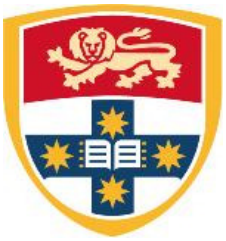
THE UNIVERSITY OF
SYDNEY

Finite Length Analysis of Rateless Codes and Their Application in Wireless Networks

Peng Wang

A thesis submitted in fulfillment of

requirements for the degree of Doctor of Philosophy



School of Electrical & Information Engineering

The University of Sydney

August 2015

Abstract

Mobile communication systems are undergoing revolutionary developments as a result of the rapidly growing demands for high data rates and reliable communication connections. The key features of the next-generation mobile communication systems are provision of high-speed and robust communication links. However, wireless communications still need to address the same challenge–unreliable communication connections, arising from a number of causes including noise, interference, and distortion because of hardware imperfections or physical limitations.

Forwarding error correction (FEC) codes are used to protect source information by adding redundancy. With FEC codes, errors among the transmitted message can be corrected by the receiver. Recent work has shown that, by applying rateless codes (a class of FEC codes), wireless transmission efficiency and reliability can be dramatically improved. Unlike traditional codes, rateless codes can adapt to different channel conditions. Rateless codes have been widely used in many multimedia broadcast/multicast applications. Among the known rateless codes, two types of codes stand out: *Luby transform* (LT) codes and Raptor codes. However, our understanding of LT codes and Raptor codes is still incomplete due to the lack of complete theoretical analysis on the decoding error performance of these codes. Particularly, this thesis focuses on the decoding error performance of these codes under *maximum likelihood* (ML) decoding, which provides a benchmark on the optimum system performance for gauging other decoding schemes. In this thesis, we discuss the effectiveness of rateless codes in

terms of the success probability of decoding. It is defined as the probability that all source symbols can be successfully decoded with a given number of successfully received coded symbols under ML decoding. This thesis provides a detailed mathematical analysis on the rank profile of general LT codes to evaluate the decoding success probability of LT codes under ML decoding. Furthermore, by analyzing the rank of the product of two random coefficient matrices, this thesis derived bounds on the decoding success probability of Raptor codes with a systematic *low-density generator matrix* (LDGM) code as the pre-code under ML decoding.

Additionally, by resorting to stochastic geometry analysis, we develop a LT codes based broadcast scheme. This scheme allows a *base station* (BS) to broadcast a given number of symbols to a large number of users, without user acknowledgment, while being able to provide a performance guarantee on the probability of successful delivery. Further, the BS has limited statistical information about the environment including the spatial distribution of users (instead of their exact locations and number) and the wireless propagation model. Based on the analysis of finite length LT codes and Raptor codes, an upper and a lower bound on the number of transmissions required to meet the performance requirement are obtained.

The technique and analysis developed in this thesis are useful for designing efficient and reliable wireless broadcast strategies. It is of interest to implement rateless codes into modern communication systems.

To my lovely parents, son, and wife:

Qingfa Wang, Hongyu Zhao, Zimo Wang and Lizhu Ouyang

Acknowledgement

There are so many names and faces that come to my mind when I think of people who have had an impact in my professional life and success, those who taught me, encouraged me, and supported me. This long journey would not have been possible at all if I did not have so many great people in my life. I am sure I will leave out the names of many of them here, but all of them will be remembered forever.

I start with my great advisors, Professor Guoqiang Mao and Doctor Zihuai Lin. I feel very fortunate to have them as my thesis advisor, and I thank them a lot for their guidance, encouragement, and support. Not only were they great mentors in my research, but they were also role models of hard work and dedication for me. I thank them for being beside me in all the ups and downs that a PhD student must surely face.

I would like to extend my special thanks to Professor Brian D. O. Anderson. I got the chance to learn about matrix theory from him, and to be inspired by his knowledge and personality. I thank him deeply for being a great mentor and support to me.

I would like to express my gratitude to Doctor Ming Ding, for his prompt response and comments on my second journal paper manuscript, and for his useful suggestions for preparing response letters.

I would like to express my gratitude to Lizhu Ouyang, for having faith in me all the time. Her understanding, her support, her encouragement and her

company are beyond words can ever describe. It is her love that gives me courage to complete the PhD study. I would like to thank my most beloved parents, Qingfa Wang and Hongyu Zhao for their constant and selfless care and love.

Thanks to all my friends for their concern, support and sharing. I thank all the labmates for making the working environment friendly.

Peng Wang

Sydney, NSW, Australia

August 2015

Statement of Originality

I hereby declare that this thesis, submitted in fulfillment of the requirements for the award of Doctor of Philosophy, in the School of Electrical and Information Engineering, the University of Sydney, is my own work unless otherwise referenced or acknowledged. The document has not been previously submitted for the award of any other qualification at any educational institution. Most of the results contained herein have been published, accepted for publication, or submitted for publication, in journals or conferences of international standing. My contribution in terms of published material is listed in "List of publications" chapter.

The original motivation to pursue research in this field was provided by thesis supervisors: Professor Guoqiang Mao from University of Technology Sydney and Dr. Zihuai Lin from University of Sydney.

List of Publications

[J1] P. Wang, G. Mao, Z. Lin, X. Ge, and B. Anderson, "Network coding based wireless broadcast with performance guarantee," IEEE Trans. Wireless Communications, vol.14, no.1, pp.532-544, Jan. 2015.

[J2] P. Wang, G. Mao, Z. Lin, M. Ding, W. Liang, X. Ge, and Z. Lin, "Performance Analysis of Raptor Codes under Maximum-Likelihood (ML) Decoding," accepted by IEEE Transactions on Communications. arXiv:1501.07323 [cs.IT], Jan2015.

[J3] M. Ding, P. Wang, D. L. Perez, G. Mao, and Z. Lin, "Performance Impact of LoS and NLoS Transmissions in Small Cell Networks," accepted by IEEE Transactions on Wireless Communications, arXiv:1503.04251 [cs.IT], Mar. 2015.

[C1] P. Wang, G. Mao, Z. Lin and X. Ge, "An Efficient Network Coding based Broadcast Scheme with Reliability Guarantee", in IEEE ICC, pp. 2879-2884, 2014.

[C2] P. Wang, G. Mao and Z. Lin, "Reliability-Constrained Broadcast using Network Coding without Feedback", IEEE WCNC, pp. 2839-2844, 2014.

[C3] M. Ding, D. L. Perez, G. Mao, P. Wang, and Z. Lin, "Will the Area Spectral Efficiency Monotonically Grow as Small Cells Go Dense?" accepted by IEEE Globecom 2015, arXiv:1505.01920 [cs.NI], May2015.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3GPP	3rd Generation Partnership Project
ARQ	Automatic Repeat reQuest
BEC	Binary Erasion Channel
BP	Belief Propagation
BS	Base Station
DVB	Digital Video Broadcasting
EM	Electromagnetic
FEC	Forward Error Correction
GE	Gaussian Elimination
GF	Galois Field
HARQ	Hybrid ARQ
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPDC	IP Datacast
IPTV	IP Television

LDGM Low-density generator matrix

LDPC Low-density parity-check

LNC Linear Network Coding

LT Luby Transform

MBMS Multimedia broadcast/multicast services

MDS Maximum distance separable

ML Maximum likelihood

NC Network Coding

PHY Physical layer

Raptor Rapid tornado

RS Reed-Solomon

TCP Transport control protocol

XOR exclusive or

Chapter 1

Introduction

This chapter describes the background and motivation for this research work by briefly introducing the field in Section 1.1. Section 1.2 explains the principal research problems, followed by a summary of the main contributions of this thesis. The outline of the thesis is provided in Section 1.3.

1.1 History

In the past century, telecommunication systems have experienced several revolutionary developments to meet the constantly rising demands for high data rates and reliable communication connections. Telecommunication systems can be divided into two categories, wired and wireless telecommunication systems. Among them, many communication channels face the same challenge–unreliable communication connections, arising from a number of causes including noise, interference, and distortion caused by hardware imperfections or physical limitations. Additionally, most applications of the modern telecommunication system can not endure erroneous transmissions.

Over the past years, several means have been proposed to address this vital challenge in telecommunication systems. Conventionally, to ensure reliable delivery of the original data, erroneous data frames or symbols need to be resent. A

renowned retransmission mechanism is *Automatic Repeat reQuest* (ARQ) [3, 4], which uses feedbacks to indicate the correct transmission or erroneous transmission of certain transmitted data frames or symbols. With ARQ, feedbacks are transmitted back to the transmitter after each transmission using either *acknowledgements* (ACKs) if the data frames or symbols are correctly received or *negative acknowledgements* (NACKs) if the data frames or symbols are deemed erroneous. If NACKs are received or ACKs are not received within a predesignated amount of time, the transmitter will retransmit the data frames or symbols. The three basic ARQ protocols are Stop-and-wait ARQ, Go-Back-N ARQ and Selective Repeat ARQ. All three ARQ protocols utilize the sliding window protocol to inform the transmitter which data frames or symbols should be retransmitted. There are also more sophisticated retransmission mechanisms, such as the Type II *hybrid ARQ* (HARQ) protocol [5]. With Type II HARQ, the transmitter will send extra redundancy on the unrecovered data frames or symbols to a particular user, instead of retransmitting the original symbols.

However, several drawbacks appear when using transmission acknowledgment. Firstly, the overhead incurred when gathering acknowledgment information from multiple receivers increases with the number of receivers. In other words, when the number of receivers is large, acknowledgement may cause significant delays and bandwidth consumption [6]. Consequently, using ARQ for wireless broadcast is not scalable [7]. Secondly, for different receivers, distinct and independent errors are often encountered. In such cases, the retransmitted data frames or symbols are only useful to a specific user and with no value for others. Hence, it is highly undesirable to send respective erroneous data frames or symbols to each user.

On the other hand, *forwarding error correction* (FEC) codes are proposed to protect the source information by adding redundancy. With FEC codes, errors among the transmitted message can be corrected by the receiver to recover the

original information. FEC codes have come to pervade every aspect of our lives. They have strongly affected not only the wireless cellular network and satellite communication systems, but also the Internet computer networks and data storage. The pioneering work of Shannon in 1948 [8] broke the ground for FEC codes. In [8], Shannon derived the theoretical limit on the transmission rate over a noisy channel, i.e., the channel capacity. Meanwhile, he introduced digital FEC codes as well, which is also called channel codes. FEC codes are capable of allowing communication with an arbitrarily small probability of error at any rate, as long as it does not exceed the channel capacity.

In the next decades, several FEC codes were proposed, such as Hamming codes [9], convolutional codes [10] and *Reed-Solomon* (RS) codes [11, 12]. These coding techniques were mostly based on the algebraic property. However, there were no FEC codes that could closely approach the theoretical performance limits proposed by Shannon until the invention of Turbo codes. In the 1990s, Berrou et al astoundingly invented turbo codes and their iterative decoder, significantly diminishing the gap to Shannon capacity [13, 14]. With the massive attention drawn by turbo codes, coding theorists redirected their research interests to the field of soft decision iterative decoders and to the search for lower complexity codes. With these efforts, *low-density parity-check* (LDPC) codes were rediscovered in the 1990s [15, 16, 17, 18, 19, 20]. These coding schemes were originally proposed by Galleger in 1963 [21]. However, back at that time, due to the insufficient computing power to implement these codes, their true power was not revealed. Nowadays, the codes mostly approaching the Shannon bound are LDPC codes, and much work has recently focused on their design and analysis.

While all the FEC codes mentioned above are designed for fixed rates, a new class of FEC codes, named rateless (fountain) codes [22], has recently been proposed. As suggested by the name, rateless codes are not designed for any rate, and their design can automatically adapt to any channel condition. Ideally, this

coding ensemble should be able to recover all k source symbols upon the reception of exactly k encoded symbols. During the transmissions, no acknowledgement or at most one feedback per user is needed. More specifically, after successfully decoding all k source symbols based on a certain number of encoded symbols that have been received, each user will send a notification to the transmitter [23]. If the transmitter only requires a statistical reliability guarantee, then no acknowledgement is needed at all. Moreover, rateless codes should be able to generate a potentially limitless stream of m_T encoded symbols out of the k source symbols.

However, the idealized digital fountain is difficult to obtain. In practice, we can only develop such codes with approximate performance. The early designs of FEC codes providing incremental redundancy [24, 25] are on the basis of *maximum distance separable* (MDS) codes [26]. MDS codes can recover a message comprising of k symbols from any set of k out of m_T encoded symbols. Yet, MDS codes do not possess rateless properties and rateless codes are not MDS. Hence, it is inevitable to relax the MDS condition to obtain practical fountain codes [28]. They should be able to recover all the k original symbols from any $k(1+\delta)$ out of the m_T encoded symbols, regardless of which $k(1+\delta)$ encoded symbols have been received, where δ is a small non-negative number.

The first class of practical rateless codes is *Luby transform* (LT) codes [29, 30, 31], which were invented by Michael Luby. LT codes are a class of random linear FEC codes based on irregular sparse graphs and random processes. They are designed to be efficiently decoded with a suboptimal decoding algorithm–*belief propagation* (BP) algorithm [32, 33]. There are other types of rateless codes, such as online codes [35] and Raptor codes [36, 37, 38] whose constructions are a combination of LT codes and one or more stages of high-rate pre-codes. In recent years, Raptor codes have been utilized in several communication standards, e.g., the *3rd Generation Partnership Project* (3GPP) *multimedia broadcast multicast*

services (MBMS) standard [39], *Internet Engineering Task Force* (IETF) *RaptorQ FEC Schemes* (RFCs) 5053 and 6330 [40, 41], the *Digital Video Broadcasting* (DVB) *Internet Protocol Datacast* (IPDC) standard [44] and the DVB *Internet Protocol television* (IPTV) standard [43].

1.2 Research Problems and Contributions in this Thesis

With the facts we have introduced previously in mind, we want to design a coding based broadcast scheme in an unreliable wireless network that a) reliably delivers information to a large number of users, b) does not rely on the user acknowledgment, and c) is able to provide a guaranteed performance on reliability, the probability of successful delivery.

Network coding (NC) has been proved to be an efficient method to significantly improve both the transmission efficiency and the reliability of transmission [45, 46, 47, 48, 49]. Several NC based broadcast schemes have been proposed in [45]. It was shown that NC based retransmission schemes perform better than their counterpart using ARQ only. However, NC based retransmission strategies rely on the use of feedback information from receivers. The drawbacks of feedback have been presented on page 2. In this work, we extend the above NC based broadcast schemes by considering other more suitable coding ensembles.

The fact that rateless codes can automatically adapt to instantaneous channel states and avoid the need for feedback channels [46, 50, 51] makes them desirable means for data transmission over lossy multicast/broadcast channels whose real-time channel erasure probability estimation might be nearly impossible to obtain. Hence, in this thesis, we mainly focus on rateless codes.

1.2.1 Fundamental Problems in Rateless Codes

Despite the successful application of rateless codes in MBMS, limited work exists on theoretically analyzing the decoding performance of rateless codes. Without analytical results, the optimization of the degree distribution as well as the parameters for rateless codes would be extremely difficult, if not impossible. Among the known rateless codes, two codes stand out: LT codes and Raptor codes.

In this thesis, we discuss the effectiveness of rateless codes in terms of the success probability of decoding. The decoding success probability is defined as the probability that a receiver can successfully decode all k source symbols given that the receiver has successfully received m_R coded symbols. However, the decoding success probability of LT codes is difficult to analyze. Since 2004 [52], coding theorists have been analyzing the decoding success probability of LT code under BP decoding. In [52], a 3-dimension state was utilized to describe the procedure of the LT decoding with BP decoding. Each state is a combination of three parameters: firstly, the number r of output symbols with degree 1 (i.e., the ripple size); secondly, the number c of symbols with degree two and above (i.e., the number of symbols in the cloud); and finally, the number u of unrecovered source symbols. Let $P(r, c, u)$ represent the probability that the LT decoding process is in state (r, c, u). The decoding is deemed to fail if it encounters a state where $u > 0$ and $r = 0$. With the degree distribution of LT codes $\Omega_d, 1 \leq d \leq k$, k source symbols and m_R received coded symbols, the decoding success probability can be expressed as

$$P_{k,m_R}^{DS} = 1 - \sum_{0<u<k,r=0} P(r, c, u). \tag{1.1}$$

In [52], the authors proposed a method as shown in (1.2) to compute the state

probability of the LT decoding procedure under BP decoding.

$$\begin{aligned} P(r,c,u-1) &= \sum_{s,t\geq 0, r\geq t-s} P(r+1+s-t, c+t, u) \\ &\quad \times \Pr[(r,c,u-1) \mid (r+1+s-t, c+t, u)]. \end{aligned} \tag{1.2}$$

This equation calculates the state probability $(r, c, u-1)$ by using the total probability theory. $\Pr[(r,c,u-1) \mid (r+1+s-t, c+t, u)]$ is the conditional probability which denotes the transition behavior from states $(r+1+s-t, c+t, u)$ to state $(r, c, u-1)$ where $t-s \leq r$ and $s, t \geq 0$. This condition event can be seen as given the state $(r+1+s-t, c+t, u)$, the LT decoder selects an output symbol from the ripple. This causes degrees of t cloud symbols to reduce to one and join the ripple. Meanwhile, among the $r+1+s-t$ symbols in the ripple s symbols are duplications of the selected symbol. This conditional probability can be calculated by

$$\begin{aligned} &\Pr[(r,c,u-1) \mid (r+1+s-t, c+t, u)] \\ &= \binom{c+t}{t} p_u^t (1-p_u)^c \binom{r+s-t}{s} \left(\frac{1}{u}\right)^s \left(1-\frac{1}{u}\right)^{r-t}, \end{aligned} \tag{1.3}$$

where p_u denotes the probability that a random output symbol is of reduced degree 1 after transition given that it was of reduced degree ≥ 2 before the transition. p_u can be expressed as

$$p_u = \frac{\sum_d \Omega_d \cdot d \cdot (d-1) \cdot \frac{(u-1)u(k-u)\cdots(k-u-d+3)}{k(k-1)\cdots(k-d+1)}}{1 - \sum_d \Omega_d \cdot \frac{(k-u)\cdots(k-u-d+1)}{k(k-1)\cdots(k-d+1)} - \sum_d \Omega_d \cdot d \cdot \frac{u(k-u)\cdots(k-u-d+2)}{k(k-1)\cdots(k-d+1)}}. \tag{1.4}$$

Details of the derivation and proof can be found in [53]. The recursion involved in the computation makes it very difficult to derive a closed-form analytical result for the decoding success probability.

In general, the BP decoding algorithm is widely used for rateless codes due to its low complexity. However, when the number of source symbols decreases,

the decoding error performance of the BP decoding algorithm suffers serious degradation. The *maximum likelihood* (ML) decoding algorithm, on the other hand, is more computationally demanding than the BP decoding for codes with a large length. Nevertheless, the ML decoding algorithm becomes affordable complexity-wise and at the same time almost imperative performance-wise for small to medium sizes. For rateless codes with limited lengths, i.e., in the order of a few thousands, Shokrollahi proposed a decoding algorithm based on the ML criterion in [54]. This will be the decoding method of choice in this thesis. Furthermore, since ML decoding is optimal in terms of the decoding error performance, the ML performance of a code can provide a benchmark on the optimum system performance for gauging the other decoding schemes.

It is worth noting that in [55, 56], a theoretical analysis was conducted on the decoding success probability of LT codes under ML decoding. However the analysis in [55] was incomplete to the extent that no rigorous analysis was presented to support some results presented in it. Furthermore, the analytical result presented on the decoding success probability was in fact an approximation only, which will be discussed in further detail in Section 3.3. In Chapter 3, we advance the work in [52, 53, 55] by providing rigorous mathematical analysis on the rank profile of a random matrix. On the basis of this analysis, we obtain the upper and lower bounds on the decoding success probability of LT codes under ML decoding.

As for Raptor codes, in [37], Shokrollahi analyzed the decoding failure probability of Raptor codes with finite length assuming the BP decoding. The analysis relies on the computation of the failure probability of the LT codes under the BP decoding, which was derived in [52]. Furthermore, in [57] a pseudo upper bound on the performance of Raptor codes under ML decoding was derived under the assumption that the number of erasures correctable by the pre-code is small. This approximation is accurate only when the rate of the pre-code is sufficiently high.

So for a more general case, the decoding failure probability of Raptor codes still needs further investigation.

1.2.2 Thesis Contributions

The main objectives of this thesis are the theoretical analysis of various types of rateless code ensembles with finite message lengths under optimal erasure decoding, i.e., *maximum likelihood* (ML) decoding. In Chapter 3, we conduct the finite length analysis of LT code ensembles under ML decoding in terms of decoding success probability. The decoding success probability of LT code is defined as the probability that a receiver can successfully decode all k source symbols given that the receiver has successfully received $m_R \geq k$ coded symbols. Specifically, if erasure channels are considered, the decoding of LT codes under ML decoding corresponds to solving a consistent system of linear equations over a binary field $GF(2)$, where the coefficients are given by the collected LT code generator matrix. Chapter 3 provides a rigorous mathematical analysis on the rank profile of a random coefficient matrix, where each row vector is independently generated by using the LT encoding process. A set of two bounds, consisting of upper and lower bounds on the decoding success probability after optimal decoding, is derived in detail. Furthermore, when binomial degree distribution introduced in Subsection 2.3.1 is applied, the upper and lower bounds merge to the exact expression. These analytical bounds are used to assess the performance of LT code ensembles or to design them efficiently without requiring extensive Monte Carlo simulations.

In Chapter 4, we provide the analytical results, i.e., an upper bound and a lower bound, on the decoding failure probability of finite length Raptor codes with a systematic *low-density generator matrix* (LDGM) code as the pre-code under ML decoding. The decoding failure probability is defined as the probability that not all k source symbols can be successfully recovered by a receiver with

a given number $m_R \geq k$ of successfully received coded symbols. The analytical results are derived by analyzing the rank of the product of two random coefficient matrices. Based on the analytical bounds on the decoding failure probability of Raptor codes, we can readily obtain analytical bounds on the decoding success probability of Raptor codes, which is unity minus decoding failure probability. Moreover, simulations are conducted to validate the accuracy of the proposed bounds. Finally, by applying binomial degree distribution into the upper and lower bounds, we simplified the general bounds with any degree distributions and any (n, k, η) LDGM codes as pre-code into a far less complex expressions. By this way, the computation complexity of derived bounds can be significantly decreased.

Furthermore, we investigate the problem of reliable and efficient broadcasting in wireless networks. The goal is to deliver a large given number of data symbols to a large number of users, without user acknowledgment, while being able to provide a performance guarantee on the probability of successful delivery. Further, the BS only has limited statistical information about the environment including the spatial distribution of users (instead of their exact locations and number) and the wireless propagation model. Our approach to tackle this problem is based on utilizing rateless codes and stochastic geometry analysis. On the basis of derived bounds on the decoding success probability of LT codes, an upper and a lower bound on the probability that all receivers in a bounded area successfully receive or decode all source symbols from the BS are derived. On the basis of the above results, the minimum number of transmissions required for a guaranteed performance on the probability of successful delivery is obtained.

1.3 Thesis Outline

The rest of the thesis is organized as follows. In Chapter 2, we briefly present the necessary background on which the thesis is based. Chapters 3, 4 and 5 comprise the major contributions of this thesis in which we investigate the decoding performance of rateless codes and its application into the wireless broadcast problems, respectively. In Chapter 6, we conclude this thesis.

Parts of the present thesis have been prepublished in the following papers which I have authored: [J1, J2, C1, C2].

Chapter 2

Background

In this chapter, we briefly present the necessary background on which the thesis is based. We begin by introducing the binary erasure channel model. Then, we provide background on network coding and related previous works. Next, we review rateless codes and important developments in these areas. Finally, we review the problem of efficient data broadcasting in wireless networks.

2.1 Binary Erasure Channel

The *binary erasure channel* (BEC), a widely used communication channel model in coding theory, was originally proposed by Elias in 1955 [10] as a simplified theoretical model. After 4 decades, due to the emergence of the Internet, the BEC model became a realistic one. Indeed, links of data networks can be modeled as erasure channels, where data is transmitted in the formation of symbols. In data networks, symbols are either received correctly or lost for certain reasons. The BEC is characterized by a parameter ε, the channel erasure probability. Specifically, a symbol is either successfully received or erased with probability $1 - \varepsilon$ or ε, respectively. Figure 2.1 depicts the BEC model. Practically, the instantaneous state of a wireless channel is difficult to obtain but the stochastic property can be obtained with less effort. In this thesis, we mostly consider

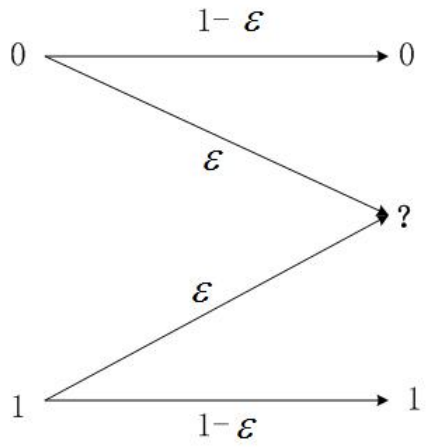


Figure 2.1: Binary erasure channel.

wireless channels as erasure channels, where the transmission is successful with a certain probability.

2.2 Network Coding

In [58], Ahlswede et al. proposed the concept of *network coding* (NC) to improve the flow of data in a network by allowing intermediate nodes to combine incoming data flows into an outgoing data flow. Recent work has shown that NC can significantly improve both the transmission efficiency and the reliability of transmission [45, 46, 47, 48, 49]. Initially, NC technique is proposed to be applied at network layer. However, its extensive benefits pushed researchers to apply it at other protocol layers. The concept of *physical-layer NC* (PNC) was originally proposed to exploit the network coding operation that occurs naturally in superimposed electromagnetic (EM) waves. In this thesis, we only consider network layer NC, where the data is transmitted in digital format. Several classes of network layer NC schemes are explained as follows.

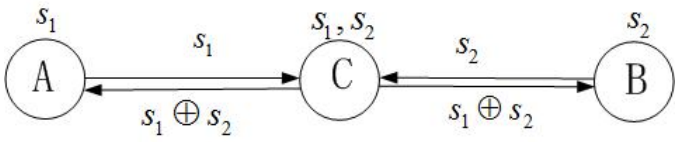


Figure 2.2: The classic two-way relaying network applying XOR coding.

2.2.1 XOR Based Network Coding

The basic idea of XOR based NC schemes is that a node encodes all or a certain set of symbols with bitwise XOR. For instance, nodes A and B exchange symbols s_1 and s_2 via a relay C, as shown in Figure 2.2. Initially, node A has the source symbol s_1 while node B has the source symbol s_2.

Both nodes deliver their source symbols to the relay node C respectively in the first step. Then, the relay node C performs XOR coding between the received symbols s_1 and s_2 to generate the coded symbols $s_1 \oplus s_2$. Then, C transmits the coded symbol $s_1 \oplus s_2$ in one transmission rather than two source symbols separately. Finally, for A, the intended symbol s_2 can be recovered by conducting $(s_1 \oplus s_2) \oplus s_1$. For node B, a similar decoding process can be done as well.

An application of XOR based NC schemes is COPE [7], which is the first practical NC scheme for wireless mesh networks.

2.2.2 Linear Network Coding

It has been proved that *linear NC* (LNC) can achieve capacity limit from the source node to each destination node in multicast networks [59]. The capacity limit is given by the max-flow min-cut bound [60]. More specifically, the maximum amount of data flows from a source node to a destination node that can pass through the network is equal to the min-cut between them [58, 61, 62, 63]. In NC, the butterfly network [58] is often used to illustrate how LNC can outperform routing. Each node generates new symbols which are linear combinations of earlier received symbols, multiplying them by coefficients chosen from a finite

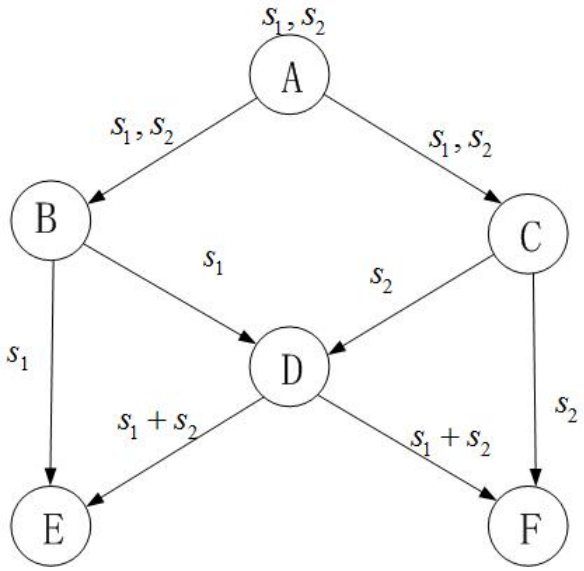


Figure 2.3: Butterfly network.

field, say Galois field $GF(q)$, where q is an arbitrary positive integer. As shown in Figure 2.3, assume that the selected finite field is $GF(2)$ and we want to broadcast two symbols s_1 and s_2 from a source A to all the nodes in the wireless network. The optimal solution for this model is provided in Figure 2.3 as well. In this scenario, node A broadcasts s_1 and s_2. Node B broadcasts only s_1, and node C broadcasts only s_2. Node D receives s_1 and s_2 and broadcasts $s_1 + s_2$. Clearly, node E can recover s_1 and s_2 by receiving both s_1 and $s_1 + s_2$. Similarly, node F can recover s_1 and s_2 by receiving s_2 and $s_1 + s_2$. Hence, $s_1 + s_2$ is beneficial to both nodes E and F. In this case, the optimal broadcasting happens with only five transmissions. In summary, local coding in the network can reduce the number of transmissions and can offer the network a better energy efficiency.

2.2.3 Network Coding Based Broadcast Schemes

In [45], NC was applied to one hop wireless broadcast problem and several NC based broadcast schemes were proposed. Here we use an example to demonstrate the NC based broadcast schemes as shown in Figure 2.4. As can be seen in Figure 2.4, the source node S broadcasts symbols P_1, P_2 and P_3 to destination

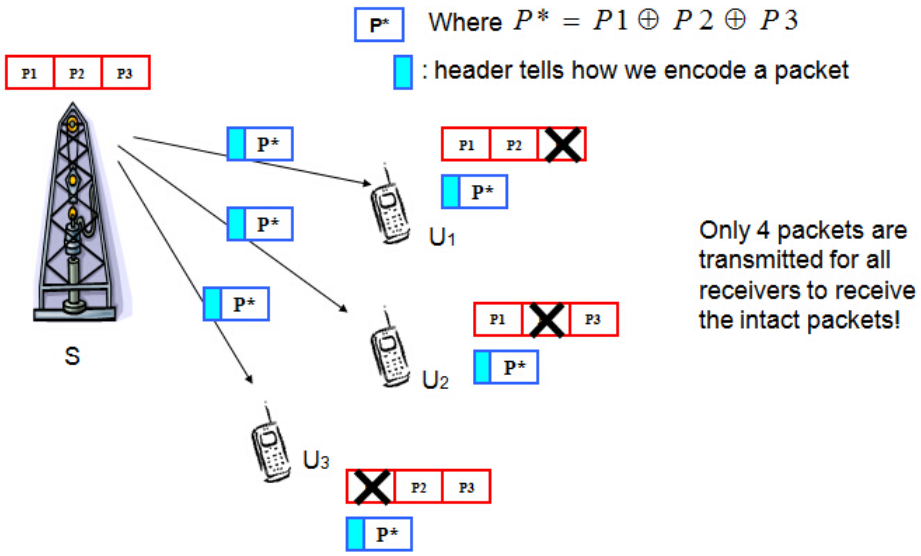


Figure 2.4: NC based broadcast schemes.

nodes U_1, U_2 and U_3. After broadcasting three original symbols P_1, P_2 and P_3, each destination node has its unique error pattern. Instead of retransmitting the original symbols that have been deemed as erroneous by destination nodes, node S broadcasts the coded symbols $P^* = P_1 \oplus P_2 \oplus P_3$. Clearly, nodes U_1, U_2 and U_3 can recover symbols that have been deemed as erroneous, respectively. Hence, 4 symbols are transmitted for all nodes to recover the intact original symbols. It was shown in [45] that NC based retransmission schemes perform better than their counterpart using *Automatic Repeat reQuest* (ARQ) only [3]. However, the NC based retransmission strategies rely on the use of feedback information from receivers. In this thesis, we want to design a coding based broadcast scheme in an unreliable wireless network that does not rely on user acknowledgment. One of the best coding schemes is rateless erasure coding [31, 37, 64, 65]. Unlike traditional codes, rateless codes are adaptable to different channel conditions and avoid the need for feedback channels [46, 50, 51].

