in *Proc. IEEE ISIT*, July 2009.

[70] R. Soundararajan and S. Vishwanath, "Hybrid coding for Gaussian broadcast channels with Gaussian sources," in *Proc. IEEE ISIT*, July 2009.

[71] C. Tian, S. Diggavi, and S. Shamai, "The achievable distortion region of sending a bivariate Gaussian source on the Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6419–6427, Oct. 2011.

[72] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1082–1102, May 2002.

[73] Z. Reznic, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3778–3788, Aug. 2006.

[74] C. Tian, S. Diggavi, and S. Shamai, "Approximate characterizations for the Gaussian source broadcasting distortion region," *IEEE Trans. Inf. Theory*, vol. 57, pp. 124–136, Jan. 2011.

[75] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. 11, pp. 558–567, Oct. 1965.

[76] V. Prabhakaran, R. Puri, and K. Ramchandran, "Colored Gaussian source-channel broadcast for heterogenous (analog/digital) receivers," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1807–1814, Apr. 2008.

[77] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, pp. 564–579, Mar. 1998.

[78] R. Puri, K. Ramchandran, and S. Pradhan, "On seamless digital upgrade of analog transmission systems using coding with side information," in *Proc. Allerton Conf.*, Sept. 2002.

[79] M. Varasteh and H. Behroozi, "Layered hybrid digital-analog codes for sending a Gaussian source over a Gaussian channel with correlated interference," 2011. Submitted to IEEE Trans. Commun.

[80] S. Bross, A. Lapidoth, and S. Tinguely, "Superimposed coded and uncoded transmissions of a Gaussian source over the Gaussian channel," in *Proc. IEEE ISIT*, pp. 2153–2155, July 2006.

[81] S. H. Lim, P. Minero, and Y.-H. Kim, "Lossy communication of correlated sources over multiple access channels," in *Proc. Allerton Conf.*, Sept. 2010.

[82] J. Mitola, "Cognitive radio," *Ph.D. dissertation*. Royal Institute of Technology (KTH), Stockholm, Sweden, 2000.

[83] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4391–4399, Nov. 2007.

[84] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3945–3958, Sept. 2009.

[85] I. Marić, R. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3536–3548, Oct. 2007.

[86] S. Rini, D. Tuninetti, and N. Devroye, "Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results," *IEEE Trans. Inf. Theory*, vol. 58, pp. 820–848, Feb. 2012.

[87] M. Gastpar, B. Rimoldi, and Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1147–1158, May 2003.

[88] N. L. Carothers, *Real Analysis*. Cambridge, U.K.: Cambridge University Press, 1999.

[89] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

APPENDIX A

APPENDIX TO CHAPTER 3

A.1 Optimal Power Allocation for $\overline{\mathcal{C}}^s$

The problem we wish to solve is given by

$$\begin{aligned} \max_{\mathcal{P}} \ & \int_{-\pi}^{\pi} \min\left\{ C\left(\frac{P_i(\theta)|h_i(\theta)|^2}{\sigma^2}\right), C\left(\frac{P_{Rj}(\theta)|g_j(\theta)|^2}{\sigma^2}\right)\right\} d\theta \\ \text{s.t.} \quad & \frac{1}{2\pi}\int_{-\pi}^{\pi} P_i(\theta)d\theta \le P_i, \ \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{Rj}(\theta)d\theta \le P_R, \\ & P_i(\theta) \ge 0, P_{Rj}(\theta) \ge 0, \quad \forall\theta, \end{aligned} \tag{A.1}$$

for $i, j \in \{A, B\}$ and $i \neq j$. Note that the objective function in (3.8) is Riemann integrable since the minimum of two Riemann integrable functions is also Riemann integrable [88]. Now, instead of solving the original optimization problem, we can partition the whole spectrum into L sub-channels and then solve the following problem instead,

$$\begin{aligned} \max_{\mathcal{P}} \ & \sum_{l=1}^{L} R_{ij}(l) \\ \text{s.t.} \quad & R_{ij}(l) \le C\left(\frac{P_i(l)|h_i(l)|^2}{\sigma^2}\right), \\ & R_{ij}(l) \le C\left(\frac{P_{Rj}(l)|g_j(l)|^2}{\sigma^2}\right), \\ & \sum_{l=1}^{L} P_i(l) \le LP_i, \ \sum_{l=1}^{L} P_{Rj}(l) \le LP_R. \\ & P_i(l) \ge 0, P_{Rj}(l) \ge 0, \quad \forall l. \end{aligned} \tag{A.2}$$

