Analysis and Design of Physical-Layer Network Coding for Relay Networks

UNSW
THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY • AUSTRALIA

Thesis by

Tao Huang

School of Electrical Engineering and Telecommunications
The University of New South Wales

August 2014

A dissertation submitted to Graduate Research School
The University of New South Wales
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Dedicated to my beloved Son, Wife and Parents

ABSTRACT

Physical-layer network coding (PNC) is a technique to make use of interference in wireless transmissions to boost the system throughput. In a PNC employed relay network, the relay node directly recovers and transmits a linear combination of its received messages in the physical layer. It has been shown that PNC can achieve near information-capacity rates. PNC is a new information exchange scheme introduced in wireless transmission. In practice, transmitters and receivers need to be designed and optimized, to achieve fast and reliable information exchange. Thus, we would like to ask: How to design the PNC schemes to achieve fast and reliable information exchange? In this thesis, we address this question from the following works:

Firstly, we studied channel-uncoded PNC in two-way relay fading channels with QPSK modulation. The computation error probability for computing network coded messages at the relay is derived. We then optimized the network coding functions at the relay to improve the error rate performance.

We then worked on channel coded PNC. The codes we studied include classical binary code, modern codes, and lattice codes. We analyzed the distance spectra of channel-coded PNC schemes with classical binary codes, to derive upper bounds for error rates of computing network coded messages at the relay. We designed and optimized irregular repeat-accumulate coded PNC. We modified the conventional extrinsic information transfer chart in the optimization process to suit the superimposed signal received at the relay. We analyzed and

designed Eisenstein integer based lattice coded PNC in multi-way relay fading channels, to derive error rate performance bounds of computing network coded messages.

Finally we extended our work to multi-way relay channels. We proposed a opportunistic transmission scheme for a pair-wise transmission PNC in a single-input single-output multi-way relay channel, to improve the sum-rate at the relay. The error performance of computing network coded messages at the relay is also improved. We optimized the uplink/downlink channel usage for multi-input multi-output multi-way relay channels with PNC to maximize the degrees of freedom capacity. We also showed that the system sum-rate can be further improved by a proposed iterative optimization algorithm.

ACKNOWLEDGMENTS

Foremost, I would like to thank my supervisor Prof. Jinhong Yuan for providing me with the opportunity to conduct my PhD research study in the Wireless Communications Laboratory (WCL) with him at The University of New South Wales, Sydney. His immense knowledge and enthusiastic in research inspired me in the past four years. He is always there to advise me and support me. All my achievements have been benefited from his knowledge, insight and inspiration.

I would like to thank my research collaborators: Dr. Tao Yang, Dr. Ingmar Land, Dr. Qifu Sun, Dr. Xiaojun Yuan, and Dr. Tiffany Jing Li, for their fruitful discussions and valuable suggestions in my research work. In addition, I would like to thank Dr. Ingmar Land from The Institute for Telecommunications Research, University of South Australia (ITR, UniSA), and Dr. Xiaojun Yuan from The Institute of Network Coding, The Chinese University of Hong Kong (INC, CUHK), for hosting my visiting research, which are joyful and great experiences in my PhD study.

I want to thank present and past members of the Wireless Communications Lab: Dr. Nan Yang, Dr. Giovanni Geraci, Yixuan (Shane) Xie, Mengyu Huang, Dr. Ido Nevat, Dr. Jun Li, Huijun Li, Dr. James Chaofeng Li, Dr. Adeel Razi, Dr. Marwan Azmi, Dr. Aiping Huang, Dr. Md. Anisul Karim, and Miaofei Xiang, for all the happiness around the WCL. Especially I would like to thank Yixuan (Shane) Xie and Dr. Jun Li for interesting discussions

on channel coding theory. I am grateful to my officemates: Long Shi, Shihao Yan, Chenxi Liu, Zhe Wang, Yi Lu, Dr. Chao Zhai, and Lu Yang, for all the happiness and friendship. A special thank should go to Joseph Yiu, who is the technical support of my lab. He is always there to help me to sort out my computer and software problems. I also need to thanks May Park, who helped me to sort out all my travel issues. I thank Dr. Rui Wang at the INC CUHK for discussions on MIMO systems, and thank all the friends I made during my visiting in the INC CUHK. I also would like to thank Ms. Lillian Lun for her administration work of my visiting at the INC CUHK.

And last, but by no means least, I would like to thank my parents for their support. I particularly thank my wife, for her heartful and unlimited supports, especially on my hardest time. It would be impossible for me to finish this work without the support from my family.

My doctoral study was funded by Australia Postgraduate Award and Engineer Research Award from The University of New South Wales. My visiting research at INC, CUHK was funded by the Endeavour Australia Cheung Kong Research Fellowship from the Australian Government. I would like to thank them for all the financial support on my research study.

Contents

List of Figures

List of Tables

List of Publications

Journal Papers

- **T. Huang**, X. Yuan, and J. Yuan, "Half-duplex MIMO Multi-way Relay Channel with Full Data Exchange: Degrees of Freedom and Sum-Rate Optimization," submitted to *IEEE Transactions on Wireless Communications.*

- Q. Sun, **T. Huang**, and J. Yuan, "On Lattice-Partition-Based Physical-Layer Network Coding over GF(4)," *IEEE Communications Letters*, vol. 10, no. 10, pp. 1988-1991, Oct. 2013.

- **T. Huang**, T. Yang, J. Yuan, and I. Land, "Design of Irregular Repeat-Accumulate Coded Physical-Layer Network Coding for Gaussian Two-Way Relay Channels," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 897-909, Mar. 2013.

- Q. Sun, J. Yuan, **T. Huang**, and W. K. Shum, "Lattice Network Codes Based on Eisenstein Integers," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2713-2725, July 2013.

- T. Yang, I. Land, **T. Huang**, J. Yuan, and Z. Chen, "Distance Spectrum and Performance of Channel-Coded Physical-Layer Network Coding for Binary-Input Gaussian Two-Way Relay Channels," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1499-1510, June 2012.

- G. Wang, W. Xiang, J. Yuan, and **T. Huang**, "Outage Analysis of Non-Regenerative Analog Network Coding for Two-Way Multi-Hop Networks", *IEEE Communications Letters*, vol. 15, no. 6, pp. 662-664, June 2011.

Conference Papers

- **T. Huang**, X. Yuan, and J. Yuan, "Degrees of Freedom of Half-Duplex MIMO Multi-Way Relay Channel with Full Data Exchange," *IEEE Global Communications Conference (GLOBECOM)*, 2014, accepted.

- **T. Huang**, J. Yuan, and Q. Sun, "Opportunistic Pair-Wise Compute-and-Forward in Multi-Way Relay Channels," in *Proceeding of the IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013.

- **T. Huang**, J. Yuan, and J. Li, "Analysis of Compute-and-Forward with QPSK in Two-Way Relay Fading Channels," in *Proceeding of the 14th Australian Communications Theory Workshop (AusCTW)*, pp. 75-80, Adelaide, Australia, Jan. 2013.

- Y. Ma, **T. Huang**, J. Li, J. Yuan, Z. Lin, and B. Vucetic, "Novel Nested Convolutional Lattice Codes for Multi-Way Relaying Systems over Fading Channels," in *Proceeding of the IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2671-2676, Shanghai, China, Apr. 2013.

- **T. Huang**, T. Yang, J. Yuan, and I. Land, "Convergence Analysis for Channel-coded Physical-Layer Network Coding in Gaussian Two-way Relay Channels," in *Proceeding of the 8th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 849-853, Aachen, Germany, Nov. 2011.

- T. Yang, I. Land, **T. Huang**, J. Yuan, and Z. Chen, "Distance Properties and Performance of Physical-Layer Network Coding with Binary Linear Codes for Gaussian Two-Way Relay Channels", in *Proceeding of IEEE International Symposium on Information Theory (ISIT)*, pp. 2070-2074, Saint Petersburg, Russia, Aug. 2011.

- G. Wang, W. Xiang, J. Yuan, and **T. Huang**, "Outage Performance of Analog Network Coding in Generalized Two-Way Multi-Hop Networks," in *Proceeding of the IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1988-1993, Quintana-Roo, Mexico, Mar. 2011. (Best Academic Paper Award)

Abbreviations

3GPP	3rd Generation Partnership Project
AF	Amplify-and-Forward
ANC	Analog Network Coding
AWGN	Additive White Gaussian Noise
BEP	Bit Error Probability
BICM	Bit-Interleaved Coded Modulation
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CPNC	Channel-Coded Physical-Layer Network Coding
CSI	Channel State Information
DF	Decode-and-Forward
EXIT	Extrinsic Information Transfer
FER	Frame Error Rate
FDMA	Frequency Division Multiple Access
GF	Galois Field
HSPA+	Evolved High-Speed Packet Access
IDMA	Interleave-Division Multiple-Access
IRA	Irregular Repeat Accumulate
LDPC	Low-Density Parity-Check
LNC	Lattice Network Coding
MD	Minimum Distance

MIMO	Multiple-Input and Multiple-Output
ML	Maximum Likelihood
MSE	Mean Square Error
MWRC	Multi-Way Relay Channel
NC	Network-Coded
OFDM	Orthogonal Frequency Division Multiplexing
OWRC	One Way Relay Channel
PAM	Pulse Amplitude Modulation
PEP	Pair-Wise Error Probability
PID	Principal Ideal Domain
PNC	Physical-Layer Network Coding
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RA	Repeat Accumulate
SI	Superimposed
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SU	Single User
TDMA	Time Division Multiple Access
TWRC	Two-Way Relay Channel
UBE	Union Bound Estimation
WEF	Weight Enumerating Function
WEP	Word Error Probability
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

Chapter 1

Introduction

1.1 Why Physical-Layer Network Coding

1.1.1 Challenge of Modern Wireless Communications

Physical-layer network coding (PNC) is a technique to make use of interference in wireless transmissions to boost the throughput of wireless networks [1]. In order to understand why PNC is an important research area in the field of wireless communications, we start with a brief introduction of wireless communications systems.

Wireless communication is the technology of transmitting and receiving information without any connected electrical conductor. In wireless communications, a typical characteristic is signal interference, which is introduced by the broadcast nature of electromagnetic waves. Interference is usually considered as a destructive phenomenon in many wireless communications standards, such as 3rd Generation Partnership Project (3GPP) standards for mobile communications, 802.11 standards for wireless local area network (WLAN). The collision of multiple signals at a receiver can result in that none of the original signals can be correctly recovered. Therefore, in conventional wireless communications systems, it is important to avoid or minimize the signal interferences. Approaches to achieve this include: time division multiple access (TDMA), frequency division multiple access (FDMA), and code division multiple access (CDMA) [2]. In these approaches, communication resources assigned to users in one system are orthogonal to each other, so that the signal from one user will not collide with the signal from another user at the receiver.

In recent years, due to the fast-paced research innovation and development, computation capability on mobile phones and tablets has dramatically increased. For example, people are able to play high-definition video, and online gaming on their smart phone these days. These applications on smart phones require large amount of data exchange in the network. Hence, how to improve the data throughput of wireless networks becomes an important issue need to be addressed.

Network coding, firstly proposed and studied by Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond W. Yeung in year 2000 [3], can improve the throughput of multicast networks. In the next two sections, we will introduce two types of network coding: straightforward network coding and physical-layer network coding. We will show how they can improve the throughput of wireless networks.

1.1.2 Straightforward Network Coding

The initial study of network coding in [3] was inspired by the multicasting problem in computer networks. In network coding, instead of simply relaying the received packets, each node take several packets and combine them together for transmission in a network. By doing so, bandwidth generally can be saved for transmission the same amount of information from source to sink. Thus, network throughput can be improved with network coding.

It has been proven that linear network coding is enough to achieve the capacity upper bound in multicast networks with one or more sources [4]. A butterfly network, as shown in Fig. 1.1, is often used to illustrate how linear network coding can outperform routing. Two source nodes, node 1 and node 2, have information w_A and w_B that must be transmitted to the two destination nodes, node 5 and node 6, and both nodes want to know w_A and w_B. Each edge can carry only single information.

We can see from Fig. 1.1 that if only routing was allowed, then the node 3 would only be able to transmit w_A or w_B to node 4, but not both. Suppose node 3 sends w_A to node 4, then node 5 would receive w_A twice and not receive w_B. Similar problem occurs at node 6 if node 3 sends w_B to node 4. However, if node 3 transmit a simple network coded information $w_A + w_B$ to node 4, then both node 5 and node 6 can receive w_A and w_B.

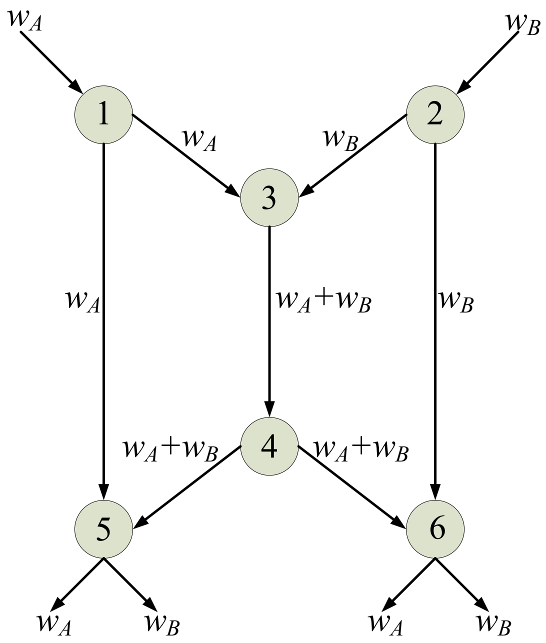


Figure 1.1A Butterfly model for wired multicast network.

Network coding, as studied in multicast computer network, can be straightforwardly applied to wireless communications [5, 6], and this application is termed as *straightforward network coding*. In order to illustrate the straightforward network coding, we consider a basic wireless communications system model: two-way relay channel (TWRC). In this model, two users A and B exchange their information w_A and w_B via an intermediate relay R, and there is no direct link between the two users. Here we consider a TDMA system, where different time slots are allocated to users and relay for transmission. All nodes are operating in half-duplex mode, i.e., each node cannot transmit and receive at the same time. For simplicity, we consider channel un-coded system.

Firstly, we show the conventional transmission scheme in a TDMA system as a comparison baseline. In a conventional TDMA system, the relay simply transmits its received information. In this case, a total of 4 time slots are required for user A and user B to exchange their message symbols w_A and w_B via the relay R, as shown in Fig. 1.2:

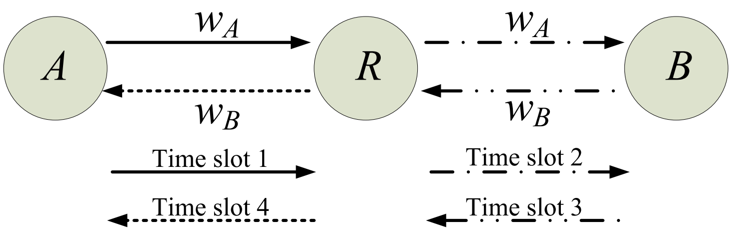


Figure 1.2A TWRC with conventional TDMA.

Time slot 1: User A transmits its message w_A to the relay R.

Time slot 2: Relay R forwards the message w_A to user B.

Time slot 3: User B transmits its message w_B to the relay R.

Time slot 4: Relay R forwards w_B to user A.

This concludes the message exchange between two users.

Now we illustrate the straightforward network coding. It has been shown in [5,6] that straightforward network coding in TWRC can reduce the above 4 time slots to 3 time slots. The corresponding time scheduling is illustrated in Fig. 1.3 and described here:

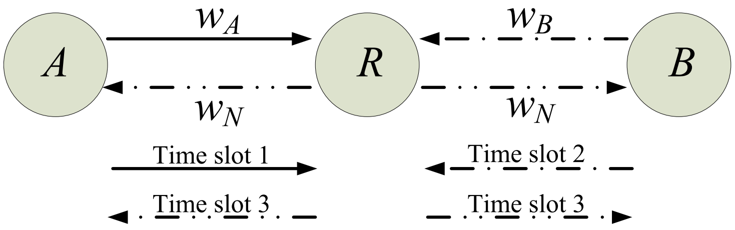


Figure 1.3A TWRC with straightforward network coding.

Time slot 1: User A transmits its message w_A to the relay R.

Time slot 2: User B transmits its message w_B to the relay R. Upon receiving both w_A and w_B, relay R computes a network coded message $w_N = f(w_A, w_B)$, where $f(\cdot)$ is a network coding function. Note that $f(\cdot)$ is known at the users.

Time slot 3: Relay R broadcasts w_N to two users. User A recovers w_B with the knowledge of self-message w_A and network coding function $f(\cdot)$. This can be denoted as $w_B = g(w_N, w_A)$, where $g(\cdot)$ is the network decoding function. Similar procedure applies at user B.

This concludes the message exchange between user A and user B.

It is worth pointing out that the network coding function $f(\cdot)$ should be selected such that $g(\cdot)$ exists, and there is no ambiguity for each user to recover the message from the other user. In other words, there is only one correct solution to $g(w_N, w_A)$ and there is only one correct solution to $g(w_N, w_B)$. For example, if the message is in binary form, then the network coding function $f(\cdot)$ can be simply as XOR operation. In this case, we have $w_N = w_A \oplus w_B$. The corresponding decoding function $g(\cdot)$ at the users is also XOR operation. After user A receives w_N, user A can deduce w_B by performing $w_A \oplus w_N$. Similar operation applies at user B.

An example study of the network coding problem for straightforward network coding employed two-way relay network can be found in [7], where Chen *et al.* investigated the network coded modulation for an asymmetric decode-and-forward two-way relay channel, and studied the methods for maximizing the corresponding Euclidean distance by jointly considering the network coding and modulation.

1.1.3 Physical-Layer Network Coding

Physical-layer network coding (PNC), firstly proposed in 2006 [1, 8], further reduces the number of time slots of one round information exchange to 2 in TWRCs. Hence, PNC improved the data throughput by 100% compared to conventional transmission scheme in TWRCs. In PNC, each relay node advocates directly recovering and transmitting the linear combinations of relay received messages in the physical layer, as shown in Fig. 1.4:

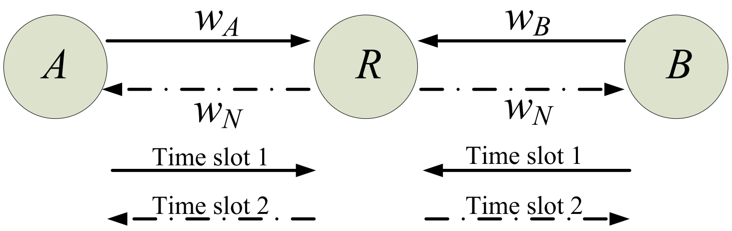


Figure 1.4TWRC with physical-layer network coding.

Time slot 1: User A and user B simultaneously transmit their messages w_A and w_B to the relay R. Let x_A be the modulated signal of w_A, and let x_B be the modulated signal of w_B, the relay received signal is

$$y_R = h_A x_A + h_B x_B + n_R, \tag{1.1}$$

where h_A and h_B are the channel coefficients and n_R is the noise at the relay. Upon receiving y_R, the relay deduces $w_N = f(w_A, w_B)$, where $f(\cdot)$ is a network coding function. Note that $f(\cdot)$ is known at the users.

Time slot 2: Relay R broadcasts w_N to two users. User A recovers w_B with the knowledge of self-message w_A and network coding function $f(\cdot)$. This can be denoted as $w_B = g(w_N, w_A)$, where $g(\cdot)$ is the network decoding function. Similar procedure applies at user B.

This concludes the message exchange between user A and user B. Note that the network coding function $f(\cdot)$ should be selected such that $g(\cdot)$ exists, and there is no ambiguity for each user to recover the message from the other user.

There are two differences between PNC and straightforward network coding:

1. Interference at the relay. In the straightforward network coding, signal collision at the relay is avoided by scheduling two users transmissions to the relay in

different time slot. However, in PNC, two users are scheduled to transmit to the relay simultaneously. Hence, relay receives superimposed signal.

2. Computation of the network coded message at the relay. In straightforward network coding, the relay is able to straightforwardly compute the network coded message. This is because the relay receives w_A and w_B in two separate time slots. However, this is not the case for PNC. In PNC, the relay receives the superimposed signal of the two users, as shown in (1.1). A key issue in PNC is how the relay can effectively compute the network coded message w_N from y_R. This issue will be addressed later in this thesis.

Example 1.1. We now illustrate a simple example of PNC. Consider an uncoded Gaussian TWRC with perfect synchronization. In this case, we have $h_A = h_B = 1$ in (1.1). In this system, binary symbol and binary phase shift keying (BPSK) modulation $(0 \mapsto -1, 1 \mapsto +1)$ are employed at all nodes. Table 1.1 shows that, with perfect synchronization, the relay sees a ternary signal constellation: $\{-2, 0, 2\}$, which is formed by superimposing BPSK constellations of the two users. A possible network coding function is also shown in Table 1.1, and it is expressed as $x_R = -x_A \times x_B$ or $w_N = w_A \oplus w_B$, where $\oplus$ is binary XOR operation. To implement this network coding function at the relay, a possible approach is that the relay uses minimum distance detection rule to find the closest superimposed constellation point to its received signal, and then uses Table 1.1 to map the estimated superimposed constellation point to the corresponding network coded symbol.

1.2 Motivations and Contributions

Since the invention of wireless communications technology, achieving reliable and robust transmission is always a challenge. When a signal is transmitted from a source to a receiver, the signal is broadcasted through a physical channel and the signal will be scattered, reflected, diffracted. The received signal are always multi-path faded, interfered signal with noise.

Table 1.1 Binary PNC with BPSK modulation

x_A (w_A)	x_B (w_B)	$x_A + x_B$	x_R (w_N)
1 (1)	1 (1)	2	-1 (0)
1 (1)	-1 (0)	0	1 (1)
-1 (0)	1 (1)	0	1 (1)
-1 (0)	-1 (0)	-2	-1 (0)

However, as aforementioned that, when PNC is employed in a network, signals from multiple transmitters are intentionally “interfered” at the receiver. The receiver is required to compute network coded messages from the superimposed signals it received. Reliable computing network coded messages at the receiver from superimposed signal introduce additional challenge than the conventional point-to-point channel. A question is raised: **How do we design the PNC schemes for relay networks to achieve fast and reliable information exchange?**

In point-to-point channel transmission, to overcome the channel impairments and improve the network performance, the following areas are usually considered and studied in the literature:

- Signal detection;
- Forward error correction;
- Network resource allocation;
- Precoder design for multi-input multi-output networks.

In this thesis, we aim to design the PNC schemes for relay networks from the above aspects. We start with the design the channel-uncoded two-way relay networks. We then move to the channel-coded two-way relay networks. After that, we study the channel-coded multi-way relay networks. Then we focus on the design of the transmission scheme for multi-way relay network. The works so far are limited to single-input single-output network. In the final work of this thesis, we study the optimization of multi-input multi-output multi-way relay network. The main contributions in this thesis are summarized below.

- **Designed an optimal network coding function at the relay for channel-uncoded PNC in two-way relay fading channels with QPSK modulation, to improve the error rate performance of detecting the corresponding network messages at the relay;**

- **Analyzed the distance spectrum of channel-coded PNC in Gaussian two-way relay channels with binary classic codes, to provides error rate upper bounds of computing the corresponding network coded messages at the relay;**

- **Designed irregular repeat-accumulate coded PNC in Gaussian two-way relay channels, to improve the error rate performance of computing the corresponding network messages at the relay;**

- **Analyzed and designed Eisenstein integer based lattice network codes for multi-way relay fading channels with PNC scheme, to provide error rate performance bounds of computing the corresponding network messages at the relay, and to provide designing methods for constructing Eisenstein integer based lattice network codes;**

- **Designed a transmission scheme for pair-wise transmission PNC in single-input single-output multi-way relay channels, to improve the network sum-rate, and to improve the error rate performance of computing the corresponding network coded messages at the relay;**

- **Optimized the uplink/downlink channel usage of multi-input multi-output multi-way relay channels with PNC to optimize the degrees of freedom capacity; Designed an iterative algorithm to optimize the precoders at the users and at the relay, to optimize the sum-rate.**

We now introduce these contributions in more details. The first main contribution of this thesis, is that **we designed an optimal network coding function at the relay for channel-uncoded PNC in two-way relay fading channels with QPSK modulation, to improve the error rate performance of detecting the corresponding network messages at the relay**. The network coding function is an important research area in PNC, and it will be introduced in more detail later in this chapter. A good network coding function at the relay can improve the error rate performance of detecting the correct network coded messages at the relay. When only considering binary input channel with binary phase shift keying (BPSK) modulation, there is no freedom to choose the network coding functions to suit the channel condition, as described in Example 1.1. When non-binary symbol with higher modulation is employed in the system, we have more flexibility of selecting good linear network coding functions depends on the channel states. Motivated by this, we investigate the network coding function at the relay for channel-uncoded PNC in two-way relay fading channels with QPSK modulation. This work falls in a research direction of finding good linear network coding functions at the relay for two-way relay channels. The detailed contributions of this work are:

- We investigated the error performance for the decoding of the network coded messages at the relay over Rayleigh fading TWRCs when QPSK modulation is employed at the users;

- We characterized the distance profile between any two immediate neighboring constellation points at the relay;

- We designed and selected the optimal computation coefficients in terms of minimizing the computation error probability at the relay;

- We derived an error performance upper bound for the decoding of the network coded messages at the relay when the optimal computation coefficients are selected;

- We showed that, by employing QPSK constellation, the diversity order of decoding the network coded messages at the relay is 1 at the high SNR region, rather than halved when 4-PAM constellation is considered in previous literature.

The related publication is:

- T. Huang, J. Yuan, and J. Li, “Analysis of Compute-and-Forward with QPSK in Two-way Relay Fading Channels,” in *Proceeding of the 14th Australian Communications Theory Workshop (AusCTW)*, pp. 75-80, Adelaide, Australia, Jan. 2013.

The second main contribution of this thesis is that **we analyzed the distance spectrum of channel-coded PNC in Gaussian two-way relay channels with binary classic codes, to provides error rate upper bounds of computing the corresponding network coded messages at the relay.** Channel coding is a vital technique for providing reliable information transfer in noisy physical channels. The research on channel coding is inspired by Claude Shannon’s work in 1948 [9]. The development of channel coding can be divided as classical coding, and modern coding [10]. Classic coding has the following two typical characteristics: classic codes are designed with large minimum distance and strong algebraic structure, and the corresponding decoding algorithms need to exploit the algebraic structure to accomplish bounded distance decoding efficiently [10]. The error rate performance of classic codes in conventional point-to-point channel is characterized by the distance spectrum. However, in PNC, the relay sees a superimposed signal from multiple transmitters, which gives a new challenge on analyzing the distance spectrum of the superimposed codewords at the relay. Motivated by this, we analyze the distance spectrum of channel-coded PNC in Gaussian two-way relay channels with binary classic codes. The detailed contributions of this work are

- We proposed a new approach to explicitly find the distance spectrum of the binary-input channel-coded PNC scheme.

- We derived an asymptotically tight performance bound for the error probability of the binary-input channel-coded PNC scheme;
- We showed that relative to the single-user scenario, the binary-input channel-coded PNC scheme exhibits the same minimum Euclidean distance but an increased number of minimum distance error events;
- We showed that at a high SNR, relative to the single-user scenario, the binary-input channel-coded PNC has an SNR penalty of at most $\ln 2$ in linear scale.

The related publications are:

- T. Yang, I. Land, T. Huang, J. Yuan, and Z. Chen, "Distance Properties and Performance of Physical Layer Network Coding with Binary Linear Codes for Gaussian Two-Way Relay Channels", in *Proceeding of IEEE International Symposium on Information Theory (ISIT)*, pp. 2070-2074, Saint Petersburg, Russia, Aug. 2011.
- T. Yang, I. Land, T. Huang, J. Yuan, and Z. Chen, "Distance Spectrum and Performance of Channel-Coded Physical-Layer Network Coding for Binary-Input Gaussian Two-Way Relay Channels," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1499-1510, June 2012.

Note that this work is a collaborated work with other researchers, and I am the sole student in this collaboration. My contribution in this work include: propose the computation functions for network coded information at the relay and compared their performance difference via simulations; analyze the structure of the superimposed codewords at the relay and formulate the distance properties; derive the union bound in a form of the distance structure; conduct simulation to verify the derived performance bound.

The third main contribution of this thesis is that **we designed irregular repeat-accumulate coded PNC in Gaussian two-way relay channels, to improve**

the error rate performance of computing the corresponding network messages at the relay. Irregular repeat-accumulate codes are a type of modern codes, which can be graphically represented [10]. The encoding and decoding of modern codes are performed locally. The local correlations among the coded bits are simple but the overall code structure is complex due to the large amount of correlations. One of the most difference between classical coding and modern coding is that the minimum Hamming distance no longer plays an important role in modern coding theory [10]. The design of modern codes for point-to-point channel is of great interest because this type of code is powerful and able to achieve capacity. However, in PNC, due to the superposition of the signals at the relay, addition challenge raised for design such modern codes with superimposed signal structure. Motivated by this, we investigate the design of irregular repeat-accumulate coded PNC in Gaussian two-way relay channels in this work. The detailed contributions of this work are

- We analyzed the component decoders of the irregular repeat-accumulate coded PNC scheme and derived the generalized update rules for these components in terms of log-likelihood ratios;

- We proposed two models for the soft information exchange among the components decoders;

- We developed upper and lower bounds on the approximation of the extrinsic information transfer (EXIT) functions of the irregular repeat-accumulate coded PNC scheme;

- We carried out an EXIT chart curve-fitting technique to construct optimized irregular repeat-accumulate codes;

- We showed that our developed irregular repeat-accumulate coded PNC schemes have significantly improved performance compared to the existing regular repeat-accumulate coded PNC schemes;

- We showed that the channel-coded PNC scheme can significantly outperform the complete decoding-based scheme if the code rate is sufficiently high.

The related publications are:

- T. Huang, T. Yang, J. Yuan, and I. Land, "Convergence Analysis for Channel-coded Physical Layer Network Coding in Gaussian Two-way Relay Channels," in *Proceeding of the 8th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 849-853, Aachen, Germany, Nov. 2011.
- T. Huang, T. Yang, J. Yuan, and I. Land, "Design of Irregular Repeat-Accumulate Coded Physical-Layer Network Coding for Gaussian Two-way Relay Channel-s," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 897-909, Mar. 2013.

The fourth main contribution of this thesis is that **we Analyzed and designed Eisenstein integer based lattice network codes for multi-way relay fading channels with PNC scheme, to provide error rate performance bounds of computing the corresponding network messages at the relay, and to provide designing methods for constructing Eisenstein integer based lattice network codes.** Lattice codes are an important class of structured modulation codes, and they are analogue of linear block codes as convolutional codes to trellis coded convolutional codes. It has been shown that lattice codes can achieve capacity on the AWGN channel [12, 13, 17–20]. At the time of this study, a general framework is developed for studying nested-lattice-based PNC schemes, termed as lattice network coding (LNC) schemes [21]. In particular, several generalized constructions of LNC schemes are given for Gaussian integer based lattice. Motivated by the work in [21], we investigate the Eisenstein integer based lattice network coding. The detailed contributions of this work are

- We presented quantization and encoding algorithms for Eisenstein integer based LNCs;

- We derived a union bound estimation of the decoding error probability;
- We generalized the Gaussian reduction algorithm for real lattices over integers [22] to be applicable for complex lattices over Eisenstein integers, and an optimal coefficient vector for Eisenstein integer based LNCs can be efficiently found in the two-transmitter single-relay system via this algorithm;
- We constructed new convolutional LNCs based on both Gaussian integers and Eisenstein integers by Complex Construction A.
- We introduced and analyzed the construction of LNCs from linear codes by Complex Construction A in a relaxed way and by Complex Construction B, with nominal coding gains and union bound estimation explicitly derived;
- We derived optimal dithering method in terms of energy efficiency for LNC over GF(4).
- We established explicit connection between parameters of the linear code and of the corresponding LNC;
- We constructed and analyzed LNCs from convolutional, BCH, and Reed-Solomon codes.

The related publications are:

- Q. Sun, J. Yuan, T. Huang, and W.-K. Shum, "Lattice Network Codes Based on Eisenstein Integers," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2713-2725, July 2013.
- Q. Sun, T. Huang, and J. Yuan, "On Lattice-Partition-Based Physical-Layer Network Coding over GF(4)," *IEEE Communications Letters*, vol. 10, no. 10, pp. 1988-1991, Oct. 2013.

Note that this work is a collaborated work with other researchers, and I am the sole student in this collaboration. My contribution in this work include: propose detailed design and construction methods for Eisenstein integer based lattice network codes; conduct actual code design and search; compute related code parameters; conduct simulations to verify the performance of the designed codes.

The fifth main contribution of this thesis is that **we designed a transmission scheme for pair-wise transmission PNC in single-input single-output multi-way relay channels, to improve the network sum-rate, and to improve the error rate performance of computing the corresponding network coded messages at the relay.** This work is motivated by [184–187]. The aforementioned research works on multi-way relay channels are limited in BPSK modulation, and the pair-wise transmission scheduling at the users side is done in a sequential order. However, higher modulation gives more flexibility of selecting the best network coding function at the relay for fading channel, as discussed in our first main contribution of this thesis. Motivated by this, we investigate the pair-wise transmission scheme PNC employed single-input single-output multi-way relay fading channels with lattice network codes. The detailed contributions of this work are

- We proposed an opportunistic pair-wise compute-and-forward employs high level modulation with nested lattice codes to improve the sum-rate of multi-user transmission;

- We demonstrated that this novel opportunistic pair-wise transmission has a 2 bits/s/Hz improvement in the sum-rate performance at signal-to-noise ratio of 30 dB for a 4-user multi-way relay channel;

- We showed that, for the same multi-way relay channel, up to 4.5 dB gain or 2.5 dB gain can be achieved for a channel-uncoded or a channel-coded system, respectively, at the frame error probability of 10^{-2}.

The related publication is:

- T. Huang, J. Yuan, and Q. Sun, "Opportunistic Pair-wise Compute-and-Forward in Multi-way Relay Channels," in *Proceeding of the IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013.

The sixth main contribution of this thesis is that **we optimized the uplink and downlink channel usage of half-duplex multi-input multi-output multi-way relay channels with PNC to optimize the degrees of freedom capacity, and we designed an iterative algorithm to optimize the precoders at the users and at the relay, to optimize the sum-rate.** So far the works in this thesis are limited to single-input single-output network. Multiple-input multiple-output (MIMO) techniques have been introduced into the study of MWRCs to allow spatial multiplexing [99–105, 109]. The degrees of freedom (DoF) is an important metric to understand the capacity behavior of the MIMO MWRC. In particular, in [99–101], the DoF analysis for MIMO MWRCs is mostly focused on pairwise data exchange. Motivated by their work, we focus on the DoF capacity and sum-rate optimization of the MIMO MWRC with full data exchange operated in half-duplex. Unlike pairwise data exchange, the uplink and downlink traffic loads are asymmetric in full data exchange. Half-duplexing allows unequal time allocation between the uplink and the downlink gives us the flexibility to optimize the uplink/downlink time allocation to maximize the DoF of the half-duplex system. The detailed contributions of this work are

- We derived the DoF capacity of the MIMO MWRC with full data exchange operated in half-duplex and full-duplex modes;
- We derived the optimal uplink/downlink time allocation to maximize the DoF of the half-duplex system;
- We showed that a significant DoF gain can be achieved by the optimal uplink/downlink time allocation, as compared with equal time allocation;

- We showed that the sum-rate is a non-convex function of the user precoders and relay precoder;
- We proposed an iterative algorithm to optimize the user precoders and the relay precoder in an alternating fashion;
- We demonstrated that the system performance can be considerably improved by a careful design of the user and relay precoders;
- We showed that the numerical results for sum-rate analysis agree with the DoF analysis.

The related publications are:

- T. Huang, X. Yuan, and J. Yuan, "Half-duplex MIMO Multi-way Relay Channel with Full Data Exchange: Degrees of Freedom and Sum-rate Optimization," submitted to *IEEE Transactions on Wireless Communications*.
- T. Huang, X. Yuan, and J. Yuan, "Degrees of Freedom of Half-duplex MIMO Multi-way Relay Channel with Full Data Exchange," *IEEE GLOBECOM*, 2014, accepted.

In addition to these main work, the following work are also related to PNC schemes, such as error performance analysis of nested convolutional lattice codes for multi-way relay fading channels, outage performance analysis of analog network coding in generalized two-way multi-hop networks. These works has been published as:

- Y. Ma, T. Huang, J. Li, J. Yuan, Z. Lin, and B. Vucetic, "Novel Nested Convolutional Lattice Codes for Multi-Way Relaying Systems over Fading Channel-s," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2671-2676, Shanghai, China, Apr. 2013.

- G. Wang, W. Xiang, J. Yuan, and T. Huang, "Outage Performance of Analog Network Coding in Generalized Two-Way Multi-Hop Networks," in *Proceeding of IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1988-1993, Quintana-Roo, Mexico, Mar. 2011. (Best Academic Paper Award)

- G. Wang, W. Xiang, J. Yuan, and T. Huang, "Outage Analysis of Non-Regenerative Analog Network Coding for Two-Way Multi-Hop Networks", *IEEE Communications Letters*, vol. 15, no. 6, pp. 662-664, June 2011.

1.3 Literature Review on PNC

Since PNC was first introduced in 2006, it has attracted much research intension. Many researches have been done in the literature to advance the understanding of PNC. In this section, we will review the research on PNC from various aspects.

1.3.1 Network Coding Functions for PNC

A key study area in PNC is about the network coding functions at the relay. As aforementioned that, in PNC, the relay directly recovers the network coded information from its received superimposed signals. A good network coding function allows the relay to effectively compute the corresponding network coded information. There are many network coding functions have been studied in the literature, such as denoise-and-forward [8], amplify-and-forward [11], and compute-and-forward [23]. Usually the study in this area is focus on channel un-coded system, and the network coding function is done at the symbol level.

In denoise-and-forward [8], the network coding function is to map the received signal into a quantized signal. The signal constellation seen at the relay is a superimposed signal constellation of the users. The superimposed signal constellation points need to be clustered into groups, and the superimposed constellation points in each group are mapped to a common network coded message. Thus, the mapping is usually many-to-one. Depends on the type of modulation used at the user side and transmission channel, the clustering and mapping technique can be quite different. The clustering rule is to try to map the closest superimposed constellation points to common network coded information, as well as to maximize the minimum distance among the clusters. For a Gaussian TWRC with BPSK modulation, the denoise mapping is XOR operation as shown in Example 1.1 [1]. However, in fading channels, the mapping can be tricky. Usually the different clustering group is formed based on the instantaneous channel state of the users. The work in [14] studies a denoise-and-forward technique, termed linear PNC scheme for real Rayleigh fading TWRCs. The authors focused on q-pulse amplitude modulation (PAM) where the relay selects some integer coefficients and computes the linear combination of two user messages. Our work in [24] extended the study of linear PNC with quadrature phase shift keying (QPSK) for real Rayleigh fading TWRCs. In the case of complex Relay fading TWRCs, the study in [26] pointed out that sometimes the relay needs to use 5-quadrature amplitude modulation (QAM) rather than the same QPSK modulation at the relay to avoid the ambiguity of decoding at the user side. The work in [25] further extended the work in [14, 24] and studies a linear PNC scheme for fading two-way relay channels. The study is limited to QAM-modulation. In this work, the relay computes the finite-set integer combinations of its received superimposed messages. It has been shown that, in order to minimize the computation error probability of network coded messages at the relay, the chosen integer coefficients actually resemble the fading channel coefficients.

The use of nested lattice codes (compute-and-forward) can also be viewed as a type of denoise-and-forward network coding function [23,27]. In their work, structured nested lattice codes are utilized. The transmitted signals from the users are lattice points in a multi-dimensional lattice over integers. The relay decodes and forwards an integer valued linear combination of transmitted signals to maximize the computation rate. It has been shown that this scheme can achieve an asymptotic gain from the information-theoretic prospective.

The amplify-and-forward scheme is also known as analog network coding (ANC) [11] in the literature. In this scheme, the relay amplifies its received superimposed signals in its transmission phase. The broadcasted signal from the relay retains the noise received at the relay. If the system is channel coded, then the relay amplifies its received signals symbol-by-symbol, and cannot exploit the correlations introduced by the channel coding. The work in [28] discussed this issue by investigating a soft-input soft-output decoder at the relay node, which computes an estimated minimum mean square error (MMSE) packet for forwarding back to the end nodes. Their work has been proven that can improve the end user performance.

From the information theoretic perspective, it has been shown in [29] that PNC with finite-field mapping schemes can achieve near information-capacity rates. Further, [27] found that PNC with finite field mapping can achieve within 1/2 bit of the capacity of the Gaussian TWRCs. It also has been found that infinite-set mapping function cannot achieve near information capacity rates [11, 30].

1.3.2 Channel-Coded PNC

In wireless communications, forward error control codes or channel codes play a vital role of achieving reliable information transmission. So far the discussions of PNC are mainly for channel un-coded systems, where the network coding function is operated at symbol level. Channel coding introduces a correlation among the transmitted symbols within a data packet. Integrating channel coding with PNC scheme is an interesting research topic. The research in this area is focused on linear codes. This is because that in the coding theory, a linear code has the characteristic that any linear combination of codewords is also a codeword. In this case, the network coding function at the relay is also a linear function.

In [46], a regular repeat-accumulate (RA) coded PNC scheme with BPSK modulation for Gaussian TWRCs is firstly investigated. Two interesting decoding approaches have been studied in their work. We now briefly describe these two decoding approaches here.

The first decoding approach can be summarized into two steps:

Step 1: PNC mapping. In this step, the relay firstly uses a minimum distance detection rule to find the closest superimposed signal constellation point to its received signal for each received symbol, and then maps it to the corresponding network coded symbol. Note that in this step, the relay can make hard decision on the network coded symbol value, or obtain a soft-information of the network coded symbol value, e.g., posterior probabilities.

Step 2: Point-to-point channel decoding. In this step, the relay feeds the outputs of Step 1, either hard decision or soft-information, into a point-to-point channel decoder to compute the network coded data packet.

This decoding approach was also studied in [31, 32]. Similar decoding approach was also studied in [27, 33, 34] where lattice coded PNC was considered.

The second decoding approach can also be summarized into two steps:

Step 1: Superimposed channel decoding. In this step, the relay firstly needs to construct a Tanner graph of a superimposed encoder. Then the relay directly decodes the superimposed data packet from its received signal on the Tanner graph of the superimposed encoder. After this step, the relay obtains a hard-decision of the superimposed data packet.

Step 2: PNC mapping. In this step, the relay maps the estimated superimposed data packet to the network coded message symbol-by-symbol.

As discussed in [46] that there exists information loss when mapping from the superimposed constellation points to the network coded symbol value, e.g., the Step 1 of the first decoding approach. We can also see this from the Example 1.1. In Example 1.1, the cardinality of superimposed two BPSK constellation points is 3, but the cardinality of the network coded symbol is 2. Thus, information loss occurs when mapping from the superimposed signal constellation to the network coded symbol value. Therefore, the second decoding approach has better performance than then the first decoding approach.

A convolutional coded PNC scheme with modified Viterbi and BCJR algorithms for TWRCs was studied in [35]. The BER performance of convolutional coded PNC at the relay can be characterized by their corresponding point-to-point channel performance, where conventional decoding algorithms was used at the relay [37, 38]. A nearly optimal decoding performance for element-wise XORed messages can be achieved at the relay node when a two-user joint trellis was used. The number of the states in the two-user joint trellis is the product of numbers of states for individual codes. Hence, the superimposed trellis of two users' codes can be very complex which makes it infeasible to practical decoders. A reduced-state trellis was proposed to reduce decoding complexity. The complexity of the reduced-state decoding is approximately the square root of that of the full-state decoding, but there was a performance loss of approximately 2 dB in SNR paying for reducing the complexity reduction. The work in [39, 40] applied the reduced-state trellis technique in turbo codes decoding at the relay.

Asynchronous PNC with Convolutional Codes was also studied in the literature [41, 42]. In [41], a channel coding scheme based on linear convolutional codes was proposed to relax the strict synchronization requirement. However, this scheme can only deal with integer symbol misalignment. Further, the framework proposed in [42] can deal with phase asynchrony and symbol arrival-time asynchrony between the signals simultaneously transmitted by multiple sources. In particular, this proposed scheme can handle both integer and fractional symbol misalignment.

Structured codes, such as the nested lattice codes, have been studied in PNC [21,23,27,119,121,126]. In [27], an achievable scheme composed of nested lattice codes for the uplink and structured binning for the downlink was proposed. Unlike conventional nested lattice codes, their codes utilize two different shaping lattices for source nodes based on a three-stage lattice partition chain. It has been shown in [27] that, the achievable rate region of this scheme is within 1/2 bit from the capacity region for each user and its sum rate is within $\log(3/2)$ bit from the sum capacity. In [23], the messages from the users were taken from a prime-sized finite field, and then mapped onto lattice points for transmission. Each relay observes a linear combination of these lattice points and attempts to decode an integer combination of them in the same filed. The underlying codes were based on lattice partitions. The achievable rates for sending equations over a finite field from transmitters to relays over real-valued channel models and complex-valued channel models were given. It also discussed the sufficient conditions on the equation coefficients so that a destination can recover one or more of the original messages. The work in [23] is more an information-theoretic approach, where the work in [21] is more an algebraic approach. In [21], a general framework is developed for studying nested-lattice-based PNC schemes, termed as lattice network coding (LNC) schemes. In particular, several generalized constructions of LNC schemes are given. Further, the performance/complexity tradeoffs of LNC schemes are discussed. Our work in [119,121] investigated the Eisenstein Integer based LNC.

The work in [51, 52] investigated LDPC coded PNC. In particular, [52] studied cyclic LDPC coded PNC, and a decoder was proposed to deal with general asynchrony. Our work [120, 127] further investigated Repeat Accumulate (RA) coded PNC [46], where convergence behaviour of the relay decoder was analyzed and a code design based on extrinsic information transfer (EXIT) chart was proposed. The designed codes significantly outperform the corresponding regular codes.

In [45], CPNC with non-binary phase shift keying (PSK) modulation was studied. Working over non-binary fields allows the relay to attempt to decode different network-coded combinations. In particular, the authors compared different mappings between selected message field and the PSK constellation, the drawn conclusion is that many mapping have identical performance in terms of frame error rate (FER).

In [91], a scheme, termed as superimposed XOR, was proposed to consider asymmetry of the channel and information flow of the TWRC in practical system. The proposed relay operation was based on both bitwise XOR and symbol-level superposition coding. This work showed that supposition coding has better performance than conventional XOR in PNC in the broadcast phase when the relay-to-user channels are asymmetrical.

1.3.3 MIMO PNC

In modern wireless communications, multiple-input and multiple-output (MIMO) is an important technique to boost the system performance by exploiting wireless multipath fading, i.e., increased data throughput (via multiplexing gain) and reliability (via diversity gain). In a MIMO system, multiple antennas are used at both the transmitter and receiver. The transmitter can transmit multiple data streams over the antennas to achieve a multiplexing gain that improves the spectral efficiency, as well as a diversity gain that improves the link reliability. MIMO as a mature technique has now been included in many standards such as IEEE 802.11n (Wi-Fi), 4G, 3GPP Long Term Evolution, WiMAX and HSPA+. To this end, MIMO PNC attracts much research attention.

In [48], the authors proposed a scheme where the relay extracts the summation and difference of the two end packets and then converts them to the network-coded form. A linear detection technique is used in this work to reduce the processing complexity. Further, this work shows that MIMO PNC significantly outperforms MIMO with straightforward network coding scheme under random Rayleigh fading channel.

In [98], a network coding scheme was proposed for MIMO TWRC with PNC, where a maximum likelihood algorithm was used to decode the XOR of the superimposed signals received at the relay. In addition, an optimized beamforming algorithm was proposed to maximize the effective channel gains for the two users. This approach was shown to outperform ANC and was able to achieve full diversity gain.

Alamouti codes are the first invented space-time block codes for a point-to-point two-transmit antenna system. The work in [49] analyzed the symbol error rate of a system in which the two end nodes are equipped with two antennas and the relay has only one antenna. With the use of Alamouti codes, it shows that a diversity order of 2 can be achieved. The work in [50] studied the Alamouti codes in a TWRC where all nodes were equipped with two antennas.

The authors in [53–55] studied linear precoding techniques for MIMO PNC. In [53], a reduced-dimension cooperative precoding scheme is proposed for MIMO TWRCs. The analytical result in [53] shows that, in the worst case, the proposed scheme is within a half bit per transmit antenna of the asymptotic sum-capacity of MIMO TWRCs. In [54, 55], an eigen-direction alignment precoding scheme is proposed for MIMO TWRCs. The proposed precoding scheme efficiently aligns the two-user's eigen-modes into the same set of orthogonal directions, and multiple independent PNC streams are implemented over the aligned eigen-modes. In [74], linear decoder of PNC with Alamouti scheme was investigated.

The ANC-based MIMO TWRC was investigated in [63–65]. In [63], each user is equipped with one antenna and the relay is equipped with multiple antennas. The capacity region at the relay was analyzed, and an optimal relay beamforming structure was given to achieve the capacity region. Further, two suboptimal beamforming schemes, based on the matched-filter and zero-forcing (ZF) techniques, were given to lower the relay complexity. In [64], a minimum mean-square-error bidirectional amplify-and-forward relaying protocol was introduced at the relay in a TWRC. In the broadcast phase, the relay selects a single antenna for downlink transmission. The work in [65] considers the application of PNC in a practical cellular system, where the base station and the relay have multiple antennas, and all mobile stations only have a single antenna. In this system, precoding can only be performed at the base station.

MIMO TWRC with multiple relays also attracted much research attention [66–72]. One strategy is to select the appropriate relay to maximize a performance metric [67,70], where [70] focused on ANC and [67] focused on decode-and-forward. Another strategy is to use the multiple relays as a distributed MIMO system [66,68,69,71,72].

The work in [73,75] investigated MIMO PNC with user transmit antenna selection strategy, where each user selects the strongest transmit antenna for the transmission. It was shown in this work that the PNC with user transmit antenna selection outperforms the PNC with space-time block codes significantly.

The degrees of freedom (DoF) is an important metric to understand the capacity behavior of the MIMO systems. The DoF analysis for PNC employed MIMO multiway relay networks has received much attention, such as the work in [99–101], which are mostly focused on pairwise data exchange. In pairwise data exchange, users exchange messages in pairwise fashion. The work in [100] considered a three-user MWRC, termed the MIMO Y channel, and the DoF capacity of the MIMO Y channel was derived under certain relay/user antenna setups. The work in [99] generalized the result of [100] to the case of an arbitrary number of users. Later, the authors in [101] considered MWRCs with clustered data exchange, i.e., the users in the network are grouped into clusters, and only the users in the same cluster communicate with each other. It's worth noting that, in pairwise data exchange, the traffic loads of the uplink and the downlink are symmetric. This uplink/downlink symmetry further implies that the signal space alignment for the uplink straightforwardly carries over to the downlink. This property is used in [99–101] to simplify the beamforming design for MIMO MWRCs with pairwise data exchange. In our work [123,124], we focus on the DoF optimization and sum-rate maximization of MIMO MWRCs with full data exchange. In full data exchange, each user broadcasts their messages to all the other users in the system through the relay, and decodes all the messages from the other users in the network. Therefore, compared to pair-wise data exchange, the uplink and downlink traffic loads are asymmetric in full data exchange. Our work [123,124] showed that a significant DoF gain can be achieved by the optimal uplink/downlink time allocation, as compared with equal time allocation.

1.3.4 Other Topics on PNC

Besides the above introduced research on PNC, there exist many other interested topics such as PNC for orthogonal frequency-division multiplexing (OFDM) schemes, synchronization, and Network topology, etc.

Implementation The very first implementation of a PNC system was introduced in [76, 78]. In this implementation, the network coding was performed in the frequency domain of an OFDM PNC system to eliminate symbol misalignment [31]. This system employed the convolutional code defined in the 802.11 standard.

OFDM-PNC Applying PNC for OFDM systems is firstly studied in [31], and implemented in [76]. The basic idea of OFDM is to carry the symbols on multiple sub-bands. If the sub-bands are smaller than coherent bandwidth of the channel, then the fading in each sub-band is flat. Apply PNC for OFDM provides the ability to deal with symbol offset and non-flat fading. An interesting research direction in OFDM PNC is channel estimation [81–84]. The work in [85] investigated the performance of OFDM systems with imperfect self-information removal at the users, which can be caused by the difference between the estimated CSI and true CSI.

Asynchrony Since the invention of PNC in [1], a question was raised about whether PNC is practical when the system is asynchronous. In reality, signals transmitted by the end users could arrive at the receiver with symbol misalignmen-t, carrier phase offset, and frequency offset. Asynchronous PNC is studied in [77, 88–90]. In [77], the authors pointed out in PNC systems operated with QPSK modulation, there is a significantly power penalty of 6 dB when the carrier phases of the two end nodes are not synchronized and offset by $\pi/4$. This issue was further studied in [90], where a general framework for decoding at the receiver based on belief propagation (BP) was proposed to effectively deal with symbol and phase asynchronies while incorporating channel coding at the same time. Their work shows that the phase penalty can be significantly reduced in channel-coded PNC with a proper receiver algorithm.

Network In general, PNC can be adopted in many types of network topologies other than TWRCs. In [1, 122, 128], the application of PNC in linear networks were investigated, where two end users exchange information through a serial of relay nodes in between. Another interested type of network is called multi-way relay channels (MWRCs) [93], where more than two users exchange information via a common relay. The works in [92, 94, 95] studied MWRCs with PNC in packet level. The works in [96, 97, 106–109] focused on the MWRCs where users equipped with single antenna and the relay equipped with multiple antennas. The works in [99–103, 123] investigated Multiple-input multiple-output (MIMO) MWRCs with PNC. Our work in [125] investigated the scheduling issue in an MWRC where only two users are allowed for transmission at one time.

Modulation In [43], a modified high-order PAMs for binary coded PNC was investigated. In particular, a non-uniform M-PAM ($M > 2$) signal constellations was proposed to lower the complexity of the PNC with high-order PAMs by utilizing binary codes. In [44], the design of modulation schemes for the PNC when the end nodes use square QAM constellation was studied. In [87], the design of relay receiver was investigated when noncoherent PNC with FSK modulation was employed.

Channel Estimation Channel state information (CSI) is an important parameter in wireless communications. This parameter measures how the transmitted signal been affected by the wireless channel, such as fading, scattering, etc. With the knowledge of CSI, the transmitter can actively adapt its transmission strategy to suit the current channel condition. The work in [79, 80] focused on the channel estimation for PNC. In [79], the channel estimation of a TWRC where each node equipped with one antenna was investigated under AF relay operation. A linear maximum SNR channel estimator was proposed and designed to maximize the average effective SNR. The work in [80] focused on the same system model as in [79]. In [80] the author suggested that the relay can actually estimate the channel first and then system power allocation can be processed. By doing so, the final data detection at the users can be optimized. Two different power allocation schemes to the training signals were proposed in their work. The corresponding optimization targets were to maximize the average effective SNR ratio of the data detection and minimize the mean-square-error (MSE) of the channel estimation, respectively. In [86], a technique termed blind known-interference cancellation was proposed to estimate the channel of the superimposed signals at the relay. It has been shown that it can almost achieve the performance of a conventional point-to-point channel.

1.4 Brief Review of Channel Coding Theory

Since most of my time in PhD study is investigate the channel-coded PNC, and the majority part of this thesis is about the channel-coded PNC, we also provide a brief review of channel coding theory here.

The research on channel coding is inspired by Claude Shannon's paper "A Mathematical Theory of Communication" in 1948 [9]. In this paper, it states that, for transmitting information in a noisy channel, there exist encoding and decoding schemes that can be used to ensure that the probability of decoding error goes to arbitrarily small for a sufficiently large code block length, provided that the code rate does not exceed the capacity of the channel.

Since the publish of [9], great efforts have been devoted by researchers to develop channel codes for point-to-point channels. Later, it was shown in [114] that simple structure of linear codes can achieve channel capacity. The development of channel coding can be divided as classical coding, and modern coding [10].

Classic coding has the following two typical characteristics: classic codes are designed with large minimum distance and strong algebraic structure, and the corresponding decoding algorithms need to exploit the algebraic structure to accomplish bounded distance decoding efficiently [10]. As pointed out in [10], classical coding cannot achieve the capacity.

Some examples of classic codes include: Hamming codes [111], BCH codes [112, 113], convolutional codes [114], Reed-Solomon codes [115], and concatenated codes [116], etc. Among them, Hamming codes, BCH codes, and Reed Solomon codes are classified as block codes due to the fact that the coding is operated block by block. On the contrary, convolutional codes can have an arbitrary length. This is because a convolutional code is generated by passing the information sequence to a finite-state filter with memory register. Each encoded output is a function of the present input and previous inputs.

In modern coding, codes can be graphically represented [10]. The encoding and decoding of modern codes are performed locally. The local correlations among the coded bits are simple but the overall code structure is complex due to the large amount of correlations. This type of code structure is powerful and able to achieve capacity. One of the most difference between classical coding and modern coding is that the minimum Hamming distance no longer plays an important role in modern coding theory [10]. Examples of modern codes are turbo codes, low-density parity-check (LDPC) codes, and repeat-accumulate (RA) codes.

Turbo codes was introduced by Berrou, Glavieux, and Thitimajshima in their paper "Near Shannon Limit Error-correcting Coding and Decoding: Turbo-codes" in 1993 [129], and this type of codes are the first practical codes to closely approach the channel capacity. A detailed review of the turbo principles, and the classic maximum *a posteriori probability* decoder can be found in [60]. The study in [60] also covers the iterative turbo receivers.

LDPC codes are a class of recently re-discovered highly efficient linear block codes [117]. LDPC codes were first introduced by Robert G. Gallager in his PhD thesis in

1960, but been ignored by the research community due to the computation complexity of implementing the encoder and decoder. LDPC codes have been shown to be able approach the channel capacity using an iterated soft-decision decoding approach [118]. The survey study in [62] extensively reviewed the literature on the family of low-density parity-check LDPC codes and their rateless relative codes. Dariush *et al.* introduced RA codes in [130]. RA codes can be viewed as special LDPC codes with a simpler encoder than general LDPC codes but with similar performance.

Polar code, firstly introduced by Erdal Arikan [131] is the first code with an explicit construction to provably capacity achieving on binary-input memoryless output-symmetric channels, and this is the most recent development in the coding theory. The minimum distance of polar codes is proportional to the square root of their block length, and hence no error floor occurs. The encoding and decoding complexity of polar codes is also low.

Lattice codes are an important class of structured modulation codes, and they are analogue of linear block codes as convolutional codes to trellis coded convolutional codes. It has been shown that lattice codes can achieve capacity on the AWGN channel [12, 13, 17–20]. The construction of lattice codes is in the Euclidean space and the construction process is algebraic in nature.

The study of coding in TWRCs also attracted much research attention. Here we give two examples of study coding in TWRCs [15, 16]. In [15], Aljohani et al. proposed and studied a joints source-coding, channel coding and modulation scheme for a two-way relaying system, and the authors have demonstrated that a significant coding gain can be achieved when comparing the proposed scheme to the conventional scheme. In [16], a novel power and bandwidth-efficient turbo trellis coded modulation assisted space division multiple access based two-way relaying scheme was proposed. This scheme was designed by Liang *et al.* for enhancing the throughput, reliability and coverage area in a cooperative communication system [16].

1.5 Outline of the Thesis

This thesis is focus on the analysis, design, and optimization of PNC employed relay networks. The aforementioned contributions will be included in this thesis. We will start with the study on the network coding function of channel-uncoded PNC employed two-way relay networks. We then focus on the error performance of the channel-coded PNC employed two-way relay networks with classic codes. We then investigate the design of the channel-coded PNC employed two-way relay networks with modern codes. We then shift our focus to the analysis and design of channel-coded PNC employed multi-way relay networks with lattice network codes. We then study the user scheduling problem in channel-uncoded/channel-coded PNC employed multi-way relay networks. At the end, we investigate the degrees of freedom and sum-rate optimization in PNC employed multi-way relay networks. The following is a description of the organization of this thesis.

In **Chapter 1**, we start by introducing the concept of PNC, and its benefit to the wireless communications. We then focus on the motivations and contributions of this thesis. At the end of this chapter, We provides a literature review of current research progress on PNC, as well as a brief literature on channel coding theory.

In **Chapter 2**, we focus on the design of optimal network coding function at the relay for channel-uncoded PNC in two-way relay fading channels with QPSK modulation, to improve the error rate performance of detecting the channel-uncoded network messages at the relay. This chapter begins by presenting the system model, including the detailed network computation process at the relay. We then derive the computation error probability for computing network coded messages at the relay. We then investigate the selection of optimal computation coefficients depends on the channel condition to maximize the interested minimum distance. At the end of this chapter, we will show the comparison between the analytical results and the simulation results.

In **Chapter 3**, we analyze the distance spectrum of channel-coded PNC in Gaussian two-way relay channels with binary classic codes, to provides error rate upper bounds of applying classic binary codes in PNC employed relay networks. This chapter begins by presenting the preliminary knowledge of classic codes. We then study the structure of the superimposed codewords, and analyze the distance spectrum. At the end, we show the error rate performance bound and demonstrate by simulations.

In **Chapter 4**, we shift our focus to the design of irregular repeat-accumulate coded PNC in Gaussian two-way relay channels, to improve the error rate performance of computing the channel-coded network messages at the relay. This chapter firstly introduces code and extrinsic information transfer (EXIT) chart. Then we move onto the irregular repeat-accumulate coded PNC scheme and analyze the corresponding component decoders and derive the generalized update rules for these components. We then develop bounds on the approximation of the EXIT functions of the irregular repeat-accumulate coded PNC and further utilize the developed bounds to optimize the irregular repeat-accumulate codes. At the end, we demonstrate the significant gain obtained via simulations.

In **Chapter 5**, we analyze and design Eisenstein integer based lattice network codes for PNC employed multi-way relay fading channels, to provide error rate performance bounds and to provide construction methods for constructing Eisenstein integer based lattice network codes. This chapter begins by visiting the basic models of compute-and-forward and lattice network coding. We then derive the error performance bound. We then focus on construction methods for Eisenstein integer based lattice network codes with general lattice partition. We then focus on construction of Eisenstein integer based lattice network codes with over GF(4). At the end of this chapter, we demonstrate the error rate bound, and show the performance of the constructed codes.

In **Chapter 6**, we design a transmission scheme for pair-wise transmission PNC in single-input single-output multi-way relay channels, to improve the network sum-rate and to improve the error rate performance at the relay. This chapter begins by introduce the background of this study. We then give detailed description of the system model. We then introduce the conventional successive pair-wise transmission scheme in the considered MWRCs. After that, we study the proposed opportunistic pair-wise transmission. In the simulation results, we show the performance improvement in both sum-rate and error rate at the relay.

In **Chapter 7**, we optimize the uplink/downlink channel usage of multi-input multi-output multi-way relay channels with PNC to maximize the degrees of freedom, and design an iterative optimization algorithm to maximize the sum-rate. This chapter begins by brief review the related work. We then introduce the detailed system model. After that, we derive the DoF capacity of the considered MIMO MWRCs with fixed channel uses. We then focus on the optimization of the DoF and corresponding uplink/downlink channel allocation. We then focus on the sum-rate optimization. In the end of this chapter, we demonstrate the optimization results via simulations.

In **Chapter 8**, we conclude this thesis and discuss possible extensions and future work.

Chapter 2

Design of Channel-Uncoded PNC in TWRCs

2.1 Introduction

In this chapter, we focus our study on the design of network coding functions in PNC employed two-way relay fading channels, where quadrature phase shift keying (QPSK) modulation is employed at the users. As introduced in Section 1.3.1 that an effective network function is important for a relay, in a physical-layer network coding (PNC) employed network, to compute reliable network coded information. This work falls in a more general linear PNC with higher modulation employed at the users and the relay. With higher modulation, the relay has more freedom to choose good linear coefficients for different users depends on their channel condition, which is the motivation of this work. Previous work in [14] investigated the asymptotically optimal error-rate performance when pulse amplitude modulation (PAM) was considered. Different from the approach in [14], we focus on the distance profile between any two immediate neighboring constellation points at the relay. Base on that, we select the optimal computation coefficients in terms of minimizing the computation error probability at the relay.

This chapter begins by presenting the system model, including the detailed network computation process at the relay. Then we derive the computation error probability for computing network coded messages at the relay, and the computation error probability is a function of a special minimum distance on the superimposed constellation seen by the relay. We then investigate how to select optimal computation coefficients depends on the channel condition to maximize the interested minimum distance. At the end of this chapter, we will show the comparison between the analytical results and the simulation results.

2.2 System Model

2.2.1 System Overview

The system model is depicted in Fig. 2.1. Two single-antenna users, denoted by A and B, exchange information via a single-antenna relay R. The users and the relay operate in a half-duplex mode and there is no direct link between the users. The complete information exchange between the users is performed in the multiple access phase followed by the broadcast phase. In the multiple access phase, two users transmit simultaneously to the relay. We assume that the relay knows the channel state information of the users. When the relay receives the superimposed signal from the two users, it computes the corresponding network coded message. In the broadcast phase, the relay broadcasts the computed message back to the users. After the users receive the network coded message from the relay, each user can retrieve each other's message by canceling its own message. The complete information exchange between the two users is accomplished in two time slots.

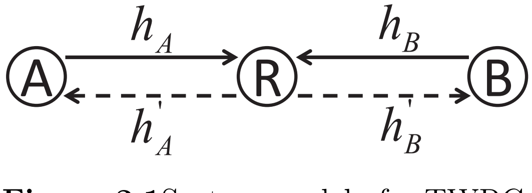


Figure 2.1 System model of a TWRC.

2.2.2 Multiple Access Phase

In this work, we consider an uncoded system with QPSK signaling. Let w_A denote the message of user A. The message is uniformly generated from a Galois field of size 4, i.e., $w_A \in GF(2^2)$. Note that when constructing the $GF(2^2)$, one may form a set of elements $\{0, 1, \alpha, \alpha^2\}$. For the ease of representation, we map α to value 2, and map α^2 to value 3. Then the message set becomes $\{0, 1, 2, 3\}$. Let $\mathcal{M}_A(\cdot)$ denote the signal constellation mapper for node A, and x_A denote the mapped signal. We have $x_A = \mathcal{M}_A(w_A)$. The same procedure also applies to user B.

Assuming perfect synchronization at the relay and equal transmission power at the two end users, the signal received by the relay is

$$y_R = h_A\sqrt{E}x_A + h_B\sqrt{E}x_B + n_R, \tag{2.1}$$

where E is the average transmission energy per symbol, h_A and h_B are the corresponding channel coefficients. In this study, we follow the study in [14], where only amplitude distribution is considered. The introducing of the phase distribution further complicates the system model, and will be studied in future. In this expression, n_R is a complex circularly-symmetric additive white Gaussian noise with zero mean and power spectrum density N_0.

The goal for the relay is to compute a network coded message from its received signal y_R, and then broadcast it to the two users. Given a computation coefficient vector $\mathbf{a} = [a_A\ a_B] \in GF(2^2) \setminus \{0\}$, i.e., $a_A, a_B \in \{1, 2, 3\}$, the *network coded message* is defined as

$$w_R^{\mathbf{a}} \triangleq (a_A \otimes w_A) \oplus (a_B \otimes w_B) \in GF(2^2) \tag{2.2}$$

where $\otimes$ and $\oplus$ denote the addition and multiplication in $GF(2^2)$, respectively. Note that w_B can be uniquely determined via (2.2) when given $\mathbf{a}$, $w_R^{\mathbf{a}}$, and w_A, so that there will be no ambiguity when recovering user B's information at user A. The same holds for user B. A computation error at the relay is declared if $\hat{w}_R^{\mathbf{a}} \neq w_R^{\mathbf{a}}$.

2.2.3 Broadcast Phase

In the downlink phase, the relay modulates the computed network coded message $x_R = \mathcal{M}_R(w_R^{\mathbf{a}})$, and then broadcasts it to the two users. The received signal at user A is $y_A = h_A^{'}\sqrt{E}x_R + n_A$, where $h_A^{'}$ is the channel coefficient from the relay to user A. We assume the users know the computation vector $\mathbf{a}$. After recovering $w_R^{\mathbf{a}}$ from the received signal, user A can extract user B's message by canceling its own message w_A. This also holds for user B.

