Interference management for multiple access relay channel in LTE-advanced using nested lattice

QIANG WEN

Master’s Degree Project
Stockholm, Sweden June 2012

School of Electrical Engineering, KTH
I would like to first express my deep-felt gratitude to my supervisor, Dr. Ming Xiao of the School of Electrical Engineering at the Royal Institute of Technology at Stockholm Sweden (KTH, Kungliga Tekniska Högskolan), for his valuable advice, warmly encouragement, enduring patience and constant support. This master thesis could not have been written without the kindly help from Dr. Ming Xiao, who not only served me as my supervisor but also encouraged and challenged me throughout my academic study. Thank you, Dr. Ming Xiao!

I also wish to thank my university, the Royal Institute of Technology, where I had an opportunity to develop the fundamental and essential academic competence. With this opportunity, I would like to thank all the teachers in KTH, without your nutritious teaching, I cannot even go any further in this thesis.

Additionally, I want to thank my colleagues in Huawei Sweden and Huawei Norway for all their hard work and dedication, from them I have a first-hand experience with live 4G network, which provides me fresh ideas from industry to complete my thesis and prepare for a career as an engineer for the newest mobile network.

And finally, I must thank my family and my girlfriend Shanshan for putting up with me during the development of this work with continuing, loving support and no complaint. Words alone cannot express how much I love them, but I do believe it is because of them that make all things possible for me.
Abstract

Although radio access technology has huge expansion in the past decades, interference management is a key concern for today’s mobile communication systems, a concern in the demand by an ever increasing range of potential applications. Future wireless network including LTE-advanced(also known as the standard for the next generation mobile communication system) will not only need to support higher data rate comparing with existing mobile network in order to meet the increasing customer demand for multimedia services, but also new techniques are under research to decrease high interference comparing with current LTE system.

Relay nodes are supposed to be supported in LTE-advanced, which will bring a huge technical innovation for current network’s structure. Network coding in relay network application is already proved to increase the throughput, improve system’s efficiency and enhance system capacity in many papers. Together with relay nodes, the new standard will also make some deep research in interference management, for example evolved inter-cell interference coordination(eICIC) and coding schemes.

This master thesis focuses on the performance acquired in multiple access relay channel(MARC) together with network coding technique, on the other hand, a new channel coding method, namely nested lattice coding, has attracts most of interests throughout this master thesis.

In this MARC network, the sources map their messages using lattice code and then broadcast them to the relay and the destination. The relay receives two independent symbols through the same channel, it will combine the two symbols using modulo lattice and then forward the new symbol to the destination node. The destination recovers those two messages using two linear equations one directly from the sources and the other one forwarded by the relay. Although this method will have some information loss, while it reduces the transmission time slots and improves the system’s efficiency.

The implementation of nested lattice in MARC network also introduces modulo lattice transformation to achieve the capacity, where the coarse lattice is used for shaping while the fine lattice serving for channel coding. It is proved that for the modulo-lattice additive noise channel, lattice decoding is optimal.

**Keywords:** interference management, LTE-advanced, multiple access relay channel, network coding, lattice code, nested lattice.
Contents

Acknowledgements i
Abstract ii
List of Figures v
Abbreviations vii

1 Introduction 2
1.1 The challenge for current LTE network 2
1.2 Interference management in LTE advanced 3
1.3 Network coding using relay 4
1.4 Lattice code for relay network 4
1.5 Structure of the report 5

2 Network coding in multiple access relay channel 6
2.1 Network coding theory 6
2.2 Relay channel 8
2.2.1 Relaying scheme 8
2.2.2 Capacity of relay channel 9
2.2.3 Relay channel applications under research 10
2.3 Network coding in relay channel 12
2.3.1 Network coding in two way relay channel 12
2.3.2 Network coding in multiple access relay channel 13

3 Lattice is everywhere 16
3.1 Introduction 16
3.2 Definition of lattice 16
3.3 Sphere packing, covering and kissing number 17
3.3.1 Sphere packing 17
3.3.2 Sphere covering 18
3.3.3 Kissing number 19
3.4 Quantizers 19
3.5 Gaussian channel coding 21
3.6 Conclusion 22
List of Figures

1.1 Inter-Cell Interference Coordination in LTE ........................................ 2

2.1 Butterfly network, network coding improve transmission efficiency. ........... 7
2.2 Relay channel ......................................................................................... 9
2.3 Three-node two way relay channel ......................................................... 11
2.4 A two-relay network system: (a) Traditional multi-hopping scheme with 4 time slots. (b) MAC-Layer XOR scheme with 3 time slots. .......... 12
2.5 Multiple access relay channel with network coding ............................. 13
2.6 Bit error rate for message A and message B through network coded MARC using MMSE .......................................................... 14

3.1 One dimensional sphere. ................................................................. 17
3.2 Sphere covering for different arrangements: (a) the square lattice, (b) the hexagonal lattice ................................................................. 19
3.3 The perfect kissing arrangement for \( n = 2 \), it is easy to prove that in two dimensions the kissing number is 6. ..................... 20
3.4 The output of the quantizer chooses the nearest center. ............ 20

4.1 Dimension 2 lattice encoder mapping scheme. ....................... 27
4.2 Lattice code based scheme showing the transmitted signal(circle) and the decoded signal(dots) .................................................. 28

5.1 Nested lattices for single-user coding. Black dots are elements of the coding lattice, and blue dots are elements of the shaping lattice. Each lattice point inside the shaded Voronoi region is a member of the codebook.................................................. 31
5.2 Multiple Access Relay Channel: (a) Traditional multi-hopping scheme with 4 time slots. (b) MAC-Layer XOR scheme with 3 time slots. (c) Lattice coding scheme with 2 time slots. ........... 32
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>AF</td>
<td>Amplify-and-Forward</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise channel</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>CF</td>
<td>Compress-and-Forward</td>
</tr>
<tr>
<td>COPE</td>
<td>Coding Opportunistically</td>
</tr>
<tr>
<td>DF</td>
<td>Decode-and-Forward</td>
</tr>
<tr>
<td>DL</td>
<td>DownLink</td>
</tr>
<tr>
<td>eICIC</td>
<td>Enhanced Inter Cell Interference Coordination</td>
</tr>
<tr>
<td>GSA</td>
<td>Global mobile Suppliers Association</td>
</tr>
<tr>
<td>GSM</td>
<td>the Global System for Mobile communications</td>
</tr>
<tr>
<td>HSPA+</td>
<td>Evolved High-Speed Packet Access</td>
</tr>
<tr>
<td>ICIC</td>
<td>Inter-Cell Interference Coordination</td>
</tr>
<tr>
<td>LDLC</td>
<td>Low Density Lattice Codes</td>
</tr>
<tr>
<td>LDPC</td>
<td>Low Density Parity Check</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAC</td>
<td>Multiple Access Channel</td>
</tr>
<tr>
<td>MARC</td>
<td>Multiple Access Relay Channel</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input and Multiple-Output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MLAN</td>
<td>Modulo-lattice additive noise</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean-Squared Error estimation</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>TWRC</td>
<td>Two Way Relay Channel</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>XOR</td>
<td>Exclusive OR</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 The challenge for current LTE network

Based on the standard of 3rd Generation Partnership Project (3GPP), the technical targets of LTE (Long Term Evolution) include peak data rates in excess of 300 Mbps, delay and latencies of less than 10 ms and manifold gains in spectrum efficiency. Unlike the previous generations, LTE first introduces orthogonal frequency division multiplexing (OFDM) for modulation. Due to its big technical advantages comparing with GSM and UMTS, it has well development after its final standard by 3GPP mainly in Release 8 with some enhancements in Release 9.

After LTE first commercial launch in 2009 by TeliaSonera, it soon attracts other operators’ attention. In this January (year 2012), the GSA (Global mobile Suppliers Association) had published a report, confirming 49 LTE operators have now launched commercial services, and 285 operators have committed to commercial LTE network deployments or are engaged in trials, technology testing or studies[28].

LTE is anticipated to become the first truly global mobile phone standard, while different frequency bands in different countries will be used. No matter some low frequency bands like 700MHZ being used by Verizon USA, it has interference from GSM (the Global System for Mobile Communications), or high

Figure 1.1: Inter-Cell Interference Coordination in LTE
“clean” frequency bands like 2.6GHz being used in Europe and Asia, it still has internal interference at cell edge, and this is the main drawback of current LTE network.

For this demand, inter-cell interference coordination (ICIC, refer to figure 1.1) has been the topic of research since GSM. In this technique, those three neighboring cells divide their total bandwidth into three parts and each cell only use one part of it, as hexagonal model is being used, those cells with the same bandwidth parts will never be neighbors. While this technique has a very clear disadvantage, which is the total bandwidth is divided by 3 and the resource blocks allocation in LTE system decreases dramatically, which means the performance of LTE, mainly the DL throughput, will decrease largely, new strategy for interference management is required for next generation mobile communication technique.

1.2 Interference management in LTE advanced

Comparing with the performance of UMTS (Universal Mobile Telecommunications System) networks, LTE Rel. 8 does not offer anything substantially unique to significantly improve spectral efficiency and interference management strategy. After LTE Advanced was standardized by 3GPP as a major enhancement of LTE standard, one of the key aspect of LTE Advanced benefits is the ability to take advantage of advanced topology networks; deployment of low power nodes in macro network, such as relays, picos and femtos.