Clearly, this optimization problem is convex and *Slater's condition* holds so that it can be solved efficiently [89] and its solution is determined by the Karush-Kuhn Tucker (KKT) conditions. The corresponding Lagrangian is given by

$$\begin{aligned}\mathcal{L} = &-\sum_{l=1}^{L} R_{ij}(l) \\ &+ \sum_{l=1} \mu_1(l) \left(R_{ij}(l) - C\left(\frac{P_i(l)|h_i(l)|^2}{\sigma^2}\right)\right) \\ &+ \sum_{l=1} \mu_2(l) \left(R_{ij}(l) - C\left(\frac{P_{Rj}(l)|g_j(l)|^2}{\sigma^2}\right)\right) \\ &- \sum_{l=1}^{L} \mu_3(l) P_i(l) - \sum_{l=1}^{L} \mu_4(l) P_{Rj}(l) \\ &+ v_1 \left(\sum_{l=1}^{L} P_i(l) - LP \right) + v_2 \left(\sum_{l=1}^{L} P_{Rj}(l) - LP \right), \end{aligned} \tag{A.3}$$

where $\mu_1(l), \mu_2(l), \mu_3(l), \mu_4(l), v_1, v_2 \geq 0$ are the dual variables. The feasible conditions together with the following conditions form the KKT conditions,

$$\begin{aligned} &a)\ \mu_1(l) \left(R_{ij}(l) - C\left(\frac{P_i(l)|h_i(l)|^2}{\sigma^2}\right)\right) = 0, \\ &b)\ \mu_2(l) \left(R_{ij}(l) - C\left(\frac{P_{Rj}(l)|g_j(l)|^2}{\sigma^2}\right)\right) = 0, \\ &c)\ \mu_3(l) P_i(l) = 0, \quad d)\ \mu_4(l) P_{Rj}(l) = 0, \\ &e)\ \mu_1(l) + \mu_2(l) = 1, \\ &f)\ \mu_1(l) \frac{|h_i(l)|^2}{\sigma^2 + P_i(l)|h_i(l)|^2} + \mu_3(l) - v_1 = 0, \\ &g)\ \mu_2(l) \frac{|g_j(l)|^2}{\sigma^2 + P_{Rj}(l)|g_j(l)|^2} + \mu_4(l) - v_2 = 0. \end{aligned}$$

Solving for the P_i and P_{Rj}, one obtains

$$P_i^{**}(l) = \left(\frac{|g_j(l)|^2}{v_1|g_j(l)|^2 + v_2|h_i(l)|^2} - \frac{\sigma^2}{|h_i(l)|^2} \right)^+ \tag{A.4}$$

$$P_{Rj}^{**}(l) = \left(\frac{|h_i(l)|^2}{v_1|g_j(l)|^2 + v_2|h_i(l)|^2} - \frac{\sigma^2}{|g_j(l)|^2} \right)^+ . \tag{A.5}$$

Now, let the number of sub-channels L go to infinity, one obtains (3.9) and (3.10) and thus completes the proof.

APPENDIX B

APPENDIX TO CHAPTER 4

B.1 Proof of Theorem 17

Without loss of generality, $\sum_{l=1}^{L} n_l \geq \sum_{l=1}^{L} m_l$ is assumed; otherwise, one can switch the role of nodes A and B. Note that, for this case, one has that $C_{ex} = \sum_{l=1}^{L} m_l$. We again define

$$R_{MAC,C}^{d} \triangleq \sum_{l=1}^{L} \min(n_l, m_l), \tag{B.1}$$

$$R_{MAC,A}^{d} \triangleq \sum_{l=1}^{L} (n_l - m_l)^{+}, \tag{B.2}$$

$$R_{MAC,B}^{d} \triangleq \sum_{l=1}^{L} (m_l - n_l)^{+}. \tag{B.3}$$

The coding scheme for the reciprocal case is described as follows. In the MAC phase, both nodes use all the bit pipes belonging to the aligned part to transmit signals. Therefore, the relays will receive a common message of $R_{MAC,C}^{d}$ bits. In addition to this, some of the bit pipes belonging to the non-aligned part are also used. At the relay, this will result in a private message sent from node A of $R_A^d \leq R_{MAC,A}^{d}$ bits and a private message sent from node B of $R_B^d \leq R_{MAC,B}^{d}$ bits.