2.3 Computation at Relay and Error Probability

2.3.1 Computation at the Relay

The received signal at the relay is a superimposed signal corrupted with noise. Let us define the *superimposed signal* as

$$x_S \triangleq h_A\sqrt{E}x_A + h_B\sqrt{E}x_B. \tag{2.3}$$

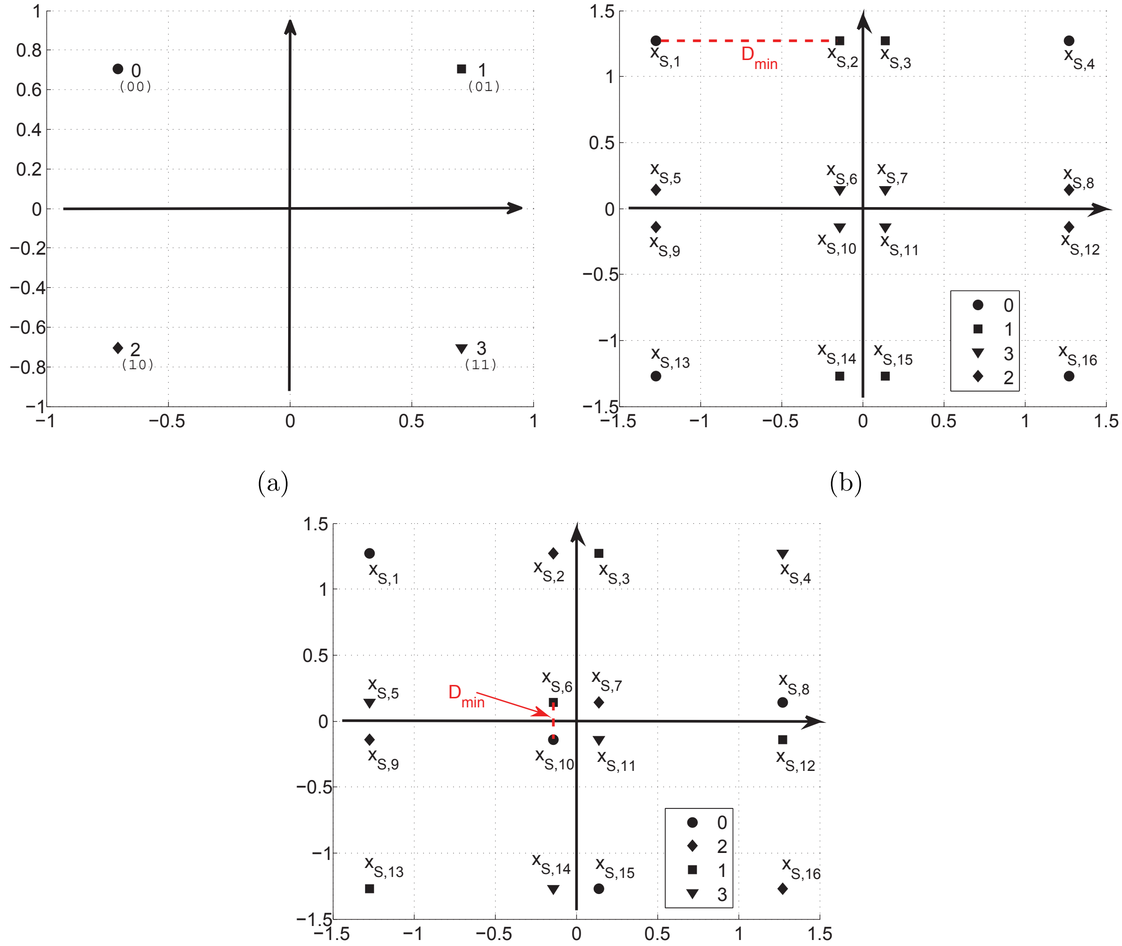
Given a channel realization vector $\mathbf{h} = [h_A \ h_B]$, a superimposed signal constellation is defined as a collection of all possible superimposed signals

$$\mathbb{X}_S \triangleq \{x_S|\mathbf{h}\}. \tag{2.4}$$

Given a QPSK constellation as shown in Fig. 2.2(a), an example of a superimposed signal constellation for QPSK is shown in Fig. 2.2(b). For the ease of referencing, we label the superimposed constellation points sequentially from $x_{S,1}$ to $x_{S,16}$. The cardinality of $\mathbb{X}_S$ is 16.

We now consider the connection between the superimposed signal constellation points and the network coded messages. For a given computation coefficient vector $\mathbf{a} = [a_A \ a_B]$, define the following set

$$\mathcal{X}_S(w_R^{\mathbf{a}}) \triangleq \{x_S : w_R^{\mathbf{a}} = (a_A \otimes w_A) \oplus (a_B \otimes w_B)\}, \tag{2.5}$$



(a)

(b)

(c)

Figure 2.2 Example of QPSK constellation and superimposed two QPSK modulation with different computation coefficient vector $\mathbf{a}$. In (a), it is a normalized QPSK constellation with unit symbol energy and Gray coded symbol value. In (b) and (c), $\mathbf{h} = [1\ 0.8]$. In (b), $\mathbf{a} = [1\ 1]$. In (c), $\mathbf{a} = [1\ 2]$.

which collects all the superimposed signal constellation points x_S corresponding to the same network coded message $w_R^{\mathbf{a}}$. We see from (2.5) that, there are only 4 possible x_S corresponding to a given network coded message $w_R^{\mathbf{a}}$. Hence, the cardinality of a set $\mathcal{X}_S(w_R^{\mathbf{a}})$ is 4, and $p(x_S|w_R^{\mathbf{a}}) = \frac{1}{4}, x_S \in \mathcal{X}_S(w_R^{\mathbf{a}})$. Note that $\mathbb{X}_S = \bigcup_{w_R^{\mathbf{a}}} \mathcal{X}_S(w_R^{\mathbf{a}})$. Fig. 2.2(b) and Fig. 2.2(c) show the examples for a superimposed QPSK constellation with computation coefficient vector $\mathbf{a} = [1\ 1]$, and $\mathbf{a} = [1\ 2]$, respectively. The superimposed constellation points x_S corresponding to the same network coded message $w_R^{\mathbf{a}}$ is marked out by the same legend. We can see from these two figures that, different computation coefficient vector can result in different mappings from the superimposed signals to the network coded messages.

After receiving y_R, the relay can employ two rules to compute the network coded message: optimal maximum likelihood computation rule and sub-optimal minimum distance computation rule, see [37]. In this paper, we consider the minimum distance computation rule. In the minimum distance computation rule, the relay firstly finds out the most likely superimposed signal which has the smallest Euclidean distance from the received signal, that is

$$\hat{x}_S = \underset{x_S \in \mathbb{X}_S}{\arg\min} |y - x_S|^2. \tag{2.6}$$

Then, the corresponding network coded message can be determined by

$$\hat{w}_R^{\mathbf{a}} = \hat{x}_S \in \mathcal{X}_S(w_R^{\mathbf{a}}). \tag{2.7}$$

2.3.2 Computation Error Probability

Given a computation coefficient vector $\mathbf{a}$ and a genuine network coded message $w_R^{\mathbf{a}}$, a decoding error occurs if the computed network coded message $\hat{w}_R^{\mathbf{a}} \neq w_R^{\mathbf{a}}$. With the MD computation rule, and $p(x_S|w_R^{\mathbf{a}}) = \frac{1}{4}, x_S \in \mathcal{X}(w_R^{\mathbf{a}})$, the pair-wise error probability between $w_R^{\mathbf{a}}$ and $\hat{w}_R^{\mathbf{a}}$ is upper bounded by

$$p_e(\hat{w}_R^{\mathbf{a}}|w_R^{\mathbf{a}}) \leqslant \sum_{\substack{x_S \in \mathcal{X}_S(w_R^{\mathbf{a}}) \\ \hat{x}_S \in \mathcal{X}_S(\hat{w}_R^{\mathbf{a}})}} \frac{1}{4} p_e(\hat{x}_S|x_S). \tag{2.8}$$

We now consider the pair-wise error probability between two superimposed signals $\hat{x}_S$ and x_S

$$p_e(\hat{x}_S|x_S) = Q\left(\sqrt{\frac{(\hat{x}_S - x_S)^2}{2N_0}}\right). \tag{2.9}$$

The average network coded symbol error probability is

$$p_e = \sum_{w_R^{\mathbf{a}}} p(w_R^{\mathbf{a}}) \sum_{\hat{w}_R^{\mathbf{a}} \neq w_R^{\mathbf{a}}} p_e(\hat{w}_R^{\mathbf{a}} | w_R^{\mathbf{a}}), \tag{2.10}$$

$$\leq \sum_{w_R^{\mathbf{a}}} p(w_R^{\mathbf{a}}) \sum_{\hat{w}_R^{\mathbf{a}} \neq w_R^{\mathbf{a}}} \sum_{\substack{x_S \in \mathcal{X}_S(w_R^{\mathbf{a}}) \\ \hat{x}_S \in \mathcal{X}_S(\hat{w}_R^{\mathbf{a}})}} \frac{1}{4} p_e(\hat{x}_S | x_S), \tag{2.11}$$

$$\overset{(a)}{=} \frac{1}{4^2} \sum_{w_R^{\mathbf{a}}} \sum_{\hat{w}_R^{\mathbf{a}} \neq w_R^{\mathbf{a}}} \sum_{\substack{x_S \in \mathcal{X}_S(w_R^{\mathbf{a}}) \\ \hat{x}_S \in \mathcal{X}_S(\hat{w}_R^{\mathbf{a}})}} Q\left(\sqrt{\frac{(\hat{x}_S - x_S)^2}{2N_0}}\right), \tag{2.12}$$

where step (a) follows from the fact $p(w_R^{\mathbf{a}}) = \frac{1}{4}$. Let $d_{\min}^2(w_R^{\mathbf{a}})$ denote the minimum inter-set squared Euclidean distance (SED) to the superimposed signals corresponding to network coded message $w_R^{\mathbf{a}}$, we have

$$d_{\min}^2(w_R^{\mathbf{a}}) \triangleq \min_{\substack{x_S \in \mathcal{X}_S(w_R^{\mathbf{a}}) \\ \hat{x}_S \in \mathbb{X}_S \backslash \mathcal{X}_S(w_R^{\mathbf{a}})}} (\hat{x}_S - x_S)^2, \tag{2.13}$$

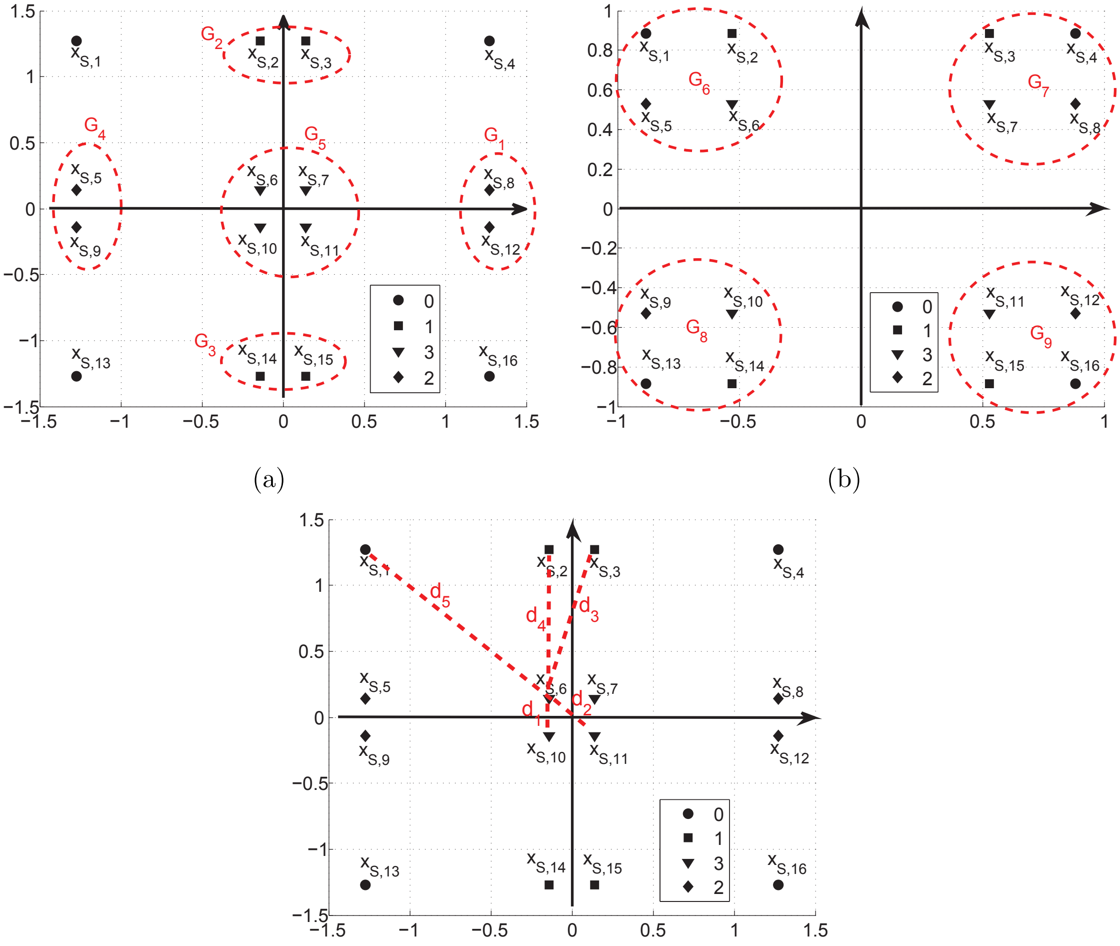
where $\mathbb{X}_S \backslash \mathcal{X}_S(w_R^{\mathbf{a}})$ denotes the set of all superimposed signals excluding the ones corresponding to the network coded message $w_R^{\mathbf{a}}$. Then the minimum inter-set SED among all $w_R^{\mathbf{a}}$ is defined as

$$D_{\min}^2 \triangleq \min_{w_R^{\mathbf{a}}} d_{\min}^2(w_R^{\mathbf{a}}). \tag{2.14}$$

In the high signal-to-noise (SNR) region, the $D_{\min}^2$ dominates the error probability of computing $w_R^{\mathbf{a}}$ [14]. Let $\mathcal{A}$ denote the multiplicity for $D_{\min}^2$. The bound in (2.12) can be simplified to

$$p_e \leq \frac{1}{4^2} \mathcal{A} Q\left(\sqrt{\frac{D_{\min}^2}{2N_0}}\right). \tag{2.15}$$

From (2.14) and (2.15), we see that this upper bound is related to the computation coefficient vector $\mathbf{a}$. This is illustrated in Fig. 2.2(b) and Fig. 2.2(c). For the given channel coefficient vector $\mathbf{h} = [1\ 0.8]$, the $D_{\min}^2$ obtained by selecting $\mathbf{a} = [1\ 1]$ is larger than the $D_{\min}^2$ obtained by selecting $\mathbf{a} = [1\ 2]$. In this case, the computation coefficient vector $\mathbf{a} = [1\ 1]$ can result in a better decoding performance than $\mathbf{a} = [1\ 2]$. This suggests that for each pair of channel realization, an optimal computation coefficient vector $\mathbf{a}_{\text{opt}}$ can be found to maximize the $D^2_{\min}$, which in turn minimizes the average computation error probability.



(a)

(b)

(c)

Figure 2.3(a) Superimposed two QPSK constellations when $\eta = 1.25$. (b) Superimposed two QPSK constellations when $\eta = 4$. (c) Interested distances in a superimposed constellation.

2.4 Distance Analysis and Performance Bound

In this work, we select the QPSK constellation with points $\frac{1}{\sqrt{2}}\{-1+i, 1+i, -1-i, 1-i\}$. The symbol value can be mapped to the constellation points by using Gray coding as shown in Fig.2.2(a).

2.4.1 Distance Analysis

Let us define $\eta \triangleq \frac{h_A}{h_B}$. We focus on the case when $h_A \geq h_B$, and this gives $\eta \geq 1$. We note that the case when $0 < \eta < 1$ is the same as $\eta > 1$. Two typical examples of the superimposed constellation are shown in Fig. 2.3(a) and Fig. 2.3(b) where $\eta = 1.25$ and $\eta = 4$, respectively. The two user's messages embedded in each superimposed constellation points are listed in Table 2.1.

Table 2.1 Embedded two users' message for each superimposed constellation points

x_S	$x_{S,1}$	$x_{S,2}$	$x_{S,3}$	$x_{S,4}$	$x_{S,5}$	$x_{S,6}$	$x_{S,7}$	$x_{S,8}$
w_A	0	0	1	1	0	0	1	1
w_B	0	1	0	1	2	3	2	3

x_S	$x_{S,9}$	$x_{S,10}$	$x_{S,11}$	$x_{S,12}$	$x_{S,13}$	$x_{S,14}$	$x_{S,15}$	$x_{S,16}$
w_A	2	2	3	3	2	2	3	3
w_B	0	1	0	1	2	3	2	3

The error probability upper bound in (2.15) is determined by the minimum inter-set SED $D^2_{\min}$. The value of $D^2_{\min}$ is determined by the distances between the neighboring points of the superimposed signal constellation. For the QPSK modulation, there are five different neighboring distances between the superimposed signal points as illustrated in Fig. 2.3(c). Now let us define these five distances: d_1 is the distance between $x_{S,6}$ and $x_{S,10}$, d_2 is the distance between $x_{S,6}$ and $x_{S,11}$, d_3 is the distance between $x_{S,6}$ and $x_{S,3}$, d_4 is the distance between $x_{S,6}$ and $x_{S,2}$, and d_5 is the distance between $x_{S,6}$ and $x_{S,1}$. Using geometry, we calculate d_1^2 as

$$d_1^2 = 2(h_A - h_B)^2 E. \tag{2.16}$$

Here, we introduce a normalization factor Eh_B^2, and divide this on both sides of (2.16), we have

$$d^2_{1,\text{norm}} = 2\eta^2 - 4\eta + 2. \tag{2.17}$$

Similarly, other interested normalized SED can be calculated:

$$d^2_{2,\text{norm}} = 4\eta^2 - 8\eta + 4, \tag{2.18}$$

$$d^2_{3,\text{norm}} = 2\eta^2 - 4\eta + 4, \tag{2.19}$$

$$d^2_{4,\text{norm}} = 2, \tag{2.20}$$

$$d^2_{5,\text{norm}} = 2\sqrt{2}. \tag{2.21}$$

Fig. 2.4 shows the normalized SED against η.

We now illustrate the procedure to find the optimum computation coefficient vector to maximize D^2_{min}. Fig. 2.4 shows that when $1 < \eta < 1 + \sqrt{\sqrt{2} - 1}$, we have $d^2_{1,\text{norm}} < d^2_{2,\text{norm}} < d^2_{4,\text{norm}} < d^2_{3,\text{norm}} < d^2_{5,\text{norm}}$, where $1 + \sqrt{\sqrt{2} - 1}$ is the η value when $d^2_{3,\text{norm}} = d^2_{5,\text{norm}}$. We firstly try to group any two superimposed signal points that have the smallest distance d_1 in Fig. 2.3(c), in order to increase the minimum inter-set distance. This suggests that $(x_{S,6},\ x_{S,10})$ should be mapped to the same network coded message. This also applies to other pairs of points that have the same distance, such as $(x_{S,6},\ x_{S,7})$, $(x_{S,10},\ x_{S,11})$, $(x_{S,7},\ x_{S,11})$, $(x_{S,2},\ x_{S,3})$, $(x_{S,8},\ x_{S,12})$, $(x_{S,14},\ x_{S,15})$, and $(x_{S,5},\ x_{S,9})$. Interestingly this means that the four points in the middle of the superimposed constellation $(x_{S,6},\ x_{S,7},\ x_{S,10},\ x_{S,11})$ should be grouped together. The exact grouping is shown as the circles with labels from G_1 to G_5 in Fig. 2.3(a). Let $w_A(x_S)$ denote the user A's embedded message in a superimposed constellation point x_S. For G_1 we have $w_A(x_{S,8}) = w_B(x_{S,12})$, and $w_B(x_{S,8}) = w_A(x_{S,12})$ as illustrated in Table 2.1. G_2, G_3, and G_4 also have the same relationship of their embedded users' messages. This requires the computation coefficients $a_A = a_B \in GF(2^2) \setminus \{0\}$ for G_1, G_2, G_3, and G_4 to map to their corresponding common network coded messages. One can verify that the same computation coefficient $a_A = a_B \in GF(2^2) \setminus \{0\}$ will also enable the superimposed points in G_5 to be mapped to a common network coded message.

After finishing grouping any two superimposed constellation points with the smallest distance d_1, we now proceed to grouping any two superimposed constellation points with the second smallest distance d_2, to further increase the minimum inter-set distance. The distance d_2 exists in the following pairs of constellation points: ($x_{S,6}$, $x_{S,11}$) and ($x_{S,7}$, $x_{S,10}$). Note that these two pairs of constellation points have been grouped in the grouping process for distance d_1. This means that d_2 is already an intra-set distance.

Next, we group any pair of superimposed constellation points with distance d_4. There are in total 16 pairs of superimposed constellation points having distance d_4, 4 pairs of superimposed constellation points in each quadrant. Take the points in the first quadrant as an example. The 4 pairs of points are: ($x_{S,3}$, $x_{S,4}$), ($x_{S,4}$, $x_{S,8}$), ($x_{S,8}$, $x_{S,7}$), ($x_{S,7}$, $x_{S,3}$). Notice that grouping each pair means that the 4 points $x_{S,3}$, $x_{S,4}$, $x_{S,7}$, and $x_{S,8}$ should be grouped and mapped to the same network coded message. The grouping for distance d_4 is shown in Fig. 2.3(b) with labels from G_6 to G_9. However, this grouping cannot be achieved. For example, for the 4 superimposed constellation points in G_7, they have a common embedded message from user A and 4 different embedded message from user B as shown in Table 2.1. In this case, a computation coefficient vector $\mathbf{a}$ cannot be found to allow these 4 superimposed constellation points to map to a common network coded message via (2.2). In fact, with any possible computation coefficient vector $\mathbf{a}$, the 4 points in G_7 always map to 4 different network coded messages. This is also true for G_6, G_8, and G_9. This suggests that any two superimposed constellation points with distance d_4 cannot be grouped and mapped to a common network coded message. Then d_4 becomes $D_{\min}$, and the corresponding multiplicity amounts to $\mathcal{A} = 32$.

With the above discussion and the observation in Fig. 2.4, we conclude the following two cases:

Case 1) $1 < \eta < 2$*:* In this case, d_1 is the smallest distances and any pair of superimposed points with this distance can be grouped and mapped to the same network coded message. This grouping process can proceed up to the distance d_4. Recall that any pair of superimposed points with distance d_4 cannot be grouped, and this suggests that $D_{\min} = d_4$. The requirement for the optimum coefficients is simply

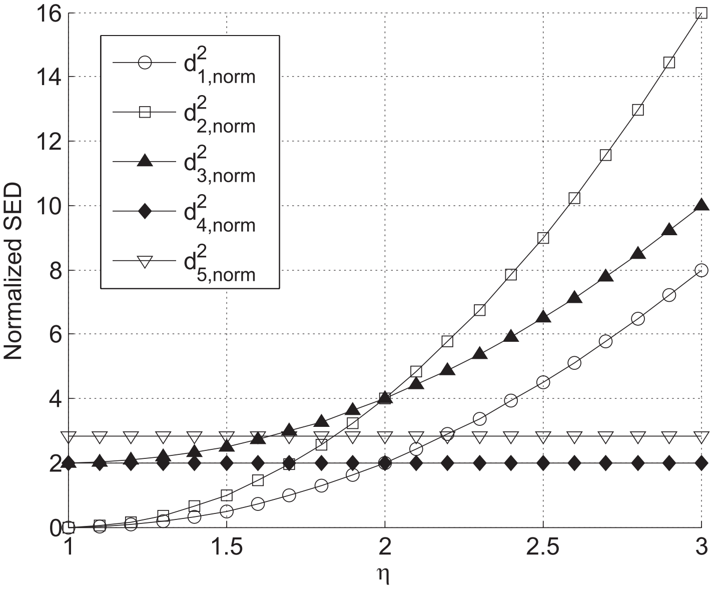


Figure 2.4 Normalized SED.

$a_A = a_B \in GF(2^2) \setminus \{0\}$. An example of the grouping result is shown in Fig. 2.3(a), where different mapped network coded messages are denoted by different marks.

Case 2) $\eta > 2$: In this case, the minimum distance is d_4. Previous analysis shows that any pair of the superimposed constellation points with distance d_4 cannot be grouped and mapped to the same network coded message. This suggests that there is no optimum computation coefficient vector to improve the computation error performance at the relay. An example of this case is shown in Fig. 2.3(b), where different mapped network coded messages are denoted by different markers.

2.4.2 Computation Performance Bound

Firstly, We derive the statistic characteristic of η. With the definition of η, its cumulative distribution function (CDF) can be written as

$$F(\eta) = \int_{h_B=0}^{h_B=\infty} \int_{h_A=0}^{h_A=\eta h_B} P(h_A, h_B) dh_A dh_B, \tag{2.22}$$

$$\overset{(b)}{=} \int_{h_B=0}^{h_B=\infty} \int_{h_A=0}^{h_A=\eta h_B} P(h_A) P(h_B) dh_A dh_B, \tag{2.23}$$

where step (b) follows the fact that any two user channels are independently faded, and $P(h_A) = 2h_A \exp(-h_A^2)$, $P(h_B) = 2h_B \exp(-h_B^2)$. Substituting $P(h_A)$, and $P(h_B)$ into (2.23), we have

$$F(\eta) = 1 - \frac{1}{1+\eta^2}, \ \eta > 0. \tag{2.24}$$

Its probability density function (PDF) can be derived as

$$f(\eta) = \frac{dF(\eta)}{d\eta} = \frac{2\eta}{(1+\eta^2)^2}. \tag{2.25}$$

We now study the average Error probability for $w_R^{\mathbf{a}_{\text{opt}}}$. We firstly consider the case when $1 < \eta < 2$. From the previous distance analysis with optimal computation coefficient $\mathbf{a}_{\text{opt}}$, we have

$$p_e(1 < \eta < 2) \leq \int_{h_B=0}^{h_B=\infty} \frac{1}{4^2} \mathcal{A} Q\left(\sqrt{\frac{2h_B^2 E}{2N_0}}\right) p(h_B) dh_B. \tag{2.26}$$

When $\eta > 2$, we have

$$p_e(2 < \eta) \leq \int_{h_B=0}^{h_B=\infty} \frac{1}{4^2} \mathcal{A} Q\left(\sqrt{\frac{2h_B^2 E}{2N_0}}\right) p(h_B) dh_B. \tag{2.27}$$

Note that for both (2.26) and (2.27) we have $\mathcal{A} = 32$.

Then the average error probability upper bound for $w_R^{\mathbf{a}_{\text{opt}}}$ when $\eta \geq 1$ can be expressed as

$$\begin{aligned} p_e(1 \leq \eta) = p_e(1 < \eta < 2) p(1 < \eta < 2 | 1 \leq \eta) \\ + p_e(2 < \eta) p(2 < \eta | 1 \leq \eta), \end{aligned} \tag{2.28}$$

where

$$p(1 < \eta < 2 | 1 \leq \eta) = \frac{F(2) - F(1)}{1 - F(1)}, \tag{2.29}$$

and

$$p(2 < \eta | 1 \leq \eta) = \frac{1 - F(2)}{1 - F(1)}, \tag{2.30}$$

Note that (2.28) follows from that $p(\eta = 1 | 1 \leq \eta) = 0$ and $p(\eta = 2 | 1 \leq \eta) = 0$.

Now, we present the computation error probability upper bound for all η value:

$$\begin{aligned} p_e &= p_e(1 \le \eta) + p_e(0 < \eta < 1), && (2.31)\\ &= 2p_e(1 \le \eta), && (2.32)\\ &\le 2\left(1 - \sqrt{\frac{\rho}{\rho+2}}\right), && (2.33) \end{aligned}$$

where $\rho = \frac{E}{N_0}$, and (2.32) follows from the fact that $0 < \eta < 1$ is the symmetric case of $\eta \ge 1$. When ρ becomes large (at high SNR region), (2.33) can be approximated to:

$$p_e \le 2\rho^{-1}. \tag{2.34}$$

Hence the diversity order is 1 at high SNR region, when the optimum compute coefficient vector $\mathbf{a}_{\text{opt}}$ is selected.

2.5 Numerical Results

We firstly compare the derived error rate upper bound to the simulation results for the decoding of the network coded messages at the relay when QPSK are considered. The optimal computation coefficient vectors $\mathbf{a}_{\text{opt}}$ in the simulation are obtained via exhaustive search. Fig. 2.5 shows that the simulation results closely match with the analytical results, and the derived upper bound is a tight upper bound.

For comparison, we show the error performance of decoding the network coded message when 4-PAM is considered, where the corresponding diversity order was reported as $\frac{1}{2}$ in [14]. We see that, QPSK outperforms 4-PAM significantly. The diversity order for decoding the network coded message at the relay is 1 with QPSK, doubling the case with 4-PAM. Our scheme is about 15 dB better than the corresponding 4-PAM scheme at the symbol error rate of 10^{-2}.

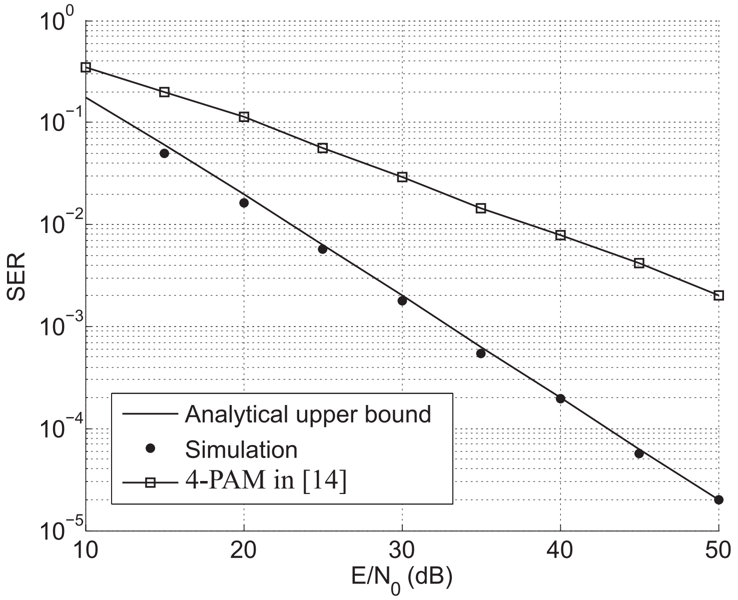


Figure 2.5 The Monte-Carlo simulation results and the analytical upper bound.

2.6 Conclusions

In this work, we designed the optimal network coding function at the relay when QPSK constellation is considered, and we analyzed the corresponding error performance at the relay. We discussed the approach to analyze the distance profile which contributes to the error events of the computer-and-forward at the relay when optimal computation coefficient vector is considered. We derived a tight computation error upper bound at the relay in a closed form. We showed that the diversity order of this scheme is 1 when optimum computation coefficients are selected. This proposed scheme outperforms the corresponding 4-PAM scheme reported in the literature by 15 dB at the symbol error rate of 10^{-2}.

Chapter 3

Analysis of CPNC in TWRCs with Binary Classic Codes

3.1 Introduction

Starting from the chapter, we investigate the channel-coded PNC (CPNC) scheme. In this chapter, we focus on the error performance of CPNC scheme for binary-input Gaussian two-way relay channels, where linear classical codes are adopted. Linear codes are important for achieving network coding function at the relay. This is because that the component-wise modulo-2 sum of two codewords is another codeword. In general, codes can be divided into two categories: block codes and convolutional codes. Block codes have fixed length due to the fact that its coding is operated block by block. Convolutional codes, on the contrary, can have arbitrary lengths. As pointed out in [10] that classic codes are designed with large minimum distance and strong algebraic structure. The error probability performance of classic codes in point-to-point channel relies on their minimum distance and distance distributions [132, 133]. However, unlike the point-to-point channel, the challenge in PNC is that the relay sees superimposed signals which can be transformed to superimposed codewords, and the relay is required to decode a network coded codewords directly from the superimposed signals it received. Thus conventional distance analysis for point-to-point channel no longer applies for PNC.

This chapter begins by presenting the preliminary knowledge of classic codes, where we focus on block codes. The introduced Hamming distance and performance upper bound are important for understanding the later analysis for PNC scheme. After that, we shift our focus on studying the structure of the superimposed codewords in PNC. We then introduce an asymptotically tight performance bound for the error probability. Finally, we demonstrate the derived performance bounds by Monte-Carlo simulations.

The work in this chapter is a collaborated work with other researchers, and I am the sole student in this collaboration. My contribution in this work includes: propose the computation functions for network coded information at the relay and compared their performance difference via simulations; analyze the structure of the superimposed codewords at the relay and formulate the distance properties; derive the union bound in a form of the distance structure; conduct simulation to verify the derived performance bound.

3.2 Preliminary: Linear Block Codes

3.2.1 Encoding of Block Codes

Block codes have fixed length due to the fact that it coding is operated block by block. An important feature of block coding is that the encoded codeword only depends on the current input information sequence of an encoder. In other words, the encoder of block codes is memoryless. Examples of block codes include Hamming codes, BCH codes, and Reed Solomon codes.

Given an (n, k) block code, the length of the input information sequence is k and the length of the output codeword is n, and $n > k$. The code rate is defined as $R = k/n$ which indicates the amount of information per coded digit. From now on, we focus our discussion on binary codes unless otherwise specified. For a binary (n, k) block code, there should be 2^k possible distinct messages and 2^k distinct codewords, since there should be a one-to-one correspondence between a message and a codeword.

A block code is linear if the following two conditions are satisfied:

1. The modulo-2 sum of two codewords is another codewords;
2. All-zero codeword is included in the code.

As a matter of fact, the first condition is important for achieving linear network coding in PNC, which we will see later in this chapter.

In linear algebra, an (n, k) linear block code is a k-dimensional subspace of the vector space V_n of all the binary n-tuples. Given k linearly independent binary n-tuples $\{\mathbf{g}_0, \mathbf{g}_1, \cdots, \mathbf{g}_{k-1}\}$, an (n, k) linear block code can be constructed by:

$$\mathbf{c} = b_0\mathbf{g}_0 + b_1\mathbf{g}_1 + \cdots + b_{k-1}\mathbf{g}_{k-1},$$

where $\mathbf{b} = (b_0, b_1, \cdots, b_{k-1})$ is the message sequence. The k linearly independent n-tuples $\{\mathbf{g}_0, \mathbf{g}_1, \cdots, \mathbf{g}_{k-1}\}$ for an (n, k) code can be organized in a matrix form

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0,n-1} \\ g_{10} & g_{11} & \cdots & g_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k-1,0} & g_{k-1,0} & \cdots & g_{k-1,n-1} \end{bmatrix}.$$

This matrix G is called the *generator matrix* of the block code. Thus, given a message sequence $\mathbf{b}$, the encoding process can be represented as

$$\mathbf{c} = \mathbf{b} \cdot \mathbf{G} = b_0\mathbf{g}_0 + b_1\mathbf{g}_1 + \cdots + b_{k-1}\mathbf{g}_{k-1}.$$

Example 3.1. A commonly used example in the literature is $(7, 4)$ code. One form of its generate matrix is

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

Thus, given a message $\mathbf{b} = (1101)$, the encoding process is as follows:

$$\begin{aligned} \mathbf{c} &= \mathbf{b} \cdot \mathbf{G} \\ &= 1 \cdot (1000101) + 1 \cdot (0100011) + 0 \cdot (0010111) + 1 \cdot (0001110) \\ &= (1101000) \end{aligned}$$

where the operations are modulo-2 addition and multiplication. Table 3.1 shows the complete messages and corresponding codewords for the $(7, 4)$ linear block code in this example, and in total there are 2^4 distinct codewords in this code.

3.2.2 The Minimum Distance of a Block Code

Given a block code, its error detection and correction capability is determined by its code minimum distance [10, 132, 133]. We start with the definition of Hamming distance. Given an (n, k) linear block code $\mathcal{C}$, the *Hamming distance* (or distance) between any two codewords $\mathbf{c}_1$ and $\mathbf{c}_2$, $d(\mathbf{c}_1, \mathbf{c}_2)$, is defined as the number of different places between these two codewords. Take the $(7, 4)$ linear block codes in Table 3.1 as an example. If $\mathbf{c}_1 = (0001110)$ and $\mathbf{c}_2 = (0010111)$, then we have $d(\mathbf{c}_1, \mathbf{c}_2) = 3$.

Another important definition in linear block code is Hamming weight (or weight). The *Hamming weight* of a codeword is defined as the total number of nonzero places in the codeword. For example, the Hamming weight of $\mathbf{c}_1 = (0001110)$ is 3, and the Hamming weight of $\mathbf{c}_2 = (0010111)$ is 4.

Recall that an important property of linear block codes is that the modulo-2 sum of two codewords is still a valid codeword. Now we can see that the Hamming distance

Table 3.1A $(7, 4)$ linear block code

Messages **b**	Codewords **c**	Hamming Weight
0000	0000000	0
0001	0001110	3
0010	0010111	4
0011	0011001	3
0100	0100011	3
0101	0101101	4
0110	0110100	3
0111	0111010	4
1000	1000101	3
1001	1001011	4
1010	1010010	3
1011	1011100	4
1100	1100110	4
1101	1101000	3
1110	1110001	4
1111	1111111	7

between two codewords of a linear block code is the Hamming weight of the modulo-2 sum of these two codewords. For example, the modulo-2 sum of $\mathbf{c}_1 = (0001110)$ and $\mathbf{c}_2 = (0010111)$ in Table 3.1 is (0011001), which is also a valid codeword with Hamming weight 3. The *minimum Hamming distance* (or minimum distance) of a code is defined as the smallest Hamming distance between any two different codewords in the code. Thus, the minimum Hamming distance of a linear block code is the smallest Hamming weight of the nonzero codewords in that code. For example, the minimum Hamming distance of the $(7, 4)$ code in Table 3.1 is 3. Ref. [132, 133] contains more details on how to derive the minimum Hamming distance from the parity-check matrix of a block code.

3.2.3 Performance Upper Bounds of Block Codes

We focus on the word error probability of block codes on AWGN channels here. Ref. [132, 133] contains more details on this topic, as well as the bit error rate upper bounds.

The word error probability of a block code on AWGN channels can be upper-bounded by a union bound which is a sum of all error events with various Hamming distances. To this end, the weight distribution of a block code is important for calculating the performance upper bound. Given an (n, k) block code, its weight distribution can be expressed by the code *weight enumerating function* [136], which is given by

$$W(X) = \sum_{i=0}^{n} W_i X^i,$$

where W_i is the number of codewords that have Hamming weight i and X is a dummy variable. For example, the weight enumerating function of the $(7, 4)$ code in Table 3.1 is

$$W(X) = 1 + 7X^3 + 7X^4 + X^7.$$

The *weight distribution* or the *weight spectrum* of a (n, k) block code with minimum Hamming distance $d_{\min}$ is given by the following set

$$\{W_{d_{\min}}, W_{d_{\min}+1}, \cdots, W_n\}.$$

We now consider the pairwise error probability. The pairwise error probability is to describe the error events when a wrong codewords is selected in the decoding process. Assume BPSK modulation is used and AWGN channel. The pairwise error probability for a binary block code is given by [137]

$$P_d = Q\left(\sqrt{2dR\frac{E_b}{N_0}}\right),$$

where d is the Hamming distance between two codewords, $R = k/n$ is the code rate, E_b is the signal energy per bit, N_0 is the single sided power spectral density of the Gaussian noise, and $Q(\cdot)$ is the complementary error function. Thus, the word error probability of union bound is a summation of all contributions from all error events with all Hamming distances, given by

$$\begin{aligned} P_e &\leq \sum_{d=d_{\min}} W_d P_d \\ &= \sum_{d=d_{\min}} W_d Q\left(\sqrt{2dR\frac{E_b}{N_0}}\right) \end{aligned}$$

where $d_{\min}$ is the minimum Hamming distance of the code, and W_d is the number of error events with Hamming distance d. Usually all-zero codeword is assumed for the ease of analysis, and W_d becomes the number of codewords with weight d in this case, which can be obtained from the weight enumerating function $W(X)$.

3.3 System Model

We consider a binary-input Gaussian two-way relay channel (TWRC) where two single-antenna users, denoted by A and B, exchange information via an intermediate single-antenna relay. The users and the relay operate in half-duplex mode and there is no direct link between the users. The transmission protocol employs two consecutive equal-duration time-slots for each round of information exchange. In the first time-slot (uplink phase), the users transmit simultaneously and the relay remains silent. In the second time-slot (downlink phase), the relay broadcasts to the two silent users. At each node, the received signal is corrupted by AWGN.

3.3.1 Transmitter Architecture of the Two Users

The transmitter architecture of the uplink phase is depicted in Fig. 3.1. Let $\mathbf{b}_A \in \{0,1\}^k$ and $\mathbf{b}_B \in \{0,1\}^k$ denote the length-k binary message sequences of user A and B, respectively. A common $\left(2^k, n\right)$ binary linear code of rate $R = k/n$ is employed

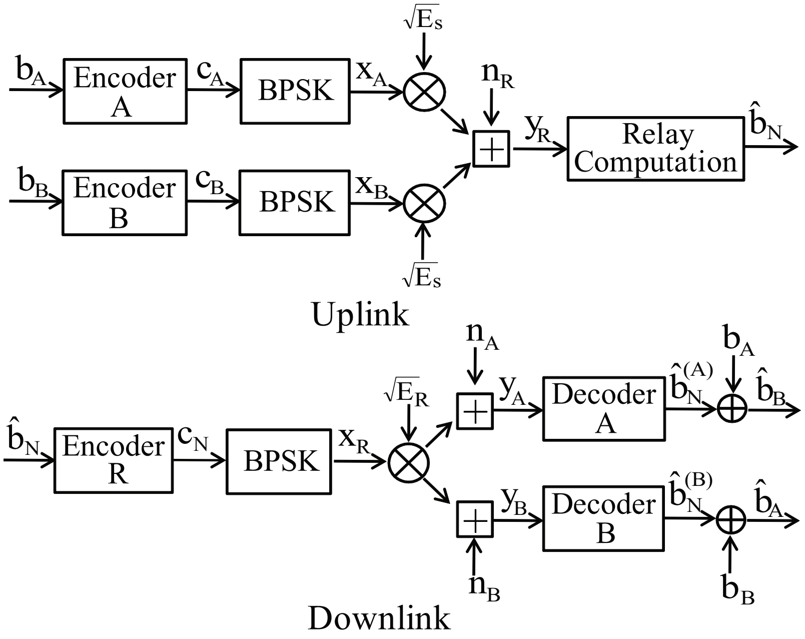


Figure 3.1 Architecture of a two-way relay system operated with channel-coded PNC. The relay computes the network-coded message $\mathbf{b}_N = \mathbf{b}_A \oplus \mathbf{b}_B$ without explicit decoding of both users' individual messages. Here, "+" denotes in the linear addition in real values and "$\oplus$" denotes the modulo-2 addition.

to encode the messages of both users, where $k = nR$. For a binary linear code, this encoding operation can be written as $\mathbf{c}_m = \mathbf{b}_m\mathbf{G}$, $m \in \{A, B\}$, where $\mathbf{G}$ is the generator matrix. We exhibit the 2^{nR} codewords as the rows of a matrix $\mathbf{C} \in \{0,1\}^{2^{nR}\times n}$ which is referred to as the "codebook" matrix. The set of 2^{nR} codewords is denoted by $\mathcal{C}$ which is referred to as the "code". The Hamming weight of a codeword $\mathbf{c} \in \mathcal{C}$ is denoted by $w_H(\mathbf{c})$. The minimum Hamming distance between any two codewords in $\mathcal{C}$ is denoted by $d_{\min}(\mathcal{C})$.

Let $a \oplus b$ denote the modulo-2 addition (XOR) of a and b for a, $b \in \{0,1\}$. For binary sequences $\mathbf{b}_A$ and $\mathbf{b}_B$, the XOR-ed message sequence is denoted by $\mathbf{b}_N \triangleq \mathbf{b}_A \oplus \mathbf{b}_B \in \{0,1\}^k$. We refer to $\mathbf{b}_N$ as a *network-coded (NC) message* [3]. Similarly, for codeword-pair $\mathbf{c}_A$ and $\mathbf{c}_B$, their XOR is $\mathbf{c}_N \triangleq \mathbf{c}_A \oplus \mathbf{c}_B \in \{0,1\}^n$ which is referred to as a *NC codeword.* Due to the linearity of the code, we have $\mathbf{c}_N \in \mathcal{C}$, that is, the NC codeword set is identical to $\mathcal{C}$. There is a one-to-one mapping between a NC codeword $\mathbf{c}_N$ and a NC message $\mathbf{b}_N$, given by $\mathbf{c}_N = \mathbf{b}_N\mathbf{G}$.

Note that in general, the relay can encode the network coded message $\mathbf{b}_N$ using a different code in the downlink transmission rather than the same code used by the users in the uplink. For example, the relay can transmit in a different code rate to suit the downlink channel condition in its transmission. In this work, we assume that the relay is using the same codebook as the user. Thus, the relay combines the two operations “Relay computation” and “Encoder” in Fig. 3.1 together and directly compute the network coded codeword from its received superimposed signal.

In the uplink phase, the codewords of the two users are modulated via BPSK ($0 \mapsto -1$, $1 \mapsto +1$), resulting in coded signal sequences

$$\mathbf{x}_m = 2\mathbf{c}_m - 1 \in \{-1, 1\}^n, m \in \{A, B\}, \tag{3.1}$$

that are simultaneously transmitted. Assuming perfect synchronization, the relay receives

$$\mathbf{y}_R = \sqrt{E_A}\mathbf{x}_A + \sqrt{E_B}\mathbf{x}_B + \mathbf{n}_R, \tag{3.2}$$

where E_m denotes the received symbol energy of user $m, m \in \{A, B\}$, and $\mathbf{n}_R$ is the AWGN sequence. The variance of the noise is $\sigma^2 = \frac{N_0}{2}$ where N_0 is the one-sided noise power spectral density. It is noteworthy that, in general, the two users in a CPNC scheme may have different data rates and different signal power. In this work, however, we will follow the pioneering work [46] by limiting our discussion to the cases where the two users have identical data rates and the same received symbol energies, i.e., $E_A = E_B = E_s$. Then, the signal received by the relay is

$$\mathbf{y}_R = \sqrt{E_A}\mathbf{x}_A + \sqrt{E_B}\mathbf{x}_B + \mathbf{n}_R = \sqrt{E_s}\mathbf{x}_s + \mathbf{n}_R, \tag{3.3}$$

where $\mathbf{x}_s \triangleq \mathbf{x}_A + \mathbf{x}_B \in \{-2, 0, 2\}^n$. We refer to $\mathbf{x}_s$ as a *superimposed* (SI) *codeword*.

Since the code $\mathcal{C}$ is linear, the SI codewords exhibit important features that will be exploited to recover the NC codeword at the relay, as we will see momentarily. To characterize those features, the following definition is required.

Definition 3.1. Given a NC codeword $\mathbf{c}_N$, we define a set $\mathcal{X}_s(\mathbf{c}_N)$ which collects all the distinct SI codewords $\mathbf{x}_s$ that correspond to $\mathbf{c}_N$, given as

$$\begin{aligned}\mathcal{X}_s(\mathbf{c}_N) \triangleq \{\mathbf{x}_s = \mathbf{x}_A + \mathbf{x}_B : \mathbf{x}_A = 2\mathbf{c}_A - 1, \mathbf{x}_B = 2\mathbf{c}_B - 1,\\ \mathbf{c}_A, \mathbf{c}_B \in \mathcal{C},\ \mathbf{c}_A \oplus \mathbf{c}_B = \mathbf{c}_N\}.\end{aligned} \tag{3.4}$$

The union of the sets $\mathcal{X}_s(\mathbf{c}_N)$ for all $\mathbf{c}_N$ is given as

$$\mathcal{X}_s = \bigcup_{\mathbf{c}_N \in \mathcal{C}} \mathcal{X}_s(\mathbf{c}_N) \tag{3.5}$$

which collects all possible sets of SI codewords. Here, (3.4) and (3.5) partition the entire SI codeword space into a number of sets $\mathcal{X}_s(\mathbf{c}_N)$, $\mathbf{c}_N \in \mathcal{C}$, where each set corresponds to a specific NC codeword $\mathbf{c}_N$. The mapping from those SI codewords $\mathbf{x}_N \in \mathcal{X}_s(\mathbf{c}_N)$ to $\mathbf{c}_N$ is multiple-to-one if $|\mathcal{X}_s(\mathbf{c}_N)| > 1$.

3.3.2 Computation Rules at the Relay

Upon receiving $\mathbf{y}_R$, the first task of the relay is to recover the NC codeword $\mathbf{c}_N = \mathbf{c}_A \oplus \mathbf{c}_B$. Since the relay does not decode both individual codewords $\mathbf{c}_A$ and $\mathbf{c}_B$, but only computes their modulo-2 sum $\mathbf{c}_N$, we consider following two computation rules.

Maximum Likelihood Computation

Previously, (3.4) and (3.5) have partitioned the entire SI codeword space into sets $\mathcal{X}_s(\mathbf{c}_N)$, $\mathbf{c}_N \in \mathcal{C}$. Upon receiving $\mathbf{y}_R$, the relay will distinguish these sets, according to their *set likelihood functions* given by

$$p\left(\mathbf{y}_R|\mathbf{c}_N\right) = p\left(\mathbf{y}_R|\mathcal{X}_s(\mathbf{c}_N)\right), \mathbf{c}_N \in \mathcal{C},$$

to determine the NC codeword $\mathbf{c}_N$. This is different from the conventional single-user decoding. The optimal maximum likelihood (ML) computation rule is performed via the following two steps:

Step 1. Calculate the set likelihood functions $p\left(\mathbf{y}_R|\mathcal{X}_s(\mathbf{c}_N)\right)$ for all $\mathbf{c}_N \in \mathcal{C}$.

Step 2. Select the most likely set and determine the estimation of the NC codeword, i.e.,

$$\widehat{\mathbf{c}}_N = \arg\max_{\mathbf{c}_N \in \mathcal{C}} p\left(\mathbf{y}_R | \mathcal{X}_s(\mathbf{c}_N)\right). \tag{3.6}$$

Minimum Distance Computation

As an alternative, we can employ a suboptimal minimum distance (MD) computation via the following two steps:

Step 1. The estimated SI codeword $\widehat{\mathbf{x}}_s$ with minimum squared Euclidean distance to $\mathbf{y}_R$ is found by

$$\widehat{\mathbf{x}}_s = \arg\max_{\mathbf{x}_s \in \mathcal{X}_s} p\left(\mathbf{y}_R | \mathbf{x}_s\right) = \arg\min_{\mathbf{x}_s \in \mathcal{X}_s} \left\|\mathbf{y}_R - \mathbf{x}_s\right\|^2. \tag{3.7}$$

Step 2. The estimation of the NC codeword is determined by finding

$$\widehat{\mathbf{c}}_N : \widehat{\mathbf{x}}_s \in \mathcal{X}_s(\widehat{\mathbf{c}}_N). \tag{3.8}$$

A *computation error* is declared if $\widehat{\mathbf{c}}_N \neq \mathbf{c}_N$.

Remark 3.1. The suboptimal MD computation, also known as "lattice decoding" [20], is of particular importance in a practical CPNC scheme. It is at the moment the only practically feasible computation method. Specifically, [46] estimates the "nearest" SI codeword using iterative believe propagation based on the "superimposed Tanner graph", and [35] estimates the nearest SI codeword via the Viterbi algorithm based on the "super-trellis". In contrast, at the moment, it is infeasible to apply the optimal ML computation in a practical CPNC scheme, due to the difficulty in calculating the set likelihood functions in (3.6).

Remark 3.2. Here, $\mathbf{c}_N$ is directly computed from the physically received signal. The complete decoding of both individual codewords $\mathbf{c}_A$ and $\mathbf{c}_B$ are circumvented and the multiplexing loss can be avoided [27].

In the downlink phase, after the NC codeword is computed, the relay broadcasts the BPSK-modulated NC codeword to the two users. Then, user A (or B) can recover its desired message $\mathbf{b}_B$ (or $\mathbf{b}_A$) with the help of the knowledge on its own message. More details about the downlink phase can be found in [46].

It is noteworthy that the operations of the uplink and downlink phases can be decoupled [46]. Since the decoding operation at each user in the downlink phase is standard, we will only focus on the probability of computation error $\Pr\left(\widehat{\mathbf{c}}_N \neq \mathbf{c}_N\right)$ at the relay. Given that, the analysis of the downlink phase is straightforward.

3.4 Structure of the Superimposed Codewords

In this section, we investigate the structural properties of the superimposed (SI) codewords $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ for $\mathbf{c}_N \in \mathcal{C}$. The results in this section are necessary to prove the distance properties and to derive the error probability of the CPNC scheme in the subsequent sections.

3.4.1 Properties of Superimposed Codewords

We first present some simple properties on the SI codewords in $\mathcal{X}_s(\mathbf{c}_N)$. Let $c_N(t), t = 1, \cdots, n$, be the t-th entry of $\mathbf{c}_N$. Define a support set

$$\mathcal{S}\left(\mathbf{c}_N\right) \triangleq \left\{t \in \{1, 2, \cdots, n\} : c_s(t) = 1\right\},$$

which collects the positions of $\mathbf{c}_N$ whose entry is 1. The complementary set of $\mathcal{S}\left(\mathbf{c}_N\right)$ is denoted by $\mathcal{S}^c\left(\mathbf{c}_N\right)$.

Property 3.1. For any NC codeword $\mathbf{c}_N$ and any SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$, we have

$$x_s(t) = \begin{cases} 0, & t \in \mathcal{S}\left(\mathbf{c}_N\right) \quad (\text{or } t : c_N(t) = 1) \\ 2 \text{ or } -2, & t \in \mathcal{S}^c\left(\mathbf{c}_N\right) \quad (\text{or } t : c_N(t) = 0) \end{cases}. \tag{3.9}$$

Explanation: For $t \in \mathcal{S}\left(\mathbf{c}_N\right)$, we have $c_N(t) = 1$ which means that $c_A(t) \neq c_B(t)$. This leads to $x_A(t) \neq x_B(t)$ and $x_s(t) = x_A(t) + x_B(t) = 0$. For $t \in \mathcal{S}^c\left(\mathbf{c}_N\right)$, we have $c_A(t) = c_B(t)$ and $x_A(t) = x_B(t)$, thus $x_s(t) = x_A(t) + x_B(t) = 2x_A(t)$. Since $x_A(t)$ is either 1 or -1, $x_s(t)$ is either 2 or -2 for $t \in \mathcal{S}^c\left(\mathbf{c}_N\right)$. ■

Property 3.2. Only $t \in \mathcal{S}^c(\mathbf{c}_N)$ are required in distinguishing different SI codewords in a given set $\mathcal{X}_s(\mathbf{c}_N)$.

Explanation: From Property 3.1, for $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$, $x_s(t) = 0$ for $t \in \mathcal{S}(\mathbf{c}_N)$. Thus, there is no difference between any two SI codewords in $\mathcal{X}_s(\mathbf{c}_N)$ for the positions $t \in \mathcal{S}(\mathbf{c}_N)$. In other words, only the positions $t \in \mathcal{S}^c(\mathbf{c}_N)$ are relevant to distinguishing different SI codewords in $\mathcal{X}_s(\mathbf{c}_N)$. ■

Property 3.3. For any $\mathbf{c}_N \neq \mathbf{c}'_N$, we have $\mathcal{X}_s(\mathbf{c}_N) \cap \mathcal{X}_s(\mathbf{c}'_N) = \emptyset$, where $\emptyset$ denotes the empty set.

Explanation: If $\mathbf{c}_N \neq \mathbf{c}'_N$, one can find at least one position, say t^*, such that $c_N(t^*) \neq c'_N(t^*)$. Let $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ and $\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)$, $\mathbf{c}_N \neq \mathbf{c}'_N$. From Property 3.1, we have $x_s(t^*) \neq x'_s(t^*)$. This yields $\mathcal{X}_s(\mathbf{c}_N) \cap \mathcal{X}_s(\mathbf{c}'_N) = \emptyset$. ■

3.4.2 Punctured Codebook

Now, we define a *punctured codebook* which will be repetitively used in this chapter. In the sequel, we use $\mathbf{a}^I$ to denote the entries of $\mathbf{a}$ indexed by I, where $I \subset \{1, \cdots, n\}$ and n is the length of the vector $\mathbf{a}$. Similarly, we use $\mathbf{A}^I$ to denote the matrix which consists of the columns of $\mathbf{A}$ that are indexed by I.

Definition 3.2 (Punctured Codebook)**.** A *punctured generator matrix* $\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}$ is obtained by removing the columns indexed by $t \in \mathcal{S}(\mathbf{c}_N)$, and keeping those indexed by $t \in \mathcal{S}^c(\mathbf{c}_N)$, from the original generator matrix $\mathbf{G}$. Similarly, a *punctured codebook* $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ is obtained by removing all the columns indexed by $t \in \mathcal{S}(\mathbf{c}_N)$, and keeping those indexed by $t \in \mathcal{S}^c(\mathbf{c}_N)$, from the original codebook $\mathbf{C}$.

Example 3.2. Consider a $(7,4)$ Hamming code with

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}. \tag{3.10}$$

Let $\mathbf{c}'_N$ be a certain codeword in $\mathcal{C}$, e.g., $\mathbf{c}'_N = \mathbf{G}\cdot [0\ 0\ 1\ 1] = [0\ 0\ 1\ 1\ 0\ 1\ 0]$. Then, $\mathcal{S}\left(\mathbf{c}'_N\right)$=$\{3, 4, 6\}$ and

$$\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{3.11}$$

obtained by deleting Column 3, 4 and 6 of $\mathbf{G}$. Then, the punctured codebook is $\mathbf{C}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)} = \mathbf{B}\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}$.

Property 3.4. The linearity remains in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$: the XOR of any two rows of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ is a row in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$.

3.4.3 Cardinality of $\mathcal{X}_s(\mathbf{c}_s)$

Now we show that the cardinalities of $\mathcal{X}_s(\mathbf{c}_N)$, $\mathbf{c}_N \in \mathcal{C}$, can be determined by using the punctured codebook. Let $r\left(\mathbf{A}\right)$ denote the number of *distinct rows* of a matrix $\mathbf{A}$ and let Rank($\mathbf{A}$) denote the *rank* of $\mathbf{A}$. Then, we have the following results.

Proposition 3.1. *Each distinct row of* $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ *yields a distinct SI codeword in* $\mathcal{X}_s(\mathbf{c}_N)$. *Identical rows in* $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ *corresponds to the same SI codeword in* $\mathcal{X}_s(\mathbf{c}_N)$.

Proof. From Property 3.1, $x_s\left(t\right) = 0$ for $t \in \mathcal{S}\left(\mathbf{c}_N\right)$. Therefore, it suffices to consider only the positions $t \in \mathcal{S}^c\left(\mathbf{c}_N\right)$, which are completely reflected in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$. By definition, we have $\mathbf{c}_N^{\mathcal{S}^c(\mathbf{c}_N)} = \mathbf{0}^{n-w_H(\mathbf{c}_N)}$ which is equivalent to $\mathbf{c}_A^{\mathcal{S}^c(\mathbf{c}_N)} = \mathbf{c}_B^{\mathcal{S}^c(\mathbf{c}_N)}$. Therefore, for $t \in \mathcal{S}^c\left(\mathbf{c}_N\right)$, we have

$$\mathbf{x}_s^{\mathcal{S}^c(\mathbf{c}_N)} = \mathbf{x}_A^{\mathcal{S}^c(\mathbf{c}_N)} + \mathbf{x}_B^{\mathcal{S}^c(\mathbf{c}_N)} = 4\mathbf{c}_A^{\mathcal{S}^c(\mathbf{c}_N)} - 2. \tag{3.12}$$

Since $\mathbf{c}_A^{\mathcal{S}^c(\mathbf{c}_N)} \in \mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$, each distinct row of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ will give a distinct $\mathbf{c}_A^{\mathcal{S}^c(\mathbf{c}_N)}$ and a distinct $\mathbf{x}_s^{\mathcal{S}^c(\mathbf{c}_N)}$ from (3.12). Moreover, identical rows of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ will map to the same $\mathbf{x}_s^{\mathcal{S}^c(\mathbf{c}_N)}$. Since the entries of $\mathbf{x}_s^{\mathcal{S}(\mathbf{c}_N)}$ are all-zero, they will map to the same $\mathbf{x}_s$. ■

Corollary 3.1. *The cardinality of $\mathcal{X}_s(\mathbf{c}_N)$ is given by*

$$|\mathcal{X}_s(\mathbf{c}_N)| = r\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right). \tag{3.13}$$

Proof. This follows from Proposition 3.1 and the definition of $r(\cdot)$. ■

Proposition 3.2. *Let $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right)$ denote the number of all-zero rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$. Then, we have*

$$r\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) = \frac{2^{nR}}{r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right)}. \tag{3.14}$$

Proof. Let the ith row of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ be denoted by $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$. Consider that $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$ is not an all-zero row. Then, besides $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$ itself, there exist other $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) - 1$ rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ whose XOR with $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$ are all-zero, since there are $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right)$ all-zero rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$. This means that there are $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) - 1$ rows identical to $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$, due to the linearity of the punctured codebook. Next, consider a row of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$, say $\mathbf{C}_{i'}^{\mathcal{S}^c(\mathbf{c}_N)}$, that is different from $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)}$, i.e., $\mathbf{C}_i^{\mathcal{S}^c(\mathbf{c}_N)} \neq \mathbf{C}_{i'}^{\mathcal{S}^c(\mathbf{c}_N)}$. Then, besides $\mathbf{C}_{i'}^{\mathcal{S}^c(\mathbf{c}_N)}$ itself, there exist other $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) - 1$ rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ which are identical to $\mathbf{C}_{i'}^{\mathcal{S}^c(\mathbf{c}_N)}$. Thus, for every distinct non-zero row of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$, there exists $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) - 1$ other rows that are identical to it. Since there are 2^{nR} rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ in total, the number of distinct rows is given by (3.14). ■

Corollary 3.2. *Assume that the codeword for each user is picked uniformly among all possible codewords. Then, for any $\mathbf{c}_N \in \mathcal{C}$, all SI codewords of $\mathcal{X}_s(\mathbf{c}_N)$ have the same probability, i.e.,*

$$p\left(\mathbf{x}_s|\mathcal{X}_s(\mathbf{c}_N)\right) = \frac{1}{|\mathcal{X}_s(\mathbf{c}_N)|}, \forall\ \mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N). \tag{3.15}$$

This follows from Proposition 3.2 and its proof.

Proposition 3.3. *The number of all-zero rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}$ is*

$$r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right) = 2^{nR-Rank\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)}. \tag{3.16}$$

Proof. To find the number of all-zero rows $r_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)}\right)$, one only needs to enumerate all the binary message vectors $\mathbf{b} \in \{0,1\}^{nR}$ satisfying

$$\mathbf{b}\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)} = \mathbf{0}^{n-w_H(\mathbf{c}_N)}. \tag{3.17}$$

Note that (3.17) can be viewed as $\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)$ independent binary equations with nR unknown variables. The number of distinct solutions satisfying (3.17) is exactly $2^{nR-\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)}$. ■

Theorem 3.1. *The cardinality of the set $\mathcal{X}_s(\mathbf{c}_N)$ is given by*

$$|\mathcal{X}_s(\mathbf{c}_N)| = 2^{Rank\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)} \tag{3.18}$$

Proof. This follows from Corollary 3.1 and Propositions 3.2 and 3.3, equations (3.13) , (3.14) and (3.16). ■

From Theorem 3.1, we have the following observation: For any $\mathbf{c}_N \in \mathcal{C}$, there are 2^{nR} codeword-pairs $(\mathbf{c}_A, \mathbf{c}_B)$ with $\mathbf{c}_A \oplus \mathbf{c}_B = \mathbf{c}_N$. In the domain of SI codewords, however, the number of distinct elements is $|\mathcal{X}_s(\mathbf{c}_N)| = 2^{\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)}$, which can be less than 2^{nR}. Moreover, from Corollary 3.2, the SI codewords of $\mathcal{X}_s(\mathbf{c}_N)$ have equal probabilities $p\left(\mathbf{x}_s|\mathcal{X}_s(\mathbf{c}_N)\right) = 1/\left|\mathcal{X}_s(\mathbf{c}_N)\right|$. Thus, there are

$$2^{nR-\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)} \tag{3.19}$$

different codeword-pairs $(\mathbf{c}_A, \mathbf{c}_B)$, whose NC codewords are equal to $\mathbf{c}_N$, overlap to the same SI codeword $\mathbf{x}_s$. Notice that this overlapping does not cause ambiguity in computing the NC codeword $\mathbf{c}_N$, since the different codeword-pairs $(\mathbf{c}_A, \mathbf{c}_B)$ have the same $\mathbf{c}_N$. This is similar to the case of uncoded PNC in the pioneering work [1].

3.4.4 Overlapping Factor

To characterize the "overlapping" described above, we will use the following definition.

Definition 3.3 (Overlapping Factor)**.** The overlapping factor w.r.t. any SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$, $\mathbf{c}_N \in \mathcal{C}$, is defined as

$$O\left(\mathbf{x}_s\right) \triangleq \left|\left\{\mathbf{c}_A, \mathbf{c}_B \in \mathcal{C} : \mathbf{c}_A \oplus \mathbf{c}_B = \mathbf{c}_N, \mathbf{x}_A + \mathbf{x}_B = \mathbf{x}_s\right\}\right|.$$

Remark 3.3. Due to Corollary 3.2, the overlapping factor is the same for all SI codewords $\mathbf{x}_s$ in the set $\mathcal{X}_s(\mathbf{c}_N)$. Thus, we may write $O\left(\mathbf{c}_N\right)$ instead of $O\left(\mathbf{x}_s\right)$. The overlapping factor $O\left(\mathbf{c}_N\right)$ is generally different for various $\mathbf{c}_N \in \mathcal{C}$.

Proposition 3.4. *For $\mathbf{c}_N \in \mathcal{C}$, from (3.19), we have*

$$O\left(\mathbf{c}_N\right) = 2^{nR - Rank\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right)}. \tag{3.20}$$

Example 3.3. For $\mathbf{c}'_N = [0, \ldots, 0]$, we have $\mathcal{S}\left(\mathbf{c}_N\right) = \emptyset$, thus $\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)} = \mathbf{G}$ and $\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right) = nR$. This results in $O\left(\mathbf{c}'_N\right) = 1$ from (3.20) which means that every codeword-pair $\left(\mathbf{c}'_A, \mathbf{c}'_B\right) : \mathbf{c}'_A \oplus \mathbf{c}'_B = \mathbf{c}'_N$ maps to a distinct SI codeword. For $\mathbf{c}'_N = [1, \ldots, 1]$, we have $\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)}\right) = 0$. This results in $O\left(\mathbf{c}'_N\right) = 2^{nR}$, which means that all codeword-pairs $\left(\mathbf{c}'_A, \mathbf{c}'_B\right) : \mathbf{c}'_A \oplus \mathbf{c}'_B = \mathbf{c}'_N$ map to the same SI codeword.

The overlapping effect described above distinguishes the CPNC scheme from the conventional complete DF-based scheme. In the complete DF-based scheme, the codeword-pair $\left(\mathbf{c}_A, \mathbf{c}_B\right)$ must be completely decoded. If different codeword-pairs are overlapped, there will be an ambiguity and the decoding of both individual codewords will fail. Therefore, multiple-access techniques such as code-division multiple-access (CDMA), or interleave-division multiple-access (IDMA) [134,135], is employed which can avoid the above ambiguity. In the CPNC scheme, however, we only need to recover the NC codeword $\mathbf{c}_N$ rather than the complete $\left(\mathbf{c}_A, \mathbf{c}_B\right)$. Thus, the overlapping effect will not affect the recovery of the NC codeword $\mathbf{c}_N$. Meanwhile, the introduction of the overlapping of the SI codewords may lead to a more efficient usage of the entire codeword space, e.g., an improved minimum distance (as we will see in the next section), in a similar fashion as in the channel un-coded PNC case [1]. The improved minimum distance of the CPNC scheme may give rise to an improved high-SNR error performance relative to the complete DF based scheme.

3.5 Distance Spectrum of CPNC

In this section, we proceed to find the distance spectrum of the CPNC scheme based on the results in the previous section. The results in this section hold for CPNC schemes with general binary linear channel codes.

3.5.1 Formulation of Pairwise Distance Spectrum

Assume that the genuine transmitted signal sequences of the two users are $\mathbf{x}_A$ and $\mathbf{x}_B$, and that their SI codeword is $\mathbf{x}_s = \mathbf{x}_A + \mathbf{x}_B \in \mathcal{X}_s(\mathbf{c}_N)$, where $\mathbf{c}_N$ is the genuine NC codeword. For an erroneous NC codeword $\mathbf{c}'_N \neq \mathbf{c}_N$, let $\mathcal{X}_s(\mathbf{c}'_N)$ be referred to as a *competing set*. Since there are $|\mathcal{X}_s(\mathbf{c}'_N)|$ SI codewords in this competing set, there will be $|\mathcal{X}_s(\mathbf{c}'_N)|$ *error events* of confusing the genuine NC codeword $\mathbf{c}_N$ with $\mathbf{c}'_N$.

We now investigate the set of squared Euclidean distances (SEDs) between the genuine SI codeword $\mathbf{x}_s$ and those in the competing set, i.e., $\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)$. To do this, we partition $\mathcal{X}_s(\mathbf{c}'_N)$ into *subsets*

$$\mathcal{X}_s^d\left(\mathbf{x}_s, \mathbf{c}'_N\right) \triangleq \{\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N) \ : \ \left\|\mathbf{x}'_s - \mathbf{x}_s\right\|^2 = d^2\}, \tag{3.21}$$

where $\min\limits_{\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)} \left\|\mathbf{x}'_s - \mathbf{x}_s\right\|^2 \leq d^2 \leq \max\limits_{\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)} \left\|\mathbf{x}'_s - \mathbf{x}_s\right\|^2$, such that all SI codewords $\mathbf{x}'_s$ in a subset $\mathcal{X}_s^d\left(\mathbf{x}_s, \mathbf{c}'_N\right)$ have the same SED d^2 to $\mathbf{x}_s$. We have the following result on the SEDs.

