Polar Lattices for Next-Generation Wireless Communications with Application to Cyber Ranges

NGUYEN Hoang Long

Graduate School of Advanced Science and Technology

Japan Advanced Institute of Science and Technology

March, 2018

Master's Thesis

Polar Lattices for Next-Generation Wireless Communications with Application to Cyber Ranges

1610060 NGUYEN Hoang Long

Supervisor : Associate Professor Razvan Beuran
Main Examiner : Associate Professor Razvan Beuran
Examiners : Professor Yasuo Tan
Associate Professor Brian Kurkoski
Associate Professor Yuto Lim

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Information Science

February, 2018

Abstract

Recently, there is a lot of interest for the application of polar codes for next-generation wireless communication networks due to its advantage compared to other well-known channel codes like turbo codes and LDPC codes. One reason is they lack mathematical proof of capacity-achieving properties and partly reveal several drawbacks for very high-speed wireless communications. Polar codes, which were proposed by Arikan, are considered to be one of the significant contributions in coding theory for recent years. They are thus a potential candidate for the next generation wireless communication such as 5G cellular systems.

In this thesis, we propose a basic scheme using polar codes for 5G systems especially for small and moderate code lengths such as 64, 128, 256. For example, the simulation results show that the polar codes for short length of 128 with successive cancellation list (SCL) decoding outperforms the turbo codes at all rates of 1/3, 1/2 and 2/3. In addition, the polar codes length of 1024 with CRC-aided SCL decoding achieve even better performance than LDPC codes in equal conditions. We then analyze the polar codes performance under Rayleigh fading and propose to apply the SCL decoding and the suitable CRC sequence as the outer channel codes to overcome the negative effect of fading channels and get SNR gain. The selection of CRC bit length must be traded off with the code length because it causes the increase in bits redundancy for moderate code length. For instance, 2.5 dB gain is obtained for polar codes 1024 with rates of 0.5 under Rayleigh fading when using our proposed design.

Lattices and *lattice codes* has received the increasing attention for their applications to wireless communications, from the research community. The advantage of a lattices based system is that it is able to merge the channel coding and the modulation as one process. It is necessary to separate the differences between lattices and lattice codes. Since a *lattice* has infinite number of number of points, a *lattice code* is generated by applying a *power constraint* to an infinite lattice. This thesis *mainly concentrates* on transmitting lattices points without power constraint over AWGN channel but this is an important preceding step for the further work on lattice codes. Various decoders have been used for lattices, where the sphere decoding algorithm is determined a maximum likelihood lattice decoding approach. And the extension part of this thesis is an effort to propose the lattice syndrome decoding that is near-optimal decoding method for small dimensions and is potential to apply to the MIMO systems.

Because the polar codes demonstrates great performance, also have a nice nested property that are suitable for constructing lattices. We therefore propose a lattices constructed from polar codes called as *polar lattices* by Construction D with modified multi-level decoding for Code formula decoding, instead of subtracting the estimated binary codeword $\hat{\mathbf{c}}_i$, Construction D decoding subtracts integer vector $\hat{\mathbf{x}}_i$. We propose to choose the code rates of polar codes on each level following the capacity rule that achieve the better performance. The simulation results show that polar lattices constructed by proposed code rate selection outperforms the previous polar lattices by the Barnes-Wall (BW) rule. For example, 2.5 dB gain is obtained at 10^{-3} of WER when we apply new approach instead of choosing the BW rule for polar lattices. The analysis is conducted for polar lattices in term of unconstrained power over AWGN channel.

Acknowledgments

In this thesis, I would like to express my deepest gratitude to my supervisor, Associate Professor Razvan Beuran at Japan Advanced Institute of Science and Technology (JAIST), Japan for his great support during my master degree. He is always ready to help me to tackle difficulties from life as well as research topics. Inspired by his patience, enthusiasm, thoroughness, and professionals, I improved my research skills, enhanced knowledge and corrected several drawbacks. There is no doubt to say that this thesis would not be possible without his guidance and support.

I would say special thanks to Associate Professor Brian Kurkoski who is my minor research advisor. I am so fortunate to study under your instruction that makes me unforgettable about your insightful advice, breakthrough ideas, and motivated research works. It is hard to demonstrate all my appreciations to you.

I gratefully appreciate supports from Professor Yasuo Tan and Associate Professor Yuto Lim, as the committee members for the valuable comments and suggestions. I am also thankful to Iida and Tojo sensei from Information Science School for unconditional support when I got the confused situation at the beginning I came JAIST. I am also grateful to other JAIST friends and seniors who helped me and inspired me a lot to solve my research issues.

A sincere thank should be sent to Erick Garcia with a patient support of polar codes and polar lattices work as well as the heartfelt sharing about private life. Mohammad Hasan is also my respected senior who is ready to help me in research. Moreover, the long days at JAIST would not be great without all my fellows from V-project especially the Hanoi-team members. I will never forget these experiences between us and even I learned a lot from my friends who are younger than me.

Finally, I want to say the sincere thanks to my family for the unconditional love and continuous support. Thanks to the non-stop encouragement from my family, I am able to continue my dream of pursuing research that I hardly decided to cease working the previous job. You are always the trust place to share my issues, available to encourage me to overcome almost pressures of research and life during almost 2 years.

Contents

List of Figures

List of Tables

Abbreviation

AWGN	*Additive White Gaussian Noise*
BEC	*Binary Erasure Channel*
BSC	*Binay Symmetric Channel*
B-DMC	*Binary-input Discete Memoryless Channel*
CRC	*Cyclic redundancy check*
LLR	*Log Likelihood Ratio*
BER	*Bit Error Rate*
BLER	*Block Error Rate*
FER	*Frame Error Rate*
WER	*Word Error Rate*
BPSK	*Binary Phase Shift Keying*
LDPC	*Low-Density Parity-Check*
OFDM	*Orthogonal frequency-division multiplexing*
MIMO	*Multiple-Input Multiple-Output*
QAM	*Quadrature-Amplitude Modulation*
VNR	*Volume to Noise Ratio*
SCD	*Successive Cancellation Decoding*
SCL	*Successive Cancellation List Decoding*

Chapter 1

Introduction

1.1 Motivation and Goals

In order to facilitate the high-level quality services, very high data rates are required for future generations of wireless communication systems. Using cellular systems has recently become the most common wireless method to access data or to perform voice communication in high speed wireless communications.

5G is the fifth-generation wireless broadband technology that supports much better data rates and capacity than the current 4G systems [1]. It is proposed to operate at 6 GHz band or millimeter waves and is set to provide peak data rates of up to 10 Gbps with 100 Mbps at cell edge. Commonly proposed cases for 5G networks are eMBB (Enhanced Mobile Broadband), Massive machine type communication (mMTC) and URLLC (Ultra Reliable and Low Latency Communications). While URLLC and nMTC are latency sensitive and need high reliabilty, eMBB supports a various ranges of Internet access with high data rates to enable huge media applications and real entertainment.

Multiple-input multiple-output (MIMO) is commonly utilized in wireless communication systems because it can enhance the channel capacity and improve end-to-end system performance without expanding the frequency bandwidth [3]. Recently, large-scale MIMO or massive MIMO systems which employ a huge number of antennas is considered to be a key technology in the fifth-generation (5G) mobile networks [1].

Channel coding is a vital part of communications systems that adds patterns of redundancy into the transmission sequence so as to achieve the better performance. Such

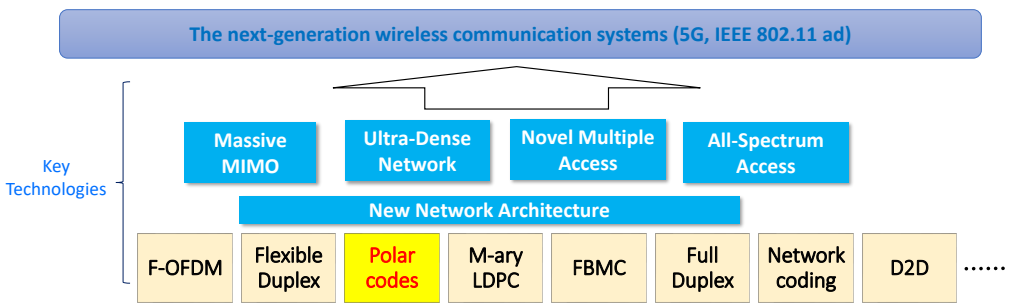

Figure 1.1: Proposed key technologies for 5G networks.

methods have played a key role in the development of the wireless communications systems. Turbo codes are a popular code which provided efficient channel coding in 3G and 4G cellular systems. Low Density Parity Check (LDPC) codes [11] are currently applied in the Wi-Fi protocol family especially the standard of IEEE 802.11ac and 802.11ad and 5G [2] because of their good performance that can achieve rates very close to channel capacity. Nevertheless, they appear to be defective in a mathematical proof of capacity-achieving properties. Polar codes [5] which are first proposed by Arikan are the capacity-achieving codes for channel of the class of binary discrete memoryless channels (B-DMCs) with reasonable complexity. Polar codes therefore are a promising candidate for high speed communication systems such as 5G and IEEE 802.11ad and is the code class which is mainly mentioned in this thesis.

Lattices are an efficient direction to solve the various problems in wireless communications due to the advantages of channel codes constructed using lattices: lattices have simple structure, their ability to obtain the capacity of the channel, and equally important, the lattices that transmitted through the medium can be decoded by different methods called as lattice decoding. In [39], the authors firstly proposed polar lattices designed by Barnes-Wall rule, but they did not mention how to investigate and select the code rate that achieves best polar lattices performance.

Recently, lattice coding has received the increasingly interests for its applications to wireless communications, from the research community. Research proposals to use lattice codes in AWGN channel, multiple-access channel (MAC), multiple-input multiple output (MIMO) broadcast channel, relay channel is increasing remarkably.

In this work, *we only consider* transmitting lattice points without power constraint over AWGN channel. Since a lattice has infinite lattice points, it is also known as the infinite constellation or coding without power constraint first mentioned in [33]. Understanding of lattices is a prerequisite to continue work on *lattice codes*. Recall that *lattices* are infinite, shaping is needed to bound the power and the *lattice codes* relates to the finite shaping region with power constraint.

As an extended discussion, physical layer security issues have received attention from researchers. Data confidentiality and security hold an increasingly important role in wireless communications. This is thus essential for cybersecurity training to help trainers better understand the attacks and master these security skills.

1.2 Contributions of Thesis

The contribution of this thesis can be summarized as follows

1. **For polar codes**

    - Propose a basic scheme for 5G system using Polar codes with performance comparisons to LDPC and Turbo codes (discussed mainly in subsection 3.6.5). In the context that there is no code among these candidates which has significant advantages over the other for all requirements for 5G. Turbo codes are currently used for 3G/4G networks but it reveals some clear drawbacks such as consuming high energy per bit, selecting only narrow range of good code rate and having a high complexity when the code length is larger. LDPC codes

demonstrate very good performance at large code length with acceptable complexity but not actually good at low code length, meanwhile polar codes with the proposed decoder can solve this problem efficiently. This study proposes to use polar codes for 5G especially for small and moderate code length (32, 64, 128, 256).

- Investigate Rayleigh fading channel scenarios with proposal to improve system performance (discussed mainly in section 3.7). The Rayleigh fading channel is considered as the worst case for wireless transmission, we therefore investigate polar codes reliability via several scenarios and compare to the ideal AWGN channel and evaluate polar codes in the fading channel.

2. **For polar lattices**

   - Introduce lattices background, propose lattices syndrome decoding method with application to MIMO systems (that is mainly mentioned in section 4.2). The sphere decoding algorithm is a maximum likelihood lattice decoding algorithm. It searches for lattice points within a fixed radius of the received signal but cost highest complexity. Our approach provides near-optimal lattice decoding with reasonable time complexity. Furthermore, the proposed MIMO detector can achieve better performance than the state-of-the-art lattice reduction-aid decoder and reach near-optimal MIMO decoder.
   - Present how to construct lattices by Construction D approach. Because polar codes have a nice nested property that are suitable for constructing lattices. We thus propose lattices constructed from polar codes called as *polar lattices* (mainly mentioned in section 4.4). The polar codes rates selection on each level impacts remarkably on the performance of polar lattices so that we can analyze the various coding design schemes which helps polar lattices achieve the best performance.

Organization of Thesis

The thesis is organized as follows

Chapter 2 provides fundamentals of the channel model, information theory and coding theory aspects. Different channels such as AWGN and Rayleigh fading channel are also introduced that are useful for next chapters.

Chapter 3 introduces polar codes, definition and the main ideas about the channel polarization phenomenon which leads to the *polar codes* definition. Two popular decoding methods such as successive cancellation (SC) decoding and successive cancellation list (SCL) decoding are also explained. One of two main contributions of this thesis is in section 3.6 and 3.7 that propose the basic scheme for 5G system using polar code.

Chapter 4 with the purpose of introducing some fundamental aspects of lattices for the AWGN channel, how to encode and decode the lattices is described. The extension part is an endeavor to apply the small dimensional lattices to MIMO system. The main second contribution is the proposed lattices constructed by polar codes that is called as *polar lattices* by Construction D. The detail of encoder and decoder are also presented.

Chapter 5 gives the discussion issues and following researches need to be extended in

the future. Also, the relation to the cybersecurity training will be mentioned as an open topic.

Chapter 6 summarizes all the important conclusions, outlines the main contributions of this thesis.

Chapter 2

Channel Coding Preliminaries

2.1 Information Theory and Coding Theory Aspects

Shannon's work [4] inspired the communication researchers to work deeply and extensively on coding theory and the past decades have seen a surge of research activities in this area. Two central topics are proposed and solved in Shannons paper: the problem of data compression, also known as source coding, and the problem of error correction, also known as channel coding. This thesis is focused on the channel coding problem.

To make the transmission of message more safe and reliable over a noisy medium, redundancy is always added to the data before transmission. At the receiver side, the redundancy efficiently helps the decoder to reconstruct the original data sequence in the presence of noise and interference. Adding redundancy is called channel coding. The set of coded words generated by the channel encoder are known as channel codes. We always call channel codes as error control codes or error correcting codes. Coding theory plays a vital role in the modern digital technology and the good ideas from the coding theory have a significant impact on practical applications. In term of performance improvement, the tendency is to design codes which have large minimum Hamming distances and then send them to modulation block so as to derive correspondingly large Euclidean distances.

The discrete memoryless channels (DMC) is considered as the simple class of channels. Symmetric B-DMCs described in the definition below is an important class of channels studied in information theory. B-DMCs consists of two common examples such as binary symmetric channels (BSC) and binary erasure channels (BEC). Then, the definition of Mutual Information and Bhattacharyya paramter is given as:

Definition 1. *(B-DMC) A binary discrete memoryless channel is a B-DMC* $W : \{0,1\} \to \mathcal{Y}$ *with the additional property that there exists a permutation over the outputs of the channel* $\pi : \mathcal{Y} \to \mathcal{Y}$ *such that* $\pi = \pi - 1$ *and* $W(y|0) = W(\pi(y)|1)$.

Definition 2. (Mutual Information). *The mutual information between the input and output of a B-DMC* W *is defined as*

$$I(W) = I(X;Y) = \sum_{y\in\mathcal{X}} \sum_{x\in\mathcal{X}} \frac{1}{2} W(y|x) \log \frac{W(y|x)}{\frac{1}{2}W(y|0) + \frac{1}{2}W(y|1)} \tag{2.1}$$

Definition 3. Bhattacharyya Parameter *The Bhattacharyya parameter are defined by*

$$Z(W) = \sum_{y\in\mathcal{Y}} \sqrt{W(y|0)W(y|1)} \tag{2.2}$$

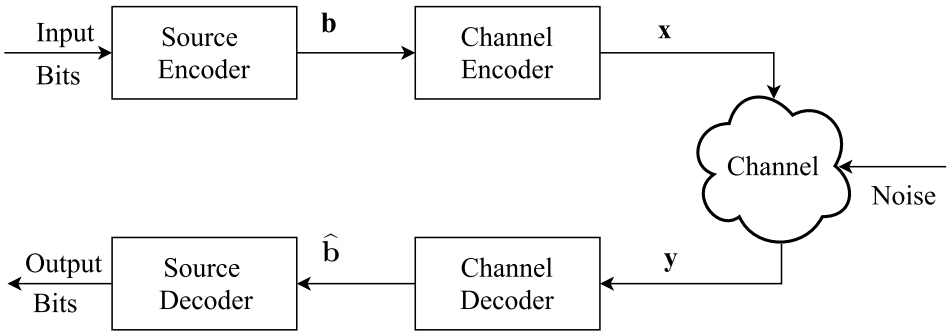

Figure 2.1: Basic communication system model.

where the the error probability when using the channel W is upper bounded by Bhattacharyya parameter.

Here, it is observed that $I(W)$ represents a measure of rate of W and $Z(W)$ denotes a measure of reliability of W. The relation between Bhattacharayya parameter the symmetric capacity is expressed by: $Z(W) = 1 - I(W)$

- **Basic notations**

In order to represent the symbos in this thesis consistently, we employ various notations shown as follows.

Denote the W as a binary input discrete memoryless channel with input alphabet X and output alphabet Y. In this thesis, both $\mathcal{X}$ and $\mathcal{Y}$ are binary sets, i.e., $\mathcal{X} = \{0, 1\}$. We use capital letters to denote random variables (RVs), such as X, Y , and normal cases to denote their realizations, such as x, y. We use $W\left(x|y\right), x \in \mathcal{X}, y \in \mathcal{Y}$ to denote the transition probabilities of a *B-DMC* W.

Also, the u_i is utilized to denote the row vector of information bits $(u_1, \ldots u_N)$. We use $\mathrm{BEC}\left(\epsilon\right)$ as abbreviation to denote the binary erasure channel with erasure probability ϵ, $\mathrm{BSC}(p)$ to denote the binary symmetric channel with crossover probability p.

In order to achieve the Shannon limit, it is necessary to use the large block lengths in practical schemes. That results in the increase of complexity of system. Therefore, in the practical applications, the low space and computational complexity coding should be first considered.

Fig. 2.1 illustrates the general model of wireless communications systems employing the source coding and channel coding block. During the transmission process, the noise is generally modeled with some probability distribution. For example, it is assumed to be Gaussian distribution that is added to the transmitted messages.

2.2 Wireless Communication over Fading Channels

The bit error probability P_b denoted as BER is a good performance measure to evaluate the systems. The BER performance system in a slow flat fading channel can be evaluated by the following integral [8]

$$P_b = \int_0^\infty P_{b,AWGN}(\gamma)\, f(\gamma) d\gamma, \tag{2.3}$$

where $P_{b,AWGN}(\gamma)$ denote the probability of error of a particular modulation scheme under the AWGN channel at a specific signal-to-noise ratio $\gamma = h^2 \frac{E_b}{N_0}$. While h denotes the channel gain which is random variable, $\frac{E_b}{N_0}$ is considered by ratio between bit energy and noise power density via AWGN scheme. Deriving from random variable h, the h^2 is also the random variable that represents the instantaneous power of the fading channel, and $f(\gamma)$ denotes the probability density function of γ on the fading channel.