In current LTE system, it improves system performance by using wider bandwidths if spectrum is available, while LTE Advanced brings the network closer to the user by adding many of these low power nodes, which is a significant performance leap in wireless networks to make the most use of topology to improve spectral efficiency and interference management.

As introduced above that the new technique ICIC has been discussed since GSM and implemented in current LTE network, while LTE Release 8 only gave a limited ICIC and does not provide mechanisms for DL control channel ICIC, and also limited number of UEs (User Equipment) can be associated with low power eNodeBs (e.g., relays, picos, etc.), which limits potential for load balancing and increase in network throughput.

LTE Advanced, in another point of view, proposed a new technique called eICIC (Enhanced Inter Cell Interference Coordination) [26] to extend ICIC to DL control in time domain effectively. The backhaul connection based on eICIC to the base station, serving as relay node in this case, can be on the same frequency (in-band) or on different frequencies (out-band) [33]. Here, the layer of planned higher-power macro eNodeBs is overlaid with layers of lower-power pico, femto or relay eNodeBs that are deployed in a less well planed or even entirely uncoordinated manner is called heterogeneous network, which has recently attracted considerable attention to cancel the high interference in the cell-edge and optimize the performance.
The relay network also brings a lot of research interests before and after it is announced to be considered in the newly LTE Advanced standard, the most attractive part is its usage in network coding and channel coding theory to enhance throughput and improve system performance.

1.3 Network coding using relay

After first brought to the world by three researchers from Chinese University of Hong Kong, network coding theory soon has a huge influence for the later network research. Network coding is designed to increase the possible network throughput, and in the multicast case can achieve the maximum data rate theoretically, as with the help of relay nodes, the number of transmissions reduces, less transmission times will bring large advantage in data throughput.

Unfortunately, however, the existing network coding approach still does not exploit the potential of wireless channel. This is because wireless environment is totally different with fixed transmission, for bad wireless channel the successful transmission times will decrease dramatically, which will bring impacts not only for this channel but also for the whole network as whether the received message can be decoded is also rely on the packets delivered from the bad wireless channel. In essence, the current network coding approach effectively forces the throughput gain bound to the capacity of the worst link, which tends to fall with the diversity of links[60].

To tackle the bottleneck problem of the wireless network coding, the relay node may not be just considered as a node to increase the total transmission rate while it should be a node which can improve the total system performance and robustness. Two way relay network and multiple access relay network are the two main relay network models under current research, and the improvement of the relay network in interference cancellation and system capacity are clear to see.

1.4 Lattice code for relay network

As the network coding has a bottleneck in wireless environment where the system performance of the whole network depends on the worst channel, many researches are mainly focus on channel coding to find new channel coding schemes to decrease the impacts of wireless network. It is, hence, of interest to investigate the maximal reliable transmission rates achievable by structured ensembles of codes.

An important class of structured codes is the class of lattice codes. Shannon’s theory suggests that the codewords of a good code should look like realizations of a zero-mean independent and identically distributed(i.i.d.) Gaussian source with power $P_X$, while De Buda’s theorem states that a code with second moment $P_X$ can approach arbitrarily closely the AWGN channel capacity. Thus Lattices is emerged as a powerful approach for the design of structured, low-complexity codes for AWGN channels. In addition to offering structure, achieving capac-
ity, and reducing complexity, lattice codes are desirable because they are the Euclidean-space analogue to linear codes.

Lattice codes have also proven useful for multi-user systems, especially in relay networks. With the help of network coding theory, Lattices have been used to establish new achievable rates in network-coded systems and achieve the channel capacity of AWGN broadcast channel[62]. To achieve the significant improvement in transmission rate and system performance using Lattices in relay network, this is also the main task of this report.

1.5 Structure of the report

Based on the introduction given above, this report is organized as follows.

Chapter 2 is dedicated to the network coding theory and relay channels. All the background about network coding and its implementation in relay networks are detailed there, since two way relay channel and multiple access relay channel are the two main relay model under research discussion, these two models are also discussed with their implementations in network coding phenomenon.

Since some basic introduction of lattices is required for the code construction, Chapter 3 recalls elementary definitions and properties of lattices.

A very important feature to consider when designing codes is their encoding and decoding. Chapter 4 gives a universal lattice decoding algorithm call Sphere Decoder.

Chapter 5 introduces the key notion of nested lattices, which gives a unifying context for understanding how lattice codes can implement in relay networks. It allows modulo-lattice additive noise channel to be the key technique to be considered for multiple access relay channel.

In Chapter 6, we give a brief overview of the whole report, and finalize the report with some conclusions. Also, some potential works for future research are also given in this chapter.
Chapter 2

Network coding in multiple access relay channel

2.1 Network coding theory

Large scale communication networks like Internet and telecommunication networks play a very important role in our daily life, at first researchers from academy and industry always tried to increase the network’s efficiency by using more valid switching theory. Until more than ten years ago, three researchers from Chinese University of Hong Kong announced a new theorem[42] which is called network coding different from the physical layer coding, it is designed to increase the possible network throughput, and in the multicast case can achieve the maximum data rate theoretically, it soon has a huge influence for the later network research.

The theory of network coding has been developed in various directions, and new applications of network coding continue to emerge[38]. Linear network coding theory is mostly considered[49], for example, if the data is moving from S source nodes to K sink nodes, so a message generated (stated as $X_k$) is a linear combination of the earlier received messages $M_i$ (considered as “evidence”) on the link by coefficient $g_{ik}$, and the relation between them is stated below:

$$X_k = \Sigma_{i=1}^{S} g_{ik} \cdot M_i \quad (2.1)$$

This equation yields a Gaussian estimation problem $X = G \cdot M$, where with the knowledge of $X$ and $M$ and the technique of Gaussian estimation, it is easy to solve the equation to obtain message $M$.

Network coding theory announces to replace routers by encoders in networks, it works by sending out the evidence of the messages(linear combination) rather than the entire messages. The evidence will be decoded at the receiver side by using the information it has[22]. Thus, coding offers the potential advantage of minimizing both latency and energy consumption, and at the same time maximizing the bit rate[38].
Figure 2.1: Butterfly network, network coding improve transmission efficiency.

The butterfly network is frequently applied as a classical example for linear network coding theory, in which there are two sources (refer to the top two nodes in figure 2.1), each has the message A or B. And there are two targets (refer to the bottom two nodes) requiring both message A and B. Each link can only carry one message at the same time which means it only transmits one bit in each time slot.

If routing is the only method to apply, then the central link would be the bottleneck which can only transmit either A or B simultaneously. In detailed circumstances, suppose the central link transmits message A at first, then the left destination would receive the message A twice while not know message B at all. The situation would also appear at the right destination node if message B is sent first. Hence, routing is insufficient because with routing scheme one more extra transmission (means one more time slot) is required to transmit message B to the left destination node or message A to the right destination node, and one redundant message is transmitted. This is network coding theory’ application to reduce the extra time slot and improve the efficiency by sending the linear combination of the messages A and B, in other words, A and B is encoded by using the formula “A ⊕ B” (exclusive OR). The left target node receives message A and combined message “A ⊕ B”, and can find B by the operation “A ⊕ (A ⊕ B)”. This is an application for linear network coding as the encoding and decoding schemes are all linear operations. While exclusive-OR(XOR) operation is a frequently used example for encoding and decoding[44], which can increase the throughput, reach the theoretical max-flow and optimize the resource utilization[49].

The above information and huge number of papers have showed the utility of network coding for multicast in wire line packet networks, when it comes to wireless network, due to its various and unique problems such as low throughput, dead spots in poor coverage area, inadequate support for mobility, while
large number of market demands, it soon attract interest in employing network coding in wireless networks. COPE was the first system architecture making network coding work in the IEEE 802.11-based wireless network\cite{44}. The main features of network coding that are most relevant to wireless networks are discussed through the paper \cite{53}, and the paper \cite{6} also explores the case for network coding as a unifying design paradigm for wireless networks, by describing how it addresses issues of throughput, reliability, mobility, and management.

Although the characteristics of wireless networks might all seem disadvantageous at the first sight, but a newer perspective reveals that some of them can be used to our advantage\cite{6}, for example broadcast, whenever one node broadcast a message, at least one nearby node receive it and forward it to the next hop, which brings spatial diversity\cite{43, 3, 1}. Wireless network also brings significant data redundancy because of multipath effects for example, while it also provides an opportunity to deal with unreliability and robustness of wireless links, for example, redundancy can be exploited to increase the information flow per transmission, and thus improve the overall network throughput\cite{6} and decrease the transmission error rate. These advantages are not escaped the notice of researchers and engineers, in increasing number of papers they explore the concept of relay channel to introduce diversity, different relay channels and their applications and advantages will be discussed in the following sections.

2.2 Relay channel

Relay was introduced to broaden coverage, enhance system capacity or improve robustness of a system. A relay channel is defined as a communication model that between a sender and a receiver one or more intermediate relay nodes is aided, it is a combination of broadcast channel(from sender to relay and receiver) and multiple access channel(from sender and relay to receiver).

