Distributed Cooperative Spatial Multiplexing System

Doctor of Philosophy

Nian Xie

University of York
Department of Electronics

August 2014

Abstract

Multiple-Input-Multiple-Output (MIMO) spatial multiplexing systems can increase the spectral efficiency manyfold, without extra bandwidth or transmit power, however these advantages are based on the assumption that channels between different transmit antenna and receive antenna are independent which requires the elements in antenna array be separated by several wavelengths. For small mobile devices, the requirement is difficult to implement in practice. Cooperative spatial multiplexing (C-SM) system provides a solution: it organizes antennas on distributed mobile stations to form a virtual antenna array (VAA) to support spatial multiplexing.

In this thesis, we propose a novel C-SM system design which includes a transmitter with two antennas, a single antenna receiver and a relay group with two single antenna relays. In this design, we assume that the transmitter tries to transmit two coded independent messages to the receiver simultaneously but the transmitter-receiver link is too weak to support the transmission. Thus a relay group is introduced to help with the transmission. After relays receive the messages from the transmitter via a 2×2 MIMO link, they first detect and quantize the received messages, then compress them independently according to the Slepian and Wolf theorem, the compressed messages are sent to the receiver simultaneously where de-compression and de-quantization are performed on the received messages. After that the resulting messages are combined to estimate the original coded messages. The estimated coded messages are decoded to produce the original messages.

The basic system structure is studied and an analytical bit error rate expression is derived. Several transmission protocols are also introduced to enhance the system BER performance.

The merit of this design is focus on the relay destination link. Because the Slepian and Wolf theorem is applied on the relay-destination link, messages at the relays can be compressed independently and de-compressed jointly at the receiver with arbitrarily small error probability but still achieve the same compression rate as a joint compression scheme does.

The Slepian and Wolf theorem is implemented by a joint source-channel code in this thesis. Several schemes are introduced and tested, the testing results and performance analysis are given in this thesis.

According to the chief executive officer (CEO) problem in the network information theory, we discover an error floor in this design. An analytical expression for this error floor is derived.

A feedback link is also introduced from the receiver to the relays to allow the relays to cooperatively adapt their compression rates to the qualities of the received messages.

Two combination schemes at the receiver are introduced, their performances are examined from the information theory point of view, the results and performance analysis are given in this thesis.

As we assume that the relay destination link is a multiple access channel (MAC) suffers from block Rayleigh fading and white Gaussian noise, the relationship between the MAC channel capacity and the Slepian and Wolf compression rate region is studied to analyse the system performance.

Contents

List of Figures

List of Tables

List of Symbols

1

$P(x)$	Probability of event x
$p(x)$	probability density function of continuous variable x or marginal probability mass function of discrete variable x
$p(x,y)$	Joint probability mass function of discrete variables x and y
σ	Noise standard deviation
N_0	Noise power spectral density
$\frac{Eb}{N_0}$	The energy per bit to noise power spectral density ratio
γ	Signal-to-Noise Ratio
$E(x)$	Expectation of variable x
$\underset{x}{E()}$	Expectation over variable x
$\int_A^B$	Integration from A to B
$\sum_{i=0}^{N} a_i$	Summation of adding a_i from index $i=0$ to $i=N$, N is an integer
$\prod_{i=0}^{N} a_i$	Product of multiplying a_i from index $i=0$ to $i=N$, N is an integer
$Q()$	The Q function
exp	The exponential function
log	The logarithmic function
N_T	Number of transmit antennas
N_R	Number of receive antennas
$\sqrt{}$	Square root
$\|\| \ \|\|$	The Euclidean distance

[1] The default definitions for commonly used symbols unless stated otherwise.

$\|\ \|$	Absolute value
$()^H$	Complex conjugate
$()^{-1}$	Inverse function
$H(x)$	Entropy of variable x
$H(x,y)$	Joint entropy of variables x and y
$H(x\|y)$	Conditional entropy of variable x given variable y
$I(x;y)$	Mutual information of variables x and y
$I(x;y\|z)$	Conditional mutual information of variables x and y given variable z
max	Maximum value
min	Minimum value
$Re()$	Real part of complex number
$Im()$	Imaginary part of complex number
i	Imaginary unit
L	log likelihood ratio

Acknowledgements

I would like to show my sincere gratitude to my supervisor, Prof. Alister G. Burr and Prof. Tad Matsumoto from Japan Advanced Institute of Science and Technology, for their support and encouragement during my Ph.D. study.

I am very grateful to my thesis advisor, Dr. Yuriy Zakharov, whose insightful discussions and suggestions have benefited me.

I would also like to thank other colleagues in the Communications Research Group.

Declaration

Some of the research presented in this thesis has resulted in some publications. These publications are listed at the end of Chapter 1.

All work presented in this thesis is original to the best knowledge of the author. References and acknowledgements to other researchers have been given as appropriate.

Chapter 1

Introduction

1.1 Synopsis and Motivation

1.1.1 MIMO System

The multiple-input-multiple-output (MIMO) systems which use multiple antennas to transmit and receive information can provide spatial diversity gain as well as spatial multiplexing capabilities at no extra bandwidth or power consumption. Therefore, this is suitable for future wireless communication environments which are highly resource-constrained, only limited and tightly regulated spectrum will be offered and energy supply on wireless terminals is limited and must be consumed wisely to gain the longest operational time possible [1] [2] [3] [4].

Motivation

However attractive, MIMO requires co-located antenna array with spacing of several wavelengths between each antenna element [5] which is critical in many practical scenarios, for space at the terminal site is limited. Another disadvantage of conventional MIMO system is that transmitted signals coming from a co-located multiple-antenna terminal may experience serious impairments at the same time due to bad channel conditions, e.g., severe shadowing and/or fading.

1.1.2 Cooperative MIMO System

The desire to overcome those problems leads to a new generation of MIMO, known as cooperative MIMO, in which antennas over different terminals (nodes) cooperative with each other to form a virtual (or distributed) array. In a wireless network, transmission of information through relaying can be energy efficient and the use of multiple relaying nodes allows high diversity gain, thus leads to improved robustness against channel impairments. [6] [7] [8] show that cooperative communications can alleviate the problem inherent in colocated MIMO systems. Specifically, in [6] [7] the outage probability of distributed antenna systems in a composite fading channel has been analysed. Substantial reduction of the outage probability has been shown for a cooperative MIMO system compared to a conventional MIMO system in [7]. In [8], MIMO systems have been outperformed by relay systems in terms of the outage capacity under correlated fading and shadowing.

One type of cooperative MIMO is focused on the cooperative diversity (C-DIV) approach, in which the source broadcasts its data to both relays and the destination. The key principle of C-DIV is that transmitted signals can be received and processed by relays which can help each other, either individually or in groups. Consequently, relays can create additional paths for information transmission from source to destination to increase diversity against fading and interference[9] [10] [11].

Another approach is cooperative spatial multiplexing (C-SM) . In a C-SM system, each relay node (referred to as multiplexing relay node) obtains different information from the source in the first phase (either by broadcasting higher order constellation modulation signal from the source or by using separate channel, time divided or frequency divided, to communication with each relay node from the source). Each relay node only transmits a sub-stream of the source information in the second phase. In this way, the transmit and receive processing requirements on the relay nodes are simplified [12] [13] compared to C-DIV because each relay only has to handle part of the original information which is preferable due to their highly limited resources in terms of energy, bandwidth, computation power, memory, and space. It is significantly effective in the high signal-to-noise ratio (SNR) regime even without any channel coding, while the efficiency in the low SNR regime can be improved with appropriate channel coding schemes. Particularly, C-SM is useful in sending high-rate data when each relay node needs to handle only low-rates.

1.2 Project Objectives

- Investigate and understand some general features of a basic C-SM model from theoretical and practical point of view.
- Based on the resulting discoveries, develop new techniques and strategies to enhance the performance of the C-SM model.
- Analyse the system performances against theoretical bounds and develop optimization schemes.
- Propose our own C-SM system design.

1.3 Contributions

- For the basic C-SM model
  - Present the bit error rate (BER) performance of the uncoded system.
  - Develop a semi-analytical method to obtain upper and lower bounds for the system's BER.
  - Investigate the impact of different modulation schemes on the overall system BER performance
  - Study the error floor in the system and derive an analytical expression for this error floor and it is verified by simulation result. This expression will be useful in the future optimization work.

- For the C-SM system with Slepian and Wolf code
  - Two Slepian and Wolf theorem implementation methods are tested, namely the super turbo code and the accumulate repeat accumulate code. Finally decide to use the ARA code because of its practicability.
  - Develop several quantization and reconstruction schemes. Simulation results and analysis show that the demodulator quantization method coupled with the log likelihood-ratio (LLR) combination de-quantization method give the best result.
  - Discover an interesting property of this system structure that lower frame error rate (FER) at relay does not correspond to lower FER at the destination: the similarity between the data at the relays also affects the FER at the destination.
  - Introduce fading relay destination link, propose two system designs, namely the separate design and the joint design. Simulations and our analysis show that the separate design outperforms the joint design from the outage probability point of view.
- For the optimization work
  - We optimize the implementation of the Slepian and Wolf code which leads to more reliable and efficient performance.
  - We also study the system from the information theoretic point of view, based on which we propose a new compression scheme which can further increase the spectral efficiency with minor BER loss.

1.4 Thesis Outline

- Chapter 2 provides background information which helps explain the concepts mentioned in the rest of the thesis.
- In Chapter 3, a basic C-SM system is studied. The structure of this system is regarded as the foundation of our final design.
- Slepian and Wolf theorem is introduced and incorporated into our system design in Chapter 4. Several implementation methods and performance analysis are also given.
- We investigate the error floor in the system in Chapter 5.
- In Chapter 6, some enhancements are made to the Slepian and Wolf code. A more efficient compression scheme is also introduced.
- We introduce a fading relay destination link into the system, develop two new system designs and according to the performance analysis, we conclude that the separate design is better than the joint design in Chapter 7,
- We conclude our contributions in Chapter 8.

1.5 Publication List

1. N. Xie and A. Burr, "Distributed cooperative spatial multiplexing with Slepian Wolf code", in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013, pp. 15.

2. N. Xie and A. Burr, "Cooperative spatial multiplexing with distributed amplify-and-forward relays", in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2012, pp. 616620.

3. N. Xie and A. Burr, "Implementation of Slepian Wolf theorem in a distributed cooperative spatial multiplexing system", present at *the 2014 European Wireless*.

4. N. Xie, A. Burr, and T. Matsumoto, "Error floor analysis for receiver with distributed relays", *Communications Letters, IEEE*, ready for submitting.

5. Nian Xie and A. Burr, "Implementation of Slepian Wolf theorem in a distributed cooperative spatial multiplexing system - Part II", Conference Paper, ready for submitting.

Chapter 2

Background

Background information which helps explain the concepts mentioned in the rest of the thesis is shown in this chapter. The structure of this chapter is given below.

Probability of error is widely used to determine and analyse the performance of our system, it can come in various forms, such as Bit Error Rate, Union bound or Outage Probability. The concepts are given in the sections from 2.1 to 2.3.

Because our project is primarily focused on cooperative MIMO systems, the fundamental information of traditional MIMO and cooperative MIMO are given in section 2.4 and 2.5. Particularly, the concept of MIMO spatial multiplexing is shown in 2.4.3.

Information theory provides guideline for system design and optimization, it is also the foundation for channel capacity which is a key aspect in examining a communication system. The information about information theory and channel capacity are introduced in 2.6 and 2.7. The network information theory is introduced in 2.10.

In this project, various coding schemes are used to implement the theorems provided by information theory, concepts of coding including channel code, source code and joint channel-source code are given in 2.8 and 2.9. The optimization for code with iterative decoding scheme is given in 2.11.

2.1 Average Bit Error Rate

BER effectively reveals the nature of the system behaviour and is often illustrated in documents containing system performance evaluations. BER is used in every step of this project from Chapter 3 to Chapter 8. The concept of average BER in fading channels is introduced below.

Assume signal-to-noise ratio (SNR) γ_s is constant during the transmission of one symbol. By integrating the BER in the additive white Gaussian noise (AWGN) channel over the fading distribution, average BER can be expressed as

$$\overline{P_s} = \int_0^\infty P_s(\gamma) p_{\gamma_s}(\gamma) d\gamma \tag{2.1}$$

where $P_s(\gamma)$ is the probability of symbol error in AWGN channel with SNR γ, $p_{\gamma_s}(\gamma)$ is the probability density function (PDF) of SNR γ.

By utilizing the relationship between SNR and signal amplitude r

$$\gamma = \frac{r^2 T_s}{2\sigma_n^2} \tag{2.2}$$

where T_s is the symbol transmission time, $\sigma_n^2 = N_0/2$ is the power spectral density (PSD) of the noise in the in-phase and quadrature branches, $p_{\gamma_s}(\gamma)$ can be calculated from the PDF of r which is usually well-known.

2.2 ML detector and its Union Bound

In Chapter 3, union bound of a maximum likelihood (ML) detector is used to calculate the upper bound of the system's BER. The derivation of the union bound is given below:

As shown below[1], symbol error probability of the ML detector is given by [5]

$$\begin{aligned} P_e &= \sum_{i=1}^{M} p(\mathbf{r} \notin Z_i | m_i \quad \textbf{sent}) p(m_i \quad \textbf{sent}) \\ &= \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r} \notin Z_i | m_i \quad \textbf{sent}) \\ &= 1 - \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r} \in Z_i | m_i \quad \textbf{sent}) \\ &= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} p(\mathbf{r} | m_i) d\mathbf{r} \\ &= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} p(\mathbf{r} = \mathbf{s}_i + \mathbf{n} | \mathbf{s}_i) d\mathbf{n} \\ &= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i - s_i} p(\mathbf{n}) d\mathbf{n} \end{aligned} \tag{2.3}$$

where r, m_i, Z_i and M are received symbol, transmitted symbol, decision region for the transmitted symbol and number of symbols on the constellation.

Because equation ((2.3)) does not have a closed form, it is approximated by a union bound which is a function of the distance between signal constellation points.

For a particular point in the constellation, calculate the probability of error when this point is transmitted but received as another point on the constellation, summarize all the possible combination. Averaging over all possible points in the constellation yields the union bound

$$P_u = \sum_{i=0}^{M} p(m_i \quad \textbf{sent}) \sum_{\substack{k=0 \\ k \neq i}}^{M} P_e(m_i \text{ is transmitted,received as } m_k) \tag{2.4}$$

where $p(m_i \quad \textbf{sent})$ is the probability that signal m_i is sent, and M is number of signals on the constellation.

Union bound is an upper bound since union bound simply adds together all the probabilities that m_i is mistakenly decoded as other point, thus some decision regions are overlapped.

[1] equal probability of transmission for all signals on the constellation is assumed in this thesis

2.3 Outage Probability

The outage probability is used in Chapter 4 and Chapter 7 as the way to evaluate system performance. Its definition is [5]

$$P_{out} = p(\gamma_s < \gamma_0) = \int_0^{\gamma_0} p_{\gamma_s}(\gamma) d\gamma \tag{2.5}$$

where γ_0 is the minimum SNR that certain performance requires, which is the cumulative distribution function (CDF) of γ_s.

Outage probability shows the probability that the instantaneous SNR drops below a specified value or equivalently the error probability exceeds a specified value.

2.4 MIMO

Multiple Input Multiple Output systems are commonly refers to the systems with multiple antennas at the transmitter and the receiver. Compared to its counterpart, Single Input Single Output (SISO) systems, MIMO systems can increase data rates through multiplexing or to improve stability through diversity without extra bandwidth or energy.

2.4.1 Diversity Gain

In order to mitigate the effects of fading on the performance of modulation over wireless channels, same data are sent over independent fading paths and combined in some way at the receiver such that the fading of the resultant signal is reduced. This technique is commonly referred as diversity-combining.

There are numerous ways to achieve diversity, however in this thesis we focus on one method called space diversity which uses spatially separated antennas in a MIMO system to create independent fading paths.

Receiver Diversity

For systems with multiple antennas at the receiver, signal received by each antenna travels through an independent fading path, signals from all antennas are combined to obtain a resultant signal which is then passed through a standard demodulator.

There are two advantages associated with receiver diversity: array gain and diversity gain. Array gain results from coherent combining of multiple receive signals and is defined as the increase in combined SNR over the average SNR.

For systems using diversity-combining techniques, according to [5] its average BER can sometimes be expressed in the form $\overline{P_s} = c\overline{\gamma}^{-M}$ where c is a constant which depends only on the modulation and coding scheme and M is called the diversity order of the system. For a system with M antennas, its maximum diversity order is M, it is said to achieve full diversity order when its diversity order equals M.

Transmitter Diversity

Transmit diversity can be observed in systems which have multiple antennas at the transmitter with the transmit power divided among these antennas. Transmit diversity can be divided into two types, depending on whether or not the channel state information (CSI) is known at the transmitter.

If path gain $r_i e^{j\theta_i}$ associated with the i_{th} antenna is known at the transmitter, a gain $\alpha_i = a_i e^{-j\theta_i}, 0 <= a_i <= 1$ for each antenna can be calculated accordingly. Signal on every channel is multiplied by its corresponding gain and is transmitted over all antennas are added "in the air", which leads to a received signal

$$r(t) = \sum_{i=1}^{M} a_i r_i s(t) \tag{2.6}$$

The weights a_i that achieve the maximum SNR are [5]

$$a_i = \frac{r_i}{\sqrt{\sum_{i=1}^{M} r_i^2}} \tag{2.7}$$

and the resulting SNR is

$$\gamma_\Sigma = \sum_{i=1}^{M} \gamma_i \tag{2.8}$$

which is the sum of SNRs on each of the individual branches.

The disadvantage of this transmit diversity is the CSI at the transmitter, which is usually done by a feedback channel from the receiver to the transmitter.

If CSI is not available at the transmitter, a naive strategy is to divide the transmit energy equally among all antennas [5].

2.4.2 Multiplexing Gain

When multiple antennas are available at both the transmitter and the receiver, multiplexing gain can be achieved by transmitting independent data simultaneously over different transmit antennas and detecting them jointly from all the receive antennas. The multiplexing gain of a MIMO system results from the fact that a MIMO channel can be decomposed into a number of parallel independent channels. By transmitting independent data over these independent channels, an increase in data rate can be get in comparison to a system with one antenna at the transmitter and receiver. This increased data rate is called the multiplexing gain.

2.4.3 Spatial Multiplexing System

MIMO spatial multiplexing (SM) system transmits independent data stream on each transmit antenna simultaneously, data steam may be independently coded or left uncoded and is received by all the receive antennas.

Assume a MIMO system with n_t transmit and n_r receive antennas communicating through a frequency flat-fading channel. A codeword $\mathbf{C} = [\mathbf{c}_0 ... \mathbf{c}_{T-1}]$ of size $n_t \times T$ contained in the

codebook $\mathscr{C}$ is transmitted over T symbol periods via n_t transmit antennas. At the k^{th} time instant, the transmitted and received signals are related by

$$\mathbf{y}_k = \sqrt{\mathbf{E}_s}\mathbf{H}_k\mathbf{c}_k + \mathbf{n}_k \tag{2.9}$$

where $\mathbf{y}_k$ is the $n_r \times 1$ received signal vector, $\mathbf{H}_k$ is the $n_r \times n_t$ channel matrix and $\mathbf{n}_k$ is a $n_r \times 1$ zero mean complex additive white Gaussian noise vector with variance σ^2. The parameter $\mathbf{E}_s$ is the energy normalization factor.

Maximum Likelihood Detector

Several detection schemes are available for the SM system, however since we only use maximum likelihood detector in our project, we will focus our attention on its introduction.

According to [5], the ML detector minimizes error probability by selecting the output $\hat{m}$ that maximum $1 - P_e = p(\hat{m} \quad \textbf{sent}|r)$, where r and P_e are received symbol and symbol error probability for the ML detector defined in section (2.2). Abbreviate $p(s_i \quad \textbf{sent}|r)$ as $p(s_i|r)$ and $p(s_i \quad \textbf{sent})$ as $p(s_i)$. Using Bayes rule [14] we have,

$$p(s_i|r) = \frac{p(r|s_i)p(s_i)}{p(r)} \tag{2.10}$$

To minimize error probability, the receiver output $\hat{m}$ must correspond to the constellation s_i that maximizes $p(s_i|r)$. Because $p(r)$ is not a function of s_i and assuming equally likely messages ($p(s_i) = 1/M$, where M is the number of symbols on the constellation), the corresponding s_i must maximize $p(r|s_i)$.

Associate $p(r|s_i)$ with a likelihood function,

$$L(s_i) = p(r|s_i) \tag{2.11}$$

The problem becomes to maximize $L(s_i)$ and since the log function increases as its argument increases, maximizing $L(s_i)$ is equivalent to maximizing $l(s_i) = \log L(s_i)$, where log is the log function. From [5] we know,

$$p(Re(r)|Re(s_{ij})) = \frac{1}{\sqrt{\frac{N_0}{2}}\sqrt{2\pi}} \exp[-\frac{1}{N_0}|Re(r_j) - Re(s_{ij})|^2] \tag{2.12}$$

where j is the index for the receive antennas, N_0 is the noise power spectral density.

$$p(Im(r)|Im(s_{ij})) = \frac{1}{\sqrt{\frac{N_0}{2}}\sqrt{2\pi}} \exp[-\frac{1}{N_0}|Im(r_j) - Im(s_{ij})|^2] \tag{2.13}$$

and because the real part and the imagine part of the receive signal are independent

$$p(r|s_{ij}) = p(Re(r)|Re(s_{ij}))p(Im(r)|Im(s_{ij})) = \frac{1}{\pi N_0} \exp[-\frac{1}{N_0}|r_j - s_{ij}|^2] \tag{2.14}$$

finally because the signals on different receive antennas are independent

$$p(r|s_i) = \frac{1}{(\pi N_0)^N} \exp[-\frac{1}{N_0}\sum_{j=1}^{N}|r_j - s_{ij}|^2] \tag{2.15}$$

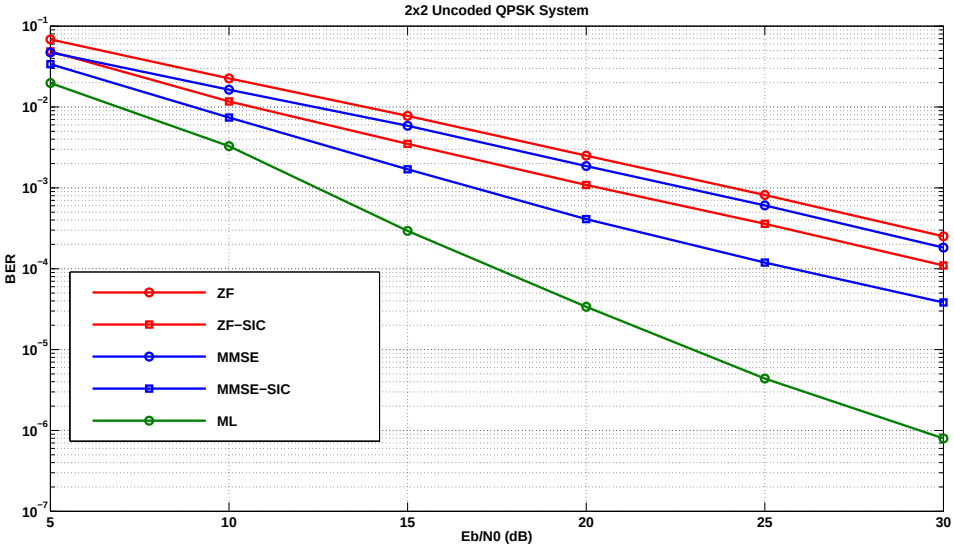

Figure 2.1: BER performance of spatial multiplexing system with various detectors in i,i,d Rayleigh slow fading channels

where N is the number of receive antennas.

So

$$l(s_i) = -\frac{1}{N_0}\sum_{j=1}^{N}(r_j - s_{ij})^2 = ||r - s_i||^2 \tag{2.16}$$

Thus the log likelihood function $l(s_i)$ depends only on the distance between the received vector r and the constellation point s_i.

Compared to other detection schemes [5], such as Zero-Forcing, Minimum Mean Square and Successive Interference Cancellation, ML detector has the best performance as shown in Figure 2.1

2.5 Cooperative MIMO

MIMO systems can increase data rate and BER performance without additional bandwidth or energy, however multiple antennas are required at both transmit and receive end, which may be difficult to implement in practice on small mobile devices. This is because space diversity requires that distance between antennas must be long enough for the fading characteristic associate with each antenna to be roughly independent. For example, in a uniform scattering environment with isotropic transmit and receive antennas, approximately one half wavelength separation between antennas is required for independent fading [5]. For directional antennas, because the multipath is confined to a small angle relative to the line of sight (LOS) path, a larger antenna separation is required for independent fading channels [15].

One solution is to place multiple spatially distributed relay stations between transmitter and receiver, which serve as virtual antenna arrays (VAA) that jointly achieve MIMO transmission [16]. VAA was introduced in [17] and another scheme for cooperative MIMO is distributed space-time block coding (DSTBC) [18] used to achieve diversity gain.

2.5.1 Performance Gain of Cooperative MIMO

Multiplexing Gain

According to [19], in cooperative MIMO systems, multiplexing gain is obtained by allowing an antenna VAA with N_t number of elements to transmit up to N_t independent data streams. At the receive VAA, it needs to solve a system of N_r equations with N_t variables. Thus, for data streams to be successfully detected, the number of nodes in the receive VAA N_r must be larger than it is in the transmit VAA N_t. The multiplexing gain of a MIMO or cooperative MIMO link is given by $\min(N_t, N_r)$ and this number can be lower if the channel gain matrix is not full-rank.

Unlike classic MIMO, the multiplexing gain in cooperative MIMO involves some cooperation overhead which can be time delay, signalling packets for coordination or interference incurred from other nodes in the VAA.

Diversity Gain

Cooperative MIMO can also achieve diversity gain by transmitting multiple correlated versions of the same signal, which improves the SNR at the receive VAA [19].

2.5.2 Application of Cooperative MIMO

Wireless Sensor Network (WSN)

In WSN, because sensors are often powered by batteries and are difficult to recharge or replace, energy-efficiency is crucial for WSN design.

1. Diversity Gain in WSNs

   In [20], the tradeoff between the transmission energy and the circuit energy for DSTBC in WSN is presented. By increasing the diversity, less transmission energy is needed, however higher diversity involves more mobile nodes, thus consuming more circuit energy. For long distance communication, transmission energy is dominant, so cooperative MIMO schemes should increase the size of VAAs to increase the diversity. On the other hand, circuit energy consumption is the major contributor in short distance communication, so lower diversity scheme should be employed.

2. Multiplexing Gain in WSNs

   [21] proposed a Vertical-Bell Laboratories Layered Space-Time (VBLAST) based scheme for cooperative MIMO, in which the receiver is equipped with multiple antennas and is assumed to have sufficient energy and computational capability. By allowing all sensors to transmit to the receiver simultaneously, the cooperation burden is shifted from the mobile nodes to the receiver. Because mobile nodes are transmitting at a higher rate than the DSTBC scheme is, the transmission duration is reduced and so is the circuit energy consumption.

3. Cluster in the WSNs

   In the context of clustered architecture, neighbouring nodes are often grouped into a cluster and administrated by a common node referred to as the cluster head.

The authors in [22] designed a clustered-based DSTBC scheme for cooperative MIMO systems. Each cluster is managed by a master cluster head (MCH) and a slave cluster head (SCH) . MCH is responsible for collecting information from members in its cluster and coordinating with its SCH to form a VAA. By adapting the transmit powers of individual nodes to the inter-cluster distance, the energy consumption is reduced.

Mobile ad-hoc Networks (MANETs)

The range extension of cooperative MIMO system reduces the number of multi-hop paths, consequently improves network throughput and reduces delay. A medium access control (MAC) and routing framework is proposed in [23] to improve a MANET's throughput. A SISO path is initially constructed using conventional routing protocol, which is gradually improved by MISO paths in which multiple nodes jointly encode the signal using DSTBC before transmitting it to a single receiver.

WLANs

In WLANs, several access points (APs) serve different devices simultaneously, the number of concurrent data steams on the uplink is constrained by the number of antennas at the APs, even if several APs are in the same area. In [24], a cooperative MIMO based scheme called interference alignment and cancellation (IAC) is proposed. It uses multi-antenna APs cooperation to solve the problem. The cooperative messages between APs are exchanged via cables or ultra-wideband transmission. A cooperative MIMO receiver with a four-element VAA can decode three concurrent streams, compared to two streams for 802.11n. Theoretically, IAC can double the throughput of a 802.11n WLAN.

Cognitive Radio Networks

Cooperative MIMO can also improve network throughput and reduce delay in MIMO cognitive radio networks[25]. In a cognitive radio network, secondary users (SUs) arrange their transmission time slots according to the traffic of primary users(PUs) . If SUs are equipped with multiple antennas, cooperative MIMO schemes can be applied, so some PU and SU transmitters (receivers) can cooperate to form a transmitting (receiving) VAA and simultaneously send (receive) PU's and SU's traffic to (from) a SU. The advantage for SUs is that they do not need to refrain from transmitting even when PUs are detected.

Cellular Systems

In cellular systems, co-channel interference among adjacent cells can severely affect the capacity of a base station (BS) , especially in smaller cells such as femto-cells. However if nearby BSs can cooperate to form a VAA to function as a giant BS [26], this problem can be solved.

2.6 Information Theory

Information theory defines the fundamental limits of data processing operations such as data compression and data transmission. Particularly for communication engineering, information

theory represents the extreme points of the set of all possible communication schemes, one side is the data compression limit and the other side is the data transmission limit, known as the channel capacity. Therefore, all modulation and data compression schemes lie between these limits.

Information theory is utilized across the whole project, in Chapter 4 we implemented the Slepian and Wolf theorem in the network information theory. In Chapter 6 we measured the mutual information between the original data and the combined received signal at the destination to determine the performance of different combination schemes. We also proved that compress beyond the classic Slepian and Wolf limit in a CSM system is possible from a information theory point of view. In Chapter 7, the majority of the analysis is based on the relationship between the Slepian and Wolf compression limit and channel capacity.

Below are some basic definitions in the information theory

2.6.1 Entropy

Entropy measures the uncertainty of a random variable, it defines the limit for lossless data compression.[2] For a discrete random variable X, its entropy is defined as [14]

$$H(X) = -\sum_{x\in\mathscr{X}} p(x)\log p(x) \tag{2.17}$$

where $p(x) = Pr\{X = x\}, x \in \mathscr{X}$ is the probability mass function. In the digital system, the log is usually to the base 2 and entropy is expressed in bits.

For any continuous random variable, its entropy is infinite, because it has infinite number of possible values.

2.6.2 Joint Entropy

Joint entropy measures the uncertainty of a pair of random variables. For a pair of discrete random variables (X,Y), their joint entropy is defined as [14]

$$H(X,Y) = -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} p(x,y)\log p(x,y) \tag{2.18}$$

where $p(x,y)$ is the joint distribution.

Conditional Entropy

For a pair of discrete random variables (X,Y), their conditional entropy is defined as [14]

$$\begin{aligned} H(Y|X) &= \sum_{x\in\mathscr{X}} p(x)H(Y|X=x) \\ &= -\sum_{x\in\mathscr{X}} p(x)\sum_{y\in\mathscr{Y}} p(y|x)\log p(y|x) \\ &= -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} p(x,y)\log p(y|x) \end{aligned} \tag{2.19}$$

[2] details see section 2.9.1

In fact, the joint entropy of a pair of random variables is the entropy of one plus the conditional entropy of the other.

$$H(X,Y) = H(X) + H(Y|X) \tag{2.20}$$

this rule can also expand to three random variables

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z) \tag{2.21}$$

2.6.3 Mutual Information

Mutual information measures the amount of information one random variable contains about another one. For a pair of discrete random variables (X,Y), their mutual information is defined as [14]

$$I(X;Y) = \sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \tag{2.22}$$

where $p(x,y)$ is the joint probability mass function and $p(x), p(y)$ are marginal probability mass functions.

Conditional Mutual Information

The conditional mutual information of random variables X and Y given Z is defined by [14]

$$\begin{aligned} I(X;Y|Z) =& H(X|Z) - H(X|Y,Z) \\ =& E_{p(x,y,z)}\log\frac{p(X,Y|Z)}{p(X|Z)p(Y|Z)} \end{aligned} \tag{2.23}$$

2.6.4 Mutual Information and Entropy

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ I(X;Y) &= H(Y) - H(Y|X) \\ I(X;Y) &= H(X) + H(Y) - H(X,Y) \\ I(X;Y) &= I(Y;X) \\ I(X;X) &= H(X) \end{aligned} \tag{2.24}$$

The relationship between $H(X), H(Y), H(X,Y), H(X|Y), H(Y|X)$ and $I(X;Y)$ is expressed in a Venn diagram Figure 2.2. Note that the mutual information $I(X;Y)$ is the intersection of the information in X with the information in Y [14].

2.7 Channel Capacity

[3] Source symbols from a finite alphabet are mapped into a sequence of channel symbols and after begin transmitted through the channel, an output sequence is produced. Each input sequence induces a probability distribution on the output sequences. In order to recover the transmitted

[3] Although channel capacity is a subject belonging to information theory, it has a fundamental influence to communication engineering, therefore it is introduced as an independent section.

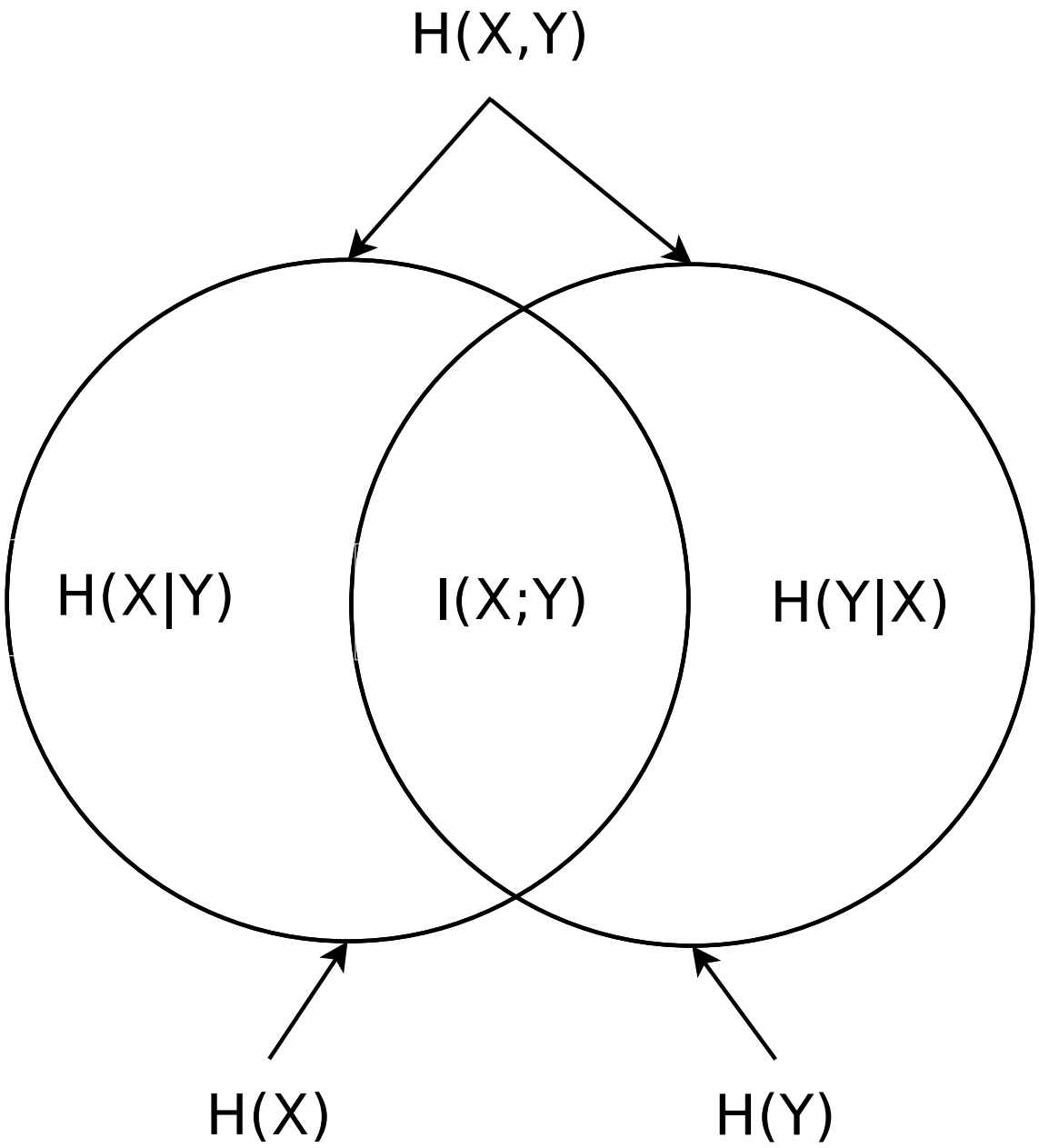

Figure 2.2: Relationship between entropy and mutual information

message from the output sequence, subsets of input sequences must be defined, so that there is only one input sequence with high probability can cause a particular output distribution. Then the message can be reconstructed with a negligible probability of error. By mapping the source into these input subsets which results in an exclusive output distribution, messages can be transmitted with very low probability of error and the maximum rate at which this can be done is called the capacity of the channel.

Defining the input alphabet $\mathscr{X}$ and output alphabet $\mathscr{Y}$, the probability of observing the output symbol y given the input symbol x is defined as $p(y|x)$. If $p(y|x)$ only depends on

the current input and is independent of previous inputs or outputs, the channel is said to be memoryless.

For each input sequence with length n, there are approximately $2^{nH(Y|X)}$ possible Y output sequences, all of them are equiprobable. In order to correctly estimate the input sequence based on the output sequence, it is essential that no two X sequences produce the same Y output sequence. The total number of possible Y sequence is $\approx 2^{nH(Y)}$, dividing this into sets of size $2^{nH(Y|X)}$ is the number of exclusive input X sequences which is $2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$.

The channel capacity is defined as

$$C = \max_{p(x)} I(X;Y) \tag{2.25}$$

where the maximum is taken over all possible input distributions $p(x)$.

Below is an example to derive a MIMO channel capacity [27].

Assume a system with N_T transmit antennas and N_R receive antennas. A constellation $\mathscr{M}$ consists of M signals.[4] s is the $N_T \times 1$ transmission vector with entries chosen from $\mathscr{M}$. H is the $N_R \times N_T$ channel matrix which is available at the receiver. y is the $N_R \times 1$ receive vector. σ is the noise standard deviation in each dimension.

According to (2.25)

$$\begin{aligned} I(s;y) =& H(y) - H(y|s) \\ =& -E[\log_2(p(y))] - (-E[\log_2(p(y|s))]) \end{aligned} \tag{2.26}$$

where $E()$ denotes expectation,because y has $2N_R$ dimensions with independent Gaussian distribution,

$$\begin{aligned} p(y|s) &= \left(\frac{1}{2\pi\sigma^2}\right)^{N_R} \exp(\frac{-||y-Hs||^2}{2\sigma^2}) \\ p(y) &= \frac{1}{M^{N_T}} \left(\frac{1}{2\pi\sigma^2}\right)^{N_R} \prod_s \exp(\frac{-||y-Hs||^2}{2\sigma^2}) \end{aligned} \tag{2.27}$$

Substituting (2.27) into (2.26),

$$\begin{aligned} I(s;y) =& -E\log_2(\frac{1}{M^{N_T}} \left(\frac{1}{2\pi\sigma^2}\right)^{N_R} \prod_s \exp(\frac{-||y-Hs||^2}{2\sigma^2})) \\ &+ \log_2 \left(\left(\frac{1}{2\pi\sigma^2}\right)^{N_R} \exp(\frac{-E||y-Hs||^2}{2\sigma^2}) \right) \end{aligned}$$

for a particular H and s, $E||y-Hs||^2 = 2N_R\sigma^2$, so

$$\begin{aligned} I(s;y) =& -E\log_2(\frac{1}{M^{N_T}} \left(\frac{1}{2\pi\sigma^2}\right)^{N_R} \prod_s \exp(\frac{-||y-Hs||^2}{2\sigma^2})) \\ &- N_R \log_2\left(2\pi\sigma^2 \exp(1)\right) \end{aligned}$$

[4] All signals in the constellation have equal probability of transmission.

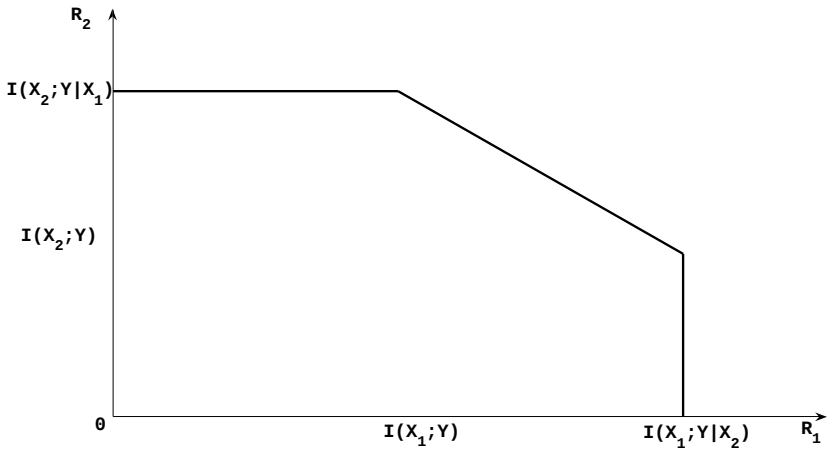

Figure 2.3: Rate region for multiple access channel

2.7.1 Multiple-Access Channel

In a Multiple-Access channel, two or more senders send information to a common receiver. Define a discrete memoryless multiple-access channel consisting of three alphabets, $\mathscr{X}_1, \mathscr{X}_2$ and $\mathscr{Y}$ and a probability mass function $p(y|x_1,x_2)$. The multiple-access channel capacity theorem [14] states that the capacity of a multiple-access channel $(\mathscr{X}_1 \times \mathscr{X}_2, p(y|x_1,x_2), \mathscr{Y})$ is the closure of the convex hull of all (R_1,R_2) satisfying

$$\begin{aligned} R_1 &< I(X_1;Y|X_2) \\ R_2 &< I(X_2;Y|X_1) \\ R_1+R_2 &< I(X_1,X_2;Y) \end{aligned} \tag{2.28}$$

for some product distribution $p_1(x_1)p_2(x_2)$ on $\mathscr{X}_1 \times \mathscr{X}_2$, as shown in Figure 2.3 where lines $R_1 = I(X_1;Y|X_2)$, $R_2 = I(X_2;Y|X_1)$, $R_1+R_2 = I(X_1,X_2;Y)$, $R_1 = 0$ and $R_2 = 0$ form a closed area, information transmission over the MAC channel using any rate pair within this area can guarantee a reliable communication.

2.8 Channel Code

A channel code is designed to detect or correct the errors introduced by the fading channel and the noise during the transmission. Normally, a channel code consists of two parts: information bits and parity bits. Parity bits are considered as the redundancy added into the encoded message, usually they are the results of a function of the original information bits. Because the length of the encoded message is longer than the original message, every bit in the original message is associated with more than one bit in the encoded message. During the transmission, some bits in the encoded message may be corrupted by the fading channel and the noise, however as long as bits associated with one particular bit in the original message are not all corrupted, there is a chance that this bit in the original message can be correctly decoded.

In a fading channel, errors occur typically in a burst and therefore cannot be corrected. However, the performance of a channel code can be improved by using interleaving. The

interleaver spreads out bursts of errors over time, so it provides a form of time diversity.

Compared to a uncoded system, a sharp drop of BER as the SNR increases may observed in a coded system, which is called "BER cliff" and this phenomenon is considered as the indicator of how close the code performance (code rate that assure arbitrarily small error rate) is to the theoretical limit (channel capacity)[5]. Another feature of a coded system is called"error floor" which shows that above a certain SNR threshold which depends on the individual code design, the BER falls off much more slowly and is caused by low-weight codewords.

2.8.1 Turbo Code

Turbo code is a widely used channel code whose performance can come within a fraction of a dB of the Shannon capacity limit on AWGN channels. It consists of two key components: parallel concatenated encoding and iterative "turbo" decoding.

Encoder

At the encoder, the information bits are encoded twice by two component convolutional encoders with an interleaver between them which makes the two encoded information bits sequences independent. Usually, each component encoder produces an information sequence and a parity sequence with a rate of 1/2. At the output of the component encoder, puncturing takes place by puncturing the whole information sequence from one encoder and half the parity sequence from both encoders in alternate positions, which results in an overall coding rate of 1/2. Figure 2.4 shows the structure of the turbo encoder and component encoder [28].

Decoder

At the decoder two component convolutional decoders are used. The soft output from two component decoders not only indicates the value of the relevant bit but also the probability that this bit is decoded correctly. The turbo decoder operates in an iterative fashion. In the first iteration, only the soft channel output information about the output of the first component encoder is provided to the first component decoder. It outputs extrinsic information which is used by the second component decoder as a priori information. This a priori information coupled with the soft channel output information about the output of the second component encoder are the input for the second component decoder. The second component decoder also gives extrinsic information. As the iterations go on, both component decoders produce extrinsic information based on the channel output and a-priori information obtained from the extrinsic information provided by the other decoder. After each iteration, the BER of the decoded sequence drops [28]. Figure 2.5 shows the structure of the turbo decoder.

2.8.2 EXIT chart

The extrinsic information transfer chart (EXIT chart) is a tool to assist the design of error correction code with iterative decoder, and was introduced by Stephan ten Brink in [29, 30].

In this technique, component decoders in an iterative decoder are characterised by their mutual information response. For each component decoder, mutual information is measured at

[5]details are shown in section 2.8.2

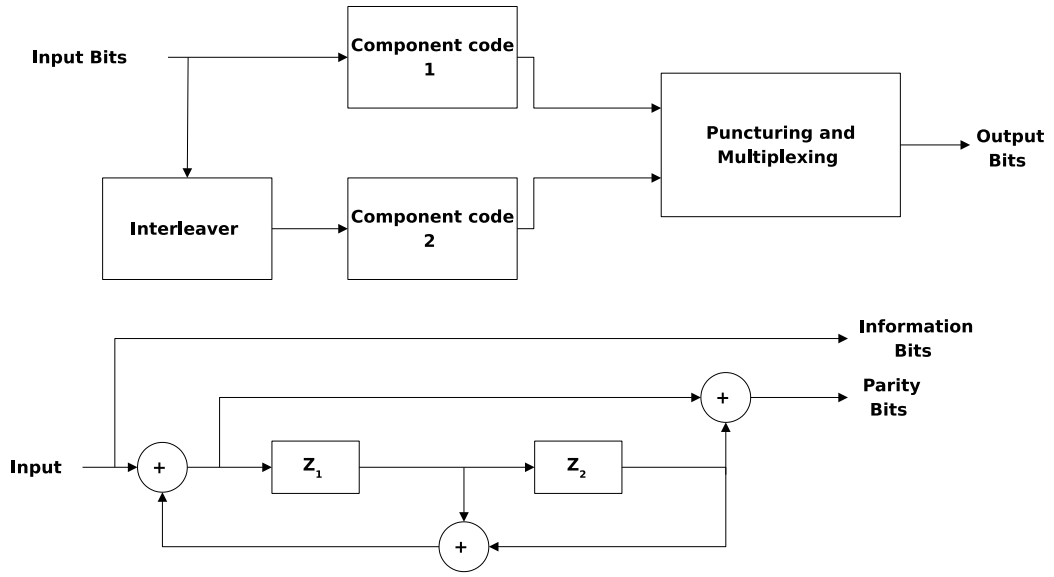

Figure 2.4: Turbo encoder and component encoder

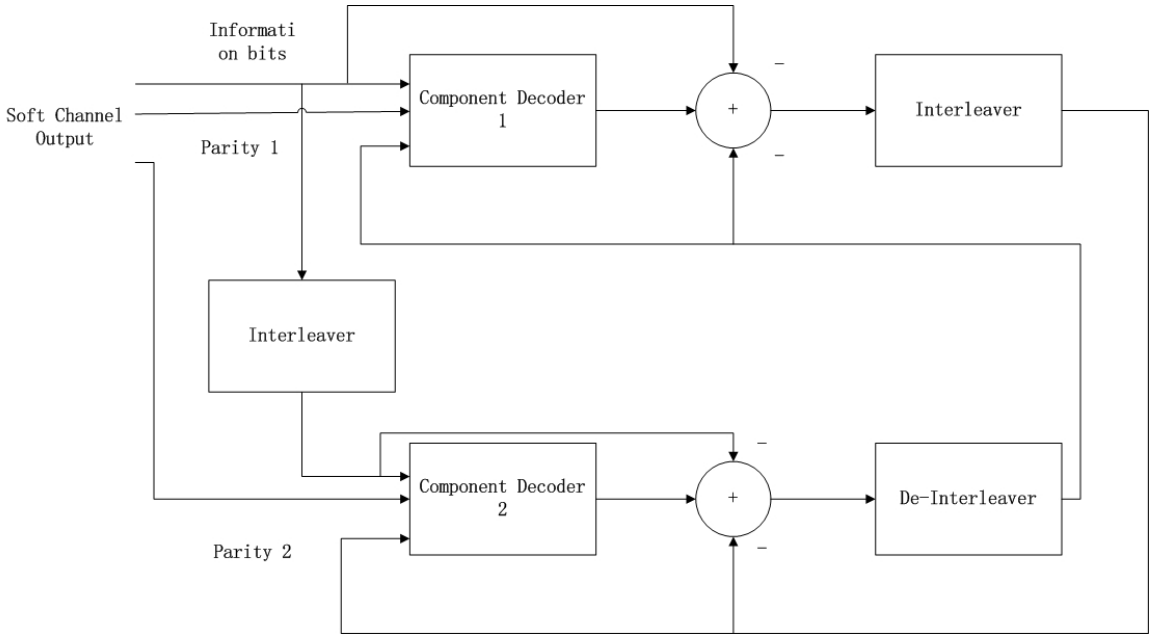

Figure 2.5: Turbo decoder

its input and output, for a series of different pseudo priori information. The mutual information response of one component decoder is then plotted on a two dimensional chart(X axis for input, Y axis for output) which will form a curve. Another component decoder's response curve is also plotted but with its input and output reversed(X axis for output, Y axis for input). A

decoding failure will be shown on the EXIT chart as two curves are intersecting. Otherwise, there would be a tunnel between them, the thinner the tunnel is, the closer the performance of the decoder is to the theoretical limit as shown in Figure 2.6. The exchange of extrinsic information between two component decoders is visualized as a step trajectory between these two curves.