2.2.1 AWGN Channel

For the sake of separation in the time domain, we specify two kinds of AWGN channels such as discrete-time (DT) and continuous-time (CT) Gaussian channel as follows:

- **Discrete-time (DT) AWGN channel**

Normally, the noise in a wireless channel is modeled as additive white Gaussian noise (AWGN):

$$y_i = x_i + n_i \tag{2.4}$$

where the noise n_i is a white Gaussian random process with mean zero and variance σ^2, $n_i \backsim N(0, \sigma^2)$. If a block code is employed subject to a power constraint $\sum_{i=1}^{N} x_i^2(m) \leq P$, $1 \leq m \leq M$, then the channel capacity is expressed as

$$C = \frac{1}{2} \log_2 \left(1 + \frac{P}{\sigma^2}\right) \text{ bits.} \tag{2.5}$$

- **Continuous-time (CT) AWGN channel**

In case of continuous-time domain, the output of waveform channel can be given by

$$y(t) = x(t) + w(t) \tag{2.6}$$

where *x(t)* is the channel input and *w(t)* is white Gaussian noise with power spectral density $N_0/2$. If signaling over the CT-AWGN channel is restricted to waveforms $x(t)$ that are time-limited to $[0,\ T]$, band-limited to $[\,W\ ,\ W\ ]$, and power-limited to P, i.e., $\int_0^T x^2(t)dt \leq PT$, then the capacity is given by

$$C = W \log_2 \left(1 + \frac{P}{N_0 W}\right) \text{bits/sec.} \tag{2.7}$$

(Signal-to-Noise Ratio and Bit/Symbol Energy): We define the received SNR as the ratio of the received signal power *Pr* to the power of the noise for transmitted signal *x(t)*. Considering the continuous-time (CT) AWGN channel $y(t) = x(t) + n(t)$, as the noise *n(t)* has uniform power spectral density (PSD) $N_0/2$, the total noise power in the bandwidth $2B$ is $N = N_0/2 \times 2B = N_0 B$. Hence the received SNR is determined as

$$SNR = \frac{P_r}{N_0 B} \tag{2.8}$$

Usually, we can use the signal energy per bit E_b (or another case per symbol, E_s) to express the SNR as

$$SNR = \frac{P_r}{N_0 B} = \frac{E_s}{N_0 B T_s} = \frac{E_b}{N_0 B T_b} \tag{2.9}$$

where T_s denotes the symbol duration and T_b is the bit duration (for binary modulation $T_s = T_b$ and $E_s = E_b$). For pulse shaping, the T_s is considered as $T_s = 1/B$. In order to simplify the notation, the SNR $= E_b/N_0$ is usually use for binary modulation scheme.

- **Error Probability for BPSK**

From [8], we get the energy depended on signal amplitude in the scheme of constellation for BPSK is given by $s_0 = \sqrt{E_b}$ and $s_1 = -\sqrt{E_b}$. The bit error probability can be calculated as

$$P_{b,BPSK,AWGN} = Q\left(\frac{2\sqrt{E_b}}{\sqrt{2N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{2.10}$$

where $Q(.)$ is the error probability function described as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy.$$

The equation 2.10 can be rewritten by

$$P_{b,BPSK,AWGN} = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{2.11}$$

where the erfc is the complementary error function that has relation to the Q function as

$$Q(x) = \frac{1}{2}\ \text{erfc}\left(\frac{x}{\sqrt{2}}\right). \tag{2.12}$$

Figure 2.2 describes the bit error probability of scheme with uncoded BPSK modulation over AWGN channel provided by equation 2.10. We can realize that there is a large gap to Shannon limit to obtain the target BER $= 10^{-5}$. As a result, it shows the disadvantage of transmission without the coding channel. With the presence of channel coding, it helps to reduce the gap coding gain at the same target BER. This is actually necessary for the practical wireless communication due to the limitation of bandwidth.

2.2.2 Rayleigh Fading Channels

In the Rayleigh fading, the channel model is assumed that the channel which will vary randomly follows the Rayleigh distribution. Where the received signal amplitude with the summation of two uncorrelated Gaussian randome variables is given as

$$p(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, r \geq 0. \tag{2.13}$$

- **Outage Probability**

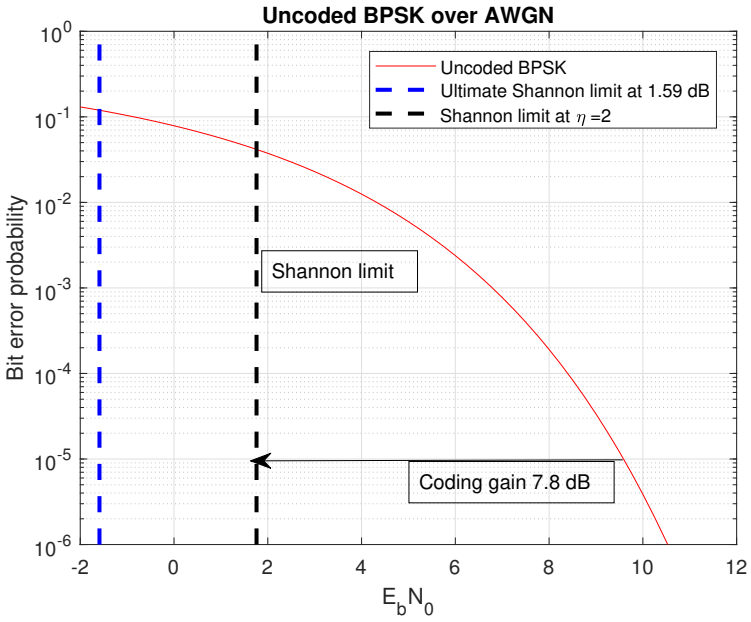

Figure 2.2: Bit error probability of Uncoded BPSK over AWGN channel.

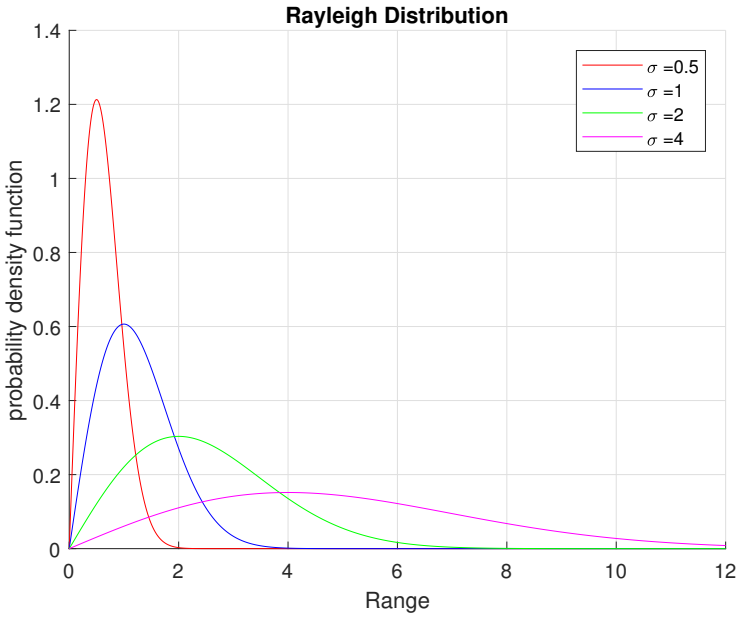

Figure 2.3: Probability Density Function of Rayleigh distribution.

The outage probability can be defined as

$$P_{out} = p(\gamma_s < \gamma_0) = \int_0^{\gamma_0} p_{\gamma_s}(\gamma) d\gamma \tag{2.14}$$

where γ_0 demonstrates the minimum SNR required for reasonable performance. The outage probability can be achieved as

$$P_{out} = \int_0^{\gamma_0} \frac{1}{\bar{\gamma_s}} e^{-\gamma_s/\bar{\gamma_s}} d\gamma_s = 1 - e^{-\gamma_s/\bar{\gamma_s}} \tag{2.15}$$

- **Average Error Probability**

The average error probability is calculated by integrating the error probability in AWGN over the fading scheme as

$$\bar{P_s} = \int_0^{\infty} P_s(\gamma) p_{\gamma_s}(\gamma) d\gamma, \tag{2.16}$$

where $P_s(\gamma)$ is the probability of symbol error in AWGN with SNR γ. Continuing some further transformation steps [8],we derive the average error probability for BPSK in Rayleigh fading as follows

$$\bar{P_b} \approx \frac{1}{4\bar{\gamma_b}} \tag{2.17}$$

Chapter 3

Polar Codes

This chapter introduces polar codes, definition and the main ideas about the channel polarization phenomenon which leads to the *polar codes* definition. Two popular decoding methods such as successive cancellation (SC) decoding and successive cancellation list (SCL) decoding are also explained. Latter part is the proposed scheme for 5G using the polar codes with small and moderate code lengths.

3.1 Channel Transformation

3.1.1 Basic Channel Transformation (Code Length N =2)

In the conventional transmission scheme, the basis for code length $N = 2$ is shown as figure 3.1. The chain rule on the joint mutual information between input and output is given as

$$\begin{aligned} I(U_1^2; Y_1^2) &= I(U_1; Y_1^2) + I(U_2; Y_1^2 | U_1) \\ &= I(U_1; Y_1) + I(U_2; Y_2) \\ &= I(W) + I(W) = 2I(W) \end{aligned} \tag{3.1}$$

Therefore, it can be seen that the conditional mutual information $I(U_1^N | Y_1^N)$ is a summation of the the mutual information of each channel W and each bit u_i is transmitted over the channel .

Now, we consider another scheme called the polarized channel transform that is different from the traditional transmission mentioned above. In this scheme, the input bits U_1^N

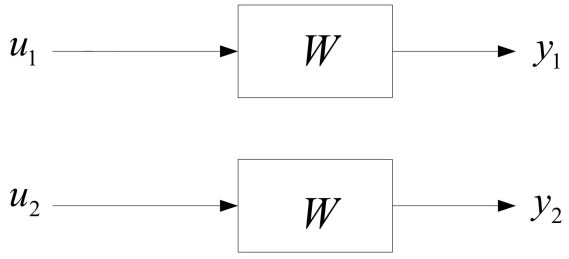

Figure 3.1: The conventional transmission scheme, length N=2

which contain the information are firstly encoded into X_1^N and then are transmitted via independent channels. Let $W : \mathcal{X} \to \mathcal{Y}$ denote a B-DMC with input alphabet $\mathcal{X} = \{0, 1\}$. X_1^2 is the input to two independent uses of W and Y_1^2 is the output. There is a linear transform between X_1^2 and $U_1^2 : X_1^2 = U_1^2 G_2$, $G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. The first step combines two independent copies of W and obtains a new channel $W_2 : \mathcal{X}^2 \to \mathcal{Y}^2$.

The transition probabilities describing the channel W_2 between U_1^2 and Y_1^2 is defined as

$$W_2(y_1^2|x_1^2) = \prod_{i=1}^{2} W(y_i|x_i) = W(y_1|u_1 \oplus u_2)W(y_2|u_2). \tag{3.2}$$

The process of generating a vector channel $W_N : \mathcal{X}^N \to \mathcal{Y}^N$ is obtained by combining the copies of W is considered channel combining [5]. Apply the chain rule of mutual information, we derive the $I(U_1^2; Y_1^2)$ as

$$I(U_1^2; Y_1^2) = I(U_1; Y_1^2) + I(U_2; Y_1^2|U_1) = I(U_1; Y_1^2) + I(U_2; Y_1^2, U_1). \tag{3.3}$$

In [5], the new terminology which corresponds to the mutual information $I(U_i; Y_1^N, U_1^i)$ as "subchannel" was first proposed. From this idea, the equation 3.2 can be evolved up to two new "subchannels" $W_2^{(1)} : \mathcal{X} \to \mathcal{Y}$ and $W_2^{(2)} : \mathcal{X} \to \mathcal{X} \times \mathcal{Y}$ with transition probabilities:

$$W_2^{(1)}\left(y_1^2|u_1^2\right) = \frac{1}{2}\sum_{u_2} W_2\left(y_1^2|u_1^2\right) = \frac{1}{2}\sum_{u_2} W\left(y_1|u_1 \oplus u_2\right) W(y_2|u_2) \tag{3.4}$$

$$W_2^{(2)}\left(y_1^2, u_1|u_2\right) = \frac{1}{2} W(y_1|u_1 \oplus u_2)W(y_2|u_2) \tag{3.5}$$

Therefore, the term $I(U_1; Y_1^2)$ is interpreted as the mutual information of subchannel $W_2^{(1)}$ which let U_1 and Y_1^2 be the input and output, considers U_2 as noise. Similarly, we can interpret $I(U_2; Y_1^2, U_1)$ as the mutual information of subchannel $W_2^{(2)}$ with input U_2 and output Y_1^2. Here, U_1 has been decoded and is known at the decoder. This phase of splitting the vector channel W_2 to series of subchannels $W_2^{(i)}, i = 1, 2$ is called channel splitting [5].

Arikan [5] proved that there is a transformation of the rate and reliability of the new subchannels $W_2^{(1)}$ and $W_2^{(2)}$ obtained from the basic transform in as shown in figure 3.2.

3.1.2 Recursive Channel Transformation (Code Length of any power of 2)

The previous subsection described the process of creating the 2 new channels $W_2^{(1)}$ and $W_2^{(2)}$. We continue to study how to generalize the number of subchannels $\{W_N^{(i)}, 1 \le i \le N\}$ for N that is any power of 2, $N = 2^n$.

The first level ($n = 1$) of the recursion combines two independent copies of W_1 as shown in Fig. 3.2 and obtains the channel $W_2 : \mathcal{X}^2 \to \mathcal{Y}^2$ with the transition probabilities

$$W_2(y_1, y_2|u_1, u_2) = W(y_1|u_1 \oplus u_2)W(y_2|u_2) \tag{3.6}$$

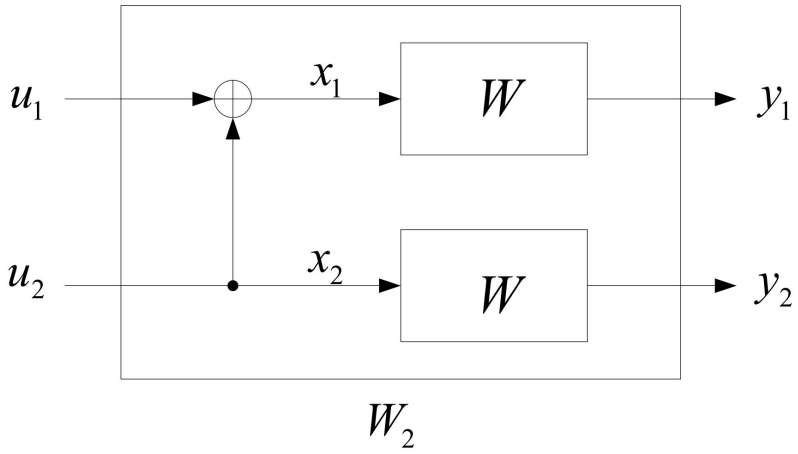

Figure 3.2: The channel W_2.

Fig. 3.3 illustrates the next level of recursion where two independent copies of W_2 are combined to create the channel $W_4 : \mathcal{X}_4 \to \mathcal{Y}_4$ with transition probabilities $W_4(y_1^4|u_1^4) = W_2(y_1^2|u_1 \oplus u_2, u_3 \oplus u_4)W_2(y_3^4|u_2, u_4)$.

In Fig. 3.3, R_4 is the permutation operation that maps an input (s_1, s_2, s_3, s_4) to $v_1^4 = (s_1, s_3, s_2, s_4)$. The mapping $u_1^4 \to x_1^4$ from the input of W_4 to the input of W_4 can be written as $x_1^4 = u_1^4 G_4$ with $G_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Thus, we have the relation $W_4(y_1^4|u_1^4) = W_4(y_1^4|u_1^4 G_4)$ between the transition probabilities of W_4 and those of W^4

In order to generalize this form, Fig. 3.4 show the general form of recursion where two independent copies of $W_{N/2}$ are combined to produce the channel W_N. The input vector u_1^N to W_N is first transformed into s_1^N so that $s_{2i1} = u_{2i-1} \oplus u_{2i}$ and $s_{2i} = u_{2i}$ for $1 \leq i \leq N/2$. In this model, R_N operates as permutation, also known as the reverse shuffle operation, and acts on its input s_1^N to form $v_1^N = (s_1, s_3, \ldots, s_{N-1}, s_2, s_4, ..., s_N)$, which becomes the input to the two copies of $W_{N/2}$ as shown in the figure.

The relationship of x_1^N and u_1^N is also demonstrated by $x_1^N = u_1^N G_N$. We call G_N the generator matrix of size N. Then, the transition probabilities of the two channels W_N and W^N are given by

$$W_N(y_1^N|u_1^N) = W^N(y_1^N|u_1^N G_N) \tag{3.7}$$

for all $y_1^N \in \mathcal{Y}^N, u_1^N \in \mathcal{X}^N$. The next part will show that G_N equals $B_N F^{\oplus n}$ for any $N = 2n$, $n \geq 0$, where B_N is a permutation matrix known as bit-reversal and $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

Lemma 1. ***(Rate and Reliability of $W_N^{(i)}$) Proof in [5].*** *Let W be a B-DMC and $W_N^{(i)}$ is defined before. For any $N = 2^n, n \geq 0, 1 \leq i \leq N$,*

$$\begin{aligned} I\left(W_{2N}^{(2i-1)}\right) &\leq I\left(W_N^{(i)}\right) \leq I\left(W_{2N}^{(2i)}\right) \\ I\left(W_{2N}^{(2i)}\right) + I\left(W_{2N}^{(2i)}\right) &= 2I\left(W_N^{(i)}\right) \end{aligned} \tag{3.8}$$

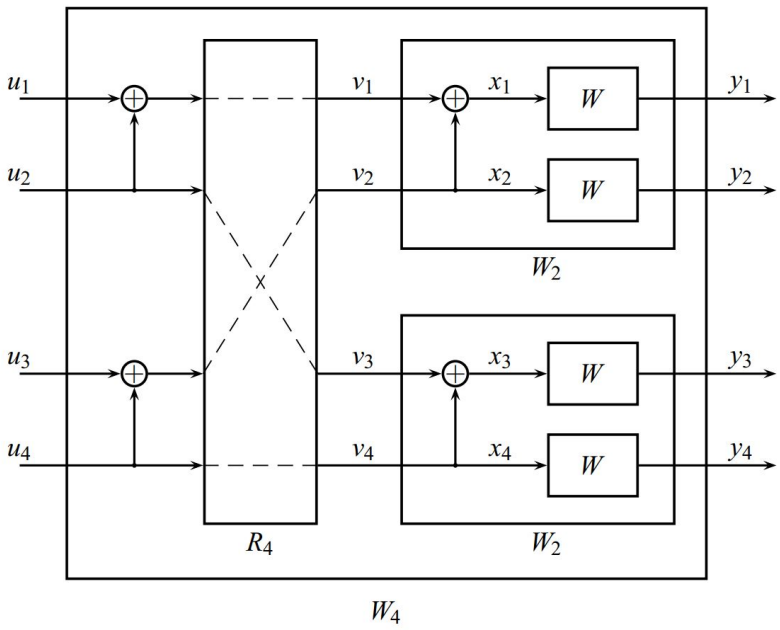

Figure 3.3: The channel W_4 and its relation to W_2 and W.

and

$$\begin{aligned} Z\left(W_{2N}^{(2i-1)}\right) &\leq 2Z\left(W_N^{(i)}\right) - Z\left(W_N^{(i)}\right)^2 \\ Z\left(W_{2N}^{(2i)}\right) &= Z\left(W_N^{(i)}\right)^2 \\ Z\left(W_{2N}^{(2i)}\right) + Z\left(W_{2N}^{(2i)}\right) &\leq 2Z\left(W_N^{(i)}\right) \end{aligned} \tag{3.9}$$

Furthermore, similar to $N = 2$, there is a special case when W is a BEC(ϵ). For BEC, the Bhattacharyya parameter of each subchannel $W_N^{(i)}$ is equal to the erasure probability of $W_N^{(i)}$ and the formula of calculating $Z(W_N^{(i)})$ is

$$\begin{aligned} Z\left(W_N^{(2i-1)}\right) &= 2Z\left(W_{N/2}^{(i)}\right) - Z\left(W_{N/2}^{(i)}\right)^2 \\ Z\left(W_N^{(2i)}\right) &= Z\left(W_{N/2}^{(i)}\right)^2 \end{aligned} \tag{3.10}$$

Example 3.1 Given the BEC(0.4) channel for $N = 8$. Consider the value of Bhattacharyya value and channel transform under BEC channel with erasure probability $\epsilon = 0.4$ shown in figure 3.5

There are three stages of the recursive transformation of $Z(W_N^{(i)})$. At each stage, the value of the Bhattacharyya parameter is computed following the rule in equation 3.10.By numerical expression at stage 1, the $Z\left(W_2^{(1)}\right) = 2 \times 0.4 - 0.4^2 = 0.64$ and $Z\left(W_2^{(2)}\right) = 0.4^2 = 0.16$. Following stages are performed with the similar rule until the stage 3.