2.2.1 Relaying scheme

In general concept, the relay can either transmit its own message or forward and amplify the message from sender to receiver, based on this idea, the relay channel can be divided into the following three relaying schemes:

1. Decode-and-Forward (DF): the relay decodes the source message in one block and transmits the re-encoded message in the following block.

2. Compress-and-Forward (CF): the relay quantizes the received signal in one block and transmits the encoded version of the quantized received signal in the following block.

3. Amplify-and-Forward (AF): the relay sends an amplified version of the received signal in the last time-slot.

Comparing with DF and CF, AF requires much less delay as the operations relay node are divided by time-slot rather than message block. Also, AF saves operation power since no decoding or quantizing operation is performed at the relay node.
There are four variables in simplest one relay network need to be considered before discussing the capacity, $X$ is the channel input and the output is $Y$; the relay’s observation is $Y_1$ and $X_1$ is the input chosen by the relay and depends only on the past observation ($Y_{11}, Y_{12}, ..., Y_{1i-1}$), please refer to figure 2.2. The capacity problem simplifies to determine the channel capacity between $X$ and $Y$ [59], which is showed in the following theorem.

**Theorem 1.** For an arbitrary relay channel, the up bound of the capacity is given by

$$ C \leq \max_{p(x,x_1)} \min \{I(X,X_1;Y), I(X;Y,Y_1|X_1)\} \quad (2.2) $$

The first term $I(X,X_1;Y)$ shows the transmission rate from the sender $X$ and the relay (send $X_1$) to receiver $Y$ (multiple access channel), the second term $I(X;Y,Y_1|X_1)$ illustrates the rate from $X$ to $Y$ and $Y_1$ (broadcast channel). Detailed proof is shown in [8].

In wireless environment, things are more complicated due to its significant characteristics like fading, here below we consider a model for wireless channel to analysis the capacity of relay channel in wireless environment. Suppose at time $t$ the terminal $i$ receives the symbol:

$$ Y_{it} = Z_{it} + \sum_{s \neq i} A_{s it} \frac{A_{si}}{\sqrt{d_{si}}} X_{st} \quad (2.3) $$

Where $d_{si}$ is the distance between terminals $s$ and $i$, $\alpha$ presents an attenuation exponent, $A_{s it}$ illustrates a complex fading random variable, and $Z_{it}$ is independent and identically distributed (i.i.d.) complex Gaussian noise with zero mean, unit variance, and i.i.d. real and imaginary parts [18].

There are two fading scenarios to be considered:
1. No fading, and $A_{s it} = 1$ for all $s$, $i$ and $t$
2. Phase fading, and $A_{s it} = e^{j\theta_{s it}}$, where $\theta_{s it}$ is uniformly distributed over $[0, 2\pi)$
No fading with one relay

Suppose no fading in only one relay network, based on Theorem 1, the up bound
for the date rate with Gaussian input distributions is showed below:

\[
R \leq \max_{0 \leq \rho \leq 1} \min\{\log(1 + \frac{P_1}{d_{I2}^2}), \log(1 + \frac{P_1}{d_{I3}^2} + \frac{P_2}{d_{23}^2} + 2\rho\sqrt{P_1P_2})\} \tag{2.4}
\]

Where \(\rho\) is the correlation coefficient of \(X\) and \(X_1\), example with plotted
figure can be found in [18].

Phase fading with one relay

When phase fading is introduced in one relay network, \(A_{s_i} = e^{j\theta_{s_i}}\), where \(\theta_{s_i}\) is
known only to terminal \(i\) for all \(s\), and the up bound of the capacity in Thereom
1 becomes:

\[
\max_{\rho(\cdot;\cdot)} \min\{I(X;X_1|Y\theta_{X,X_1}), I(X,Y;Y_1|\theta_{XY}\theta_{X_1,Y})\} \tag{2.5}
\]

Based on the procedure in [18], when \(\rho = 0\), the above equation can be
simplified and maximized as:

\[
\min\{\log(1 + \frac{P_1}{d_{I2}^2}), \log(1 + \frac{P_1}{d_{I3}^2} + \frac{P_2}{d_{23}^2})\} \tag{2.6}
\]

It shows that in a multi-hopping with phase fading system, it will achieve
the channel capacity if the relay is in the region near the source terminal, and
the capacity at that situation can be:

\[
C = \log(1 + \frac{P_1}{d_{I3}^2} + \frac{P_2}{d_{23}^2}) \tag{2.7}
\]

Phase fading with many relays

From the results in the above section for phase fading with only one relay, we
can conclude that in the situation when phase fading with many relays, the
channel capacity can be maximized when all the relays are near the source, and
the corresponding capacity can be:

\[
C = \log(1 + \sum_{i=1}^{l-1} \frac{P_i}{d_{II}^2}) \tag{2.8}
\]

2.2.3 Relay channel applications under research

There are two main relay channels frequently using for research due to their
simplicity and typicality, those two examples are two way relay channel(TWRC)
and multiple access relay channel(MARC).

Two way relay channel

Two way relay channel is the simplest three-node relay channel with Gaussian
additive noise in which two end nodes exchange information via a relay node, it
is attracting increasing attention in research area, figure 2.3 shows a structure
of three-node two way relay channel.

The capacity for two way relay channel is discussed in the paper[67], which compute the maximum information exchange rate under all the possible transmission strategies. From Theorem 1, the capacity of the two way relay channel is defined as:

$$C = \max_{s \in \{ \text{all possible schemes} \}} \min \{ R_{X,Y}(s), R_{Y,X}(s) \}$$ (2.9)

Based on the paper [67], the up bound of the capacity is given as:

$$C \leq \frac{1}{2} \log_2(1 + \min(SNR_1, SNR_2)) \log_2(1 + SNR_3) + \log_2(1 + SNR_4)$$ (2.10)

Where $\log_2(1 + SNR_i)$ is the Shannon channel capacity for a Gaussian channel with $SNR_i$, detailed proof can be found in the paper [67].

**Multiple access relay channel**

Multiple access relay channel (MARC) is another popular relay network topology drawing research’s interests, where multiple sources communicate with a single destination in the presence of a relay node. This network model is very common in our daily life, for example wireless ad hoc and sensor networks, an intermediate relay node is used to aid communication between several sources and the destination.

The relay initial concept was to step up the spectral efficiency of mobile radio networks by allowing each mobile station to act as a relay for one other mobile station, while the multiple access relay channel is introduced by quantifying the improvement of this concept by the discussion of capacity[23]. Based on the Theorem 1, suppose a set of $M$ source node $G \subset S = \{1, 2, ..., M\}$, the input is defined as $X_G = \{X_i : i \in G\}$, $Y = \{Y_{M+1}, Y_{M+2}\}$ indicates the outputs from the relay and the sources simultaneously, and $G^c$ to be the complement of $G$ in $S$, the up bound of the set of transmission rates is given in [45] as:

$$\sum_{i \in G} R_i \leq \min \left( I(X_G; Y|X_{G^c}, X_{M+1}, U), I(X_G, X_{M+1}; Y_{M+2}|X_{G^c}, U) \right)$$ (2.11)
2.3 Network coding in relay channel

It has been discussed above and recognized that the wireless relay networks represent a fertile ground for devising communication nodes based on network coding, especially particular for applications in two way relay channel and multiple access relay channel.

2.3.1 Network coding in two way relay channel

One suitable application of the network coding arises for the two way relay channels, where two nodes A and B exchange their information with each other assisted by using a relay node in the middle, please refer to figure 2.4.

In traditional wireless communication, it would require four transmission time slots to exchange two packets: Node A to the relay, and the relay to Node B, and vice versa. With network coding technique, on the other hand, Node A and Node B temporarily store their transmitted packets for later decoding.
After two time slots, the relay has received the packets, encodes (e.g. XORs) and broadcasts them back to Node A and Node B within one time slot. Node A and Node B each recover their packets by decoding (e.g. XORing) the received packet with the stored one. The number of transmission time slots reduces to three, one less than in the traditional transmission. From this point of view, the throughput will arise around 25%, while this will sacrifice the performance of the channel as the bit error rate will increase.

2.3.2 Network coding in multiple access relay channel

As discussed above that MARC is based on the relay system where multiple sources (mainly two) use a common relay. In the realization of MARC assisted by network coding, the relay forwards the network-coded message instead of two separate messages received from the two sources, still achieving diversity gain, for example decreasing transmission bit error rate.

Figure 2.5 shows the application of network coding in a decode-and-forward (DF) MARC system, where the relay node forwards the messages $X_A$ and $X_B$ received from two sources, while the modulo-2 summation implements the network coding. In a conventional MARC network, the relay would transmit the decoded messages $\hat{X}_{RA}$ and $\hat{X}_{RB}$ received from A and B by using two orthogonal channels, the destination can either recover $X_A$ by the message received from the direct transmission or the one forwarded by the relay, identical operation will also recover B’s message [58]. Of course this will increase information redundancy, while the paper [21] shows that the redundancy which is contained in the transmission of the relay can be exploited more efficiently with joint network-channel coding.