In [31], two methods to measure the mutual information for BPSK modulated system are introduced.

For a general case with sufficient samples, mutual information can be calculated as:

$$I(L;X) = 1 - E\{\log_2(1+\exp^{-L})\} \approx 1 - \frac{1}{N}\sum_{n=1}^{N}\log_2(1+\exp^{-x_n L_n}) \tag{2.29}$$

where N is the number of samples, L is LLR, X is bit value $(0,1)$.

However, if the samples have a symmetric probability density function(PDF) $p(-y|x=+1) = p(y|x=-1)$ and satisfy the consistency condition $p(-y|x) = \exp^{-L_c xy} p(y|x)$, the calculation of mutual information can be simplified as:

$$I(L;X) \approx 1 - \frac{1}{N}\sum_{n=1}^{N} H_b(P_{e_n}) = 1 - \frac{1}{N}\sum_{n=1}^{N} H_b(\frac{\exp^{+|L_n|/2}}{\exp^{+|L_n|/2}+\exp^{-|L_n|/2}})$$

where $H_b()$ is the binary entropy function. Compared with the general algorithm (2.29), knowledge of X is not required.

The a priori information used in the EXIT chart comes from a pseudo component decoder, it models the non-linear performance of the mutually other decoder in the iterative decoding scheme. According to ten Brink's original assumption, it is modelled as an AWGN channel.

For systems with other modulations, mutual information can be calculated using the algorithm in [31].

$$\begin{aligned} I(x,y) =& H(x) - H(x|y) \\ =& 1 - \sum_x \int p(x,y)\log_2(1/p(x|y))\,dy \\ =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(1/p(x|y))] \end{aligned}$$

where x, y represent the original message and the decoded message in the LLR form. I, H, E, $p(x,y)$ and $p(x|y)$ indicate mutual information, entropy, expectation, joint probability density function and conditional probability density function respectively.

By using Bayes'law, the equation can be rewritten as

$$\begin{aligned} I(x,y) =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(\frac{p(y)}{p(x)p(y|x)})] \\ =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(\frac{\sum_x p(x)p(y|x)}{p(x)p(y|x)})] \\ =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(1+\frac{p(y|-x)}{p(y|x)})] \end{aligned} \tag{2.30}$$

where $p(y|-x)$ and $p(y|x)$ can be obtained empirically from the decode message.

Figure2.6 shows the EXIT chart for a $\frac{1}{2}$ rate repeat accumulator code in the AWGN channel with 5 dB SNR.

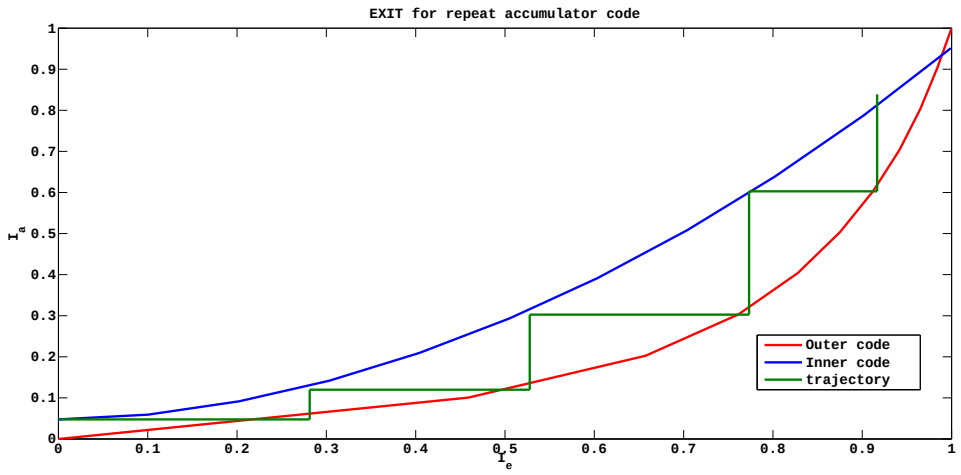

Figure 2.6: EXIT for repeat accumulator code

2.9 Source coding

Source coding is used to compress the source data by identifying and eliminating the redundancy in the data. In practice, it is done by assigning shorter codewords to the more frequent outcomes of the data source, and longer descriptions to the less frequent outcomes.

2.9.1 Lossless source coding

The source sequence X^n is encoded into an index M at rate R bits per source symbol, and the receiver decodes the index to estimate the source sequence $\hat{X}^n$. The problem for lossless source coding is to find the lowest compression rate in bits per source symbol such that the probability of decoding error decays asymptotically to zero with the code block length n.

The Lossless Source Coding Theorem proves that the optimal rate for lossless source coding of a discrete memoryless source X is $R^* = H(x)$ which is the entropy of the source [5].

2.9.2 Lossy Source Coding

According to [5], if the source is continuous, then lossless reconstruction of the source sequence would require an infinite transmission rate![6] So for a continuous source, the reconstruction is only required to be close to the source sequence according to some fidelity criterion.

Unlike the lossless source coding where there is an optimal compression rate, the lossy source coding involves a tradeoff between the rate and the desired distortion.

Let $\hat{X}^n$ be the reconstructed alphabet: an average distortion between x^n and $\hat{x}^n$ is defined as

$$d(x^n,\hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n} d(x_i,\hat{x}_i) \tag{2.31}$$

and the expected distortion is defined as

$$E(d(X^n,\hat{X}^n)) = \sum_{x^n} p(x^n)d(x^n,\hat{x}^n) \tag{2.32}$$

[6]entropy of a continuous source is infinite, see section 2.6.1

where $p(x^n)$ is the probability of source sequence x^n been sent.

A rate-distortion pair (R,D) is said to be achievable if there exists a sequence $(2^{nR},n)$ codes with

$$\lim_{x\to\infty}\sup E(d(X^n,\hat{X}^n)) \leq D \tag{2.33}$$

the rate-distortion function $R(D)$ is the infimum of rates R such that (R,D) is achievable.

According to the Lossy Source Coding Theorem [5], the rate-distortion function for a discrete memoryless source X and a distortion measure $d(x,\hat{x})$ is

$$R(D) = \min_{p(\hat{x}|x):E(d(X,\hat{X}))\leq D} I(X;\hat{X}) \tag{2.34}$$

for $D \geq D_{min} = \min_{\hat{x}(x)} E[d(X,\hat{x}(X))]$.

2.9.3 Joint Source-Channel Coding

Shannon determined that the source and channel codes can be designed separately on an AWGN channel without any loss if there are infinite source code dimensions, channel code length and complexity. The Source-Channel Separation Theorem [32] states that for a discrete memoryless source U and a distortion measure $d(u,\hat{u})$ with rate-distortion function $R(D)$ and a discrete memoryless channel $p(y|x)$ with capacity C

$$\begin{aligned} &\text{If } rR(D) < C, \text{ then } (r,D) \text{ is achievable.} \\ &\text{If } (r,D) \text{ is achievable, then } rR(D) \leq C \end{aligned} \tag{2.35}$$

where r is the transmit rate.

Some joint source channel coding schemes have been developed to provide alternative approaches under more realistic system assumption, such as finite source code dimensions and channel code length, etc [5].

Source-Optimized Channel Coding

In source-optimized channel coding, the source code is designed for a noiseless channel, the channel code is then designed for this source to minimize end-to-end distortion over the given channel based on the distortion associated with corruption of the different bits. The unequal error protection (UEP) code which matches the bit error probability of the different component channel codes to the bit priorities associated with the source code is an example [5].

Channel-Optimized Source coding

Another approach is to optimize the source code based on the error probability associated with the channel code, where the channel code is designed independently of the source. An example of this category includes the channel-optimized vector quantizer and its scalar variation [5].

2.10 Network Information Theory

2.10.1 Network Information Flow Problem

A networked system consists of a set of information sources and communication nodes connected by a network. Each node observes one or more sources and wishes to reconstruct other

sources or to compute a function based on all the sources. In order to achieve the required task, the nodes need to communicate with each other over the network.

Network information theory tries to answer the following questions [33]:

- What is the limit on the amount of communication needed?
- How can this limit be achieved?

2.10.2 MAX-FLOW MIN-CUT Theorem

Consider a wired network, such as the Internet or a distributed storage system which can be modelled by a directed graph $(\mathcal{N}, \varepsilon)$ with link capacities C_{jk} bits between node j to node k as depicted in Figure 2.7. Assume a R-bit message is needed to transmit from source node 1 to the destination node N. What is the network capacity C that allows a reliable communication?

The answer is given by the *max − flow min − cut theorem*. It shows that the capacity is equal to the minimum cut capacity [33],

$$C = \min C(S) \tag{2.36}$$

where $C(S) = \sum_{j \in S, k \in T} C_{jk}$ is the capacity of the cut (S,T), where cut (S,T) is a partition of $\mathcal{N}$ such that $s \in S$, $t \in T$. The cut-set of (S,T) is the set $\{(u,v) \in \varepsilon : u \in S, v \in T\}$. It also shows that simple routing at the intermediate (relay) nodes is enough to achieve the capacity without errors. In other words, each node forwards its incoming bits over its outgoing links. Therefore, under this networked system model, information can be regarded as a commodity to be shipped over a transportation network.

2.10.3 Network Information Theory

The max-flow min-cut theorem and Shannon's channel capacity theory have influenced communication and networking fundamentally. However the model it uses has only a single source-destination pair communicating over a noisy networked channel which does not fit many important aspects of real-world networks [33]:

- Networked system have multiple sources and destinations
- The task of the network is often to compute a function or to make a decision
- Wireless communication uses a shared broadcast medium
- Networked systems involve complex tradeoffs between competition for resources and cooperation for common good.
- Many networks allow for feedback and interactive communication
- Source-channel separation does not hold for networks in general.
- Network security is often a primary concern.
- Data from the sources is often bursty and network topology evolves dynamically.

Network information theory tries to answer the information flow questions while capturing some of these real-world networks features.

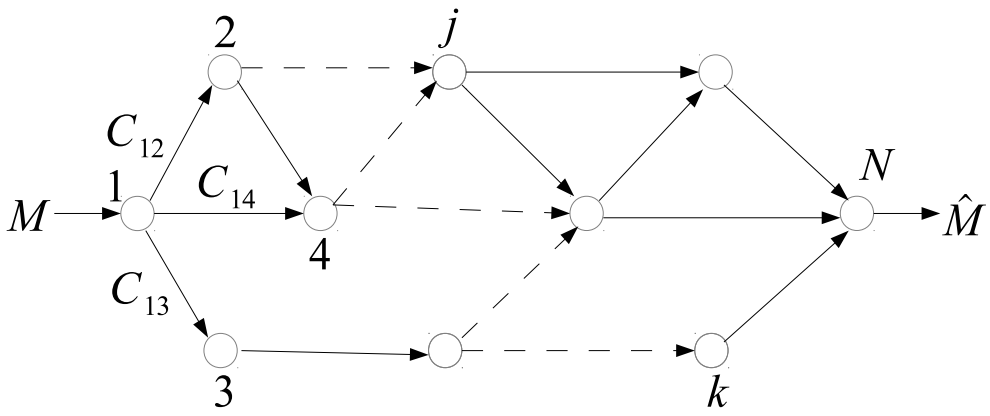

Figure 2.7: Graphical single-source single-destination network

Multiple Sources and Destinations

Routing and point-to-point source/channel coding cannot satisfy the need for networks with many sources and destinations.

1. **Graphical multicast network** [33]

   Suppose we wish to send a message to multiple destinations. Unlike the unicast scenario, routing is not optimal in general even if we treat the system as a graphical network. Instead, we need to encode the incoming packets at the relay nodes.

   Assume there is a network as shown in Figure 2.8, where function $\bigoplus()$ denotes modulo-2 plus. Source node 1 wishes to send a 2-bit message $(M_1, M_2) \in 0, 1^2$ to destination nodes 6 and 7. And assume capacity on all links is 1 bit. Therefore, by using routing only, message M_1 and M_2 can not be sent over edge (4,5) simultaneously and reach both destinations.

   However, if relay node 4 can perform modulo-2 sum operation, 2-bit message can be compressed to 1 bit and communicate to nodes 6 and 7 where another modulo-2 sum operation is performed and therefore the 2-bit message can be communicated to both destinations.

2. **Distributed Compression**

   The Slepian-Wolf theorem was introduced by David Slepian and Jack Wolf in 1973 in their paper [34]. This theorem deals with the lossless compression of distributed corre-

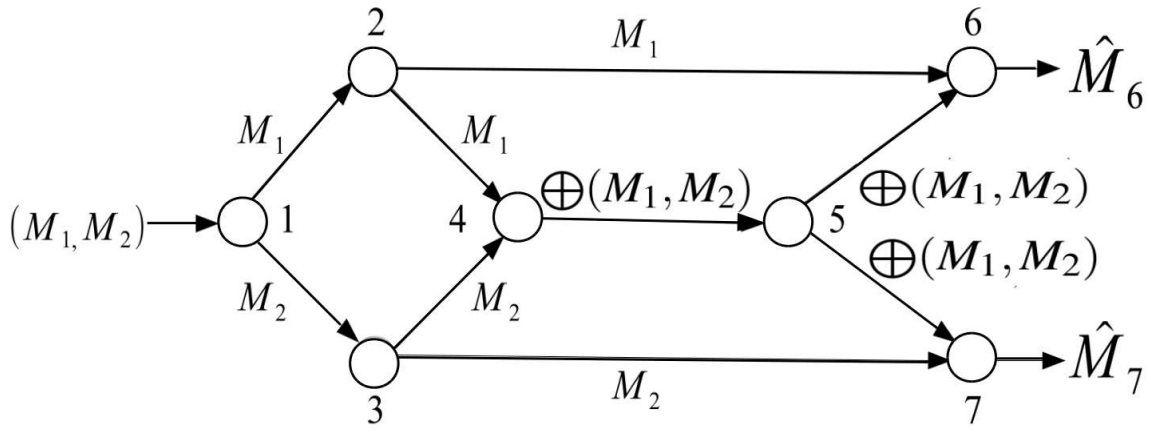

Figure 2.8: The butterfly network: 2-bit message can be sent to destination nodes 6 and 7 using linear network coding

lated sources. It indicates that each of the correlated sources can be encoded separately without the knowledge of other sources and the compressed data from all these sources can be jointly decoded with arbitrary small error probability. Theoretically proved, Slepian-Wolf coding can achieve the same compression rate as the optimal joint compression does. The concept of Slepian-Wolf theorem is shown in Figure 2.9. Source X_1 and X_2 are two correlated sources, and their joint entropy and conditional entropies are $H(X_1,X_2)$ and $H(X_1|X_2)$,$H(X_2|X_1)$. If we encode X_1,X_2 separately without exploiting the correlation between them, the achievable rate pair (R_1,R_2) is constrained by

$$\begin{aligned} R_1+R_2 &\geq H(X_1)+H(X_2) \\ R_1 &\geq H(X_1) \\ R_2 &\geq H(X_2) \end{aligned}$$

The corresponding rate region is shown in Figure 2.10 as the nonSW region. However, if we implement Slepian-Wolf code which takes into account the correlation between sources, the achievable rate region is

$$\begin{aligned} R_1+R_2 &\geq H(X_1,X_2) \\ R_1 &\geq H(X_1|X_2) \\ R_2 &\geq H(X_2|X_1) \end{aligned}$$

The corresponding rate region is shown in Figure 2.10 as the Slepian-Wolf region. By implementing a Slepian-Wolf code, less transmit rate is required, therefore less channel capacity is required.

Below is an example to demonstrate the Slepian-Wolf theorem. Assume binary sources X_1 and X_2 each contains 8 equiprobable symbols, then their entropies are $H(X_1) = H(X_2) = 3$. The correlation between two sources is that they only differ in one position at most. For example if X_1 is $\{000\}$, $X_2 \in \{000,100,010,001\}$. If encoding these two sources separately, without exploiting the correlation between them, at least $H(X_1)+H(X_2) = 6$ bits are needed. However, if taking into account their correlation,

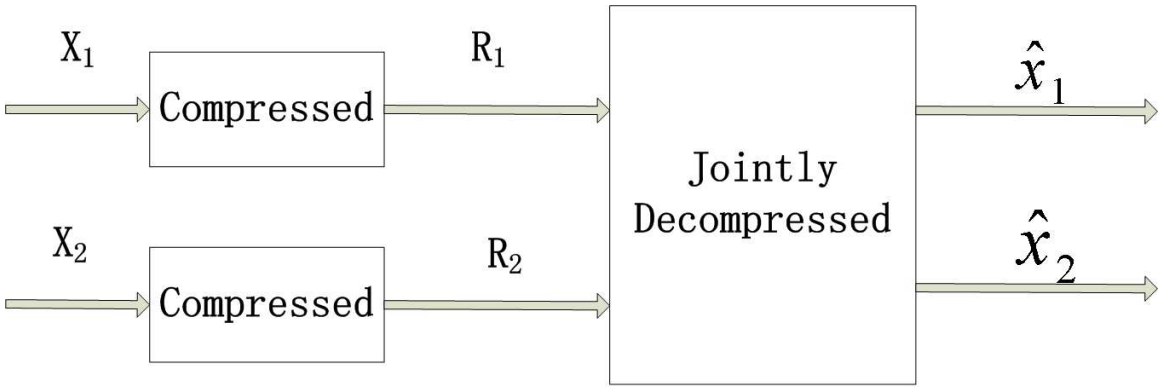

Figure 2.9: Slepian-Wolf concept, x_1 and x_2 are the estimation of source X_1 and X_2.

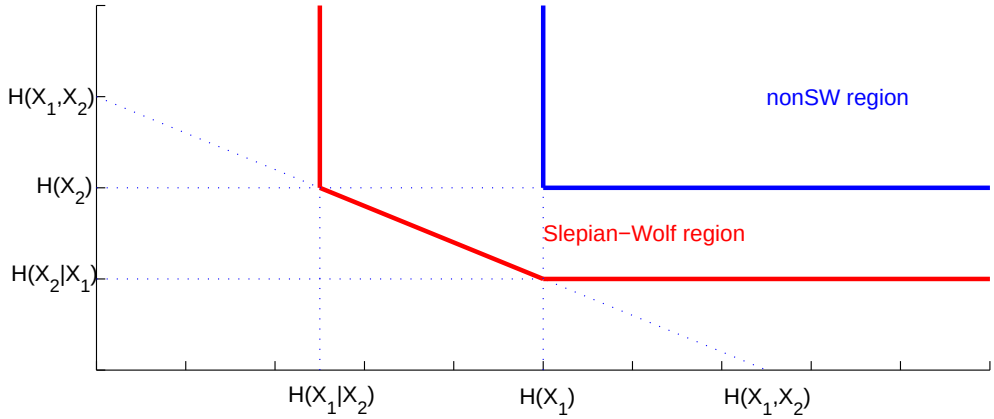

Figure 2.10: Slepian-Wolf rate region

only $H(X_1,X_2) = H(X_1) + H(X_2|X_1) = 5$ bits are needed. This can be done by first dividing X_2 into 4 subsets, namely $S_0 = \{000, 111\}, S_1 = \{001, 110\}, S_2 = \{010, 101\}$ and $S_3 = \{011, 100\}$, all symbols in each subset differ in all three positions. When sending the information of X_1 and X_2, 3 bits are required for X_1 and 2 bits are required for the indices of the subsets in X_2, which makes the total transmit rate 5 bits. If $\{010\}$ and $\{11\}$ are received from X_1 and X_2 respectively, X_1 can be perfectly decoded as $\{010\}$, whereas X_2 can be $\{011, 100\}$ with equal probability. However, since the joint decoder knows the correlation between the two sources, it will pick $\{011\}$ as the estimation of X_2 accordingly.

Wireless Networks

Network information theory plays an important role in dealing with wireless channels.

Traditionally a wireless network can be treated as a set of separate point-to-point channels via time or frequency division. This approach however, can not take full advantage of the wireless medium.

1. **Gaussian broadcast channel**

The sender wishes to communicate with two receivers via a Gaussian broadcast channel.

$$\begin{aligned} Y_1 &= g_1 X + Z_1 \\ Y_2 &= g_2 X + Z_2 \end{aligned} \tag{2.37}$$

where $Z_1 \sim N(0, N_0/2)$ and $Z_2 \sim N(0, N_0/2)$ are the receiver noise components, and $g_1^2 > g_2^2$ which means the channel to receiver 1 is stronger than the channel to receiver 2.

If we send messages M_1 and M_2 separately via different time intervals or frequency bands, then we can treat the broadcast channel as separate point-to-point channels. However, we may use a superposition coding scheme [35], namely add the codewords of the two messages together and send the sum to both receivers. The stronger receiver decodes both codewords while the weaker receiver treats the other codeword as noise and only decodes its own codeword. The reliable rate pair region for this scheme is larger than the rate pair region for the separate point-to-point channels scheme.

2. **Gaussian Vector Broadcast Channel** If links within the network contain MIMO channels, then the coding techniques for single-antenna (scalar) channels cannot satisfy the need. For each receiver in the network, the received signal consists of three parts: the desired message, receiver noise and interference which is independent of the Gaussian noise.

   When the interference sequence is available at the receiver, it can be subtracted from the received message and hence the channel capacity will not be affected by the interference. Assume the interference is only available at the transmitter, because of the power constraint, it is not always possible to subtract the interference before transmitting. However, by applying some judicious precoding scheme, even though the interference can not be completely removed, the channel capacity will not be affected.

   The scheme is applied as the following

   - To communicate M_2 to receiver 2, the received message is $Y_2 = g_2 X_2 + g_2 X_1 + Z_2$ with desired message X_2, interference $g_2 X_1$ and AWGN Z_2. Receiver 2 recovers M_2 while treats the interference $g_2 X_1$ as part of the AWGN.
   - To communicate M_1 to receiver 1, the received message is $Y_1 = g_1 X_1 + g_1 X_2 + Z_1$ with desired message X_1, interference $g_1 X_2$ and AWGN Z_1. The interference $g_1 X_2$ is assumed available at the transmitter. By applying the dirty paper scheme, the transmission rate can be as close as that for the channel $Y' = g_1 X_1 + Z_1$ which has no interference.

3. **Gaussian Relay Channel** In Gaussian relay networks, some nodes can act as relays to help other nodes communicate their message. Treating a network like this as a set of separate point-to-point links cannot fully exploit its advantage because cooperation between nodes can increase the transmit rate greatly.

   Assume there is a node between the transmitter and the receiver, functioning as a relay to help the transmitter communicate message to the receiver. In the commonly used scheme, the relay can receive the message that is sent to the receiver by the transmitter. The relay recovers the message and re-sends it to the receiver which may become an interference to the original message. If the distance between the transmitter and the

receiver is large, this scheme performs well because the interference due to the concurrent transmission is weak, while if the distance is not large, the interference can adversely affect the communication of the message.

- **Decode and Forward Relay** [33]
  For the relay in decode and forward (DF) mode, it detects and decodes[7] the information in the received signal, re-modulates the decoded information and transmits the signal to the next node. Due to the complexity of the detecting and the decoding process, DF relay can consume considerable amount of energy. By exploiting the modulation and the code, DF relay can reduce the errors caused by the fading channel and the noise, however if there are remaining errors after the decoding process, they will be passed on to the destination and cannot be detected. This disadvantage leads to a variant of DF mode, called Soft Decode and Forward Relay
  For soft decode and forward relay, instead of taking the hard decision at the output of the decoder, soft information is taken and quantized, then the quantized information is re-modulated and transmitted to the next node. At the destination, soft information from all relays is combined in some way before taking the final hard decision. During the combination at the destination, soft information of the unreliable bits will be corrected by their reliable counterparts, so the overall bit error rate will be decreased.
- **Amplify and Forward Relay** [33]
  In amplify and forward (AF) relaying, the relay receives a signal from the other node and re-transmits an amplified version to the next node. Because it does not need to detect or decode the information in the signal, the computation complexing is low, therefore less energy is needed. CSI may need to calculate the amplification factor at each relay. Because AF mode operates on the symbol level, all the undesirable fading and noise information are also forwarded which decreases the system's spectrum efficiency.
- **Compressed and Forward Relay** [33]
  As an alternative to the decode and forward relay, compress and forward relaying not only decodes the received message but also compresses the decoded message and sends the compression index to the receiver. Note that some compress and forward relays quantize the received message instead of decoding it and compress the quantization index. The advantage of this scheme is obvious, it can save bandwidth. However this scheme is sensitive to error on the relay destination link.

4. **Wireless Fading Channels** Wireless channels gains are sometime time varying because of the signal scattering over multiple paths and user mobility. Network information theory also studies channel models and their corresponding capacity that allow the gains in the Gaussian channels to vary randomly with time.

Joint Source-Channel Coding

For point-to-point communication, Shannon shows that separate source and channel coding is asymptotically optimal. However, it turns out that such separation does not hold in general for

[7] If there is no code used in the system, only detecting is performed.

sending correlated sources over multiuser networks. Thus in some multiuser settings it is advantageous to perform joint source-channel coding. Network information theory also discusses the natural definitions of various notions of common information between two sources.

Secure Communication

Confidentiality of information is a vital requirement in networking applications such as online-purchasing. Assume a legitimate costumer (A) wants to communicate a message to a service provider (B) while keeping it secret from a eavesdropper (C). When the link between A and B is stronger than that to C, a confidential message can be communicated reliably even without a shared secret key between A and B. On the other hand, if the link between A and B is weaker than that to C, no unprotected confidential message can be communicated reliably. Network information theory shows that A and B can agree on a secret key through a public communication channel that C has access to. This key can then be used to communicate a confidential message.

2.11 Design of Capacity-Achieving Channel Code

Among the family of concatenated error correction codes, some have potential to yield a decoding performance close to theoretical limits (Shannon Bound). However efficient decoding was considered extremely complex. [36] shows that iterative decoding scheme allows a concatenated code to decode almost optimally without high computation burden.

Generally, there are two methods to design an iterative decoded concatenated code that can approach the theoretical limits, namely curve-fitting [37] and density evolution [38, 39], however because we decide to use the curve-fitting method for the optimization in the future, we will focus our attention to this method here.

2.11.1 Curve-Fitting Method

According to [30], the size of the area between two decoding curves in an EXIT chart is proportional to the performance loss of the corresponding decoder. Therefore, code design problem becomes curves fitting problem on the EXIT chart.

Below is an example of using curve-fitting method to design a repeat-accumulate code [37].

Consider a repeat-accumulate code with rate $R = k/n$ as its encoder and decoder structure are shown in Figure 2.11 and Figure 2.12. The encoder consists of four main parts: k repetition codes with variable rates, an edge interleaver, $n-k$ single parity check codes and a differential encoder. The repetition codes are regarded as variable nodes, the single parity check codes as check nodes and the differential encoder as accumulator.

The node degrees are defined as the number of edges from an interleaver. Thus the ith variable node has degree of $d_v^{(i)}$, though it has $d_v^{(i)}+1$ edges. Similarly, the ith check node has degree of $d_c^{(i)}$. The rate of the ith repetition code is $1/d_v^{(i)}$ and the ith single parity check code is $d_c^{(i)}$.

For the decoder, during the decoding process, all information is represented in log-likelihood ratio (LLR) form. The decoding procedures are applied as follows: $n-k$ parity bits from the

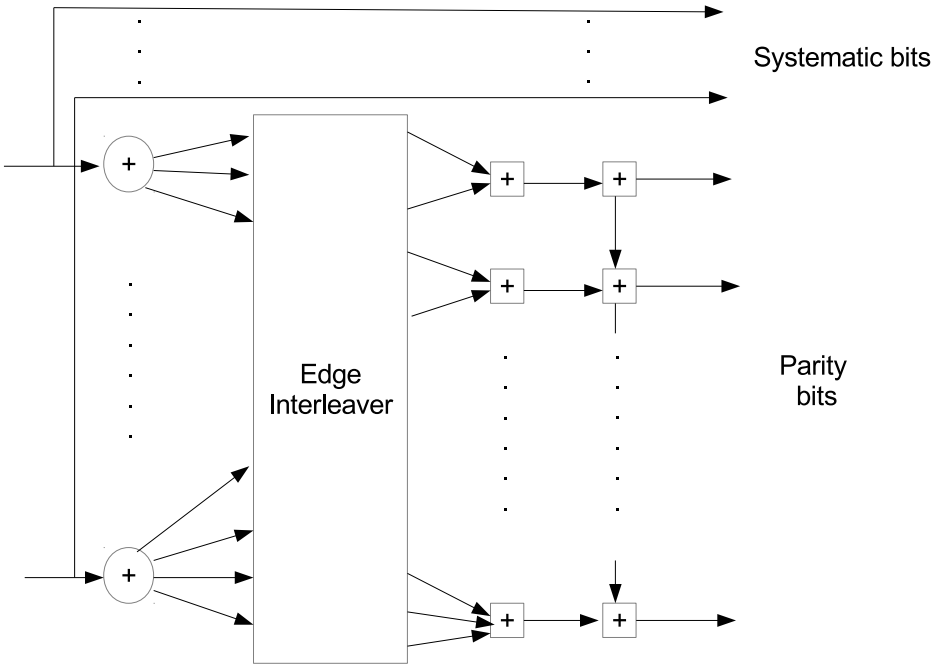

Figure 2.11: Encoder for a systematic RA code

channel in form of the LLR are taken by the inner decoder which performs the BCJR decoding over the memory-one trellis of the accumulator. The soft results which are then input into the check node decoder (CND) consist of the outputs of $n-k$ check nodes. The soft results generated by the CND are forwarded through the edge interleaver to the outer variable node decoder (VND) which contains k variable nodes. Besides the input from the edge interleaver, VND also takes k systematic bits from the channel in the form of the LLR. The soft outputs produced by the VND are fed back through the edge interleaver to the CND. One iteration is completed when the CND generates soft outputs and inputs them back to the de-accumulator. The De-accumulator and the CND are often considered as one single decoder treated as the inner decoder whereas the VND is treated as the outer decoder.

1. **EXIT Curve of the Outer Variable Node Decoder**

   A variable node of degree d_v has d_v+1 incoming messages of which d_v come from the edge interleaver and one from the channel. The decoding process of the variable node decoder is

   $$L_{i,out} = L_{ch} + \sum_{j \neq i} L_{j,in} \tag{2.38}$$

   which simply adds all incoming LLRs together. $L_{j,in}$ is the jth a priori LLR going into the VND, $L_{i,out}$ is the ith extrinsic LLR coming out of the VND and L_{ch} is the LLR of the systematic bit from the channel.

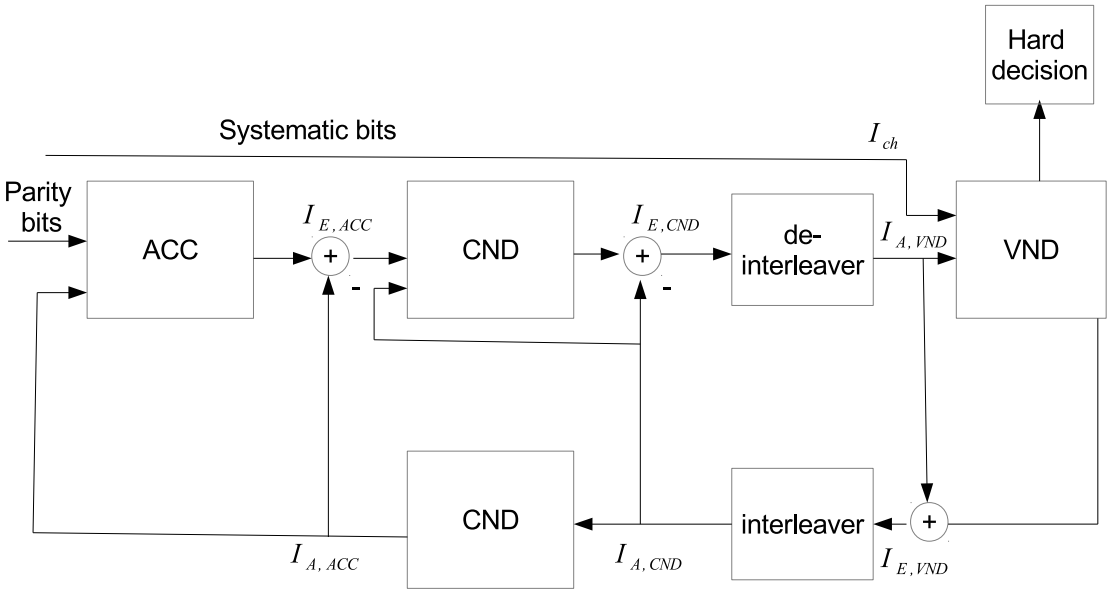

Figure 2.12: Decoder for a systematic RA code

Consider an AWGN channel with BPSK modulation and noise variance σ_n^2.

$$Y = X + N \tag{2.39}$$

where $Pr(X = \pm 1) = 1/2$ and N is zero-mean Gaussian noise with variance σ_n^2.
The LLR from the channel is

$$L_{ch}(y) = \log \frac{p(y|x=+1)}{p(y|x=-1)} = \frac{2}{\sigma_n^2} y \tag{2.40}$$

where $p(y|x=\pm 1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y\mp x)^2}{2\sigma^2}}$. The mean and variance of $L_{ch}(Y)$ conditioned on $X = \pm 1$ are

$$\begin{aligned} \mu_{ch} &= \frac{\pm 2}{\sigma_n^2} = \pm 4R\frac{E_b}{N_0} \\ \sigma_{ch}^2 &= \frac{4}{\sigma_n^2} = 8R\frac{E_b}{N_0} \end{aligned} \tag{2.41}$$

where R is the code rate.

Assume $L_{j,in}$ is the channel output of the jth interleaved bit transmitted using BPSK through an AWGN channel. Hence, the EXIT function of a VND with degree d_v is

$$I_{E,VND}(I_A, d_v, \frac{E_b}{N_0}, R) = J\left(\sqrt{(d_v - 1)[J^{-1}(I_A)]^2 + \sigma_{ch}^2}\right) \tag{2.42}$$

where $J(\sigma_{ch})$ is the mutual information $I(X;L_{ch}(Y))$.

$$\begin{aligned} J(\sigma_{ch}) &= H(X) - H(X|L_{ch}(Y)) \\ &= 1 - \int_{-\infty}^{\infty} \frac{e^{-(\xi-\sigma_{ch}^2/2)^2/2\sigma_{ch}^2}}{\sqrt{2\pi\sigma_{ch}^2}} \log_2[1+e^{-\xi}]d\xi \end{aligned} \tag{2.43}$$

where $H(x)$ is the entropy of X and $H(X|L_{ch}(Y))$ is the entropy of X conditioned on $L_{ch}(Y)$. Because LLR contains the same amount of information as the signal itself does, $I(X;L_{ch}(Y))$ is the same as $I(X;Y)$. The capacity of this channel is, therefore, $J(\sigma_{ch}) = J(2/\sigma_n)$.

As shown in [40], for practical application, split $J()$ into two parts, one for the interval $0 \le \sigma \le \sigma^*$ and another one for the interval $\sigma^* \le \sigma$ where $\sigma^* = 1.6363$. Use a polynomial fit for the left interval and an exponential fit for the right interval. After applying the Marquardt-Levenberg algorithm,

$$J(\sigma) \approx \begin{cases} a_{J,1}\sigma^3 + b_{J,1}\sigma^2 + c_{J,1}\sigma & 0 \le \sigma \le \sigma^* \\ 1 - e^{a_{J,2}\sigma^3 + b_{J,2}\sigma^2 + c_{J,2}\sigma + d_{J,2}} & \sigma^* \le \sigma \le 10 \\ 1 & \sigma \ge 10 \end{cases} \tag{2.44}$$

where

$$\begin{aligned} &a_{J,1} = -0.0421061 \quad b_{J,1} = 0.209252 \quad c_{J,1} = -0.00640081 \\ &a_{J,2} = 0.00181491 \quad b_{J,2} = -0.142675 \quad c_{J,2} = -0.0822054 \quad d_{J,2} = 0.0549608 \end{aligned}$$

For the inverse $J^{-1}()$ function, it is also split into two intervals

$$J^{-1}(I) \approx \begin{cases} a_{\sigma,1}I^2 + b_{\sigma,1}I + c_{\sigma,1}\sqrt{I} & 0 \le I \le I^* \\ -a_{\sigma,2}\ln[b_{\sigma,2}(1-I)] - c_{\sigma,2}I & I^* \le I \le 1 \end{cases} \tag{2.45}$$

where

$$\begin{aligned} &I^* = 0.3646 \\ &a_{\sigma,1} = 1.09542 \quad b_{\sigma,1} = 0.214217 \quad c_{\sigma,1} = 2.33727 \\ &a_{\sigma,2} = 0.706692 \quad b_{\sigma,2} = 0.386013 \quad c_{\sigma,2} = -1.75017 \end{aligned}$$

Because a non-systematic RA code does not have systematic bits from the channel, $\sigma_{ch} = 0$. Hence, the EXIT function of a non-systematic VND with degree d_v is

$$I_{E,VND}(I_A, d_v, \frac{E_b}{N_0}, R) = J\left(\sqrt{(d_v-1)[J^{-1}(I_A)]^2}\right) \tag{2.46}$$

Figure 2.13 shows some EXIT chart curves for the VND with rate $R = 1/2$ at $E_b/N_0 = 0.5$ dB. Systematic nodes are plotted in solid lines while non-systematic nodes are plotted in dashed lines. As the degree increases, the EXIT chart shows that the VND has strong error correction capability. The array of degree shown on this plot is [2 3 4 6 8 16]. When the VND is non-systematic, its EXIT chart curve starts at the origin, and if the inner decoder also starts at the origin, the decoding process will not begin to converge. Hence, for non-systematic code, a biregular RA code is used whose inner decoder EXIT chart curve does not start at the origin [41].

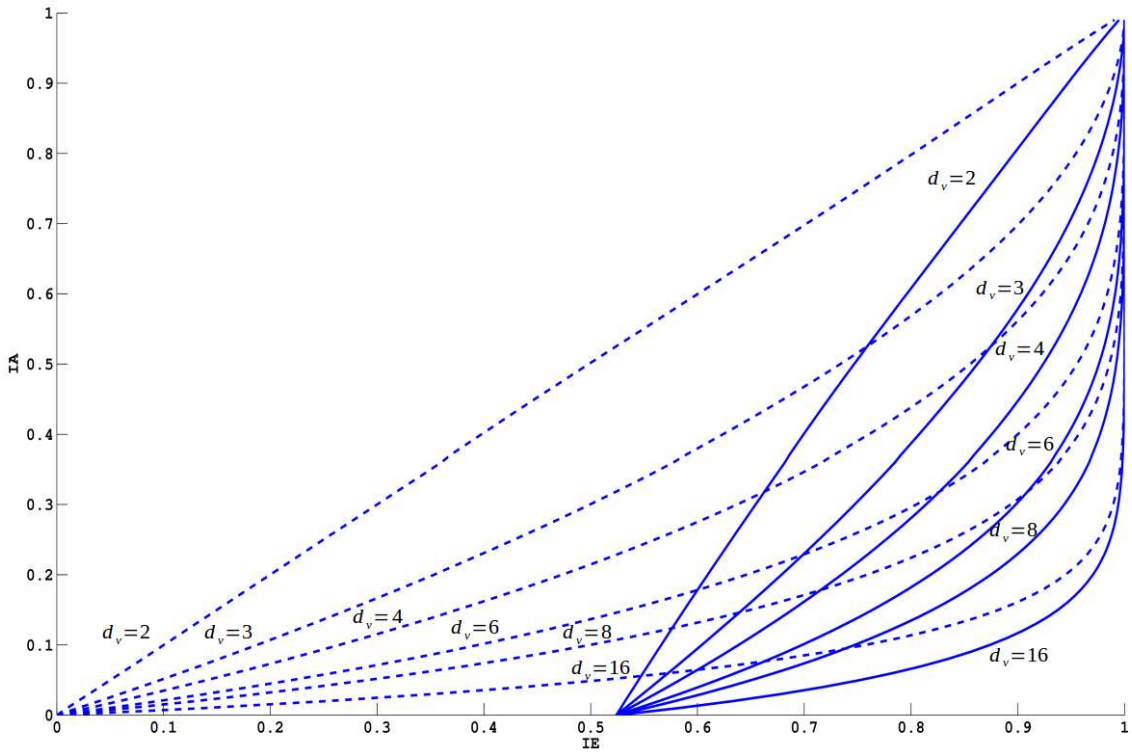

Figure 2.13: EXIT chart curves for the VND with rate $R = 1/2$ at $E_b/N_0 = 0.5$ dB. Systematic nodes are plotted in solid lines while non-systematic nodes are plotted in dashed lines. The array of degree shown on this plot is $[2 \quad 3 \quad 4 \quad 6 \quad 8 \quad 16]$. For both systematic and non-systematic node, as the degree increases, the corresponding curves are moving downwards.

2. **EXIT Curve of the Inner Check Node Decoder**

   Consider a CND of degree d_c. It has $d_c + 1$ incoming messages, in which d_c from the edge interleaver and one from the accumulator. The decoding equation of a CND is [42]

$$L_{i,out} = \ln \frac{1 - \prod_{j \neq i} \frac{1 - e^{L_{j,in}}}{1 + e^{L_{j,in}}}}{1 + \prod_{j \neq i} \frac{1 - e^{L_{j,in}}}{1 + e^{L_{j,in}}}} \tag{2.47}$$

   where the indexes i and j include all $d_c + 1$ messages from the edge interleaver and the accumulator.

   Assume $L_{j,in}$ is the channel output of the jth bit from the check node transmitted using BPSK through an AWGN channel. The analysis of the CND is split into two parts based on the direction of the information flow. One is for CND-to-ACC direction, where edge interleaver feeds information back to the accumulator. The other is for CND-to-interleaver direction, where information is fed through the edge interleaver to the VND.

   - *CND-to-ACC direction*

     The closed form of the CND EXIT chart curves are shown in [43, 44]. Those curves can also be obtained by simulation. Alternatively, for the binary erasure channel,

a duality property exists [45, 46] which can represent the EXIT chart curve of the degree d_c single parity-check code in terms of the EXIT chart curve of a repetition code with degree $1/d_c$.

$$I_{E,SPC(I_A,d_c)} = 1 - I_{E,REP}(1 - I_A, d_c) \tag{2.48}$$

Although, this property does not exist for BPSK/AWGN a priori inputs, the distortion is negligible [43, 44]. Hence

$$I_{A,ACC}(I_A, d_c) \approx 1 - I_{E,REP}(1 - I_A, d_c + 1) = 1 - J(\sqrt{d_c} J^{-1}(1 - I_A)) \tag{2.49}$$

- *CND-to-interleaver direction*

  Each CND uses one LLR from the accumulator and $d_c - 1$ LLRs from the edge interleaver to compute its extrinsic information. By applying the duality, the EXIT chart curve of the CND-to-interleaver direction can be expressed approximately as

$$\begin{aligned} I_{E,CND}(I_A, I_E, d_c) &\approx 1 - I_{E,REP}(1 - I_A, 1 - I_E, d_c) \\ &= 1 - J(\sqrt{(d_c - 1)[J^{-1}(1 - I_A)]^2 + [J^{-1}(1 - I_E)]^2}) \end{aligned} \tag{2.50}$$

  where $I_A = I_{A,CND}$ and $I_E = I_{E,ACC}$.

Figure 2.14 shows some CND EXIT chart curves with various degrees at $E_b/N_0 = 0.5$ dB and $R = 1/2$. CND-to-ACC direction curves are plotted as dashed lines, whereas CND-to-interleaver direction curves are represented by solid lines. As the degree increases, the EXIT shows that the error correction capability of the CND decreases. The degree array shown on this plot is $[1 \quad 2 \quad 3 \quad 4 \quad 6 \quad 8 \quad 16]$. For the case when $d_c = 1$, there is one edge to the edge interleaver. Hence for the CND-to-ACC direction, $I_{A,ACC} = I_{A,CND}$ and for the CND-to-interleaver direction, $I_{E,CND} = I_{E,ACC}$. Therefore, the EXIT chart curves are a diagonal line and a horizontal line respectively.

3. **EXIT Curve of the Inner Accumulator Decoder**

   The EXIT curve of the accumulator decoder is calculated by simulation and represented as

$$I_{E,ACC}(I_{A,ACC}, \frac{E_b}{N_0}, R) \tag{2.51}$$

   Figure 2.15 shows some accumulator EXIT chart curves under various $\frac{E_b}{N_0}$, as the $\frac{E_b}{N_0}$ increases, the curve moves upwards. The $\frac{E_b}{N_0}$ array shown on this plot is $[-2 \quad -1 \quad 1 \quad 2]$ dB.

   [46] gives the accumulator curves for BECs that have the same capacity as the AWGN.

$$I_{E,ACC}(I_{A,ACC}, q) = \left[\frac{1 - q}{1 - qI_{A,ACC}}\right]^2 \tag{2.52}$$

   where $1 - q = C(E_b/N_0)$ and $C(E_b/N_0)$ is the capacity in bits per channel use with binary antipodal modulation for the give E_b/N_0.

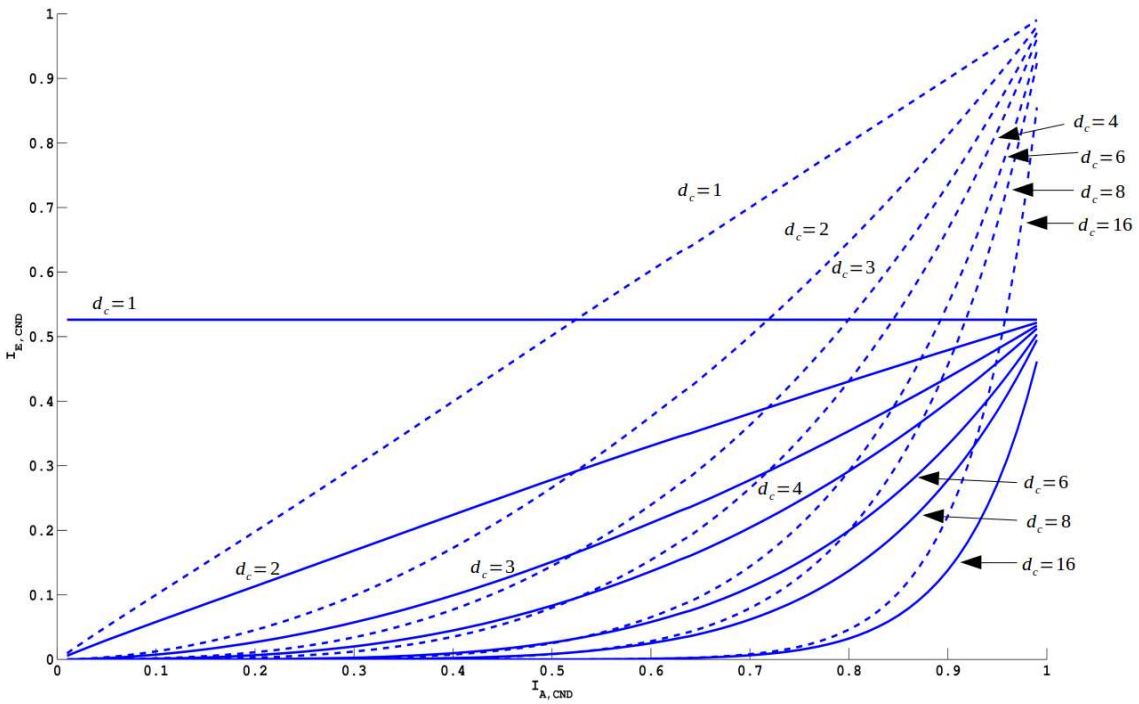

Figure 2.14: CND EXIT chart curves with various degrees at $E_b/N_0 = 0.5$ dB and $R = 1/2$. CND-to-ACC direction curves are plotted as dashed lines, whereas CND-to-interleaver direction curves are represented by solid lines. The degree array shown on this plot is $[1 \quad 2 \quad 3 \quad 4 \quad 6 \quad 8 \quad 16]$. For both directions, as the degree increases, the corresponding curves are moving downwards.