2.3 Rateless Codes

Rateless (Digital fountain) codes are a new class of forwarding error correction (FEC) codes. They were first characterized in [22], where no actual coding construction was proposed but some application scenarios were suggested. The reason why this type of coding ensembles is named "digital fountain" is because of the similarities between a water fountain filling a cup and the original message being able to be recovered. More specifically, a water fountain which can be seen as an unlimited waterdrops can fill a cup by gathering a sufficient number of waterdrops. Similarly, the original message is able to be successfully decoded after collecting a sufficient number of encoded symbols. Hence, this thesis will utilize the terms "fountain code" and "rateless code" synonymously. Initially, rateless codes were invented for the BEC as a replacement for retransmission schemes such as ARQ [4] to combat the challenge of unreliable transmission.

Rateless codes have been widely used in the broadcast/multicast application, a scenario in which the information with common interest is broadcasted by a transmitter to multiple users spontaneously and in which the users experience various channel states and distinct losses. Specifically, rateless codes can generate a potentially limitless stream of coded symbols. A sufficient number of successfully received coded symbols can lead to successfully decoding of all k source symbols with high probability and this sufficient number can be slightly more than k.

In this thesis, applying rateless codes in the broadcast scenario is considered. To clearly distinguish quantities related to transmitters and quantities related to receivers in this thesis, the symbols "T" and "R" are used.

2.3.1 Luby Transform (LT) Codes

The first class of practical rateless codes is *Luby transform* (LT) codes [29, 30, 31], which were invented by Luby. In LT codes, the source symbol length can be

arbitrary. A symbol consists of l $GF(2)$-elements. However, this number l does not affect the decoding error performance of a fixed but arbitrary code [31]. Therefore, $l = 1$ is assumed throughout this thesis. To transmit a traffic session containing k source symbols, each coded symbol is independently generated by the transmitter, and the entire session can be recovered from any $m_R = k + O(k \log(k/\delta))$ coded symbols with a probability of $1 - \delta$, where δ is a small positive constant.

LT Codes Construction and the Degree Distribution

The encoding process of an LT code is a linear map $GF(2)^k \rightarrow GF(2)^{m_T}$ and is represented by an $m_T \times k$ generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_T \times k}$ over $GF(2)$, i.e., $\mathbf{G}^{\mathrm{LT}}_{m_T \times k} \in GF(2)^{m_T \times k}$, where $m_T \geq k$. The k source symbols $\mathbf{s} = (s_1, \ldots, s_k) \in GF(2)^k$ are mapped to m_T coded symbols $\mathbf{y} = (y_1, \ldots, y_{m_T}) \in GF(2)^{m_T}$ by

$$\mathbf{G}^{\mathrm{LT}}_{m_T \times k} \mathbf{s}^T_{k \times 1} \quad = \quad \mathbf{y}^T_{m_T \times 1}. \tag{2.1}$$

Contrary to traditional block codes, the matrix $\mathbf{G}^{\mathrm{LT}}_{m_T \times k}$ is generated online and can differ for each data traffic session. After the transmissions, a certain receiver can correctly receive m_R coded symbols. The LT code generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ describes the edges of a bipartite graph [27] that link the input nodes to the output nodes. The input nodes represent source symbols and the output nodes represent coded symbols. Figure 2.5 depicts an example of an LT code, where $k = 5$ and $m_R = 6$. Circular nodes represent the source symbols, and the rectangular nodes correspond to the received coded symbols. The decoder is assumed to know all the connections between each correctly received coded and source symbol, i.e. the generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is known by the receiver. This can be achieved by gathering the coding information contained in the head of the coded symbols to produce $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$.

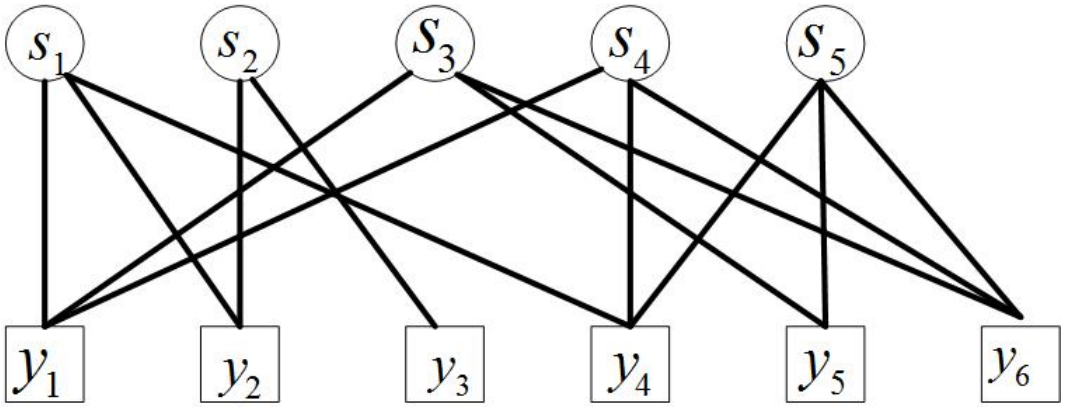


Figure 2.5: An example of an LT code, where $k = 5$ and $m_R = 6$.

The decoding error performance of LT codes mainly relies on the probability mass function (pmf) on the degree of output nodes, which is also called degree distribution. This degree distribution $\Omega_d, d \in \{1, ..., k\}$ is defined as the probability that a coded symbol links to d distinct source symbols, chosen uniformly at random. And $\sum_{d=1}^{k} \Omega_d = 1$. Generally, the degree distribution is expressed in terms of a generator polynomial

$$\Omega(x) \quad = \quad \sum_{d=1}^{k} \Omega_d x^d. \tag{2.2}$$

In the transmitter's generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_T \times k}$ and the receiver's generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$, the d non-zero entries in a row correspond to the d connections between a coded symbol and d source symbols. The value of the coded symbol is determined by the summation of the connected d source symbols over $GF(2)$.

Decoding Algorithms

For BECs, there are two distinct decoding algorithms: the *belief propagation* (BP) decoding algorithm [32] which is efficient but suboptimal [55, 56] and the *maximum-likelihood* (ML) decoding algorithm [67] which is optimal but computationally more demanding. These two decoding algorithms will be briefly introduced as follows.

Belief Propagation Decoding Under BECs, the BP decoding algorithm is also known as LT process or peeling decoding [31, 34]. The BP decoding algorithm can be best explained by using the decoding graph, i.e. the bipartite graph that represents the relationship between the input nodes and the output nodes. A step-by-step example of BP decoding of an LT code over $GF(2)$ can be found in Figure 2.6, which demonstrates the decoding process of the BP algorithm in detail. The BP algorithm can be expressed as follows [32, 33].

1. At least one output node of degree 1 needs to be found to start the decoding process. If none can be found, the decoding process fails and additional output nodes need to be collected to restart the decoding process.

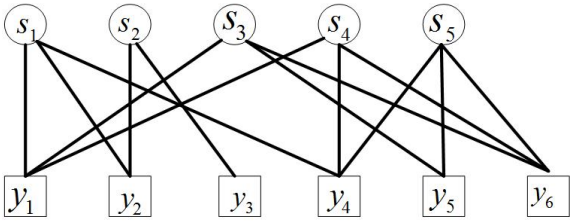
2. Select one output node of degree 1 and disseminate the value of the selected output node to the linked input node.

3. Remove the used output node and its edge from the decoding graph.

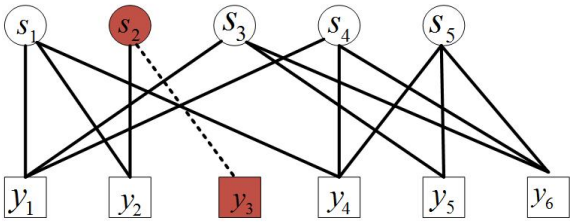
4. Disseminate the value of the recovered input node to all linked output nodes. These output nodes add the value of the recovered input node to there value over $GF(2)$.

5. If all input nodes have been decoded, the decoding process ends successfully. If there still exist unrecovered input nodes, continue with step 2.

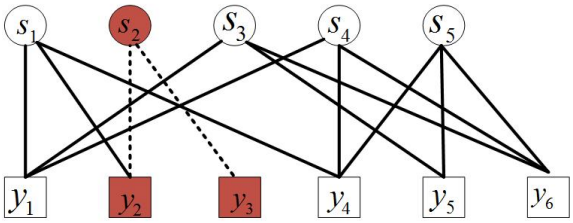
For a large number of input nodes, the suboptimal BP algorithm has excellent performance. Nevertheless, for a small to medium number of input nodes, the decoding process frequently fails due to the lack of output nodes of degree 1. In such cases, more additional output nodes are required for successful decoding. Therefore, for small to medium input sizes, the ML decoding algorithm is a desirable choice.



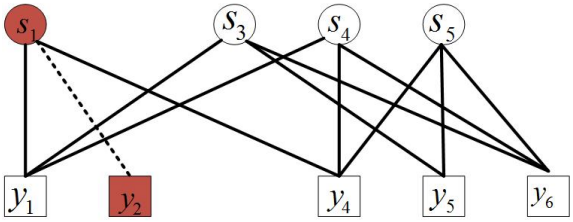
(a) Encoding Graph



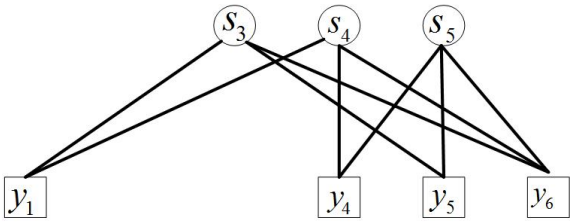
(b) Find an output node of degree 1 and disseminate its value to the linked input node.



(c) Decoded input node disseminate its value to all linked output nodes. Remove the used edges



(d) Find another output node of degree 1 and repeat procedure in step (b)



(e) Repeat procedure in step (c). Since no degree 1 output node has been created in the decoding process, the decoding process fails and additional output nodes need to be collected

Figure 2.6: Exemplary belief propagation decoding of an LT code

Maximum Likelihood Decoding The ML decoding algorithm is the optimal decoding algorithm in terms of decoding success probability. Over BECs, ML decoding of LT codes corresponds to solving a system of m_R consistent linear equations in k unknowns over a binary field $GF(2)$. If the generator matrix $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ has full column rank, i.e. $rank(\mathbf{G}^{\mathrm{LT}}_{m_R \times k}) = k$, all k source symbols can be uniquely determined. If $rank(\mathbf{G}^{\mathrm{LT}}_{m_R \times k}) < k$, the solution of $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}\mathbf{s}^{T}_{k \times 1} = \mathbf{y}^{T}_{m_R \times 1}$ spans a $(k - rank(\mathbf{G}^{\mathrm{LT}}_{m_R \times k}))$-dimensional vector space.

We can solve such a problem with the *Gaussian elimination* (GE) algorithm. Due to the relatively high computational complexity of GE, ML decoding is practically only suitable for codes with small to medium input sizes. In this thesis, we focus on the finite length analysis of LT codes and Raptor codes under ML decoding.

Efficient Maximum Likelihood Decoding Algorithms Apart from GE, there are a number of other algorithms (e.g. [67, 68, 69, 70, 71, 72, 73, 74, 75]) that achieve the ML erasure correction performance and meanwhile decrease the computational complexity. For instance, a distinct decoding algorithm has been proposed in [75], which is called inactivation decoding. Its basic idea is the consecutive use of BP decoding followed by ML decoding. More specifically, if degree 1 output nodes exist, the BP decoding algorithm will be firstly used. When the BP decoding algorithm fails to find an output node of degree 1, an unrecovered input node will be selected and declared as inactivated. Then, the inactivated input node is seen as recovered, and the decoding process continues. The values of the inactivated input nodes are recovered at the end using GE on a matrix where the number of rows and columns are roughly equal to the number of inactivations. Although this concatenation of BP and ML decoding may not be the fastest algorithm, it is the method of choice in this thesis. So far, in telecommunication standards such as the *the 3rd Generation Partnership Project* (3GPP) *Multime-*

dia Broadcast/Multicast Service (MBMS) [39], Qualcomm's Raptor10^{TM} [40] and RaptorQ^{TM} [41, 42, 76] codes are used. Meanwhile, the previously mentioned inactivation decoding [75] is implemented. A detailed explanation of inactivation decoding can be found in [75, 76].

Special Degree Distributions

Given a certain number m_R of coded symbols, the number k of source symbols, the binary field $GF(2)$ and the above mentioned decoding algorithms, the degree distribution $\Omega(x)$ is the only factor that affects the decoding error performance of an LT code.

For BP decoding, several degree distributions have been proposed whose objectives are to optimize the size of the so-called ripple. The ripple is defined as the set of output nodes with degree 1 during the BP decoding process. Since decoding failure is caused by the ripple running empty, it is extremely important to ensure that the ripple size stays non-empty throughout the whole decoding process. On the other hand, over-sized ripples are undesirable and should be avoided as well. These degree distributions are introduced as follows.

The Soliton Distributions In [31], Luby proposed two degree distributions. One is ideal soliton distribution

$$\Omega_d = \begin{cases} \frac{1}{k} & \text{if } d = 1 \\ \\ \frac{1}{d(d-1)} & \text{if } 2 \leq d \leq k, \end{cases} \tag{2.3}$$

which can theoretically achieve the expected ripple size of one. However, the actual ripple size in practice fluctuates around the expected ripple size. It is highly likely that, during the decoding process, ripple sizes become empty before all source symbols have been recovered. In such cases, the decoding process fails.

The other is robust soliton distribution, which is a more advanced version of

the ideal soliton distribution in terms of stability due to its higher expected ripple size. The robust soliton distribution is defined as follows. Let $L = c\log(k/\delta)\sqrt{k}$ for some suitable constants $c, \delta > 0$. Define that

$$\tau(i) = \begin{cases} \frac{L}{ik}, & \text{if } 1 \leq i \leq \frac{k}{L} - 1 \\ L\log\left(\frac{L}{\delta}\right), & \text{if } i = \frac{k}{L} \\ 0 & \text{if } \frac{k}{L} + 1 \leq i \leq k, \end{cases} \tag{2.4}$$

and

$$\rho(i) = \begin{cases} \frac{1}{k} & \text{if } i = 1 \\ \frac{1}{i(i-1)} & \text{if } 2 \leq i \leq k. \end{cases} \tag{2.5}$$

Adding the ideal soliton distribution $\rho(i)$ to $\tau(i)$, the degree distribution Ω_d is obtained by applying normalization, that is

$$\Omega_d = \frac{\rho(i) + \tau(i)}{\beta}, \tag{2.6}$$

where $\beta = \sum_{i=1}^{k} [\rho(i) + \tau(i)]$.

A Degree Distribution Optimized for BP Decoding In [37], Shokrollahi proposed a degree distribution for precoded LT codes, i.e. Raptor codes, for BP decoding, which is expressed as

$$\begin{aligned} \Omega(x) &= 0.007969x + 0.49357x^2 + 0.16622x^3 + 0.072646x^4 \\ &+ 0.082558x^5 + 0.056058x^8 + 0.037229x^9 \\ &+ 0.05559x^{19} + 0.025023x^{65} + 0.003135x^{66}. \end{aligned} \tag{2.7}$$

This degree distribution is obtained by optimizing degree distribution for precoded LT codes under BP decoding with a semi-heuristic method. Moreover,

this degree distribution has been frequently used as a reference to illustrate the accuracy of the derived bound in literature [66, 77].

A Degree Distribution of Standardized Raptor codes In 3GPP standard [39], Raptor codes have been standardized for MBMS. The degree distribution of the LT codes is set as

$$\begin{aligned}\Omega(x) &= 0.0099x + 0.4663x^2 + 0.2144x^3 + 0.1152x^4 \\ &+ 0.1131x^{10} + 0.0811x^{11}. \end{aligned} \tag{2.8}$$

This degree distribution is frequently utilized as a reference in this thesis to verify the accuracy of the analytical bounds developed in this thesis. Meanwhile, it is used to compare with degree distributions that are more suitable for ML decoding.

The Standard and the Sparse Random Ensemble and the Expurgated Random Ensembles In [57, 79, 80, 81, 82, 83], Schotsch summarized the random matrices that are constructed by an entry-wise independent random process, i.e. a Bernoulli process. The standard random ensemble [57] is generated such that each entry in the matrix is chosen independently and uniformly at random from $GF(2)$. The degree distribution of the standard random ensemble can be expressed as

$$\begin{aligned}\Omega(x) &= \sum_{d=0}^{k} \binom{k}{d} \left(\frac{1}{2}\right)^d \left(\frac{1}{2}\right)^{k-d} x^d \\ &= \left(\frac{1}{2}\right)^k \sum_{d=0}^{k} \binom{k}{d} x^d. \end{aligned} \tag{2.9}$$

The sparse random ensemble [57] is created by adjusting the probability that each entry samples to be zero. For an element $[\mathbf{G}^{\mathrm{LT}}_{m_T \times k}]_{i,j}$ in $\mathbf{G}^{\mathrm{LT}}_{m_T \times k}$, we define

this probability as

$$P_0 \triangleq \Pr\left[[\mathbf{G}^{\mathrm{LT}}_{m_T\times k}]_{i,j}=0\right]. \tag{2.10}$$

The degree distribution of the sparse random ensemble is thus

$$\Omega(x) = \sum_{d=0}^{k}\binom{k}{d}(P_0)^{k-d}(1-P_0)^d x^d. \tag{2.11}$$

In the above two random ensembles, Ω_0 is apparently not zero. As coded symbols do not encode any source symbol are redundant and should be avoided. By setting $\Omega_0 = 0$ and normalizing all other probabilities, the degree distribution of expurgated random ensembles can be obtained. The degree distribution of the expurgated standard random ensemble, also named binomial degree distribution, can be shown as

$$\begin{aligned}\Omega(x) &= \frac{1}{1-\left(\frac{1}{2}\right)^k}\sum_{d=1}^{k}\binom{k}{d}\left(\frac{1}{2}\right)^d\left(\frac{1}{2}\right)^{k-d}x^d \\ &= \frac{1}{2^k-1}\sum_{d=0}^{k}\binom{k}{d}x^d.\end{aligned} \tag{2.12}$$

For the expurgated sparse random ensemble, the degree distribution can be expressed as

$$\Omega(x) = \frac{1}{1-(P_0)^k}\sum_{d=1}^{k}\binom{k}{d}(P_0)^{k-d}(1-P_0)^d x^d, \tag{2.13}$$

where P_0 is the probability of sampling a non-zero element prior to setting $\Omega_0 = 0$.

In terms of decoding error performance, both just introduced expurgated random ensembles have excellent performance under ML decoding [57]. Although there is no rigorous theoretical proof, the binomial degree distribution is still generally considered to be the optimal degree distribution for LT codes under ML decoding [57]. Its excellent performance has been verified in [57, 65, 64] by means

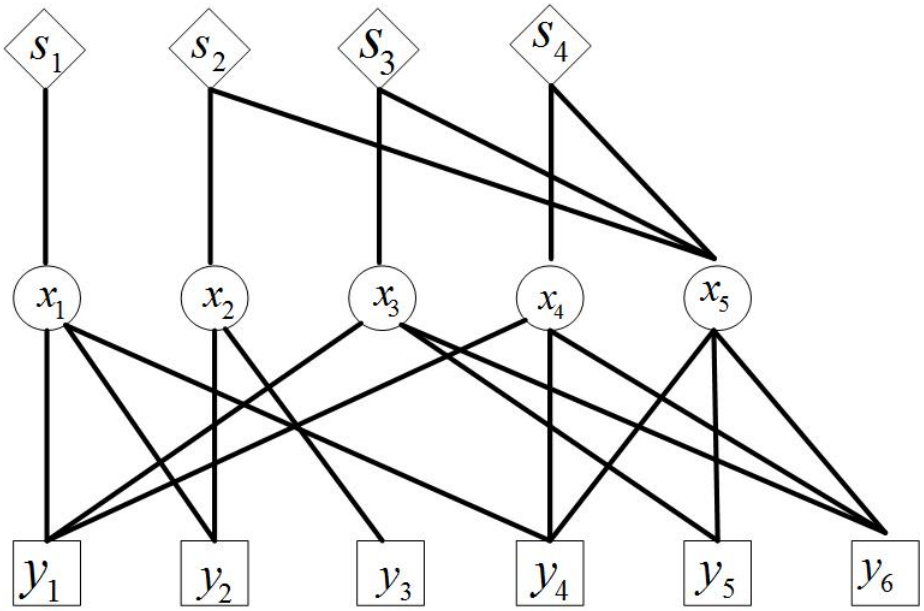


Figure 2.7: An example of a Raptor code with a systematic pre-code, where $k = 4$, $n = 5$ and $m_R = 6$.

of Monte Carlo simulations and tight performance bounds. Further details will be provided in Chapters 3 and 4.

2.3.2 Raptor Codes

Raptor codes are concatenated codes [37], which combine traditional FEC codes with LT codes. They can relax the condition that all input (source) symbols need to be recovered in an LT decoder. The name Raptor is a portmanteau word made of rapid and Tornado. In [84], classical Tornado codes are proposed, which are a class of erasure-resilient codes based on irregular bipartite graphs. An example of a Raptor code encoding graph is depicted in Figure 2.7. Rhombus nodes denote the source symbols, circular nodes represent the intermediate symbols, and the rectangular nodes correspond to the coded symbols. The source symbols $\mathbf{s} = (s_1, \ldots, s_4) \in GF(2)^4$ are first encoded with a pre-code, such as a Hamming code, an LDPC code or an LDGM code. In this way, the intermediate symbols $\mathbf{x} = (x_1, \ldots, x_5) \in GF(2)^5$ are generated. Then, we encode intermediate symbols with an LT code to get the final coded symbols. In such case, even though the

LT decoder cannot recover all intermediate symbols, all source symbols can still be successfully decoded with high probability.

The reason to develop Raptor codes is that LT codes usually have a rather limited performance. The limitation appears in terms of a high error floor, since LT code ensembles include some codewords with very few non-zero entries. However, a pre-code can dramatically improve the decoding error performance by lowering the error floor [37, 57]. Therefore, for BP decoding, LT codes are not intended to be used stand-alone but only in combination with a pre-code. Note that Raptor codes have already been standardized in 3GPP to efficiently disseminate data over a broadcast/multicast network to provide MBMS [39].

2.4 Data Broadcast in Wireless Networks

Broadcasting has been widely used in wireless networks to disseminate information of common interest, e.g. safety warning messages, emergency information and weather information, to a large number of users [1, 2]. There are two major challenges in wireless broadcast. The first one is the unreliable nature of wireless communications. The second one is acknowledging the correct reception of every broadcast symbol by every receiver, particularly when the number of receivers is large.

Due to the unreliable nature of wireless communications, qualities of wireless links often vary temporally and spatially. ARQ is a common solution to combat the challenge of unreliable wireless communications. The drawbacks of transmission acknowledgment have been summarized on page 2. Moreover, the instantaneous state of a wireless channel is difficult to obtain. This is particularly true for highly dynamic networks where the user population and the users' locations change dramatically with time. Take vehicular networks as an example, due to the mobility of vehicles, it is difficult to obtain the exact location of each

vehicle and the exact channel state of each vehicle-*base station* (BS) channel. But the density of vehicles at a particular time period of a day can typically be obtained with much less effort. Therefore, it is highly desirable to design a wireless broadcast scheme that a) uses minimal information about network environment, not relying on information such as the exact number of receivers, the exact location of each receiver and the channel state of each receiver-BS channel, b) reliably delivers information to a large number of users, c) does not rely on user acknowledgment, and d) is able to provide a guaranteed performance on the probability of successful delivery.

In this thesis we tackle the above challenges by resorting to the NC technique [7, 45, 85] and stochastic geometry analysis. NC based broadcast schemes have been introduced in Subsection 2.2.3. However, their NC based retransmission strategy relies on the use of feedback information from receivers. Other coding techniques can be implemented at BS to meet the requirements mentioned above. One of the most suitable options is rateless (Fountain) erasure coding [31, 37, 64, 65]. The numerous advantages of rateless codes have been demonstrated in Section 2.3. In this thesis we develop a random network coding (rateless erasure coding) based broadcast scheme. This scheme allows a BS to broadcast a given number of symbols to an unknown number of receivers without requiring the receivers to acknowledge the correct receipt of broadcast symbols. In the meantime it is able to provide a performance guarantee on the probability of successful delivery.

Chapter 3

Finite-Length Analysis of LT Codes

In this chapter, we investigate the decoding success probability of finite-length LT codes under *maximum likelihood* (ML) decoding. The decoding success probability is defined as the probability that a receiver can successfully decode all k source symbols with ML decoding given that the receiver has successfully received a certain number of coded symbols. Specifically, if erasure channels are considered, the decoding of LT codes under ML decoding corresponds to solving a consistent system of linear equations over a binary field $GF(2)$, where the coefficients are given by the collected LT code generator matrix. In this chapter, we provide a rigorous mathematical analysis on the rank profile of a random coefficient matrix, where each row vector is independently generated by using the LT encoding process, which is explained in detail in Section 3.2. On the basis of this analysis, we derive upper and lower bounds on the decoding success probability of finite-length LT codes under ML decoding over *binary erasure channel* (BEC). The results of this chapter appear in [J1, C1, C2].

3.1 Introduction

Rateless codes, such as *Luby transform* (LT) codes, were developed to improve the transmission efficiency [90, 50, 31]. The advantages of rateless codes are summarised in Section 2.3. Due to these salient advantages of rateless codes, rateless codes have drawn a lot of attention from industry and academia. The first class of practical digital rateless codes is LT codes [31], which were invented by Luby. LT codes were reviewed in detail in Subsection 2.3.1.

It was shown in [37] that LT codes can deliver excellent performance when the value of k is large. In reality, a traffic session may contain a small numbers of symbols only. Under this scenario, a large symbol overhead, which is defined as $\gamma_R = \frac{m_R}{k}$ and is a key parameter related to the error-performance of LT codes, is however reported [91]. Hyytia et al. [91] optimized the configuration of the degree distribution for LT codes when the number of symbols is small. However, as presented in [91], their proposed methods are not scalable and can only handle the situation when the number of source symbols k is around 10. The authors in [55] proposed a new algorithm for decoding. Using this algorithm, the symbol overhead γ_R is reduced.

A major challenge in analyzing the performance of LT codes is that the decoding success probability of LT codes is difficult to analyze. In this chapter, we investigate the performance of LT codes in terms of the success probability of decoding. In [52], the authors proposed a method to recursively compute the decoding success probability of LT codes under *belief propagation* (BP) decoding. Details of the derivation can be found in [53]. If erasure channels are considered, the decoding of LT codes under *maximum likelihood* (ML) decoding corresponds to solving a consistent system of linear equations over a binary field $GF(2)$, where the coefficients are given by the collected LT code generator matrix. It is worth noting that in [55, 56], a theoretical analysis was conducted on the decoding success probability of LT codes under ML decoding. However the analysis in [55, 56]

was incomplete to the extent that no rigorous analysis was presented to support some results presented in the chapter 3. The analytical result presented on the decoding success probability was in fact an approximation only, which will be discussed in further details in the analysis of Section 3.3. In this chapter, we advance the work in [52, 53, 55, 56] by providing rigorous mathematica analysis on the rank profile of a random coefficient matrix. On the basis of this analysis, we derive upper and lower bounds on the decoding success probability of LT codes under ML decoding over BEC.

Our major contributions can be summarized as follows:

- Firstly, in this chapter, we derive the analytical results, i.e., an upper bound and a lower bound, on the decoding success probability of finite-length LT codes under ML decoding, which is defined as the probability that all source symbols can be successfully decoded by a receiver with a given number of successfully received coded symbols. The analytical results are obtained by conducting an analysis of the rank profile of a random coefficient matrix.

- Secondly, simulations are conducted to validate the accuracy of the proposed bounds. More specifically, LT codes with different degree distributions are evaluated to measure the accuracy of the derived bounds.

The rest of the chapter is organized as follows. Section 3.2 reviews encoding and decoding process of LT codes. In Section 3.3, we analyze the probability that a receiver can successfully decode all source symbols conditioned on the event that the receiver has successfully received a known number of coded symbols. In Section 3.4, we validate our analytical results using simulations. Section 3.5 concludes the chapter.

3.2 Preliminaries

In this section, we review encoding and decoding process of LT code.

When LT codes are used by the transmitter to deliver k source symbols, the following encoding rule is utilized to generate each coded symbol: firstly a positive integer d (often referred to as the "degree" [31] of coded symbols) is drawn from the set of integers $\{1, ..., k\}$ according to a probability distribution $\mathbf{\Omega} = (\Omega_1, ..., \Omega_k)$ where Ω_d is the probability that d is picked and $\sum_{d=1}^{k} \Omega_d = 1$. Then, d distinct source symbols are selected randomly and independently from the k source symbols, where each source symbol is selected with equal probability. These d source symbols are then network encoded using XOR operation to generate the coded symbol [31, 37]. Finally, the coded symbol is transmitted to all receivers.

A typically used decoding process for LT codes is the so-called "LT process" [31], but it is well known that the LT process is not able to decode all decodable source symbols from the successfully received coded symbols. Therefore in this chapter, we use a different decoding algorithm called the full-rank decoding [55] to decode the source symbols. More specifically, let $m_R (m_R \geq k)$ be the number of coded symbols that have already been successfully received by a receiver. We use a $1 \times k$ row vector to represent the information contained in a coded symbol, where the j^{th} entry of the row vector is 1 if the corresponding coded symbol is a result of XOR operation on the j^{th} source symbol (and other source symbols); otherwise the j^{th} entry equals to 0. Thus, a random row vector in this chapter refers to the row vector of a randomly chosen coded symbol where the coded symbol is generated using the LT codes encoding process. In this way, the information contained in the m_R coded symbols can be represented by a $m_R \times k$ matrix, denoted by $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$.

Recall that ML decoding of an LT code over BEC corresponds to solving a consistent system of m_R random linear equations in k unknowns over a binary field $GF(2)$. The probability that the system is solvable is equal to the probability that the decoding matrix $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ at the receiver has rank k. Hence, the decoding success probability of LT codes after ML decoding equals the probability that

$\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ has rank k.

There have been a large number of works (e.g. [92, 93, 94, 95, 96, 97, 98, 99, 100]) examining rank properties of random matrices. However, all the works consider the random matrices that contain certain restrictions on either the randomness or on the dimensions of the matrix. In terms of randomness, the restrictions are that only element-wise uniform randomness is considered, i.e. each element of the random matrix is sampled uniformly from $(0, 1)$ (standard random ensemble). In terms of the matrix dimensions, the restrictions are only considering square matrices, i.e., $k \times k$ matrices or deriving only asymptotic expressions. Nevertheless, for the analysis of LT codes, with a finite length and a row-wise random matrix construction, the mature results from the previous literature are far less sufficient. In this chapter, we derive analytical results on the probability that $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is a full rank matrix.

3.3 Analysis on the Decoding Success Probability of LT Codes

Denote by $R^k_{m_R}$ the event that a receiver can successfully decode all k source symbols conditioned on the event that the receiver has successfully received m_R coded symbols. In this section, we shall analyze the probability of $R^k_{m_R}$. Particularly an upper and a lower bound on $\Pr\left[R^k_{m_R}\right]$ will be derived.

We say that the receiver can recover all k source symbols from the m_R coded symbols if and only if $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is a full rank matrix, i.e. its rank equals to k. Note that in this chapter, all algebraic operations and the associated analysis are conducted in a binary field. Obviously the event that $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is a full rank matrix is equivalent to the event $R^k_{m_R}$.