Lemma 3.1. *All possible SEDs between the genuine transmitted SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ and any SI codeword in a given competing set $\mathcal{X}_s(\mathbf{c}'_N)$, denoted by $\left\|\mathbf{x}'_s - \mathbf{x}_s\right\|^2$, $\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)$, are given by*

$$d_i^2 = 4E_s d_H\left(\mathbf{c}_N, \mathbf{c}'_N\right) + i \cdot 16E_s, \tag{3.22}$$

where i is an integer of $0 \leq i \leq \left|\mathcal{S}^c\left(\mathbf{c}_N\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)\right|$ and $d_H\left(\mathbf{c}_N, \mathbf{c}'_N\right)$ is the Hamming distance between $\mathbf{c}_N$ and $\mathbf{c}'_N$. In particular, we have

$$d_0^2 = \min_{\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)} \left\|\mathbf{x}'_s - \mathbf{x}_s\right\|^2 = 4E_s d_H\left(\mathbf{c}_N, \mathbf{c}'_N\right). \tag{3.23}$$

Proof of Lemma 3.1. For $t \in \mathcal{S}^c(\mathbf{c}_N) \cap \mathcal{S}(\mathbf{c}'_N)$ and $t \in \mathcal{S}(\mathbf{c}_N) \cap \mathcal{S}^c(\mathbf{c}'_N)$, we have $c_N(t) \neq c'_N(t)$ and $x_s(t) \neq x'_s(t)$, Thus, $|x_s(t) - x'_s(t)|^2 = 4$. Since $|\mathcal{S}^c(\mathbf{c}_N) \cap \mathcal{S}(\mathbf{c}'_N)| + |\mathcal{S}(\mathbf{c}_N) \cap \mathcal{S}^c(\mathbf{c}'_N)| = d_H(\mathbf{c}_N, \mathbf{c}'_N)$, the total contribution from these positions to the squared Euclidean distance is $4E_s d_H(\mathbf{c}_N, \mathbf{c}'_N)$. Note that the contribution from $d_H(\mathbf{c}_N, \mathbf{c}'_N)$ to $\|\mathbf{x}_s - \mathbf{x}'_s\|^2$ is exclusively in this case. For positions $t \in \mathcal{S}(\mathbf{c}_N) \cap \mathcal{S}(\mathbf{c}'_N)$, we have $c_N(t) = c'_N(t) = 1$ and thus $x_s(t) = x'_s(t) = 0$. The contribution from these positions to the squared Euclidean distance is zero. For positions $t \in \mathcal{S}^c(\mathbf{c}_N) \cap \mathcal{S}^c(\mathbf{c}'_N)$ and $t : x_s(t) = x'_s(t)$, the contribution to the squared Euclidean distance is zero. Finally, for position t such that $c_N(t) = c'_N(t) = 0$ and $x_s(t) = -x'_s(t) = \pm 2$, we have $|x_s(t) - x'_s(t)|^2 = 16$. Let i be the number of such positions, then,

$$\|\mathbf{x}_s - \mathbf{x}'_s\|^2 = 4E_s d_H(\mathbf{c}_N, \mathbf{c}'_N) + i \cdot 16E_s. \tag{3.24}$$

The parameter i is upper-bounded by the maximum number of such positions,

$$|\mathcal{S}^c(\mathbf{c}_N) \cap \mathcal{S}^c(\mathbf{c}'_N)|.$$

■

Using (3.22), the subsets of the competing set $\mathcal{X}_s(\mathbf{c}'_N)$ w.r.t. $\mathbf{x}_s$ can also be written as

$$\mathcal{X}_s^{d_i}(\mathbf{x}_s, \mathbf{c}'_N) \triangleq \left\{\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N) \ : \ \|\mathbf{x}_s - \mathbf{x}'_s\|^2 = d_i^2\right\} \tag{3.25}$$

where $i = 0, \cdots, N \triangleq |\mathcal{S}^c(\mathbf{c}_N) \cap \mathcal{S}^c(\mathbf{c}'_N)|$. We refer to the subset $\mathcal{X}_s^{d_0}(\mathbf{x}_s, \mathbf{c}'_N)$, whose elements have the minimum SED to $\mathbf{x}_s$, as the *minimum distance subset* w.r.t. $\mathbf{x}_s$ and $\mathcal{X}_s(\mathbf{c}'_N)$. The corresponding error events are referred to as the *minimum distance error events* w.r.t. $\mathbf{x}_s$ and $\mathcal{X}_s(\mathbf{c}'_N)$.

Remark 3.4. In conventional single-user point-to-point transmission over AWGN channel, the distance between the BPSK-modulated codewords $\mathbf{x}_s$=$2\mathbf{c}_N$$-1$ and $\mathbf{x}'_s$=$2\mathbf{c}'_N$$-1$ is determined by the Hamming distance $d_H(\mathbf{c}_N, \mathbf{c}'_N)$ between the binary codewords. In the CPNC scheme for a TWRC, for NC codewords $\mathbf{c}_N$ and $\mathbf{c}'_N$, the effective distances between the genuine SI codeword $\mathbf{x}_s$ and $\mathbf{x}'_s$, $\mathbf{x}'_s \in \mathcal{X}_s(\mathbf{c}'_N)$, are determined by a set of Euclidean distances as shown in (3.22). In particular, the set of Euclidean distances is not solely determined by the Hamming distance $d_H\left(\mathbf{c}_N, \mathbf{c}'_N\right)$ and can not be described with a single value.

To characterize the set of distances between the genuine SI codeword $\mathbf{x}_s$ and those in the competing set $\mathcal{X}_s(\mathbf{c}'_N)$, we will use the following definition.

Definition 3.4 (Pair-wise Distance Spectrum)**.** Denote by

$$\mathcal{J}^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right) \triangleq \left|\mathcal{X}_s^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right)\right|, i = 0, 1, \cdots, N.$$

The cardinalities of the subsets $\mathcal{X}_s^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right), i = 0, 1, \cdots, N$, are collected as

$$\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right) \triangleq \left[\mathcal{J}^{d_0}\left(\mathbf{x}_s, \mathbf{c}'_N\right), \cdots, \mathcal{J}^{d_N}\left(\mathbf{x}_s, \mathbf{c}'_N\right)\right]. \tag{3.26}$$

We refer to $\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$ as the *pair-wise distance spectrum (PDS)* between the genuine transmitted SI codeword $\mathbf{x}_s$ and the SI codewords in the competing set $\mathcal{X}_s(\mathbf{c}'_N)$.

Here, the PDS $\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$ specifies the SED spectrum w.r.t. all error events of sending $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ but the receiver recovers $\mathbf{c}'_N \neq \mathbf{c}_N$.

3.5.2 Determination of the Pair-Wise Distance Spectrum

Now, we show that our previously proposed punctured codebook method can be utilized to determine the PDS. To this end, the following corollary is needed.

Corollary 3.3. *For $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$, the PDS $\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$ can be completely determined by evaluating the positions $t \in \mathcal{S}^c\left(\mathbf{c}_N\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)$ only.*

Corollary 3.3 follows from Lemma 3.1 (and its proof). It suggests that we can focus on the positions $t \in \mathcal{S}^c\left(\mathbf{c}_N\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)$ to find the PDS. Let $\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ be the punctured generator matrix determined by deleting the columns of $\mathbf{G}$ indexed by $t \in \mathcal{S}\left(\mathbf{c}_N\right) \cup \mathcal{S}\left(\mathbf{c}'_N\right)$ and keeping those indexed by $t \in \mathcal{S}^c\left(\mathbf{c}_N\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)$. Let $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ be the associated punctured codebook. Define

$$\begin{aligned}&\mathbb{A}\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)\\ &\triangleq \left[A_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right), \cdots, A_N\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)\right]\end{aligned} \tag{3.27}$$

where $A_i\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)$ denotes the number of rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ with Hamming weight i. The result on the PDS is given in the following theorem.

Theorem 3.2. *The PDS w.r.t. the genuine SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ and $\mathcal{X}_s(\mathbf{c}'_N)$ is given by*

$$\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right) = \frac{\mathbb{A}\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}{O\left(\mathbf{c}'_N\right)}. \tag{3.28}$$

where $O\left(\mathbf{c}'_N\right)$ is the overlapping factor given in (3.20).

Proof. For $t \in \mathcal{S}^c\left(\mathbf{c}_N\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)$, we have $c_N(t) = c'_N(t) = 0$. Thus $\mathbf{c}_A^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_s)} = \mathbf{c}_B^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)}$ and $\left(\mathbf{c}'_A\right)^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} = \left(\mathbf{c}'_B\right)^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)}$. From this, we obtain

$$\mathbf{x}_s^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} = 4\mathbf{c}_A^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} - 2, \tag{3.29}$$

which is certain given the genuine SI codeword $\mathbf{x}_s$. In addition,

$$\left(\mathbf{x}_N^{'}\right)^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} = 4\left(\mathbf{c}'_A\right)^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} - 2 \tag{3.30}$$

which has $|\mathcal{X}_N(\mathbf{c}'_N)|$ events. Since $\left(\mathbf{c}'_A\right)^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)} \in \mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$, there are $A_i\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)$ rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ which differs from $\mathbf{c}_A^{\mathcal{S}^c\left(\mathbf{c}'_N\right)\cap\mathcal{S}^c(\mathbf{c}_N)}$ in i positions, where we have used the linearity of the punctured codebook. Notice that these $A_i\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)$ rows are mapped to $A_i\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)/O\left(\mathbf{c}'_N\right)$ distinct SI codewords in $\mathcal{X}_N(\mathbf{c}'_N)$, due to the overlapping effect. Thus, we have

$$\mathcal{J}^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right) \triangleq \left|\mathcal{X}_s^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right)\right| = A_i\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)/O\left(\mathbf{c}'_N\right),$$

where the division of $O\left(\mathbf{c}'_N\right)$ is due to the overlapping effect. From (3.26), (3.28) is obtained. ■

Corollary 3.4. *Theorem 3.2 suggests that the PDS between the SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ and those in $\mathcal{X}_s(\mathbf{c}'_N)$ depends only on $\mathbf{c}_N$ and $\mathbf{c}'_N$, not $\mathbf{x}_s$. Thus, we may write $\mathcal{J}\left(\mathbf{c}_N, \mathbf{c}'_N\right)$ instead of $\mathcal{J}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$ for all $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$. Also, we may write $\mathcal{J}^{d_i}\left(\mathbf{c}_N, \mathbf{c}'_N\right)$ instead of $\mathcal{J}^{d_i}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$, $i = 0, \cdots, N$, for the terms in the PDS.*

Example 3.4. We again use the $(7,4)$ Hamming code to show how to utilize the punctured codebook to determine the PDS. Assume that the genuine SI codeword is $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$, where $\mathbf{c}_N = \mathbf{G}\cdot [1\ 0\ 0\ 0] = [1\ 0\ 0\ 0\ 1\ 1\ 0] \in \mathcal{C}$. Let $\mathcal{X}_s(\mathbf{c}'_N)$, $\mathbf{c}'_N = \mathbf{G}\cdot [0\ 0\ 1\ 1] = [0\ 0\ 1\ 1\ 0\ 1\ 0]$, be the competing set (as in Example 3.2. Then, $\mathcal{S}\left(\mathbf{c}_N\right) \cup \mathcal{S}\left(\mathbf{c}'_N\right) = \{1, 3, 4, 5, 6\}$ and we obtain

$$\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}^T \tag{3.31}$$

by removing Column 1, 3, 4, 5, 6 from $\mathbf{G}$. Now, one can easily find $\mathbb{A}\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)$ by examining the rows of $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)} = \mathbf{B}\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$. For this example, $\mathbb{A}\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) = [4, 8, 4]$. Finally, given $O\left(\mathbf{c}'_N\right) = 2$ (see Example 3.4), the PDS is given by

$$\mathcal{J}\left(\mathbf{c}_N, \mathbf{c}'_N\right) = \frac{\mathbb{A}\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}{O\left(\mathbf{c}'_N\right)} = [2, 4, 2]. \tag{3.32}$$

The above procedure can be used to find the PDS $\mathcal{J}\left(\mathbf{c}_N, \mathbf{c}'_N\right)$ for any pair of codewords in $\mathcal{C}$.

We next show some results on the minimum distance subset, which will be useful in the performance analysis in the next section.

Corollary 3.5. *The cardinality of the minimum distance subset* $\mathcal{X}_s^{d_0}\left(\mathbf{x}_s, \mathbf{c}'_N\right)$*, i.e., the first term of the PDS* $\mathcal{J}\left(\mathbf{c}_N, \mathbf{c}'_N\right)$*, is given by*

$$\mathcal{J}^{d_0}\left(\mathbf{c}_N, \mathbf{c}'_N\right) = 2^{Rank\left(\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) - Rank\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}. \tag{3.33}$$

Proof: From Proposition 3.3, the number of all-zero rows in $\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ is $A_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) = 2^{nR - \mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}$. Then, from (3.28), we obtain

$$\begin{aligned} \mathcal{J}^{d_0}\left(\mathbf{c}_N, \mathbf{c}'_N\right) &= \frac{A_0\left(\mathbf{C}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}{O\left(\mathbf{c}'_N\right)} \\ &= 2^{\mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) - \mathrm{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right)}. \end{aligned} \tag{3.34}$$

Lemma 3.2. *Let* $d_{10}\left(\mathbf{c}_N,\mathbf{c}'_N\right) \triangleq \left|\mathcal{S}\left(\mathbf{c}_N\right)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)\right|$. *We have the following upper bound on* $\mathcal{J}^{d_0}\left(\mathbf{c}_N,\mathbf{c}'_N\right)$

$$\mathcal{J}^{d_0}\left(\mathbf{c}_N,\mathbf{c}'_N\right) \le 2^{d_{10}\left(\mathbf{c}_N,\mathbf{c}'_N\right)}. \tag{3.35}$$

Proof. First of all, note that $\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ can be obtained by puncturing

$$\left|\mathcal{S}\left(\mathbf{c}_N\right)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)\right|$$

columns from $\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}$. Thus, the reduction in the rank is at most $\left|\mathcal{S}\left(\mathbf{c}_N\right)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)\right|$, that is,

$$\text{Rank}\left(\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) - \text{Rank}\left(\mathbf{G}^{\mathcal{S}^c(\mathbf{c}_N)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)}\right) \le \left|\mathcal{S}\left(\mathbf{c}_N\right)\cap\mathcal{S}^c\left(\mathbf{c}'_N\right)\right| \tag{3.36}$$

where the equality is satisfied if $\mathbf{G}^{\mathcal{S}^c\left(\mathbf{c}'_N\right)}$ has full column rank. Using (3.36) in (3.34), we obtain (3.35). ■

3.5.3 Overall Distance Spectrum

So far, we have obtained the distance spectrum w.r.t. the genuine NC codeword $\mathbf{c}_N$ and a specific competing set $\mathcal{X}_s(\mathbf{c}'_N)$. Based on that, the overall distance spectrum w.r.t. $\mathbf{c}_N$ and all competing sets $\mathcal{X}_s(\mathbf{c}'_N), \mathbf{c}'_N \in \mathcal{C}, \mathbf{c}'_N \ne \mathbf{c}_N$, can be straightforwardly found. In particular, from Lemma 3.1, the minimum SED between the genuine transmitted SI codeword $\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)$ and any erroneous SI codeword is given by

$$\min_{\substack{\mathbf{x}'_s\in\mathcal{X}_s(\mathbf{c}'_N),\\ \mathbf{c}'_N\in\mathcal{C},\mathbf{c}'_N\ne\mathbf{c}_N}} \left\|\mathbf{x}_s-\mathbf{x}'_s\right\|^2 = 4E_s d_{\min}\left(\mathcal{C}\right),$$

where $d_{\min}\left(\mathcal{C}\right)$ is the minimum Hamming distance of the underlying channel code $\mathcal{C}$. This suggests that the minimum Euclidean distance of the CPNC scheme is the same as that of the conventional single-user case, which extends the minimum distance property of a channel un-coded PNC scheme [1] to channel-coded cases. Yet, relative to the single-user case, there is an extra multiplicity, i.e., $\mathcal{J}^{d_0}\left(\mathbf{c}_N,\mathbf{c}'_N\right)$, in counting the minimum distance error events. This effect may lead to an SNR penalty relative to the single-user scenario, as we will see in the next section.

3.6 Performance Analysis at the Relay

In previous sections, we have shown the pairwise distance spectrum of a CPNC scheme with a general binary linear code. Now, we are ready to analyze the computation error probability at the relay.

3.6.1 Pair-wise Error Probability

Consider the genuine SI codeword is $\mathbf{x}_s^* = \mathbf{x}_A^* + \mathbf{x}_B^* \in \mathcal{X}_s(\mathbf{c}_N^*)$. With the optimal ML computation, the pair-wise error probability (PEP) that the relay recovers $\overline{\mathbf{c}}_N = \mathbf{c}'_N$, $\mathbf{c}'_N \neq \mathbf{c}_N^*$, is given by

$$\begin{aligned}
&P_e^{\mathrm{ML}}\left(\mathbf{x}_s^* \in \mathcal{X}_s(\mathbf{c}_N^*) \to \mathbf{c}'_N\right) = \Pr\left[p\left(\mathbf{y}_R|\mathbf{c}_N^*\right) \leq p\left(\mathbf{y}_R|\mathbf{c}_N^{'}\right)\right] \\
&\overset{(a)}{=} \Pr\left[p\left(\mathbf{c}_N^*|\mathbf{y}_R\right) \leq p\left(\mathbf{c}_N^{'}|\mathbf{y}_R\right)\right] \\
&\overset{(b)}{=} \Pr\left[\sum_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^*)} p\left(\mathbf{x}_s|\mathbf{y}_R\right) \leq \sum_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^{'})} p\left(\mathbf{x}_s|\mathbf{y}_R\right)\right] \\
&\overset{(c)}{=} \Pr\left[\sum_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^*)} \frac{p\left(\mathbf{y}_R|\mathbf{x}_s\right)p\left(\mathbf{x}_s\right)}{p\left(\mathbf{y}_R\right)} \leq \sum_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^{'})} \frac{p\left(\mathbf{y}_R|\mathbf{x}_s\right)p\left(\mathbf{x}_s\right)}{p\left(\mathbf{y}_R\right)}\right] \\
&\overset{(d)}{=} \Pr\left[\frac{\sum\limits_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^*)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)}{|\mathcal{X}_s(\mathbf{c}_N^*)|} \leq \frac{\sum\limits_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}'_N)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)}{|\mathcal{X}_s(\mathbf{c}'_N)|}\right].
\end{aligned} \tag{3.37}$$

In the above "(a)" uses the fact that $p(\mathbf{c}_N) = \frac{1}{2^{nR}}$ for all $\mathbf{c}_N \in \mathcal{C}$; "$(b)$" is from the "total probability" rule; "(c)" follows from the Bayes' rule;"(d)" is from Corollary 3.2.

With the suboptimal MD computation rule, the PEP is given by

$$\begin{aligned}
&P_e^{\mathrm{MD}}\left(\mathbf{x}_s^* \in \mathcal{X}_s(\mathbf{c}_N^*) \to \mathbf{c}'_N\right) \\
&= \Pr\left[\max_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^*)} p\left(\mathbf{y}_R|\mathbf{x}_s\right) \leq \max_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N^{'})} p\left(\mathbf{y}_R|\mathbf{x}_s\right)\right].
\end{aligned} \tag{3.38}$$

Comparing the above two PEPs, we see that the sub-optimal MD computation rule uses the likelihood function of the nearest SI codeword, given by $\max\limits_{\mathbf{x}_s \in \mathcal{X}_s(\mathbf{c}_N)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)$

in (3.38), to approximate the average likelihood function $\frac{1}{|\mathcal{X}_s(\mathbf{c}_N)|}\sum\limits_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)$ in (3.37). Then we have

$$P_e^{\mathrm{ML}}\left(\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)\to\mathbf{c}_N'\right)\leq P_e^{MD}\left(\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)\to\mathbf{c}_N'\right), \tag{3.39}$$

since the PEP with the sub-optimal MD computation can not be smaller than that with the optimal ML computation. The PEP with the MD computation yields an upper bound on the PEP of the CPNC scheme.

3.6.2 Union Bound

Recall (3.38), the PEP of the CPNC scheme is upper-bounded by

$$\begin{aligned}
&P_e\left(\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)\to\mathbf{c}_N'\right)\\
&\overset{(a)}{\leq}\Pr\left[\max_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N^*)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)\leq\max_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N')} p\left(\mathbf{y}_R|\mathbf{x}_s\right)\right]\\
&\overset{(b)}{\leq}\Pr\left[p\left(\mathbf{y}_R|\mathbf{x}_s^*\right)\leq\max_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N')} p\left(\mathbf{y}_R|\mathbf{x}_s\right)\right]\\
&\overset{(c)}{\leq}\sum_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N')}\Pr\left[p\left(\mathbf{y}_R|\mathbf{x}_s^*\right)\leq p\left(\mathbf{y}_R|\mathbf{x}_s\right)\right]\\
&\overset{(d)}{=}\sum_{i=0}^{N}\mathcal{J}^{d_i}\left(\mathbf{c}_N^*,\mathbf{c}_N'\right)Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*,\mathbf{c}_N'\right)+i\cdot 4E_s}{\sigma^2}}\right)
\end{aligned} \tag{3.40}$$

where $N=\left|\mathcal{S}^c\left(\mathbf{c}_N^*\right)\cap\mathcal{S}^c\left(\mathbf{c}_N'\right)\right|$. In the above, "$(a)$" follows from (3.39); "$(b)$" follows from the fact that $p\left(\mathbf{y}_R|\mathbf{x}_s^*\right)\leq\max\limits_{\mathbf{x}_s\in\mathcal{X}_s(\mathbf{c}_N^*)} p\left(\mathbf{y}_R|\mathbf{x}_s\right)$; "$(c)$" is from the union bound; and "(d)" has used the definition of $\mathcal{J}^{d_i}\left(\cdot\right)$ and Corollary 3.4. From (3.40), we can see that the PEP $P_e\left(\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)\to\mathbf{c}_N'\right)$ is only dependent on $\mathbf{c}_N^*$ and $\mathbf{c}_N'$ for all $\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)$. Thus, $P_e\left(\mathbf{x}_s^*\in\mathcal{X}_s(\mathbf{c}_N^*)\to\mathbf{c}_N'\right)$ will be replaced by $P_e\left(\mathbf{c}_N^*\to\mathbf{c}_N'\right)$ in the sequel.

Remark 3.5. From (3.40), it is clear that the PEP $P_e\left(\mathbf{c}_N^*\to\mathbf{c}_N'\right)$ is determined by both the Hamming distance $w_H\left(\mathbf{c}_N^*,\mathbf{c}_N'\right)$ and the PDS $\mathcal{J}\left(\mathbf{c}_N^*,\mathbf{c}_N'\right)$. This is different from the conventional single-user scenario and the PDS of the CPNC scheme generally results in an *extra multiplicity* in the PEP (3.40).

Given the PEP, the word error probability (WEP) conditioned on $\mathbf{c}_N^*$ is upper bounded by

$$\begin{aligned} P_e\left(\mathbf{c}_N^*\right) &\leq \sum_{\mathbf{c}'_N \in \mathcal{C}, \mathbf{c}'_N \neq \mathbf{c}_N^*} P_e\left(\mathbf{c}_N^* \to \mathbf{c}'_N\right) \\ &\leq \sum_{\substack{\mathbf{c}'_N \in \mathcal{C}, \\ \mathbf{c}'_N \neq \mathbf{c}_N^*}} \sum_{i=0}^{N} \mathcal{J}^{d_i}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) + i4E_s}{\sigma^2}}\right) \end{aligned} \tag{3.41}$$

and the average WEP is

$$P_e = \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c}_N \in \mathcal{C}} P_e\left(\mathbf{c}_N\right). \tag{3.42}$$

To compute the WEP (3.42), the PDS $\mathcal{J}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ is required. For a short code, $\mathcal{J}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ can be determined by using the proposed punctured codebook method together with Theorem 3.2, as in Example 3. However, as the codeword length n increases, the number of distinct rows in $\mathbf{C}^{\mathcal{S}^c\left(\mathbf{c}_N^*\right) \cap \mathcal{S}^c\left(\mathbf{c}'_N\right)}$ increases exponentially with n and the task becomes prohibitive. To simplify this task, we next provide an approximate performance bound for the CPNC scheme.

3.6.3 An Approximate Performance Bound

As mentioned previously, determining the complete $\mathcal{J}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ is not practical for a medium-to-large n. To solve this problem, we derive an upper bound for the high SNR case by using the following lemma. Here, the SNR under consideration is the per-user SNR of the uplink phase, given by $\frac{E_s}{\sigma^2}$.

Lemma 3.3. *For a finite codeword length n, we have*

$$\lim_{\sigma^2 \to 0} \frac{\sum\limits_{i=1}^{N} \mathcal{J}^{d_i}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) + i \cdot 4E_s}{\sigma^2}}\right)}{\mathcal{J}^{d_0}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)}{\sigma^2}}\right)} = 0. \tag{3.43}$$

Given Lemma 3.3, we have the following theorem.

Theorem 3.3. *At high SNRs, i.e., $\sigma^2 \to 0$, we have*

$$\lim_{\sigma^2 \to 0} P_e\left(\mathbf{c}_N^* \to \mathbf{c}_N'\right) \leq 2^{d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)} Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{\sigma^2}}\right) \tag{3.44}$$

where $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right) \triangleq \left|\mathcal{S}\left(\mathbf{c}_N^*\right) \cap \mathcal{S}^c\left(\mathbf{c}_N'\right)\right|.$

For convenience, (3.44) is written as

$$P_e\left(\mathbf{c}_N^* \to \mathbf{c}_N'\right) \overset{\sigma^2 \to 0}{\lesssim} 2^{d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)} Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{\sigma^2}}\right). \tag{3.45}$$

Proof. From Lemma 3.3, we have

$$\lim_{\sigma^2 \to 0} P_e\left(\mathbf{c}_N^* \to \mathbf{c}_N'\right) \leq \mathcal{J}^{d_0}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right) Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{\sigma^2}}\right). \tag{3.46}$$

Using (3.35) of Lemma 3.2 in (3.46), (3.44) is obtained. ■

From (3.46), we notice that only the events corresponding to the minimum distance set, whose cardinality is $\mathcal{J}^{d_0}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)$, are relevant to the PEP. The error probabilities from other events vanish at high SNRs. The cardinality of the minimum distance subset give rises to a *multiplicity* of $\mathcal{J}^{d_0}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right) \leq 2^{d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}$ in the error probability in (3.45).

Remark 3.6. From Theorem 3.3, at high SNRs, (3.45) is an upper bound of $\Pr\left(\mathbf{c}_N^* \to \mathbf{c}_N'\right)$. Later in Section V, numerical results will demonstrate that (3.45) leads to a tight upper bound at high SNRs. Furthermore, we will see from numerical results that (3.45) is an upper bound for the entire range of SNRs.

For more insight, consider a single-user (SU) one-way relay (OWRC) case, where the relay is required to decode the message. The PEP (at the relay) of this SU case, $\Pr\left(\mathbf{c}_N^* \overset{\mathrm{SU}}{\to} \mathbf{c}_N'\right)$, is given by

$$\Pr\left(\mathbf{c}_N^* \overset{\mathrm{SU}}{\to} \mathbf{c}_N'\right) = Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{\sigma^2}}\right). \tag{3.47}$$

Comparing (3.45) and (3.47), at high SNRs, the PEP (upper bound) of the CPNC two-way relay scheme is approximately increased by a factor of $2^{d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}$ relative to that of the single-user OWRC case.

Next, given the PEP of the CPNC scheme, the WEP conditioned on $\mathbf{c}_N$ is approximated as

$$P_e\left(\mathbf{c}_N^*\right) \overset{\sigma^2\to 0}{\lesssim} \sum_{\mathbf{c}'_N \in \mathcal{C}, \mathbf{c}'_N \neq \mathbf{c}_N^*} 2^{d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)} Q\left(\sqrt{\frac{E_s d_H\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)}{\sigma^2}}\right). \tag{3.48}$$

At this stage, we need to determine $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ which is dependent on $\mathbf{c}_N^*$. For a short code, this can be found by examining all codewords in $\mathcal{C}$. For long codes, evaluating $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ for all codewords is prohibitive and we will use the following approximation for $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$.

Let us consider a capacity-achieving random code with a large codeword length n. We choose two codewords at random, selecting each digit of each word independently as 0 or 1 with equal probability. Then for large n the two codewords will, with high probability, differ in about half the positions in the block. This was shown in [138, pp. 134]. Now let $\mathbf{c}_N$ and $\mathbf{c}'_N$ denote these two codewords. Then we have $\Pr\{|d_H(\mathbf{c}_N^*, \mathbf{c}'_N) - n/2| < \varepsilon_1\} \overset{n\to\infty}{\to} 1$ for an arbitrarily small $\varepsilon_1 > 0$. Similarly, for large n, the two codewords will, with high probability, have a quarter of the position in the block that $\mathbf{c}_N^*$ is 1 and $\mathbf{c}'_N$ is 0. Then we have $\Pr\{|d_{10}(\mathbf{c}_N^*, \mathbf{c}'_N) - n/4| < \varepsilon_2\} \overset{n\to\infty}{\to} 1$ for an arbitrarily small $\varepsilon_2 > 0$. This will give rise to

$$\Pr\left\{\left|\frac{d_{10}(\mathbf{c}_N^*, \mathbf{c}'_N)}{d_H(\mathbf{c}_N^*, \mathbf{c}'_N)} - \frac{1}{2}\right| < \varepsilon\right\} \overset{n\to\infty}{\to} 1$$

for an arbitrarily small $\varepsilon > 0$. This means that for random codes, $d_{10}(\mathbf{c}_N^*, \mathbf{c}'_N)$ approaches $d_H(\mathbf{c}_N^*, \mathbf{c}'_N)/2$ with probability 1, as n approaches infinity. We assume that this behavior also exists for practical long linear codes and use the following approximation

$$d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right) \approx \frac{d_H\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)}{2}. \tag{3.49}$$

Later, we will see from numerical results that this approximation characterizes $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}'_N\right)$ very well.

Now, substituting $d_{10}\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)$ with $\frac{d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{2}$ in (3.48), we obtain

$$\begin{aligned} P_e\left(\mathbf{c}_N^*\right) &\overset{\sigma^2 \to 0}{\lesssim} \sum_{\mathbf{c}_N' \in \mathcal{C}, \mathbf{c}_N' \neq \mathbf{c}_N^*} 2^{\frac{d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{2}} Q\left(\sqrt{\frac{E_N d_H\left(\mathbf{c}_N^*, \mathbf{c}_N'\right)}{\sigma^2}}\right) \\ &= \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} A_d\left(\mathbf{C}\right) 2^{\frac{d}{2}} Q\left(\sqrt{\frac{E_s d}{\sigma^2}}\right) \end{aligned} \tag{3.50}$$

where $A_d\left(\mathbf{C}\right)$, $d = d_{\min}, \cdots, d_{\max}$, denotes the number of codewords in $\mathbf{C}$ with Hamming weights d, i.e., the weight enumerating function (WEF) [136, 139, 140].

It is clear that, with the approximation (3.49), $P_e\left(\mathbf{c}_N\right)$ is now only dependent on the codebook $\mathbf{C}$ and independent of $\mathbf{c}_N$. The averaged WEP is

$$\begin{aligned} P_e = \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c}_N \in \mathcal{C}} P_e\left(\mathbf{c}_N\right) &\overset{\sigma^2 \to 0}{\lesssim} \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} A_d\left(\mathbf{C}\right) 2^{\frac{d}{2}} Q\left(\sqrt{\frac{E_s d}{\sigma^2}}\right) \\ &\leq \frac{1}{2} \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} A_d\left(\mathbf{C}\right) \exp\left(-\frac{d}{2}\left(\frac{E_s}{\sigma^2} - \ln 2\right)\right) \end{aligned} \tag{3.51}$$

where we have used $Q\left(\sqrt{x}\right) \leq \frac{1}{2}\exp\left(-\frac{x}{2}\right)$. The bit error probability (BEP) can be given as

$$\begin{aligned} P_b &\overset{\sigma^2 \to 0}{\lesssim} \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} B_d\left(\mathbf{C}\right) 2^{\frac{d}{2}} Q\left(\sqrt{\frac{E_s d}{\sigma^2}}\right) \\ &\leq \frac{1}{2} \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} B_d\left(\mathbf{C}\right) \exp\left(-\frac{d}{2}\left(\frac{E_s}{\sigma^2} - \ln 2\right)\right) \end{aligned} \tag{3.52}$$

where $B_d\left(\mathbf{C}\right)$ is the average information weight w.r.t. all codewords of weight d [132].

Comparing (3.52) with the BEP of the single-user OWRC case

$$P_b^{\mathrm{SU}} \leq \frac{1}{2} \sum_{d=d_{\min}(\mathbf{C})}^{d_{\max}(\mathbf{C})} B_d\left(\mathbf{C}\right) \exp\left(-\frac{d}{2}\frac{E_s}{\sigma^2}\right), \tag{3.53}$$

at high SNRs, the CPNC scheme exhibits an SNR degradation relative to the single-user case. This SNR degradation is at most $\ln 2$ in linear scale. A look-up table between this SNR degradation of $\ln 2$ in linear scale and that in logarithmic (dB) scale is presented in Table 3.2 for several values of SNR in dB. This loss is essentially due to the multiple error events w.r.t. the minimum distance subset in (3.40).

3.7 Numerical Results

In this section, we show the error-rate performance of CPNC schemes, based on Monte-Carlo simulations, to verify the analysis given in previous sections. Two types of binary linear codes are considered: Hamming codes and convolutional codes. For reference purpose, the performance of a single-user OWRC case with the same codes are also included. Note that all the error rates refer to those at the relay, which follows [46].

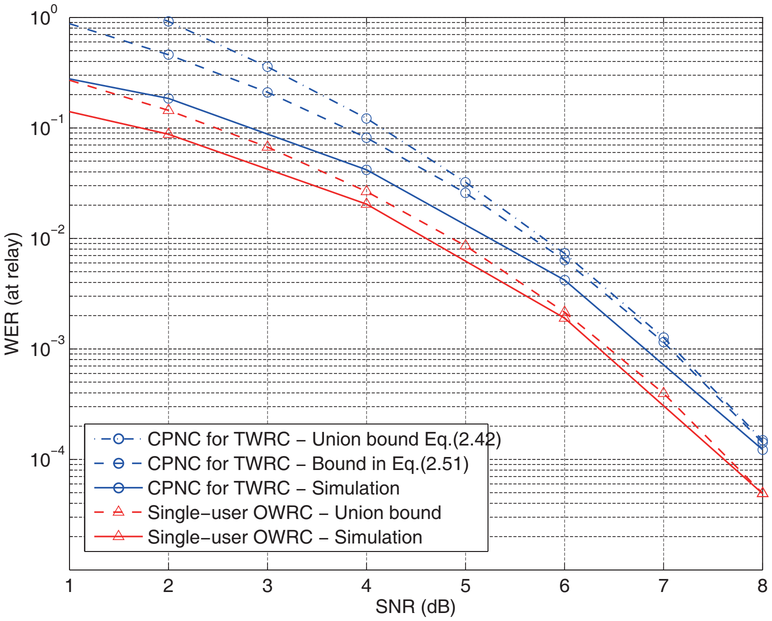


Figure 3.2 Performance of a CPNC scheme with a (7, 4) Hamming Code.

Fig. 3.2 presents the word error rate (WER) of a CPNC scheme with $\mathcal{C}$ being a (7,4) Hamming code. The union bound (3.41) and (3.42), the approximated bound (3.51), and the simulated results are depicted. It is shown that, as SNR increases, the simulated curve asymptotically approaches both (3.42) and (3.51), which are tight upper bounds. It is also observed that the two-user CPNC is inferior to the single-user case with a SNR degradation less than $\ln 2$ (linear scale) at high SNRs. For example, for SNR=6 dB of the single-user case, to achieve the same WER, the two-user CPNC scheme carries an SNR penalty of 0.56 dB. From Table 3.2 at SNR=6 dB, an SNR

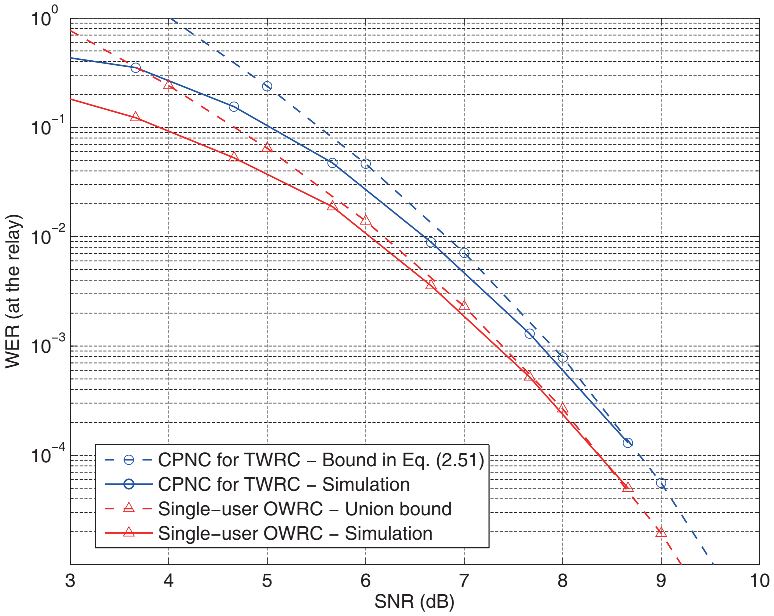


Figure 3.3 Performance of a CPNC scheme with a (15, 11) Hamming Code.

loss of ln 2 (linear scale) is equivalent to a SNR loss of 0.70 dB. This suggests the SNR loss in linear scale is less than ln 2, which validates our analysis. The performance of a CPNC scheme with $\mathcal{C}$ being a (15,11) Hamming code is shown in Fig. 3.3. Similar observations are made which lead to the similar conclusions to those w.r.t. Fig. 3.2. In addition, we observe that the SNR loss is (almost) identical to ln 2 for the scenario with a (15,11) Hamming code at a high SNR.

Table 3.2 Relationship between an SNR loss in linear scale and that in Decibel

$10\log_{10}\text{SNR}$	$10\log_{10}\left(\frac{\text{SNR}+\ln 2}{\text{SNR}}\right)$
0 dB	2.29 dB
1 dB	1.90 dB
2 dB	1.58 dB
⋮	⋮
6 dB	0.7 dB
10 dB	0.29 dB

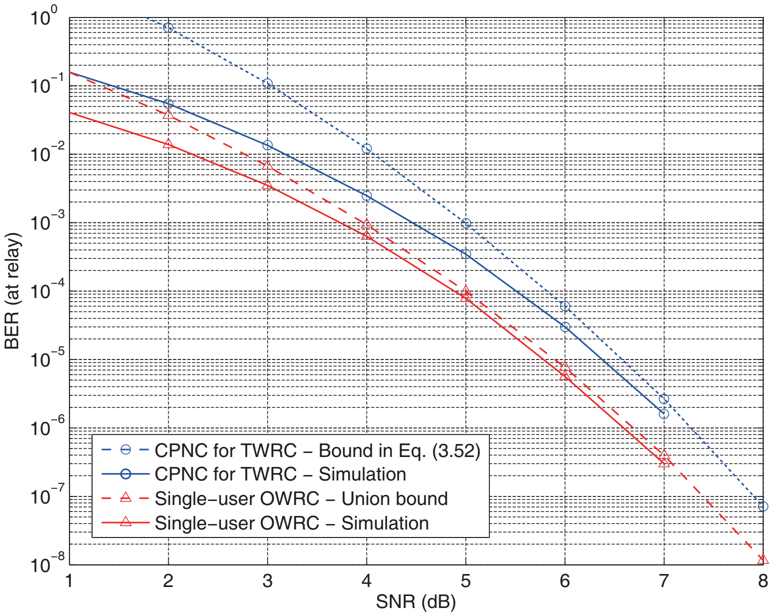


Figure 3.4 Performance of a CPNC scheme with a $[5,7]_8$ convolutional code.

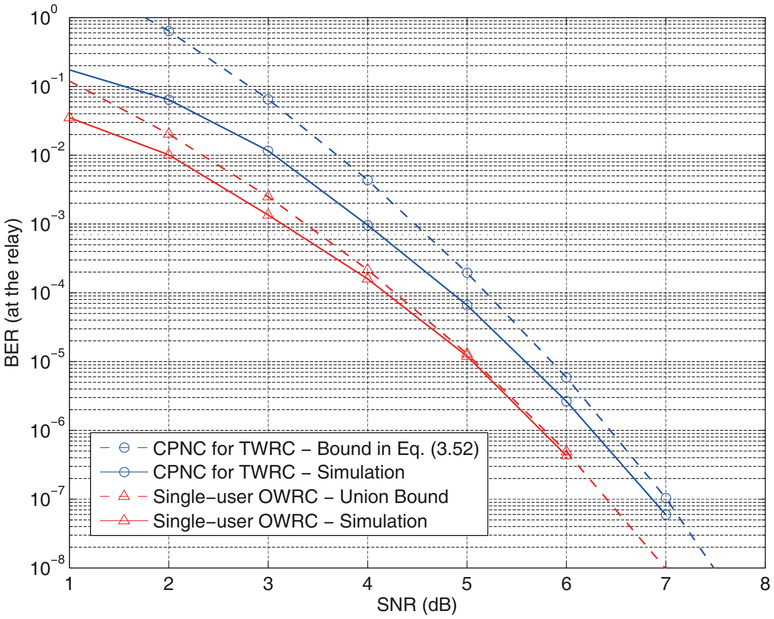


Figure 3.5 Performance of a CPNC scheme with a $[23,35]_8$ convolutional code.

In Fig. 3.4, the BER of the CPNC scheme is plotted where $\mathcal{C}$ is a convolutional code with generator polynomials $[5,7]_8$. In the simulations, we construct a "super trellis" and use the Viterbi algorithm to compute $\overline{\mathbf{c}}_s$ at the relay. The details can be found in [35]. Here, we use (3.52) as the performance bound where $d_{\max}(\mathbf{C})$ is set to 9 according to [141]. Fig. 3.4 shows that the bound (3.52) is asymptotically tight at high SNRs. Again, the SNR loss of the CPNC scheme relative to single-user case is shown to follow the analysis in Section IV. In Fig. 3.5, similar comparisons are carried out with a "stronger" convolutional code with generator polynomials $[23, 35]_8$. The same conclusions can be drawn in terms of the tightness of bound (3.52), as well as the SNR loss of the two-user CPNC scheme relative to the single-user case. In Fig. 3.6,

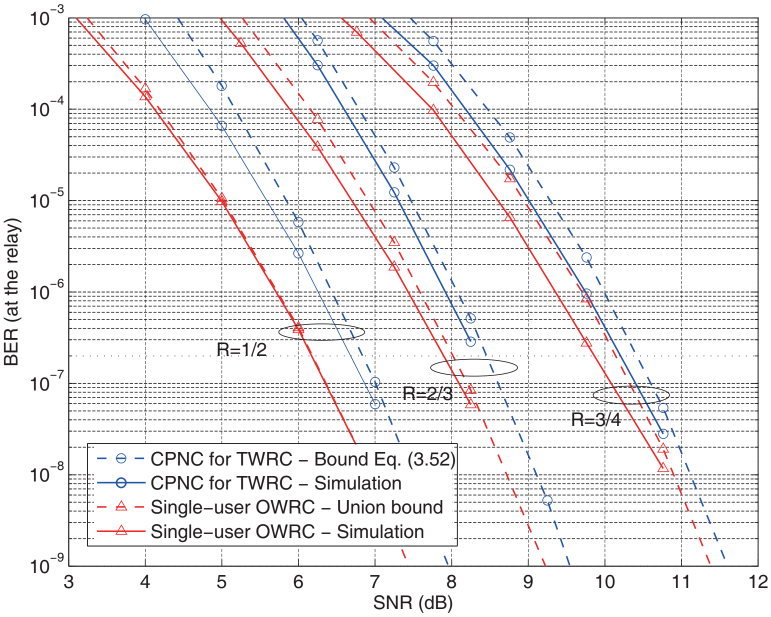


Figure 3.6 Performance of a CPNC scheme with convolutional codes with various rates.

we plot the BERs of CPNC schemes with convolutional codes of rates 1/2, 2/3 and 3/4, with the generator polynomials $[23, 35]_8$, $[27, 75, 72]_8$, $[36, 14, 32, 07]_8$, respectively [142]. For each code, the numerical result of the CPNC scheme is compared to the performance bound as well as to the performance of the single-user case. We observe that the analytical results match very well with the BERs for all code rates under consideration. The SNR degradation of the CPNC scheme relative to the single-user case is less than $\ln 2$ (linear scale) for these convolutional codes, and the SNR loss approaches $\ln 2$ as the code rate increases. In summary, the performance bounds developed have been substantiated in Figures 3.2-3.6 to be asymptotically tight ones for Hamming codes and convolutional codes. In addition, the analytical result on the SNR loss of the two-user CPNC scheme relative to the single-user case is also shown to be accurate. These results, in turn, confirm the distance property analysis of a CPNC scheme in this chapter.

3.8 Conclusions

In this work, we found that the minimum Euclidean distance of the CPNC scheme remains the same as that of the single-user case. This extends the minimum distance property of a channel un-coded PNC scheme to channel-coded cases. Yet, our analysis showed that the CPNC scheme is subjected to an increased multiplicity of minimum distance error events. At a high SNR, this leads to an SNR penalty of at most $\ln 2$ in linear scale, relative to the single-user scenario. The findings in this work suggest that, when designing a channel-coded PNC in a Gaussian two-way relay channel with binary classic codes, the error performance of computing the network coded information at the relay can be predicted by the error performance of the same codes in a conventional point-to-point channel.

Chapter 4

Design of CPNC in TWRCs with Binary Modern Codes

4.1 Introduction

In previous chapter, we analyzed the error performance of the channel-coded PNC (CPNC) for binary input Gaussian two-way relay channels (TWRCs) with classic codes. In this chapter, we investigate the design of modern channel codes coded PNC scheme. In particular, we study the design of Irregular repeat-accumulate (IRA) coded PNC in Gaussian two-way relay channels. There are many different ways to design IRA codes for conventional point-to-point channels. For example, the IRA code can be optimized using density evolution technique, or using extrinsic information transfer (EXIT) chart. In this chapter, we focus on the design of IRA codes with EXIT chart. Our goal is to design IRA codes for the CPNC scheme such that it performs close to the capacity limit of a binary-input Gaussian TWRC. In the IRA-PNC scheme under consideration, the relay computes a binary network codeword, from its received noisy ternary superimposed signal sequence, which is then forwarded to the users. To carry out this computation, it is required to extend the conventional Tanner graph [154] to an equivalent Tanner graph (ETG), defined over a ternary superimposed signal domain [46]. The presence of the ternary signal leads to the challenges in the convergence analysis and design of a capacity-approaching IRA-PNC scheme.

This chapter begins by presenting the background of channel coding theory. We then give preliminary knowledge of IRA codes, and EXIT charts. We will focus on the IRA-PNC scheme, and analyze the corresponding component decoders and derive the generalized update rules for these components in terms of log-likelihood ratios (LLRs). Two models for the soft information exchanged among the components decoders will be discussed and bounds on the approximation of the EXIT functions of the IRA-PNC scheme will be developed. Based on that, we will carry out an EXIT chart curve-fitting technique to construct optimized IRA codes. In the simulation section, we will compare our developed to the existing regular RA coded PNC schemes. We will also compare our developed IRA-PNC scheme with a complete decoding-based network coding scheme, in which the relay completely decodes both users' messages, using iterative multi-user detection and decoding, and then form the network-coded message.

4.2 Preliminary: Repeat-Accumulate Codes and EXIT Charts

4.2.1 Repeat-Accumulate Codes

RA codes is firstly introduced in [130], and can be graphically presented by Tanner graph [154]. RA codes can be classified as systematic RA codes and non-systematic RA codes. Non-systematic RA codes transmit only the parity bits, and can be seen as a class of serially concatenated codes, where outer code is repetition code and inner code is an accumulator. Systematic RA codes transmit both the message bits and parity bits, and they are not serially concatenated codes because both the inner and outer decoders receive channel information.

The encoding process for a rate $1/q$ non-systematic RA code is shown in Fig. 4.1, where q is the number of repetition at the outer code. A higher rate RA code can be obtained by placing a combiner just before the accumulator, or puncturing the output of the accumulator [155]. The encoding process is described as follow: The information sequence $b_1, \cdots, b_k$, are repeated q times after the repeater. The repeated qk bits are then randomly interleaved. The randomly interleaved qk bits are then passed through a $\frac{1}{1+D}$ accumulator, which can also be called a differential encoder. A non-systematic RA code only transmits the parity bits, where a systematic RA code transmits both information bits and parity bits.

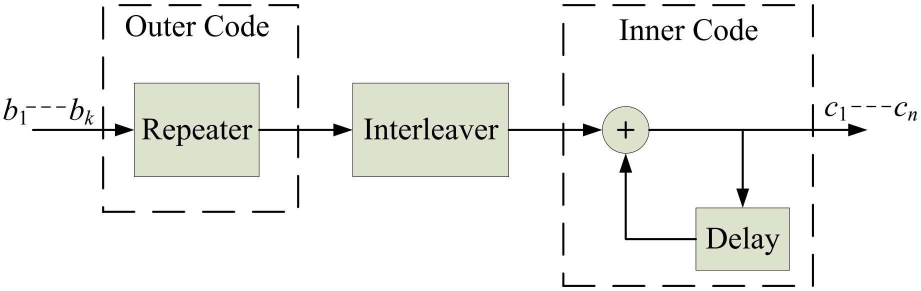


Figure 4.1 Block diagram of the encoding process of non-systematic RA codes.

Irregular RA (IRA) codes were firstly described in [153]. The difference between regular RA codes and IRA codes is that the outer code of an IRA code is a mixture of repetition codes. Compare to the regular RA codes, IRA codes are more flexible to design, and are able to achieve channel capacity. Compare to low-density parity-check (LDPC) codes, IRA codes offers competitive performance, but with much more simpler encoder. It is easier to present IRA codes using Tanner graph [154]. An example of Tanner graph of an IRA code is shown in Fig. 4.2.

The message sequence bits b_i, $i = 1, \cdots, k$, are repeated d_v times, where $d_v \in \{2, 3, \cdots\}$ specifies the length of repetition. The repetition, or variable node, degree distribution is given by $\lambda(d_v)$, $d_v \in \{2, 3, \cdots\}$ where $\lambda(d_v)$ is the portion of message

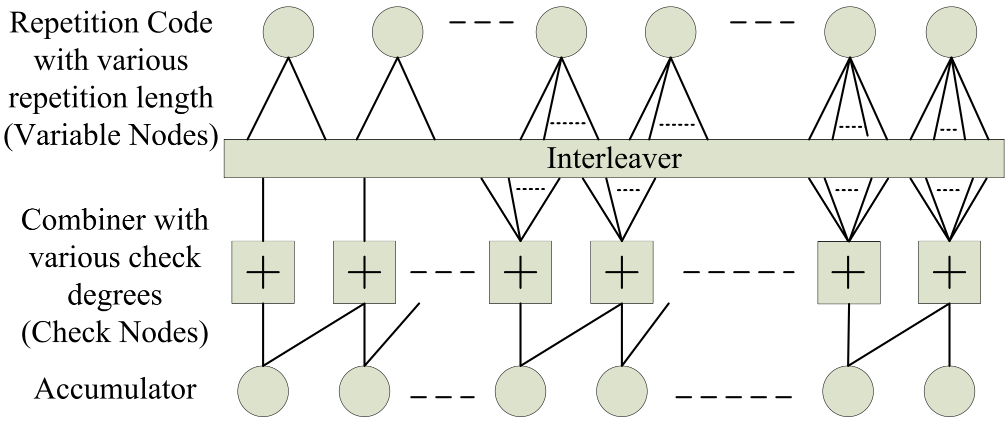

Repetition Code with various repetition length (Variable Nodes)
Interleaver
Combiner with various check degrees (Check Nodes)
Accumulator

Figure 4.2 Tanner graph presentation of an IRA code.

bits with repetition length d_v. Notice that $\lambda(d_v) \geq 0, \sum_{d_v=2}^{\infty} \lambda(d_v) = 1$. The repeated bit sequence is permuted by a random interleaver. The interleaved sequence is encoded by a series of parity-check codes (combiner) of degrees d_c, where $d_c \in \{1, 2, \cdots\}$. The check node degree distribution is given by $\rho(d_c)$, $d_c \in \{1, 2, \cdots\}$ where $\rho(d_c)$ is the portion of CNs whose number of connected edges is $d_c + 1$, which includes d_c edges connected to the interleaver and 1 edge connected to the Accumulator. The parity-check coded bits are then passed through an accumulator, generating the coded sequence c_j, $j = 1, \cdots, n$.

The decoding of IRA codes in AWGN channel can be performed on their Tanner graph with Belief propagation technique, which also known as sum-product message passing decoding algorithm. More details related to decoding of IRA codes can be found in [155].

4.2.2 Extrinsic Information Transfer Chart

Extrinsic Information Transfer (EXIT) Chart is firstly introduced by Stephan ten Brink in 1999 [156]. The motivation of introducing EXIT chart was to visualizing the convergence behaviour of the iterative decoding process of a given code. The EXIT chart is a technique that can simplify the design and construction of good, and capacity achieving codes with iterative decoding algorithm [143, 144, 146, 150].

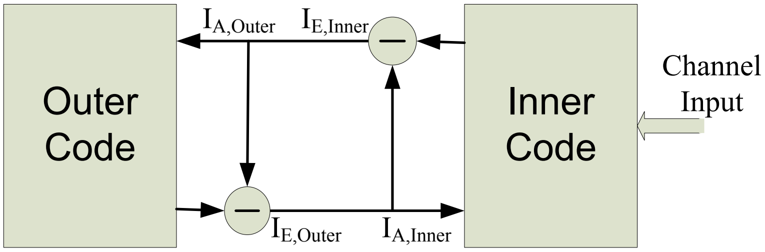


Figure 4.3 Information exchange between constitute decoders, where I_{A} denotes the a priori input and I_{E} denotes the extrinsic output.

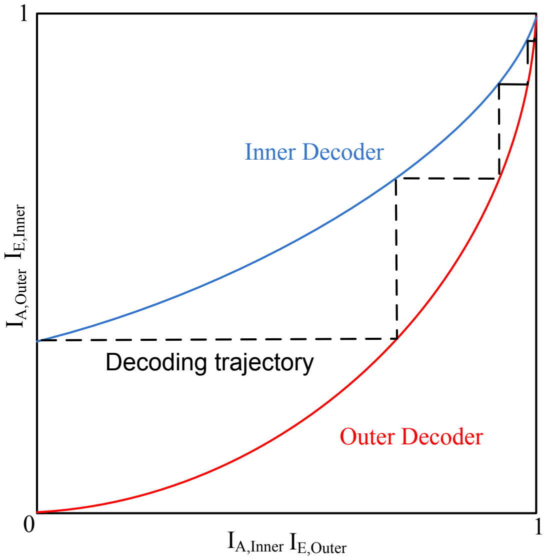


Figure 4.4 Example of an EXIT chart for an non-systematic binary IRA code.

Fig. 4.3 illustrates the information exchange of the iterative decoding process of a concatenate code. The extrinsic output of a constitute decoder becomes the a priori input of the other constitute decoder. Fig. 4.4 shows an example of an EXIT chart for the concatenated code in Fig. 4.3. The EXIT curve of the outer decoder starts from 0. This is because the outer decoder has no channel input. On the contrary, the start point of inner decoder EXIT curve is a non-zero point, since it has channel input. We can see that the decoding path is like a stepping function between the EXIT curves of the inner decoder and the outer decoder. Therefore, an open tunnel is required between the two EXIT function curves for a code can be decoded. To design an optimal code, the two EXIT curves need to lie close to each other, and a large number of iterations will be required. More detail of EXIT chart and designing codes with EXIT chart for point-to-point channel can be found in [15,56–59,61,132,143,144,146,150,155]. In addition, Hanzo *et al.* studied an extensive range of channel codes in [59], such as convolutional codes, block codes, and turbo codes. The authors also compared the following coded modulations under various channel conditions: Trellis Coded Modulation, Turbo Trellis Coded Modulation, Bit-Interleaved Coded Modulation (BICM) and Iterative BICM.

4.3 System Model

We consider a Gaussian TWRC where two single-antenna users, denoted by A and B, exchange information via an intermediate single-antenna relay. The users and the relay operate in half-duplex mode and there is no direct link between the users. The transmission protocol under consideration employs two time-slots for each round of information exchange. In the first time-slot (uplink phase), the users transmit their signals to the relay. In the second time-slot (downlink phase), the relay broadcasts to the two users. At each node, the received signal is corrupted by additive white Gaussian noise (AWGN).

Now we illustrate the CPNC scheme for the binary-input Gaussian TWRC. The block diagrams of the CPNC scheme is depicted in Fig. 4.5, which follows [46]. We first consider the uplink phase. Let $\mathbf{b}_A = [b_A(1), \cdots, b_A(k)] \in \{0,1\}^k$ denote the length-k binary message sequences of user A. This message sequence is encoded with a binary linear channel code and the resulting codeword is denoted by $\mathbf{c}_A = [c_A(1), \cdots, c_A(n)] \in \{0,1\}^n$, where n denotes the length of the codeword. The code rate per user is given by $R = k/n$. The users' codewords are modulated via binary phase shift keying (BPSK) ($0 \mapsto -1, 1 \mapsto +1$), resulting in the signal sequences $\mathbf{x}_A = 2\mathbf{c}_A - 1 \in \{-1,+1\}^n$. The encoder and modulation for user B are the same as user A, with similar notations. The signal sequences of user A and user B are

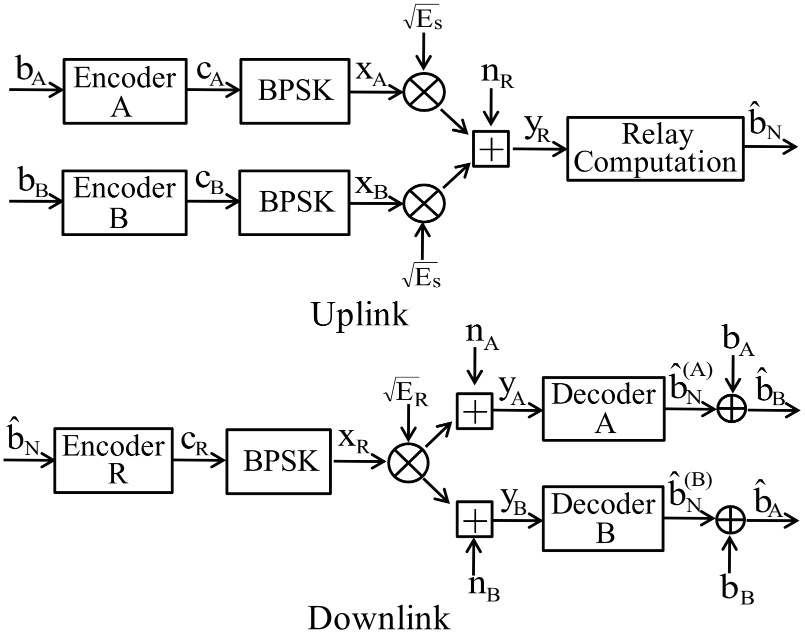


Figure 4.5 Architecture of a two-way relay system operated with CPNC. The relay computes the network-coded message $\mathbf{b}_N = \mathbf{b}_A \oplus \mathbf{b}_B$ without explicit decoding of both users' individual messages. Here, "+" denotes in the linear addition in real values and "$\oplus$" denotes the modulo-2 addition.

transmitted simultaneously.

It is noteworthy that, in general, the two users in a CPNC scheme of a TWRC may have different data rates and different signal power. In this chapter, however, we will follow the pioneering work [46] by limiting our discussion to the cases where the two users have identical data rates and the same received symbol energies. Assuming perfect synchronization [46], the signal received by the relay is

$$\mathbf{y}_R = \sqrt{E_s}\mathbf{x}_A + \sqrt{E_s}\mathbf{x}_B + \mathbf{n}_R = \sqrt{E_s}(\mathbf{x}_A + \mathbf{x}_B) + \mathbf{n}_R, \tag{4.1}$$

where E_s is the received symbol energy per-user and $\mathbf{n}_R$ is the AWGN vector at the relay. The variance of the noise is σ^2 and the per-user SNR in the uplink phase is given by E_s/σ^2. The perfect synchronization assumption here means that the phase and amplitude are fully synchronized at the receiver side. This assumption serves as a start point on the researching of modern channel coding in PNC.

In light of the notion of network coding [3], the relay needs to deliver a network-coded message to the two users. In this chapter, we define the *network-coded (NC) message* sequence as $\mathbf{b}_N \triangleq \mathbf{b}_A \oplus \mathbf{b}_B$, where "$\oplus$" denotes the element-wise modulo-2 addition operation. Upon receiving $\mathbf{y}_R$, the task of the relay is to *compute* an estimate of the NC message sequence $\mathbf{b}_N$, given as

$$\widehat{\mathbf{b}}_N = F_R\left(\mathbf{y}_R\right). \tag{4.2}$$

A *computation error* at the relay is declared if $\widehat{\mathbf{b}}_N \neq \mathbf{b}_N$.

In the downlink phase, as shown in Fig. 4.5, the relay re-encodes the computed NC message sequence $\widehat{\mathbf{b}}_N$ into a codeword $\mathbf{c}_R$, which is then BPSK-modulated, resulting signal sequence $\mathbf{x}_R$. This signal is broadcast to the two users. Then, user A first decodes the NC message $\widehat{\mathbf{b}}_N$. If the NC message is correctly recovered by both the relay and user A, user A can correctly recover user B's message by performing

$$\widehat{\mathbf{b}}_B = \mathbf{b}_A \oplus \widehat{\mathbf{b}}_N, \tag{4.3}$$

with the help of its own knowledge of $\mathbf{b}_A$. In contrast, if a computation error happens at the relay, a decoding error will happen. Note that there could be a minor case where the NC message $\mathbf{b}_N$ is wrongly computed by the relay but the final decoding result at a user is correct. However, the probability of such a case vanishes as k increases. We will not consider this trivial case in this chapter. The operation at user B is similar to that of user A. This completes one round of information exchange. More details about the downlink phase operation can be found in [46].

In the above CPNC scheme, the operation of decoding the NC message $\mathbf{b}_N$ at each user in the downlink phase is a standard single-user decoding. Thus, the key issue in the decoding of the CPNC scheme is to efficiently compute the NC message $\mathbf{b}_N$ at the relay in the uplink phase, i.e., Eq. (4.2). In this chapter, we will only investigate the computation of the NC message $\mathbf{b}_N$ at the relay (as in [46]).

4.4 Irregular Repeat-Accumulate Coded PNC

In general, any binary linear channel code, such as a convolutional code [35], a turbo code and a low-density parity-check (LDPC) code [51], can be employed in the CPNC scheme. In this work, we consider IRA codes, since their encoding is simpler than that of LDPC codes, and their structure allows a more flexible design than Turbo-codes. In particular, IRA codes have a flexible code structure, defined by the degree distribution of the variable nodes and check nodes, which allows for convenient design by curve-fitting in EXIT charts [143]. Here we consider non-systematic IRA codes. Our analysis and design also apply to a CPNC scheme with systematic IRA codes or other types of codes.

4.4.1 Encoding with an IRA Code

Consider the system in this chapter. In the uplink phase, user A's message bits $b_A(t)$, $t = 1, \cdots, k$, are repeated d_v times, where $d_v \in \{2, 3, \cdots\}$ specifies the length of repetition. The repetition, or variable node (VN), degree distribution is given by $\lambda(d_v)$, $d_v \in \{2, 3, \cdots\}$ where $\lambda(d_v)$ is the portion of message bits with repetition length d_v. Notice that $\lambda(d_v) \geq 0, \sum_{d_v=2}^{\infty} \lambda(d_v) = 1$. The repeated bit sequence is permuted by a random interleaver, denoted by $\pi(\cdot)$. The interleaved sequence is encoded by a series of parity-check codes of degrees d_c, where $d_c \in \{1, 2, \cdots\}$. The check node (CN) degree distribution is given by $\rho(d_c)$, $d_c \in \{1, 2, \cdots\}$ where $\rho(d_c)$ is the portion of CNs whose number of connected edges is $d_c + 1$. We denote the average CN and VN degrees by $\bar{d}_c$ and $\bar{d}_v$, respectively. The parity-check coded bits are then passed through an accumulator (ACC), generating the coded sequence $\mathbf{c}_A$. The same operation is performed at user B. The irregularity of this code resides in various repetition lengths (VN degrees) and various CN degrees.

It is noteworthy that when the two users transmit with the same rate, the same code is employed [46]. Then, the modulo-sum of the two users' codewords is still a codeword of the code. This is to ensure that the relay is able to compute the linear network coded codeword without decoding individual user's codeword, as we will see later.

4.4.2 Iterative Computation of the NC message at the Relay

The algorithm in [46] only applies to VNs of degree 3 and CNs of degree 2. Here, we develop a computation algorithm that applies to VNs and CNs of any degree. Then, we represent this algorithm in a log-likelihood ratio (LLR) format which will be required in the subsequent EXIT chart analysis.

Let us define $\mathbf{b}_s \triangleq \mathbf{b}_A + \mathbf{b}_B \in \{0,1,2\}^k$ and $\mathbf{c}_s \triangleq \mathbf{c}_A + \mathbf{c}_B \in \{0,1,2\}^n$ as a *superimposed message sequence* and a *superimposed codeword*, respectively. Consider the following linear processing

$$\mathbf{y}'_R = \mathbf{y}_R + 2\sqrt{E_s} = 2\sqrt{E_s}(\mathbf{c}_A + \mathbf{c}_B) + \mathbf{n}_R, \tag{4.4}$$

where $\mathbf{y}'_R$ is equivalent to $\mathbf{y}_R$ for the purpose of computing $\mathbf{b}_N$. From (4.4) we see that the signal $\mathbf{y}'_R$ is a noisy copy of the superimposed codeword $\mathbf{c}_s$. To compute the desired NC message, a "virtual encoding" process [46], which maps each superimposed message $\mathbf{b}_s$ to a superimposed codeword $\mathbf{c}_s$, is required. For an IRA coded PNC scheme, specifically, this "virtual encoding" process can be described via an equivalent Tanner graph (ETG), formed by superimposing two conventional Tanner graphs [46] of the same single-user IRA code, as shown in Fig. 4.6. The structure of the ETG resembles that of the single-user IRA code. However, there are two major differences:

1. The inputs and outputs have ternary symbols, i.e., $\mathbf{b}_s \in \{0,1,2\}^k$ and $\mathbf{c}_s \in \{0,1,2\}^n$. The message exchanged between the component nodes consists of the probabilities for the ternary symbols.

2. The ETG features an *equivalent CN function* and an *equivalent VN function*, denoted by $f_{\text{CN}}(\cdot)$ and $f_{\text{VN}}(\cdot)$, respectively, which are different from those of the single-user case.

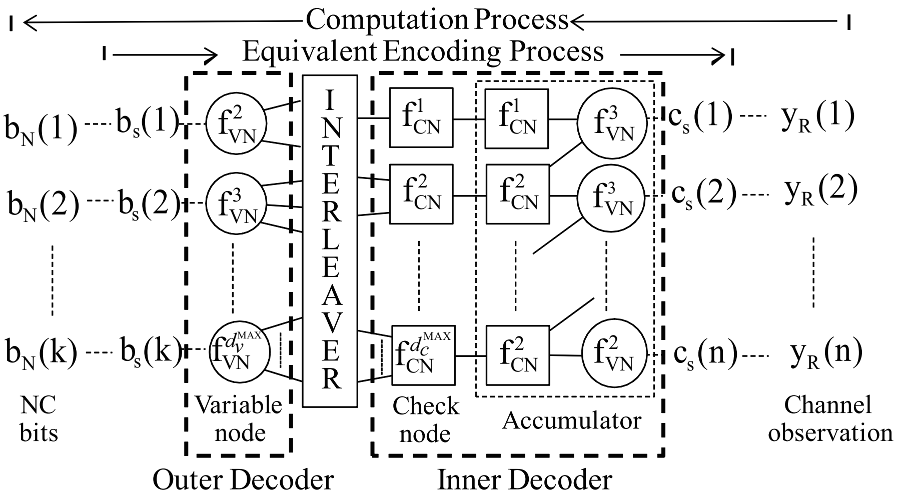


Figure 4.6 Equivalent Tanner graph of an IRA-PNC scheme.

Given $\mathbf{y}'_R$, the relay first exploits the ETG to compute an estimate of the ternary superimposed message sequence, denoted by $\widehat{\mathbf{b}}_s$. Next, given $\widehat{\mathbf{b}}_s$, the estimated NC message sequence $\widehat{\mathbf{b}}_N$ is obtained by calculating the modulo-2 of $\widehat{\mathbf{b}}_s$, i.e.,

$$\widehat{b}_N(t) = \begin{cases} 0 & \text{if } \widehat{b}_s(t) = 0 \text{ or } 2, \\ 1 & \text{if } \widehat{b}_s(t) = 1, \end{cases} \tag{4.5}$$

for $t = 1, \cdots, k$.

Now we briefly illustrate how to iteratively compute the superimposed message sequence $\mathbf{b}_s$, based on the ETG given above. Consider a node in the ETG which has L edges. The ternary *a priori* message to the lth edge of this node, $l = 1, \cdots, L$, is denoted by $P^{(l)} = [p_0^{(l)},\ p_1^{(l)},\ p_2^{(l)}]$, in which $p_\theta^{(l)}$ is the probability that the lth edge takes on the value of θ, $\theta \in \{0, 1, 2\}$. The collection of $P^{(l)}$ of all edges is denoted by $P \triangleq \{P^{(1)}, \cdots, P^{(L)}\}$. In the iterative computation process, a node takes the *a priori* probabilities P to calculate the *extrinsic probabilities*, according to its update rule. For the lth edge, $l = 1, \cdots, L$, the ternary extrinsic probabilities are denoted by $Q^{(l)} = [q_0^{(l)},\ q_1^{(l)},\ q_2^{(l)}]$, and the collection of them for all edges is denoted by $Q \triangleq \{Q^{(1)}, \cdots, Q^{(L)}\}$. The update rule can then be generally written as

$$Q = f(P). \tag{4.6}$$

Here, we use “P” and “Q” to distinguish the *a priori* probabilities from the extrinsic probabilities.

