2014

Physical layer security against pollution attack in wireless relay networks using random network coding

Duk Hee Yoon
Iowa State University

Physical layer security against pollution attack
in wireless relay networks using random network coding

by

Duk Hee Yoon

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Computer Engineering

Program of Study Committee:
Sang Wu Kim, Major Professor
Zhengdao Wang
Ahmed Kamal
Yong Guan
Tanya Rosenblat

Iowa State University
Ames, Iowa
2014

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First of all, I am deeply grateful to my major professor, Dr. Sang Wu Kim, for his patience and encouragement. From him, I have learned how to conduct research with diligence and endurance. In addition, his insightful feedbacks have been great guidance whenever I was stuck on my research problems. I also offer many thanks to my committee members: Dr. Zhengdao Wang, Dr. Ahmed Kamal, Dr. Yong Guan, and Dr. Tanya Rosenblat. Their feedbacks have been invaluable to improve my research. I would also like to thank my friends whom I met at Iowa State University: Young Jin Chun, Taha Khalaf, Navneet Malani, Hien Ta, Mahdi Zamanighomi, Titus Rotich, Songtao Lu, Abdulkadir Celik, Hye Won Lee, Chang Geun Yoo, Jungmin Park, Jongho Im, and Jungwook Paek. They have encouraged me throughout this long journey. Finally, I express my sincere appreciation to my family for their endless and unconditional love throughout my life.

ABSTRACT

Network coding can remarkably improve the network capacity by combining incoming packets at intermediate nodes. However, the packet combining also causes the network to be particularly vulnerable to the pollution attack that injects false data into the information flow of the network. This dissertation includes two studies on mitigating pollution attack in two-hop wireless relay network that employs random network coding.

First, we investigate how the finite field size affects the network coding performance in terms of the probability of symbol error and the throughput in adversarial networks where the false data is injected by the malicious attackers at source nodes and/or relay nodes. Also, we examine how the optimal field size that minimizes the probability of symbol error or that maximizes throughput changes as the trustworthiness of node or the number of combined packets changes.

Second, we propose two schemes for detecting the polluted packets and discarding them before decoding by exploiting physical layer information which is directly overheard from the source nodes. The proposed scheme I applies the threshold-based method to detect the presence of falsely injected data within each packet, while the proposed scheme II compares all received network-coded packets and selects the most trustable ones. Unlike many existing signature-based detection schemes, the proposed schemes do not require that additional information bits are attached into each packet.

CHAPTER 1. INTRODUCTION

In this chapter, we introduce backgrounds of channel coding, network coding, and network coding security. Then, we review the advanced related literatures for security issues on network coding. Finally, we discuss the contributions and the outline of this dissertation.

1.1 Channel Coding (Reed-Solomon Codes)

The wireless communication system applies channel encoding and decoding techniques, in order to prevent the transmitted data from being received incorrectly due to wireless channel impairments. Let

$$\mathbf{m} = \{m_1, \cdots, m_K\} \tag{1.1}$$

be the message word that the transmitter node wants to send to the receiver node where m_k is the kth symbol over finite field $GF(q)$. Fig.1.1 shows the overall data transmission model with channel encoding and decoding procedures. At first, the transmitter node encodes (transforms) the message word $\mathbf{m}$ to the corresponding channel codeword. If systematic Reed-Solomon codes [13],[23] are used, the codeword encoded from $\mathbf{m}$ is given by

$$\mathbf{x} = \{x_1, \cdots, x_N\} \tag{1.2}$$

$$= \{\underbrace{m_1, \cdots, m_K}_{\mathbf{m}}, \underbrace{w_1, \cdots, w_{N-K}}_{\text{parity symbols}}\} \tag{1.3}$$

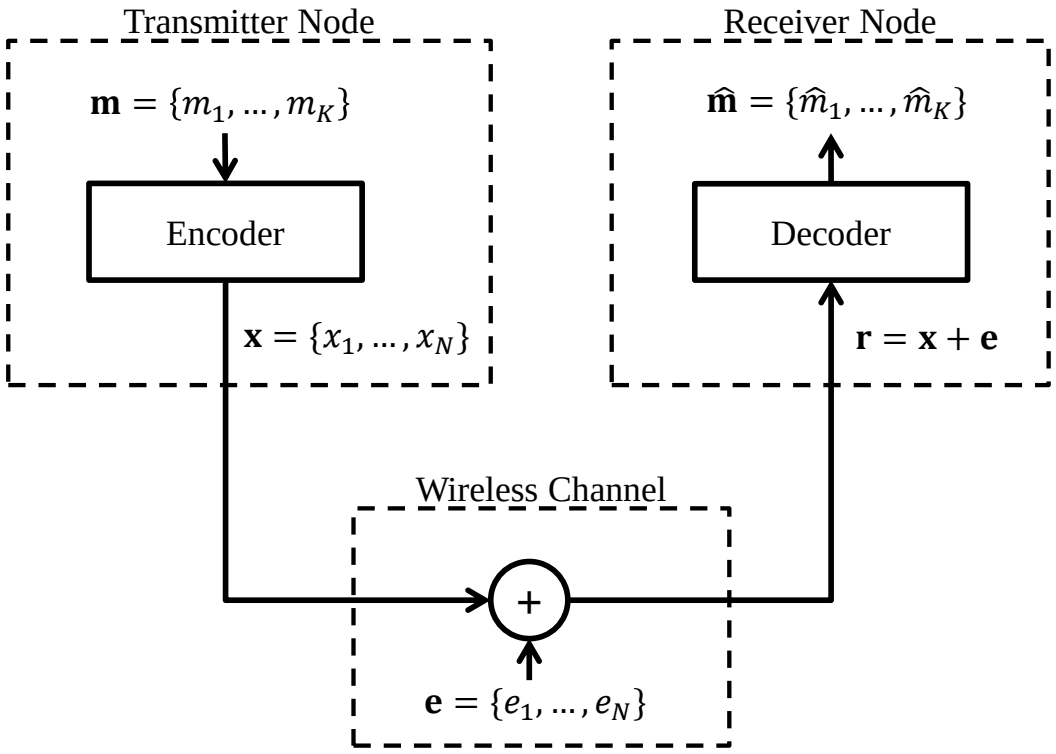

Figure 1.1 Data transmission model with channel encoding and decoding.

where x_n denotes the nth symbol over $GF(q)$ and $N = q - 1$ denotes the length of a codeword. Since the codeword is systematic, the first K symbols of the codeword are equal to message symbols and the following $N - K$ symbols $w_1, \cdots, w_{N-K}$ are parity symbols which are calculated from $\mathbf{m}$ by Reed-Solomon encoder. When the transmitter node sends $\mathbf{x}$ through wireless channel to the receiver node, the channel error vector

$$\mathbf{e} = \{e_1, \cdots, e_N\} \tag{1.4}$$

is added to $\mathbf{x}$ where e_n is over $GF(q)$, so that the receiver node obtains

$$\mathbf{r} = \mathbf{x} + \mathbf{e} \tag{1.5}$$

which can be seen as the noisy version of $\mathbf{x}$. And then, Reed-Solomon decoder at the receiver node decodes $\mathbf{r}$ and obtains

$$\hat{\mathbf{m}} = \{\hat{m}_1, \cdots, \hat{m}_K\} \tag{1.6}$$

as the decoder output.

If the codeword length N is desired to be smaller than $q-1$, shortened Reed-Solomon codes can be applied. In this case, the transmitter node generates a new message word

$$\mathbf{m}^{'} = \{\underbrace{0, \cdots, 0}_{\substack{q-N-1 \\ \text{zeros}}}, \underbrace{m_1, \cdots, m_K}_{\mathbf{m}}\} \tag{1.7}$$

by inserting $q-N-1$ zeros before the original message word $\mathbf{m}$. The systematic codeword encoded from $\mathbf{m}^{'}$ is given by

$$\mathbf{x}^{'} = \{\overbrace{0, \cdots, 0, \underbrace{m_1, \cdots, m_K}_{\mathbf{x}}}^{\mathbf{m}^{'}}, w_1^{'}, \cdots, w_{N-K}^{'}\} \tag{1.8}$$

where $w_1^{'}, \cdots, w_{N-K}^{'}$ are parity symbols calculated from $\mathbf{m}^{'}$ by Reed-Solomon encoder. As a result, the length of $\mathbf{x}^{'}$ is $q-1$. Then, the desired codeword

$$\mathbf{x} = \{m_1, \cdots, m_K, w_1^{'}, \cdots, w_{N-K}^{'}\} \tag{1.9}$$

with the length N is obtained by removing the first $q-N-1$ zero symbols from $\mathbf{x}^{'}$. The transmitter node sends $\mathbf{x}$ and the receiver nodes receives $\mathbf{r} = \mathbf{x} + \mathbf{e}$. Then, the receiver node generates

$$\mathbf{r}^{'} = \{\underbrace{0, \cdots, 0}_{\substack{q-N-1 \\ \text{zeros}}}, \underbrace{r_1, \cdots, r_N}_{\mathbf{r}}\} \tag{1.10}$$

by adding $q-N-1$ zero symbols before $\mathbf{r}$ and put it into the Reed-Solomon decoder. Then,

$$\hat{\mathbf{m}}^{'} = \{\hat{m}_1^{'}, \cdots, \hat{m}_{q-N-1}^{'}, \underbrace{\hat{m}_{q-N}^{'}, \cdots, \hat{m}_{q-N+K-1}^{'}}_{\hat{\mathbf{m}}}\} \tag{1.11}$$

is obtained as the decoder output. By removing the first $q-N-1$ symbols from $\hat{\mathbf{m}}^{'}$, the reconstructed message word $\hat{\mathbf{m}}$ is obtained.

By the property of Reed-Solomon codes, the decoder successfully reconstructs $\mathbf{m}$ (i.e., $\hat{\mathbf{m}} = \mathbf{m}$) if there are at most

$$t = \left\lfloor \frac{N-K}{2} \right\rfloor \tag{1.12}$$

different symbols between $\mathbf{r}$ and $\mathbf{x}$. In other words, $\hat{\mathbf{m}}$ is not equal to $\mathbf{m}$ if the channel error vector $\mathbf{e}$ has more than t nonzero symbols. In (1.12), $\lfloor a \rfloor$ denotes the largest integer which is not larger than a. For example, $\lfloor 1.9 \rfloor = 1$ and $\lfloor 2 \rfloor = 2$.

1.2 Network Coding

Network coding is an innovative relaying strategy that remarkably advances throughput, reliability, and capacity of the network by a simple and powerful idea that each intermediate node (i.e., relay node) in the network combines multiple packets into a packet called network-coded packet and send it [1]. A network-coded packet which is successfully received at the destination node is represented as

$$\mathbf{p}_r = \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s, \ r = 1, \cdots, R \tag{1.13}$$

where R denotes the number of received network-coded packets, r denotes the packet index, and S denotes the number of combined message packets. The network coding coefficient vector

$$\mathbf{c}_r = \{c_{r,1}, \cdots, c_{r,S}\} \tag{1.14}$$

where $c_{r,s} \in GF(q)$ denotes the network coding coefficient corresponding to the sth message packet $\mathbf{x}_s$ is attached in the header of the network-coded packet [4]. If the destination node receives S network-coded packets $\mathbf{p}_1, \cdots, \mathbf{p}_S$ whose coding coefficient vectors $\mathbf{c}_1, \cdots, \mathbf{c}_S$ are linearly independent, the original message packets are recovered

by

$$\underbrace{\begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1}}_{\mathbf{C}} \begin{bmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_S \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} \tag{1.15}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix}. \tag{1.16}$$

The basic idea of random network coding [2] is to randomly select the network coding coefficients from the finite field $GF(q)$. By using random network coding, relay nodes do not need to communicate each other to decide linearly independent network coding coefficient vectors. In other words, random network coding enhances the distributed characteristics of the network than the conventional deterministic network coding. Instead, there is the nonzero probability that network-coded packets received at the destination node are not linearly independent, so that the original message packets cannot be recovered. However, this additional communication overhead can be negligibly small by applying large field size (e.g., $q = 256$), at the cost of higher computational complexity to find the inverse matrix of $\mathbf{C}$.

1.3 Security Issues on Network Coding

In this section, we introduce several kinds of security attacks for wireless relay networks using network coding, focusing on the pollution attack that we address throughout this dissertation.

1.3.1 Eavesdropping

In eavesdropping, the malicious adversaries try to secretly overhear transmitted packets to understand the content of those packets. It is known that network coding has better

robustness against eavesdropping than the conventional packet forwarding, because network coding basically combines the original message packets [5]. In order to understand the content of the original message packets in network coding, the eavesdroppers must successfully listen the enough number of linearly independent network-coded packets. If they fail in this, they cannot understand any of the message packets.

1.3.2 Entropy Attack

In entropy attack [30],[34],[37], the malicious adversaries intentionally transmit linearly dependent (i.e., non-innovative) network-coded packets, in order to waste the network resources to send the useless packets. That is, entropy attackers badly use the property that destination nodes in network coding require the enough number of linearly independent network-coded packets in order to reconstruct the original message packets.

1.3.3 Pollution Attack

In pollution attack, the malicious relays intentionally send false packets which are different from true packets, in order to prevent destinations from receiving correct information [38]. It is also called false packet injection attack or Byzantine attack. Considering pollution attack, a network-coded packet which is received at the destination is represented as

$$\mathbf{p}_r = \underbrace{\sum_{s=1}^{S} c_{r,s}\mathbf{x}_s}_{\text{true packet}} + \mathbf{f}_r, \ r = 1, \cdots, R \tag{1.17}$$

where $\mathbf{f}_r$ denotes the falsely injected vector capturing any type of modification that causes $\mathbf{p}_r$ to be different from the true network-coded packet by [19]. If $\mathbf{f}_r$ is a nonzero vector, the packet $\mathbf{p}_r$ is called polluted. Otherwise, $\mathbf{p}_r$ is called unpolluted.