The main result of this section is summarized in the following theorem:

Theorem 3.1. *When the transmitter generates coded symbols using the LT codes*

and the coded symbols received at a receiver are decoded using the full-rank decoding, the probability that a receiver can successfully decode all k source symbols from m_R received coded symbols with $m_R \geq k$, denoted by $R^k_{m_R}$, satisfies

$$\Pr\left[R^k_{m_R}\right] \leq \mathbf{e}_k \left(\mathbf{X}\right)^{m_R-1} \mathbf{e}_1^T, \tag{3.1}$$

where $\mathbf{e}_k$ is a $1 \times k$ row vector with the k^{th} entry equal to 1 and all other entries equal to 0,

$$\mathbf{X} = \begin{pmatrix} 1-O_1^1 & 0 & \cdots & 0 & 0 \\ O_1^1 & 1-O_2^2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-O_{k-1}^{k-1} & 0 \\ 0 & 0 & \cdots & O_{k-1}^{k-1} & 1-O_k^k \end{pmatrix}$$

and

$$O_m^m = \frac{\Pr\left[R_{m+1}^{m+1}\right]}{\Pr\left[R_m^m\right]}.$$

Further,

$$\Pr\left[R_m^m\right] = \prod_{q=2}^{m} \left[(1-I_q)^{\binom{m}{q}}\right],$$

where I_q is given by:

$$I_{q,\ q\geq 2} = (Q_{10}, Q_{20}, \ldots, Q_{k0})\mathbf{Tr}^{q-2}(\Omega_1, \Omega_2, \ldots, \Omega_k)^T$$

and $\mathbf{Tr}$ *in the above equation is given by*

$$\mathbf{Tr} = \begin{pmatrix} Q_{11} & \cdots & Q_{(k-1)1} & Q_{k1} \\ Q_{12} & \cdots & Q_{(k-1)2} & Q_{k2} \\ \vdots & \ddots & \vdots & \vdots \\ Q_{1k} & \cdots & Q_{(k-1)k} & Q_{kk} \end{pmatrix}$$

and

$$Q_{ij} = \begin{cases} \sum\limits_{\substack{0 \le a \le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i < j \\ \sum\limits_{\substack{1 \le a \le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i = j \\ \sum\limits_{\substack{i-j \le a \le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i > j. \end{cases}$$

In addition to the above upper bound, a lower bound of $\Pr\left[R_{m_R}^k\right]$ *can also be obtained:*

$$\Pr\left[R_{m_R}^k\right] \ge \mathbf{e}_k \begin{pmatrix} 1-u_1 & \cdots & 0 & 0 \\ u_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1-u_{k-1} & 0 \\ 0 & \cdots & u_{k-1} & 1-u_k \end{pmatrix}^{m_R-1} \mathbf{R}(1), \tag{3.2}$$

where

$$u_z = \max_{0\le i\le k-z}\{\sum_{d=0}^{z-1}[\binom{z-1}{d}P_g(d+i-z+1)] + \sum_{d=1}^{z-1}[\binom{z-1}{d}P_g(d)]\}$$

and $P_g(d) = \frac{\Omega_d}{\binom{k}{d}}$.

The rest of this section is devoted to the proof of Theorem 3.1. Because of the close connection between the event $R^k_{m_R}$ and the event that $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ is a full rank matrix, the analysis of $\Pr\left[R^k_{m_R}\right]$ is conducted by analyzing the rank of $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$.

3.3.1 Analysis of the Rank of a Random Matrix

In this subsection, we give procedure on computing the probability that $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ is a full rank matrix, where $m_R \geq k$.

Let $R^r_{m_R}$ be the event that the rank of the encoding coefficient matrix $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ is r and let $\Pr\left[R^r_{m_R}\right]$ be its probability. Define the rank profile of $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ to be a vector $\mathbf{R}(m_R) = \left(\Pr\left[R^1_{m_R}\right], \Pr\left[R^2_{m_R}\right], \ldots, \Pr\left[R^k_{m_R}\right]\right)^T$. Noting that the decoding success probability is equal to the probability that the rank of the encoding coefficient matrix $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ equals k, i.e. $\Pr\left[R^k_{m_R}\right]$, our analysis on the decoding success probability relies on a recursive computation of $\mathbf{R}(m_R)$ as m_R increases.

When $m_R = 1$, it can be readily shown that $\mathbf{R}(1) = (\Pr[R^1_1], \Pr[R^2_1], \ldots, \Pr\left[R^k_1\right])^T = (1, 0, \ldots, 0)^T$. For $m_R > 1$, the rank profile of $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ can be obtained from the rank profile of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ recursively. Particularly, $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ can be considered as $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ with an additional row $\mathbf{x}$ added into $\mathbf{G}_{(n-1)\times k}$. The degree of $\mathbf{x}$, i.e. the number of non-zero elements of $\mathbf{x}$, is chosen according to the pre-defined degree distribution $\mathbf{\Omega} = (\Omega_1, ..., \Omega_k)$ and each non-zero element is

then placed randomly and uniformly into $\mathbf{x}$. Let $rk(\mathbf{G})$ be the rank of the matrix $\mathbf{G}$ and let $Im(\mathbf{G})$ be the row vector space generated by a matrix $\mathbf{G}$. That is, $Im(\mathbf{G})$ is the vector space formed by all linear combinations of the rows of $\mathbf{G}$. Note that it may possibly occur that $Im(\mathbf{G}_{n\times k}) = Im(\mathbf{G}_{m\times k})$ where $m \neq n$. If a row vector $\mathbf{x}$ can be expressed as a linear combination of the row vectors of $\mathbf{G}$, we say that $\mathbf{x} \in Im(\mathbf{G})$; otherwise $\mathbf{x} \notin Im(\mathbf{G})$. For $k \geq r \geq 2$, it can be shown that

$$\begin{aligned}
&\Pr\left[rk(\mathbf{G}^{\mathrm{LT}}_{m_R\times k}) = r\right] \\
=&\Pr\left[rk(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = r\right] \times \\
&\Pr\left[\mathbf{x} \in Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) \mid rk(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = r\right] \\
&+\Pr\left[rk(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = r-1\right] \times \\
&\Pr\left[\mathbf{x} \notin Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) \mid rk(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = r-1\right]. \qquad (3.3)
\end{aligned}$$

For convenience let $O^{r-1}_{m_R-1} = \Pr\left[\mathbf{x} \notin Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) \mid R^{r-1}_{n-1}\right]$. It follows from the equation (3.3) that:

$$\Pr\left[R^r_{m_R}\right] = \Pr\left[R^r_{m_R-1}\right](1 - O^r_{m_R-1}) + \Pr\left[R^{r-1}_{m_R-1}\right] O^{r-1}_{m_R-1}. \qquad (3.4)$$

Based on (3.4), the following equation can be obtained by recursion:

$$\begin{aligned}
&\mathbf{R}(m_R) \\
=&\begin{pmatrix}
1 - O^1_{m_R-1} & \cdots & 0 & 0 \\
O^1_{m_R-1} & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 1 - O^{k-1}_{m_R-1} & 0 \\
0 & \cdots & O^{k-1}_{m_R-1} & 1 - O^k_{m_R-1}
\end{pmatrix} \mathbf{R}(m_R - 1) \\
=&(\prod_{l=1}^{m_R-1} \mathbf{X}_l)\mathbf{R}(1), \qquad (3.5)
\end{aligned}$$

where

$$\mathbf{X}_l = \begin{pmatrix} 1-O_l^1 & 0 & \cdots & 0 & 0 \\ O_l^1 & 1-O_l^2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-O_l^{k-1} & 0 \\ 0 & 0 & \cdots & O_l^{k-1} & 1-O_l^k \end{pmatrix}.$$

The probability that $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is of full rank, hence all k source symbols can be successfully decoded, can be calculated by:

$$\begin{aligned} \Pr\left[R^k_{m_R}\right] &= \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \mathbf{R}(m_R) \\ &= \mathbf{e}_k\Big(\prod_{l=1}^{m_R-1} \mathbf{X}_l\Big)\mathbf{R}(1), \end{aligned} \tag{3.6}$$

where $\mathbf{e}_i$, $1 \le i \le k$, is a $1 \times k$ row vector with the i^{th} entry equal to 1 and all other entries equal to 0.

The above recursive way of computing the rank profile of $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ and the probability that $\mathbf{G}^{\mathrm{LT}}_{m_R \times k}$ is a full rank matrix relies on the knowledge of the parameters $O^z_{m_R-1} = \Pr\left[\mathbf{x} \notin Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) \mid R^z_{m_R-1}\right], 1 \le z \le k$. In the following paragraphs, we give analysis on the computation of $\Pr\left[\mathbf{x} \notin Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) \mid R^z_{m_R-1}\right]$.

For convenience let A_{m_R-1} be the event that $\mathbf{x} \notin Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k})$ and $\overline{A_{m_R-1}}$ be the complement of event A_{m_R-1}. Temporarily assuming that $rk(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = z, 1 \le z \le k$ and noting that $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ is a random matrix, under the above two conditions, let V^z be a row vector space formed by all linear combinations of the rows of *an instance* of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$. Of course the dimension of V^z equals to z, hence the superscript. Further, let $\mathcal{E}^z$ be the set of *all possible* and *distinct* V^zs: $\mathcal{E}^z \triangleq \{V^z\}$. When $z = k$, the row vector space whose dimension is k is unique. However when $1 \le z < k$, there are multiple distinct row vector spaces with dimension z. For convenience, we number the elements of $\mathcal{E}^z$ sequentially

and denote by Γ_v^z be the set of indices of all V^z satisfying $V^z \in \mathcal{E}^z$. Denote by V_i^z the i^{th} element of $\mathcal{E}^z$. As noted in the last paragraph, the coding coefficient matrix $\mathbf{G}$ and the vector space formed by the row vectors of $\mathbf{G}$ have independent significance in the sense that for two positive integers $m, n \geq z$ and $m \neq n$, it may happen that $V_i^z = Im(\mathbf{G}_{n\times k}) = Im(\mathbf{G}_{m\times k})$. That is, the vector space and its existence does not depend on some details of the coding coefficient matrix, e.g. number of rows in the coding coefficient matrix and a particular instance of the coding coefficient matrix.

Let $F_{i,n-1}^z$ be the event $Im(\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}) = V_i^z$. It can be readily shown that: 1) $R_{m_R-1}^z = \cup_{i\in\Gamma_v^z} F_{i,m_R-1}^z$, i.e. event that the rank of the encoding coefficient matrix $\mathbf{G}_{m_R\times k}^{\mathrm{LT}}$ is z equals to the joint events that $Im(\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}) = V_i^z$ for all $i, i \in \Gamma_v^z$; 2) $F_{i,m_R-1}^z \cap F_{j,m_R-1}^z = \tilde{\mathrm{A}}^{\check{}}$ for $i \neq j$. Using the definitions of the two events $R_{m_R}^z$ and F_{i,m_R-1}^z, Bayes' formula and the above two results, we have

$$\begin{aligned}
&\Pr\left[\mathbf{x} \in Im(\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}) \mid rk(\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}) = z\right] \\
=&\Pr\left[\overline{A_{m_R-1}} \mid R_{m_R-1}^z\right] = \frac{\Pr\left[\overline{A_{m_R-1}} \cap R_{m_R-1}^z\right]}{\Pr\left[R_{m_R-1}^{r-1}\right]} \\
=&\frac{\Pr\left[\overline{A_{m_R-1}} \cap (\cup_{i\in\Gamma_v^z} F_{i,m_R-1}^z)\right]}{\Pr\left[\cup_{i\in\Gamma_v^z} F_{i,m_R-1}^z\right]} = \frac{\sum_{i\in\Gamma_v^z}\Pr\left[\overline{A_{m_R-1}} \cap F_{i,m_R-1}^z\right]}{\sum_{i\in\Gamma_v^z}\Pr\left[F_{i,m_R-1}^z\right]} \\
=&\frac{\sum_{i\in\Gamma_v^z}\Pr\left[\overline{A_{m_R-1}} \mid F_{i,m_R-1}^z\right]\Pr\left[F_{i,m_R-1}^z\right]}{\sum_{i\in\Gamma_v^z}\Pr\left[F_{i,m_R-1}^z\right]}.
\end{aligned} \tag{3.7}$$

Let $\overline{B_i^z}$ be the event that $\mathbf{x} \in V_i^z$. Conditioned on the event F_{i,m_R-1}^z and noting that $\mathbf{x}$ is drawn randomly and independently of the row vectors of $\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}$, we have

$$\overline{A_{m_R-1}} \mid F_{i,m_R-1}^z \Leftrightarrow \overline{B_i^z} \mid F_{i,m_R-1}^z. \tag{3.8}$$

Because each row vector is drawn *independently* of other row vectors, the two events $\mathbf{x} \in V_i^z$ and $Im(\mathbf{G}_{(m_R-1)\times k}^{\mathrm{LT}}) = V_i^z$ are independent. It follows using the definitions of $\overline{B_i^z}$ and F_{i,m_R-1}^z that $\Pr\left[\overline{B_i^z} \mid F_{i,m_R-1}^z\right] = \Pr\left[\overline{B_i^z}\right] = \Pr\left[\mathbf{x} \in V_i^z\right]$.

For the other term $\Pr\left[F^z_{i,m_R-1}\right]$ in (3.7), we recall that F^z_{i,m_R-1} is the event $Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}) = V^z_i$. Let E^z_{i,m_R-1} be the event $V^z_i \subseteq Im(\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k})$ and obviously $F^z_{i,m_R-1} \subseteq E^z_{i,m_R-1}$. Conditioned on the event E^z_{i,m_R-1}, without loss of generality, let $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_z\}$ be the row vectors of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ that forms a basis of V^z_i. The set of row vectors of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ that forms a basis of V^z_i may not be unique. Let $\{\mathbf{w}_1, \mathbf{w}_1, ..., \mathbf{w}_{m_R-z-1}\}$ be the remaining row vectors of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$. Further note that each row vector of $\mathbf{G}^{\mathrm{LT}}_{(m_R-1)\times k}$ is formed *independently* of other row vectors. Noting that $F^z_{i,m_R-1} \subseteq E^z_{i,m_R-1}$, it can be shown that

$$\begin{aligned}
&\Pr\left[F^z_{i,m_R-1}\right] \\
=&\Pr\left[F^z_{i,m_R-1}|E^z_{i,m_R-1}\right]\Pr\left[E^z_{i,m_R-1}\right] \\
=&\Pr\left[\mathbf{w}_1 \in V^z_i \cap \cdots \cap \mathbf{w}_{m_R-z-1} \in V^z_i|E^z_{i,m_R-1}\right]\Pr\left[E^z_{i,m_R-1}\right] \\
=&\left(\Pr\left[\mathbf{w}_1 \in V^z_i|E^z_{i,m_R-1}\right]\right)^{m_R-z-1}\Pr\left[E^z_{i,m_R-1}\right] \\
=&\left(\Pr\left[\overline{B^z_i}\right]\right)^{m_R-z-1}\Pr\left[E^z_{i,m_R-1}\right], \qquad (3.9)
\end{aligned}$$

where the last step result is because the two events $\mathbf{w}_1 \in V^z_i$ and E^z_{i,m_R-1} are independent. Combining the three equations (3.7), (3.8), and (3.9), conclusion follows that

$$\begin{aligned}
&\Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right] \\
=&\frac{\sum_{i\in\Gamma^z_v}\Pr\left[\overline{A_{m_R-1}} \mid F^z_{i,m_R-1}\right]\Pr\left[F^z_{i,m_R-1}\right]}{\sum_{i\in\Gamma^z_v}\Pr\left[F^z_{i,m_R-1}\right]} \\
=&\frac{\sum_{i\in\Gamma^z_v}\Pr\left[\overline{B^z_i}\right]\left(\Pr\left[\overline{B^z_i}\right]\right)^{m_R-z-1}\Pr\left[E^z_{i,m_R-1}\right]}{\sum_{i\in\Gamma^z_v}\left(\Pr\left[\overline{B^z_i}\right]\right)^{m_R-z-1}\Pr\left[E^z_{i,m_R-1}\right]} \\
=&\frac{\sum_{i\in\Gamma^z_v}\left(\Pr\left[\overline{B^z_i}\right]\right)^{m_R-z}\Pr\left[E^z_{i,m_R-1}\right]}{\sum_{i\in\Gamma^z_v}\left(\Pr\left[\overline{B^z_i}\right]\right)^{m_R-z-1}\Pr\left[E^z_{i,m_R-1}\right]}. \qquad (3.10)
\end{aligned}$$

As manifested in equation (3.10), the computation of $\Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right]$, which is required for computing the rank profile of $\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ and the probability that

$\mathbf{G}^{\mathrm{LT}}_{m_R\times k}$ is a full rank matrix, relies on the knowledge of $\Pr\left[\overline{B_i^z}\right]$ and $\Pr\left[E^z_{i,m_R-1}\right]$. These parameters can be difficult to obtain when k is large. Therefore in the rest of this section, we devote our efforts to finding an upper and a lower bound of $\Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right]$, which will be shown later using simulations to be reasonably tight.

Derivation of An Upper Bound of $\Pr\left[R^k_{m_R}\right]$

Let $a_{i,m_R-1} = \Pr\left[E^z_{i,m_R-1}\right]$ and $b_{i,z} = \Pr\left[\overline{B_i^z}\right]$ for notational convenience. Equation (3.10) can be rewritten as:

$$\Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right] = \frac{\sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z}}{\sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}}. \tag{3.11}$$

Next we shall evaluate the monotonicity of $\Pr\left[\overline{A_{m_R-1}} \mid R^{r-1}_{m_R-1}\right]$ as a function of m_R. It can be shown that :

$$\begin{aligned}
&\Pr\left[\overline{A_{m_R}} \mid R^z_{m_R}\right] - \Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right] \\
=&\frac{\sum_{i\in\Gamma_v^z} a_{i,n} b_{i,z}^{m_R-z+1}}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z}} - \frac{\sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z}}{\sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}} \\
=&\frac{\sum_{i\in\Gamma_v^z} a_{i,m_R} a_{i,m_R-1} b_{i,z}^{2m_R-2z} - \sum_{i\in\Gamma_v^z} a_{i,m_R} a_{i,m_R-1} b_{i,z}^{2m_R-2z}}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z} \sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}} \\
&+ \frac{\sum_{j\in\Gamma_v^z}\sum_{i\in\Gamma_v^z} a_{i,m_R} a_{j,m_R-1} b_{i,z}^{m_R-z+1} b_{j,z}^{m_R-z-1}}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z} \sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}} \\
&- \frac{\sum_{j\in\Gamma_v^z}\sum_{i\in\Gamma_v^z} a_{i,m_R} a_{j,m_R-1} b_{i,z}^{m_R-z} b_{j,z}^{m_R-z}}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z} \sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}} \\
=&\frac{\sum_{j\in\Gamma_v^z}\sum_{i\in\Gamma_v^z} a_{i,m_R} a_{j,m_R-1} b_{i,z}^{m_R-z-1} b_{j,z}^{m_R-z-1} (b_{i,z}^2 - 2b_{i,z}b_{j,z} + b_{j,z}^2)}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z} \sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_{i,z}^{m_R-z-1}} \\
=&\frac{\sum_{j\in\Gamma_v^z}\sum_{i\in\Gamma_v^z} a_{i,m_R} a_{j,m_R-1} b_{i,z}^{m_R-z-1} b_{j,z}^{m_R-z-1} (b_{i,z} - b_{j,z})^2}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_i^{m_R-z} \sum_{i\in\Gamma_v^z} a_{i,m_R-1} b_i^{m_R-z-1}} \geq 0.
\end{aligned} \tag{3.12}$$

As a result of the above analysis, we can conclude that the conditional probability $\Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R}\right]$ is a monotonically increasing function with m_R and $\Pr\left[\overline{A_{m_R}} \mid R^z_{m_R}\right] \geq \Pr\left[\overline{A_{m_R-1}} \mid R^z_{m_R-1}\right] \geq \cdots \geq \Pr\left[\overline{A_z} \mid R^z_z\right]$.

Let

$$\mathbf{X}=\begin{pmatrix} 1-O_1^1 & 0 & \cdots & 0 & 0 \\ O_1^1 & 1-O_2^2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-O_{k-1}^{k-1} & 0 \\ 0 & 0 & \cdots & O_{k-1}^{k-1} & 1-O_k^k \end{pmatrix}. \tag{3.13}$$

We can then obtain that

$$\begin{aligned} \mathbf{e}_k(\prod_{l=1}^{m_R-1}\mathbf{X}_l)\mathbf{R}(1) &\leq \mathbf{e}_k(\mathbf{X})^{m_R-1}\mathbf{R}(1) \\ \Pr\left[R_{m_R}^k\right] &\leq \mathbf{e}_k(\mathbf{X})^{m_R-1}\mathbf{R}(1). \end{aligned} \tag{3.14}$$

Now an upper bound of the decoding success probability is derived and this relies on the knowledge of $O_z^z, 1 \leq z \leq k$. In the following paragraphs, we present analysis leading to the computation of $O_z^z, 1 \leq z \leq k$. Noting that when $1 \leq z \leq k$, $\mathbf{x} \notin Im(\mathbf{G}_{z\times k}^{\mathrm{LT}}) \cap rk(\mathbf{G}_{z\times k}^{\mathrm{LT}}) = z \Leftrightarrow rk(\mathbf{G}_{(z+1)\times k}^{\mathrm{LT}}) = z+1$, it can be shown that

$$\begin{aligned} O_z^z &= \Pr\left[\mathbf{x} \notin Im(\mathbf{G}_{z\times k}^{\mathrm{LT}}) \mid rk(\mathbf{G}_{z\times k}^{\mathrm{LT}}) = z\right] \\ &= \frac{\Pr\left[\mathbf{x} \notin Im(\mathbf{G}_{z\times k}^{\mathrm{LT}}) \cap rk(\mathbf{G}_{z\times k}^{\mathrm{LT}}) = z\right]}{\Pr\left[rk(\mathbf{G}_{z\times k}^{\mathrm{LT}}) = z\right]} \\ &= \frac{\Pr\left[rk(\mathbf{G}_{(z+1)\times k}^{\mathrm{LT}}) = z+1\right]}{\Pr\left[rk(\mathbf{G}_{z\times k}^{\mathrm{LT}}) = z\right]} = \frac{\Pr\left[R_{z+1}^{z+1}\right]}{\Pr\left[R_z^z\right]}, \end{aligned} \tag{3.15}$$

where $\Pr[R_z^z]$ represents the probability that a random (encoding coefficient) matrix $\mathbf{G}_{z\times k}^{\mathrm{LT}}$, $z \leq k$, is of full rank. The method to calculate $\Pr[R_z^z]$ is provided in the following lemma.

Lemma 3.2. *Let $\mathbf{v}_i$ be the i^{th} row vector of $\mathbf{G}_{z\times k}^{LT}$. Denote by I_q (whose value will be determined later in Lemma 3.3) the probability of the event that $\sum_{i=1}^{q}\mathbf{v}_i = \mathbf{0}$,*

conditioned on that the summation of any w row vectors of $\mathbf{G}_{z\times k}^{LT}$ is not equal to $\mathbf{0}$*, where* $\mathbf{0}$ *is a* $1\times k$ *row vector with all elements equal to* 0*,* $w \in \mathbb{Z}^+,\ 1 < w < q$*.* $\Pr[R_z^z]$ *can be determined by:*

$$\Pr[R_z^z] \;=\; \prod_{q=2}^{z}\left[(1-I_q)^{\binom{z}{q}}\right]. \tag{3.16}$$

Proof. We observe that $\mathbf{G}_{z\times k}^{\mathrm{LT}}$ being full rank implies that there does *not* exist a set of coefficients $c_1,\ldots,c_r$ such that $\sum_{i=1}^{r} c_i\mathbf{v}_i = 0$. Further, since we are working in a binary field, c_i can be either 1 or 0. It follows that $\mathbf{G}_{z\times k}^{\mathrm{LT}}$ being full rank is a sufficient and necessary condition for that for every integer $2 \le q \le r$, the summation of any q row vectors of $\mathbf{G}_{z\times k}^{\mathrm{LT}}$ is not equal to $\mathbf{0}$. This observation forms the basis of the proof.

Let $NZ(q)$ be the event that the summation of any q row vectors in $\mathbf{G}_{z\times k}^{\mathrm{LT}}$ are not equal to $\mathbf{0}$. The probability of $NZ(2)$ can be expressed as $\Pr[NZ(2)] = (1-I_2)^{\binom{r}{2}}$. Further, for every integer q satisfying $3 \le q \le r$,

$$\Pr[\cap_{i=2}^{q} NZ(i)] = \Pr\left[NZ(q) \mid \cap_{i=2}^{q-1} NZ(i)\right]\Pr\left[\cap_{i=2}^{q-1} NZ(i)\right]. \tag{3.17}$$

With the recursive application of equation (3.17), we can conclude that the probability that $\mathbf{G}_{z\times k}^{\mathrm{LT}}$, $z \le k$, is of full rank can be obtained as

$$\Pr[R_z^z] \;=\; \Pr(\cap_{i=2}^{z} NZ(i)) = \prod_{q=2}^{z}\left[(1-I_q)^{\binom{z}{q}}\right]. \tag{3.18}$$

□

Now we shall derive I_q which is required in Lemma 3.2. To obtain I_q, we must first evaluate the degree transition probability Q_{ij}, i.e. the probability that the row vector $\mathbf{S}_q$ produced by summing q row vectors has degree j given that the row vector $\mathbf{S}_{q-1}$ generated by summing the first $q-1$ row vectors of the above q

row vectors has degree i. We can derive Q_{ij}[55] as:

$$Q_{ij} = \begin{cases} \sum\limits_{\substack{0\le a\le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i < j \\ \sum\limits_{\substack{1\le a\le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i = j \\ \sum\limits_{\substack{i-j\le a\le \min(k-j,i) \\ b=j-i+a}} \Omega_{a+b} \frac{\binom{i}{a}\binom{k-i}{b}}{\binom{k}{a+b}}, & i > j \end{cases}, \tag{3.19}$$

where Ω_d, $1 \le d \le k$ is the degree distribution of LT codes, which is defined in Section 3.2.

Now we are ready to analyze I_q.

Lemma 3.3. *Let* $\mathbf{Tr}$ *be a* $k \times k$ *transition matrix with dimension* $k \times k$ *whose* $(j,i)^{th}$ *element equal to* Q_{ij}*. The matrix* $\mathbf{Tr}$ *can be expressed as:*

$$\mathbf{Tr} = \begin{pmatrix} Q_{11} & \cdots & Q_{(k-1)1} & Q_{k1} \\ Q_{12} & \cdots & Q_{(k-1)2} & Q_{k2} \\ \vdots & \ddots & \vdots & \vdots \\ Q_{1k} & \cdots & Q_{(k-1)k} & Q_{kk} \end{pmatrix},$$

the probability I_q *is given by:*

$$I_{q,\ q\ge 2.} = (Q_{10}, Q_{20}, \ldots, Q_{k0})\mathbf{Tr}^{q-2} \cdot (\Omega_1, \Omega_2, \ldots, \Omega_k)^T. \tag{3.20}$$

Proof. To obtain I_q, we analyze the degree distribution of row vector $\mathbf{S}_w$ which is the sum of w row vectors. Note that the degree of $\mathbf{S}_w$ should not equal to 0.

Let $\mathbf{D}^w = (D_1^w, \ldots, D_k^w)^T$ be the degree distribution of the sum of w (random) row vectors and $w \geq 1$, where D_i^w is the probability that the degree of the row vector $\mathbf{S}_w$ is $i,\ 1 \leq i \leq k$. When $w = 1$, the degree distribution $\mathbf{D}^1$ is obviously $(\Omega_1, \Omega_2, \ldots, \Omega_k)^T$. For $w \geq 2$, the relationship can be analytically described as :

$$D_m^w = (Q_{1m}, Q_{2m}, \ldots, Q_{km})(D_1^{w-1}, \ldots, D_k^{w-1})^T. \tag{3.21}$$

From the equation (3.21), it follows that:

$$\begin{aligned}
\mathbf{D}^w &= (D_1^w, \ldots, D_k^w)^T \\
&= \begin{pmatrix} Q_{11} & \cdots & Q_{(k-1)1} & Q_{k1} \\ \vdots & \ddots & \vdots & \vdots \\ Q_{1(k-1)} & \cdots & Q_{(k-1)(k-1)} & Q_{k(k-1)} \\ Q_{1k} & \cdots & Q_{(k-1)k} & Q_{kk} \end{pmatrix} \begin{pmatrix} D_1^{w-1} \\ \vdots \\ D_{k-1}^{w-1} \\ D_k^{w-1} \end{pmatrix} \\
&= \mathbf{Tr}^{w-1} \cdot (\Omega_1, \Omega_2, \ldots, \Omega_k)^T.
\end{aligned} \tag{3.22}$$

As an easy consequence of equation (3.22), I_q can be obtained:

$$\begin{aligned}
I_q &= D_0^q = \sum_{i=1}^{k} D_i^{q-1} Q_{i0} = (Q_{10}, Q_{20}, \ldots, Q_{k0})\mathbf{D}^{q-1} \\
&= (Q_{10}, Q_{20}, \ldots, Q_{k0})\mathbf{Tr}^{q-2} \cdot (\Omega_1, \Omega_2, \ldots, \Omega_k)^T.
\end{aligned} \tag{3.23}$$

□

Using (3.14), (3.15) and Lemmas 3.2 and 3.3, an upper bound on $\Pr\left[R_{m_R}^k\right]$ can be computed, which completes the first part of the proof of Theorem 3.1 on the upper bound.

Derivation of A Lower Bound of $\Pr\left[R_{m_R}^k\right]$

In addition to the upper bound derived earlier in the section, a lower bound on the decoding success probability can also be obtained:

$$\begin{aligned}\Pr\left[\overline{A_{m_R}} \mid R_{m_R}^z\right] &= \frac{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z+1}}{\sum_{i\in\Gamma_v^z} a_{i,m_R} b_{i,z}^{m_R-z}} \leq \max_{i\in\Gamma_v^z}\{b_{i,z}\} \\ &\leq \max_{i\in\Gamma_v^z}\{\Pr\left[\overline{B_i^z}\right]\}. \end{aligned} \tag{3.24}$$

Thus we can obtain that

$$\begin{aligned}\mathbf{e}_k(\mathbf{X}_{min})^{m_R-1}\mathbf{R}(1) &\leq \mathbf{e}_k(\prod_{l=1}^{m_R-1}\mathbf{X}_l)\mathbf{R}(1) \\ \Pr\left[R_{m_R}^k\right] &\geq \mathbf{e}_k(\mathbf{X}_{min})^{m_R-1}\mathbf{R}(1), \end{aligned} \tag{3.25}$$

where $\mathbf{X}_{min}$ is given in (3.26).

$$\begin{aligned}&\mathbf{X}_{min} \\ =&\begin{pmatrix} 1-\max_{i\in\Gamma_v^1}\{\Pr\left[\overline{B_i^1}\right]\} & \cdots & 0 & 0 \\ \max_{i\in\Gamma_v^1}\{\Pr\left[\overline{B_i^1}\right]\} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1-\max_{i\in\Gamma_v^{k-1}}\{\Pr\left[\overline{B_i^{k-1}}\right]\} & 0 \\ 0 & \cdots & \max_{i\in\Gamma_v^{k-1}}\{\Pr\left[\overline{B_i^{k-1}}\right]\} & 1-\max_{i\in\Gamma_v^k}\{\Pr\left[\overline{B_i^k}\right]\} \end{pmatrix} \end{aligned} \tag{3.26}$$

The above lower bound relies on the knowledge of $\max_{i\in\Gamma_v^z}\{\Pr\left[\overline{B_i^z}\right]\}, 1 \leq z \leq k$. In the following analysis, we give analysis that leads to the computation of $\max_{i\in\Gamma_v^z}\{\Pr\left[\overline{B_i^z}\right]\}$.