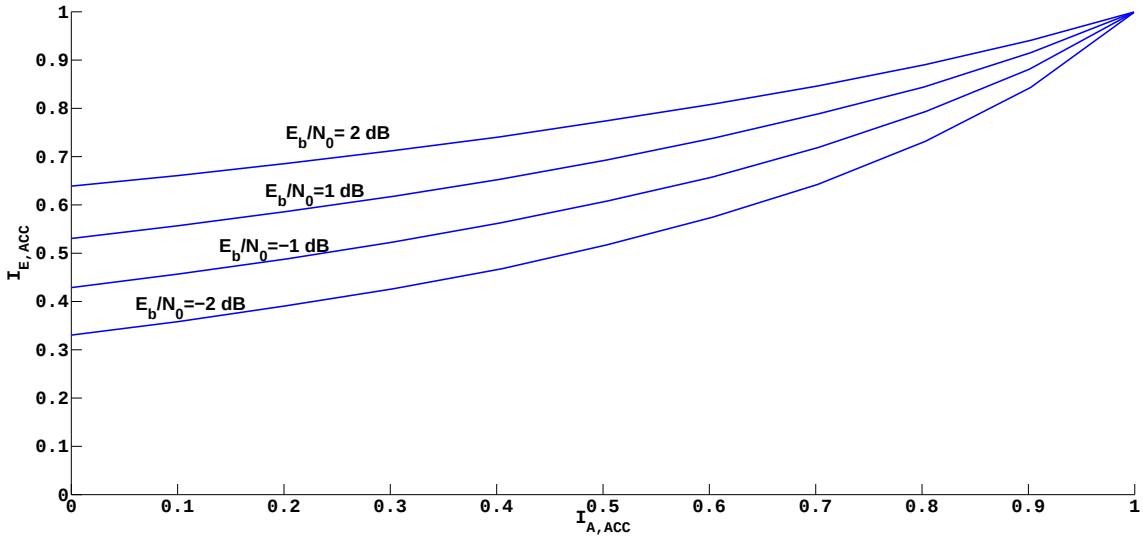

Figure 2.15: Accumulator EXIT chart curves under various $\frac{E_b}{N_0}$, as the $\frac{E_b}{N_0}$ increases, the curve moves upwards. The $\frac{E_b}{N_0}$ array shown on this plot is $[-2 \quad -1 \quad 1 \quad 2]$ dB.

4. **EXIT Curve of the Combined Inner ACC and CND**

   Combining the EXIT chart function of the accumulator and the check node decoder, the result is

$$\begin{aligned} I_{E,ACC\&CND} &= \left(I_{A,CND}, d_c, \frac{E_b}{N_0}, R\right) \\ &= I_{E,CND}\left(I_{A,CND}, I_{E,ACC}\left(I_{A,ACC}, \frac{E_b}{N_0}, R\right), d_c\right) \\ &= I_{E,CND}\left(I_{A,CND}, I_{E,ACC}\left(I_{A,ACC}(I_{A,CND}, d_c), \frac{E_b}{N_0}, R\right), d_c\right) \end{aligned} \tag{2.53}$$

   Figure 2.16 shows some EXIT chart curves for the combined inner accumulator and CND decoder with various degrees. When the degree increases, the EXIT chart curves of the combined decoder shows that its error correction capability decreases. The degree array shown on this plot is $[3 \quad 4 \quad 6 \quad 8 \quad 16]$. For the case of $d_c = 1$, the combined decoder is effectively simplified to an accumulator decoder as shown in Figure 2.15. Note that all curves on the plot with $d_c > 1$ start at the origin, which also occurs for the non-systematic VND curves in Figure 2.13, which means that the decoder cannot even begin to converge, hence the non-systematic RA code with $d_c > 1$ is useless. In order to allow the non-systematic RA code to work, a fraction of the check nodes that have degree one must not be empty. This technique is call code doping [41], it ensures that the decoder can make progress in the first iteration.

   Instead of measuring $I_{E,ACC}$ and computing $I_{E,ACC\&CND}$, $I_{E,ACC\&CND}$ can be measured directly. Because the latter curve stays fixed, it is more practical for the curve fitting procedure.

5. **EXIT Curve for Code Mixtures**

   - *Systematic RA Code*

     Consider a systematic RA code consists of k variable nodes and $n-k$ check nodes. Assume this RA code is check-regular, which means all of them have same degree d_c. Hence the code design only involves specifying the variable node degrees $d_v^{(i)}, i = 1, \ldots, k$. Let D_v be the number of different variable node degrees, and denote these degrees by $\tilde{d_{v,i}}, i = 1, \ldots, D_v$. Therefore, the average variable node degree is

$$\bar{d_v} = \sum_{i=1}^{D_v} a_{v,i}\tilde{d_{v,i}} \tag{2.54}$$

     where $a_{v,i}$ is the fraction of nodes that have degree $\tilde{d_{v,i}}$ and intuitively, $\sum_{i=1}^{D_v} a_{v,i} = 1$. And furthermore,

$$\begin{aligned} k\bar{d_v} &= (n-k)d_c \\ \bar{d_v} &= \frac{1-R}{R}d_c \end{aligned} \tag{2.55}$$

     since VND and CND have the same number of edges to the edge interleaver.

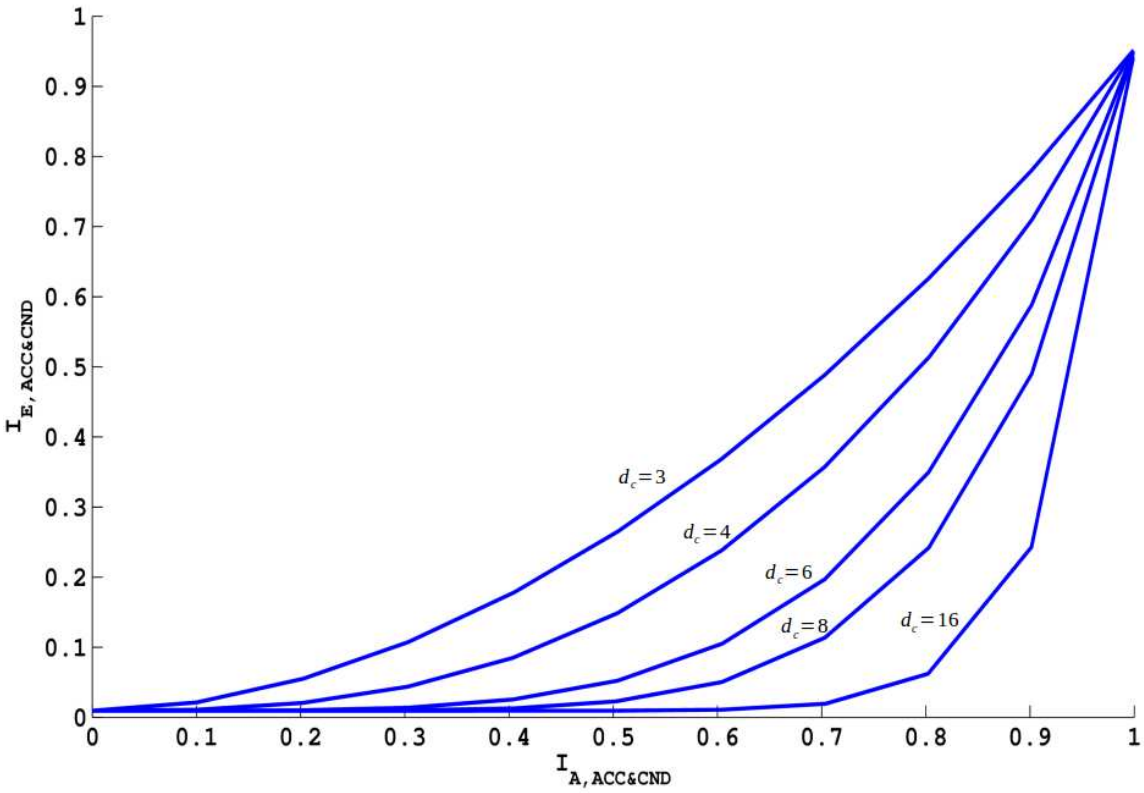

Figure 2.16: EXIT chart curves for the combined inner accumulator and CND decoder with various degrees. The degree array shown on this plot is [3 4 6 8 16]. As the degree increases, the corresponding curve is moving downwards.

Let $b_{v,i}$ be the fraction of edges connected to the variable nodes with degree $\tilde{d_{v,i}}$. The total number of edges connected to those variable nodes with degree $\tilde{d_{v,i}}$ is $(ka_{v,i})\tilde{d_{v,i}}$. Therefore,

$$b_{v,i} = ((ka_{v,i})\tilde{d_{v,i}})/(k\bar{d_v}) = \frac{R\tilde{d_{v,i}}}{(1-R)d_c}a_{v,i} \tag{2.56}$$

[47, 46] shows that the EXIT curve of a mixture of codes is the average of the component EXIT curves. Because the extrinsic information on each edge is different, by averaging over $b_{v,i}$, the VND EXIT curve is

$$I_{E,VND}\left(I_A, \frac{E_b}{N_0}, R\right) = \sum_{i=1}^{D_v} b_{v,i} I_{E,VND}\left(I_A, \tilde{d}_{v,i}, \frac{E_b}{N_0}, R\right) \tag{2.57}$$

In order to satisfy 2.55 and $\sum_{i=1}^{D_v} a_{v,i} = 1$, only $D_v - 2$ edge fractions can be adjusted, in other words, in order to perform curve fitting, $D_v > 3$ must be satisfied.

- *Non-systematic RA Code*

  Now consider a non-systematic RA code. which has n check nodes. As mentioned above, in order to trigger the convergence, an irregular CND is used. Assume $D_c =$

$\tilde{d}_{v,1} = 2$	$\tilde{a}_{v,1} = 0.0631$	$\tilde{b}_{v,1} = 0.0210$
$\tilde{d}_{v,2} = 3$	$\tilde{a}_{v,2} = 0.6306$	$\tilde{b}_{v,2} = 0.3153$
$\tilde{d}_{v,13} = 13$	$\tilde{a}_{v,3} = 0.3063$	$\tilde{d}_{v,3} = 0.6637$

Table 2.1: Curve fitting parameters for systematic RA code

2 and the average check node degree is

$$\bar{d_c} = \sum_{i=1}^{D_c} a_{c,i}\tilde{d}_{c,i} = a_{c,1} + a_{c,2}\tilde{d}_{c,2} \tag{2.58}$$

where $a_{c,i}$ is the fraction of the check nodes with degree $\tilde{d}_{c,i}$ and $\tilde{d}_{c,1} = 1$. Since VND and CND have the same number of edges to the edge interleaver,

$$\begin{aligned} k\bar{d_v} &= n\bar{d_c} \\ \bar{d_v} &= \frac{\bar{d_c}}{R} \end{aligned} \tag{2.59}$$

Let $b_{v,i}$ be the fraction of edges connected to the variable nodes with degree $\tilde{d_{v,i}}$, it is now

$$b_{v,i} = (R\tilde{d}_{v,i})/(\bar{d_c})a_{v,i} \tag{2.60}$$

The EXIT curves of the irregular CND can be obtained either by simulation or by using

$$I_{E,ACC\&CND}\left(I_A, \frac{E_b}{N_0}, R\right) = \sum_{i=1}^{2} b_{c,i} I_{E,ACC\&CND}\left(I_A, \tilde{d}_{c,i}, \frac{E_b}{N_0}, R\right) \tag{2.61}$$

where $b_{c,i} = (\tilde{d}_{c,i}/\bar{d_c})a_{c,i}$.

Figure 2.17 shows the EXIT chart curve of combined ACC and CND with mixed degrees under $E_b/N_0 = 5$ dB. As $a_{c,1}$ increases, the EXIT chart shows that the error correction capability of the combined decoder increases. The degree array is set to $\tilde{d}_{c,1} = 1$ and $\tilde{d}_{c,2} = 3$, the fraction array is set to $a_{c,1} = [0.5 \quad 0.2 \quad 0]$;

6. **Design Examples**

   Consider a SISO AWGN channel with a half rate RA code and BPSK modulation. Set $D_v = 3$ and fit the EXIT curve of the VND to that of the ACC & CND.

   Figure 2.18 shows the EXIT chart curves of a systematic RA code with $d_c = 6$ and the curve fitting parameters in Table 2.1 under $E_b/N_0 = 0.5$dB.

   Figure 2.19 shows the EXIT chart curves of a non-systematic RA code with the curve fitting parameters in Table 2.2 under $E_b/N_0 = 0.5$dB.

   Figure 2.20 shows the BER performance of a systematic RA code with the curve fitting parameters as shown in Table 2.1. The BER cliff occurs around -2.8dB, within 0.2dB, BER drops from 10^{-2} to 10^{-5}.

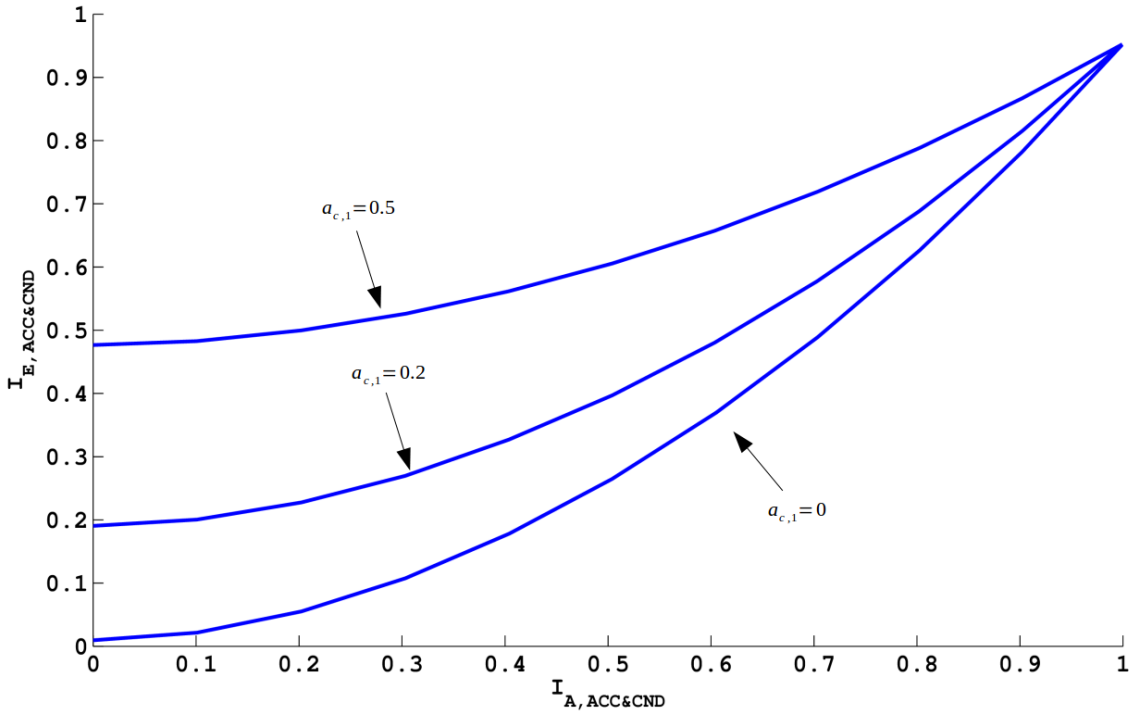

Figure 2.17: EXIT chart curve of combined ACC and CND with mixed degrees under $E_b/N_0 =$ 5 dB. As $a_{c,1}$ increases, the corresponding curve is moving upwards. The degree array is set to $\tilde{d}_{c,1} = 1$ and $\tilde{d}_{c,2} = 3$, the fraction array is set to $a_{c,1} = [0.5 \quad 0.2 \quad 0]$

$a_{c,1} = 0.2$	$\tilde{d}_{c,1} = 1$	$\tilde{d}_{c,2} = 3$
$\tilde{d}_{v,1} = 2$	$\tilde{a}_{v,1} = 0.2789$	$\tilde{b}_{v,1} = 0.1073$
$\tilde{d}_{v,2} = 4$	$\tilde{a}_{v,2} = 0.5258$	$\tilde{b}_{v,2} = 0.4044$
$\tilde{d}_{v,13} = 13$	$\tilde{a}_{v,3} = 0.1953$	$\tilde{d}_{v,3} = 0.4883$

Table 2.2: Curve fitting parameters for non-systematic RA code

A BER error floor is observed near 10^{-5}. There are several reasons which contribute to this error floor: degree 2 variable node can weaken the code's distance spectrum especially when it takes a large percentage of all variable nodes. Hence the EXIT convergence tunnel becomes narrow near $(I_A, I_E) = (1, 1)$. The limited girth of the interleaver also contributes to the error floor. It introduces correlations between the extrinsic LLR, and the BEC stopping sets [48, 49] cause the convergence to slow down.

7. **Curve Fitting for MIMO Communication Link**

   For a systematic RA code, the detector must be taken into account for both the inner and outer EXIT curves, because both of them receive LLR from the channel. Hence the design of an iterative detecting and decoding scheme for systematic RA code becomes complex. On the other hand, for a non-systematic RA code, only the inner decoder is associated with the detector and furthermore, it can perform as well as the systematic RA code, thus a non-systematic code is considered in this part. The new inner decoder

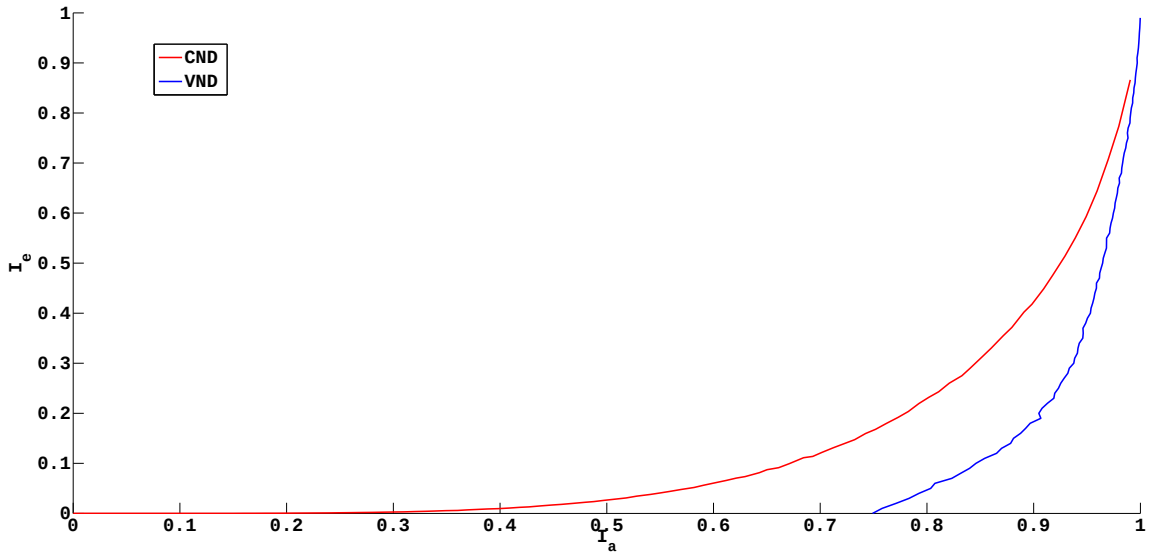

Figure 2.18: EXIT chart curves of a systematic RA code with $d_c = 6$ and the curve fitting parameters in Table 2.1 under $E_b/N_0 = 0.5$dB.

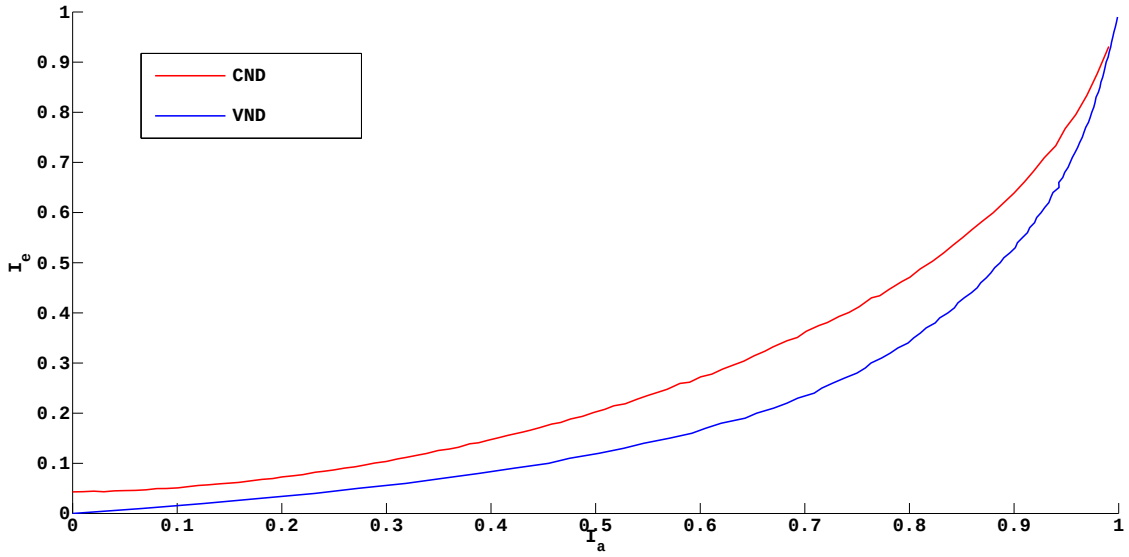

Figure 2.19: EXIT chart curves of a non-systematic RA code with the curve fitting parameters in Table 2.2 under $E_b/N_0 = 0.5$dB.

consists of the detector, the ACC decoder and the CND. The transfer curve of the outer decoder, VND, is designed to match the compound inner decoder.

- *MIMO Communication Link*

  The system structure of a RA code with MIMO link is shown in Figure 2.21. Assume there are M transmit and N receive antennas. The transmit symbol is thus a vector $s = [s_1, \ldots, s_M]^T$ of size $M \times 1$ and all entries are taken from a complex value constellation set of size 2^{M_c}. Hence there are $M \cdot M_c$ bits in each transmitted

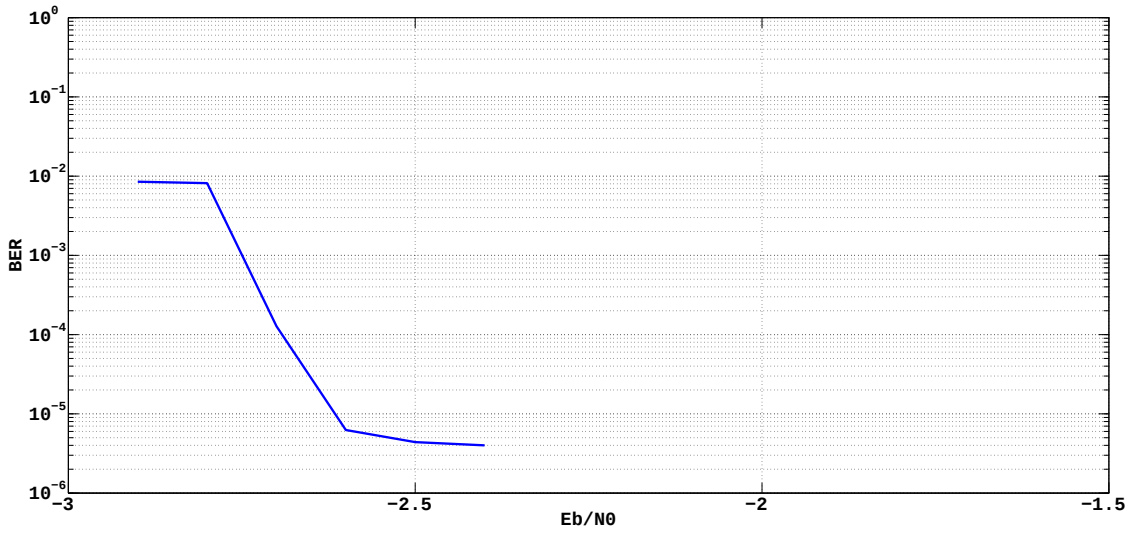

Figure 2.20: BER performance for a systematic half rate RA code with curve fitting

symbol. The average energy per transmit symbol is E_s and $E[||s_m||^2] = E_s/M$. The received vector $y = Hs + n$ has size $N \times 1$, where H is the $N \times M$ channel matrix, and n is the noise vector of size $N \times 1$. Each entry of N is an independent, complex value, zeros mean, Gaussian random variable with variance $\sigma^2 = N_0/2$ on both real and imaginary parts. The normalized SNR E_b/N_0 is defined as

$$\left.\frac{E_b}{N_0}\right|_{dB} = \left.\frac{E_s}{N_0}\right|_{dB} + 10\log_{10}\frac{N}{RMM_c} \tag{2.62}$$

Each entry of H is an independent, complex value, zero mean, Gaussian random variable with variance $1/2$ on both real and imaginary parts. And H is assumed to be known at the destination only. H is also assumed to be ergodic which means it changes for every symbol s instead of being quasistatic which means it remains the same over long time intervals. The ergodic model capacity is provided in [50, 51, 52]

$$C = E\left[\log_2 \det\left(I + \frac{E_s}{N_0}\frac{1}{M}HH^{\dagger}\right)\right] \tag{2.63}$$

where I is an identity matrix, and $H^{\dagger}$ represents the complex-conjugate transpose of H. This capacity can only be achieved by using Gaussian constellation, however, in this section Gray-mapped QPSK is considered.

- *EXIT Curve of the MIMO Detector*

  The MIMO detector calculates LLR for each bit by considering all 2^{MM_c} possible hypotheses [53, 54, 55]. There is no closed form expression for the MIMO detector, hence it is measured by Monte Carlo simulation and denoted as

$$I_{E,DET}\left(I_{A,DET}, \frac{E_b}{N_0}\right) \tag{2.64}$$

  The output of the detector is then fed into the decoder as a priori information.

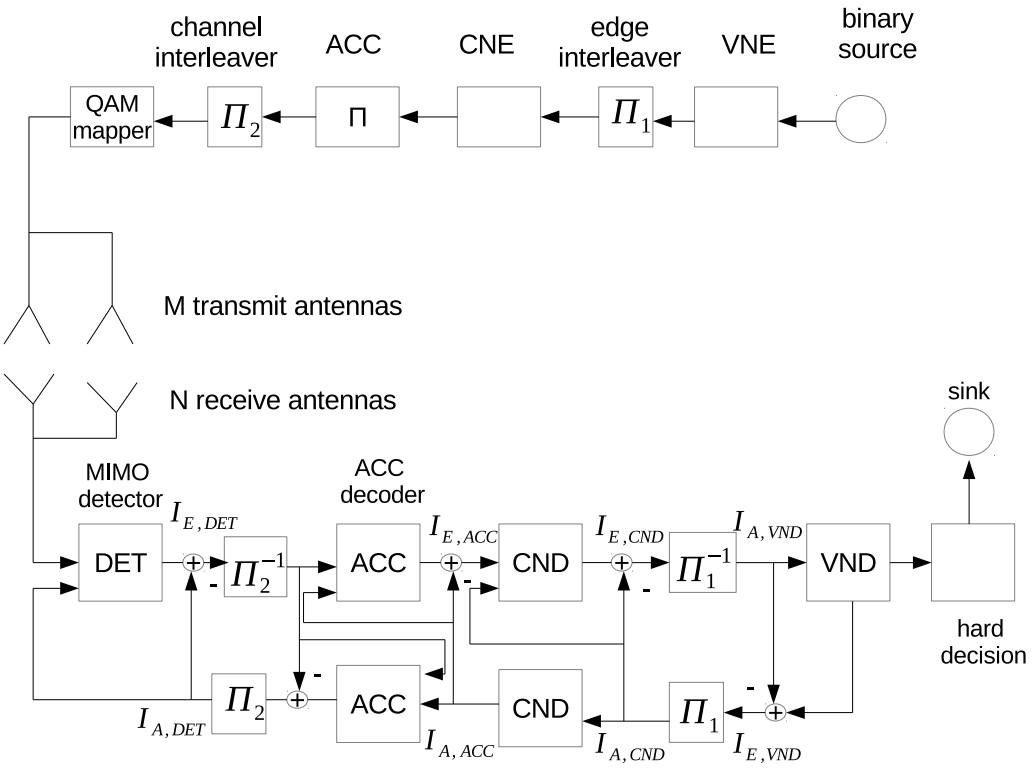

Figure 2.21: MIMO model with RA code

- *EXIT Curve of the Combined Detector, ACC and CND*

  The output of the MIMO detector is deinterleaved before being sent into the inner decoder comprising the ACC decoder and CND. Both the channel interleaver and the edge interleaver are designed to break the dependencies within the LLRs.

  The receiver structure shown in Figure 2.21 is a double serial concatenation [56], which has a special feature that it requires an inner detection loop between the MIMO detector and the ACC decoder. This loop enhances the extrinsic information $I_{E,ACC}$ for a given a priori information $I_{A,ACC}$. Examples in [37] shows that for a 4×1 MIMO detector, saturation is achieved after three iterations as shown in Figure 2.22 and for other scenario like 2×1 or 4×2 MIMO detectors, only two iterations are required.

  An alternative to the inner loop is to perform the detection and accumulator decoding together in a single trellis which would have two states with 2^{MM_c} edges for each state. Each edge corresponds to one symbol. The next state is determined by inputting MM_c bits into the ACC, with 2^{MM_c-1} transitions arriving in state 0 and state 1 respectively. For detecting/decoding, BJCR [57] algorithm is applied to the trellis.

  The code design for this system goes like this: run simulations to obtain the DET& ACC curve first, then fit the VND curve to the DET& ACC curve. There are two advantages for the trellis detection scheme over the inner loop scheme. First without inner loop, decoding complexing is reduced. Second, simulations shows that trellis

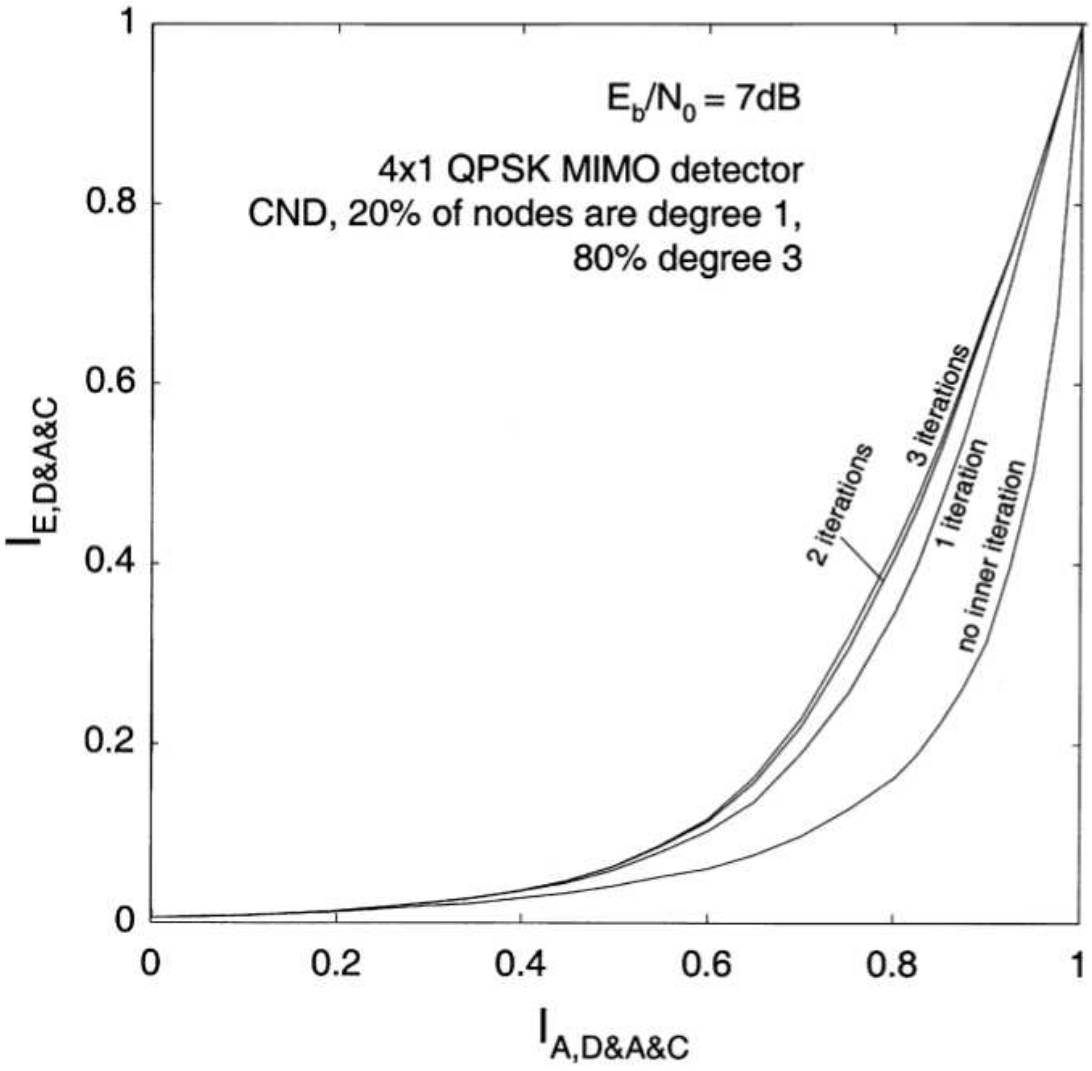

Figure 2.22: EXIT curves for combined 4×1 MIMO detector and the inner decoder, copied from [37]. Parameters: $R = 1/2$, $n = 10^5$, $a_{c,1} = 0.2$, $\tilde{d}_{c,1} = 1$, $\tilde{d}_{c,2} = 3$

detection is more robust to a code parameters mismatches [58].

- *Design Example*

  [37] provides a design example of curve fitting for an RA code with 4×1 MIMO link. For check node, it sets $a_{c,1} = 0.2$, $\tilde{d}_{c,1} = 1$ and $\tilde{d}_{c,2} = 3$. Hence the combined detector and inner decoder curve only depends on the channel. The remaining parameters are set as follows: QPSK modulation, $R = 1/2$, $D_v = 3$, $n = 10^5$ and the parameters for the variable node is in Table 2.3. And the resulting EXIT chart is shown in Figure 2.23

$$\begin{array}{lll} \tilde{d}_{v,1} = 2 & \tilde{a}_{v,1} = 69.05 & \tilde{b}_{v,1} = 26.56 \\ \tilde{d}_{v,2} = 5 & \tilde{a}_{v,2} = 25.13 & \tilde{b}_{v,2} = 24.16 \\ \tilde{d}_{v,13} = 44 & \tilde{a}_{v,3} = 5.82 & \tilde{d}_{v,3} = 49.28 \end{array}$$

Table 2.3: Curve fitting VNE parameters for RA code with 4×1 MIMO link

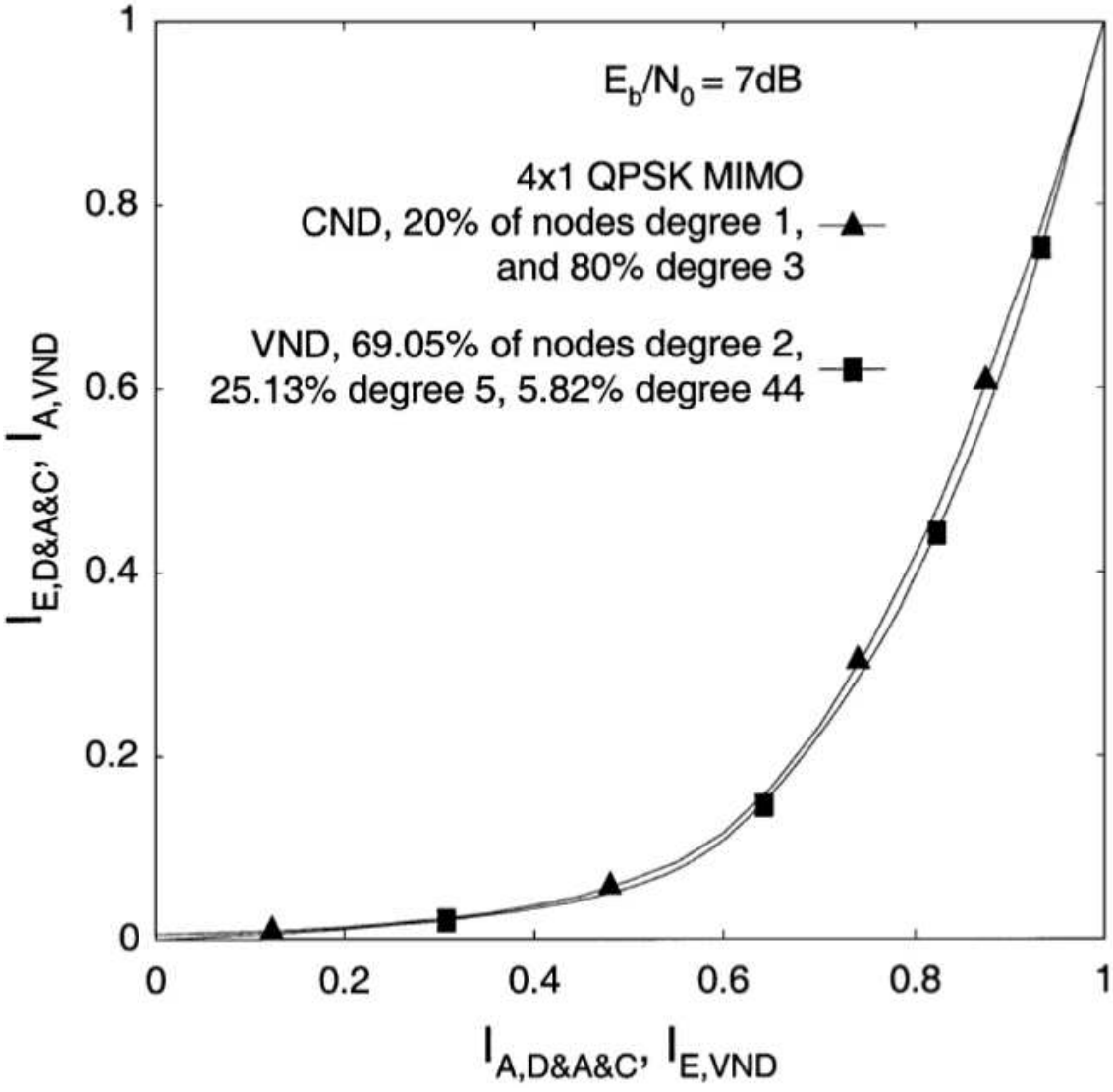

Figure 2.23: Curve fitting for non-systematic RA code with 4×1 MIMO link and QPSK modulation, copied from [37].

Chapter 3

Cooperative Spatial Multiplexing

3.1 Introduction

In this chapter we study a basic C-SM system model, which will be the foundation for our future design.

We propose an amplify-and-forward cooperative spatial multiplexing (C-SM) scheme in which the transmitter (source), equipped with a single antenna, forms a virtual antenna array from a collection of nearby distributed antennas, each belonging to different single-antenna wireless terminals and utilizes them as analogue non-regenerative relays. The source then transmits different signals to each nearby relay by using a different time slot. Each relay amplifies and forwards the received signal with a gain factor to the relays near the receiver, then those relays amplify and forward the received signal to the receiver. The receiver, also equipped with a single antenna, receives different versions of the signals from different relays via separate time slots and then uses a Maximum likelihood detector to decode the information. This combination of transmitter, relays, and receiver forges a virtual MIMO system and realizes its performance in single-antenna wireless terminals environment.

To simplify the problem and provide a clear presentation of the proposed system, we assume that all channels experience flat fading in one symbol period and the relay nodes are perfectly synchronized.

In this chapter, we focus our attention on two main objectives:

- We try to derive a mathematical expression for the overall systematic BER, from which we expect to understand the overall system BER performance quantitatively.

  With respect to the BER derivation, the typical approach is to use the Moment Generating Function (MGF) method, derive BER from the probability density function (PDF) of the received SNR with the help of MGF, as in [59] and [60]. However, because our system involves multiple continuous Rayleigh fading, it may not be amenable to have a closed-form PDF or MGF and it may also be difficult to estimate asymptotic behavior using Monte Carlo simulation, since long simulations are often required to estimate BER at high SNR. For this reason we use a semi-analytical approach based on bounding techniques which can greatly reduce the simulation length at high SNR.

- We try to understand the impact of different modulations on the overall system performances.

We assume relays are either near the transmitter or near the receiver and different relay groups are far away from each other so the channels between relay groups are weaker than the channels between the transmitter and its nearby relays (also the channels between the receiver and its nearby relays). We deploy different modulation schemes across the system to exploit the strong channels but also try to compensate for the poor channels.

In this chapter our focus is on true distributed spatial multiplexing techniques which can be exploited by terminals and relays with one antenna only and for simplicity reason all nodes in this system (includes transmitter, receiver and relays) only have single antenna. To implement spatial multiplexing in such a scheme requires that the source first distributes its information to a group of nearby relays, which then each transmit independent data to another group of relays close to the destination. These relays then forward their received signals to the destination, where the original source information is detected. This necessarily involves at least two stages of relaying; here we also generalize to the case in which further intermediate clusters of relays are employed, resulting in more than two stages of relaying. In this chapter, again for simplicity, we wish to evaluate BER performance of uncoded transmission. We assume maximum likelihood (ML) detection at the destination, which optimizes performance. Note however that we restrict our attention to spatial multiplexing using no more than two independent streams and QPSK modulation, so that detection complexity is low.

3.2 Literature Review

There is by now a substantial literature on cooperative spatial multiplexing, but very little of it addresses the system architecture we wish to focus on. [16] discuss a system named as MIMO tunnel in the paper. It consists of single antenna transmitter, receiver and relays. There are several relay clusters between transmitter and receiver, each one containing an equal number of relay stations. Transmission starts with K independent signals being transmitted from the transmitter to the first relay cluster in a interference free channel, every relay station in the first relay cluster receives one of the K signals (there are K relay stations in each relay cluster). Then signals are jointly transmitted over the same physical channel from the first relay cluster to the last relay cluster, while intermediate nodes store, amplify and forward the received signals simultaneously such that signals received by the current relay station are linear superpositions of the signals forwarded by relay stations in the previous relay cluster. Note that [16] gives only the capacity analysis of the system.

[61] also discusses amplify-and-forward multiple hop cooperative spatial multiplexing, however the major difference is that its system has multiple antennas at the receiver and applies a minimum mean square error (MMSE) detector.

[62] discusses amplify-and-forward cooperative spatial multiplexing system for multiple single-antenna user pairs and for multiple-antenna transmitter, relay and receiver system.

In [63], the system has a multiple-antenna transmitter and receiver with single-antenna relays. The number of antennas at the transmitter and relay is greater than the number of relays in order to enable V-BLAST detection algorithms. Furthermore, the transmitter can know the channel state information of the first and the second hops via feedback channels.

3.3 Chapter Outline

The remainder of this chapter is organized as follows.

In Section 3.4 the system structure is shown in general form, along with transmission pattern and corresponding mathematical model.

A mathematical derivation of the system BER performance in the form of a union bound using our semi-analytical approach is offered in section 3.5. Details of the derivation are given in section A.

Section 3.6 shows the numerical results and the analysis of the system BER performance in Monte Carlo simulation and in our semi-analytical method.

In Section 3.7, two alternative system structures are proposed, along with their corresponding transmission pattern and graphic results of their system BER performance in Monte Carlo simulation. Furthermore, performance comparison and analysis are also shown.

Conclusions for our work so far is given in section 3.8.

3.4 System Model

In this chapter we consider a multi-hop spatial multiplexing communication system with a single antenna at transmitter and receiver and groups of AF relays between them. Each relay group contains two distributed relays with a single antenna each. Figure 3.1 shows the system structure. There are two independent data streams transmitted by each group of relays. A maximum likelihood detector is implemented at the receiver to detect the receive symbols and perfect channel knowledge is assumed at the receiver.

The transmission of the signal from transmitter to receiver consists of at least three phases depends on the number of relay groups. In the first half of phase one, the transmitter transmits the signal corresponding to one stream to one of the relays in group 1, and in the second half of the phase, the other stream is similarly transmitted to the other relay of group 1. In phase two, the two relays in group 1 each simultaneously transmit the amplified signal of the corresponding independent data stream to the subsequent group, and then in phase i, $i = 3 \cdots n-1$ the i_{th} group transmits to the $(i+1)_{th}$. Finally in the first half of phase n, the first relay of group n transmits its amplified signal to the receiver, and in the second half the second relay does so. Figure 3.2 shows the transmission time line.

$\mathbf{H}_1$ and $\mathbf{H}_{n+1}$ are 2×2 diagonal matrices (all the non-diagonal entries are zero) whose diagonal entries are independent complex Gaussian channel fading coefficients with unity variance corresponding to the links between the transmitter and its neighbouring relay group in phase one and the links between the receiver and its neighbouring relay group in phase n respectively. $\mathbf{H}_2$ to $\mathbf{H}_n$ are 2×2 channel matrices in phases two to $n-1$, which again have independent complex Gaussian entries with unity variance. $\mathbf{D}$ represents the transmitted symbol vector whose entries are drawn from the QPSK signal constellation with average transmit power unity. $\mathbf{N}_1$ to $\mathbf{N}_{n+1}$ are 2×1 matrices containing complex noise with variance σ^2 at the relay groups and the

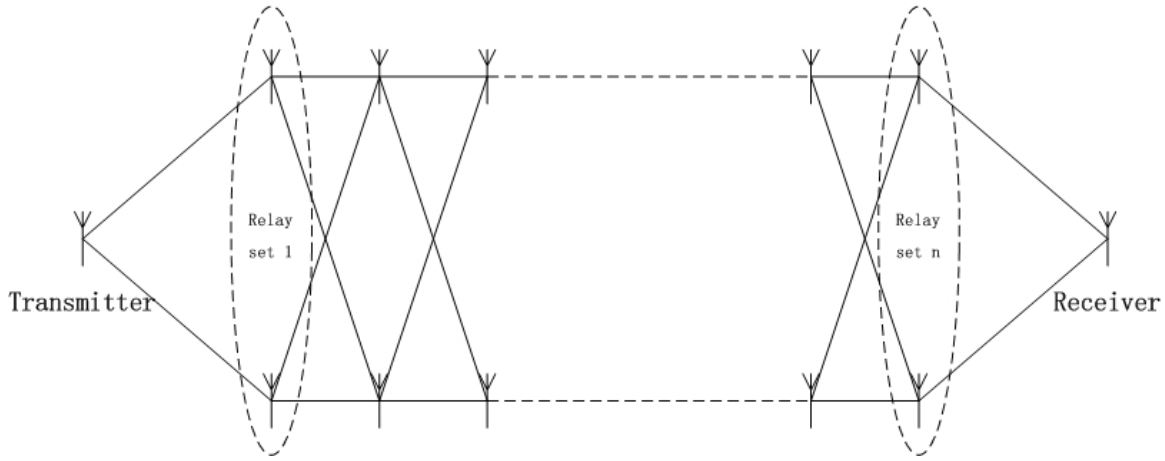

Figure 3.1: Distributed spatial multiplexing system with n relay sets

receiver respectively. Therefore the output $\mathbf{R}$ from the receive antenna is,

$$\begin{aligned}\mathbf{R} =&\mathbf{H}_{n+1}(g_n(\mathbf{H}_n(\cdots(g_2(\mathbf{H}_2(g1(\mathbf{H}_1\mathbf{D}+\mathbf{N}_1))+\mathbf{N}_2))\cdots)\\ &+\mathbf{N}_n))+\mathbf{N}_{n+1}\\ =&\mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots g_2\mathbf{H}_2g_1\mathbf{H}_1\mathbf{D}+\mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots\\ &g_2\mathbf{H}_2g_1\mathbf{N}_1+\mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots g_3\mathbf{H}_3g_2\mathbf{N}_2+\cdots\\ &+\mathbf{H}_{n+1}g_n\mathbf{H}_ng_{n-1}\mathbf{N}_{n-1}+\mathbf{H}_{n+1}g_n\mathbf{N}_n+\mathbf{N}_{n+1}\end{aligned} \tag{3.1}$$

Here g_k $(1 \leq k \leq n)$ is the constant amplification factors at relay group k $(1 \leq k \leq n)$. The value of g_k $(1 \leq k \leq n)$ is set to be

$$g_k = \begin{cases} \sqrt{(1+2\sigma^2)^{-1}} & \text{if } k = 1; \\ \sqrt{(2+2\sigma^2)^{-1}} & \text{if } 2 \leq k \leq n. \end{cases} \tag{3.2}$$

which will ensure that the transmit power at each relay is the same as it is at the transmitter. The overall channel is

$$\mathbf{H}_o = \mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots g_2\mathbf{H}_2g_1\mathbf{H}_1 \tag{3.3}$$

and the overall noise is

$$\begin{aligned}\mathbf{N}_o =& \mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots g_2\mathbf{H}_2g_1\mathbf{N}_1+\mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots\\ &g_3\mathbf{H}_3g_2\mathbf{N}_2+\cdots+\mathbf{H}_{n+1}g_n\mathbf{H}_ng_{n-1}\mathbf{N}_{n-1}\\ &+\mathbf{H}_{n+1}g_n\mathbf{N}_n+\mathbf{N}_{n+1}\end{aligned} \tag{3.4}$$

$\mathbf{R}$ can be rewritten as

$$\mathbf{R} = \mathbf{H}_o\mathbf{D}+\mathbf{N}_o \tag{3.5}$$

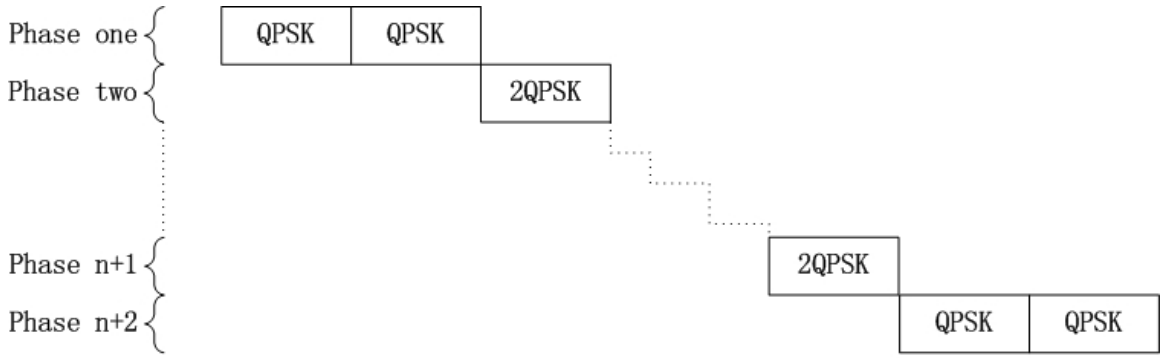

Figure 3.2: Transmission time line

3.5 Bit Error Rate Analysis

We use the union bound to analyze the BER performance of the system. According to 5.1.5 in [5], the symbol error rate between message m_i and m_k can be represented as

$$\begin{aligned} &P_e(m_i \text{ is transmitted,received as } m_k) \\ &= \mathrm{P}(\|\mathbf{R}-\mathbf{H}_o\mathbf{M}_k\| < \|\mathbf{R}-\mathbf{H}_o\mathbf{M}_i\|) \end{aligned} \tag{3.6}$$

where $P()$ is probability of the event, $||\ \ ||$ is the Euclidean distance, $\mathbf{M}_i$ is the transmitted symbol vector and $\mathbf{M}_k$ is the symbol vector detected in error. Equation (3.6) can be further modified to

$$\begin{aligned} &P_e(\mathbf{M}_i \text{ is transmitted,received as } \mathbf{M}_k) \\ &= \mathrm{P}(\|\mathbf{R}-\mathbf{H}_o\mathbf{M}_k\|^2 < \|\mathbf{R}-\mathbf{H}_o\mathbf{M}_i\|^2) \\ &= \mathrm{P}(\|\mathbf{H}_o\mathbf{M}_i+\mathbf{N}_o-\mathbf{H}_o\mathbf{M}_k\|^2 \\ &< \|\mathbf{H}_o\mathbf{M}_i+\mathbf{N}_o-\mathbf{H}_o\mathbf{M}_i\|^2) \\ &= \mathrm{P}(\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)+\mathbf{N}_o\|^2 < \|\mathbf{N}_o\|^2) \\ &= \mathrm{P}(\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|^2+\|\mathbf{N}_o\|^2 \\ &+ 2\mathrm{Re}(\mathbf{N}_o^H\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)) < \|\mathbf{N}_o\|) \\ &= \mathrm{P}(\mathrm{Re}(\mathbf{N}_o^H\frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}) \\ &> \frac{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}{2}) \end{aligned} \tag{3.7}$$

where superscripts H denote conjugate transpose and Re() means real part. For a particular set of channels and fixed i, k, the variable $\frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}$ is a 2×1 matrix with constant entries. $\mathbf{N}_o^H$ is a 1×2 matrix with independent complex Gaussian variables. So the variable $\mathrm{Re}(\mathbf{N}_o^H\frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|})$ is two independent Gaussian variables added together which is also a Gaussian variable [64], therefore

$$\begin{aligned} &P_e(\mathbf{M}_i \text{ is transmitted,received as } \mathbf{M}_k) \\ &= \mathrm{Q}(\frac{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}{2\tilde{\sigma}}) \end{aligned} \tag{3.8}$$

where $Q()$ is the Q function [64], and $\tilde{\sigma}$ is the standard deviation of $\mathrm{Re}(\mathbf{N}_o^H \frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|})$.