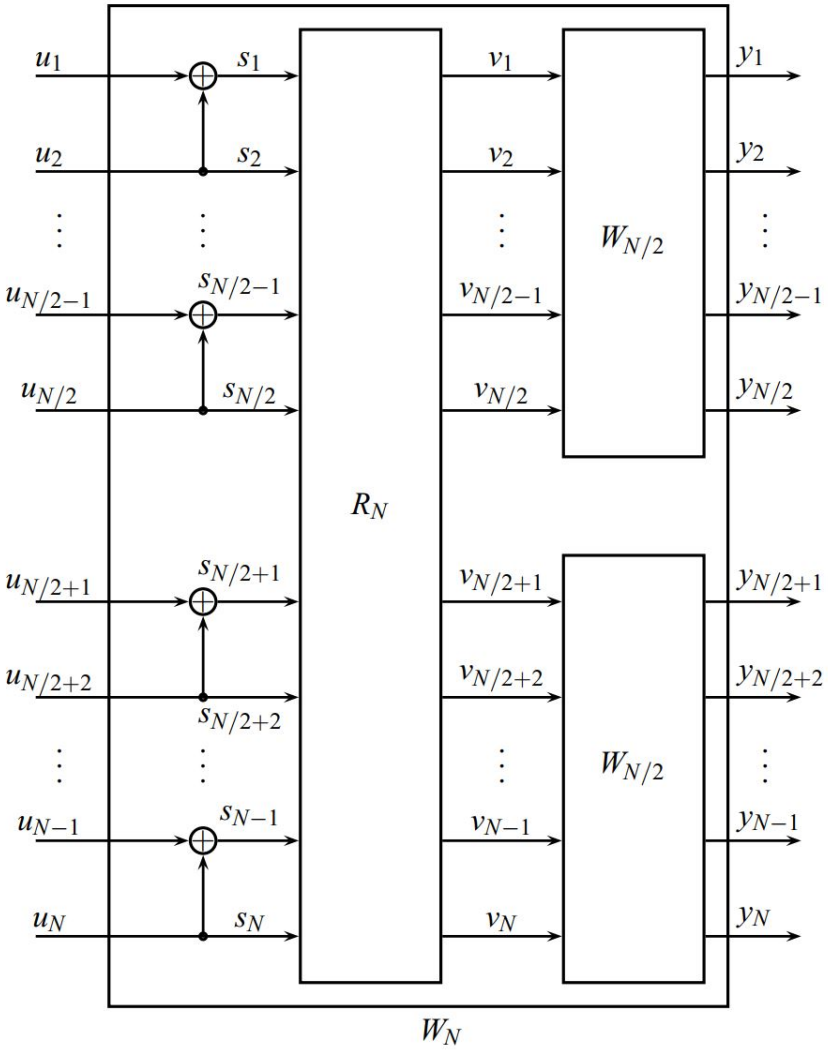

Figure 3.4: Recursive construction of W_N from two copies of $W_{N/2}$.

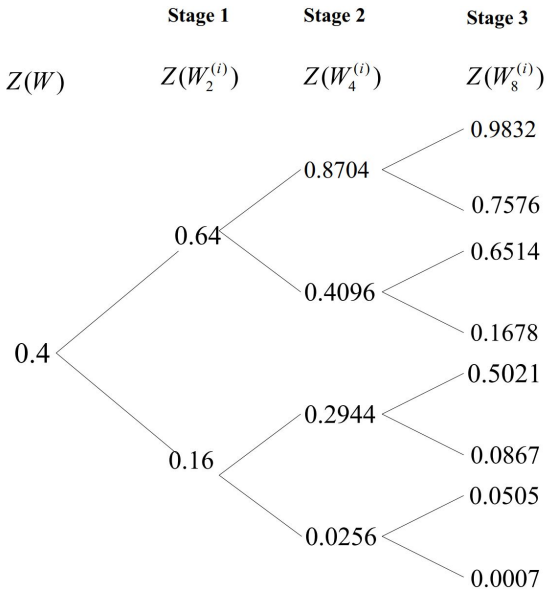

Figure 3.5: The evolution of Bhattacharyya parameter of code length $N = 8$.

3.2 Channel Polarization

Channel polarization is a transformation by which one manufactures out of N independent copies of a given B-DMC W a second set of N channels $\{W_N^{(i)} : 1 \leq i \leq N\}$ such that, as N becomes large, the symmetric capacity terms $\{I(W_N^{(i)})\}$ tend towards 0 or 1 for all but a vanishing fraction of indices i. Channel polarization contains two operations, the *channel combination phase* and the *channel splitting phase*.

Combination Two independent copies of $W_{\frac{N}{2}}$ are combined to produce the channel W_N, which is represented by

$$W_N : X^N \to Y^N,$$

exemplified on Figure 3.7, where N can be any power of two, $N = 2^n$, $n \geq 0$. The recursion begins at the 0-th level ($n = 0$) with only one copy of W and we set $W_1 = W$.

Splitting W_N is split back into a set of N binary-input coordinate channels, represented by

$$W_N^{(i)} : X^N \to Y^N \times X^{i-1}, 1 \leq i \leq N, \tag{3.11}$$

defined by the transition probabilities

$$W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) = \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(y_1^N|u_1^N) \tag{3.12}$$

where (y_1^N, u_1^{i-1})denotes the output of $W_N^{(i)}$ and u_i is input,

$$W_N\left(y_1^N|u_1^N\right) = \prod_{i=1}^{N} W(y_i|x_i),$$

and $x_1^N = u_1^N G_N$.

It can be realized that, when the value of N reaches large enough, the channels are then polarized and so the symmetric capacity $I(W)$ which is the highest rate at reliable communication is close to either 0 or 1.

Figure 3.6 also describes the change of symmetric capacity when the code length increases significantly. The proportion of subchannels with capacity 0 or 1 are dominant and the histograms are also illustrate these subchannels. The BEC channel capacity [7] is $1-P_e$. In this case, the erasure probability is set as 0.4 and when N reaches $2^{20} = 1048576$, the proportion of subchannels with capacity is very close to the capacity 0.6. We can easily realize that these histograms indicate the trend that the they approach the channel capacity in theory when N is infinity. Therefore, the polarization phenomenon is also explained visually.

Definition 4. *Frozen bits are the indices for which symmetric capacity $I(W_N^{(i)})$ are closer to 0. Frozen bits are the set of $N - K$ elements.*

In general, by applying the channel combining and channel splitting operation, N identical channel W turn into the set of $W_N^{(i)}$ with different properties. The phenomenon definition of polarization is given as below.

Mastering the polarization effect is the key step to construct the polar codes that is mentioned in the following section.

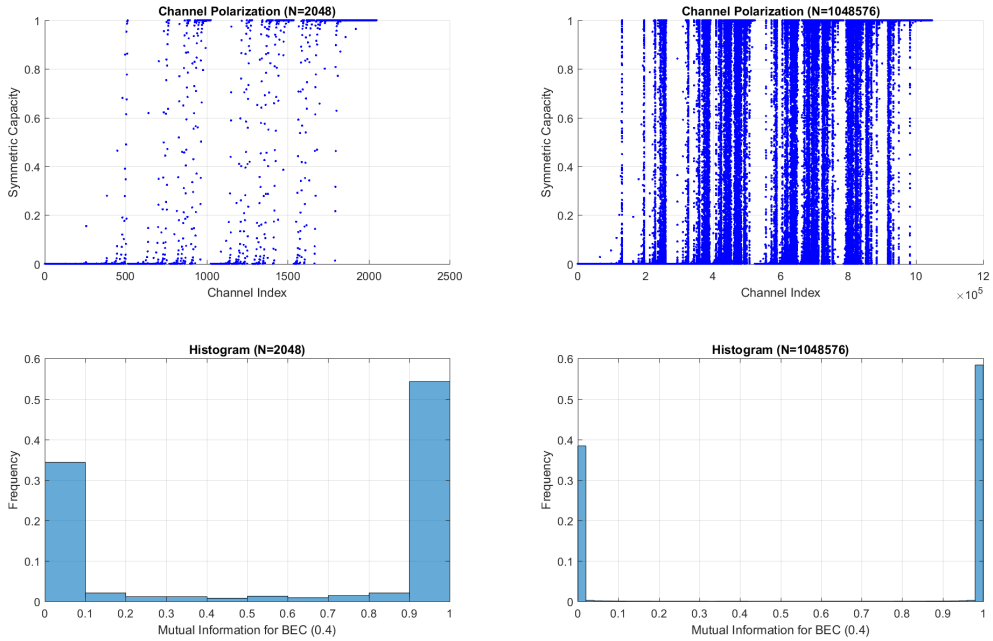

Figure 3.6: Basic model of construction and transmission of polar codes.

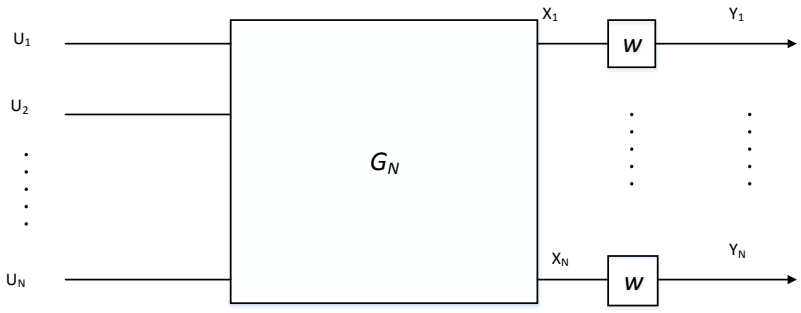

Figure 3.7: Basic model of construction and transmission of polar codes.

3.3 Definition of Polar Codes

The idea of polar codes is based on a method, called channel polarization. Channel polarization refers to the phenomenon that after implementing a linear transform to the channel inputs, the effective channels seen by some of the bits are better than the original channel W and others get worse. In other words, these channels are polarized. Interestingly, as the code length N increases, these effective channels tend towards either a perfect channel (with capacity 1) or a completely noisy one (with capacity 0). If N approaches infinity, the fraction of perfect channels approaches the channel capacity of W.

Definition 5. (Polar Codes) [5]. *For a given B-DMC, W, a code $P(N, K, \mathcal{A}, u_{\mathcal{A}^c})$ is a polar code for W if the selection of the elements of information set $\mathcal{A}$ follows a specified rule. $\mathcal{A}$ with K elements is chosen as a subset of $\{1, 2, ..., N\}$ such that $Z(X_N^{(i)}) \leq Z(X_N^{(j)}), i \in \mathcal{A}, j \in \mathcal{A}^c$.*

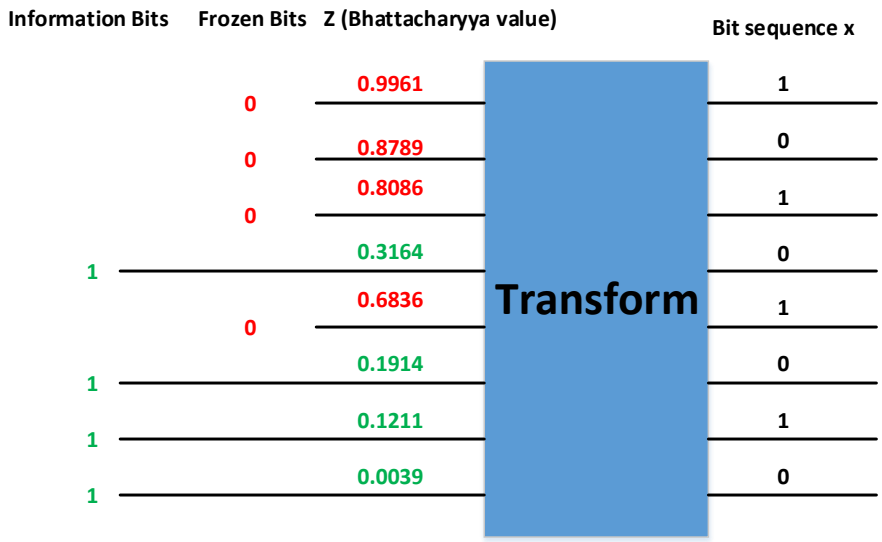

Figure 3.8: Polar codes with code length 8, code rate 4/8 under the BEC (0.5). The Bhattacharyya parameters are computed. The frozen bits are denoted by the "0s"

Example 3.2 For a BEC(0.5), the polar code with code length N = 8, code rate 4/8 is designed as shown in Figure 3.8. After calculating the Bhattacharyya parameter, we choose the indices with the small $Z(W_N^{(i)})$ as the elements of information set. Therefore, information set $\mathcal{A} = \{4, 6, 7, 8\}$. Other positions with larger $Z(W_N^{(i)})$ are used to put the auxiliary bits. The transform block can be considered as the matrix G_N with the output $x_i^8 = u_i^8 G_8$.

3.4 Polar Encoding

For a $Polar(N, K, \mathcal{A})$ polar code or simplified denote $P(N, K)$, in this section we present how to transform the information vector $\mathbf{u}$ of length K into vector $\mathbf{x}$ of length N. That is, the code rate will be $R = K/N$.

To construct a polar encoder, K of information bits are selected first and then put into these inputs, while the remaining $N - K$ inputs are frozen. The frozen bits are normally set to 0, and it is a fact both encoder and the decoder may have know the value of frozen bits in advance. An (8,4) polar encoder with frozen bits u_1, u_2, u_3, and u_5 is demonstrated by Fig. 3.8

shows an (8; 4) polar encoder in which the frozen bits are u_1, u_2, u_3, and u_5. The rest of bits sequence consisting of $\{4; 6; 7; 8\}$ is denoted as $\mathcal{A}$ and is called the information set.

- Choosing $\mathcal{A}$: We know that the set $\mathcal{A}$ should be carefully selected to get the good codes. The method for choosing $\mathcal{A}$ is that we imagine decoding all N inputs of G_N with no frozen bits, and determine the probability of decoding error for each input. These probabilities depend on the channel W. We optimize the polar code for W by choosing $\mathcal{A}$ as the set of inputs with the lowest error probabilities.

3.4.1 Encoding with $F^{\otimes}$

The generator matrix of G_N is presented in the previous section. We should make the detail expression of process of creating the generator matrix G^P for polar codes given by Plotkin construction:

$$G_N^P = B_N F^{\otimes n}, \tag{3.13}$$

where G_N^P is used to specify the block length N, $F^{\otimes n}$ is the Plotkin matrix, $F = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$, $N = 2^n$, $\otimes$ is the Kronecker product and B is called the permutation matrix that has relation with the permutation matrix:

$$B_N = R_N(I_2 \otimes B_{\frac{N}{2}}) \tag{3.14}$$

where I_2 is the 2×2 identity matrix and R_N is the reverse shuffle permutation matrix which was mentioned in Fig. 3.4. G_N^P can be expanded as

$$G_N^P = R_N(F \otimes I_{\frac{N}{2}}) \cdot (I_2 \otimes G_{\frac{N}{2}}^P), \tag{3.15}$$

where the I is the identity matrix. The detail of mathematical proof is given in [5]

3.5 Relation to the Reed-Muller Codes

At the first glance, the polar codes seems to be somewhat similar to Reed-Muller (RM) codes. This part shows the relations as well as the differences between two codes. Several numerical results have been given in [6] to shed light the performance comparison.

Let $G_N(N = 2^n)$ denote the generator matrix of RM(N, K) where N and K denote the block lengths and information length respectively.

The RM codes have several construction methods. One of those methods is making use of generator matrix that is quite similar to that of polar codes. The generator matrix is based on a submatrix of $F^{\otimes n}$ that is obtained by the choosing the rows of $F^{\otimes n}$ with Hamming weights (number of $1s$ in that row) that are as large as possible.

In contrast, the generator matrix of polar codes Polar(N,K) is selected according to channel polarization using the Bhattacharyya parameter or symmetric capacity. The difference in the selection of rows is the main reason why polar codes perform better than Reed-Muller codes.

For example, we construct the RM (8,5) code from the

$$F^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \tag{3.16}$$

and select 5 of its heaviest rows to obtain

$$G_{RM}(8,5) = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{3.17}$$

Meanwhile, the polar codes generation from the Bhattacharyya parmeter in 3.10 produces the values $z_8 = (0.996, 0.684, 0.809, 0.121, 0.879, 0.191, 0.316, 0.004)$, which gives $\pi_8 = (8, 4, 6, 7, 2, 3, 5, 1)$, and the generator matrix is uniquely determined as

$$G_P(8,5) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{3.18}$$

3.6 Polar Decoding

Decoding part is used to recover the ruined received signal to the original codeword. For polar codes, there exist various techniques to realize the decoding process. The first decoding method is successive cancellation (SC) decoding [5]. This approach is believed to get the symmetric capacity for polar codes with both encoding and decoding complexities $O(N\log N)$, however, for the moderate code length, the error performance of polar code is under the expectation which is lower than emerging channel codes like turbo codes or LDPC codes.The successive cancellation list (SCL) decoding which is modified from SC decoding is proposed to approach the (Maximum Likelihood) ML decoding performance with reasonable complexity [13]. Moreover, various decoding schemes such as belief propagation [14], improved stack successive cancellation decoder [15], [16], sequential successive cancellation decoder [17], [25] have also been proposed. Especially, the decoding aided by a cyclic redundancy check (CRC) for the SCL decoder named as CRC-aided SCL decoding offers comparable performance or even superior to low-density parity-check codes (LDPC) and Turbo codes [13].

This section concentrates on describing the two popular decoding schemes of successive cancellation (SC) and successive cancellation list decoding (SCL).

3.6.1 Successive Cancellation Decoding

As can be seen by its name, a SC decoder means the bits are decoded in order from u_1 to u_N. As stated before, the mutual information $I(u_i; y_1^N, u_1^{i-1})$ or Bhattacharyya parameter $Z(u_i; y_1^N, u_1^{i-1})$ is essential to know the channel performance, then the decoder can decode the u_i when anticipating the u_1^{i-1}. In other word, the decoder *recognizes the frozen bits values* $\{u_j, j \in \mathcal{A}^c\}$ in advance.