Note that a particular row vector with degree d occurs with probability

$$P_g(d) = \frac{\Omega_d}{\binom{k}{d}}, \tag{3.27}$$

where Ω_d is the probability that a (any) row vector with degree d is chosen and $\binom{k}{d}$ is the total number of degree d vectors among all $1 \times k$ binary vectors. Recall that the degree of a vector is the number of non-zero elements in it. Recall that $\mathbf{e}_i$ is a $1 \times k$ row vector with the i^{th} entry equal to 1 and all other entries equal to 0. Obviously $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ forms a set of orthogonal basis vectors where *any row vector,* hence a row vector in any V_i^z, $i \in \Gamma_v^z$, in the coding coefficient matrix can be represented as a linear combination of these basis vectors. Let us focus now on a z dimensional subspace formed by $\{\mathbf{e}_1, \ldots, \mathbf{e}_z\}$, denoted by $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_z\}}$. Using some straightforward combinatorial argument and further noting that we are working in a binary field, it can be shown that the number of degree d, $d \leq z$, vectors in $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_z\}}$ is given by $\binom{z}{d}$. Therefore $\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_z\}}\right] = \sum_{d=1}^{z} \left[\binom{z}{d} P_g(d)\right]$. Denote by Ω_i^z any other z dimensional vector space whose basis vectors are the row vectors of a matrix obtainable by reshuffling the columns of the matrix $\{\mathbf{e}_1, \ldots, \mathbf{e}_z\}^T$ (or equivalently any other z dimensional vector space whose basis vectors are obtained by randomly choosing z vectors from $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$). Because the number of non-zero elements are uniformly and independently distributed in a row vector, it follows that $\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_z\}}\right] = \Pr\left[\mathbf{x} \in \Omega_i^z\right]$.

Now let us consider a z dimensional vector space formed by the basis vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_{z-1}, \mathbf{e}_z + \mathbf{e}_{z+1}\}$. Except for the last basis vector which has degree 2, all other basis vectors have degree 1 only. Using some straightforward combinatorial argument, the number of vectors in $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\mathbf{e}_{z+1}\}}$ containing $\mathbf{e}_z + \mathbf{e}_{z+1}$ *and* having a degree $d+2$ is given by $\binom{z-1}{d}$; the number of vectors in $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\mathbf{e}_{z+1}\}}$ *not* containing $\mathbf{e}_z + \mathbf{e}_{z+1}$ *and* having a degree d is given by

$\begin{pmatrix} z-1 \\ d \end{pmatrix}$. Therefore

$$\begin{aligned} &\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\mathbf{e}_{z+1}\}}\right] \\ =&\sum_{d=0}^{z-1}\left[\begin{pmatrix} z-1 \\ d \end{pmatrix} P_g(d+2)\right] + \sum_{d=1}^{z-1}\left[\begin{pmatrix} z-1 \\ d \end{pmatrix} P_g(d)\right]. \end{aligned} \tag{3.28}$$

Similarly, denote by Ω_i^z any other z dimensional vector space whose basis vectors are the row vectors of a matrix obtainable by reshuffling the columns of the matrix $\{\mathbf{e}_1, \ldots, \mathbf{e}_{z-1}, \mathbf{e}_z + \mathbf{e}_{z+1}\}^T$. It can be shown that $\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\mathbf{e}_{z+1}\}}\right] = \Pr\left[\mathbf{x} \in \Omega_i^z\right]$. Continuing with the above discussion for $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\mathbf{e}_{z+1}+\mathbf{e}_{z+2}\}}$, ……, $V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\cdots+\mathbf{e}_k\}}$, it can be shown that

$$\begin{aligned} &\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\cdots+\mathbf{e}_i\}}\right] \\ =&\sum_{d=0}^{z-1}[\begin{pmatrix} z-1 \\ d \end{pmatrix} P_g(d+i-z+1)] + \sum_{d=1}^{z-1}[\begin{pmatrix} z-1 \\ d \end{pmatrix} P_g(d)], \end{aligned} \tag{3.29}$$

where $0 \leq i \leq k-z$. Because we are working in the binary field, it can be shown that the above discussion covers all occurrences of z dimensional spaces.

Summarizing the above discussion, it follows that

$$\max_i\{\Pr\left[\overline{B_i^z}\right]\} = \max_{0\leq i\leq k-z} \Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\cdots+\mathbf{e}_{z+i}\}}\right], \tag{3.30}$$

where the values of $\Pr\left[\mathbf{x} \in V_{\{\mathbf{e}_1,\ldots,\mathbf{e}_{z-1},\mathbf{e}_z+\cdots+\mathbf{e}_{z+i}\}}\right]$ is given by (3.29).

Combining equations (3.25), (3.27), (3.29) and (3.30), the second part of the proof of Theorem 3.1 on the lower bound is also completed.

3.4 Simulation Results

In this section, we use simulations to validate the accuracy of the analytical results and the tightness of the bounds by plotting the decoding failure probability which is one minus the decoding success probability. The simulations are conducted in a simulator written in MATLAB. Each point shown in the figures is the average value obtained from 10000 simulations. The 95% confidence interval is shown in the figures too. For clarity, the simulation parameters adopted in this section are summarized in Table I.

Table 3.1: Simulation parameters

Rateless codes encoding parameters	
Number of source symbols k	$5, 10, 20, 40$ and 80
The degree distributions for LT codes	
Ideal soliton degree distribution	$\Omega_d = \frac{1}{d(d-1)}, 2 \leq d \leq k$ and $\Omega_1 = \frac{1}{k}$
Robust soliton degree distribution	$c = 0.1$ and $\delta = 0.05$
Expurgated sparse random LT code ensemble	$\Omega_d^{Sparse}, 1 \leq d \leq k$
Binomial degree distribution	$\Omega_d = \frac{\binom{k}{d}}{(2^k-1)}, 1 \leq d \leq k$

Analytical and simulation results are presented in Fig. 3.1 , 3.2 , 3.3 and 3.4 on the probability that not all 5 source symbols can be successfully received/decoded by a receiver as a function of reception overhead $\gamma_R = m_R/k$. The degree distributions of LT codes are chosen as the widely used ideal soliton degree distribution [31], robust soliton degree distribution [31] described on page 20 of this thesis the binomial degree distribution [89] and the expurgated sparse random LT code ensemble [57], which is expressed as:

$$\Omega_d^{Sparse} = \frac{\binom{k}{d}(P_0)^{k-d}(1-P_0)^d}{1-(P_0)^k}, 1 \leq d \leq k, \tag{3.31}$$

where P_0 is the probability of having a zero element in the generator matrix $\mathbf{G}^{LT}$ and is set as 0.45 in this chapter. The upper bound is calculated by using Eq. 3.2 and Eq. 4.2 with $n = k$. And the lower bound is calculated by using Eq. 3.1

and Eq. 4.29 with $n = k$. As shown in Fig. 3.1 , 3.2 , 3.3 and 3.4 for different degree distributions, our analytical results match the simulation results very well, which validate the accuracy of the analysis in this chapter. When the overhead is small, our proposed analytical bounds demonstrate better accuracy than the bounds proposed by Schotsch et al. in [57]. When the reception overhead γ_R is set to 1.4, for ideal soliton distribution, the decoding failure probability of LT codes equals 0.214; for robust soliton distribution, the decoding failure probability is 0.254; for expurgated sparse random LT code ensemble, the decoding failure probability increases to 0.19; for binomial degree distribution, the decoding failure probability becomes 0.184. The performance of LT codes with the binomial degree distribution outperforms those obtained with the other three degree distributions in terms of decoding failure probability. Furthermore, the analytical bounds of the decoding success probability of LT codes with the binomial degree distribution merge to the exact expression of decoding failure probability. Therefore, we will use LT codes with the binomial degree distribution in the following simulations of this chapter.

When the number of source symbols k varies from 5 to 80, our analytical results still match the simulation results very well as shown in Fig. 3.5(a) and 3.5(b). When the number of source symbols increases, the conclusion that LT codes can significantly reduce the overhead of reception required to meet the same performance objective.

3.5 Summary

In this chapter, we investigated the the decoding success probability of finite-Length LT codes under ML decoding. The decoding success probability is the probability that a receiver can successfully decode all k source symbols with ML decoding conditioned on the event that the receiver has successfully received a

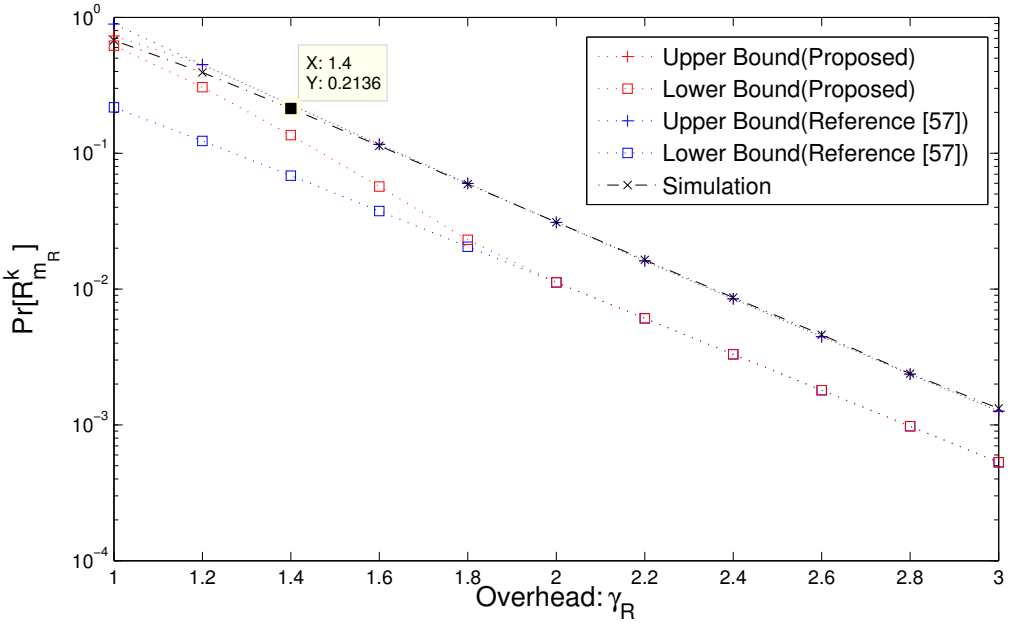


Figure 3.1: The decoding failure probabilities of LT codes with ideal soliton degree distribution [31] versus overhead γ_R.

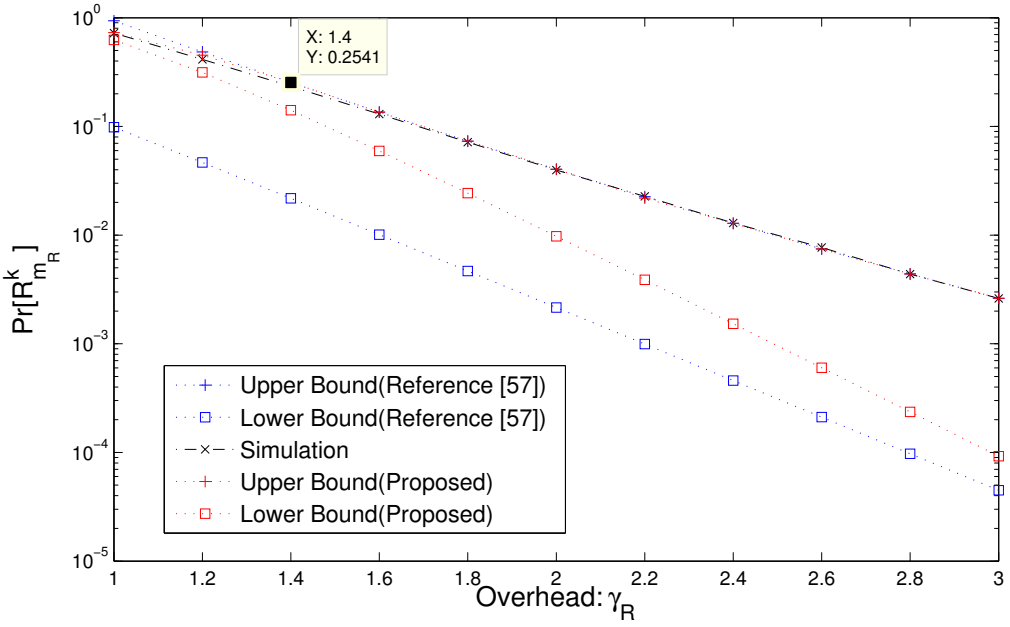


Figure 3.2: The decoding failure probabilities of LT codes with robust soliton degree distribution [31] versus overhead γ_R.

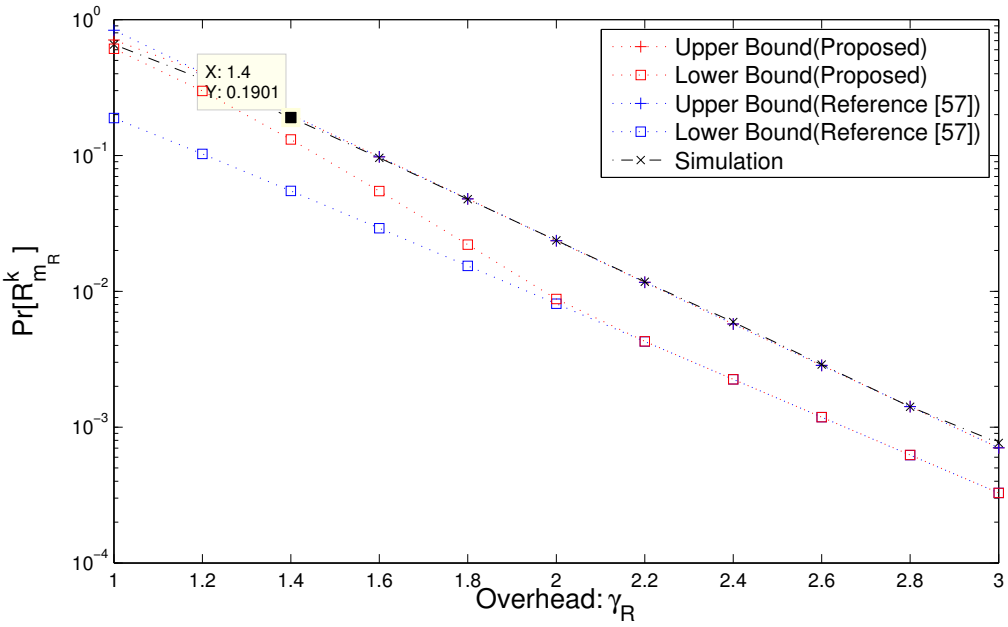


Figure 3.3: The decoding failure probabilities of LT codes with expurgated sparse random LT code ensemble [57] versus overhead γ_R.

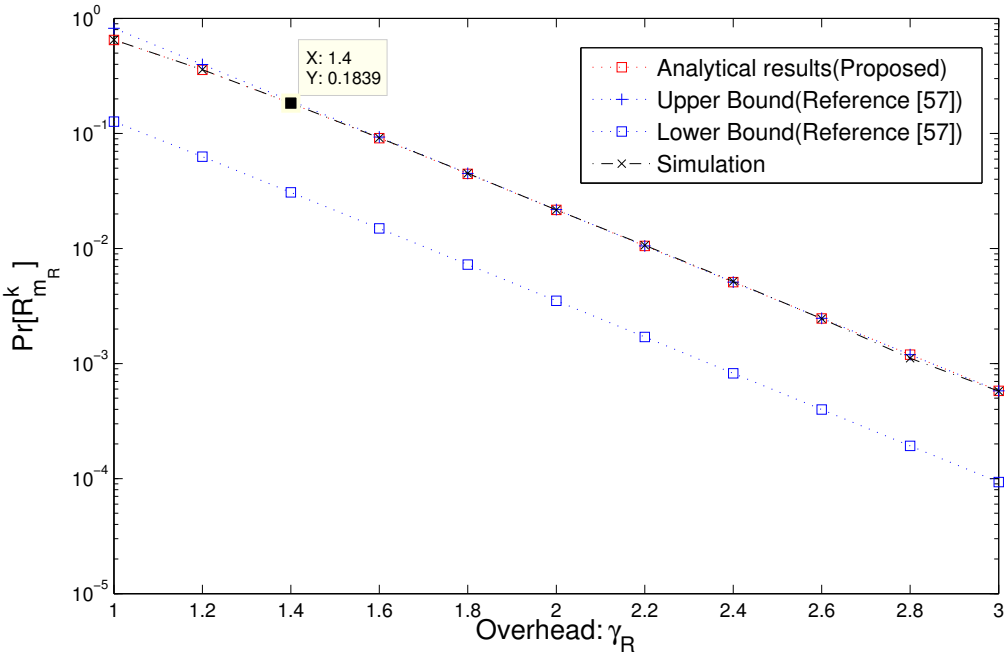
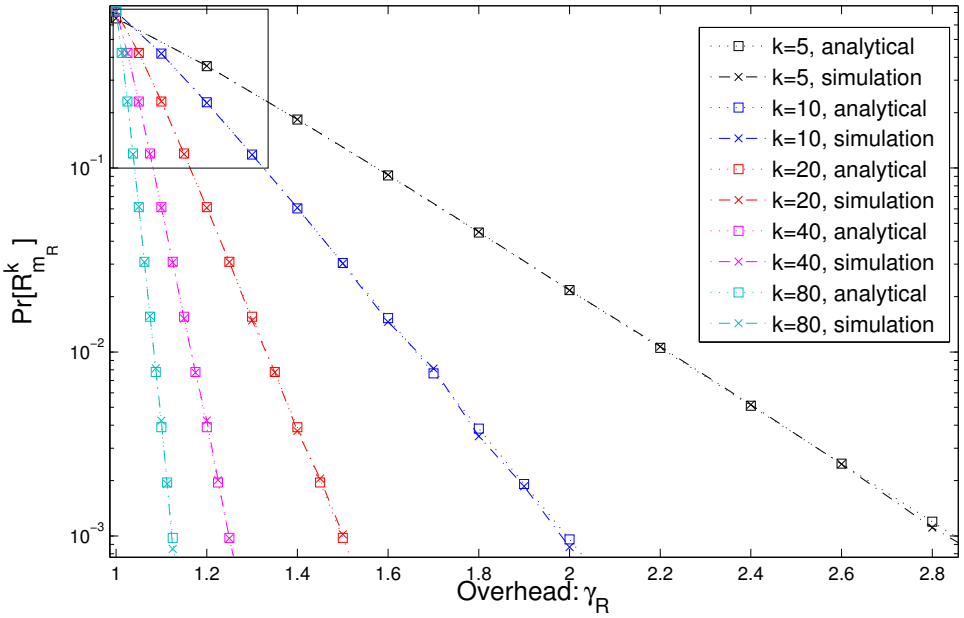


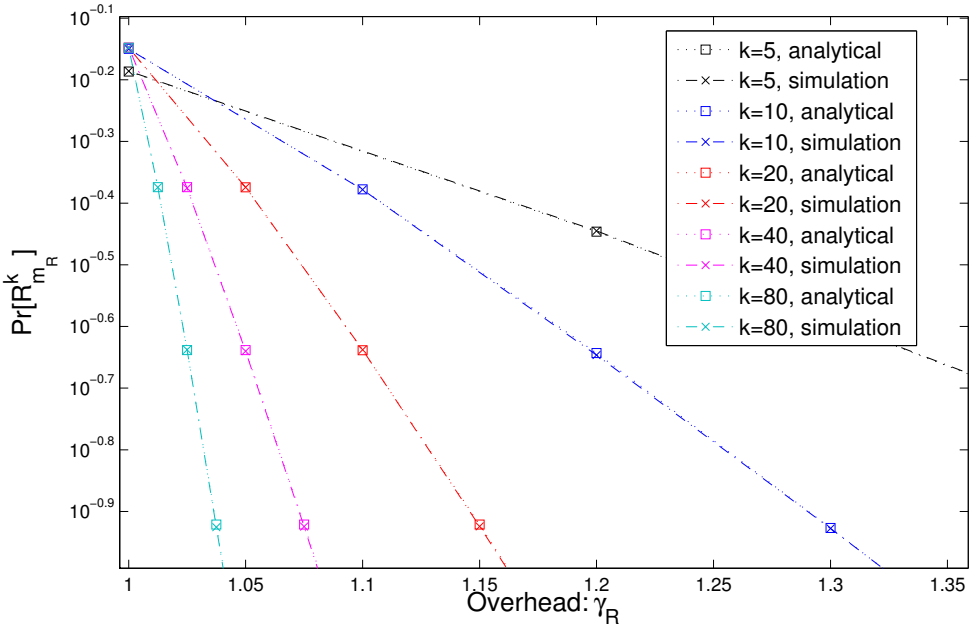
Figure 3.4: The decoding failure probabilities of LT codes with binomial degree distribution [89] versus overhead γ_R.

certain number of coded symbols. Specifically, if erasure channel is considered, the decoding of LT codes under ML decoding corresponds to solving a consistent system of linear equations over a binary field $GF(2)$, where the coefficients are given by the collected LT code generator matrix. In this chapter, we provide rigorous mathematical analysis on the rank profile of a random coefficient matrix, where each row vector is independently generated by using the LT encoding process. On the basis of this analysis, we derive upper and lower bounds on the decoding success probability of LT codes under ML decoding over BEC.

LT codes can be applied to wireless broadcast scenario. By using the analytical bounds on the decoding success probability of LT code under ML decoding, an upper and a lower bound on the probability that all receivers successfully decode all source symbols from the BS can be derived, which will be presented in Chapter 5. The technique and analysis developed in this chapter can be useful for designing broadcast strategies to deliver information of common interest to a large number of users efficiently and reliably.



(a) Full Scale



(b) Zoom of the rectangular box in (a)

Figure 3.5: The decoding failure probabilities of LT codes with the binomial degree distribution at different values of the overhead γ. The number of source symbols k is set to be 5, 10, 20, 40 and 80 respectively.

Chapter 4

Finite-Length Analysis of Raptor Codes

In the preceding chapter, we have investigated the decoding success probability of finite-length LT codes under *maximum likelihood* (ML) decoding. In this chapter, we take a further step by studying the decoding success probability of finite-length Raptor codes with a systematic *low-density generator matrix* (LDGM) code as the pre-code under ML decoding. Different from previous studies which rely on the use of approximation to obtain the pseudo upper bound on the performance of Raptor codes under ML decoding, this chapter provides analytical bounds on the decoding failure probability of Raptor codes under ML decoding. The decoding failure probability is defined as the probability that not all source symbols can be successfully decoded with a given number of successfully received coded symbols. These analytical bounds are derived by conducting a detailed mathematical analysis on the rank of the product of two random coefficient matrices. Based on analytical bounds on the decoding failure probability of Raptor codes, we can readily obtain analytical bounds on the decoding success probability of Raptor codes, which is unity minus decoding failure probability. Simulations are conducted to validate the accuracy of the analysis. More specifically, Raptor

codes with different degree distributions and pre-codes, are evaluated using the derived bounds with high accuracy. The results of this chapter appear in [J2].

4.1 Introduction

Rateless codes have been briefly introduced in Section 2.3. Because of the above mentioned advantages, rateless codes have the potential to replace the conventional *automatic repeat request* (ARQ) mechanism as a new mechanism of *transmission control protocol* (TCP) [66].

Among the known rateless codes, two codes stand out. One is the *Luby transform* (LT) codes, whose performance has been investigated in Chapter 3. The other one is the Raptor codes, which are the first class of fountain codes with linear time encoding and decoding complexities. Moreover, Raptor codes only require $O(1)$ time to generate a coded symbol [37]. Note that Raptor codes have already been standardized in *the 3rd Generation Partnership Project* (3GPP) to efficiently disseminate data over a broadcast/multicast network to provide *multimedia broadcast multicast services* (MBMS) [39].

Despite the successful application of Raptor codes in 3GPP, our understanding of Raptor codes is still incomplete due to the lack of complete theoretical analysis on the decoding error performance of Raptor codes. Without analytical results, the optimization of the degree distribution as well as the parameters for Raptor codes would be extremely difficult, if not impossible.

In this chapter, we investigate the performance of Raptor codes in terms of the decoding success probability. Without loss of generality, we investigate the decoding failure probability of Raptor codes under ML decoding first. The decoding failure probability is the probability that not all source symbols can be decoded by ML decoding with a given number of successfully received coded symbols. It is a commonly used performance metric in the performance analysis

of rateless codes. In [37], Shokrollahi analyzed the decoding failure probability of Raptor codes with finite length assuming the *belief propagation* (BP) decoding. The analysis relies on the computation of the failure probability of the LT codes under the BP decoding, which was derived in [52]. ML decoding, on the other hand, is more computationally demanding than the BP decoding for codes with a large length. The derivation of bounds on the decoding failure probability assuming ML decoding is however a significant problem, because it provides a benchmark on the optimum system performance for gauging the other decoding schemes. Furthermore, it is worth noting that in [57] a pseudo upper bound on the performance of Raptor codes under ML decoding is derived under the assumption that the number of erasures correctable by the pre-code is small. This approximation is accurate only when the rate of the pre-code is sufficiently high. So for a more general case, the decoding failure probability of Raptor codes still needs further investigation.

In this paper, we further treat Raptor codes by analyzing the decoding failure probability of Raptor codes, i.e., not all source symbols can be successfully decoded with ML decoding by a receiver with a given number of successfully received coded symbols, and verifying the derived results via simulations. The contributions of this work are summarized in the following:

- Firstly, this chapter provides the analytical results, i.e., an upper bound and a lower bound, on the decoding failure performance of Raptor codes with a systematic LDGM code as the pre-code under ML decoding, which is measured by the probability that not all source symbols can be successfully decoded by a receiver with a given number of successfully received coded symbols. The analytical results are derived by conducting an analysis on the rank of the product of two random coefficient matrices.

- Based on the upper and lower bounds on the decoding failure probability of Raptor codes, we can readily obtain the lower and upper bounds on the

decoding success probability of Raptor codes, which is unity minus decoding failure probability.

- Moreover, simulations are conducted to validate the accuracy of the proposed bounds. More specifically, Raptor codes with different degree distributions and pre-codes are evaluated, which establishes the accuracy of the bounds.

The rest of the chapter is organized as follows. In Section 4.2, a brief review of the encoding and decoding process of Raptor codes is given. In Section 4.3, a performance analysis of Raptor code is conducted by deriving an upper bound and a lower bound on the probability that not all source symbols can be successfully decoded by a receiver with a given number of successfully received coded symbols. Section 4.4 validates the analytical results through simulations, followed by concluding remarks in Section 4.5.

4.2 An Introduction to Raptor Codes

This section is provided to familiarize the readers with the basic idea of Raptor codes, and their efficient encoding and decoding algorithms.

The encoding process of Raptor codes is carried out in two phases: a) Encode k source symbols with a (n, k) error correction code, which is referred to as pre-code $\mathcal{C}$, to form n intermediate symbols; b) Encode the n intermediate symbols with an LT code. Each coded symbol is generated by the following encoding rules of LT codes. Firstly, a positive integer d (often referred to as the "degree" [31] of coded symbols) is drawn from the set of integers $\{1, ..., n\}$ according to a probability distribution $\mathbf{\Omega} = (\Omega_1, ..., \Omega_n)$, where Ω_d is the probability that d is picked and $\sum_{d=1}^{k} \Omega_d = 1$. Then, d distinct intermediate symbols are selected randomly and independently from the n intermediate symbols to form the coded symbol to be transmitted using the XOR operation [37, 31], where each interme-

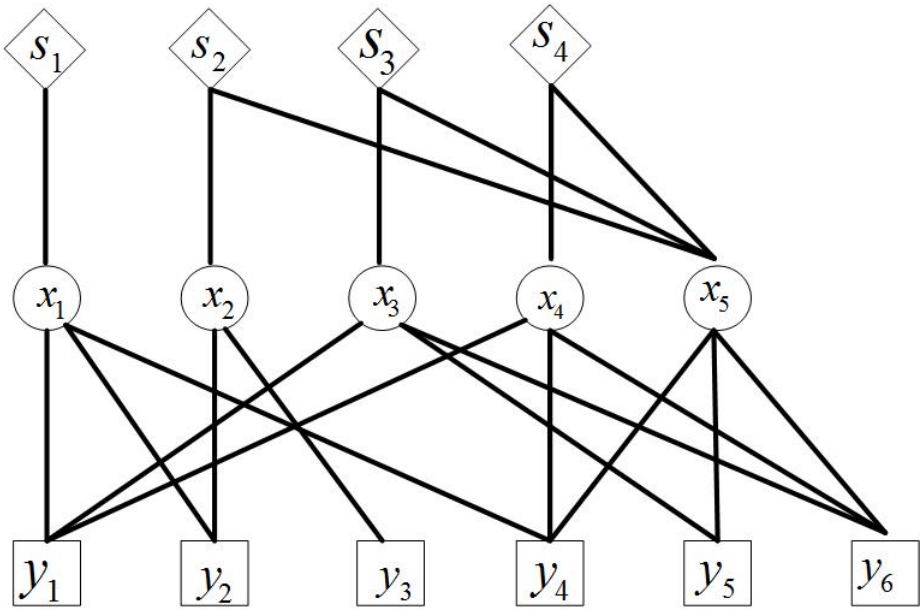


Figure 4.1: Two-stage structure of a Raptor code with a systematic pre-code.

diate symbol is selected with equal probability. A Raptor code with parameters $(k, \mathcal{C}, \mathbf{\Omega})$ is an LT code with distribution $\mathbf{\Omega} = (\Omega_1, ..., \Omega_n)$ on n symbols that are the coded symbols of the pre-code $\mathcal{C}$. An illustration of a Raptor code is given in Figure 4.1. In practice, the parity check matrix of the pre-code of Raptor codes is a deterministic matrix. For example, in 3GPP standard [39], the parity check matrix of the pre-code of the standardized Raptor codes is a systematic deterministic matrix. Using a systematic deterministic matrix as the pre-code of the standardized Raptor codes ensures that the parity check matrix of the pre-code is a full-rank matrix. However, it is difficult to obtain tractable analytical results of the decoding performance for such Raptor codes. Therefore, in this chapter we adopt a Raptor code ensemble with a semi-random (n, k, η) LDGM code as the pre-code for analytical tractability while ensuring that the parity check matrix of the pre-code is a full-rank matrix. The generator matrix of the pre-code, $\mathbf{G}^{\text{pre}}_{n\times k}$, can be written as $\mathbf{G}^{\text{pre}}_{n\times k} = [\mathbf{I}_k|\mathbf{P}_{k\times(n-k)}]^T$, where $\mathbf{I}_k$ is an identity matrix of size k, and $\mathbf{P}_{k\times(n-k)}$ is a k by $(n-k)$ matrix whose entries are independent and identically distributed (i.i.d) Bernoulli random variables with parameter η. Such

a code is denoted as an (n, k, η) LDGM code. Furthermore, we can obtain the parity check matrix of this LDGM code as $\mathbf{H}_{(n-k)\times n} = [\mathbf{P}_{(n-k)\times k}|\mathbf{I}_{(n-k)}]_{(n-k)\times n}$.

Let m_R, $(m_R \geq k)$, be the number of coded symbols that have already been successfully received by a receiver and $\gamma_R = \frac{m_R}{k}$, $(\gamma_R \geq 1)$ be the overhead of reception. When a coded symbol is received by a receiver, we use a $1 \times k$ binary row vector $\mathbf{g}_i^{\mathrm{LT}}\mathbf{G}^{\mathrm{pre}}$ to represent the coding information contained in the coded symbol, where $\mathbf{G}^{\mathrm{LT}}$ is a $k\gamma_R \times n$ binary matrix and $\mathbf{g}_i^{\mathrm{LT}}$ is the i^{th} row vector of $\mathbf{G}^{\mathrm{LT}}$ and $\mathbf{G}^{\mathrm{pre}}$ is a $n \times k$ binary matrix. Let $[\mathbf{G}]_{i,j}$ be the entry of the i^{th} row and the j^{th} column of a matrix $\mathbf{G}$. Particularly, $\left[\mathbf{g}_i^{LT}\right]_{1,j}$ is 1 if the coded symbol is a result of the XOR operation on the j^{th} intermediate symbol (and other intermediate symbols); otherwise $\left[\mathbf{g}_i^{LT}\right]_{1,j}$ equals 0. For $[\mathbf{G}^{\mathrm{pre}}]_{i,j}$, it is 1 if the i^{th} intermediate symbol is a result of the XOR operation on the j^{th} source symbol (and other source symbols); otherwise $[\mathbf{G}^{\mathrm{pre}}]_{i,j}$ equals 0. Therefore, a random row vector in this paper refers to the row vector of a randomly chosen coded symbol where the coded symbol is generated using the Raptor encoding process described above. Recall that $\mathbf{s} = (s_1, s_2, ..., s_k)$ represents the k source symbols to be transmitted. The coded symbol can be expressed as: $y_i = \mathbf{g}_i^{\mathrm{LT}}\mathbf{G}^{\mathrm{pre}}\mathbf{s}^T$, where "$\mathbf{s}^T$" is the transpose of $\mathbf{s}$.