Because a closed form for the probability density function of $\frac{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}{2\tilde{\sigma}}$ is hard to obtain, methods like the MGF approach[65] were implausible, we therefore use a semi-analytical approach to calculate the BER, using the formula[66]

$$\begin{aligned} &P_e(\mathbf{M}_i \text{ is transmitted,received as } \mathbf{M}_k) \\ &= \mathop{\mathrm{E}}_{\mathbf{H}_o}\{\mathrm{Q}(\frac{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|}{2\tilde{\sigma}})\} \end{aligned} \tag{3.9}$$

$\mathop{\mathrm{E}}_{\mathbf{H}_o}\{\}$ is expectation over variable $\mathbf{H}_o$ and $P_e(\mathbf{M}_i$ is transmitted,received as $\mathbf{M}_k)$ is obtained by averaging the results of the Q function over a number of randomly generated $\mathbf{H}_o$. Note that the effective standard deviation $\tilde{\sigma}$ of the noise can also be obtained analytically in terms of the channel matrices $\mathbf{H}_1$ to $\mathbf{H}_n$ and $\mathbf{M}_i$ and $\mathbf{M}_k$. For the derivation see Appendix A.

Finally the union bound can be expressed as

$$P_u = \sum_{i=0}^{M} P(\mathbf{M}_i) \sum_{\substack{k=0 \\ k\neq i}}^{M} P_e(\mathbf{M}_i \text{ is transmitted,received as } \mathbf{M}_k) \tag{3.10}$$

where $P(\mathbf{M}_i)$ is the probability that signal $\mathbf{M}_i$ is sent, and M is the number of combinations of constellation points possible in the transmitted signal vector. In this chapter we assume all combinations are equiprobable and hence $P(\mathbf{M}_i) = 1/M$.

In addition to the union bound, we also develop and implement a lower bound using the formula

$$P_m = \sum_{i=0}^{M} P(\mathbf{M}_i) P_m(\mathbf{M}_i \text{ is transmitted}) \tag{3.11}$$

which is called minimum distance [1] bound in this thesis.

In this formula $P_e(\mathbf{M}_i$ is transmitted,received as $\mathbf{M}_k)$ in equation (3.10) is replaced with $P_m(\mathbf{M}_i$ is transmitted) which is the error rate for the minimum value of $\|\mathbf{M}_i - \mathbf{M}_k\|$ over all the combinations of (i,k) where $0 \leq k \leq M$ and $k \neq i$. In other word, $P_m(\mathbf{M}_i$ is transmitted) $=$ $\min\{P_e(\mathbf{M}_i$ is transmitted,received as $\mathbf{M}_k), k = 0, ...i-1, i, ...M\}$

3.6 Numerical Results

Figure 3.3 shows the result of system simulation (blue broken line) with two relay sets, corresponding union bound (red continuous line) and minimum distance bound (green continuous line). From Figure 3.3, it is clear, as expected, that the union bound is an upper bound and the minimum distance bound is a lower bound. The looseness of the union bound is due to the fact that the second minimum distance (sometimes even the third and the fourth minimum distances) is similar to the minimum distance.

The overall system BER performance has first order diversity due to the errors on the links $\mathbf{H}_1$ between the transmitter and relay group 1. Since on $\mathbf{H}_1$ independent streams are transmitted

[1] distance between the received symbols at the receiver

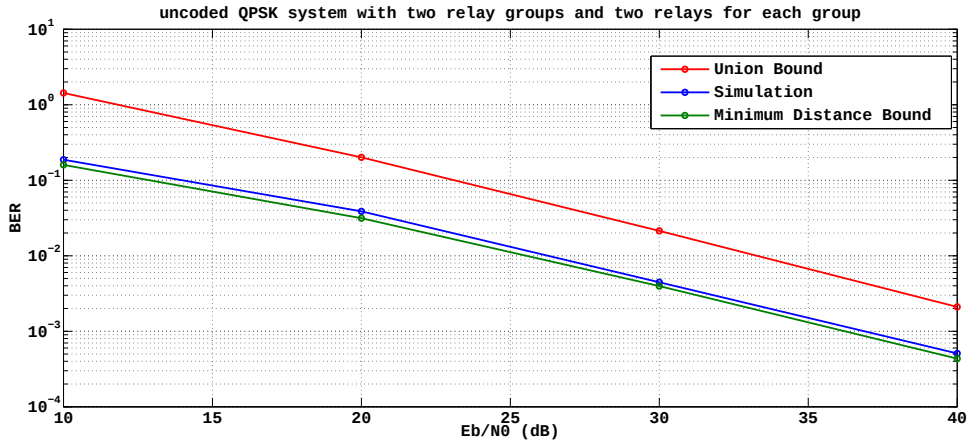

Figure 3.3: Exact BER, union bound and minimum distance bound for system with two relay groups

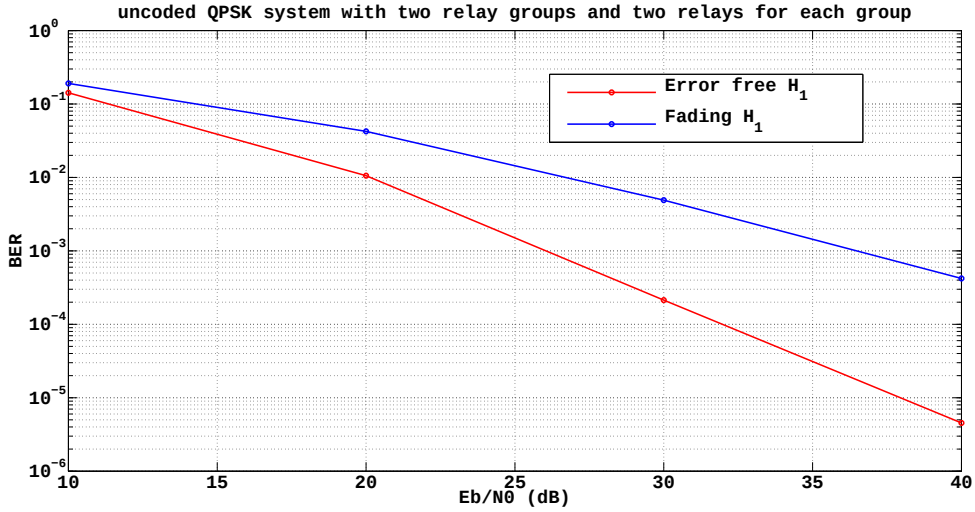

Figure 3.4: System performance with a fading $\mathbf{H}_1$ (first order diversity) and system performance with an error free $\mathbf{H}_1$ (second order diversity)

over different links, there is no redundant route here, whereas in other parts of the system, redundant routes are available. If we replace $\mathbf{H}_1$ by an error free channel, the overall system BER performance will become second order diversity as shown in Figure 3.4. The blue line is the overall system BER performance when $\mathbf{H}_1$ is fading, red line is the BER when $\mathbf{H}_1$ is replaced by an error-free channel.

3.7 Alternative System

For simplicity reason, systems in this chapter have only two relay groups and two relay nodes in every relay group.

3.7.1 Soft-Decoding-And-Forward System

We further introduce free-space path loss (FSPL) into our system by assuming that the distance between transmitter and its nearest relay group is much shorter than the distance between relay groups. The same also goes to the distance between receiver and its nearest relay group. Therefore signal received by relay group close to receiver is weaker than the signal received by relay group close to transmitter and signal received by receiver. In order to cope with the weak signal but also take advantage of the strong signal, we decide to use different modulations across the system. The whole process goes like: the transmitter first sends out two different QPSK signals to the two relay nodes in relay group one, then each relay node will transform the received QPSK signal to two BPSK signals, after that two relay nodes in relay group one transmit one of their two BPSK signals simultaneously to the next relay group. After the first two BPSK signals reach the receiver, the second two BPSK signals will be transmitted by relay group one. Finally, two relays in relay group two transmit their received signals to the receiver via different time slots.

Transformation Mechanism

This mechanism takes place at the relays in relay groups one to separate the received QPSK signal to two BPSK signals.

The relay node divides all signals on the constellation into two groups according to the information they carry (signal carries bit 1 as its first bit and signal carries bit 0 as its first bit will be separate into different group). After a QPSK signal is received, the relay node will find one signal from each group who is closest (has the minimum distance) to the received signal. Assuming the distances between the received signal and its closest two points from each group whose first bit is 1 and 0 are d_{11} and d_{10}, the log likelihood-ratio of the first bit is

$$LLR_1 = \frac{-d_{11}^2 + d_{10}^2}{2\sigma^2} \tag{3.12}$$

Applying the same procedure to the second bit of the received signal, the log likelihood-ratio of the second bit is

$$LLR_2 = \frac{-d_{21}^2 + d_{20}^2}{2\sigma^2} \tag{3.13}$$

In order to constrain the transmit power, hyperbolic function tanh() is used. Finally, the two transformed BPSK signals are

$$\tanh(LLR_1) \tag{3.14}$$

and

$$\tanh(LLR_2) \tag{3.15}$$

At high SNR, because of the small noise variance σ^2 and large distance between different points, *LLR* becomes large. Consequently, $\tanh(LLR)$ is close to or almost equals to 1, so this mechanism produces a signal very similar to standard BPSK with average transmit power to 1.

The advantage of this transformation mechanism is that it applies soft decision which will give better BER performance than the hard decision system used in the decode-and-forward system.

3.7.2 Simple Amplify-and-Forward System

The original system and soft-decoding-and-forward system provide only first order diversity; in order to achieve higher order diversity we develop a new system in which the transmitter first broadcasts one QPSK signal to relays in relay group one. The relays amplify and forward the received signals to relays in relay group two where the received signals are stored. Next, the another QPSK is sent by the transmitter and received by relay group two via relay group one using the same procedure as the first QPSK signal does. In relay group two, the received signals are combined with the previously stored signals and transmitted to the receiver.

Combining Mechanism

The signals received by every relay node in the last relay group consist of two QPSK signals superposed together which contain the same information but via different rout.

The QPSK constellation the transmitter use has unity transmit power, so if we scale the power of one received signal by half and add it to the other one, a 16 QAM signal with average transmit power 2.5 will be formed.

Note that the combining mechanism here also affects the diversity of the system. If we use the same combining style on every relay node in the last relay group, the system will not have full receive diversity. For example, if on one relay node the first signal is scaled down by half and added to the second signal, then on the other relay node the procedure must be performed the other way around. This prevent the situation in which the weaker signal is scaled down on both relay nodes.

Figure 3.5 is the comparison between different combining styles. The system has two relay groups with three relay nodes each. The blue line which approaches third order diversity corresponds to the combining style that weights each received signal differently in every relay node, the red line which is second order diversity corresponds to the combining style that weights each received signal in the same way in every relay nodes.

3.7.3 Performance Comparison

In order to make these three systems comparable, we set the average transmit energy per bit in every phase across these three systems equal to 0.5. We assume every system uses the same symbol period T to transmit one signal, so the through-put of these three systems are: the original system uses 5 symbol periods to transmit 4 bits (two QPSK signals), the soft-decode-and-forward system uses 6 symbol periods to transmit 4 bits (two QPSK signals), the simple amplify-and-forward system uses 6 symbol periods to transmit 4 bits (two QPSK signals). Note that the number of symbol periods the original system uses does not match those used by the other two systems, but the difference is not very significant.

Figure 3.6 shows the BER performance of the original system (blue), the soft-decode-and-forward system (red) and the simple amplify-and-forward system (green). It is obvious that the simple amplify-and-forward system outperforms other two significantly due to its higher diversity order.

Figure 3.7 shows the throughput of these three systems, every block is one symbol period, phase one is the transmission from transmitter to relay group one, phase two is from relay group one to relay group two, phase three is from relay group two to the receiver.

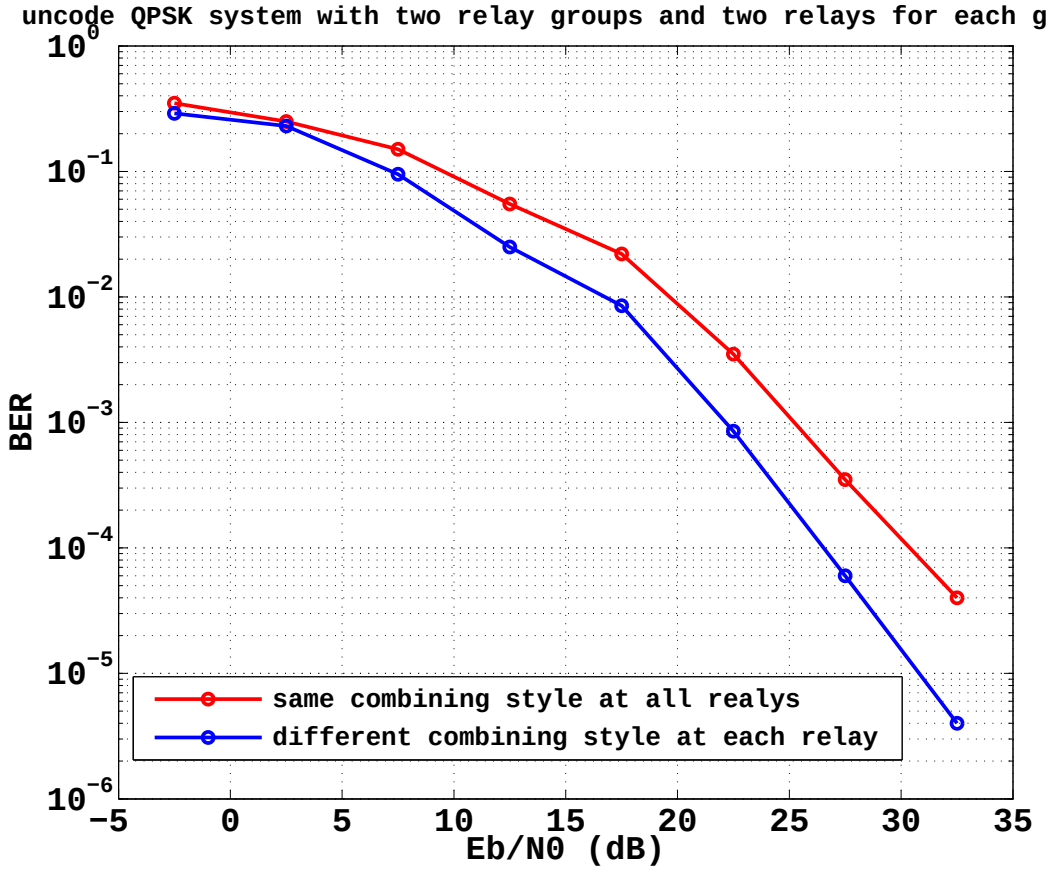

Figure 3.5: performances of different combining mechanisms in section 3.7.2

3.8 Conclusion

We introduce a distributed spatial multiplexing system which utilizes virtual antenna arrays forged by distributed single antenna relays. We have also described a semi-analytical method to determine upper and lower bounds on BER. Simulation results have confirmed this method, it has the advantage over full Monte Carlo simulation that since the BER bounds for each randomly-selected channel are obtained analytically rather than by counting errors, much shorter simulations are required, particularly at high bit energy to noise density ratio.

The main contributions of this chapter are:

- Present the BER performance of a multi-hop distributed spatial multiplexing system with single-antenna terminals and relays as shown in Figure 3.3.
- Develop a semi-analytical method to obtain a upper and a lower bounds for the system's BER as shown in Figure 3.3.
- Investigate the impact of different modulation schemes on the overall system BER performance and find out that the simple amplify-and-forward system outperforms others significantly due to its higher diversity order according to Figure 3.6.

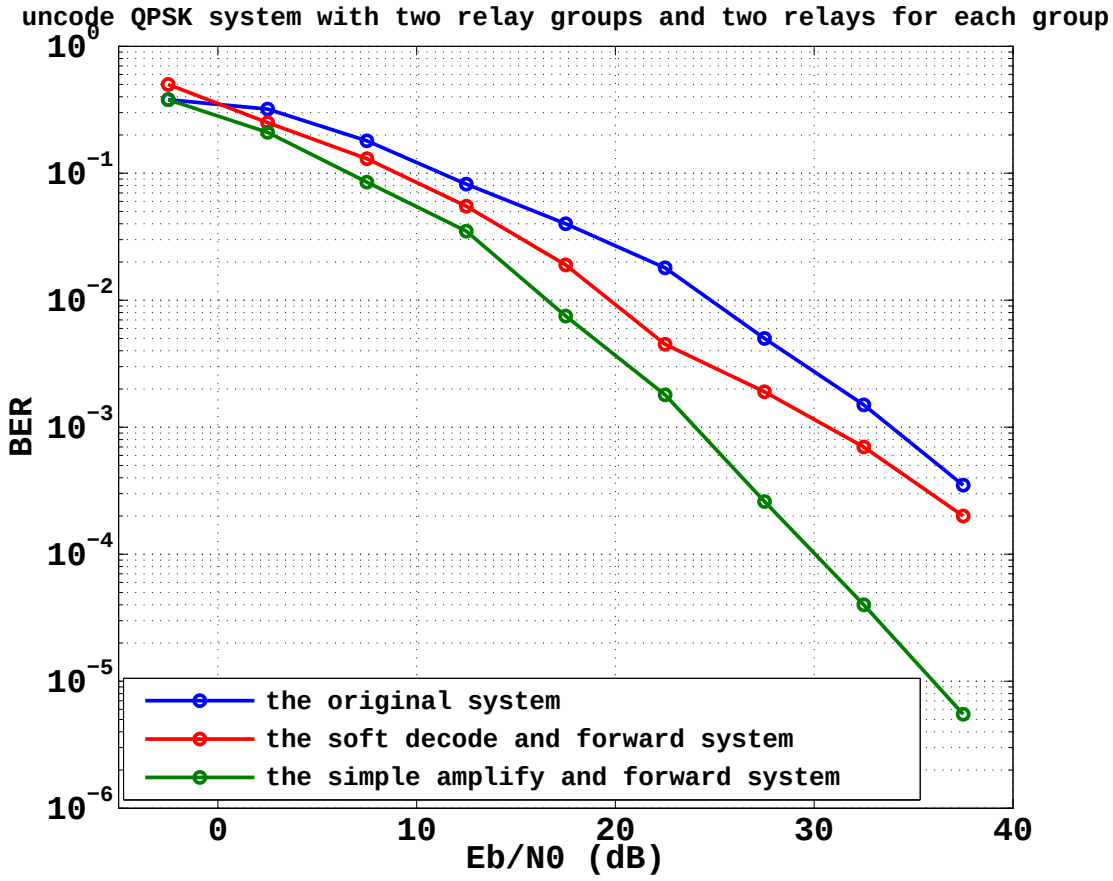

Figure 3.6: Comparison between original system, soft-decoding-and-forwarding system and simply amplify-and-fording system

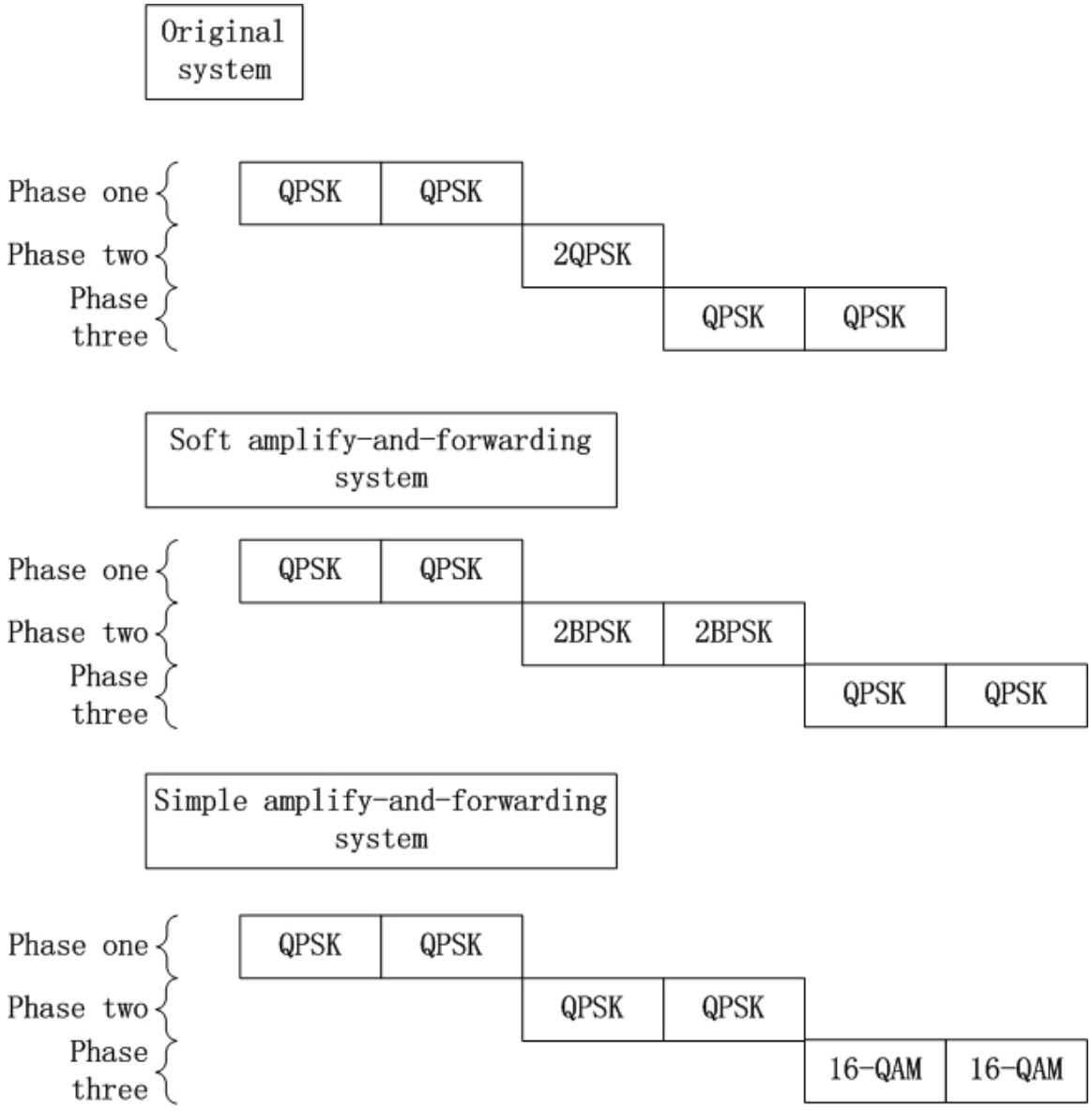

Figure 3.7: System throughput time line

Chapter 4

Cooperative Spatial Multiplexing with Slepian and Wolf Code - Blind Compression

4.1 Introduction

When spatially adjacent mobile stations communicate with each other, they form a mutually communicating entity, termed a Virtual Antenna Array (VAA)[67, 68, 69], in which the wired links of traditional MIMO systems are replaced by wireless links. The VAA concept is illustrated in Figure 4.1. The transmitter sends information to the receive VAAs via the MIMO channel (green, purple, orange, black lines). The receiver then gathers information via intra-links (red lines) from all the relays in the receive VAAs.

Based on our study of the VAA in our basic C-SM model in chapter 3, we discovered that correlation exists between information on different relays, and therefore we try to implement a distributed source code on the relays to increase the spectral efficiency on the relay-destination link by compressing the correlated information on relays.

The correlation between the information at each relay allows the Slepian-Wolf theorem to be applied in the system shown in Figure 4.2. Signals sent out by the transmitter are superposed at the relays. Since the information originates from the same source, the signals at the relays are correlated.

In this chapter we try to tackle the following problems:

- Implement a Slepian-Wolf code on the relay-destination link, which increases the spectral efficiency by compressing the correlated data on relays. In order to support the Slepian-Wolf coding at the relay, several quantization methods are developed. At the destination, corresponding de-quantizers are also applied to reconstruct the signal.

- Test two Slepian-Wolf code implementations, namely the Super-Turbo code[70, 71, 72] and the Accumulate-Repeat-Accumulate code (ARA) [37, 40, 73], and determine which one is more suitable for the system.

In this chapter, because of the complexity and sensitivity to CSI quality of such systems, we consider only fixed gain relaying and perfect CSI at the destination.

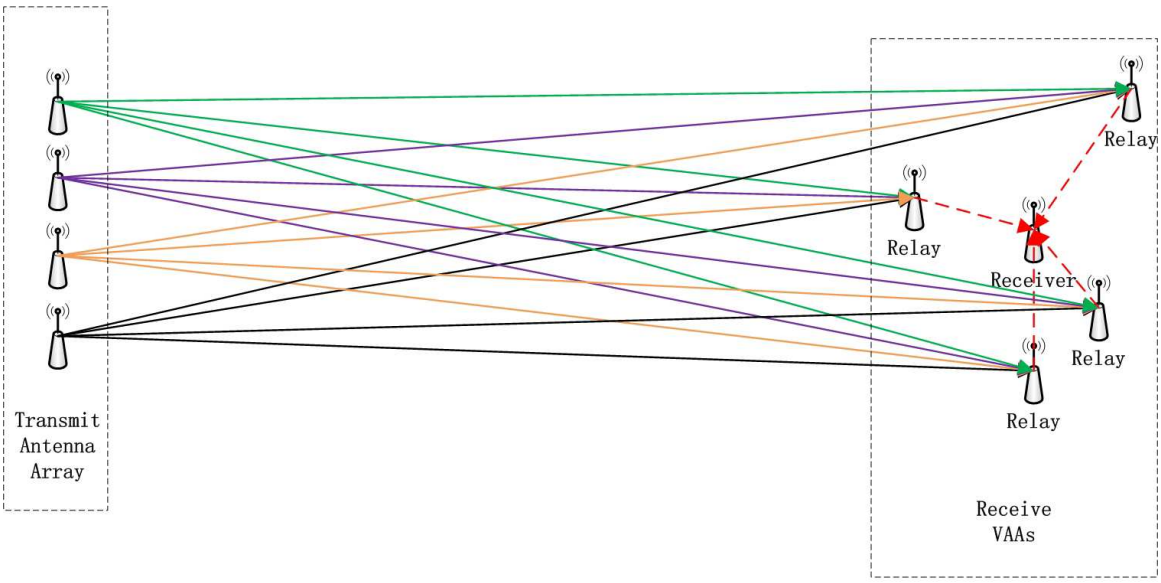

Figure 4.1: Virtual antenna array concept: broken arrow lines denote links within virtual antenna arrays, solid arrow lines represent links between different VAAs.

4.1.1 Chapter Outline

The chapter is set out as follows. In the next section we state the full system model, then in section 4.3 we describe two Slepian-Wolf implementation methods. Three Vector-Quantization methods and two De-Quantization methods are introduced in section 4.4. Simulation result and analysis of system properties are given in section 4.5 and we conclude the chapter in section 4.6.

4.2 System Scenario and Structure

The system structure is shown in Figure 4.2. Data is transmitted from Source to Destination via relays (R21, R22). Every relay node and the receiver is fitted with only one antenna, the transmitter has an antenna array with two elements (A1,A2) and no direct link between Source and Destination is available. Additive White Gaussian Noise (AWGN) with variance σ^2 is added to every node in this system, however, because We focus our attention on the relay-destination link and the correlation between relays depends on the source-relay link, only the 2×2 channel between the transmitter and the relay is assumed to be independent Rayleigh fading with unity variance, while the remaining channels are non-fading AWGN channel. Data from the source is modulated to a 16 QAM signal and transmitted to the relays via the 2×2 MIMO fading channels. The signals received at the relays are

$$\begin{aligned} \mathbf{S_{rec}} &= \mathbf{SH} + \mathbf{N_2} \\ &= \begin{bmatrix} s & s \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} + \begin{bmatrix} n_{21} & n_{22} \end{bmatrix} \\ &= \begin{bmatrix} (h_1+h_3)s+n_{21} \\ (h_2+h_4)s+n_{22} \end{bmatrix}^T \end{aligned} \tag{4.1}$$

where h_1 to h_4 are independent complex Gaussian channel fading coefficients with unit variance for the 2×2 channel H between the transmitter and the relays, n_{21}, n_{22} are AWGN at relay, superscripts T denotes transpose, s represents the transmitted message (or symbol) which is drawn from the 16 QAM signal constellation

An alternative way is to send independent data from each antenna of the source using QPSK signals to the relays via the MIMO channel. The signals received at the relays are

$$\begin{aligned} \mathbf{S_{rec}} &= \mathbf{SH} + \mathbf{N_2} \\ &= \begin{bmatrix} s_1 & s_2 \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} + \begin{bmatrix} n_{21} & n_{22} \end{bmatrix} \\ &= \begin{bmatrix} h_1 s_1 + h_3 s_2 + n_{21} & h_2 s_1 + h_4 s_2 + n_{22} \end{bmatrix} \end{aligned} \tag{4.2}$$

In equation (4.1), it is seen that the received signal at each relay is a standard 16 QAM signal with Rayleigh fading, whereas in (4.2), the received signal at each relay is two QPSK signals with independent Rayleigh fading superposed together, which results an equivalent 16 QAM signal but not with a standard shape.

The received signals at the relays are first quantized then compressed by the Slepian-Wolf encoder. Signals from the relays are transmitted simultaneously to the destination via a Multiple Access Channel (MAC) (see section 2.7.1).

At the destination, the compressed information are first detected and decompressed by the Slepian-Wolf decoder, then according to different quantizing schemes, the decompressed information is either used to reconstruct the received signals at the relays (LBG quantization scheme, see section 4.4) or to provide soft information directly for the outer decoder (demapper-decoder scheme and demodulator scheme, see section 4.4).

The outer code serves two purposes: first, to correct errors due to the noise and the channel fading, and secondly to remove remaining errors after Slepian-Wolf decoding using the ARA decoder [37].

4.3 Implementation of Slepian-Wolf code

The Slepian-Wolf code generally includes following structure (see [70, 71, 72, 37, 40, 73] for details)

- Different sources are randomly interleaved before being independently encoded. Illustrate in Figure 4.3
- A joint decoder performs iterative decoding with soft-in/soft-out decoders [74]
- Each source has its own iterative decoder, and extrinsic information is exchanged between these decoders via the Log Likelihood Ratio (LLR) updating function (see equation 4.3). This is illustrated in Figure 4.4

Among these features, the LLR updating function is one of the most crucial steps in Slepian-Wolf code implementation. After each individual iterative decoder works out the extrinsic information for its own source, this extrinsic information will be put into the LLR updating

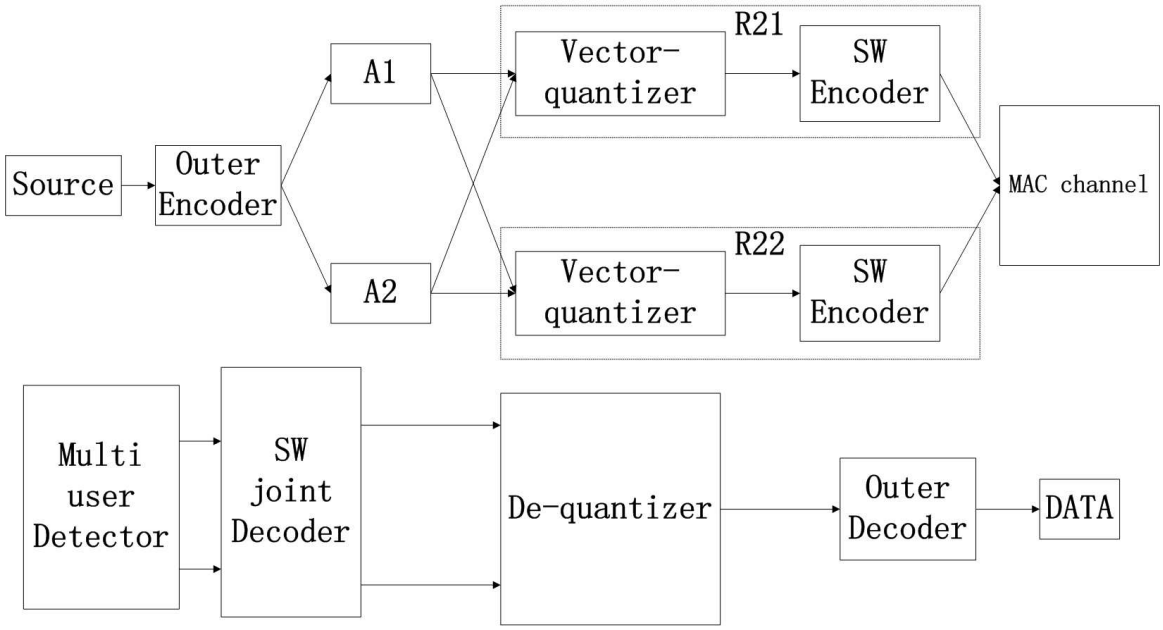

Figure 4.2: Overall system structure

function, which will produce the *a* priori information for the other source. The algorithms for the LLR updating function are [73]

$$Pr(\hat{b}=0) = (1-p)Pr(b=0) + pPr(b=1)$$
$$Pr(\hat{b}=1) = (1-p)Pr(b=1) + pPr(b=0)$$

where

$$p = \frac{\text{Numer of bits that are different between detected messages from the two relays}}{\text{length of the overall message}}$$

is the bit difference probability between two sources, $Pr(b=0)$ and $Pr(b=1)$ are the probability of a particular bit being 0 and 1 in one source, $Pr(\hat{b}=0)$ and $Pr(\hat{b}=1)$ are the probability of the corresponding bit being 0 and 1 in the other source.

For the iterative soft-in/soft-out decoders, the LLR updating function can be rewritten as [73]

$$L^{\hat{b}} = \ln \frac{(1-p)\exp(L^b)+p}{(1-p)+p\exp(L^b)} \tag{4.3}$$

where $L^{\hat{b}}$ and L^b are bit log likelihood ratio from each source.

The probability of difference p in equation (4.3) can be estimated by the following algorithm

$$p = \frac{1}{N}\sum_{n=1}^{N} \frac{\exp(L^b)+\exp(L^{\hat{b}})}{(1+\exp(L^b))(1+\exp(L^{\hat{b}}))} \tag{4.4}$$

where N is the number of bits in one frame.

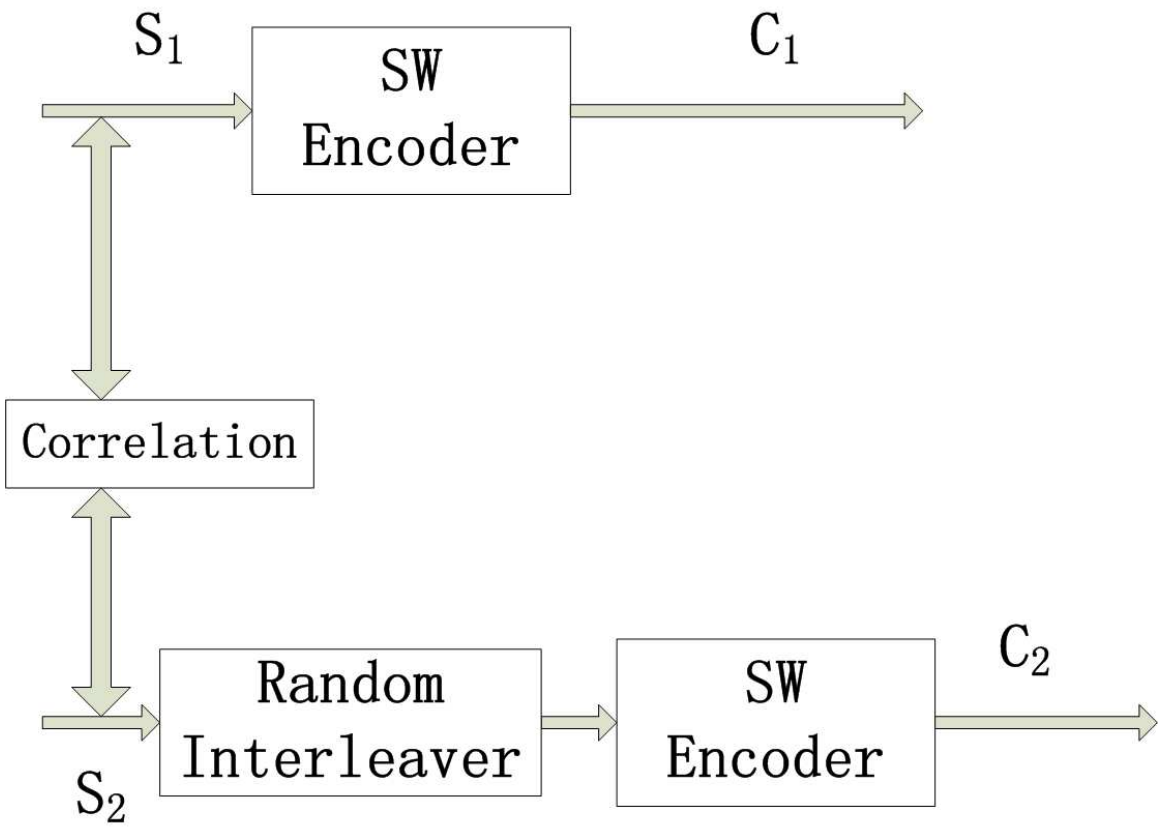

Figure 4.3: Slepian-Wolf encoder implementation

There are techniques that can improve the efficiency of the iterative decoding process, such as Perfect Doping [41]. This randomly places pilot bits into the coded message during the encoding procedure. The location and the bit values of the pilot bits are known by the decoder, so in the decoder procedure the corresponding LLR of those pilot bits are set to a large value, and the confidences of these bits are then spread across the whole frame which in turn helps to bring the decoding procedure to convergence.

4.3.1 Super-Turbo Code

The Super-Turbo code was introduced by Javier Garcia-Frias in his papers [70, 71, 72].

Encoder

Both sources are encoded by turbo encoder separately: the two encoders do not communicate with each other. The information bits generated from the second source are interleaved before they go into the encoder. The output of each turbo encoder is punctured to achieve the desired overall transmit rate. After puncturing every turbo encoder only transmits half the information bits and the same amount of parity bits. Two turbo encoders perform puncturing on even and odd position of the information bits respectively. Figure 4.5 shows the structure of the super-turbo code encoder.

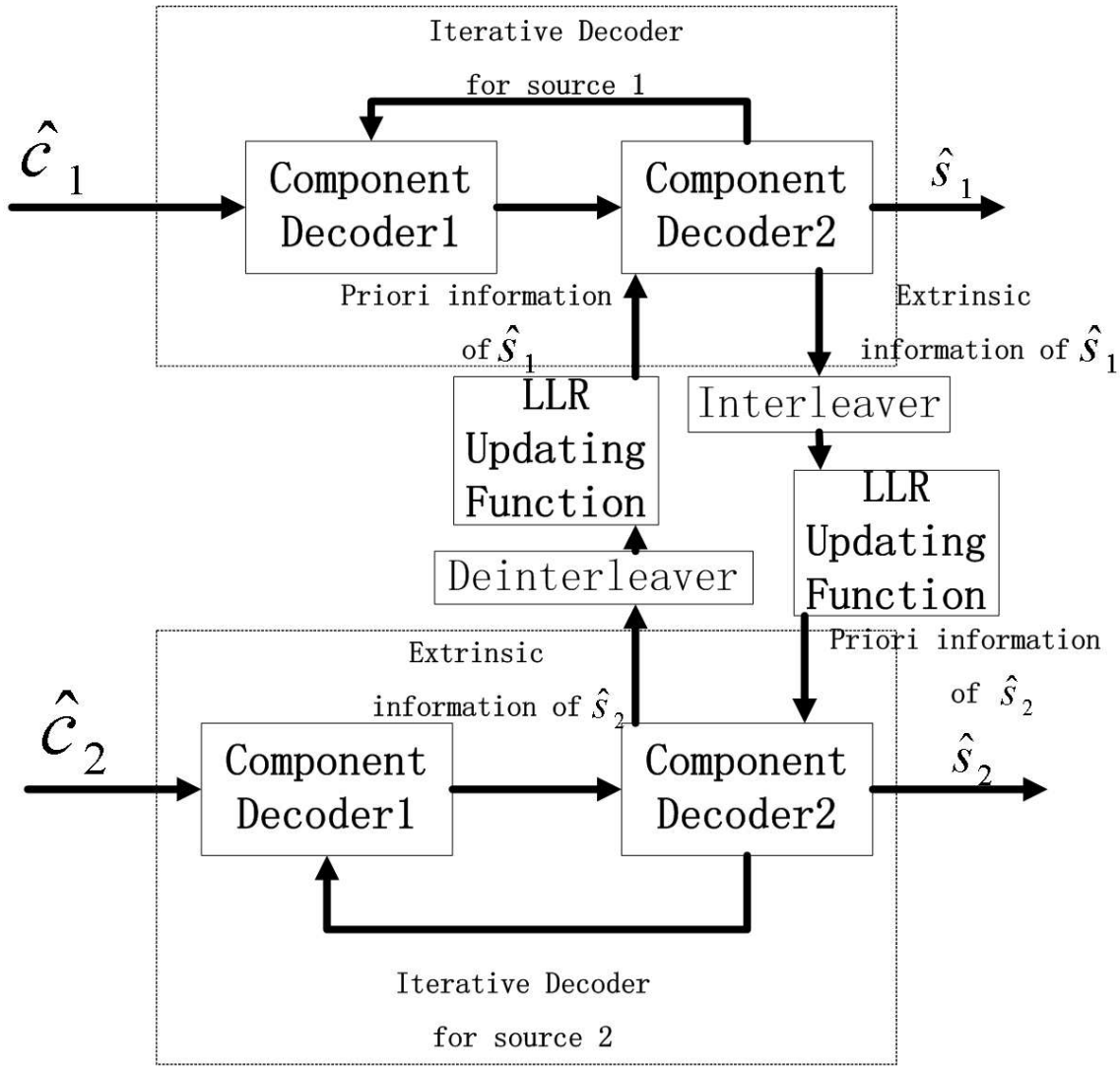

Figure 4.4: Slepian-Wolf decoder implementation, $\hat{c_1}$ and $\hat{c_2}$ are the LLR of the coded messages C_1 and C_2 in Figure 4.3, $\hat{s_1}$ and $\hat{s_2}$ are the estimations of data S_1 and S_2 in Figure 4.3

Decoder

The Slepian-Wolf decoder consists of two turbo decoders, each one associates with a source. Two turbo decoders perform iterative decoding with additional extrinsic information exchanged between them. Figure 4.6 shows the structure of the Slepian-Wolf encoder.

4.3.2 Accumulate-Repeat-Accumulate Code

Accumulate-Repeat-Accumulate(ARA) code is based on Repeat-Accumulate(RA) code[37, 40] and is introduced in paper[73] as a way to implement the Slepian-Wolf theorem. Its encoder and decoder structures are shown in Figure 4.7 and Figure 4.8. The ARA code is a concatenated code, in which normally the doped accumulator and the variable node encoder (VNE) are considered as the outer code and the check node encoder(CNE) and the other doped accumulator are considered as the inner code. The outer code and the inner code are connected

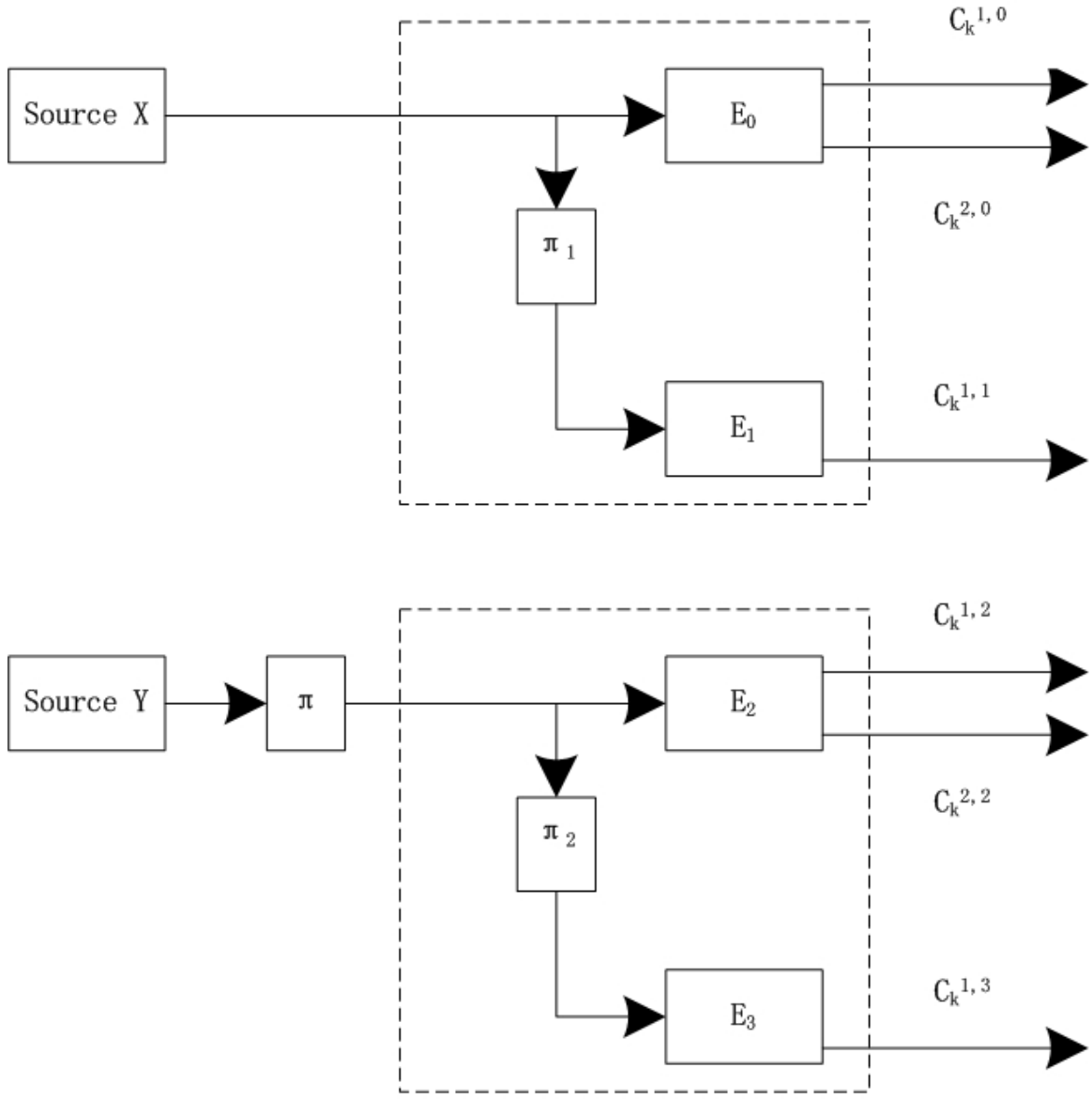

Figure 4.5: Super-turbo code encoder

by an interleaver.