In the BC phase, the relay first re-routes the bits in R_A^d and R_B^d to other sub-channels that can support the transmission of the private messages to the intended destination. After that, each sub-channel independently performs the BBC coding [55], i.e., separate coding. Since the uplink channel and the downlink channel are reciprocal, it is guaranteed that the sub-channel l can reliably transmit exactly

$\min(n_l, m_l)$ bits. So all the functions can get through the channel and

$$R_C^d = \sum_{l=1}^{L} \min(n_l, m_l) = R_{MAC,C}^d. \tag{B.4}$$

In addition to that, the sub-channel l can also support the transmission of private messages of

$$n_l - \min(n_l, m_l) = (n_l - m_l)^+, \tag{B.5}$$

$$m_l - \min(n_l, m_l) = (m_l - n_l)^+, \tag{B.6}$$

bits to nodes A and B, respectively. Therefore, the private message sent from node A that is decodable at node B can have a rate up to

$$\begin{aligned} R_A^d &= \min\left\{\sum_{l=1}^{L}(m_l - n_l)^+, R_{MAC,A}^d\right\} \\ &= \min\left\{\sum_{l=1}^{L}(m_l - n_l)^+ + \min(n_l, m_l), \sum_{l=1}^{L}(n_l - m_l)^+ + \min(n_l, m_l)\right\} - \min(n_l, m_l) \\ &= \min\left\{\sum_{l=1}^{L} m_l, \sum_{l=1}^{L} n_l\right\} - \min(n_l, m_l) \\ &\overset{(a)}{=} \sum_{l=1}^{L} m_l - \min(n_l, m_l), \end{aligned} \tag{B.7}$$

where (a) is due to the assumption that $\sum_{l=1}^{L} n_l \geq \sum_{l=1}^{L} m_l$. Also, the private

message sent from node B that is decodable at node A can have a rate up to

$$\begin{aligned}
R_B^d &= \min\left\{\sum_{l=1}^{L}(n_l - m_l)^+, R_{MAC,B}^d\right\} \\
&= \min\left\{\sum_{l=1}^{L}(n_l - m_l)^+ + \min(n_l, m_l), \sum_{l=1}^{L}(m_l - n_l)^+ + \min(n_l, m_l)\right\} - \min(n_l, m_l) \\
&= \min\left\{\sum_{l=1}^{L} n_l, \sum_{l=1}^{L} m_l\right\} - \min(n_l, m_l) \\
&\overset{(a)}{=} \sum_{l=1}^{L} m_l - \min(n_l, m_l),
\end{aligned} \tag{B.8}$$

where again in (a) we use the assumption that $\sum_{l=1}^{L} n_l \geq \sum_{l=1}^{L} m_l$. Therefore, this scheme achieves the rate pair given by

$$\begin{aligned}
R_{AB}^d &= R_C^d + R_A^d = \sum_{l=1}^{L} m_l, \\
R_{BA}^d &= R_C^d + R_B^d = \sum_{l=1}^{L} m_l.
\end{aligned} \tag{B.9}$$

Thus, there always exists a routing strategy that can achieve the capacity region.

B.2 Proof of Theorem 18

Similar to the proof of Theorem 17, without loss of generality, we assume

$$\sum_{l=1}^{L} \log\left(1 + P|h_{Al}|^2\right) \geq \sum_{l=1}^{L} \log\left(1 + P|h_{Bl}|^2\right) = C_{ex}. \tag{B.10}$$

For the sub-channels having a equal gain, i.e., $|h_{Al}|^2 = |h_{Bl}|^2$, we set $\alpha_{Al} = 1$, $\alpha_{Bl} = 1$, $R_{Al}^{(2)} = 0$, and $R_{Bl}^{(2)} = 0$. The relay directly decodes the received signal to

the modulo-sum of two lattice codewords. This results in

$$R_{Al}^{(1)} = R_{Bl}^{(1)} \geq \log\left(1 + P|h_{Al}|^2\right) - 1. \tag{B.11}$$

For the sub-channels with $|h_{Al}|^2 > |h_{Bl}|^2$, node A can have one extra codeword in the MAC phase. We therefore set $\alpha_{Bl} = 1$ and $R_{Bl}^{(2)} = 0$ and choose the power allocation at node A such that the two lattices are perfectly aligned at the relay as

$$\alpha_{Al} = \frac{|h_{Bl}|^2}{|h_{Al}|^2}. \tag{B.12}$$

The relay first decodes the extra codeword $\mathbf{x}_{Al}^{(2)}$ by treating the lattice part as noise. It then subtracts the decoded codeword and computes the lattice function. Again, this results in

$$\begin{aligned} R_{Al}^{(1)} + R_{Al}^{(2)} &= \log\left(\frac{1}{2} + P|h_{Bl}|^2\right)^+ + \log\left(1 + \frac{P|h_{Al}|^2(1-\alpha_{Al})}{1 + 2P|h_{Bl}|^2}\right) \\ &\geq \log\left(1 + P|h_{Al}|^2\right) - 1, \end{aligned} \tag{B.13}$$

and

$$R_{Bl}^{(1)} + R_{Bl}^{(2)} \geq \log\left(1 + P|h_{Bl}|^2\right) - 1. \tag{B.14}$$