Consider a polar code with the parameters $(N, K, A, u_{\mathcal{A}^c})$. From u_1^N and information of y_1^N, $\mathcal{A}$ and $u_{\mathcal{A}^c}$, an estimated $\hat{u}_1^N$ is generated by the decoder. Where the likelihood ratio (LR) is given by

$$L_N^{(i)}\left(y_1^N, u_1^{i-1}\right) = \frac{W_N^{(i)}\left(y_1^N, u_1^{i1}|u_i = 0\right)}{W_N^{(i)}\left(y_1^N, u_1^{i1}|u_i = 1\right)} \tag{3.19}$$

Commonly, the $0s$ is used to represent frozen bits. Then a SC decoder take the following steps to decode the source bits. The decoding process is expressed below with i from 1 to N:

- if $i \in \mathcal{A}^c, \hat{u}_i = u_i$.
- if $i \in \mathcal{A}$, calculate the LR $L_N^{(i)}(\mathbf{y}_1^N, \hat{u}^{i-1})$ and make the decision as

$$\hat{u}_i = \begin{cases} 0 & \text{if } L_N^{(i)}(\mathbf{y}_1^N, \hat{u}^{i-1}) \geq 1 \\ 1 & \text{if } otherwise \end{cases}$$

.

In [12], LLRs values can be initialized with

$$L_N^{(i)}(\mathbf{y}, \hat{u}^{i-1}) = -2\sqrt{\frac{2E_c}{N_0}}y \tag{3.20}$$

where $E_c = \frac{K}{N}E_b$ and the design-SNR is $E_c/N_0 = \frac{K}{N}E_b/N_0$.

SC Decoding complexity: Arikan [5] also stated that traditional SC decoding, the LRs of frozen bits do not need to be calculated. It is straightforward to observe that the total number of LRs that need to be calculated is $N(1 + \log N)$. Therefore, the SC decoding complexity should be $O(N \log N)$.

3.6.2 Successive Cancellation List Decoding

As mentioned in the beginning of this chapter, the fundamental idea of the SCL decoder is that instead of retaining only one survival path in the SC decoder, the SCL decoder keeps L survival paths with higher or more probable path metrics. Lastly, the path with the best path metric is chosen as the final decoding result. List decoding has increased complexity but efficiently combats error propagation in the simple SC decoding and thus leads to superior error-rate performance.

The SCL decoder performs splitting each decoding path into two paths including in both $\hat{u}_i = 0$ and $\hat{u}_i = 1$ (if u_i is an unfrozen bit). In other words, the SCL decoder defines specific number of L best paths and it will eliminate the other ones with least probable paths. This is essential to avoid the previous hard decision errors. Finally, the decoder selects the best metric path from the list as the estimated codeword.

We assume to send a sequence u_0^{N-1} and then receive a sequence y_0^{N-1} correspondingly, the log-likelihood ratio of the estimation $\hat{u}_i$ of information bits u_i can be defined as [19]

$$L_N^{(i)}\left(y_0^{N-1}, u_0^{i-1}\right) \triangleq \ln \frac{W_N^{(i)}\left(y_0^{N-1}, u_0^{i-1}|0\right)}{W_N^{(i)}\left(y_0^{N-1}, u_0^{i-1}|1\right)} \tag{3.21}$$

where $W_N^{(i)}\left(y_0^{N-1}, u_0^{i-1}|u_i\right)$ is the transition probability of the i-th subchannel. Therefore $\hat{u}_i = \delta(L_N^{(i)})$, where $\delta(x) = \frac{1}{2}(1 - sign(x))$.

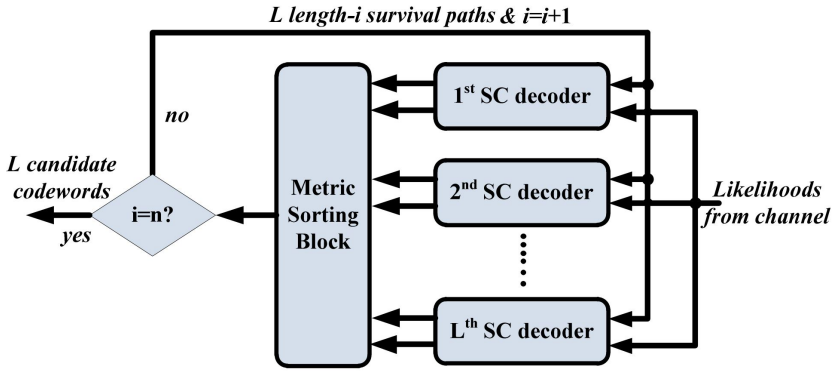

Figure 3.9: Block diagram of L-size SCL decoder.

For each path $l \in [0 : L-1]$ in decoding step i, the SCL decoder keeps track of L most likely paths with metrics of

$$\begin{aligned} PM_l^i &\triangleq -\ln\left(\mathbb{P}\left[U^i = \hat{u}_0^i[l] | Y = y_0^{N-1}\right]\right) \\ &= \sum_{j=0}^{i} \ln\left(1 + e^{-\left(1-2\hat{u}_i[l]\cdot L_N^{(j)}[l]\right)}\right) \\ &= PM_l^{(i-1)} + \ln\left(1 + e^{-\left(1-2\hat{u}_i[l]\cdot L_N^{(j)}[l]\right)}\right) \end{aligned} \tag{3.22}$$

the equation 3.22 can be well-approximated in [20] as

$$PM_l^i \approx \begin{cases} PM_l^i & \text{if } \hat{u}_i[l] = \delta\left(L_N^{(i)}[l]\right), \\ PM_l^i + \left|L_N^{(i)}[l]\right| & \text{otherwise.} \end{cases} \tag{3.23}$$

During decoding $\hat{u}_i$, the surviving paths $\hat{u}_0^{i-1}[l]$, $l \in [0 : L-1]$, are splitted into $2L$ paths with decoding $\hat{u}_i$ as 0 and 1, respectively. Afterward, the selected codewords calculated by 3.22 or 3.23 are sorted. The selected codeword with the best metric is then return at the last stage.

In order to better illustrate the SCL decoder, figure 3.9 is given to describe the general diagram of the SCL decoder. From viewpoint of each element (N, K) SC decoder, an SCL decoder with L-size (N, K) can be considered as the combination of L duplicates of (N, K) SC element decoders.

In [18], the author summarized the shortened version of SCL decoder as below. Let $S^{(i)}$ denote all the paths in i-th bit, L denotes the max number of paths. The SCL decoding algorithm is as:

1. **Initialization**: Let $S^{(0)}$= (Φ) , P $\{\Phi\}$=1, (Φ denotes empty set).

2. $\hat{u}_i$ **estimation**
   (a) For $i = 1, 2, \ldots, N$
   With all path set $S^{(i-1)}$, for $u^i = 0$ and $u^i = 1$, let

$$S^{(i)} = \{u_1^i | u_1^{i-1} \in S^{(i-1)}, \hat{u}_i \in \{0,1\}\}$$

then compute path probability

$$P\left(u_1^i\right) = P\left(u_1^{i-1}\right) P\left(u_1^i = b | u_1^{i-1} = \hat{u}_1^{i-1}\right)$$

where $b \in \{0,1\}$

$$P\left(u_1^i = b | \hat{u}_1^{i-1} = \hat{u}_1^{i-1}\right) \triangleq \begin{cases} \frac{W_N^i\left(y_1^N, u_i^{i-1} | \hat{u}_i = b\right)}{\sum_{b' \in \{0,1\}} W_N^i\left(y_1^N, \hat{u}_i^{i-1} | \hat{u}_i = b'\right)} & \text{if } i \in \mathcal{A}^c \\ 1_{b=0} & \text{if } i \in \mathcal{A} \end{cases}$$

(b) If $\left|S^{(i)}\right| \leq L$, then skip this step. Else, save the L most probability paths from $2L$ paths and drop the remaining.

(c) Judge path: when we accomplish all bits estimation, we select the max likelihood rate path from L paths.

$$\hat{u}_1^N = \arg_{\hat{\nu}_1^N \in S^{(N)}} \max \prod_{i=1}^{N} W\left(y_i | x_i = \left(\nu_1^N \cdot G_N\right)\right)$$

For a direct implementation, the time of $O(LN^2)$ and space of $O(LN \log N)$ is taken by SCL decoder. In [13], authors proposed to use the space-efficient structure and the memory sharing that help to decrease the complexity to $O(LN \log N)$ and $O(LN)$ respectively.

3.6.3 CRC-Aided Successive Cancellation List Decoding

Cyclic redundancy check (CRC) is the most commonly error detection technique in the area of information and coding theory. Since the CRCs are flexible to apply in hardware, simple to analyze in mathematical issue and very useful for detecting errors interfered by the channel noises. CRC codes therefore are the good support for improving performance of polar codes.

Revisit the construction of polar codes with k free unfrozen bits. The channel is able to deploy the concatenation scheme instead it sets all the information bit for transmission.

In this case, CRC code operates as an outer code and then the polar code is used as an inner code. Numerous studies has been proposed to use the effectively concatenating CRC with polar codes that improve the decoding performance called as CRC-Aided Successive Cancellation List Decoding [13],[15].

Considering an additional scheme for polar encoding as in figure 3.10, the CRC sequence of length m bits is added to the information sequence. This CRC code is rate $k/(k+m)$ and thus the effective rate of the polar code is $K/N = (k+m)/N$ but only k bits represent information. When we increase m to improve error correction performance, the code rate is also increased and may impacts on the decoding performance of of polar codes. This effect is more obvious when the code length is small or moderate. Hence, we should take this into consideration in using CRC with polar codes.

We denote generator polynomial as $g(x) = g_m x^m + \cdots + g_1 x + g_0$, and $\mathbf{g} = [g_m; \ldots; g_0]$ is the coefficient row vector of generator polynomial of CRC code $g(x)$. Here, the $k+m$ bits are transmitted through $k+m$ subchannels.

When the L paths are determined at the final stage, the path with the best metric is selected as the most reliable path and then is tested by the CRC sequence. If the path satisfies the CRC condition, the path is chosen. Otherwise, the CRC continues to test with the second highest likelihood path and this process is implemented iteratively until one path passes the CRC checker.

3.6.4 Theoretical Bound for Error Performance of Polar Codes

For polar codes, given any fixed $0 < \beta < 1/2$, and any code rate $R < I(W)$. The block error probability $P_e(\mathcal{A})$ is bounded by

$$P_e(\mathcal{A}) \leq 2^{-N^{\beta}} \tag{3.24}$$

where the equation 3.24 indicates that polar codes achieve Shannon capacity of W asymptotically as N tends to infinity. In the literature on polar coding, the (upper) limit value on β is usually called the *error exponent* of the coding scheme

Lemma 2. Bound on Block Error Probability [5]. *For any given $B-DMC, W$, the bound on the average block error probability $P_e(\mathcal{A})$ of polar codes with parameter $(N, K, \mathcal{A}, u_{\mathcal{A}^c})$ is given by*

$$\max_{i \in \mathcal{A}} \frac{1}{2}\left(1 - \sqrt{\left(1 - Z\left(W_N^{(i)}\right)^2\right)}\right) \leq P_e(\mathcal{A}) \leq \sum_{i \in (A)} Z\left(W_N^{(i)}\right)^2 \tag{3.25}$$

Because $Z\left(W_N^{(i)}\right)$ upper bounds the error probability of decoding bit u_i, then the block error probability is bounded by sum of $Z\left(W_N^{(i)}\right)$ of all information bits. The right side of expression 3.25 holds. As mentioned by [23] that the error probability P_e of each decoding bit u_i is lower bounded by $Pe \geq \frac{1}{2}\left(1 - \sqrt{1 - Z^2}\right)$, where Z denotes the Bhattacharyya parameter of W on which u_i is transmitted.

3.6.5 Designs of Proposed 5G Polar Codes with Aid from CRC

The basic framework of polar codes encoding and decoding for 5G is shown in Fig. 3.10. At the transmitter side, the polar codes is used as channel coding scheme. Similar to the turbo code module, function blocks like segmentation of Transmission Block (TB) into multiple Code Block (CBs) are also employed when using polar codes at transmitter. At the receiver, the system implements CB blocks and concatenating CB blocks into one TB block. The SCL decoding scheme is proposed to decode each CB block and the specific list size $L = 32$, with the aid of 16-bit CRC which is described in the above part.

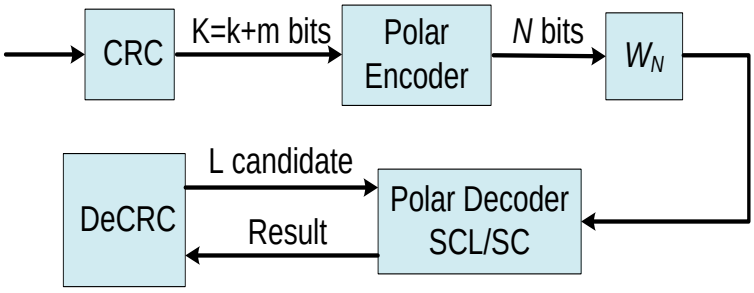

Figure 3.10: The framework of Polar codes in the 5G trial system [27]. (k = info. block length, m = crc bits length, $K = k + m$, N = encoded block length after rate-matching) .

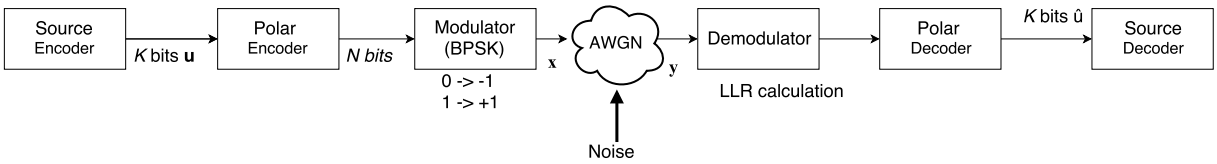

Figure 3.11: Basic schematic of Polar SC decoder via the AWGN channel.

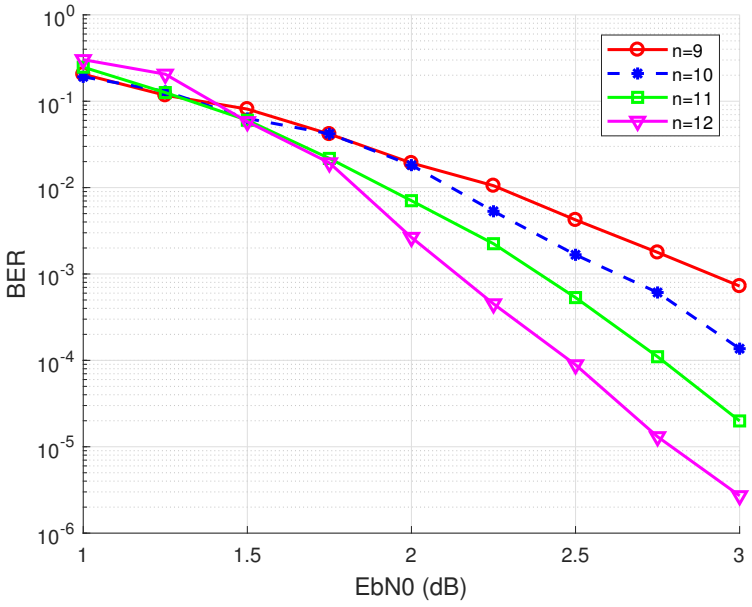

Figure 3.12: Bit error rate of Polar Successive Cancellation Decoder with different code length.

3.7 Performance Evaluation of Polar Codes

3.7.1 Polar Codes in AWGN Channel

Fig. 3.11 describes the basic schematic for polar encoder and decoder via the AWGN channel. First, bit sequence $\mathbf{u}$ is created randomly of length K. The polar encoder block encodes K bits $\mathbf{u}$ into N bits $\mathbf{x}$ and then they are modulated by the BPSK method. In general, we assume the channel is an additive White Gaussian Noise Channel (AWGN). The receiver side performs decoding the N received noisy message $\mathbf{y}$ into K estimated codeword $\hat{\mathbf{u}}$. The error is recognized when any $\hat{\mathbf{u}} \neq \mathbf{u}$.

Table 3.1 shows the general complexity comparison between these well-known channel codes including turbo, LDPC and polar codes, where N, R, M denotes the code length, code rate and number of parity bits, respectively. Furthermore, m is the memory length of component code of turbo code, d_v is the average variable degree of LDPC parity check matrix and d_c is the average check degree of parity check matrix also. The parameter I_{max} is the maximum number of iterations. As a result, table 3.2 represents the numerical complexity of these channel codes schemes for the specific numbers of code rates. 8 iterations are used for the turbo decoder and 50 iterations for LDPC decoder. The list size for polar codes is 8, it is about 10 times the complexity of SC decoder.

Fig. 3.12 illustrates the BER performance of polar codes under different code lengths for the AWGN channels. Here, SC decoding is applied for polar codes that is graphically explained by the Fig. 3.11. The polar codes with code length of $2^n = 2^9, 2^{10}, 2^{11}, 2^{12}$ with the code rate $R = 0.5$ are investigated. Increasing the code length of polar codes leads to performance improvement. However, the gap to the Shannon limit is still considerable.

The performance of polar SC decoding indicates that this approach can not be comparable to LDPC or turbo code. The Fig. 3.13 shows the simulation result of the proposed candidate for 5G scheme that uses the SCL decoding (L =32) and CRC precoding. The results were obtained by Monte-Carlo simulation 500000 trials via JAIST parallel computer system. The comparison shows that the proposed candidate achieves better performance to the LDPC codes in the fair condition.

Fig. 3.14 investigates the comparison performance of polar codes and turbo codes with small block length of 128 and various code rates such as 1/3, 1/2 and 2/3. Polar codes are decoded by the SCL decoder with list size = 8 and CRC size is 8 bits. This result shows that polar codes for short length outperforms the turbo codes in all rates. However, the significant advantages are disappeared when the block length increases to 1024 as displayed by Fig. 3.15. The polar decoder $N = 1024$ with list size of 32 and CRC of 16 bits only gets the comparable performance to CTC WiMax turbo codes length of 960. Only when increasing the list size to 1024 does polar SC list decoder obtain the significant performance gain in term of block error rate (BLER) compared to turbo code. But the trade-off paid is that the decoder complexity is much higher than other decoders.