Raptor codes can be decoded by using a variety of decoding algorithms. A typically used decoding algorithm for Raptor codes is the so-called "LT process" [31], but it is well known that the LT process is unable to decode all the source symbols which can be possibly recovered from information contained in the received coded symbols. For example, the LT process relies on the existence of at least one degree-one coded symbol to be received in order to start the decoding process. For Raptor codes with limited lengths, i.e. in the order of a few thousand, *maximum likelihood* (ML) decoding [54] has been used to replace the LT process to decode the source symbols. The performance of the ML decoding algorithm is the same as the Gaussian elimination. One way to apply Gaussian

elimination on Raptor code is to solve a system of linear equations given in the following [37, 86].

$$\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k}\mathbf{s}^{T}_{k\times 1} \;=\; \mathbf{y}_{k\gamma_R\times 1}, \tag{4.1}$$

where $\mathbf{y}_{k\gamma_R\times 1} = (y_1, y_2, ..., y_{k\gamma_R})^T$. Additionally, we can obtain the following Lemma:

Lemma 4.1. *A receiver can recover all k source symbols from the $k\gamma_R$ coded symbols using the ML decoding algorithm if and only if $(\mathbf{G}^{LT}_{k\gamma_R\times n}\mathbf{G}^{pre}_{n\times k})_{k\gamma_R\times k}$ is a full-rank matrix, i.e. its rank equals k [37].*

Note that in this paper, all algebraic operations and the associated analysis are conducted in a binary field. Denote by $A^k_{k\gamma_R}$ the event that a receiver can successfully decode all k source symbols conditioned on the event that the receiver has successfully received $k\gamma_R$ coded symbols. Obviously the event that $(\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k})_{k\gamma_R\times k}$ is a full-rank matrix is equivalent to the event $A^k_{k\gamma_R}$ happening. Let $\overline{A^k_{\gamma_R k}}$ be the complement of event $A^k_{k\gamma_R}$. The main result of this paper is summarized in Theorems 4.2 and 4.3.

4.3 Performance Analysis of Raptor Codes

In this subsection, we shall analyze the probability $\Pr\left[\overline{A^k_{\gamma_R k}}\right]$. Because of the equivalence between the event $A^k_{\gamma_R k}$ and the event that $(\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k})_{\gamma_R k\times k}$ is a full-rank matrix, the analysis of the decoding failure probability $P^{DF}_{k,n,\gamma_R} = \Pr\left[\overline{A^k_{\gamma_R k}}\right]$ is conducted by analyzing the probability that the rank of $(\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k})_{k\gamma_R\times k}$ is not k.

4.3.1 Upper Bound on the Decoding Failure Probability of Raptor Codes

In this subsection, we will derive an upper bound on the decoding failure probability of Raptor codes with a systematic (n, k, η) LDGM code as the pre-code, which is presented in the following theorem:

Theorem 4.2. *When a receiver successfully received $k\gamma_R$ coded symbols generated by using the Raptor code $(k, \mathcal{C}, \Omega(x))$ where $\mathcal{C}$ is an (n, k, η) LDGM code and the coded symbols received at a receiver are decoded using ML decoding, the probability that a receiver cannot successfully decode all k source symbols with $k\gamma_R, k\gamma_R \geq k$, received coded symbols, denoted by P^{DF}_{k,n,γ_R}, is upper bounded by*

$$P^{DF}_{k,n,\gamma_R} \leq \sum_{i=1}^{k} \binom{k}{i} \sum_{r=i}^{n-k+i} \left(J\left(r\right)\right)^{k\gamma_R} D\left(i, r\right), \tag{4.2}$$

where

$$J(r) = \sum_{d=1}^{n} \Omega_d \frac{\sum_{s=0,2,\ldots,2\left\lfloor \frac{d}{2} \right\rfloor} \binom{r}{s}\binom{n-r}{d-s}}{\binom{n}{d}}$$

and

$$D(i,r) = \binom{n-k}{r-i} \left[\frac{1+(1-2\eta)^i}{2}\right]^{n-k-r+i} \times \left[\frac{1-(1-2\eta)^i}{2}\right]^{r-i}$$

and Ω_d is the degree distribution of LT codes.

Proof. Our proof relies on the use of the union bound of the independent events that vectors in the column vector space of $\mathbf{G}^{\text{pre}}_{n\times k}$ are in the null space of $\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}$.

According to the property of the matrix product [87, Eq. (4.5.1)], we have

$$rank(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k})$$

$$= rank(\mathbf{G}^{\text{pre}}_{n\times k}) - \dim\{N(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}) \cap R(\mathbf{G}^{\text{pre}}_{n\times k})\}, \tag{4.3}$$

where $N(\bullet)$ is the right-hand null space of a matrix, $R(\bullet)$ is the column vector space generated by a matrix and $\dim\{\mathcal{V}\}$ represents the number of vectors in any basis for a vector space $\mathcal{V}$. It follows from the definition of $\mathbf{G}^{\text{pre}}_{n\times k}$ given earlier that the rank of $\mathbf{G}^{\text{pre}}_{n\times k}$ surely is k. It can be readily obtained that:

$$\begin{aligned} P^{DF}_{k,n,\gamma_R} &= \Pr[rank(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}) \neq k] \\ &= \Pr[\dim\{N(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}) \cap R(\mathbf{G}^{\text{pre}}_{n\times k})\} \neq 0]. \end{aligned} \tag{4.4}$$

For convenience let $W_{k\gamma_R,n,k}$ represent the event that $\dim\{N(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n})\cap R(\mathbf{G}^{\text{pre}}_{n\times k})\} \neq 0$. Now we need to analyze $P^{DF}_{k,n,\gamma_R} = \Pr[W_{k\gamma_R,n,k}]$. Provided that $\mathbf{G}^{\text{pre}}_{n\times k}$ is a systematic (n,k,η) LDGM code, the event $\dim\{N(\mathbf{G}^{LT}_{k\gamma_R\times n}) \cap R(\mathbf{G}^{pre}_{n\times k})\} \neq 0$, i.e., $W_{k\gamma_R,n,k}$, is equivalent to the event that at least one column vector from $R(\mathbf{G}^{\text{pre}}_{n\times k})$ is among $N(\mathbf{G}^{\text{LT}}_{k\gamma_R\times n})$, i.e., $\cup_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})}\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}$, where $\mathbf{x}$ is a column vector of $R(\mathbf{G}^{\text{pre}}_{n\times k})$. It can be readily shown that:

$$\begin{aligned} \Pr[W_{k\gamma_R,n,k}] &= \Pr\left[\cup_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})}\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}\right] \\ &\leq \sum_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}\right]. \end{aligned} \tag{4.5}$$

The column vector space $R(\mathbf{G}^{\text{pre}}_{n\times k})$ is partitioned into k subspace $(\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_k)$ and $\mathcal{V}_i$ is the subspace that contains all the column vectors which are a summation of i column vectors of $\mathbf{G}^{\text{pre}}_{n\times k}$. We denote by Γ_i as the set of indices of the column vectors in $\mathcal{V}_i$ and there are $\binom{k}{i}$ indices in Γ_i. Let $\mathbf{x}^i_a$ represent the $a^{th}, a \in \Gamma_i$ column vector in $\mathcal{V}_i$. It can be readily shown that:

$$\sum_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}] = \sum_{i=1}^{k}\sum_{a\in\Gamma_i} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a = \mathbf{0}]. \tag{4.6}$$

We can observe that $\mathbf{x}^i_a = \mathbf{G}^a_{n\times i}\mathbf{1}_i$ where $\mathbf{G}^a_{n\times i}$ is the matrix formed by i selected

column vectors from k column vectors of $\mathbf{G}^{\text{pre}}_{n\times k}$ and $\mathbf{1}_i$ represents the $i\times 1$ all one column vector. Let $|\mathbf{x}^i_a|$ represent the weight of column vector $\mathbf{x}^i_a$, considering the law of total probability, we have

$$\begin{aligned}&\Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a=\mathbf{0}]\\ =&\sum_{r=0}^{n}\Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a=\mathbf{0}\middle|\left|\mathbf{x}^i_a\right|=r\right]\Pr\left[\left|\mathbf{x}^i_a\right|=r\right].\end{aligned}\tag{4.7}$$

Firstly, we need to calculate $\Pr[|\mathbf{x}^i_a|=r]$. Provided $\mathbf{G}^{\text{pre}}_{n\times k}=[\mathbf{I}_k|\mathbf{P}_{k\times(n-k)}]^T$, in the first k entries of $\mathbf{G}^a_{n\times i}\mathbf{1}_i$ there are i ones. If $|\mathbf{x}^i_a|=r$, then there are $r-i$ ones in the last $n-k$ entries of $\mathbf{G}^a_{n\times i}\mathbf{1}_i$, .i.e, $\mathbf{P}^a_{(n-k)\times i}\mathbf{1}_i$. Hence we can obtain that

$$\Pr\left[\left|\mathbf{x}^i_a\right|=r\right] = \Pr\left[\left|\mathbf{P}^a_{(n-k)\times i}\mathbf{1}_i\right|=(r-i)\right],\tag{4.8}$$

and $i\le r\le n-k+i$. The rows of $\mathbf{P}^a_{(n-k)\times i}$, i.e., $\mathbf{p}_j, 1\le j\le(n-k)$, are random binary row vectors, which are generated independently. Each entry of $\mathbf{P}^a_{(n-k)\times i}$ is an independent and identically distributed (i.i.d) Bernoulli random variable with parameter η. Therefore, $\Pr[\mathbf{p}_j\mathbf{1}_i=0]=\Pr[\mathbf{p}_{k,k\neq j}\mathbf{1}_i=0]$. The event that the j^{th} entry in $\mathbf{x}^i_a$ is zero is equivalent to the event that there are even number of ones in row vector $\mathbf{p}_j$. We have

$$\begin{aligned}\Pr[\mathbf{p}_j\mathbf{1}_i=0] &= \Pr\left[|\mathbf{p}_j| \text{ is even }\right]\\ &= \sum_{s=0,2,\ldots,2\left\lfloor\frac{i}{2}\right\rfloor}\binom{i}{s}\eta^s(1-\eta)^{(i-s)}\\ &= \frac{[(\eta+(1-\eta))^i+(-\eta+(1-\eta))^i]}{2}\\ &= \frac{1+(1-2\eta)^i}{2}.\end{aligned}\tag{4.9}$$

There are $\binom{n-k}{r-i}$ possible *combinations* for $r-i$ ones in the last $n-k$ entries. It

can be readily shown that:

$$\begin{aligned} & \Pr\left[\left|\mathbf{P}^a_{(n-k)\times i}\mathbf{1}_i\right| = (r-i)\right] \\ = \; & \tbinom{n-k}{r-i}\{\Pr[\mathbf{p}_j\mathbf{1}_i = 0]\}^{n-k-r+i} \\ & \times\{1-\Pr[\mathbf{p}_j\mathbf{1}_i = 0]\}^{r-i}. \end{aligned} \tag{4.10}$$

Combining equations (4.8), (4.9) and (4.10), we can obtain that

$$\begin{aligned} & D(i,r) = \Pr\left[\left|\mathbf{x}^i_a\right| = r\right] \\ = \; & \tbinom{n-k}{r-i}\left[\frac{1+(1-2\eta)^i}{2}\right]^{n-k-r+i} \\ & \times\left[\frac{1-(1-2\eta)^i}{2}\right]^{r-i}. \end{aligned} \tag{4.11}$$

For $\mathbf{x}^i_a, \mathbf{x}^i_{b,b\neq a} \in \mathcal{V}_i$, $\mathbf{P}^a_{(n-k)\times i}$ and $\mathbf{P}^b_{(n-k)\times i}$ have the same probability to form the same matrix formation. So we can obtain that $\Pr\left[\left|\mathbf{P}^a_{(n-k)\times i}\mathbf{1}_i\right| = (r-i)\right] = \Pr\left[\left|\mathbf{P}^b_{(n-k)\times i}\mathbf{1}_i\right| = (r-i)\right]$, in turn $\Pr\left[|\mathbf{x}^i_a| = r\right] = \Pr\left[|\mathbf{x}^i_b| = r\right]$. Now, we calculate $\Pr\left[\mathbf{G}^{LT}_{k\gamma_R\times n}\mathbf{x}^i_a = 0 \mid |\mathbf{x}^i_a| = r\right]$. The rows of $\mathbf{G}^{\mathrm{LT}}_{\gamma_R k\times n}$, i.e., $\mathbf{g}^{\mathrm{LT}}_j, 1 \leq j \leq k\gamma_R$, are random binary row vectors, which are generated independently. We have

$$\begin{aligned} & \Pr\left[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a = \mathbf{0}\middle|\left|\mathbf{x}^i_a\right| = r\right] \\ = & \left\{\Pr\left[\mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^i_a = 0\middle|\left|\mathbf{x}^i_a\right| = r\right]\right\}^{k\gamma_R}. \end{aligned} \tag{4.12}$$

The degree of $\mathbf{g}^{\mathrm{LT}}_j$, i.e. the number of non-zero elements of $\mathbf{g}^{\mathrm{LT}}_j$, is chosen according to the pre-defined degree distribution $\mathbf{\Omega} = (\Omega_1, ..., \Omega_n)$ and each non-zero element is then placed randomly and uniformly into $\mathbf{g}^{\mathrm{LT}}_j$. It can be readily obtained that

$$\begin{aligned} & \Pr\left[\mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^i_a = 0\middle|\left|\mathbf{x}^i_a\right| = r\right] \\ = \; & \sum_{d=1}^{n}\Omega_d\Pr\left[\mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^i_a = 0\middle|\left|\mathbf{x}^i_a\right| = r, \left|\mathbf{g}^{\mathrm{LT}}_j\right| = d\right]. \end{aligned} \tag{4.13}$$

Let $\mathbf{r}_j^i = (\mathbf{g}_{j1}^{\mathrm{LT}}\mathbf{x}_{a1}^i, \mathbf{g}_{j2}^{\mathrm{LT}}\mathbf{x}_{a2}^i, ..., \mathbf{g}_{jn}^{\mathrm{LT}}\mathbf{x}_{an}^i)$, where $\mathbf{g}_{jk}^{\mathrm{LT}}$ is $\left[\mathbf{g}_j^{\mathrm{LT}}\right]_{1,k}$ and $\mathbf{x}_{ak}^i$ is $[\mathbf{x}_a^i]_{k,1}$. Then, we can obtain that

$$\begin{aligned} & \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}_a^i = 0 \middle| \left|\mathbf{x}_a^i\right| = r, \left|\mathbf{g}_j^{\mathrm{LT}}\right| = d\right] \\ = \; & \Pr\left[\left|\mathbf{r}_j^i\right| \text{ is even } \middle| \left|\mathbf{x}_a^i\right| = r, \left|\mathbf{g}_j^{\mathrm{LT}}\right| = d\right] \\ = \; & \frac{\sum_{s=0,2,...,2\left\lfloor\frac{d}{2}\right\rfloor}\binom{r}{s}\binom{n-r}{d-s}}{\binom{n}{d}}. \end{aligned} \tag{4.14}$$

Combining equations (4.13) and (4.14), we can obtain that

$$\begin{aligned} J(r) &= \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}_a^i = 0 \middle| \left|\mathbf{x}_a^i\right| = r\right] \\ &= \sum_{d=1}^{n} \Omega_d \frac{\sum_{s=0,2,...,2\left\lfloor\frac{d}{2}\right\rfloor}\binom{r}{s}\binom{n-r}{d-s}}{\binom{n}{d}}. \end{aligned} \tag{4.15}$$

Inserting equation (4.12) into (4.15), it can be obtained that

$$\Pr\left[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = \mathbf{0} \middle| \left|\mathbf{x}_a^i\right| = r\right] = [J(r)]^{k\gamma_R}. \tag{4.16}$$

We can obtain that $\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = \mathbf{0} \mid |\mathbf{x}_a^i| = r]$ is only determined by the weight of $\mathbf{x}_a^i$ rather than which i column vectors are chosen from $\mathbf{G}_{n\times k}^{\mathrm{pre}}$ to obtain the summation $\mathbf{x}_a^i$. So we can conclude that $\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = \mathbf{0}] = \Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_b^i = \mathbf{0}]$. Recall that there are $\binom{k}{i}$ indices in $\hat{\mathrm{I}}“_i$. Inserting equations (4.11) and (4.16) into (4.7) and combining with equation (4.6), yields the following results

$$\begin{aligned} & P_{k,n,\gamma_R}^{DF} = \Pr[W_{k\gamma_R,n,k}] \\ \leq \; & \sum_{i=1}^{k}\sum_{a\in\hat{\mathrm{I}}“_i} \Pr\left[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = \mathbf{0}\right] \\ = \; & \sum_{i=1}^{k}\binom{k}{i}\sum_{r=i}^{n-k+i}\left[\sum_{d=1}^{n}\Omega_d\frac{\sum_{s=0,2,...,2\left\lfloor\frac{d}{2}\right\rfloor}\binom{r}{s}\binom{n-r}{d-s}}{\binom{n}{d}}\right]^{k\gamma_R} \\ & \times\binom{n-k}{r-i}\left[\frac{1+(1-2\eta)^i}{2}\right]^{n-k-r+i}\left[\frac{1-(1-2\eta)^i}{2}\right]^{r-i}, \end{aligned} \tag{4.17}$$

which proves the assertion. □

4.3.2 Lower Bound on the Decoding Failure Probability of Raptor Codes

In addition to the above upper bound, we can also derive lower bounds on the decoding failure probability of Raptor codes with a systematic (n, k, η) LDGM code as the pre-code, which are presented in the following theorems:

Theorem 4.3. *When a receiver successfully received $k\gamma_R$ coded symbols generated by using the Raptor code $(k, \mathcal{C}, \Omega(x))$ where $\mathcal{C}$ is an (n, k, η) LDGM code and the coded symbols received at a receiver are decoded using ML decoding, the probability that a receiver cannot successfully decode all k source symbols with $k\gamma_R, k\gamma_R \geq k$, received coded symbols, denoted by P^{DF}_{k,n,γ_R}, is lower bounded by:*

$$\begin{aligned}
& P^{DF}_{k,n,\gamma_R} \\
\geq & \sum_{i=1}^{k} \binom{k}{i} \sum_{r=i}^{n-k+i} (J(r))^{k\gamma_R} D(i,r) \\
& - \frac{1}{2} \sum_{i=1}^{k} \binom{k}{i} \sum_{w_0=0}^{i} \sum_{w_1=i-w_0} \sum_{w_2=0}^{k-i} \mathbf{1}(w_0+w_2)\mathbf{1}(w_1+w_2) \\
& \times \binom{i}{w_0}\binom{k-i}{w_2} \{ \sum_{r_0=w_0}^{n-k+w_0} \sum_{r_1=w_1}^{n-k+w_1} \sum_{r_0=w_2}^{n-k+w_2} D(w_0,r_0) D(w_1,r_1) \\
& \times D(w_2,r_2) [J(r_0)J(r_1)J(r_2) + \overline{J}(r_0)\overline{J}(r_1)\overline{J}(r_2)] \}^{k\gamma_R}, \qquad (4.18)
\end{aligned}$$

where

$$\mathbf{1}(x) := \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise}, \end{cases}$$

$\overline{J}(\cdot) = 1 - J(\cdot)$, *$D(w_0, r_0)$ is defined in equation (4.11) and $J(r_0)$ is defined in equation (4.15).*

Proof. Similar to [66, Lemma 10], by using the Bonferroni inequality [88], we can

obtain a lower bound of $\Pr[W_{k\gamma_R,n,k}]$ as:

$$\begin{aligned} & P^{DF}_{k,n,\gamma_R} = \Pr[W_{k\gamma_R,n,k}] \\ = \; & \Pr[\cup_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}] \\ \overset{(a)}{\geq} \; & \sum_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0}] \\ & -\frac{1}{2}\sum_{\mathbf{x},\mathbf{y}\in R(\mathbf{G}^{\text{pre}}_{n\times k}),\mathbf{x}\neq\mathbf{y}} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0} \,\&\, \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{y} = \mathbf{0}]. \end{aligned} \tag{4.19}$$

where $\mathbf{x} = \mathbf{G}^{\text{pre}}_{n\times k}\mathbf{a}, \mathbf{a} \in GF(2)^k$ and $\mathbf{y} = \mathbf{G}^{\text{pre}}_{n\times k}\mathbf{b}, \mathbf{b} \in GF(2)^k \backslash \mathbf{a}$. The first term can be calculated by using Theorem 4.2. Recall that $\mathcal{V}_i$ is a subspace that contains all the column vectors which are summation of i column vectors of $\mathbf{G}^{\text{pre}}_{n\times k}$, $\hat{\text{I}}“_i$ is the set of indices of the column vectors in $\mathcal{V}_i$ and $\mathbf{x}^i_a$ represents the $a^{th}, a \in \hat{\text{I}}“_i$ column vectors in $\mathcal{V}_i$. It can be readily shown that:

$$\begin{aligned} & \sum_{\mathbf{x},\mathbf{y}\in R(\mathbf{G}^{\text{pre}}_{n\times k}),\mathbf{x}\neq\mathbf{y}} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0} \,\&\, \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{y} = \mathbf{0}] \\ = & \sum_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \sum_{\mathbf{y}\in R(\mathbf{G}^{\text{pre}}_{n\times k})\backslash\mathbf{x}} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x} = \mathbf{0} \,\&\, \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{y} = \mathbf{0}] \\ = & \sum_{i=1}^{k} \sum_{a\in\hat{\text{I}}“_i} \sum_{\mathbf{y}\in R(\mathbf{G}^{\text{pre}}_{n\times k})\backslash\mathbf{x}^i_a} \Pr[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a = \mathbf{0} \,\&\, \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{y} = \mathbf{0}], \end{aligned} \tag{4.20}$$

where $\mathbf{x}^i_a = \mathbf{G}^{\text{pre}}_{n\times k}\mathbf{a}, |\mathbf{a}| = i$. Recall that $\mathbf{y} = \mathbf{G}^{\text{pre}}_{n\times k}\mathbf{b}, \mathbf{b} \in GF(2)^k$. We define three binary vectors $\mathbf{z}_0$, $\mathbf{z}_1$, and $\mathbf{z}_2 \in GF(2)^k$ such that for $t = 1, ..., k, \mathbf{z}_0(t) = 1$ if and only if $\mathbf{a}(t) = 1$ and $\mathbf{b}(t) = 1$, $\mathbf{z}_1(t) = 1$ if and only if $\mathbf{a}(t) = 1$ and $\mathbf{b}(t) = 0$, and $\mathbf{z}_2(t) = 1$ if and only if $\mathbf{a}(t) = 0$ and $\mathbf{b}(t) = 1$. Let w_0, w_1 and w_2 be the weights of vectors $\mathbf{z}_0$, $\mathbf{z}_1$, and $\mathbf{z}_2$, respectively. For $\mathbf{x}^i_a$, we have $\mathbf{z}_0 + \mathbf{z}_1 = \mathbf{a}$ and $\mathbf{z}_0 + \mathbf{z}_2 = \mathbf{b}$. Hence we can obtain:

$$\begin{aligned} & \Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}^i_a = \mathbf{0} \,\&\, \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{y} = \mathbf{0}\right] \\ = & \Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_0 = \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_1 = \mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_2\right. \end{aligned}$$

$$\left| |\mathbf{z}_0| = w_0 \& |\mathbf{z}_1| = w_1 \& |\mathbf{z}_2| = w_2 \right]. \tag{4.21}$$

Let $I_{\mathbf{z}} = \{i_{\mathbf{z}1}, i_{\mathbf{z}2}, ..., i_{\mathbf{z}\tau}\}$ be the set of indices such that $t \in I_{\mathbf{z}}$ for $\mathbf{z}(t) = 1$, we can obtain the sets of indices of vectors $\mathbf{z}_0$, $\mathbf{z}_1$, and $\mathbf{z}_2$ as $I_{\mathbf{z}_0}$, $I_{\mathbf{z}_1}$ and $I_{\mathbf{z}_2}$. Corresponding to the three sets $I_{\mathbf{z}_0}$, $I_{\mathbf{z}_1}$ and $I_{\mathbf{z}_2}$, each column of the matrix $\mathbf{G}^{pre}_{n\times k}$, $\mathbf{g}^{pre}_i, 1 \leq i \leq k$, can be divided into four mutually exclusive parts, $\mathbf{g}_{\mathbf{z}_0}$, $\mathbf{g}_{\mathbf{z}_1}$, $\mathbf{g}_{\mathbf{z}_2}$ and $\cup_{1\leq i\leq k}\mathbf{g}^{pre}_i \backslash (\mathbf{g}_{\mathbf{z}_0} \cup \mathbf{g}_{\mathbf{z}_1} \cup \mathbf{g}_{\mathbf{z}_2})$, i.e., $\mathbf{g}_{\mathbf{z}_0} \cap \mathbf{g}_{\mathbf{z}_1} = \{0\}$. Let $\mathbf{g}_{\mathbf{z}_0}$ be the subset of $\cup_{1\leq i\leq k}\mathbf{g}^{pre}_i$ such that all the elements of this subset are selected from $\cup_{1\leq i\leq k}\mathbf{g}^{pre}_i$ according to the indices in set $I_{\mathbf{z}_0}$ and $\mathbf{G}^{pre}_{\mathbf{z}_0}$ be the matrix whose columns are elements of $\mathbf{g}_{\mathbf{z}_0}$. The length of $\mathbf{g}_{\mathbf{z}_0}$ is w_0. The same operation is applied to the formation of $\mathbf{g}_{\mathbf{z}_1}$ and $\mathbf{g}_{\mathbf{z}_2}$, in which the elements are selected according to the indices in the set $I_{\mathbf{z}_1}$ and $I_{\mathbf{z}_2}$, and have length w_1 and w_2, respectively. Let $\mathbf{x}^{w_0} = \mathbf{G}^{pre}_{\mathbf{z}_0}\mathbf{1}_{w_0}$, $\mathbf{x}^{w_1} = \mathbf{G}^{pre}_{\mathbf{z}_1}\mathbf{1}_{w_1}$ and $\mathbf{x}^{w_2} = \mathbf{G}^{pre}_{\mathbf{z}_2}\mathbf{1}_{w_2}$. Equivalently, equation (4.27) can be rewritten as,

$$\begin{aligned}
&\Pr\Big[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k}\mathbf{z}_0 = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k}\mathbf{z}_1 = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{G}^{\mathrm{pre}}_{n\times k}\mathbf{z}_2 \\
&\quad \Big| |\mathbf{z}_0| = w_0, |\mathbf{z}_1| = w_1, |\mathbf{z}_2| = w_2\Big] \\
&= \Pr[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_0} = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_1} = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_2}].
\end{aligned} \tag{4.22}$$

Recall that the rows of $\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}$, i.e., $\mathbf{g}^{\mathrm{LT}}_j, 1 \leq j \leq k\gamma_R$, are random binary row vectors, which are generated independently. We have

$$\begin{aligned}
&\Pr\left[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_0} = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_1} = \mathbf{G}^{\mathrm{LT}}_{k\gamma_R\times n}\mathbf{x}^{w_2}\right] \\
&= \left\{\Pr\left[\mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_0} = \mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_1} = \mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_2}\right]\right\}^{k\gamma_R}.
\end{aligned} \tag{4.23}$$

According to the law of total probability, we have

$$\Pr\left[\mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_0} = \mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_1} = \mathbf{g}^{\mathrm{LT}}_j\mathbf{x}^{w_2}\right]$$

$$
\begin{aligned}
&= \sum_{r_0=w_0}^{n-k+w_0} \sum_{r_1=w_1}^{n-k+w_1} \sum_{r_0=w_2}^{n-k+w_2} \Pr[|\mathbf{x}^{w_0}| = r_0] \\
&\times \Pr[|\mathbf{x}^{w_1}| = r_1] \Pr[|\mathbf{x}^{w_2}| = r_2] \\
&\times \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} \right. \\
&\quad \left. \Big| \, |\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2 \right]. \qquad (4.24)
\end{aligned}
$$

For $\Pr[|\mathbf{x}^{w_0}| = r_0]$, this can be calculated by using equation (4.11). Because all algebraic operations are conducted in a binary field, $\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0}$ can only be 1 or 0. Equation (4.23) can be further written as :

$$
\begin{aligned}
&\Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} \right. \\
&\quad \left. \Big| \, |\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2 \right] \\
&= \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = 0, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = 0, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} = 0 \right. \\
&\quad \left. \Big| \, |\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2 \right] \\
&+ \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = 1, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = 1, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} = 1 \right. \\
&\quad \left. \Big| \, |\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2 \right]. \qquad (4.25)
\end{aligned}
$$

Recall that $\mathbf{x}^{w_0} = \mathbf{G}_{\mathbf{z}_0}^{\mathrm{pre}}\mathbf{1}_{w_0}$, $\mathbf{x}^{w_1} = \mathbf{G}_{\mathbf{z}_1}^{\mathrm{pre}}\mathbf{1}_{w_1}$, $\mathbf{x}^{w_2} = \mathbf{G}_{\mathbf{z}_2}^{\mathrm{pre}}\mathbf{1}_{w_2}$ and the columns of $\mathbf{G}_{\mathbf{z}_0}^{\mathrm{pre}}$, $\mathbf{G}_{\mathbf{z}_1}^{\mathrm{pre}}$, $\mathbf{G}_{\mathbf{z}_2}^{\mathrm{pre}}$ are mutually exclusive to each other. So the event that $|\mathbf{x}^{w_0}| = r_0$ is independent of the event that $|\mathbf{x}^{w_1}| = r_1$ or $|\mathbf{x}^{w_2}| = r_2$ and the event that $\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = 1$ is independent of the event that $\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = 1$ or $\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} = 1$. Conditioned on $|\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2$, the first part in equation (4.25) can be expressed as:

$$
\begin{aligned}
&\Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = 0, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = 0, \mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} = 0 \right. \\
&\quad \left. \Big| \, |\mathbf{x}^{w_0}| = r_0, |\mathbf{x}^{w_1}| = r_1, |\mathbf{x}^{w_2}| = r_2 \right] \\
&= \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0} = 0 \Big| \, |\mathbf{x}^{w_0}| = r_0 \right] \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_1} = 0 \Big| \, |\mathbf{x}^{w_1}| = r_1 \right] \\
&\quad \Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_2} = 0 \Big| \, |\mathbf{x}^{w_2}| = r_2 \right]. \qquad (4.26)
\end{aligned}
$$

Based on the *previous* analysis, we know that $\Pr[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0}=0\Big||\mathbf{x}^{w_0}|=r_0]$ only relates to parameter r_0. Let $D(w_0,r_0)=\Pr[|\mathbf{x}^{w_0}|=r_0]$ and $J(r_0)=\Pr[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}^{w_0}=0|\,|\mathbf{x}^{w_0}|=r_0]$. For $J(r_0)$, it can be calculated by using equations (4.13) and (4.14). Based on the *previous* analysis, we know that $J(r_0)$ only relates to parameter r_0 and $D(w_0,r_0)$ is affected by parameter r_0 and w_0. Hence for the same parameters w_0, w_1 and w_2, equation (4.22) has the same result. Because $\mathbf{x}_a^i\neq\mathbf{y}$, we can obtain that $w_1+w_2\neq 0$ and $w_0+w_2\neq 0$. For $\mathbf{x}_a^i$, when $|\mathbf{z}_0|=w_0$, we have $w_1=i-w_0$ and there are $\binom{i}{w_0}$ possible combinations of $\mathbf{z}_0$. For $\mathbf{z}_2$, there are $\binom{k-i}{w_2}$ possible combinations of $\mathbf{z}_2$ when $|\mathbf{z}_2|=w_2$. Inserting equations (4.22), (4.24), (4.23), (4.25) and (4.26) into (4.21), we can obtain:

$$\begin{aligned}
&\sum_{\mathbf{y}\in R(\mathbf{G}_{n\times k}^{\mathrm{pre}})\backslash\mathbf{x}_a^i}\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}]\\
&=\sum_{w_0=0}^{i}\sum_{w_1=i-w_0}\sum_{w_2=0}^{k-i}\mathbf{1}(w_0+w_2)\mathbf{1}(w_1+w_2)\tbinom{i}{w_0}\tbinom{k-i}{w_2}\\
&\times\{\sum_{r_0=w_0}^{n-k+w_0}\sum_{r_1=w_1}^{n-k+w_1}\sum_{r_0=w_2}^{n-k+w_2}D(w_0,r_0)D(w_1,r_1)D(w_2,r_2)\\
&\quad[J(r_0)J(r_1)J(r_2)+\overline{J}(r_0)\overline{J}(r_1)\overline{J}(r_2)]\}^{\gamma_R k},
\end{aligned}\tag{4.27}$$

where $\mathbf{1}(x):=\begin{cases}0 & \text{if } x=0\\ 1 & otherwise\end{cases}$. For $\mathbf{x}_a^i,\mathbf{x}_{b,b\neq a}^i\in\mathcal{V}_i$, the probability $\sum_{\mathbf{x}_a^i\neq\mathbf{y}}\Pr\left[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}\right]$ is affected by parameter i. So we can obtain that $\sum_{\mathbf{x}_a^i\neq\mathbf{y}}\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}]=\sum_{\mathbf{x}_b^i\neq\mathbf{y}}\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}]$. Recall that there are $\binom{k}{i}$ indices in $\hat{\mathrm{I}}^{i}$. We can obtain that

$$\begin{aligned}
&\sum_{\mathbf{x},\mathbf{y}\in R(\mathbf{G}_{n\times k}^{\mathrm{pre}}),\mathbf{x}\neq\mathbf{y}}\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}\,]\\
&=\sum_{i=1}^{k}\sum_{a\in\hat{\mathrm{I}}^{i}}\sum_{\mathbf{y}\in R(\mathbf{G}_{n\times k}^{\mathrm{pre}})\backslash\mathbf{x}_a^i}\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i=\mathbf{0}\,\&\,\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{y}=\mathbf{0}]\\
&=\sum_{i=1}^{k}\tbinom{k}{i}\sum_{w_0=0}^{i}\sum_{w_1=i-w_0}\sum_{w_2=0}^{k-i}\mathbf{1}(w_0+w_2)\mathbf{1}(w_1+w_2)
\end{aligned}$$

$$\times \binom{i}{w_0}\binom{k-i}{w_2}\{\sum_{r_0=w_0}^{n-k+w_0}\sum_{r_1=w_1}^{n-k+w_1}\sum_{r_0=w_2}^{n-k+w_2} D(w_0,r_0)D(w_1,r_1)$$
$$\times D(w_2,r_2)[J(r_0)J(r_1)J(r_2)+\overline{J}(r_0)\overline{J}(r_1)\overline{J}(r_2)]\}^{\gamma_R k}. \quad (4.28)$$

The proof of Theorem 4.3 is completed. □

The computation complexity of the above equation is high, i.e., $O(\frac{1}{8}n^6k^3(n-k)^3)$. We derive another lower bound whose computation complexity is decreased significantly.