It has been shown that packet-mixing property of network coding might cause the network to be particularly weak to pollution attack [10]. This is because all message

packets can be incorrectly reconstructed due to even one polluted network-coded packet. If the destination receives network-coded packets $\mathbf{p}_1, \cdots, \mathbf{p}_S$ and their coefficient vectors $\mathbf{c}_1, \cdots, \mathbf{c}_S$ are linearly independent, the reconstructed message packets are given by

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \vdots \\ \hat{\mathbf{x}}_S \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1}}_{\mathbf{C}} \begin{bmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_S \end{bmatrix} \tag{1.18}$$

$$= \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} + \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_S \end{bmatrix} \tag{1.19}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} + \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_S \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_S \end{bmatrix} \tag{1.20}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} + \underbrace{\begin{bmatrix} \tilde{c}_{1,1} & \cdots & \tilde{c}_{1,S} \\ \vdots & \ddots & \vdots \\ \tilde{c}_{S,1} & \cdots & \tilde{c}_{S,S} \end{bmatrix}}_{\mathbf{C}^{-1}} \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_S \end{bmatrix} \tag{1.21}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} + \begin{bmatrix} \tilde{c}_{1,1}\mathbf{f}_1 + \cdots + \tilde{c}_{1,S}\mathbf{f}_S \\ \vdots \\ \tilde{c}_{S,1}\mathbf{f}_1 + \cdots + \tilde{c}_{S,S}\mathbf{f}_S \end{bmatrix}. \tag{1.22}$$

From (1.22), we can see that each false injection vector $\mathbf{f}_r$ is multiplied by each element in $\mathbf{C}^{-1}$ and is added to all of the original message packet $\mathbf{x}_1, \cdots, \mathbf{x}_S$. If q is not small, it is likely that all elements of $\mathbf{C}^{-1}$ are nonzero. Therefore, if $\mathbf{f}_1$ is a nonzero vector, it is likely that all of $\mathbf{x}_1, \cdots, \mathbf{x}_S$ are affected by $\mathbf{f}_1$. As a result, one polluted network-coded packet can cause that all message packets are reconstructed wrongly.

1.4 Related Literatures

Security issues of network coding have been addressed by many researchers in order to make the network coding more practical and reliable on the adversarial environment. The mixing characteristic of the network coding makes the network robust against the eavesdropping attack [5], but vulnerable to the false injection attack because even one polluted network-coded packet can cause all message packets to be incorrectly recovered [6, 10]. Several approaches have been suggested to overcome this drawback of network coding. Kim *et al.* [7] proposed a watchdog scheme that examines whether a relay node is malicious or not by comparing packets which are received from different nodes. Ho *et al.* [8] proposed a scheme that attaches the additional information into a network-coded packet, in order for the destination node to detect if the received network-coded packet is modified by malicious adversary. Jaggi *et al.* [11] proposed a signature-based distributed scheme to detect Byzantine attacks. Kim *et al.* [9] proposed packet recycling scheme that restores the true network-coded packet from the polluted network-coded packet by exploiting the physical layer information. Tran *et al.* [28] proposed an authentication method for multi-casting using network coding, based on its null space characteristics. Yu *et al.* [31] suggested a new homomorphic signature scheme for network coding, not to allow malicious relays to make a new signatures for their falsely injected packets. Gkantsidis *et al.* [30] proposed a scheme that nodes collaborate to inform each other about polluted data blocks, for peer-to-peer networks with network coding. Oggier *et al.* [33] proposed an authentication scheme that polluted packets can be discarded through verification at relays, as well as at destinations. Kehdi *et al.* [36] suggested a new detection scheme for random network coding against pollution attack, by exploiting subspace characteristics of random linear network coding. Yu *et al.* [32] proposed a scheme employing MAC and key-predistribution for XOR-ing network coding against pollution attack. Authors in [24], [25], [26], [27], [29], [35] proposed homomorphic MAC

(Message Authentication Code) schemes to detect pollution attacks for network coding.

1.5 Dissertation Contributions and Outline

In this dissertation, we consider the two-hop wireless relay network in which malicious adversaries can access network nodes and inject false data into true packets. As a result, the destination nodes receive those polluted packets. In the pursuit of mitigating the damage from pollution attack, we address the following two problems.

1. How parameters related to network-coding affect the performance, under pollution attack?

2. How can the destination node detect polluted packets?

Chapter 2 investigates the first problem. We study how the finite field size q affects the throughput and the probability of symbol error, on wireless two-hop relay networks employing random network coding in which each node might intentionally inject false data and the transmitted data is subject to wireless channel error. We show that the probability of correct decoding exponentially decreases as the number of combined packets increases. We also present that the throughput is proportional to $\frac{\log_2 q}{q}$ bits per transmitted symbol. We investigate the optimum field size to minimize the probability of symbol error or to maximize the throughput, in terms of the number of combined packets and the trustworthiness of node. This chapter was modified from a paper published in [20].

Chapter 3 addresses the second problem. We suggest two physical-layer schemes to detect the integrity of received network-coded packets and exploit the detection result for decoding the original messages, in the presence of false injection attacks. The integrity is determined by utilizing the physical layer information that is directly overheard from source nodes to the destination node. The proposed scheme I uses the threshold-based approach to detect the presence of attack upon arrival of each network-coded packet,

while the proposed scheme II compares all received network-coded packets and selects the most trustable packets. We show that the proposed schemes provide the significantly smaller probability of decoding error than the conventional random selection scheme which does not exploit the physical-layer information and that they perform close to the cryptographic scheme that requires computational and bandwidth overheads. We also analyze the average delay showing that the proposed scheme I reconstructs the original messages earlier than the proposed scheme II at the expense of the higher probability of decoding error. We also provide the analysis and numerical results of the average throughput. Part of this work has been published in [21].

Finally, Chapter 4 discusses future research works and concludes the dissertation.

CHAPTER 2. FIELD SIZE OF RANDOM NETWORK CODING IN UNTRUSTWORTHY NETWORKS

2.1 Introduction

In this chapter, we analyze the probability of symbol error and the throughput with random network coding in untrustworthy networks, where the malicious attackers access network nodes and purposely inject false data and data transmissions are subject to wireless channel errors. We show that there exists an optimal field size that maximizes the throughput (bits/channel use) or minimizes the probability of symbol error. We examine the optimal field size in terms of the trustworthiness of node and the number of combined packets in untrustworthy network. We also derive the asymptotic throughput and asymptotic probability of decoding error as the number of combined packets becomes large and examine how they change depending on the field size, the number of combined packets, and node trustworthiness. We show that the probability of correct decoding exponentially decreases with the increasing number of combined packets and that the maximum throughput scales as $\frac{\log_2 q}{q}$ bits per symbol transmission where q is the finite field size. The former suggests to restrict the number of combined packets when node trustworthiness is low.

The remainder of this chapter is organized as follows. Section 2.2 describes the system model. Section 2.3 provides the derivation and the numerical results of the probability of symbol error. Section 2.4 presents the derivation and the numerical results of the throughput. Finally, Section 2.5 concludes the chapter.

2.2 System Model

We consider a two-hop wireless network consisting of S source nodes, R relay nodes, and multiple destination nodes. Source nodes generate and send the new information to the destination nodes through relay nodes where the packets from S source nodes are linearly combined and sent to the destination nodes, as shown in Fig. 2.1.

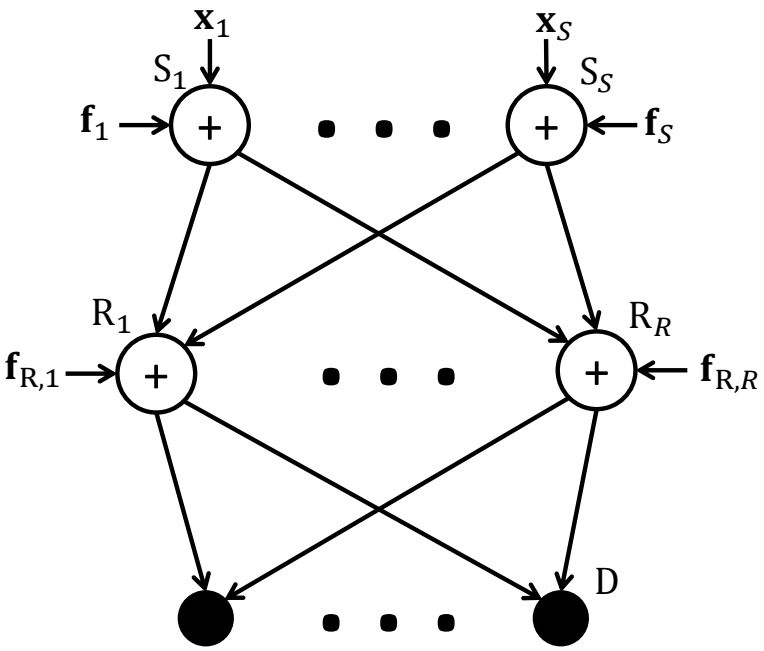

Figure 2.1 Two-hop network model.

The sth source node S_s sends a packet

$$\tilde{\mathbf{x}}_s = \mathbf{x}_s + \mathbf{f}_s, \ \ s = 1, \cdots, S \tag{2.1}$$

where $\mathbf{x}_s$ is the true message packet and $\mathbf{f}_s$ captures the intentional false packet injected by S_s (pollution attack). We assume that $\mathbf{x}_s$ and $\mathbf{f}_s$ are vectors of elements from a Galois field $GF(q)$. We also assume that the elements of $\mathbf{f}_s$ are independently chosen and that nonzero values of f_s are equiprobable. If $\mathbf{f}_s = 0$, then no false packet is injected by S_s.

The *trustworthiness of a node* is measured by the probability that the transmitted packet from the node is equal to its true message packet, i.e., $P(\mathbf{f}_s = \mathbf{0}) = P(\tilde{\mathbf{x}}_s = \mathbf{x}_s)$.

The rth relay node R_r receives

$$\dot{\mathbf{x}}_{s,r} = \tilde{\mathbf{x}}_s + \mathbf{e}_{s,r},\ r = 1, \cdots, R \tag{2.2}$$

where $\mathbf{e}_{s,r}$ denotes the packet error between S_s and R_r. If $\mathbf{e}_{s,r} = \mathbf{0}$, R_r receives $\tilde{\mathbf{x}}_s$ correctly and, otherwise, R_r receives $\tilde{\mathbf{x}}_s$ incorrectly. We assume that the elements of $\mathbf{e}_{s,r}$ (symbol errors) are independent. Then, the rth relay node linearly combines $\dot{\mathbf{x}}_{s,r}$ to produce a coded packet

$$\tilde{\mathbf{x}}_{\mathrm{R},r} = \sum_{s=1}^{S} c_{r,s}\dot{\mathbf{x}}_{s,r} + \mathbf{f}_{\mathrm{R},r},\ r = 1, \cdots, R \tag{2.3}$$

where the coefficients $\{c_{r,s}\}$ are randomly chosen with equal probability from $GF(q)$ and $\mathbf{f}_{\mathrm{R},r}$ captures the intentionally or unknowingly injected false packet by R_r. The multiplications and summations in (2.3) are over $GF(q)$. The probability $P(\mathbf{f}_{\mathrm{R},r} = \mathbf{0})$ represents the trustworthiness of node R_r. We assume that the coefficients are contained in the packet header and are known by all destination nodes. We assume that $\mathbf{f}_{\mathrm{R},r}$'s are independent across relays.

The reference destination node D receives

$$\mathbf{y}_r = \tilde{\mathbf{x}}_{\mathrm{R},r} + \mathbf{e}_{\mathrm{R},r},\ r = 1, \cdots, R \tag{2.4}$$

$$= \sum_{s=1}^{S} c_{r,s}\dot{\mathbf{x}}_{s,r} + \mathbf{f}_{\mathrm{R},r} + \mathbf{e}_{\mathrm{R},r} \tag{2.5}$$

$$= \sum_{s=1}^{S} c_{r,s}(\tilde{\mathbf{x}}_s + \mathbf{e}_{s,r}) + \mathbf{f}_{\mathrm{R},r} + \mathbf{e}_{\mathrm{R},r} \tag{2.6}$$

$$= \sum_{s=1}^{S} c_{r,s}(\mathbf{x}_s + \mathbf{f}_s + \mathbf{e}_{s,r}) + \mathbf{f}_{\mathrm{R},r} + \mathbf{e}_{\mathrm{R},r} \tag{2.7}$$

where $\mathbf{e}_{\mathrm{R},r}$ denotes the packet error between R_r and D.