To sum up, from these above results. Although turbo codes are being used in various applications including 3G/4G, LTE standards, it may not satisfy the performance requirement for eMBB for all code rates and block lengths because the complexity is too high for higher data rates. In addition, an error floor appears in turbo code BER. Beside from that, the modern LDPC decoders use soft decision algorithms which help to improve the decoder gain, reduce the complexity and latency. We can observe that LDPC codes perform very well at longer code length and a wide range of code rates with reasonable

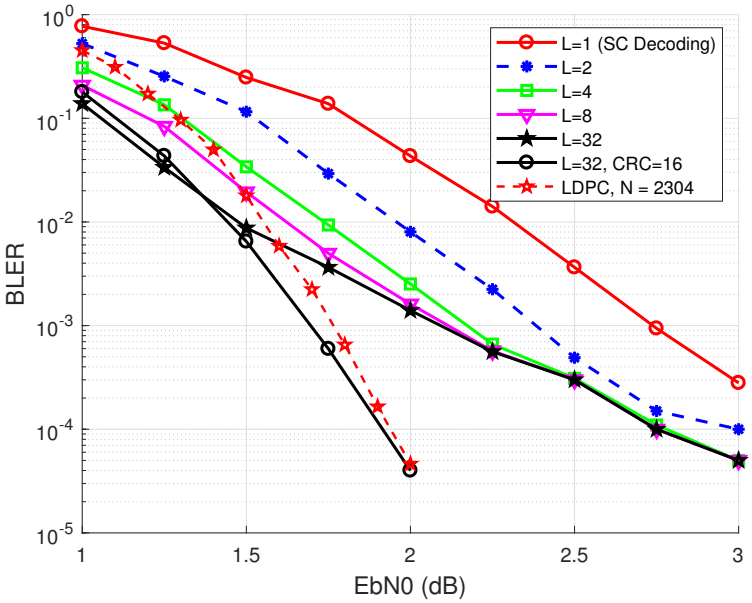

Figure 3.13: Block error rate comparison of proposed system using Polar codes, rate 1/2 with CRC aided and LDPC codes, rate 1/2 , where the LDPC simulation result is extracted from [26].

Table 3.1: Complexity comparison of coding schemes [21]

Channel Code	Addition	Multiplications
Turbo (Max-log-MAP)	$I_{max} \cdot 16 \cdot RN2^m$	$2(I_{max} - 1)K$
LDPC (min-sum)	$I_{max}(2Nd_v + 2M)$	$I_{max}(N - K)(2d_c + 3)$
LDPC (Sum-Product)	$I_{max}(2Nd_v + M(2d_c - 1))$	-
Polar (SC List)	$LN \cdot \log N + (N - M)L \log(2L)$	$LN \log N$
Polar SC	-	$N \log N$

complexity. Due to their excellent ability to achieve theoretical limits of channel capacity, LDPC codes are utilized in many advanced communication systems such as DVB-S2, 802.11n or proposed 802.11ad, etc. The polar codes have excellent performance at small and moderate code lengths compared to all other codes. The state-of-the-art polar decoder for the large code length only gets performance comparable to LDPC codes and has the higher complexity with big list size. Therefore, we strongly propose to apply the polar codes for 5G in the small and moderate code length, that is the consensus with the content in 3GPP radio access network (RAN) #87 meeting in November, 2016 [28]. Furthermore, polar+CRC codes with list decoding provides excellent performance at large code length with the large list size. The cost is very high complexity, but with the adaptive list decoding [19], [24] that allows the polar decoder adjust the list size according to its computational power, the polar codes are expected to apply in a wide range of code lengths, code rates with the reasonable decoding complexity.

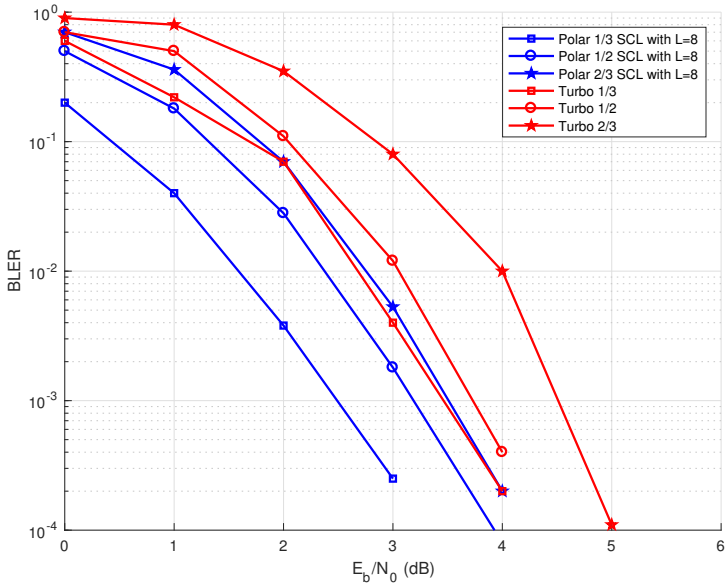

Figure 3.14: Block error rate comparison of polar codes block length of 128 and turbo codes length of simulation result is extracted from [26].

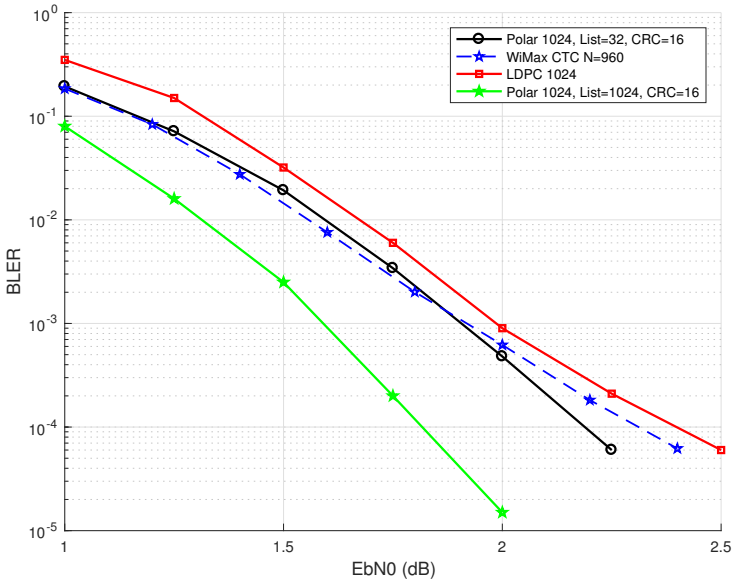

Figure 3.15: Block error rate comparison of polar codes block length of 1024 with list size = 32; 1024 and turbo codes used for WiMax CTC length of 960, iteration 8, along with LDPC 1024 (LDPC and turbo codes results is extracted from [26]). The same codes rate at $R = 1/2$ and modulation scheme BPSK.

Table 3.2: Numerical complexity of decoding schemes [22]

Block length	Coding Scheme	Complexity ($\times 10^3$)			Percentage		
		Rate 1/3	Rate 1/2	Rate 2/3	Rate 1/3	Rate 1/2	Rate 2/3
128	Turbo	65.5	98.3	131.1	100 %	100 %	100 %
	LDPC	66.0	57.2	48.5	100.7 %	58 %	37 %
	Polar SC	1.0	1.0	1.0	1.5 %	1.0 %	0.8 %
	Polar SCL	11.0	11.0	11.0	16.8 %	11.2 %	8.4 %
1024	Turbo	1048.6	1572.9	2097.9	100 %	100 %	100 %
	LDPC	1056	916	776	100.7 %	58.2 %	37.0 %
	Polar SC	24.6	24.6	24.6	2.3 %	1.6 %	1.2 %
	Polar SCL	245.5	245.5	245.5	23.4 %	15.6 %	11.7 %

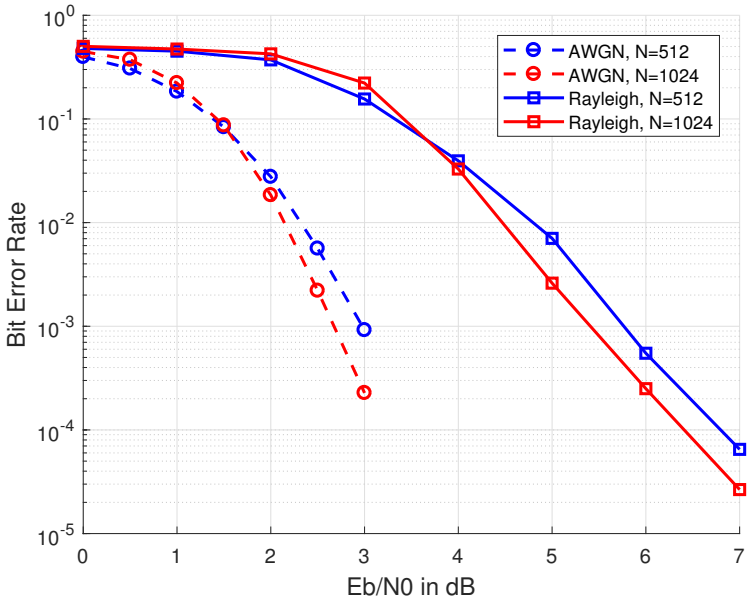

Figure 3.16: Bit error rate comparison of Polar codes via the Rayleigh fading and AWGN channel.

3.7.2 Polar Codes over Rayleigh Fading Channels

We consider the fading channel model, where the output is given by

$$y_i = h_i s_i + n_i, i = 1, \dots N, \tag{3.26}$$

where N is the frame size. In the i-th channel, $s_i = \pm 1$ is the channel input after BPSK modulation $s_i = (-1)^{x_i}$, y_i is the channel output and n_i is the zero mean independent Gaussian noise $n_i \sim \mathcal{N}(0, \sigma^2)$ and h_i is the channel gain. In this case, we assume h_i follows the Rayleigh distribution $h_i \sim \frac{h_i}{\sigma_h^2} e^{-h_i^2/2\sigma_h^2}, h \geq 0$, where σ_h is the scale parameter For simplicity, we assume that the channel state information (CSI) is known at the receiver.

In the *block fading* channel model, a transmission frame of N symbols is affected by $1 \leq B \leq N$ independent fading realizations, leading to a block of $n = N/B$ symbols which are effected by the same fading realization. Changing the value of B makes the different types of fading. Example, for B =1, we consider as the *fast fading* and the *slow fading* for $B = N$.

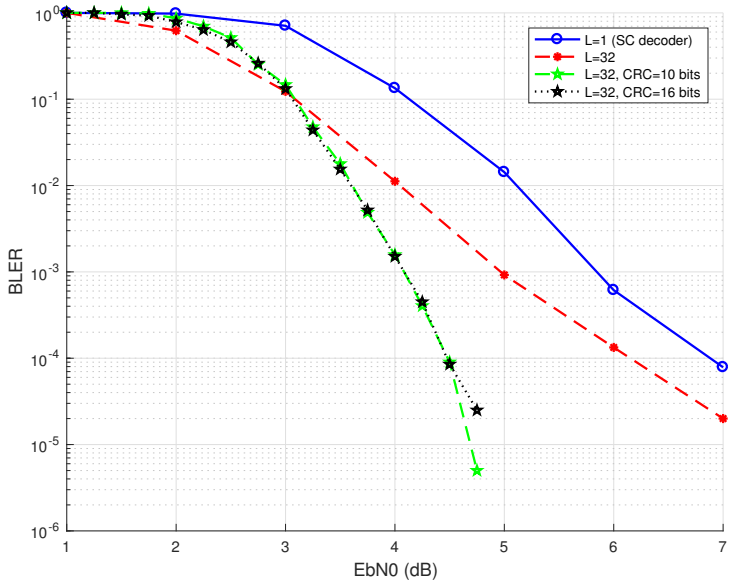

Figure 3.17: Block error rate of CRC-aided (CRC size = 10 bits and 16 bits) SCL polar decoder via the Rayleigh fading.

With the received signal scaled by the channel gain, in [29], the log-likelihood ratios of received symbols are given as $LR(y,h) = \exp\left(\frac{2yh}{\sigma^2}\right)$.

Fig. 3.16 shows the performance of polar codes via the Rayleigh fading scenario in comparison to AWGN channel in term of bit error rate. In the Rayleigh scheme, we choose the scale parameter $\sigma_h = 1/\sqrt{2}$ for the fading coefficient. We easily observe that about 2.7 dB is lost due to the negative effect of block fading condition to get the BER of 10^{-3} compared to AWGN case.

Fig. 3.17 expresses the block error rate of proposed method for the polar codes length 1024 bits at rate 0.5. The SCL decoding was implemented with aid from CRC. As mentioned at previous section, when the code length is small or moderate, if m large CRC bits sequence are padded, the effective rate of polar code is then $K/N = (k+m)/N$. This figure also indicates that adding the suitable CRC bits like 10 can achieve better performance than 16 bits CRC. Thanks to this improvement, 2.5 dB gain can be achieved in case of Rayleigh fading channel.

Chapter 4

Lattices and Polar Lattices

This chapter presents about *lattices* that differ from the *lattice codes*, the application of lattice syndrome decoding to MIMO system and then, a proposed polar lattices transformed from polar codes by Construction D will be presented.

Lattice codes are applied in many communication scenarios with continuous-output channels, such as the AWGN channel. The very important thing is to separate the difference between lattices and lattice codes. In practice, only a finite set of points of a lattice Λ can be used as a signal constellation in a communication system. Since a *lattice* has infinite lattice points, a *lattice code* is generated by applying the *power constraint* to an infinite lattice. In term of applied lattice codes, we have to consider both the (packing problem, coding gain) and (covering problem, shaping gain).

4.1 Lattices Definition

Definition An n-dimensional *lattice* Λ is an discrete additive subgroup of $\mathbb{R}^n$.

Property: a lattice Λ is a subset of $\mathbb{R}^n$ with the property that Λ forms a group under addition. That is , if $\mathbf{x}, \mathbf{y} \in \Lambda$ then $\mathbf{x}+ \mathbf{y} \in \Lambda$ as well. As a result, Λ is an infinite collection of points. Let take the example from the 2-dimentional lattices which are helpful for illustrating the concepts.

A lattice can be described by n basis vectors, $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n$, where $\mathbf{g}_i$ is a column vector representing a point in $\mathbb{R}^n$. Then a point $\mathbf{x}$ is the linear combinations of the $\mathbf{g}_i$:

$$\mathbf{x} = \mathbf{g}_1 b_1 + \mathbf{g}_2 b_2 + ... + \mathbf{g}_n b_n = \sum_{a}^{b} \mathbf{g}_i b_i \tag{4.1}$$

where the b_i are integers. The lattice point $\mathbf{x}$ can be expressed using a generator $\mathbf{x} = \mathbf{Gb}$ and $\mathbf{G}_i$ is the n-by-n matrix that consists of the n column vectors $\mathbf{g}_i$ (for convenience, we represent the vectors as the columns),

$$\begin{bmatrix} | & | & & | \\ g_1 & g_2 & . \ . \ . & g_n \\ | & | & & | \end{bmatrix} \tag{4.2}$$

Check matrix: We already know that a lattice is written as $\mathbf{x} = \mathbf{Gb}$. We define a matrix $\mathbf{H} = \mathbf{G}^{-1}$ so that $\mathbf{Hx} = \mathbf{b}$. $\mathbf{H}$ is a check matrix because if $\mathbf{Hx}$ is an integer vector,

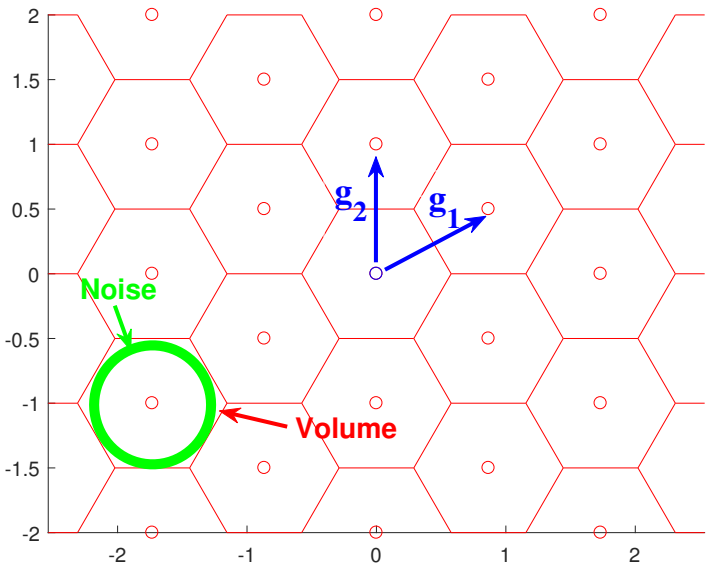

Figure 4.1: Illustration of hexagonal lattice formed by the basic vectors $\mathbf{g}_1$ and $\mathbf{g}_2$ in equation 4.3.

then $\mathbf{x}$ is a lattice points. We also have the matrix property: $\mathbf{HG}^{-1} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix.

Example: the hexagonal matrix illustrated in Fig. 4.3 has a generator matrix:

$$\mathbf{G} = \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & 1 \end{bmatrix}. \tag{4.3}$$

A lattice Λ is a discrete additive subgroup of $\mathbb{R}^n$ [32]. For an n-dimensional lattice Λ and $\mathbf{y} \in \mathbb{R}^n$, the problem of decoding, or shortest-distance quantization is to find the the element of Λ closest to $\mathbf{y}$ in the Euclidean-distance sense:

$$\mathbf{x} = \arg\max_{\mathbf{u}\in\mathbf{\Lambda}} \left\|\mathbf{u} - \mathbf{y}\right\|^{\mathbf{2}} \tag{4.4}$$

Then, some concepts that appears repeatly in this work will be introduced as follows.

The **Voronoi region** for any point $\mathbf{x}$ is the set of all points in $\mathbb{R}^n$ that are closer to $\mathbf{x}$ than to any other element of Λ. Mathematically, the volume of the Voronoi region $V(\Lambda)$ is given by

$$V(\Lambda) = |det(\mathbf{G})|. \tag{4.5}$$

Volume-to-Noise-Ratio As introduced in [33], the *unconstrained power communication* AWGN channel is a useful theoretical tool to analyze the coding aspect of lattices without considering the shaping region. Here, the definition of volume $V \triangleq V(\Lambda) = |\det \mathbf{G}|$ constrains the transmit power by the lattice density, because there is no transmit power constraint.

We could scale the lattice to be arbitrarily large, so that the lattice points are far apart and the noise ineffective. So, we constrain the system using the density of lattice points. The volume of the fundamental region $V(\Lambda)$ measures the density of points in n

dimensions. A rough approximation of the length of one side is $\sqrt[n]{V(\Lambda)}$, and power is the length (or energy) squared. So we use signal power measured as $V(\Lambda)^{2/n}$ and the noise power is σ^2. The VNR is the measure of the density of the lattice and the normalized VNR (NVNR) is given by

$$\text{VNR} = \frac{|V(\Lambda)|^{2/n}}{2\pi e\sigma^2}. \tag{4.6}$$

The Voronoi region of lattices as well as the Volume of lattices are illustrated in Λ Fig. 4.1 for the simple case of hexagonal lattices.

4.2 Lattice Syndrome Decoding Application to MIMO Systems

The use of lattice viewpoint is remarkably relevant for the MIMO detection problem. In this section, we propose a new approach called lattice syndrome decoding, and apply it to MIMO detection. We describe four algorithms based on storing error vectors in a syndrome lookup table, which is feasible for the number of antennas typically used in MIMO detection. This work is partly demonstrated in [37].

Numerous MIMO detection techniques have been introduced [30]. Minimum mean-squared error (MMSE) and zero-forcing (ZF) have low complexity, but a large performance gap with repsect to the optimal maximum likelihood (ML) detector. ML detection provides optimal performance, but the detection process of ML schemes is performed by an exhaustive search over all the possible transmitted symbol vectors, hence the complexity increases exponentially with the number of antennas.

From the lattice viewpoint, MIMO detection can be viewed as the problem of finding the closest lattice point. The sphere decoding algorithm is a maximum likelihood lattice decoding algorithm [34]. It searches for lattice points within a fixed radius of the received signal.