Initially, the relay calculates the ternary *intrinsic probabilities* based on the channel observation $\mathbf{y}'_R$:

$$\begin{aligned} p_\theta^{\mathrm{CH}} &= p(c_s = \theta | y'_R), \\ &= \begin{cases} \gamma \exp\left(-\frac{(y'_R - \theta\cdot(2\sqrt{E_s}))^2}{2\sigma^2}\right), \ \theta = 0, 2, \\ 2\gamma \exp\left(-\frac{(y'_R - \theta\cdot(2\sqrt{E_s}))^2}{2\sigma^2}\right), \ \theta = 1, \end{cases} \end{aligned} \tag{4.7}$$

where we have omitted the time-index, and γ is a normalization factor to ensure that $p_0^{\mathrm{CH}} + p_1^{\mathrm{CH}} + p_2^{\mathrm{CH}} = 1$. These intrinsic probabilities are collected as $P^{\mathrm{CH}} = [p_0^{\mathrm{CH}},\ p_1^{\mathrm{CH}},\ p_2^{\mathrm{CH}}]$, and they are only available to the accumulator. For the component decoders, the initialized *a priori* probabilities of each edge are $P^{(l)} = [1/4,\ 1/2,\ 1/4]$, $l = 1, \cdots, L$ [46].

In the process of computing $\mathbf{b}_s$, the ternary messages are iteratively exchanged among the component nodes in the ETG, in a similar fashion as the conventional iterative decoding of the single-user IRA code. As the receiver iterates, the ternary messages are refined using the update rules (4.6) of the component nodes, which will be detailed next. After a number of iterations, the computation process converges and a decision is made towards the estimated superimposed message sequence $\widehat{\mathbf{b}}_s$. Then the estimated NC message $\widehat{\mathbf{b}}_N$ can be obtained according to (4.5).

4.4.3 Update Rules with Probabilities

Let us first consider a CN with degree $d_c = 2$. There are $L = 3$ edges connected to this CN. Following the common approach in literature for IRA codes [143], a CN with degree d_c has d_c edges connected to the interleaver and one additional edge connected to the ACC. A VN with degree d_v has d_v edges connected to the interleaver. Recall that the *a priori* messages available to the first and second edge are given by $P^{(1)} = \left[p_0^{(1)}, p_1^{(1)}, p_2^{(1)}\right]$ and $P^{(2)} = \left[p_0^{(2)}, p_1^{(2)}, p_2^{(2)}\right]$, respectively. The extrinsic message of the third edge, denoted by $Q^{(3)} = \left[q_0^{(3)}, q_1^{(3)}, q_2^{(3)}\right]$, can be obtained as

$Q^{(3)} = f^2_{\text{CN}}\left(P^{(1)}, P^{(2)}\right)$, where the update rule $f^2_{\text{CN}}(\cdot)$ for a CN of degree 2 is given by [46]

$$q_0^{(3)} = \gamma\left(p_0^{(1)}p_0^{(2)} + \frac{p_1^{(1)}p_1^{(2)}}{2} + p_2^{(1)}p_2^{(2)}\right), \tag{4.8}$$

$$q_1^{(3)} = \gamma\left(p_1^{(1)}p_2^{(2)} + p_2^{(1)}p_1^{(2)} + p_1^{(1)}p_0^{(2)} + p_0^{(1)}p_1^{(2)}\right), \tag{4.9}$$

$$q_2^{(3)} = \gamma\left(p_0^{(1)}p_2^{(2)} + \frac{p_1^{(1)}p_1^{(2)}}{2} + p_2^{(1)}p_0^{(2)}\right). \tag{4.10}$$

Here, γ is a normalization factor to ensure that $q_0^{(3)} + q_1^{(3)} + q_2^{(3)} = 1$.

In general, for a CN with a degree $d_c > 2$, the update function $f^{d_c}_{\text{CN}}(\cdot)$ can be obtained by successively utilizing the degree-2 CN update rule, given by

$$\begin{aligned}\Gamma^{(2)} &= f^2_{\text{CN}}(P^{(1)}, P^{(2)}),\\ &\vdots\\ \Gamma^{(l)} &= f^2_{\text{CN}}(\Gamma^{(l-1)}, P^{(l)}),\\ &\vdots\\ Q^{(d_c+1)} = \Gamma^{(d_c)} &= f^2_{\text{CN}}(\Gamma^{(d_c-1)}, P^{(d_c)}).\end{aligned}$$

We refer to the above approach as a *successive update*.

4.4.4 Update Rules with LLRs

The ternary probabilities exchanged in the CPNC decoders put challenges on the analysis and design of the scheme. We next represent the update rule in terms of LLRs, which will be required in our subsequent EXIT chart analysis. For the lth edge of a component node in the ETG, the LLR couple associated with the *a priori* (ternary) probabilities are defined as

$$\Lambda_P^{(l)} \triangleq \log\left(\frac{p_0^{(l)} + p_2^{(l)}}{p_1^{(l)}}\right) \text{ and } \Omega_P^{(l)} \triangleq \log\left(\frac{p_0^{(l)}}{p_2^{(l)}}\right), \tag{4.11}$$

which are sufficient statistics of p_0, p_1, p_2.

From (4.5), we see that values $b_s(t) = 0$ and $b_s(t) = 2$ of the superimposed message are both mapped to the NC message bit $b_N(t) = 0$. Therefore, $\Lambda_P^{(l)}$ is related to the LLR of the binary NC message bit, and it has a pivotal role in the iterative computation process. To distinguish the two LLR values in (4.11), we refer to $\Lambda_P^{(l)}$ as the *primary* LLR and $\Omega_P^{(l)}$ as the *secondary* LLR. Similarly, the primary and secondary LLRs associated with the extrinsic probabilities are defined as

$$\Lambda_Q^{(l)} \triangleq \log\left(\frac{q_0^{(l)} + q_2^{(l)}}{q_1^{(l)}}\right) \text{ and } \Omega_Q^{(l)} \triangleq \log\left(\frac{q_0^{(l)}}{q_2^{(l)}}\right). \tag{4.12}$$

Consider a CN with $d_c = 2$. The primary extrinsic LLR of the third edge is calculated by

$$\begin{aligned}\Lambda_Q^{(3)} &= \log\left(\frac{q_0^{(3)} + q_2^{(3)}}{q_1^{(3)}}\right)\\ &\overset{(a)}{=} \log\left(\frac{p_0^{(1)}p_0^{(2)} + \frac{p_1^{(1)}p_1^{(2)}}{2} + p_2^{(1)}p_2^{(2)} + p_0^{(1)}p_2^{(2)} + \frac{p_1^{(1)}p_1^{(2)}}{2} + p_2^{(1)}p_0^{(2)}}{p_1^{(1)}p_2^{(2)} + p_2^{(1)}p_1^{(2)} + p_1^{(1)}p_0^{(2)} + p_0^{(1)}p_1^{(2)}}\right),\\ &= \log\left(\frac{1 + \exp\left(\Lambda_P^{(1)}\right)\exp\left(\Lambda_P^{(2)}\right)}{\exp\left(\Lambda_P^{(1)}\right) + \exp\left(\Lambda_P^{(2)}\right)}\right).\end{aligned} \tag{4.13}$$

where $\overset{(a)}{=}$ follows from (4.8)-(4.10). The secondary extrinsic LLR is calculated as

$$\begin{aligned}\Omega_Q^{(3)} &= \log\left(\frac{q_0^{(3)}}{q_2^{(3)}}\right),\\ &= \log\left(\frac{1 + \exp\left(\Omega_P^{(1)}\right)\exp\left(\Omega_P^{(2)}\right) + K_{\text{CN}}}{\exp\left(\Omega_P^{(2)}\right) + \exp\left(\Omega_P^{(1)}\right) + K_{\text{CN}}}\right),\end{aligned} \tag{4.14}$$

where

$$K_{\text{CN}} = \frac{\left[1 + \exp\left(\Omega_P^{(1)}\right)\right]\left[1 + \exp\left(\Omega_P^{(2)}\right)\right]}{2\exp\left(\Lambda_P^{(1)}\right)\exp\left(\Lambda_P^{(2)}\right)}. \tag{4.15}$$

Now, the update rule in terms of LLRs of a CN of $d_c = 2$ is given by

$$\begin{aligned}&\left[\Lambda_Q^{(3)}, \Omega_Q^{(3)}\right] = f_{\text{CN}}^2\left(\left[\Lambda_P^{(1)}, \Omega_P^{(1)}\right], \left[\Lambda_P^{(2)}, \Omega_P^{(2)}\right]\right),\\ &= \left[\log\left(\frac{1 + \exp\left(\Lambda_P^{(1)}\right)\exp\left(\Lambda_P^{(2)}\right)}{\exp\left(\Lambda_P^{(1)}\right) + \exp\left(\Lambda_P^{(2)}\right)}\right), \log\left(\frac{1 + \exp\left(\Omega_P^{(1)}\right)\exp\left(\Omega_P^{(2)}\right) + K_{\text{CN}}}{\exp\left(\Omega_P^{(2)}\right) + \exp\left(\Omega_P^{(1)}\right) + K_{\text{CN}}}\right)\right].\end{aligned}$$

The update rule of a CN with $d_c > 2$ can be calculated using the successive update approach described previously. In the sequel, we will use $f_{\mathrm{CN}}^{d_c}(\cdot)$ to denote the update rule of a CN of degree d_c in LLRs. A property of the update rule of a CN is presented next, which will be used later in the next section.

Property 4.1. For a CN with degree d_c, we have the output secondary LLR $\Omega_Q^{(l)} = 0$ as long as there exists an edge l', $l' \neq l$, such that the input secondary LLR $\Omega_P^{(l')} = 0$.

Explanation: In (4.14), if any of $\Omega_P^{(1)}$ or $\Omega_P^{(2)}$ equals to zero, the term $\Omega_Q^{(3)}$ will be zero. Consider the successive update rule, we obtain Property 4.1. ■

The derivation of the VN update rule $f_{\mathrm{VN}}(\cdot)$ in LLRs is similar and it is given by

$$\begin{aligned}
&\left[\Lambda_Q^{(l)}, \Omega_Q^{(l)}\right] \\
&= f_{\mathrm{VN}}^{d_v}\left(\left[\Lambda_P^{(1)}, \Omega_P^{(1)}\right], \cdots, \left[\Lambda_P^{(l-1)}, \Omega_P^{(l-1)}\right], \left[\Lambda_P^{(l+1)}, \Omega_P^{(l+1)}\right], \cdots, \left[\Lambda_P^{(d_v+1)}, \Omega_P^{(d_v+1)}\right]\right), \\
&= \left[(d_v - 2)\log 2 + \sum_{l'=1, l' \neq l}^{d_v} \Lambda_P^{(l')} + K_{\mathrm{VN}}, \sum_{l'=1, l' \neq l}^{d_v} \Omega_P^{(l')}\right]. \qquad (4.16)
\end{aligned}$$

where

$$K_{\mathrm{VN}} = \log\left(\frac{1 + \prod_{l'=1, l' \neq l}^{d_v} \exp\left(\Omega_P^{(l')}\right)}{\prod_{l'=1, l' \neq l}^{d_v}\left(1 + \exp\left(\Omega_P^{(l')}\right)\right)}\right). \qquad (4.17)$$

The detailed derivation is given in the Appendix B.

4.5 Convergence Behavior Analysis and Optimization of IRA-PNC

It is well-known that in the conventional single-user AWGN channel, the performance of an iteratively decoded IRA code is largely affected by its VN degree distribution $\lambda(d_v)$, and the CN degree distribution $\rho(d_c)$. The optimal performance can be achieved using the EXIT chart curve-fitting technique [143]. Now, we adopt this methodology in designing the IRA-PNC scheme, so as to approach the capacity limit of the binary-input Gaussian TWRC. However, for the two-user CPNC scheme, there lacks a method to characterize the EXIT behaviors w.r.t. the ternary probabilities that are exchanged in the iterative computation process.

In this section, we first propose a method to model the soft information exchanged among the components of the IRA-PNC scheme. This will enable us to obtain upper and lower bounds on the approximation of the EXIT functions. Based on that, we design optimal component codes via curve-fitting.

4.5.1 Modeling of EXIT Functions

To carry out convergence behavior analysis, we partition the ETG of the IRA-PNC scheme into two parts: an *inner component decoder* consisting of the combined CN and ACC (CN-ACC) decoder, and an *outer component decoder* consisting of the VN decoder. The idea of the EXIT chart technique is to predict the behavior of the iterative process by solely looking at the input/output mutual information of the two individual component decoders of the CPNC scheme.

Unlike the decoding of a conventional binary IRA code where *binary* probabilities are exchanged between the component decoders, the soft information exchanged between the CPNC component decoders have a ternary form. The ternary probabilities (or soft information) of the CPNC scheme can also be written in terms of the primary LLR Λ and the secondary LLR Ω, as in the previous section. In particular, the primary LLR Λ is related to the NC message $\widehat{b}_N$ to be computed. For simplicity, we omit the time index here. We denote b_N the random variable w.r.t. the NC message bit and denote Λ the random variable w.r.t. the primary LLR. Thus, the mutual information between b_N and the input primary LLR Λ_P, $I_A = I\left(b_N; \Lambda_P\right)$, will be used for tracking the *a priori* information of a component decoder of the CPNC scheme. Similarly, the mutual information between b_N and the output primary extrinsic LLR Λ_Q, $I_E = I\left(b_N; \Lambda_Q\right)$, will be used for tracking the corresponding extrinsic informa-

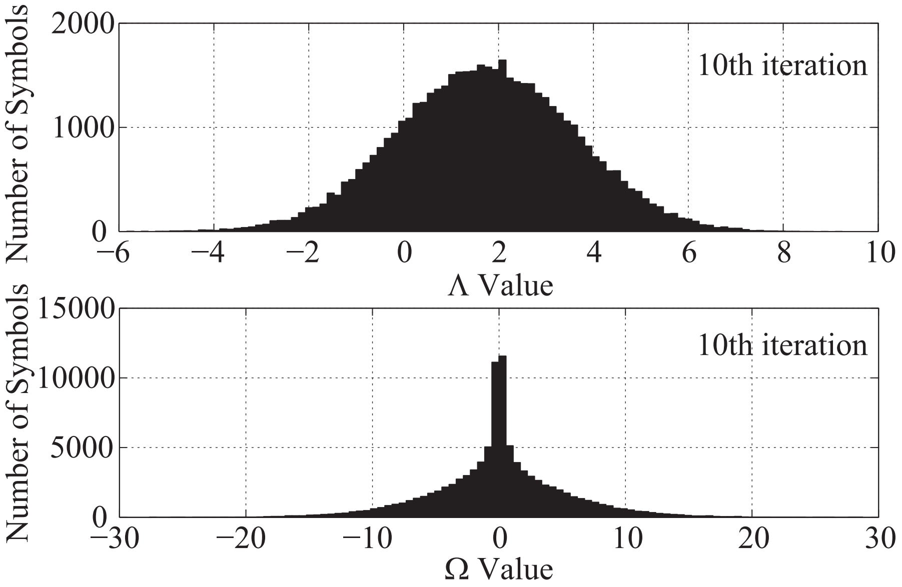


Figure 4.7 Histogram of the one-side extrinsic primary LLR output and Histogram of the extrinsic secondary LLR output from the VN decoder during iterative computation process.

tion. An output mutual information of $I_E = 1$ means that all NC message bits b_N can be decoded error free.

The relationship of the input-output mutual information, i.e., the EXIT function, of the inner component decoder (CN-ACC decoder) with CN degree distribution $\rho(d_c)$ can be written as

$$I_E = T_{\text{Inner}}\left(I_A, \mathbb{P}(\Omega_P), \rho(d_c), E_s/\sigma^2\right), \tag{4.18}$$

where $\mathbb{P}(\Omega_P)$ denotes the probability density function (PDF) of the secondary LLR Ω_P. We remark that, unlike the conventional single-user case, the EXIT function of the CPNC scheme is also affected by the PDF of the secondary LLR Ω_P. Similarly, the EXIT function of the outer component decoder (VN decoder) with VN degree distribution $\lambda(d_v)$ can be written as

$$I_E = T_{\text{Outer}}\left(I_A, \mathbb{P}(\Omega_P), \lambda(d_v)\right). \tag{4.19}$$

Notice that the EXIT function of the VN decoder is not affected by the SNR, since it is not directly connected to the channel observation.

Numerical results show that the PDF of the primary LLR approaches a consistent Gaussian-like distribution [144] with its mean equal to half of its variance, with the increasing number of iterations, as shown in Fig. 4.7. Thus, similar to [144], we can approximate the primary *a priori* LLR as

$$\Lambda_P = \frac{\sigma_\Lambda^2}{2}(1 - 2b_N) + n_\Lambda, \tag{4.20}$$

where n_Λ is a Gaussian random variable with variance σ_Λ^2, and we omit the time-index here for simplicity.. However, the PDF of the secondary LLR, as shown in Fig. 4.7, is not a Gaussian-like distribution. This makes the analytical treatment of the EXIT functions difficult. In order to tackle this problem, we propose two models for the secondary *a priori* information.

Model I: We assume that perfect secondary LLR is available in this model, that is,

$$\dot{\Omega}_P = \begin{cases} +\Psi & \text{if } b_s = 0, \\ 0 & \text{if } b_s = 1, \\ -\Psi & \text{if } b_s = 2, \end{cases} \tag{4.21}$$

where Ψ denotes a large positive value, e.g., 30, used in our simulation. Since the actual decoding process does not have perfect *a priori* information on the secondary LLR Ω_P for the component decoders, we have

$$I_E = T_{\text{Inner}}\left(I_A, \mathbb{P}(\Omega_P), \rho(d_c), E_s/\sigma^2\right) \leq T_{\text{Inner}}\left(I_A, \mathbb{P}(\dot{\Omega}_P), \rho(d_c), E_s/\sigma^2\right)$$

for the inner component decoder (CN-ACC decoder). Thus, we can obtain an *upper bound* for the approximation of the EXIT function of the inner component decoder by using Model I. Similarly, we can obtain an upper bound on the approximation of the EXIT function of the outer component decoder using Model I.

Model II: We assume the *a priori* secondary LLR Ω_P is completely absent, i.e., $\ddot{\Omega}_P = 0$.

As the actual decoding retains certain *a priori* information on the secondary LLR Ω_P for the component decoders, setting Ω_P to zero will result in an information loss.

Following the data processing inequality [145], we have $I_E = T_{\text{Inner}}\left(I_A, \mathbb{P}(\Omega_P), \rho(d_c), E_s/\sigma^2\right) \geq T_{\text{Inner}}\left(I_A, \mathbb{P}(\ddot{\Omega}_P), \rho(d_c), E_s/\sigma^2\right)$ for the inner component decoder and this also applies to the outer component decoder. Thus, a *lower bound* of the approximation of the EXIT function can be obtained using Model II.

4.5.2 EXIT Charts

We next show the EXIT functions of the component decoders of the IRA-PNC scheme using the two *a prior* information models developed earlier. The EXIT functions of the inner CN-ACC decoder with CN degrees $d_c = 1, \cdots, 5$ are shown in Fig. 4.8. These EXIT functions are obtained via simulations where Model I and Model II are used to construct the *a priori* information. Clearly, the EXIT function obtained with Model I is always higher than with Model II. This suggests that the availability of the secondary LLR Ω_P can contribute to a higher output extrinsic information. From Fig. 4.8, we also observe that the gap between the EXIT functions with Model I and with Model II diminishes as the CN degree d_c increases. When $d_c \geq 2$, the gap is almost unnoticeable. Here we give an intuitive explanation for this behavior. Let us consider the *a priori* information model I, the probability of $\Omega_P = 0$ is 0.5 since 50% of message bits have $b_N = 1$. Recall Property 4.1 which states that the output secondary LLR of a CN is zero as long as one of its input edges has the secondary LLR equal to zero. As the CN degree d_c increases, the probability of this event (there exists one edge whose secondary LLR is zero) also increases. As a result, there will be more zero-secondary LLR at the output of the CN nodes. This will restrain the propagation of the secondary LLR from the CN to the ACC of the inner decoder. As the CN degree becomes very high, the propagation of the secondary LLR becomes minimal and Model I and Model II tends to be identical.

From [143], it is known that to minimize the area between the EXIT functions of the component decoders, a capacity achieving IRA code tends to have a fairly large average CN degree, i.e., $\bar{d}_c > 2$. In this circumstance, for inner component decoder,

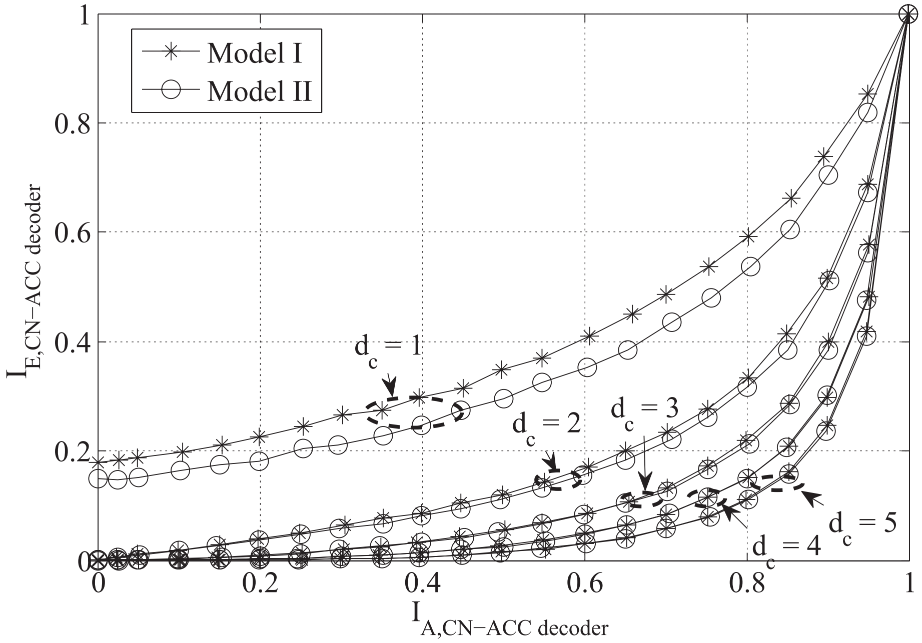


Figure 4.8 Comparison between Model I and Model II for inner CN-ACC decoder with various CN degrees. The code rate is 1/3, and the per-user SNR is $E_b/N_0 = 2.2$ dB.

the upper bound (Model I) and lower bound (Model II) of the approximation of the EXIT functions overlaps with each other. Therefore, either Model I or II can be used to obtain the approximation of the EXIT function of the inner component. Similarly, we can also use Model I or Model II to obtain the upper and lower bounds on the approximation of the EXIT function of the outer component decoder. We remark that since Model II gives a lower bound on the approximation of the EXIT function for either the inner or the outer component decoder, an optimal code based on Model II can always have its convergence guaranteed when the SNR is above its designed threshold.

Example 4.1. We consider an IRA-PNC scheme with per-user code rate of $R = 1/3$. In particular, the IRA under consideration has an average CN degree $\bar{d}_c = 2.4$ and an average VN degree $\bar{d}_v = 7.2$. The code parameters are given in Table A.1. In Fig. 4.9, we plot the EXIT function obtained by using the *a priori* information Model II and the actual decoding trajectory obtained from simulation. We observe that using

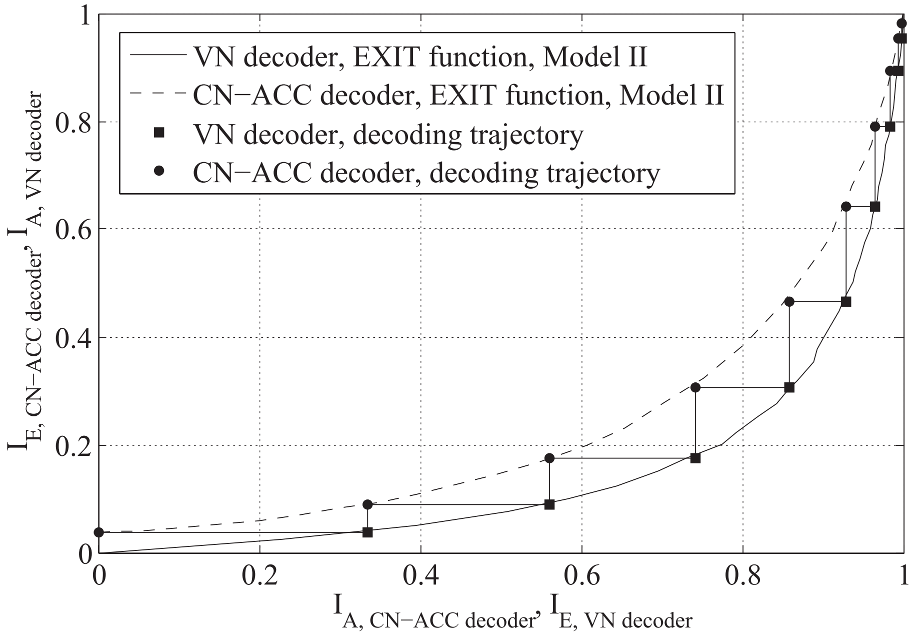


Figure 4.9 Comparison between EXIT Model II and actual decoding trajectory for an IRA code at $R = 1/3$ and $E_b/N_0 = 4$ dB.

the *a priori* information Model II, the EXIT functions of the component decoders of the IRA-PNC scheme can be accurately characterized. In the sequel, we will focus on using *a priori* information Model II for the design of the IRA-PNC scheme.

4.5.3 Code Optimization

Based on the developed EXIT functions of the component decoders of the IRA-PNC scheme, we now adopt the EXIT chart curve-fitting technique to design optimal IRA codes. The goal is to find CN and VN degree distributions such that the gap between the EXIT curves of the inner component decoder and the outer component decoder is minimized. Similar to [143], we first select an appropriate CN degree distribution. Then, we fit the EXIT curve of VN decoder to that of the CN-ACC decoder, by optimizing the degree distribution of the VN decoder via linear programming.

We next show two examples of the code design via EXIT curve-fitting for the IRA-PNC scheme. To avoid redundancy, the implementation details of the curve-fitting are omitted and can be found in [143]. In this chapter, we restrict ourselves to the

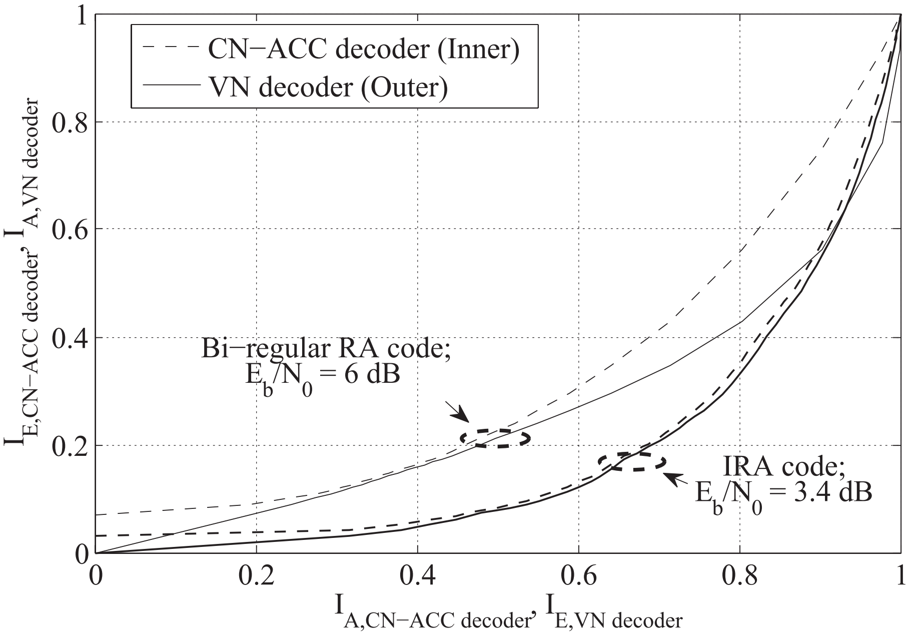


Figure 4.10EXIT charts of the bi-regular coded PNC scheme and the optimized IRA-PNC scheme, where $R = 3/4$.

commonly used concentrated check degree distributions [10, Section 3.17].

Example 4.2. We consider an IRA-PNC scheme where the code rate of each user is $R = 3/4$. In a conventional single-user case, it is known that, for a non-systematic RA code, a non-zero fraction of the CNs should have degree one to ensure that its decoder makes progress in the first iteration [143, 146]. From the check node update rule described in Section 4.4.4, we notice that in the IRA-PNC scheme, the CN degree distribution should contain a non-zero fraction for degree one CNs, similar to the conventional single-user RA codes case. In this design example, the choice of the portion of $d_c = 1$ CNs follows from the convention in [143]. The details are provided in Table A.1. In addition, to carry out EXIT curve-fitting, flexibility of the VN nodes are required so that the average VN degree cannot be too small, e.g., $\bar{d}_v > 3$. Then, for a code rate of 3/4, the average CN degree should be large enough, e.g., $\bar{d}_c > 2$. In this setting, the EXIT function of the bi-regular code can be characterized by Model II.

Fig. 4.10 shows the EXIT chart of a bi-regular RA coded PNC scheme whose de-

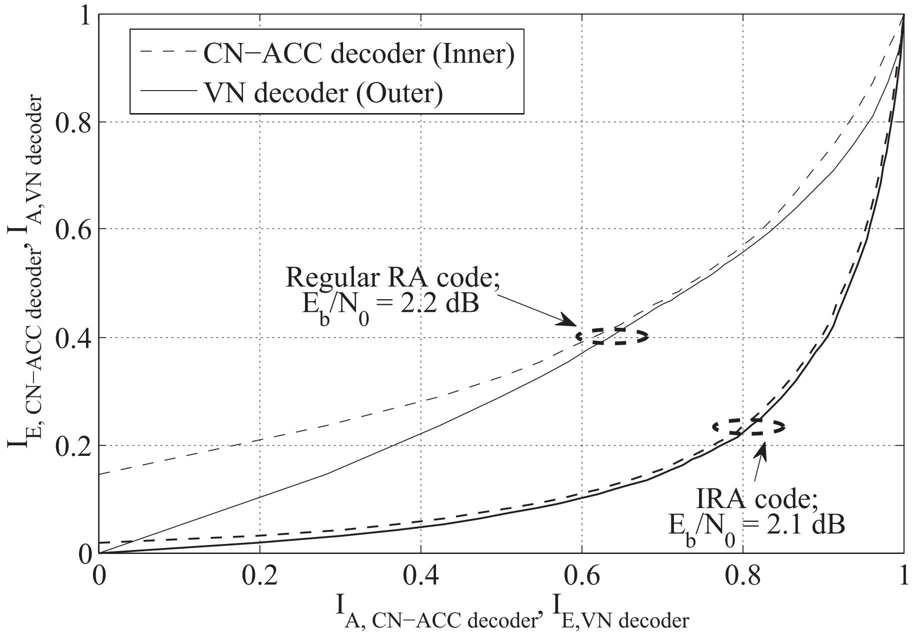


Figure 4.11EXIT charts of the regular coded PNC scheme and the optimized IRA-PNC scheme, where $R = 1/3$.

gree distributions are given in Table A.1. The decoding threshold for this benchmark scheme is found to be at $E_b/N_0 = 6$ dB. The EXIT chart of our optimized IRA-PNC scheme is also shown in Fig. 4.10, whose degree distributions are given in Table A.1. The decoding threshold of our optimized IRA-PNC scheme is found to be at $E_b/N_0 = 3.4$ dB. This shows that our developed IRA-PNC scheme can significantly outperform the bi-regular RA coded PNC scheme. The performance improvement is obtained from fitting the EXIT functions, and we refer to this performance improvement as a *curve-fitting gain*. In this example, the curve-fitting gain is about 2.6 dB.

Example 4.3. We consider another case where $R = 1/3$. For a regular RA coded PNC scheme, the threshold is found to be at $E_b/N_0 = 2.2$ dB, as shown in Fig. 4.11. We construct an IRA code for the CPNC scheme, using the curve-fitting technique based on our developed EXIT functions. The decoding threshold is found to be at $E_b/N_0 = 2.1$ dB. The degree distributions of our designed IRA code for the CPNC scheme are also given in Table A.1. We see that for the case of $R = 1/3$, the CPNC scheme with the designed IRA code only slightly outperforms that with a regular code. We next explain why the performance improvement is slight in this case.

Consider a simplified computation approach in which the secondary LLR Ω_P is always set to zero in the iterative computation process. Then, from (4.11), the soft information exchanged in the CPNC components can be completely specified by $[p_0+p_2, p_1]$, which has two elements. Here, we refer to this simplified approach as iterative computation with *binary information exchange*. In contrast, we refer to the approach utilizing both the primary and secondary LLRs as iterative computation with *ternary information exchange*, since the exchanged soft information has three elements, see (4.11). The performance improvement of using ternary information exchange (which utilizes the secondary LLR), over that with binary information exchange (which does not use the secondary LLR), is referred to as the *secondary LLR gain*. For the CPNC scheme of per-user coding rate $R = 1/3$, it is shown that the secondary LLR gain is as much as 0.5 dB when a regular RA code is utilized (see Fig. 4.12). In the process of optimizing the degree distributions of the IRA code to obtain the curve-fitting gain, the inner component decoder tends to have a relatively large average CN degree. This results in a reduced secondary LLR gain, as discussed in Section IV.B. For a relatively large CN degree, the secondary LLR gain vanishes. Finally, the combined effect of increased curve-fitting gain and reduced secondary LLR gain leads to only a slight performance improvement.

In contrast, in Example 4.2 where $R = 3/4$, the average degree of the CN of the CPNC scheme with a regular/biregular RA code is already relatively large, e.g., $\bar{d}_c \geq 2$. In this case, the secondary gain is already fairly small, as we can see in Fig. 4.8. Therefore, as we carry out the curve-fitting, there is no loss in the secondary LLR gain and the curve-fitting gain leads to a significant performance improvement.

It is noteworthy that IRA codes are special LDPC codes with a simpler encoder than general LDPC codes but with similar performance. The code optimization technique proposed in this chapter may apply to general LDPC codes.

4.6 Simulation Results

In the previous section, we have designed IRA-PNC schemes based on EXIT chart analysis and curve-fitting techniques. In this section, we present numerical results to show the benefits of our designed IRA-PNC schemes for finite code lengths. Specifically, we first compare the bit-error rate (BER) performance of our developed IRA-PNC schemes to the existing CPNC schemes with regular (or bi-regular) RA codes. Next, we compare the performance of our developed IRA-PNC scheme to the capacity limits, as well as to the complete decoding-based scheme.

In the simulations, we consider the BER performance of computing the NC message $\mathbf{b}_N$ at the relay. In all simulations, the length of the binary message sequence of each user is set to $k = 32768$. In the iterative computation process, the maximum number of iterations is set to 200.

4.6.1 IRA-PNC versus Regular/Bi-regular Coded PNC

Per-user Code Rate $R = 3/4$

The BER simulation results of the CPNC scheme with this code rate are shown in Fig. 4.12. At a BER of 10^{-4}, our developed IRA coded PNC scheme performs about 2.6 dB better than the bi-regular RA coded PNC scheme. This is in line with our EXIT chart analysis. From this result, we can conclude that IRA codes designed based on our EXIT analysis can significantly improve the performance of the CPNC scheme. We also notice that there is no performance degradation when the iterative computation with ternary information exchange is replaced by that with the binary information exchange.

Per-user Code Rate $R = 1/3$

The BER simulation results of the CPNC scheme with this code rate are shown in Fig. 4.12. When the iterative computation with binary information exchange is utilized,

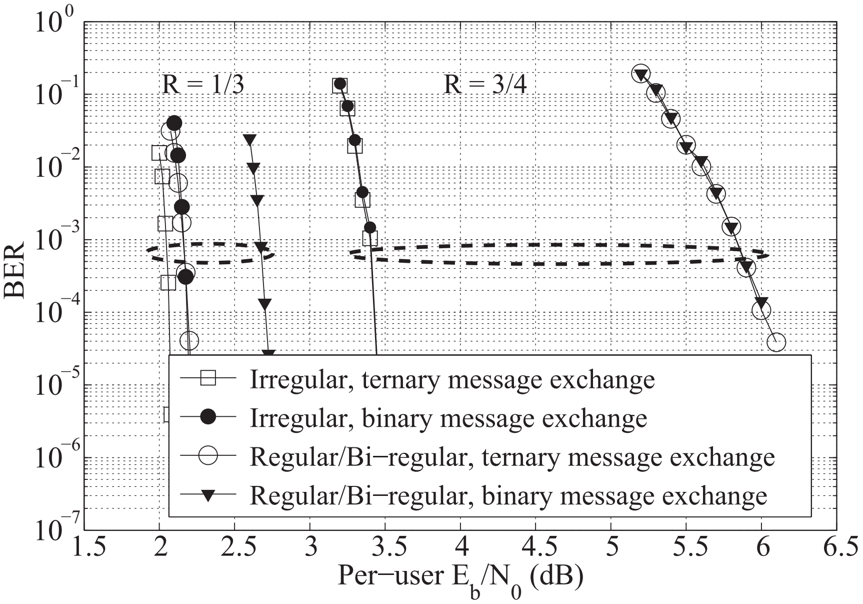


Figure 4.12 Simulation results of the designed IRA-PNC scheme where $R = 3/4, 1/3$.

our IRA-PNC scheme is about 0.5 dB better than the existing PNC scheme with the regular code in [46]. The performance improvement is from the full realization of the curve-fitting gain. When ternary information exchange is utilized, the designed IRA-PNC scheme is about 0.1 dB better than that with the regular RA code. These results are also in line with our EXIT chart analysis.

Other Code Rates

Fig. 4.13 shows the performance of the optimized IRA-PNC scheme with various code rates, where ternary information exchange is utilized. For code rates of 1/2 and 2/3, at a BER of 10^{-4}, we observe that the performance improvement over regular RA-PNC schemes are 1.6 dB and 1.9 dB, respectively. The code parameters are given in Table A.1.

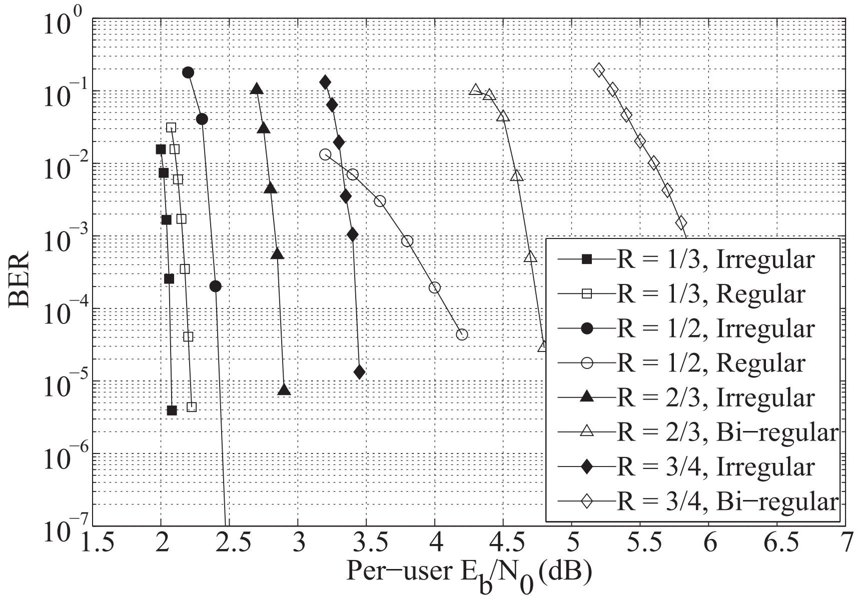


Figure 4.13 Performance of optimized IRA-PNC schemes at various code rates, where ternary information exchange decoding is utilized.

4.6.2 CPNC Versus Complete Decoding-Based Network Coding

Now, we compare the performance of the CPNC scheme to the scheme which performs complete decoding to generate the NC message at the relay. For a fair comparison, the *optimized* IRA-PNC scheme and the *optimized* complete decoding-based scheme are considered. In particular, given a total power constraint, equal power allocation is the best for a PNC scheme, see [14,46]. On the other hand, unequal power allocation is optimal for the complete decoding-based scheme [147,149], as this facilitates the complete separation of two users' codewords.

For the scheme with complete decoding, we employ iterative multi-user detection and decoding (IDD) [147] to fully decode both user A and user B's messages $\widehat{\mathbf{b}}_A$ and $\widehat{\mathbf{b}}_B$, and then determine the NC message as $\widehat{\mathbf{b}}_A \oplus \widehat{\mathbf{b}}_B$. In an IDD algorithm, soft information is iteratively exchanged between multiple single user decoders and a multi-user detector; details for code optimization using IDD algorithm can be found in [149, 150]. We emphasize that the "CNC1" scheme in [46] is equivalent to the

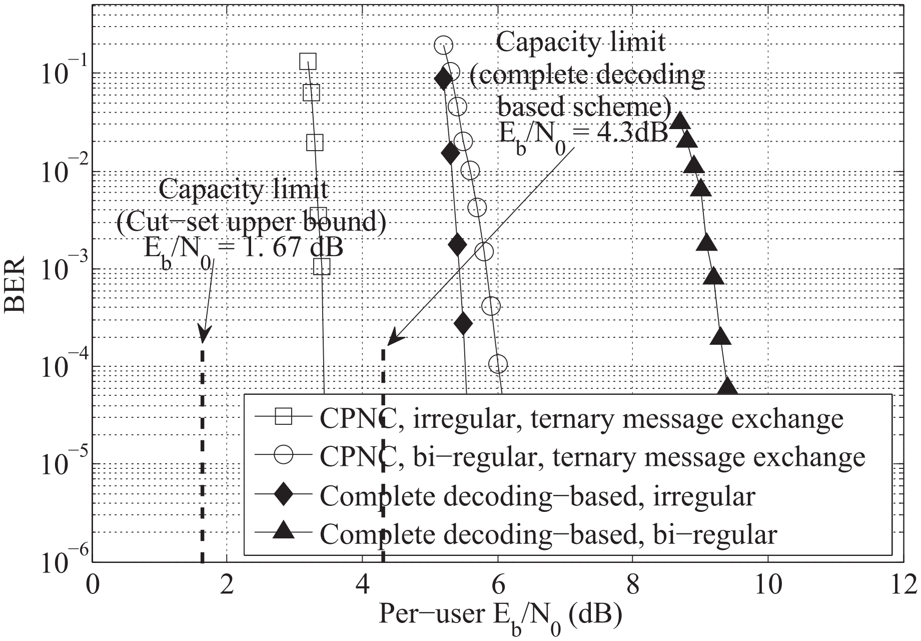


Figure 4.14 Comparison between the CPNC scheme and the complete decoding-based scheme where $R = 3/4$.

complete decoding scheme (considered in this chapter) with no iteration between the multi-user detector and the decoder. The performance loss of not using the IDD, however, can be up to several dB in power efficiency. For a fair comparison, in the scheme with complete decoding at the relay, we use the optimal power allocation between the two users [145, 149] and optimize its IRA code for degree distributions. The optimized code is provided in Table A.1.

In Fig. 4.14, we compare the performance of our IRA-PNC scheme and the complete decoding-based scheme, with $R = 3/4$. The optimized power allocation ratio for the complete decoding-based scheme is 3.2 at this code rate. The capacity limit for the complete decoding-based scheme is found to be at $E_b/N_0 = 4.3$ dB. The limit from the cut-set capacity upper bound[1] [27] of a Gaussian TWRC with binary inputs is found to be at $E_b/N_0 = 1.67$ dB. Note that the capacity limits for both schemes are for binary inputs with BPSK modulation. At BER = 10^{-4}, our

[1] The actual capacity limit of a binary-input Gaussian TWRC is still an open problem. Therefore, we use the upper bound as a reference.

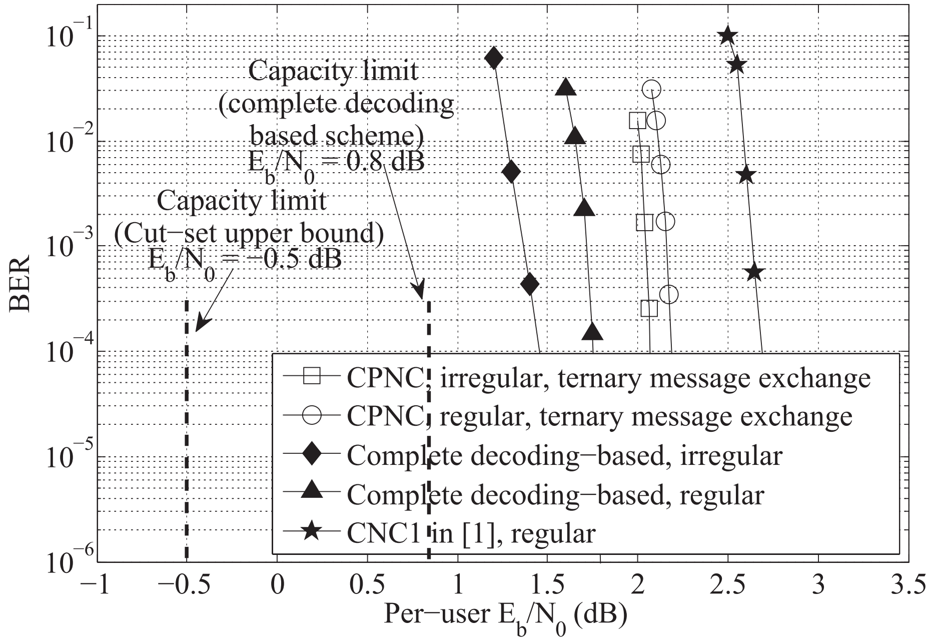


Figure 4.15 Comparison between CPNC scheme the complete decoding-based scheme where $R = 1/3$.

developed IRA-PNC scheme is about 1.7 dB away from the capacity upper bound. At BER = 10^{-4}, the IRA-PNC scheme is about 2 dB better than the optimized complete decoding-based scheme. Note that, for this case, our designed IRA-PNC scheme clearly outperforms the capacity limit of the complete decoding-based scheme.

In Fig. 4.15, we compare the performance of the IRA-PNC scheme and the complete decoding-based scheme where $R = 1/3$. The optimized power allocation ratio for the complete decoding-based scheme is 1.6 at this code rate. At BER = 10^{-4}, the optimized complete decoding-based scheme with IDD is about 0.6 dB better than the optimized IRA-PNC scheme. This shows that PNC with compute-and-forward may not be a good choice when the code rate is low. This is in line with the information theoretic result [27, 33], which shows that complete decoding-based scheme can outperform the CPNC scheme in terms of their achievable rates, as the SNR or coding rate becomes small. In Fig. 4.15, we also include the performance of the CNC1 scheme discussed in [46], which is equivalent to the complete decoding-based scheme but without iteration between the detector and decoders. The CNC1 suffers from a loss of about 1 dB relative to the complete decoding-based scheme with IDD. Due to this loss, the CNC1 scheme performs worse than the IRA-PNC scheme. We emphasize that when the complete decoding-based scheme is properly designed, it outperforms the CPNC scheme at $R = 1/3$.

4.7 Conclusion

In this chapter, we developed an IRA coded PNC scheme for binary input Gaussian TWRCs. We extended the EXIT chart technique to analyze the convergence behavior of the iterative computation process of the IRA-PNC scheme. Based on that, we optimized the degree distributions of the components of the IRA-PNC scheme. Our optimized IRA-PNC scheme significantly outperforms existing regular (or bi-regular) RA coded PNC schemes. We also showed that a CPNC scheme has the most significant benefit when the code rate is high. Interestingly, in a high coding rate regime, the performance improvement of using our EXIT curve-fitting to design an IRA coded PNC scheme is most significant. We also noted that, for a very low SNR or code rate, CPNC scheme is worse than the complete decoding based scheme with iterative multi-user detection and decoding. This agrees with existing information theoretic analysis results.

Chapter 5

Design of CPNC in MWRCs with Lattice Codes

5.1 Introduction

In Chapter 3 and Chapter 4, we focused on the analysis and design of binary coded PNC in two-way relay channels. In this chapter, we extend our study to design non-binary codes for a more general network, namely, multi-way relay channels (MWRC). In particular, we investigate the error performance of lattice network coded PNC scheme and its construction methods.

Nazer and Gastpar generalized the PNC scheme to a *compute-and-forward* scheme for multi-way relay networks [23]. This novel scheme utilizes structured nested lattice codes. The transmitted signals are lattice points in a multi-dimensional lattice over integers. Based on noisy observations of transmitted signals, the relay decodes and forwards an integer valued linear combination of transmitted signals to maximize the computation rate. It is shown that an asymptotic gain can be achieved from the information-theoretic perspective by a compute-and-forward scheme based on an infinite sequence of structured nested lattices.

In order to design and implement practical compute-and-forward schemes, a general algebraic framework, called *lattice network coding*, is developed in [21] for lattice partition based PNC schemes. Based on this framework, a variety of lattice network coding schemes can be constructed based on a principal ideal domain R in the complex field $\mathbb{C}$ [21, 165–167]. For a lattice network code (LNC) with a hypercube shaping region, such as in the Gaussian integer case, the probability of decoding error is derived in [21] by using the union bound estimation (UBE). This UBE implies that the choice of an optimal compute-and-forward coefficient vector over R abides by the *minimum variance criterion* of effective noise [21]. Consequently, the lattice reduction algorithms over Gaussian integers, such as the ones in [168] and [169] can be applied to find an optimal coefficient vector.

Motivated by the work in [21], we investigate the practical LNC design and decoding error performance analysis to Eisenstein integer based LNCs in this chapter. We start with the preliminary of the compute-and-forward model proposed in [23]. We then introduce the Eisenstein integer based lattice network codes. In the performance analysis, we will discuss the decoding error probability of Eisenstein integer based lattice network codes, and how to find the optimal scaling factor and optimal coefficient vector. Following that, we will discuss various designs of Eisenstein integer based lattice network codes. In the latter part of this chapter, we will focus on the construction of Eisenstein integer based lattice network codes over GF(2^2), which is more practical for real world implementation. We will show the simulation results in the end of this chapter for the error performance of our constructed codes.

The work in this chapter is a collaborated work with other researchers, and I am the sole student in this collaboration. My contribution in this work includes: propose detailed design and construction methods for Eisenstein integer based lattice network codes; conduct actual code design and search; compute related code parameters; conduct simulations to verify the performance of the designed codes.

5.2 Preliminaries: Computer-and-Forward

5.2.1 System Model

In this chapter, we consider a compute-and-forward scheme for a single relay system with L transmitters. We adopt the system model as in [23] and it is shown in Fig. 5.1. In the compute-and-forward scheme, each of L independent transmitters sends a message via a Gaussian multiple-access channel (MAC). Each message belongs to the message space W defined over a subring R of $\mathbb{C}$. For the transmitter l, it first maps the message $\mathbf{w}_l \in W$ to an n-dimensional complex-valued signal $\mathbf{x}_l$ by the *encoding function*

$$\mathcal{E} : W \to \mathbb{C}^n,$$

and then transmits it through the Gaussian MAC. The transmitted signal $\mathbf{x}_l$ is under the average energy constraint

$$\frac{1}{n} E[\|\mathcal{E}(\mathbf{w}_l)\|^2] \leq P,$$

where the message $\mathbf{w}_l$ is assumed to be uniformly selected in the message space W.

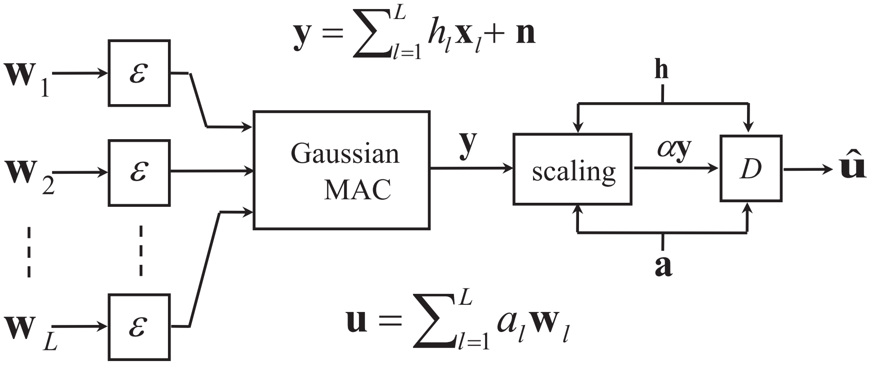


Figure 5.1 The compute-and-forward model

The received signal at the relay is given by a row vector

$$\mathbf{y} = \sum\nolimits_{l=1}^{L} h_l \mathbf{x}_l + \mathbf{n} \tag{5.1}$$

where h_l is the fading coefficient for the channel from transmitter l to the receiver at the relay and it follows a complex Gaussian distribution with mean zero and variance one. The channel vector $\mathbf{h} = (h_1, \cdots, h_L) \in \mathbb{C}^L$ is assumed to be known at the receiver but unknown at the transmitters. In (5.1), $\mathbf{n}$ represents a complex circularly-symmetric additive white Gaussian noise (AWGN) vector with zero mean and power spectrum density N_0. Here we define the signal-to-noise-ratio (SNR) as P/N_0, where P is the symbol energy at each transmitter.

Given a coefficient vector $\mathbf{a} \triangleq (a_1, \cdots, a_L)$ over R and a scaling factor $\alpha \in \mathbb{C}$, the goal of the relay receiver is to reliably decode an R-linear combination of the transmitted messages $\mathbf{u} = \sum_{l=1}^{L} a_l \mathbf{w}_l$ based on the scaled version of the received signal $\alpha\mathbf{y}$. Let $\mathcal{D}$: $\mathbb{C}^n \times \mathbb{C}^L \times \mathbb{C}^L \to W$ denote the decoding function and $\hat{\mathbf{u}}$ denote the decoded message $\mathcal{D}(\alpha\mathbf{y} \mid \mathbf{h}, \mathbf{a})$. Then, a pairwise decoding error occurs when $\hat{\mathbf{u}} \neq \mathbf{u}$. We denote the conditional pairwise decoding error probability by $P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a})$. Since the scaled received signal can be written as

$$\alpha\mathbf{y} = \sum_{l=1}^{L} a_l \mathbf{x}_l + \sum_{l=1}^{L} (\alpha h_l - a_l)\mathbf{x}_l + \alpha\mathbf{n},$$

the R-linear combination $\sum_{l=1}^{L} a_l \mathbf{x}_l$ has an *effective noise*

$$\mathbf{n}_{\text{eff}} \triangleq \sum_{l=1}^{L} (\alpha h_l - a_l)\mathbf{x}_l + \alpha\mathbf{n}.$$

5.2.2 Lattice Network Coding

We now give a brief review of the basic concept of lattice network codes (LNCs) [21].

Let $R \subset \mathbb{C}$ be a principal ideal domain (PID), which is a commutative ring such that

- Whenever $ab = 0$ for $a, b \in R$, either $a = 0$ or $b = 0$;
- Every ideal in R can be written as $aR = \{ar : r \in R\}$ for some $a \in R$.

Note that an ideal in a commutative ring R means a set of elements in R that is closed under addition and under multiplication by an arbitrary element in R. Well-known PIDs in $\mathbb{C}$ include the ring of integers $\mathbb{Z}$ and the ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$.

Definition 5.1. Let $N \leq n$. A subset Λ of $\mathbb{C}^n$ is called an N-dimensional R-lattice if it forms an R-module of rank N, that is, Λ is closed under addition and under multiplication by scalars in the ring R, and there are N linearly independent vectors $\mathbf{b}_1, ..., \mathbf{b}_N \in \Lambda$ such that $\Lambda = \{\sum_{1 \leq j \leq N} r_j \mathbf{b}_j : r_j \in R \ \ \forall j\}$. A subset Λ' of Λ is called a sublattice of Λ if it is an R-module.

Given an R-lattice Λ and a sublattice Λ' of Λ, the quotient group $\Lambda/\Lambda' = \{\boldsymbol{\lambda} + \Lambda' : \boldsymbol{\lambda} \in \Lambda\}$ naturally forms a partition of Λ. For an LNC, the message space is $W = \Lambda/\Lambda'$, which can also be regarded as an R-module. As a simple example, consider the PID $\mathbb{Z}$ of integers, which itself can be regarded as a 1-dimensional $\mathbb{Z}$-lattice. Every integer corresponds to a lattice point. The set $2\mathbb{Z}$ of even integers forms a sublattice of $\mathbb{Z}$, but the set of odd integers is not a sublattice of $\mathbb{Z}$ since it is not closed under multiplication by an even integer. The quotient group $\mathbb{Z}/2\mathbb{Z} = \{2\mathbb{Z}, 2\mathbb{Z}+1\}$ forms a partition of $\mathbb{Z}$ into two sets of lattice points, *i.e.*, the set of even integers and the set of odd integers.

Throughout this chapter, we shall assume that $|\Lambda/\Lambda'|$ has finite cardinality. The implementation of both encoding and decoding of an LNC involves a *lattice quantizer*. A lattice quantizer of a lattice Λ means a mapping $\mathcal{D}_\Lambda : \mathbb{C}^n \to \Lambda$, which sends a vector $\mathbf{x} \in \mathbb{C}^n$ to a nearest lattice point in Λ in terms of the Euclidean distance, that is,

$$\mathcal{D}_\Lambda(\mathbf{x}) \triangleq \operatorname{argmin}_{\boldsymbol{\lambda} \in \Lambda} \|\mathbf{x} - \boldsymbol{\lambda}\|. \tag{5.2}$$

The set of points in $\mathbb{C}^n$ that are mapped to a lattice point $\boldsymbol{\lambda} \in \Lambda$ by $\mathcal{D}_\Lambda$ is called the *Voronoi region* of $\boldsymbol{\lambda}$. The Voronoi region of the origin $\mathbf{0}$ of Λ is called the *fundamental Voronoi region* of Λ and it is denoted by $\mathcal{V}(\Lambda)$.

The encoding function $\mathcal{E} : W \to \Lambda$ of an LNC Λ/Λ' maps each coset $\boldsymbol{\lambda} + \Lambda'$ in Λ/Λ' to a coset leader, which is a lattice point contained both in the coset $\boldsymbol{\lambda} + \Lambda'$ and in the Voronoi region $\mathcal{V}(\Lambda')$ of the origin of Λ'. This Voronoi region $\mathcal{V}(\Lambda')$ is also called the *shaping region* of the LNC. Let $\bar{\varphi}$ be an embedding mapping from Λ/Λ' into Λ such that $\bar{\varphi}(\boldsymbol{\lambda} + \Lambda')$ and $\boldsymbol{\lambda}$ are in the same coset $\boldsymbol{\lambda} + \Lambda'$. Then, $\mathcal{E}$ can be represented as

$$\mathbf{x}_l = \mathcal{E}(\mathbf{w}_l) = \bar{\varphi}(\mathbf{w}_l) - \mathcal{D}_{\Lambda'}(\bar{\varphi}(\mathbf{w}_l)).$$

The image of $\mathcal{E}$ is also referred to as *constellations* of the LNC.

The decoding function of an LNC is described by

$$\hat{\mathbf{u}} = \mathcal{D}(\alpha\mathbf{y} \mid \mathbf{h}, \mathbf{a}) = \varphi(\mathcal{D}_\Lambda(\alpha\mathbf{y})), \tag{5.3}$$

where φ is the natural projection mapping from Λ onto Λ/Λ' via $\varphi(\boldsymbol{\lambda}) = \boldsymbol{\lambda} + \Lambda'$. It has been shown in [21] that the decoding error probability $P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) = \Pr[\mathcal{D}_\Lambda(\mathbf{n}) \notin \Lambda']$. When the shaping region of an LNC is a (rotated) hypercube in $\mathbb{C}^n$, a union bound estimation (UBE) of $P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a})$ is derived in [21] in terms of the minimum inter-coset distance $d(\Lambda/\Lambda')$, which is equal to the length of shortest vectors in the complement $\Lambda \backslash \Lambda'$ of Λ' in Λ, and the number $K(\Lambda/\Lambda')$ of these shortest vectors. One consequence of this UBE is that in hypercube shaped LNCs, the choice of both the optimal scaling factor α_{opt} and an optimal coefficient vector $\mathbf{a}_{\text{opt}}$ is prescribed by the *minimum variance criterion* of effective noise [21].

5.3 Fundamentals on Eisenstein Integer Based Lattice Network Codes

5.3.1 Eisenstein Integers

In this section we first introduce some basic algebraic and geometric properties of Eisenstein integers. Let $\omega = \frac{-1+\sqrt{-3}}{2}$. A complex number is called an Eisenstein integer when it can be written in the form of $a + b\omega$, where a, b are integers. The ring $\mathbb{Z}[\omega]$ of Eisenstein integers not only forms a PID, but also a Euclidean domain [171]. The norm of an Eisenstein integer $a + b$ is equal to $|a + b\omega|^2$, i.e., the squared absolute value, and it can be represented by $a^2 + b^2 - ab$. In consequence, we have the following two properties [172].

- There are six units in $\mathbb{Z}[\omega], i.e., \pm 1, \pm\omega, \pm(1+\omega)$;
- An Eisenstein integer $a+b\omega$ is prime in $\mathbb{Z}[\omega]$, iff it is the product of an Eisenstein unit and a rational prime congruent to 2 modulo 3, or $|a+b\omega|^2$ is a rational prime.

Note that an element p in a PID R is said to be prime if it is not a unit, and for any $r, s \in R$ such that p divides rs, p divides either r or s.

The reasons for investigation of Eisenstein integer based LNCs are highlighted as follows. First, since the Voronoi region of $\mathbb{Z}[\omega]$ is a regular hexagon, the analysis in [21] based on hypercube shaping does not apply to the case of $\mathbb{Z}[\omega]$. In addition, note that if a finite field $\mathbb{F}_q$ is representable by lattice partitions over $\mathbb{Z}[\omega]$, then q is either equal to 3, or congruent to 1 modulo 6, or the square of a rational prime that is congruent to 2 modulo 3. On the other hand, if $\mathbb{F}_q$ can be represented by lattice partitions over $\mathbb{Z}[i]$, then q is either equal to 2, or congruent to 1 modulo 4, or the square of a rational prime that is congruent to 3 modulo 4. Therefore, lattice partitions over $\mathbb{Z}[\omega]$ enrich the candidates of finite fields for LNC design. For example, $\mathbb{F}_4$ can be represented by the lattice partition $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ but not $\mathbb{Z}[i]/2\mathbb{Z}[i]$. Actually, $\mathbb{F}_2 \cong \mathbb{Z}/2\mathbb{Z}$ is the only field of characteristic 2 that can be represented by $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$, where $\beta \in \mathbb{Z}[i]$.

Furthermore, as illustrated in the next example, even if the $\mathbb{Z}[i]$- and $\mathbb{Z}[\omega]$-based message spaces are (field) isomorphic to each other, their constellations do not form a linear bijection. Therefore, even if a same linear code is adopted to construct a $\mathbb{Z}[i]$-based and a $\mathbb{Z}[\omega]$-based LNC, different constellations of the message spaces need to be considered.

Example 5.1. A finite field $\mathbb{F}_q$ is isomorphic to both $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ and $\mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$ for some $\beta \in \mathbb{Z}[i]$ and $\gamma \in \mathbb{Z}[\omega]$ only if q is congruent to 1 modulo 12. Write $\beta = 2+3i$ and $\gamma = 4+3\omega$. Consider a message space $W = \mathbb{F}_{13} \cong \mathbb{Z}[i]/\beta\mathbb{Z}[i] \cong \mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$. Fig. 5.2(a) and Fig. 5.2(b) depict the sublattices $\beta\mathbb{Z}[i]$ of $\mathbb{Z}[i]$ and $\gamma\mathbb{Z}[\omega]$ of $\mathbb{Z}[\omega]$, respectively. The fundamental Voronoi regions of $\beta\mathbb{Z}[i]$ or $\gamma\mathbb{Z}[\omega]$ and constellations of

$\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ or $\mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$ are also shown in Fig. 5.2(a) or Fig. 5.2(b). Observe that

$$\mathcal{E}_{\mathbb{Z}[i]}(W) = \{0, \pm 1, \pm i, \pm(1+i), \pm(1-i), \pm 2, \pm 2i\}$$

and

$$\mathcal{E}_{\mathbb{Z}[\omega]}(W) = \{0, \pm 1, \pm\omega, \pm(1+\omega), \pm(1-\omega), \pm(1+2\omega), \pm(2+\omega)\}.$$

Since both elements 1 and 2 are in $\mathcal{E}_{\mathbb{Z}[i]}(W)$ whereas there are no two elements in $\mathcal{E}_{\mathbb{Z}[\omega]}(W)$ such that one is twice the other, for any linear bijection $\phi : \mathbb{Z}[i] \to \mathbb{Z}[\omega]$, $\phi(\mathcal{E}_{\mathbb{Z}[i]}(W)) \neq \mathcal{E}_{\mathbb{Z}[\omega]}(W)$.

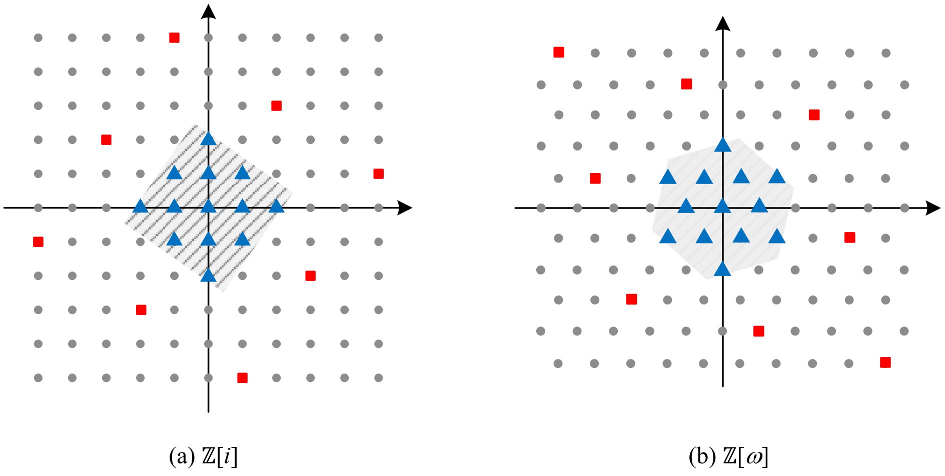


Figure 5.2 Different constellations of a message space $W = \mathbb{F}_{13}$. In diagram (a), $W \cong \mathbb{Z}[i]/\beta\mathbb{Z}[i]$, where $\beta = 2 + 3i$. In diagram (b), $W \cong \mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$, where $\gamma = 4 + 3\omega$. In either diagram, squares represent elements in the sublattice, triangles represent the constellation points, and the shaded region is the Voronoi region of the origin of the sublattice.

Since $\mathbb{Z}[\omega]$ is a PID, every sublattice of $\mathbb{Z}[\omega]$, which is essentially an ideal, is generated by a nonzero Eisenstein integer. Since the six Eisenstein units are those nonzero elements in $\mathbb{Z}[\omega]$ closest to the origin, the Voronoi region of every sublattice of $\mathbb{Z}[\omega]$ is also a hexagon.

An R-lattice Λ_1 is said to be equivalent to another R-lattice Λ_2 when Λ_1 can be obtained from Λ_2 by possibly rotation, reflection and change of scale [170]. In the remainder of the paper, for all Λ/Λ', the lattice Λ' is assumed to be equivalent to either $\mathbb{Z}[i]^N$ or $\mathbb{Z}[\omega]^N$, that is, the shaping region of the LNC is either a hypercube or a product of regular hexagons.

Proposition 5.1. *Assume that Λ' is equivalent to $\mathbb{Z}[\omega]^N$ with a scaling factor γ. The volume $V(\Lambda')$ of $\mathcal{V}(\Lambda')$ is $(\frac{\sqrt{3}}{2}|\gamma|^2)^N$. Moreover, via continuous approximation [173], the average power for the constellation of Λ/Λ' is $\frac{5N}{36n}|\gamma|^2$.*

Proof. $V(\Lambda') = \int_{\mathcal{V}((\gamma\mathbb{Z}[\omega])^N)} d\mathbf{x} = \left(\int_{\mathcal{V}(\gamma\mathbb{Z}[\omega])} dx\right)^N = (\frac{\sqrt{3}}{2}|\gamma|^2)^N$. Via continuous approximation, the average power for Λ/Λ' can be approximated by the average power of a random vector $\mathbf{x}$ uniformly distributed over $\mathcal{V}(\Lambda')$. Thus,

$$\begin{aligned}
\frac{1}{n}E[\|\mathcal{E}(\mathbf{w}_l)\|^2] &= \frac{\int_{\mathcal{V}(\gamma\mathbb{Z}[\omega])^N} \|\mathbf{x}\|^2 d\mathbf{x}}{nV(\Lambda')} \\
&= \frac{N\int_{\mathcal{V}(\gamma\mathbb{Z}[\omega])} |x|^2 dx \int_{\mathcal{V}(\gamma\mathbb{Z}[\omega])^{N-1}} d\mathbf{x}}{n\int_{\mathcal{V}(\gamma\mathbb{Z}[\omega])^N} d\mathbf{x}} \\
&= \frac{N\frac{5\sqrt{3}}{72}|\gamma|^4}{n\frac{\sqrt{3}}{2}|\gamma|^2} \\
&= \frac{5N}{36n}|\gamma|^2.
\end{aligned}$$

■

The two basic attributes of LNCs over $\mathbb{Z}[\omega]$ in Proposition 5.1 are important for the analysis of decoding error probability in Section 5.4. In comparison, for an LNC Λ/Λ' over $\mathbb{Z}[i]$, where Λ' is equivalent to $\mathbb{Z}[i]^N$ with a scaling factor γ, the volume of $\mathcal{V}(\Lambda')$ is 1 and the average power for the message space Λ/Λ' is $\frac{N}{6n}|\gamma|^2$.