Encoder

The encoder of the ARA code consists of the following components:

- doped accumulator
- variable node encoder(VNE)
- check node encoder(CNE)

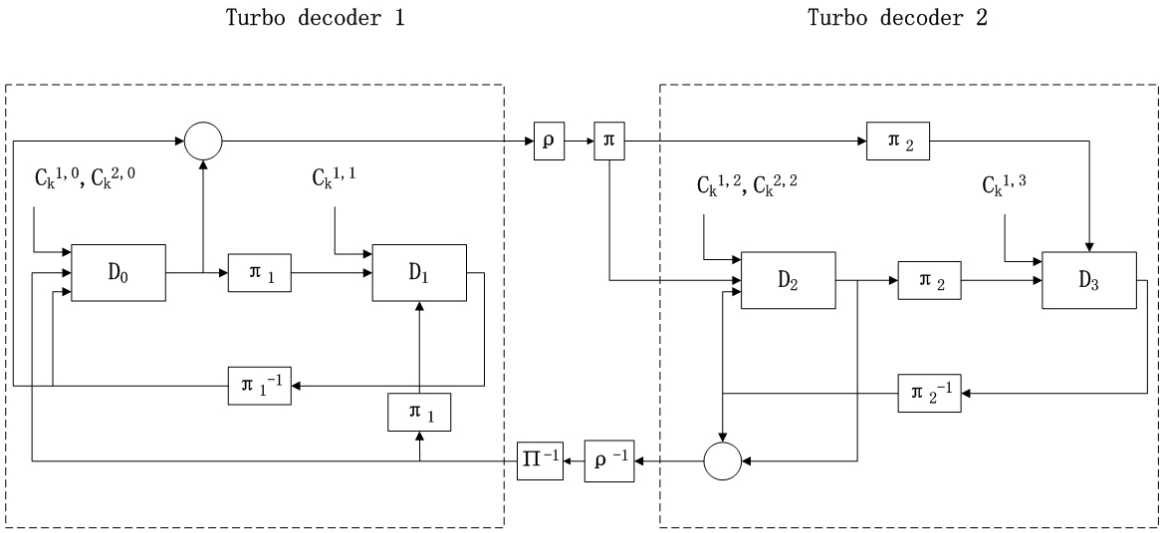

Figure 4.6: Super-turbo code decoder: component decoders D_0 D_1 D_2 D_3 correspond to component encoders E_0 E_1 E_2 E_3 respectively in the Slepian-Wolf encoder. Interleaver π π_1 π_2 and deinterleaver π^{-1} π_1^{-1} π_2^{-1} correspond to the interleaver π π_1 π_2 respectively in the Slepian-Wolf encoder. ρ and ρ^{-1} denote the LLR updating function, see equations (4.3) (4.4) in section 4.3

- interleaver

The accumulator is a memory one recursive convolutional encoder [28]. Doping is a technique to enhance the decoding convergence in iterative decoding procedure[41]. The structure of the doped accumulator is shown in Figure 4.9. The output of a doped accumulator consists of the original information bits from S and coded bits from C, if doping period is P(positive integer), the output is S with its every P_{th} bit replaced by the corresponding bit in C. This doping technique is referred as systematic doping in [41]. There are two doped accumulators in the ARA encoder, the one in the outer code, attached to the source, is the outer accumulator, which shall be activated only when the correlation of two sources is low (the two sources are more independent). The other one in the inner code, attached to the antenna, is the inner accumulator, which shall be activated only when signal to noise power ratio(SNR) is low.

The VNE is a repetition code, which simply repeats the input bits several times and outputs them. The number of times it repeats the input bits is called degree. If a different portion of the input bits are repeated different number of times, this kind of VNE is called an irregular VNE.

The CNE is a parity check code. It takes input bits and perform modulo 2 addition on them, then outputs the result. The number of bits that participate in the modulo 2 addition is called the degree. If a different portion of the input bits have different degrees, this kind of CNE is referred to as an irregular CNE.

The interleaver used in the ARA code is a random interleaver, which takes the output of the VNE and randomly shuffles the input bits, then outputs them to the CNE.

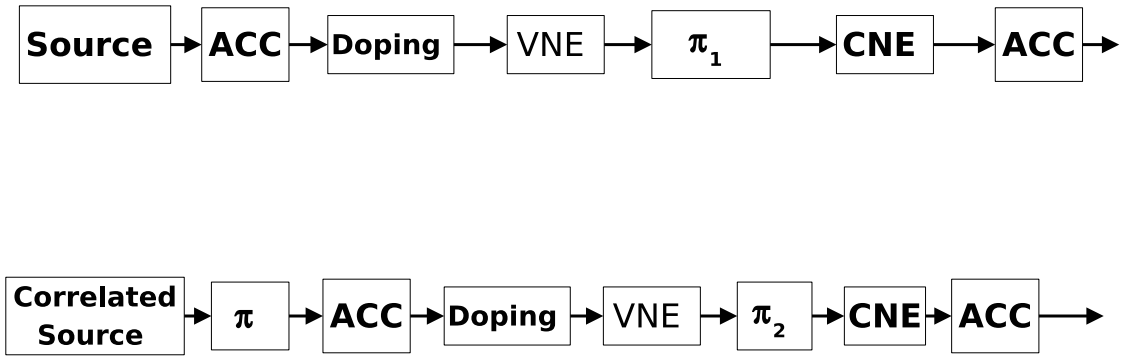

Figure 4.7: Accumulate-Repeat-Accumulate encoder

Decoder

The decoder of the ARA code consists of the following components:

- De-accumulator
- variable node decoder (VND)
- check node decoder (CND)
- LLR updating function, see equations (4.3) (4.4) in section 4.3

Each of these component decoders takes the LLR as input, and outputs LLR.

The de-accumulator is the decoder for the accumulator. It can be a BCJR decoder[57] or, since the accumulator is just a memory one convolutional code, de-accumulator can also use the simplified version of the BCJR algorithm [75] Algorithm 8.4.

A degree n variable node decoder takes $n+1$ inputs, one information bit and n coded bits. The corresponding output value $L_{i,out}$ is to add the rest of the input values $L_{j,in}$ together.

$$L_{i,out} = \sum_{j \neq i} L_{j,in}$$

A degree n check node decoder take $n+1$ inputs, one coded bit and n information bits. The corresponding output $L_{i,out}$ is to "box plus" the rest of the input values $L_{j,in}$

$$L_{i,out} = \sum_{j \neq i} \boxplus L_{j,in}$$

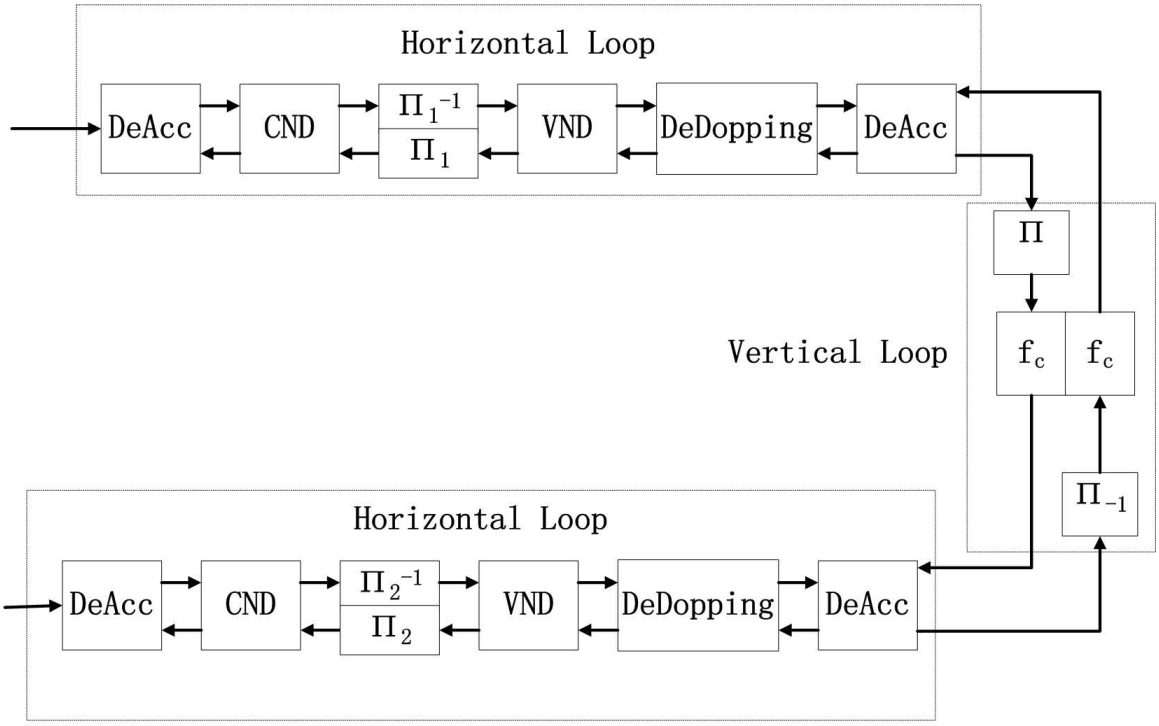

Figure 4.8: Accumulate-Repeat-Accumulate decoder

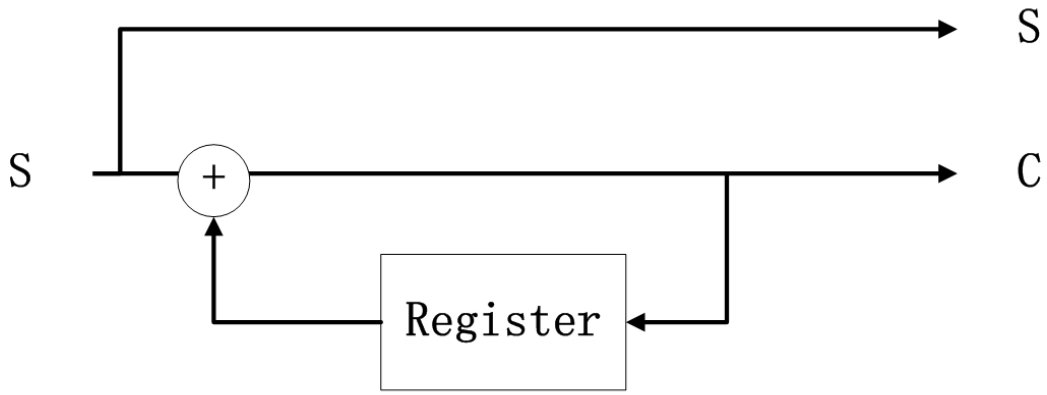

Figure 4.9: Doped accumulator

The rules for "box plus" are defined below [42]:

- $L_u \boxplus \infty = L_u$
- $L_u \boxplus -\infty = -L_u$
- $L_u \boxplus 0 = 0$

- $\sum_{j=1}^{J} \boxplus L(u_j) = log \frac{1+\prod_{j=1}^{J} \tanh(L(u_j)/2)}{1-\prod_{j=1}^{J} \tanh(L(u_j)/2)}$

As shown in Figure 4.8, the horizontal loop represents separate decoding for each source without knowledge from the other sources, whereas the vertical loop represents the exchange of extrinsic information between individual decoders.

In the joint decoding process, each source first does several horizontal loops, then one vertical loop takes place between them: this is regarded as one joint decoding loop. Normally, in order to successfully decode a frame several joint decoding loops are needed, and the number of loops depends on the SNR and the compression rate.

4.3.3 Performance Analysis

The performances of the Super-Turbo code is compared with that of the ARA code in this section in various aspects.

Complexity

The complexity of the Super-Turbo code is considered greater than that of the ARA code. Taking the one presented in [70, 71, 72] as an example, it uses recursive convolutional code with generator matrix $G(D) = \frac{1+D+D^2+D^3}{1+D^2+D^3}$ as its component code, whereas in the ARA code, its component codes are just memory one recursive convolutional codes (Accumulator) with generator matrix $G(D) = \frac{1+D}{1}$, a repetition code and a parity check code. If we use the simplified version of BCJR algorithm [75] Algorithm 8.4 in the Accumulator decoder, the complexity of decoding procedure can be further reduced.

Efficiency

Because the Super-Turbo code uses more complex generator matrix than the ARA code does, its error correction capacity is better than that of the ARA code. In the simulation conducted in this project, the Super-Turbo code normally only needs around 5 joint decoding loops to decode a frame, but it costs the ARA code 60 or even more than 100 loops to decode the same frame.

There are techniques that can improve the decoding efficiency of the ARA code, such as Perfect Doping [41]. Unlike Systematic Doping (see section 4.3.2), Perfect Doping introduces pilot bits into the coded during the encoding procedure. The location and the bit value of the pilot bits are known by the decoder, so in the decoder procedure the corresponding LLR of those pilot bits are set to a large value. The confidence of these bits is then spread across the whole frame which in turn helps to bring the decoding procedure to convergence.

Compression Rate

Table 4.1 shows the compression rate comparison between the two schemes.

p	0.025	0.05	0.1	0.2
$H(S_1, S_2)$	1.17	1.29	1.47	1.72
Compression rate achieved by ARA code	1.41	1.46	1.59	1.85
Compression rate achieved by Super-Turbo code	1.31	1.435	1.63	1.89

Table 4.1: Compression rate comparison between ARA code and Super-Turbo code, where $H(S_1, S_2)$ is the joint entropy of the correlated sources, and p is the probability of difference between the correlated sources, cited from [73]

Disadvantages

For the Super-Turbo code, because it uses puncturing to reduce its transmit rate, the puncture pattern is crucial to its performance, especially for those parity bits which are punctured by more than 50%. In [70, 71, 72], the description of the puncture pattern is extremely limited. We tried "uniform style" and "pseudorandom style" as mentioned in these papers, however both of them do not work very well. So far we have found several puncture patterns that work, however the principle for the puncture pattern design for the Super-Turbo code is still an open problem.

For the ARA code, because it uses CNE to reduce the transmit rate, puncture pattern design is not a problem at all. However, according to our simulation and [37], this code has an error floor at around 10^{-5} level.

4.3.4 Exit Chart Assisted Code Design

Figure 4.10 and 4.11 show the EXIT chart for the ARA decoding process.

Because in the ARA code, the outer code (the VNE and the outer accumulator) has two inputs (one from the inner code, the other from the other component decoder via the vertical loop) and one output, the EXIT chart has to be in 3 dimensions. In Figure 4.10, the blue surface represents the inner code (the inner Accumulator and the CNE), and its input is plotted on the Y axis and output is plotted on the Z axis. The red surface represents the outer code, the input from the inner code is plotted on the Z axis, the input from the vertical loop is plotted on the X axis and the output is plotted on the Y axis which is reversed compared to the inner code.

Several features of this chart need to be pointed out.

- Because the inner code is attached to channel directly, its surface(blue) does not start from $(0,0,0)$ point. The signal from channel gives it some initial information.
- Because the inner code does not receive information from vertical loop, so its surface(blue) is parallel to X axis which represents information from vertical loop.
- When the input from the inner code is zero (Z axis), even if the vertical loop gives full information (mutual information equals one on X axis), the output of the outer code (Y axis) still cannot reach the (1,1,1) point. That because, according to LLR updating algorithm (4.3), the output from the vertical loop is only an estimate of the data, the accuracy depends on the probability of difference (the proportion of the received information on the two relays that are different from each other 4.4). As long as this probability does not

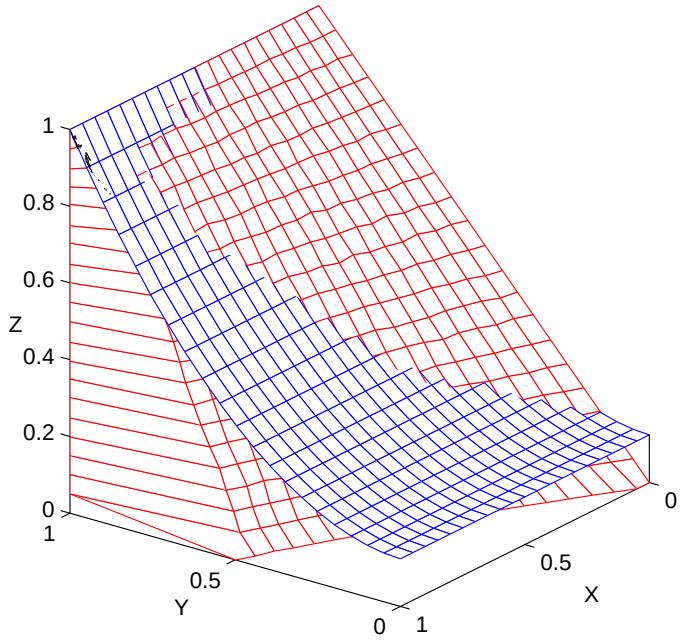

Figure 4.10: Exit chart for ARA code

equal 1 or 0, the mutual information of the input from vertical loop will be less than 1 (full information).

Figure 4.11 gives an alternative view of the EXIT chart which reveals the trajectory(black line) between the two surfaces. It starts at the $(0,0,0)$ point, and steps up all the way to the $(1,1,1)$ point by iteration.

The scenario for this particular decoding process is:

- Frame length 16380
- SNR is 10dB
- Probability of difference is 10%, joint entropy is 1.468996
- Transmit rate is 1.659585
- Number of iterations is 406, because no Perfect Doping is used.

4.4 Vector-Quantization

As shown in Figure 4.2, the vector quantizer converts the analogue signals to their digital representations, which are input to the Slepian-Wolf encoder. The performance of the vector quantizer has a significant impact on the overall system performance. The more similar the outputs of the vector quantizers at the two relays, the more the Slepian-Wolf code can compress. And the more accurate the reconstructed signal, the higher the probability that the outer decoder can correct the error.

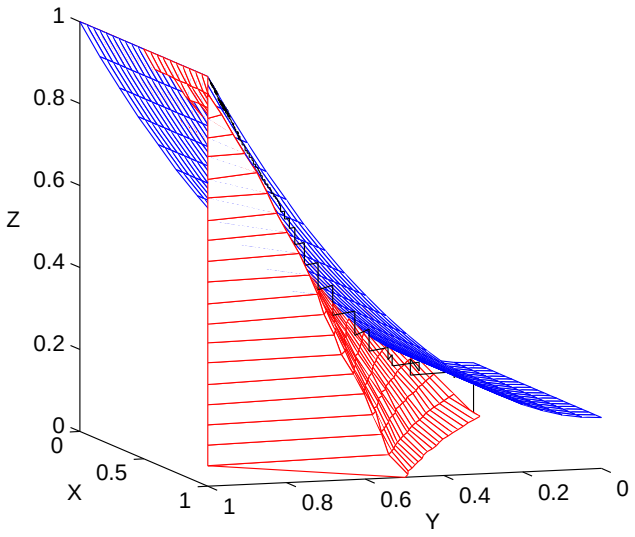

Figure 4.11: Trajectory for ARA code

4.4.1 Quantization

Several quantizers have been developed in this chapter, which are discussed in following sections.

Demapper and Outer Decoder

One option for the vector quantization at the relay is to use a soft demapper and the outer decoder, which is introduced in [27]. The algorithm for the demapper is

$$\begin{aligned} L_E(x_k|y) \approx & \frac{1}{2}\max_{X\in\mathbb{X}_{k,+1}}\{-\frac{1}{\sigma^2}\parallel y-H\cdot s\parallel^2+X_{[k]}^T\cdot L_{A,[k]}\} \\ & -\frac{1}{2}\max_{X\in\mathbb{X}_{k,-1}}\{-\frac{1}{\sigma^2}\parallel y-H\cdot s\parallel^2+X_{[k]}^T\cdot L_{A,[k]}\} \end{aligned} \tag{4.5}$$

where $\mathbb{X}_{k,+1}$ denotes the subset in which the k_{th} bit is $+1$ and $\mathbb{X}_{k,-1}$ denotes the subset in which the k_{th} bit is -1. Vector $L_{A,[k]}$ denotes the feedback LLRs from the decoder without the one for the k_{th} bit and vector $X_{[k]}^T$ denotes the corresponding bit values without the one for the k_{th} bit. y, H, s, σ denote received signal, channel state information, transmitted symbol and noise standard deviation. The use of the approximately equal sign in equation (4.5) is due to the fact that this algorithm uses the Max-log approximation.

The initial information available to the demapper is

- Channel State Information (CSI) H
- Noise variance σ^2
- Constellation used by the transmitter s

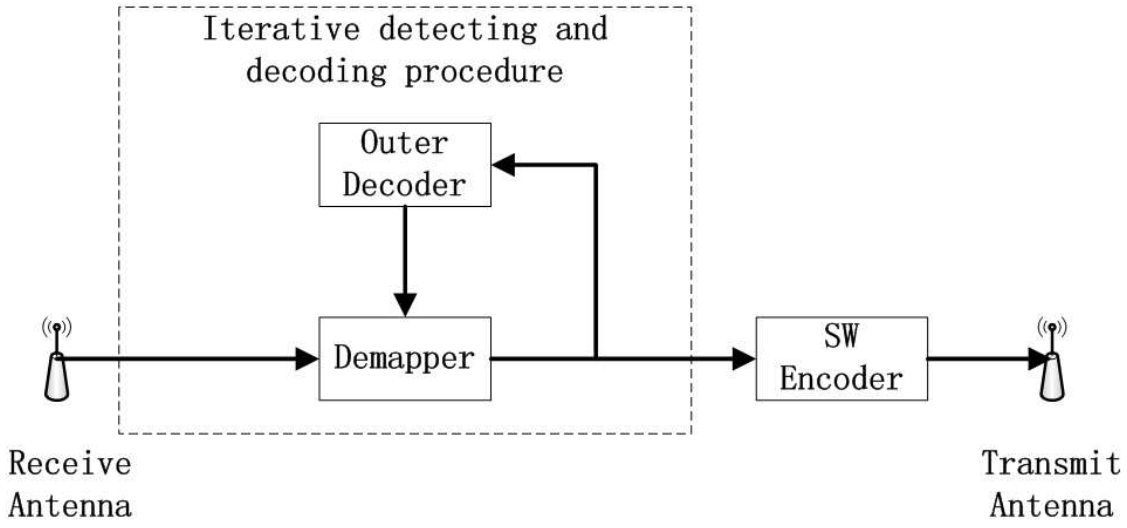

Figure 4.12: Demapper and decoder quantization algorithm

For each symbol, the demapper takes the complex analogue received signal value y, and the feedback soft information from the outer decoder $L_{A,[k]}$ as input (in the first iteration $L_{A,[k]}$ is set to 0). For each bit in this symbol, the demapper considers all possible symbols in the constellation, and calculates the symbol-wise value for each $-\frac{1}{\sigma^2} \parallel y - H \cdot s \parallel^2$ plus bitwise value $X_{[k]}^T \cdot L_{A,[k]}$, where vector $X_{[k]}^T$ represents the bits value for the corresponding symbol without the current bit, $L_{A,[k]}$ represents the corresponding LLR which also excludes the current bit. All these values are then divided into two groups according to the current bit value, and the largest from each group is chosen to form the final value for this bit. After several iterations between demapper and the outer decoder a hard decision is taken at the output of the demapper. The quantization results are then input into the Slepian-Wolf encoder. Figure 4.12 shows the quantizer structure.

Demodulator

The demodulator uses the Maximum Likelihood algorithm. For every received symbol, the demodulator calculates its distances to all symbols in the constellation, then chooses the one with the shortest distance to the received symbol as the estimated output.

$$\hat{s} = \arg\min_{s \in S} \{ \parallel y - H \cdot s \parallel^2 \}$$

where S contains all symbols in the constellation, H is CSI, y is the received symbol complex value.

LBG Algorithm

Vector quantization (VQ) is a vector quantization scheme that allocates quantization points (codewords in the codebook) on the constellation according to the distribution of the samples to minimize the averaged mean square quantization error.

Details of the LBG algorithm are illustrated below[76].

1. Given τ (the threshold to confirm the convergence). Fix $\varepsilon > 0$ (the range to split the quantization points) to be a "small" number, we use $\varepsilon = 0.01$ in our system.

2. Let $N = 1$ (number of quantization points) and

$$c_1^* = \frac{1}{M}\sum_{m=1}^{M} x_m$$

where x_m are quantization samples. Calculate

$$D_{ave}^* = \frac{1}{Mk}\sum_{M}^{m=1} \| x_m - c_1^* \|^2$$

where M is number of quantization samples, k is number of sample dimensions.

3. Splitting: For $i = 1, 2, \ldots, N$, set

$$c_i^{(0)} = (1+\varepsilon)c_i^*$$
$$c_{N+i}^{(0)} = (1-\varepsilon)c_i^*$$

Set $N = 2N$

4. Iteration: Let $D_{ave}^{(0)} = D_{ave}^*$. Set the iteration index $i = 0$.

i. For $m = 1, 2, \ldots, M$, find the minimum value of

$$\| x_m - c_n^{(i)} \|^2$$

over all $n = 1, 2, \ldots, N$. Let n^* be the index which achieves the minimum. Set

$$Q(x_m) = c_{n^*}^{(i)}$$

ii. For $n = 1, 2, \ldots, N$, update the code vector

$$c_n^{(i+1)} = \frac{\sum_{Q(x_m)=c_n^{(i)}} x_m}{\sum_{Q(x_m)=c_n^{(i)}} 1}$$

iii. Set $i = i + 1$

iv. Calculate

$$D_{ave}^{(i)} = \frac{1}{MK}\sum_{m=1}^{M} \| x_m - Q(x_m) \|^2$$

v. If $(D_{ave}^{i-1} - D_{ave}^{(i)})/D_{ave}^{(i-1)} > \varepsilon$, go back to Step (i).

vi. Set $D_{ave}^* = D_{ave}^{(i)}$. For $n = 1, 2, \ldots, N$, set

$$c_n^* = c_n^{(i)}$$

as the final code vectors.

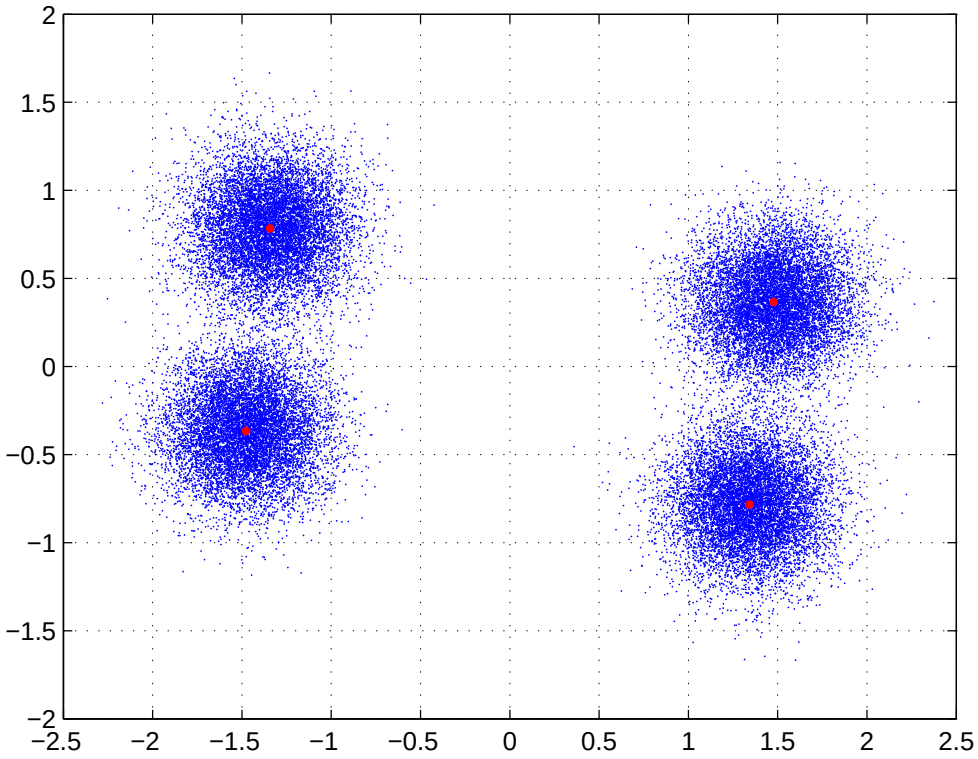

Figure 4.13: Vector quantization result for two dimensions random variable. Blue dots are samples and red dots are quantization points.

5. Repeat Steps 3 and 4 until the desired number of code vectors is obtained.

Because the LBG algorithm does not need CSI or modulation constellation, its implementation is more flexible, however its output is less stable.

The LBG algorithm differs fundamentally from the Demapper and decoder and the Demodulator algorithms. For both the latter algorithms an estimate is formed of the data symbol itself, so their outputs are the estimates of the coded message at the output of the outer encoder. However for the LBG algorithm, its output is a digital representation of the received analogue signal, rather than the data symbol. Example of the LBG algorithm is shown in Figure 4.13 in which samples (blue dots) can be generally divided into four groups on the constellation, the result of the LBG algorithm is that it allocates one quantization point (red dot) to each group which minimized the averaged mean square quantization error.

4.4.2 De-quantization

De-quantization processes the output of the Slepian-Wolf decoder and forms the input for the outer decoder. Two methods can be used in this system.

LLR Combination

This method works only for approaches that produce an estimate of the coded message at the relay, namely the Demapper-decoder and the Demodulator.

The outputs of the Slepian-Wolf decoder are two arrays of LLR, each of which represents the estimated coded message at the corresponding relay. They are then combined to form one array of LLR which is then input into the outer decoder. [77]

$$L_{combined} = \ln(\frac{1-p_1}{p_1})sign(L^1) + \ln(\frac{1-p_2}{p_2})sign(L^2)$$

where p_1, p_2 are the probability of error at each relay and L^1, L^2 are the output LLR from the Slepian-Wolf decoder, the function sign() takes the sign (positive or negative) of its argument. Note that this method requires the destination to know the probability of error at the relay (the method to obtain p_1, p_2 at the relays has not yet been published by the author of [77]) and since it directly produces the soft information for the outer decoder, the MIMO detector is not required.

Signal Reconstruction

This method, developed in this project, can be implemented more flexibly than the LLR combination method, so it can be used by the Demapper-decoder, the demodulator and the LBG algorithm. The first step is to reconstruct the quantized signal: each of the Slepian-Wolf decoder outputs is converted back to an analogue complex signal using

$$\hat{s} = \sum_{m=1}^{M} s_m \prod_{n=1}^{N} P(s_m^n)$$

where M is the number of symbols on the constellation, N is the number of bits on one symbol, S_m^n represents the n_{th} bit on the m_{th} symbol (the subscript th represents the ordinal number), $P(S_m^n)$ represents the probability of that bit being 1 or 0 according to S_m^n, S_m represents the analogue complex value of the m_{th} symbol, $\hat{s}$ is the reconstructed signal.

Then these signals are fed to a MIMO detector which uses the same algorithm as the Demapper and the outputs are then input into the outer decoder.

Note that because the MIMO detector also takes feedback from the outer decoder, the whole process can be an iterative detection and decoding process.

4.5 System Performance and Analysis

Simulated end-to-end BER results are shown in Figure 4.14. The blue line represents the configuration that uses the Demapper-decoder for quantizing and LLR Combination for dequantizing. The purple line shows the result of the demodulator coupled with Signal Reconstruction and the red line corresponds to the Demodulator and LLR Combination set. As shown in Table 4.2 transmit rates from relay group two to destination are adjusted according to the SNR

EbN0	10	20	30	40
transmit rate	2	1.904800	1.801800	1.739200

Table 4.2: Compressed transmission rates versus SNR that result first order diversity FER

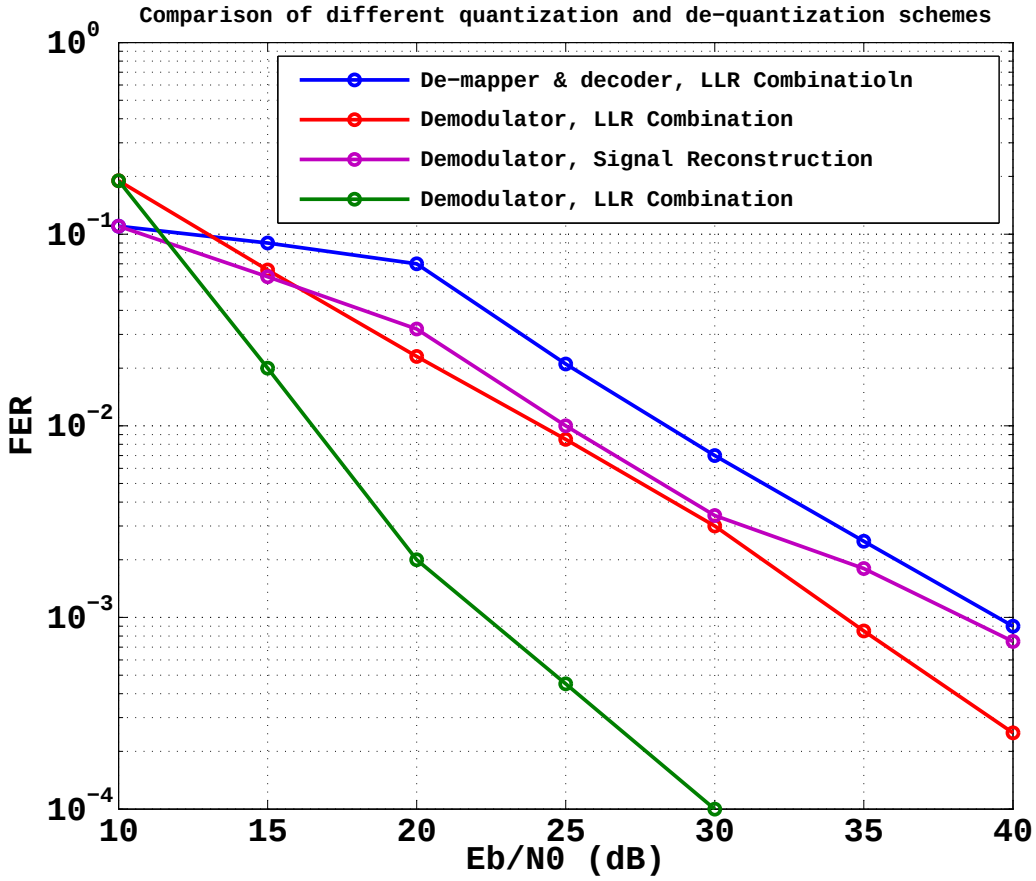

Figure 4.14: Simulation results for different Quantization and De-quantization configurations, blue line for (Demapper-decoder, LLR Combination), purple line for (Demodulator, Signal Reconstruction), red line for (Demodulator, LLR Combination), green line for (Demodulator, LLR Combination, uncompressed), frame length is 5000 bits.

4.5.1 Analysis

The Demapper and decoder algorithm, takes into account not only the CSI and modulation constellation, but also the LLR feedback from the decoder, so it has the greatest error correction capability among these three. It has been demonstrated by simulation of BER at the relay (not end-to-end BER) that this algorithm has the lowest Bit Error Rate (BER) at the relay (comparing the quantization results with the output of the outer encoder at the transmitter) among these three schemes.

For the Demodulator, because it only uses CSI and modulation constellation, it has poorer error correction capability at the relay than the Demapper and decoder algorithm. This likewise

has been confirmed by simulation.

However, for the LBG algorithm, because its output is not the code message but the digital representation of the received analogue signal, there is no way to measure the BER.

Besides the BER at the relay, another crucial issue is the difference between the outputs at the two relays after quantization. A larger difference results in larger joint entropy, and if the transmit rate from the relays to the destination is fixed, when the joint entropy is greater than the transmit rate, decoding will fail.

The simulation reported in Figure 4.15 shows that the demapper-decoder algorithm is more likely to result in an excessive difference between the decoded coded message at the relays, and hence a joint entropy greater than the transmit rate, which will cause decoding failure at the Slepian-Wolf decoder. The demapper-decoder algorithm (right hand side of Figure 4.15) gives a higher proportion of detected coded message at the relays which are the same (as shown by the height of the histogram bar at 0) than does the demodulator algorithm (left hand side). However, when the detected coded message at the relays are not exactly the same, the demodulator algorithm results in fewer bit differences between relays than the demapper-decoder algorithm, and of those coded message that are different, for the demapper-decoder algorithm about 40% of the bits in the frame are different while for demodulator only 20% on average are. This occurs due to error extension effects in the recursive convolutional decoder when the signal level at a relay is insufficient to allow correct decoding.

This effect of the demapper and decoder algorithm means that more frames are decoded incorrectly than for the demodulator algorithm. This is justified by Figure 4.14, as the red line is lower than the blue line.

For the LBG algorithm, because its output is less stable than the Demapper and decoder and the Demodulator algorithms, it normally results in the largest joint entropy among these three, so results for LBG algorithm are not shown in Figure 4.14.

For the two de-quantizing schemes, the LLR combination scheme takes the error probability on each relay as an extra information source, so it gives better performance, however, this extra information increases the complexity of the relay since it has to estimate this probability of error and this scheme works only for quantizers that output estimates of the code message (it is not compatible with the LBG algorithm).

The signal reconstruction method on the other hand, has a more flexible application and requires less information, but this advantage comes at a cost to its performance. This is shown in Figure 4.14 as the red line is lower than the purple line.

4.5.2 Trade-off Between Error Correction Capability and Channel Capacity

There is a trade-off in this system between error correction capability and channel spectral efficiency, depending on the number of quantization points at the relay. More quantization points result in more accurate reconstruction of the received signal, but also requires more bits to represent the quantization points, so the data rate after the quantization is increased. Furthermore, more accurate quantization increases the probability of different quantization results at the two relays, which increases the joint entropy.

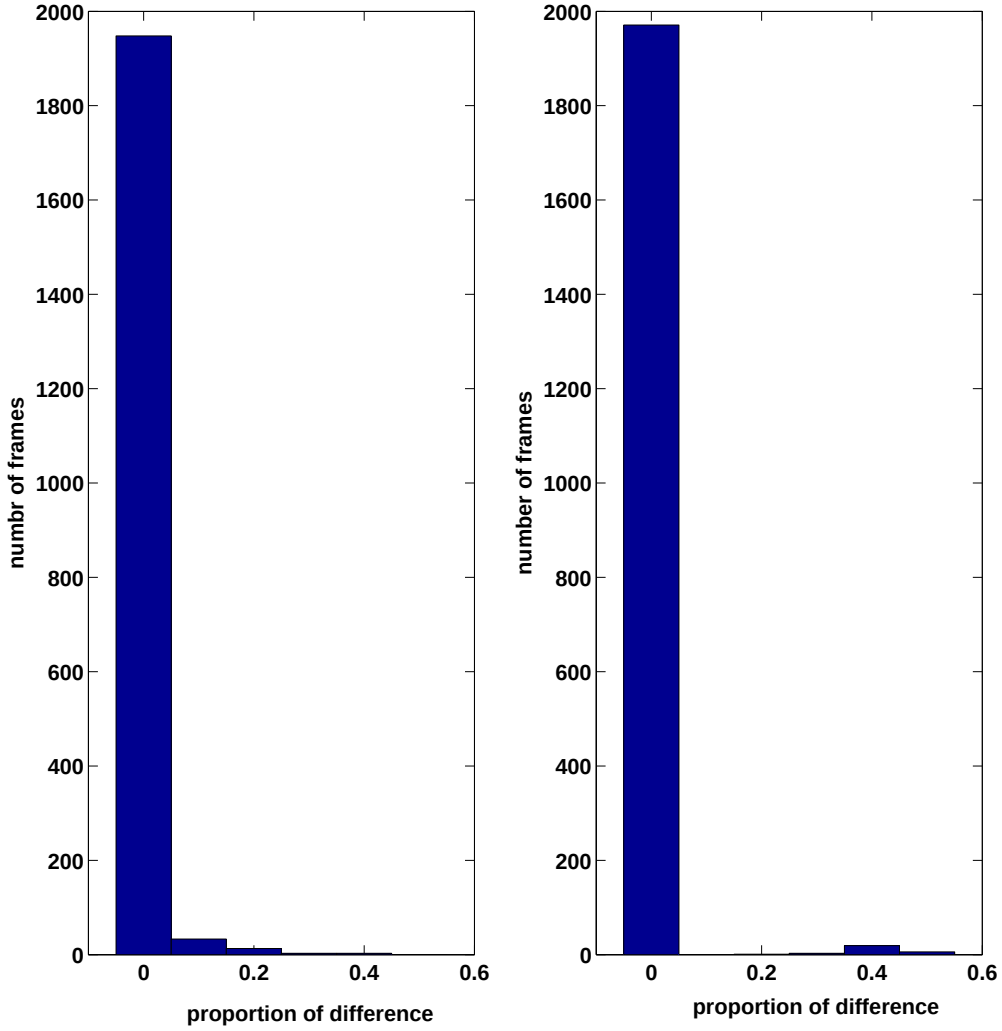

Figure 4.15: Histogram of the difference between the detected coded messages at the relays (p in equation (4.3)) for 2000 frames at 30dB SNR. Y axise denotes the number of frames, X axise shows the proportion of bits that are different between the detected coded messages at the relays. The left hand graph corresponds to the demodulator algorithm, right hand to the demapper and decoder algorithm.

4.5.3 Diversity Order and Transmit Rates

There are two factors that can cause frame errors in this system fading and noise at the relays; and joint entropy in excess of the transmit rate. When the channel to only one of the relays is in a deep fade, this will cause a large number errors in the detected coded message at this relay, which will in turn cause a large joint entropy. However this frame can still be decoded correctly if the transmit rate is greater than this joint entropy because the other relay provides reliable information to the destination. Thus if the transmit rate is sufficient, full diversity order is provided, but for greater compression only first order is possible. Table 4.2 shows the transmit

rates that achieve first order diversity for system using Demodulator and LLR combination, and this is demonstrated by the corresponding FER curves(red and green line) in Figure 4.14.

4.6 Conclusion

Based on our study in Chapter 3, we propose an advanced version of the distributed cooperative spatial multiplexing system . By implementing the Slepian and Wolf theorem, the system has higher spectral efficiency but still lower BER. Our contributions are listed below:

- We test two Slepian and Wolf theorem implementation methods, namely the super turbo code and the accumulate repeat accumulate code. Finally decide to use the ARA code because of its practicability.
- We develop several quantization and reconstruction schemes. Simulation results and analysis show that the demodulator quantization method coupled with the LLR combination de-quantization method give the best result.
- We discover an interesting property of this system structure that lower FER at relay does not correspond to lower BER at the destination: the similarity between the data at the relays also affects the FER at the destination.

Chapter 5

Error Floor Analysis for Receiver with Distributed Relays

5.1 Introduction

In this chapter we analyse system performances based on the design in chapter 4.

The system concept is illustrated in Figure 5.1. The transmitter first broadcasts information to the receive VAAs (green, purple, orange, black lines). Then, the receiver gathers information via intra-links (red lines) from all the relays in the receive VAAs. All relays in this system use decoder-and-forward protocol with fixed gain.

In this chapter, uncompressed information is transmitted from the relays in order to simplify the analysis.

We assumed that the relays do not perform full decoding, but they only perform demodulation which leads to a situation where the message from the source can not always be detected correctly, thus although the information at the two relays in the receive VAA are originated from the same source (information B and C are originated from A in Figure 5.2), the demodulated information at each relay (information B and C) contains random errors. Therefore we put our attention on the chief executive officer (CEO) problem [33] point of view in this chapter. Information D and E (link D_2 further introduces errors into information B and C, resulting information D and E.) are two noisy observations of the source information A, the combiner at the receiver needs to estimate the information A from these two noisy observations. Since errors may be included in the information part before combining, there is an inevitable error floor between the estimated information F and the source information A. In this chapter, we derived an analytical expression for the error floor, which will be useful in optimization (such as code rate design for the outer code). Our work is verified through Monte-Carlo simulations.

The chapter is set out as follows. In section 5.2, we introduced our system structure, transmission protocol and the assumption we made. Analytical derivation for the error floor is shown in section 5.3, followed by simulation results in section 5.4 which verified our work. Finally we conclude our contribution in section 5.5.

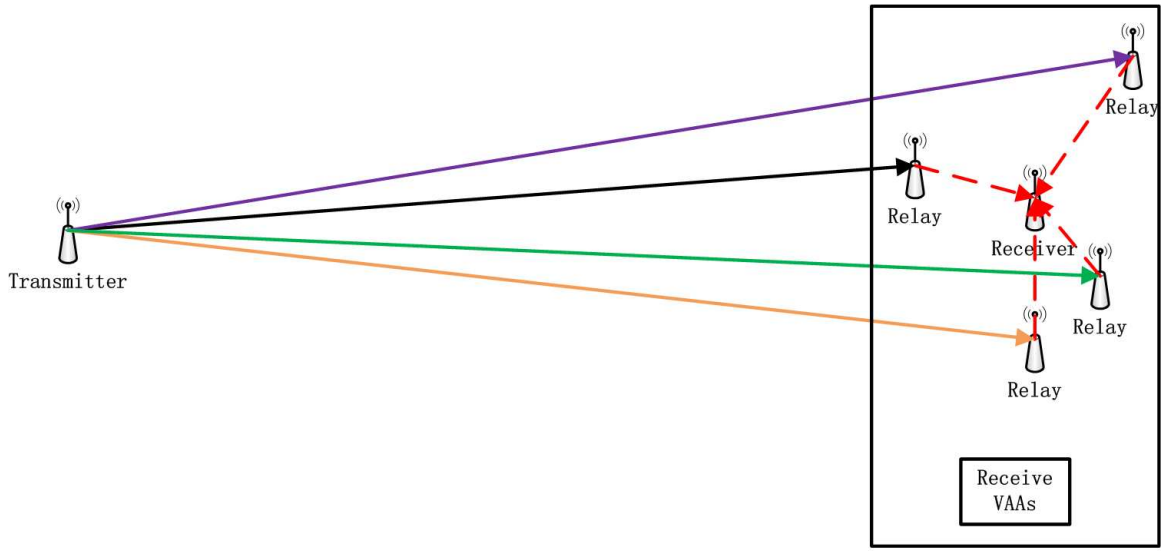

Figure 5.1: Virtual antenna array concept: broken arrow lines denote links within virtual antenna arrays, solid arrow lines represent links between transmitter and VAAs.

5.2 System Model

The system structure is shown in Figure 5.2. Data is first encoded by a convolutional code followed by an accumulator (outer code) before modulation. The transmitter broadcasts the signals to the relays in receiver relay group via the block Rayleigh fading channel with unity variance, the received signals at each relay are demodulated and encoded again by a convolutional code followed by an accumulator. The encoded messages from each relay in the receiver relay group are forwarded to the receiver via a multiple access channel(MAC). At the receiver, messages from each relay in the receiver relay group are first separated by a multi-user detector and then decoded jointly by the joint decoder which includes two convolutional decoders (corresponding to the two convolutional encoders at the relays) and a vertical loop between these two decoder to exchange the correlation information. The decoded information from each relay in the receiver relay group are combined before sent into the outer code decoder to estimate the data.

Following are assumptions made in this system.

- Link D1 suffers from independent block Rayleigh fading. Link D2 is an AWGN channel.
- The receiver relay group is composed of two relays. Every node(the transmitter, the receiver and all the relay nodes) in this system has only one antenna attached.
- Multi-user detector at the receiver can separate the signal perfectly.