Theorem 4.4. *When a receiver successfully received $k\gamma_R$ coded symbols generated by using the Raptor code $(k, \mathcal{C}, \Omega(x))$ where $\mathcal{C}$ is an (n, k, η) LDGM code and the coded symbols received at a receiver are decoded using ML decoding, the probability that a receiver cannot successfully decode all k source symbols with $k\gamma_R, k\gamma_R \geq k$, received coded symbols, denoted by P^{DF}_{k,n,γ_R}, is lower bounded by:*

$$\begin{aligned} & P^{DF}_{k,n,\gamma_R} \\ \geq & \sum_{i=1}^{k}\binom{k}{i}\sum_{r=i}^{n-k+i}\left[\sum_{d=1}^{n}\Omega_d\frac{\binom{n-r}{d}}{\binom{n}{d}}\right]^{k\gamma} \\ & \times\binom{n-k}{r-i}\left[(1-\eta)^i\right]^{n-k-r+i}\left[1-(1-\eta)^i\right]^{r-i}., \end{aligned} \quad (4.29)$$

Proof. Similar as that in [57, Theorem 3.18], by using the idea that the k source symbols cannot be recovered if at least one *source node* (SN) cannot be recovered. A lower bound on the of $\Pr[W_{k\gamma,n,k}]$ is therefore given by the probability that there exist SNs are not connected to any of the $k\gamma$ independent output nodes (ONs) through the n intermediate nodes (INs)

$$\begin{aligned} & P^{DF}_{k,n,\gamma} = \Pr[W_{k\gamma,n,k}] \\ \geq & \Pr[\cup_{i\in\{1,\ldots,k\}}\text{i SNs are not connected to the k}\gamma\text{ ONs}]. \end{aligned} \quad (4.30)$$

The probability that the i particular (fixed but arbitrary) SNs are connected to

some r of the n INs, i.e., the probability that the $r-i$ particular rows of $\mathbf{P}_{(n-k)\times k}$ have i all-zero columns, is given by

$$\begin{aligned} & \Pr[\text{i SNs are connected to the r INs}] \\ = \ & \binom{n-k}{r-i}\left[(1-\eta)^i\right]^{n-k-r+i}\left[1-(1-\eta)^i\right]^{r-i}. \end{aligned} \tag{4.31}$$

The probability that r particular (i fixed and $r-i$arbitrary) INs who have links to the i particular SNs are not connected to the $k\gamma$ ONs, i.e. the probability that the r particular columns of $\mathbf{G}^{\mathrm{LT}}_{k\gamma\times n}$ are all-zero columns, is given by

$$\begin{aligned} & \Pr\left[\text{ith SN cannot be recovered by the k}\gamma\text{ ONs}\right] \\ = \ & \left[\sum_{d=1}^{n}\Omega_d\frac{\binom{n-r}{d}}{\binom{n}{d}}\right]^{k\gamma}. \end{aligned} \tag{4.32}$$

Recall that there are $\binom{k}{i}$ possible combinations of i particular SNs. Combining Eq. (4.31) and (4.32) with Eq. (4.30), yields the following results

$$\begin{aligned} & P^{DF}_{k,n,\gamma} = \Pr[W_{k\gamma,n,k}] \\ \geq \ & \Pr[\cup_{i\in\{1,\ldots,k\}}\text{i SNs are not connected to the k}\gamma\text{ ONs}] \\ = \ & \sum_{i=1}^{k}\binom{k}{i}\sum_{r=i}^{n-k+i}\left[\sum_{d=1}^{n}\Omega_d\frac{\binom{n-r}{d}}{\binom{n}{d}}\right]^{k\gamma} \\ & \times\binom{n-k}{r-i}\left[(1-\eta)^i\right]^{n-k-r+i}\left[1-(1-\eta)^i\right]^{r-i}. \end{aligned} \tag{4.33}$$

The proof of Theorem 4.4 is completed. □

4.3.3 A Special Case of the Derived Bounds

In this subsection, we consider a special degree distribution – binomial degree distribution (the expurgated standard random ensemble), which is studied in [78, 89]. When we apply the binomial degree distribution (the expurgated standard random ensemble) into Theorem 4.2, we can simplify equation (4.2) into a far

less complex expression. The simplification procedure is shown in the following Corollary.

Corollary 4.5. *When a receiver successfully received $k\gamma_R$ coded symbols generated by using the Raptor code $(k, \mathcal{C}, \Omega(x))$ where $\mathcal{C}$ is an (n, k, η) LDGM code, $\Omega(x) = \sum_{d=1}^{n} \frac{\binom{n}{d} x^d}{(2^n-1)}$ and the coded symbols received at a receiver are decoded using ML decoding, the probability that a receiver cannot successfully decode all k source symbols with $k\gamma_R, k\gamma_R \geq k$, received coded symbols, denoted by P_{k,n,γ_R}^{DF}, satisfies*

$$P_{k,n,\gamma_R}^{DF} \leq (2^k - 1)(\frac{(2^{n-1}-1)}{(2^n-1)})^{k\gamma_R}. \tag{4.34}$$

Proof. When the binomial degree distribution (the expurgated standard random ensemble) [78, 89], i.e., $\Omega_d = \frac{\binom{n}{d}}{(2^n-1)}, 1 \leq d \leq n$, is inserted into equation (4.13), we can obtain that

$$\begin{aligned} & \Pr[\mathbf{g}_j^{\mathrm{LT}} \mathbf{x}_a^i = 0 \mid \left|\mathbf{x}_a^i\right| = r] \\ = \; & (2^n - 1)^{-1} \sum_{d=1}^{n} \sum_{s=0,2,\ldots,2\left\lfloor \frac{d}{2} \right\rfloor} \binom{r}{s}\binom{n-r}{d-s}. \end{aligned} \tag{4.35}$$

Similar to [89, Lemma 2], when the upper limit of the inner summation is changed from $2\left\lfloor \frac{d}{2} \right\rfloor$ to $2\left\lfloor \frac{n}{2} \right\rfloor$, it will not affect the result of equation (4.35). This is because $\binom{n-r}{d-s}$ with $s > 2\left\lfloor \frac{d}{2} \right\rfloor$ equals 0.

$$\begin{aligned} & \Pr\left[\mathbf{g}_j^{\mathrm{LT}} \mathbf{x}_a^i = 0 \mid \left|\mathbf{x}_a^i\right| = r\right] \\ = \; & (2^n - 1)^{-1} \sum_{d=1}^{n} \sum_{s=0,2,\ldots,2\left\lfloor \frac{n}{2} \right\rfloor} \binom{r}{s}\binom{n-r}{d-s} \\ = \; & (2^n - 1)^{-1} \sum_{s=0,2,\ldots,2\left\lfloor \frac{n}{2} \right\rfloor} \binom{r}{s} \sum_{d=1}^{n} \binom{n-r}{d-s}. \end{aligned} \tag{4.36}$$

The reason why the order of the two summations can be exchanged is because the inner summation variable s is now independent of the outer summation variable

d. Note that $1 \leq d \leq n$. Now we want to change the lower limit of the inner summation of equation (4.36) from 1 to 0 without affecting its result.

$$\begin{aligned}
&\Pr\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}_a^i = 0 \mid \left|\mathbf{x}_a^i\right| = r\right] \\
&= (2^n-1)^{-1}\{\sum_{s=0,2,\ldots,2\lfloor\frac{n}{2}\rfloor}\binom{r}{s}[\sum_{d=0}^{n}\binom{n-r}{d-s} - \binom{n-r}{d-s}_{d=0}]\} \\
&= (2^n-1)^{-1}\{[\sum_{s=0,2,\ldots,2\lfloor\frac{n}{2}\rfloor}\binom{r}{s}\sum_{d=0}^{n}\binom{n-r}{d-s}] - \binom{r}{s}\binom{n-r}{d-s}_{s=d=0}\}. \qquad (4.37)
\end{aligned}$$

This is because the term $\binom{n-r}{d-s}_{d=0}$ equals 0 for $s \neq 0$. Hence, only the case $s = 0$ needs to be considered. The term $\binom{n-r}{d-s}$ restricts d to $s \leq d \leq n-r+s$, such that

$$\sum_{d=0}^{n}\binom{n-r}{d-s} = \sum_{d=s}^{n-r+s}\binom{n-r}{d-s} = \sum_{d=0}^{n-r}\binom{n-r}{d} = 2^{n-r}. \qquad (4.38)$$

Combining this term with the last expression for $\Pr[\mathbf{g}_j^{LT}\mathbf{x}_a^i = 0 \mid |\mathbf{x}_a^i| = r]$ yields

$$\begin{aligned}
&\left[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}_a^i = 0 \mid \left|\mathbf{x}_a^i\right| = r\right] \\
= \;& (2^n-1)^{-1}\left(2^{n-r}\sum_{s=0,2,\ldots,2\lfloor\frac{n}{2}\rfloor}\binom{r}{s} - 1\right) \\
= \;& (2^n-1)^{-1}(2^{n-r}2^{r-1}-1) \\
= \;& \frac{(2^{n-1}-1)}{(2^n-1)}, \qquad (4.39)
\end{aligned}$$

where we have used identity $\sum_{s \text{ even}}\binom{r}{s} = 2^{r-1}$. We can observe that $\Pr[\mathbf{g}_j^{\mathrm{LT}}\mathbf{x}_a^i = 0 \mid |\mathbf{x}_a^i| = r]$ is independent from the weight of $\mathbf{x}_a^i$, hence $\Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = 0 | \, |\mathbf{x}_a^i| = r] = \Pr[\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x}_a^i = 0]$. Combining equations (4.12), (4.39), (4.6) and (4.4), we can obtain that

$$\begin{aligned}
&P_{k,n,\gamma_R}^{DF} = \Pr[W_{k\gamma_R,n,k}] \\
= \;& \Pr\left[\cup_{\mathbf{x}\in R(\mathbf{G}_{n\times k}^{\mathrm{pre}})}\mathbf{G}_{k\gamma_R\times n}^{\mathrm{LT}}\mathbf{x} = 0\right]
\end{aligned}$$

$$\begin{aligned} &\leq \sum_{\mathbf{x}\in R(\mathbf{G}^{\text{pre}}_{n\times k})} \Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}=0\right] \\ &= (2^k-1)\Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{x}=0|\,|\mathbf{x}|=r\right] \\ &= (2^k-1)(\frac{(2^{n-1}-1)}{(2^n-1)})^{k\gamma_R}. \end{aligned} \tag{4.40}$$

□

As for Theorem 4.3, we can simplify the lower bound into a far less complex expression as well. This is summarized in the following Corollary.

Corollary 4.6. *When a receiver successfully received $k\gamma_R$ coded symbols generated by using the Raptor code $(k,\mathcal{C},\Omega(x))$ where $\mathcal{C}$ is an (n,k,η) LDGM code and the coded symbols received at a receiver are decoded using ML decoding, the probability that a receiver cannot successfully decode all k source symbols with $k\gamma_R, k\gamma_R \geq k$, received coded symbols, denoted by P^{DF}_{k,n,γ_R}, satisfies*

$$\begin{aligned} &P^{DF}_{k,n,\gamma_R} \\ \geq\ &(2^k-1)\left[\frac{(2^{n-1}-1)}{(2^n-1)}\right]^{k\gamma_R} - (2^k-1)(2^{k-1}-1) \\ &\times\left\{\left[\frac{(2^{n-1}-1)}{(2^n-1)}\right]^3+\left[1-\frac{(2^{n-1}-1)}{(2^n-1)}\right]^3\right\}^{k\gamma_R}. \end{aligned} \tag{4.41}$$

Proof. The binomial degree distribution [89], i.e., $\Omega_d = \frac{\binom{n}{d}}{(2^n-1)}, 1 \leq d \leq n$, is inserted into equation (4.9), by using the result of equation (4.39), we can obtain that

$$\begin{aligned} J(r_0) &= \Pr[\mathbf{g}^{\text{LT}}_j\mathbf{x}^{w_0}=0|\,|\mathbf{x}^{w_0}|=r_0] \\ &= \frac{(2^{n-1}-1)}{(2^n-1)}. \end{aligned} \tag{4.42}$$

Insert equation (4.42) into equation (4.22), we can obtain that

$$\Pr\left[\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_0=\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_1=\mathbf{G}^{\text{LT}}_{k\gamma_R\times n}\mathbf{G}^{\text{pre}}_{n\times k}\mathbf{z}_2\right.$$

$$
\begin{aligned}
&| \, |\mathbf{z}_0| = w_0 \& \, |\mathbf{z}_1| = w_1 \& \, |\mathbf{z}_2| = w_2] \\
&= \sum_{r_0=w_0}^{n-k+w_0} \sum_{r_1=w_1}^{n-k+w_1} \sum_{r_0=w_2}^{n-k+w_2} D(w_0, r_0) D(w_1, r_1) D(w_2, r_2) \\
&\times \{ [\frac{(2^{n-1}-1)}{(2^n-1)}]^3 + [1 - \frac{(2^{n-1}-1)}{(2^n-1)}]^3 \}^{k\gamma_R} \\
&= \{ [\frac{(2^{n-1}-1)}{(2^n-1)}]^3 + [1 - \frac{(2^{n-1}-1)}{(2^n-1)}]^3 \}^{k\gamma_R}. \qquad (4.43)
\end{aligned}
$$

Insert equation (4.43) into equation (4.27), we can obtain that

$$
\begin{aligned}
&\sum_{\mathbf{x}_a^i \neq \mathbf{y}} \Pr[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R \times n} \mathbf{x}_a^i = 0 \& \mathbf{G}^{\mathrm{LT}}_{k\gamma_R \times n} \mathbf{y} = 0] \\
&= \sum_{w_0=0}^{i} \sum_{w_1=i-w_0} \sum_{w_2=0}^{k-i} \mathbf{1}(w_0 + w_2) \mathbf{1}(w_1 + w_2) \binom{i}{w_0} \binom{k-i}{w_2} \\
&\times \{ [\frac{(2^{n-1}-1)}{(2^n-1)}]^3 + [1 - \frac{(2^{n-1}-1)}{(2^n-1)}]^3 \}^{k\gamma_R} \\
&= (2^k - 2) \{ [\frac{(2^{n-1}-1)}{(2^n-1)}]^3 + [1 - \frac{(2^{n-1}-1)}{(2^n-1)}]^3 \}^{k\gamma_R}. \qquad (4.44)
\end{aligned}
$$

Combining equation (4.44), (4.20) and (4.19), we can obtain that

$$
\begin{aligned}
&P^{DF}_{k,n,\gamma_R} = \Pr[W_{k\gamma_R,n,k}] \\
&\geq \sum_{\mathbf{x} \in R(\mathbf{G}^{\mathrm{pre}}_{n \times k})} \Pr[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R \times n} \mathbf{x} = 0] \\
&\quad - \frac{1}{2} \sum_{\mathbf{x},\mathbf{y} \in R(\mathbf{G}^{\mathrm{pre}}_{n \times k}), \mathbf{x} \neq \mathbf{y}} \Pr[\mathbf{G}^{\mathrm{LT}}_{k\gamma_R \times n} \mathbf{x} = 0 \& \mathbf{G}^{\mathrm{LT}}_{k\gamma_R \times n} \mathbf{y} = 0] \\
&= (2^k - 1)(\frac{(2^{n-1}-1)}{(2^n-1)})^{k\gamma_R} - \frac{1}{2} \sum_{i=1}^{k} \binom{k}{i} (2^k - 2) \\
&\quad \times \{ [\frac{(2^{n-1}-1)}{(2^n-1)}]^3 + [1 - \frac{(2^{n-1}-1)}{(2^n-1)}]^3 \}^{k\gamma_R} \\
&= (2^k - 1) \left[\frac{(2^{n-1}-1)}{(2^n-1)} \right]^{k\gamma_R} - (2^k - 1)(2^{k-1} - 1) \\
&\quad \times \left\{ \left[\frac{(2^{n-1}-1)}{(2^n-1)} \right]^3 + \left[1 - \frac{(2^{n-1}-1)}{(2^n-1)} \right]^3 \right\}^{k\gamma_R}. \qquad (4.45)
\end{aligned}
$$

□

Compared with the general expressions in Theorems 4.2 and 4.3, in the simplified expressions of Corollaries 4.5 and 4.6, we can easily observe the relationship between the decoding failure probability and the parameters of the encoding rules, i.e., k, n and γ_R. Additionally, the computation complexity of the derived upper bound can be reduced from $O(\frac{1}{2}n^2k(n-k))$ to $O(1)$. As for the derived lower bound, the computation complexity can be reduced from $O(\frac{1}{8}n^6k^3(n-k)^3)$ to $O(1)$.

4.4 Simulation Results

In this section, we use MATLAB based simulations to validate the accuracy of the analytical results and the tightness of the proposed performance bounds. Each point shown in the figures is the average result obtained from 10^6 simulations. For clarity, the simulation parameters adopted in this section are summarized in Table 4.1.

Table 4.1: Simulation parameters

Rateless codes encoding parameters	
Number of source symbols k	$20, 40, 70$ and 100
Number of intermediate symbols n	$21, 41, 71$ and 102
Parameter for Bernoulli random variables η	0.3, 0.7
Pre-code $\mathcal{C}$	(n, k, η) LDGM code
The degree distributions for LT codes	
Standard degree distribution	$\Omega^{3GPP}(x)$
Binomial degree distribution	$\Omega_d = \frac{\binom{n}{d}}{(2^n-1)}, 1 \leq d \leq n$
Ideal soliton degree distribution	$\Omega_d = \frac{1}{d(d-1)}, 2 \leq d \leq n$ and $\Omega_1 = \frac{1}{n}$

4.4.1 Verification of the Derived Bounds

In this subsection, the number of source symbols is set to be $k = 20$ and the degree distribution of Raptor codes follows the widely used ideal soliton degree distribution [31]. Besides, the pre-code $\mathcal{C}$ is assumed to be $(21, 20, 0.3)$ and $(21, 20, 0.7)$

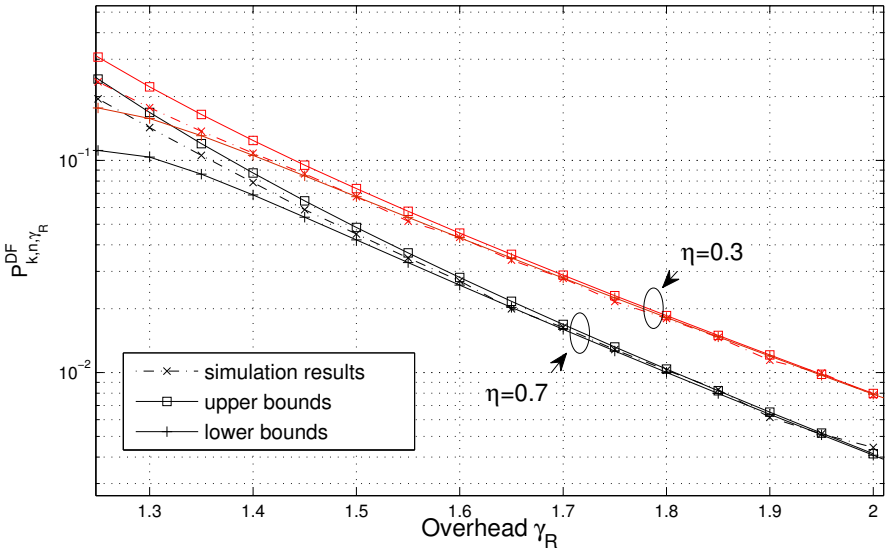


Figure 4.2: The decoding failure probabilities of Raptor codes with ideal soliton degree distribution and (21, 20, η) LDGM codes as the pre-code versus overhead γ_R. Parameter for Bernoulli random variables η is set as 0.3 and 0.7.

LDGM codes.

In Fig. 4.2, both analytical and simulation results are presented on P^{DF}_{k,n,γ_R}, the probability that a receiver cannot successfully decode all $k = 20$ source symbols, for different values of the reception overhead $\gamma_R = m_R/k$. As shown in Fig. 4.2, our analytical results, i.e., the upper bound and the lower bound, are consistent with the simulation results very well. This validates the accuracy of the analysis in this paper. However, when the overhead γ_R is small, there is still a gap between the upper (lower) bound and simulation results in Fig. 4.2. The gap between the exact value and the upper (lower) bound is caused by the approximation used in equation (4.2), and the gap between the exact value and the lower bound is caused by equation (4.29).

4.4.2 Investigation of the Impact of Degree Distribution on the Decoding Failure Probability of Raptor Codes

In this subsection, we investigate the performance for different degree distributions of LT codes when we fix the Pre-code C as (21, 20, 0.7). The investigated degree distributions of LT codes are represented by 3 cases. Case 1 uses the binomial degree distribution [89]. Case 2 investigates the widely used ideal soliton degree distribution [31]. Case 3 is the standardized degree distribution in 3GPP [39, Annex B]:

$$\begin{aligned}\Omega^{3GPP}(x) &= 0.0099x + 0.4663x^2 \\ &\quad +0.2144x^3 + 0.1152x^4 \\ &\quad +0.1131x^{10} + 0.0811x^{11}. \end{aligned} \tag{4.46}$$

As shown in Fig. 4.3, for different degree distributions, our analytical bounds agree very well with the simulation results. The performance of Raptor codes with the binomial degree distribution outperforms those obtained with the other three degree distributions. Furthermore, the expressions of the decoding failure probability of Raptor codes with the binomial degree distribution in Corollaries 4.5 and 4.6 are less computationally demanding compared with the expressions in Theorems 4.2 and 4.3. Therefore, we use Raptor codes with the binomial degree distribution in the following simulations.

4.4.3 Investigation of the Impact of k on the Decoding Failure Probability of Raptor Codes

When the number of source symbols k varies from 20 to 100, our analytical results still match the simulation results very well. From the figures we can see that the derived upper and lower bounds are asymptotically tight as the overhead grows.

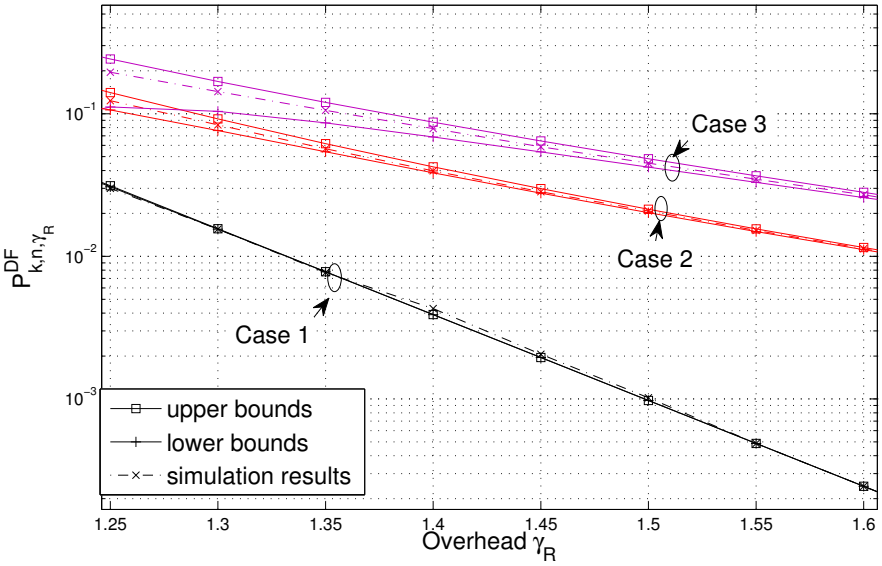


Figure 4.3: The decoding failure probabilities of Raptor codes with $(21, 20, 0.7)$ LDGM code as the pre-code and different degree distributions versus overhead γ_R. The degree distributions of Raptor codes are chosen as ideal soliton degree distribution [31], the standardized degree distribution in 3GPP [39, Annex B] and binomial degree distribution [89].

However, when the overhead is small, the gaps between the bounds and the simulated values are still visible. This is caused by the union bound in equation (4.5). Additionally, as shown in Figs. 4.4(a) and 4.4(b), at a larger number of the source symbols, less reception overhead $\gamma_R = m_R/k$ is required to achieve the same performance on the decoding failure probability.

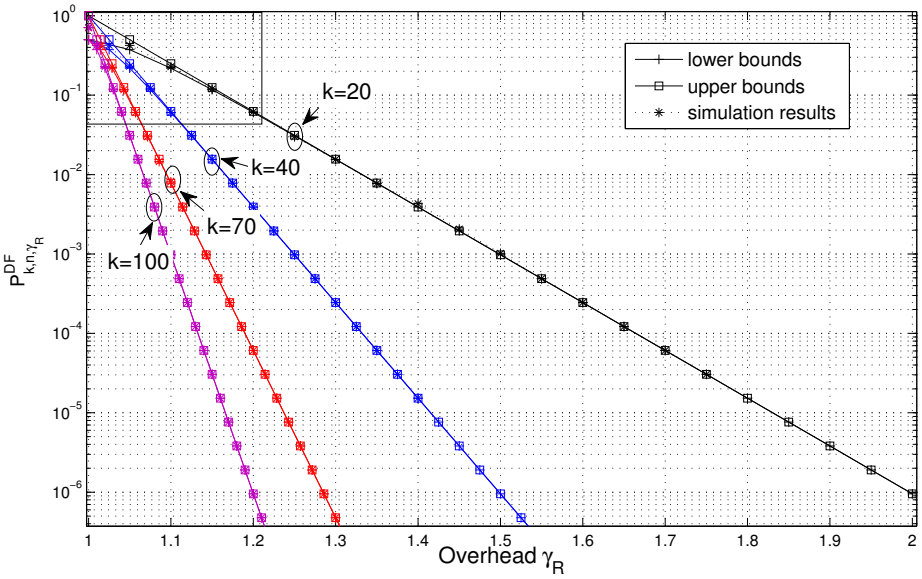
4.4.4 Investigation of the Impact of m on the Decoding Failure Probability of Raptor Codes

When the number of intermediate symbols m is set to be k, the special case that no precode is used, we can get another set of upper and lower bounds on the decoding failure probability of LT Codes. In this subsection, we compare the performance of LT and Raptor codes. As shown in Fig. 4.5, as we expected, Raptor codes can achieve lower decoding failure probabilities than LT codes.

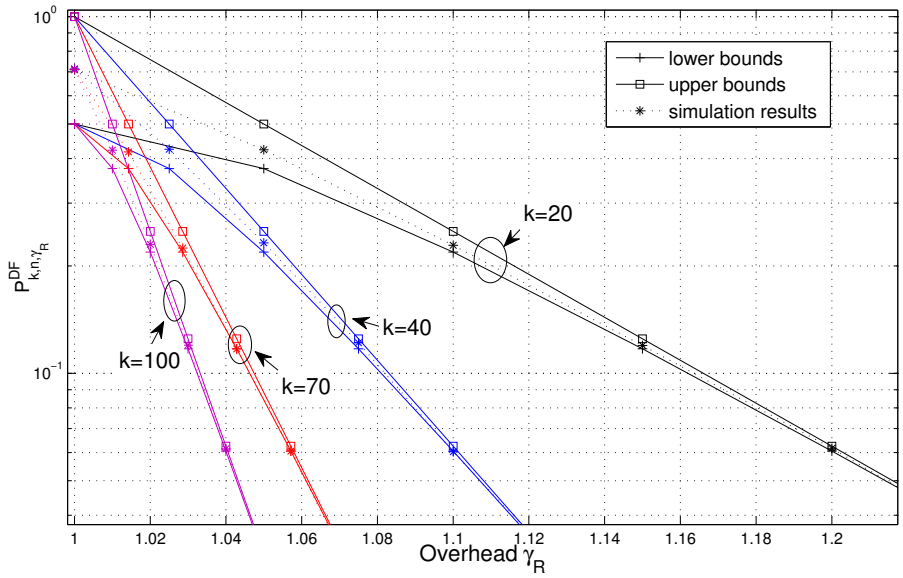
4.5 Summary

In this paper we focused on finite-length Raptor codes with a systematic LDGM code as pre-code and derived upper and lower bounds on the decoding failure probabilities of Raptor codes under ML decoding, which is measured by the probability that not all source symbols can be successfully decoded by a receiver with a given number of successfully received coded symbols. ML decoding ensures successful decoding when a full-rank matrix is received. Due to the concatenated coding structure of Raptor codes, we have analyzed the rank behavior of the product of two random matrices. Finally, by applying a special degree distribution–binomial degree distribution [89] into the upper and lower bound, we simplified the general bound with any degree distributions and any (n, k, η) LDGM codes as pre-code into a far less complex expression. The computation complexity of the derived bounds can be significantly decreased.

On the basis of the results presented in the paper, in the future, we plan to explore the optimum degree distribution and optimal parameter of Raptor codes in different channels.



(a) Full Scale



(b) Zoom of the rectangular box in (a)

Figure 4.4: The decoding failure probabilities of Raptor codes with the binomial degree distribution and $(n, k, 0.7)$ LDGM codes as the pre-code at different values of the overhead γ_R. The number of source symbols k is set to be 20, 40, 70 and 100 respectively.