5.3.2 Lattice Quantization and Encoding Over $\mathbb{Z}[\omega]$

In the design of an efficient encoder and decoder for a general $\mathbb{Z}[\omega]$-based LNC scheme, the 1-dimensional baseline case plays a fundamental role. In the following, we will focus on the 1-dimensional baseline system. In particular, we now introduce the quantization over Λ and Λ', as well as encoding of Λ/Λ', where $\Lambda = \mathbb{Z}[\omega]$ and $\Lambda' = \gamma\mathbb{Z}[\omega]$, $\gamma \in \mathbb{Z}[\omega]$.

The quantization $\mathcal{D}_\Lambda$ of a complex value x to an Eisenstein integer can be done as follows. Note that the lattice points in $\mathbb{Z}[\omega]$ can be divided into two sets:

1. $\mathbb{Z}[\sqrt{-3}] = \{a + \sqrt{-3}b : a, b \in \mathbb{Z}\}$;

2. $\omega + \mathbb{Z}[\sqrt{-3}] = \{\omega + a + \sqrt{-3}b : a, b \in \mathbb{Z}\}$.

Let

$$\beta_1 = \lfloor \mathrm{Re}\{x\} \rceil + \sqrt{-3} \lfloor \mathrm{Im}\{x\}/\sqrt{3} \rceil \tag{5.4}$$

$$\beta_2 = \lfloor \mathrm{Re}\{x - \omega\} \rceil + \sqrt{-3} \lfloor \mathrm{Im}\{x - \omega\}/\sqrt{3} \rceil + \omega \tag{5.5}$$

where $\lfloor \cdot \rceil$ denotes rounding to nearest integer. The Eisenstein integers β_1 and β_2 are, respectively, a nearest point in $\mathbb{Z}[\sqrt{-3}]$ and $\omega + \mathbb{Z}[\sqrt{-3}]$ to x in terms of the Euclidean distance. The quantizer $\mathcal{D}_\Lambda$ then maps x to the one in $\{\beta_1, \beta_2\}$ which is closer to x.

Based on the quantizer $\mathcal{D}_\Lambda$, the quantization of a complex value x over Λ' can be realized by

$$\mathcal{D}_{\Lambda'} = \gamma \mathcal{D}_\Lambda(\gamma^{-1} x),$$

where the inverse of γ is taken in $\mathbb{C}$.

Since $\mathbb{Z}[\omega]$ is a Euclidean domain, it has a division algorithm such that every $\lambda \in \mathbb{Z}[\omega]$ can be written as $q\gamma + r$ with $q, r \in \mathbb{Z}[\omega]$ and $|r| < |\gamma|$. Thus, the design of the encoding function $\mathcal{E} : W \to \Lambda$ is equivalent to propose an appropriate division algorithm such that

1. There is a unique output remainder for all elements in a same coset divided by γ;

2. The remainder is in the fundamental Voronoi region of $\gamma\mathbb{Z}[\omega]$.

We next propose one such possible division algorithm, which is adapted from the one in [174].

Algorithm 5.1 Division Algorithm

```
Initialization: Given $\lambda, \gamma \in \mathbb{Z}[\omega]$, the following routine outputs a unique remainder
    $r \in \mathbb{Z}[\omega]$ of $\lambda'$ divided by $\gamma$ for any $\lambda' \in \lambda + \gamma\mathbb{Z}[\omega]$.
1: Compute the rational number $x = \lambda/\gamma$;
2: Compute the nearest Eisenstein integer $\beta_1$ (resp. $\beta_2$) in $\mathbb{Z}[\sqrt{-3}]$ (resp. $\omega + \mathbb{Z}[\sqrt{-3}]$) to $x$ by (5.4) (resp. by (5.5));
3: Let $r_1$ denote $\lambda - \beta_1\gamma$ and $r_2$ denote $\lambda - \beta_2\gamma$;
4: $r = r_1$ if either $|r_1| < |r_2|$ or $|r_1| = |r_2|$ and $\mathrm{Re}\{\beta_1\} < \mathrm{Re}\{\beta_2\}$;
5: $r = r_2$ if either $|r_2| < |r_1|$ or $|r_1| = |r_2|$ and $\mathrm{Re}\{\beta_2\} < \mathrm{Re}\{\beta_1\}$.
```

After Step (3) of Algorithm 5.1, the computed r_1 and r_2 are, respectively, the Eisenstein integers in $\gamma\mathbb{Z}[\sqrt{-3}]$ and in $\gamma(\mathbb{Z}[\sqrt{-3}]+\omega)$ closest to λ. Because $|r_1|, |r_2| < |\gamma|$ and $|\beta_1 - \beta_2| \geq |\gamma|$ for all $\beta_1 \in \gamma[\mathbb{Z}\sqrt{-3}]$ and $\beta_2 \in \gamma(\mathbb{Z}[\sqrt{-3}]+\omega)$, it can be shown that when $|r_1| \neq |r_2|$, either r_1 or r_2 is in $\mathcal{V}(\gamma\mathbb{Z}[\omega])$, and when $|r_1| = |r_2|$, both r_1 and r_2 are in $\mathcal{V}(\gamma\mathbb{Z}[\omega])$. Thus, when $|r_1| \neq |r_2|$, the one in r_1 and r_2 with smaller norm is the unique output r for all inputs belonging to $\lambda + \gamma\mathbb{Z}[\omega]$. When $|r_1| = |r_2|$, since $\mathrm{Re}\{\beta_1\} \neq \mathrm{Re}\{\beta_2\}$ for all $\beta_1 \in \mathbb{Z}[\sqrt{-3}]$ and $\beta_2 \in \mathbb{Z}[\sqrt{-3}]+\omega$, the selection rules in steps (iv) and (v) of the algorithm guarantee the unique output r for all inputs belonging to $\lambda + \gamma\mathbb{Z}[\omega]$.

It is interesting to point out that the computational complexity of quantization and constellations over $\mathbb{Z}[\omega]$ is in the same order as those over $\mathbb{Z}[i]$.

For a $\mathbb{Z}[\omega]$-based LNC, after the decoding error probability is derived in next section, we shall see that the minimum variance criterion of effective noise also applies to choosing the optimal scaling factor α_{opt} and an optimal coefficient vector $\mathbf{a}_{\text{opt}}$.

5.4 Performance Analysis

5.4.1 Decoding Error Probability

In this section, we analyze the probability of decoding error for the compute-and-forward scheme with $\mathbb{Z}[\omega]$-Based LNCs. Recall that the error probability of a general LNC system is shown in [21] and it is equal to $\Pr[\mathcal{D}_\Lambda(\mathbf{n}) \notin \Lambda']$, where $\mathcal{D}_\Lambda$ is a lattice quantizer and $\mathbf{n}_{\text{eff}}$ is the effective noise $\sum_{l=1}^{L} (\alpha h_l - a_l)\mathbf{x}_l + \alpha\mathbf{n}$. As a counterpart of the error probability upper bound of hypercube shaped LNCs derived in [21], we have the following theorem for the error probability upper bound of LNCs with shaping regions to be product of regular hexagons.

Theorem 5.1. *Consider an LNC Λ/Λ' in which $\Lambda' \subset \mathbb{C}^n$ is equivalent to $\mathbb{Z}[\omega]^N$. The union bound estimation (UBE) on the probability of decoding error for Λ/Λ' is*

$$P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \lessapprox K(\Lambda/\Lambda') \exp\left(-\frac{d^2(\Lambda/\Lambda')}{4N_0 Q(\alpha, \mathbf{a})}\right) \tag{5.6}$$

where $d(\Lambda/\Lambda')$ is the length of shortest vectors in $\Lambda\backslash\Lambda'$, $K(\Lambda/\Lambda')$ is the number of these shortest vectors, and $Q(\alpha, \mathbf{a}) = |\alpha|^2 + \frac{n}{N}\text{SNR}\|\alpha\mathbf{h} - \mathbf{a}\|^2$. Moreover, when $n = N$, $N_0 Q(\alpha, \mathbf{a})$ represents the variance of uncorrelated components in the effective noise vector $\mathbf{n}_{\text{eff}}$.

The proof of the theorem is given in Appendix C.1. Its main flow is analogous to the one used in [21] but the detailed derivations require careful manipulation of the algebraic and geometric properties of Eisenstein integers. It is worthwhile to note that the UBE in this theorem holds when $N \leq n$. In comparison, the UBE on the decoding error probability for hypercube shaped LNCs in [21] is under the assumption $N = n$. The UBE for hypercube shaped LNCs over $\mathbb{Z}[i]$ could be extended to be applicable over the case $N < n$ by the same technique adopted here. Moreover, to show the variance of each component in $\mathbf{n}_{\text{eff}}$ equal to $N_0 Q(\alpha, \mathbf{a})$ when $N = n$, the theorem above makes use of the uncorrelated property. When $N < n$, the components in $\mathbf{n}_{\text{eff}}$ are not necessarily uncorrelated, and hence $N_0 Q(\alpha, \mathbf{a})$ does not necessarily represent the variance of components in $\mathbf{n}_{\text{eff}}$. It is also intriguing to note that the UBE on the decoding error probability is in the same form for the hypercube shaped case and for the case with shaping region to be product of regular hexagons.

As the constellations in a $\mathbb{Z}[\omega]$-based LNC is different from that in a $\mathbb{Z}[i]$-based one, in the following, we shall give a general representation of the UBE on the decoding error probability. This general representation will highlight the performance improvement of $\mathbb{Z}[\omega]$-based LNCs over $\mathbb{Z}[i]$-based ones in terms of the coding gain and shaping gain.

Define the *nominal coding gain* of an LNC in a similar way as in [173] for a lattice code:

$$\gamma_c(\Lambda/\Lambda') = \frac{d^2(\Lambda/\Lambda')}{V(\Lambda)^{1/n}}$$

where $V(\Lambda)$ is the volume of the fundamental Voronoi region $\mathcal{V}(\Lambda)$. Let $\mathbf{M}$ be an $N \times n$ generator matrix for Λ. According to [170], when Λ is a $\mathbb{Z}[i]$-lattice, $V(\Lambda) = \det(\mathbf{M}\mathbf{M}^{\mathrm{H}})$, and when Λ is a $\mathbb{Z}[\omega]$-lattice, $V(\Lambda) = \left(\frac{\sqrt{3}}{2}\right)^N \det(\mathbf{M}\mathbf{M}^{\mathrm{H}})$. The nominal coding gain measures the increase in density of Λ over the baseline Gaussian integer lattice $\mathbb{Z}[i]$. As Λ' is a sublattice of Λ, we may consider Λ' to be a coarse lattice of the fine lattice Λ. The second moment or average energy per dimension of a uniform distribution over the fundamental Voronoi region $\mathcal{V}(\Lambda')$ is $P(\Lambda')$, and its normalized second moment is $G(\Lambda') = P(\Lambda')/[V(\Lambda')^{1/n}]$. For the Voronoi region of $\mathbb{Z}[i]$, the normalized second moment is $1/6$. The *shaping gain* of the region $\mathcal{V}(\Lambda')$ is defined as [173] $\gamma_s(\Lambda') = \frac{1/6}{G(\Lambda')} = \frac{V(\Lambda')^{1/n}}{6P(\Lambda')}$. By continuous approximation, $P(\Lambda')$ is regarded as the average power for the message space Λ/Λ', so we can write $\gamma_s(\Lambda/\Lambda') = \gamma_s(\Lambda')$. The shaping gain measures how much less is the average energy of $\mathcal{V}(\Lambda')$ relative to a hypercube centered at the origin.

Now the UBE on the probability of decoding error for the LNC Λ/Λ' can be approximated using the following corollary.

Corollary 5.1. *Let R denote either $\mathbb{Z}[\omega]$ or $\mathbb{Z}[i]$. Consider an LNC Λ/Λ' in which the N-dimensional lattice Λ' is equivalent to R^N. We have the UBE for the LNC*

Λ/Λ':

$$P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \lessapprox K(\Lambda/\Lambda') \exp\left(-\frac{\gamma_c(\Lambda/\Lambda')\gamma_s(\Lambda/\Lambda')3\mathrm{SENR}_{\mathrm{norm}}}{2}\right)$$

where the ***normalized signal-to-effective-noise ratio*** *(*$\mathrm{SENR}_{\mathrm{norm}}$*) is defined as*

$$\mathrm{SENR}_{\mathrm{norm}} \triangleq \frac{\mathrm{SENR}}{2^\rho} \triangleq \frac{\frac{1}{n}E[\|\mathbf{x}_l\|^2]}{2^\rho N_0 Q(\alpha, \mathbf{a})} \approx \frac{P(\Lambda')}{2^\rho N_0 Q(\alpha, \mathbf{a})}$$

and ρ measures the spectral efficiency *or the* rate *of the code, and it is given by*

$$\rho = \frac{1}{n}\log_2 |\Lambda/\Lambda'| = \frac{1}{n}\log_2 \frac{V(\Lambda')}{V(\Lambda)}.$$

For the $\mathbb{Z}[i]$-based baseline LNC, that is, $\Lambda = \mathbb{Z}[i]^n$ and $\Lambda' = (\beta\mathbb{Z}[i])^n$, where β is a nonzero Gaussian integer, both the nominal coding gain and the shaping gain are equal to 1 (0 dB). In comparison, for the $\mathbb{Z}[\omega]$-based baseline LNC, that is, $\Lambda = \mathbb{Z}[\omega]^n$ and $\Lambda' = (\beta\mathbb{Z}[\omega])^n$, where β is a nonzero Eisenstein integer, by Proposition 5.1, the nominal coding gain can be calculated to be $2\sqrt{3}/3$ (0.625 dB), and the shaping gain $\gamma_s(\Lambda/\Lambda') = \frac{1/6}{G(\Lambda')} = 3\sqrt{3}/5$ (0.167 dB).

5.4.2 Optimal Scaling Factor and Optimal Coefficient Vector

According to Theorem 5.1, for an LNC Λ/Λ' over $\mathbb{Z}[\omega]$, different selection of scaling factor α and coefficient vector $\mathbf{a}$ at the relay will yield different decoding error probabilities. Therefore, in order to obtain the optimal performance of Λ/Λ' in terms of decoding error probability, we discuss the optimal selection of α and $\mathbf{a}$ in this subsection.

For hypercube shaped LNCs, in order to minimize the decoding error probability, the choice of the scaling factor α and coefficient vector $\mathbf{a}$ is prescribed by the *minimum variance criterion* [21], that is, to minimize the variance $N_0Q(\alpha, \mathbf{a})$ of components in the effective noise vector $\mathbf{n}_{\mathrm{eff}}$. Consequently, the optimal scaling factor is equal to

$$\alpha_{\mathrm{opt}} = \frac{\mathbf{a}\mathbf{h}^{\mathrm{H}}\mathrm{SNR}}{\mathrm{SNR}\|\mathbf{h}\|^2 + 1}, \tag{5.7}$$

which was first derived in [23] as the MMSE coefficient to maximize the computation rate, and finding an optimal coefficient vector $\mathbf{a}_{\text{opt}}$ is equivalent to solving the shortest vector problem

$$\mathbf{a}_{\text{opt}} = \arg\min_{\mathbf{a}\neq\mathbf{0}} \|\mathbf{a}\mathbf{L}\|, \tag{5.8}$$

where $\mathbf{L}$ is a lower triangular matrix (over $\mathbb{C}$) such that

$$\mathbf{L}\mathbf{L}^{\text{H}} = \text{SNR}\mathbf{I}_L - \frac{\text{SNR}^2}{\text{SNR}\|\mathbf{h}\|^2+1}\mathbf{h}^{\text{H}}\mathbf{h}.$$

(See [21]) Thus, we can apply any lattice reduction algorithms, such as the Gaussian reduction algorithm over $\mathbb{Z}[i]$ in [175] for 2-dimensional case as well as the approximate suboptimal algorithms in [168] and [169] for higher dimensional cases, to solving (5.8) over $\mathbb{Z}[i]$.

Consider an LNC Λ/Λ' in which the N-dimensional lattice Λ' is equivalent to $\mathbb{Z}[\omega]^N$. Justified by Theorem 5.1, the minimum variance criterion also applies to Λ/Λ' so as to minimize the decoding error probability. Hence, for a $\mathbb{Z}[\omega]$-based LNC, α_{opt} can also be calculated by (5.7), and $\mathbf{a}_{\text{opt}}$ can be found by solving (5.8) too.

To the best of our knowledge, there is no efficient lattice reduction algorithm to find an exact solution for the shortest vector problem (5.8) over $\mathbb{Z}[\omega]$. The generalized LLL algorithm proposed by Napias in [176] can be applied to yield an approximate solution for (5.8) over $\mathbb{Z}[\omega]$. For the special case $L = 2$, since an exact solution of (5.8) can be efficiently found by the Gaussian reduction algorithm for lattices over $\mathbb{Z}[i]$ (See [175]), it is natural to ask whether an exact solution of (5.8) can also be efficiently found for lattices over $\mathbb{Z}[\omega]$. In the remaining part of this section, we shall give an affirmative answer by extending the Gaussian reduction algorithm to work for $\mathbb{Z}[\omega]$-based lattices.

Recall that given two vectors $\mathbf{v}_1, \mathbf{v}_2$ over the field $\mathbb{R}$ of real numbers, the Gaussian reduction algorithm for real lattices over $\mathbb{Z}$ (See [22] for example) will yield a shortest vector $\mathbf{u}_1$ in the $\mathbb{Z}$-lattice generated by $\mathbf{v}_1$ and $\mathbf{v}_2$, as well as another vector $\mathbf{u}_2$ which has the shortest length in the same lattice excluding those vectors generated by $\mathbf{u}_1$. Thus, if the lower triangular matrix $\mathbf{L}$ in (5.8) is over $\mathbb{R}$, then an optimal solution

$\mathbf{a}_{\text{opt}}$ of (5.8) over $\mathbb{Z}$ could be found by applying the Gaussian reduction algorithm to row vectors in $\mathbf{L}$ and then taking $\mathbf{a}_{\text{opt}} = \mathbf{u}_1 \cdot \mathbf{L}^{-1}$.

To the best of our knowledge, there is no reference with sufficient details to justify whether the Gaussian reduction algorithm can be adapted for complex lattices over Eisenstein integers. With formal justification, we next generalize the Gaussian reduction algorithm for real lattices over $\mathbb{Z}$ to be applicable to complex lattices over $\mathbb{Z}[\omega]$, so that it can induce an optimal coefficient vector $\mathbf{a}_{\text{opt}}$ over $\mathbb{Z}[\omega]$. Recall that $\mathcal{D}_\Lambda$ denotes a quantizer of a complex lattice Λ subject to (5.2).

Algorithm 5.2 Gaussian Reduction for a Complex $\mathbb{Z}[\omega]$-Lattice

Initialization: Given a basis $(\mathbf{v}_1, \mathbf{v}_2)$ of a 2-dimensional complex lattice Λ over $\mathbb{Z}[\omega]$, where $\|\mathbf{v}_1\| \leq \|\mathbf{v}_2\|$, the following routine returns a shortest nonzero vector $\mathbf{u}_1$ in Λ and a shortest vector $\mathbf{u}_2$ in $\Lambda \backslash \{\beta \mathbf{u}_1 : \beta \in \mathbb{Z}[\omega]\}$.

```
Set $\mathbf{u}_1 := \mathbf{v}_1$, $\mathbf{u}_2 := \mathbf{v}_2$, and finished := 0;
while finished == 0 do
    Set $\mathbf{u}'_2 := \mathbf{u}_2 - \mathcal{D}_{\mathbb{Z}[\omega]}(\mathbf{u}_1\mathbf{u}_2^{\text{H}}/\|\mathbf{u}_1\|^2)\mathbf{u}_1$;
    if $\|\mathbf{u}_1\| > \|\mathbf{u}'_2\|$ then
        Set $\mathbf{u}_2 := \mathbf{u}_1$, $\mathbf{u}_1 := \mathbf{u}'_2$;
    else
        Set $\mathbf{u}_2 := \mathbf{u}'_2$, finished := 1;
    end if
end while
```

In the algorithm above, the magnitude of vector $\mathbf{u}_1$ is strictly decreasing in each iteration. As there are finitely many lattice points in Λ with magnitude smaller than or equal to $\|\mathbf{v}_1\|$, the algorithm terminates in finitely many steps. Moreover, $\{\mathbf{u}_1, \mathbf{u}_2\}$ keeps to be a basis of Λ throughout the algorithm, because each iteration updates $\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]$ by either setting

$$\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} := \begin{bmatrix}-\mathcal{D}_{\mathbb{Z}[\omega]}(\mathbf{u}_1\mathbf{u}_2^{\text{H}}/\|\mathbf{u}_1\|^2) & 1\\ 1 & 0\end{bmatrix} \cdot \begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix}$$

or

$$\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} := \begin{bmatrix} 1 & 0 \\ -\mathcal{D}_{\mathbb{Z}[\omega]}(\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2) & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

where the 2 × 2 matrices have determinants ±1.

It will be justified in Appendix C.2 that the output $\mathbf{u}_1$ of Algorithm 5.2 is a shortest nonzero vector in Λ and the output vector $\mathbf{u}_2$ is a shortest vector in $\Lambda\backslash\{\beta\mathbf{u}_1 : \beta \in \mathbb{Z}[\omega]\}$.

5.5 Construction of $\mathbb{Z}[\omega]$-Based LNCs with General Lattice Partition

For both $\mathbb{Z}[\omega]$-based and $\mathbb{Z}[i]$-based LNCs Λ/Λ', the nominal coding gain

$$\gamma_c(\Lambda/\Lambda') = \frac{d^2(\Lambda/\Lambda')}{V(\Lambda)^{1/n}}$$

and the shaping gain

$$\gamma_s(\Lambda/\Lambda') = \frac{V(\Lambda')^{1/n}}{6P(\Lambda')}$$

are two important parameters in UBE of decoding error probabilities. In $\mathbb{Z}[\omega]$ case, the nominal coding gain and the shaping gain of a baseline LNC are, respectively, 0.625 dB and 0.167 dB, in contrast to the $\mathbb{Z}[i]$ case, where both gains are 0 dB. In this section, we will discuss constructions of LNCs from linear codes. Let π be a prime element in a PID $R \subset \mathbb{C}$. Given an $[n,\, k]$ linear code C over the field $R/\pi R$, denote by $w_E^{min}(C)$ the minimum squared Euclidean norm of non-zero codewords of C.

5.5.1 LNC From a Linear Code by Complex Construction A

Complex Construction A [170] is a method to construct a complex R-lattice. It is adapted in [21] to design an LNC Λ/Λ' as follows: Λ is constructed from C by Complex Construction A and Λ' is set to be $(\pi R)^n$. In our opinion, to adapt Complex Construction A in the framework of lattice network coding, it would be natural to set the coarse lattice Λ' in a more relaxed manner, that is, in the form $(\pi^r R)^n$, where $r \geq 1$.

Algorithm 5.3 Complex Construction A

```
1: Consider a linear code $C$ of length $n$ over the finite field $R/\pi R$. An LNC $\Lambda/\Lambda'$
   can be constructed by Complex Construction A via
       $\Lambda = \{\boldsymbol{\lambda} \in R^n : \sigma(\boldsymbol{\lambda}) \in C\}$
   where $\sigma$ is the natural projection from $R^n$ to $(R/\pi R)^n$, and $\Lambda' = (\pi^r R)^n$, where
   $r \geq 1$.
```

Algorithm 5.3 gives a lattice partition Λ/Λ' from a linear code C. It is worthwhile to note that this lattice partition can also be got from Complex Construction D in [21] by setting a sequence of nested linear codes $C_0 \supseteq C_1 \supseteq ... \supseteq C_{r-1} \supseteq C_r$ over $R/\pi R$ with $C_0, ..., C_{r-1}$ as $[n, n]$ trivial codes and C_r as the given code C in Algorithm 5.3. Now we specify the generator matrices for lattices Λ and Λ' and their relationship.

Proposition 5.2. *Let Λ/Λ' be the LNC constructed by Algorithm 5.3 from a linear code C over $R/\pi R$. Let $[I_k \ B_{k\times(n-k)}]$ be a $k \times n$ matrix over R such that $\sigma([I_k \ B_{k\times(n-k)}])$ is a generator matrix for C. The respective generator matrices $\mathbf{M}_\Lambda$ for Λ and $\mathbf{M}_{\Lambda'}$ for Λ' can be described by*

$$\mathbf{M}_\Lambda = \begin{bmatrix} I_k & B_{k\times(n-k)} \\ \mathbf{0} & \pi I_{n-k} \end{bmatrix}$$

and

$$\mathbf{M}_{\Lambda'} = \begin{bmatrix} \pi^r I_k & \pi^r B_{k\times(n-k)} \\ \mathbf{0} & \pi^r I_{n-k} \end{bmatrix}.$$

Since

$$\mathbf{M}_{\Lambda'} = \begin{bmatrix} \pi^r I_k & \mathbf{0} \\ \mathbf{0} & \pi^{r-1} I_{n-k} \end{bmatrix} \mathbf{M}_\Lambda,$$

we have

$$\Lambda/\Lambda' \cong (R/\pi^r R)^k \oplus (R/\pi^{r-1} R)^{n-k},$$

where $\oplus$ represents the direct sum of two R-modules. Moreover, in the special case $R = \mathbb{Z}[i]$ or $\mathbb{Z}[\omega]$, $d^2(\Lambda/\Lambda')$ and $K(\Lambda/\Lambda')$ can be respectively expressed as

$$d^2(\Lambda/\Lambda') = \begin{cases} w_E^{min}(C), & \text{when } r = 1 \\ d^2(\Lambda) = \min(|\pi|^2, w_E^{min}(C)), & \text{when } r > 1 \end{cases},$$

$K(\Lambda/\Lambda') = K(\Lambda)$ when $r > 1$, where $K(\Lambda)$ is the number of shortest nonzero vectors in Λ.

When $r = 1$, the constructed LNC Λ/Λ' is isomorphic to the linear code C. In this case, instead of decoding $\hat{\mathbf{u}}$ by quantization over the fine lattice Λ as in formula (5.3), we can find $\hat{\mathbf{u}}$ by a generally suboptimal but more efficient hard-decision method as follows. The scaled received signal $\alpha\mathbf{y}$ is first quantized to a vector $\boldsymbol{\lambda} \in R^n$ by symbol-based quantization over R. Then, $\hat{\mathbf{u}}$ is set to be the element in C found by any decoding algorithm of C with input $\sigma(\boldsymbol{\lambda}) \in (R/\pi R)^n$. This decoding method has been considered in [177] for decoding integer-based compute-and-forward schemes.

Now we discuss the UBE of the decoding error probability for the $\mathbb{Z}[\omega]$-based LNCs constructed by Algorithm 5.3.

Corollary 5.2. *Let Λ/Λ' be the LNC constructed by Algorithm 5.3 from a linear code C over $R/\pi R$. When $R = \mathbb{Z}[\omega]$, the nominal coding gain of Λ/Λ' is*

$$\gamma_c(\Lambda/\Lambda') = \begin{cases} \frac{w_E^{min}(C)}{\frac{\sqrt{3}}{2}|\pi|^{2(1-k/n)}}, & \text{when } r = 1 \\ \frac{\min(|\pi|^2, w_E^{min}(C))}{\frac{\sqrt{3}}{2}|\pi|^{2(1-k/n)}}, & \text{when } r > 1 \end{cases}. \tag{5.9}$$

The UBE on the decoding error probability can be written as

$$\begin{aligned} &P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \\ &\lessapprox \begin{cases} K(\Lambda/\Lambda') \exp\left(-\frac{9}{5}\frac{w_E^{min}(C)}{|\pi|^{2(1-k/n)}}\mathrm{SENR}_{\mathrm{norm}}\right), & \text{when } r = 1 \\ K(\Lambda) \exp\left(-\frac{9}{5}\frac{\min(|\pi|^2, w_E^{min}(C))}{|\pi|^{2(1-k/n)}}\mathrm{SENR}_{\mathrm{norm}}\right), & \text{when } r > 1 \end{cases}. \end{aligned}$$

When $R = \mathbb{Z}[i]$, the nominal coding gain of Λ/Λ' is

$$\gamma_c(\Lambda/\Lambda') = \begin{cases} \frac{w_E^{min}(C)}{|\pi|^{2(1-k/n)}}, & \text{when } r = 1 \\ \frac{\min(|\pi|^2, w_E^{min}(C))}{|\pi|^{2(1-k/n)}}, & \text{when } r > 1 \end{cases}.$$

The UBE on the decoding error probability can be written as

$$P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \lessapprox \begin{cases} K(\Lambda/\Lambda') \exp\left(-\frac{3}{2}\frac{w_E^{min}(C)}{|\pi|^{2(1-k/n)}}\mathrm{SENR}_{\mathrm{norm}}\right), & \text{when } r = 1 \\ K(\Lambda) \exp\left(-\frac{3}{2}\frac{\min(|\pi|^2, w_E^{min}(C))}{|\pi|^{2(1-k/n)}}\mathrm{SENR}_{\mathrm{norm}}\right), & \text{when } r > 1 \end{cases}.$$

The above UBE provides design criteria for constructing optimal LNCs or lattice partitions. More specifically, in order to minimize the decoding error probability, one needs to (1) maximize $w_E^{min}(C)$; and (2) minimize $K(\Lambda/\Lambda')$ or $K(\Lambda)$.

5.5.2 Convolutional LNCs by Complex Construction A

We now discuss construction of LNCs by Algorithm 5.3 from rate-1/2 feed-forward convolutional codes over $\mathbb{F}_{13}$ with memory order v, $1 \leq v \leq 5$. Write $\beta = 2 + 3i$ and $\gamma = 4 + 3\omega$. The message space $W = \mathbb{F}_{13}$ can be represented by either $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ or $\mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$. We adopt the encoding functions $\mathcal{E}_{\mathbb{Z}[i]} : \mathbb{Z}[i]/\beta\mathbb{Z}[i] \to \mathbb{Z}[i]$ and $\mathcal{E}_{\mathbb{Z}[\omega]} : \mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega] \to \mathbb{Z}[\omega]$, whose constellations are depicted in Fig. 5.2 in Section 5.3. The corresponding mappings from $\mathbb{F}_{13}$ into $\mathbb{Z}[i]$ and into $\mathbb{Z}[\omega]$ are summarized in Table D.1. By computer search, we find generator polynomials $\mathbf{g}(D)$ to generate rate-1/2 feed-forward convolutional codes C over $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ and over $\mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$ with maximum $w_E^{min}(C)$ and smallest $K(\Lambda/\Lambda')$, and list them respectively in Table D.2 and Table D.3. Due to computational limitation, the search for the cases $v = 4$ and 5 are not exhaustive and hence the parameters listed in Table D.2 and D.3 are sub-optimal. Both tables also list the nominal coding gain $\gamma_c(\Lambda/\Lambda')$ and the number $K(\Lambda/\Lambda')$ of shortest vectors in $\Lambda\backslash\Lambda'$ of the corresponding LNCs Λ/Λ' constructed from C by applying Algorithm 5.3 with $r = 1$. We can see that when the memory order of the convolutional code C increases from 1 to 4, the minimum squared Euclidean norm $w_E^{min}(C)$ of non-zero codewords in C and the nominal coding gain also increase under both $\mathbb{Z}[\omega]$-based and the $\mathbb{Z}[i]$-based constellations. Moreover, the minimum squared norm $w_E^{min}(C)$ is same under both constellations for all the cases $1 \leq v \leq 4$. Thus, according to Corollary 5.2, a $\mathbb{Z}[\omega]$-based LNC can be correspondingly constructed from a rate-1/2 $\mathbb{F}_{13}$-convolutional code such that its nominal coding gain is 0.625 dB higher than the $\mathbb{Z}[i]$-based one constructed from an $\mathbb{F}_{13}$ convolutional code of same rate and memory. With the additional 0.167 dB shaping gain, Corollary 5.1 asserts that the $\mathbb{Z}[\omega]$-based LNC will have a better performance than the $\mathbb{Z}[i]$-based one, despite of the larger number $K(\Lambda/\Lambda')$ as listed in the tables. This will be illustrated by simulation in next section.

The rate of an LNC Λ/Λ' is equal to $\frac{1}{n}\log_2|\Lambda/\Lambda'|$. Assume that Λ/Λ' is a $\mathbb{Z}[\omega]$-based LNC constructed by Algorithm 5.3 from a linear code C over $\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega]$. Thus, a larger r could yield a higher rate LNC. On the other hand, when $r \geq 2$, the length $d(\Lambda/\Lambda')$ of the shortest vectors in $\Lambda\backslash\Lambda'$ is equal to the length $d(\Lambda)$ of the shortest nonzero vectors in Λ, which is upper bounded by $|\pi|$. Since the decoding error probability will decrease exponentially with respect to the nominal coding gain, which is the ratio of $d^2(\Lambda/\Lambda')$ over the n^{th} root of the volume of Λ, the increase of the code rate by increasing r is at the cost of code performance. To compensate this to some extent, we are motivated to design LNCs by Complex Construction B, which is another method to construct a complex R-lattice [170].

5.5.3 LNC From a Linear Code by Complex Construction B

We now describe a lattice partition algorithm based on Complex Construction B. Now we obtain another lattice partition Λ/Λ' from a linear code C. In terms of Complex Construction D in [21], this lattice partition can also be regarded as constructed from a sequence of nested linear codes $C_0 \supseteq C_1 \supseteq ... \supseteq C_{r-1} \supseteq C_r$ over $R/\pi R$ with $C_0, ..., C_{r-2}$ as $[n, n]$ trivial codes, C_{r-1} as the $[n, n-1]$ single-parity-check code, and C_r as the given code C in Algorithm 5.4. Next, we show the generator matrices for lattices Λ and Λ' and their relationship.

Algorithm 5.4 Complex Construction B

1: Consider a linear code C of length n over $R/\pi R$ subject to $\sum_{1\leq i\leq n} c_i = 0$ for each $(c_1, \cdots, c_n) \in C$. Define

$$\Lambda = \{\boldsymbol{\lambda} \triangleq (\lambda_1, \cdots, \lambda_n) \in R^n : \sigma(\boldsymbol{\lambda}) \in C, \sum\nolimits_{i=1}^{n} \lambda_i \equiv 0 \mod \pi^2\} \tag{5.10}$$

where σ is the natural projection from R to $(R/\pi R)^n$, and $\Lambda' = (\pi^r R)^n$, where $r \geq 2$. In this way, Λ is an n-dimensional R-lattice and Λ' is a sublattice of Λ. An LNC Λ/Λ' is thus constructed from C by Complex Construction B.

Theorem 5.2. *Let Λ/Λ' be an LNC constructed from an $[n, k]$ linear code C over $R/\pi R$ by Algorithm 5.4. There exists a generator matrix $\mathbf{M}_\Lambda$ for Λ and $\mathbf{M}_{\Lambda'}$ for Λ' in the form*

$$\mathbf{M}_\Lambda = \begin{bmatrix} I_k & & & B_{k\times(n-k)} & & \\ & \pi & -\pi & 0 & \dots & 0 \\ & & & \ddots & & \\ \mathbf{0} & & & & & \\ & 0 & \dots & 0 & \pi & -\pi \\ & 0 & 0 & \dots & 0 & \pi^2 \end{bmatrix}, \mathbf{M}_{\Lambda'} = \begin{bmatrix} \pi^r I_k & & & \pi^r B_{k\times(n-k)} & & \\ & \pi^r & -\pi^r & 0 & \dots & 0 \\ & & & \ddots & & \\ \mathbf{0} & & & & & \\ & 0 & \dots & 0 & \pi^r & -\pi^r \\ & 0 & 0 & \dots & 0 & \pi^r \end{bmatrix} \tag{5.11}$$

Consequently,

$$\mathbf{M}_{\Lambda'} = \begin{bmatrix} \pi^r I_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \pi^{r-1} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \pi^{r-2} \end{bmatrix} \mathbf{M}_\Lambda,$$

and hence

$$\Lambda/\Lambda' \cong (R/\pi^r R)^k \oplus (R/\pi^{r-1} R)^{n-k-1} \oplus (R/\pi^{r-2} R).$$

Moreover, in the special case that $R = \mathbb{Z}[i]$ or $\mathbb{Z}[\omega]$,

$$\begin{aligned} d^2(\Lambda/\Lambda') &= d^2(\Lambda) = \min(2|\pi|^2, w_E^{min}(C)) \\ K(\Lambda/\Lambda') &= K(\Lambda), \text{ when } |\pi|^2 \neq 2 \end{aligned}$$

Proof. See Appendix C.3. ■

We next give explicit UBE of the decoding error probability for $\mathbb{Z}[\omega]$-based LNCs constructed by Complex Construction B.

Corollary 5.3. *When $R = \mathbb{Z}[\omega]$, the nominal coding gain of the LNC Λ/Λ' constructed from a linear code over $R/\pi R$ by Algorithm 5.4 is*

$$\gamma_c(\Lambda/\Lambda') = \frac{\min(2|\pi|^2, w_E^{min}(C))}{\frac{\sqrt{3}}{2}|\pi|^{2(1-(k-1)/n)}}.$$

The UBE on the decoding error probability can be written as

$$P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \lessapprox K(\Lambda) \exp\left(-\frac{9}{5}\frac{\min(2|\pi|^2, w_E^{min}(C))}{|\pi|^{2(1-(k-1)/n)}}\mathrm{SENR}_{\mathrm{norm}}\right).$$

This UBE also gives design guidelines for constructing optimal LNCs in terms of the minimal error probability: the LNC should have maximum $w_E^{min}(C)$ and minimum $K(\Lambda)$.

It is of particular interest to discuss Complex Construction B over Eisenstein integers. For instance, the well-known complex Leech lattice can be constructed based on it (See Example 12 of Chapter 7 in [170].) As discussed in the next example, the associated LNC of complex Leech lattice has a good tradeoff between code rate and nominal coding gain among the LNCs constructed from ternary Golay code.

Example 5.2. Let $\pi = 1 + 2\omega$ and C represent the [12, 6, 6] extended Golay code over $\mathbb{F}_3 \cong \mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega]$. Table D.4 compares several attributes of LNCs Λ/Λ' with the 12-dimensional lattice Λ constructed from C by different methods. When Λ/Λ' is constructed by Complex Construction A, there is a tradeoff between the rate and the nominal coding gain of the LNC. For example, assume that Λ is constructed by Complex Construction A (Formula (5.9)). When Λ' is changed from $(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])^{12}$ to $(\mathbb{Z}[\omega]/\pi^2\mathbb{Z}[\omega])^{12}$, the rate of the LNC Λ/Λ' can be increased by three times from $\frac{1}{2}\log_2 3$ to $\frac{3}{2}\log_2 3$, but the nominal coding gain is decreased from 6.02 dB to 3.01 dB. However, when Λ' is kept to be $(\mathbb{Z}[\omega]/\pi^2\mathbb{Z}[\omega])^{12}$ while Λ is constructed by Complex Construction B (Formula (5.10)), the LNC Λ/Λ' has rate $\frac{17}{12}\log_2 3$, which is close to $\frac{3}{2}\log_2 3$, and the nominal coding gain 5.62 dB, which is much better than 3.01 dB obtained by Complex Construction A at rate $\frac{3}{2}\log_2 3$. If we further choose Λ to be the complex Leech lattice, which is constructed based on complex Construction B, and set Λ' to be $(\mathbb{Z}[\omega]/\pi^3\mathbb{Z}[\omega])^{12}$, the rate of the LNC can be increased to $\frac{3}{2}\log_2 3$ and the nominal coding gain to 6.02 dB.

The example above demonstrates that Complex Construction B is possible to yield LNCs with higher rates without sacrificing the nominal coding gain much.

5.6 Construction of $\mathbb{Z}[\omega]$-Based LNCs over GF(4)

5.6.1 Why Interested in GF(4)

We now focus on the construction of $\mathbb{Z}[\omega]$-Based LNCs over GF(4). The motivation of studying this particular lattice partition is that signal constellations of size to be a power of 2 are always preferred in real world implementation. Thus, for the practical design of LNCs, it will be more convenient to adopt a lattice partition $R/\pi R$ that is isomorphic to a finite field of size 2^m. Unfortunately, the only finite fields of characteristic 2 that can be represented by $R/\pi R$ for some $\pi \in R$ when $R = \mathbb{Z}, \mathbb{Z}[i], \text{or } \mathbb{Z}[\omega]$ include

- $\text{GF}(2) \cong \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}[i]/(1+i)\mathbb{Z}[i]$;
- $\text{GF}(2^2) \cong \mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$.

Since GF(2) can be represented over $\mathbb{Z}$, it is no longer desirable to consider its constellation over the complex plane. Moreover, compared with over GF(2), the flexibility is higher over GF(4) to design LNCs with a large minimum inter-coset distance. In this letter, we aim at examining the practical design of GF(4)-based LNCs from linear codes over GF(4^m), where $m \geq 1$.

In this section, we will incorporate the dithering method at the users and show how to design the optimal dither at the transmitters to save the average transmission power.

5.6.2 Dithering in the System Model

With dithering is considered in the user transmission, the encoding function $\mathcal{E}$ becomes

$$\mathcal{E}(\mathbf{w}_l) \triangleq \bar{\varphi}(\mathbf{w}_l) + \mathbf{d} - \mathcal{D}_{\Lambda'}(\bar{\varphi}(\mathbf{w}_l) + \mathbf{d}),$$

where

$$\bar{\varphi} : \Lambda/\Lambda' \to \Lambda$$

is an embedding mapping, $\mathbf{d} \in \mathbb{C}^n$ is a *dither vector*, and $\mathcal{D}_{\Lambda'} : \mathbb{C}^n \to \Lambda'$ is the lattice quantizer of the coarse lattice Λ'. The estimated R-linear combination of the transmitted messages at the relay becomes

$$\hat{\mathbf{u}} = \varphi(\mathcal{D}_\Lambda(\alpha\mathbf{y} - \sum\nolimits_{l=1}^{L} \mathbf{a}_l\mathbf{d})),$$

where φ is the natural projection from Λ onto Λ/Λ' via $\varphi(\boldsymbol{\lambda}) = \boldsymbol{\lambda} + \Lambda'$, and the $\mathcal{D}_\Lambda$ is the lattice quantizer of the fine lattice Λ.

5.6.3 Baseline LNCs over GF(4)

The set $\mathbb{Z}[\omega]$ of Eisenstein integers is depicted in Fig. 5.3(a). It is naturally qualified as a lattice in the complex plane. We next examine the field structure of $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$. The four cosets in $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ are labeled by differently shaped points in Fig. 5.3(a), each of which can be assigned a coset leader 0, 1, ω and ω^2 respectively. Noting that $\omega^2 + \omega + 1 = 0$, the arithmetic of addition and multiplication among these four cosets can be easily computed and is summarized in Table 5.1, where each coset is represented by the assigned coset leader. It is easy to check that equipped with this arithmetic, $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ is indeed isomorphic to the finite field GF(4).

The message space of the baseline LNC over GF(4) is $W = \gamma(\mathbb{Z}[\omega]/2\mathbb{Z}[\omega])^n$, where γ is a scaling factor to control the transmission power. It can be regarded as an n-dimensional vector space over GF(4). If the encoding function $\mathcal{E}$ maps a message $\mathbf{w}_l \in W$ to its coset leader whose entry elements are in $\{0, \gamma, \gamma\omega, \gamma\omega^2\}$, then the average power of the message space is $\frac{1}{n}E[\|\mathcal{E}(\mathbf{w}_l)\|^2] = \frac{3}{4}|\gamma|^2$. From an

Table 5.1 The arithmetic in cosets of $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$

$+$	0	1	ω	ω^2
0	0	1	ω	ω^2
1	1	0	ω^2	ω
ω	ω	ω^2	0	1
ω^2	ω^2	ω	1	0

$\times$	0	1	ω	ω^2
0	0	0	0	0
1	0	1	ω	ω^2
ω	0	ω	ω^2	1
ω^2	0	ω^2	1	ω

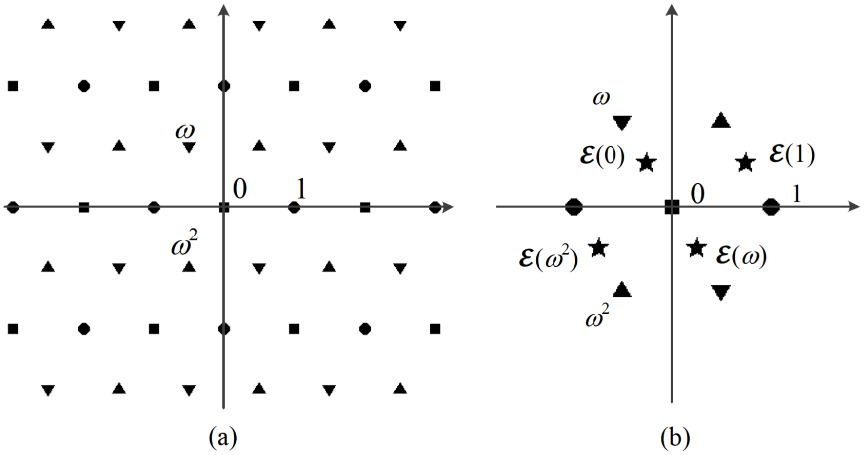


Figure 5.3(a) The set $\mathbb{Z}[\omega]$ of Eisenstein integers in the complex plane is partitioned into four cosets by modulo 2, each of which is depicted in a different shape and is labeled with a coset leader. (b) A constellation of $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ to achieve the optimal average power $\frac{1}{2}$ is depicted by star points.

information theoretical point of view, random dither is required in encoding in order to make the quantization noise to be uniformly distributed and independent of encoded signals [178]. For the sake of energy-efficiency in practical systems, a fixed dither is required in order to minimize the average transmission power or error probability under a transmission power constraint [21]. We next design an optimal dither vector for the encoding function $\mathcal{E}$ in terms of energy efficiency.

Proposition 5.3. *The optimum average power for the baseline LNC is* $\frac{1}{2}|\gamma|^2$, *which can be obtained by dither vectors* $\mathbf{d} = \gamma(d_1, \ \cdots, d_n)$, *where* $d_j \in \{\pm\frac{\omega}{2}, \pm\frac{\omega^2}{2}, \pm\frac{1}{2}\}$ $\forall 1 \leq j \leq n$.

Proof. It suffices to show the case when $\gamma = 1$ and $n = 1$. Write $d = d_R + d_I i$, where d_R and d_I are real numbers. By symmetry, without loss of generality, assume $-1 \leq d_R \leq 0 \leq d_I \leq \sqrt{3}$. Based on different d, the encoding function $\mathcal{E}$ maps a coset into different values:

1. $\mathcal{E}(0 + 2\mathbb{Z}[\omega])$ is d when $d_I \leq \frac{\sqrt{3}}{3}d_R + \frac{2\sqrt{3}}{3}$ and is $-2\omega + d$ otherwise;

2. $\mathcal{E}(1 + 2\mathbb{Z}[\omega])$ is $1 + d$ when $d_I \leq -\frac{\sqrt{3}}{3}d_R + \frac{\sqrt{3}}{3}$ and is $-1 - 2\omega + d$ otherwise;

3. $\mathcal{E}(\omega + 2\mathbb{Z}[\omega])$ is $-\omega + d$;

4. $\mathcal{E}(\omega^2 + 2\mathbb{Z}[\omega])$ is $\omega^2 + d$ when $d_I \geq -\frac{\sqrt{3}}{3}d_R$ and $d_R \geq -\frac{1}{2}$, is $1 - \omega + d$ when $d_I \geq \frac{\sqrt{3}}{3}d_R + \frac{\sqrt{3}}{3}$ and $d_R \leq -\frac{1}{2}$, and is $-\omega^2 + d$ otherwise.

After enumerating all the possibilities above, we can find that $E[\|\mathcal{E}(\mathbb{Z}[\omega]/2\mathbb{Z}[\omega])\|^2]$ is minimized to be $\frac{1}{2}$ only when $d = -\frac{1}{2}$ and $d = -\frac{1}{4} + \frac{\sqrt{3}}{4}i$. Say, for example, when $-\frac{1}{2} \leq d_R \leq 0$ and $\frac{-\sqrt{3}}{3}d_R \leq d_I \leq \frac{\sqrt{3}}{3}d_R + \frac{\sqrt{3}}{3}$, $\mathcal{E}(\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]) = \frac{1}{4}(|d|^2 + |1 + d|^2 + |-\omega + d|^2 + |\omega^2 + d|^2) = 4d_R^2 + 4d_I^2 + 2d_R - 2\sqrt{3}d_I + 3$, which is minimized to be $\frac{1}{2}$ with $d_R = -\frac{1}{4}$ and $d_I = \frac{\sqrt{3}}{4}$. ■

It is worthwhile to note that both the Conway-Sloane (CS) method [179] and the maximally biased method [180] can be applied to find a good dither vector for a general real lattice partition in terms of low average power. When $\mathbb{Z}[\omega]$ is regarded as a two-dimensional real lattice, both methods can yield an optimal dither vector for the baseline LNC. Fig. 5.3(b) depicts an optimally dithered constellation of $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ in the complex plane, according to $d = \frac{\omega}{2}$. When an LNC has a large constellation, its average power is always computed by continuous approximation [121], which assumes the encoded messages continuously uniformly distributed over the fundamental Voronoi region of the coarse lattice. In comparison, the average power of the baseline LNC over GF(4) computed by continuous approximation is equal to $\frac{5}{9}|\gamma|^2$, which underestimates the average power $\frac{3}{4}|\gamma|^2$ in the non-dithered case and overestimates the average power $\frac{1}{2}|\gamma|^2$ in the optimally dithered case. Therefore, the optimal dither reduces the average power of a baseline LNC over GF(4) by 1/3, that is, from $\frac{3}{4}|\gamma|^2$ to $\frac{1}{2}|\gamma|^2$.

5.6.4 Design of LNCs over GF(4)

Let C be an $[n, k]$ linear code over GF(4). We now investigate the rate-$\frac{2k}{n}$ LNC Λ/Λ' constructed from different classes of C by Complex Construction A [21]:

$$\Lambda = \{\boldsymbol{\lambda} \in \gamma\mathbb{Z}[\omega]^n : \sigma(\gamma^{-1}\boldsymbol{\lambda}) \in C\}, \Lambda' = \gamma(2\mathbb{Z}[\omega])^n \tag{5.12}$$

where the scaling factor γ controls the transmission power, and σ is the natural projection from $\mathbb{Z}[\omega]^n$ onto $(\mathbb{Z}[\omega]/2\mathbb{Z}[\omega])^n$. An important property of the LNC thus constructed is

$$\Lambda/\Lambda' \cong (\mathbb{Z}[\omega]/2\mathbb{Z}[\omega])^k \cong C.$$

Consequently, the optimal dither vector derived in Proposition 5.3 for baseline LNCs is also optimal here.

Proposition 5.4. *The optimum average power for the LNC Λ/Λ' constructed from C by (5.12) is $\frac{1}{2}|\gamma|^2$, which can be obtained by the dither vectors in Proposition 5.3.*

When the LNC is constructed by complex Construction A from a linear code C over a field represented by a large lattice partition in the complex plane, $d^2(\Lambda/\Lambda')$ is equal to the minimum squared Euclidean weight of nonzero codewords in C [21], which is in general larger than the minimum Hamming distance $w_H(C)$ of C. In the special case of GF(4), however, since every nonzero coset in $\mathbb{Z}[\omega]/2\mathbb{Z}[\omega]$ contains exactly two units in $\mathbb{Z}[\omega]$, which have Euclidean norm 1, we have the following simple characterization.

Proposition 5.5. *For the LNC constructed by (5.12), we have*

$$d^2(\Lambda/\Lambda') = |\gamma|^2 w_H(C) \quad \text{and} \quad K(\Lambda/\Lambda') = 2^{w_H(C)}K(C),$$

where $K(C)$ is the number of codewords with weight $w_H(C)$.

Thus, the characterization of $d^2(\Lambda/\Lambda')$ and $K(\Lambda/\Lambda')$ of LNCs Λ/Λ' over GF(4) is much easier to determine than other lattice partitions with large constellation. Here, we only need to determine the minimum Hamming distance $w_H(C)$ and its multiplicity $K(C)$ of the linear code C, rather than its Euclidean distance of the code constellations.

Example 5.3. In practice, the LNC should be such constructed that the relay receiver is able to decode a linear combination of transmitted messages not only reliably but also relatively efficiently. Since both convolutional codes and BCH codes are known to have efficient hard-decision and soft-decision decoding algorithms, we next provide a parameter characterization for C being a convolutional code and a BCH code, as a reference for practical design of LNCs by (5.12).

The parameters of rate-1/2 feed-forward convolutional codes over GF(4) with memory order v, $1 \leq v \leq 5$, are obtained by computer search and summarized in Table 5.2. The coefficient vectors of the generator polynomials for those convolutional codes C with maximum $w_H(C)$ are listed in the table in descending order of associated degrees. The table also lists the nominal coding gains $\gamma_c(\Lambda/\Lambda') = \frac{d^2(\Lambda/\Lambda')}{\frac{\sqrt{3}}{2}4^{1-k/n}}$ (See [121]) of the LNCs Λ/Λ' constructed from these convolutional codes via (5.12). The nominal coding gain measures the increase in density of Λ/Λ' over the baseline LNC. Compared with the convolutional LNCs over GF(13) which have larger constellation in [121], the convolutional LNCs over GF(4) have similar nominal coding gains.

Table 5.3 summarizes the parameters of several BCH codes over GF(4) of length 15 and 63. The coefficient vectors of the generator polynomials $g(X)$ are listed in the table in descending order of associated degrees. We obtain $w_H(C)$ and $K(C)$ by computer enumeration (with the use of MacWilliams identity in the case $n = 63$.) It is worthwhile to point out that the real minimum distance $w_H(C)$ is same as the designed distance of all BCH codes listed in the table, but this equivalence is not generally true.

Besides the design of an LNC from a linear code over GF(4), more generally, we

Table 5.2 Parameters of rate-1/2 convolutional codes over GF(4)

v	$g(D)$	$w_H(C)$	$\gamma_c(\Lambda/\Lambda')$
1	$[1\ 1]$, $[\omega\ 1]$	4	3.63 dB
2	$[1\ 1\ 1]$, $[1\ \omega\ 1]$	6	5.40 dB
3	$[1\ \omega^2\ \omega\ \omega^2]$, $[\omega\ \omega^2\ \omega^2\ \omega^2]$	8	6.65 dB
4	$[\omega\ \omega^2\ \omega^2\ \omega\ \omega^2]$, $[\omega^2\ 0\ 1\ \omega^2\ \omega^2]$	9	7.16 dB
5	$[\omega\ 0\ 1\ \omega^2\ \omega^2\ 1]$, $[\omega\ \omega^2\ \omega^2\ \omega^2\ \omega\ 1]$	11	8.03 dB

Table 5.3 Parameters of BCH codes over GF(4)

n	k	$g(X)$	$w_H(C)$	$\gamma_c(\Lambda/\Lambda')$	$K(C)$
15	9	$[1\ \omega^2\ 1\ 1\ \omega\ \omega\ 1]$	5	5.21 dB	189
	7	$[1\ 0\ 1\ \omega^2\ \omega^2\ 1\ \omega^2\ 0\ \omega]$	7	5.86 dB	405
63	54	$[1\ 0\ \omega^2\ 1\ 0\ 1\ 1\ \omega^2\ \omega\ 1]$	5	6.76 dB	8505
	50	$[1\ \omega\ \omega\ 1\ \omega\ \omega^2\ 0\quad \omega^2\ \omega^2\ \omega^2\ 0\ 1\ 0\ \omega^2]$	7	7.83 dB	3591

can also construct an LNC from a linear code C over GF(4^m) for $m > 1$, by adapting C's expanded linear code C_{e} over GF(4) to formula (5.12). Specifically, let β denote a primitive element of GF(4^m), that is, $\beta, \cdots, \beta^{4^m-1}$ constitute all nonzero elements in GF(4^m). Then, $\{1, \beta, \cdots, \beta^{m-1}\}$ forms a natural basis of GF(4^m) over the subfield GF(4) $= \{0, 1, \beta^{\frac{4^m-1}{3}}, \beta^{\frac{2(4^m-1)}{3}}\} \subset$ GF(4^m). Denote by ϕ the natural mapping from GF(4^m) onto the m-dimensional vector space GF(4^m) via

$$\phi(\sum\nolimits_{j=0}^{m-1} c_j\beta^j) = (c_0, \cdots, c_{m-1}).$$

Applying component-wise, the bijection ϕ also extends to a bijection from GF$(4^m)^n$ onto GF$(2)^{mn}$. An $[n, k]$ linear code C over GF(4^m) can then be expanded to an $[mn, mk]$ code C_{e} over GF(4) in terms of the basis $\{1, \beta, \cdots, \beta^{m-1}\}$ by $C_{\mathrm{e}} = \{\phi(\mathbf{c}) : \mathbf{c} \in C\}$. It is easy to verify that the expanded code C_{e} is linear. Concomitantly, an

mn-dimensional, rate-$\frac{2k}{n}$ LNC Λ/Λ' can be constructed from C by

$$\begin{aligned} \Lambda &= \{\boldsymbol{\lambda} \in \gamma\mathbb{Z}[\omega]^{mn} : \sigma(\gamma^{-1}\boldsymbol{\lambda}) = \phi(\mathbf{c}) \text{ for some } \mathbf{c} \in C\} \\ \Lambda' &= \gamma(2\mathbb{Z}[\omega])^{mn} \end{aligned} \tag{5.13}$$

Proposition 5.6. *For the LNC constructed from a linear code C over GF(4^m) by (5.13), the following propositions hold:*

- $\Lambda/\Lambda' \cong \mathrm{GF}(4)^{mk} \cong \phi(C)$
- *The optimum average power for Λ/Λ' is $\frac{1}{2}|\gamma|^2$, which can be obtained by the dither vectors in Proposition 5.3.*
- $|\gamma|^2 w_H(C) \leq d^2(\Lambda/\Lambda') = |\gamma|^2 w_H(C_{\mathrm{e}}) < |\gamma|^2 m w_H(C)$
- $K(\Lambda/\Lambda') = 2^{w_H(C_{\mathrm{e}})} K(C_{\mathrm{e}})$

An application of (5.13) is to construct Λ/Λ' from Reed-Solomon (RS) codes over GF(4^m). As an example, we design a length-30 rate-$\frac{22}{15}$ LNC Λ/Λ' constructed by (5.13) from a [15, 11, 5] RS code C over GF(16). If the generator polynomial for C is selected to be $(X-\beta)\cdots(X-\beta^4)$, that is, C is a narrow-sense RS code, then the LNC has parameters $d^2(\Lambda/\Lambda') = 5$ and $K(\Lambda/\Lambda') = 2^5 \cdot 648$. On the other hand, if the generator polynomial for C is changed to $(X-\beta^2)\cdots(X-\beta^5)$, $d^2(\Lambda/\Lambda')$ will be increased to 6, but $K(\Lambda/\Lambda')$ is also increased to $2^6 \cdot 9480$. Thus, the LNC parameters depend on the selection of generator polynomials of the RS code. Since expanded codes C_e can be treated as generalized concatenated codes [181], some lower bounds can be derived for $w_H(C_e)$ based on this structure (See, for example, [182]). Along this line, it will be an interesting future work to derive new bounds on $d(\Lambda/\Lambda')$ for the LNCs constructed by (5.13).

5.7 Simulation Results

5.7.1 Construction with General Lattice Partition

In this section, we evaluate via simulations the error performance of LNCs to decode a linear combination of the messages sent from L transmitters, where $L = 2$. In this case, the optimal coefficient vector $\mathbf{a}$ can be efficiently found by the generalized Gaussian reduction algorithm in Algorithm 5.2.

We first evaluate the performance of three baseline LNC schemes designed over $\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega] \cong \mathbb{F}_{|\pi|^2}$, where π is an Eisenstein prime with norm 13, 37 and 61. We compare it with the baseline LNC schemes over $\mathbb{Z}[i]/\pi\mathbb{Z}[i] \cong \mathbb{F}_{|\pi|^2}$, where π is a Gaussian prime with norm 13, 37 and 61. Fig. 5.4 plots the average symbol error probability (SEP) of the schemes as a function of SNR, where the channel gains are fixed and the receiver chooses a single linear function. For better comparison, it also depicts the SEP derived by UBE in (5.6) for baseline LNCs over $\mathbb{Z}[\omega]$ with constellation sizes 13 and 61. The results show that for the same constellation size, the LNC schemes over $\mathbb{Z}[\omega]$ are about 0.5-0.6 dB better than the ones over $\mathbb{Z}[i]$, and their UBE gives a good upper bound approximate over high SNR for the decoding error probability. These results are consistent with the analysis in Section 5.4.1.

Now we consider a time-varying fading channel in the simulation. In Fig. 5.5, we show the average SER of the baseline LNC schemes $\mathbb{Z}[i]/(2 + 3i)\mathbb{Z}[i] \cong \mathbb{F}_{13}$ and $\mathbb{Z}[\omega]/(4 + 3\omega)\mathbb{Z}[\omega] \cong \mathbb{F}_{13}$ over a Rayleigh fading channel. Here each scheme adopts an optimal coefficient vector $\mathbf{a}_{\text{opt}}$, which can be found based on [175] in the $\mathbb{Z}[i]$ case and on Algorithm 5.2 of Section 5.4.2 in the $\mathbb{Z}[\omega]$ case. It is clear to see that the SNR of the $\mathbb{Z}[\omega]$-based LNC outperforms the $\mathbb{Z}[i]$-based LNC by about 0.6 dB at the SER of 10^{-2}, consistent with the analysis in Section 5.4.1.

We next compare the performance of LNCs constructed from convolutional codes over $\mathbb{F}_{13} \cong \mathbb{Z}[i]/(2+3i)\mathbb{Z}[i] \cong \mathbb{Z}[\omega]/(4+3\omega)\mathbb{Z}[\omega]$ with memory order $v = 1, 2$ as listed in Table D.2 and Table D.3 in Appendix D. The length of the information sequence is set

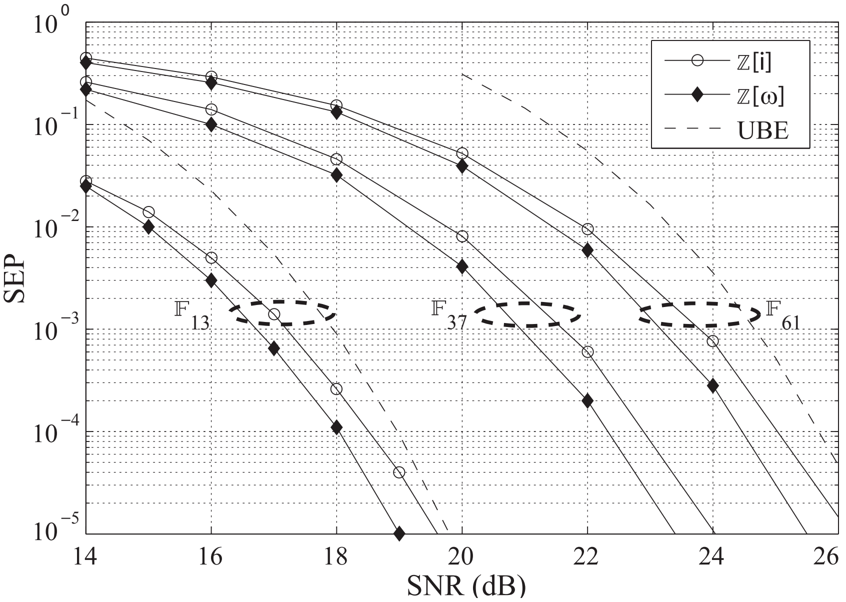


Figure 5.4 Average SEP of baseline LNC schemes under a fixed channel gain.

to 99 when $v = 1$ and to 98 when $v = 2$. The code rate of the LNCs thus constructed is $\frac{99}{200}\log_2 13$ when $v = 1$ and $\frac{98}{200}\log_2 13$ when $v = 2$. Fig. 5.6 depicts the average frame error rate (FER) under the fixed channel gain $\mathbf{h} = [-1.17 + 2.15i\ 1.25 - 1.63i]$ as a function of $\text{SENR}_{\text{norm}}$. This fixed channel gain has been adopted in [23] and [21] for evaluating the performance of different hypercube shaped LNC schemes. We can see that for both memory order $v = 1, 2$, the $\mathbb{Z}[\omega]$-based LNC has better frame error performance, which is about 0.2 to 0.3 dB at FER $= 10^{-4}$. Fig. 5.7 compares the average FER of these LNCs over a Rayleigh fading channel, in which an optimal coefficient vector is adopted. At FER $= 10^{-2}$, the performance gap in terms of SNR between $\mathbb{Z}[\omega]$-based and $\mathbb{Z}[i]$-based LNCs is about 1 dB for memory order $v = 1$ and 0.5 dB for memory order $v = 2$. Fig. 5.7 also compares these LNCs with the generic compute-and-forward schemes, denoted by 'Nazer-Gastpar', constructed by the approach proposed in [23] for $\mathbb{Z}[i]$ case and then extended in [188] to $\mathbb{Z}[\omega]$ case. Note that the Nazer-Gastpar schemes involve an infinite sequence of structured lattice partitions. They are not implementable in practice but can be regarded as a theoretical guideline for the error probability at a given achievable computation rate

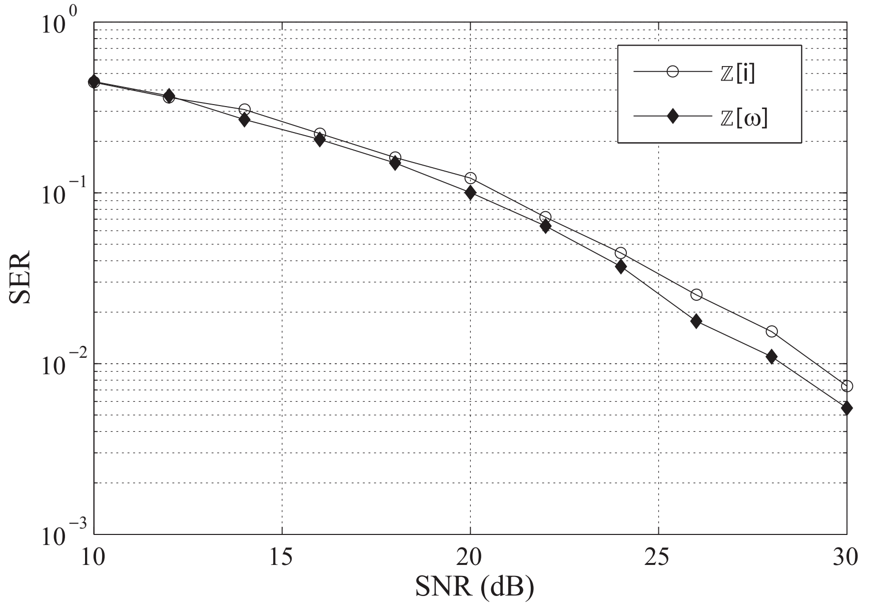


Figure 5.5 Average symbol error probability of different baseline LNC schemes over Rayleigh fading channels.

of compute-and-forward schemes. This shows how practical schemes can perform compared to theoretical benchmark. Both $\mathbb{Z}[i]$-based and $\mathbb{Z}[\omega]$-based Nazer-Gastpar schemes are set to have rate $\frac{1}{2}\log_2 13$ and adopt the decoding criterion that a frame error occurs iff $\frac{1}{2}\log_2 13 \geq \log_2 \frac{\text{SNR}}{Q(\alpha,\mathbf{a})}$. It is clear that the $\mathbb{Z}[\omega]$-based Nazer-Gastpar scheme is better than the $\mathbb{Z}[i]$-based one, and that for both $\mathbb{Z}[i]$ and $\mathbb{Z}[\omega]$ cases, the memory order-2 convolutional LNC is just about 4 dB away from the Nazer-Gastpar scheme.

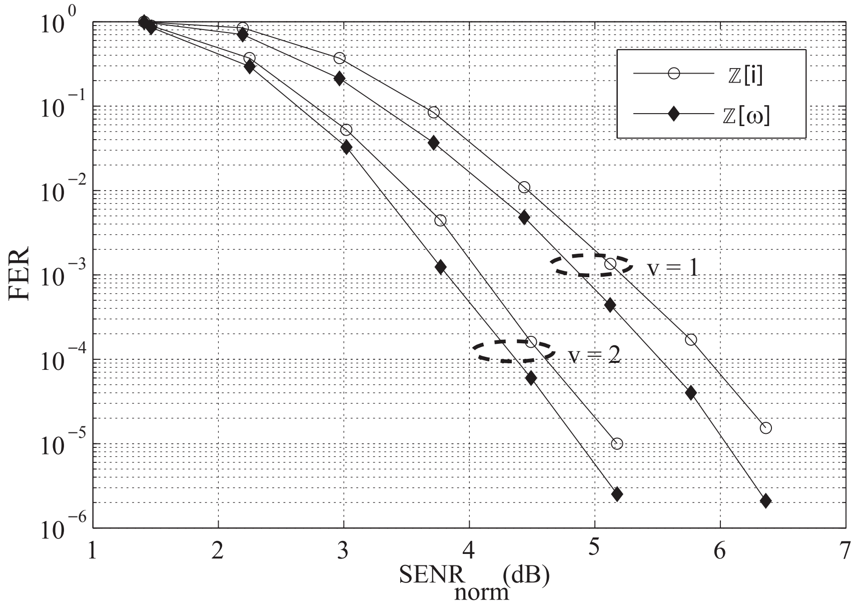


Figure 5.6FER of the LNC schemes constructed from convolutional codes listed in Table D.2 and Table D.3 with $v = 1, 2$, under a fixed channel gain.