Inspired by syndrome decoding for finite-field codes, the lattice syndrome decoder attempts to find an estimated codeword closest to a received sequence. Syndrome decoding is based on storing error vectors in a lookup table, however some modifications are needed so lattice syndrome decoding can handle soft-input vectors. Four lattice syndrome decoding algorithms are presented, progressively solving a shortcoming of the previous one. The fourth algorithm, a tabular lattice syndrome decoding algorithm called Algorithm D, is an efficient and promising candidate as a general lattice decoding algorithm. MIMO detection, one of the main applications of lattice decoding, can be efficiently performed.

While *lattice syndromes* have appeared previously in the literature, for example in the context of coded modulation [36] and low-density lattice codes [31], this work represents the first consideration of *lattice syndrome decoding*.

Some important aspects of lattice syndrome decoding are:

- Syndrome decoding requires another algorithm, preferably an optimal one, to generate the syndrome lookup table; for lattice syndrome decoding, we use the Schnorr-Euchner algorithm.

- Numerical results show that lattice syndrome decoding has negligible performance loss with respect to optimal decoding, for the MIMO channel. However, Algorithm D is not optimal in general.

- Lattice syndrome decoding is performed using table lookup operations. While lattice syndrome decoding has some initialization complexity to generate the syndrome lookup table, operations are very efficient. Once lookup tables are generated, they can be used many times, suitable for stable MIMO channels. We can think of this as a fast implementation of the Schnorr-Euchner algorithm.

In order to emphasize the parallelism between syndrome decoding of finite-field codes and syndrome decoding of lattices, syndrome decoding of codes is reviewed. Matrices representing codes and lattices are assumed to be full rank.

4.2.1 Syndrome Decoding of Finite-Field Codes

Let $\mathbb{F}$ be a finite field of arbitrary size so $\mathbb{F}^n$ is an n-dimensional vector space. A finite code C is a a k-dimensional vector subspace of $\mathbb{F}^n$. Since C is a subspace and thus is a subgroup of $\mathbb{F}^n$, the quotient group $\mathbb{F}^n/C$ is formed. The coset of $\mathbf{a} \in \mathbb{F}^n$ is the set $\mathbf{a} + C$. There are $\mid \mathbb{F} \mid^{n-k}$ cosets. The coset leaders are a single representative element from each coset, chosen to be a coset member of lowest Hamming weight.

Let $\mathbf{H}_c$ be an $(n-k) \times n$ parity-check matrix for $\mathbf{C}$. The syndrome $\mathbf{s}$ of any sequence $\mathbf{y} \in \mathbb{F}^n$ is $\mathbf{s} = \mathbf{H}_c\mathbf{y}$; the syndrome of a codeword is a vector of zeros. If $\mathbf{e}$ is the coset leader of coset containing $\mathbf{y}$, then there is a unique $\mathbf{c} \in C$ such that $\mathbf{y} = \mathbf{c} + \mathbf{e}$. The syndrome of $\mathbf{y}$ is :

$$\mathbf{s} = \mathbf{H}_c\mathbf{y} = \mathbf{H}_c\left(\mathbf{c} + \mathbf{e}\right) = \mathbf{H}_c\mathbf{e}. \tag{4.7}$$

That is, all element of a coset have the same syndrome, that of the coset leader.

Syndrome decoding for a finte-field code dinds the estimated codeword $\hat{\mathbf{x}} \in C$ closest to a received sequence $\mathbf{y} \in \mathbb{F}^n$. It uses the syndrome of $\mathbf{y}$ to find an estimated error vector $\hat{e}$ and thus achieve the estimated codeword $\hat{\mathbf{c}}$. The estimated error is the coset leader for the syndrome, which is stored in a syndrome table ψ. Syndrome decoding has an input received sequence $\mathbf{y}$ and output $\hat{\mathbf{x}}$:

1. Compute syndrome $\mathbf{s} = \mathbf{H}_c\mathbf{y}$
2. Look up estimated error $\hat{\mathbf{e}} = \psi\left(\mathbf{s}\right)$.
3. Output nearest codeword $\hat{\mathbf{x}} = \mathbf{y} - \hat{\mathbf{e}}$

4.2.2 Lattice Syndrome Decoding (Algorithm A)

- **Cosets** The lattice shift or coset is defined as

$$\Lambda_x = \mathbf{x} + \Lambda = \left\{x + \lambda : \lambda \in \Lambda\right\}. \tag{4.8}$$

  A coset is a discrete set of points such that the difference vector between every pain of points belongs to the lattice. However, the coset itself is, in general , not a lattice, as it is not closed under addition; it does not contain the origin.

Let $\mathbb{R}$ be the set of real numbers, so that $\mathbb{R}^n$ is an n-dimensional vector space. A lattice Λ is a vector subspace of $\mathbb{R}^n$.

Further, let Λ' be a superlattice of Λ with $\Lambda' = \frac{1}{m}\Lambda$ for m=2, 3, ...; Λ' and $\frac{1}{m}\Lambda$ are used interchangeably. Since Λ is a sublattice and thus a subgroup of Λ', the quotient group Λ'/Λ is created. The coset of $\mathbf{a} \in \Lambda'$ is the set $a + \Lambda$. There are m^n cosets. The coset leaders are a single representative element from each coset, chosen to be a coset member of smallest norm.

Let G be an $n \times n$ generator matrix for Λ, and let $\mathbf{G}_c = \mathbf{G}^{-1}$ be the corresponding check matrix. Furthermore, Λ' has generator matrix $\frac{1}{m}\mathbf{G}$ and check matrix $m\mathbf{G}_c$.

Definition 6. *The lattice syndrome* $\boldsymbol{s}$ *of* $\boldsymbol{z} \in \frac{1}{m}\Lambda$ *with respect to a lattice* Λ *having check matrix* $\boldsymbol{G}_c$*, is*

$$\boldsymbol{s} = m\boldsymbol{H}\boldsymbol{z} \bmod m \tag{4.9}$$

so that $\boldsymbol{s} \in \{0, \ldots, m-1\}^n$.

The syndrome of a lattice point is a vector of integers. If $\mathbf{e}$ is the coset leader of the coset containing $\mathbf{y}$, then there is a unique $\mathbf{x} \in \Lambda$ such that $\mathbf{z} = \mathbf{x} + \mathbf{e}$ where the channel output $\mathbf{z}$ is an element of a superlattice. The syndrome of $\mathbf{z}$ is:

$$\begin{aligned} \mathbf{s} &= m\mathbf{G}_c\mathbf{z} \bmod m \\ &= m\mathbf{G}_c(\mathbf{x} + \mathbf{e}) \bmod m \\ &= m\mathbf{G}_c\mathbf{e} \bmod m \end{aligned} \tag{4.10}$$

That is, all elements of a coset have the same syndrome, that of the coset leader.

Lattice syndrome decoding finds the estimated lattice point $\hat{\mathbf{x}} \in \Lambda$ closest to a received sequence $\mathbf{z} \in \Lambda'$. It uses the syndrome of $\mathbf{z}$ to find an estimated error $\hat{\mathbf{e}}$, and thus the estimated lattice point $\hat{\mathbf{x}}$. The estimated error is the coset leader for the syndrome, which is stored in a syndrome table ϕ, found by Algorithm S1, next. Syndrome decoding has received sequence $\mathbf{z}$ as input, and estimated lattice point $\hat{\mathbf{x}}$ as output:

1. Compute syndrome $s = m\mathbf{G}_c\mathbf{z} \mod m$.
2. Look up estimated error $\hat{\mathbf{e}} = \phi(\mathbf{s})$.
3. Output estimated lattice $\hat{\mathbf{x}} = \mathbf{z} - \hat{\mathbf{e}}$.

Algorithm S1: Generation of syndrome table ϕ. The coset leaders are the codewords of a nested lattice code Λ'/Λ. Because Λ and Λ' are self-similar lattices, these codewords can easily be found. Let $Q_\Lambda(\mathbf{y})$ be the element of Λ closest to $\mathbf{y} \in \mathbb{R}^n$. For syndrome $s \in \{0, ..., m-1\}^n$ find the corresponding coset leader $\mathbf{e}$:

$$\mathbf{e} = \frac{1}{m}\mathbf{G}\mathbf{s} - Q_\Lambda\left(\frac{1}{m}\mathbf{G}\mathbf{s}\right), \tag{4.11}$$

and the syndrome table entry is thus:

$$\phi(\mathbf{s}) = \mathbf{e} \tag{4.12}$$

This lattice syndrome decoding, including generation of the syndrome table, is described in Algorithm A.

Algorithm S1 Syndrome Table Generation.

$$\phi = \text{SyndromeTable}(\mathbf{G}, m)$$

Input: Generator matrix $\mathbf{G}$ for an n-dimentional lattice. Scaling integer m.
Step For each $s \in \{0, 1 \ldots, m-1\}^n$, compute:

$$\phi(\mathbf{s}) = \mathbf{e} = \frac{1}{m}\mathbf{G}\mathbf{s} - Q_\Lambda\left(\frac{1}{m}\mathbf{G}\mathbf{s}\right).$$

Ouput Syndrome decoding table ϕ.

Algorithm A Lattice Syndrome Decoding

$$\hat{\mathbf{x}} = \text{AlgorithmA}(\phi, \mathbf{z})$$

Input: Syndrome table ϕ, from Algorithm S1. n-dimentional lattice Λ with generator matrix $\mathbf{G}$ and check matrix $\mathbf{G}_c$; scaling integer m; decoder input $\mathbf{z} \in \frac{1}{m}\Lambda$.
Step 1 Compute syndrome $\mathbf{s} = m\mathbf{G}_c\mathbf{z} \bmod m$
Step 2 Look up estimated error $\hat{\mathbf{e}} = \phi(\mathbf{s})$
Ouput nearest lattice point: $\hat{\mathbf{x}} = \mathbf{z} - \hat{\mathbf{e}}$.

4.2.3 Advanced Lattice Syndrome Decoding (Algorithm B, C, D)

- Algorithm A can be improved by using a second input to break ties (the decoder that generates the syndrome table implicitly breaks ties). We know that the algorithm has input $\mathbf{z} \in \Lambda'$. An additional input $\mathbf{y} \in \mathbb{R}^n$ is a point near $\mathbf{z}$ used in breaking ties, rather than breaking ties arbitrarily. Now for each $\mathbf{z} \in \Lambda'$ having syndrome $\mathbf{s}$, the syndrome table entry $\phi(\mathbf{s})$ is the set of all elements $\mathbf{x} \in \Lambda$ that are at the same minimum distance from $\mathbf{z}$. Given a generator matrix $\mathbf{G}$ and scaling m. In other words, Algorithm B is similar to Algorithm A, but in addition to $\mathbf{z}$, a real input $\mathbf{y} \in \mathbb{R}^n$ which is near the superlattice is available to break ties.

- **Algorithm C** Using the nesting property of lattices, syndrome decoding is applied recursively. Beginning with a fine lattice, a real input $\mathbf{y}$ is quantized to the superlattice using some suboptimal but efficient technique. Because the superlattice is fine, the error due to suboptimality may be made as small as desired. The output at one iteration step, quantized to a point in a lattice, is the input to the next iteration step.

- **Algorithm D** A tabular approach decodes in multiple superlattices, selects the best solution, then proceeds iteratively.

In the trade-off between space complexity and time complexity, Algorithms A–D obtain fast decoding (low time complexity) at the expense of a large amount of memory (high space complexity). The memory requirements for the lookup table are exponential in n, making this technique attractive for decoding known-good lattices of modest dimension, or for detection in stable MIMO channels.

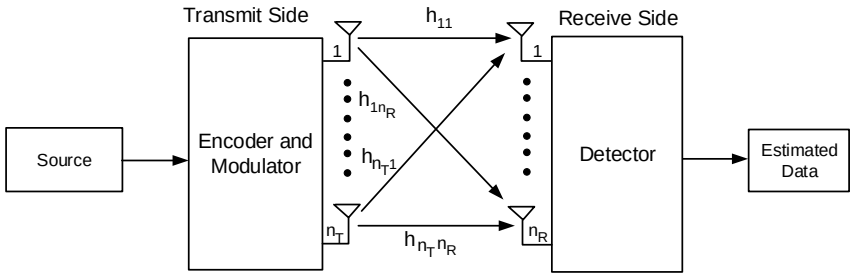

Figure 4.2: Diagram of MIMO system with n_T transmitters and n_R receivers.

4.2.4 Performance Evaluation of Lattice Syndrome Decoding and MIMO System

We consider the MIMO system in Fig. 4.13, with n_T transmitters and n_R receivers. The received signal vector $\mathbf{u}$ depends on the transmitted vector $\mathbf{v}$ as

$$\mathbf{u} = \mathbf{H}\mathbf{v} + \mathbf{w}, \tag{4.13}$$

where $\mathbf{v}$ is a vector representing the transmitted signals. $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n]$ is an $n_R \times n_T$ complex-valued matrix which contains channel coefficients distributed as $\mathcal{CN}(0, 1)$, and $\mathbf{w}$ is an additive noise vector, independent Gaussian random variables with zero mean and variance σ^2.

Let $\mathbf{H}^R$ and $\mathbf{H}^I$ denote the real and imaginary part of channel vector $\mathbf{H}$, and the same for $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{u}$. Equation (4.13) can be represented as

$$\begin{pmatrix} \mathbf{u}^R \\ \mathbf{u}^I \end{pmatrix} = \begin{pmatrix} \mathbf{H}^R & -\mathbf{H}^I \\ \mathbf{H}^I & \mathbf{H}^R \end{pmatrix} \begin{pmatrix} \mathbf{v}^R \\ \mathbf{v}^I \end{pmatrix} + \begin{pmatrix} \mathbf{w}^R \\ \mathbf{w}^I \end{pmatrix}. \tag{4.14}$$

In the conversion from complex to real, from (4.14) we can see the size of channel matrix has been increased to $2n_R \times 2n_T$.

At the receiver side, the maximum-likelihood detector (MLD), which is an optimal detector, detects the transmitted vectors

$$\hat{\mathbf{v}} = \arg\min_{v \in \mathcal{V}^m} \| \mathbf{u} - \mathbf{H}\mathbf{v} \|^2, \tag{4.15}$$

where $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{w} \in \mathbb{R}^n$, and $\mathbf{v} \in \mathcal{V}^m$ where $\mathcal{V}$ denotes the finite set of real-valued transmitted signal. The MLD computes the Euclidean distance between the received vector and all possible transmitted vectors via a given channel $\mathbf{H}$.

In case of lattices, as shown in Fig. 4.3, we consider the vector $\mathbf{x}$ transmitted via the AWGN channel with additive noise $\mathbf{w}$, then the received sequence $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ is described by $y_i = x_i + w_i$ with $w_i \sim \mathcal{CN}(0, \sigma^2)$. The vector $\mathbf{u}$ and matrix $\mathbf{H}$ play a similar role to the vector $\mathbf{y}$ and channel matrix $\mathbf{G}$, respectively, for lattices. As a result, the channel matrix $\mathbf{H}$ can be viewed as the basis for a discrete lattice.

The role of matrix of basis generators is mentioned both in the MIMO channel matrix and in lattices; in order to avoid confusion, we denote this matrix by the same symbol $\mathbf{H}$ for both cases. In case of the generator matrix $\mathbf{H}$ of lattices, all elements are independent random variables distributed as $\mathcal{CN}(0, 1)$. We assume that the channel state information (CSI) is known at the receiver.

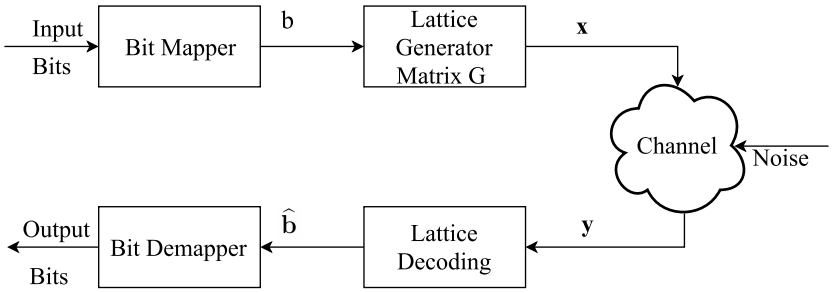

Figure 4.3: Basic diagram of system employing lattice with generator matrix **G**.

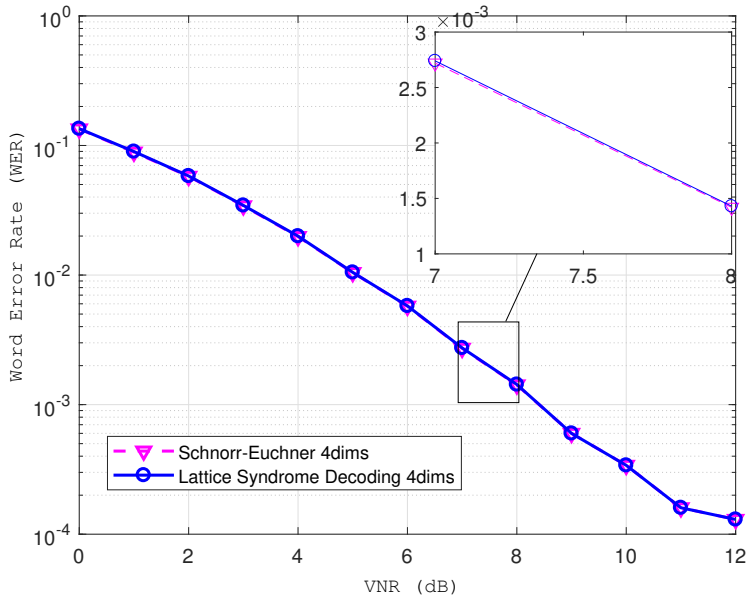

Figure 4.4: Word error rate of lattice decoders versus VNR in dB

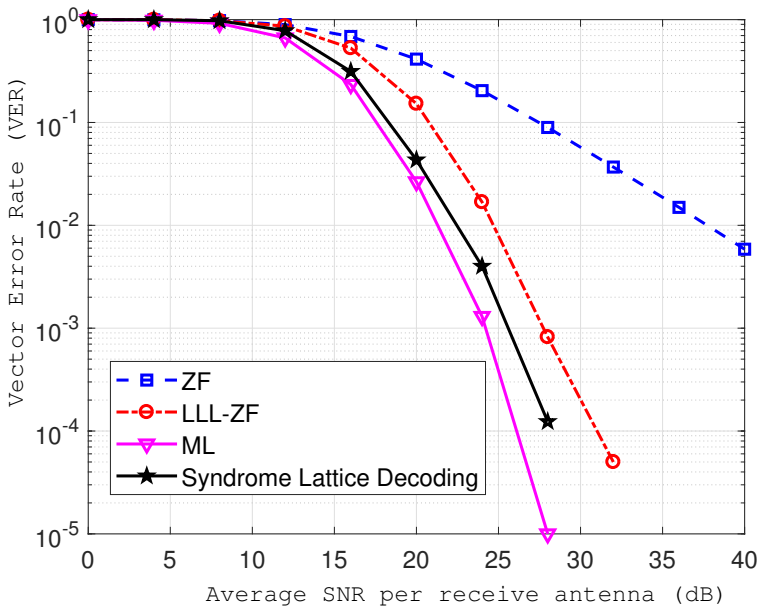

Figure 4.5: The comparison of vector-error-rate (VER) of different detectors in 2×2 MIMO system with 16-QAM

Fig. 4.4 illustrates the performance analysis of two lattice decoders using Schnorr-Euchner (S-E) algorithm and new proposed lattice syndrome decoder by the word-error-rate versus the VNR in dB. The implementation of a closest point search algorithm mainly based on the S-E strategy presented by Agrell, et al. in [35], which is regarded as the optimal search method in lattice decoding. Numerically, this figure also plots the lattice syndrome decoding when the generator matrix is chosen randomly with size of 4×4. It is also shown that the great advantage of employing the lattice syndrome decoding as an alternative to the optimal lattice decoding.