The linear independence (innovativeness) of coded packets can be checked by examining the coding coefficients that are placed in the packet header. Given R coded packets,

the probability that there exists a set of S coded packets that are linearly independent is given by

$$P_I(R,S) = \prod_{m=R-S+1}^{R} \left(1 - \frac{1}{q^m}\right) \tag{2.8}$$

The probability of receiving such a set of linearly independent vectors given R packet transmissions is $P_I(R,S)$. The destination may recover the message packets by calculating

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \vdots \\ \hat{\mathbf{x}}_S \end{bmatrix} = \mathbf{C}^{-1} \begin{bmatrix} \mathbf{y}_{[1]} \\ \vdots \\ \mathbf{y}_{[S]} \end{bmatrix} \tag{2.9}$$

$$= \mathbf{C}^{-1} \left(\mathbf{C} \begin{bmatrix} \mathbf{x}_1 + \mathbf{f}_1 + \mathbf{e}_{1,[1]} \\ \vdots \\ \mathbf{x}_S + \mathbf{f}_S + \mathbf{e}_{S,[S]} \end{bmatrix} + \begin{bmatrix} \mathbf{f}_{\mathrm{R},[1]} + \mathbf{e}_{\mathrm{R},[1]} \\ \vdots \\ \mathbf{f}_{\mathrm{R},[S]} + \mathbf{e}_{\mathrm{R},[S]} \end{bmatrix} \right) \tag{2.10}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} + \begin{bmatrix} \mathbf{f}_1 + \mathbf{e}_{1,[1]} \\ \vdots \\ \mathbf{f}_S + \mathbf{e}_{S,[S]} \end{bmatrix} + \mathbf{C}^{-1} \begin{bmatrix} \mathbf{f}_{\mathrm{R},[1]} + \mathbf{e}_{\mathrm{R},[1]} \\ \vdots \\ \mathbf{f}_{\mathrm{R},[S]} + \mathbf{e}_{\mathrm{R},[S]} \end{bmatrix} \tag{2.11}$$

where $[i] \in \{1, \cdots, R\}$ denotes the index of the selected coded packet, and

$$\mathbf{C} = \begin{bmatrix} c_{[1],1} & \cdots & c_{[1],S} \\ \vdots & \ddots & \vdots \\ c_{[S],1} & \cdots & c_{[S],S} \end{bmatrix} \tag{2.12}$$

is the $S \times S$ encoding matrix, and

$$\mathbf{C}^{-1} = \begin{bmatrix} \tilde{c}_{1,1} & \cdots & \tilde{c}_{1,S} \\ \vdots & \ddots & \vdots \\ \tilde{c}_{S,1} & \cdots & \tilde{c}_{S,S} \end{bmatrix} \tag{2.13}$$

is the inverse matrix of $\mathbf{C}$. Then,

$$\hat{\mathbf{x}}_s = [\tilde{c}_{s,1}, \cdots, \tilde{c}_{s,S}] \begin{bmatrix} \mathbf{y}_{[1]} \\ \vdots \\ \mathbf{y}_{[S]} \end{bmatrix}$$
$$= \mathbf{x}_s + \mathbf{f}_s + \mathbf{e}_{s,[s]} + \sum_{i=1}^{S} \tilde{c}_{s,i}(\mathbf{f}_{\mathrm{R},[i]} + \mathbf{e}_{\mathrm{R},[i]}) \tag{2.14}$$

denotes the sth recovered message packet.

From (2.14), the message packet $\mathbf{x}_s$ can be successfully recovered, (i.e., $\hat{\mathbf{x}}_s = \mathbf{x}_s$) if

$$\mathbf{f}_s + \mathbf{e}_{s,[s]} + \sum_{i=1}^{S} \tilde{c}_{s,i}(\mathbf{f}_{\mathrm{R},[i]} + \mathbf{e}_{\mathrm{R},[i]}) = \mathbf{0}. \tag{2.15}$$

It should be noted from (2.15) that decoding of $\mathbf{x}_s$ is not affected by the falsely injected packets at other sources $\{\mathbf{f}_j, j \neq s\}$ and the channel errors on other source-to-relay links $\{\mathbf{e}_{j,[s]}, j \neq s\}$, while the falsely injected packets at the relay and the channel errors on the relay-to-destination links affect the decoding of all $\mathbf{x}_1, \cdots, \mathbf{x}_S$.

We consider q-ary quadrature amplitude modulation (QAM) over a Rayleigh fading channel with additive white Gaussian noise (AWGN). QAM is currently used for higher order modulation in several wireless standards. In this setting, the symbol error probability between S_s and R_r is given by [22]

$$P(e_{s,r} \neq 0) = \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{q}}\right) \int_0^{\pi/2} \left(1 + \frac{1.5h\gamma_b}{(q-1)\sin^2\phi}\right)^{-1} d\phi$$
$$- \frac{4}{\pi}\left(1 - \frac{1}{\sqrt{q}}\right)^2 \int_0^{\pi/4} \left(1 + \frac{1.5h\gamma_b}{(q-1)\sin^2\phi}\right)^{-1} d\phi \tag{2.16}$$
$$\approx \frac{2(q-1)}{3h\gamma_b} \tag{2.17}$$

where $e_{s,r}$ is a symbol of $\mathbf{e}_{s,r}$ and

$$h = \frac{S \log_2 q}{S + R} \tag{2.18}$$

and γ_b is the received SNR per information bit at a relay.

2.3 Probability of Symbol Error

In this section we derive the probability of symbol error after decoding

$$P\left(f_s + e_{s,[s]} + \sum_{i=1}^{S} \tilde{c}_{s,i}(f_{\mathrm{R},[i]} + e_{\mathrm{R},[i]}) = 0\right) \tag{2.19}$$

where f_s, $e_{s,[s]}$, $f_{\mathrm{R},[i]}$, $e_{\mathrm{R},[i]}$ denotes a symbol of $\mathbf{f}_s$, $\mathbf{e}_{s,[s]}$, $\mathbf{f}_{\mathrm{R},[i]}$, and $\mathbf{e}_{\mathrm{R},[i]}$, respectively. Let

$$\begin{aligned} u_{[i]} &= f_{\mathrm{R},[i]} + e_{\mathrm{R},[i]} \\ v_s &= f_s + e_{s,[s]} \end{aligned} \tag{2.20}$$

denote a symbol of $\mathbf{f}_{\mathrm{R},[i]} + \mathbf{e}_{\mathrm{R},[i]}$ and $\mathbf{f}_s + \mathbf{e}_{s,[s]}$, respectively. Then, it follows from (2.15) that a symbol x_s in $\mathbf{x}_s$ can be correctly decoded, (i.e., $\hat{x}_s = x_s$) if

$$v_s + \sum_{i=1}^{S} \tilde{c}_{s,i} u_{[i]} = 0. \tag{2.21}$$

It can be shown that the conditional probability that x_s is correctly decoded given

$$\mathbf{u} := (u_{[1]}, \cdots, u_{[S]}) = \mathbf{0} \tag{2.22}$$

is given by

$$P(\hat{x}_s = x_s | \mathbf{u} = \mathbf{0}) = P(v_s = 0) \tag{2.23}$$

$$= P(f_s = 0)P(e_{s,[s]} = 0) + \frac{(1 - P(f_s = 0))(1 - P(e_{s,[s]} = 0))}{q - 1} \tag{2.24}$$

while that given $\mathbf{u} \neq \mathbf{0}$ is given by

$$P(\hat{x}_s = x_s | \mathbf{u} \neq \mathbf{0}) = \frac{q^{S-1} - P(v_s = 0)}{q^S - 1}. \tag{2.25}$$

Proof of (2.24) and (2.25) is provided in Appendix A.

Since the elements of $\mathbf{u}$ are independent, the probability of $\mathbf{u} = \mathbf{0}$ is given by

$$P(\mathbf{u} = \mathbf{0}) = \prod_{i=1}^{S} P(u_{[i]} = 0) \tag{2.26}$$

where

$$P(u_{[i]}=0)=P(f_{\mathrm{R},[i]}=0)P(e_{\mathrm{R},[i]}=0)+\frac{(1-P(f_{\mathrm{R},[i]}=0))(1-P(e_{\mathrm{R},[i]}=0))}{q-1}. \quad (2.27)$$

Therefore, the conditional probability of correct decoding of a symbol given that S linearly independent coded packets are received is given by

$$\begin{aligned} P_C &= P(\hat{x}_s = x_s) \\ &= P(\hat{x}_s = x_s|\mathbf{u}=\mathbf{0})P(\mathbf{u}=\mathbf{0}) + P(\hat{x}_s = x_s|\mathbf{u}\neq\mathbf{0})P(\mathbf{u}\neq\mathbf{0}). \end{aligned} \quad (2.28)$$

Then, the average probability of symbol error given R coded packets are available is $1-P_C P_I(R,S)$.

2.3.1 Trustworthy Source

If all source nodes are trustworthy (i.e., $f_s=0$) and the source-to-relay channels are error-free (i.e., $e_{s,r}=0$), then

$$P(\hat{x}_s = x_s|\mathbf{u}=\mathbf{0}) = P(v_s=0) \quad (2.29)$$

$$= 1 \quad (2.30)$$

for $s=1,\cdots,S$. Therefore, it follows from (2.25) and (2.28)

$$P_C = P(\mathbf{u}=\mathbf{0}) + \frac{q^{S-1}-1}{q^S-1}P(\mathbf{u}\neq\mathbf{0}) \quad (2.31)$$

$$= P(\mathbf{u}=\mathbf{0}) + \frac{q^{S-1}-1}{q^S-1}(1-P(\mathbf{u}=\mathbf{0})) \quad (2.32)$$

$$= \frac{q^{S-1}-1}{q^S-1} + \frac{q^S-q^{S-1}}{q^S-1}P(\mathbf{u}=\mathbf{0}) \quad (2.33)$$

$$= \frac{q^{S-1}-1+q^{S-1}(q-1)P(\mathbf{u}=\mathbf{0})}{q^S-1} \quad (2.34)$$

$$\approx \frac{1+(q-1)P(\mathbf{u}=\mathbf{0})}{q}. \quad (2.35)$$

2.3.2 Homogeneous Network

If the trustworthinesses of all nodes (source and relay) are the same, i.e.,

$$P(f_s \neq 0) = P(f_{\mathrm{R},r} \neq 0) := p_f \tag{2.36}$$

for all $s = 1, \cdots, S$ and $r = 1, \cdots, R$, and the error probabilities on all links in the network are the same, i.e.,

$$P(e_{s,r} \neq 0) = P(e_{\mathrm{R},r} \neq 0) := p_e \tag{2.37}$$

for all $s = 1, \cdots, S$ and $r = 1, \cdots, R$, then we obtain from (2.24) that

$$P(v_s = 0) = P(u_{[i]} = 0) \tag{2.38}$$

$$= 1 - p_f - p_e + \frac{q p_f p_e}{q-1}. \tag{2.39}$$

Therefore, it follows from (2.26)-(2.28) that

$$P_C = \alpha \cdot \alpha^S + \frac{q^{S-1} - \alpha}{q^S - 1}(1 - \alpha^S) \tag{2.40}$$

$$= \frac{\alpha^{S+1}(q^S - 1) + q^{S-1} - \alpha - q^{S-1}\alpha^S + \alpha^{S+1}}{q^S - 1} \tag{2.41}$$

$$= \frac{q^{S-1}(\alpha^{S+1} q + 1 - \alpha q^{-(S-1)} - \alpha^S)}{q^{S-1}(q - q^{-(S-1)})} \tag{2.42}$$

$$= \frac{1 - \alpha q^{-(S-1)} + (\alpha q - 1)\alpha^S}{q - q^{-(S-1)}} \tag{2.43}$$

$$\approx \frac{1 + (\alpha q - 1)\alpha^S}{q} \tag{2.44}$$

where

$$\alpha := 1 - p_f - p_e + \frac{q p_f p_e}{q-1} \tag{2.45}$$

represents the trustworthiness of received symbol. From the expression in (2.44) we can see that the probability of correct decoding decreases exponentially with the increasing number of combined packets and that the decreasing rate depends on the trustworthiness of received symbols. As the trustworthiness of symbol (α) decreases, the probability of correct decoding decreases faster. This suggests to limit the number of combined nodes when their trustworthiness are small or channels are erroneous.

2.3.3 Numerical Results

Fig. 2.2 shows the average probability of symbol error versus the field size in log scale for different values of node trustworthiness in homogeneous network. The number R of transmissions of coded packets by the relays is fixed (delay-limited scenario). The trustworthiness of source node is determined by $P(f_s = 0)$ and that of relay node is determined by $P(f_{\mathrm{R},r} = 0)$. We can see that there exists an optimal field size that minimizes the average probability of symbol error. This follows from the tradeoff between the probability of linear independence of coded packets $P_I(R, S)$ in (2.8) and the conditional probability of correct decoding P_C in (2.28). As the field size q gets larger, $P_I(R, S)$ increases while P_C decreases. Since the average probability of symbol error is $1 - P_C P_I(R, S)$, these two conflicting effects result in an optimal q that minimizes the average probability of symbol error. We also find that the optimal field size decreases as the trustworthiness of node decreases. This follows from the higher probability of f_s (false injection) being canceled by $e_{s,[s]}$ (channel error) for smaller q, suggesting the use of small field size when the trustworthiness of node is low.

Fig. 2.3 shows the average probability of symbol error versus log of the field size for different numbers of combined packets S for the case of γ_b=25dB and $P(f_s = 0)$=$P(f_{\mathrm{R},r} = 0)$=0.99, with the fixed rate of $\frac{S}{R} = 0.8$. We note that there is an optimal field size to minimize the average probability of symbol error for each pair of S and R. We also find that the optimal field size decreases with the increasing S.

Fig. 2.4 shows the average probability of symbol error versus the number of combined packets for different levels of node trustworthiness for the case of q=32 and γ_b=30dB, and $\frac{S}{R}$=0.8. We can see that the average probability of symbol error increases as the number of combined packets S increases.