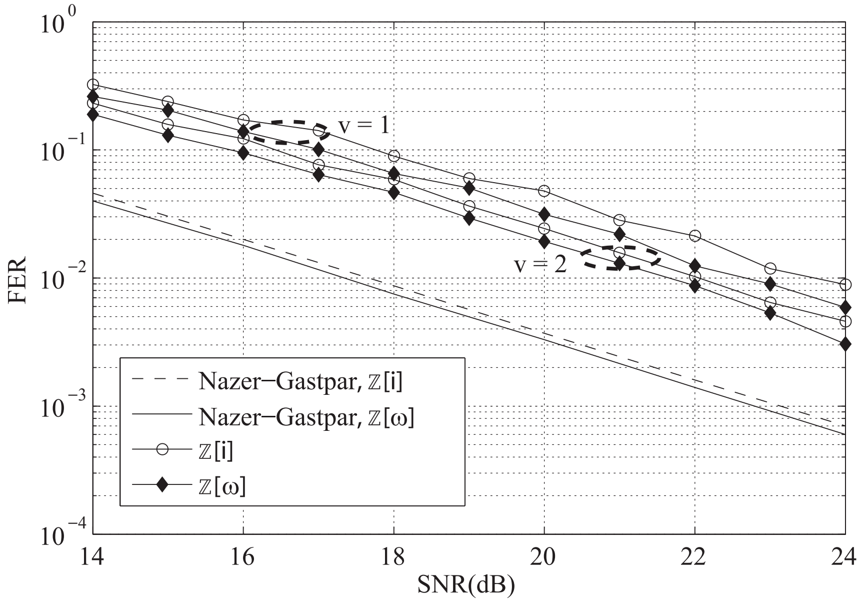


Figure 5.7FER of the LNC schemes constructed from convolutional codes listed in Table D.2 and Table D.3 with $v = 1, 2$, over a Rayleigh fading channel.

5.7.2 Construction over GF(4)

We now evaluate the error performance of LNCs over GF(4) via simulations. Again, we assume the number of user is 2.

Fig. 5.8 depicts the average frame error rate (FER) of four 15-dimensional LNCs as a function of normalized SNR $\triangleq \frac{\frac{1}{n}E[\|\mathbf{x}\|^2]}{2^\rho N_0}$, where $\rho = \frac{1}{n}\log_2|\Lambda/\Lambda'|$ represents the spectral efficiency or the rate of the LNC. We can first observe that the performance of the optimally dithered baseline LNC is around 3.5 dB better than the non-dithered baseline LNC at the FER 10^{-2}. For the LNCs constructed by (5.12) from the 15-dimensional BCH codes listed in Table 5.3, optimal dither is also taken into account, and the hard-decision decoding (and thus suboptimal) is adopted at the relay. At the FER 10^{-2}, a performance gain of around 3 dB for the rate-$\frac{18}{15}$ LNC and 4 dB for rate-$\frac{14}{15}$ LNC can be observed.

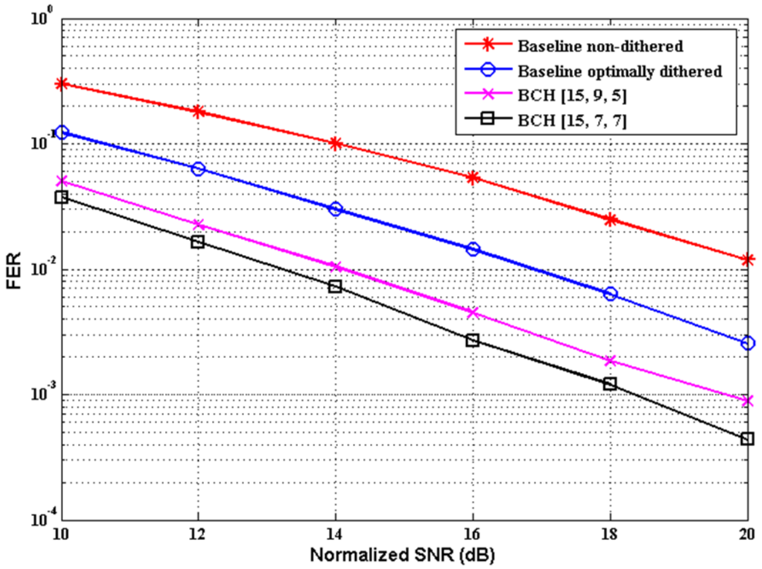


Figure 5.8 Average frame error rate of 15-dimensional LNCs over GF(2^2)

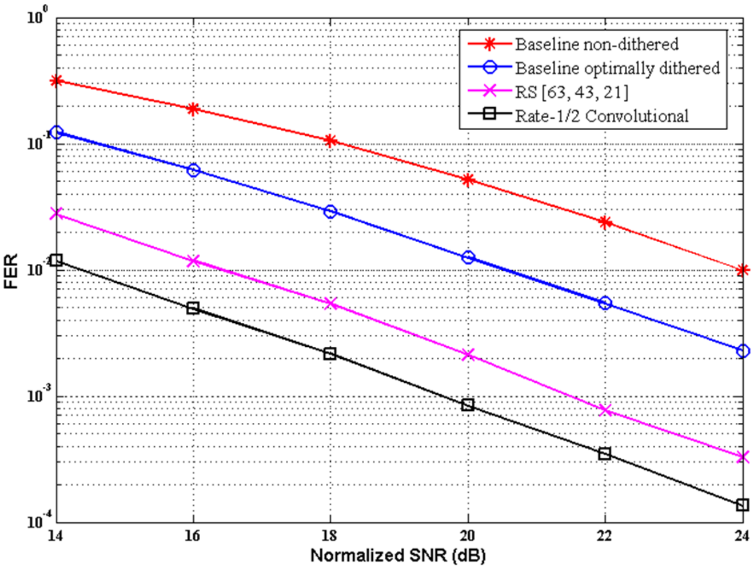


Figure 5.9 Average frame error rate of 189-dimensional LNCs over GF(2^2)

Fig. 5.9 depicts the FER of four 189-dimensional LNCs. Compared with the non-dithered baseline LNC, at the FER 10^{-2}, the rate-$\frac{86}{63}$ LNC constructed by (5.13) from the [63, 43] RS code and the rate-1 LNC constructed by (5.12) from the convolutional code in Table 5.2 with $v = 1$, both optimally dithered, can respectively obtain around 7.5 dB and 9.5 dB performance gain. Suboptimal hard-decision decoding is adopted for the LNC constructed from the RS code whereas is the Viterbi decoding algorithm is adopted for the convolutional LNC.

5.8 Conclusion

In this chapter, we focused on Eisenstein integers based lattice network codes (LNCs). We first present $\mathbb{Z}[\omega]$-based quantization and encoding algorithms, whose computational complexity is in the same order as the case over the PID $\mathbb{Z}[i]$ of Gaussian integers. Then, a union bound estimation (UBE) of the decoding error probability of $\mathbb{Z}[\omega]$-based LNC is derived. Next, we generalize the Gaussian reduction algorithm to be applicable for complex lattices over $\mathbb{Z}[\omega]$, such that it is able to find an optimal coefficient vector for a $\mathbb{Z}[\omega]$-based LNC in the two-transmitter single-relay system. In addition, based on the UBE, design criteria for optimal LNCs with minimum decoding error probability are also formulated and applied to construct both Gaussian integer based and Eisenstein integer based good rate-1/2 convolutional LNCs by Complex Construction A. The constructed LNCs provide up to 7.65 dB nominal coding gains over Rayleigh fading channels. Furthermore, we introduce the construction of LNCs from linear codes by Complex Construction B, and explicitly formulate the nominal coding gains and error performance of the LNCs thus constructed. We also illustrate by examples that the LNCs constructed by Complex Construction B provide a better tradeoff between code rate and nominal coding gain.

In addition to the construction of Eisenstein integers based LNCs for general lattice partition, we also particularly interested in constructing LNCs over the finite field GF(2^2), whose quaternary constellation has practical interests. We derived the optimal dither method for LNCs over GF(4), so that 1/3 average transmission power can be saved. Construction methods of LNCs from linear codes over GF(4^m), where $m \geq 1$, are also introduced with explicit parameter characterization provided. As design examples, parameters of LNCs constructed from convolutional, BCH, and Reed-Solomon codes are presented and analyzed. Numerical results illustrate that at the frame error rate 10^{-2}, these LNCs, after optimally dithered, can provide up to 9.5 dB gain in terms of normalized SNR compared with the baseline LNCs over GF(4).

Chapter 6

Design of Pair-Wise Transmission PNC in SISO MWRCs

6.1 Introduction

So far we have studied the design of PNC from the signal detection and forward error correction perspective. In this chapter, we study the transmission scheduling problem in single-input single-output (SISO) multi-way relay channels (MWRCs). An MWRC is a multi-cast network where all users exchange their information via a single relay, without any direct link among the users. A typical example is a satellite communication system where multiple ground stations that geographically located far away, communicate with each other via a common satellite. In this work, we aim to design a user pair-wise transmission scheme in the uplink phase (from users to the relay) to improve the computation rate of the network coded information at the relay.

This chapter begins by introducing the background of this study. We then give detailed description of the system model, and we introduce the computation of network coded information at the relay when linear lattice codes are employed. Then we introduce the conventional successive pair-wise transmission scheme in the considered MWRCs. After that, we study the proposed opportunistic pair-wise transmission. In the simulation results, we show the performance improvement in both sum-rate and error rate at the relay. Then we conclude this chapter.

6.2 Background

In [183], the MWRC on Gaussian channel is firstly investigated. The achievable rate region with amplify-and-forward, decode-and-forward, and compress-and-forward has been characterized. In [184], the authors derived the capacity of the binary MWRC, where multiple users exchange information at a common rate via a relay. A pair-wise time division multiple access (TDMA) functional-decoding-forward coding strategy is proposed. It has been shown that this approach achieves the common-rate capacity [184]. The error propagation effect at the relay decoding with binary phase shift keying (BPSK) modulation has been analyzed in [185, 186]. The work in [187] proposed a joint decoding strategy at the relay through the belief propagation algorithm. This proposed technique utilized the correlation between the adjacent network coded symbols to mitigate the error propagation as investigated in [185,186]. This work has been further extended to other channel imperfection cases.

The aforementioned research works on MWRCs are limited in BPSK modulation. The pair-wise transmission scheduling at the users side is done in a sequential order. In this chapter, we investigate the pair-wise compute-and-forward transmission for multi-way relay fading channels. In order to improve the sum-rate of multi-user transmission, we consider high level modulation with nested lattice codes, where the relay computes integer linear combinations of the users' messages rather than decoding individual messages. In addition, we propose an *opportunistic* pair-wise compute-and-forward by exploiting the multi-user fading channels. We show that by choosing the pair-wise transmission scheduling appropriately, we can achieve a significant improvement for the sum-rate of the multi-user transmission. We further demonstrate the practical benefit of this proposed scheme by using both uncoded and channel-coded systems for small scale MWRCs.

6.3 System Model of a MWRC

6.3.1 System Overview

We consider an MWRC with L users and a single relay, as shown in Fig. 6.1. The L users exchange their information via the relay, and there is no direct link among the users [183]. The complete information exchange among the users is performed via multiple access phase and broadcast phase. We investigate the case where two users transmit simultaneously at one time in the multiple access phase as shown in Fig. 6.1 (a). We also assume that the relay knows all the channel state information of all the users. This transmission model is similar to that of [184, 185]. When the relay receives the superimposed signal from each pair of users, it computes their corresponding network coded messages. In the broadcast phase, the relay broadcasts the computed messages to the users as shown in Fig. 6.1 (b). After the users receive all the network coded messages from the relay, the relay can retrieve all other users' messages by canceling its own message. In the following, we will focus on the user transmission in the multiple access phase.

In order to let each user decode all other users' messages, an L-user MWRC requires at least $L - 1$ pair-wise uplink transmission in the multiple access phase. The number of downlink transmissions in the broadcast phase is same as the number of uplink transmissions in the multiple access phase. We assume that the channel is in block fading. This means that the channels among the users and the relay remain unchanged during the multiple access phase and the broadcast phase.

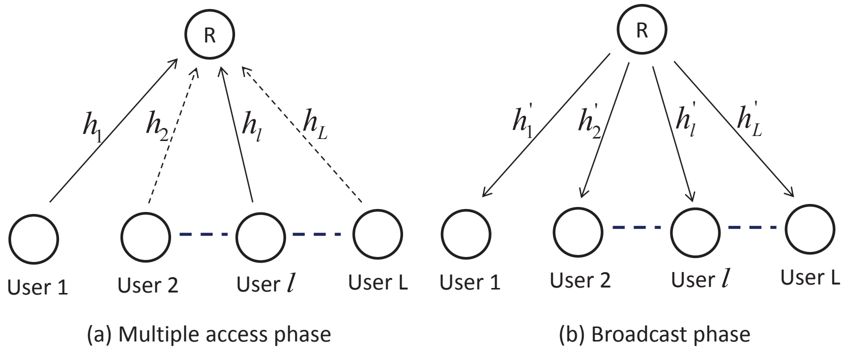


Figure 6.1 System model for both multiple access phase and broadcast phase. In diagram (a), only 2 users transmit simultaneously at one time in the multiple access phase. The solid arrows represent the transmitting user-pair, and the dashed arrows represent that the other users are silent. In diagram (b), the broadcasted messages can be received by all the L users.

6.3.2 Compute-and-Forward in a Pair-Wise MWRC

In this work, we employ nested lattice codes at the users, which follow Section 5.2. Given a transmission user-pair (j, k) in the multiple access phase, let the message for user j be $\boldsymbol{w}_j$, where bold letters here are used to represent row vectors. The corresponding transmitted signal be $\boldsymbol{x}_j = \mathcal{E}(\boldsymbol{w}_j)$, with an average power constraint $\frac{1}{n}E[||\mathcal{E}(\boldsymbol{w}_j)||^2] \leq P$. The received signal at the relay is

$$\boldsymbol{y}_{(j,k)} = h_j\boldsymbol{x}_j + h_k\boldsymbol{x}_k + \boldsymbol{n} \tag{6.1}$$

where $\boldsymbol{n}$ is a complex circularly-symmetric additive white Gaussian noise vector with zero mean and power spectrum density N_0. Let $\boldsymbol{a} = [a_j, a_k]$ denote the computation vector for user-pair (j, k), $a_j, a_k \in \Lambda$ and $a_j, a_k \notin \Lambda'$. The goal for the relay is to decode an R-linear combination of transmitted message

$$\boldsymbol{w}_{(j,k)} = a_j\boldsymbol{w}_j + a_k\boldsymbol{w}_k. \tag{6.2}$$

This computation is based on a scaled version of the received signal $\alpha\boldsymbol{y}$. Let us define φ as the natural projection mapping from Λ onto Λ/Λ' via $\varphi(\boldsymbol{\lambda}) = \boldsymbol{\lambda} + \Lambda'$. The decoder of the LNC can be described by

$$\hat{\boldsymbol{w}}_{(j,k)} = \mathcal{D}(\alpha\boldsymbol{y}|\boldsymbol{h}, \boldsymbol{a}) = \varphi(\mathcal{D}_\Lambda(\alpha\boldsymbol{y})) \tag{6.3}$$

where $\boldsymbol{h} = [h_j, h_k]$, and $\mathcal{D}$ is the lattice quantizer. It has been shown in [23] that the optimum scaling factor is the minimum mean square error (MMSE) coefficient given by

$$\alpha_{\text{MMSE}} = \frac{\boldsymbol{a}\boldsymbol{h}^{\text{H}}\rho}{(1 + ||\boldsymbol{h}||^2\rho)} \tag{6.4}$$

where ρ is the SNR defined as P/N_0, and "H" denotes Hermitian transpose. For a computation vector $\boldsymbol{a}$, the corresponding computation rate is [21, 23]

$$R_{j,k}^{C}(\boldsymbol{h}, \boldsymbol{a}) = \log_2^+ \left(\left(||\boldsymbol{a}||^2 - \frac{\rho|\boldsymbol{a}\boldsymbol{h}^{\text{H}}|^2}{1 + \rho||\boldsymbol{h}||^2} \right)^{-1} \right). \tag{6.5}$$

6.4 Successive Pair-Wise Transmission

Conventional pair-wise transmission in an MWRC is done in a sequential order [184–187]. We term this type of transmission as *successive pair-wise transmission.* We now briefly discuss this conventional transmission scheduling scheme.

Given an L-user single relay MWRC, the users are labeled from 1 to L. At the i-th time slot of the multiple access phase, the scheduled transmission user-pair is $(i, i+1)$. The received signal at time slot i is

$$\boldsymbol{y}_{(i,i+1)} = h_i \boldsymbol{x}_i + h_{i+1} \boldsymbol{x}_{i+1} + \boldsymbol{n}.$$

Then the relay computes the R-linear combination of transmitted message for user-pair $(i, i+1)$ as

$$\boldsymbol{w}_{(i,i+1)} = a_i \boldsymbol{w}_i + a_{i+1} \boldsymbol{w}_{i+1}.$$

There will be in total $L-1$ time slots in the multiple access phase. We can define a pair-wise transmission scheduling matrix $\mathbb{S}$ with size $(L-1) \times L$ to represent this scheme

$$\mathbb{S} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{bmatrix}$$

where the (i, j)-th element $s_{i,j}$ denotes whether user j is activated for transmission at time slot i. Here, $s_{i,j} = 0$ represent silent, and $s_{i,j} = 1$ represent active. After the relay broadcasts all $L - 1$ network coded messages to the users, each user is able to decode all other users' messages.

For the ease of representation, in the following, we omit $\boldsymbol{h}$ and $\boldsymbol{a}$ in the notation of the computation rate for each pair of users. For this transmission scheduling scheme, we can obtain each user's transmission rate. That is

$$R_l < \begin{cases} R^C_{1,2} & \text{if } l = 1 \\ \min\{R^C_{l-1,l}, R^C_{l,l+1}\} & \text{if } l = 2, \cdots, (L-1) \\ R^C_{L-1,L} & \text{if } l = L \end{cases} \tag{6.6}$$

where R_l is the transmission rate for user l, and $R^C_{j,k}$ is given in (6.5). The explanation of (6.6) is as follows: User 1 only transmits in time slot 1 with user 2, so we have $R_1 < R^C_{1,2}$. User L only transmits in time slot $L-1$ with user $L-1$, so we have $R_L < R^C_{L-1,L}$. For user $l \in \{2, 3, \cdots, L-1\}$, it transmits in the $(l-1)$-th and l-th time slots, with previous user and next user respectively. So we have $R_l < \min\{R^C_{l-1,l}, R^C_{l,l+1}\}$. The sum-rate for an MWRC uplink can be expressed as

$$R_{\text{sum}} = \sum_{l=1}^{L} R_l. \tag{6.7}$$

6.5 Opportunistic Pair-Wise Transmission

Successive pair-wise transmission is very simple. However, it does not consider the effect of time-vary fading channel. In this section, we present an opportunistic pair-wise transmission. The key idea is that at each time slot, a pair of users, which has the maximum computation rate, is selected for transmission. In order for the users to recover all others messages from $L - 1$ network codewords forwarded from the relay, the scheduled user-pairs during $L - 1$ time slots should be linearly independent with each other.

We now present the opportunistic pair-wise scheduling algorithm for an L-user single relay MWRC, which is given in Algorithm 6.1.

After $L-1$ times user-pair selection, we can construct a pair-wise transmission scheduling matrix

$$\mathbb{S} = [\boldsymbol{s}_1 \ \boldsymbol{s}_2 \ \cdots \ \boldsymbol{s}_{L-1}]^{\mathrm{T}} \tag{6.15}$$

where T denotes transpose operation. This matrix has a rank of $L-1$ due to the subjected conditions (6.11) and (6.14) during the selection process. This ensures that the users can decode all other users' messages after received these $L-1$ network coded messages.

By employing this opportunistic pair-wise transmission, the transmission rate for user l must be less than the minimum computation rate, of which the l-th user was scheduled for transmission. That is,

$$R_l < \min\{R^C_{j_1,l}, R^C_{j_2,l}, \cdots, R^C_{l,k_1}, R^C_{l,k_2}, \cdots\} \tag{6.16}$$

where $1 \leq j_1, j_2, \cdots \leq l-1$, and $l+1 \leq k_1, k_2, \cdots \leq L$. The users' sum-rate can be expressed as

$$R_{\text{sum}} = \sum_{l=1}^{L} R_l. \tag{6.17}$$

This opportunistic pair-wise transmission can be further extended to scheduling more than two users to transmit in each time slot, to take the advantage of the compute-and-forward scheme for multi-user systems.

6.6 Simulation Results

6.6.1 Linear Network Codes

In this part, we consider two linear network codes: Gaussian integer based LNC and Eisenstein integer based LNC. A Gaussian integer is such a complex number that its real and imaginary part are both integers. It can be represented as

$$\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\} \tag{6.18}$$

Algorithm 6.1 Opportunistic pair-wise transmission

```
We label the users from 1 to $L$.
For time slot 1, we select user-pair $l_1$ and $l_2$ such that
    $(l_1, l_2) = \arg\max\{R^C_{l_1,l_2} | l_1, l_2 \in \{1, \cdots, L\}, l_1 < l_2\}$    (6.8)
    where $R^C_{l_1,l_2}$ is given in (6.5). Note that for each user-pair $(l_1, l_2)$, the computation rate $R^C_{l_1,l_2}$ should be also maximized by choosing optimal $\boldsymbol{a} = [a_{l_1}, a_{l_2}]$. This can be effectively done via Gaussian reduction algorithm [21].
We form a row vector with length $L$
    $\boldsymbol{s}_1 = [0 \ \cdots \ 0 \ \underset{\overset{\uparrow}{l_1}}{1} \ 0 \ \cdots \ 0 \ \underset{\overset{\uparrow}{l_2}}{1} \ 0 \ \cdots \ 0]$    (6.9)
    where $\boldsymbol{s}_1[l_1] = \boldsymbol{s}_1[l_2] = 1$ and $\boldsymbol{s}_1[l] = 0$, $l \in \{1, \cdots, L\}$, $l \neq l_1$, $l \neq l_2$. Note that this row vector is binary, which is referred to as *pair-wise user selection vector* (PSV).
For time slot 2, we select user-pair $l_3$ and $l_4$ such that
    $(l_3, l_4) = \arg\max\{R^C_{l_3,l_4} | l_3, l_4 \in \{1, \cdots, L\}, l_3 < l_4\}$    (6.10)
    s.t.
    $\boldsymbol{s}_2 = [0 \ \cdots \ 0 \ \underset{\overset{\uparrow}{l_3}}{1} \ 0 \ \cdots \ 0 \ \underset{\overset{\uparrow}{l_4}}{1} \ 0 \ \cdots \ 0]$ and $\boldsymbol{s}_2 \neq \boldsymbol{s}_1$.    (6.11)
We now define the following set
    $\mathcal{F}(\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_i) = \left\{ \sum_{t=1}^{i} \kappa_t \boldsymbol{s}_t, \kappa_t \in \{0, 1\} \right\}$    (6.12)
    where the summation is modulo-2 addition. This set forms an $i$-dimensional subspace over binary field, which is spanned by the binary PSVs $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\cdots$, $\boldsymbol{s}_i$.
For the $i$-th time slot, $2 < i \leq L - 1$, we select the pair of users, such that
    $(j, k) = \arg\max\{R^C_{j,k} | j, k \in \{1, \cdots, L\}, j < k\}$    (6.13)
    s.t.
    $\boldsymbol{s}_i \notin \mathcal{F}(\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{i-1})$.    (6.14)
    The constraint in (6.14) ensures that the PSV of the user-pair selected in the $i$-th time slot can not belong to the subspace spanned by $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\cdots$, $\boldsymbol{s}_{i-1}$. That is, new message will be received by the relay for the $i$-th time slot.
After $L - 1$ selection, stop.
```

where $i = \sqrt{-1}$. The norm of a Gaussian integer can be calculated as $a^2 + b^2$. A $\mathbb{Z}[i]$-based LNC is hypercube shaped.

An Eisenstein integer can be written as

$$\mathbb{Z}[\omega] = \{a + b\omega | a, b \in \mathbb{Z}\} \tag{6.19}$$

where $\omega = \frac{-1+\sqrt{-3}}{2}$. The norm of an Eisenstein integer is $a^2 + b^2 - ab$. A $\mathbb{Z}[\omega]$-based LNC is hexagonal shaped. More details of $\mathbb{Z}[i]$-based and $\mathbb{Z}[\omega]$-based LNC and their quantization operation can be found in [21].

In this part, we consider a message space $W = \mathbb{F}_{13} \cong \mathbb{Z}[i]/\beta\mathbb{Z}[i] \cong \mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$, where $\beta = 2+3i$ and $\gamma = 4+3\omega$. The corresponding lattice, sublattice, and fundamental Voronoi regions of $\beta\mathbb{Z}[i]$ or $\gamma\mathbb{Z}[\omega]$ and constellations of $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ or $\mathbb{Z}[\omega]/\gamma\mathbb{Z}[\omega]$ are shown in Fig. 5.2(a) or 5.2(b). In this case, we have

$$\mathcal{E}_{\mathbb{Z}[i]}(W) = \{0, \pm 1, \pm i, \pm(1+i), \pm(1-i), \pm 2, \pm 2i\}$$

and

$$\mathcal{E}_{\mathbb{Z}[\omega]}(W) = \{0, \pm 1, \pm \omega, \pm(1+\omega), \pm(1-\omega), \pm(1+2\omega), \pm(2+\omega)\}.$$

6.6.2 Users Transmission Sum-Rate

We compare the users' average uplink transmission sum-rate between successive pair-wise transmission and proposed opportunistic pair-wise transmission for 3-user and 4-user cases. We assume Rayleigh faded channels. At each SNR, the sum-rate is averaged over 10^5 channel realizations. Fig. 6.2 and Fig. 6.3 show the sum-rate for $\mathbb{Z}[i]$-based and $\mathbb{Z}[\omega]$-based LNCs in MWRCs, respectively, with 3 and 4 users. We can see that the proposed opportunistic pair-wise transmission has significantly improved sum-rate compared to the successive pair-wise transmission. For example, for a 4-user MWRC, a 2 bits/s/Hz improvement is observed at 30 dB, and a 1.25 bits/s/Hz improvement is observed for a 3-user MWRC. Larger improvements can be achieved at higher SNR.

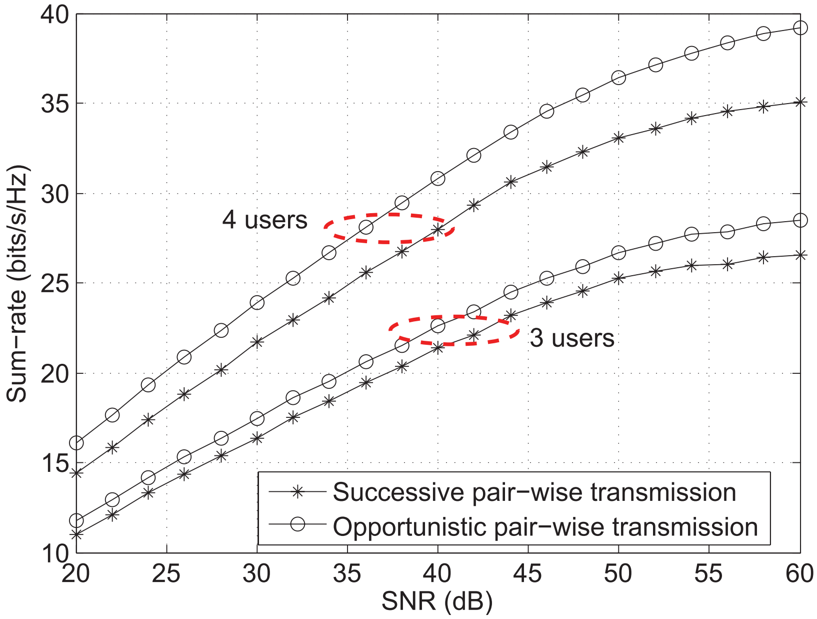


Figure 6.2 Users transmission sum-rate for the $\mathbb{Z}[i]$-based LNC in MWRCs.

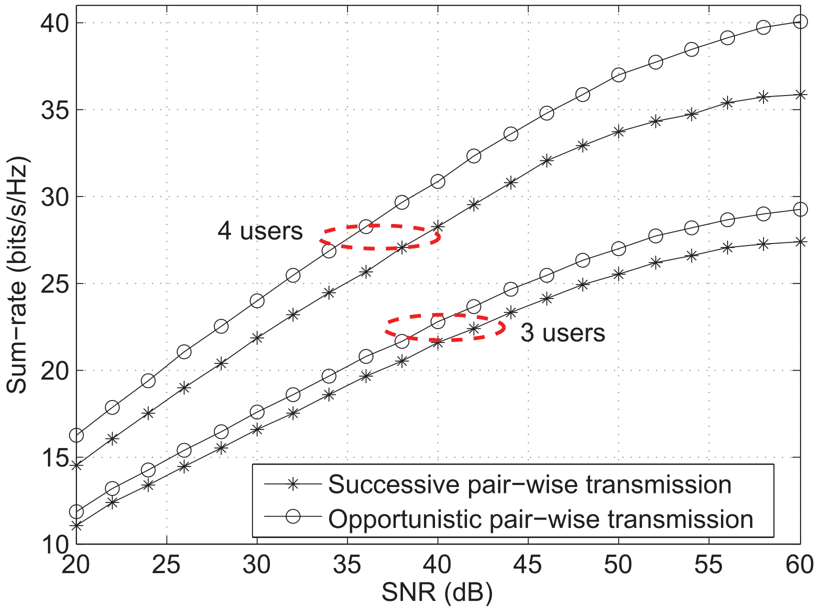


Figure 6.3 Users transmission sum-rate for the $\mathbb{Z}[\omega]$-based LNC in MWRCs.

6.6.3 Error Performance

We firstly compare uncoded MWRCs. We compare the symbol-error-rate (SER) performance between the opportunistic pair-wise transmission and the successive pair-wise transmission at the relay. A computation error for the user-pair (j, k) is declared if the computation result $\widehat{\boldsymbol{w}}_{j,k} \neq \boldsymbol{w}_{j,k}$. For each SNR, we collect 200 symbol errors. Note that in this paper we consider a disjointed decoding strategy at the relay. If a joint decoding strategy is employed, better performance may be achieved [187].

Fig. 6.4 and Fig. 6.5 show the SER performance for $\mathbb{Z}[i]$-based LNC in uncoded 3-user and 4-user MWRCs, respectively. The figures illustrate that the opportunistic pair-wise transmission has a 3 dB gain for an uncoded 3-user system and a 4.5 dB gain for an uncoded 4-user system at the 10^{-2} SER level, when compared to successive pair-wise transmission. Fig. 6.6 and Fig. 6.7 show the SER performance for $\mathbb{Z}[\omega]$-based LNC in uncoded 3-user and 4-user MWRCs, respectively. Similar performance improvements are observed for $\mathbb{Z}[\omega]$-based LNC.

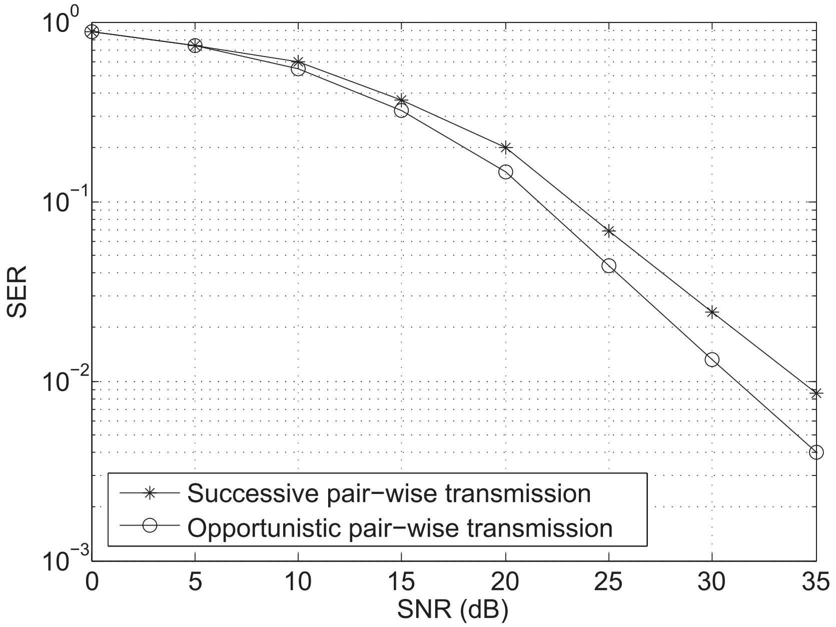


Figure 6.4SER for the $\mathbb{Z}[i]$-based LNC in an uncoded 3-user MWRC.

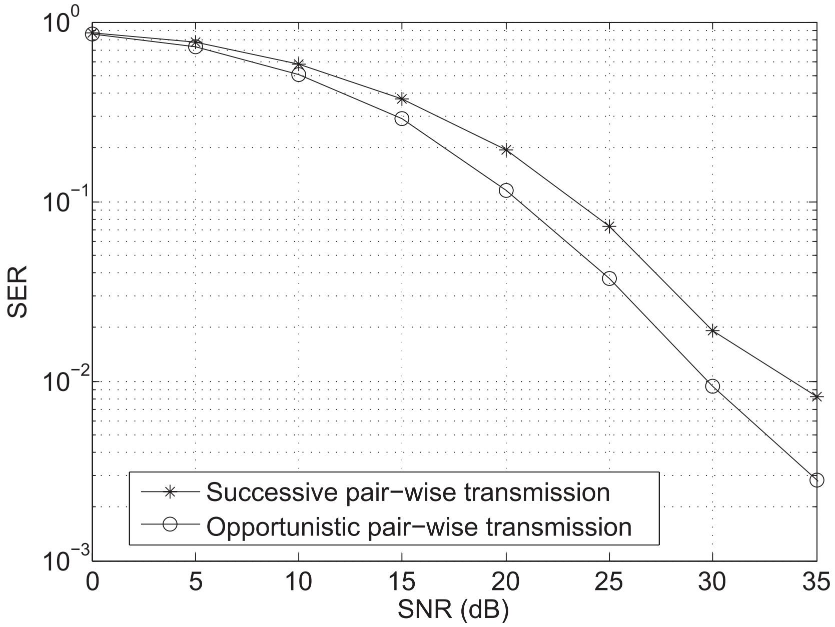


Figure 6.5SER for the $\mathbb{Z}[i]$-based LNC in an uncoded 4-user MWRC.

We now incorporate the LNCs into this system. We construct rate-1/2 convolutional codes $\mathcal{C}$ over $\mathbb{Z}[i]/(2+3i)\mathbb{Z}[i] \cong \mathbb{F}_{13}$ and over $\mathbb{Z}[\omega]/(4+3\omega)\mathbb{Z}[\omega] \cong \mathbb{F}_{13}$ with memory order $v = 1$ and 2. The construction method is based on Complex Construction A. The constructed codes have maximized minimum squared Euclidean norm of non-zero codewords of $\mathcal{C}$ ($w_E^{min}(\mathcal{C})$) and minimized K, representing the number of codewords with $w_E^{min}(\mathcal{C})$. The codes parameters are listed in Table 6.1.

We now compare the frame error rate (FER) when the designed convolutional LNCs with memory order-1 are employed. The information sequence length is set to 99. We collect 200 frame errors at each SNR. Fig. 6.8 and Fig. 6.9 show the FER performance for $\mathbb{Z}[i]$-based and $\mathbb{Z}[\omega]$-based LNCs, respectively, in a channel-coded 4-user MWRC. We can see that even the memory order-1 linear lattice network codes can provide a significant performance improvement. Specifically, a 2.5 dB gain at the 10^{-2} FER level is observed for the 4-user system.

Table 6.1 Rate-1/2 convolutional codes over $\mathbb{F}_{13}$

v	**Type**	$\mathbf{g}(D)$	$w_E^{min}(\mathcal{C})$	K
1	$\mathbb{Z}[i]$	$(1) + (2i)D$	8	4
		$(1 - i) + (i)D$		
	$\mathbb{Z}[\omega]$	$(-1 + w) + (w)D$	8	12
		$(1 + w) + (2 + w)D$		
2	$\mathbb{Z}[i]$	$(1 - i) + (-2)D + (-i)D^2$	12	4
		$(i) + (2)D + (1 + i)D^2$		
	$\mathbb{Z}[\omega]$	$(-2 - \omega) + (2 + \omega)D + (-1 - 2\omega)D^2$	12	24
		$(-1) + (-\omega)D + (-1 - \omega)D^2$		

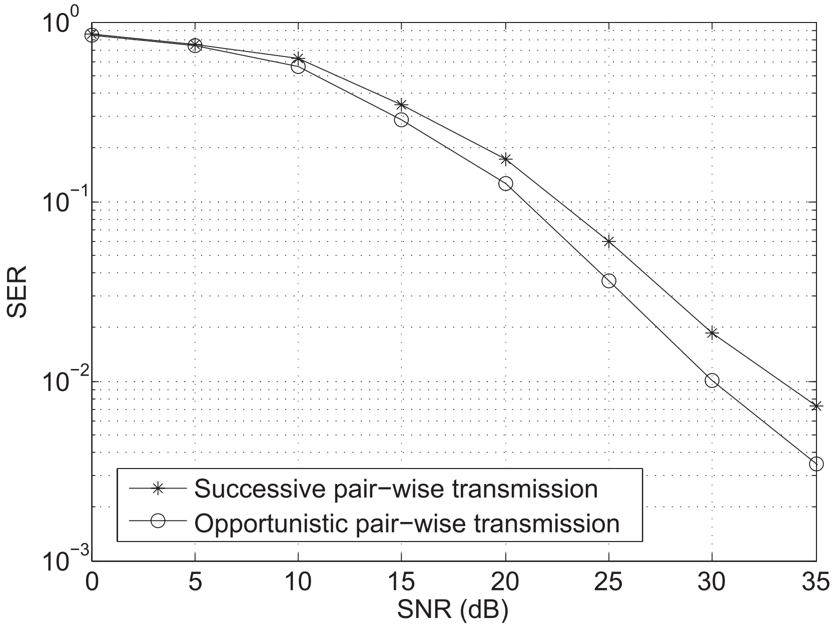


Figure 6.6SER for the $\mathbb{Z}[\omega]$-based LNC in an uncoded 3-user MWRC.

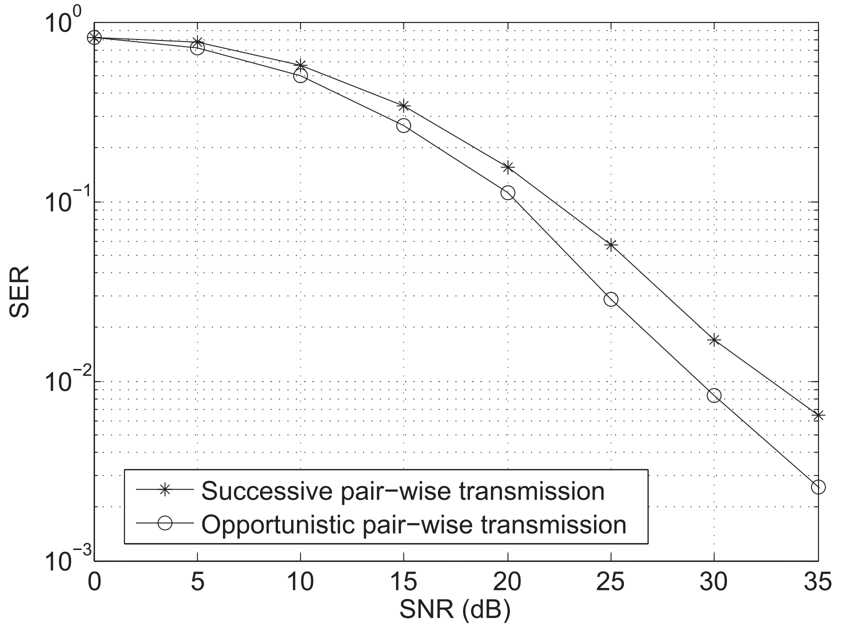


Figure 6.7SER for the $\mathbb{Z}[\omega]$-based LNC in an uncoded 4-user MWRC.

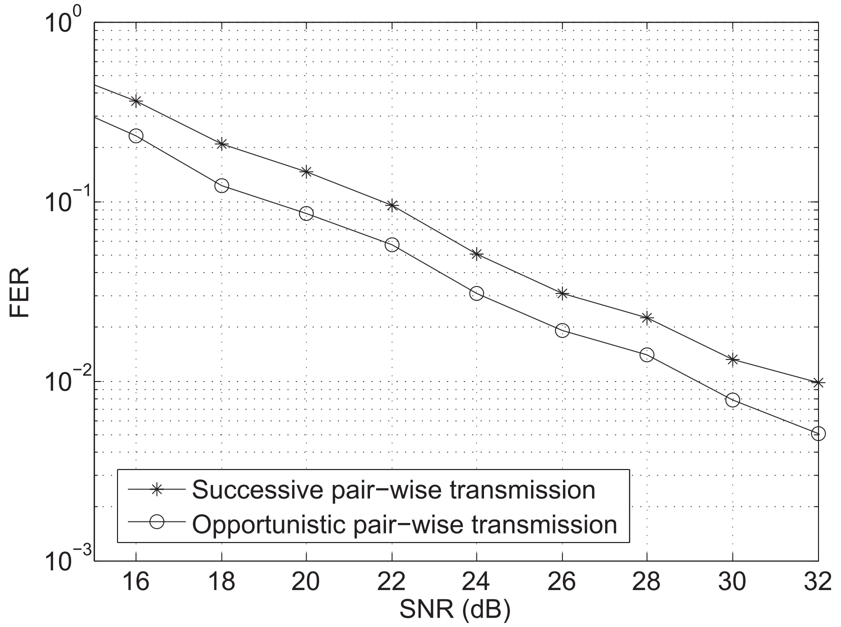


Figure 6.8FER for the $\mathbb{Z}[i]$-based LNC in a channel-coded 4-user MWRC.

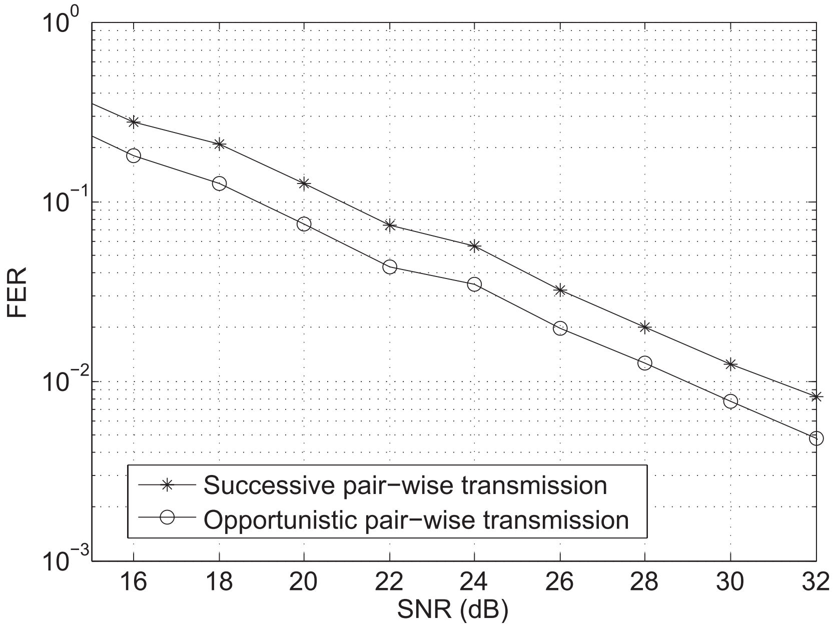


Figure 6.9FER for the $\mathbb{Z}[\omega]$-based LNC in a channel-coded 4-user MWRC.

6.7 Conclusion

In this chapter, we focus on design a transmission scheme that can improve the system sum-rate and error-rate performance in the uplink phase. We proposed an opportunistic pair-wise transmission in MWRCs based on pair-wise compute-and-forward relaying. High level modulation with nested lattice codes is considered to improve the sum-rate of multi-user transmission. The comparison between the opportunistic pair-wise transmission and the successive pair-wise transmission shows a 1.25 bits/s/Hz and a 2 bits/s/Hz sum-rate improvement at 30 dB for 3-user and 4-user MWRCs respectively. We also demonstrated that, this novel transmission scheme can achieve $3 \sim 4.5$ dB gain at the 10^{-2} SER level for uncoded small scale MWRCs with 3 to 4 users, and 2.5 dB gain at the 10^{-2} FER level for channel-coded 4-user MWRCs.

Chapter 7

Optimization of MIMO MWRCs with PNC

7.1 Introduction

In this chapter, we extend our study in Chapter 6 from single-input single-output multi-way relay channels (MWRCs) to multi-input multi-output (MIMO) MWRCs with PNC. In particular, we investigate the degrees of freedom and sum-rate optimization of half-duplex MIMO MWRC with full data exchange.

This chapter starts by introducing the background of this work. We then describe the detailed system model. After that, we focus on the DoF capacity of the considered MIMO MWRCs with fixed channel uses. We optimize the DoF results with channel use allocations for uplink phase and downlink phase. Following the DoF analysis, we shift our focus to the system sum-rate optimization. Finally, we conclude this chapter with some highlights on our results.

7.2 Background

As studied in previous section, in an MWRC, multiple users with no direct links exchange information with the help of a single relay node. In [93], two data exchange models, namely, full data exchange and pairwise data exchange were introduced in MWRCs. In full data exchange, each user wants to decode all the messages in the system. In pairwise data exchange, users in the system are grouped into pairs, and the two users in each pair exchange private messages via the relay node. These data exchange models have been studied, e.g., in [92, 189, 190].

Multiple-input multiple-output (MIMO) techniques have been introduced into MWRCs to allow spatial multiplexing [99–105, 109]. The degrees of freedom (DoF) is an important metric to understand the capacity behavior of the MIMO MWRC. The DoF analysis for MIMO MWRCs in the existing literature [99–101] is mostly focused on pairwise data exchange. In particular, the authors in [100] considered a three-user MWRC, termed the MIMO Y channel, and the DoF capacity of the MIMO Y channel was derived under certain relay/user antenna setups. The work in [99] generalized the result of [100] to the case of an arbitrary number of users. Later, the authors in [101] considered MWRCs with clustered data exchange, i.e., the users in the network are grouped into clusters, and only the users in the same cluster communicate with each other. It's worth noting that, in pairwise data exchange, the traffic loads of the uplink and the downlink are symmetric. This uplink/downlink symmetry further implies that the signal space alignment for the uplink straightforwardly carries over to the downlink. This property is used in [99–101] to simplify the beamforming design for MIMO MWRCs with pairwise data exchange.

In this chapter, we focus on the design of communication mechanisms over MIMO MWRCs with full data exchange. We derive the DoF capacity of the MIMO MWRC with full data exchange operated in half-duplex and full-duplex modes. For many communication networks, the half-duplex DoF capacity is simply one half of the full-duplex capacity; see, e.g., [99–101]. Interestingly, this is not the case for MIMO

MWRCs with full data exchange. The fundamental reason is that, unlike pairwise data exchange, the uplink and downlink traffic loads are asymmetric in full data exchange. More specifically, in full data exchange, as every user wants to learn all the messages from the other users, the downlink is usually the throughput bottleneck. Half-duplexing allows unequal time allocation between the uplink and the downlink, and therefore enables a DoF higher than half of the corresponding full-duplex DoF capacity. We derive the optimal uplink/downlink time allocation to maximize the DoF of the half-duplex system. We show that a significant DoF gain can be achieved by the optimal uplink/downlink time allocation, as compared with equal time allocation.

Further, we investigate the achievable rates of the considered MIMO MWRCs. We show that the sum-rate is a non-convex function of the user precoders and relay precoder. Thus, the sum-rate maximization problem cannot be tackled trivially. In this chapter, we propose an iterative algorithm to optimize the user precoders and the relay precoder in an alternating fashion. Numerical results demonstrate that the system performance can be considerably improved by a careful design of the user and relay precoders. We also show that the numerical results agree with the DoF analysis obtained in this chapter.

7.3 System Model

7.3.1 System Overview

In this chapter, we consider a MIMO MWRC, in which K users exchange information with the help of a single relay, as illustrated in Fig. 7.1. There is no direct link between any two users. Each user is equipped with M antennas, and the relay with N antennas. The information exchange model is assumed to be full data exchange, i.e., each user broadcasts its message to all the other users in the network, and wants to decode all the messages in the system. Throughout this chapter, we assume that perfect channel state information (CSI) is available at all nodes.

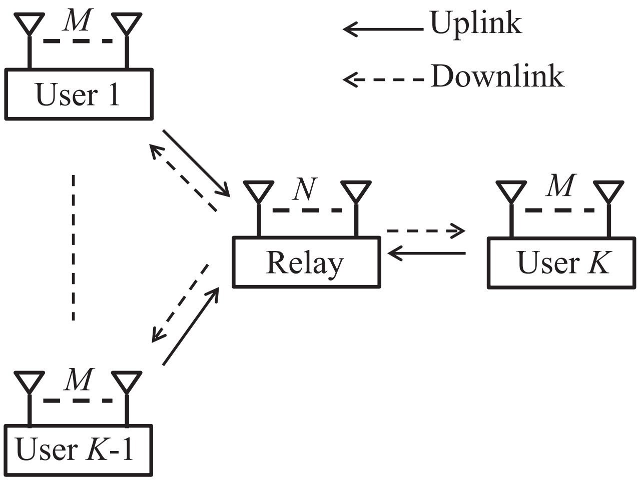


Figure 7.1A MIMO MWRC with full data exchange, where each user wants to learn all the messages in the system.

The system operates in the half-duplex mode, i.e., a node cannot transmit and receive signals simultaneously. Each round of information exchange consists of two phases, namely, an uplink phase and a downlink phase. In the uplink phase, all users transmit to the relay simultaneously using a common frequency band. In the downlink phase, the relay broadcasts to the users. More details are described below.

7.3.2 Uplink Phase

Assume that the uplink phase consists of T_{u} channel uses. In the t_{u}-th channel use, the received signal vector $\mathbf{y}_{\mathrm{R}}(t_{\mathrm{u}})$ at the relay is given by

$$\mathbf{y}_{\mathrm{R}}(t_{\mathrm{u}}) = \sum_{i=1}^{K} \mathbf{H}_i(t_{\mathrm{u}})\mathbf{x}_i(t_{\mathrm{u}}) + \mathbf{n}_{\mathrm{R}}(t_{\mathrm{u}}),\ t_{\mathrm{u}} = 1, \cdots, T_{\mathrm{u}},$$

where $\mathbf{H}_i(t_{\mathrm{u}}) \in \mathbb{C}^{N\times M}$ denotes the channel matrix from user i to relay R in the t_{u}-th uplink channel use and its elements are i.i.d. drawn from $\mathcal{CN}(0,1)$, $\mathbf{x}_i(t_{\mathrm{u}}) \in \mathbb{C}^{M\times 1}$ is the transmitted signal from user i, and $\mathbf{n}_{\mathrm{R}}(t_{\mathrm{u}})$ is an additive circularly-symmetric Gaussian noise vector drawn from $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$.

Denote $\mathbf{y}_{\mathrm{R}} = [\mathbf{y}_{\mathrm{R}}^{\mathrm{T}}(1)\cdots\mathbf{y}_{\mathrm{R}}^{\mathrm{T}}(T_{\mathrm{u}})]^{\mathrm{T}}$, $\mathbf{x}_i = [\mathbf{x}_i^{\mathrm{T}}(1)\cdots\mathbf{x}_i^{\mathrm{T}}(T_{\mathrm{u}})]^{\mathrm{T}}$, $\mathbf{n}_{\mathrm{R}} = [\mathbf{n}_{\mathrm{R}}^{\mathrm{T}}(1)\cdots\mathbf{n}_{\mathrm{R}}^{\mathrm{T}}(T_{\mathrm{u}})]^{\mathrm{T}}$, and

$$\begin{aligned}\mathbf{H}_i &= \mathrm{diag}\{\mathbf{H}_i(1), \cdots, \mathbf{H}_i(T_\mathrm{u})\}\\ &= \begin{bmatrix} \mathbf{H}_i(1) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{H}_i(T_\mathrm{u}) \end{bmatrix}.\end{aligned}$$

Then

$$\begin{aligned}\mathbf{y}_\mathrm{R} &= \sum_{i=1}^{K} \mathbf{H}_i \mathbf{x}_i + \mathbf{n}_\mathrm{R} && (7.1)\\ &= \mathbf{H}\mathbf{x} + \mathbf{n}_\mathrm{R}, && (7.2)\end{aligned}$$

where $\mathbf{x} = [\mathbf{x}_1^\mathrm{T} \cdots \mathbf{x}_K^\mathrm{T}]^\mathrm{T}$, and $\mathbf{H} = [\mathbf{H}_1 \cdots \mathbf{H}_K]$. The transmission power of each user is limited to P_U per channel use, i.e.

$$\frac{1}{T_\mathrm{u}} \mathrm{Tr}(\mathbf{x}_i \mathbf{x}_i^\mathrm{H}) \leq P_\mathrm{U},\ i \in \{1, \cdots, K\}. \qquad (7.3)$$

7.3.3 Downlink Phase

We now consider the downlink phase which, without loss of generality, consists of T_d channel uses. In the t_d-th downlink channel use, the received signal at user i, $i = 1, \cdots, K$, is

$$\mathbf{y}_i(t_\mathrm{d}) = \mathbf{G}_i(t_\mathrm{d})\mathbf{x}_\mathrm{R}(t_\mathrm{d}) + \mathbf{n}_i(t_\mathrm{d}),\ t_\mathrm{d} = 1, \cdots, T_\mathrm{d},$$

where $\mathbf{x}_\mathrm{R}(t_\mathrm{d})$ is the signal transmitted at the relay in the t_d-th downlink channel use, $\mathbf{G}_i(t_\mathrm{d}) \in \mathbb{C}^{M\times N}$ denotes the channel matrix from the relay to user i in the t_d-th downlink channel use and its entries are i.i.d. drawn from $\mathcal{CN}(0, 1)$, and $\mathbf{n}_i(t_\mathrm{d})$ is an additive circularly-symmetric Gaussian noise vector drawn from $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$.

Let $\mathbf{y}_i$ be the received signal at user i in the downlink phase, $i = 1, \cdots, K$. It can be written as

$$\mathbf{y}_i = \mathbf{G}_i \mathbf{x}_\mathrm{R} + \mathbf{n}_i,\ i = 1, \cdots, K, \qquad (7.4)$$

where

$$\mathbf{x}_{\mathrm{R}} = [\mathbf{x}_{\mathrm{R}}^{\mathrm{T}}(1) \cdots \mathbf{x}_{\mathrm{R}}^{\mathrm{T}}(T_{\mathrm{d}})]^{\mathrm{T}},$$

and

$$\mathbf{y}_i = [\mathbf{y}_i^{\mathrm{T}}(1) \cdots \mathbf{y}_i^{\mathrm{T}}(T_{\mathrm{d}})]^{\mathrm{T}}.$$

$\mathbf{G}_i = \mathrm{diag}\{\mathbf{G}_i(1), \cdots, \mathbf{G}_i(T_{\mathrm{d}})\}$ is a block diagonal matrix which contains all the channel matrixes from the relay to the user i in the downlink phase. The $\mathbf{n}_i = [\mathbf{n}_i^{\mathrm{T}}(1) \cdots \mathbf{n}_i^{\mathrm{T}}(T_{\mathrm{d}})]^{\mathrm{T}}$ is the noise vector at user i in the downlink phase. The transmission power of the relay is limited to P_{R} per channel use, i.e.

$$\frac{1}{T_{\mathrm{d}}} \mathrm{Tr}(\mathbf{x}_{\mathrm{R}} \mathbf{x}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}.$$

At the end of the downlink phase, each user i, $i = 1, \cdots, K$, firstly removes self-information from its received signal $\mathbf{y}_i$. Let $\mathbf{y}_{\bar{i}}$ be the received signal at user i with its self-information removed. Let $\mathbf{x}_{\bar{i}} = [\mathbf{x}_1^{\mathrm{T}} \cdots \mathbf{x}_{i-1}^{\mathrm{T}} \mathbf{x}_{i+1}^{\mathrm{T}} \cdots \mathbf{x}_K^{\mathrm{T}}]^{\mathrm{T}}$ be the true signal vector that targeted to decode at user i. Each user i decodes an estimated signal $\widehat{\mathbf{x}}_{\bar{i}} = [\widehat{\mathbf{x}}_1^{\mathrm{T}} \cdots \widehat{\mathbf{x}}_{i-1}^{\mathrm{T}} \widehat{\mathbf{x}}_{i+1}^{\mathrm{T}} \cdots \widehat{\mathbf{x}}_K^{\mathrm{T}}]^{\mathrm{T}}$ of all the other users from $\mathbf{y}_{\bar{i}}$. The detailed information flows are illustrated in Fig. 7.2, where virtual antennas are used at each node to represent multiple channel uses in the uplink phase and the downlink phase.

Note that all entries of the uplink/downlink channel matrices are independently drawn from $\mathcal{CN}(0, 1)$. This ensures that all the channel matrices are of full rank with probability one. For example, $\mathrm{rank}(\mathbf{H}_i(t_{\mathrm{u}})) = \min\{M, N\}$.

7.3.4 Linear Precoding Techniques

For ease of further exposition, we focus on the linear precoding techniques for the transmitters. We firstly consider the linear precoding at the users. Let $\mathbf{s}_i \in \mathbb{C}^{L \times 1}$ be the information vector transmitted from user i, $i = 1, \cdots, K$, in the uplink phase, where $L \leq MT_{\mathrm{u}}$. Assume the entries of the information vector for each user are independent and identically distributed (i.i.d.) drawn from $\mathcal{CN}(0, 1)$. Let $\mathbf{P}_i \in$

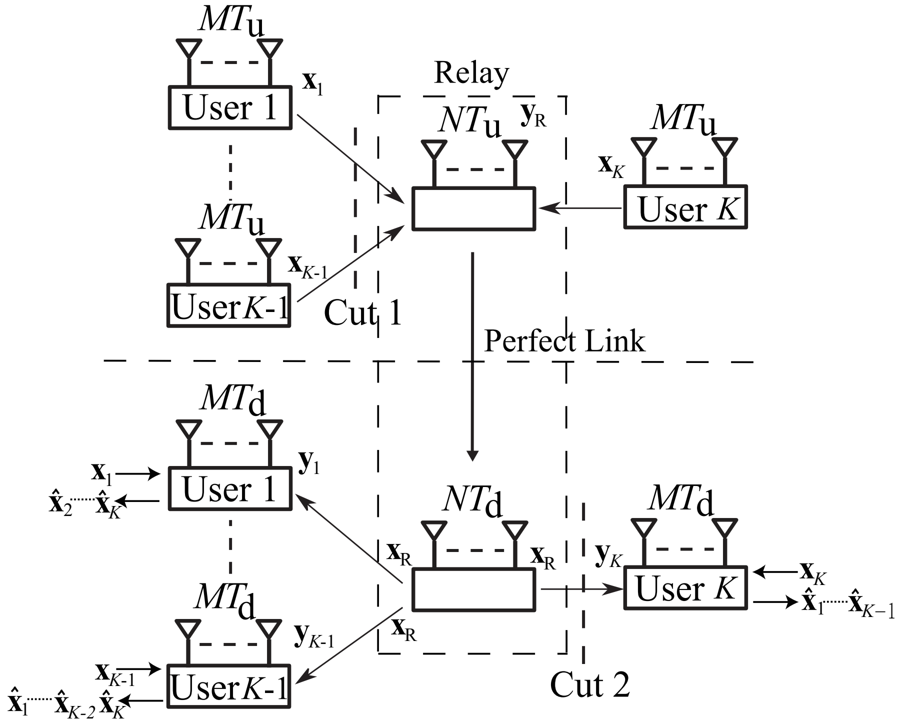


Figure 7.2 Information flows in the uplink phase and the downlink phase. In this figure, the relay node is artificially separated into 2 nodes, with a perfect link (i.e., the output of the perfect link is equal to the input of this link).

$\mathbb{C}^{MT_{\rm u}\times L}$ be the linear precoder at user i, $i = 1, \cdots, K$, in the uplink phase. Then the signal $\mathbf{x}_i$ transmitted at each user i for the uplink phase is

$$\mathbf{x}_i = \mathbf{P}_i\mathbf{s}_i. \tag{7.5}$$

In this chapter, we assume amplify-and-forward (AF) relaying. Later we will show that AF relay operation achieves the optimal DoF for the MIMO MWRC with full data exchange. With AF relaying, the transmitted signal $\mathbf{x}_{\rm R}$ at the relay in the downlink phase is given by

$$\mathbf{x}_{\rm R} = \mathbf{F}\mathbf{y}_{\rm R}, \tag{7.6}$$

where $\mathbf{F} \in \mathbb{C}^{NT_{\rm d}\times NT_{\rm u}}$ is the relay precoder. Substitute (7.6) into (7.4), the received signal of user i, $i = 1, \cdots, K$, is written as

$$\begin{aligned}\mathbf{y}_i &= \mathbf{G}_i\mathbf{F}\mathbf{H}\mathbf{x} + \mathbf{G}_i\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i \\ &= \underbrace{\mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{x}_{\bar{i}}}_{\text{Target Signal}} + \underbrace{\mathbf{G}_i\mathbf{F}\mathbf{H}_i\mathbf{x}_i}_{\text{Self interference}} + \underbrace{\mathbf{G}_i\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i}_{\text{Effective noise}}, \qquad (7.7)\end{aligned}$$

where $\mathbf{H}_{\bar{i}} = [\mathbf{H}_1 \cdots \mathbf{H}_{i-1}\mathbf{H}_{i+1} \cdots \mathbf{H}_K]$, and $\mathbf{x}_{\bar{i}} = [\mathbf{x}_1^\mathrm{T} \cdots \mathbf{x}_{i-1}^\mathrm{T}\mathbf{x}_{i+1}^\mathrm{T} \cdots \mathbf{x}_K^\mathrm{T}]^\mathrm{T}$. The signal received at user i after self interference cancellation can be written as

$$\mathbf{y}_{\bar{i}} = \mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{x}_{\bar{i}} + \mathbf{G}_i\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i. \qquad (7.8)$$

Substitute (7.5) into (7.8), we have

$$\begin{aligned}\mathbf{y}_{\bar{i}} &= \mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}\mathbf{s}_{\bar{i}} + \mathbf{G}_i\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i \qquad (7.9) \\ &= \sum_{j=1, j\neq i}^{K} \mathbf{G}_i\mathbf{F}\mathbf{H}_j\mathbf{P}_j\mathbf{s}_j + \mathbf{G}_i\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i, \qquad (7.10)\end{aligned}$$

where $\mathbf{P}_{\bar{i}} = \mathrm{diag}\{\mathbf{P}_1, \cdots, \mathbf{P}_{i-1}, \mathbf{P}_{i+1}, \cdots, \mathbf{P}_K\}$ and $\mathbf{s}_{\bar{i}} = [\mathbf{s}_1^\mathrm{T} \cdots \mathbf{s}_{i-1}^\mathrm{T}\mathbf{s}_{i+1}^\mathrm{T} \cdots \mathbf{s}_K^\mathrm{T}]^\mathrm{T}$.

7.3.5 Achievable Rate Region

It is worth noting that, the overall channel seen at user i in one round of information exchange, $i = 1, \cdots, K$, is an equivalent multiple access channel, since each user wants to decode all the messages from the other $K - 1$ users, as shown in (7.10). Thus, there are $2^{K-1} - 1$ rate constraints for each index i. In total, there are $K \cdot (2^{K-1} - 1)$ rate constraints in the network. The corresponding achievable rate region is detailed as follows.

At user i, $i = 1, \cdots, K$, define

$$\mathcal{I}_i \triangleq \{1, \cdots, i-1, i+1, \cdots, K\}.$$

For a subset $\mathcal{S} \subseteq \mathcal{I}_i$, its complementary set is denoted by $\mathcal{S}^\mathrm{c}$. Then, the achievable rate region at user i , $i = 1, \cdots, K$, is given by

$$\begin{aligned}\mathcal{R}_i : \sum_{j, j\in\mathcal{S}} R_j &\leq \frac{1}{T} I\left(\{\mathbf{s}_j | j \in \mathcal{S}\}; \mathbf{y}_{\bar{i}} | \{\mathbf{s}_{j'} | j' \in \mathcal{S}^{\mathrm{c}}\}\right) \\ &= \frac{1}{T}\log\left(\frac{\left|\sum\limits_{j,j\in\mathcal{S}} \mathbf{G}_i\mathbf{F}\mathbf{H}_j\mathbf{P}_j\mathbf{P}_j^{\mathrm{H}}\mathbf{H}_j^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \sigma^2(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I})\right|}{|\sigma^2(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I})|}\right), \ \mathcal{S} \subseteq \mathcal{I}_i,\end{aligned} \tag{7.11}$$

where $T = T_{\mathrm{u}} + T_{\mathrm{d}}$ is the total number of channel uses in one round of information exchange.

An achievable rate region $\mathcal{R}$ for the considered system is then given as

$$\mathcal{R} : \bigcap_{i=1}^{K} \mathcal{R}_i. \tag{7.12}$$

7.3.6 Degrees of Freedom (DoF)

For notational convenience, we assume $P_{\mathrm{U}} = P_{\mathrm{R}} = P$ and denote the signal-to-noise ratio (SNR) as $\rho = P/\sigma^2$ without compromising the generality of the DoF results in this chapter.

Denote the decoding rate at user i, $i = 1, \cdots, K$, at SNR ρ, as

$$R_{\mathrm{r},i}(\rho) \triangleq \sum_{j=1,\ j\neq i}^{K} R_j(\rho),$$

where $R_j(\rho)$ is the rate of user j. The decoding rate $R_{\mathrm{r},i}(\rho)$ is achievable if user i decodes the messages from all the other $K-1$ users with vanishing error probability, i.e., $\Pr(\widehat{\mathbf{s}}_{\bar{i}} \neq \mathbf{s}_{\bar{i}}) \to 0$ when $T \to \infty$, $i \in \{1, \cdots, K\}$, where $\widehat{\mathbf{s}}_{\bar{i}} = [\widehat{\mathbf{s}}_1^{\mathrm{T}} \cdots \widehat{\mathbf{s}}_{i-1}^{\mathrm{T}} \widehat{\mathbf{s}}_{i+1}^{\mathrm{T}} \cdots \widehat{\mathbf{s}}_K^{\mathrm{T}}]^{\mathrm{T}}$ is the estimated version of $\mathbf{s}_{\bar{i}}$ at user i. In other words, $R_{\mathrm{r},i}(\rho)$ falls into the rate region $\mathcal{R}_i$ defined in (7.11).

The sum-rate per channel use is defined as

$$R(\rho) \triangleq \sum_{i=1}^{K} R_{\mathrm{r},i}(\rho), \tag{7.13}$$

where achievable rate tuple $(R_{\mathrm{r},1}, \cdots, R_{\mathrm{r},K})$ falls into the rate region $\mathcal{R}$ defined in (7.12). Define an achievable total DoF as

$$d \triangleq \lim_{\rho\to\infty} \frac{R(\rho)}{\log \rho}. \tag{7.14}$$

The maximum of d in (7.14) over all achievable schemes is referred to as the sum-DoF capacity, or simply, the DoF capacity.

7.4 Degrees of Freedom with Fixed Channel Uses

In this section, we present the DoF analysis of the half-duplex MIMO MWRC with fixed uplink/downlink time allocation, i.e., T_{u} and T_{d} are fixed.

7.4.1 Main Result

Theorem 7.1. *For the K-user half-duplex MIMO MWRC described in Section 7.3, the DoF capacity is given by*

$$d = \begin{cases} \dfrac{K}{T}\min\{(K-1)MT_{\mathrm{u}}, MT_{\mathrm{d}}\}, & \dfrac{M}{N} \in \left(0, \dfrac{1}{K-1}\right] & (7.15)\\ \dfrac{K}{T}\min\{NT_{\mathrm{u}}, MT_{\mathrm{d}}\}, & \dfrac{M}{N} \in \left(\dfrac{1}{K-1}, 1\right) & (7.16)\\ \dfrac{K}{T}\min\{NT_{\mathrm{u}}, NT_{\mathrm{d}}\}, & \dfrac{M}{N} \in [1, \infty). & (7.17) \end{cases}$$

Remark 7.1. With slight modification, Theorem 7.1 can be applied to the full-duplex case. As a key difference from the half-duplex case, a full-duplex node transmits and receives signals simultaneously, which implies that the uplink and downlink phases always have same time duration. Therefore, for the full duplex MIMO MWRC with full data exchange, we only need to set $T = T_{\mathrm{u}} = T_{\mathrm{d}}$ in Theorem 7.1. The corresponding DoF is given as

$$d = K\min\{M, N\}. \tag{7.18}$$

The Remark 7.18 can be obtained from the following two factors: Firstly, the system is assumed to perform full data exchange in this work, meaning that each user will need to recover all the information transmitted from all other users in the system. Secondly, in a full-duplex system, the uplink transmission and download transmission can be performed in different frequency band, to make the uplink and downlink interference free. In half-duplex system, the maximum DoF is the minimum of the DoF of the uplink transmission and downlink transmission, as shown in Theorm 7.1. Thus, in full-duplex transmission, the DoF is obtained by setting equal number of uplink transmission and downlink transmission.

Remark 7.2. It is interesting to compare the DoFs of the MIMO MWRCs operating in full data exchange and in pairwise data exchange (by assuming that both are full-duplex systems). From Remark 7.1, the DoF for full data exchange is given by (7.18). The DoF capacity for pairwise data exchange is in general unknown. A DoF outer bound is given in [99, 100] as

$$d \leq \min\{MK, 2N\}.$$

That is, the DoF for pairwise data exchange is always limited by $2N$. This is different from the DoF for full data exchange in (7.18) which is unbounded as $K \to \infty$. The reason is that, in full data exchange, each spatial stream going through the relay is forwarded to $K-1$ users, and thus achieves a DoF of $K-1$; however, in pairwise data exchange, each spatial stream through the relay is desirable only by one particular user, and therefore, achieves only a DoF of one.