While in network coding scenario, the relay node would forward the network coded message to the destination, and the destination can use this message with those two messages received from the direct transmission to recover $X_A$ and $X_B$
by implementing the following decode strategy, this strategy can enhance the transmission efficiency and increase the throughput.

As it’s known that the destination will receive three kinds of message $\tilde{X}_A$ and $\tilde{X}_B$ from the sources and $\tilde{X}_R$ from the relay, if there is no channel distortion, there is a relation between the message from the relay and the messages from the sources which is $X_R = X_A \oplus X_B$, so the minimum mean-squared error estimation (MMSE) may apply to recover the corresponding messages, while this estimation operation has already introduced in two way relay channel by [32]. First construct a message block $\tilde{U} = \{\tilde{X}_A, \tilde{X}_B, \tilde{X}_R\}$, there will be four different possible results based on the relation $X_R = X_A \oplus X_B$:

$$\tilde{U} = \{\tilde{X}_A, \tilde{X}_B, \tilde{X}_R\}$$

$$U_1 = \{0, 0, 0\}$$

$$U_2 = \{0, 1, 1\}$$

$$U_3 = \{1, 1, 0\}$$

$$U_4 = \{1, 0, 1\}$$  \hspace{1cm} (2.13)

Suppose $U = \{U_1, U_2, U_3, U_4\}$, so the symbol $k \in [1, 4]$ parallels each message block in $U$, and $\hat{U}$ is the recovered signal in the same structure with $\tilde{U}$, the distortion is computed as:

$$D(k) = E(||\tilde{U} - U_k||^2)$$  \hspace{1cm} (2.14)

Once the expected distortions for all $k$ are computed, the index with the minimum expected distortion can be chosen by the following argument:

$$v = \arg \min_k (D(k))$$  \hspace{1cm} (2.15)

Consider the case that plain earth model[25] is used for each channel in MARC system with transmit power 0 dB, Gaussian additive noise with power
−90 dB is added to each channel, the distances of each two node for this simulation are \( AR = 50 m, BR = 75 m, RD = 100 m, AD = 120 m, BD = 140 m \), please refer to figure for the exact position of each nodes. Then the bit error rate(BER) for message A and message B by using MMSE strategy for network coding MARC is shown in the figure 2.6.
Chapter 3

Lattice is everywhere

3.1 Introduction

Lattices have many significant applications in geometry and mathematics, particularly in connection with number theory, sphere packing and sphere covering, they also arise in applied physics and chemistry in connection with mineralogy and crystallography.

The main application of lattices other than geometry is in engineering, especially in the channel coding problem, i.e. the design of codes for a band-limited channel with white Gaussian noise[7]. While sphere packings also give a way to design optimal codes for band-limited channel, as the theoretical investigation of band-limited channels’ information capacity is equivalent the requirement for the best sphere packings in high dimensions. For the properties of sphere packings in low dimensions, they are frequently used in the design of practical signaling systems, i.e. the Trellis coded modulation schemes.

3.2 Definition of lattice

The lattice has a property that zero vector is a center and if $\mu$ and $\nu$ are centers of spheres, then $\mu + \nu$ and $\mu - \nu$ are also centers of existing spheres, and a center is always called a lattice point. So in general, if $n$ sphere centers $\nu_1, \nu_2, ..., \nu_n$ of an $n$-dimensional lattice are exist, the set of all centers consists of the sums $\Sigma k_i \nu_i$, where $k_i$ are integers, while the set of vectors $\nu_1, \nu_2, ..., \nu_n$ is a basis for the lattice, which means the lattice $\Lambda$ is composed of all integral combinations of the basis vectors.

A lattice fundamental region is defined as a building block which when repeated many times to fill the whole space with just one lattice point in each copy, it is also an example of fundamental parallelotope which consists of the points:

$$\theta_1\nu_1 + \theta_2\nu_2 + ... + \theta_n\nu_n \quad (0 \leq \theta_i < 1)$$

(3.1)
Let the coordinates of the basis vectors be

\[ \nu_1 = (\nu_{11}, \nu_{12}, \ldots, \nu_{1m}) \]
\[ \nu_2 = (\nu_{21}, \nu_{22}, \ldots, \nu_{2m}) \]
\[ \vdots \]
\[ \nu_n = (\nu_{n1}, \nu_{n2}, \ldots, \nu_{nm}) \]  

So the matrix

\[ M = \begin{pmatrix}
\nu_{11} & \nu_{12} & \ldots & \nu_{1m} \\
\nu_{21} & \nu_{22} & \ldots & \nu_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
\nu_{n1} & \nu_{n2} & \ldots & \nu_{nm}
\end{pmatrix} \]  

is called generator matrix, it is also denoted as \( M = [\nu_1 | \nu_2 | \ldots | \nu_n] \), the lattice \( \Lambda \) can also be denoted using generator matrix:

\[ \Lambda = \{ \nu = M \cdot i : i \in \mathbb{Z}^n \}, \mathbb{Z} = \{0, \pm1, \pm2, \ldots\} \]

The fundamental Voronoi region of \( \Lambda \) is defined as

\[ \mathcal{V} = \{ x \in \mathbb{R}^n : \| x \| \leq \| x - \nu \|, \forall \nu \in \Lambda \} \]

where \( \| . \| \) denotes Euclidean norm, and \( \mathbb{R}^n \) shows the Euclidean space, its relation with lattice and fundamental Voronoi region is:

\[ \mathbb{R}^n = \Lambda + \mathcal{V} \]

### 3.3 Sphere packing, covering and kissing number

#### 3.3.1 Sphere packing

The sphere packing solves the problem with how densely a large number of identical spheres can be packed together in n-dimensional space. Figure 3.1 shows an example of one-dimensional sphere and its packing.

Assume a point \( x \) in Euclidean space, which can me simplified as a string of \( n \) real numbers:

\[ x = (x_1, x_2, \ldots, x_n) \]

So the point in a sphere with center \( \nu = (\nu_1, \nu_2, \ldots, \nu_n) \) and radius \( \rho \) should satisfy:

\[ \| x_1 - \nu_1 \|^2 + \| x_2 - \nu_2 \|^2 + \ldots + \| x_n - \nu_n \|^2 = \rho^2 \]
In another point of view, a sphere packing can be specified by the centers \( \nu \) and the radius \( \rho \). Suppose a lattice \( \Lambda \) and Voronoi region \( \mathcal{V} \), with a given radius \( \rho \), a sphere packing can be denoted as the set \( \Lambda + \rho \mathcal{B} \) in Euclidean space, where the lattice is defined as the center of the sphere in section 3.2 and \( \mathcal{B} \) is an unit sphere. The spheres have no intersection areas, which means for any lattice points \( x, y \in \Lambda (x \neq y) \), it should follow the condition:

\[
(x + \rho \mathcal{B})(y + \rho \mathcal{B}) = \emptyset
\]

(3.9)

and based on[12] [64], the packing radius \( \rho^\text{pack}_\Lambda \) of the lattice is defined as:

\[
\rho^\text{pack}_\Lambda = \sup\{\rho : \Lambda + \rho \mathcal{B} \text{ is a packing}\}
\]

(3.10)

where \( \sup\{,\} \) denotes the minimum distance between the sphere’s center to the sphere’s boarder.

Similarly, the “effective radius” of the Voronoi region \( \rho^\text{effec}_\Lambda \) is defined as the radius of a sphere which has the same volume with the sphere with packing radius \( \rho^\text{pack}_\Lambda \). The packing efficiency \( \gamma^\text{pack}(\Lambda) \) is denoted as the ratio between the packing radius and the effective radius:

\[
\gamma^\text{pack}(\Lambda) = \frac{\rho^\text{pack}_\Lambda}{\rho^\text{effec}_\Lambda}
\]

(3.11)

### 3.3.2 Sphere covering

Comparing with sphere packing problem, covering problem tries to find the most economical way to cover n-dimensional Euclidean space with equal overlapping spheres. Similarly, the set \( \Lambda + \rho \mathcal{B} \) is a covering of Euclidean space when:

\[
\mathbb{R}^n \subseteq \Lambda + \rho \mathcal{B}
\]

(3.12)

The covering radius \( \rho^\text{cov}_\Lambda \) of the lattice is defined as:

\[
\rho^\text{cov}_\Lambda = \min\{\rho : \Lambda + \rho \mathcal{B} \text{ is a covering}\}
\]

(3.13)

where \( \min\{,\} \) shows the minimum radius to cover the sphere, which is also the maximum distance between the center to the sphere’s boarder. And the covering efficiency \( \gamma^\text{cov}(\Lambda) \) can be defined as:

\[
\gamma^\text{cov}(\Lambda) = \frac{\rho^\text{cov}_\Lambda}{\rho^\text{effec}_\Lambda}
\]

(3.14)

Comparing with the packing efficiency which should be no more than 1, while the covering efficiency should be no less than 1[41], while the optimized solution is to find both \( \gamma^\text{pack}(\Lambda) \) and \( \gamma^\text{cov}(\Lambda) \) is equal 1[40], which is:

\[
\{\gamma^\text{pack}(\Lambda)\}_\text{optimized} = \{\gamma^\text{cov}(\Lambda)\}_\text{optimized} = 1
\]

(3.15)

Thus the covering problem consists of finding the most efficient covering spheres of \( n \)-dimensional space, in figure 3.2, it gives two different covering arrangements, the square lattice and the hexagonal lattice. Assume the distance
between each two spheres' centers in these two arrangements are 2 for both, the effective radius are the same, while for covering radius, the square lattice has $\rho_{cov}^{sq} = \sqrt{2}$, and the covering radius for the hexagonal lattice is $\rho_{cov}^{hex} = \frac{2\sqrt{3}}{3}$, so the covering efficiency can be calculated as:

$$\gamma_{cov}(\Lambda_{sq}) = \frac{\rho_{cov}^{sq}}{\rho_{eff}^{sq}} = \frac{\sqrt{2}}{\frac{\sqrt{2}}{2}} = 1$$

$$\gamma_{cov}(\Lambda_{hex}) = \frac{\rho_{cov}^{hex}}{\rho_{eff}^{hex}} = \frac{\frac{2\sqrt{3}}{3}}{\frac{\sqrt{3}}{3}} = \frac{2\sqrt{3}}{\sqrt{3}} = 2 > 1$$

So the second covering with hexagonal lattice is more efficient, as $\gamma_{cov}(\Lambda_{sq}) < \gamma_{cov}(\Lambda_{hex})$, and the spheres don’t overlap as much as in the first one (the square lattice).