The comparison between the different MIMO detection methods including (Zero-Forcing) ZF, (Maximum Likelihood Detector) MLD, (Lenstra-Lenstra-Lovász) LLL-ZF detector, and lattice decoding using new proposed lattice syndrome decoding is shown in Fig. 4.5 in terms of VER versus the average signal-to-noise ratio per receive antenna. The MIMO system with 16-QAM input symbols are transmitted through 4×4 antennas without channel coding or space-time coding. As mentioned, the lattice-reduction-aided detector using LLL-ZF can achieve the considerable improvement compared with the linear detector like ZF whose poor performance is due to the noise enhancement. For the sake of improvement, the reliability of lattice syndrome decoder are also exposed. Since the 4×4 complex channel matrix can be transformed to a lattice of 8 dimensions, and all other conditions the same as the MIMO channel, the comparison is fair. The lattice syndrome decoder outperforms the LLL-ZF decoder, for example, we obtain a gain of 2 dB at a vector-error rate (VER) of 10^{-3}. Furthermore, its curve is close to the curve of ML detector.

4.3 Lattice Construction

There are various techniques that construct lattices from finite-field codes. Lattice Construction maps symbols from a finite field code to lattice points. Since lattice points are real numbers, but the code symbols are not real numbers, we should be careful on how to

map.

There exists several constructions including Constructions A, B, C, D, D' and E. The Construction A was considered as a simple version one that convert a linear binary code to the Euclidean space. We can use $\mathbf{x} \mod 2 = (x_1 \mod 2, ..., x_n \mod 2)$ to denote a modulo-2 reduction of each of the components of $\mathbf{x} \in \mathbb{R}^N$.

Construction B was proposed by Conway and Sloane in 1999 as a way to make a connection between Reed Muller codes and Barnes-Wall lattices. However, Construction D is more general, and is only introduced later when describing the Barnes-Wall lattices. Construction C is also formed from binary codes, but in general forms a sphere packing but not a lattice. In the cases where Construction C forms a lattice, it coincides with Construction D.

Construction D is a generalization of Construction A. While Construction A uses a single code $\mathcal{C}_1$, Construction D uses a sequence of a nested binary codes: $\mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq ... \subseteq \mathcal{C}_{a-1}$. Both Construction D and D' form lattices from multiple binary codes but Construction D uses the codes generator matrix and Construction D' uses the code parity-check matrix.

Due to the huge recent interest in the Construction D, this research thus mainly concentrates on the Construction D with application to the lattices from polar codes.

$*$ Construction D

Presented by Barnes and Sloane [38], Construction D is generated by a set of nested binary linear codes $\mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq ... \subseteq \mathcal{C}_a$ with parameters $[N, K_i, d_i]$ for each binary linear code $\mathcal{C}_i$. The minimum distance is constrained by $d_i \geq \frac{4^i}{\gamma}$, where $\gamma = 1$ or 2, for $i = 1, ..., a$.

The *minimum distance* of a lattice is the minimum Euclidean distance between any pair of lattice points, namely,

$$d_{\min}(\Lambda) = \min_{x \neq y} \{d(\mathbf{x}, \mathbf{y}) | \mathbf{x}, \mathbf{y} \in \Lambda\}, \tag{4.16}$$

where $d(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|^2$, and $\|\|$ is the Euclidean norm.

Definition 7. *(Nested Binary Linear Codes)Let $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$ be a basis for $\mathbb{F}_2^n$. Let $a \geq 1$, for $\mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq ... \subseteq \mathcal{C}_a = \mathbb{F}_2^n$ are nested linear codes if $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{k_i}$ span $\mathcal{C}_i$, for $i = 0, 1, \ldots, a-1$.*

Nesting means that for code $\mathcal{C}_0$, generator vectors $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{k_0}$ are included in those for code $\mathcal{C}_1$ with generator vectors $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{k_1}$. This nesting holds for any code which is the subcode of another code. The n-by-n matrix consisting of the basis vectors as columns $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$ is denoted $\tilde{\mathbf{G}}$, written as:

$$\tilde{\mathbf{G}} = \begin{bmatrix} | & | & & | & & | & & | \\ \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_{k_0} & \cdots & \mathbf{g}_{k_1} & \cdots & \mathbf{g}_n \\ | & | & & | & & | & & | \end{bmatrix} \tag{4.17}$$

The n-by-k_i generator matrix for code $\mathcal{C}_i$ is $\tilde{\mathbf{G}}_i$ for $i = 0, 1, \ldots, a$, which consists of spell 1 to k_i of $\tilde{\mathbf{G}}$

Definition 8. ***(Construction D)*** *Let $\mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq \ldots \subseteq \mathcal{C}_a = \mathbb{F}_2^n$ be nested binary linear codes with generator matrices $\tilde{\boldsymbol{G}}_0, \ldots, \tilde{\boldsymbol{G}}_{a-1}, \tilde{\boldsymbol{G}}$, respectively. Then a Construction D lattice consists of all vectors of the form*

$$\boldsymbol{x} = \sum_{i=0}^{a-1} 2^i \tilde{\boldsymbol{G}}_i \cdot \boldsymbol{u}_i + 2^a \tilde{\boldsymbol{G}} \cdot \boldsymbol{z} \tag{4.18}$$

where $\boldsymbol{z} \in \mathbb{Z}^n$ and $\boldsymbol{u}_i = (u_{0,1}, \ldots, u_{0,k_0})^t$ for $i = 0, 1, \ldots, a-1$ are binary vectors. The binary matrices $\tilde{\boldsymbol{G}}_i$ are taken as real-valued.

∗ Construction D Generator Matrix

The generator matrix for a lattice is not unique. However for Construction D lattices with a specific basis $\tilde{\mathbf{G}}$ and $k_0, k_1, ..., k_{a-1}$, the *Construction D generator matrix* is the specific lattice generator matrix $\mathbf{G}$ given by

$$\mathbf{G} = \tilde{\mathbf{G}} \cdot \mathbf{D}^{-1} \tag{4.19}$$

where $\mathbf{D}$ is a diagonal matrix with diagonal entries d_{ii}

$$d_{ii} = 2^{-k} \text{ for } r_{k-1} \leq i \leq r_k \tag{4.20}$$

with with $k = \{0, 1, \ldots, a\}$

In Definition 8, a lattice point $\mathbf{x}$ is found by selecting integers $\mathbf{z}$ and binary vectors $\mathbf{u}_0$ to $\mathbf{u}_{a-1}$. A lattice point is found by selecting integers $\mathbf{b} \in \mathbb{Z}^n$ so that $\mathbf{x} = \mathbf{G} \cdot \mathbf{b}$. Considering the example of $a = 3, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and $\mathbf{z}$ are related to the integers $\mathbf{b}$ by:

$$\begin{aligned} b_i &= u_{0,i} + 2u_{1,i} + 4u_{2,i} + 8z_i \text{ for } 1 \leq i \leq k_0 \\ b_i &= u_{1,i} + 2u_{2,i} + 4z_i \text{ for } k_0 < i \leq k_1 \\ b_i &= u_{2,i} + 2z_i \text{ for } k_1 \leq i \leq k_2 \\ b_i &= z_i \text{ for } k_2 \leq i \leq n \end{aligned} \tag{4.21}$$

∗ Properties of Construction D Lattices

For Construction D lattices, the volume is given by

$$V(\Lambda) = 2^{an - \sum_{i=0}^{a-1} k_i}, \tag{4.22}$$

The following lemma relates the minimum distance of the binary codes to the lattice squared minimum distance $d_{\min}^2$. If $\gamma = 1$, then the binary codes have minimum distance 4, 16, 64, ... If $\gamma = 2$, then the binary codes have minimum distance 2, 8, 32, ...

Lemma 4.1 Let code $\mathcal{C}_i$ have minimum distance $d_i \geq 4^{a-i}/\gamma$ for $i = \{0, 1, \ldots, a-1\}$, where $\gamma = \{1, 2\}$. Then the Construction D lattice has squared minimum distance

$$d_{\min}^2 \geq 4^a/\gamma \tag{4.23}$$

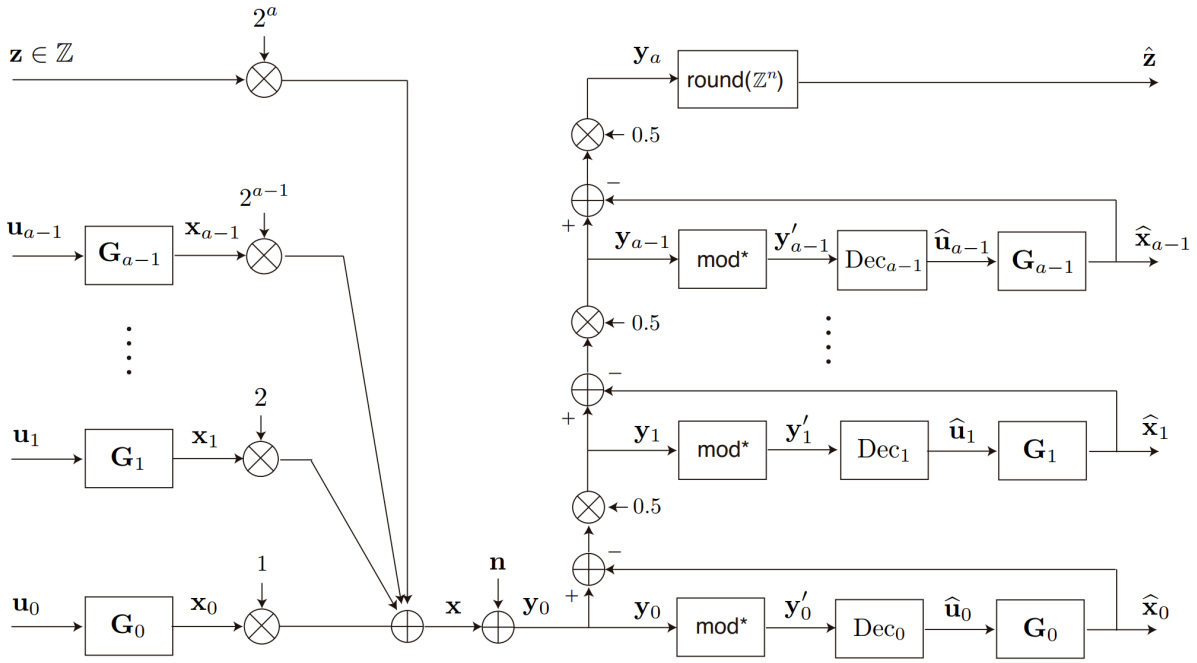

Figure 4.6: Encoder and Decoder structures of Construction D lattices.

$*$ Encoding and Decoding of Construction D Model

Fig. 4.6 describes the encoder, channel and successive cancellation decoder structures of Construction D lattices.

We consider the transmitted lattice point expressed as

$$\mathbf{x} = \mathbf{G} \cdot \mathbf{b}, \tag{4.24}$$

which may be decomposed as:

$$\mathbf{x} = \tilde{\mathbf{G}}_0 \cdot \mathbf{u}_0 + 2\tilde{\mathbf{G}}_1 \cdot \mathbf{u}_1 + \cdots + 2^a \tilde{\mathbf{G}} \cdot \mathbf{z} \tag{4.25}$$

Define $\mathbf{x}_i$ as $\mathbf{x}_i = \mathbf{G}_i \cdot \mathbf{u}_i$, for $i = \{0, 1, \ldots, a-1\}$. Then the transmitted lattice point is expressed as

$$\mathbf{x} = \mathbf{x}_0 + 2\mathbf{x}_1 + \cdots + 2^a \mathbf{z} \tag{4.26}$$

And the received point is given by

$$\mathbf{y}_0 = \mathbf{x} + \mathbf{n} \tag{4.27}$$

where $\mathbf{n}$ is noise.

Consider using $\mathcal{C}_0$ to decode $\mathbf{x}_0$. Apply a modulo-2 operation to the received sequence, so that the contribution of $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are removed as

$$\begin{aligned} \mathbf{y}'_0 &= \mathbf{y}_0 \mod 2 \\ &= \mathbf{x}_0 \mod 2 + \mathbf{n}' \\ &= \tilde{\mathbf{G}}_0 \cdot \mathbf{u}_0 \mod 2 + \mathbf{n}', \end{aligned} \tag{4.28}$$

Algorithm 1 Decoding Construction D Lattice.

Input: noisy input $\mathbf{y}$, generator matrices $\tilde{\mathbf{G}}_0, \tilde{\mathbf{G}}_1, \ldots \tilde{\mathbf{G}}_{a-1}$ for Λ
Ouput estimated lattice point $\hat{\mathbf{x}}$

1: $\mathbf{y}_0 = \mathbf{y}$
2: **for** $i = 0, 1, \ldots, a-1$ **do**
3: $\quad \mathbf{y}'_i = |\mathrm{mod}_2(\mathbf{y}_i + 1) - 1|$
4: $\quad \hat{\mathbf{c}}_i = \mathrm{Dec}_i(\mathbf{y}'_i)$ or $\hat{\mathbf{u}}_i = \mathrm{Dec}_i(\mathbf{y}'_i)$
5: $\quad \hat{\mathbf{x}}_i = \tilde{\mathbf{G}}_i \cdot \hat{\mathbf{u}}_i$ or $\hat{\mathbf{x}}_i = \tilde{\mathbf{G}}_i \cdot (\mathbf{E}_i \odot \hat{\mathbf{c}}_i)$
6: $\quad \mathbf{y}_{i+1} = \frac{\mathbf{y}_i - \hat{\mathbf{x}}_i}{2}$
7: **end for**
8: $\hat{\mathbf{z}} = \lfloor \mathbf{y_a} \rceil$
9: $\hat{\mathbf{x}} = \hat{\mathbf{x}}_0 + 2\hat{\mathbf{x}}_1 + \ldots + 2^{a-1}\hat{\mathbf{x}}_{a-1} + 2^a\hat{\mathbf{z}}$

where $\mathbf{n}'$ is the noise after the modulo operation. The decoder Dec_0 expects $\mathbf{c}_0$ plus noise, but is provided with $\mathbf{x}$ mod 2 plus noise. The modulo-2 value of the lattice point x is $\tilde{\mathbf{G}}_0 \cdot \mathbf{u}_0$ mod 2, and

$$\tilde{\mathbf{G}} \cdot \mathbf{u}_0 \mod 2 = \tilde{\mathbf{G}} \odot \mathbf{u}_0 \tag{4.29}$$

Thus, excluding the noise component, the lattice point after modulo-2 is equal to the codeword $\mathbf{c}_0 = \tilde{\mathbf{G}} \odot \mathbf{u}_0$, and we are justified in using the binary decoder.

Reencoding is the key step in Construction D lattice decoding but is optional for several decoders. In reencoding, $\hat{\mathbf{x}}_i$ is obtained from $\hat{\mathbf{u}}_i$ as $\hat{\mathbf{u}}_i = \mathbf{E}_i \odot \hat{\mathbf{c}}_i$. If the decoder produces the estimated information $\hat{\mathbf{u}}_i$ directly, then this step can be omitted. The estimate $\hat{\mathbf{x}}_i$ is obtained by reencoding as

$$\hat{\mathbf{x}}_i = \tilde{\mathbf{G}}_i \cdot \hat{\mathbf{u}}_i = \tilde{\mathbf{G}}_i \cdot (\mathbf{E}_i \odot \hat{\mathbf{c}}_i) \tag{4.30}$$

Then, the estimate $\hat{\mathbf{x}}_0$ is subtracted from the input, and this is divided by 2 to obtain $\mathbf{y}_1$

$$\mathbf{y}_1 = \frac{\mathbf{y}_0 - \hat{\mathbf{x}}_0}{2}, \tag{4.31}$$

as the input of the next level. This process continues recursively, until $\hat{\mathbf{x}}_{a-1}$ is obtained.

This is the key distinction from Code formula decoding, instead of subtracting the estimated binay codeword $\hat{\mathbf{c}}_i = \tilde{\mathbf{G}}_i \odot \hat{\mathbf{u}}_i$, Construction D decoding subtracts $\hat{\mathbf{x}}_i = \tilde{\mathbf{G}}_i \cdot \hat{\mathbf{u}}_i$.

Successive cancellation proceeds by estimating $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots$, subtracting this from the received signal. Finally, the sequence $\mathbf{y}_a$ is rounded to the nearest integer sequence to estimated $\hat{\mathbf{z}}$. These procedures are depicted in Fig 4.6. Consequently, the estimated lattice point is given by

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_0 + 2\hat{\mathbf{x}}_1 + \ldots + 2^{a-1}\hat{\mathbf{x}}_{a-1} + 2^a\hat{\mathbf{z}} \tag{4.32}$$

The decoding algorithm of general Construction D lattices is described in Algorithm 1. The notation $\lfloor \mathbf{x} \rceil$ in line 8 indicates the integer sequence nearest $\mathbf{x}$. Decoder $\mathcal{C}_i$ finds the binary codeword $\{0, 1\}$ closest to $\mathbf{y}'_i$. The modulo operation applies to the noise as well, and distance to (0, 1) should be preserved. The following "triangle function" preserves these distances, and performs the modulo- 2 operation as

$$\mathrm{mod}^*(y) = |\mathrm{mod}_2(\mathbf{y}_i + 1) - 1| \tag{4.33}$$

where mod_2 indicates a component-wise modulo-2 function.

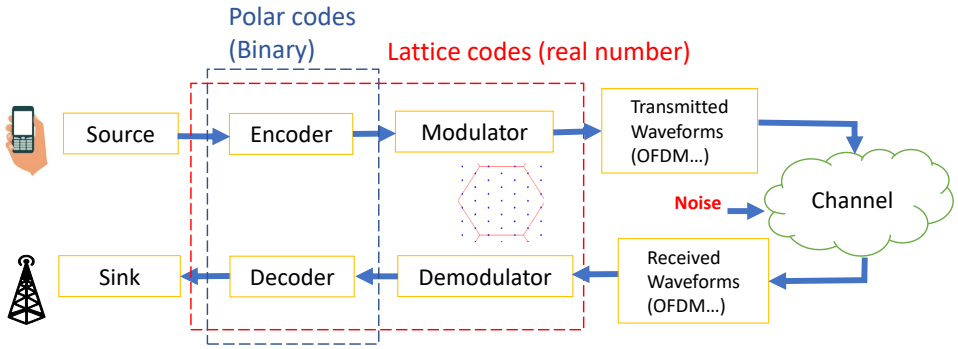

Figure 4.7: Proposed model of Polar lattices in wireless communication systems

4.4 Proposed Polar Lattices

The proposed generalized model for the polar lattices is illustrated in Fig. 4.7. The waveforms block is optional and not analyzed in this study and the simulation results as well.

Since polar lattices Λ_P are constructed from polar codes, they are also specified by the code lengths N and information lengths K of each polar codes components.

A sequence $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_N)$ is the estimate of the lattice point, which is the output of the decoder following the multilevel lattices rule.