Remark 7.3. From Theorem 7.1, we can see the DoF bottleneck of the network in different antenna configurations. In the case of $\frac{M}{N} \in \left(0, \frac{1}{K-1}\right]$, we see from (7.15) that the DoF is bottlenecked at the users. The reason is that (7.15) depends on the user's antenna number M, but not on the relay's antenna number N. In this case, to increase the number of antenna at the relay cannot further increase the system DoF capacity. On the other hand, in the case of $\frac{M}{N} \in [1, \infty)$, we see from (7.17) that the DoF depends on N but not on M, implying that the DoF is bottlenecked at the relay. In this case, to increase the number of antenna at the user cannot further increase the system DoF capacity. In the remaining case of $\frac{M}{N} \in \left(\frac{1}{K-1}, 1\right)$, we see from (7.16) that the DoF depends on both M and N. We also see that the DoF depends on the uplink/downlink time durations T_u and T_d.

7.4.2 Achievability Proof of Theorem 7.1

A key difficulty in the proof is that it is not easy to determine the boundary of the achievable rate region $\mathcal{R}$ in (7.12). This complicates the analysis of the achievable DoF. To circumvent this difficulty, we first establish a lower bound on the achievable sum-rate by replacing the downlink channel matrix $\mathbf{G}_i$ of each user i with a common "worse" channel matrix $\mathbf{G}_0$, satisfying

$$\mathbf{G}_0\mathbf{G}_0^\mathrm{H} \preceq \mathbf{G}_i\mathbf{G}_i^\mathrm{H},\ i \in \{1, \cdots, K\}, \tag{7.19}$$

and $\mathbf{G}_0$ is a full rank matrix. Clearly, as K is finite and $\{\mathbf{G}_i | i = 1, \cdots, K\}$ are random, such a $\mathbf{G}_0$ exists with probability one. With the above downgrading operation, the received signal at each user i becomes the same, $i \in \{1, \cdots, K\}$, and can be written as

$$\mathbf{y}_0 = \sum_{j=1}^{K} \mathbf{G}_0\mathbf{F}\mathbf{H}_j\mathbf{P}_j\mathbf{s}_j + \mathbf{G}_0\mathbf{F}\mathbf{n}_\mathrm{R} + \mathbf{n}_i. \tag{7.20}$$

The corresponding achievable rate region $\mathcal{R}_0$ of the MIMO multi-way channel in (7.20) is given as follow. Denote the index set $\mathcal{I} \triangleq \{1, \cdots, K\}$. Then, an achievable rate region is given by

$$\begin{aligned}\mathcal{R}_0 : \sum_{j, j\in\mathcal{S}} R_j &\leq \frac{1}{T} I\left(\{\mathbf{s}_j | j \in \mathcal{S}\}; \mathbf{y}_0 | \{\mathbf{s}_{j'} | j' \in \mathcal{S}^\mathrm{c}\}\right) \\ &= \frac{1}{T}\log\left(\frac{\left|\sum_{j,j\in\mathcal{S}} \mathbf{G}_0\mathbf{F}\mathbf{H}_j\mathbf{P}_j\mathbf{P}_j^\mathrm{H}\mathbf{H}_j^\mathrm{H}\mathbf{F}^\mathrm{H}\mathbf{G}_0^\mathrm{H} + \sigma^2(\mathbf{G}_0\mathbf{F}\mathbf{F}^\mathrm{H}\mathbf{G}_0^\mathrm{H} + \mathbf{I})\right|}{|\sigma^2(\mathbf{G}_0\mathbf{F}\mathbf{F}^\mathrm{H}\mathbf{G}_0^\mathrm{H} + \mathbf{I})|}\right),\ \mathcal{S} \subset \mathcal{I}.\end{aligned} \tag{7.21}$$

From (7.19), we see that the original channel in (7.7) for each user i, $i = 1, \cdots, K$, is always better than the one in (7.20), and therefore $\mathcal{R}_0 \subseteq \mathcal{R}$. Note that $\mathcal{R}_0$ in (7.21) is actually the rate region of the K-user multiple access channel in (7.20) by excluding the sum-rate constraint (i.e., the inequality in (7.21) by letting $\mathcal{S} = \mathcal{I}$).

It can be shown that the maximum sum-rate of $\mathcal{R}_0$ is achieved at the interception of the hyperplanes specified by the inequalities in (7.21) by letting $\mathcal{S} \in \{\mathcal{I}_i | i = 1, \cdots, K\}$. Therefore, noting $\mathcal{R}_0 \subseteq \mathcal{R}$, we obtain

$$\sum_{i=1}^{K} R_{\mathrm{r},i} \geq \sum_{\mathcal{S},\ \mathcal{S} \in \{\mathcal{I}_i | i=1,\cdots,K\}} \frac{1}{T} \log \left(\frac{\left| \sum_{j,\ j \in \mathcal{S}} \mathbf{G}_0 \mathbf{F} \mathbf{H}_j \mathbf{P}_j \mathbf{P}_j^{\mathrm{H}} \mathbf{H}_j^{\mathrm{H}} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \sigma^2 (\mathbf{G}_0 \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \mathbf{I}) \right|}{\left| \sigma^2 (\mathbf{G}_0 \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \mathbf{I}) \right|} \right)$$

$$= \sum_{i=1}^{K} \frac{1}{T} \log \left(\frac{\left| \frac{P}{M} \mathbf{G}_0 \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{H}_{\bar{i}}^{\mathrm{H}} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \sigma^2 (\mathbf{G}_0 \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \mathbf{I}) \right|}{\left| \sigma^2 (\mathbf{G}_0 \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_0^{\mathrm{H}} + \mathbf{I}) \right|} \right), \tag{7.22}$$

where the last step follows by setting $\mathbf{P}_i = \sqrt{P/M}\mathbf{I}$, $i = 1, \cdots, K$, without affecting the achievability proof.

With the definition of the DoF in (7.14), we have

$$d \geq \sum_{i=1}^{K} \frac{1}{T} \operatorname{rank}(\mathbf{G}_0 \mathbf{F} \mathbf{H}_{\bar{i}}). \tag{7.23}$$

Recall that $\mathbf{G}_0$ is chosen of full rank, and $\mathbf{H}_{\bar{i}}$ are randomly generated. Also, we randomly choose the precoder $\mathbf{F} \in \mathbb{C}^{NT_{\mathrm{d}} \times NT_{\mathrm{u}}}$ at the relay. Then, we have

$$\operatorname{rank}(\mathbf{G}_0 \mathbf{F} \mathbf{H}_{\bar{i}}) = \min\{MT_{\mathrm{d}}, NT_{\mathrm{d}}, NT_{\mathrm{u}}, (K-1)MT_{\mathrm{u}}\}$$

with probability one.

Case 1) $\frac{M}{N} \leq \frac{1}{K-1}$: By noting $(K-1)M \leq N$ and $M < N$, we obtain

$$\operatorname{rank}(\mathbf{G}_0 \mathbf{F} \mathbf{H}_{\bar{i}}) = \min\{(K-1)MT_{\mathrm{u}}, MT_{\mathrm{d}}\} \tag{7.24}$$

with high probability.

Case 2) $\frac{1}{K-1} < \frac{M}{N} < 1$: In this case, $(K-1)M > N$ and $M < N$. Then, (7.24) reduces to

$$\operatorname{rank}(\mathbf{G}_0 \mathbf{F} \mathbf{H}_{\bar{i}}) = \min\{NT_{\mathrm{u}}, MT_{\mathrm{d}}\}. \tag{7.25}$$

Case 3) $\frac{M}{N} \geq 1$: In this case, $(K-1)M > N$ and $M \geq N$, and hence

$$\operatorname{rank}(\mathbf{G}_0\mathbf{F}\mathbf{H}_{\bar{i}}) = \min\{NT_{\mathrm{u}}, NT_{\mathrm{d}}\} \tag{7.26}$$

with high probability.

Substitute the three cases (7.24) (7.25), and (7.26) into (7.23), we completes the proof of the achievability.

7.4.3 Converse Proof of Theorem 7.1

The converse of Theorem 7.1 is given in this subsection. The proof is built upon the cut-set theorem in [145]. Without loss of generality, we firstly consider the information flow from users $1, \cdots, K-1$ to user K, as illustrated in Fig. 7.2. We also assume the elements of the signal vector $\mathbf{x}_i$ transmitted from each user i and the relay are Gaussian distributed and are independently drawn from $\mathcal{CN}(0, P)$. In this case, the received signal at the relay for the uplink phase can be written as

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}_{\bar{K}}\mathbf{x}_{\bar{K}} + \mathbf{H}_K\mathbf{x}_K + \mathbf{n}_{\mathrm{R}}, \tag{7.27}$$

where $\mathbf{H}_{\bar{K}} = [\mathbf{H}_1 \cdots \mathbf{H}_{K-1}]$, $\mathbf{x}_{\bar{K}} = [\mathbf{x}_1^{\mathrm{T}} \cdots \mathbf{x}_{K-1}^{\mathrm{T}}]^{\mathrm{T}}$. The received signal at user K in the downlink phase is

$$\mathbf{y}_K = \mathbf{G}_K\mathbf{x}_{\mathrm{R}} + \mathbf{n}_K. \tag{7.28}$$

Applying the cut-set theorem in [145] on the considered information flow, as shown in Fig. 7.2, the cut-set bounds of this user cooperation scenario in this MIMO MWRC are given as

$$\text{Cut 1}: R_{\mathrm{r},K} \leq \frac{1}{T} I(\mathbf{x}_{\bar{K}}; \mathbf{y}_{\mathrm{R}} | \mathbf{x}_K), \tag{7.29}$$

$$\text{Cut 2}: R_{\mathrm{r},K} \leq \frac{1}{T} I(\mathbf{x}_{\mathrm{R}}; \mathbf{y}_K), \tag{7.30}$$

where

$$R_{\mathrm{r},K} = \sum_{i=1}^{K-1} R_i$$

is the decoding rate at the user K, and R_i is the information rate of user i per channel use. The $I(\cdot;\cdot)$ stands for the mutual information function. Therefore, we have the decoding rate at user K

$$R_{\mathrm{r},K} \leq \frac{1}{T} \min \left\{ I(\mathbf{x}_{\bar{K}}; \mathbf{y}_{\mathrm{R}} | \mathbf{x}_K), I(\mathbf{x}_{\mathrm{R}}; \mathbf{y}_K) \right\}. \tag{7.31}$$

For Cut 1, we write the mutual information part of the right hand side of (7.29) as

$$I(\mathbf{x}_{\bar{K}}; \mathbf{y}_{\mathrm{R}} | \mathbf{x}_K) = h(\mathbf{y}_{\mathrm{R}} | \mathbf{x}_K) - h(\mathbf{y}_{\mathrm{R}} | \mathbf{x}_K, \mathbf{x}_{\bar{K}}), \tag{7.32}$$

where $h(\cdot)$ stands for the entropy function. Since we assume that perfect CSI is available at all nodes, the first term of (7.32) can be written as

$$h(\mathbf{y}_{\mathrm{R}} | \mathbf{x}_K) = h(\mathbf{H}_{\bar{K}} \mathbf{x}_{\bar{K}} + \mathbf{n}_{\mathrm{R}}). \tag{7.33}$$

Similarly, we have the second term in (7.32) written as

$$h(\mathbf{y}_{\mathrm{R}} | \mathbf{x}_K, \mathbf{x}_{\bar{K}}) = h(\mathbf{n}_{\mathrm{R}}). \tag{7.34}$$

From the results of (7.33) and (7.34), we have

$$\begin{aligned} I(\mathbf{x}_{\bar{K}}; \mathbf{y}_{\mathrm{R}} | \mathbf{x}_K) &= h(\mathbf{H}_{\bar{K}} \mathbf{x}_{\bar{K}} + \mathbf{n}_{\mathrm{R}}) - h(\mathbf{n}_{\mathrm{R}}) \\ &= \log \left(\det \left(\mathbf{I} + \frac{(K-1)T_{\mathrm{u}}\rho}{\min\{(K-1)MT_{\mathrm{u}}, NT_{\mathrm{u}}\}} \mathbf{H}_{\bar{K}} \mathbf{H}_{\bar{K}}^{\mathrm{H}} \right) \right) \\ &= \sum_{j=1}^{\min\{(K-1)MT_{\mathrm{u}}, NT_{\mathrm{u}}\}} \log \left(1 + \frac{(K-1)T_{\mathrm{u}}\rho\kappa_j}{\min\{(K-1)MT_{\mathrm{u}}, NT_{\mathrm{u}}\}} \right), \end{aligned} \tag{7.35}$$

where κ_j is the j-th eigenvalue of the matrix $\mathbf{H}_{\bar{K}} \mathbf{H}_{\bar{K}}^{\mathrm{H}}$.

Similarly, for Cut 2, we have

$$\begin{aligned} I(\mathbf{x}_{\mathrm{R}}; \mathbf{y}_K) &= h(\mathbf{G}_K \mathbf{x}_{\mathrm{R}} + \mathbf{n}_K) - h(\mathbf{n}_K) \\ &= \log \left(\det \left(\mathbf{I} + \frac{T_{\mathrm{d}}\rho}{\min\{NT_{\mathrm{d}}, MT_{\mathrm{d}}\}} \mathbf{G}_{\bar{K}} \mathbf{G}_{\bar{K}}^{\mathrm{H}} \right) \right) \\ &= \sum_{j=1}^{\min\{MT_{\mathrm{d}}, NT_{\mathrm{d}}\}} \log \left(1 + \frac{T_{\mathrm{d}}\rho\gamma_j}{\min\{MT_{\mathrm{d}}, NT_{\mathrm{d}}\}} \right), \end{aligned} \tag{7.36}$$

where γ_j is the j-th eigenvalue of the matrix $\mathbf{G}_{\bar{K}}\mathbf{G}_{\bar{K}}^{\mathrm{H}}$.

Using the results in (7.35) and (7.36), together with (7.31), we have the DoF for information transfer from user group consisting of $\{1, \cdots, K-1\}$ to user K as

$$\begin{aligned} d_K &\le \frac{1}{T} \lim_{\rho\to\infty} \frac{\min\{I(\mathbf{x}_{\bar{K}}; \mathbf{y}_{\mathrm{R}}|\mathbf{x}_K), I(\mathbf{x}_{\mathrm{R}}; \mathbf{y}_K)\}}{\log\rho} \\ &= \frac{1}{T}\min\{\min\{(K-1)MT_{\mathrm{u}}, NT_{\mathrm{u}}\}, \min\{MT_{\mathrm{d}}, NT_{\mathrm{d}}\}\} \\ &= \frac{1}{T}\min\{(K-1)MT_{\mathrm{u}}, NT_{\mathrm{u}}, MT_{\mathrm{d}}, NT_{\mathrm{d}}\}. \end{aligned}$$

The above DoF upper bound is for user K. From the symmetry of the users, we immediately obtain (7.15).

7.5 DoF Optimization with Uplink/Downlink Time Allocation

We now consider optimizing the uplink/downlink time allocation to maximize the DoF of the MIMO MWRC. Given the total time duration T, this problem is formulated as

$$\underset{T_{\mathrm{u}},T_{\mathrm{d}}}{\text{maximize}} \quad d \tag{7.37a}$$

$$\text{subject to} \quad T_{\mathrm{u}} + T_{\mathrm{d}} = T, \tag{7.37b}$$

where d is given in (7.15) of Theorem 7.1.

7.5.1 Optimal Uplink/Downlink Time Allocation

The solution to (7.37) is given by the Theorem below.

Theorem 7.2. *For the half-duplex MWRC (M, N, K), the DoF capacity with optimal* *uplink/downlink time allocation is given by*

$$d_{\max} = \begin{cases} (K-1)M, & \frac{M}{N} \in \left(0, \frac{1}{K-1}\right] \quad (7.38) \\ \frac{KMN}{M+N}, & \frac{M}{N} \in \left(\frac{1}{K-1}, 1\right) \quad (7.39) \\ \frac{KN}{2}, & \frac{M}{N} \in [1, \infty). \quad (7.40) \end{cases}$$

The corresponding optimal uplink/downlink time allocation is

$$\frac{T_{\mathrm{u}}}{T_{\mathrm{d}}} = \begin{cases} \frac{1}{K-1}, & \frac{M}{N} \in \left(0, \frac{1}{K-1}\right] \quad (7.41) \\ \frac{M}{N}, & \frac{M}{N} \in \left(\frac{1}{K-1}, 1\right) \quad (7.42) \\ 1, & \frac{M}{N} \in [1, \infty). \quad (7.43) \end{cases}$$

Note that T is selected such that T_{u} and T_{d} computed from (7.41) are integers.

Proof. Case 1) $\frac{M}{N} \leq \frac{1}{K-1}$: In this case, d is given by (7.15). We substitute $T_{\mathrm{u}} = T - T_{\mathrm{d}}$ into (7.15), yielding

$$d = \frac{K}{T} \min\{(K-1)M(T - T_{\mathrm{d}}), MT_{\mathrm{d}}\}, \quad (7.44)$$

where d is a piecewise linear function of T_{d} for $1 \leq T_{\mathrm{d}} < T$. Specifically, d in (7.44) is monotonically increasing in $T_{\mathrm{d}} \in \left[1, \frac{K-1}{K}T\right]$, and monotonically decreasing in $T_{\mathrm{d}} \in \left[\frac{K-1}{K}T, T\right)$. Therefore, the maximum of d in this case is given by (7.38), which is achieved at $T_{\mathrm{d}} = \frac{K-1}{K}T$, with the corresponding ratio $\frac{T_{\mathrm{u}}}{T_{\mathrm{d}}}$ given by (7.41).

Case 2) $\frac{1}{K-1} < \frac{M}{N} < 1$: In this case, d is given by (7.16). We substitute $T_{\mathrm{u}} = T - T_{\mathrm{d}}$ into (7.16), yielding

$$d = \frac{K}{T} \min\{N(T - T_{\mathrm{d}}), MT_{\mathrm{d}}\}, \quad (7.45)$$

where d is a piecewise linear function of T_{d} for $1 \leq T_{\mathrm{d}} < T$. Clearly, d in (7.45) is maximized when $N(T - T_{\mathrm{d}}) = MT_{\mathrm{d}}$, with the corresponding ratio $\frac{T_{\mathrm{u}}}{T_{\mathrm{d}}}$ given by (7.42).

Case 3) $\frac{M}{N} \geq 1$: In this case, d is given by (7.17). We substitute $T_{\mathrm{u}} = T - T_{\mathrm{d}}$ into (7.17), yielding

$$d = \frac{K}{T} \min\{N(T - T_{\mathrm{d}}), NT_{\mathrm{d}}\}. \quad (7.46)$$

Note that d is a piecewise linear function of $T_{\rm d}$ for $1 \leq T_{\rm d} < T$. The maximum of d in (7.46) is achieved when $N(T - T_{\rm d}) = NT_{\rm d}$, and the corresponding ratio $\frac{T_{\rm u}}{T_{\rm d}}$ is given by (7.43). This completes the proof. ■

Remark 7.4. In (7.38), $d_{\max}/N$ is not a piecewise linear function of $\frac{M}{N} \in \left(\frac{1}{K-1}, 1\right)$. This non-linearity is due to the uplink/downlink time allocation (which is a function of $\frac{M}{N}$ given in (7.42)).

7.5.2 Alternative Approaches

For comparison, we consider the case of uplink/downlink equal time allocation. By letting $T_{\rm u} = T_{\rm d} = \frac{T}{2}$ in Theorem 7.1, we obtain that the DoF with equal time allocation is given by

$$d_{\rm eq} = \begin{cases} \dfrac{KM}{2}, & \dfrac{M}{N} \in (0,1) & (7.47) \\ \dfrac{KN}{2}, & \dfrac{M}{N} \in [1,\infty)\,. & (7.48) \end{cases}$$

Define the DoF gain of the proposed scheme over the scheme with uplink/downlink equal time allocation as

$$\Delta d \triangleq d_{\max} - d_{\rm eq}, \tag{7.49}$$

where $d_{\max}$ is given in (7.38). We have following result.

Proposition 7.1. *For the half-duplex MWRC (M, N, K), the maximal DoF gain is achieved at $\frac{M}{N} = \frac{1}{2}$ for $K = 3$, and achieved at $\frac{M}{N} = \sqrt{2} - 1$ for $K \geq 4$.*

Proof. Substituting (7.38) and (7.47) into the (7.49) we obtain the DoF gain per relay dimension as

$$\frac{\Delta d}{N} = \begin{cases} \left(\dfrac{K}{2} - 1\right)\dfrac{M}{N}, & \dfrac{M}{N} \in \left(0, \dfrac{1}{K-1}\right] & (7.50) \\ \dfrac{K}{2}\dfrac{\frac{M}{N} - \left(\frac{M}{N}\right)^2}{1 + \frac{M}{N}}, & \dfrac{M}{N} \in \left(\dfrac{1}{K-1}, 1\right) & (7.51) \\ 0, & \dfrac{M}{N} \in [1, \infty)\,. & (7.52) \end{cases}$$

Eqn. (7.50) is a continuous function when $\frac{M}{N} > 0$. More specifically, (7.50) is a monotonically increasing function of $\frac{M}{N}$ since $\left(\frac{K}{2} - 1\right) > 0$; (7.51) is a concave function within the range of $[0, 1]$, and its maximum is achieved at $\frac{M}{N} = \sqrt{2} - 1$. As (7.50) is a continuous function, the maximal DoF gain is achieved at $\frac{M}{N} = \frac{1}{2}$ for $K = 3$. For other cases when $K \geq 4$, we have $\frac{1}{K-1} \leq \frac{1}{3} < \sqrt{2} - 1$. The maximal DoF gain is achieved at $\frac{M}{N} = \sqrt{2} - 1$. ■

Remark 7.5. Alternatively, we can quantify the relative DoF gain by evaluating $\frac{\Delta d}{d_{\text{eq}}}$. For $K = 3$, the relative DoF gain is 33.33% at $\frac{M}{N} = \frac{1}{2}$, and for $K \geq 4$, the relative DoF gain is 41.42% at $\frac{M}{N} = \sqrt{2} - 1$. We see that significant DoF improvement is obtained by the optimal time allocation.

Remark 7.6. For asymptotic cases when there is a large number of users $K \to \infty$ and the relay has a much larger number of antennas than the user's $\frac{M}{N} \to 0$, the relative DoF gain $\frac{\Delta d}{d_{\text{eq}}}$ can achieve its maximum value of 100%. To see this, we first note that $\frac{1}{K-1} \to 0$ as $K \to \infty$. Thus, Δd is given by (7.51) for $\frac{M}{N} \in (0, 1)$. Letting $\frac{M}{N} \to 0$, we obtain from (7.47) and (7.51) that

$$\lim_{\frac{M}{N} \to 0} \frac{\Delta d}{d_{\text{eq}}} = \lim_{\frac{M}{N} \to 0} \frac{1 - \frac{M}{N}}{1 + \frac{M}{N}} = 100\%.$$

An intuitive explanation is as follows. When $K \to \infty$, the optimal time allocation is given by $\frac{T_{\text{u}}}{T_{\text{d}}} = \frac{M}{N}$ in (7.42) for any $\frac{M}{N} \in (0, 1)$. Together with the condition of $\frac{M}{N} \to 0$, we see that the optimal time allocation strategy is to allocate almost all the time for downlink transmission, which leads to a DoF gain of 100%, as compared with equal time allocation.

Another approach for comparison is TDMA. In a conventional TDMA system, K users are scheduled to transmit in K orthogonal time slots. Hence, the DoF capacity with TDMA scheme is given by

$$d_{\text{TDMA}} = \begin{cases} \dfrac{(K-1)M}{2}, & \dfrac{M}{N} \in (0, 1) \qquad (7.53) \\ \dfrac{(K-1)N}{2}, & \dfrac{M}{N} \in [1, \infty)\,. \qquad (7.54) \end{cases}$$

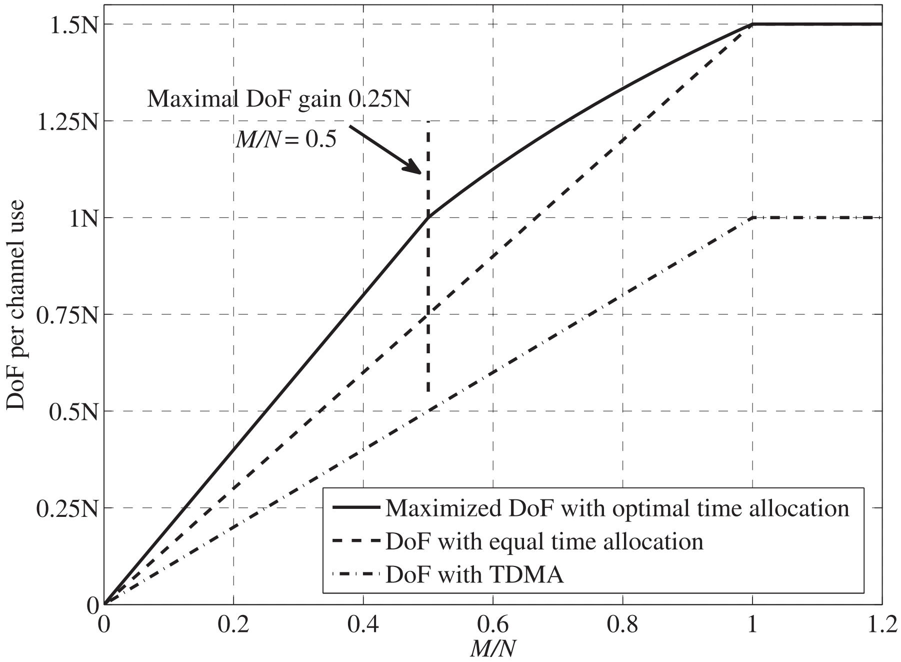


Figure 7.3 Comparison of the DoF capacity for the MWRC with $K = 3$ users.

Eqn. (7.53) can be intuitively explained as follows. For example, when $M < N$, the relay can receive at most M independent signal streams from each user. On the other hand, the relay can broadcast at most M signal streams to the users. Thus, the total independent signal streams received in the system is at most $(K-1)M$. Due to the symmetric characteristic of the system, the DoF result when $M < N$ is given by (7.53).

7.5.3 Numerical Comparisons

In this subsection, we illustrate numerical results for the DoF analysis. Fig. 7.3 shows the DoF per channel use of the MIMO MWRC with $K = 3$ against the user/relay antenna ratio $\frac{M}{N}$. We see that the maximal DoF per channel use is not a piece-wise linear function of $\frac{M}{N}$, as discussed in Remark 7.4. The performance gain is achieved when the number of user antenna is less than the relay antenna. When $\frac{M}{N} \geq 1$, the DoF per channel use remains constant, and equal time allocation is optimal. This is because when $\frac{M}{N} \geq 1$, the bottleneck of the information flow is the relay, as discussed in Remark 7.3. In this case, further increasing the number of user antennas cannot increase the DoF. We also see from Fig. 7.3 that the maximum DoF gain is obtained at $\frac{M}{N} = \frac{1}{2}$, which is consistent with Proposition 7.1. We also illustrate the DoF of a conventional TDMA system in Fig. 7.3 for the case of $K = 3$. We can see the significant DoF improvement obtained by allowing multiple users transmitting simultaneously in the uplink phase and performing network coding at the relay.

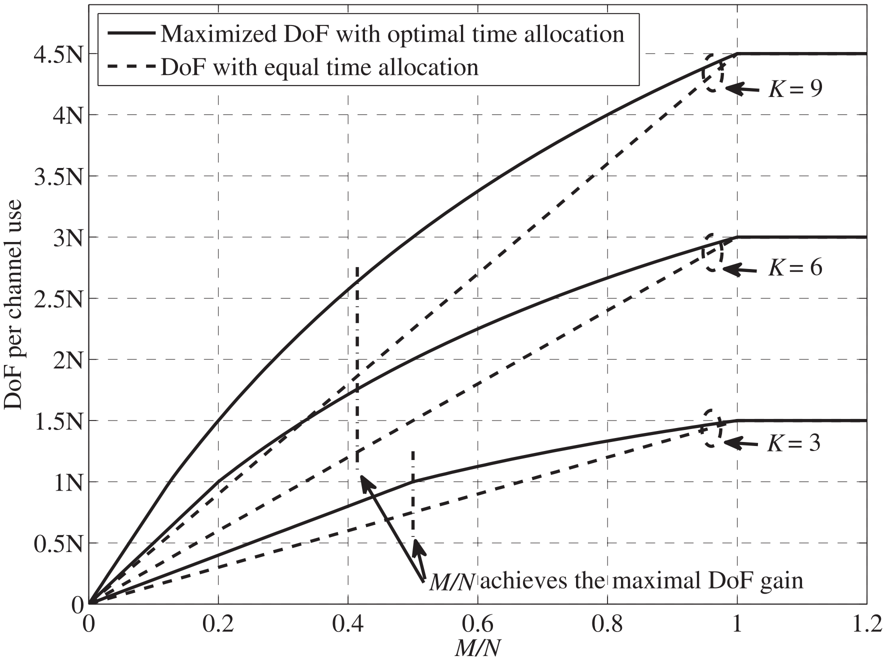


Figure 7.4 The DoF capacity of the MWRC against the antenna ratio $\frac{M}{N}$. The number of users is set as $K = 3, 6, 9$.

We next examine a specific antenna setup of this 3-user system, i.e., $\frac{M}{N} = \frac{1}{2}$, or equivalently, $N = 2M$. In each uplink channel use, the relay receives $2M$ linear equations of $3M$ signal streams from the users. In each downlink channel use, each user only receives M linear equations from the relay due to the limited number of receiving antennas. In order for each user to decode the messages of the other two users with the cancellation of their self-message, each user needs at least $2M$ linear equations. In this case, two downlink channel uses are required for each uplink channel

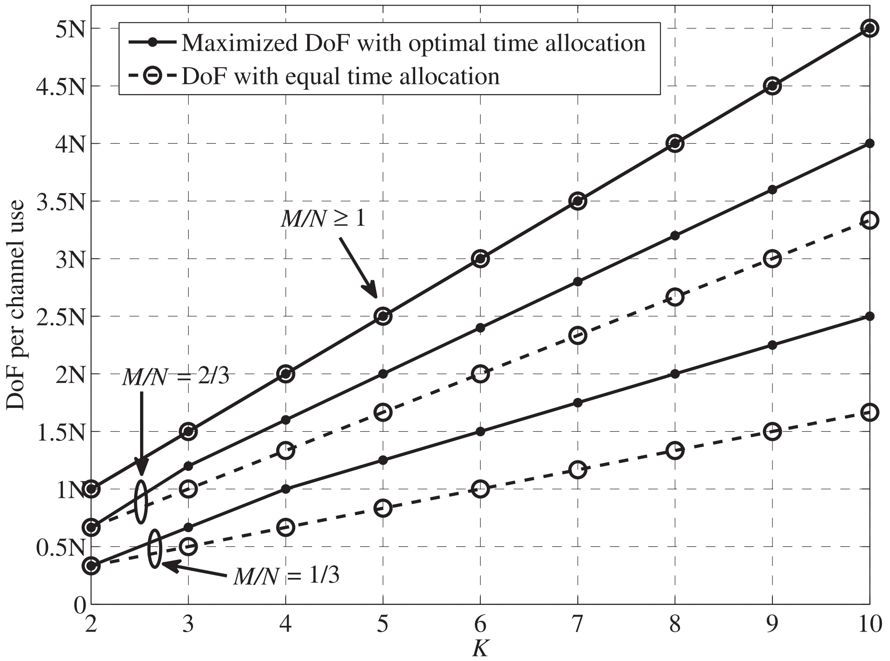


Figure 7.5 The DoF capacity of the MWRC with different antenna ratio $\frac{M}{N}$, against the number of users K. The antenna ratio is set as $\frac{M}{N} = \frac{1}{3}, \frac{2}{3}$, and $\frac{M}{N} \geq 1$. When $\frac{M}{N} \geq 1$, the equal time allocation becomes optimal for all integers $K \geq 2$.

use. This gives 3 channel use in total for one round of information exchange. Thus, $d = 2M \times (3 \text{ users})/(3 \text{ channel uses}) = 2M = N$, which agrees with the Fig. 7.3.

We now compare the DoF with different number of users. Fig. 7.4 shows the comparison of the DoF per channel use with $K = 3,\ 6,$ and 9. We see that the total DoF per channel use increases with the increasing number of users. The antenna configuration that achieves the maximal DoF gain for each K is marked out in Fig. 7.4, which agrees with the analytical results in Proposition 7.1. Fig. 7.5 shows the DoF capacity against the number of users K for given antenna ratios of $\frac{M}{N} = \frac{1}{3}, \frac{2}{3}$ and $\frac{M}{N} \geq 1$. Fig. 7.6 shows the specific case when $\frac{M}{N} = \frac{1}{3}$ in Fig. 7.5, and we added the DoF with TDMA scheme for comparison. When $\frac{M}{N} < 1$, the optimized DoF is equal to the DoF obtained with equal time allocation for $K = 2$. This is because MWRC becomes TWRC for $K = 2$, and the information flows for uplink and downlink are symmetric. In addition, Fig. 7.5 suggests that the optimized DoF is a piecewise linear

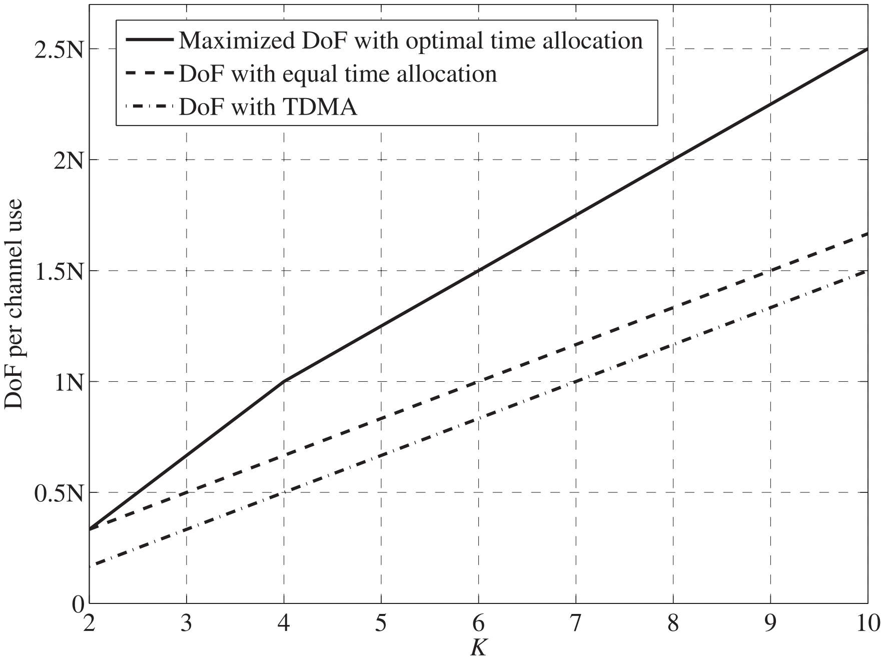


Figure 7.6 The DoF capacity of the MWRC with fixed antenna ratio $\frac{M}{N} = \frac{1}{3}$, against the number of users K.

function of K when $\frac{M}{N} < 1$. This can be seen from (7.38) and (7.39). We can also see from Fig. 7.5 and Fig. 7.6 that, the DoF with equal time allocation and with TDMA are both linear functions of K for any given antenna ratio $\frac{M}{N}$, and they have the same slope. To achieve the same DoF, equal time allocation always requires one less user in the system than the TDMA scheme. This can be seen from (7.47) and (7.53). Furthermore, we see from Fig. 7.5 that the DoF gain increases with the increase of K when $\frac{M}{N} < 1$. In the case when $\frac{M}{N} \geq 1$, Fig. 7.5 shows that the optimized DoF is a linear function of K. This can be seen from (7.40). In addition, we see from Fig. 7.5 that equal time allocation becomes optimal for all integers $K \geq 2$. This can be seen from (7.40) and (7.48).

7.6 Sum-Rate Maximization

In this section, we study the sum-rate optimization. We show that, the system sum-rate can be maximized by iteratively updating the precoder at the relay and the precoders at the users.

7.6.1 Problem Formulation

The achievable sum-rate R defined in (7.13) is confined within the rate region $\mathcal{R}$ in (7.12) which consists of $K(2^{K-1}-1)$ rate constraints. Hence, it is difficult to obtain an explicit expression of the achievable sum-rate. Following [148], we only consider those rate constraints involving $K-1$ users. Then, an upper bound of the sum-rate can be written as

$$R^{\mathrm{UB}} = \frac{1}{T}\sum_{i=1}^{K} \log\left(\frac{\left|\sum\limits_{j=1, j\neq i}^{K} \mathbf{G}_i\mathbf{F}\mathbf{H}_j\mathbf{Q}_j\mathbf{H}_j^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \sigma^2\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right)\right|}{\left|\sigma^2\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right)\right|}\right), \qquad (7.55)$$

where $\mathbf{Q}_j = \mathbf{P}_j\mathbf{P}_j^{\mathrm{H}}$. The sum-rate maximization problem can be formulated as

$$\begin{aligned}
&\underset{\{\mathbf{Q}_i|i\in\{1,\cdots,K\}\},\mathbf{F}}{\text{maximize}} && R^{\mathrm{UB}} && (7.56\text{a})\\
&\text{subject to} && \frac{1}{T_{\mathrm{u}}}\operatorname{Tr}\left(\mathbf{Q}_i\right) \leq P, \quad i \in \{1,\cdots,K\} && (7.56\text{b})\\
& && \frac{1}{T_{\mathrm{d}}}\operatorname{Tr}\left(\sum_{i=1}^{K}\mathbf{F}\mathbf{H}_i\mathbf{Q}_i\mathbf{H}_i^{\mathrm{H}}\mathbf{F}^{\mathrm{H}} + \sigma^2\mathbf{F}\mathbf{F}^{\mathrm{H}}\right) \leq P, && (7.56\text{c})
\end{aligned}$$

where R^{UB} is given in (7.55).

The optimization problem in (7.56) involves the set of user precoders $\{\mathbf{Q}_i|i \in \{1,\cdots,K\}\}$, and the relay precoder $\mathbf{F}$. A general approach to this problem is to iteratively optimize one group of parameters while fixing the other group of parameters; see, e.g., [109]. In our problem (7.56), we can iteratively updating the set of user precoders $\{\mathbf{Q}_i|i \in \{1,\cdots,K\}\}$, with fixed relay precoder $\mathbf{F}$, and then updating relay precoder $\mathbf{F}$ with fixed set of user precoders $\{\mathbf{Q}_i|i \in \{1,\cdots,K\}\}$.

7.6.2 Optimizing $\{\mathbf{Q}_i\}$ for Fixed $\mathbf{F}$

We first consider fixing the relay precoder $\mathbf{F}$ and optimizing the user precoders. For a fixed relay precoder $\mathbf{F}$, R^{UB} in (7.56) is a concave function of $\mathbf{Q}_i$, $i \in \{1,\cdots,K\}$

[151]. Thus, standard convex optimization tools can be used to find the optimal user precoders $\{\mathbf{Q}_i | i \in \{1, \cdots, K\}\}$. We omit the details for brevity.

7.6.3 Optimizing F for Fixed $\{\mathbf{Q}_i\}$

We now consider the optimization of $\mathbf{F}$ in the problem (7.56) for fixed $\{\mathbf{Q}_i | i \in \{1, \cdots, K\}\}$. This problem is in general difficult to solve directly. In the following, we propose an iterative algorithm to approximately solve this problem.

We first convert the optimization of relay precoder $\mathbf{F}$ in our problem (7.56) with fixed set of user precoders $\{\mathbf{Q}_i | i \in \{1, \cdots, K\}\}$, into a form that is suitable for iterative optimization, following the approach in [109].

Recall that all transmitted signals from the users are i.i.d. drawn from $\mathcal{CN}(0, 1)$. At user i, $i \in \{1, \cdots, K\}$, the intended message is $\mathbf{s}_{\bar{i}}$. Let $p(\mathbf{s}_{\bar{i}})$ be the distribution of $\mathbf{s}_{\bar{i}}$, it is given by

$$p(\mathbf{s}_{\bar{i}}) = \pi^{-(K-1)L} \exp\left(-\mathbf{s}_{\bar{i}}^{\mathrm{H}} \mathbf{s}_{\bar{i}}\right). \tag{7.57}$$

Let $p(\mathbf{s}_{\bar{i}} | \mathbf{y}_{\bar{i}})$ be the *a posterior* distribution of $\mathbf{s}_{\bar{i}}$ given $\mathbf{y}_{\bar{i}}$. From [164], this *a posterior* distribution follows $\mathcal{CN}(\mathbf{\Omega}_i \mathbf{y}_{\bar{i}}, \mathbf{\Delta}_i)$, and can be written as

$$p(\mathbf{s}_{\bar{i}} | \mathbf{y}_{\bar{i}}) = \pi^{-(K-1)L} \left|\mathbf{\Delta}_i\right|^{-1} \exp\Big(-(\mathbf{s}_{\bar{i}} - \mathbf{\Omega}_i \mathbf{y}_{\bar{i}})^{\mathrm{H}} \mathbf{\Delta}_i^{-1} \left(\mathbf{s}_{\bar{i}} - \mathbf{\Omega}_i \mathbf{y}_{\bar{i}}\right)\Big), \tag{7.58}$$

where

$$\begin{aligned} \mathbf{\Omega}_i &= \mathrm{Cov}(\mathbf{s}_{\bar{i}}, \mathbf{y}_{\bar{i}}) \mathrm{Cov}^{-1}(\mathbf{y}_{\bar{i}}, \mathbf{y}_{\bar{i}}) \\ &= (\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})^{\mathrm{H}} \big((\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})(\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})^{\mathrm{H}} + \sigma^2 (\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I})\big)^{-1}, && \text{(7.59a)} \\ \mathbf{\Delta}_i &= \mathbf{I} - \mathbf{\Omega}_i \mathrm{Cov}(\mathbf{y}_{\bar{i}}, \mathbf{s}_{\bar{i}}) \\ &= \mathbf{I} - \mathbf{\Omega}_i \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}}. && \text{(7.59b)} \end{aligned}$$

The $\mathrm{Cov}(\cdot, \cdot)$ in (7.59) stands for covariance matrix.

Let $p(\mathbf{s}_{\bar{i}}, \mathbf{y}_{\bar{i}})$ be the joint distribution of $\mathbf{s}_{\bar{i}}$ and $\mathbf{y}_{\bar{i}}$. Then the decoding rate upper bound at user i, $i \in \{1, \cdots, K\}$, can be written as

$$\begin{aligned} R_{\mathrm{r},i}^{\mathrm{UB}} &= \frac{1}{T} I(\mathbf{s}_1 \cdots, \mathbf{s}_{i-1}, \mathbf{s}_{i+1}, \cdots, \mathbf{s}_K; \mathbf{y}_{\bar{i}}) \\ &= \frac{1}{T \ln 2} \iint p(\mathbf{s}_{\bar{i}}, \mathbf{y}_{\bar{i}}) \ln \frac{p(\mathbf{s}_{\bar{i}}|\mathbf{y}_{\bar{i}})}{p(\mathbf{s}_{\bar{i}})} d\mathbf{s}_{\bar{i}} d\mathbf{y}_{\bar{i}} \\ &= \frac{1}{T \ln 2} \iint p(\mathbf{s}_{\bar{i}}) p(\mathbf{y}_{\bar{i}}|\mathbf{s}_{\bar{i}}) \ln \frac{p(\mathbf{s}_{\bar{i}}|\mathbf{y}_{\bar{i}})}{p(\mathbf{s}_{\bar{i}})} d\mathbf{s}_{\bar{i}} d\mathbf{y}_{\bar{i}}, \end{aligned} \tag{7.60}$$

where

$$p(\mathbf{y}_{\bar{i}}|\mathbf{s}_{\bar{i}}) = \pi^{-MT_{\mathrm{d}}} |\mathbf{\Psi}_i|^{-1} \exp\left(-(\mathbf{y}_{\bar{i}} - \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}} \mathbf{s}_{\bar{i}})^{\mathrm{H}} \mathbf{\Psi}_i^{-1} (\mathbf{y}_{\bar{i}} - \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}} \mathbf{s}_{\bar{i}})\right),$$

and

$$\mathbf{\Psi}_i = \sigma^2 \left(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right).$$

Substitute (7.57) and (7.58) into (7.60), we can rewrite the decoding rate upper bound at user i, $i = 1, \cdots, K$, as

$$\tilde{R}_{\mathrm{r},i}^{\mathrm{UB}} = -\frac{1}{T \ln 2} \left(\ln |\mathbf{\Delta}_i| - (K-1) L + \mathrm{Tr}\left(\mathbf{\Delta}_i^{-1} \mathbf{D}_i\right)\right), \tag{7.61}$$

where

$$\mathbf{D}_i = (\mathbf{I} - \mathbf{\Omega}_i \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}}) (\mathbf{I} - \mathbf{\Omega}_i \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})^{\mathrm{H}} + \sigma^2 \mathbf{\Omega}_i \left(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right) \mathbf{\Omega}_i^{\mathrm{H}}. \tag{7.62}$$

Thus, the sum-rate in (7.55) can be rewritten as

$$\tilde{R}^{\mathrm{UB}} = -\frac{1}{T \ln 2} \sum_{i=1}^{K} \left(\ln |\mathbf{\Delta}_i| - (K-1) L + \mathrm{Tr}\left(\mathbf{\Delta}_i^{-1} \mathbf{D}_i\right)\right), \tag{7.63}$$

where $\mathbf{D}_i$ is given in (7.62). A verification of this derived sum-rate upper bound is shown in Appendix E.

Now the optimization of relay precoder $\mathbf{F}$ in problem (7.56) with fixed set of user precoders $\{\mathbf{Q}_i | i \in \{1, \cdots, K\}\}$ can be rewritten as

$$\underset{\mathbf{F}, \{\mathbf{\Delta}_i, \mathbf{\Omega}_i | i \in \{1, \cdots, K\}\}}{\text{maximize}} \quad \tilde{R}^{\mathrm{UB}} \tag{7.64a}$$

$$\text{subject to} \quad \frac{1}{T_{\mathrm{d}}} \mathrm{Tr}\left(\sum_{i=1}^{K} \mathbf{F} \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^{\mathrm{H}} \mathbf{F}^{\mathrm{H}} + \sigma^2 \mathbf{F} \mathbf{F}^{\mathrm{H}}\right) \leq P. \tag{7.64b}$$

The optimization problem in (7.64) is a quadratically constrained quadratic program (QCQP) [151] with respect to $\mathbf{F}$. The Lagrangian is given by

$$\mathcal{L}\left(\mathbf{F}, \{\mathbf{\Delta}_i, \mathbf{\Omega}_i | i \in \{1, \cdots, K\}\}\right) = \tilde{R}^{\mathrm{UB}} + \lambda \left(\frac{1}{T_{\mathrm{d}}} \operatorname{Tr}\left(\sum_{i=1}^{K} \mathbf{F}\mathbf{H}_i\mathbf{Q}_i\mathbf{H}_i^{\mathrm{H}}\mathbf{F}^{\mathrm{H}} + \sigma^2\mathbf{F}\mathbf{F}^{\mathrm{H}} \right) - P \right), \tag{7.65}$$

where λ stands for the Lagrangian multiplier. Letting $\partial\mathcal{L}/\partial\mathbf{F} = 0$, we obtain the optimal $\mathbf{F}$ in vector form as

$$\operatorname{vec}(\mathbf{F}) = \left(\mathbf{\Phi}_1 + \mathbf{\Phi}_2\right)^{-1} \cdot \operatorname{vec}\left(\sum_{i=1}^{K} \left(\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}\mathbf{\Delta}_i^{-1}\mathbf{\Omega}_i\mathbf{G}_i \right)^{\mathrm{H}} \right) \tag{7.66}$$

where

$$\mathbf{\Phi}_1 = \sum_{i=1}^{K} \left(\left(\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}\mathbf{P}_{\bar{i}}^{\mathrm{H}}\mathbf{H}_{\bar{i}}^{\mathrm{H}} + \sigma^2\mathbf{I} \right)^{\mathrm{T}} \otimes \left(\mathbf{G}_i^{\mathrm{H}}\mathbf{\Omega}_i^{\mathrm{H}}\mathbf{\Delta}_i^{-1}\mathbf{\Omega}_i\mathbf{G}_i \right) \right),$$

$$\mathbf{\Phi}_2 = \left(\sum_{i=1}^{K} \mathbf{H}_i\mathbf{P}_i\mathbf{P}_i^{\mathrm{H}}\mathbf{H}_i^{\mathrm{H}} + \sigma^2\mathbf{I} \right)^{\mathrm{T}} \otimes (\lambda\mathbf{I}),$$

and $\operatorname{vec}(\cdot)$ stands for the vectorization function, $\otimes$ stands for Kronecker product, and λ is selected to satisfy the power constraints in (7.64b). The relay precoder $\mathbf{F}$ is obtained by converting the $\operatorname{vec}(\mathbf{F})$ computed in (7.66) to its matrix form.

7.6.4 Overall Algorithm

The sum-rate maximization problem in (7.56) can be solved by iteratively updating the user precoders and the relay precoder. The overall procedure is outlined in the following algorithm. Note that the result obtained by the proposed iterative optimization procedure is not the global optimum result. The optimization results from the iterative optimization procedure are sensitive to the initial precoder, especially at the high SNR region.

Algorithm 7.1 Iterative Sum-rate Maximization

```
Initialization:
    Set $\mathbf{P}_i = \alpha \cdot \mathrm{rand}(MT_\mathrm{u}, L)$, $i \in \{1, \cdots, K\}$, and set $\mathbf{F} = \beta \cdot \mathrm{rand}(NT_\mathrm{d}, NT_\mathrm{u})$,
    where $\alpha$ and $\beta$ are chose to meet the user and relay power constraint. Set $\varepsilon$ to
    an arbitrary small real number.
Iteration:
while the $R^{\mathrm{UB}}$ can be increased by more than $\varepsilon$ do
    Calculate relay precoder $\mathbf{F}$ base on (7.66).
    Optimize $\{\mathbf{P}_i | i \in \{1, \cdots, K\}\}$ using CVX optimization tool.
    Compute $\{\mathbf{\Delta}_i, \mathbf{\Omega}_i | i \in \{1, \cdots, K\}\}$ based on (7.59).
end while
```

7.6.5 Numerical Results

In this subsection, we illustrate the numerical results of the rate optimization. We consider two systems. Fig. 7.7 shows various sum-rates for network where $K = 3$, $M = 2$, $N = 4$, and Fig. 7.8 shows various sum-rates for network where $K = 4$, $M = 2$, $N = 6$.

We firstly consider the DoF. For the network $K = 3$, $M = 2$, $N = 4$, the optimal DoF calculated based on (7.38) is 4. The DoF with equal time allocation calculated from (7.47) is 3. The DoF with TDMA scheme calculated from (7.53) is 2. For the network $K = 4$, $M = 2$, $N = 6$, the optimal DoF calculated based on (7.38) is 6. The DoF with equal time allocation calculated from (7.47) is 4. The DoF with TDMA scheme calculated from (7.53) is 3.

In Fig. 7.7, we can see that the sum-rates of the system $K = 3$, $M = 2$, $N = 4$ with proposed time optimization grow linearly with the slope of 4; the sum-rate of the system with equal time allocation grows linearly with the slope of 3; and the sum-rate of the system in TDMA scheme grows linearly with the slope of 2. These results match well with the DoF analysis in this paper. Similarly, we can see numerical

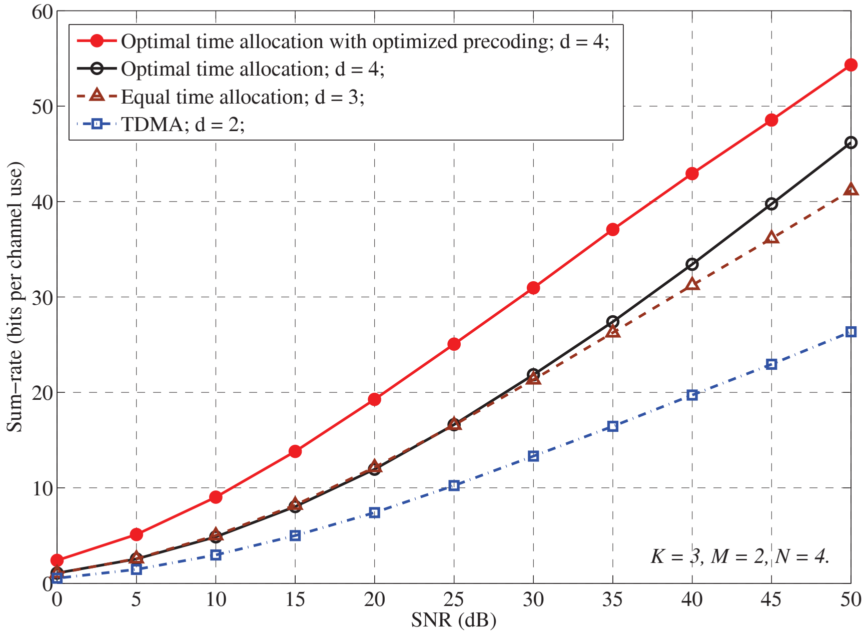


Figure 7.7 The comparison of sum-rate of a MIMO MWRC: $K = 3$, $M = 2$, $N = 4$.

results match well with the DoF analysis for system $K = 4$, $M = 2$, $N = 6$ in Fig. 7.8.

We now compare the sum-rates of the considered system. Fig. 7.7 and Fig. 7.8 demonstrated a significant sum-rate improvement by the proposed iterative sum-rate optimization algorithm in Subsection 7.6.4. The gain is more than 10 bits per channel use in the medium to high SNR region. It is worth pointing out that the optimization results are sensitive to the initial precoder in the high SNR region. In our simulation, we randomly initialize the precoder 200 times at each SNR and select the best one. Further, Fig. 7.7 and Fig. 7.8 compared the sum-rate of various schemes. It also shows that with network coding, the sum-rate improves significantly, compared to the TDMA scheme.

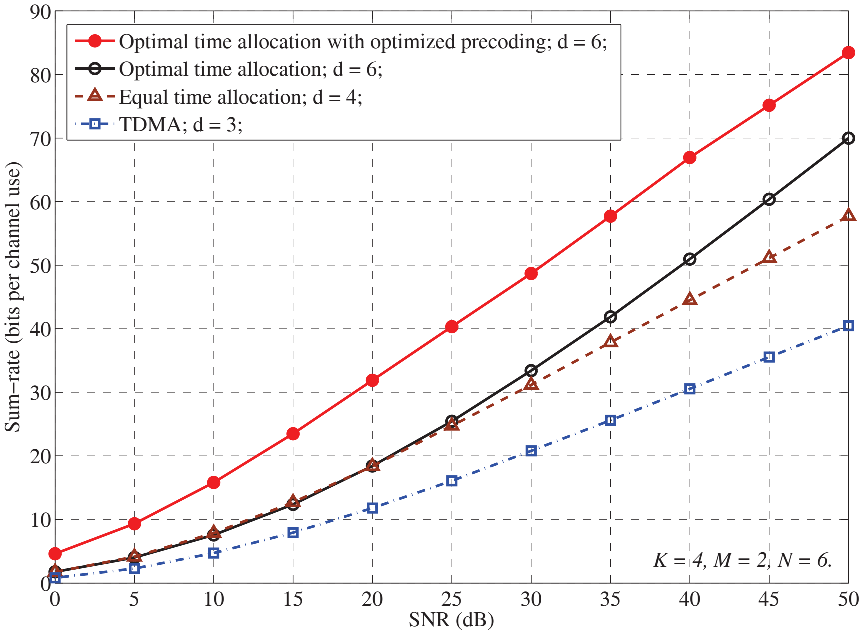


Figure 7.8 The comparison of sum-rate of a MIMO MWRC: $K = 4$, $M = 2$, $N = 6$.

7.7 Conclusion

In this chapter, we derived the DoF capacity of the MWRC with full data exchange. Our derived result generally applies to any antenna and user configurations, as well as to both full-duplex and half-duplex communications. We showed that, unlike other MWRCs with pairwise data exchange, the number of time slots for the uplink and downlink phases are asymmetric when full data exchange is considered. As a result, we optimized the uplink/downlink time allocation to enhance the DoF of the half-duplex network. Both our analysis and numerical results showed that a significant DoF improvement can be achieved by the proper uplink/downlink time allocation. Further, we proposed an iterative algorithm to optimize the user precoders and the relay precoder for sum-rate maximization. The numerical results of the sum-rate match well with the DoF analysis in this chapter, and they also demonstrated that the system sum-rate performance can be considerably improved by a careful design of the user and relay precoders.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

This thesis is aim to investigate the question “How to design PNC schemes for various relay networks to achieve fast and reliable information exchange?”. As a start, Chapter 2 is focused on the channel-uncoded PNC system in a TWRC. In particular, we considered relay fading channel with phase variation onlyand each user is using QPSK modulation. In PNC scheme, relay sees an superimposed constellation. The problem of minimizing the detection error for network-coded information at the relay becomes a grouping problem. The grouping rule is to maximize the minimum distance between each group. The relay can select different network coding coefficients to achieve various grouping on the superimposed constellation. Thus, we proposed an optimum grouping strategy to achieve linear network coding at the relay with maximized minimum distance. The performance bound was derived and numerical simulation was used to show the tightness of the derived performance bound.

In Chapter 3, the study moved to analyzing the performance of binary classic codes coded PNC system in a TWRC. The challenge here in this work is to characterize the distance spectrum of the network coded codewords at the relay from the superimposed codewords at the relay. It has been found out that, different from the conventional single-user scenario, the effective distance spectrum of the PNC scheme is determined by both the Hamming distance spectrum of the underlying channel code, and an extra distance spectrum of the superimposed codewords. Base on that, an asymptotically tight performance bound for the error probability of the channel-coded scheme at high SNR region was derived. The derivation showed that the extra distance spectrum of the superimposed codewords leads to an SNR penalty of at most $\ln 2$ in linear scale.

Chapter 4 studied the design of more industrial practical modern codes in PNC systems. The IRA code is selected due to its easy encoding process. Other than the conventional analysis in point-to-point channel, this study revealed that the presence of the ternary signal leads to the challenges in its convergence analysis. We analyzed the component decoders and derived the generalized update rules in terms of log-likelihood ratios. We then proposed two models for the soft information exchanged among the components decoders, and developed bounds on the approximation of the extrinsic information transfer (EXIT) functions. Based on that, we carried out an EXIT chart curve-fitting technique to construct optimized codes. Our developed irregular repeat-accumulate coded PNC schemes have significantly improved performance compared to the existing regular repeat-accumulate coded PNC schemes.

In Chapter 5, the research of channel-coded PNC is focused on more complicated system: multi-way relay networks. In existing literature, lattice codes have been found good characteristics in network coding. In this chapter, the Eisenstein integer based lattice network codes are studied for PNC employed multi-way relay fading channels. A union bound estimation of the decoding error probability has been derived. The design criteria for optimal lattice network codes with minimum decoding error probability were formulated in this work. Furthermore, the construction of lattice network codes from linear codes by Complex Construction B is studied. The nominal coding gains and error performance of for various LNCs are also analyzed in this work. It has been demonstrated in this work that the lattice network codes constructed by Complex Construction B provide a better tradeoff between code rate and nominal coding gain. Further, the optimal dither method in terms of energy efficiency for LNC over GF(4) is derived, and this optimal dither can save on average 1/3 transmission power. Next, the lattice network code design from linear codes over GF(4) was investigated, and explicit connection between parameters of the linear code and of the corresponding lattice network code was established. Moreover, linear lattice codes from convolutional, BCH, and Reed-Solomon codes over GF(4) was constructed in this study.

Chapter 6 focused on improving the sum-rate performance of PNC employed SISO MWRCs. It has been shown that by choosing the pair-wise transmission scheduling appropriately, we can achieve a significant improvement for the sum-rate of the multiuser transmission. High level modulation with nested lattice codes was adopted in this study. An opportunistic pair-wise transmission scheme was proposed to exploit the multi-user fading channels. In addition, it has been shown that the proposed opportunistic pair-wise transmission can provide up to 4.5 dB gain for an uncoded 4-user MWRC, and up to 2.5 dB gain for a rate-1/2 memory order-1 convolutional coded 4-user MWRC at the frame error rate of 10^{-2}, when compared to the conventional pair-wise transmission.

Chapter 7 extended the study in MIMO MWRCs. In particular, this chapter investigated the problem of maximizing the DoF of MIMO MWRCs with full data exchange, and an iterative algorithm was proposed to maximize the sum-rate. The DoF capacity of considered network is derived, and the results can be generally applied to both full-duplex and half-duplex systems. Further, optimization of the uplink/downlink time allocation to enhance the DoF of the half-duplex network was studied. Both analysis and numerical results showed that a significant DoF improvement can be achieved by the proper uplink/downlink time allocation. Further, an iterative algorithm to maximize system sum-rate was proposed, and the optimization parameters were the system precoders. It has been demonstrated in numerical simulation that the system sum-rate performance can be considerably improved by a careful design of the user and relay precoders.

8.2 Future Work

The problems studied in this thesis are important in understanding the PNC. These problems can be further investigated in following aspects:

- In Chapter 2, the current work was carried out under amplitude fading only condition. This work can be further extended to more realistic fading channel model. The phase fading further increase the system complexity by bringing more types of clustering cases. Incorporate the phase fading in the system model is of great interest, since it simulates the real communication channel condition.

- In Chapter 3 and Chapter 4, the analysis and design of the channel codes in PNC is limited in binary code. In practical communication systems, higher modulation is always in great interest. Extend the current work from binary codes to higher modulation codes are challenging due to the fact that the increased modulation level will create more complicated mapping between the superimposed codewords and the network coded codewords at the relay. This change will not only complicate the performance analysis process, but also will complicate the optimization operation using EXIT chart.

- In Chapter 7, the study of the DOF of MWRCs with full data exchange is limited in the case where all system users exchange information via the common relay. A more general system model can be a multi-cluster system [103], where multiple user groups are using one common relay, and the information exchange only occurs within each user group. This system model is more likely to happen in real word communication system. For example, grouped users share files with each other, but each group does not share with other group members. This generalized system model requires that signal alignments should be also considered among the user groups. In other words, each group members can only receive signals within their group, and cannot receive signals from other groups.

Therefore, the received signals from different groups should be in orthogonal signal space. In the future, I will focus on the DOF analysis in multi-cluster system model.

Appendix A

Code Parameters in Chapter 4

Table A.1 Code Parameters

Scheme	**Code Type**	**CN**		**VN**		$\bar{d}_c$	$\bar{d}_v$	**Threshold**
		d_c	ρ	d_v	λ			E_b/N_0
CPNC	**Regular** $R = 1/3$	1	1	3	1	1	3	2.2 dB
	Regular $R = 1/2$	1	1	2	1	1	2	4 dB
	Bi-regular $R = 2/3$	1	0.65	4	1	2.67	4	4.8 dB
		4	0.1444					
		7	0.2056					
	Bi-regular $R = 3/4$	1	0.2288	4	1	3	4	6 dB
		3	0.5424					
		5	0.2288					
	Irregular $R = 1/3$	1	0.30	2	0.1542	2.4	7.2	2.1 dB
		3	0.70	3	0.3353			
				7	0.1375			
				8	0.2237			
				21	0.1493			
	Irregular $R = 1/2$	1	0.30	3	0.3612	3.1	6.2	2.4 dB
		4	0.70	4	0.4282			
				16	0.1778			
				17	0.0328			
	Irregular $R = 2/3$	1	0.20	2	0.2243	3.4	5.1	2.9 dB
		4	0.80	3	0.4322			
				6	0.1823			
				7	0.1073			
				28	0.0539			
	Irregular $R = 3/4$	1	0.20	2	0.3221	4.2	5.6	3.4 dB
		5	0.80	3	0.3297			
				6	0.2272			
				7	0.0478			
				31	0.0732			
Complete decoding	**Irregular** $R = 1/3$	1	0.20	3	0.4963	2.6	7.8	1.5 dB
		3	0.80	4	0.1144			
				9	0.0829			
				10	0.2004			
				29	0.0870			
				30	0.0190			
	Irregular $R = 3/4$	1	0.10	2	0.2672	2.8	3.73	5.6 dB
		3	0.90	3	0.5915			
				7	0.0493			
				8	0.0610			
				19	0.0310			

Appendix B

Derivations of Node Update Rules in Chapter 4

B.1 Derivation of VN update rule

Consider a VN with degree d_v. Without loss of generality, we consider the update of the extrinsic information on the first edge, based on the *a priori* information from the edges with index 2 to d_v.

Borrowing the result Eq. (13) of [46] and extending it to $d_v - 1$ edges, we have

$$q_0^{(1)} = \gamma \cdot 4^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_0^{(l)},$$
$$q_1^{(1)} = \gamma \cdot 2^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_1^{(l)},$$
$$q_2^{(1)} = \gamma \cdot 4^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_2^{(l)},$$

where γ is for normalization purpose. The primary LLR $\Lambda_Q^{(1)}$ is calculated by

$$\begin{aligned} \Lambda_Q^{(1)} &= \log\left(\frac{q_0^{(1)} + q_2^{(1)}}{q_1^{(1)}}\right) = \log\left(\frac{4^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_0^{(l)} + 4^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_2^{(l)}}{2^{(d_v-2)} \cdot \prod_{l=2}^{d_v} p_1^{(l)}}\right) \\ &= (d_v - 2)\log 2 + \sum_{l=2}^{d_v} \Lambda_P^{(l')} + K_{\mathrm{VN}}. \end{aligned} \tag{B.1}$$

where

$$K_{\mathrm{VN}} = \log \left(\frac{1 + \prod_{l=2}^{d_v} \exp\left(\Omega_P^{(l')}\right)}{\prod_{l=2}^{d_v} \left(1 + \exp\left(\Omega_P^{(l')}\right)\right)} \right). \tag{B.2}$$

The derivation of the secondary LLR $\Omega_Q^{(1)}$ can be carried similarly.

B.2 Derivation of CN update rule

The CN update function is calculated based on its mapping table. There are two steps to prove the successive update approach for the CN in the CPNC scheme. First step, we generalize the CN mapping table for any degree d_c. Second step, we prove that the generalized CN mapping table can be derived by successively using degree-2 mapping table.

In the ETG, the value on each edge is taken from $\{0, 1, 2\}$. For a CN of degree d_c, let N_θ denote the number of input edges with value $\theta \in \{0, 1, 2\}$, $N_0 + N_1 + N_2 = d_c$. The generalized mapping rule is given as

Case 1: If $N_1 \mod 2 \neq 0$, then the CN outputs 1;

Case 2: If $N_1 \mod 2 = 0$, $N_1 \neq 0$, then the CN outputs 0 or 2;

Case 3: If $N_1 = 0$, we have two subcases:

- Case 3 a) $N_2 \mod 2 \neq 0$, the CN outputs 2;
- Case 3 b) $N_2 \mod 2 = 0$, then the CN outputs 0;

The above cases cover all combinations of the input value for a given CN. We only consider the first case while the manipulation for the other cases can be performed in a similar way. Let us organize the input edges by their superimposed value. Then each user's value relates to the superimposed value as

$$\text{Superimposed value} \ : \ 0\cdots0 \; 1\cdots1 \; 2\cdots2 \tag{B.3}$$

$$\text{User A}'\text{s value} \ : \ 0\cdots0 \; \kappa_1\cdots\kappa_\eta \; 1\cdots1 \tag{B.4}$$

$$\text{User B}'\text{s value} \ : \ 0\cdots0 \; \omega_1\cdots\omega_\eta \; 1\cdots1 \tag{B.5}$$

where $\kappa, \omega \in \{0, 1\}$ and η is the number of edges whose superimposed value is 1.

We now consider the edges with superimposed value of 1 in (B.3) whose total number is N_1. Note that $N_1 = \eta$. For the corresponding edges of user A and user B, we use $N_1^{(A)}$ denotes the number of edges for user A that have bit value $\kappa = 1$, and we use $N_1^{(B)}$ denotes the number of edges for user B that have bit value $\omega = 1$. Note that $N_1^{(A)} + N_1^{(B)} = N_1$. The total number of bits with value 1 in (B.4) is $N_1^{(A)} + N_2$, and the total number of bits with value 1 in (B.5) is $N_1^{(B)} + N_2$. In Case 1, N_1 is an odd number, which means that $N_1^{(A)}$ and $N_1^{(B)}$ can only be one even number and one odd number. This means that $N_1^{(A)} + N_2$ or $N_1^{(B)} + N_2$ can only be one odd number and one even number. Thus, the check node outputs for each user are different, e.g., if user A's check node output is 1, then user B's check node output is 0. This leads to an output superimposed value 1 in the ETG.

We now show that the successive update approach is valid for the first case. Let us temporally use $\mathcal{F}^2_{\text{CN}}(\cdot)$ denotes the mapping function for degree-2 mapping function. From (10) in [46], we have

$$\mathcal{F}^2_{\text{CN}}(1, 1) = \tilde{\theta} \tag{B.6}$$

$$\mathcal{F}^2_{\text{CN}}(1, \tilde{\theta}) = 1 \tag{B.7}$$

where $\tilde{\theta} \in \{0, 2\}$. Note that in (B.6), $\tilde{\theta}$ has equal probability to be 0 or 2. We next consider two scenarios.

Scenarios I: $N_1 = 1$. The successive update function for this case can be written as

$$\mathcal{F}^2_{\text{CN}}(\tilde{\theta}, \cdots \mathcal{F}^2_{\text{CN}}(\tilde{\theta}, \mathcal{F}^2_{\text{CN}}(\tilde{\theta}, \mathcal{F}^2_{\text{CN}}(\tilde{\theta}, 1)))), \tilde{\theta} \in \{0, 2\}. \tag{B.8}$$

The calculation of (B.8) only needs (B.7) and it is straight forward to see the result is 1.