For the sub-channels with $|h_{Al}|^2 < |h_{Bl}|^2$, we switch the role of nodes A and B and hence have $\alpha_{Al} = 1$, $R_{Al}^{(2)} = 0$, and

$$\alpha_{Bl} = \frac{|h_{Al}|^2}{|h_{Bl}|^2}. \tag{B.15}$$

For this case, the relay again first decodes the extra codeword $\mathbf{x}_{Bl}^{(2)}$ by treating the lattice part as noise, subtracts it out, and then decodes the lattice function. Again,

one can bound the sum rates as

$$R_{Al}^{(1)} + R_{Al}^{(2)} \geq \log\left(1 + P|h_{Al}|^2\right) - 1, \tag{B.16}$$

and

$$\begin{aligned} R_{Bl}^{(1)} + R_{Bl}^{(2)} &= \log\left(\frac{1}{2} + P|h_{Al}|^2\right)^+ + \log\left(1 + \frac{P|h_{Bl}|^2(1-\alpha_{Bl})}{1+2P|h_{Al}|^2}\right) \\ &\geq \log\left(1 + P|h_{Bl}|^2\right) - 1. \end{aligned} \tag{B.17}$$

In the BC phase, the relay first performs bit loading for assigning the bits in $\mathbf{x}_{Al}^{(2)}$ and $\mathbf{x}_{Bl}^{(2)}$ to other sub-channels that can support the transmission of the private message to the intended destination. This mimics the behavior of routing in the deterministic model. After that, each sub-channel independently performs the BBC coding [55], i.e., separate coding. Since the uplink channel and the downlink channel are reciprocal, it is guaranteed that all the decoded lattice functions can be reliably transmitted inside the same sub-channel. Hence, the functions can have

$$R_C = \sum_{l=1}^{L} R_{Al}^{(1)} = \sum_{l=1}^{L} R_{Bl}^{(1)}, \tag{B.18}$$

bits. In addition to that, the sub-channel l can also support the transmission of a private message of a rate $\log(1 + P|h_{Al}|^2) - R_{Bl}^{(1)}$ to node A, and that of a rate $\log(1 + P|h_{Bl}|^2) - R_{Al}^{(1)}$ to node B. Therefore, the private message sent from node A

that is decodable at node B can have a rate up to

$$\begin{aligned}
R_A &= \min\left\{\sum_{l=1}^{L} R_{Al}^{(2)}, \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2) - R_{Al}^{(1)}\right\} \\
&= \min\left\{\sum_{l=1}^{L} R_{Al}^{(1)} + R_{Al}^{(2)}, \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2)\right\} - R_C \\
&\geq \min\left\{\sum_{l=1}^{L}\log(1+P|h_{Al}|^2) - 1, \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2)\right\} - R_C \\
&\geq \min\left\{\sum_{l=1}^{L}\log(1+P|h_{Al}|^2), \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2)\right\} - R_C - L \\
&\overset{(a)}{=} \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2) - R_C - L
\end{aligned} \tag{B.19}$$

where (a) follows from the assumption that $\sum_{l=1}^{L}\log(1+P|h_{Al}|^2) \geq \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2)$. Also, the private message sent from node B that is decodable at node A can have a rate up to

$$\begin{aligned}
R_B &= \min\left\{\sum_{l=1}^{L} R_{Bl}^{(2)}, \sum_{l=1}^{L}\log(1+P|h_{Al}|^2) - R_{Bl}^{(1)}\right\} \\
&= \min\left\{\sum_{l=1}^{L} R_{Bl}^{(1)} + R_{Bl}^{(2)}, \sum_{l=1}^{L}\log(1+P|h_{Al}|^2)\right\} - R_C \\
&\geq \min\left\{\sum_{l=1}^{L}\log(1+P|h_{Bl}|^2) - 1, \sum_{l=1}^{L}\log(1+P|h_{Al}|^2)\right\} - R_C \\
&\geq \min\left\{\sum_{l=1}^{L}\log(1+P|h_{Bl}|^2), \sum_{l=1}^{L}\log(1+P|h_{Al}|^2)\right\} - R_C - L \\
&\overset{(a)}{=} \sum_{l=1}^{L}\log(1+P|h_{Bl}|^2) - R_C - L,
\end{aligned} \tag{B.20}$$