### 3.3.3 Kissing number

The associated object for the kissing number, comparing with sphere packing and covering, is to find out how many spheres touch another sphere, this number is denoted as the kissing number $\tau$, for a lattice packing, $\tau$ is the same for every sphere.

It is proved that the hexagonal packing is indeed an optimal sphere packing for 2-dimensional space [20], so it is obvious that hexagonal packing of equal-sized disks (2-dimensional circle) in the plane is the optimal lattice packing [34] with kissing number $k(2) = 6$, please refer the figure 3.3.

### 3.4 Quantizers

Suppose there are $M$ points $P_1, P_2, \ldots, P_M$ in Euclidean space $\mathbb{R}^n$, the input $x$ is an arbitrary point of $\mathbb{R}^n$, after the quantizer the output $y$ chooses the nearest $P_i$, please refer the figure. So the quantizers should be designed to minimize
Figure 3.3: The perfect kissing arrangement for $n = 2$, it is easy to prove that in two dimensions the kissing number is 6.

Figure 3.4: The output of the quantizer chooses the nearest center.

mean square error (MMSE), i.e. the average of $\|x - P_i\|^2$.

Assume the nearest neighbor quantizer is $Q_\Lambda(.)$, for the quantization, it has the definition:

$$Q_\Lambda(x) = y, y \in \Lambda, \text{if } \|x - y\| \leq \|x - z\|, \forall z \in \Lambda$$ (3.18)

If $x$ is uniformly distributed in $\mathbb{R}^n$, the lattice quantizer problem is to find $n$-dimensional $\Lambda$ to minimize the normalized second moment of the lattice that are congruent to its Voronoi regions [7]. Based on [64], the second moment $\sigma^2_\Lambda$ of the lattice $\Lambda$ is defined as the second moment per dimension of a uniform distribution over the fundamental Voronoi region $V$:

$$\sigma^2_\Lambda = \frac{1}{Vol(V)} \cdot \frac{1}{n} \int_V \|x\|^2 dx$$ (3.19)

where $Vol(V)$ indicates the volume of the fundamental Voronoi region, let $V = Vol(V)$, the normalized second moment $G(\Lambda)$ is given as:

$$G(\Lambda) = \frac{\sigma^2}{V^\frac{2}{n}}$$ (3.20)

Suppose $G_n$ denotes the minimum possible value of $G(\Lambda_n)$ over all lattices in the Euclidean $\mathbb{R}^n$, which is also the solution of minimizing the normalized
second moment in equation 3.20. While $G^*_n$, the normalized second moment of a sphere, approaches $\frac{1}{2\pi}$ as the dimension $n$ goes to infinity[64], it also gives that $G_n > G^*_n > \frac{1}{2\pi e}$ for all $n$. The paper [65] indicates the quantization noise of a lattice achieving $G_n$ is ”white”, and it also shows that

$$\lim_{n \to \infty} G_n = \frac{1}{2\pi e} \quad (3.21)$$

3.5 Gaussian channel coding

The additive white Gaussian noise (AWGN) channel can be denoted by using the relation between the input $X$ and output $Y$:

$$Y = X + Z \quad (3.22)$$

where $Z$ is i.i.d. Gaussian noise with $N(0, \sigma^2)$, and define the ”effective radius” of the noise is given as:

$$\rho_N = \sqrt{nP_N} \quad (3.23)$$

where $P_N$ is the power of the noise.

The reason that lattice codes were introduced to AWGN channel is due to the codes can AWGN channel’s capacity[55]. The lattice version of Gaussian channel coding problem is to find an $n$-dimensional lattice that minimizes the error probability $P_e$, while this coding problem was first considered by Poltyrev[36] for unconstrained AWGN channel, so in this point of view any lattice point can be transmitted with infinite power and transmission rate. For a given lattice, the role of decoder is try to find the nearest lattice point to the received signal, so the error probability $P_e$ is the probability that the decoder chooses the wrong lattice point or the probability that the noise leaves the Voronoi region of the transmitted lattice point[64]:

$$P_e(\Lambda, \rho_N) = Pr\{N \notin V\} \quad (3.24)$$

From the above definition, it is clear that the probability of decoding error can be subjected to the ratio of the radius of the Voronoi region and the “effective radius” of the noise, based on [12] this ratio can be defined as:

$$\gamma_{AWGN}(\Lambda, \rho_N) = \frac{\rho^{effec}_\Lambda}{\rho_N} \quad (3.25)$$

where $\rho^{effec}_\Lambda$ is the ”effective radius” of the noise, it is given in [65] which yields:

$$\rho^{effec}_\Lambda = \sqrt{(n + 2)G^*_n(Vol(V))^\frac{1}{2}} \quad (3.26)$$

substitute it to 3.25, with the result in 3.21, then the equation 3.25 can be stated as:

$$\gamma_{AWGN}(\Lambda, \rho_N) = \frac{\rho^{effec}_\Lambda}{\rho_N} = \frac{(Vol(V))^\frac{1}{2}}{2\pi e P_N} + \Delta \quad (3.27)$$

where $\lim_{n \to \infty} \Delta = 0$, so the problem to minimize the decoding error probability can be simplified to minimize the radius ratio $\gamma_{AWGN}(\Lambda, \rho_N)$. 

3.6 Conclusion

From the discussion in the above sections, we can summarize the four problems that involve lattices in the following list.

1. Sphere packing, maximizing the packing radius $\rho_\Lambda^{\text{pack}}$ of Voronoi region, the best bound is given by Minkowski [27].

2. Sphere covering, minimizing the covering radius $\rho_\Lambda^{\text{cov}}$ of Voronoi region, and it is the Rogers bound [39].

3. Quantizing, minimizing the normalized second moment $G(\Lambda)$.

4. Gaussian channel coding, minimizing the decoding error probability $P_e$, which is also applied for the ratio of the Voronoi region radius and the noise ”effective radius” $\gamma_{\text{AWGN}}(\Lambda, \rho_N)$, and the bound for unconstrained AWGN channel is given in [36] and [65], for the channel with restrictions on power and transmission rate, the bound is given by Shannon in [47] and [48].

It is well known that a lattice is good when it is close to the Voronoi region’s sphere no matter in which criteria listed above, which is also approved that a lattice is optimal in all four senses[12].
Chapter 4

Lattice encoder and decoder

4.1 Introduction

Shannon theory suggests the fundamental limits of data compression and reliable communication [56], the goal of each encoding and decoding for the additive white Gaussian noise (AWGN) channel is to find the codes whose transmission rates can approach the channel capacity [17] [5],

\[ C = \frac{1}{2} \log(1 + SNR) \quad (4.1) \]

Where \( SNR = \frac{P_X}{P_N} \) is the signal-to-noise ratio. Shannon’s work had indicated that there must be sphere packings in spaces of high dimension \( n \) with sufficiently high density to approach channel capacity [16]. A concept called groups codes for AWGN channel was considered in [50], where the codewords lie on the surface of the sphere with radius \( \sqrt{n}P_X \).

Under Shannon theory, the codewords of a good code should look like realizations of a zero-mean independent and identically distributed (i.i.d.) Gaussian source with power \( P_X \) [14]. Based on this conclusion, the applications of lattices for the AWGN channel was first discussed by de Buda in his paper [11] and corrected theorem proof in [54]. de Buda’s paper demonstrated that optimal codes need not be random, but rather that some of them have structures, e.g., the structure of a lattice code.

In addition to offering structure, achieving capacity, and reducing complexity, lattice codes are desirable because they are analogous with linear codes in Euclidean space. Many researches have been sortie into constructing block [16] and trellis codes [4] for AWGN channel by using lattice, inspired by LDPC codes, low density lattice codes (LDLC) were proposed in [51]. Lattice codes are also proved powerful in multi-user systems, it was shown that lattice can achieve the capacity of AWGN broadcast channel [62] and the capacity of AWGN multiple access channel [29].
When a lattice code is defined in this manner, the optimality of the coding is relying on the maximum likelihood (ML) decoding, a frequently used decoding scheme for lattice codes, which requires to find the nearest lattice point inside the sphere to the received signal. While in the paper [2], the authors use lattice decoding to find the closest lattice point, ignoring the boundary of the code, which preserves the lattice symmetry in the decoding process and saves complexity.