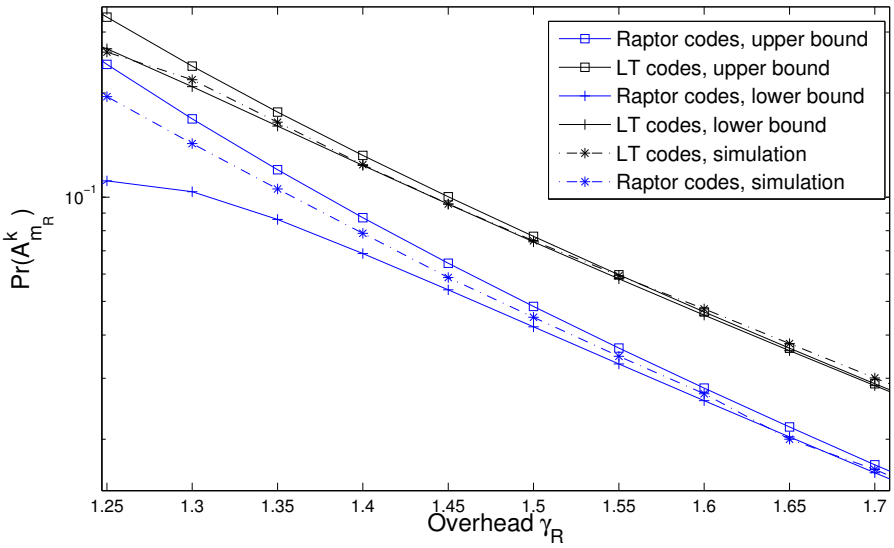


Figure 4.5: The decoding failure probabilities of Raptor codes with ideal soliton degree distribution and (21, 20, 0.7) LDGM codes as the pre-code and LT codes with ideal soliton degree distribution versus overhead γ_R.

Chapter 5

LT Codes based Wireless Broadcast Scheme

In the preceding two chapters, we have investigated the decoding success probability of finite-length LT codes and finite-length Raptor codes with a systematic *low-density generator matrix* (LDGM) code as the pre-code under *maximum-likelihood* (ML) decoding. Different from traditional FEC codes, rateless codes are adaptable to different channel conditions and avoid the need for feedback channels [46, 50, 40]. In this chapter we develop a LT codes based broadcast scheme that allows a *base station* (BS) to broadcast a given number of symbols to an unknown number of receivers, without requiring the receivers to acknowledge the correct receipt of broadcast symbols and in the meantime being able to provide a performance guarantee on the probability of successful delivery. Further, the BS only has limited statistical information about the environment including the spatial distribution of users (instead of their exact locations and number) and the wireless propagation model. Performance analysis is conducted. On that basis, an upper and a lower bound on the number of symbol transmissions required to meet the performance guarantee are obtained. Simulations are conducted to validate the accuracy of the theoretical analysis. The technique and analysis

developed in this chapter are useful for designing efficient and reliable wireless broadcast strategies. The results of this chapter appear in [J1].

5.1 Introduction

An important issue in wireless networks is efficient and reliable broadcasting. In this chapter, we want to design a wireless broadcast scheme that a) uses minimal information about network environment, b) can deliver information to a large number of users, c) does not rely on user acknowledgment, and d) is able to provide a guaranteed performance on the probability of successful delivery.

The advantages of rateless codes have been demonstrated in detail in Chapter 1 and 2. Due to these salient advantages of rateless codes, in this chapter we choose LT codes for use in our broadcast strategy design.

In [101], Tukmanov et al. studied the effect of cooperation on broadcast and derived analytical results characterizing the performance of a non-cooperative broadcast scheme and a cooperative broadcast scheme respectively. In their schemes, the network coding technique was not employed. In [45], Dong et al. compared the efficiency of the network coding based broadcast scheme and traditional ARQ based schemes. Their network coding based broadcast scheme relied on the feedback information provided by the receivers. In [46], Nguyen et al. investigated the benefits of applying rateless (fountain) codes on improving the transmission efficiency of broadcast without considering the decoding success probability.

In this chapter, we develop a LT codes based broadcast scheme that allows a BS to broadcast a given number of symbols to an unknown number of receivers, without requiring the receivers to acknowledge the correct receipt of broadcast symbols and in the meantime being able to provide a performance guarantee on the probability of successful delivery. Further, we assume that the BS only

has limited prior knowledge about the network environment, which includes the spatial distribution of the receivers, i.e. the receiver density λ, and the wireless propagation model. However the BS may not know the exact number of receivers and their locations. The above assumption is due to the consideration that in some highly dynamic networks, particularly vehicular networks, the receiver density in the coverage area of a BS is relatively stable and easy to estimate however the receivers in the coverage area may be changing quickly. Compared with the broadcast scheme without coding, the LT codes technique can facilitate information dissemination by reducing the minimum number of transmissions while providing a guaranteed performance on the probability of successful delivery. In this chapter, we apply the theoretical analysis on the decoding success probability for a single transmitter and receiver pair using LT codes from Chapter 3 in an one to all broadcast scenario. The performance of the proposed LT codes based broadcast scheme is validated both analytically and via simulations. The following is a detailed summary of our contributions:

- A LT codes based broadcast scheme is proposed, which broadcasts a given number of symbols from a BS to a large number of users with a priori knowledge about the spatial distribution of the receivers and the wireless propagation model only. The scheme does not need users' acknowledgment and is able to provide a performance guarantee on the probability of successful delivery.

- The performance of the proposed scheme is analyzed. On the basis of the analytical bounds of decoding success probability of finite-length LT codes under *maximum likelihood* (ML) decoding derived in Chapter 3, the upper and lower bounds on the probability that all receivers in a bounded area successfully receive or decode all source symbols from the BS are derived.

- On the basis of the above results, the minimum number of transmissions

required for a guaranteed performance on the probability of successful delivery is obtained.

- Simulations are conducted which validate both the accuracy of the analysis and the performance improvement of the proposed scheme.

The technique and analysis presented in this chapter can be useful for designing broadcast strategies to deliver information of common interest to a large number of users efficiently and reliably.

The rest of the chapter is organized as follows. Section 5.2 describes the system model and problem formulation. In Section 5.3, we carry out performance analysis of the proposed LT codes based broadcast scheme and present a technique to estimate the number of transmissions required to meet the performance objective on the probability of successful delivery. In Section 5.4, we validate our analytical results using simulations. Section 5.5 concludes the chapter.

5.2 System model and Problem Formulation

5.2.1 System Model

In this chapter, a cellular network with one BS and an unknown number of receivers is considered. Receivers are distributed across a two dimensional disk, denoted by $D(o, R)$, according to a homogeneous *Poisson point process* (PPP) Φ with intensity λ where $D(o, R)$ represents a disk centered at the origin o and with a radius R. The BS is located at the origin. Let $\{\boldsymbol{x}_i\}$ denote the set of receivers on $D(o, R)$ and we refer to a receiver by its location $\boldsymbol{x}_i$.

We assume that the channels from the BS to different receivers are independent [1]. For the data transmission from the BS located at o to a receiver located

[1]The assumption of channel independence has been widely used and is also supported by some measurement studies although we acknowledge that in some environment channel correlations can be a major concern. For example, in [102] it was shown that the coherence distance in an omnidirectional Rayleigh channel is: $\frac{9\lambda}{16\pi}$ [102, Eq. (5.116)] where λ is the wavelength

at $\boldsymbol{x}_i$, the SNR of the received signal is written as:

$$\mathbf{SNR}_i \;=\; \frac{P_t h_i \left\|\boldsymbol{x}_i\right\|^{-\alpha}}{N_o}, \tag{5.1}$$

where P_t is the transmitting power of the BS, N_o is the background noise power, α is the path loss exponent and $\|\boldsymbol{x}_i\|$ represents the Euclidean norm of $\boldsymbol{x}_i$. Parameter h_i is a random positive number modeling the small scale fading and shadowing between the BS and $\boldsymbol{x}_i$ and is assumed to be exponentially distributed with a mean value of 1 [101].

The BS broadcasts coded symbols to all receivers where the source symbols are coded using LT codes. A (coded) symbol is considered to be successfully delivered from the BS to the receiver $\boldsymbol{x}_i$ when the instantaneous SNR is greater than or equal to a designated threshold δ. Denote by P_i the probability of successful symbol delivery for the receiver $\boldsymbol{x}_i$. It follows that

$$P_i \;=\; \Pr[\mathbf{SNR}_i \geq \delta]. \tag{5.2}$$

Further, for each receiver, we assume that the event that a (coded) symbol is successfully received and the event that another (coded) symbol is received are independent.

5.2.2 Problem Formulation

The metric of interest is the number of transmissions by the BS, denoted by L, required to deliver k source symbols of equal length to all receivers in $D\,(o, R)$ such that the probability of successful delivery of all k symbols to all receivers is

and the value for a non-omnidirectional channel is only slightly different [102, Eq. (5.117)]. In a more recent work it was shown [103] that if a pair of receivers are separated by more than λ, their received signals from a common transmitter can be considered independent [102, p. 243] (with a correlation coefficient less than 0.15). At 800 MHz $\lambda = 0.375$ m, thus the requirement on the separation of receivers (in order for the channels to be considered independent) can be easily met.

above a predesignated threshold $1-\epsilon$, where ϵ is a small positive constant.

Denote by η_i the event that all k source symbols have been received, i.e. successfully decoded from the coded symbols received from the BS, by receiver $\boldsymbol{x}_i$. Let

$$\eta \triangleq \bigcap\nolimits_{i\in\Gamma}\eta_i, \tag{5.3}$$

where Γ denotes the set of indices of all the receivers and η represents the event that all k source symbols have been received, i.e. successfully decoded from the coded symbols received from the BS, by all the receivers. Obviously $\Pr(\eta)$ depends on the number of (coded) symbols broadcast by the BS. Denote by m the number of symbols broadcast from the BS and we also write η as $\eta(m)$ to emphasize the dependence of η on m when necessary. Parameter L can be defined more rigorously as:

$$L \triangleq \arg\min_{m} \Pr(\eta(m)) \geq 1-\epsilon. \tag{5.4}$$

In this chapter, we shall quantitatively characterize the value of L. This is done by first deriving the upper and lower bounds on the decoding success probability $\Pr(\eta_i)$ for a single BS and receiver pair using finite-length LT codes. On that basis, the upper and lower bounds on the probability $\Pr(\eta)$ that all receivers successfully decode all source symbols from the BS are derived. Consequently, the upper and lower bounds on L are obtained which allows us to draw conclusion on the number of (coded) symbols that the BS needs to transmit with LT codes to guarantee that $\Pr(\eta) \geq 1-\epsilon$.

Fig. 5.1 illustrates the system model.

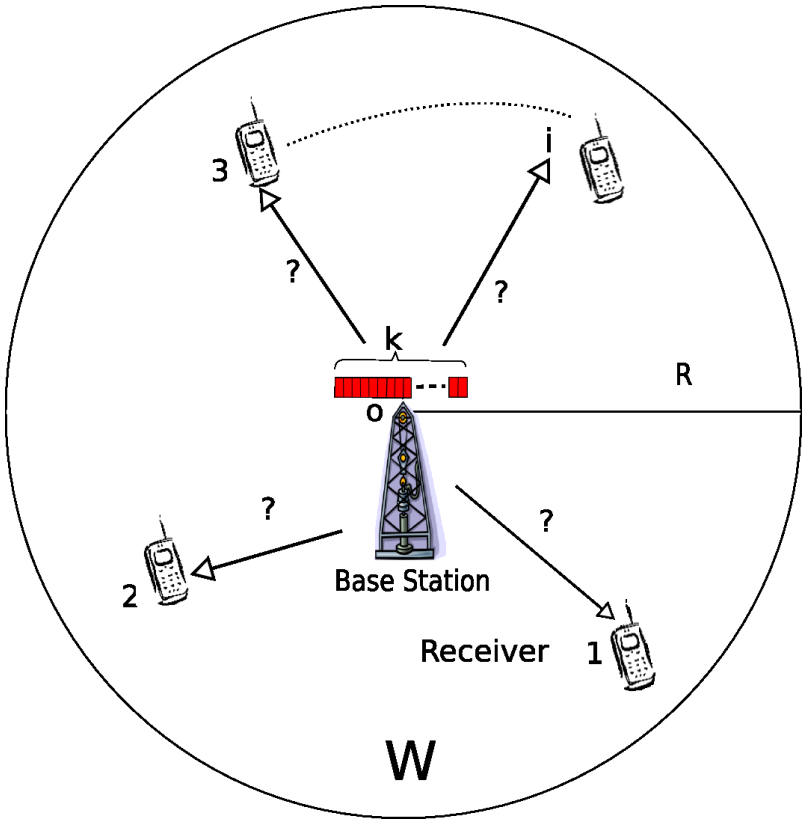


Figure 5.1: An illustration of the system model

5.3 Analysis on the Overall Success Probability for Multiple Receivers

On the basis of the analysis in the last section, which investigated the decoding success probability of a single receiver who have successfully received m_R coded symbols from the BS, in this section, we continue to analyze the overall success probability that all receivers have successfully received all k symbols, i.e. $\Pr(\eta)$ where the event η is defined in equation (5.3).

For convenience, let $\phi(m_R)$, $\phi_l(m_R)$ and $\phi_u(m_R)$ be the exact value, the upper and the lower bound of $\Pr\left[R^k_{m_R}\right]$ as suggested in Theorem 3.1 respectively. According to Theorem 3.1, $\phi(m_R) \geq \phi_l(m_R) = \mathbf{e}_k(\mathbf{X}_{min})^{(m_R-1)}\mathbf{R}(1)$ and $\phi(m_R) \leq \phi_u(m_R) = \mathbf{e}_k(\mathbf{X})^{(m_R-1)}\mathbf{R}(1)$. Denote by L the total number of transmissions required on the BS in order to meet the objective $\Pr(\eta) \geq 1-\epsilon$. Let

$\Pr(\eta(m_T))$ denote the probability that all the k source symbols can be successfully decoded by all the receivers after m_T transmissions by the BS. It can be expressed as:

$$\Pr(\eta(m_T)) = \sum_{j=0}^{\infty} \Pr[\eta(m_T, j) \mid N = j] \Pr[N = j], \tag{5.5}$$

where $\eta(m_T, j)$ is the event that all k source symbols have been successfully received/decoded from the m_T coded symbols broadcast by the BS, by all j receivers in the coverage area of the BS $D(o, R)$. Additionally, N is the total number of receivers in $D(o, R)$. Parameter N is a Poissonly distributed non-negative integer with mean $\lambda\pi R^2$:

$$\Pr(N = j) = \frac{(\lambda\pi R^2)^j \exp(-\lambda\pi R^2)}{j!}. \tag{5.6}$$

As an easy consequence of the Poisson distribution of receivers [104] and the independence of channels between the BS and the receivers, it can be obtained that

$$\Pr[\eta(m_T, j) \mid N = j] = \prod_{i=1}^{j} \Pr[\eta_i(m_T)] = (\Pr[\eta_i(m_T)])^j, \tag{5.7}$$

where $\eta_i(m_T)$ represents the event that the i^{th} receiver (which is randomly drawn from the set of all receivers) can successfully decode all k source symbols when the BS broadcasts m_T coded symbols.

For the same receiver, the received coded symbols broadcast by the BS are independent of each other. Let r_i be the (random) distance between the i^{th} receiver and the BS and $r_i = \|\boldsymbol{x}_i\|$. It readily follows that

$$\begin{aligned} &\Pr[\eta_i(m_T) \mid r_i = y] \\ = &\sum_{m_R=k}^{m_T} \binom{m_T}{m_R} \{P_i(y)\}^{m_R} \{1 - P_i(y)\}^{m_T - m_R} \phi(m_R), \end{aligned} \tag{5.8}$$

where the term $\binom{m_T}{m_R}\{P_i(y)\}^{m_R}\{1-P_i(y)\}^{m_T-m_R}$ represents the probability that out of m_T coded symbols broadcast by the BS, m_R coded symbols are received by the i^{th} receiver. Here $P_i(y)$ represents the probability that a coded symbol is successfully received by the i^{th} receiver conditioned on that $r_i = y$. According to the definition in Section 5.2, $P_i(y)$ can be expressed as:

$$P_i(y) \quad = \quad \Pr[\mathbf{SNR}_i(y) \geq \delta], \tag{5.9}$$

where $\mathbf{SNR}_i(y)$ is instantaneous SNR of the channel between the BS and the i^{th} receiver. Using equation (5.1) and note that h_i is exponentially distributed with mean value 1, equation (5.9) can be rewritten as:

$$P_i(y) \quad = \quad \Pr[h_i \geq \frac{N_o\delta y^\alpha}{P_t}] = \exp(-\frac{N_o\delta y^\alpha}{P_t}). \tag{5.10}$$

Inserting equation (5.10) into equation (5.8) we obtain:

$$\begin{aligned}
&\Pr\left[\eta_i\left(m_T\right) \mid r_i = y\right] \\
&= \sum_{m_R=k}^{m_T} \binom{m_T}{m_R}\{P_i(y)\}^{m_R}\{1-P_i(y)\}^{m_T-m_R}\phi(m_R) \\
&= \sum_{m_R=k}^{m_T} \binom{m_T}{m_R}\phi(m_R)\left[\exp(-\frac{N_o\delta y^\alpha}{P_t})\right]^{m_R} \\
&\times\left[1-\exp(-\frac{N_o\delta y^\alpha}{P_t})\right]^{m_T-m_R} \\
&= \sum_{m_R=k}^{m_T} \binom{m_T}{m_R}\phi(m_R)\sum_{i=0}^{m_T-m_R}\binom{m_T-m_R}{i}(-1)^{(m_T-m_R-i)} \\
&\times\left[\exp(-\frac{N_o\delta y^\alpha}{P_t})\right]^{(m_T-m_R-i)}\left[\exp(-\frac{N_o\delta y^\alpha}{P_t})\right]^{(m_R)} \\
&= \sum_{m_R=k}^{m_T} \binom{m_T}{m_R}\phi(m_R)\sum_{i=0}^{m_T-m_R}\binom{m_T-m_R}{i}(-1)^{(m_T-m_R-i)}\left[\exp(-\frac{N_o\delta y^\alpha}{P_t})\right]^{(m_T-i)}
\end{aligned} \tag{5.11}$$

Owing to the property of Poisson process, conditional on the number of receivers $N = j$, all j receivers are independent and identically distributed (i.i.d.)

on $D(o,R)$ following a uniform distribution. Therefore the cumulative distribution function of r_i can be easily obtained:

$$\Pr[r_i \leq y] = \frac{y^2}{R^2}, y \in [0,R] \tag{5.12}$$

and the probability density function of r_i is given by $\frac{2y}{R^2}$.

Using the total probability theorem, we can now derive $\Pr[\eta_i(m_T)]$ as:

$$\begin{aligned}\Pr[\eta_i(m_T)] &= \int_{y=0}^{y=R} \Pr[\eta_i(m_T) \mid r_i = y] \frac{2y}{R^2} dy \\ &= \sum_{m_R=k}^{m_T} [\frac{2\binom{m_T}{m_R}\phi(m_R)}{R^2}] \sum_{i=0}^{m_T-m_R} \binom{m_T-m_R}{i}(-1)^{(m_T-m_R-i)} \\ &\times \int_{y=0}^{y=R} y \left[\exp(-\frac{(m_T-i)N_o\delta y^\alpha}{P_t})\right] dy. \end{aligned} \tag{5.13}$$

Further, the integral inside equation (5.13) can be computed:

$$\begin{aligned}&\int_{y=0}^{y=R} y \left[\exp(-\frac{(m_T-i)N_o\delta y^\alpha}{P_t})\right] dy \\ =& \left[\frac{\Gamma[\frac{2}{\alpha}, \frac{(m_T-i)N_o\delta y^\alpha}{P_t}]}{\alpha\left(\frac{(m_T-i)N_o\delta}{P_t}\right)^{\frac{2}{\alpha}}}\right]_0^R \\ =& \frac{\Gamma[\frac{2}{\alpha}, \frac{(m_T-i)N_o\delta R^\alpha}{P_t}] - \Gamma[\frac{2}{\alpha}, 0]}{\alpha\left(\frac{(m_T-i)N_o\delta}{P_t}\right)^{\frac{2}{\alpha}}}, \end{aligned} \tag{5.14}$$

where $\Gamma(n,x)$ is the incomplete Gamma function.

Inserting equations (5.6), (5.7), (5.13) and (5.14), into the equation (5.5), we can obtain an upper bound and a lower bound on $\Pr(\eta(m_T))$, which are given by (5.15) and (5.16), respectively. Particularly, using the lower bound on $\Pr(\eta(m_T)$ in (5.16), the minimum number of transmissions required by the BS in order to meet the performance guarantee that $\Pr(\eta) \geq 1-\epsilon$ can be determined.

$$\Pr(\eta(m_T)) \leq \exp\left\{\lambda 2\pi \sum_{m_R=k}^{m_T} \binom{m_T}{m_R}\phi_u(m_R) \sum_{i=0}^{m_T-m_R} \binom{m_T-m_R}{i}(-1)^{(m_T-m_R-i)}\right.$$

$$\times \left[\frac{\Gamma[\frac{2}{\alpha}, \frac{(m_T-i)N_o\delta R^\alpha}{P_t}] - \Gamma[\frac{2}{\alpha}, 0]}{\alpha \left(\frac{(m_T-i)N_o\delta}{P_t}\right)^{\frac{2}{\alpha}}} \right] - \lambda\pi R^2 \Bigg\}, \tag{5.15}$$

$$\Pr\left(\eta\left(m_T\right)\right) \geq \exp\left\{ \lambda 2\pi \sum_{m_R=k}^{m_T} \binom{m_T}{m_R} \phi_u(m_R) \sum_{i=0}^{m_T-m_R} \binom{m_T-m_R}{i} (-1)^{(m_T-m_R-i)} \right.$$
$$\times \left[\frac{\Gamma[\frac{2}{\alpha}, \frac{(m_T-i)N_o\delta R^\alpha}{P_t}] - \Gamma[\frac{2}{\alpha}, 0]}{\alpha \left(\frac{(m_T-i)N_o\delta}{P_t}\right)^{\frac{2}{\alpha}}} \right] - \lambda\pi R^2 \Bigg\}. \tag{5.16}$$

5.4 Simulation Results

In this section, we use simulations to validate the accuracy of the analytical results and the tightness of the bounds. The simulations are conducted in a simulator written in Matlab. Each point shown in the figures is the average value obtained from 10^5 simulations. The 95% confidence interval is shown in these figures as well. The radius R is chosen to be 2.5 km. The receiver density is varied from $\lambda = 10$ nodes/km^2 to $\lambda = 100$ nodes/km^2. The number of source symbols is chosen to be 5. The degree distribution of LT codes follows the widely used Luby's Ideal Soliton distribution [31]. Path-loss exponent is set to be $\alpha = 2$. The transmitting power of the transmitter (BS) P_t is set to be 10 dBm and the thermal noise power density N_o is -80 dBm. The SINR threshold δ is set to be 0 dB. For comparison, the scenario that the BS broadcasts without using network coding is also shown in some figures. When the BS broadcasts without using network coding, the BS broadcasts the k source symbols sequentially and repeat the process when the last source symbol is broadcast. Theoretical analysis for the scenario that the BS broadcasts without using network coding is trivial compared with that using LT codes and hence is not presented in the thesis.

Analytical and simulation results are presented in Fig. 5.2 on the probability that all receivers successfully receive all 5 source symbols as a function of the number of transmissions by the BS. As shown in Fig. 5.2, our analytical results, i.e., upper and lower bounds, match the simulation results very well, which val-

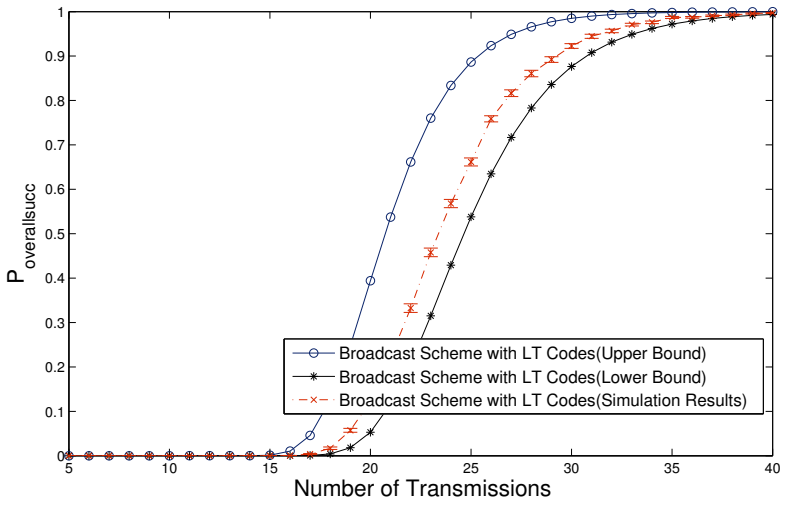


Figure 5.2: The probability of successfully decoding all 5 source symbols by all receivers versus the number of coded symbols broadcast by the BS.

idate the accuracy of the analysis in this paper. However there is still a gap between the upper (lower) bounds and simulation results in the figures. The gap between the exact value and the upper bound is caused by the approximation used in equation (3.1) and the gap between the exact value and the lower bound is caused by equation (3.2).

In Fig. 5.3, we further compare the success probabilities of broadcast using LT codes and without using network coding. As shown in Fig. 5.3, it can be seen that the use of LT codes yields much better performance in terms of the number of transmitted symbols required to meet the same performance objective on the probability of successful delivery (i.e. all receivers receive all source symbols). In comparison, without using network coding, the BS needs to transmit more symbols to meet the performance objective. For example, when the probability of successful delivery is set to be 0.947, at most 33 transmissions is needed when LT codes are used, while 50 broadcasts are required when coding is not used, which represents a saving of 50% transmissions when using LT codes.

Fig. 5.4 shows the overall success probabilities of the proposed LT coding based broadcast scheme as a function of the node density when the number of

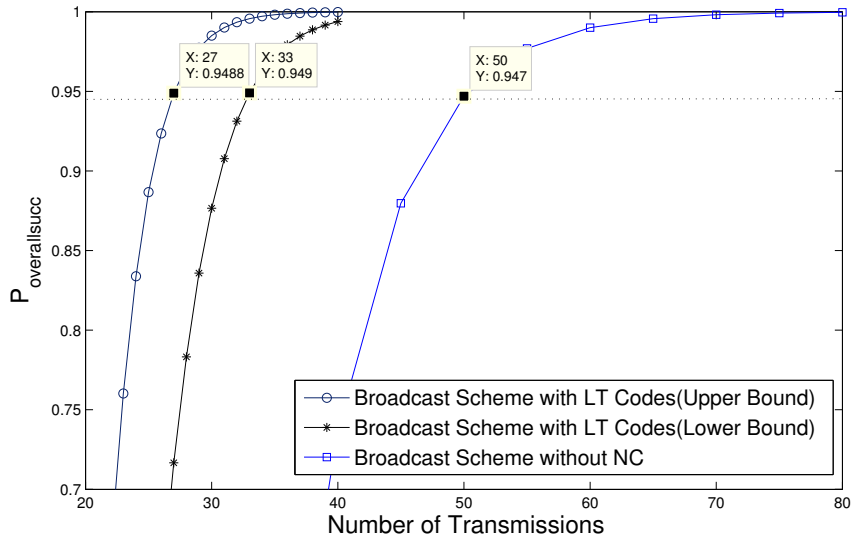


Figure 5.3: The probabilities of successfully decoding all 5 source symbols by all receivers for broadcast scheme using LT codes and that without NC as a function of the number of transmissions by the BS

broadcast from the BS is fixed at 35. We can see that the simulation results match well with the theoretical results. Further, for all values of the node density, broadcasting using LT codes offers better performance than broadcasting without using coding. We also can observe that as the node density increases the gaps between the upper and the lower bounds become bigger. This is because the differentiation of the gap of the bounds is a positive value when $0 \leq \lambda \leq \lambda_c$, where λ_c is a positive number and can be easily calculated.

The variation of the system overall success probabilities of the proposed LT codes based broadcast scheme with the path loss exponent is demonstrated in Fig. 5.5(a) and Fig. 5.5(b). The number of source symbols and the number of broadcast from the BS are set to be 15 and 75, respectively. The radius R is chosen to be 400 m. The receiver density is set to be $\lambda = 10$ nodes/km^2. The transmitting power of the BS P_t is set to be -18 dBm and the thermal noise power density N_o is -80 dBm. We can observe that the simulation results lie between the upper and lower bound, i.e., are consistent with the theoretical results. Further, for all values of the path loss exponent, broadcasting using LT

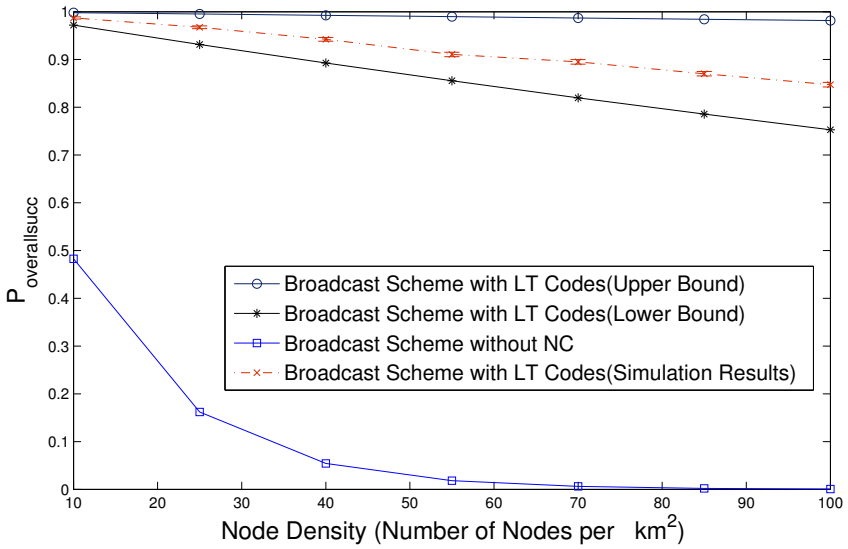
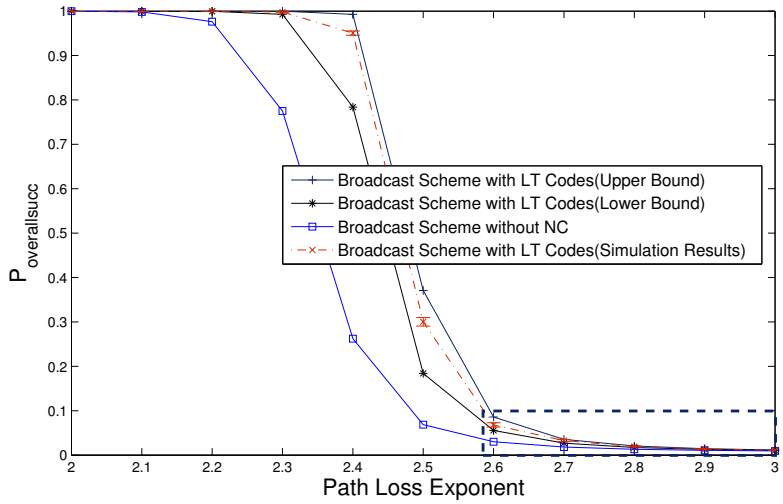


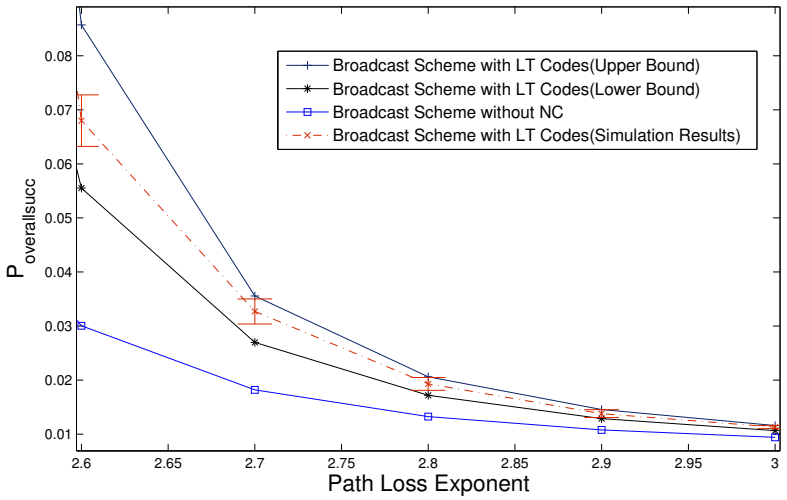
Figure 5.4: The probabilities of successfully decoding all 5 source symbols by all receivers for broadcast scheme using LT codes and that without coding as a function of the node density.

codes outweighs the performance of broadcasting without using coding.