where again (a) is due to the assumption that $\sum_{l=1}^{L} \log(1+P|h_{Al}|^2) \geq \sum_{l=1}^{L} \log(1+P|h_{Bl}|^2)$. Therefore, this scheme achieves the rate pair given by

$$\begin{aligned} R_{AB} &= R_C + R_A \\ &\geq \sum_{l=1}^{L} \log(1+P|h_{Bl}|^2) - L \\ &= C_{ex} - L, \end{aligned} \tag{B.21}$$

and that from node B to A is given by

$$\begin{aligned} R_{BA} &= R_C + R_B \\ &\geq \sum_{l=1}^{L} \log(1+P|h_{Bl}|^2) - L \\ &= C_{ex} - L. \end{aligned} \tag{B.22}$$

This shows that there always exists a bit loading strategy such that together with separate BBC coding at each sub-channel, it achieves the exchange capacity to within L bits.

APPENDIX C

APPENDIX TO CHAPTER 6

C.1 Equivalence of (6.17) and (6.19)

In this appendix, we verify that with the knowledge of actual channel SNR, two proposed schemes perform exactly the same. For fixed γ and $P_h = P - P_a$, the second term in (6.19) becomes

$$\Gamma^T \Lambda_{UY}^{-1} \Gamma = \frac{\mathbb{E}[VU]^2 \mathbb{E}[Y^2] - 2\mathbb{E}[VU]\mathbb{E}[VY]\mathbb{E}[UY] + \mathbb{E}[VY]^2 \mathbb{E}[U^2]}{\mathbb{E}[U^2]\mathbb{E}[Y^2] - \mathbb{E}[UY]^2}, \tag{C.1}$$

where

$$\mathbb{E}[VU] = \alpha \mathbb{E}[S'V] + \kappa \sigma_V^2, \tag{C.2}$$

$$\mathbb{E}[VY] = \mathbb{E}[S'V], \tag{C.3}$$

$$\mathbb{E}[U^2] = P_h + \alpha^2 \mathbb{E}[S'^2] + \kappa^2 \sigma_V^2 + 2\alpha\kappa \mathbb{E}[S'V], \tag{C.4}$$

$$\mathbb{E}[Y^2] = P_h + \mathbb{E}[S'^2] + N, \tag{C.5}$$

$$\mathbb{E}[UY] = P_h + \alpha \mathbb{E}[S'^2] + \kappa \mathbb{E}[S'V], \tag{C.6}$$

with α and κ^2 determined by (6.22) and

$$\begin{aligned} \mathbb{E}[S'^2] = a\gamma^2 \sigma_V^2 + [1 + \sqrt{a}(1-\gamma)]^2 \sigma_S^2 \\ + 2\sqrt{a}\gamma[1 + \sqrt{a}(1-\gamma)]\rho\sigma_V\sigma_S, \end{aligned} \tag{C.7}$$

$$\mathbb{E}[S'V] = \sqrt{a}\gamma\sigma_V^2 + [1 + \sqrt{a}(1-\gamma)]\rho\sigma_V\sigma_S. \tag{C.8}$$

After some algebra, we can rewrite the numerator and denominator in (C.1) as, respectively,

$$\frac{P_h\left\{\mathbb{E}[S'V]^2ND^* + \sigma_V^2 P_h(\sigma_V^2\mathbb{E}[Y^2] - \mathbb{E}[S'V]^2)\right\}}{(P_h+N)D^*}, \tag{C.9}$$

and

$$\frac{P_h\left\{(ND^* + P_h\sigma_V^2)\mathbb{E}[Y^2] - P_h\mathbb{E}[S'V]^2\right\}}{(P_h+N)D^*}. \tag{C.10}$$

Thus, we can rewrite (6.19) as

$$\begin{aligned} D_{hda} &= \sigma_V^2 - \Gamma^T\Lambda_{UY}^{-1}\Gamma = \sigma_V^2 - \frac{\text{(C.9)}}{\text{(C.10)}} \\ &= D^*\frac{N(\sigma_V^2\mathbb{E}[Y^2] - \mathbb{E}[S'V]^2)}{(ND^* + P_h\sigma_V^2)\mathbb{E}[Y^2] - P_h\mathbb{E}[S'V]^2} \\ &\overset{(a)}{=} \frac{ND^*(\sigma_V^2\mathbb{E}[Y^2] - \mathbb{E}[S'V]^2)}{(P_h+N)(\sigma_V^2\mathbb{E}[Y^2] - \mathbb{E}[S'V]^2)} \\ &= \frac{D^*}{1+\frac{P_h}{N}}, \end{aligned} \tag{C.11}$$

where (a) follows from that $D^* = \sigma_V^2 - \mathbb{E}[VY]^2/\mathbb{E}[Y^2]$ and $\mathbb{E}[VY] = \mathbb{E}[S'V]$. This completes the proof.