4.2 Definitions

Based on the definition of lattice in section 3.2, the lattice can be defined as:

$$\Lambda = \{v = \lambda \cdot M | \lambda \in \mathbb{Z}^n \}$$  \hfill (4.2)

where $M$ is its generator matrix which is defined in equation 3.3 as:

$$M = \begin{pmatrix}
\nu_{11} & \nu_{12} & \ldots & \nu_{1m} \\
\nu_{21} & \nu_{22} & \ldots & \nu_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
\nu_{n1} & \nu_{n2} & \ldots & \nu_{nm}
\end{pmatrix}$$ \hfill (4.3)

while the Gram matrix is defined as $G = MM^T$ for the lattice, where $T$ denotes transposition. As the generator matrix contains the basis vectors $\{\nu_i\}_{i=1}^n$ of the lattice, the $(i,j)$th entry of $G$ is the inner product $\langle \nu_i, \nu_j \rangle = \nu_i \cdot \nu_j^T$.

The determinant of the lattice $\Lambda$ is defined to be the determinant of the Gram matrix $G$

$$\det(\Lambda) = \det(G)$$ \hfill (4.4)

For full-rank lattices, i.e. $m = n$, where the generator matrix $M$ is a square matrix, and then the determinant of $\Lambda$ is

$$\det(\Lambda) = (\det(M))^2$$ \hfill (4.5)

For full-rank lattices, the square root of the determinant is the volume of the fundamental parallelepiped or Voronoi region $V$, also called volume of the lattice, which is denoted as $\text{vol}(\Lambda)$.

Define a lattice $\nu \in \Lambda$, and $r$ is in the fundamental Voronoi region $r \in V$, for every $x \in \mathbb{Z}^n$, it can be uniquely written as:

$$x = \nu + r$$ \hfill (4.6)

Then $\nu$ can be the nearest neighbor of $x$ in $\Lambda$ with $\nu = Q_V(x)$, and $r = x \mod \Lambda$ is the apparent error $x - Q_V(x)$. A lattice $\Lambda$ has many possible basic Voronoi cells, it is common to use the notation $x \mod \Lambda$ for the modulo lattice operation.

Refer to the equation 3.18, the nearest neighbor quantizer associated with any fundamental Voronoi region $V$ of $\Lambda$ is defined as

$$Q_V(x) = \nu, \text{ if } x \in \nu + V$$ \hfill (4.7)
Clearly the quantization error $Q_V(x) - x$ depends on the source vector $x$, in fact, it is a deterministic function of it.

If one lattice can be obtained from another by a rotation, reflection and change of scale, then these two lattices are equivalent[30]. There are two famous lattices which are useful for our later discussion, one is integer lattices $\mathbb{Z}^n$, and the other is lattices $A_n$.

### 4.2.1 Integer lattices $\mathbb{Z}^n$

A lattice $\Lambda$ is called an integer lattice if its Gram matrix has coefficients in $\mathbb{Z}$, where $\mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$, this is also the simplest lattice which can be denoted as:

$$\mathbb{Z}^n = \{(\nu_1, \nu_2, \cdots, \nu_n), \nu_i \in \mathbb{Z}\} \quad (4.8)$$

For integer lattice, both generator matrix and Gram matrix are the identity matrix.

### 4.2.2 Lattices $A_n$

Two dimensional lattice of this type is called hexagonal lattice, for general dimensions, the definition is given as:

$$A_n = \{(\nu_0, \nu_1, \cdots, \nu_n) \in \mathbb{Z}^{n+1}, \sum_{i=0}^{n} = 0\} \quad (4.9)$$

the Gram matrix is given as:

$$G = \begin{pmatrix}
2 & -1 & 0 & \cdots & 0 \\
-1 & 2 & -1 & \cdots & 0 \\
0 & -1 & 2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 2
\end{pmatrix} \quad (4.10)$$

### 4.3 Dither

Dither is an intentional noise which is added to the source before the quantization and subtracted after the quantization in dithered lattice quantization[63]. If the dither is uniform over the basic Voronoi region of the lattice $\Lambda$, then the resulting error is independent of the source, the term “dithering” refers to intentional randomization aimed to improve the perceptual effect of the lattice quantization[19].

Although it seems strange that adding noise can improve quantization performance, while one has to admit that dither is independent of the source statistics to guarantee a desired distortion level[68]. Here below a general example will be given to better understand why adding noise can improve performance, which is also introduced in [63].

Suppose a team sitting around a meeting table to compute the sum of the ages of all team members, but without revealing the individual age of any one of
them. The team leader decides to play one trick that he would write a random number $U$ on a paper and pass it on to his right neighbor, who will add to it his/her own age and pass the new number to his/her right neighbor, and so on. When the last team member has finished his/her age addition and pass the number to the team leader, he/she will add his/her own age and subtract $U$ to get the sum of the ages. So if the number $U$ is large enough, the individual age is kept secret, the number each team member gets is statistically independent of the individual ages of the preceding team member.

The idea of quantization with dither is similar with the example above, suppose the dither is $U$, the reconstruction of the message $x$ after encoder and decoder is given by

$$\hat{x} = Q_{V}(x + U) - U$$  \hspace{1cm} (4.11)

It indicates that the encoder adds the dither $U$ to the message before the quantization, while the decoder subtracts the dither from the associated lattice point. The quantization error with dither becomes

$$e_Q = \hat{x} - x = Q_{V}(x + U) - U - x$$  \hspace{1cm} (4.12)

or follow the definition in section 4.2, the quantization error is equivalently $-((x + U) \mod \Lambda)$.

As it is shown in [65] and [66] that a uniform distribution over its support Voronoi region $V$ remains uniform after the mapping from $V$ to another Voronoi region, so if the dither $U$ is uniformly distributed over the fundamental Voronoi region $V$, then the quantization error with dither in equation 4.12 is uniform over $-V$ (the reflection of $V$), which is stated below

$$[Q_{V}(x + U) - U - x] \sim \text{Unif}(-V), \text{ for any } x$$  \hspace{1cm} (4.13)

Based on the paper [63], the MSE of the dithered quantizer is equal to the second moment of the dither which can be derived as

$$\frac{1}{n} E \| \hat{x} - x \|^2$$  \hspace{1cm} (4.14)

Under the fundamental Voronoi region $V$, the second moment of the dither becomes the lattice second moment which is given in equation 3.19 as

$$\frac{1}{n} E \| U \|^2 = \frac{1}{Vol(V)} \cdot \frac{1}{n} \int_{V} ||x||^2 dx = \sigma_{\Lambda}^2$$  \hspace{1cm} (4.15)

### 4.4 Lattice encoding

As we all know that the binary bits are the basic transmitted messages all over the wireless transmission channel, based on the discussion above, the main concern about lattice encoding is how to find a way to map binary bits into lattice points.

Based on the information theory, Gaussian sources and channels should be encoded using “Gaussian codebooks”, which means the codewords should be
selected from a Gaussian generating distribution\cite{64}. The resulting codebook in $\mathbb{Z}^n$ ($n$ being the code dimension) has a spherical shape (refer to figure 4.1), with roughly evenly spaced points (lattice points) as codewords.

The codebook was unbounded lattice and not shaped to fit the source variance, while the lack of shaping is compensated for by entropy coding\cite{64}. In details, the lattice points which fall inside the spherical source region will get a shorter binary representation and dominate the coding rate, and on the other hand, those points other side the spherical region will have negligible contributions.

One lattice codebook can be constructed based on the figure 4.1, those codewords are lattice points in $\mathbb{Z}^n$ and have a lattice structure, which have roots in De Buda’s spherical lattice codes. In another way of speaking, each lattice points can be presented in numeric values at each dimension, so collecting all the values in each dimension can be used to construct a binary sequence. In this way all the binary sequence can be mapped into lattice points.

### 4.5 Lattice decoding

A lattice decoder is simply an Euclidean quantizer, or more generally, a quantizer with respect to a fundamental region $\mathcal{V}$ \cite{14}, which means the decoder quantizes the received vector to obtain hypothesized codeword, and solves the closest lattice point problem. As lattice decoder is to find the closest lattice point to a given received point, it is also called sphere decoder.

The closest lattice point algorithm was first present in \cite{35}, then the Finke-Pohst algorithm was given in \cite{15} to enumerate all lattice points within a sphere centered at the origin, a recent work \cite{10} shows its relation to other well-known detection algorithms. The key idea which makes the sphere decoder efficient is that the number of lattice points which are found inside a sphere is significantly smaller than the number of points within a hypercube containing the hyper-
Figure 4.2: Lattice code based scheme showing the transmitted signal(circle) and the decoded signal(dots)

sphere as the dimension of the space grows[30].