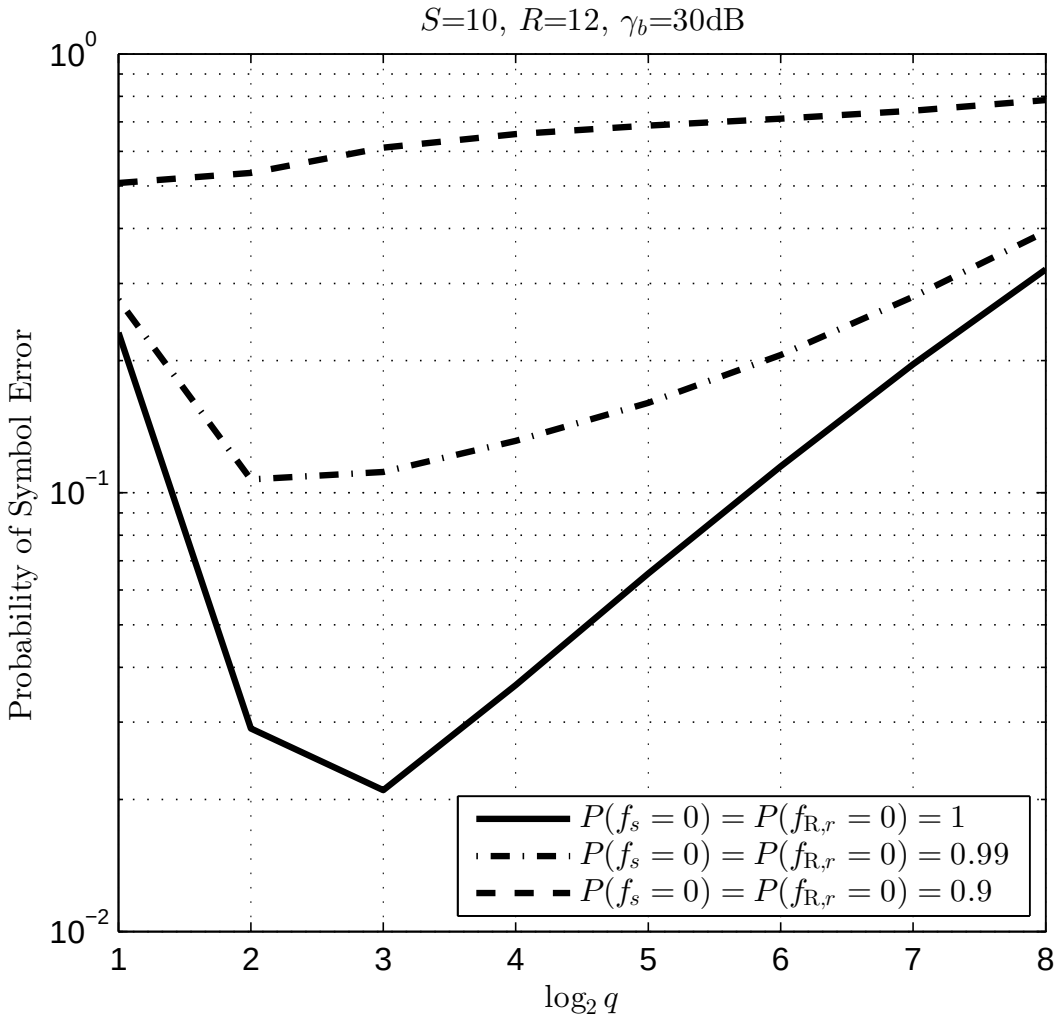

Figure 2.2 Probability of symbol error versus log of the field size for different levels of node trustworthiness; S=10, R=12, γ_b=30dB.

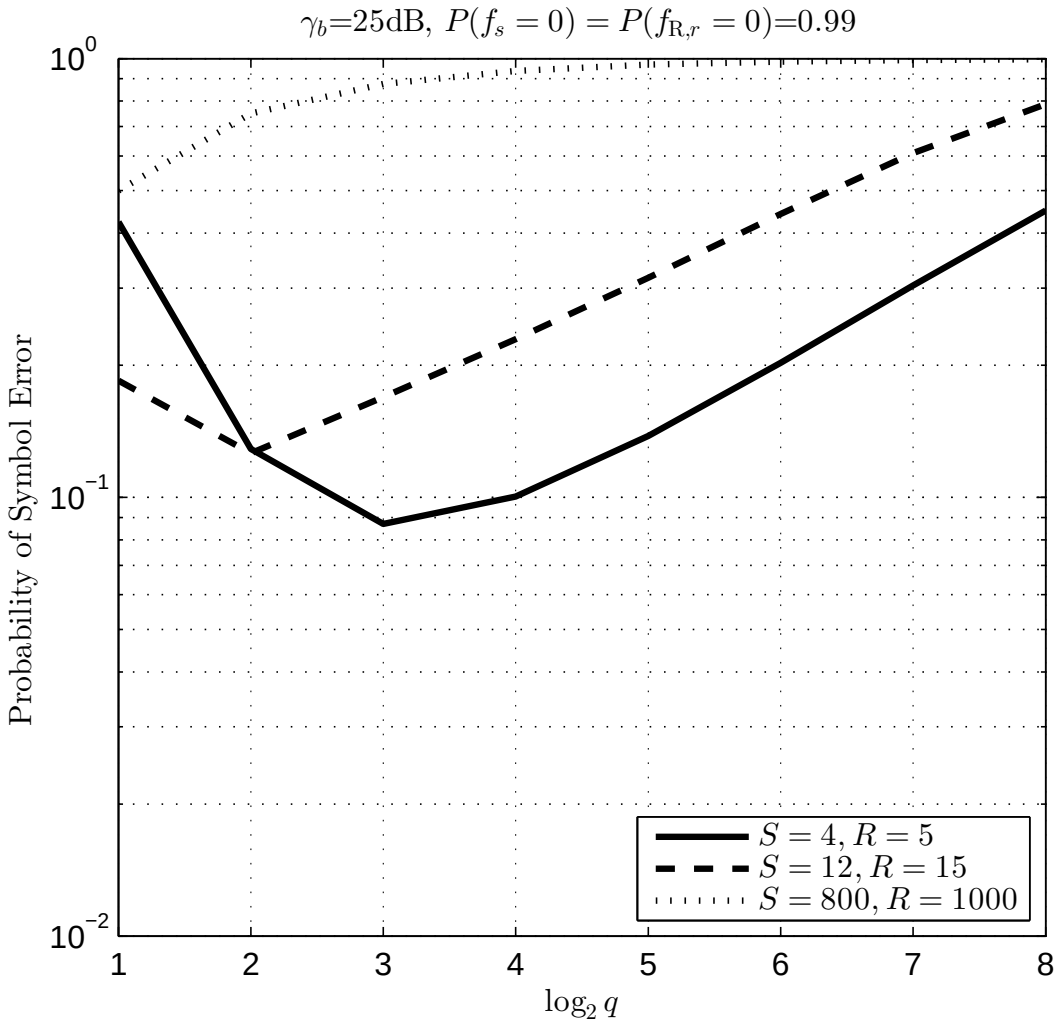

Figure 2.3 Probability of symbol error versus log of the field size for different numbers of combined packets; γ_b=25dB, $P(f_s = 0)$=$P(f_{\mathrm{R},r} = 0)$=0.99.

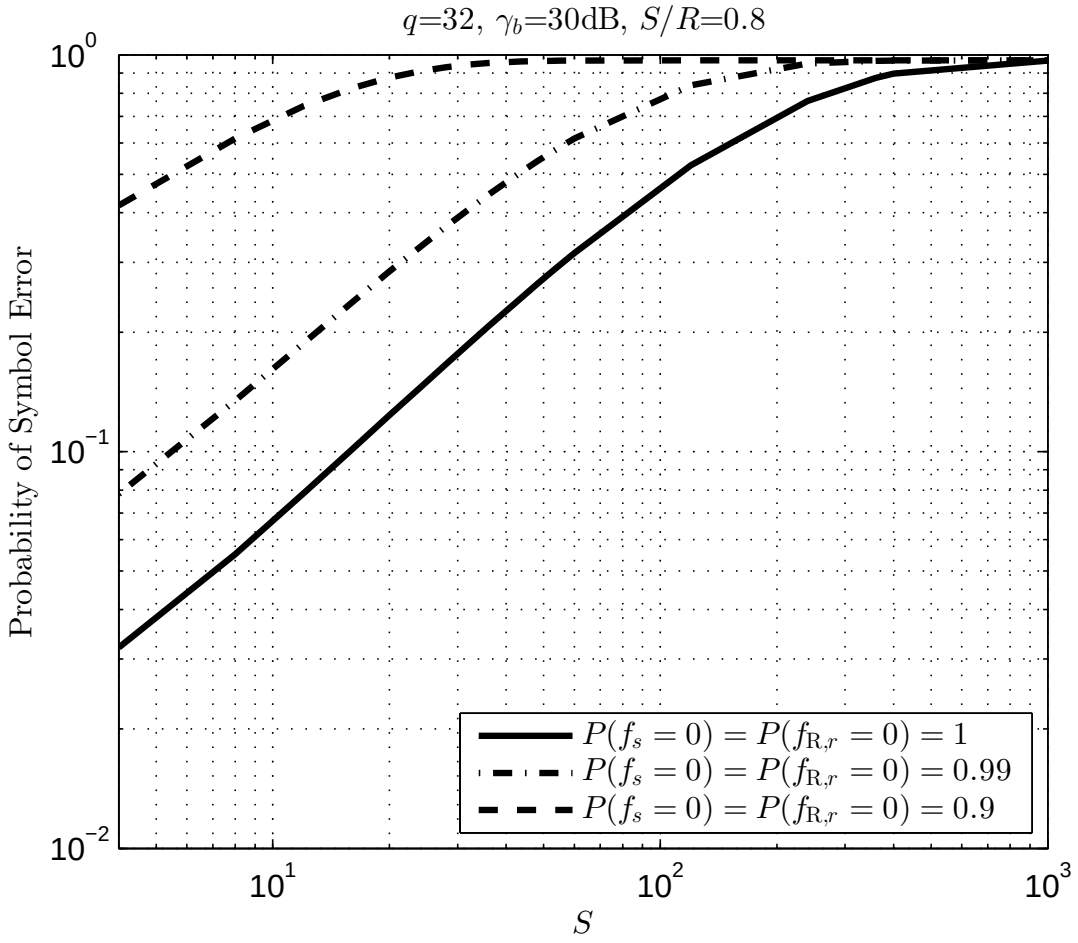

Figure 2.4 Probability of symbol error versus the number of combined packets for different levels of node trustworthiness; q=32, γ_b=30dB, $\frac{S}{R}$=0.8.

2.4 Throughput

The throughput W is defined as the average number of correctly decoded bits per symbol transmission (channel use). For a given $\mathbf{u}$, the decoding error events across the source nodes are independent. This follows from the assumption that f_s's ($e_{s,r}$'s) are independent. Hence, the conditional average number of correctly decoded symbols for a given $\mathbf{u}$ is $SP(\hat{x}_s = x_s|\mathbf{u})$, and averaging it over $\mathbf{u}$ yields SP_C.

With each random linear combination (a coded packet of length L bits) transmitted, the relay appends a packet header identifying the encoding coefficients $c_{r,s}, s = 1, \cdots, S$, which requires an additional $S \log_2 q$ bits of overhead with every $L \log_2 q$ bits transmitted. Therefore, the throughput is given by

$$W = \frac{L}{L+S} \frac{SP_C P_I(R,S) \log_2 q}{S+R} \tag{2.46}$$

$$\approx \frac{SP_C P_I(R,S) \log_2 q}{S+R} \tag{2.47}$$

where $S+R$ is to account for S channel uses by S sources and R channel uses by R relays, and $\log_2 q$ is to account for the number of bits per q-ary symbol. In this chapter we assume $L \gg S$ such the factor $\frac{L}{L+S}$ is close to one. In homogeneous networks, it follows from (2.8), (2.44), (2.46), that the throughput is given by

$$W = \frac{S[1 + (\alpha q - 1)\alpha^S] P_I(R,S) \log_2 q}{(S+R)q}. \tag{2.48}$$

2.4.1 Large-Scale Homogeneous Network

For large S and R while $\beta = \frac{S}{R}$ fixed, it can be shown that $P_I(R,S)$ can be made arbitrarily close to 1. Therefore, it follows from (2.43) and (2.46) that the throughput approaches to

$$\lim_{S,R\to\infty} W \approx \begin{cases} \frac{\beta \log_2 q}{1+\beta}, & \text{if } p_f + p_e = \frac{q p_f p_e}{q-1} \\ \frac{\beta \log_2 q}{(1+\beta)q}, & \text{if } p_f + p_e > \frac{q p_f p_e}{q-1} \end{cases} \tag{2.49}$$

where $p_f = P(f_s \neq 0) = P(f_{\mathrm{R},r} \neq 0)$ and $p_e = P(e_{s,r} \neq 0) = P(e_{\mathrm{R},r} \neq 0)$. Proof of (2.49) is provided in Appendix B. Since $\frac{\beta}{1+\beta}$ is an increasing function of β, the maximum throughput of $\frac{\log_2 q}{2}$ and $\frac{\log_2 q}{2q}$, depending on p_f and p_e, are achieved when $\beta = 1$, i.e., $S = R$. We can see from (2.49) that the convergence rate is faster with smaller α, i.e., larger p_f or p_e.

The condition $p_f + p_e = \frac{qp_f p_e}{q-1}$ corresponds to the case of $\alpha = 1$ and is satisfied if $p_f = p_e = 0$ (trustworthy network). In error-free, attack-free scenario, all symbols are received correctly, hence the throughput should increase on the order of $\log_2 q/2$.