5.3 Error Floor Analysis

At the receiver after decoding the information from each relay in the receiver relay group, the combiner needs to estimate the original information (information A) from two erroneous

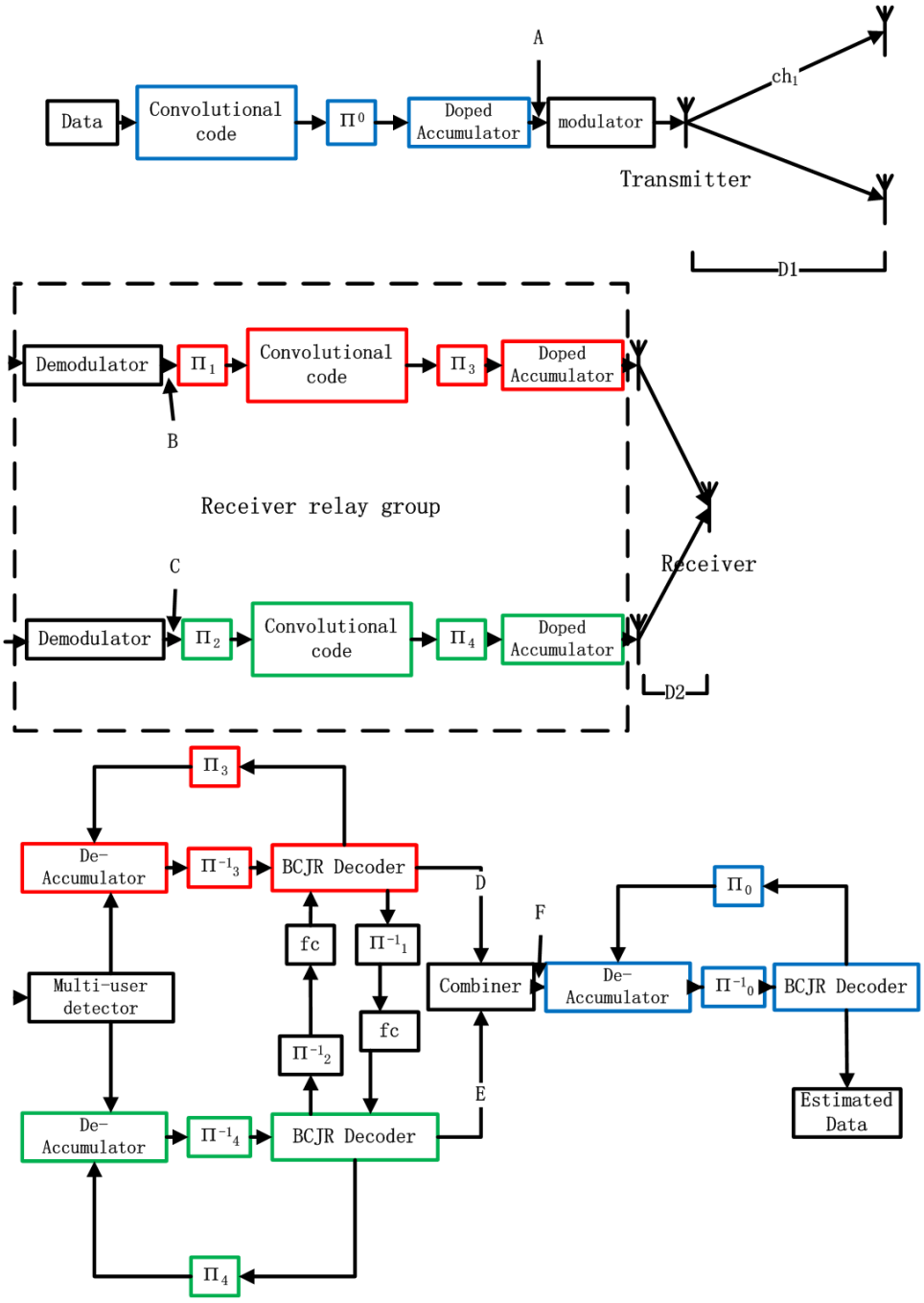

Figure 5.2: System structure: the upper part of the figure shows the system structure from the source to the receive antenna of the receiver relay group, the middle part is from relays in the receiver relay group to the receive antenna of the receiver, the lower part is the receiver structure. The corresponding encoder and decoder are marked in the same colour. The fc function is defined in [77].

copies (information D and E, which are the estimation of information B and C respectively) of it which can be interpreted as a CEO problem (the output of the combiner, information F, is the estimation of the information A).

Because the information B and C are erroneous copies of the information A, even if the joint decoder can recover the information B and C perfectly (information D and E are exact the same as the information B and C respectively), the combiner is still unable to reproduce the information A completely, as a consequence there is an error floor when measure the bit error rate(BER) of the information F against the information A.

5.3.1 Error Floor for Fixed BER

We assume that the BER of information B and C against information A are p_1 and p_2 respectively. So for any bit in information A, the probability that its corresponding bits in B and C are all wrong is p_1p_2 and the probabilities that only one of them goes wrong are $(1-p_1)p_2$ (corresponding bit in C goes wrong) and $(1-p_2)p_1$ (corresponding bit in B goes wrong). And because the combiner uses the algorithm in equation (4.4.2)

$$L_{combined} = \ln(\frac{1-p_1}{p_1})sign(L^1) + \ln(\frac{1-p_2}{p_2})sign(L^2)$$

where $L_{combined}$, L^1, L^2 are the LLR representing F, D and E, the function sign() takes the sign (positive or negative) of its argument [77], when $p_1 > p_2$, $sign(L_{combined}) = sign(L^2)$ and the probability that the corresponding bit in F goes wrong is $p_1p_2 + (1-p_1)p_2 = p_2$ and when $p_1 < p_2$, $sign(L_{combined}) = sign(L^1)$ and the probability of error in F is $p_1p_2 + (1-p_2)p_1 = p_1$, which can be summarized as

$$min(p_1, p_2) \tag{5.1}$$

The situation of $p_1 = p_2 = p$ rarely happens, we ignore it in this chapter.

5.3.2 Probability Density Function (PDF) of the BER at the Relay

Because of the independent block Rayleigh fading channel in link D1 and the white Gaussian noise at the relays in the receiver relay group, BER at each relay in the receiver relay group, instead of being a constant, is a random variable .

According to [78], we use the probability transformation method [79] to calculate the PDF of the BER at the relay.

Regarding the BER at the relay, because the demodulator at the relay uses the maximum likelihood (ML) algorithm and we assume that errors only happens between the symbol and its nearest neighbour, the symbol error rate (SER) for each individual symbol on the constellation can be approximately calculated as

$$P_s \approx Q(\frac{d_{min}}{2\sqrt{N_0/2}}) \tag{5.2}$$

where $Q()$ denotes the Q-function, d_{min} represents the distance between the symbol and its nearest neighbour on the constellation and $N_0/2$ is the noise power spectral density on each dimension.

Furthermore we assume that every symbol on the constellation is equiprobable, therefore the overall SER is

$$P_s = \sum_{i=1}^{m} \frac{1}{n} N_i Q(\frac{d^i_{min}}{2\sqrt{N_0/2}}) \tag{5.3}$$

where m is the number of symbols on the constellation, N_i is the number of the nearest neighbours for the ith symbol and d^i_{min} is the distance between the ith symbol and its nearest neighbour. Because Gray mapping is applied, the overall BER is [80]

$$P_e = \frac{1}{N_b} P_s \tag{5.4}$$

where N_b is the number of bits per symbol.

Regarding the distribution of d_{min}, we find out it follows the Rayleigh distribution with $\sigma = \frac{d}{\sqrt{2}}$ where d is the distance between symbols on the constellation (see section "PDF of the Minimum Distance" below). So the PDF of d_{min} is

$$p_{d_{min}}(x) = \frac{x}{\sigma^2} \exp(\frac{-x^2}{2\sigma^2}) \tag{5.5}$$

According to [79], PDF of the BER is

$$p_e = \frac{1}{|P_e(P_e^{-1})'|} p_{d_{min}}(P_e^{-1}) \tag{5.6}$$

where P'_e is the derivative of P_e and P_e^{-1} is the real root [81] of P_e.

PDF of the Minimum Distance

Assume coefficients for channels ch_1 on D1 link (as shown in Figure 5.2) is $a_1 + b_1 i$, where a_1, b_1 are normally distributed independent random variables with mean $\mu = 0$ and standard deviation $\sigma_2 = \frac{1}{\sqrt{2}}$, i is imaginary unit. The distance between the transmitted signal $s_1 = c_1 + d_1 i$ and $s_2 = c_2 + d_2 i$, wherec_1, d_1, c_2, d_2 are constants, at the relay in the receiver relay group is

$$\begin{aligned} d_{min} =& ||ch_1 s_1 - ch_1 s_2|| \\ & ||((a_1 + b_1 i)(c_1 + d_1 i)) - ((a_1 + b_1 i)(c_2 + d_2 i))|| \\ =& ||a_1(c_1 - c_2) + b_1(d_2 - d_1) \\ & + (b_1(c_1 - c_2) + a_1(d_1 - d_2))i|| \\ =& \sqrt{(a_1^2 + b_1^2)((c_1 - c_2)^2 + (d_1 - d_2)^2)} \\ =& \sqrt{(da_1)^2 + (db_1)^2} \end{aligned} \tag{5.7}$$

where $||\ \ ||$ is the Euclidean distance, $d = \sqrt{(c_1 - c_2)^2 + (d_1 - d_2)^2}$, is the distance between symbols on the constellation. According to the property of the normal distribution [64], variable da_1 and db_1 follows $\mathcal{N}(\mu_2 = 0, \sigma_3 = d\sigma_2 = \frac{d}{\sqrt{2}})$. Finally according to the property of the Rayleigh distribution [64], d_{min} is a Rayleigh random variable with parameter $\sigma_4 = \frac{d}{\sqrt{2}}$.

5.3.3 Error Floor for Variable BER

From equation (5.6), it's clear that if the SNR on D1 link is fixed then σ in equation (5.5) and (5.6) is fixed, so the PDF of the BER at the relays in relay VAAs is fixed. Then the error floor is the average of (5.1) over variables p_1 and p_2 whose PDF are given by equation (5.6).

According to [79], PDF of $min(x,y)$, where x and y are two independent random variables, is

$$\begin{aligned} &p(w = min(x,y)) = f_x(w)R_y(w) + f_y(w)R_x(w) \\ &R_x(x) = 1 - F_x(x) \\ &R_y(x) = 1 - F_y(x) \end{aligned} \tag{5.8}$$

where $f_x()$ and $F_x()$ are the PDF and the Cumulative distribution function (CDF) of x. Same rules also apply to y.

The PDF of the BER is (5.6), then the CDF of the BER is

$$F(x) = \int_0^x p_e(u)\,\mathrm{d}u \tag{5.9}$$

so the PDF of $min(p_1, p_2)$ is

$$p_{min}(x) = 2(1 - F(x))p_e(x) \tag{5.10}$$

Finally the average of the $min(p_1, p_2)$ is calculated as

$$E[P_{errorfloor}] = \int x p_{min}(x)\,\mathrm{d}x \tag{5.11}$$

Because it's unable to work out the closed form expression for (5.11), we use the numerical integration to calculate the average error floor value.

5.4 Simulation

In this section, we show the simulation results to verify our derivations in section 5.3.

We set the parameters as follow: all convolution codes are half rate, doping rate for all the accumulators is 2 and we use 16 QAM modulation on D1 link with the distance between symbols $d = 2$.

For 16 QAM modulation format, there are four symbols on the constellation that have two nearest neighbours, four symbols that have four nearest neighbours and eight symbols that have three nearest neighbours. Therefore the overall SER in (5.3) is

$$\begin{aligned} P_s =& \frac{4}{16} * 2Q(\frac{d_{min}}{2\sqrt{N_0/2}}) + \frac{8}{16} * 3Q(\frac{d_{min}}{2\sqrt{N_0/2}}) \\ &+ \frac{4}{16} * 4Q(\frac{d_{min}}{2\sqrt{N_0/2}}) \\ &= 3Q(\frac{d_{min}}{2\sqrt{N_0/2}}) \end{aligned} \tag{5.12}$$

and the overall BER in (5.4) is

$$P_e = \frac{1}{4}P_s = \frac{3}{4}Q(\frac{d_{min}}{2\sqrt{N_0/2}}) \tag{5.13}$$

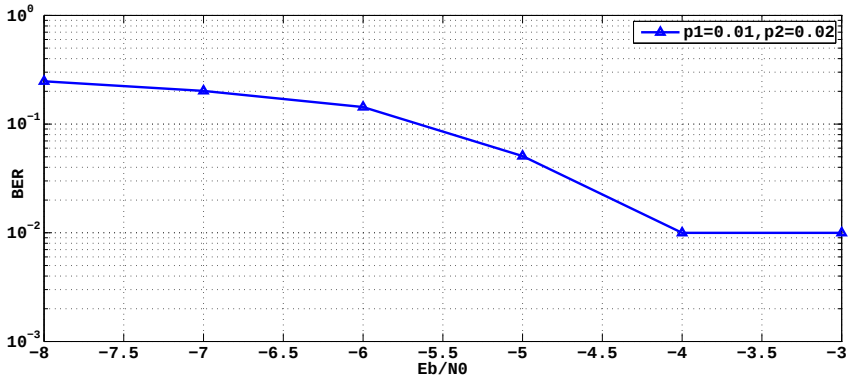

Figure 5.3: Error floor when p_1 and p_2 are fixed

then the PDF of the BER in (5.6) is

$$p_e = \frac{8}{3}\sqrt{N_0\pi}\exp(\frac{(P_e^{-1})^2}{4N_0})\frac{P_e^{-1}}{\sigma^2}\exp\frac{-(P_e^{-1})^2}{2\sigma^2} \tag{5.14}$$

where $P_e^{-1} = 2\sqrt{N_0}e^{-1}(1-\frac{8}{3}x)$ and $e^{-1}()$ is the inverse error function.

Furthermore, the CDF of the BER in (5.9) is

$$\begin{aligned} F(x) &= \int_0^x p_e(u)\,\mathrm{d}u \\ &= \exp(-N_0e^{-1}(1-\frac{8}{3}x)^2) \end{aligned} \tag{5.15}$$

and finally $p_{min}(x)$ in (5.11) is

$$\begin{aligned} p_{min} =& 2(\frac{8}{3}\sqrt{N_0\pi}\exp(\frac{(P_e^{-1})^2}{4N0})\frac{P_e^{-1}}{\sigma^2}\exp\frac{-(P_e^{-1})^2}{2\sigma^2}) \\ & (1-\exp(-N_0e^{-1}(1-\frac{8}{3}x)^2)) \end{aligned} \tag{5.16}$$

note that the integration interval of x in (5.11) is defined by P_e^{-1} in (5.14), which is $[0,\ \frac{3}{8}]$

5.4.1 Error Floor for Fixed BER

The simulation to verify the error floor derivation is conducted like this: random information frame is generated to simulate information A, then random bit-flipping errors with probability p_1 and p_2 are introduced into A to generate information B and C respectively before feed them into corresponding encoders at the relays. At the destination, information D and E, estimation of B and C, are worked out then fed into the combiner to estimate information F. BER of F against A is measured. Figure 5.3 shows the simulation results and the error floor matches the result from equation(5.1).

5.4.2 PDF of the BER at the Relay

For the simulation in this section, random information frame is generated to simulate information A, then A goes through link D1, after the demodulator at the relay, information B and C

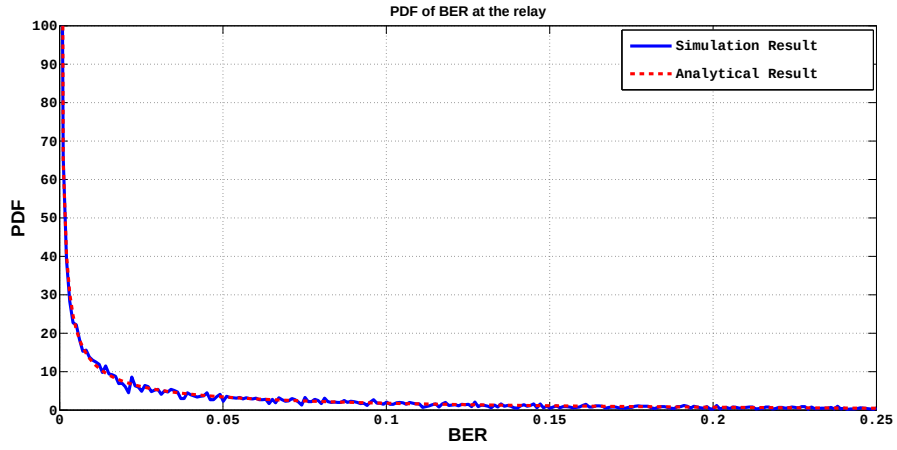

Figure 5.4: PDF of the BER at the relay, Eb/N0 on D1 link is 10dB.

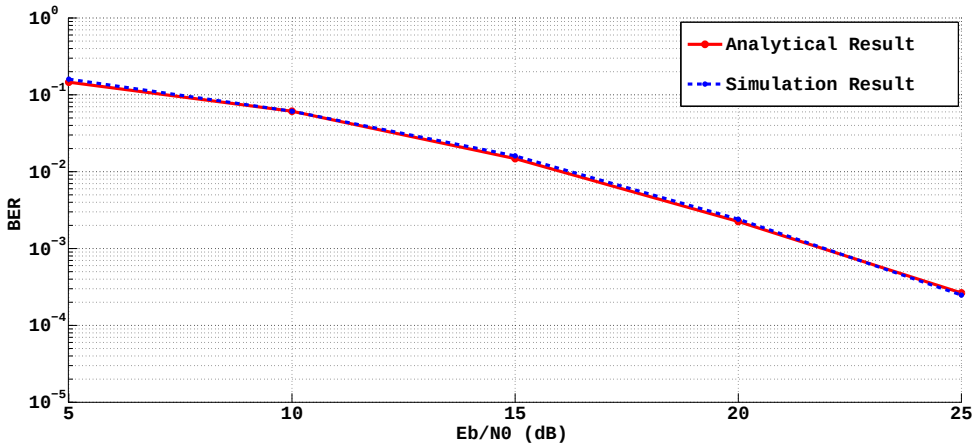

Figure 5.5: Error floor value against Eb/N0 on D1 link

are generated. BER of B (or C) against A is measured and plotted along with the analytical result as shown in Figure 5.4.

5.4.3 Error Floor for Variable BER

Simulation in this section covers the majority of the system structure as shown in Figure 5.2 except for the outer code (parts in blue). So information A is randomly generated as above and we measure the BER of information F against A. In order to fully reveal the error floor, we set a very high value (20dB) to the Eb/N0 on D2 link, so that we can ignore the errors happen on this link. Result is shown in Figure 5.5

5.5 Conclusion

We proposed a system with distributed relays assisted receiver, the relay-destination link can be viewed as a CEO problem. Therefore it has an inevitable error floor. We derive an analytical

expression for this error floor and it is verified by simulation result. This expression will be useful in the future optimization work.

Chapter 6

Cooperative Spatial Multiplexing with Slepian and Wolf Code - adaptive Compression and Further Compression

6.1 Introduction

Based on our design in chapter Chapter 4, we further study some system properties from information theory point of view and implement optimization schemes accordingly to further increase the spectral density on the relay destination link.

In this chapter we refer to mobile stations as relays. In a C-SM system there are several ways to organize relays, some systems[82] only select relays that can fully decode or detect the received message to participate in the next phase transmission, some[83, 84] only allow relays to receive or transmit a selected portion of a signal and others[85] let all relays transmit whatever they receive indiscriminately. In [86, 87], errors at the relays are considered in the system design, however since every relay receives different information the discarded information can only be corrected by an outer code. None of them jointly and efficiently exploit the diversity and spectral efficiency benefits available.

The Slepian and Wolf (SW) [34] theorem proves that it is possible to compress the correlated information at distributed sources independently and jointly recover each of them with arbitrarily small error probability at the destination as long as the total number of bits on all sources is bigger than the joint entropy of all the information and the number of bits at each source is bigger than its entropy conditioned on the other source.

Our contributions in this chapter are

- We make some optimizations to the implementation of the Slepian and Wolf code which increase its efficiency and robustness.
- We develop an algorithm that can identify and forward information at each relay which is useful in decoding the original data, while rejecting information which does not contribute towards this goal, which further increases the system efficiency and allows us to compress the information beyond the limit defined in the Slepian and Wolf theorem with small distortion.
- We study the system properties from the information theory point of view and point out

some common features for MIMO cooperative systems which can be very crucial for performance optimizations, such as relay selection, distributed compression for correlated information.

6.1.1 The CEO problem

A decentralized observation of the state of a system is advantageous in many situations. In decentralized detection, agents only observe partial information about the source, so they do not consume much energy as centralized observation does and all agents directly communicate to a decision fusion center without exchanging information among them, which makes the deployment even cheaper and flexible. The decision fusion center can compute a global decision of the state of the source based on the partial observations from by all the agents.

The CEO problem [88], which is similar to a decentralized detection, estimates the state of a source from an information theoretic point of view. It can be regarded as a multi-terminal source coding problem. However, unlike a decentralized detection, it can produce a sequence of decisions based on a set of successive observations.

The problem can be described as a distributed source coding problem where the goal is to reconstruct a source from coded noisy observations rather than reconstructing the noisy observations themselves. As shown in Figure 6.1, source X^n is observed through a AWGN broadcast channel, where $Z_1^n \sim \mathcal{N}(0,N_0)$ and $Z_2^n \sim \mathcal{N}(0,N_0)$ are independent noise components. The observed sequences $Y_j^n, j = 1,2$, are separately encoded and transmitted to the decoder where an estimation $\hat{X}^n$ of X^n with mean squared error distortion D is produced.

A $(2^{nR_1}, 2^{nR_2}, n)$ code for the CEO problem consists of

1. two encoders, where encoder 1 assigns an index $M_1^n(Y_1^n) \in [1 : 2^{nR_1})$ to each sequence Y_1^n and encoder 2 assigns an index $M_2(Y_2^n) \in [1 : 2^{nR_2})$ to each sequence Y_2^n
2. a decoder that assigns an estimate $\hat{x}^n$ to each index pair $(m_1,m_2) \in [1 : 2^{nR_1}) \times [1 : 2^{nR_2})$

A rate-distortion triple (R_1,R_2,D) is said to be achievable if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with

$$\lim_{n\to\infty} \sup E(\frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{X}_i)^2) \leq D$$

The rate-distortion region $\mathscr{R}_{CEO}(D)$ for the quadratic Gaussian CEO problem is the closure of the set of all rate pair (R_1,R_2) such that (R_1,R_2,D) is achievable.

This problem is closely related to the quadratic Gaussian distributed source coding problem and the rate-distortion function is given by the following [33]

Theorem 1. *The rate-distortion region $\mathscr{R}_{CEO}(D)$ for the quadratic Gaussian CEO problem is the set of rate pairs (R_1,R_2) such that*

$$\begin{aligned} R_1 &\geq r_1 + R\left(\frac{1}{D}\left(\frac{1}{P} + \frac{1-2^{-2r_2}}{N_2}\right)^{-1}\right) \\ R_2 &\geq r_2 + R\left(\frac{1}{D}\left(\frac{1}{P} + \frac{1-2^{-2r_1}}{N_1}\right)^{-1}\right) \\ R_1 + R_2 &\geq r_1 + r_2 + R(\frac{P}{D}) \end{aligned} \tag{6.1}$$

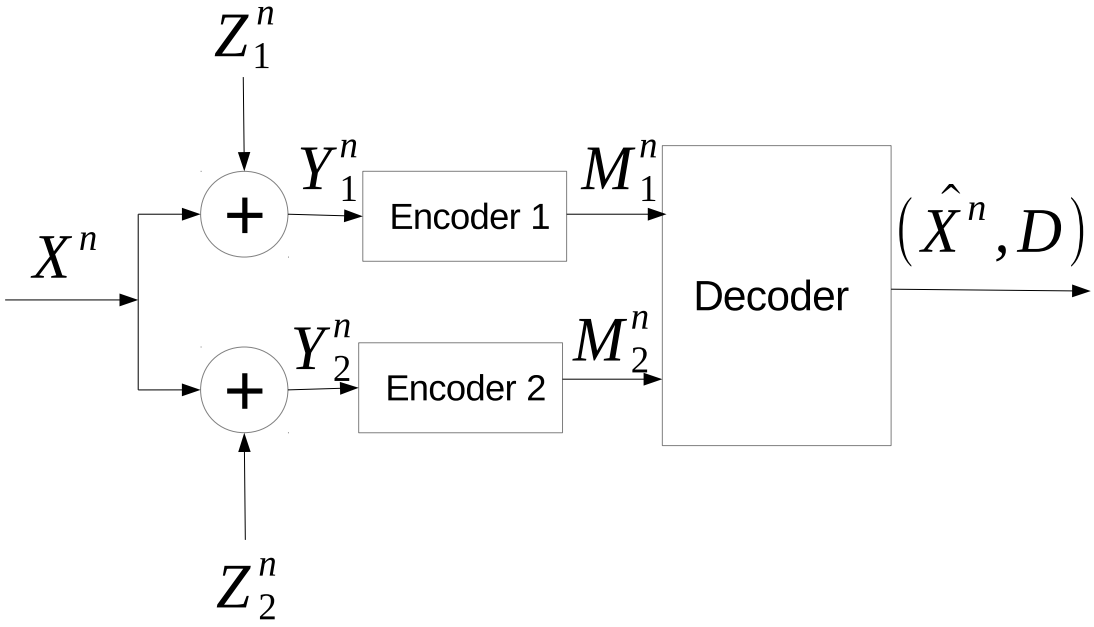

Figure 6.1: Quadratic Gaussian CEO problem

for some $r_1, r_2 \geq 0$ *that satisfy the condition*

$$D \geq \left(\frac{1}{P} + \frac{1-2^{-2r_2}}{N_2} + \frac{1-2^{-2r_1}}{N_1} \right)^{-1} \tag{6.2}$$

where P, $N_i, i = 1,2$, $r_i, i = 1,2$ *are transmitted symbol power, noise power spectral density and transmit rate without distortion.*

Achievability is proved using the Berger-Tung coding scheme with distributed Gaussian test channels as for quadratic Gaussian distributed lossy source coding [89, 88].

6.1.2 Chapter Outline

This chapter is set out as follows. In the next section we state the full system model and point out the difference between the system in this chapter and the ones in the previous chapters. Then in section 6.3 implementation details are given. Simulation parameters, results and analysis are introduced in section 6.4. Conclusion and future work are given in section 6.5.

6.2 System Model and Slepian and Wolf Theorem

The system structure is shown in Figure 6.2. The proposed system has no direct link between the transmitter and the receiver: instead there is a relay group (the receive group) that assists the

receiver. We assume that the relays and the receiver have only one antenna, that the transmitter has two antennas and that there are only two relays in the receive group. For simplicity, we also assume that link between the receive group and the destination is error free (marked in blue). The link between the transmitter and the receive group is a 2×2 MIMO channel suffering from independent block Rayleigh fading (marked in red) and additive white Gaussian noise. Data is encoded by the outer encoder, then the codeword is split into two parts, one for each antenna, and signals are forwarded simultaneously to the relays in the receive group. Relays there first detect the codeword and express it in log-likelihood ratio (LLR) form. The LLRs are then quantized and compressed before being sent simultaneously to the destination. At the destination, the compressed codeword from the receive group is first detected and sent to the decompresser to recover the original codeword, and the combined codeword is then sent to the outer decoder to estimate the original data.

Part of the system, from the receive group to the destination, can be modelled as a "Chief Executive Officer" (CEO) problem [77]. In Chapter 5 an error floor is observed in the overall system performance for this structure, therefore we include an outer code in the system to eliminate it.

In this chapter, we focus our effort on the relay-destination link. The same codeword is transmitted to the relays in the receive group, however because of the fading link and the noise random errors are introduced, so each relay receives a different version of the codeword. This means that the codewords at each relay are correlated, so the SW theorem can be applied.

Two circumstances that are preconditioned in the SW theorem dramatically affect the implementation of our system. First it assumes that before compression each relay must know the overall joint entropy and as a consequence is aware of the targeted compression ratio. Since relays do not have information about each other, it is not possible for them to calculate the joint entropy before the compression. What they can do, however, is to estimate the mutual information value between the received codeword and the original one (details of which are given in section 6.3), and forward this value to the destination before compression, and receive the feedback about the mutual information values of other relays from the destination. Based on the feedback and its own value, relays can estimate the joint entropy which determines the compression ratio. The method of estimation is given in section 6.3. Compared to the system in Chapter 4 which does not have a feedback channel, system proposed in this chapter can adapt its compression ratio to the joint entropy of the information on relays, thus it can use the channel on the relay destination link more efficiently.

Second it assumes that codewords on all relays are useful and should be recovered at the destination losslessly. We draw a Venn diagram Figure 6.3 to clarify this concept. Circles R_1, R_2 and S represent the codeword at the relays and the source. Each capital letter represents the surrounding enclosed area. The SW theorem states that we can only losslessly compress $R_1 + R_2$ to $A + B + D + E + F + G$, which is the joint entropy of R_1 and R_2. However, from the system point of view, not all this joint entropy represents useful information. Only $E + F + G$ which represents information about the source is relevant (calculation of $E + F + G$ is in Appendix B): the rest of the joint entropy only contains information about the noise and the fading channel. Hence in theory, it is possible to do lossy compression on R_1, R_2 (not being able to recover them completely at the destination) but still maintain the system performance. Take R_2 for example, further compression is possible by identifying the information corresponding to $D + B$, which can be done by examining the amplitude of the codeword in LLR form. Bits which have low amplitude LLR (unreliable bits) are more likely be associated with $D + B$,

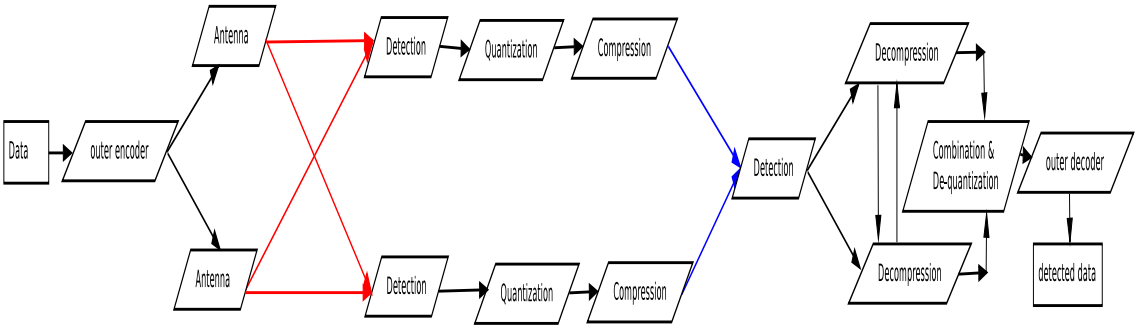

Figure 6.2: System structure

whereas those which have high amplitude (reliable bits) probably belong to $F+G$. However, according to our study, unreliable bits tend to occur evenly across the whole codeword. Hence identifying these bits typically require more data than is saved by omitting to transmit them. A simulation result is given in section 6.4. However one suboptimal but practical way to perform lossy further compression is to detect the unreliable codeword instead of unreliable bits. In such words, $F+G$ is fairly small, and hence the entire codeword may be dropped. Details are given in section 6.3.

6.3 Implementation

In section 6.2, we mention that based on its own mutual information and that at other relays, each relay in the receive group should have the ability to estimate the joint entropy. In this system we use a lookup table to achieve this, and assume that this table is available to all the relays in the receive group. The lookup table contains the relationship between the mutual information on each relay and the joint entropy. Then once every relay in the receive group knows the others' mutual information at all other relays it can estimate joint entropy using the table. The lookup table is generated using the information recorded in simulation, and assuming the simulation runs are reasonably long, the lookup table can be very accurate. A typical lookup table is in Figure 6.4.

For calculating the mutual information at the relay, we use the algorithm in [31]

$$\begin{aligned} I(x,y) =& H(x) - H(x|y) \\ =& 1 - \sum_{x} \int p(x,y) \log_2(1/p(x|y))\, dy \\ =& 1 - \underset{x,y}{E}[\log_2(1/p(x|y))] \end{aligned}$$

where x, y represent the original codeword and the received codeword in the LLR form. $I(x,y)$, $H()$, $E()$, $p(x,y)$ and $p(x|y)$ indicate mutual information, entropy, expectation, joint probability density function and conditional probability density function respectively.

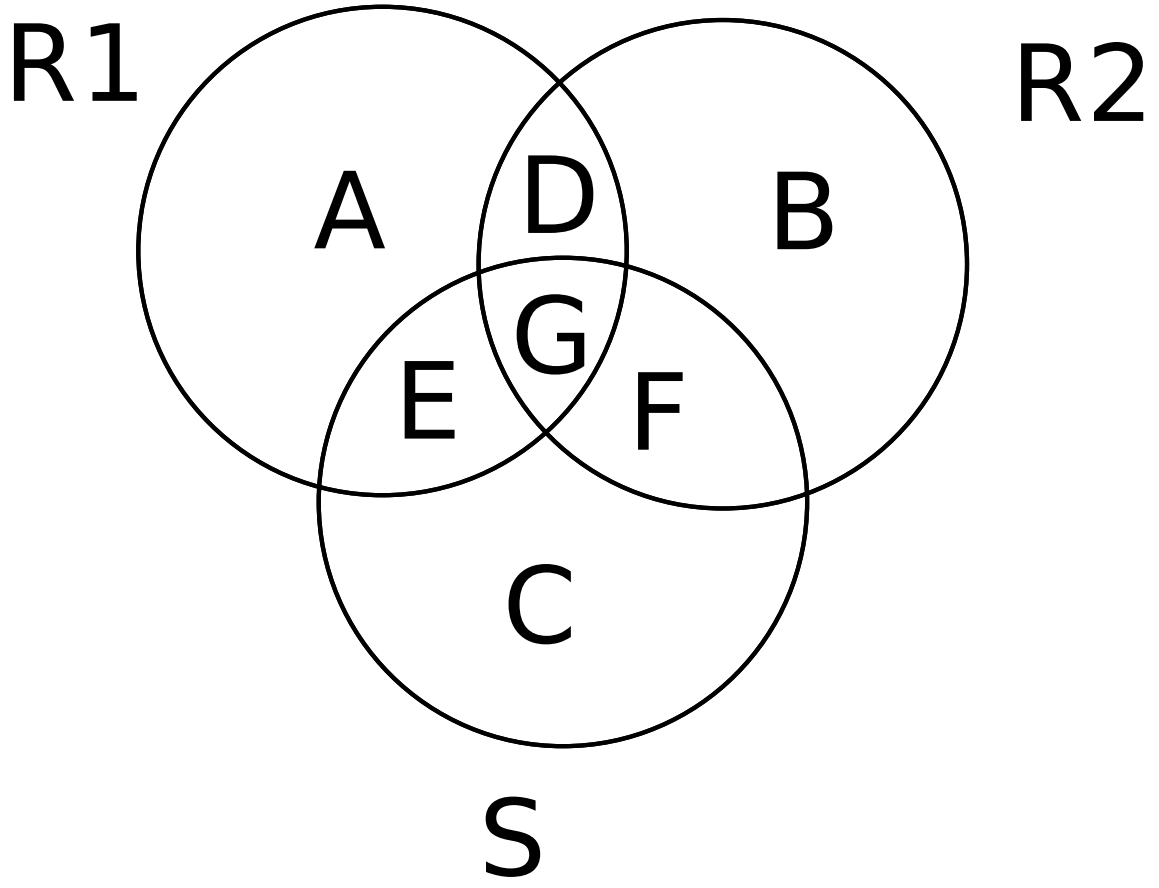

Figure 6.3: Relationship between information at relays and the source

By using Bayes'law [14], the equation can be rewritten as

$$\begin{aligned} I(x,y) =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(\frac{p(y)}{p(x)p(y|x)})] \\ =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(\frac{\sum_x p(x)p(y|x)}{p(x)p(y|x)})] \\ =& 1 - \mathop{\mathrm{E}}_{x,y}[\log_2(1+\frac{p(y|-x)}{p(y|x)})] \end{aligned} \tag{6.3}$$

where $p(y|-x)$ and $p(y|x)$ can be obtained empirically from the detected codeword. In practice, because the original codeword x is not available at the relay, $\frac{p(y|-x)}{p(y|x)}$ can be approximated as $\frac{p(y|-1)}{p(y|1)}p(x=1)+\frac{p(y|1)}{p(y|-1)}p(x=-1)$, where $p(x=1)$ and $p(x=-1)$ are the probabilities of one particular bit in the codeword obtained from its LLR. The derivation from the second line to the third line in equation (6.3) is based on the assumption that for the original codeword x, $p(x=1)=p(x=-1)=1/2$.

Assuming the codeword is long enough, the mutual information on each relay in the receive group is determined by the fading channel on the source relay link (red links in Figure 6.2). As

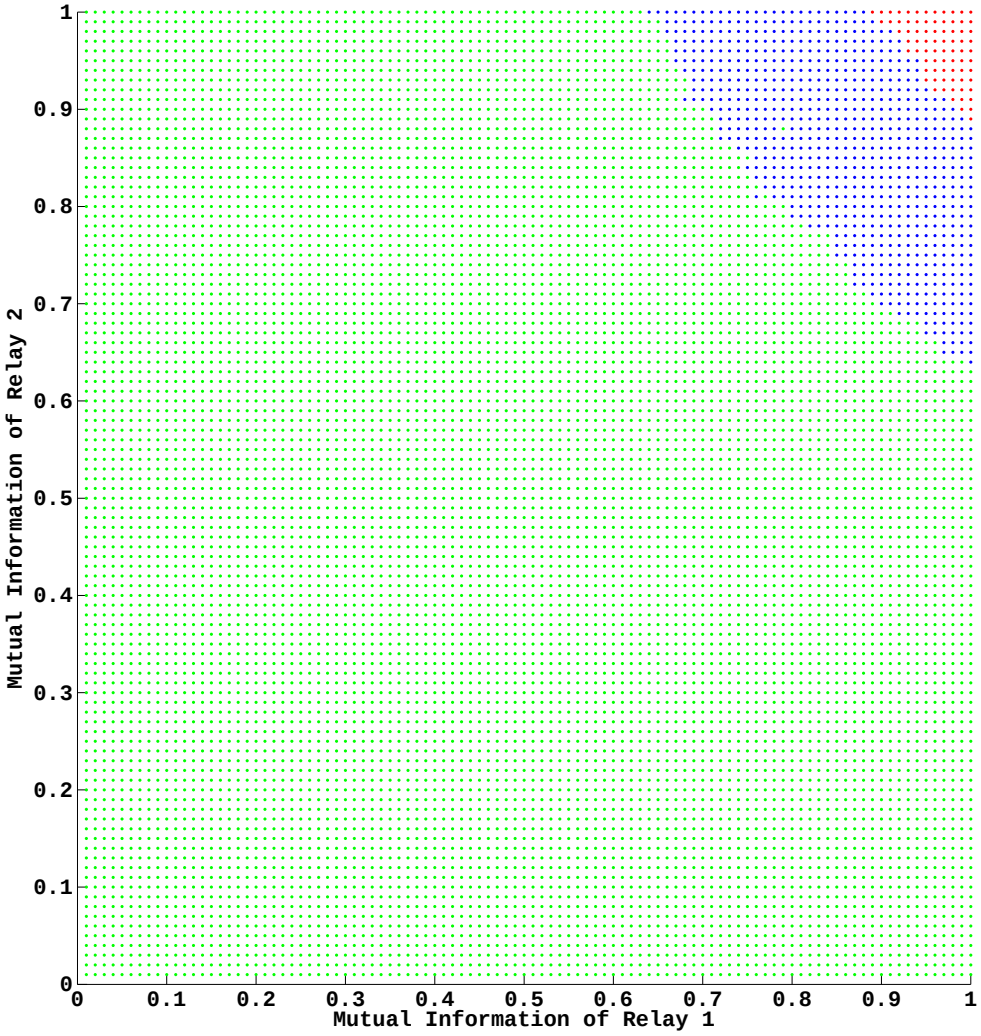

Figure 6.4: Optimum rate lookup table, green dots mean compression rate $R >= 1.7$, blue means $R >= 1.4$ and $R < 1.7$, red means $R < 1.4$

a consequence, relays in the receive group only have to estimate and exchange their mutual information when the fading channel on the source relay link is changed (assume that the fading channel changes slowly at keeps stable at least for one frame length). And furthermore, we can use another lookup table to replace the calculation of the joint entropy at the relay. This lookup table contains the relationship between the fading channel and the mutual information at the relay. Assuming that the transmitter uses BPSK, then the received constellation on each relay in the receive group is a parallelogram which can be defined by three parameters, namely the lengths of the two neighbouring sides and the angle between them as shown in Figure 6.5. We assume that relays in the receive group have the channel state information (CSI) hence know these three parameters. So every time when the fading channel is changed, relays in the receive group will input the three new corresponding parameters to the table and obtain

EbN0	-5	-2.5	0	2.5	5
average joint entropy difference	0.0033	0.0039	0.0042	0.0038	0.0032

Table 6.1: Average joint entropy differences between random channel sets and their equivalent sets

R_2 \ R_1	$< T$	$>= T$
$< T$	$R_1 = d$,$R_2 = d$	$R_1 = nc$,$R_2 = d$
$>= T$	$R_1 = d$,$R_2 = nc$	$R_1 = c$,$R_2 = c$

Table 6.2: Frame dropping rule, where T, d, nc and c represent threshold, drop, no compression and compression according to the SW theorem respectively.

the estimated mutual information. This lookup table is also generated using the information recorded in simulation, however this table is not included in the simulation presented in this chapter, because the calculation load is beyond the processing limitations of our computer. We however conduct a simulation to compare the estimated mutual information to the mutual information calculated from equation (6.3), which shows that the difference is infinitesimal as shown in Table 6.1.

Assume channels from transmitter to one relay is $\begin{pmatrix} a+bj & c+dj \end{pmatrix}$, distance between sig1 and sig2 is d_1, distance between sig1 and sig3 is d_2, distance between sig2 and sig3 is d_3 as shown in Figure 6.5. The relationship between channel parameters and the received signals is:

$$\begin{aligned} a^2+b^2 &= \frac{d_1^2}{4} \\ c^2+d^2 &= \frac{d_2^2}{4} \\ (a-c)^2+(b-d)^2 &= \frac{d_3^2}{4} \end{aligned} \tag{6.4}$$

There are not sufficient equations to solve the problem, so one additional equation $a = b$ is added deliberately to help solve equation (6.4).

The simulation that prove the feasibility of the channel mutual information table is conducted as follow: first randomly generate a set of channels $\begin{pmatrix} a+bj & c+dj \end{pmatrix}$, then using equation (6.4) to calculate its equivalent set of channels $\begin{pmatrix} a'+b'j & c'+d'j \end{pmatrix}$, using both sets of channels to run simulation and compare the corresponding mutual information.

For further compression, as stated in section 6.2, the relay detects the unreliable codeword and does not transmit it at all. We use the mutual information of the codewords to identify which are unreliable one: a mutual information threshold is set and all relays in the receive group follow the rule in table 6.2. When the system is operating under the frame dropping scheme, the SW scheme is not activated. If codewords on all relays are below the threshold, nothing is transmitted.

The implementations for other parts of the system in Figure 6.2 are similar to those in our previous paper [90]. The outer code is a concatenated code with two convolutional codes as its component codes. The outer decoder uses the BCJR algorithm [57]. The detector uses a soft de-mapper [27] and the outer decoder recursively to estimate the LLRs for the received

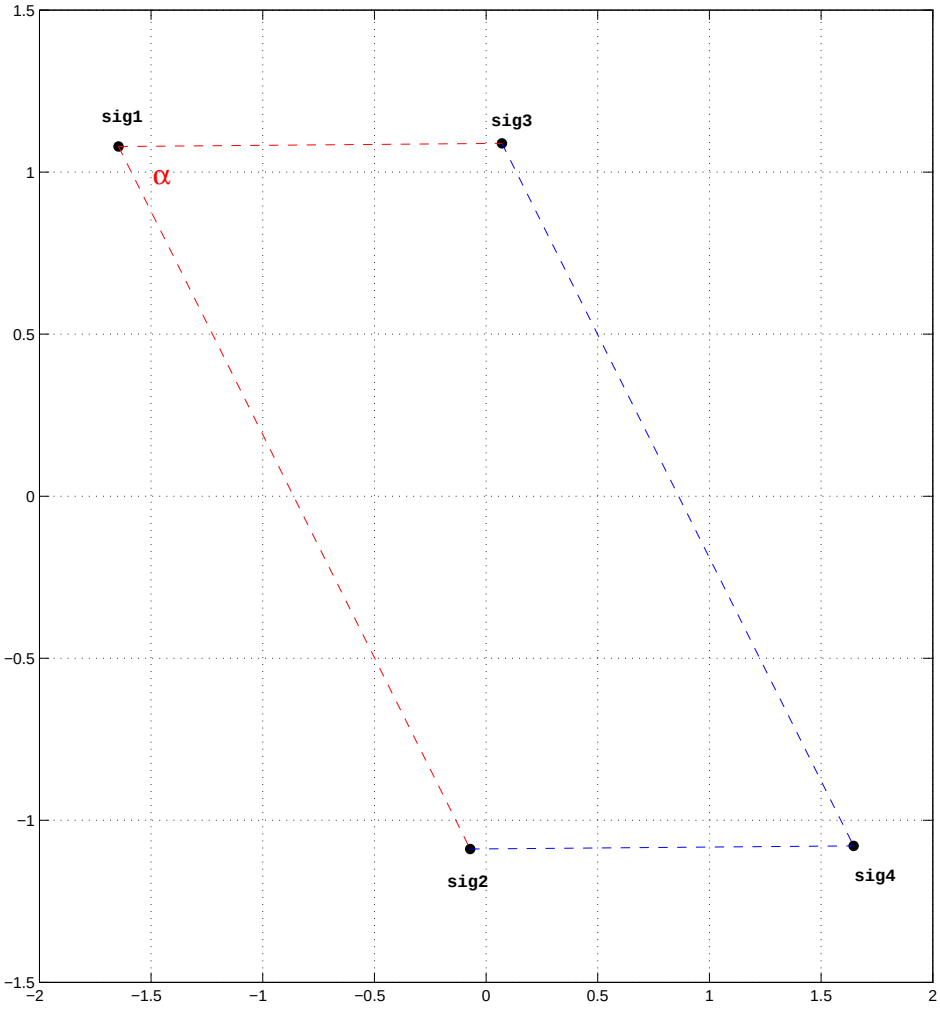

Figure 6.5: Three parameters used to define a channel, distance between sig1 and sig3, distance between sig1 and sig2 and angle α

codeword. The quantizer uses the LBG algorithm [76] to quantize the LLRs as shown in Figure 6.6. For each codeword at the relay in the receive group, because parts of it go through different independent fading channels, the distribution of the amplitude of the LLRs in one codeword is not consistent, therefore every part of the codeword needs to be quantized separately. The combiner sums the corresponding LLRs from each codeword.

Optimization for the Slepian and Wolf code implementation

The compressor uses the Accumulate-Repeat-Accumulate (ARA) code [73] with a slight difference from that in [90]. In the previous chapter, the compressor at each relay processes the codeword in the same manner, whereas in this chapter, the compressor on each relay first di-

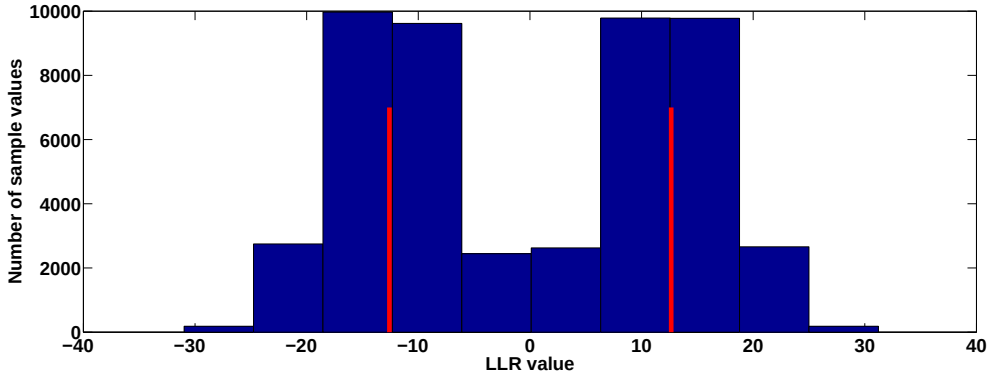

Figure 6.6: Vector quantization result for one dimension random variable. Blue bars are histogram of LLR samples from the relays and red lines are quantization points.

vides the codeword into two equal parts, of which one compresses the first part while the other compresses the second part. These setting give us two advantages. When the de-compressor at the destination tries to recover the codeword for each relay it uses the information of one codeword as the extrinsic information to decode the other codeword, and since each relay compresses a different part of the codeword, the extrinsic information from the other codeword is uncompressed which makes it more reliable for decoding, and hence the overall system performance is more reliable and better than that in [90]. And although the compressor on each relay treats the codeword differently, the frame lengths of the compressed codewords at all relays are still the same, which is preferable when we use the multiple access channel (MAC) on the relay destination link in the future.

6.4 Simulation Result and Analysis

The parameters for the simulation are set as follows: the frame length for the codeword is 5×10^4, BPSK modulation is used on the fading link which is changed for every codeword, the rate of the outer code is $\frac{1}{2}$, threshold for frame dropping scheme is 0.8. The quantizer uses only two quantization points.

Figure 6.8 and Figure 6.9 show the overall BER performance and the rate on the relay destination link. If the system is not compressed, each relay in the receive group has a whole codeword to transmit, because there are two relays in the group which makes the total rate 2. For the case where the system uses only Slepian and Wolf compression, Figure 6.8 shows that there is virtually no loss for the compressed system (blue broken line) compared with the uncompressed system (red line), and Figure 6.9 shows that the actual rate for the compressed system (blue line) saves more bandwidth as the SNR increases, however the lower bound for SW compression which is indicated as the red line is not very close to the blue line. This is caused by the performance limitation of the ARA code which is out of the scope of this chapter.