The lattice decoding algorithm searches all the points in lattice $\Lambda$ which are found inside a sphere of given covering radius $\rho$, refer to figure 4.2 where all the black dots present the lattice points while the dot marked with a circle, the decoded signal chooses the lattice point which is the nearest to it. This guarantees that only the lattice points within the squared distance $\rho^2$ from the received point are considered in the metric minimization.

The lattice decoding algorithm can be described in the following key steps:

1. Map the received signals into points $r$ in $\mathbb{Z}^n$ which is showing in figure 4.2
2. Choose the lattice $\Lambda = \{v = \lambda \cdot M | \lambda \in \mathbb{Z}^n\}$
3. The quantization function is defined as: $Q(v) = \|v\|^2 = vv^T = \lambda G \lambda^T$, where $G = MM^T$ is the Gram matrix.
4. Find all points in the sphere with squared distance $\rho^2$, which can be done by solving the inequality $Q(v) \leq \rho^2$.
5. Choose the lattice point $v$ minimizing $\|r - v\|^2$

The problem of the lattice decoding is to solve the following equation:

$$\min_{v \in \Lambda} \|r - v\|^2 = \min_{w \not\in r - \Lambda} \|w\|^2$$

(4.16)

The problem changes to search the shortest vector $w$ in the transferred lattice $r - \Lambda$ in the $n$-dimensional Euclidean space $\mathbb{Z}^n$.

The decoding error is denoted as

$$D_e(v) = \|r - v\|^2$$

(4.17)
Since the decoding algorithm is designed in this manner, the decoding error probability for any codeword is given by:

\[ P_e = \Pr(D_e(v) \neq \min_{w \in r-A} ||w||^2) \]  \hspace{1cm} (4.18)

The advantage of this sphere decoding algorithm is that we vectors with a norm greater than the given radius will never be tested, as every tested vector requires the computation of its norm, which will increase number of operations largely.

In order to be sure to always find a lattice point inside the sphere that the received point centered, the selected sphere’s radius should be equal to the covering radius of the lattice.
Chapter 5

Nested lattice codes and modulo-lattice additive noise channel

5.1 Introduction

Structure codes gives a way for the network information theory with simple point-to-point communication techniques while very high complexity, but current applications for example binning scheme[9] are only for noiseless channel network coding until the creation of the idea for nested codes[62].

A binning scheme divides a set of codewords into subsets, so in each subsets the codewords are far apart as possible. Although the binning scheme has structure which is shown in [62], the source coding is always lossy in general applications. The idea of nested codes, no matter nested linear codes [31] for discrete cases or nested lattices [61] for the continuous case, is came out for extending the idea of coset-code binning to noisy channel network coding applications. Nested linear/lattice code are useful because in many communication problems, specially multi-terminal systems, such codes can bring large advantages in average performance compared to random codes[24].

In general, the idea of nested codes is to generate diluted version of the original coset code[62], refer to figure 5.1. Nested lattices are also applied as a unifying model for some classical point-to-point coding techniques, for example like constellation shaping for the additive white Gaussian noise (AWGN) channel, and combined shaping and precoding for the intersymbol interference(ISI) channel[62].

5.2 Modulo-Lattice additive noise channel

It was shown in the previous chapters that the use of linear/lattice are particularly well suited for additive noise channels. The advantages by using lattice in multiple-access relay channel are shown in figure 5.2, in traditional wireless
communication for multiple-access relay channel, it would require transmission time slots to transmit two packets from source nodes A and B to the destination node D. While by using normal network coding which is implemented in chapter 2, this strategy reduces the number of time slots to 3, and it can enhance the transmission efficiency and increase the throughput at the same time. The number of time slots can be decreased to 2 when lattice is using for this type of channel, although there will be some information loss comparing with the traditional multiple-access relay channel communication, the transmission rate will be doubled.

How to implement lattice in multiple-access relay channel became a research interest until the paper [13] derived a transformation technique that transforms a power-constrained multiple-access channel (MAC) into an modulo-lattice additive noise (MLAN) channel, at the price of some information loss. For some channels, the rate increase offered by lattice coding overcome the information loss during the transformation process, and more over, for a “good” lattice, the information loss goes to zero as the dimension of the lattice goes to infinity.

It is given in the previous paragraph that the transmission time only consists of two time slots for lattice implementation in MARC, refer to figure 5.2, which is an advantage with the use of modulo-lattice transmission. During the first time slot, node A and B broadcast their information to the relay and the destination. The destination receives two different signals from node A and B, it decodes a linear combination of these two messages which can be presented as $t_A v_A + t_B v_B$. 
where $t_A$ and $t_B \in \mathbb{Z}$ are the lattice coefficients and $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. The relay also receives two signals from node A and B which can be decoded as a linear combination of them given as $\nu_R = k_A \nu_A + k_B \nu_B$, where $k_A$ and $k_B \in \mathbb{Z}$ are the lattice coefficients. During the second time slot, the relay performs network coding scheme using modulo lattices and transmits a new message $x_R$ to the destination. Finally, the destination recovers the two messages using the two equations decoded during the two time slots.

5.3 Nested lattices for shaping and coding used in multiple access relay channel

It is possible to achieve capacity using linear codes instead of a code drawn at random for MLAN channel using a nested lattice code[14], the coarse lattice is used for shaping so it is a good quantizer, and the fine lattice defines the codewords so it is a good channel code.

Let $\Lambda$ be a coarse lattice, and $\mathcal{V}$ is the fundamental Voronoi region of the fine lattice $\Lambda_f$, $\nu_i$ is a fine lattice point where $i \in \{A, B, R\}$. $U_i$ is the dither signal, and it is a random independent variable uniformly distributed over the fundamental Voronoi region $\mathcal{V}_i$, we assume that the dither $U_i$ is known to transmitter $i$ as well as to the receiver (the relay node or the destination node in MARC).
Hence the transmitted vector $x_i$, $i \in \{A, B, R\}$ can be written as:

$$x_i = (\nu_i + U_i) \mod \Lambda \quad (5.1)$$

The received signal at the relay node is transformed by using modulo lattice additive channel operation, which can be given as:

$$y_R = (\alpha_R(h_{AR}x_A + h_{BR}x_B + z_R) + k_AU_A + k_BU_B) \mod \Lambda \quad (5.2)$$

where $\alpha_R$ is the minimum mean square error factor that minimizes the effective noise\[14], $h_{AR}$ and $h_{BR}$ denote the sources-relay channel gains.

For the received signal $y_R$, it can be simplified as:

$$y_R = (\alpha_R(h_{AR}x_A + h_{BR}x_B + z_R) - k_AU_A - k_BU_B) \mod \Lambda$$

$$= (\alpha_R(h_{AR}x_A + h_{BR}x_B + z_R) - k_Ax_A - k_Bx_B + k_Ax_A + k_Bx_B - k_AU_A - k_BU_B) \mod \Lambda$$

$$= (h_{AR}x_A(\alpha_R - \frac{k_A}{h_{AR}}) + h_{BR}x_B(\alpha_R - \frac{k_B}{h_{BR}}) + \alpha_Rz_R)$$

$$+ k_Ax_A + k_Bx_B - k_AU_A - k_BU_B) \mod \Lambda$$

$$= (z_{R,eff} + k_A(\nu_A + U_A) \mod \Lambda + k_B(\nu_B + U_B) \mod \Lambda$$

$$= (z_{R,eff} + k_A\nu_A + k_B\nu_B) \mod \Lambda$$

$$= (\nu_R + z_{R,eff}) \mod \Lambda \quad (5.3)$$

where $z_{R,eff} = h_{AR}x_A(\alpha_R - \frac{k_A}{h_{AR}}) + h_{BR}x_B(\alpha_R - \frac{k_B}{h_{BR}}) + \alpha_Rz_R$ is the effective noise and $\nu_R = (k_A\nu_A + k_B\nu_B)$ is a lattice point\[57].