Scenarios II: $N_1 > 1$, and note that $N_1 \mod 2 \neq 0$. In this case, we have $(N_1-1)/2$ pairs of input edges with value 1, and an extra input edge with value 1. Each pair of the edges with value 1 can be updated firstly by using (B.6). By doing this, the CN will have an equivalent $N_0 + (N_1 - 1)/2 + N_2$ number of input edges with value $\tilde{\theta}$ and an extra input edge with value 1. Thus, this case becomes Scenarios I, and the CN outputs 1. The manipulation for Case 2 and Case 3 can also be done in a similar way.

Appendix C

Proofs and Justifications in Chapter 5

C.1 Proof of Theorem 5.1

By the same analysis as in [21], we can show the following inequality for any positive real value ν in the high signal-to-noise ratios regime,

$$P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) \leq \sum_{\boldsymbol{\lambda} \in \mathrm{Nbr}(\Lambda \backslash \Lambda')} K(\Lambda/\Lambda') \exp\left(-\frac{1}{2}\nu\|\boldsymbol{\lambda}\|^2 + \frac{1}{4}\nu^2\|\boldsymbol{\lambda}\|^2|\alpha|^2 N_0\right) \cdot \prod\nolimits_l E\left[\exp\left(\nu \mathrm{Re}\{(\alpha h_l - a_l)\mathbf{x}_l \boldsymbol{\lambda}^{\mathrm{H}}\}\right)\right]$$

where $\mathrm{Nbr}(\Lambda\backslash\Lambda')$ means the set of neighbors of the origin in $\Lambda\backslash\Lambda'$.

Since Λ' is equivalent to $\mathbb{Z}[\omega]^N$ with a scaling factor $\gamma \in \mathbb{C}$, there exists an $n \times n$ unitary matrix $\mathbf{U}$ such that $\mathbf{x}_l\mathbf{U}$ is a vector with the first N entries in the Voronoi region $\mathcal{V}(\gamma\mathbb{Z}[\omega])$ and the last $n - N$ entries equal to 0. Denote by $\mathbf{x}'_l$ the N-dimensional random row vector with i.i.d. components uniformly distributed over $\mathcal{V}(\gamma\mathbb{Z}[\omega])$, and by $\boldsymbol{\lambda}'$ the vector consisting of the first N components in $\boldsymbol{\lambda}\mathbf{U}$. By continuous approximation, $\mathrm{Cov}(\mathbf{x}_l) = \mathrm{Cov}(\mathbf{x}'_l)$. Moreover, since $\mathbf{x}_l\boldsymbol{\lambda}^{\mathrm{H}} = (\mathbf{x}_l\mathbf{U})(\boldsymbol{\lambda}\mathbf{U})^{\mathrm{H}} = \mathbf{x}'_l\boldsymbol{\lambda}'$, Lemma

C.1 in the sequel asserts that

$$\begin{aligned}
&E[\exp(\nu\mathrm{Re}\{(\alpha h_l - a_l)\mathbf{x}_l\boldsymbol{\lambda}^{\mathrm{H}}\})] = E[\exp(\nu\mathrm{Re}\{(\alpha h_l - a_l)\mathbf{x}_l'\boldsymbol{\lambda}'^{\mathrm{H}}\})]\\
&\leq \exp(\frac{5\gamma^2}{144}\nu^2|\alpha h_l - a_l|^2\|\boldsymbol{\lambda}\|^2) = \exp(\frac{nP}{4N}\nu^2|\alpha h_l - a_l|^2\|\boldsymbol{\lambda}\|^2),
\end{aligned}$$

where the last equality holds due to $P = \frac{5N}{36n}|\gamma|^2$ according to Proposition 5.1. Subsequently,

$$\begin{aligned}
&P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a})\\
&\leq \sum_{\boldsymbol{\lambda}\in\mathrm{Nbr}(\Lambda\backslash\Lambda')} \exp(-\frac{1}{2}\nu\|\boldsymbol{\lambda}\|^2 + \frac{1}{4}\nu^2\|\boldsymbol{\lambda}\|^2|\alpha|^2N_0)\\
&\qquad \cdot \prod\nolimits_l \exp(\frac{nP}{4N}\nu^2|\alpha h_l - a_l|^2\|\boldsymbol{\lambda}\|^2)\\
&\leq \sum_{\boldsymbol{\lambda}\in\mathrm{Nbr}(\Lambda\backslash\Lambda')} \exp(-\frac{1}{2}\nu\|\boldsymbol{\lambda}\|^2 + \frac{1}{4}\nu^2\|\boldsymbol{\lambda}\|^2|\alpha|^2N_0\\
&\qquad + \frac{nP}{4N}\|\nu\boldsymbol{\lambda}\|^2\|\alpha\mathbf{h} - \mathbf{a}\|^2)\\
&< \sum_{\boldsymbol{\lambda}\in\mathrm{Nbr}(\Lambda\backslash\Lambda')} \exp(-\frac{1}{2}\nu\|\boldsymbol{\lambda}\|^2 + \frac{1}{4}\nu^2\|\boldsymbol{\lambda}\|^2N_0Q(\alpha, \mathbf{a}))\\
&\leq \sum_{\boldsymbol{\lambda}\in\mathrm{Nbr}(\Lambda\backslash\Lambda')} \exp\left(-\frac{\|\boldsymbol{\lambda}\|^2}{4N_0Q(\alpha, \mathbf{a})}\right),
\end{aligned}$$

where the last inequality is obtained by setting $\nu = \frac{1}{N_0Q(\alpha,\mathbf{a})}$. For high SNRs,

$$\begin{aligned}
P_e(\mathbf{u} \to \hat{\mathbf{u}} \mid \mathbf{h}, \mathbf{a}) &\leq \sum_{\boldsymbol{\lambda}\in\mathrm{Nbr}(\Lambda\backslash\Lambda')} \exp\left(-\frac{\|\boldsymbol{\lambda}\|^2}{4N_0Q(\alpha, \mathbf{a})}\right)\\
&\approx K(\Lambda/\Lambda')\exp\left(-\frac{d^2(\Lambda/\Lambda')}{4N_0Q(a, \mathbf{a})}\right).
\end{aligned}$$

We last show that when $N = n$, every component in $\mathbf{n}$ is uncorrelated with each other and has variance $N_0Q(\alpha, \mathbf{a})$. It suffices to prove the covariance matrix of $\mathbf{x}_l$ equal to $P\mathbf{I}_n$, which is obviously true for $P\mathbf{I}_n = \mathrm{Cov}(\mathbf{x}_l') = \mathrm{Cov}(\mathbf{x}_l\mathbf{U}) = \mathbf{U}^{\mathrm{H}}\mathrm{Cov}(\mathbf{x}_l)\mathbf{U}$.

Lemma C.1. *Let $\mathbf{x} = (x_1, x_2, \cdots, x_N) \in \mathbb{C}^N$ be a row vector uniformly distributed over $\mathcal{V}(\gamma\mathbb{Z}[\omega])^N$, where $\gamma \in \mathbb{C}$. For an arbitrary N-dimensional row vector $\mathbf{v}$ over $\mathbb{C}$,*

$$E[\exp(\mathrm{Re}\{\mathbf{x}\mathbf{v}^{\mathrm{H}}\})] \leq \exp(\frac{5\gamma^2}{144}\|\mathbf{v}\|^2)$$

Proof. It suffices to show the case $\gamma = 1$. In this case, $\mathbf{x}$ is uniformly distributed over the direct product of N regular hexagons with edge length $\sqrt{3}/3$ and centered at the origin. Then,

$$\begin{aligned}
E[\exp(\mathrm{Re}\{\mathbf{x}\mathbf{v}^{\mathrm{H}}\})] =& E[\exp(\mathrm{Re}\{\mathbf{x}\}\mathrm{Re}\{\mathbf{v}\}^T + \mathrm{Im}\{\mathbf{x}\}\mathrm{Im}\{\mathbf{v}\}^T)] \\
=& \prod\nolimits_{j=1}^{N} E[\exp(\mathrm{Re}\{x_j\}\mathrm{Re}\{v_j\} + \mathrm{Im}\{x_j\}\mathrm{Im}\{v_j\})] \\
=& \frac{2\sqrt{3}}{3}\prod\nolimits_j \int_{\mathcal{V}(\mathbb{Z}[\omega])} (\exp(\mathrm{Re}\{x_j\}\mathrm{Re}\{v_j\} + \mathrm{Im}\{x_j\}\mathrm{Im}\{v_j\}) d\mu(x_j)
\end{aligned}$$

where μ is the complex Lebesgue measure. Write $v_j = a + bi$ and $x_j = x + yi$ for each j. Then,

$$\begin{aligned}
&\frac{2\sqrt{3}}{3}\int_{\mathcal{V}(\mathbb{Z}[\omega])} (\exp(\mathrm{Re}\{v_j\}\mathrm{Re}\{x_j\} + \mathrm{Im}\{v_j\}\mathrm{Im}\{x_j\}) d\mu(x_j) \\
&= \frac{2\sqrt{3}}{3}\int_{-\frac{1}{2}}^{0}\int_{-\frac{\sqrt{3}}{3}(1+x)}^{-\frac{\sqrt{3}}{3}(1+x)} (\exp(ax+by) + \exp(-ax+by))dydx \\
&= \frac{8}{b(3a^2-b^2)}(\sqrt{3}a\sinh\frac{a}{2}\sinh\frac{\sqrt{3}b}{6} - b\sinh\frac{\sqrt{3}b}{6}\sinh\frac{\sqrt{3}b}{6} \\
&\quad + b\cosh\frac{a}{2}\cosh\frac{\sqrt{3}b}{6} - b\cosh\frac{\sqrt{3}b}{6}\cosh\frac{\sqrt{3}b}{6}) \\
&= \frac{2}{3b'(a'^2-b'^2)}(a'\sinh a'\sinh b' - b'\sinh b'\sinh b' \\
&\quad + b'\cosh a'\cosh b' - b'\cosh b'\cosh b') \ // \ a' = \frac{a}{2}; b' = \frac{\sqrt{3}b}{6} \\
&= \frac{2}{3}(\frac{\sinh b'}{b'}\frac{a'\sinh a' - b'\sinh b'}{a'^2-b'^2} + \cosh b'\frac{\cosh a' - \cosh b'}{a'^2-b'^2}) \\
&= \frac{2}{3}\frac{\sinh b'}{b'}(\frac{a'\sinh a' - b'\sinh b' + \cosh a' - \cosh b'}{a'^2-b'^2}) \\
&\quad + \frac{2}{3}(\cosh b' - \frac{\sinh b'}{b'})(\frac{\cosh a' - \cosh b'}{a'^2-b'^2}) \triangleq A + B
\end{aligned}$$

Next we shall upper bound A and B respectively:

$$\begin{aligned}
A \le\ & \frac{2\sinh b'}{3b'}(1+\frac{1}{2!}+\frac{a'^2+b'^2}{3!}+\frac{a'^2+b'^2}{4!}\\
&+\frac{a'^4+b'^4+a'^2b'^2}{5!}+\frac{a'^4+b'^4+a'^2b'^2}{6!}+\ldots)\\
\le\ & \frac{\sinh b'}{b'}\cdot(1+\frac{5}{36}(a'^2+b'^2)+\frac{7}{1080}(a'^4+b'^4+a'^2b'^2)+\ldots)\\
\le\ & \frac{\sinh b'}{b'}\cdot\exp(\frac{5}{36}(a'^2+b'^2))\\
B =\ & \frac{2}{3}(\frac{2b'^2}{3!}+\frac{4b'^4}{5!}+\frac{6b'^6}{7!}+\cdots)(\frac{1}{2!}+\frac{a'^2+b'^2}{4!}\\
&+\frac{a'^4+b'^4+a'^2b'^2}{6!}+\ldots)\\
\le\ & \frac{1}{3}(\frac{2b'^2}{3!}+\frac{4b'^4}{5!}+\frac{6b'^6}{7!}+\cdots)\exp(\frac{a'^2+b'^2}{12})
\end{aligned}$$

Thus, $A+B \le \exp(\frac{5}{36}(a'^2+b'^2))[1+\frac{b'^2}{3!}+\frac{b'^4}{5!}+\cdots+\frac{1}{3}(\frac{2b'^2}{3!}+\frac{4b'^4}{5!}+\cdots)] \le \exp(\frac{5}{36}(a'^2+b'^2))\exp(\frac{5b'^2}{18})) = \exp(\frac{5}{144}(a^2+b^2))$. Consequently, $E[\exp(\mathrm{Re}\{\mathbf{x}\mathbf{v}^{\mathrm{H}}\})] \le \exp(\frac{5\gamma^2}{144}\|\mathbf{v}\|^2)$.

■

C.2 Justification of Algorithm 5.2

We shall first show that the output vector $\mathbf{u}_1$ of the algorithm is a shortest nonzero vector in Λ. It is equivalent to show that for any nonzero lattice point $\mathbf{v}$ in Λ, $\|\mathbf{v}\| \ge \|\mathbf{u}_1\|$.

As stated in the remark following Algorithm 5.2, $\{\mathbf{u}_1, \mathbf{u}_2\}$ keeps to be a basis of Λ. Thus, the vector $\mathbf{v}$ can be written as $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$, where $\alpha_1, \alpha_2 \in \mathbb{Z}[\omega]$. Correspondingly,

$$\|\mathbf{v}\|^2 = |\alpha_1|^2 \cdot \|\mathbf{u}_1\|^2 + |\alpha_2|^2 \cdot \|\mathbf{u}_2\|^2 + 2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \tag{C.1}$$

Because $\alpha_1\alpha_2^*$ is also an element in $\mathbb{Z}[\omega]$, $|\alpha_1\alpha_2^*|^2$ is a rational nonnegative integer. We shall divide the discussion into the following three cases:

(i) $|\alpha_1\alpha_2^*|^2 = 0$. In this case, we have either $\alpha_1 = 0$ or $\alpha_2 = 0$. If $\alpha_1 = 0$, then $\alpha_2 \ne 0$ and thus $|\alpha_2|^2 \ge 1$. Consequently, $\|\mathbf{v}\|^2 \ge |\alpha_2|^2\|\mathbf{u}_2\|^2 \ge |\alpha_2|^2\|\mathbf{u}_1\|^2 \ge \|\mathbf{u}_1\|^2$. Likewise, if $\alpha_2 = 0$, we have $\|\mathbf{v}\|^2 \ge |\alpha_1|^2\|\mathbf{u}_1\|^2 \ge \|\mathbf{u}_1\|^2$.

The proof in the other two cases needs the following two properties of the two output vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ of the algorithm: (a) $\|\mathbf{u}_1\| \leq \|\mathbf{u}_2\|$; (b) $\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2$ is always in the Voronoi region $\mathcal{V}(\mathbb{Z}[\omega])$. The second property holds because when the algorithm terminates, $\mathbf{u}_2$ is set to be equal to $\mathbf{u}_2'$, which is calculated from $\mathbf{u}_1$ and non-updated $\mathbf{u}_2$ in Step 3 of the algorithm, and $\mathbf{u}_1{\mathbf{u}_2'}^{\mathrm{H}}/\|\mathbf{u}_1\|^2 = (\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}} - \|\mathbf{u}_1\|^2\mathcal{D}_{\mathbb{Z}[\omega]}(\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2))/\|\mathbf{u}_1\|^2 = \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2 - \mathcal{D}_{\mathbb{Z}[\omega]}(\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2) \in \mathcal{V}(\mathbb{Z}[\omega])$.

(ii) $|\alpha_1\alpha_2^*|^2 = 1$. Since the norm of every nonzero Eisenstein integer is a positive integer, $|\alpha_1|^2 = |\alpha_2|^2 = |\alpha_1\alpha_2^*|^2 = 1$, *i.e.*, all of α_1, α_2 and $\alpha_1\alpha_2^*$ are units in $\mathbb{Z}[\omega]$. Since $\mathcal{V}(\mathbb{Z}[\omega])$ is a regular hexagon, $\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2 \in \mathcal{V}(\mathbb{Z}[\omega])$ implies $\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2 \in \mathcal{V}(\mathbb{Z}[\omega])$. Using the fact that all points in $\mathcal{V}(\mathbb{Z}[\omega])$ have real parts no less than $-1/2$, we get

$$2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \geq -\|\mathbf{u}_1\|^2. \tag{C.2}$$

Continuing the argument in (C.1), $\|\mathbf{v}\|^2 = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + 2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \geq \|\mathbf{u}_2\|^2 \geq \|\mathbf{u}_1\|^2$.

(iii) $|\alpha_1\alpha_2^*|^2 > 1$. Because every point in $\mathcal{V}(\mathbb{Z}[\omega])$ has Euclidean norm no larger than $1/\sqrt{3}$, $\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}/\|\mathbf{u}_1\|^2 \in \mathcal{V}(\mathbb{Z}[\omega])$ implies $|\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}| \leq \frac{1}{\sqrt{3}}\|\mathbf{u}_1\|^2$, and thus

$$2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \geq -2|\alpha_1\alpha_2^*| \cdot |\mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}| \geq -\frac{2}{\sqrt{3}}|\alpha_1\alpha_2^*| \cdot \|\mathbf{u}_1\|^2 \tag{C.3}$$

Continuing the argument in (C.1), we have $\|\mathbf{v}\|^2 = |\alpha_1|^2 \cdot \|\mathbf{u}_1\|^2 + |\alpha_2|^2 \cdot \|\mathbf{u}_2\|^2 + 2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \geq (|\alpha_1|^2 + |\alpha_2|^2 - \frac{2}{\sqrt{3}}|\alpha_1\alpha_2^*|) \cdot \|\mathbf{u}_1\|^2 \geq \|\mathbf{u}_1\|^2$, where the last inequality holds because at least one of α_1, α_2 is a non-unit, and every nonzero non-unit Eisenstein integer has norm no smaller than 3.

This completes the proof that the output $\mathbf{u}_1$ is indeed a shortest vector in Λ. Now we justify that the output vector $\mathbf{u}_2$ of the algorithm is a shortest vector in $\Lambda\backslash\{\beta\mathbf{u}_1 : \beta \in \mathbb{Z}[\omega]\}$.

Let $\mathbf{v}$ be a nonzero lattice point in Λ which is $\mathbb{Z}[\omega]$-linearly independent of $\mathbf{u}_1$, and denote by α_1, α_2 the two Eisenstein integers so that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$. Note that $\alpha_2 \neq 0$. We shall show $\|\mathbf{v}\| \geq \|\mathbf{u}_2\|$ by considering the following four cases.

1. $|\alpha_1|^2 = 0$. In this case, $\|\mathbf{v}\|^2 = |\alpha_2|^2 \cdot \|\mathbf{u}_2\|^2 \geq \|\mathbf{u}_2\|^2$

2. $|\alpha_1|^2 = |\alpha_2|^2 = 1$. In this case, $\|\mathbf{v}\|^2 = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + 2\mathrm{Re}(\alpha_1\alpha_2^* \cdot \mathbf{u}_1\mathbf{u}_2^{\mathrm{H}}) \geq \|\mathbf{u}_2\|^2$, where the inequality follows from (C.2).

3. $|\alpha_1|^2 > 1$, $|\alpha_2|^2 = 1$. By combining (C.3) into (C.1), $\|\mathbf{v}\|^2 \geq (|\alpha_1|^2 - \frac{2}{\sqrt{3}}|\alpha_1|) \cdot \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 \geq \|\mathbf{u}_2\|^2$ where the last inequality is based on the fact that $|\alpha_1| \geq \sqrt{3}$.

4. $|\alpha_2|^2 > 1$. By combining (C.3) into (C.1), $\|\mathbf{v}\|^2 \geq |\alpha_1|^2 \cdot \|\mathbf{u}_1\|^2 + |\alpha_2|^2 \cdot \|\mathbf{u}_2\|^2 - \frac{2}{\sqrt{3}}|\alpha_1| \cdot |\alpha_2| \cdot \|\mathbf{u}_1\|^2 = (|\alpha_1| - \frac{1}{\sqrt{3}})^2\|\mathbf{u}_1\|^2 - \frac{1}{3}\|\mathbf{u}_1\|^2 + |\alpha_2|^2 \cdot \|\mathbf{u}_2\|^2 \geq (|\alpha_1| - \frac{1}{\sqrt{3}})^2\|\mathbf{u}_1\|^2 + (|\alpha_2|^2 - \frac{1}{3})\|\mathbf{u}_2\|^2 \geq \|\mathbf{u}_2\|^2$, where the last two inequalities follow from $\|\mathbf{u}_1\| \leq \|\mathbf{u}_2\|$ and $|\alpha_2|^2 \geq 3$ respectively.

C.3 Proof of Theorem 5.2

Let $\mathbf{c}_1, \cdots, \mathbf{c}_k$ be such a set of basis for the linear code C that their juxtaposition in rows forms a generator matrix over $R/\pi R$ in the systematic form. Correspondingly, let $\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_k$ be a set of row vectors over R such that:

1. $\sigma(\boldsymbol{\lambda}_i) = \mathbf{c_1}$;

2. $\sum_{j=1}^{n} \lambda_{ij} = 0$, where λ_{ij} is the j^{th} entry in $\boldsymbol{\lambda}_i$;

3. The juxtaposition of $\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_k$ by rows forms a $k \times n$ matrix over R in which the first k columns form an identity matrix.

Due to the condition $\sum_{i=1}^{n} c_i = 0$ for each $(c_1, \cdots, c_n) \in C$, such a choice of $\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_k$ always exists.

Now establishing the matrices $\mathbf{M}_\Lambda$ and $\mathbf{M}_{\Lambda'}$ in (5.11) by describing B_{n-k} to be the last $n-k$ columns in the juxtaposition of $\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_k$ by rows. Obviously $\mathbf{M}_{\Lambda'}$ is a generator matrix for Λ'. For $\mathbf{M}_\Lambda$, observe that it is of full rank and every row in it is a vector in the R-lattice Λ defined in (5.10). Thus, the R-lattice generated by $\mathbf{M}_\Lambda$ is a sublattice of Λ. Reversely, consider an arbitrary vector $\boldsymbol{\lambda}$ in Λ. Then,

$\sigma(\boldsymbol{\lambda}) = \sum_{i=1}^{k} a_i \mathbf{c}_i$ for some $a_i \in R/\pi R$, and the sum of components in $\boldsymbol{\lambda}$ is congruent to 0 modulo π^2. Let $\boldsymbol{\lambda}'$ be an arbitrary R-linear combination $\sum_{i=1}^{k} \alpha_i \boldsymbol{\lambda}_i$ subject to $\sigma(\alpha_i) = a_i$. In order to prove that $\boldsymbol{\lambda}$ is in the R-lattice generated by $\mathbf{M}_\Lambda$, it remains to show that so is $\boldsymbol{\lambda} - \boldsymbol{\lambda}'$. Since

$$\sigma(\boldsymbol{\lambda} - \boldsymbol{\lambda}') = \sigma(\boldsymbol{\lambda}) - \sigma(\sum_{i=1}^{k} \alpha_i \boldsymbol{\lambda}_i) = \mathbf{0}$$

and the sum of components in $\boldsymbol{\lambda} - \boldsymbol{\lambda}'$ is congruent to 0 modulo π^2, the vector $\boldsymbol{\lambda} - \boldsymbol{\lambda}'$ can be written as an R-linear combination of the rows in the $n \times n$ matrix:

$$\begin{bmatrix} \pi & -\pi & 0 & \ldots & 0 \\ & & \ddots & & \\ 0 & \ldots & 0 & \pi & -\pi \\ 0 & 0 & \ldots & 0 & \pi^2 \end{bmatrix}.$$

Because each row in this matrix is in turn an R-linear combination of the rows in $\mathbf{M}_\Lambda$, the vector $\boldsymbol{\lambda} - \boldsymbol{\lambda}'$, and thus $\boldsymbol{\lambda}$ is in the R-lattice generated by $\mathbf{M}_\Lambda$. This verifies that $\mathbf{M}_\Lambda$ is a generator matrix for Λ.

Now consider the special case $R = \mathbb{Z}[\omega]$ or $\mathbb{Z}[i]$. Since $(\pi, -\pi, 0, \cdots, 0)$ is a vector in $\Lambda \backslash \Lambda'$ with squared length $2|\pi|^2$, $d^2(\Lambda/\Lambda') = \min(2|\pi|^2, w_E^{min}(C))$. On the other hand, $(\pi^2, 0, \cdots, 0)$ is a shortest nonzero vector in Λ' with squared length $|\pi^2|^2 = |\pi|^4$. Thus, when $|\pi|^4 > 2|\pi|^2$, there is no shortest nonzero vector in Λ that is also in Λ', which implies $d^2(\Lambda/\Lambda') = d^2(\Lambda)$ and $K(\Lambda/\Lambda') = K(\Lambda)$. However, the only case for $|\pi|^4 \leq 2|\pi|^2$ is a Gaussian prime π with norm 2. In this case, $d^2(\Lambda/\Lambda') = \min(2|\pi|^2, w_E^{min}(C)) = \min(|\pi|^4, w_E^{min}(C)) = d^2(\Lambda)$. The proof is complete.

Appendix D

Tables in Chapter 5

Table D.1 The mappings from $\mathbb{F}_{13} \cong \mathbb{Z}[i]/(2+3i)\mathbb{Z}[i] \cong \mathbb{Z}[\omega]/(4+3\omega)\mathbb{Z}[\omega]$ to $\mathbb{Z}[i]$ and $\mathbb{Z}[\omega]$ by $\mathcal{E}_{\mathbb{Z}[i]} : \mathbb{Z}[i]/(2+3i)\mathbb{Z}[i] \to \mathbb{Z}[i]$ and $\mathcal{E}_{\mathbb{Z}[\omega]} : \mathbb{Z}[\omega]/(4+3\omega\mathbb{Z})[\omega] \to \mathbb{Z}[\omega]$.

$\mathbb{F}_{13}$	0	1	2	3	4	5	6	7	8	9	10	11	12
$\mathcal{E}_{\mathbb{Z}[i]}(\mathbb{F}_{13})$	0	1	2	$2i$	$-1-i$	$-i$	$1-i$	$-1+i$	i	$1+i$	$-2i$	-2	-1
$\mathcal{E}_{\mathbb{Z}[\omega]}(\mathbb{F}_{13})$	0	1	$-1+\omega$	ω	$1+\omega$	$2+\omega$	$-1-2\omega$	$1+2\omega$	$-2-\omega$	$-1-\omega$	$-\omega$	$1-\omega$	-1

Table D.2 Rate-1/2 convolutional codes C over $\mathbb{Z}[i]/(2+3i)\mathbb{Z}[i]$ ($\cong \mathbb{F}_{13}$) with maximum $w_E^{min}(C)$, and the corresponding LNCs Λ/Λ' with Λ constructed from C by (5.9) and $\Lambda' = ((2+3i)\mathbb{Z}[i])^n$

v	$\mathbf{g}(D)$	$\gamma_c(\Lambda/\Lambda')$	$w_E^{min}(C)$	$K(\Lambda/\Lambda')$
1	$1+2D$ $2+(1+i)D$	2.22 (3.46dB)	8	4
2	$1+D+(2i)D^2$ $(-1-i)+(-1+i)D+(-1-i)D^2$	3.33 (5.22dB)	12	4
3	$2+(1-i)D+(2i)D^2+(-2)D^3$ $1+(1+i)D+iD^2+iD^3$	4.44 (6.47dB)	16	8
4*	$(-2i)+(-i)D+(2i)D^2+(-1)D^3+(i)D^4$ $(-1)+2D+0D^2+(-1+i)D^3+(1-i)D^4$	4.99 (6.98dB)	18	4
5*	$(-2)+(-i)D^2+(-1)D^3+(1-i)D^4+D^5$ $(-1+i)+(2i)D+(-2)D^3+(-1+i)D^4+(-1-i)D^5$	5.82 (7.65dB)	21	16
$*$ not exhaustive search				

Table D.3 Rate-1/2 convolutional codes C over $\mathbb{Z}[\omega]/(4+3\omega)\mathbb{Z}[\omega](\cong \mathbb{F}_{13})$ with maximum $w_E^{min}(C)$, and the corresponding convolutional LNCs Λ/Λ' with Λ constructed from C by (5.9) and $\Lambda' = ((4+3\omega)\mathbb{Z}[\omega])^n$

v	$\mathbf{g}(D)$	$\gamma_c(\Lambda/\Lambda')$	$w_E^{min}(C)$	$K(\Lambda/\Lambda')$
1	$1+D$	2.56 (4.09dB)	8	12
	$(-1+w)+(2+w)D$			
2	$1+D+(-1+w)D^2$	3.85 (5.85dB)	12	24
	$(-1+w)+(1-w)D+(1+w)D^2$			
3	$(2+w)+(1+2w)D+(1+2w)D^2+(-1-2w)D^3$	5.13 (7.10dB)	16	96
	$(-w)+(w)D+(w)D^2+(1+w)D^3$			
4*	$(-1)+(-w)D+(1-w)D^2+(-2-w)D^3+(1-w)D^4$	5.76 (7.61dB)	18	30
	$(2+w)+(1+w)D+(-1-w)D^2+(-1-2w)D^3+D^4$			
5*	$(1+w)+(1+w)D+(-1+w)D^2+(-1-2w)D^3+(-1-2w)D^4+(-w)D^5$	5.76 (7.61dB)	18	6
	$(1+w)+(-1+w)D+(0)D^2+(1+w)D^3+(-2-w)D^4+(1-w)D^5$			
$*$ not exhaustive search				

Table D.4 Parameters in various LNCs Λ/Λ' constructed from $[12, 6, 6]$ ternary Golay code by different methods.

Λ	By Complex Construction A (Formula (5.9) in Algorithm 5.3)		By Complex Construction B (Formula (5.10) in Algorithm 5.4)	Complex Leech Lattice
$\Lambda' =$	$(\mathbb{Z}[w]/\pi\mathbb{Z}[w])^{12}$	$(\mathbb{Z}[w]/\pi^2\mathbb{Z}[w])^{12}$	$(\mathbb{Z}[w]/\pi^2\mathbb{Z}[w])^{12}$	$(\mathbb{Z}[w]/\pi^3\mathbb{Z}[w])^{12}$
$\rho =$	$\frac{1}{2}\log_2 3$	$\frac{3}{2}\log_2 3$	$\frac{17}{12}\log_2 3$	$\frac{3}{2}\log_2 3$
$d^2(\Lambda/\Lambda') =$	6	3	6	3
$\gamma_c(\Lambda/\Lambda') =$	4 (6.02 dB)	2 (3.01 dB)	3.65 (5.62 dB)	4 (6.02 dB)

Appendix E

Verification of Derived Sum-Rate in Chapter 7

In this appendix, we verify the derived sum-rate upper bound in (7.63), against the original sum-rate upper bound in (7.55). The sum-rate upper bound in (7.55) can be written as:

$$\begin{aligned} R^{\mathrm{UB}} &= \sum_{i=1}^{K} \frac{1}{T} R_i^{\mathrm{UB}} \\ &= \sum_{i=1}^{K} \frac{1}{T} \log \left(\frac{\left| \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}} \mathbf{P}_{\bar{i}}^{\mathrm{H}} \mathbf{H}_{\bar{i}}^{\mathrm{H}} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \sigma^2 (\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I}) \right|}{\left| \sigma^2 (\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I}) \right|} \right), \end{aligned} \tag{E.1}$$

We repeat $\mathbf{\Omega}_i$ and $\mathbf{\Delta}_i$ here:

$$\begin{aligned} \mathbf{\Omega}_i &= \mathrm{Cov}(\mathbf{s}_{\bar{i}}, \mathbf{y}_{\bar{i}}) \mathrm{Cov}^{-1}(\mathbf{y}_{\bar{i}}, \mathbf{y}_{\bar{i}}) \\ &= (\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})^{\mathrm{H}} \big((\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}}) (\mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}})^{\mathrm{H}} + \sigma^2 (\mathbf{G}_i \mathbf{F} \mathbf{F}^{\mathrm{H}} \mathbf{G}_i^{\mathrm{H}} + \mathbf{I}) \big)^{-1}, \end{aligned} \tag{E.2a}$$

$$\begin{aligned} \mathbf{\Delta}_i &= \mathbf{I} - \mathbf{\Omega}_i \mathrm{Cov}(\mathbf{y}_{\bar{i}}, \mathbf{s}_{\bar{i}}) \\ &= \mathbf{I} - \mathbf{\Omega}_i \mathbf{G}_i \mathbf{F} \mathbf{H}_{\bar{i}} \mathbf{P}_{\bar{i}}. \end{aligned} \tag{E.2b}$$

We now show that $\log |\mathbf{\Delta}_i| = -R_i^{\mathrm{UB}}$ in (E.1), for $i = 1, \cdots, K$. Let $\mathbf{U}_i =$

$\mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}$, and let $\mathbf{V}_i = \mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I}$, we have

$$\begin{aligned}
\log|\boldsymbol{\Delta}_i| &= \log|\mathbf{I} - \boldsymbol{\Omega}_i\mathbf{U}_i| \\
&= \log\left|\mathbf{I} - {\mathbf{U}_i}^{\mathrm{H}}(\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i)^{-1}\mathbf{U}_i\right| \\
&\overset{(a)}{=} \log\left|\mathbf{I} - \mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}}(\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i)^{-1}\right| \\
&= \log\left(\left|\mathbf{I} - \mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}}(\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i)^{-1}\right|\frac{\left|\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i\right|}{\left|\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i\right|}\right) \\
&= \log\left|\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i - \mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}}\right| - \log\left|\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i\right| \\
&= -\log\left(\frac{\left|\mathbf{U}_i{\mathbf{U}_i}^{\mathrm{H}} + \sigma^2\mathbf{V}_i\right|}{|\sigma^2\mathbf{V}_i|}\right) \\
&= -R_i^{\mathrm{UB}}, \qquad \text{(E.3)}
\end{aligned}$$

where step (a) use the fact that $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$ when $\mathbf{AB}$ and $\mathbf{BA}$ are both square matrices.

Next, we show that actually in (7.63), we have

$$\mathrm{Tr}\left(\boldsymbol{\Delta}_i^{-1}\mathbf{D}_i\right) = (K-1)\,L,$$

where

$$\mathbf{D}_i = \left(\mathbf{I} - \boldsymbol{\Omega}_i\mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}\right)\left(\mathbf{I} - \boldsymbol{\Omega}_i\mathbf{G}_i\mathbf{F}\mathbf{H}_{\bar{i}}\mathbf{P}_{\bar{i}}\right)^{\mathrm{H}} + \sigma^2\boldsymbol{\Omega}_i\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right)\boldsymbol{\Omega}_i^{\mathrm{H}}.$$

Thus, we have

$$\begin{aligned}
\mathrm{Tr}\left(\boldsymbol{\Delta}_i^{-1}\mathbf{D}_i\right) &= \mathrm{Tr}\left(\boldsymbol{\Delta}_i^{-1}\left(\boldsymbol{\Delta}_i\boldsymbol{\Delta}_i^{\mathrm{H}} + \sigma^2\boldsymbol{\Omega}_i\left(\mathbf{G}_i\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{G}_i^{\mathrm{H}} + \mathbf{I}\right)\boldsymbol{\Omega}_i^{\mathrm{H}}\right)\right) \\
&= \mathrm{Tr}\left(\boldsymbol{\Delta}_i^{\mathrm{H}}\right) + \mathrm{Tr}\left(\sigma^2\boldsymbol{\Delta}_i^{-1}\boldsymbol{\Omega}_i\mathbf{V}_i\boldsymbol{\Omega}_i^{\mathrm{H}}\right) \\
&= \mathrm{Tr}(\mathbf{I}) - \mathrm{Tr}\left((\boldsymbol{\Omega}_i\mathbf{U}_i)^{\mathrm{H}}\right) + \mathrm{Tr}\left(\sigma^2\boldsymbol{\Delta}_i^{-1}\boldsymbol{\Omega}_i\mathbf{V}_i\boldsymbol{\Omega}_i^{\mathrm{H}}\right) \\
&\overset{(b)}{=} (K-1)L + \mathrm{Tr}\left(\sigma^2\boldsymbol{\Delta}_i^{-1}\boldsymbol{\Omega}_i\mathbf{V}_i\boldsymbol{\Omega}_i^{\mathrm{H}} - (\boldsymbol{\Omega}_i\mathbf{U}_i)^{\mathrm{H}}\right), \qquad \text{(E.4)}
\end{aligned}$$

where step (b) follows the fact that the $\mathbf{I}$ in $\boldsymbol{\Delta}_i$ has the size $(K-1)L \times (K-1)L$. Now we only need to show that in (E.4)

$$\sigma^2\boldsymbol{\Delta}_i^{-1}\boldsymbol{\Omega}_i\mathbf{V}_i\boldsymbol{\Omega}_i^{\mathrm{H}} = (\boldsymbol{\Omega}_i\mathbf{U}_i)^{\mathrm{H}}. \qquad \text{(E.5)}$$

From (E.2a), we can obtain

$$\mathbf{V}_i = \frac{\mathbf{\Omega}_i^{-1}\mathbf{U}_i^{\mathrm{H}} - \mathbf{U}_i\mathbf{U}_i^{\mathrm{H}}}{\sigma^2}. \tag{E.6}$$

Substitute (E.2b) and (E.6) into the left hand side of (E.5), we immediate obtain the right hand side of (E.5). Thus we have Eqn. (E.4) $= (K-1)L$.

Bibliography

[1] S. Zhang, S. C. Liew, P. P. Lam, “Hot topic: physical-layer network coding,” in *Proc. 12th ACM MobiCom*, pp. 358-365, New York, NY, USA, 2006.

[2] A. Goldsmith, “Wireless Communications,” Cambridge University Press, 2005.

[3] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204-1216, July 2000.

[4] S.-Y. R. Li, R.W. Yeung and N.Cai, ”Linear Network Coding,” IEEE Trans. Inform. Theory, vol. 49, no.2, pp. 371-381, Feb. 2003.

[5] Y. Wu, P.A. Chou, S.Y. Kung, “Information exchange in wireless networks with network coding and physical layer broadcast”, in *Proc. 39th Annual Conf. Inform. Sci. and Systems* (CISS), 2005.

[6] C. Fragouli, J.Y. Boudec, J. Widmer, “Network coding: an instant primer”, *ACM SIGCOMM Comput. Commun.* Rev. 36 (1), pp. 63-68, 2006

[7] W. Chen, Z. Cao, and L. Hanzo, “Maximum euclidean distance network coded modulation for asymmetric decode-and-forward two-way relaying,” *IET Communications*, vol. 7, no. 10, pp.988-998, July 2013.

[8] P. Popovski, H. Yomo, “The anti-packets can increase the achievable throughput of a wireless multi-hop network,” in *Proc. IEEE International Conference on Communication (ICC)*, pp. 3885-3890, Istanbul, Turkey, Jun. 2006.

[9] C. Shannon, “The mathematical theory of communication (parts 1 and 2),” *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, 1948.

[10] T. Richardson, and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[11] S. Katti, S.S. Gollakota, D. Katabi, “Embracing wireless interference: analog network coding,” in *Proc ACM SIGCOMM*, pp. 397-408, Aug. 2007.

[12] R. de Buda, “The upper error bound of a new near-optimal code,” in *IEEE Trans. Inf. Theory*, vol. 21, pp. 441-445, July 1975.

[13] R. de Buda, “Some optimal codes have structure,” in *IEEE J. Select. Areas Commun.*, vol. 7, pp. 893-899, Aug. 1989.

[14] T. Yang and I. B. Collings, “Asymptotically optimal error-rate performance of linear physical-layer network coding in Rayleigh fading two-way relay channels,” *IEEE Commun. Letters*, vol. 16, no. 7, pp. 1068-1071, July 2012.

[15] A. J. Aljohani, S. X. Ng, R. G. Maunder and L. Hanzo, “EXIT-chart aided joint source-coding, channel-coding and modulation design for two-way relaying,” *IEEE Trans. Vehicular Technol.*, vol. 62, no. 6, pp.2496-2506, July 2013.

[16] W. Liang, S. X. Ng, and L. Hanzo, “TTCM-Aided SDMA-Based Two-Way Relaying,” In *Proceeding of IEEE Vehicular Technology Conference (VTC Fall)*, San Francisco, USA, Sept. 2011.

[17] T. Linder, Ch. Schlegel, and K. Zeger, “Corrected proof of de Buda’s Theorem,” in *IEEE Trans. Inf. Theory*, pp. 1735-1737, Sept. 1993.

[18] H. A. Loeliger, “Averaging bounds for lattices and linear codes,” in *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767-1773, Nov. 1997.

[19] R. Urbanke and B. Rimoldi, “Lattice codes can achieve capacity on the AWGN channel,” in *IEEE Trans. Inf. Theory*, pp. 273-278, Jan. 1998.

[20] U. Erez and R. Zamir, “Achieving 1/2 log(1 + SNR) on the AWGN channel with lattice encoding and decoding,”in *IEEE Trans. Inf. Theory*, vol. 50, pp. 2293-2314, Oct. 2004.

[21] C. Feng, D. Silva, and F. R. Kschischang, “An algebraic approach to physical-layer network coding,” *IEEE Trans. Inform. Theory*, vol. 59, no. 11, pp. 7576-7596, Nov. 2013.

[22] M. R. Bremner, *Lattice basis reduction – an introduction to the LLL algorithm and its applications*, CRC Press, Boca Raton, 2012.

[23] B. Nazer and M. Gastpar, “Compute-and-forward: harnessing interference through structured codes,” *IEEE Trans. Inform. Theory*, vol. 57, no. 10, p-p. 6463-6486, Oct. 2011.

[24] T. Huang, J. Yuan, and J. (Tiffany) Li, “Analysis of Compute-and-Forward with QPSK in Two-way Relay Fading Channels,” in *Proc. of the 14th Australian Communications Theory Workshop (AusCTW)*, pp. 75-80, Adelaide, Australia, Jan. 2013.

[25] T. Yang and I. Collings, “On the Optimal Design and Performance of Linear Physical-Layer Network Coding for Fading Two-Way Relay Channels,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp.956-967, Feb. 2014.

[26] T. Koike-Akino, P. Popovski, V. Tarokh, “Optimized constellations for two-way wireless relaying with physical network coding,” *IEEE J. Select. Areas Commun.*, vol. 27, no. 5, pp. 773-787, June 2009.

[27] W. Nam, S. Y. Chung, and Y. H. Lee, “Capacity of the Gaussian two-way relay channel to within 1/2 bit,” *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488-5494, Nov. 2010.

[28] S. Zhang, S. C. Liew, Q. Zhou, L. Lu, W. Wang, "Non-memoryless analog network coding in two-way relay channel," *Proc. IEEE International Conference on Communication (ICC)*, June 2011.

[29] R. Knopp, "Two-way wireless communication via relay station," *GDR-ISIS Meeting*, ENST, Paris, Mar. 2007.

[30] S. C. Liew, S. Zhang, L. Lu, "Physical-layer Network Coding: Tutorial, Survey, and Beyond," Invited Paper, Elsevier Phycom, Special Issue of Physical Communication on *Network Coding and Its Applications to Wireless Communications*, vol. 6, pp. 4-42, Mar. 2013.

[31] F. Rossetto and M. Zorzi, "On the design of practical asynchronous physical layer network coding," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications* (SPAWC), pp. 469-473, Perugia, Italy, June 2009.

[32] P. Popovski, and H. Yomo, "Physical network coding in two-way wireless relay network," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 707-712, June 2007.

[33] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5641-5654, Nov. 2010.

[34] B. Nazer, and M. Gastpar, "Lattice coding increases multicast rates for gaussian multiple-access networks," in *Proc. Annual Allerton Conference*, pp. 1089-1096, Sept. 2007

[35] D. To and J. Choi, "Convolutional codes in two-way relay networks with physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2724-2729, Sept. 2010.

[36] S. Zhang, S. C. Liew, L. Lu, “Physical layer network coding schemes over finite and infinite fields”, in *Proc. IEEE Globecom*, New Orleans, LA, Nov. 2008.

[37] T. Yang, I. Land, T. Huang, J. Yuan, and Z. Chen, “Distance Spectrum and Performance of Channel-Coded Physical-Layer Network Coding for Binary-Input Gaussian Two-Way Relay Channels,” *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1499-1510, June 2012.

[38] T. Yang, I. Land, T. Huang, J. Yuan, and Z. Chen, “Distance Properties and Performance of Physical Layer Network Coding with Binary Linear Codes for Gaussian Two-Way Relay Channels”, in *Proceeding of IEEE International Symposium on Information Theory (ISIT)*, pp. 2070-2074, Saint Petersburg, Russia, Aug. 2011.

[39] S. Lu, Y. Li, and J. Cheng, “Low-complexity turbo decoding scheme for two-way relay network, in *Proc. International Conference on Wireless Communications and Signal Processing*, 2010.

[40] D. To, J. Choi, “Reduced-state decoding with two-way relay networks with physical-layer network coding”, in *Proc. IEEE Information Theory Workshop (ITW)*, Aug. 2010.

[41] D. Wang, S. Fu, and K. Lu, “Channel coding design to support asynchronous physical layer network coding,” in *Proceedings of Globecom.* Honolulu, Hawaiian, Nov. 2009.

[42] Q. Yang and S. C. Liew, “Asynchronous Convolutional-Coded Physical-Layer Network Coding,” Available: http://arxiv.org/abs/1312.1447.

[43] H. J. Yang, Y. Choi, J. Chun, “Modified higher-order PAMS for binarycoded physical-layer network coding,” *IEEE Commun. Letters*, vol. 14, no. 8, pp. 689-691, Aug. 2010.

[44] V. Namboodiri, K. Venugopal, B.S. Rajan, “Physical Layer Network Coding for Two-Way Relaying with QAM,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5074-5086, Oct. 2013.

[45] Z. Faraji-Dana, and P. Mitran, “On Non-Binary Constellations for Channel-Coded Physical-Layer Network Coding,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 312-319, Jan. 2013.

[46] S. Zhang and S. C. Liew, “Channel coding and decoding in a relay system operated with physical-layer network coding,” *IEEE Jour. Select Area. Commun.*, vol. 27, pp. 788-796, June 2009.

[47] M. C. Castro, B. F. Uchoa-Filho, T. T. V. Vinhoza, M. Noronha-Neto, and J. Barros, “Improved joint turbo decoding and physical-layer network coding,” in *Proc. IEEE Information Theory Workshop (ITW)*, pp. 532-536, Sept. 2012.

[48] S. Zhang and S. C. Liew, “Physical layer network coding with multiple antennas,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Sydney, Australia, Apr. 2010.

[49] D. To, J. Choi, I. M. Lim, “Error probability analysis of bidirectional relay systems using alamouti scheme,” *IEEE Commun. Lett.*, , vol. 14, no.8, pp. 758-760, Aug. 2010.

[50] N. Xu and S. Fu, On the performance of two-way relay channels using space-time codes, *Int. J. Commun. Syst.*, vol. 24, no. 8, pp. 1002-1014, Jan., 2011

[51] D. Wübben and Y. Lang, “Generalized sum-product algorithm for joint channel decoding and physical-layer network coding in two-way relay systems,” in *Proceedings of Globecom*, Miami, USA, Dec. 2010.

[52] X. Wu, C. Zhao, and X. You, “Joint LDPC and physical-layer network coding for asynchronous bi-directional relaying,” *IEEE Jour. Select Area. Commun.*, vol. 31, no. 8, pp. 1446-1454, July 2013.

[53] T. Yang, X. Yuan, and I. B. Collings, “Reduced-dimension cooperative precoding for MIMO two-way relay channels,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4150-4160, Nov. 2012.

[54] T. Yang, X. Yuan, L. Ping, I.B. Collings and J. Yuan, “Eigen-Direction Alignment Aided Physical Layer Network Coding for MIMO Two-Way Relay Channels,” in *Proc. of the IEEE Int. Symp. on Information Theory (ISIT)*, Saint-Petersburg, Russia, pp. 2253-2257, July 2011.

[55] T. Yang, X. Yuan, L. Ping, I. B. Collings, and J. Yuan, “A new physical-layer network coding scheme with eigen-direction alignment precoding for MIMO two-way relaying,” *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 973-986, Mar. 2013.

[56] R. G. Maunder and L. Hanzo, “Iterative decoding convergence and termination of serially concatenated codes,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 216-224, Jan. 2010.

[57] R. G. Maunder and L. Hanzo, “Extrinsic Information Transfer Analysis and Design of Block-Based Intermediate Codes,” *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp.762-770, March 2011.

[58] R. G. Maunder and L. Hanzo, “Near-capacity irregular variable length coding and irregular unity rate coding,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5500-5507, Nov. 2009

[59] L. Hanzo, T. H. Liew, B. Yeap, and R. Y. S. Tee, *Turbo Coding, Turbo Equalisation and Space-Time Coding: EXIT-Chart Aided Near-Capacity Designs for Wireless Channels*, 2nd edition, Wiley, 2010.

[60] L. Hanzo, J. P. Woodard, and P. Robertson, “Turbo Decoding and Detection for Wireless Applications,” *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1178-1200, June 2007

[61] R. G. Maunder and L. Hanzo, "Genetic Algorithm Aided Design of Component Codes for Irregular Variable Length Coding," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1290-1297, May 2009.

[62] N. Bonello, S. Chen and L. Hanzo, "Low-density parity-check codes and their rateless relatives," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 3-26, 2011.

[63] R. Zhang, Y. C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699-712, June 2009.

[64] R. Annavajjala, A. Maaref, and J. Zhang, "Multiantenna analog network coding for multihop wireless networks," *Int. J. Digital Multimedia Broadcasting*, no. 368562, 2010.

[65] Z. Ding, I. Krikidis, J. Thompson, and K. K. Leung, "Physical layer network coding and precoding for the two-way relay channel in cellular systems," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 696-712, Feb. 2011.

[66] S. A. K. Tanoli, I. Khan, N. Rajatheva, F. Atachi, "Advances in relay networks: performance and capacity analysis of spaceCtime analog network coding," *Eurasip J. Wireless Commun. Netw.*, 2010.

[67] Q. F. Zhou, Y. Li, F.C.M. Lau, B. Vucetic, "Decode-and-forward two-way relaying with network coding and opportunistic relay selection," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3070-3076, Nov. 2010.

[68] Z. Yi, and I. M. Kim, "Optimum beamforming in the broadcasting phase of bidirectional cooperative communication with multiple decodeand-forward relays," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5806-5812, Dec. 2009.

[69] P. Hu, C.W. Sung, K.W. Shum, "Joint channel-network coding for the Gaussian two-way two-relay network," *Eurasip J. Wireless Commun. Netw.*, 2010.

[70] L. Song, G. Hong, B. Jiao, M. Debbah, "Joint relay selection and analog network coding using differential modulation in two-way relay channels," *IEEE Trans Vehicular Tech.*, vol. 59, no. 6, pp. 3070-3076, July. 2010.

[71] S. Kim, and J. Chun, "Network coding with MIMO pre-equalizer using modulo in two-way channel," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2008.

[72] B. Rankov, and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 379-389, 2007.

[73] M. Huang, J. Yuan, and T. Yang, "Error probability of physical-layer network coding in multiple-antenna two-way relay channel," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2012.

[74] M. Huang, and J. Yuan, "Physical-layer network coding with Alamouti scheme for the TWRC with linear decoder," in *Proc. Australian Communications Theory Workshop (AusCTW)*, 2014.

[75] M. Huang, and J. Yuan, " Error Performance of Physical-Layer Network Coding in Multiple-Antenna TWRC," *IEEE Trans Vehicular Tech.*, to appear, 2014.

[76] L. Lu, T. Wang, S.C. Liew, S. Zhang, "Implementation of physical-layer network coding," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 4734-4740, Ottawa, Canada, June 2012.

[77] Y. Hao, D. Goeckel, Z. Ding, D. Towsley, and K. K. Leung, "Achievable rates for network coding on the exchange channel," in *Proc. IEEE MILCOM*, Oct. 2007.

[78] L. Lu, L. You, Q. Yang, T. Wang, M. Zhang, S. Zhang, and S. C. Liew, "Real-time implementation of physical-layer network coding," in *Proceedings of the*

2nd Workshop on Software Radio Implementation Forum. ACM, pp. 71-76, 2013.

[79] F. Gao, R. Zhang and Y. C. Liang, “Optimal channel estimation and training design for two-way relay networks,” *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 3024-3033, Oct. 2009.

[80] B. Jiang, F. Gao, X. Gao, A. Nallanathan, “Channel estimation and training design for two-way relay networks with power allocation,” *IEEE Trans. Wireless Commun.*, vol.9, no.6, pp. 2022-2032, June 2010.

[81] F. Gao, R. Zhang, Y.C. Liang, “Channel estimation for OFDM modulated two-way relay networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4443-4455, Nov. 2009.

[82] T. Sjodin, G. Gacacin, F. Adachi, “Two-slot channel estimation for analog network coding based on OFDM in a frequency-selective fading channel,” *in Proc. 71st IEEE Vehicular Technology Conference*, May 2010.

[83] W. Yang, Y. Cai, J. Hu, W. Yang, “Channel estimation for two-way relay OFDM networks,” *Eurasip J. Wirel. Commun. Netw.*, 2010.

[84] X. Liao, L. Fan, F. Gao, “Blind channel estimation for OFDM modulated two-way relay network,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2010.

[85] H. Gacanin, F. Adachi, “The performance of network coding at the physical layer with imperfect self information removal,” *Eurasip J. Wirel. Commun. Netw.*, 2010.

[86] S. Zhang, S. C. Liew, H. Wang, “Blind known interference cancellation,” *IEEE Jour. Select Area. Commun.*, vol.31, no.8, pp.1572-1582, Aug. 2013.

[87] M. C. Valenti, D. Torrieri, T. Ferrett, "Noncoherent physical-layer network coding with FSK modulation: relay receiver design issues," IEEE Trans. Commun., vol.59, no.9, pp. 2595-2604, Sept. 2011.

[88] S. Zhang, S. C. Liew, P. P. Lam, "On the synchronization of physicallayer network coding," *in Proc. IEEE ITW*, pp. 404-408, Oct. 2006.

[89] L. Lu, S. C. Liew, S. Zhang, "Optimal decoding algorithm for asynchronous physical-layer network coding," in *Proc. IEEE International Conference on Communications (ICC)*, June 2011.

[90] L. Lu, S. C. Liew, "Asynchronous physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 819-831, Feb. 2012.

[91] J. Liu, M. Tao, Y. Xu, and X. Wang, "Superimposed XOR: a new physical layer network coding scheme for two-way relay channels," in *Proc. IEEE Globecom*, Nov. 2009.

[92] M. Chen, A. Yener, "Multiuser two-way relaying: detection and interference management strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4296-4305, Aug. 2009.

[93] D. Gündüz, A. Yener, A. Goldsmith, and H. Poor, "The multiway relay channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 51-63, Jan. 2013.

[94] Y. E. Sagduyu, D. Guo, R. Berry, "Throughput optimal control for relay-assisted wireless broadcast with network coding," in *Proc. 5th IEEE Annual Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops*, June 2008.

[95] Y. E. Sagduyu, D. Guo, R. Berry, "On the delay and throughput of digital and analog network coding for wireless broadcast," in *Proc. 42nd Annual Conference on Information Sciences and Systems (CISS)*, pp. 534-539, Mar. 2008

[96] F. Gao, T. Cui, B. Jiang, X. Gao, "On communication protocol and beamforming design for amplify-and-forward N-way relay networks," in *Proc. 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 109-112, Dec. 2009.

[97] A. U. T. Amah, and A. Klein, "Beamforming-based physical-layer network coding for non-regenerative multi-way relaying," *Eurasip J. Wirel. Commun. Netw.*, 2010.

[98] Z. Zhou, B. Vucetic, "An optimized network coding scheme in twoway relay channels with multiple relays," in *Proc. IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1717C1721, Sept. 2009.

[99] K. Lee, N. Lee, and I. Lee, "Achievable degrees of freedom on K-user Y channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1210-1219, Mar. 2012.

[100] N. Lee, J. B. Lim, and J. Chun, "Degrees of freedom of the MIMO Y channel: Signal space alignment for network coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3332-3342, July 2010.

[101] Y. Tian and A. Yener, "Degrees of freedom for the MIMO multi-way relay channel," submitted to *IEEE Trans. Inf. Theory*, Available: http://arxiv.org/abs/1308.1389.

[102] H. Xin, X. Yuan, and S. Liew, "Space-division approach for multi-pair MIMO two-way relaying: A principal-angle perspective, in *Proc. IEEE Globecom*, Atlanta, USA, Dec. 2013.

[103] X. Yuan, "MIMO multiway relaying with clustered full data exchange: Signal space alignment and degrees of freedom," to appear in *IEEE Trans. Wireless Commun.*.

[104] R. Wang and X. Yuan, “MIMO multiway relaying with pair-wise data exchange: A degrees of freedom perspective,” submitted to *IEEE Trans Sig. Process.*

[105] Z. Fang, X. Yuan, and X. Wang, “DEBIT: Distributed energy beamforming and information transfer for multiway relay channels,” submitted to *IEEE Trans. Wireless Commun.*.

[106] F. Wang, S.C. Liew, D. Guo, “Wireless MIMO switching with zero-forcing relaying,” in *Proc. 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 551-558, Sept. 2011

[107] F. Wang, S.C. Liew, D. Guo, Wireless MIMO switching with zeroforcing relaying and network-coded relaying, *IEEE J. Sel. Areas Commun.*, to appear, Available at: http://arxiv.org/abs/1104.4035v1.

[108] F. Wang, S.C. Liew, D. Guo, “Wireless MIMO switching with MMSE relaying,” in *Proc. IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 1127-1131, July 2012.

[109] F. Wang, X. Yuan, S. Liew, and D. Guo, “Wireless MIMO switching: Sum mean square error and sum rate optimization, in *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5297-5312, Sept. 2013.

[110] P. Elias, “Coding for noisy channels,” in *IRE Conv. Rec.*, vol. 3, pt. 4, pp. 37-46. Mar. 1955.

[111] R. Hamming, “Error detecting and correcting codes,” *Bell System Tech. J.*, pp. 147-160, April 1950.

[112] A. Hocquenghem, “Codes correcteurs derreurs,” *Chiffres*, vol. 2, pp. 147-156, 1959.

[113] R. S. Bose and D. K. Ray-Chaudhuri, On a class of error correcting binary group codes, *Information Control*, vol. 3, pp. 68-79, March 1960.

[114] P. Elias, "Coding for noisy channels," *IRE Conv. Rep.*, pp. 37-47, 1955.

[115] I. S. Reed and G. Solomon, "Polynomials codes over certain finite fields," *J. Soc. Indust. Appl. Math*, vol. 8, no. 2, pp. 300-304, June 1960.

[116] G. D. Forney, *Concatenated Codes.* Cambridge, MA, MIT Press, 1966.

[117] D. MacKay and R. Neal, "Near shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 32, no. 18, p. 1645, August 1996.

[118] G. D. Forney, F. R. Kschischang, R. J. McEliece, and D. A. Spielman, "Introduction to the Special Issue on Codes on Graphs and Iterative Algorithms," in *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 493-497, 2001.

[119] Q. Sun, T. Huang, and J. Yuan, "On Lattice-Partition-Based Physical-Layer Network Coding over GF(4)," *IEEE Communications Letters*, vol. 10, no. 10, pp. 1988-1991, Oct. 2013.

[120] T. Huang, T. Yang, J. Yuan, and I. Land, "Design of Irregular Repeat-Accumulate Coded Physical-Layer Network Coding for Gaussian Two-way Relay Channels," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 897-909, Mar. 2013.

[121] Q. Sun, J. Yuan, T. Huang, and W. K. Shum, "Lattice Network Codes Based on Eisenstein Integers," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2713-2725, July 2013.

[122] G. Wang, W. Xiang, J. Yuan, and T. Huang, "Outage Analysis of Non-Regenerative Analog Network Coding for Two-Way Multi-Hop Networks", *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 662-664, June 2011.

[123] T. Huang, X. Yuan, and J. Yuan, "Degrees of Freedom of Half-duplex MIMO Multi-way Relay Channel with Full Data Exchange," *IEEE GLOBECOM*, 2014, accepted.

[124] T. Huang, X. Yuan, and J. Yuan, "Half-duplex MIMO Multi-way Relay Channel with Full Data Exchange: Degrees of Freedom and Sum-rate Optimization," submitted to *IEEE Transactions on Wireless Communications*.

[125] T. Huang, J. Yuan, and Q. Sun, "Opportunistic Pair-wise Compute-and-Forward in Multi-way Relay Channels," *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013.

[126] Y. Ma, T. Huang, J. Li, J. Yuan, Z. Lin, and B. Vucetic, "Novel Nested Convolutional Lattice Codes for Multi-Way Relaying Systems over Fading Channel-s," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2671-2676, Shanghai, China, Apr. 2013.

[127] T. Huang, T. Yang, J. Yuan, and I. Land, "Convergence Analysis for Channel-coded Physical Layer Network Coding in Gaussian Two-way Relay Channels," in *Proceeding of the 8th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 849-853, Aachen, Germany, Nov. 2011.

[128] G. Wang, W. Xiang, J. Yuan, and T. Huang, "Outage Performance of Analog Network Coding in Generalized Two-Way Multi-Hop Networks," in *Proceeding of IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1988-1993, Quintana-Roo, Mexico, Mar. 2011.

[129] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proc. IEEE International Conference on Communications (ICC)*, May 1993.

[130] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. 36th Annu. Allerton Conf. Commun., Control, Computing*, pp. 201- 210, 1998.

[131] E. Arikan, “Channel Polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051-3073, July 2009.

[132] L. Shu, and D. J. Costello, *Error Control Coding: Fundamentals and Applications*, 2nd edition, Prentice-Hall, 2004.

[133] B. Vucetic, and J. Yuan, *Turbo Codes: Principles and Applications*, Kluwer Academic Publishers, May 2001.

[134] T. Yang and J. Yuan, “Performance of iterative decoding for superposition modulation-based cooperative transmission”, *IEEE Tran. on Wirless Comm.,* vol 9, no.1, pp. 51-59, Jan. 2010 .

[135] T. Yang, J. Yuan and Z. Shi, “Rate optimization for IDMA systems with iterative joint multi-user decoding”, *IEEE Tran. on Wirless Comm.,* vol 8, no.3, pp. 1148-1153, Mar. 2009.

[136] S. Benedetto and G. Montorsi, “Unveiling turbo-codes: Some results on parallel concatenated coding schemes,” IEEE Trans. Inform. Theory, vol. 42, no. 2, pp. 409-428, Mar. 1996.

[137] J. G. Proakis, *Digital Communications*, 4th Ed. McGraw-Hill, New York, 2001.

[138] Robert G. Gallager, *Information theory and reliable communications*, JOHN WILEY & SONS, 1968.

[139] T. M. Duman and M. Salehi, “New performance bounds for turbo codes”, *IEEE Tran. on Comm.,* vol 46, no.6, pp. 717-723, June 1998.

[140] D. Divsalar, H. Jin and R. J. McEliece, “Coding theorems for “Turbo-Like” codes”, Proc. of Allerton, 1998.

[141] A. Burr, *Modulation and Coding for Wireless Communications*, Prentice Hall, 2003.

[142] D. G. Daut, J. W. Modestino, L. D. Wismer, "New short constraint length convolutional code constructions for selected rational rates", *IEEE Tran. on Inform. Theory*, vol 28, no.5, pp. 794-800, Sept. 1982.

[143] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2764-2772, Nov. 2003.

[144] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727-1737, Oct. 2001.

[145] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[146] S. ten Brink, "Code doping for triggering iterative decoding convergence," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Washington, DC, USA, p. 235, July 2001.

[147] T. Yang, and J. Yuan, "Performance of iterative decoding for superposition modulation-based cooperative transmission," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 51-59, Jan. 2010.

[148] T. Cui, J. Kliewer, and T. Ho, "Communication protocols for N-way All-Cast relay networks," *IEEE Trans. Commun.*, vol. 60, no. 11, pp.3239-3251, Nov. 2012

[149] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 938-947, Apr. 2006.

[150] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670-678, Apr. 2004.

[151] S. Boyd, and L. Vandenberghe, *Convex Optimization,* Cambridge University Press, 2004.

[152] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning”, *IEEE Trans. Inf. Theory*, vol. 48, no.6, pp. 1250-1276, June 2002.

[153] H. Jin, A. Khandekar, and R. McEliece, ”Irregular repeat-accumulate codes” in *Proc. 2nd Int. Symp. Turbo Codes Related Topics*, 2000.

[154] R. M. Tanner, “A recursive approach to low complexity codes,” *IEEE Trans. Inform. Theory*, vol. 27, no.5, pp. 533-547, Sept. 1981.

[155] S. J. Johnson, *Iterative Error Correction: Turbo, Low-Density Parity-Check and Repeat-Accumulate Codes*, Cambridge University Press, 2010.

[156] S. T. Brink, “Convergence of iterative decoding,” *Electronics Letters* , vol.35, no.10, pp.806-808, 13 May 1999.

[157] L. Wei and W. Chen, “Compute-and-forward network coding design over multi-source multi-relay channels,” *IEEE Trans. Wireless Comm.*, vol. 11, no. 9, pp. 3348-3357, Sept. 2012.

[158] L. Wei and W. Chen, “Efficient compute-and-forward network codes search for two-way relay channels,” *IEEE Comm. Letters*, vol. 16, no. 8, pp. 1204-1207, Aug., 2012.

[159] S. H. Lim, Y.-H. Kim, A. El Gamal and S.-Y. Chung, “Noisy network coding,” *IEEE Trans. Inf. Theory*, vol. 57, No. 5, pp. 3132-3152, May, 2011.

[160] J. Du, M. Xiao, M. Skoglund and S. Shamai, “Short-Message Noisy Network Coding with Partial Source Cooperation,” *IEEE ITW*, Lausanne, Switzerland, Sept. 2012.

[161] B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proc. IEEE*, vol. 99, no. 3, pp. 438-460, Mar., 2011.

[162] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "A network coding approach to cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3714-3722, Oct., 2007.

[163] M. Xiao and M. Skoglund, "Multiple-user cooperative communications based on linear network coding," *IEEE Trans. Comm.*, vol. 58, no. 12, pp. 3345-3351, Dec., 2010.

[164] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, 2001.

[165] C. Feng, D. Silva, and F. R. Kschischang, "Lattice network coding via signal codes," *IEEE ISIT*, St. Petersburg, Russia, Aug., 2011.

[166] C. Feng, D. Silva, and F. R. Kschischang, "Lattice network coding over finite rings," *12th Canadian Workshop on Inf. Theory*, 2011.

[167] C. Feng, D. Silva, and F. R. Kschischang, "Blind computer-and-forward," *IEEE ISIT*, Cambridge, MA, Jul., 2012.

[168] Y. H. Gan, C. Ling and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2701-2710, July 2009.

[169] A. Sakzad, E. Viterbo, Y. Hong and J. Boutros, "On the Ergodic Rate for Compute-and-Forward," *Netcod*, Cambridge, MA, June 2012.

[170] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer-Verlag, 1999.

[171] R. A. Mollin, *Advanced Number Theory with Applications*, New York: CRC Press, 2011.

[172] J. H. Conway and R. Guy, *The Book of Numbers*, New York: Springer-Verlag, 1996.

[173] G. D. Forney, MIT lecture notes on Introduction to Lattice and Trellis Codes.

[174] K. Jarvis, *NTRU over the Eisenstein integers*, Master thesis, University of Ottawa, 2011.

[175] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," *IEEE Globecom*, Taipei, Nov., 2002.

[176] H. Napias, "A generalization of the LLL-algorithm over euclidean rings or orders," *Journal de Théorie des Nombres de Bordeaux*, vol. 8, pp. 387-396, 1996.

[177] S.-H. Hong and G. Caire, "Quantized compute and forward: a low-complexity architecture for distributed antenna systems," *IEEE ITW*, Paraty, Brazil, Oct., 2011.

[178] R. M. Gray and T. G. Stockham, "Dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 805-812, 1993.

[179] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 820-824, Nov., 1983.

[180] G. D. Forney, "Multidimensional Constellations–Part II: Voronoi Constellation-s," *IEEE J. Select. Areas Commun.*, vol. 7, no. 6, pp. 941-958, Aug., 1989.

[181] I. Dumer, "Concatenated codes and their multilevel generalizations," *Handbook of coding theory*, vol. 2, pp. 1911-1988, North Holland, 1998.

[182] K. Sakakibara and M. Kasahara, "On the minimum distance of a q-ary image of a q^m-ary cyclic code," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1631-1635, Sept. 1996.

[183] D. Gündüz, A. Yener, A. Goldsmith, and H. Poor, "The multi-way relay channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, pp. 339-343, Seoul, Korea, July 2009.

[184] L. Ong, S. Johnson, and C. Kellett, "An optimal coding strategy for the binary multi-way relay channel," *IEEE Communications Letters*, vol. 14, no. 4, pp.330-332, Apr. 2010.

[185] S. N. Islam and P. Sadeghi, "Error propagation in a multi-way relay channel," *in Proc. International Conference on Signal Processing and Communication Systems (ICSPCS)*, pp. 1-8, Honolulu, Hawaii, Dec. 2011.

[186] S. N. Islam and P. Sadeghi, and S. Durrani. "Error performance analysis of decode-and-forward and amplify-and-forward multi-way relay networks with binary phase shift keying modulation," *IET Commun.*, vol. 7, no. 15, pp. 1605-1616, Oct. 2013.

[187] S. N. Islam, and P. Sadeghi, "Joint decoding: extracting the correlation among user pairs in a multi-way relay channel," *in Proc. 23rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 54-59, Sydney, Australia, Sept. 2012.

[188] E. Tunali, K. Narayanan, J. Boutros and Y.-C. Huang, "Lattice Codes Based on Eisenstein Integers for Compute-And-Forward," *50th Annual Allerton Conference*, Urbana-Champaign, IL, Oct. 2012.

[189] T. Cui, T. Ho, and J. Kliewer, "Space-time communication protocols for n-way relay networks," in *Proc. IEEE Global Commun. Conf.*, New Orleans, LA, Nov. 2008.

[190] L. Ong, S. J. Johnson, and C. M. Kellett, "The capacity region of multiway relay channels over finite fields with full data exchange," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3016-3031, May 2011.