C.2 Digital Wyner-Ziv Scheme

In this appendix, we summarize the digital Wyner-Ziv scheme for lossy source coding with side-information V' ($V = V' + W$ with $W \sim \mathcal{N}(0, D^*)$) at receiver. Similar to the previous sections, we omit all the ε and/or δ intentionally for the sake of convenience and to maintain clarity.

Suppose the side-information is available at both sides, the least required rate R_{WZ} for achieving a desired distortion D is [64]

$$R_{WZ} = \frac{1}{2}\log\frac{D^*}{D}. \tag{C.12}$$

Let us set this rate to be arbitrarily close to the rate given in (6.16), the rate that the channel can support with arbitrarily small error probability. The best possible distortion one can achieve for this setup is then given as

$$D = \frac{D^*}{1 + \frac{P - P_a}{N}}. \tag{C.13}$$

This distortion can be achieved as follows [64],

1. Let T be the auxiliary random variable given by

$$T = \alpha_{sep} V + B, \tag{C.14}$$

where

$$\alpha_{sep} = \sqrt{\frac{D^* - D}{D^*}} \tag{C.15}$$

and $B \sim \mathcal{N}(0, D)$. Generate a length n i.i.d. Gaussian codebook $\mathcal{T}$ of size $2^{nI(T;V)}$ and randomly assign the codewords into 2^{nR} bins with R chosen from (6.16). For each source realization $\mathbf{v}$, find a codeword $\mathbf{t} \in \mathcal{T}$ such that $(\mathbf{v}, \mathbf{t})$ is jointly typical. If none or more than one are found, an encoding failure is declared.

2. For each chosen codeword, the encoder transmit the bin index of this codeword by the DPC with rate given in (6.16).

3. The decoder first decodes the bin index (the decodability is guaranteed by the rate we chose) and then looks for a codeword $\hat{\mathbf{t}}$ in this bin such that $(\hat{\mathbf{t}}, \mathbf{v}')$ is jointly typical. If this is not found, a dummy codeword is selected. Note that as $n \to \infty$, the probability that $\hat{\mathbf{t}} \neq \mathbf{t}$ vanishes. Therefore, we can assume that $\hat{\mathbf{t}} = \mathbf{t}$ from now on.

4. Finally, the decoder forms the MMSE from $\mathbf{t}$ and $\mathbf{v}'$ as $\hat{\mathbf{v}} = \mathbf{v}' + \hat{\mathbf{w}}$ with

$$\hat{\mathbf{w}} = \frac{\alpha_{sep} D^*}{\alpha_{sep}^2 D^* + D}(\mathbf{t} - \alpha_{sep}\mathbf{v}'). \tag{C.16}$$

It can be verified that for the choice of α the required rate is equal to (C.12) and the corresponding distortion is

$$\begin{aligned} &\mathbb{E}[(V - \widehat{V})^2] = \mathbb{E}[(W - \widehat{W})^2] \\ &= D^* \left(1 - \frac{\alpha_{sep}^2 D^*}{\alpha_{sep}^2 D^* + D}\right) = D. \end{aligned} \tag{C.17}$$

C.3 Wyner-Ziv with Mismatched Side-Information

In this appendix, we calculate the expected distortion of the digital Wyner-Ziv scheme in the presence of side-information mismatch. Specifically, we consider the Wyner-Ziv problem with an i.i.d. Gaussian source and the MSE distortion measure. Let us assume that the best achievable distortion in the absence of side-information mismatch to be D. The encoder believes that the side-information is V', and $V = V' + W$ with $W \sim N(0, D^*)$. However, the side-information turns out to be V_a' and has the relation $V = V_a' + W_a$ with $W_a \sim N(0, D_a^*)$. Under the same rate, we want to calculate the actual distortion D_a suffered by the decoder.