In section 6.2, we suggested that further compression can only be implemented at codeword level rather than on the bit level and ran a simulation to demonstrate this. A LLR amplitude threshold is set (in our case $abs(LLR) < 1$), and if the amplitude of a bit is below the threshold, this bit will be marked as unreliable. After identifying unreliable bits, the joint entropy is

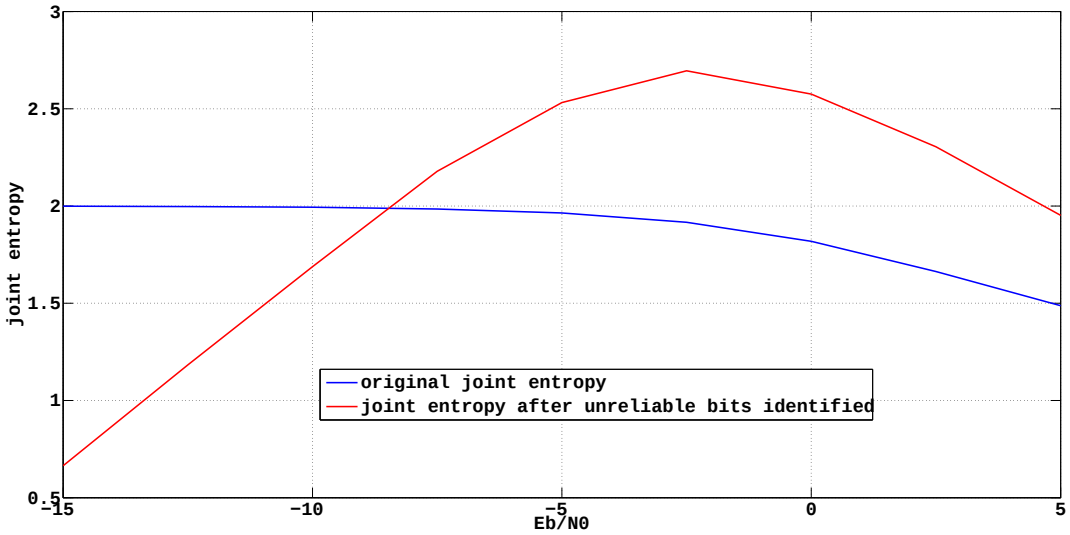

Figure 6.7: Joint entropy between detected messages on relays after unreliable bits identified at low Eb/N0

measured to examine the feasibility of further lossless compression, and is shown as the dark green line in Figure 6.9. As the plot indicates, identifying unreliable bits actually increases the average joint entropy. According to our study, two reasons contribute to this increase, first the identification increases the cardinality of a code symbols, since they become ternary rather than binary. Second we discover that unreliable bits tend to spread evenly across an entire codeword, and hence their location can not efficiently be encoded using run length coding. We also measure joint entropy at even lower SNR in Figure 6.7 and discover that only when SNR is smaller than -10dB will the joint entropy after the unreliable bit identification be smaller than the original joint entropy. The result of this simulation shows that we can only perform further (lossy) compression at the codeword rather than the bit level.

The light green line in Figure 6.8 and Figure 6.9 shows the performance of the frame dropping scheme. Because it is a lossy compression scheme, in Figure 6.8, a minor loss is expected. However in Figure 6.9, it shows that this scheme allows further compression. Note that, especially in low SNR region, the system can compress below the joint entropy without significantly increasing end-to-end BER. The difference between the frame dropping rate and the SW compression rate (blue line) decreases as the SNR increases, this is because the number of unreliable codewords decreases as the SNR increases, so fewer frames are dropped.

The purple line in Figure 6.9 indicates the lower bound for lossless compression generally. It represents the area $E+F+G$ in Figure 6.3. Note that none of the current compression schemes, namely the SW scheme and the frame dropping scheme, can remove the redundant areas of this diagram completely and exclusively. The SW scheme can only remove $D+G$ whereas frame dropping scheme drops R_1 or R_2 or both. The undesirable areas A and B are not removed efficiently. That is the reason this lower bound is far away from the actual performance. The method to calculate the purple line is given in the Appendix B.

We also measured the impact of different LLR schemes on the system performance from the information theory point of view. As shown in Figure 6.10, the blue line (red line) represents the mutual information between the source and the overall received signal (overall received

LLRs) of all relays in the receive group. In these two scenarios all relays share information with each other, so the mutual information performance is optimum.

For one particular channel, the conditional probability density function for the overall received signal is

$$p(y|x) = \frac{1}{2\pi\sigma^2}\exp\left[\frac{-1}{2\sigma^2}||y-Hx||^2\right] \tag{6.5}$$

where σ indicates noise standard deviation on one dimension. For the overall received signal y and x become 2×1 arrays representing the overall received signal on both relays and their corresponding transmit signals on both transmit antennas, and H is a 2×2 matrix representing the fading channel.

The overall LLR is

$$LLR = \log\Big(\frac{\sum\limits_{s_i\in S_0}\exp\big(-\frac{||y-s_i||^2}{N0}\big)}{\sum\limits_{s_i\in S_1}\exp\big(-\frac{||y-s_i||^2}{N0}\big)}\Big) \tag{6.6}$$

where S_0 and S_1 are the subsets of the received constellation points on both relays that have 0 or 1 on a certain position. The conditional probability density function of the overall LLR can be obtained empirically. Putting equation [6.5] and the empirical conditional probability density function of equation [6.6] into equation [6.3] and averaging over different channel samples we obtain the blue and the red lines in Figure 6.10, which almost completely overlap with each other.

As indicated in the plot, representing information in signal form or in LLR form is equivalent, in that the mutual information is the same for both form. However, when information is extracted by the demapper and the outer decoder from each relay separately and summed together by the combiner at the destination, some information is lost, as is indicated by the green line in the plot. Furthermore quantization also causes information loss. Although introducing more quantization points can reduce the information loss, there is a trade off between the fidelity of the quantizer and the spectral efficiency of the relay-destination link. Another interesting fact is that the LLR combination scheme used in our previous chapter [77] which requires relays to estimate the error probability has almost the same performance as two points quantization LLR summation scheme does. Whereas the LLR summation scheme does not need to estimate the probability of error and can outperform the LLR combination scheme by increasing the quantization points.

6.5 Conclusion and Future Work

In this chapter we have implemented Slepian and Wolf compression in a typical cooperative spatial multiplexing system.

The conclusions of this chapter are

- We optimize the implementation of the Slepian and Wolf code which leads to more reliable and efficient performance than it can achieve in the previous chapters, details is given section 6.3.

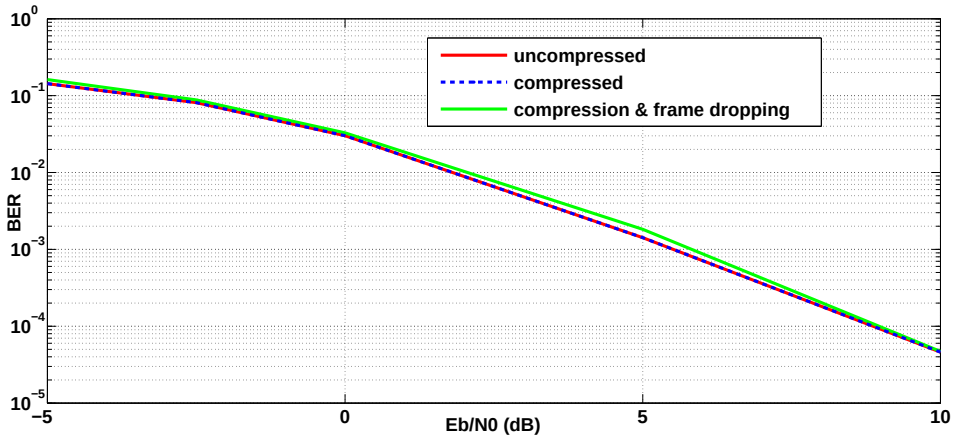

Figure 6.8: BER of different compression schemes and uncompressed scheme

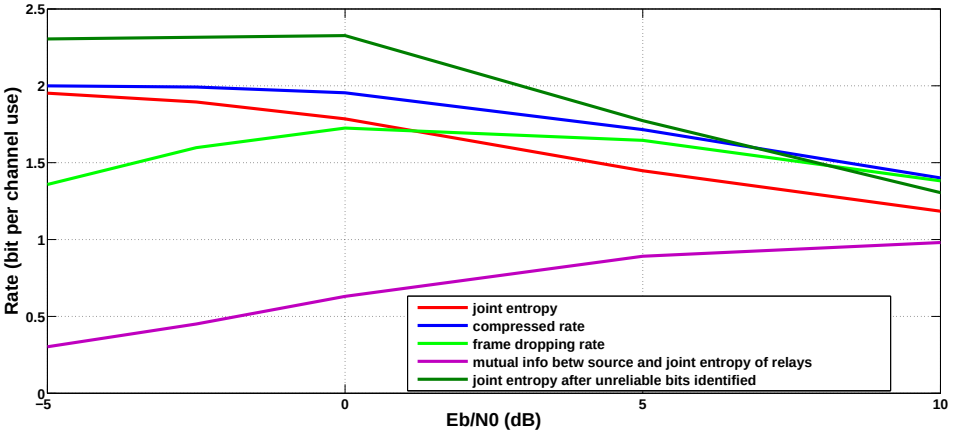

Figure 6.9: On the relay destination link, achievable rate of different compression schemes, limit defined in the Slepian and Wolf theorem (the joint entropy) and theoretical minimum achievable rate (mutual information between source and joint entropy of relays)

- We also study the system from the information theoretic point of view, based on which we propose a new compression scheme which can further increase the spectral efficiency beyond the limit defined in the Slepian and Wolf theorem (joint entropy of the information on the relays) with minor BER loss as shown in in Figure 6.8 and Figure 6.9.

Several means of further optimization may be possible. First the performance of the ARA code which implements the SW theorem needs further improvement. Second, the MAC channel can be applied on the relay destination link which can double the spectral efficiency. Last, a more efficient lossy compression scheme maybe developed to increase compression without increasing BER.

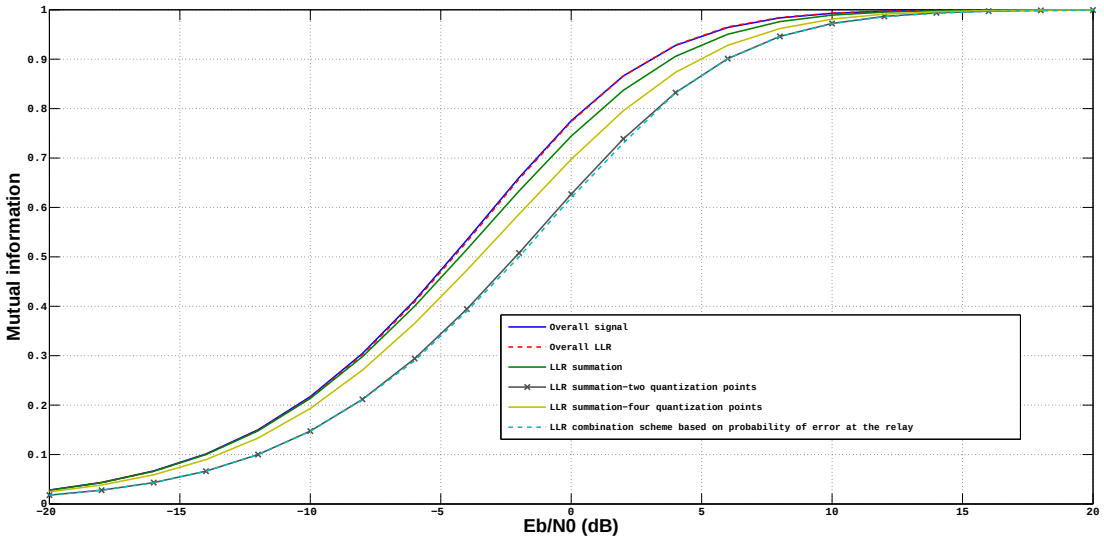

Figure 6.10: Mutual information of different LLR schemes

Chapter 7

Cooperative Spatial Multiplexing with Slepian and Wolf Code and Multiple Access Channel

7.1 Introduction

In Chapter 6 we studied the implementation of the Slepian and Wolf theorem in a cooperative spatial multiplexing system, assuming that the channels between the relays and the destination are error free. However a more practical assumption would be that the link is a multiple access channel suffering from block Rayleigh fading and white Gaussian noise.

In this chapter, we modify the system in the previous chapter to enable it to work with a fading multiple access relay-destination link. System performance is investigated and simulation results are analysed from a theoretical point of view using network information theory. Two system designs are proposed, one follows source channel code separation theorem [32], while the other is guided by the joint source channel code theorem [91]. From simulation and analysis result we conclude that for the systems proposed in this chapter the outage probability of the separate design is better than that of the joint design, which coincides with the results in [92].

Similar work has been conducted extensively in wireless sensor network (WSN) scenarios, as in [93, 94] where systems transmitting correlated sources over a MAC channel are discussed and in [95] where system transmitting correlated sources over independent fading channels is presented. In WSN certain number of sensors are deployed in the area of interest with a common data sink deployed nearby. Each sensor observes the source from a different point of view, however because each point of view is not totally independent, the data produced by all sensors are correlated. The degree of correlation increases as the density of the sensors increase. The correlation among data can lead to considerable redundancy in the sensor network, if not carefully exploited, it can cause unnecessary energy consumption or performance loss. On the other hand, the energy provided for a wireless sensor may not be renewed, also the computational capacity of individual sensor is very limited.

Shannon proved that a reliable point-to-point transmission is possible only if the entropy of the source is less than the capacity of the memoryless channel [32]. Hence, a simple method to determine whether reliable communication is feasible or not is a simple comparison of rates of the source code with the rates of the channel code. However, for communication over networks,

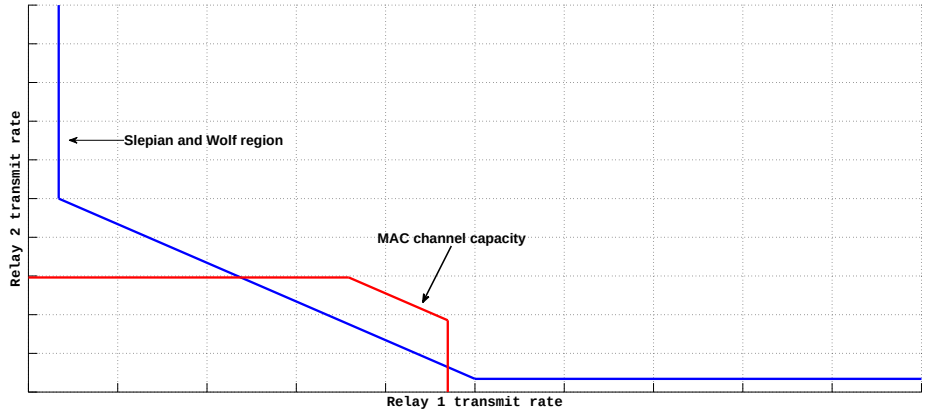

Figure 7.1: Slepian and Wolf compression rate region and MAC capacity region sample for reliable communication

[33] (chapter 14) concludes that source channel separation does not always hold in general for communicating correlated sources over multiuser channels.

This chapter is organized as the following: section **??** introduces the rate regions associated with our system design, namely the multiple access rate region and the Slepian and Wolf rate region, and the relationship between them. The system structure is explained in section 7.3. The implementation that enables the system to work with a fading MAC channel on the relay destination link is shown in section 7.4. In section 7.5 methods to determine reliable communication are discussed. Simulation results are shown and discussed in section 7.6. We concludes our contribution in section 7.7.

7.2 Rate Region for Reliable Transmission

[96] proves the rate region for reliable transmission of correlated sources over a MAC channel.

For transmitting correlated sources over a multiple access channel, if and only if the Slepian and Wolf data compression region and multiple access channel capacity region have an intersection as shown in Figure 7.1, reliable communication is feasible.

7.2.1 Tradeoff in Transmitting Correlated Sources over MAC

For any pair of rates (R_1, R_2) in the intersection, they indicate the compression rate for the Slepian and Wolf code as well as the code rate for the channel code. Take relay 1 for example, assume the frame length in symbols at relay 1 is L, after Slepian and Wolf compression, the frame length is $L \times R_1$. Because the code rate for the channel code is also R_1, the frame length after the channel encoding is $\frac{L \times R_1}{R_1} = L$.

The challenge for transmitting a pair of correlated sources over a MAC is the tradeoff between compression code and channel code. The compression code reduces the redundancy at each sources, the nature of this redundancy at the relays is due to the correlation. The channel code on the other hand, introduces redundancy into the compressed information and the purpose

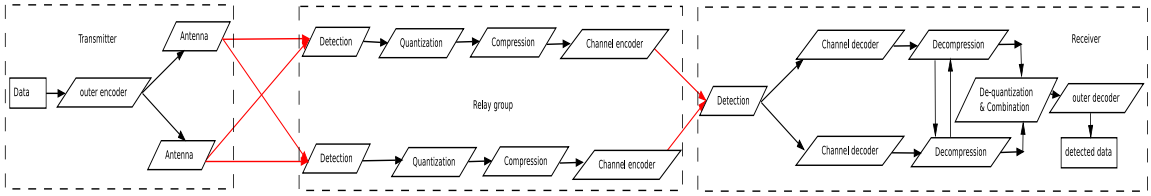

Figure 7.2: System structure for separate source channel code design

is to combat the fading on the channel and noise at the receiver. The amount of redundancy (number of bits in a frame) α_1, which is bounded by the Slepian and Wolf region(blue line in Figure 7.1), removed by the compression code has to be the same as the amount α_2, which is bounded by the MAC capacity region(red line in Figure 7.1), introduced by the channel code. If α_1 is too large, although α_2 is large enough to combat the fading channel, decoding process of the compression code can not recover the original data, whereas if α_1 is too small, there is not enough redundancy to combat the fading channel, hence errors happen at the output of the channel decoder, decoding process of the compression code still can not recover the original data.

7.3 System Model

The system structure for the separate design is shown in Figure 7.2. The proposed system has no direct link between the transmitter and the receiver, instead there is a relay group (the receive group) that assists the receiver. We assume that the relays and the receiver have only one antenna, the transmitter has two antennas and there are only two relays in the receive group. The relay-destination link is a multiple access channel suffering from block Rayleigh fading and white Gaussian noise. The link between the transmitter and the receive group is a 2×2 MIMO channel suffering from independent block Rayleigh fading and additive white Gaussian noise. In the separate design, Data is encoded by the outer encoder, then the codeword is split into two parts, one for each antenna, and BPSK signals are forwarded simultaneously to the relays in the receive group. The relays first detect the codewords and express them in log-likelihood ratio (LLR) form. The LLRs are then quantized and compressed by the Slepian and Wolf compressor before being sent to the channel encoder where redundancy is added into the compressed codewords to combat the fading on the relay-destination link. Codewords from all relays are sent simultaneously to the destination via a multiple access channel. At the destination, soft demapper and channel code decoders first try to estimate the compressed codewords, then the detected compressed codewords are sent to the decompressors to recover the original codewords at relays, and the combined codeword is then sent to the outer decoder to estimate the original data.

We assume channel state information of the MAC and the joint entropy of the information on the two relays are available at the destination, so the destination can draw a rate-region figure like Figure 7.1 and decide which pair of rates (R_1, R_2) should be used for the current frame. Rate information is sent to each relay via the feedback channel from the destination. In theory, any pair of rates in the intersection area are feasible, however because of the performance limitation of the Slepian and Wolf code and the channel code, the pair of rates that allows reliable communication needs to be chosen carefully (details are given later in this chapter).

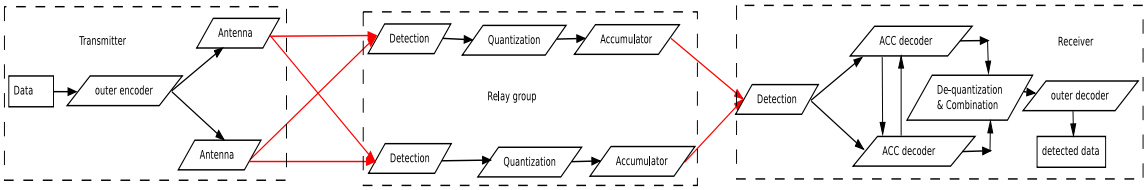

Figure 7.3: System structure for source channel code joint design

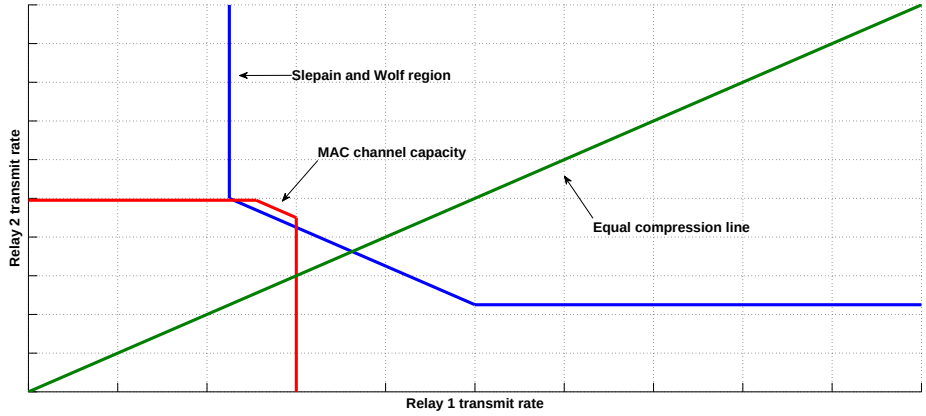

Figure 7.4: Rate region that does not allow equal compression

The system structure for the joint design is shown in Figure 7.3. The only difference between the joint design and the separate design is we use an accumulator to replace the Slepian and Wolf code and the channel code. At the destination, two accumulator decoders also exchange correlation information between relays (Chapter 4, Equation 4.3 and Equation 4.4) to help with the estimation. In the separate design, Slepian and Wolf code removes the correlation redundancy in the relays while channel code adds redundancy to the compressed codewords to combat the fading. Whereas, in the joint design, accumulator does not remove nor add any redundancy at all. Since the codewords are not compressed, correlation redundancy can be used to combat fading instead of decompression.

7.4 Implementation

For separate design, the accumulate repeat accumulate (ARA) code is used to implement the Slepian and Wolf theorem. It is worth noticing that equal compression is not possible in this system when the diagonal line of the whole rate region does not go through the intersection area as shown in Figure 7.4. Therefore, when equal compression (described in chapter 6) is not feasible, the ARA code can only do unequal compression.

The turbo code is used as the channel code in the separate design. Equal space puncturing is applied at the output of the turbo code to adjust the code rate.

For joint design, the ARA code and Turbo code are replaced by an accumulator, which is a memory one recursive convolutional code as shown in Figure 7.5. At the destination, two accumulators also exchange correlation information follow the rule in equation (4.3).

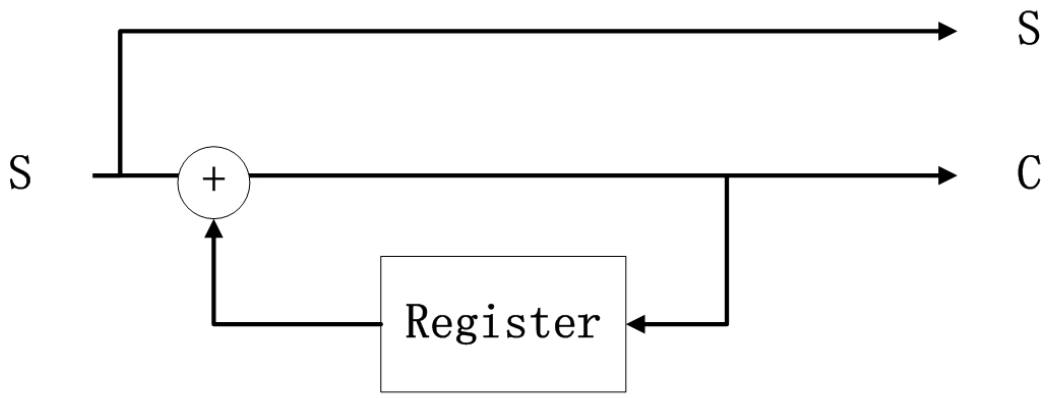

Figure 7.5: Accumulator

7.5 Performance Evaluation

As mentioned above, in theory any pair of rates in the intersection area should guarantee reliable communication however because of the performance limitation of the Slepian and Wolf code and the channel code, not all pairs of rates are feasible.

Figure 7.6 shows the EXIT chart of the detector (see equation 4.5) and the component turbo decoders for the separate design. The EXIT chart is drawn according to Figure 7.7, where the two component turbo decoders are treated as one decoder (blue box) and the detector is treated separately (red box).

Figure 7.6 shows the EXIT curves for three different code rates, namely $[\frac{1}{3}\ \frac{1}{2}\ 1]$. For each code rates, the curve underneath represents the joint turbo codes and the curve above represents the detector. The areas between the two curves for each code rates are around $[0.1\ 0.12\ 0.19]$ respectively and are considered as an approximation of the performance loss [46] (average distance between the MAC channel capacity and the pair of achievable code rates). Since in our system the minimum and the maximum code rates for the turbo code are $\frac{1}{2}$ and 1, the performance loss between 0.12 and 0.19 is expected.

Figure 7.8 shows the achievable code rates for a MAC channel sample in the simulations.

Figure 7.9 shows the EXIT curves of the joint accumulator decoder for the joint design. They are drawn according to Figure 7.11 where the detector and the joint decoder are treated separately.

Figure 7.9 shows the EXIT curves of the joint accumulator decoder with various degrees of correlation between the relays, namely the probability of difference $p = [0.2\ 0.1\ 0.05\ 0.01\ 0]$. According to the LLR updating function 4.3, stronger correlation (smaller p) leads to a larger LLR amplitude (larger mutual information) which is shown in Figure 7.10. Because the joint accumulator decoder exchanges information between its components decoders via the LLR updating function, the performance of the joint decoder deteriorates as the correlation decreases. In Figure 7.9 this feature is shown as the EXIT curves terminate further away from the $(1,1)$ point as the correlation decreases (bigger p). Under this circumstance, it is hard to find an open tunnel between the joint decoder EXIT curve and the detector EXIT curve (however one pos-

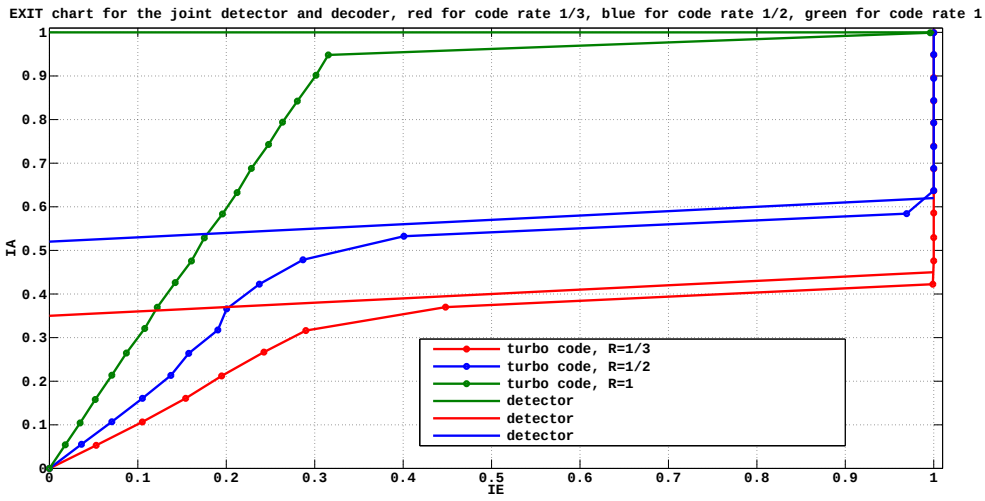

Figure 7.6: EXIT chart of the detector and the component turbo decoders for the separate design, code rates of $[\frac{1}{3}\ \frac{1}{2}\ 1]$ are marked in red, blue and green respectively.

sible way is that the detector EXIT curve is almost horizontal and very close to the top of the EXIT chart which is shown in Figure 7.12). Noting that according to the LLR updating function, when $p = 0$ the updating function does not change LLR at all, hence the joint accumulator decoder is equivalent to a turbo decoder with the accumulator as its component decoders (its EXIT curve is the rightmost one in Figure 7.9, which reaches the (1,1) point). Its EXIT chart is shown in Figure 7.13 with an area between two curves equals to 0.19 approximately.

As stated above, the joint design works only when there is a full correlation or the MAC channel is perfect. Otherwise, it will suffer inevitable errors. Hence according to our discovery the correlation redundancy used by the joint design is not as good as the redundancy generated by the channel code in the separate design when dealing with the fading channel and the noise.

Figure 7.14 shows the performance limitation of the ARA code. There are two compression schemes for the ARA code, one compresses the frame at each relay independently as shown in the green line. The advantage of this scheme is that relays can compress their frames with different rates.The other scheme first splits the frame at each relay into two equal parts. Then different relays compress different parts of their frames. The advantage of this scheme is that at the destination, when the decompressor tries to exchange the information between frames from different relays, every relay has a reliable (uncompressed) part to share. The performance of this scheme is shown as the red line. The differences between the achievable compression rates and the joint entropies are the performance limitation of the ARA code, which are also the minimum distances between the rates and the boundary of the Slepian and Wolf region that allow the compression code to work.

Figure 7.15 shows the EXIT chart for the ARA code with $p = 0.15$ (joint entropy is 1.6097). It achieves compression rate of 1.7757. The area between the two curves is about 0.09 which is close to the average performance loss of 0.166/2=0.084. The EXIT chart is drawn according to Figure 7.16 where the accumulator decoders and the check node decoders (blue box) are treated as one decoder and two variable node decoders are treated as another decoder (red box).

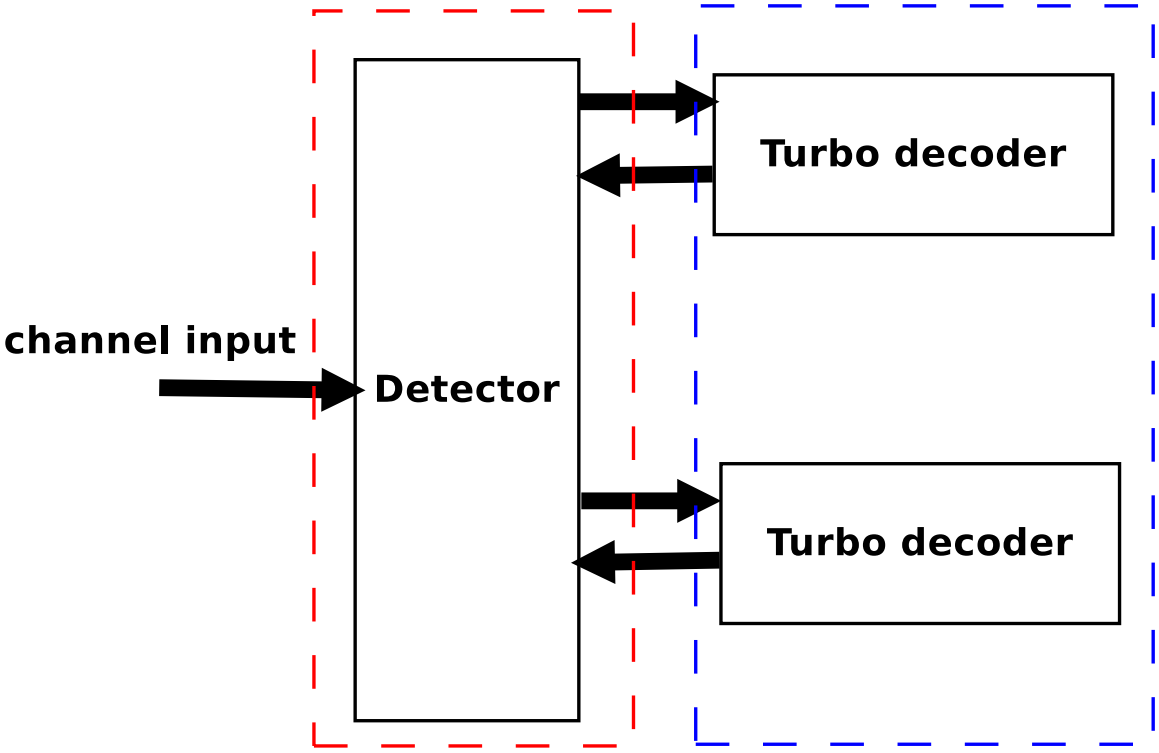

Figure 7.7: EXIT chart of the detector and the component turbo decoders for the separate design, the two component decoders are treated as one decoder.

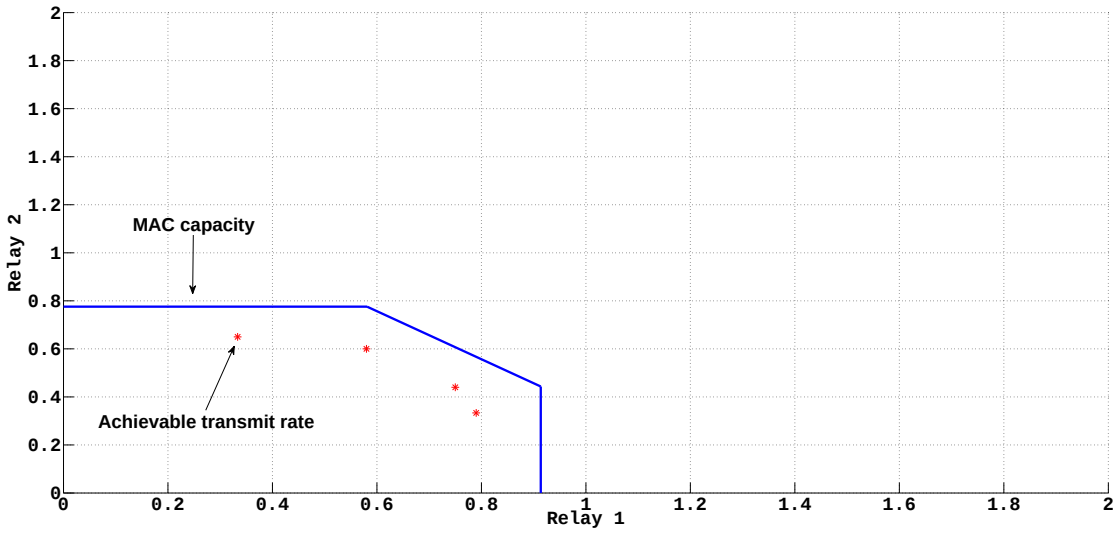

Figure 7.8: Achievable rates for the detector and the turbo decoders

7.6 Simulation Results

The parameters used in the simulation are shown below: frame length is 5×10^4, unpunctured code rate for the turbo code is $\frac{1}{3}$, number of quantization points for LLR quantization is 2. For the ARA code, the parameter for the variable nodes is fixed as in Table 7.1, parameters for the

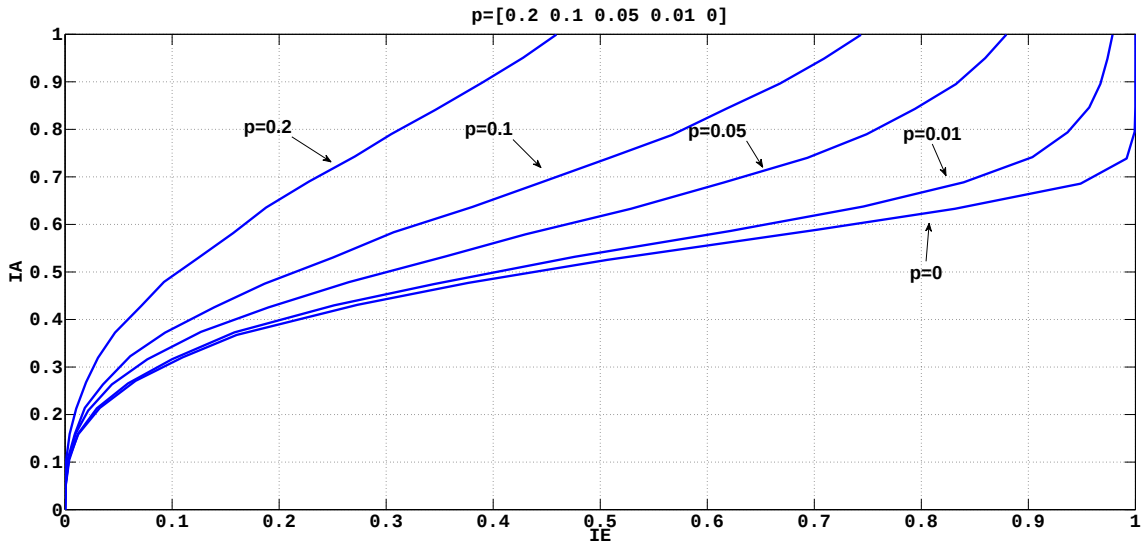

Figure 7.9: EXIT chart of the joint accumulator decoder with various P (probability of difference)

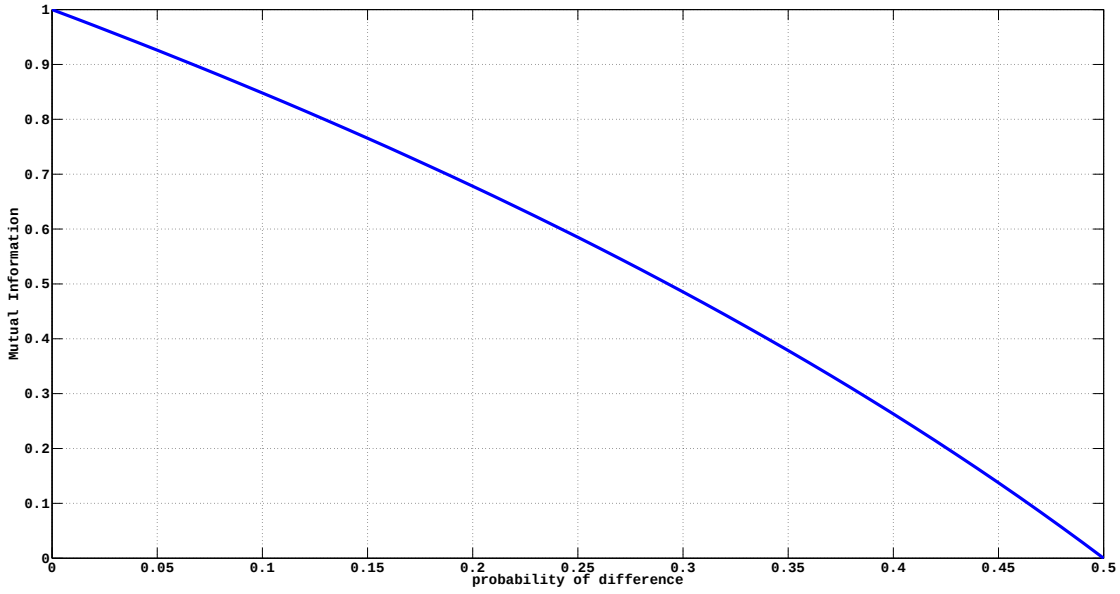

Figure 7.10: Mutual information of the correlation information

check nodes vary according to the rate region. Eb/N0 between transmitter and relay is 10 dB. Generation polynomial for the turbo code is $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$.

In the simulation, for each pair of channel samples (source-relay link and relay-destination link), rates region like Figure 7.1 is drawn first. If there is no intersection, then a theoretic outage is observed. On the other hand, if there is an intersection, but the rate pair that satisfies the conditions in section 7.5 does not exist, a small intersection outage is observed. The theoretical outage and the small intersection outage are compared with the outage from the simulation (frame error rate) for the joint design and the separate design, as shown in Figure 7.17.

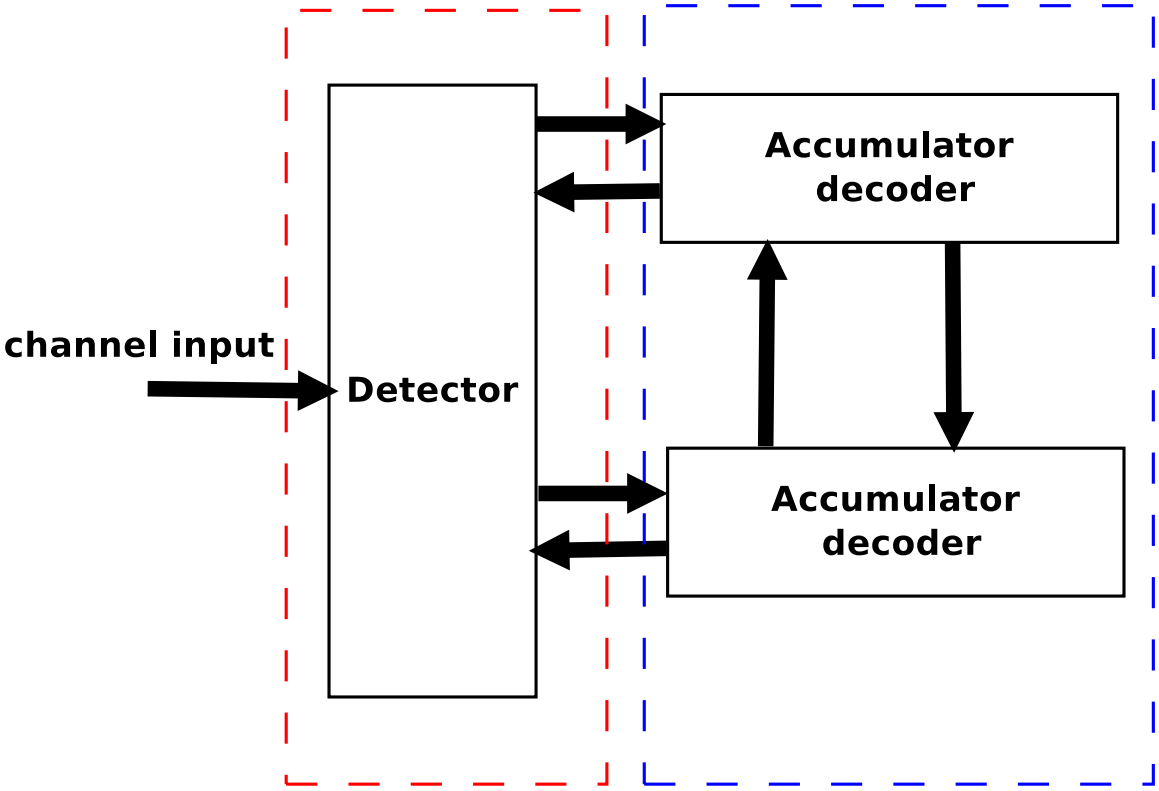

Figure 7.11: EXIT chart of the joint accumulator decoder and the detector

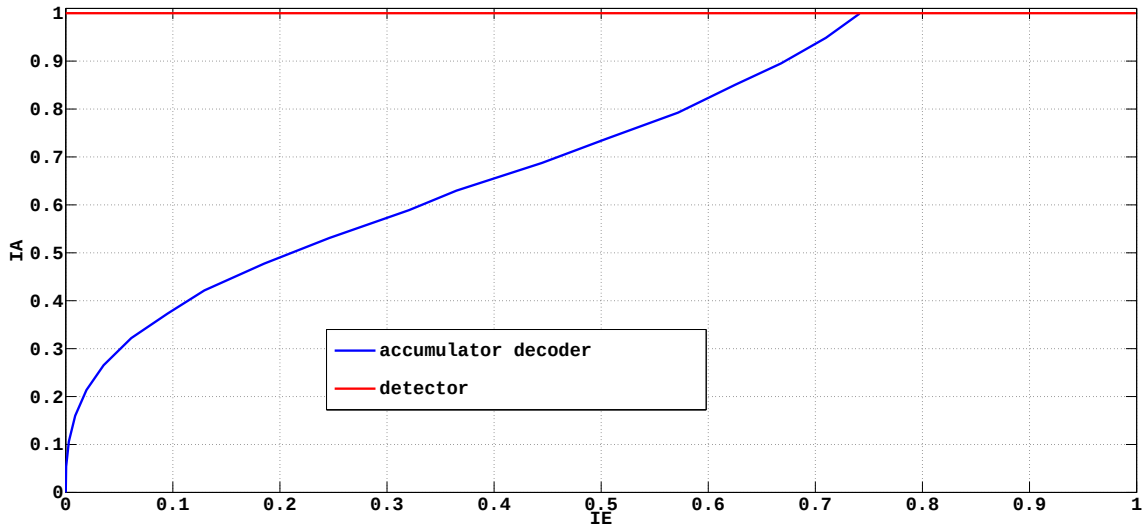

Figure 7.12: EXIT chart of the joint accumulator decoder and the detector with $p = 0.1$ and horizontal detector EXIT curve

Figure 7.17 shows that from the outage point of view, separate design outperforms the joint design which coincides with the results in [92] and our discovery stated above. Both separate design (red line) and joint design (green line) suffer a loss when compared with the theoretical outage (blue line). However, when small intersection area outage is considered (blue broken line), the results of both designs are justified.

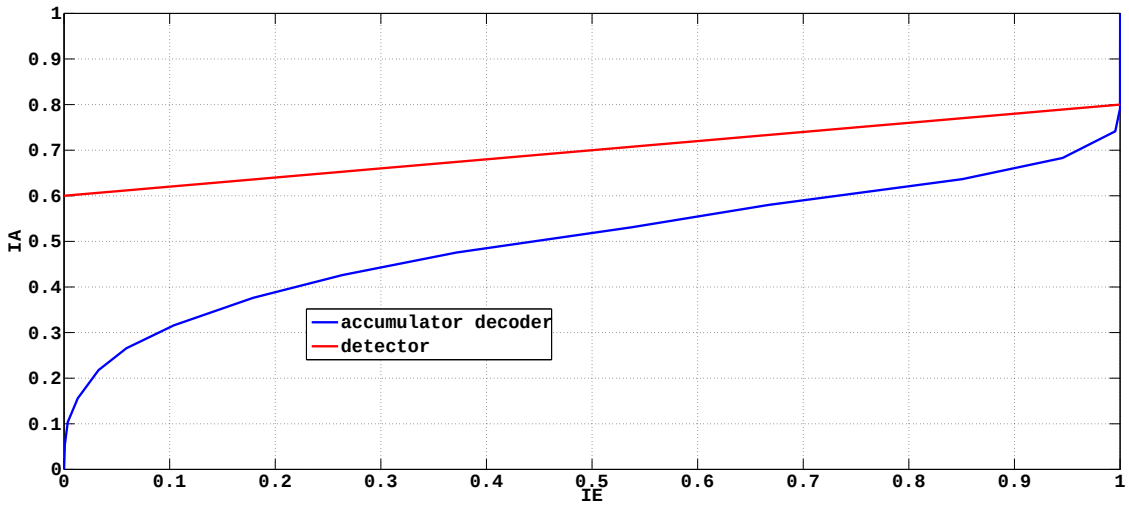

Figure 7.13: EXIT chart of the joint accumulator decoder and the detector with $p = 0$

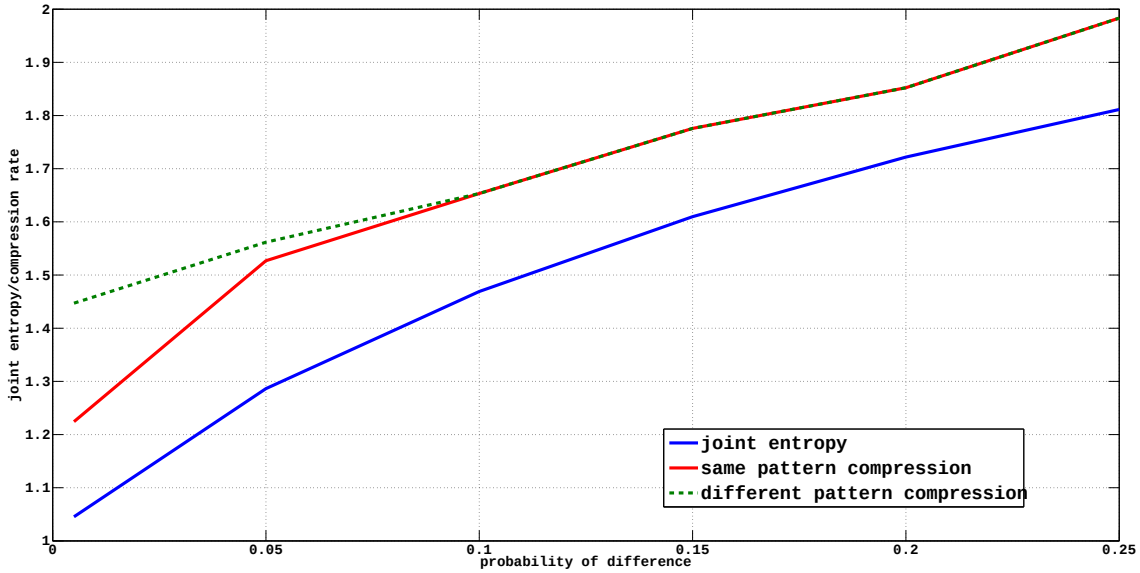

Figure 7.14: Comparison of compression rates for different compression schemes.

Figure 7.18 and Figure 7.19 show the relationship between the intersection area and the number of bits in error for the joint design and the separate design. As the intersection area is a trapezoid, we measure its height (relate to the distance between the joint entropy and the sum rate of the MAC channel) and the average length of the two parallel sides (longer side is the joint entropy line, shorter side is the MAC channel sum rate line). Two figures all show that as the height increases, number of frames with errors decreases. For frames with the same relative short height, when the average length of the two parallel sides increases, number of frames in errors also increases. Simulations show that frames with longer average sides often have smaller joint entropy and the sum rate (weaker MAC channel) which tends to cause errors.

The average overall BER (outer code included)for the joint design and the separate design

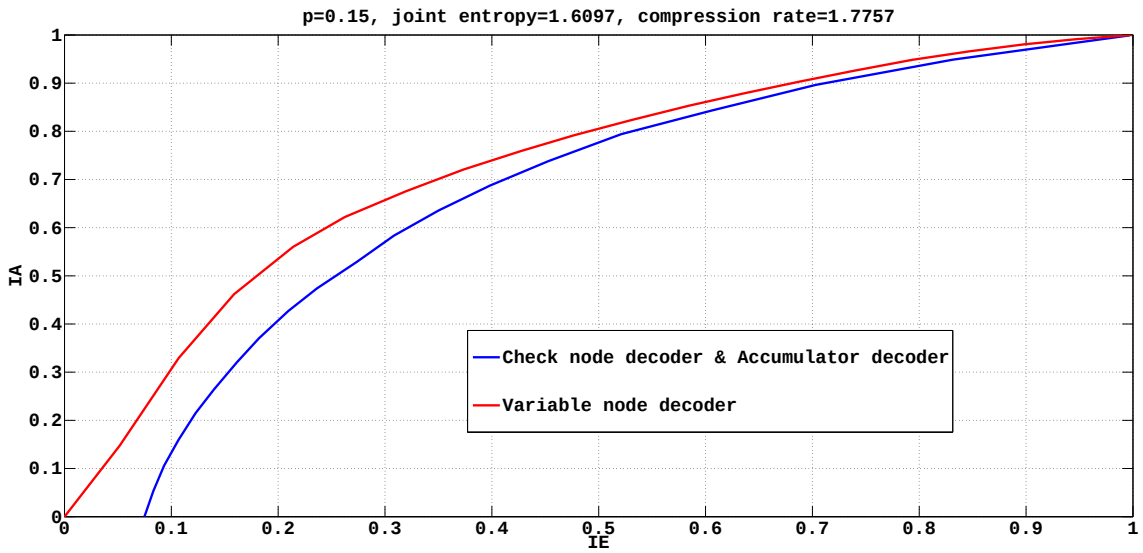

Figure 7.15: EXIT chart for the ARA code, probability of difference is p=0.15, joint entropy is 1.6097, compression rate is 1.7757.