The condition $p_f + p_e > \frac{qp_f p_e}{q-1}$ corresponds to the case of $\alpha < 1$, i.e., untrustworthy network. The throughput converges to $\frac{\log_2 q}{2q}$ and the optimal field size that maximizes the throughput is 2 or 4. The maximum achievable throughput with the optimal choice of field size is 0.25.

2.4.2 Homogeneous Network with Large q

For large q, $P_C \to \alpha^{S+1}$ where $\alpha \to (1-p_f)(1-p_e)$. Therefore, the throughput converges to

$$\lim_{q\to\infty} W \approx \frac{\beta}{1+\beta}[(1-p_f)(1-p_e)]^{S+1}\log_2 q \tag{2.50}$$

It should be noted that the symbol error probability p_e is an increasing function of q and approaches to 1 as q approaches to ∞. Therefore, the throughput will be close to 0 for large q.

2.4.3 Numerical Results

Fig. 2.5 shows the throughput (bits/symbol transmission) versus the field size in log scale for different values of node trustworthiness. We find that the optimal field size that maximizes the throughput decreases as the trustworthiness of node decreases. This is similar to the optimal field size in Fig. 2.2 that minimizes the average probability of symbol error when the number of transmissions is limited.

Fig. 2.6 shows the throughput versus the field size in log scale for the different values of S and R. We find that the optimal field size that maximizes the throughput decreases as S increases and that the asymptotic maximum throughput of $\frac{\beta \log_2 q}{(1+\beta)q} = \frac{2}{9}$ is achieved when $q = 2$ or 4. This matches well with the asymptotic result in (2.49). Figs. 2.5 and 2.6 indicate that the field size has to be decreased when the nodes are not trustworthy or the number of nodes that are combined is large such as in large-scale networks.

Fig. 2.7 shows the throughput versus the number of combined packets S for different levels of node trustworthiness. We can see that the throughput converges to $\frac{\beta \log_2 q}{(1+\beta)q} = \frac{1}{9}$ as S and R increase and that the convergence rate is faster when the nodes are less trustworthy.

2.5 Conclusion

In this chapter, we showed that there exists an optimal field size that minimizes the probability of symbol error or that maximizes the throughput. We found that the optimal field size that minimizes the probability of symbol error decreases as the trustworthiness of node decreases, suggesting the use of a smaller field size as the trustworthiness of node decreases. We also found that the probability of correct decoding of packet decreases exponentially with the increasing number of packets that are combined and that the decreasing rate is faster when the trustworthiness of node is smaller. The asymptotic result gives an insight to performance of random network coding in large-scale networks, such as wireless sensor networks where the network nodes are not trustworthy and the wireless channels are erroneous.

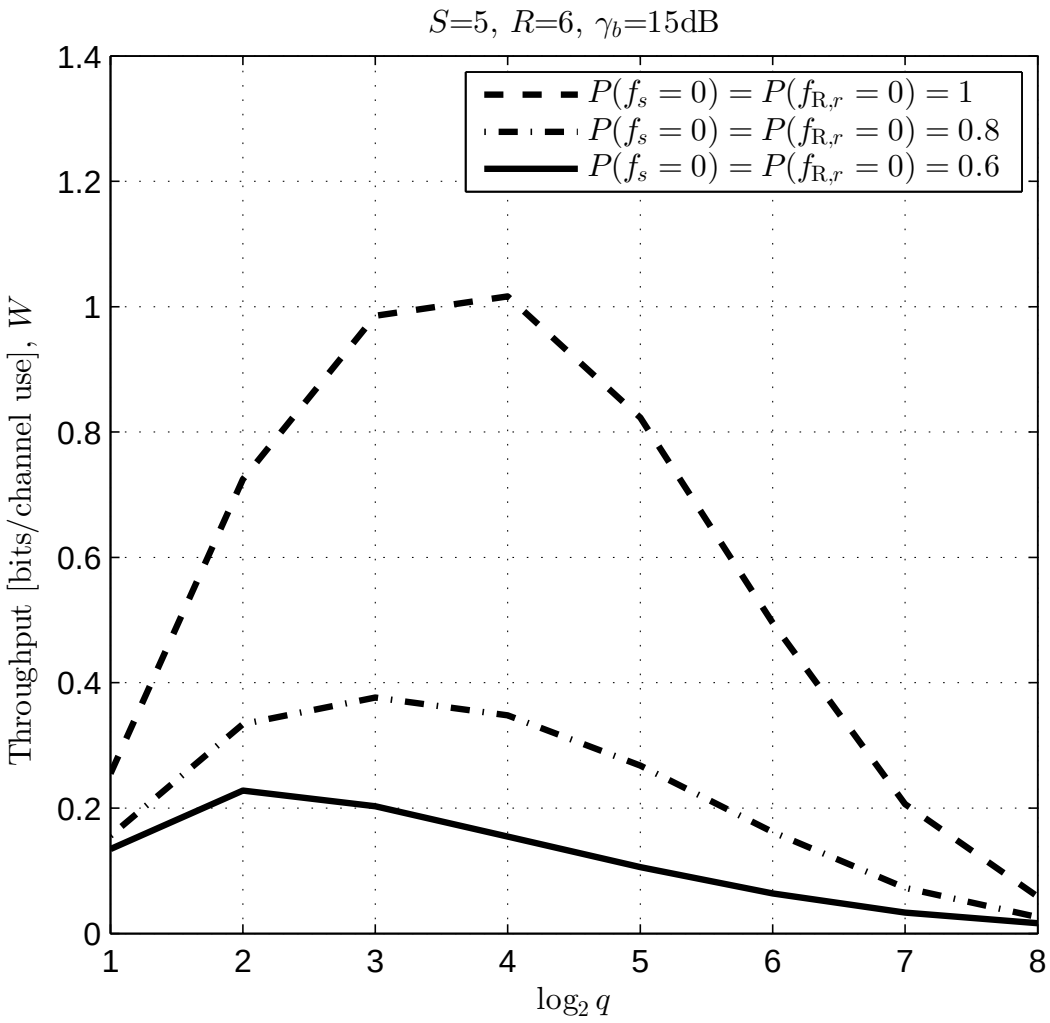

Figure 2.5 Throughput versus log of the field size for different levels of node trustworthiness; S=5, R=6, γ_b=15dB.

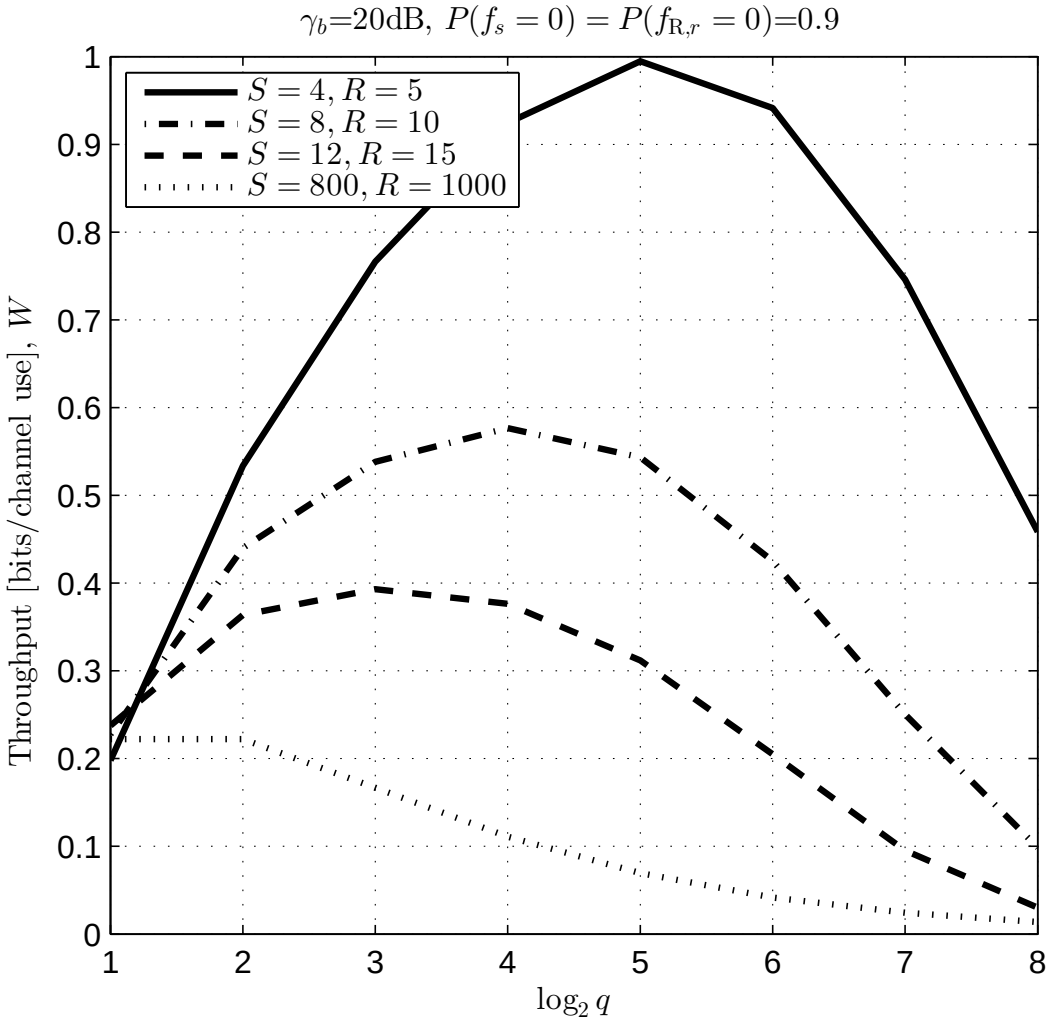

Figure 2.6 Throughput versus log of the field size for different numbers of combined packets; γ_b=20dB, $P(f_s = 0)$=$P(f_{\mathrm{R},r} = 0)$=0.9.

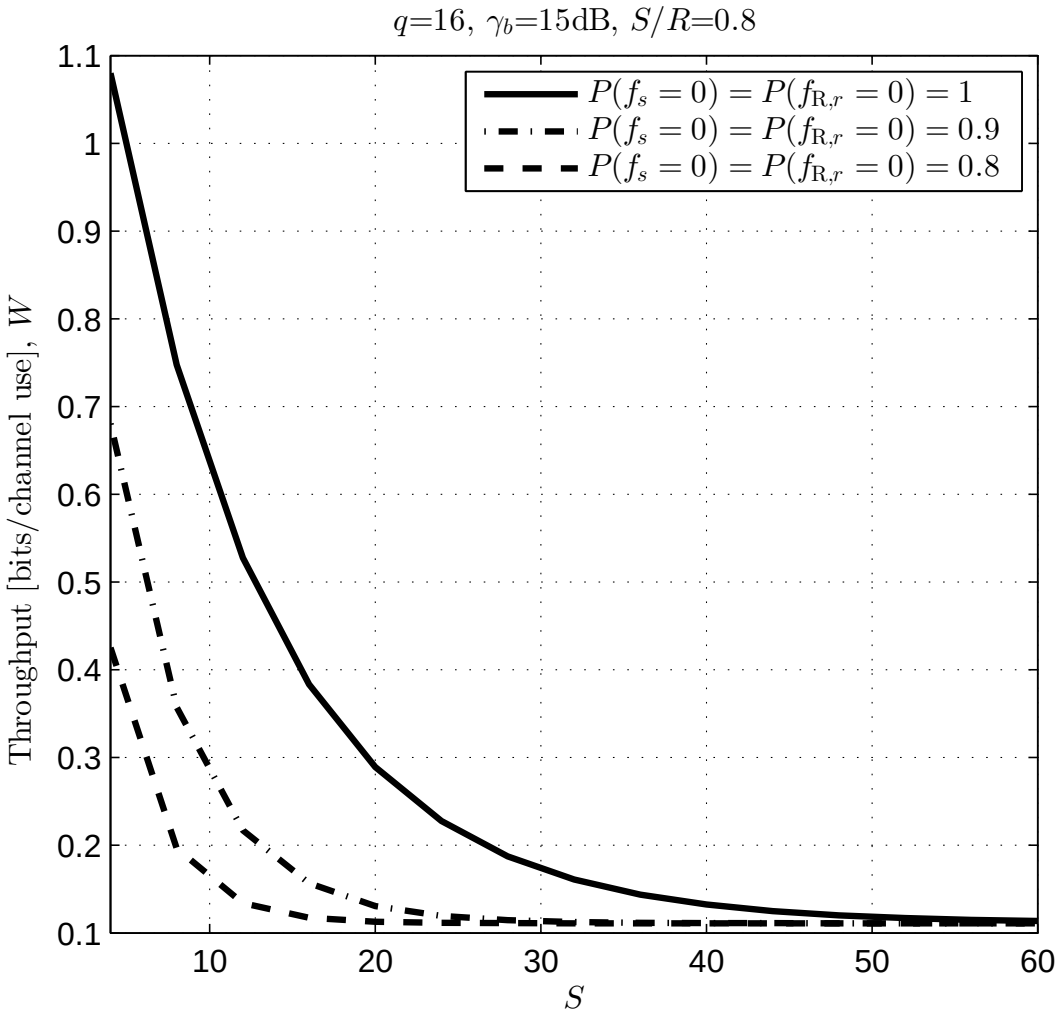

Figure 2.7 Throughput versus the number of combined packets for different levels of node trustworthiness; q=16, γ_b=15dB, $\frac{S}{R}$=0.8.