Since the encoder has been fixed to deal with the side-information, V', at decoder, the auxiliary random variable is as in (C.14) with the coefficient given in (C.15). Since the decoder knows the actual side-information, V_a', perfectly, it only has to estimate W_a. By the orthogonality principle, the MMSE estimate $\widehat{W_a}$ can be obtained as

$$\widehat{W_a} = \frac{\alpha_{sep} D_a^*}{\alpha_{sep}^2 D_a^* + D}(T - \alpha_{sep} V_a') \tag{C.18}$$

Therefore, the estimate of the source is $\widehat{V} = V'_a + \widehat{W}_a$. The corresponding distortion is given as

$$\begin{aligned} D_a &= \mathbb{E}[(V - \widehat{V})^2] = \mathbb{E}[(W_a - \widehat{W}_a)^2] \\ &= \frac{D^* D_a^*}{D^* D_a^* + (D^* - D_a^*)D} D \end{aligned} \tag{C.19}$$

Here, we give an example in Fig. C.1 to see the performance improvement through having the access of a better side-information. In this figure, we plot the $-10\log_{10} D_a$ as $-10\log_{10} D_a^*$ increases, i.e., as the actual side-information improves. The outer bound is obtained by assuming the transmitter always knows the distribution of actual side-information at decoder and the distortion of the HDA scheme is computed through derivations in section 6.4.2. The parameters are set to be $P = N = 1$ and $D^* = 0.1$. One can observe in the figure that both the schemes benefit from a better side-information at decoder. Moreover, it can be seen that these two schemes provide the same performance under side-information mismatch.

C.4 Discussions for SNR Mismatch Cases

As discussed previously, both the digital DPC scheme and the HDA scheme benefit from a better SNR. Here, we wish to analyze and compare the performance for these two schemes under SNR mismatch. Since the digital DPC scheme makes estimate from T (see Appendix C.2) and V' (which is a function of Y) and the HDA scheme makes estimate from U and Y, it suffices to compare $I(V;T,Y)$ with $I(V;U,Y)$. By the chain rule of mutual information, we have

$$I(V;T,Y) = I(V;Y) + I(V;T|Y), \tag{C.20}$$

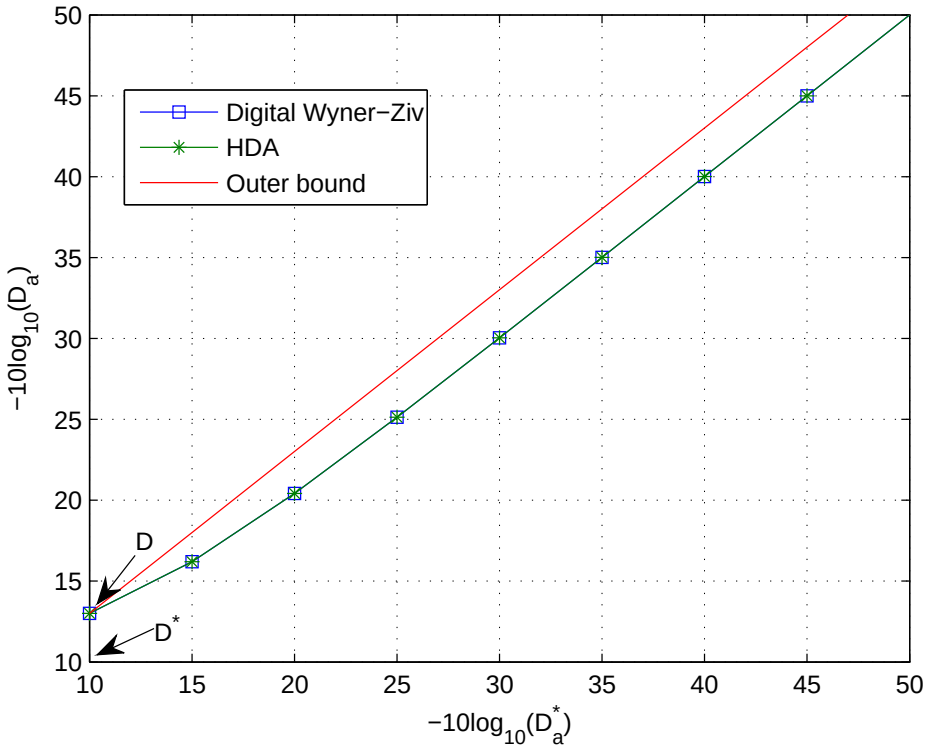

Figure C.1: Wyner-Ziv problem with side-information mismatch.