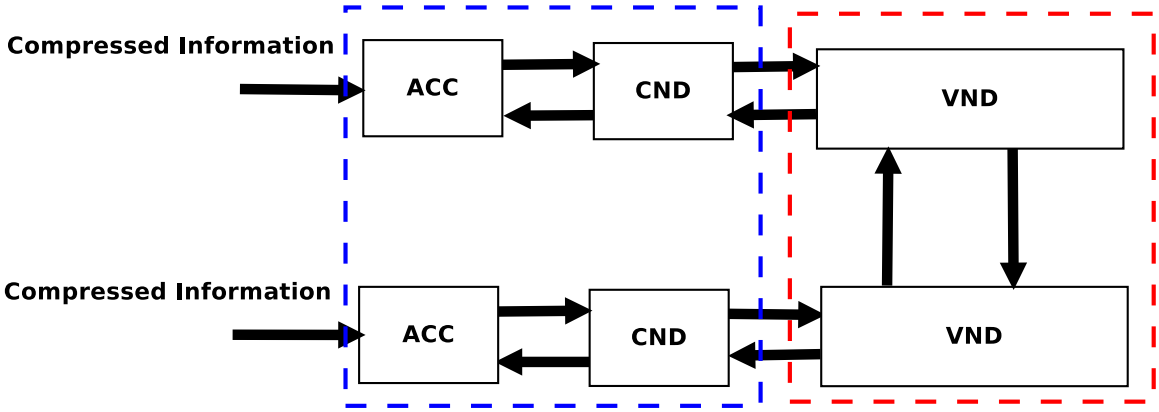

Figure 7.16: EXIT chart of the joint ARA decoder, treat blue box and red box as two component decoders.

is shown in Figure 7.20. As is shown, the joint design has a better BER than the separate design does, this is because the structure of the joint design is simpler and more robust than that of the separate design. Hence when rates mismatched, less bits will be in error.

7.7 Conclusion

We consider a distributed cooperative spatial multiplexing system with a fading multiple access channel on the relay-destination channel and propose two system designs, namely the separate design and the joint design. Simulations show that the separate design outperforms the joint design from the outage probability point of view which coincides with the results in [92] and our analysis of Figure 7.9, Figure 7.12 and Figure 7.13 (paragraph 6 in section 7.5). The reason behind this result is that the joint design works only when there is a full correlation or the MAC

time of repetition	2	3	9
percentage of frame	0.25	0.7	0.05

Table 7.1: VNE parameters: 25% of frame is repeated twice, 70% of frame is repeated thrice and 5% is repeated nine times.

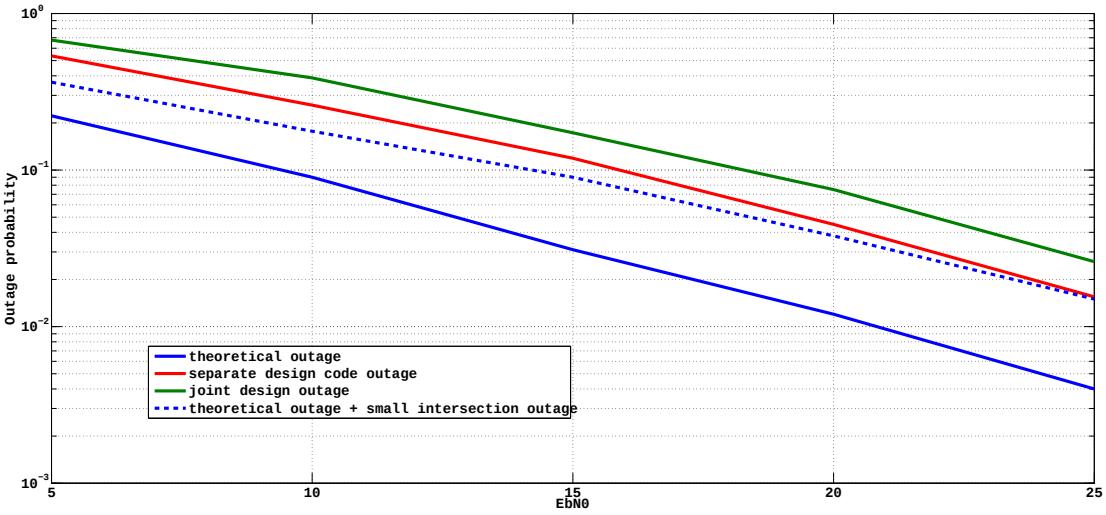

Figure 7.17: Outage probability of the joint design and the separate design

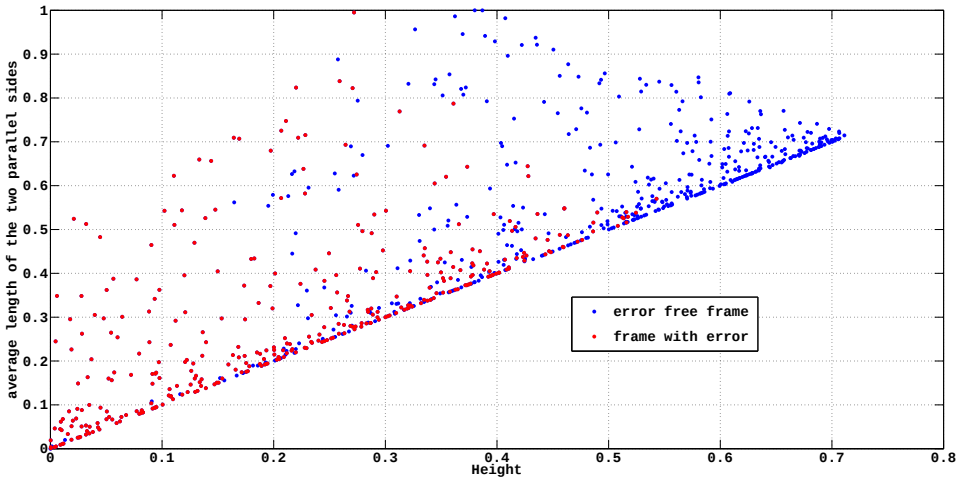

Figure 7.18: Relationship between area of intersection and the number of bits in error for the joint design

channel is perfect. Otherwise, it will suffer inevitable errors. Hence according to our discovery the correlation redundancy used by the joint design is not as good as the redundancy generated by the channel code in the separate design when dealing with the fading channel and the noise.

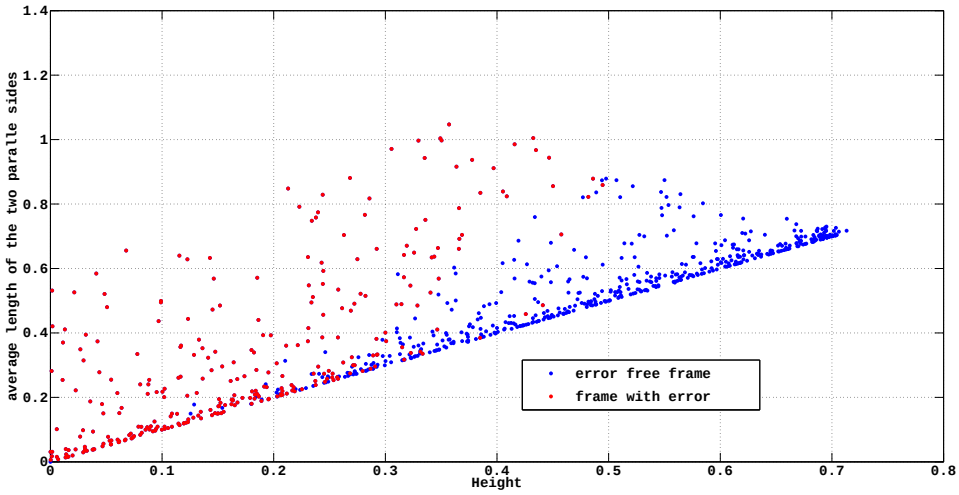

Figure 7.19: Relationship between area of intersection and the number of bits in error for the separate design

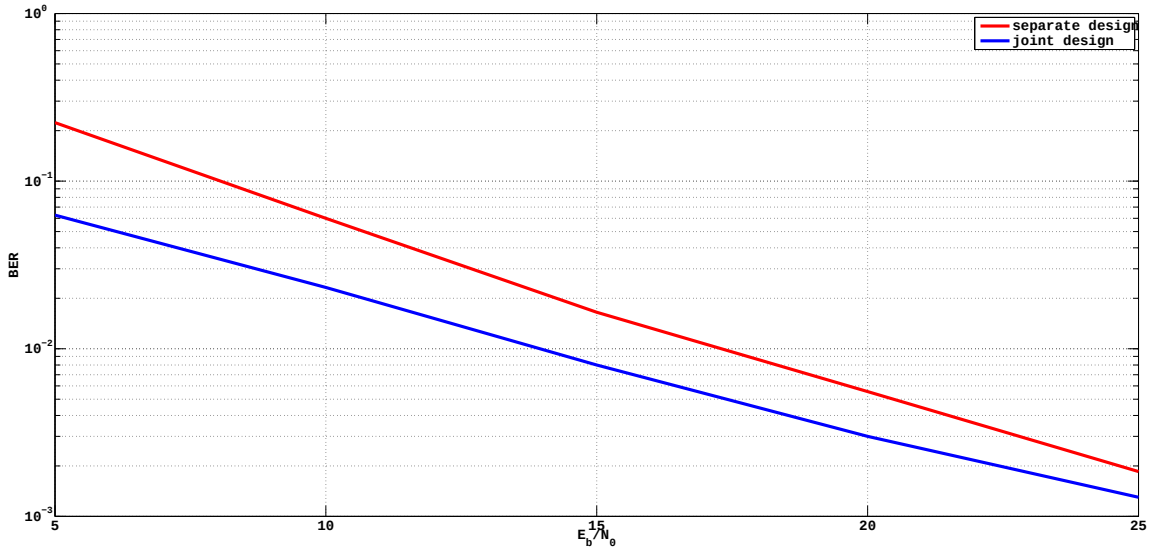

Figure 7.20: Overall system BER of the joint design and the separate design

Chapter 8

Conclusion

8.1 Summary of Achievements

In this chapter we summarize the achievements of our work for this project as well as the overview of our final design.

- In Chapter 3, we have studied a cooperative spatial multiplexing system with a very basic structure. The bit error rate performance is evaluated analytically and the result is justified by simulations. Alternative transmission protocols are also developed to deal with the assumption that different links suffer different path losses across the whole system. Alternative protocols try to compensate for the weak link but also take advantage of the strong link.

- In Chapter 4, we have incorporated the Slepian and Wolf theorem (compression of distributed correlated sources) into our system design. Different Slepian and Wolf code implementation schemes are analysed and compared. Finally, we decided to use accumulate repeat accumulate code to implement the Slepian and Wolf theorem and discovered that for system with fixed compression rate, full diversity order cannot be achieved.

- In Chapter 5, we have analysed our design from the network information theory point of view, and realized that our design results an inevitable error floor. This error floor is studied analytically and the result is justified by simulations. Since the only way to eliminate this error floor is to apply an outer code, the study of this error floor helps to design the outer code.

- In Chapter 6, we have revealed the fundamental challenges of implementing the Slepian and Wolf theorem in the cooperative spatial multiplexing system. We provide our solutions to these difficulties and the results are promising.

- In Chapter 7, we have added more realistic assumption that the relay destination link suffers from block Rayleigh fading and white Gaussian noise and analyse the system performances from the network information theory point of view. We provide solutions for our system to adapt to this new assumption and the results are justified by simulations.

We propose a novel design of a cooperative spatial multiplexing system which outperforms that of others, furthermore we justify the performances of our design by theoretical and practical analysis. However there are still some aspects of this design that can be improved.

8.2 Future Work

- Applying curve fitting technique introduced in section 2.11 to the codes used in our design, compression code and channel code as well, to push the performance of the codes to their theoretical limits, namely capacity-achieving for the channel codes and entropy-achieving for the source codes.

  By applying the curve fitting techniques, we expect the tunnel in the EXIT charts Figure 7.6, Figure 7.13 and Figure 7.15 becomes narrower which indicates the decreasing of the rate loss [46].

- Further developing of the analysis in Chapter 5, extend it to the system with relays which perform soft decision.

  In Chapter 5, we assume that relays do the hard decision on the received signals and the receiver uses equation (4.4.2) to combine the information that comes from the relays (assume error probabilities on every relays are available at the receiver). Whereas in chapter 7 the final design assumes that relays do the soft decision and forward the quantized soft information to the receiver where information from relays is simply added together.

- Applying Berger-Tung coding scheme [89, 88] (lossy compression of distributed correlated sources with given distortion) in our system design.

  Theorem 2. ***(Berger-Tung Inner Bound)*** *Let (X_1, X_2) be a 2 discrete memoryless source and $d_1(x_1, \hat{x}_1)$ and $d_2(x_2, \hat{x}_2)$ be two distortion measures. A rate pair (R_1, R_2) is achievable with distortion pair (D_1, D_2) for distributed lossy source coding if*

$$\begin{aligned} R_1 &> I(X_1; U_1 | U_2, Q) \\ R_2 &> I(X_2; U_2 | U_1, Q) \\ R_1 + R_2 &> I(X_1, X_2; U_1, U_2 | Q) \end{aligned} \tag{8.1}$$

  for some conditional probability mass function $p(q)p(u_1|x_1,q)p(u_2|x_2,q)$ with $|U_j| <= |X_j|+4, j = 1,2$, and functions $\hat{X}_1(u_1,u_2,q)$ and $\hat{X}_2(u_1,u_2,q)$ such that $E(d_j(X_j,\hat{X}_j)) <= D_j, j = 1,2$

  U is a auxiliary random variable that does not correspond to any of the channel variables. Q is a time-sharing random variable. The Berger-Tung inner bound reduces to the Slepian-Wolf region when d_1 and d_2 are Hamming distortion measures and $D_1 = D_2 = 0$.

  Theorem 3. ***(Berger-Tung Outer Bound)*** *Let (X_1, X_2) be a 2 discrete memoryless source and $d_1(x_1, \hat{x}_1)$ and $d_2(x_2, \hat{x}_2)$ be two distortion measures. If a rate pair (R_1, R_2) is achievable with distortion pair (D_1, D_2) for distributed lossy source coding, then it must satisfy the inequalities*

$$\begin{aligned} R_1 &>= I(X_1, X_2; U_1 | U_2) \\ R_2 &>= I(X_1, X_2; U_2 | U_1) \\ R_1 + R_2 &>= I(X_1, X_2; U_1, U_2) \end{aligned} \tag{8.2}$$

  for some conditional probability mass function $p(u_1,u_2|x_1,x_2)$ and function $\hat{x}_1(u_1,u_2)$ and $\hat{x}_2(u_1,u_2)$ such that $U_1 \to X_1 \to X_2$ and $X_1 \to X_2 \to U_2$ from Markov chains and $E(d_j(X_j,\hat{X}_j)) < D_j, j = 1,2$

This outer bound is similar to the Berger-Tung inner bond except that the region is convex without the use of a time-sharing random variable and the Markov conditions are weaker than the condition $U_1 \rightarrow X_1 \rightarrow X_2 \rightarrow U_2$ in the inner bound.

Appendix A

Derivation of $\tilde{\sigma}$

The derivation of $\tilde{\sigma}$, standard deviation of $\mathrm{Re}(\mathbf{N}_o^H \frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|})$ is shown here. All the symbolic calculation in this section has been carried out using Mathematica [97].

From (3.4), the overall noise can be expressed as

$$\begin{aligned}\mathbf{N}_o = \mathbf{H}_{n+1} g_n(\mathbf{H}_n \cdots \mathbf{H}_3 g_2(\mathbf{H}_2 g_1 \mathbf{N}_1 + \mathbf{N}_2) + \mathbf{N}_3 + \cdots \\ + \mathbf{N}_n) + \mathbf{N}_{n+1} = \begin{pmatrix} n_{o1} \\ n_{o2} \end{pmatrix}\end{aligned} \tag{A.1}$$

Assuming

$$\begin{aligned}\hat{\mathbf{N}} &= \mathbf{H}_2 g_1 \mathbf{N}_1 + \mathbf{N}_2 \\ &= \begin{pmatrix} h_1 & h_2 \\ h_3 & h_4 \end{pmatrix} \begin{pmatrix} n_1 \\ h_2 \end{pmatrix} + \begin{pmatrix} n_3 \\ h_4 \end{pmatrix} \\ &= \begin{pmatrix} \hat{n}_{1p} + \hat{n}_{1q} i \\ \hat{n}_{2p} + \hat{n}_{2q} i \end{pmatrix}\end{aligned} \tag{A.2}$$

where subscripts p and q denote the real and imaginary parts of a complex number, and i is the imaginary unit and $h_1 = h_1' + h_1'' i$, $n_1 = n_1' + n_1'' i$ the rest of other entries in channel and noise matrices follow the same rule. Note that channel matrix is a particular sample of the fading channel, whereas noise matrices contain noise random variables.

After applying basic mathematical manipulation, it is shown that

$$\begin{aligned}\hat{n}_{1p} &= h_1' n_1' - h_1'' n_1'' + h_2' n_2' - h_2'' n_2'' + n_3' \\ \hat{n}_{1q} &= h_1'' n_1' + h_1' n_1'' + h_2'' n_2' + h_2' n_2'' + n_3'' \\ \hat{n}_{2p} &= h_3' n_1' - h_3'' n_1'' + h_4' n_2' - h_4'' n_2'' + n_4' \\ \hat{n}_{2q} &= h_3'' n_1' + h_3' n_1'' + h_4'' n_2' + h_4' n_2'' + n_4''\end{aligned} \tag{A.3}$$

and

$$\begin{aligned}E(\hat{n}_{1p}\hat{n}_{2p}) &= h_1' h_3' + h_1'' h_3'' + h_2' h_4' + h_2'' h_4'' \\ E(\hat{n}_{1q}\hat{n}_{2q}) &= h_1' h_3' + h_1'' h_3'' + h_2' h_4' + h_2'' h_4'' \\ E(\hat{n}_{1p}\hat{n}_{2q}) &= h_1' h_3'' - h_1'' h_3' + h_2' h_4'' - h_4' h_2'' \\ E(\hat{n}_{1q}\hat{n}_{2p}) &= h_1'' h_3' - h_3'' h_1' + h_2'' h_4' - h_4'' h_2'\end{aligned} \tag{A.4}$$

where E denotes expectation. Note that during the calculation of (A.4), two properties of Gaussian random variables are used,

1. *Property 1* The mean of a squared standard normally distributed random variable is 1
2. *Property 2* The mean of the product of two independent standard normally distributed random variables is 0

It is easy to see that

$$E(\hat{n}_{1p}\hat{n}_{2p}) = E(\hat{n}_{1q}\hat{n}_{2q}) \tag{A.5}$$

$$E(\hat{n}_{1p}\hat{n}_{2q}) = -E(\hat{n}_{1q}\hat{n}_{2p}) \tag{A.6}$$

.

Applying Theorem 4 (defined and proofed below) to (A.5) and (A.6), it can be further proved that

$$E(\mathrm{Re}(n_{o1})\mathrm{Re}(n_{o2})) = E(\mathrm{Im}(n_{o1})\mathrm{Im}(n_{o2})\ ) \tag{A.7}$$

$$E(\mathrm{Re}(n_{o1})\mathrm{Im}(n_{o2})) = -E(\mathrm{Im}(n_{o1})\mathrm{Re}(n_{o2})\ ) \tag{A.8}$$

where Im() and Re() represent imaginary and real part respectively of a complex number. Because n_{o1} and n_{o2} are circularly symmetric, then

$$E(\mathrm{Re}(n_{o1})\mathrm{Im}(n_{o1})) = E(\mathrm{Re}(n_{o2})\mathrm{Im}(n_{o2})\ ) = 0 \tag{A.9}$$

according to the Bernstein's theorem [98].

Assume

$$E(\mathrm{Re}(n_{o1})^2) = E(\mathrm{Im}(n_{o1})^2) = \sigma 1 \tag{A.10}$$

$$E(\mathrm{Re}(n_{o2})^2) = E(\mathrm{Im}(n_{o2})^2) = \sigma 2 \tag{A.11}$$

$$\frac{\mathbf{H}_o(\mathbf{M}_k - \mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i - \mathbf{M}_k)\|} = \begin{pmatrix} x_{1p} + x_{1q}i \\ x_{2p} + x_{2q}i \end{pmatrix} = \mathbf{X} \tag{A.12}$$

$$E(\mathrm{Re}(n_{o1})\mathrm{Re}(n_{o2})) = \rho_a \tag{A.13}$$

$$E(\mathrm{Re}(n_{o1})\mathrm{Im}(n_{o2})) = \rho_b \tag{A.14}$$

then the variance of $\mathrm{Re}(\mathbf{N}_o^H \frac{\mathbf{H}_o(\mathbf{M}_k-\mathbf{M}_i)}{\|\mathbf{H}_o(\mathbf{M}_i-\mathbf{M}_k)\|})$ is

$$E(Re\left(\begin{pmatrix} n_{o1} \\ n_{o2} \end{pmatrix}^H \begin{pmatrix} x_{1p} + x_{1q}i \\ x_{2p} + x_{2q}i \end{pmatrix}\right)^2) \tag{A.15}$$

Combining (A.7),(A.8),(A.9),(A.13),(A.14),(A.10) and (A.11) (details are given below), (A.15) can be modified to

$$\frac{1}{2}\mathrm{Tr}((E(\mathbf{N}_o\mathbf{N}_o^H))(\mathbf{X}\mathbf{X}^H)) \tag{A.16}$$

where $Tr()$ is the trace of a matrix.

Therefore the standard deviation

$$\tilde{\sigma} = \sqrt{\frac{1}{2}\mathrm{Tr}((E(\mathbf{N}_o\mathbf{N}_o^H))(\mathbf{X}\mathbf{X}^H))} \tag{A.17}$$

where

$$\begin{aligned}
E(\mathbf{N}_o\mathbf{N}_o^H) =& E((\mathbf{H}_{n+1}g_n(\mathbf{H}_n\cdots\mathbf{H}_3g_2(\mathbf{H}_2g_1\mathbf{N}_1+\mathbf{N}_2)+\mathbf{N}_3+\cdots+\mathbf{N}_n)+\mathbf{N}_{n+1}))\\
& E((\mathbf{H}_{n+1}g_n(\mathbf{H}_n\cdots\mathbf{H}_3g_2(\mathbf{H}_2g_1\mathbf{N}_1+\mathbf{N}_2)+\mathbf{N}_3+\cdots+\mathbf{N}_n)+\mathbf{N}_{n+1})^H)\\
=& 2\sigma^2\mathbf{H}_{n+1}g_n\mathbf{H}_n\cdots\mathbf{H}_3g_2\mathbf{H}_2g_1\mathbf{H}_{n+1}^Hg_n\mathbf{H}_n^H\cdots\mathbf{H}_3^Hg_2\mathbf{H}_2^Hg_1\\
& +2\sigma^2\mathbf{H}_n\cdots\mathbf{H}_3g_2\mathbf{H}_n^H\cdots\mathbf{H}_3^Hg_2\\
& +\cdots\\
& +2\sigma^2\mathbf{H}_{n+1}g_n\mathbf{H}_{n+1}^Hg_n\\
& +2\sigma^2\mathbf{I}
\end{aligned} \tag{A.18}$$

and $\mathbf{I}$ is a 2×2 identity matrix

The proof that A.16 is equal to A.15 is shown below

Assume $\begin{pmatrix} n_{o1} \\ n_{o2} \end{pmatrix}$ equals to $\begin{pmatrix} n_1'+n_1''i \\ n_2'+n_2''i \end{pmatrix}$, A.15 equals to

$$\begin{aligned}
& Re\left[\begin{pmatrix} n_1'-n_1''i & n_2'-n_2''i \end{pmatrix}\begin{pmatrix} x_{1p}+x_{1q}i \\ x_{2p}+x_{2q}i \end{pmatrix}\right]^2\\
=&[n_1'x_{1p}+n_1''x_{1q}+n_2'x_{2p}+n_2''x_{2q}]^2\\
=&(n_1'x_{1p})^2+(n_1''x_{1q})^2+(n_2'x_{2p})^2+(n_2''x_{2q})^2\\
&+2n_1'x_{1p}n_1''x_{1q}+2n_1'x_{1p}n_2'x_{2p}+2n_1'x_{1p}n_2''x_{2q}+2n_1''x_{1q}n_2'x_{2p}+2n_1''x_{1q}n_2''x_{2q}+2n_2'x_{2p}n_2''x_{2q}
\end{aligned} \tag{A.19}$$

Applying (A.13),(A.14),(A.10) and (A.11), A.15 is equals to

$$2\rho_a(x_{1p}x_{2p}+x_{1q}x_{2q})+2\rho_b(x_{1p}x_{2q}-x_{1q}x_{2p})+\sigma 1^2(x_{1p}^2+x_{1q}^2)+\sigma 2^2(x_{2p}^2+x_{2q}^2) \tag{A.20}$$

A.16 is equal to

$$\begin{aligned}
& \frac{1}{2}Tr(\left(\begin{pmatrix} n_1'+n_1''i \\ n_2'+n_2''i \end{pmatrix}\begin{pmatrix} n_1'-n_1''i & n_2'-n_2''i \end{pmatrix}\right)\\
& \left(\begin{pmatrix} x_{1p}+x_{1q}i \\ x_{2p}+x_{2q}i \end{pmatrix}\begin{pmatrix} x_{1p}-x_{1q}i & x_{2p}-x_{2q}i \end{pmatrix}\right))\\
=&\frac{1}{2}Tr(\begin{pmatrix} n_1'^2+n_1''^2 & n_1'n_2'+n_1''n_2''+(n_1''n_2'-n_1'n_2'')i \\ n_2'n_1'+n_2''n_1''+(n_2''n_1'-n_2'n_1'')i & n_2'^2+n_2''^2 \end{pmatrix}\\
& \begin{pmatrix} x_{1p}^2+x_{1q}^2 & x_{1p}x_{2p}+x_{1q}x_{2q}+(x_{1q}x_{2p}-x_{1p}x_{2q})i \\ x_{2p}x_{1p}+x_{2q}x_{1q}+(x_{2q}x_{1p}-x_{2p}x_{1q})i & x_{2p}^2+x_{2q}^2 \end{pmatrix})\\
=&\frac{1}{2}Tr(\begin{pmatrix} 2\sigma 1^2 & 2\rho_a-2\rho_bi \\ 2\rho_a+2\rho_bi & 2\sigma 2^2 \end{pmatrix}\\
& \begin{pmatrix} x_{1p}^2+x_{1q}^2 & x_{1p}x_{2p}+x_{1q}x_{2q}+(x_{1q}x_{2p}-x_{1p}x_{2q})i \\ x_{2p}x_{1p}+x_{2q}x_{1q}+(x_{2q}x_{1p}-x_{2p}x_{1q})i & x_{2p}^2+x_{2q}^2 \end{pmatrix})\\
=&4\rho_a(x_{1p}x_{2p}+x_{1q}x_{2q})+4\rho_b(x_{1p}x_{2q}-x_{1q}x_{2p}))+2\sigma 1^2(x_{1p}^2+x_{1q}^2)+2\sigma 2^2(x_{2p}^2+x_{2q}^2)
\end{aligned} \tag{A.21}$$

Compare A.21 to A.20, A.16 is proofed.

Theorem 4. *If matrix* $\mathbf{N} = \begin{pmatrix} n_{1p} + n_{1q}i \\ n_{2p} + n_{2q}i \end{pmatrix}$*, whose entries are white Gaussian but not necessarily circularly symmetric noises, has following properties,*

$$E(n_{1p}n_{2p}) = E(n_{1q}n_{2q}) \tag{A.22}$$

$$E(n_{1p}n_{2q}) = -E(n_{1q}n_{2p}) \tag{A.23}$$

then $\mathbf{R} = \mathbf{HN} + \mathbf{Z} = \begin{pmatrix} r_{1p} + r_{1q}i \\ r_{2p} + r_{2q}i \end{pmatrix}$ *also has the same properties, that is*

$$E(r_{1p}r_{2p}) = E(r_{1q}r_{2q}) \tag{A.24}$$

$$E(r_{1p}r_{2q}) = -E(r_{1q}r_{2p}) \tag{A.25}$$

for any given $\mathbf{H}$*, where* $\mathbf{H}$ *is a* 2×2 *matrix whose entries are fixed channel coefficients and* $\mathbf{Z} = \begin{pmatrix} z_{1p} + z_{1q}i \\ z_{2p} + z_{2q}i \end{pmatrix}$ *is a matrix whose entries are white circularly symmetric Gaussian noises and are independent of* $\mathbf{N}$*. (Note that the expectations here are taken over the noise components, not the channel coefficients)*

Proof of Theorem 4. Assume that

$$E(n_{1p}^2 = n_{1q}^2) = \hat{\sigma}_1^2 \tag{A.26}$$

$$E(n_{2p}^2 = n_{2q}^2) = \hat{\sigma}_2^2 \tag{A.27}$$

Since the entries of $\mathbf{Z}$ are circularly symmetric,

$$E(z_{1p}z_{2p}) = E(z_{1q}z_{2q}) = E(z_{1p}z_{2q}) = E(z_{1q}z_{2p}) = E(z_{1p}z_{1q}) = E(z_{2p}z_{2q}) = 0 \tag{A.28}$$

$$E(z_{1p}^2) = E(z_{1q}^2) = E(z_{2q}^2) = E(z_{2p}^2) = \hat{\sigma}_3^2 \tag{A.29}$$

Because $\mathbf{N}$ and $\mathbf{Z}$ are independent, the expectation of the product of any combination of n_x, $x \subset (1p, 1q, 2p, 2q)$ and z_x, $x \subset (1p, 1q, 2p, 2q)$ is equal to zero.

Combining (A.22),(A.23),(A.26),(A.27),(A.28),(A.29) and the independence between $\mathbf{N}$ and $\mathbf{Z}$, we establish by expending the expression and taking expectations that $\mathbf{R} = \mathbf{HN} + \mathbf{Z} = \begin{pmatrix} r_{1p} + r_{1q}i \\ r_{2p} + r_{2q}i \end{pmatrix}$ has the following properties

$$E(r_{1p}r_{2p}) = E(r_{1q}r_{2q}) \tag{A.30}$$

$$E(r_{1p}r_{2q}) = -E(r_{1q}r_{2p}) \tag{A.31}$$

□

Appendix B

Calculation of mutual information between source and the joint entropy of relays

As shown in Figure 6.3, the joint entropy of the relays is $A+B+D+E+F+G$ and the mutual information between source and the joint entropy of relays is $E+F+G$ which represents the useful information about the source contained in the relays signals and the theoretical minimum achievable rate on the relay destination link.

$$E+F+G=(E+G)+(F+G)-G$$

where $E+G$ and $F+G$ are the mutual information between the source and each relay respectively and G is the mutual information between the source and both relays. According to [99]

$$\begin{aligned} G &= I(S,R_1,R_2) \\ &= I(R_1,R_2)-I(R_1,R_2|S) \\ &= \sum_{s\in S}\sum_{r_1\in R_1}\sum_{r_2\in R_2} p_{S,R_1,R_2}(s,r_1,r_2) \\ &\log\frac{p_S(s)p_{S,R_1,R_2}(s,r_1,r_2)}{p_{R_1,S}(r_1,s)p_{R_2,S}(r_2,s)} \end{aligned}$$

where $p()$ is marginal or joint probability mass function according to its subscript.

List of Abbreviations

AF Amplify and Forward

AP Access Point

ARA Accumulate-Repeat-Accumulate

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BS Base Station

C-DIV Cooperative Diversity

C-SM Cooperative Spatial Multiplexing

CEO problem Chief Executive Officer problem

CND Check Node Decoder

CNE Check Node Encoder

CSI Channel State Information

DF Decode and Forward

DSTBC Distributed Space-Time Block Coding

EXIT chart Extrinsic Information Transfer chart

FER Frame Error Rate

FSPL Free-Space Path Loss

IAC Interference Alignment and Cancellation

LLR Log Likelihood-Ratio

LOS Line Of Sight

MANETs Mobile ad-hoc Networks

MCH Master Cluster Head

MGF Moment Generating Function

MIMO Multiple-Input-Multiple-Output

ML Maximum Likelihood

MMSE Minimum Mean Square Error

PDF Probability Density Function

PSD Power Spectral Density

PU Primary User

SCH Slave Cluster Head

SISO Single-Input-Single-Output

SM Spatial Multiplexing

SNR Signal-to-Noise Ratio

SU Secondary User

SW theorem Slepian and Wolf theorem

UEP Unequal Error Protection

VAA Virtual Antenna Array

VBLAST Vertical-Bell Laboratories Layered Space-Time

VND Variable Node Decoder

VNE Variable Node Encoder

VQ Vector Quantization

WSN Wireless Sensor Network

Bibliography

[1] M. Dohler, E. Lefranc, and H. Aghvami, "Space-time block codes for virtual antenna arrays," in *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002.*, vol. 1, sept. 2002, pp. 414 – 417 vol.1.

[2] G. Golden, C. Foschini, R. Valenzuela, and P. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14 –16, jan 1999.

[3] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744 –765, mar 1998.

[4] S. Alamouti, "A simple transmit diversity technique for wireless communications ," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451 –1458, oct 1998.

[5] A. Goldsmith, *Wireless Communications*. UK: Cambridge University Press, 2005.

[6] W. Roh and A. Paulraj, "Outage performance of the distributed antenna systems in a composite fading channel," in *2002 IEEE 56th Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall.*, vol. 3, 2002, pp. 1520 – 1524 vol.3.

[7] V. Emamian, P. Anghel, and M. Kaveh, "Multi-user spatial diversity in a shadow-fading environment," in *2002 IEEE 56th Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall.*, vol. 1, 2002, pp. 573 – 576 vol.1.

[8] H. Dai, "Distributed Versus Co-Located MIMO Systems with Correlated Fading and Shadowing," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.*, vol. 4, may 2006, p. IV.

[9] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415 – 2425, oct. 2003.

[10] G. Wang, Y. Zhang, and M. Amin, "Cooperation diversity using differential distributed space-time codes," in *2004 the 5th International Symposium on Multi-Dimensional Mobile Communications Proceedings. The 2004 Joint Conference of the 10th Asia-Pacific Conference on*, vol. 1, aug.-1 sept. 2004, pp. 287 – 291 vol.1.

[11] Y. Hua, Y. Mei, and Y. Chang, "Parallel wireless mobile relays with space-time modulations," in *2003 IEEE Workshop on Statistical Signal Processing*, sept.-1 oct. 2003, pp. 375 – 378.

[12] S. Kim and R. Cherukuri, "Cooperative spatial multiplexing for high-rate wireless communications," in *2005 IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, june 2005, pp. 181 – 185.

[13] S. W. Kim and R. Cherukuri, "Adaptive forwarding and coding in cooperative spatial multiplexing system," in *2005 IEEE 62nd Vehicular Technology Conference, 2005. VTC-2005-Fall.*, vol. 4, sept., 2005, pp. 2418 – 2422.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[15] W. Lee, *Mobile Communications Engineering*. USA: New York: McGraw-Hill, 1982.

[16] B. Rankov and A. Wittneben, "Distributed spatial multiplexing in a wireless network," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, vol. 2, nov. 2004, pp. 1932 – 1937 Vol.2.

[17] M. Dohler and A. Aghvami, "Distributed Antennas: The Concept of Virtual Antenna Arrays," in *Cooperation in Wireless Networks: Principles and Applications*, F. Fitzek and M. Katz, Eds. Springer Netherlands, 2006, pp. 421–461.

[18] J. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, 2003.

[19] D. Nguyen and M. Krunz, "Cooperative MIMO in wireless networks: recent developments and challenges," *IEEE Network.*, vol. 27, no. 4, pp. 48–54, July 2013.

[20] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1098, 2004.

[21] S. Jayaweera, "Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 984–989, 2006.

[22] M. Siam, M. Krunz, and O. Younis, "Energy-Efficient Clustering/Routing for Cooperative MIMO Operation in Sensor Networks," in *IEEE INFOCOM 2009*, 2009, pp. 621–629.

[23] G. Jakllari, S. Krishnamurthy, M. Faloutsos, P. Krishnamurthy, and O. Ercetin, "A Cross-Layer Framework for Exploiting Virtual MISO Links in Mobile Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, pp. 579–594, 2007.

[24] S. Gollakota, S. D. Perli, and D. Katabi, "Interference Alignment and Cancellation," *SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 4, pp. 159–170, Aug. 2009. [Online]. Available: http://doi.acm.org/10.1145/1594977.1592588

[25] S. Hua, H. Liu, M. Wu, and S. Panwar, "Exploiting MIMO antennas in cooperative cognitive radio networks," in *IEEE INFOCOM, 2011 Proceedings*, 2011, pp. 2714–2722.

[26] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, 2010.

[27] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications.*, vol. 51, no. 3, pp. 389 – 399, march 2003.

[28] L. Hanzo, T. Liew, and B. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*. Wiley, 2002. [Online]. Available: http://books.google.co.uk/books?id=3mOhrvMWTwQC

[29] S. ten Brink, "Convergence of iterative decoding," *Electronics Letters*, vol. 35, no. 10, pp. 806 –808, may 1999.

[30] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727 –1737, oct 2001.

[31] J. Hagenauer, "The EXIT Chart - Introduction to Extrinsic Information Transfer," in *12th Europ. Signal Proc. Conf (EUSIPCO) in Iterative Processing, In Proc.*, 2004, pp. 1541–1548.

[32] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July, October 1948. [Online]. Available: http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf

[33] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. UPH, Shaftesbury Road, Cambridge, CB2 8BS, United Kingdom: Cambridge University Press, January 2012.

[34] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471 – 480, jul 1973.

[35] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory.*, vol. 18, no. 1, pp. 2–14, Jan 1972.

[36] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *IEEE International Conference on Communications, 1993. ICC '93 Geneva. Technical Program, Conference Record.*, vol. 2, May 1993, pp. 1064–1070 vol.2.

[37] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Transactions on Signal Processing.*, vol. 51, no. 11, pp. 2764 – 2772, nov 2003.

[38] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Transactions on Information Theory.*, vol. 47, no. 2, pp. 599–618, Feb 2001.

[39] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Transactions on Information Theory.*, vol. 47, no. 2, pp. 619–637, Feb 2001.

[40] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Transactions on Communications.*, vol. 52, no. 4, pp. 670 – 678, april 2004.

[41] S. ten Brink, "Code doping for triggering iterative decoding convergence," in *2001 IEEE International Symposium on Information Theory, 2001. Proceedings.*, 2001, p. 235.

[42] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory.*, vol. 42, no. 2, pp. 429 –445, mar 1996.

[43] A. A. E. Sharon and S. Litsyn, "EXIT functions for the Gaussian channel," in *40th Annu. Allerton Conf. Communication, Control, Computers*, Oct 2003, p. 972981.

[44] ——, "EXIT functions for continuous channelsPart I: Constituent codes," in *IEEE Trans. Commun*.

[45] S.-Y. Chung, J. Forney, G.D., T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Communications Letters.*, vol. 5, no. 2, pp. 58–60, Feb 2001.

[46] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Transactions on Information Theory.*, vol. 50, no. 11, pp. 2657–2673, Nov 2004.

[47] M. Tchler and J. Hagenauer, "EXIT charts of irregular codes ," 2002.

[48] T. Richardson and R. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Transactions on Information Theory.*, vol. 47, no. 2, pp. 638–656, Feb 2001.

[49] C. Di, D. Proietti, I. Telatar, T. Richardson, and R. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Transactions on Information Theory.*, vol. 48, no. 6, pp. 1570–1579, Jun 2002.

[50] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[51] E. T. Ar and I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, 1999.

[52] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory.*, vol. 44, no. 6, pp. 2619–2692, Oct 1998.

[53] A. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *IEEE-VTS Fall VTC 52nd Vehicular Technology Conference, 2000.*, vol. 1, 2000, pp. 473–478 vol.1.

[54] A. van Zelst, R. Van Nee, and G. Awater, "Turbo-BLAST and its performance," in *IEEE VTS 53rd Vehicular Technology Conference, 2001. VTC 2001 Spring.*, vol. 2, 2001, pp. 1282–1286 vol.2.

[55] A. Stefanov and T. Duman, "Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: system model, decoding approaches, and practical considerations," *IEEE Journal on Selected Areas in Communications.*, vol. 19, no. 5, pp. 958–968, May 2001.

[56] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Analysis, design, and iterative decoding of double serially concatenated codes with interleavers," *IEEE Journal on Selected Areas in Communications.*, vol. 16, no. 2, pp. 231–244, Feb 1998.

[57] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284 – 287, mar 1974.

[58] S. ten Brink and G. Kramer, "Turbo processing for scalar and vector channels," *3rd Int. Symp. Turbo Codes Related Topics*, pp. 23–30, Sept 2003.

[59] A. Abdaoui, S. Ikki, and M. Ahmed, "Performance Analysis of MIMO Cooperative Relaying System Based on Alamouti STBC and Amplify-and-Forward Schemes," in *IEEE International Conference on Communications (ICC), 2010*, may 2010, pp. 1 –6.

[60] J.-B. Kim and D. Kim, "End-to-End BER Performance of Cooperative MIMO Transmission with Antenna Selection in Rayleigh Fading," in *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006. ACSSC '06.*, 29 2006-nov. 1 2006, pp. 1654 –1657.

[61] Y. Zhang, G. Wang, and M. Amin, "Cooperative Spatial Multiplexing in Multi-Hop Wireless Networks," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.*, vol. 4, may 2006, p. IV.

[62] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 50 –56, july 2009.

[63] C. Wang, T. Yuan, and D. Yang, "Cooperative Relay Network Configuration with Spatial Multiplexing and Beamforming," in *International Conference on Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007.*, sept. 2007, pp. 137 –140.

[64] M. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*, ser. International Series in Engineering and Computer Science. Springer, 2007. [Online]. Available: http://books.google.co.uk/books?id=zjJdP0CJUAYC

[65] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*. The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England: John Wiley and Sons, Inc, 2000.

[66] G. Lindell, "Some exact union bound results for maximum likelihood detection in MIMO systems," in *International Symposium on Information Theory, 2005. ISIT 2005. Proceedings.*, sept. 2005, pp. 2223 –2227.

[67] M. Dohler, "Virtual Antenna Arrays," Ph.D. dissertation, Univ. of London, Kings College London, University of London, Strand, London WC2R 2LS, November 2003. [Online]. Available: http://www.cttc.es/resources/doc/080516-virtual-antenna-arrays-phd-thesis-23653.pdf

[68] M. Dohler, J. Dominguez, and H. Aghvami, "Link capacity analysis for virtual antenna arrays," in *2002 IEEE 56th Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall.*, vol. 1, 2002, pp. 440 – 443 vol.1.

[69] M. Dohler, A. Gkelias, and H. Aghvami, "Resource allocation for FDMA-based regenerative multihop links," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1989 – 1993, nov. 2004.

[70] J. Garcia-Frias, "Compression of correlated binary sources using turbo codes," *IEEE Communications Letters*, vol. 5, no. 10, pp. 417 –419, oct. 2001.

[71] ——, "Joint source-channel decoding of correlated sources over noisy channels," in *Data Compression Conference, 2001. Proceedings. DCC 2001.*, 2001, pp. 283 –292.

[72] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 555 – 559, april 2005.

[73] V. Tervo, T. Matsumoto, and P.-S. Lu, "Distributed Joint Source-Channel Coding for Correlated Sources Using Non-systematic Repeat-Accumulate Based Codes," *Wireless Personal Communications*, pp. 1–15, 2012. [Online]. Available: http://dx.doi.org/10.1007/s11277-012-0579-5

[74] J. Lodge, R. Young, P. Hoeher, and J. Hagenauer, "Separable MAP ldquo;filters rdquo; for the decoding of product and concatenated codes," in *IEEE International Conference on Communications, 1993. ICC '93 Geneva. Technical Program, Conference Record.*, vol. 3, May 1993, pp. 1740–1745 vol.3.

[75] H. Wymeersch, *Iterative Receiver Design*. UPH, Shaftesbury Road, Cambridge, CB2 8BS, United Kingdom: Cambridge University Press, September 2007.

[76] Y. Linde, A. Buzo, and R. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications.*, vol. 28, no. 1, pp. 84 – 95, jan 1980.

[77] K. A. X. He, X. Zhou and T. Matsumoto, "Wireless Mesh Networks Allowing Intra Link Errors: CEO Problem Viewpoint," in *2012 International Symposium on Information Theory and its Applications (ISITA2012)*, 2012, under review.

[78] M. Massaro, "The Distribution of Error Probability for Rayleigh Fading and Gaussian Noise," *IEEE Transactions on Communications.*, vol. 22, no. 11, pp. 1856 – 1858, nov 1974.

[79] A. Papoulis, *Probability, Random Variables And Stochastic Processes - Athanasios Papoulis [3rd Edition]*. New York:McGraw-Hill: Mcgraw-Hill College, February 1, 1991.

[80] I. Glover and P. Grant, *Digital Communications*. Prentice Hall, 2010. [Online]. Available: http://books.google.co.uk/books?id=qaUUCXBSBEwC

[81] G. Arfken, H. Weber, and F. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*. Elsevier, 2012. [Online]. Available: http://books.google.co.uk/books?id=qLFo_Z-PoGIC

[82] S. Zhang and V. K. N. Lau, "Design and Analysis of Multi-Relay Selection for Cooperative Spatial Multiplexing," in *IEEE International Conference on Communications, 2008. ICC '08.*, 2008, pp. 1129–1133.

[83] A. Darmawan, S. Kim, and H. Morikawa, "Amplify-and-Forward Scheme in Cooperative Spatial Multiplexing," in *16th IST Mobile and Wireless Communications Summit, 2007.*, 2007, pp. 1–5.

[84] T. Duong and H. Zepernick, "Performance analysis of cooperative spatial multiplexing with amplify-and-forward relays," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, 2009*, 2009, pp. 1963–1967.

[85] Y. Zhang, G. Wang, and M. Amin, "Cooperative Spatial Multiplexing in Multi-Hop Wireless Networks," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.*, vol. 4, 2006, pp. IV–IV.

[86] S. W. Kim and R. Cherukuri, "Adaptive forwarding and coding in cooperative spatial multiplexing system," in *2005 IEEE 62nd Vehicular Technology Conference, 2005. VTC-2005-Fall.*, vol. 4, 2005, pp. 2418–2422.

[87] K. Lee and L. Hanzo, "Optimal decoding for hard-decision forwarding aided cooperative spatial multiplexing systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4416–441, 2009.

[88] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[89] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory.*, vol. 43, no. 5, pp. 1549–1559, Sep 1997.

[90] N. Xie and A. Burr, "Distributed Cooperative Spatial Multiplexing with Slepian Wolf Code," in *IEEE 77th Vehicular Technology Conference (VTC Spring), 2013*, June 2013, pp. 1–5.

[91] M. Fresia, F. Perez-Cruz, H. Poor, and S. Verdu, "Joint Source and Channel Coding," *IEEE Signal Processing Magazine.*, vol. 27, no. 6, pp. 104–113, Nov 2010.

[92] K. de Bruyn, V. Prelov, and E. Van Der Meulen, "Reliable transmission of two correlated sources over an asymmetric multiple-access channel (Corresp.)," *IEEE Transactions on Information Theory.*, vol. 33, no. 5, pp. 716–718, Sep 1987.

[93] A. Jeyaraj and M. Zarki, "A distributed constant bit rate Slepian-Wolf based multiple access scheme with nodes having common and correlated information," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008.*, Sept 2008, pp. 1–6.

[94] Y. Murin, R. Dabora, and D. Gunduz, "Source-Channel Coding Theorems for the Multiple-Access Relay Channel," *IEEE Transactions on Information Theory.*, vol. 59, no. 9, pp. 5446–5465, Sept 2013.

[95] A. Yedla, H. Pfister, and K. Narayanan, "Code Design for the Noisy Slepian-Wolf Problem," *IEEE Transactions on Communications.*, vol. 61, no. 6, pp. 2535–2545, June 2013.

[96] T. Cover, A. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Transactions on Information Theory.*, vol. 26, no. 6, pp. 648–657, Nov 1980.

[97] (2011) The Mathematica Website. [Online]. Available: http://www.wolfram.com/mathematica/

[98] E. Lukacs and E. P. King, "A Property of the Normal Distribution," *The Annals of Mathematical Statistics*, vol. 25, no. 2, pp. 389–394, 06 1954. [Online]. Available: http://dx.doi.org/10.1214/aoms/1177728796

[99] T. Van de Cruys, "Two multivariate generalizations of pointwise mutual information," in *Proceedings of the Workshop on Distributional Semantics and Compositionality*, ser. DiSCo '11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 16–20. [Online]. Available: http://dl.acm.org/citation.cfm?id=2043121.2043124