CHAPTER 3. PHYSICAL-LAYER APPROACH TO DETECT POLLUTION ATTACK IN WIRELESS NETWORK CODING

3.1 Introduction

In this chapter, we propose two physical layer approaches to filter out polluted packets, thereby improving the reliability of decoding, in the given two-hop wireless relay networks where some coded packets might be polluted at the relay. The integrity of each received coded packet is detected at the destination based on the Hamming distance between the coded packet and the corresponding linear combination of noisy message packets, which are directly overheard from the sources. Hence, the proposed schemes do not require any overhead unlike traditional cryptographic schemes which require additional bits attached in transmitted packets. The proposed scheme I uses the predetermined threshold to detect the presence of pollution attack within each coded packet upon its arrival, while the proposed scheme II compares all coded packets to select the coded packets having the highest integrity. Hence, the proposed scheme I spends the less time to reconstruct the original messages than the proposed scheme II, while the proposed scheme II provides the higher reliability of decoding. It is shown that both of the proposed schemes provide the significantly lower probability of decoding error than the traditional random selection scheme which does not exploit the physical-layer data and performs close to the cryptographic schemes that need bandwidth and computational overheads.

The remaining part of this chapter is structured as follows. In Section 3.2, the system model is presented. In Section 3.3, the mechanisms that the proposed schemes detect the polluted coded packets are described. In Section 3.4, we provide how the attacker generates the polluted coded packets. The analytical derivation and numerical results of the probability of decoding error, average delay, and average throughput for the proposed schemes and other compared schemes are provided in Section 3.5, Section 3.6, and Section 3.7, respectively. Finally, we conclude the chapter in Section 3.8.

3.2 System Model

We consider a two-hop wireless network in which S sources transmit independent message packets $\mathbf{x}_1, \cdots, \mathbf{x}_S$ to the destination via a relay as depicted in Fig. 3.1. Each message packet

$$\mathbf{x}_s = \{x_{s,1}, \cdots, x_{s,N}\}, \ s = 1, \cdots, S \tag{3.1}$$

is an (N, K) systematic Reed-Solomon codeword[13, 17] and each symbol $x_{s,n}$ is an element of finite field $GF(q)$.

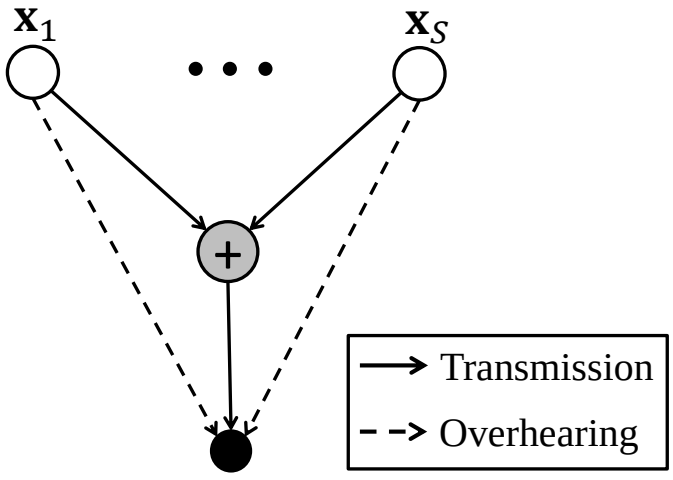

Figure 3.1 Two-hop network model with S sources, single relay, and single destination.

In phase 1, each source sends its message packet to the relay on its own orthogonal channel. We assume that all message packets are correctly decoded by the relay. Due

to the broadcast nature of wireless medium, the destination may overhear the message packets that arrive directly from the source nodes (dotted lines in Fig. 3.1) possibly with some errors. Let

$$\mathbf{r}_s = \mathbf{x}_s + \mathbf{e}_s, \ \ s = 1, \cdots, S \tag{3.2}$$

denote the received packet at the destination where

$$\mathbf{e}_s = \{e_{s,1}, \cdots, e_{s,N}\}, \ \ s = 1, \cdots, S \tag{3.3}$$

is the channel error vector between the sth source and the destination before channel decoding. The channel between each source and the destination is modeled by q-ary symmetric channel with crossover probability p, i.e.,

$$P(e_{s,n} = i) = \begin{cases} 1-p, & i = 0 \\ \frac{p}{q-1}, & i = 1, \cdots, q-1 \end{cases} \tag{3.4}$$

where $e_{s,n} \in GF(q)$ is the nth symbol of $\mathbf{e}_s$. After channel decoding, the correctness of the decoded word $\hat{\mathbf{x}}_s$ is examined by CRC (Cyclic Redundancy Check) code[15] which is attached to $\mathbf{x}_s$. We assume that CRC check always detects the presence of decoding error. Then, the destination stores

$$\mathbf{y}_s = \begin{cases} \mathbf{x}_s, & \text{if } \hat{\mathbf{x}}_s = \mathbf{x}_s \\ \mathbf{r}_s, & \text{if } \hat{\mathbf{x}}_s \neq \mathbf{x}_s \end{cases} \tag{3.5}$$

In phase 2, the relay combines the received packets to produce a coded packet

$$\mathbf{t}_r = \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s, \ \ r = 1, \cdots, R \tag{3.6}$$

where the coefficients $\{c_{r,s}\}$ are randomly selected from nonzero elements of finite field $GF(q)$. We assume that the field size q is sufficiently large, so that coded packets are linearly independent with probability close to one[2, 3, 12]. We consider the scenario where the malicious adversary may access the relay and modify $\mathbf{t}_r$ into

$$\mathbf{p}_r = \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r, \ \ r = 1, \cdots, R \tag{3.7}$$

where $\mathbf{f}_r$ is the falsely injected packet that captures the modification on the rth coded packet. The modified packet $\mathbf{p}_r$ is then sent to the destination. We assume that the relay-destination channel is error-free.

Let Λ be a random variable denoting the number of incorrectly channel-decoded message packets at the destination during phase 1. i.e.,

$$\Lambda = \Big|\{s \mid \mathbf{y}_s \neq \mathbf{x}_s, s = 1, \cdots, S\}\Big| \tag{3.8}$$

where $|\cdot|$ denotes the cardinality (i.e., the number of elements) of a set.

If $\Lambda = 0$, the destination does not need coded packets which are transmitted from the relay during phase 2, because it already obtained all correct message packets $\mathbf{x}_1, \cdots, \mathbf{x}_S$ during phase 1. If $\Lambda > 0$, the destination needs S linearly independent packets consisting of all $S - \Lambda$ correctly channel-decoded message packets during phase 1 and Λ coded packets which are chosen from all R coded packets received during phase 2, in order to find the reconstructed message packets $\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_S$ given by

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \\ \vdots \\ \\ \hat{\mathbf{x}}_S \end{bmatrix} = \left[\begin{array}{c} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(S-\Lambda)} \\ \hline \mathbf{c}_{[1]} \\ \vdots \\ \mathbf{c}_{[\Lambda]} \end{array}\right]^{-1} \left[\begin{array}{c} \mathbf{y}_{(1)} \\ \vdots \\ \mathbf{y}_{(S-\Lambda)} \\ \hline \mathbf{p}_{[1]} \\ \vdots \\ \mathbf{p}_{[\Lambda]} \end{array}\right] \tag{3.9}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \\ \vdots \\ \\ \mathbf{x}_S \end{bmatrix} + \left[\begin{array}{c} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(S-\Lambda)} \\ \hline \mathbf{c}_{[1]} \\ \vdots \\ \mathbf{c}_{[\Lambda]} \end{array}\right]^{-1} \left[\begin{array}{c} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \hline \mathbf{f}_{[1]} \\ \vdots \\ \mathbf{f}_{[\Lambda]} \end{array}\right] \tag{3.10}$$

where $(1), \cdots, (S-\Lambda) \in \{1, \cdots, S\}$ are the indices of the successfully channel-decoded message packets during phase 1, while $[1], \cdots, [\Lambda] \in \{1, \cdots, R\}$ are the indices of selected coded packets from $\mathbf{p}_1, \cdots, \mathbf{p}_R$. Note that how to select Λ coded packets from R coded packets depends on the selection rule of the utilized scheme, which is described in the section 3.3. A length-S vector $\mathbf{c}_{[r]} = \{c_{[r],1}, \cdots, c_{[r],S}\}$ is the coefficient vector associated with the coded packet $\mathbf{p}_{[r]}$, while $\mathbf{b}_{(j)}$ is a unit vector having all zeros but unique 1 at the (j)th element. For example, in the case of $S = 3$ and $R = 5$, if we assume that $\mathbf{y}_1$ and $\mathbf{y}_3$ are correctly channel-decoded and the destination selects $\mathbf{p}_2$ from coded packets $\mathbf{p}_1, \cdots, \mathbf{p}_5$, the reconstructed message packets $\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_S$ are given by

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \\ \hat{\mathbf{x}}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{(1)} \\ \mathbf{b}_{(2)} \\ \mathbf{c}_{[1]} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}_{(1)} \\ \mathbf{y}_{(2)} \\ \mathbf{p}_{[1]} \end{bmatrix} \tag{3.11}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ c_{2,1} & c_{2,2} & c_{2,3} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_3 \\ \mathbf{p}_2 \end{bmatrix} \tag{3.12}$$

$$= \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ c_{2,1} & c_{2,2} & c_{2,3} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{f}_2 \end{bmatrix} \tag{3.13}$$

where $\Lambda = 1$, $\mathbf{y}_{(1)} = \mathbf{y}_1 = \mathbf{x}_1$, $\mathbf{y}_{(2)} = \mathbf{y}_3 = \mathbf{x}_3$, $\mathbf{p}_{[1]} = \mathbf{p}_2 = \sum_{s=1}^{3} c_{2,s}\mathbf{x}_s + \mathbf{f}_2$, $\mathbf{b}_{(1)} = \mathbf{b}_1 = \{1, 0, 0\}$, $\mathbf{b}_{(2)} = \mathbf{b}_3 = \{0, 0, 1\}$, and $\mathbf{c}_{[1]} = \mathbf{c}_2 = \{c_{2,1}, c_{2,2}, c_{2,3}\}$.

From above description, we define that decoding is *successful*,

1. when $\Lambda = 0$.

2. when $\Lambda > 0$, Λ coded packets are selected, and all S reconstructed message packets

are equal to all S original message packets that sources sent. i.e.,

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \vdots \\ \hat{\mathbf{x}}_S \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_S \end{bmatrix} \tag{3.14}$$

In Appendix C, we prove that (3.14) is equivalent to $\mathbf{f}_{[1]} = \mathbf{0}, \cdots, \mathbf{f}_{[\Lambda]} = \mathbf{0}$, which denotes that all Λ coded packets which are determined to be unpolluted are actually unpolluted. Therefore, decoding is *failed*,

1. when $\Lambda > 0$ and less than Λ coded packets are selected.

2. when $\Lambda > 0$, Λ coded packets are selected, and at least one of Λ selected coded packets is polluted.

3.3 Proposed Detection Schemes

In this section, we describe two proposed schemes to detect the polluted packets.

3.3.1 Scheme I

Upon receiving a coded packet during phase 2, the destination checks if it is linearly independent with the successfully received message packets $\mathbf{y}_{(1)}, \cdots, \mathbf{y}_{(S-\Lambda)}$ during phase 1, and the previously selected coded packets. If not, it is discarded. If so, the destination checks if the coded packet is polluted or not, by the detection scheme described below. If it is determined to be unpolluted, it is selected. As soon as the destination selects Λ coded packets, it stops receiving the coded packet and recovers message packets based on $S - \Lambda$ message packets $\mathbf{y}_{(1)}, \cdots, \mathbf{y}_{(S-\Lambda)}$ and Λ selected coded packets $\mathbf{p}_{[1]}, \cdots, \mathbf{p}_{[\Lambda]}$. If the number of selected coded packets is less than Λ by the end of phase 2, then the decoding fails.

Given $\mathbf{y}_1, \cdots, \mathbf{y}_S$ and $\mathbf{p}_r$, a sufficient statistic for $\mathbf{f}_r$ is

$$\mathbf{z}_r = \mathbf{p}_r - \sum_{s=1}^{S} c_{r,s}\mathbf{y}_s \tag{3.15}$$

$$= \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r - \sum_{s=1}^{S} c_{r,s}\left(\mathbf{x}_s + \dot{\mathbf{e}}_s\right) \tag{3.16}$$

$$= \underbrace{\mathbf{f}_r - \sum_{s=1}^{S} c_{r,s}\dot{\mathbf{e}}_s}_{\mathbf{g}_r} \tag{3.17}$$

where

$$\dot{\mathbf{e}}_s = \begin{cases} \mathbf{0}, & \text{if } \hat{\mathbf{x}}_s = \mathbf{x}_s \\ \mathbf{e}_s, & \text{if } \hat{\mathbf{x}}_s \neq \mathbf{x}_s \end{cases} \tag{3.18}$$

That is, $\mathbf{z}_r$ contains all the information necessary to make a decision on $\mathbf{f}_r$ given $\mathbf{y}_1, \cdots, \mathbf{y}_S$ and $\mathbf{p}_r$. This follows from the equality $I(\mathbf{f}_r; \mathbf{y}_1, \cdots, \mathbf{y}_S) = I(\mathbf{f}_r; \mathbf{z}_r)$ which is proved in Appendix D.