When the number of source symbols increases, the conclusion that the use of LT codes can significantly reduce the number of transmissions required to meet the same performance objective, compared with that without using coding still hold. As demonstrated in Figs. 5.6, 5.7, 5.8 and 5.9, compared with broadcasting without using coding, the BS can reduce the number of transmissions required to meet the same performance objective, which leads to reduced transmission latency and energy consumption. When the performance objective, i.e., the probability of successful delivery, is set to 0.954, for k=10, the ratio of the number of symbols transmitted without using coding to that using LT codes equals 2.037; for k=20, the ratio is 2.5; for k=50, the ratio increases to 3.095; for k=100, the ratio becomes 3.5. It seems that the ratio increases as the number of source symbols increases.



(a) Full Scale



(b) Zoom of the dotted rectangular box in (a)

Figure 5.5: The probabilities of successfully decoding all 15 source symbols by all receivers for broadcast scheme using LT codes and that without coding ***vs*** the path loss exponent.

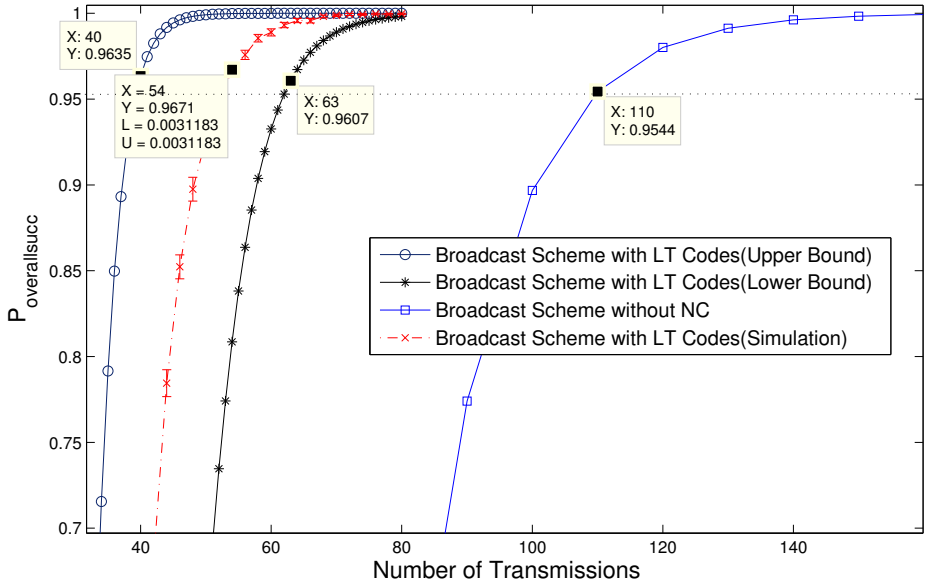


Figure 5.6: The probabilities of successfully decoding all $k = 10$ source symbols by all receivers for broadcast scheme using LT codes and that without coding as a function of the number of transmissions by the BS

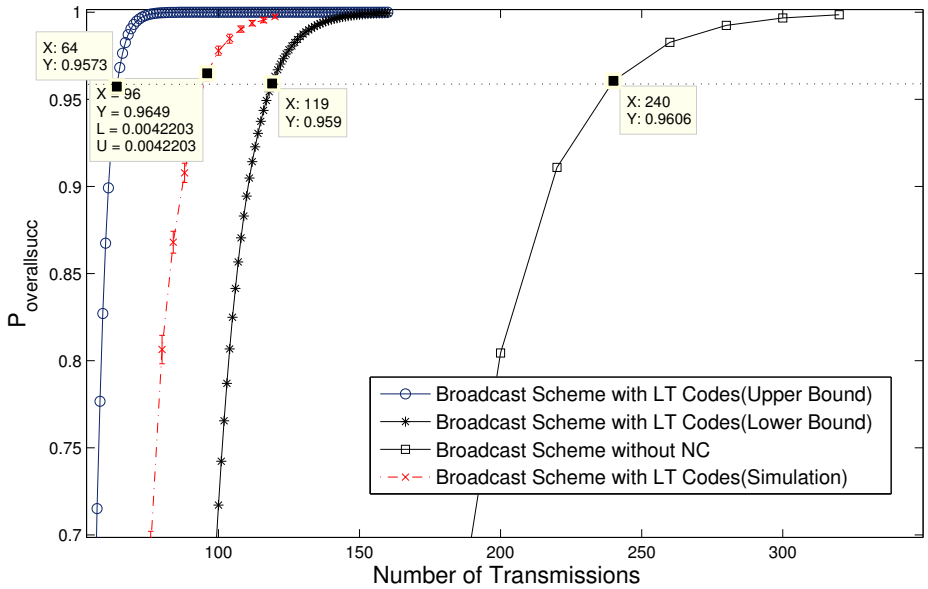


Figure 5.7: The probabilities of successfully decoding all $k = 20$ source symbols by all receivers for broadcast scheme using LT codes and that without coding as a function of the number of transmissions by the BS

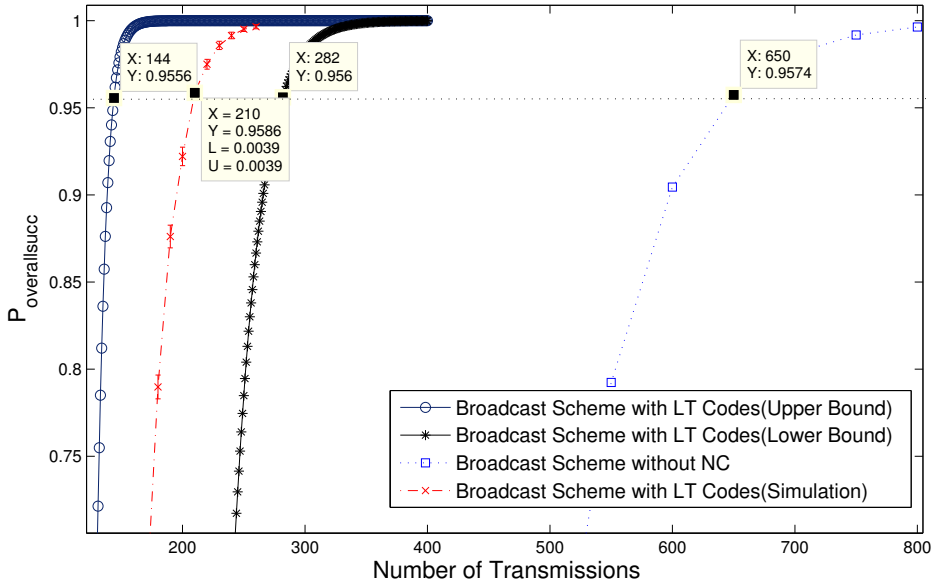


Figure 5.8: The probabilities of successfully decoding all $k = 50$ source symbols by all receivers for broadcast scheme using LT codes and that without coding as a function of the number of transmissions by the BS

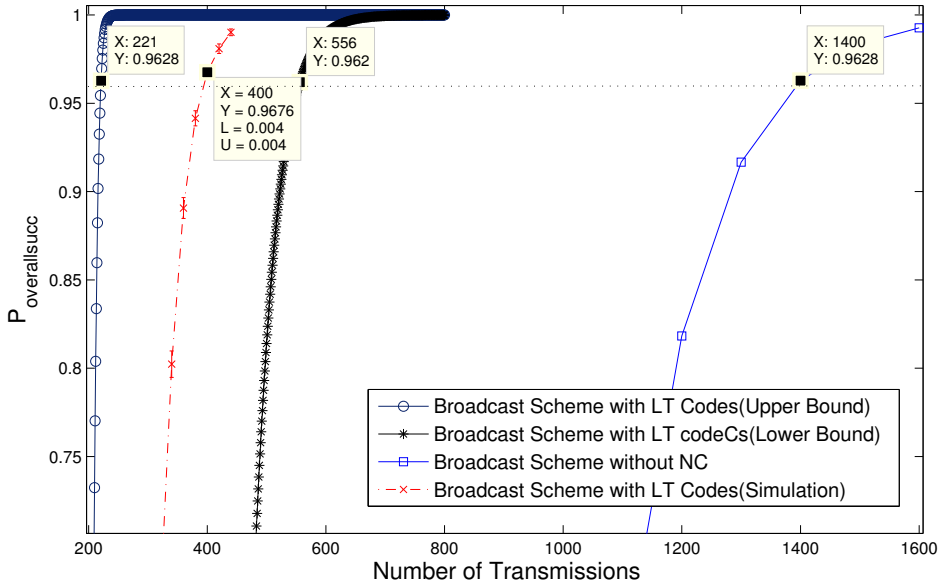


Figure 5.9: The probabilities of successfully decoding all $k = 100$ source symbols by all receivers for broadcast scheme using LT codes and that without coding as a function of the number of transmissions by the BS

5.5 Summary

In this chapter we studied reliable broadcast in a wireless network with a BS and a number of receivers. More specifically, we assume that the BS only has limited statistical information about the environment including the spatial distribution of users (instead of their exact locations and number) and the wireless propagation model. By resorting to stochastic geometry analysis, a LT codes based broadcast scheme was designed that allows the BS to broadcast a given number of source symbols to a large number of users, without user acknowledgment, while being able to provide a performance guarantee on the probability of successful delivery. The scheme is based on a rigorous analysis on the probability of successful delivery using LT codes, conducted in Chapter 3. The upper and lower bounds on the probability that all receivers successfully decode all source symbols from the BS were derived in Section 5.3. On that basis, the upper and lower bounds of the number of transmissions required for a guaranteed performance on the probability of successful delivery were obtained. Simulations were conducted to validate the accuracy of the theoretical analysis. It was shown that the use of LT codes can significantly reduce the number of transmissions required to meet the same performance objective, compared with that without using network coding. The technique and analysis developed in this paper can be useful for designing broadcast strategies to deliver information of common interest to a large number of users efficiently and reliably.

Chapter 6

Conclusion and Future Work

In this thesis, we considered random network coding (rateless erasure coding) as the desirable means for wireless broadcast problems because of the advantages of rateless codes: avoiding the need for feedback channels and being able to adapt to different channel conditions. Two types of rateless code, i.e., *Luby transform* (LT) codes and Raptor codes, were focused on in this thesis. The decoding success probabilities of finite-length LT Codes and Raptor codes under *maximum likelihood* (ML) decoding were investigated.

In this chapter, we conclude the thesis by summarizing our contributions.

6.1 Finite-Length Analysis of LT Codes

In Chapter 3 we studied the decoding success probability of LT codes under ML decoding over BEC, i.e., the probability that all source symbols can be successfully decoded by a receiver with a given number of successfully received coded symbols under ML decoding. Since ML decoding of an LT code is equivalent to solving a consistent system of linear equations, where the coefficients are given by an LT code generator matrix created according to some specifically designed random processes. In Chapter 3, we provided rigorous mathematical analysis on the rank profile of a random coefficient matrix. On the basis of this analysis, we

derived upper and lower bounds on the decoding success probability of LT codes under ML decoding. Moreover, simulations were conducted to validate the accuracy of the proposed bounds. More specifically, LT codes with different degree distributions were evaluated which establishes the accuracy of the bounds. We showed that when binomial degree distribution introduced in Subsection 2.3.1 is applied, the upper and lower bounds merge to the exact expression. Moreover, its performance outperforms the other degree distributions [31] in terms of decoding success probability.

6.2 Finite-Length Analysis of Raptor Codes

In Chapter 4 we investigated the decoding success probability of Raptor codes with *low-density generator matrix* (LDGM) codes as the pre-code under ML decoding over BEC. The decoding success probability of these compound codes is equivalent to the probability that the product of two random matrices has full rank. In Chapter 4, we firstly provided the analytical results, i.e., an upper bound and a lower bound, on the decoding failure probability of Raptor codes with a systematic LDGM code as the pre-code under ML decoding. The decoding failure probability is the probability that not all source symbols can be successfully recovered by a receiver with a given number of successfully received coded symbols under ML decoding. The analytical results are derived by analyzing the rank of the product of two random coefficient matrices. Based on the analytical bounds on the decoding failure probability of Raptor codes, we can readily obtain the analytical bounds on the decoding success probability of Raptor codes, which is unity minus decoding failure probability. Moreover, simulations were conducted to validate the accuracy of the proposed bounds. More specifically, Raptor codes with different degree distributions and pre-codes, were evaluated which establishes the accuracy of the bounds. Finally, by applying binomial de-

gree distribution [89] into the upper and lower bounds, we simplified the general bounds with any degree distributions and any (n, k, η) LDGM codes as pre-code into a far less complex expressions. By this way, the computation complexity of derived bounds can be significantly decreased.

The developed bounds enable a quick assessment of the decoding error properties of a coding ensemble without the need for time-consuming Monte Carlo simulations. They can be used to find the optimum degree distribution and parameters of Raptor codes.

6.3 LT Codes based Wireless Broadcast Scheme

In Chapter 5, we developed a LT codes based broadcast scheme that allows a base station (BS) to broadcast a given number of symbols to an unknown number of receivers, without requiring the receivers to acknowledge the correct receipt of broadcast symbols and in the meantime being able to provide a performance guarantee on the probability of successful delivery. Further, the BS only has limited statistical information about the environment including the spatial distribution of users (instead of their exact locations and number) and the wireless propagation model. Based on the decoding success probability of LT codes under ML decoding derived in Chapter 3, the performance of the proposed scheme was analyzed. On that basis, an upper and a lower bound on the number of symbol transmissions required to meet the performance guarantee were obtained. Simulations were conducted to validate the accuracy of the theoretical analysis. The analytical bounds developed in Chapter 5 are useful for designing efficient and reliable wireless broadcast strategies. The scheme proposed in Chapter 5 is expected to be also helpful to set the corresponding parameters of wireless broadcast in a more realistic setting.

6.4 Future Work

In this section, some of the interesting open directions for future research are listed below:

- This thesis analyzed the performance for rateless code under erasure channels. It would be interesting to consider other channel models such as AWGN channels and fading channels.
- Implementing rateless codes into modern communication systems is an important research topic.
- It is worthwhile to explore the optimum degree distribution and parameters design of the finite-length rateless codes with the ML decoding bounds derived in this thesis.
- An interesting research direction could be to develop new practical *transmission control protocol* (TCP) based on rateless codes.
- A straightforward extension of the proposed network coding based broadcast schemes is the Multimedia broadcasting/multicasting in wireless cellular networks.
- It would be interesting to extend the application of rateless codes based broadcasting to *Device to Device* (D2D) networks.

Bibliography

[1] A. F. Molisch, *Wireless Communications* - Second Edition. Wiley, 2011.

[2] X. Li, *Wireless Ad Hoc and Sensor Networks*. Cambridge University Press, 2008.

[3] S. Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes," *IEEE Commun. Magazine*, vol. 22, no. 12, pp. 5-17, 1984.

[4] L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, Third Edition, 2003

[5] P. Frenger, S. Parkvall and E. Dahlman, "Performance comparison of HARQ with Chase combining and incremental redundancy for HSDPA," in *Proceedings of IEEE VTC 54th*, vol.3, pp.1829-1833, 2001

[6] I. H. Hou and P. R. Kumar, "Broadcasting delay-constrained traffic over unreliable wireless links with network coding," in *Proceedings of the 12th ACM MobiHoc*, pp. 1-10, 2011.

[7] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "Xors in the air: practical wireless network coding," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 497-510, 2008.

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

[9] R. W. Hamming, "Error detecting and error correcting codes," *The Bell Sys. Tech. Jour.*, vol. 29, pp. 147-160, April 1950.

[10] P. Elias, "Coding for noisy channels," in *IRE Convention Record*, Part IV, pp. 37-46, 1955.

[11] K. A. Bush, "Orthogonal arrays of index unity," *Annals of Math. Stati.*, vol. 23, pp. 426-434, 1952.

[12] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Jour. of the Society for Indus. and Appl. Math. (SIAM)*, vol. 8, no. 2, pp. 300-304, 1960.

[13] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes (1)," in *Proceedings of IEEE ICC*, pp. 1064-1070, May 1993.

[14] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261-1271, October 1996.

[15] N. Wiberg, H.-A. Loeliger, and R. Kotter. "Codes and Iterative Decoding on General Graphs," *Europ. Trans. Telecommun. (ETT)*, vol. 6, no. 5, pp. 513-525, 1995.

[16] N. Wiberg, H.-A. Loeliger, and R. Kotter. "Codes and Iterative Decoding on General Graphs," in *Proceedings of IEEE ISIT*, p. 468, Whistler, Canada, September 1995.

[17] D. J. MacKay and R. M. Neal. "Good Codes based on Very Sparse Matrices," in *Proceedings of Cryptography and Coding, 5th IMA Conference, vol. 1025 of Lecture Notes in Computer Science*, pp. 100-111, Cirencester, UK, December 1995.

[18] D. J. C. MacKay and R. M. Neal. "Near Shannon Limit Performance of Low Density Parity Check Codes," *IEE Electronics Letters*, vol. 32, no. 18, pp. 1645-1646, July 1996.

[19] M. Sipser and D. A. Spielman, "Expander codes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1710-1722, 1996.

[20] D. A. Spielman, "Linear-time encodable and decodable error- correcting codes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 388-397, 1996.

[21] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: M.I.T. Press, 1963.

[22] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," *SIGCOMM Comput. Commun. Rev.*, vol. 28, no. 4, pp. 56-67, 1998.

[23] W. Xiao, S. Agarwal, D. Starobinski and A. Trachtenberg, "Reliable Rateless Wireless Broadcasting With Near-Zero Feedback," *IEEE/ACM Trans. Netw.*, vol. 20, no. 6, pp. 1924-1937, Dec. 2012.

[24] D. M. Mandelbaum, "An adaptive-feedback coding scheme using incremental redundancy," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 388-389, 1974.

[25] B. Dorsch, "Successive check digits rather than information repetition," in *Proceedings of IEEE ICC*, pp. 323-327, Boston, MA, USA, June 1983.

[26] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North-Holland, 1977.

[27] R. M. Tanner, "A Recursive Approach to Low Complexity Codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533-547, 1981.

[28] N. Alon and M. Luby, "A linear time erasure-resilient code with nearly optimal recovery," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1732-1736, 1996.

[29] M. Luby, "Information additive code generator and decoder for communication systems," in U.S. Patent 6,307,487, October 2001.

[30] M. Luby, "Information additive code generator and decoder for communication systems," in U.S. Patent 6,373,406, April 2002.

[31] M. Luby, "LT codes," in *Proceedings of the 43rd IEEE FOCS*, pp. 271-280, 2002.

[32] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman, "Efficient erasure correcting codes," *IEEE Trans. Inf. Theory*, vol. 47, pp. 569-584, Feb. 2001.

[33] A. Shokrollahi, "LDPC codes: An introduction," http://algo.epfl.ch/contents/output/pubs/ldpc-intro.pdf. 2002.

[34] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, Mar. 2008.

[35] P. Maymounkov, "Online codes," Technical Report TR2002-833, Secure Computer Systems Group, New York University, Tech. Rep., November 2002.

[36] A. Shokrollahi, "Raptor Codes," in *Proceedings of IEEE ISIT*, p. 36, Chicago, IL, USA, June 2004.

[37] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551-2567, 2006.

[38] A. Shokrollahi, S. Lassen, and M. Luby, "Multi-Stage Code Generator and Decoder for Communication Systems," U.S. Patent 7,068,729, June 2006.

[39] "3GPP TS 26.346 v6.1.0 (2005-06) Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Protocols and Codecs," Tech. Rep. 2006.

[40] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, "Raptor forward error correction scheme for object delivery," Internet Engineering Task Force RFC 5053, Tech. Rep., October 2007.

[41] M. Luby, A. Shokrollahi, M. Watson, T. Stockhammer, and L. Minder, "Raptorq forward error correction scheme for object delivery," Internet Engineering Task Force RFC 6330, Tech. Rep., August 2011.

[42] QUALCOMM Incorporated, "RaptorQTM Technical Overview," 2010.

[43] ETSI Technical Specification 102.472 V1.3.1., "Digital video broadcasting DVB; IP datacast over DVB-H: Content delivery protocols," Tech. Rep., June 2009.

[44] ETSI Technical Specification 102.034 V1.4.1., "Digital video broadcasting DVB; transport of MPEG-2 TS based DVB services over IP based networks," Tech. Rep., August 2009.

[45] N. Dong, T. Tuan, N. Thinh, and B. Bose, "Wireless broadcast using network coding," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 914-925, 2009.

[46] H. D. T. Nguyen, T. Le-Nam, and H. Een-Kee, "On transmission efficiency for wireless broadcast using network coding and fountain codes," *IEEE Commu. Letters*, vol. 15, no. 5, pp. 569-571, 2011.

[47] C. Fragouli, J. Widmer, and J. Y. Le Boudec, "Efficient broadcasting using network coding," *IEEE/ACM Trans. Netw.*, vol. 16, no. 2, pp. 450-463, 2008.

[48] C. Fragouli, J.-Y. L. Boudec, and J. Widmer, "Network coding: an instant primer," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 63-68, 2006.

[49] A. Eryilmaz, A. Ozdaglar, M. Medard, and E. Ahmed, "On the delay and throughput gains of coding in unreliable networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5511-5524, 2008.

[50] J. W. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE Journal on Selected Areas in Comm.*, vol. 20, pp. 1528-1540, Oct. 2002.

[51] M. Luby, T. Gasiba, T. Stockhammer, and M. Watson, "Reliable multimedia download delivery in cellular broadcast networks," IEEE Trans. Broadcast., vol. 53, no. 1, pp. 235-246, Mar. 2007.

[52] R. Karp, M. Luby, and A. Shokrollahi, "Finite length analysis of LT codes," in *Proceedings of IEEE ISIT*, p. 39, 2004.

[53] A. Shokrollahi, *Theory and applications of Raptor codes*. Springer- Verlag, 2009.

[54] A. Shokrollahi, S. Lassen, and R. Karp, "Systems and processes for decoding chain reaction codes through inactivation," US Patent 6,856,263. [Online]. Available: http://www.google.com/patents/US6856263. Feb. 15 2005.

[55] F. Lu, C. H. Foh, C. J. Cai, and L. Chia, "LT codes decoding: Design and analysis," in *Proceedings of IEEE ISIT*, pp. 2492-2496, 2009.

[56] H. Lu, F. Lu, J. Cai, and C. H. Foh, "LT-W: Improving LT Decoding With Wiedemann Solver," *IEEE Trans. Inf. Theorys*, vol. 59, no. 12, pp. 7887-7897, Dec. 2013.

[57] B. Schotsch, "Rateless coding in the finite length regime," Ph.D. dissertation, Inst. of Commun. Systems and Data Proc., RWTH Aachen, Aachen, Germany, Jul 2014.

[58] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204-1216, 2000.

[59] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371-381, 2003.

[60] P.Elias,A.Feinstein,andC.Shannon, "A note on the maximum flow through a network," *IRE Trans. Inf. Theory*, vol. 2, no. 4, pp. 117-119, 1956.

[61] D. S. Lun, M. Medard, and M. Effros, "On coding for reliable communication over packet networks," in *Proceedings of Allerton conference on commun., control, and computing*, 2004.

[62] D. S. Lun, M. Medard, and R. Koetter, "Efficient operation of wireless packets networks using network coding," in *Proceedings of International Workshop on Convergent Technology*, 2005.

[63] D. S. Lun, N. Ratnakar, M. Medard, R. Koetter, D. R. Karger, T. Ho, and E. Ahmed, "Minimum-cost multicast over coded packet networks," *IEEE Trans. Inf. Theory*, vol. 20, pp. 1528-1540, 2006.

[64] P. Wang, G. Mao, Z. Lin, X. Ge, and B. Anderson, "Network coding based wireless broadcast with performance guarantee," *IEEE Trans. Wireless Commun.*, vol.14, no.1, pp.532-544, Jan. 2015.

[65] P. Wang, G. Mao, Z. Lin, and X. Ge, "An efficient network coding based broadcast scheme with reliability guarantee," in *Proceedings of IEEE ICC*, pp. 2879-2884, 2014.

[66] N. Rahnavard, B. Vellambi, and F. Fekri, "Rateless codes with unequal error protection property," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1521-1532, April 2007.

[67] B. LaMacchia and A. Odlyzko, "Solving Large Sparse Linear Systems over Finite Fields," *Advances in Cryptology-CRYPTO '90, vol. 537 of Lecture Notes in Computer Science*, pp. 109-133. Springer, 1991.

[68] C. Pomerance and J. W. Smith, "Reduction of Huge, Sparse Matrices over Finite Fields Via Created Catastrophes," *Experimental Math.*, vol. 1, pp. 89-94, 1992.

[69] D. Coppersmith, "Solving Linear Equations over GF(2): Block Lanc- zos Algorithm," *Linear Algebra and its Applications*, vol. 192, no. 0, pp. 33-60, 1993.

[70] D. Coppersmith, "Solving Homogeneous Linear Equations over GF(2) via Block Wiedemann Algorithm," *Math of Computation*, vol. 62, no. 205, pp. 333-350, 1994.

[71] D. Burshtein and G. Miller, "Efficient Maximum-Likelihood Decoding of LDPC Codes over the Binary Erasure Channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2837-2844, 2004.

[72] H. Pishro-Nik and F. Fekri, "On Decoding of Low-Density Parity- Check Codes over the Binary Erasure Channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 439-454, 2004.

[73] G. Liva, B. Matuz, E. Paolini, and M. Chiani, "Pivoting Algorithms for Maximum Likelihood Decoding of LDPC Codes over Erasure Chan- nels," in *Proceedings of IEEE GLOBECOM*, pp. 1-6, 2009.

[74] E. Paolini, G. Liva, B. Matuz, and M. Chiani, "Maximum Likelihood Erasure Decoding of LDPC Codes: Pivoting Algorithms and Code Design," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3209-3220, 2012.

[75] A. Shokrollahi, S. Lassen, and R. Karp, "Systems and processes for decoding chain reaction codes through inactivation," in U.S. Patent Number 6856263, 2005.

[76] A. Shokrollahi and M. Luby, *Raptor Codes*, vol. 6, pp. 213-322, 2011.

[77] J. Yue, Z. Lin, B. Vucetic, G. Mao, and T. Aulin, "Performance analysis of distributed performance analysis of distributed raptor codes in wireless sensor networks," *IEEE Trans. Commun*, vol. 61, no. 10, pp. 4357-4368, 2013.

[78] D. MacKay, "Fountain Codes," *IEE Proceedings Communications*, vol. 152, no. 6, pp. 1062-1068, 2005.

[79] B. Schotsch, H. Schepker, and P. Vary, "The Performance of Short Random Linear Fountain Codes under Maximum Likelihood Decoding," in *Proceedings of IEEE ICC*, pp. 1-5, Kyoto, Japan, June 2011.

[80] B. Schotsch, R. Lupoaie, and P. Vary, "The Performance of Low- Density Random Linear Fountain Codes over Higher Order Galois Fields under Maximum Likelihood Decoding," *Annual Allerton Conference on Commun., Control and Computing*, pp. 1004-1011, Monticello, IL, USA, September 2011.

[81] B. Schotsch and P. Vary, "Design of Unequally Error Protecting Low-Density Random Linear Fountain Codes," in *Proceedings of IEEE PIMRC*, Sydney, Australia, September 2012.

[82] B. Schotsch and R. Lupoaie, "Finite Length LT Codes over Fq for Unequal Error Protection with Biased Sampling of Input Nodes," in *Proceedings of IEEE ISIT*, Cambridge, MA, USA, July 2012.

[83] B. Schotsch, G. Garrammone, and P. Vary, "Analysis of LT Codes over Finite Fields under Optimal Erasure Decoding," *IEEE Commun. Letters*, vol. 17, no. 9, pp. 1826-1829, September 2013.

[84] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proceedings of 29th ACM Symposium on Theory of Computing*, El Paso, Texas, USA, 1997.

[85] M. Ghaderi, D. Towsley, and J. Kurose, "Reliability gain of network coding in lossy wireless networks," in *Proceedings of 27th IEEE INFOCOM*, pp. 13-18, April 2008.

[86] T. Stockhammer, A. Shokrollahi, M. Watson, M. Luby, and T. Gasiba, "Application Layer Forward Error Correction for Mobile Multimedia Broadcasting," in Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T and Media FLO, B. Furhet and S. Ahson, Eds. Boca Raton, FL: CRC Press, pp. 239-280, 2008.

[87] C. D. Meyer, *Matrix analysis and applied linear algebra.* SIAM, 2000.

[88] L. Comtet, *Advanced Combinatorics.* Reidel, 1974.

[89] B.E. Schotsch, H. Schepker, and P. Vary, "The Performance of Short Random Linear Fountain Codes under Maximum Likelihood Decoding," in *Proceedings of IEEE ICC*, pp. 1-5, Kyoto, Japan, June 2011.

[90] M. Mitzenmacher, "Digital fountains: A survey and look forward," in *Proceedings of Inf. Theory workshop*, pp. 271-276, Oct. 2004.

[91] E. Hyytia, T. Tirronen, and J. Virtamo, "Optimal degree distribution for LT codes with small message length," in *Proceedings of 26th IEEE INFOCOM*, pp. 2576-2580, 2007.

[92] G. Landsberg, "Ueber eine anzahlbestimmung und eine damit zusammenhaÌˆngende reihe," *Journal fuÌˆr diereine und angewandte Mathematik*, vol. 111, pp. 87-88, 1893.

[93] P. Erdos and A. Renyi, "On random matrices," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences (Series A)*, vol. 8, pp. 455-461, 1963.

[94] A. M. Odlyzko, "On the ranks of some (0, 1)-matrices with constant row sums," *Journal of the Australian Mathematical Society (Series A)*, vol. 21, no. 2, pp. 193-201, 1981.

[95] V. F. Kolchin, "Random graphs and systems of linear equations in finite fields," *Random Structures and Algorithms*, vol. 5, no. 1, pp. 135-146, 1994.

[96] J. Blomer, R. M. Karp, and E. Welzl, "The rank of sparse random matrices over finite fields," *Random Structures and Algorithms*, vol. 10, no. 4, pp. 407-419, July 1997.

[97] V. F. Kolchin, *Random Graphs*, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999.

[98] C. Cooper, "On the distribution of rank of a random matrix over a finite field," *Random Structures and Algorithms*, vol. 17, no. 3-4, pp. 197-212, October-December 2000.

[99] C. Cooper, "On the rank of random matrices," *Random Structures and Algorithms*, vol. 16, no. 2, pp. 209-232, March 2000.

[100] A. A. Levitskaya, "Systems of random equations over finite algebraic structures," *Cybernetics and Systems Analysis*, vol. 41, pp. 67-93, 2005.

[101] A. Tukmanov, D. Zhiguo, S. Boussakta, and A. Jamalipour, "On the broadcast latency in finite cooperative wireless networks," *IEEE Trans. Wireless Commun*, vol. 11, no. 4, pp. 1307-1313, 2012.

[102] T. Rappaport, *Wireless Communications: Principles and Practice,* ser. Prentice Hall Communication Engineering and Emerging Technologies. Prentice Hall PTR, 2002.

[103] S. Rajabi, M. Shahabadi, and M. ArdebiliPoor, "Modeling of the correlation coefficients of a receive antenna array in a MIMO multipath channel," in *Proceedings of 2nd IEEE/IFIP ICI*, pp. 1-4, 2006.

[104] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes,* 2nd ed., ser. Probability and its Applications. Verlag, vol. I, 2003.