Polar lattices produced by Construction D are represented by $\Lambda_P(K^1, K^2, \ldots, K^i)$, which are lattices formed by a group of nested binary codes, $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq ... \subseteq \mathcal{C}_a$.

Columns from the subcode generator matrix are the information bits of a polar code component, after combining the generator matrix form component generator matrix, polar lattices by construction D is obtained when the polar subcodes are identified.

4.4.1 Multi-level Encoder and Decoder

In the multi-level polar lattices, the number of level can be seen as the number of nested polar codes elements.

The generator matrix for polar lattice codes $\mathbf{G}^{\Lambda_P}$ is constructed, where $\mathbf{G}_i$ is the polar subcode generator matrix.

The generator matrix G^{Λ_P} produces lattice points $\mathbf{x}$, which are the input to the channel.

$$\mathbf{x} = \mathbf{G}^{\Lambda_P} \cdot \mathbf{b}, \tag{4.34}$$

where the $\mathbf{b}$ is computed from the summation of the component binary information vectors $\mathbf{u}_i$ in equation 4.21. We can see that both $\mathbf{b}$ and $\mathbf{x}$ are integer vectors.

The multi-level decoder implements decoding the received vectors $\mathbf{y}$ as the input. The fundamental idea is to find the closest lattice point from the noisy channel output sequence $\mathbf{y}$. The main idea of this part is consistent with the Algorithm 1 for the Construction D lattice.

Output of this decoder is the estimated codeword $\hat{\mathbf{x}}$ and it should be compared to the original transmitted codeword $\mathbf{x}$ to compute the reliability of decoder.

4.4.2 On the Design of Polar Lattices with Capacity Rule

Selecting the polar subcodes to construct the polar lattices plays a vital role in error performance.

In [39], the authors proposed the polar lattices which is analogous to the Barnes-Wall lattices from Reed-Muller codes. The basic idea is that they realized several similarities between polar codes and Reed-Muller codes and it is natural to construct the polar lattices from polar codes one each level and use Construction D to cooperate these polar codes together to produce a lattice. This construction results in better performance of polar lattices, compared to Barnes-Wall lattices.

Using a multistage decoding approach, incurred by the union bound, the overall error probability is upper bounded by the sum of the block error probabilities at individual levels

$$P_e(\mathcal{C}, \sigma^2) \leq P_e(\mathcal{C}_0, \sigma^2) + P_e(\mathcal{C}_1, (\sigma/2)^2) + \ldots + P_e(\mathcal{C}_{a-1}, ((\sigma/2^{a-1})^2) \tag{4.35}$$

Based on union bound. It is straightforward that the noise power of next level is reduced by $\sigma_i = \sigma_{i-1}/2$. Yan *et al.*'s start to compute with 2^{nd} level with the target error probability as 10^{-5} and thus compute the values of $\sigma_3 \approx 0.08719$, $\sigma_2 = 2 * \sigma_3 = 0.08719$ and $\sigma_1 = 2 * \sigma_2 = 0.3488$ respectively. Correspondingly, the code rates are also given as $K_1/N = 0.22$ and $K_2/N = 0.9$ for $N = 1024$ in simulation. This approach is useful for design of polar lattices by the *capacity rule* mentioned in [41].

4.5 Performance Evaluation of Polar Lattices

In order to analyze polar lattices system performance, the simulation result was conducted under the Monte-Carlo simulation method versus the change of VNR in dB. There are 2 main parameters taken into account such as the word error rate (WER) and symbol error rate (SER) expressed by

$$\text{SER} = \frac{\text{Errors}}{\text{Total number of symbols}} \tag{4.36}$$

$$\text{WER} = \frac{\text{any error on the codewords}}{\text{Total number of transmitted codewords}} \tag{4.37}$$

where the "word error" occurs if the estimated codeword $\hat{\mathbf{x}} \neq \mathbf{x}$ and the "symbol error" is $\hat{x}_i \neq x$, where x_i is an integer in integer vector $\mathbf{x}$.

Fig. 4.8 shows the word error rate comparison of lattices with code length of $N = 128$ by 3 different construction approaches. The first two results (red and green lines) are extracted from [39] for Barnes-Wall (BW) lattices and polar lattices constructed with BW rule. In this simulation, the Belief-Propagation (BP) decoder is utilized and the polar lattices outperformes the BW lattices versus VNR in dB.

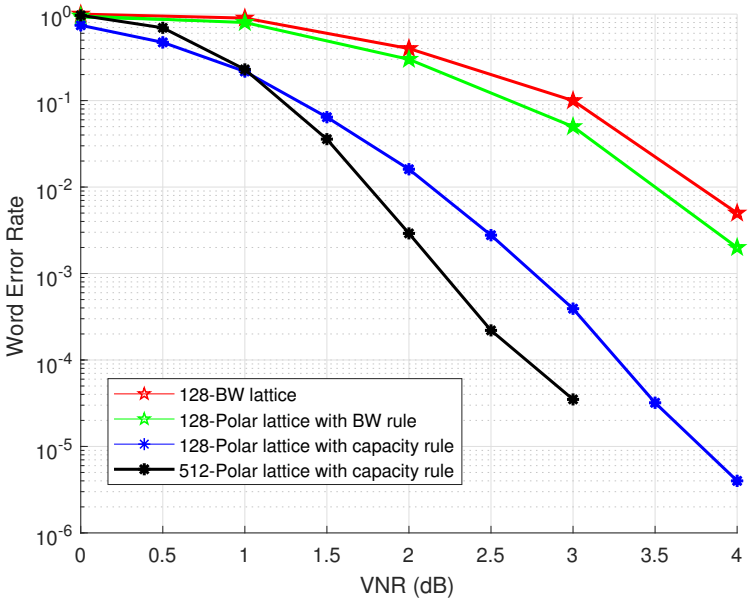

Figure 4.8: Performance comparison of polar lattices length of 128 and 512 by different code rates selection

For the sake of improvement, we also employ the polar lattices by construction D. As mentioned from previous section, the capacity rule is applied for computing the code rates and then $K_1 = \lfloor 0.22 \cdot 128 \rceil = 28$ bits and $K_2 = \lfloor 0.9 \cdot 128 \rceil = 115$ bits. The successive cancellation decoder is used as well for each level of polar codes. Under fair conditions, the proposed polar lattices are superior to both previous lattices. For example, 2.5 dB gain is obtained at 10^{-3} of WER when we apply new approach instead of choosing the BW rule for polar lattices.

Chapter 5

Discussion and Conclusion

5.1 Future work

Due to the time limitation of master period, there exits a couple of open topics that need to continue in the future

∗ **For polar codes**

- Polar codes have been shown to achieve Shannon capacity. Their applications is being discussing tremendously for 5G systems. This thesis only mentioned the low order of modulation scheme (BPSK), in order to achieve very high data and adapt to other waveform method (OFDM, FBMC ...), the higher constellation such as 64-QAM, 128-QAM should be considered in design.
- The rate matching with puncturing bits is a useful approach to increase performance of high speed transmission without redundancy of information bits. The future study should take into account this technique.

∗ **For polar lattices**

- One interesting topic is the Gaussian shaping technique for polar lattices in the power-constrained AWGN channel. This is based on source polarization. In the further research, we will be able to achieve the capacity $1/2\log(1+SNR)$ with low-complexity multistage successive cancellation (SC) decoding for any given signal-to-noise ratio (SNR). In order to achieve this, we should investigate various shaping schemes over polar lattices.
- To make polar lattices more competitive in practice, it is important to improve their finite-length performance. One way is to use more sophisticated decoding algorithms for the component polar code at each level. It is a fact that for short to moderate block lengths, the performance of polar SC decoding have a big gap to other well-known codes such as LDPC or Turbo codes. For this reason, we used cancellation list decoding (SCL), which is an enhanced version of SC decoding, and approaches the Maximum-Likelihood (ML) performance at high SNR region. Furthermore, when an outer cyclic redundancy check (CRC) code is concatenated, the SCL decoder can compete with LDPC codes. In the future, we hope to further improve the polar

lattices decoder by applying these advantages that can be competitive with LDPC lattices.

- We should investigate more approaches in selecting the code rates and the number of level in construction D for polar lattices.

5.2 Relation to the Cybersecurity Training

Data security is extremely important in common communications. it is reasonable to argue that security measures should be implemented at all network layers. In addition, with the development of ad-hoc and decentralized networks, higher-layer techniques, such as encryption and key distribution, are complex and difficult to implement. The cybersecurity, also known as IT security, focuses on protecting networks, computers, programs and data from attack, damage, unintended or unauthorized access.

Wireless network needs security especially at the physical layer due to its broadcast nature of wireless medium. In general, some aspects of cybersecurity for low level can be categorized as

- Achieving the secrecy of the physical layer of communication networks.
- Solving the the security challenges in large distributed networks.
- Assuring information privacy in data networks, databases and more broadly, big data across applications.

In this approach, the polar and lattices are considered to offer well properties for the high secrecy in the physical layer to prevent various kinds of attacks.

Cyber attacks are determined as the critical threats to any communication systems due to its variety, very short time spread and serious levels. The attacks from physical layer may come from Wormhole Attack (relay attack), location spoofing by radio signal interception and relaying, Information Leakage Attack or the key extraction attack, etc.

Cybersecurity training seems to be an efficient approach to enhance the prevention of cyber breaches effectively. Correspondingly, there exists various available training programs. The largest source for information security training in the world which is known as SANS [43], has been providing variety of training programs in term of cybersecurity major. Another well-known online learning platform is Udemy that supports a course with name of Cyber Security for presenting cybersecurity concepts in a wide range from threat analysis, risk management, encryption, and firewalls.

For the purpose of providing the cybersecurity training to everyone, the CyTrONE (Cybersecurity Training and Operation Network Environment)[42] has been recently created to automate the training content and environment setup by Cyber Range Organization and Design (CROND) NEC-endowed chair at Japan Advanced Institute of Science and Technology (JAIST). This system helps to generate a cyber range from a description in text format get by the organizer. Thanks to this framework, the cyber range creation task is much simpler and is expected to easier setup automatically in near future . Therefore, the cybersecurity training for the physical layer security should be one of the important elements.

5.3 Conclusion

In this work, we have presented the theoretical and practical aspects of polar codes versus the power constraint (denoted by the signal-to-noise ratio or E_b/N_0) and extended to the polar lattices by Construction D versus the unconstrainted power (denoted by the volume-to-noise ratio VNR).

The main issues of this study can be separated into 2 parts as follows

- In term of polar codes: we presented the fundamentals of polar codes, how to construct the encoders and decoders and then proposed a basic scheme for 5G system using Polar codes with performance comparison to LDPC and turbo codes. That is, polar codes performed very well in range of small and moderate code lengths with reasonable complexity. In additions, with the modifications of adaptive list decoding that allows the list size to adjust according to its computational power, the polar codes are expected to apply in a wide range of code lengths, code rates with the flexible decoding complexity. Furthermore, we also investigated the performance or polar codes under the Rayleigh fading conditions and apply the advanced SCL decoder to these schemes to reduce the bad effects of fading schemes.

- In term of polar lattices: One of the useful applications from lattices has been proposed in this work, we show how good the lattices syndrome decoding is, and thus it can be promising candidate for the MIMO system that reach the Near-Maximum-Likelihood approach.

  We have proposed a polar lattices by Construction D with modified multi-level decoding for Code formula decoding, instead of subtracting the estimated binary codeword $\hat{\mathbf{c}}_i$, Construction D decoding subtracts integer vector $\hat{\mathbf{x}}_i$. The algorithm decoding also is demonstrated with detailed explanation. Since the polar lattices are derived from the polar codes, their construction is equally efficient. The numerical results also indicates that the polar lattices constructed by proposed code rate selection outperforms the polar lattices by the Barnes-Wall rule in [39]. As the lattices dimensions N increases, this approach is expected to get the capacity-achieving and can be competitive with LDPC lattices. In the further researches with the Gaussian shaping, polar lattices may achieve the capacity of the power-constrained AWGN channel [40].

Publication

1. Long H. Nguyen, Brian M. Kurkoski, "General-Purpose Lattice Decoding Using Lookup Tables" in submitted to *IEEE International Symposium on Information Theory*, June. 2018. Acceptance notification is sent out by March 31, 2018.

Bibliography

[1] IMT-2020 (5G) Promotion Group, "5G vision and demand", May. 2014

[2] IEEE 802.11-2012-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY), 2012.

[3] N. Costa, and S. Haykin, *Multiple-Input, Multiple-Output Channel Models: Theory and Practice,* Wiley-Interscience, 2010.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.,* vol. 27, pp. 379-423, 623-656, July and October 1948.

[5] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory,* vol. 55, pp. 3051-3073, July 2009.

[6] E. Arkan et al., "A performance comparison of polar codes and ReedMuller codes," *IEEE Commun. Lett,*vol. 12, no. 6, pp. 447-449, 2008

[7] Thomas M. Cover, Joy A. Thomas, "Elements of information theory," 1st Edition. New York: Wiley-Interscience, 1991.

[8] Andrea Goldsmith, "Wireless Communications," Cambridge University Press,. 2005.

[9] H. Si, O. O. Koyluoglu, and S. Vishwanath, "Polar coding for fading channels: binary and exponential channel cases," *IEEE Trans. Commun.,* vol. 62, no. 8, pp. 2638-2650, 2014.

[10] J. J. Boutros and E. Biglieri, "Polarization of quasi-static fading channels, in *International Symposium on. Information Theory Proceedings (ISIT), IEEE, 2013,* pp. 769-773.

[11] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory,* vol. IT-8, PP. 21-28, Jan. 1962.

[12] H. Vangala, E. Viterbo, and Y. Hong, "A comparative study of polar code constructions for the AWGN channel, *arXiv preprint arXiv:1501.02473,* 2015.

[13] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Transactions on Information Theory,* vol. 61, no. 5, pp. 2213-2226, 2015.

[14] Y. Zhang, A. Liu, X. Pan, Z. Ye, and C. Gong, "A modified belief propagation polar decoder," *IEEE Commun. Lett.,* vol. 18, no. 7, pp. 1091-1094, Jul. 2014.

[15] K. Niu and K. Chen, "Stack decoding of polar codes," *Electron. Lett.,* vol. 48, no. 12, pp. 695696, 2012.

[16] K. Chen, K. Niu, and J. Lin, "Improved successive cancellation decoding of polar codes," *IEEE Trans. Commun.,* vol. 61, no. 8, pp. 31003107, Aug, 2013.

[17] V. Miloslavskaya and P. Trifonov, "Sequential decoding of polar codes," *IEEE Communications Letters,* vol. 18, no. 7, pp. 1127-1130, 2014.

[18] Lin Qi, Yu Xu, Tong Liu, Zheng Dou, "An improved successive cancellation decoder for polar codes" in *IEEE International Conference on Electronic Information and Communication Technology (ICEICT)*, Aug. 2016.

[19] Sha Shi, Bing Han, Jing-Liang Gao, and Yun-Jiang Wang, "Enhanced Successive Cancellation List Decoding of Polar Codes," *IEEE Communications Letters,*, Vol. 21, No. 6, pp. 1233-1236, 2017.

[20] A. Balatsoukas-Stimming, M. B. Parizi, and A. Burg, "LLR-based successive cancellation list decoding of polar codes," *IEEE Trans. Signal Process.,* vol. 63, no. 19, pp. 5165-5179, Oct. 2015.

[21] M. Sybis, K. Wesolowski, K. Jayasinghe, V. Venkatasubramanian and V. Vukadinovic, "Channel Coding for Ultra-Reliable Low-Latency Communication in 5G Systems," *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall),* Montreal, QC, 2016, pp. 1-5.

[22] H. Gamage, N. Rajatheva and M. Latvaaho, "Channel coding for enhanced mobile broadband communication in 5G systems," *2017 European Conference on Networks and Communications (EuCNC)*, Oulu, 2017, pp. 1-6.

[23] Satish Babu Korada, Eren Sasoglu, and Rudiger Urbanke, "Polar codes: characterization of exponent, bounds, and constructions," *IEEE Trans. Inform. Theory,* vol. 56, no. 12, Dec, 2010.

[24] Hao Liang, Aijun Liu, Yingxian Zhang, and Qingshuang Zhang, "Analysis and Adaptive Design of Polar Coded HARQ Transmission Under SC-List Decoding," *IEEE Wireless Communications Letters,* vol. 6, no. 6, Dec. 2017.

[25] P. Trifonov, "Efficient Design and Decoding of Polar Codes," *IEEE Trans. Commun,* vol. 60, no. 11, pp. 3221-3227, 2012.

[26] Simulations by Iterative Solutions Coded Modulation Library, 2007. http://www.iterativesolutions.com/Matlab.htm

[27] B. Zhang, H. Shen, B. Ying, "A 5G Trial of Polar Code", *2016 IEEE Globecom Workshops (GC Wkshps), Washington, DC,* 2016, pp. 1-6.

[28] Session Chairman (Nokia), "Chairmans Notes of Agenda Item 7.1.5 Channel coding and modulation," *3GPP TSG RAN WG1 Meeting 87,* R1-1613710, Reno, USA, November 2016.

[29] L. Liu and C. Ling, "Polar codes and polar lattices for independent fading channels," *IEEE Transactions on Communications,* vol. 64, no. 12, pp. 4923-4935, 2016

[30] H. Yao and G. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Conf. Global Commun.,* , Nov. 2002, vol. 1, pp. 424-428.

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

[32] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups,* 3rd edition. Springer Verlag, 1999.

[33] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Transactions on Information Theory,* vol. 40, pp. 409-417, March. 1994.

[34] W. H. Mow, "Universal lattice decoding: Principle and recent advances," *Wireless Commun. Mobile Comput., Special Issue on Coding and Its Appl. Wireless CDMA Syst.,* vol. 3, pp. 553-569, Aug. 2003.

[35] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, pp. 2202-2214, Aug. 2002.

[36] G. D. Forney, Jr., "Multidimensional constellations—Part II: Voronoi constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 941-958, August 1989.

[37] Long H. Nguyen, Brian M. Kurkoski, "Lattice Syndrome Decoding for MIMO Detection" in proceeding of *technical research report, JAIST*, December. 2017.

[38] E. S. Barnes and N. J. A. Sloane, "New lattice packings of spheres," *Canadian Journal of Mathematics,* vol. XXXV, no. 1, pp. 117-130, 1983.

[39] Y. Yan and C. Ling, "A construction of lattices from polar codes," in Proceedings of the *IEEE Information Theory Workshop,* pp. 124-128, Lausanne, Switzerland, 2012.

[40] Y. Yan, L. Liu, C. Ling, and X. Wu, "Construction of capacity-achieving lattice codes: Polar lattices," *ArXiv e-prints,* vol. abs/1411.0187, November 2014. [Online]. Available: http://arxiv.org/abs/1411.0187.

[41] G. D. Forney Jr., M. Trott, and S.-Y. Chung, "Spherebound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inform. Theory,* vol. 46, no. 3, pp. 820 -850, May 2000.

[42] R. Beuran, C. Pham, D. Tang, K. Chinen, Y. Tan, Y. Shinoda, "CyTrONE: An Integrated Cybersecurity Training Framework," *International Conference on Information Systems Security and Privacy (ICISSP 2017), Porto, Portugal,* February 19-21, 2017.

[43] SANS Information Security Training — Cyber Certifications — Research. Retrieved on August 4th, 2016 from https://www.sans.org/.