The Hamming weight of $\mathbf{z}_r$ is given by

$$W_H(\mathbf{z}_r) = W_H(\mathbf{f}_r + \mathbf{g}_r) \tag{3.19}$$

where

$$\mathbf{g}_r = -\sum_{s=1}^{S} c_{r,s}\dot{\mathbf{e}}_s. \tag{3.20}$$

Since it is likely that $W_H(\mathbf{z}_r)$ with $\mathbf{f}_r = \mathbf{0}$ is smaller than that with $\mathbf{f}_r \neq \mathbf{0}$, a coded packet with smaller $W_H(\mathbf{z}_r)$ is more likely to be unpolluted. Based on this observation, the coded packet is regarded as unpolluted if $W_H(\mathbf{z}_r)$ is less than or equal to the predetermined threshold η and polluted otherwise. i.e.,

$$W_H(\mathbf{z}_r) \underset{\hat{H}_1}{\overset{\hat{H}_0}{\lessgtr}} \eta \tag{3.21}$$

where $\hat{H}_0$ and $\hat{H}_1$ denotes that the coded packet is detected as unpolluted and polluted, respectively. By exploiting p and $\Lambda = \lambda$ that the destination observes at the end of phase

1, the threshold η is given by

$$\eta = \frac{E[W_H(\mathbf{g}_r)]_{p,\lambda} + E[W_H(\mathbf{f}_r + \mathbf{g}_r)]_{p,\lambda}}{2} \tag{3.22}$$

where

$$E[W_H(\mathbf{g}_r)]_{p,\lambda} = \sum_{g=0}^{N} g \cdot \underbrace{P(G = g|\Lambda = \lambda)}_{(3.47)} \tag{3.23}$$

denotes the expected value of $W_H(\mathbf{z}_r)$ of an unpolluted coded packet given p and λ, and

$$E[W_H(\mathbf{f}_r + \mathbf{g}_r)]_{p,\lambda} = \sum_{j=0}^{N}\sum_{g=0}^{N} j \cdot \underbrace{P(W_H(\mathbf{f}_r + \mathbf{g}_r) = j|\mathbf{f}_r \neq \mathbf{0}, G = g)}_{(3.59)} \cdot \underbrace{P(G = g|\Lambda = \lambda)}_{(3.47)} \tag{3.24}$$

denotes the expected value of $W_H(\mathbf{z}_r)$ of a polluted coded packet given p and λ with the assumption that $W_H(\mathbf{f}_r)$ is equal to d_{min} and those d_{min} nonzero symbols in $\mathbf{f}_r$ are randomly located. $P(G = g|\Lambda = \lambda)$ and $P(W_H(\mathbf{f}_r + \mathbf{g}_r) = j|\mathbf{f}_r \neq \mathbf{0}, G = g)$ are derived in Section 3.5.

3.3.2 Scheme II

Basic idea of the proposed scheme II is to select linearly independent coded packets having the smallest $W_H(\mathbf{z}_r)$'s from all R coded packets. After all R coded packets are received, the destination finds $W_H(\mathbf{z}_1), \cdots, W_H(\mathbf{z}_R)$ by (3.15). Then, it checks if the coded packet having the smallest $W_H(\mathbf{z}_r)$ is linearly independent with the $S-\Lambda$ message packets $\mathbf{y}_{(1)}, \cdots, \mathbf{y}_{(S-\Lambda)}$ which are successfully received during phase 1. If so, the coded packet is selected. Otherwise, it is discarded. Suppose that it is selected. Then, the destination checks if the coded packet having the second smallest $W_H(\mathbf{z}_r)$ is linearly independent with the $S - \Lambda$ message packets and the one previously selected coded packet. If so, it is selected, and so on. If multiple coded packets have the same $W_H(\mathbf{z}_r)$'s, they are checked in random order. If Λ coded packets are selected until all R coded packets are checked, the destination tries to reconstruct all message packets with them and $S - \Lambda$ successfully received message packets. Otherwise, decoding fails.

3.4 False Injection Vector

In this section, we describe how the attacker generates the false injection vector $\mathbf{f}_r$ which is added to the true coded packet $\mathbf{t}_r$ as shown in (3.7). We assume that the attacker has whole knowledge about the proposed schemes described in Section 3.3. The attacker wants to maximize the probability that polluted coded packets are not detected (i.e., misdetection) thus selected for decoding (i.e., included among $\mathbf{p}_{[1]}, \cdots, \mathbf{p}_{[\Lambda]}$ in (3.9)), thereby maximizing the probability of decoding error. There are three conditions that $\mathbf{f}_r$ needs to meet, in order to maximize the probability of misdetection.

1. The true coded packet $\mathbf{t}_r$ in (3.7) is a codeword because it is a linear combination of codewords $\mathbf{x}_1, \cdots, \mathbf{x}_S$. Since relay-to-destination channel is error-free as assumed in section 3.2 and we assume that the destination knows it, the received (polluted) coded packet $\mathbf{p}_r$ is suspicious to the destination if it is not a codeword. Hence, the attacker makes $\mathbf{p}_r$ a different codeword from $\mathbf{t}_r$ by adding another nonzero codeword $\mathbf{f}_r$ to $\mathbf{t}_r$. Therefore, $\mathbf{f}_r$ must be a nonzero codeword.

2. We assume that the attacker cannot see the value of $\mathbf{g}_r$ in (3.17), although it knows detection principles of the proposed schemes. Therefore, the attacker cannot find a certain $\mathbf{f}_r$ that cancels out $\mathbf{g}_r$ and results in $W_H(\mathbf{f}_r + \mathbf{g}_r) < W_H(\mathbf{g}_r)$. Hence, the attacker makes $W_H(\mathbf{f}_r)$ as small as possible, because the smaller $W_H(\mathbf{f}_r)$ causes the higher probability of $W_H(\mathbf{f}_r + \mathbf{g}_r) = W_H(\mathbf{g}_r)$, where the destination cannot distinguish the polluted coded packets from unpolluted coded packets. Hence, the attacker selects a nonzero codeword $\mathbf{f}_r$ with the smallest Hamming weight, $d_{min} = N - K + 1$.

3. We also assume that the attacker cannot know the positions of nonzero symbols in $\mathbf{g}_r$. Thus, the attacker cannot find the positions of nonzero symbols in $\mathbf{f}_r$ causing the smallest $W_H(\mathbf{f}_r + \mathbf{g}_r)$. Hence, nonzero symbols in $\mathbf{f}_r$ are randomly located.

From above conditions, $\mathbf{f}_r$ should be a codeword having d_{min} nonzero symbols which are randomly located among N symbols. By [23] (Theorem 8-5 at page 189), for any combinations of d_{min} out of N coordinates in a MDS(Maximum Distance Separable) codeword, there exist $q-1$ codewords each of which has d_{min} nonzero symbols only at those coordinates. Therefore, it is theoretically feasible to find such a codeword. In order to generate it, the attacker repeats following steps for every polluted coded packet.

1. Randomly determine d_{min} coordinates among $1, \cdots, N$.
2. Select one of d_{min}-weight codewords having nonzero symbols at those coordinates.

3.5 Probability of Decoding Error

In this section, we derive the probability of decoding error for the two proposed schemes, the random selection scheme, and the cryptographic scheme, averaged over Λ which denotes the number of unsuccessfully channel-decoded message packets in phase 1. According to the definition of decoding success given in Section 3.2, the conditional probability of decoding success given $\Lambda = \lambda$ is given by

$$P(\text{decoding success}|\Lambda=\lambda) = \begin{cases} 1, & \text{if } \lambda = 0 \\ P(\underbrace{\text{Select } \Lambda \text{ coded packets}}_{\hat{\mathcal{F}}_\Lambda}, \underbrace{\text{Those are all unpolluted}}_{\mathcal{F}_\Lambda} | \Lambda=\lambda), & \text{if } \lambda > 0 \end{cases} \tag{3.25}$$

For brief representation, let $\hat{\mathcal{F}}_\Lambda$ and $\mathcal{F}_\Lambda$ be an event that Λ coded packets are selected for decoding and an event that all those coded packets are actually unpolluted, respectively.

Then, the probability of decoding error is given by

$$P_E = 1 - \sum_{\lambda=0}^{S} P(\text{decoding success}|\Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.26}$$

$$= 1 - P(\Lambda = 0) - \sum_{\lambda=1}^{S} P(\hat{\mathcal{F}}_\Lambda, \mathcal{F}_\Lambda|\Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.27}$$

$$= 1 - P(\Lambda = 0) - \sum_{\lambda=1}^{S} P(\mathcal{F}_\Lambda|\hat{\mathcal{F}}_\Lambda, \Lambda = \lambda) \cdot P(\hat{\mathcal{F}}_\Lambda|\Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.28}$$

where

$$P(\Lambda = \lambda) = \binom{S}{\lambda} P_e^\lambda (1 - P_e)^{S-\lambda} \tag{3.29}$$

and

$$P_e = \sum_{j=t+1}^{N} \binom{N}{j} p^j (1-p)^{N-j} \tag{3.30}$$

is the probability that an overheard message packet is decoded incorrectly, where $t = \left\lfloor \frac{d_{min}-1}{2} \right\rfloor = \left\lfloor \frac{N-K}{2} \right\rfloor$ is the error correction capability of Reed-Solomon code [17].

3.5.1 Random Selection

Random selection scheme selects the first Λ linearly independent coded packets. Since the field size q is sufficiently large, each coded packet is innovative. Therefore, the conditional probability that Λ coded packets are selected given $\Lambda = \lambda$ is given by

$$P(\hat{\mathcal{F}}_\Lambda|\Lambda = \lambda) = \begin{cases} 1, & 0 < \lambda \leq R \\ 0, & \text{otherwise.} \end{cases} \tag{3.31}$$

The conditional probability that all Λ selected coded packets are actually unpolluted given $\Lambda = \lambda$ is given by

$$P(\mathcal{F}_\Lambda|\hat{\mathcal{F}}_\Lambda, \Lambda = \lambda) = (1 - p_f)^\lambda. \tag{3.32}$$

It follows from (3.28), (3.31), and (3.32) that the average probability of decoding error is given by

$$P_E = 1 - P(\Lambda = 0) - \sum_{\lambda=1}^{\min(S,R)} (1-p_f)^{\lambda} \cdot P(\Lambda = \lambda) \tag{3.33}$$

$$= 1 - \sum_{\lambda=0}^{\min(S,R)} (1-p_f)^{\lambda} \cdot P(\Lambda = \lambda). \tag{3.34}$$

If $S \leq R$, then

$$P_E = 1 - \sum_{\lambda=0}^{S} \binom{S}{\lambda} \{(1-p_f)P_e\}^{\lambda} (1-P_e)^{S-\lambda} \tag{3.35}$$

$$\stackrel{(a)}{=} 1 - (1 - p_f \cdot P_e)^S \tag{3.36}$$

where (a) is because $\sum_{j=0}^{c} \binom{c}{j} \alpha^j \beta^{c-j} = (\alpha+\beta)^c$ by [18].

3.5.2 Cryptographic Scheme

We assume that the cryptographic scheme perfectly detects the polluted coded packets. That is,

$$P(\mathcal{F}_\Lambda | \hat{\mathcal{F}}_\Lambda, \Lambda = \lambda) = 1. \tag{3.37}$$

If we let R_f denote the number of polluted coded packets, then the conditional probability that Λ coded packets are selected for decoding given $\Lambda = \lambda$ is given by

$$P(\hat{\mathcal{F}}_\Lambda | \Lambda = \lambda) = \sum_{r_f=0}^{R} P(\hat{\mathcal{F}}_\Lambda | R_f = r_f, \Lambda = \lambda) \cdot P(R_f = r_f | \Lambda = \lambda) \tag{3.38}$$

where

$$P(\hat{\mathcal{F}}_\Lambda | R_f = r_f, \Lambda = \lambda) = \begin{cases} 1, & 0 < \lambda \leq R - r_f \\ 0, & \text{otherwise} \end{cases} \tag{3.39}$$

is the conditional probability that Λ coded packets are selected given $R_f = r_f$ and $\Lambda = \lambda$, and

$$P(R_f = r_f | \Lambda = \lambda) = P(R_f = r_f) \tag{3.40}$$

$$= \binom{R}{r_f} p_f^{r_f} (1-p_f)^{R-r_f}. \tag{3.41}$$

Therefore, it follows from (3.38) that the probability of decoding error is given by

$$P_E = 1 - \sum_{r_f=0}^{R} \sum_{\lambda=0}^{\min(S,R-r_f)} P(R_f = r_f|\Lambda = \lambda) \cdot P(\Lambda = \lambda). \tag{3.42}$$

3.5.3 Scheme I

The joint conditional probability $P(\hat{\mathcal{F}}_\Lambda, \mathcal{F}_\Lambda|\Lambda = \lambda)$ can be expressed as

$$\begin{aligned}
&P(\hat{\mathcal{F}}_\Lambda, \mathcal{F}_\Lambda|\Lambda = \lambda) \\
&= \sum_{g=0}^{N} P(\mathcal{F}_\Lambda, \hat{\mathcal{F}}_\Lambda, G = g|\Lambda = \lambda) && (3.43)\\
&= \sum_{g=0}^{N} \underbrace{P(\mathcal{F}_\Lambda|\hat{\mathcal{F}}_\Lambda, G = g, \Lambda = \lambda)}_{(3.60)} \cdot \underbrace{P(\hat{\mathcal{F}}_\Lambda|G = g, \Lambda = \lambda)}_{(3.51)} \cdot \underbrace{P(G = g|\Lambda = \lambda)}_{(3.47)} && (3.44)
\end{aligned}$$

where G denotes the number of nonzero column vectors in

$$\dot{\mathbf{E}} = \begin{bmatrix} \dot{\mathbf{e}}_1 \\ \vdots \\ \dot{\mathbf{e}}_S \end{bmatrix} = \begin{bmatrix} \dot{e}_{1,1} & \cdots & \dot{e}_{1,N} \\ \vdots & \ddots & \vdots \\ \dot{e}_{S,1} & \cdots & \dot{e}_{S,N} \end{bmatrix}. \tag{3.45}$$