and

$$I(V;U,Y) = I(V;Y) + I(V;U|Y). \tag{C.21}$$

Thus, we only have to compare $I(V;T|Y)$ to $I(V;U|Y)$. Let us consider $\rho = 0$ case for example,

$$\begin{aligned} I(V;T|Y) &= h(T|Y) - h(T|V,Y) \\ &= h(\alpha_{sep}V + B|Y) - h(\alpha_{sep}V + B|V,Y) \\ &= h(\alpha_{sep}V - \alpha_{sep}\beta_a Y + B|Y) - h(\alpha_{sep}B|V,Y) \\ &= h(\alpha_{sep}W_a + B|Y) - h(B) \\ &\overset{(a)}{=} h(\alpha_{sep}W_a + B) - h(B) \\ &= \frac{1}{2}\log\frac{\alpha_{sep}^2 D_a^* + D}{D}, \end{aligned} \tag{C.22}$$

where α_{sep} and W_a are defined in Appendix C.3 and (a) follows from the orthogonality principle.

$$\begin{aligned} I(V;U|Y) &= h(U|Y) - h(U|V,Y) \\ &= h(U|Y) - h(X_h + \alpha S' + \kappa V|V,Y) \\ &= h(U|Y) - h\left((1-\alpha)X_h - \alpha Z_a|V,Y\right) \\ &\overset{(a)}{\geq} h(U|Y) - h\left((1-\alpha)X_h - \alpha Z_a\right) \\ &= \frac{1}{2}\log\frac{\mathbb{E}[U^2] - \mathbb{E}[UY]^2/\mathbb{E}[Y^2]}{(1-\alpha)^2 P_h + \alpha^2 N_a}. \end{aligned} \tag{C.23}$$

where (a) follows from that conditioning reduces entropy and the equality occurs if there is no SNR mismatch.

Two examples are given in Fig. C.2 to compare these two quantities with and

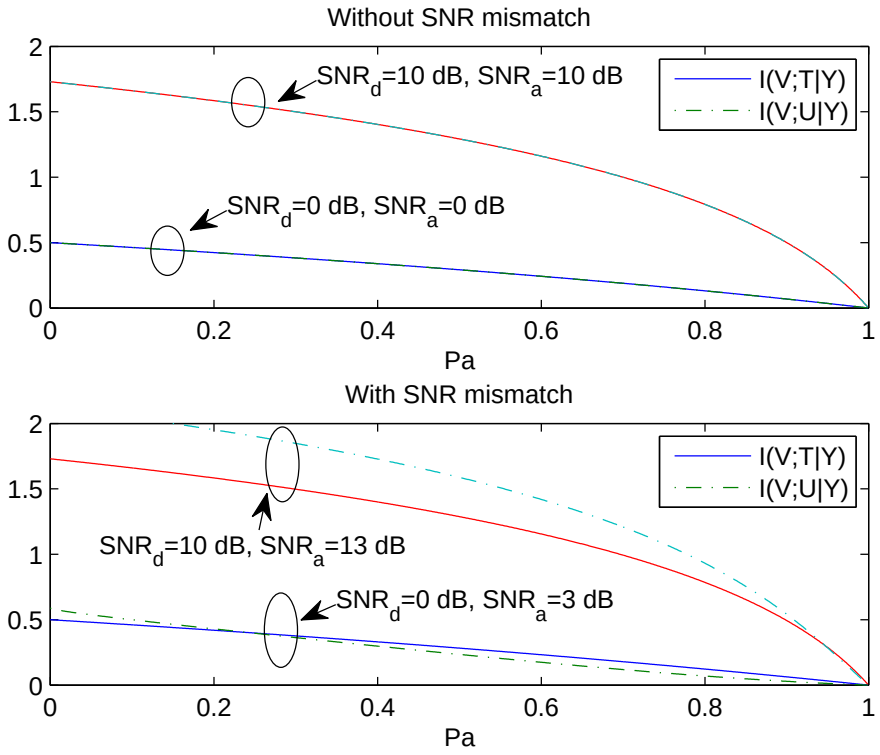

Figure C.2: Comparisons of (C.22) and (C.23).

without SNR mismatch for a small and a large designed SNR, respectively. One can observe that without SNR mismatch, these two quantities coincide with each other for all choices of P_a. This implies the result in section 6.4 that without mismatch the digital DPC scheme and the HDA scheme provide exactly the same distortion. However, with SNR mismatch, we can observe that which quantity is larger really depends on P_a for the small designed SNR case. On the other hand for designed SNR = 10 dB case, we have $I(V;U|Y) > I(V;T|Y)$ for a wide range of P_a (except for some P_a close to 1). This explains the results in section 6.5 that, for large designed SNRs, the HDA scheme has better results than the digital DPC scheme does while for small designed SNRs we cannot make this conclusion easily.