The optimal value of $\alpha_R$ is calculated by minimizing the power of the effective noise which is given as:

$$\alpha_R = \arg \min_{\alpha_R} E\{|z_{R,eff}|^2\}$$

$$= \arg \min_{\alpha_R} E\{|\alpha_R|^2 + |h_{AR}|^2x_A^2(\alpha_R - \frac{k_A}{h_{AR}})^2 + |h_{BR}|^2x_B^2(\alpha_R - \frac{k_B}{h_{BR}})^2\}$$

$$= \arg \min_{\alpha_R} E\{|\alpha_R|^2P_N + |h_{AR}|^2P_A(\alpha_R - \frac{k_A}{h_{AR}})^2 + |h_{BR}|^2P_B(\alpha_R - \frac{k_B}{h_{BR}})^2\} \quad (5.4)$$

where $P_N$, $P_A$ and $P_B$ are the power of noise and signals from node A and B respectively, thus, to minimize the above argument, the optimal value of $\alpha_R$ can be derived as:

$$\alpha_R = \frac{|h_{AR}|^2P_Ak_A + |h_{BR}|^2P_Bk_B}{|h_{AR}|^2P_A + |h_{BR}|^2P_B + P_N} \quad (5.5)$$

Thus, $\nu_R$ can be decoded at the relay node, and it transmits $x_R$ to the destination as:

$$x_R = (\nu_R + U_R) \mod \Lambda \quad (5.6)$$
During the first time slot, the destination node also receives two different messages directly from node A and B, where the received signal $y_{D1}$ is given as:

$$y_{D1} = (\alpha_{D1}(h_{AD}x_A + h_{BD}x_B + z_{D1}) + t_AU_A + t_BU_B) \mod \Lambda \quad (5.7)$$

Based on the equation 5.3, the above equation can be simplified as:

$$y_{D1} = (t_A\nu_A + t_B\nu_B + z_{D_{eff}1}) \mod \Lambda \quad (5.8)$$

where the effective noise $z_{D_{eff}1} = h_{AD}x_A(\alpha_{D1} - \frac{k_A}{h_{AD}}) + h_{BD}x_B(\alpha_{D1} - \frac{k_B}{h_{BD}}) + \alpha_{D1}z_{D1}$, to minimize the power of the effective noise, we can derive the optimal value of $\alpha_{D1}$ which is given as:

$$\alpha_{D1} = \frac{|h_{AD}|^2P_A\frac{k_A}{h_{AD}} + |h_{BD}|^2P_B\frac{k_B}{h_{BD}}}{|h_{AD}|^2P_A + |h_{BD}|^2P_B + P_N} \quad (5.9)$$

During the second time slot, the signal received at the destination node from the relay after the modulo-lattice transformation is given as:

$$y_{D2} = (\alpha_{D2}(h_{RD}x_R + z_{D2}) + mRU_R) \mod \Lambda \quad (5.10)$$

Based on the equation 5.3, the above equation can be simplified as:

$$y_{D2} = (mR\nu_R + z_{D_{eff}2}) \mod \Lambda \quad (5.11)$$

where $z_{D_{eff}2} = (h_{RD}x_R(\alpha_{D2} - \frac{m_R}{h_{RD}}) + \alpha_{D2}z_{D2})$ is the effective noise, minimize it we can get the optimal value of $\alpha_{D2}$ as:

$$\alpha_{D2} = \frac{|h_{RD}|^2P_R\frac{m_R}{h_{RD}}}{|h_{RD}|^2P_R + P_N} \quad (5.12)$$

Hence, the destination node can recover the two messages $\hat{\nu}_A$ and $\hat{\nu}_B$ sent by the source nodes A and B, by decoding the received signals $y_{D1}$ and $y_{D2}$.

### 5.4 Conclusion and discussion

It has been demonstrated in [14] that using nested lattice codes for modulo additive noise channel, lattice codes can achieve capacity. While this transformation of the original power-constrained channel into a modulo additive noise channel is largely increase the throughput to achieve the Shannon capacity, it is not strictly information lossless. As the error exponent is lower-bounded by the Poltyrev exponent for coding in power unconstrained AWGN channel[37].

A nested lattice code is a lattice code whose boundary region is the Voronoi region of a sublattice. The coarse lattice(shaping lattice) is defining the MLAN channel, this lattice is selected to meet the condition that its normalized second moment approaches that of a sphere when the dimension goes to infinity, which is given as $\frac{1}{2^n}$. While the fine lattice should be good for channel coding, it should approach the Poltyrev exponent to decrease the error exponent. The selection of a “good” nested lattices pair(coarse lattice and fine lattice) is very
important for the final coding and decoding performance.

The lattice coefficients \((t_A, t_B, k_A, k_B)\) must be integers and the matrix \(M\) should be full row rank which is given as:

\[
M = \begin{pmatrix}
  t_A & t_B \\
  k_A & k_B
\end{pmatrix}
\]  

(5.13)

One way to get the lattice coefficients is introduced in [52] where a rounding method for the channel coefficients is using to assign their values to the lattice coefficients.
Chapter 6

Conclusion and future work

6.1 Conclusion

This work focuses on the application of interference management by using nested lattice coding for multiple access relay channel, which can be a potential technique to be considered in future radio access technology, for example LTE-Advanced. However, there are other contexts where this technique proves to be useful in LTE-Advanced. In the last chapter, we give a brief summation of the following topics which have already been discussed in the previous chapters:

1. Network coding for relay networks
2. Modulo lattice additive noise channel
3. Nested lattice gives a way for interference management.

Among these three topics, nested lattice coding in modulo lattice additive noise channel is the most promising area for further research work. Recently, network coding in relay networks has been shown to improve the system's performance.

6.1.1 Relay networks with network coding

Relay networks constitute a promising concept to improve system performance, to mitigate coverage holes, to provide a uniform user experience over the entire cell area and last but not least to satisfy high traffic demands in hot-zones. Besides those advantages that relay brings to the network system, the idea of bring relay nodes in next mobile network will have a huge innovation for the network structure.

Network coding in relay network application is already proved to increase the throughput, improve system’s efficiency and enhance system capacity in many papers. In this master thesis, we implement the scenario that using network coding in multiple access relay channel, in which we can see that the network coding theory does bring large advantages no matter to improve throughput, but also to largely decrease the bit error rate from previous transmission schemes.
6.1.2 Modulo lattice additive noise channel

Although the network coding in multiple access relay channel has already shown its improvements, the separation of message transmissions from different source nodes in current applications still has large potential to improve the transmission scheme with acceptable system performance. Lattice coding has lightened the sky for this application.

The idea of bring lattice in multiple access relay channel derive a transformation technique that transforms a power-constrained multiple-access channel into an modulo-lattice additive noise (MLAN) channel, at the price of some information loss. While for some channels, the rate increase offered by lattice coding overcome the information loss during the transformation process, and more over, for a "good" lattice, the information loss goes to zero as the dimension of the lattice goes to infinity.

In this work, we proposed the technique of modulo lattice additive noise channel for multiple access relay network. Comparing with the transmission time by using network coding in multiple access relay channel, which requires 3 time slots to transmit messages from both two source nodes to the destination node, while only consists of two time slots for lattice implementation, where increases 33\% of the total throughput.

6.1.3 Nested lattice gives a way for interference management

Interference management shows its importance in current live mobile network, as every transmitter causes a degree of interference into every receiver, no matter it is the base station or the mobile phone. From the previous research and live radio access technologies(e.g. GSM, WCDMA etc.), interference is attenuated by the separation between the transmitter and receiver in the geographic(e.g. GSM) and frequency domains(e.g. WCDMA).

In this master thesis, we have presented the paradigm that using nested lattice code in modulo lattice additive noise channel. The coarse lattice(shaping lattice) is defining the MLAN channel, this lattice is selected to meet the condition that its normalized second moment approaches that of a sphere when the dimension goes to infinity, which is given as $\frac{1}{27\pi^2}$. While the fine lattice should be good for channel coding, it should approach the Poltyrev exponent to decrease the error exponent. The selection of a “good” nested lattices pair(coarse lattice and fine lattice) is very important for the final coding and decoding performance.

6.2 Future work

We should note that this work is an intermediate step toward more general relay networks, for example multiple relays. It is proved in many applications that the relay nodes will bring large advantages to the network, while most current researches are focused on one relay node network, how to coordinate with other nodes if more relay nodes are introduced for the system attracts more interests
in the future research.

As in mobile network, unlike the fixed cable transmission, fading is always an interesting topic to be discussed. Upon this point, we haven’t discussed the situation when fading is introduced in multiple access relay channel with lattice coding, as fading will largely influence the channel capacity which has already given for relay channel in chapter 2.

MIMO(multiple-input and multiple-output) technology has implemented in modern wireless communications standards such as LTE, WiMAX and HSPA+, due to its significant improvement in data throughput and spectral efficiency. Because of these properties, MIMO is an important part to be introduced in future mobile communication standards, to consider MIMO using lattice will attract more attentions in the future researches.
Bibliography


length in a lattice, including a complexity analysis. Mathematics of Com-

[16] G David Forney. Coset codes-part i : Introduction and geometrical classi-

[17] G David Forney and Gottfried Ungerboeck. Modulation and coding for
linear gaussian channels. IEEE Transactions on Information Theory,

for wireless relay channels. In the 41st Allerton Conference on Communi-
cation, Control, and Computer, Monticello,USA, October 1’C3, 2003.


[21] Christoph Hausl and Philippe Dupraz. Joint network-channel coding for
the multiple-access relay channel. 3rd Annual Conf on Sensor and Ad Hoc

[22] Network World Jim Duffy. Network coding: networking’s next rev-

[23] G Kramer and A J Van Wijngaarden. On the white gaussian multiple-
access relay channel. 2000 IEEE International Symposium on Information
Theory Cat No00CH37060, 25:40, 2000.


[25] S B Slimane Lars Ahlin and Jens Zander. Principles of Wireless Commu-

[26] Lars Lindbom, Robert Love, Sandeep Krishnamurthy, Chunhai Yao,
Nobuhiko Miki, and Vikram Chandrasekhar. Enhanced inter-cell interfer-
ence coordination for heterogeneous networks in lte-advanced : A survey.
Access, pages 1–18, 2011.


[28] Global mobile Suppliers Association. 49 commercial lte networks confirmed
by gsa, January 2012.


[30] Frédérique Oggier and Emanuele Viterbo. Algebraic number theory and
code design for rayleigh fading channels. Foundations and Trends in Com-