For sufficiently large q, G is close to $W_H(\mathbf{g}_r)$ for $r = 1, \cdots, R$ because sum of two nonzero symbols is nonzero with high probability. If we let G_λ be G given $\Lambda = \lambda$, the conditional probability of $G = g$ given $\Lambda = \lambda$ is recursively given by

$$\begin{aligned}
&P(G = g|\Lambda = \lambda) \\
&= P(G_\lambda = g) && (3.46)\\
&= \begin{cases} P(G_1 = g), & \text{if } \lambda = 1 \\ \sum_{g_{\lambda-1}=0}^{N} \sum_{g_1=0}^{N} P(G_\lambda = g|G_{\lambda-1} = g_{\lambda-1}, G_1 = g_1) \cdot P(G_{\lambda-1} = g_{\lambda-1}) \cdot P(G_1 = g_1), & \text{if } \lambda > 1 \end{cases} && (3.47)
\end{aligned}$$

where

$$P(G_1 = g_1) = P(W_H(\dot{\mathbf{e}}_s) = g_1 | \dot{\mathbf{e}}_s \neq \mathbf{0}) \tag{3.48}$$

$$= \begin{cases} \dfrac{\binom{N}{g_1} p^{g_1}(1-p)^{N-g_1}}{\sum_{j=t+1}^{N} \binom{N}{j} p^j (1-p)^{N-j}}, & g_1 > t \\ 0, & g_1 \leq t \end{cases} \tag{3.49}$$

The conditional probability of $G_\lambda = g$ given $G_{\lambda-1} = g_{\lambda-1}$ and $G_1 = g_1$ is given by

$$P(G_\lambda = g | G_{\lambda-1} = g_{\lambda-1}, G_1 = g_1)$$

$$= \begin{cases} \dfrac{\binom{g_{\lambda-1}}{g_{\lambda-1}+g_1-g}\binom{N-g_{\lambda-1}}{g-g_{\lambda-1}}}{\binom{N}{g_1}}, & \text{if } \max(g_{\lambda-1}, g_1) \leq g \leq \min(N, g_{\lambda-1}+g_1) \\ 0, & \text{otherwise} \end{cases} \tag{3.50}$$

Let $\dot{\mathbf{e}}_{<1>}, \dot{\mathbf{e}}_{<2>}, \cdots, \dot{\mathbf{e}}_{<\lambda>}$ denote the nonzero error vectors given $\Lambda = \lambda$ in $\dot{\mathbf{E}}$. In order to have $G_\lambda = g$, $g - g_{\lambda-1}$ nonzero column vectors of $\begin{bmatrix} \dot{\mathbf{e}}_{<2>} \\ \vdots \\ \dot{\mathbf{e}}_{<\lambda>} \end{bmatrix}$ should not overlap with the nonzero symbols of $\dot{\mathbf{e}}_{<1>}$, while $g_{\lambda-1} - (g - g_1)$ nonzero column vectors of $\begin{bmatrix} \dot{\mathbf{e}}_{<2>} \\ \vdots \\ \dot{\mathbf{e}}_{<\lambda>} \end{bmatrix}$ should overlap with the nonzero symbols of $\dot{\mathbf{e}}_{<1>}$. The probability of such that is given by (3.50). Fig.3.2 shows an example of $\dot{\mathbf{E}}$ in the case of $g_1 = 3$ and $g_{\lambda-1} = 3$.

The conditional probability that Λ coded packets are selected given $G = g$ and $\Lambda = \lambda$ is given by

$$P(\hat{\mathcal{F}}_\Lambda | G = g, \Lambda = \lambda)$$

$$= \begin{cases} \displaystyle\sum_{c=0}^{R-\lambda} \binom{c+\lambda-1}{c} P(\hat{H}_1 | G = g)^c P(\hat{H}_0 | G = g)^\lambda, & \text{if } 0 < \lambda \leq R \\ 0, & \text{otherwise} \end{cases} \tag{3.51}$$

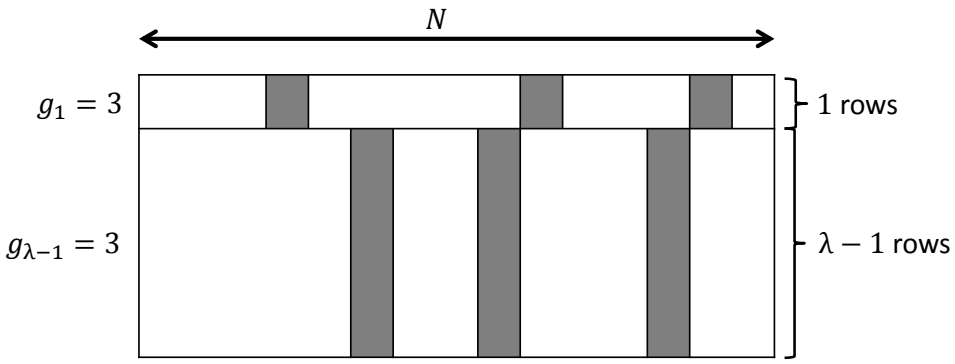

Figure 3.2 An example of $\dot{\mathbf{E}}$

where c is the number of coded packets that are decided as polluted until the destination selects the Λth coded packet, and

$$P(\hat{H}_1|G=g) = P_{FA}(g)(1-p_f) + (1-P_{MD}(g))p_f \tag{3.52}$$

is the conditional probability that each coded packet is decided as polluted given $G=g$. In (3.52),

$$P_{FA}(g) = P(W_H(\mathbf{f}_r+\mathbf{g}_r) > \eta|\mathbf{f}_r=\mathbf{0}, G=g) \tag{3.53}$$

$$= P(W_H(\mathbf{g}_r) > \eta|G=g) \tag{3.54}$$

is the conditional probability of false alarm given $G=g$. Since $W_H(\mathbf{g}_r)=g$, we obtain

$$P_{FA}(g) = \begin{cases} 0, & \text{if } g \le \eta \\ 1, & \text{if } g > \eta \end{cases} \tag{3.55}$$

Similarly, the conditional probability of misdetection given $G=g$ is given by

$$P_{MD}(g) = P(W_H(\mathbf{f}_r+\mathbf{g}_r) \le \eta|\mathbf{f}_r \neq \mathbf{0}, G=g) \tag{3.56}$$

$$= \sum_{j=0}^{\lfloor\eta\rfloor} P(W_H(\mathbf{f}_r+\mathbf{g}_r)=j|\mathbf{f}_r\neq\mathbf{0}, G=g) \tag{3.57}$$

$$= \sum_{j=\max(d_{min},g)}^{\min(\lfloor\eta\rfloor,g+d_{min})} \frac{\binom{g}{g+d_{min}-j}\binom{N-g}{j-g}}{\binom{N}{d_{min}}}. \tag{3.58}$$

The conditional probability of $W_H(\mathbf{f}_r + \mathbf{g}_r) = j$ given $\mathbf{f}_r \neq \mathbf{0}$ and $G = g$ is given by

$$\begin{aligned} &P(W_H(\mathbf{f}_r + \mathbf{g}_r) = j | \mathbf{f}_r \neq \mathbf{0}, G = g) \\ &= \begin{cases} \frac{\binom{g}{g+d_{min}-j}\binom{N-g}{j-g}}{\binom{N}{d_{min}}}, & \max(d_{min}, g) \leq j \leq \min(N, g + d_{min}) \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3.59}$$

which follows from the same argument used in obtaining (3.50).

Next, the conditional probability that all selected coded packets are actually unpolluted given $G = g$, $\Lambda = \lambda$ is given by

$$P(\mathcal{F}_\Lambda | \hat{\mathcal{F}}_\Lambda, G = g, \Lambda = \lambda) = P(H_0 | \hat{H}_0, G = g)^\lambda. \tag{3.60}$$

Therefore, it follows from (3.27), (3.44), (3.47), (3.51), and (3.60) that the average probability of decoding error is given by

$$\begin{aligned} P_E = 1 - P(\Lambda = 0) - \sum_{\lambda=1}^{\min(S,R)} \sum_{g=0}^{N} \sum_{c=0}^{R-\lambda} \binom{c + \lambda - 1}{c} P(\hat{H}_1 | G = g)^c P(\hat{H}_0, H_0 | G = g)^\lambda \\ \cdot P(G = g | \Lambda = \lambda) \cdot P(\Lambda = \lambda) \end{aligned} \tag{3.61}$$

where

$$P(\hat{H}_0, H_0 | G = g) = P(\hat{H}_0 | H_0, G = g) P(H_0 | G = g) \tag{3.62}$$

$$= (1 - P_{FA}(g))(1 - p_f). \tag{3.63}$$

3.5.3.1 Scheme I with Reference Threshold η_{opt}

As explained in Section 3.3, the destination finds the threshold η by substituting observed p and $\Lambda = \lambda$ into (3.22). If we assume that the destination knows the value of p_f, it can finds the reference threshold

$$\eta_{opt} = \underset{\eta' \in \{0, \cdots, N\}}{\arg\max} \; P(\text{decoding success} | \Lambda = \lambda) \tag{3.64}$$

by comparing $P(\text{decoding success} | \Lambda = \lambda)$'s for $\eta' = 0, 1, \cdots, N$. Although this might be practically infeasible because it requires knowledge of p_f, η_{opt} could be a good reference to compare with η. Fig.3.3 and Fig.3.4 show that η is close to η_{opt}.

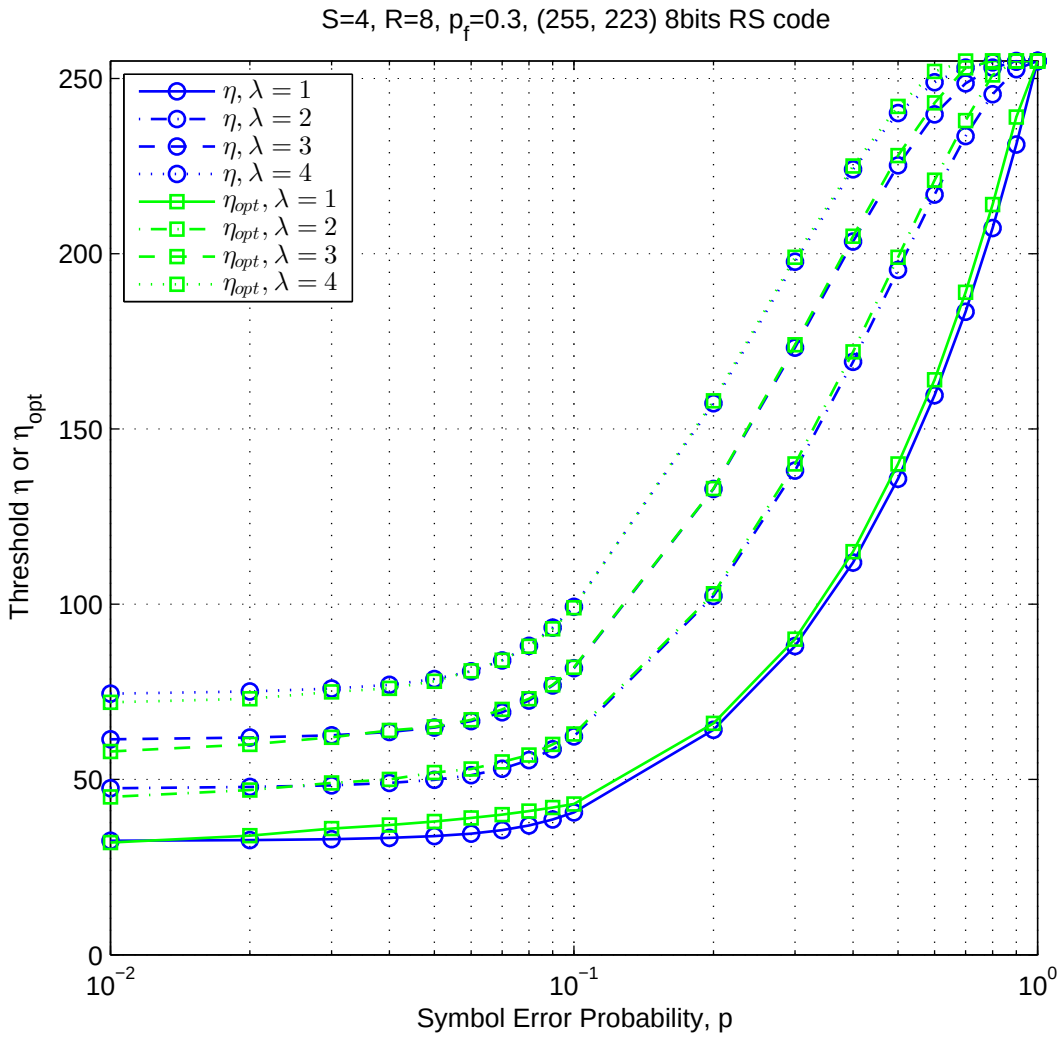

Figure 3.3 Threshold η or η_{opt} versus the symbol error probability p; $S = 4, R = 8, p_f = 0.3, N = 255, K = 223$.

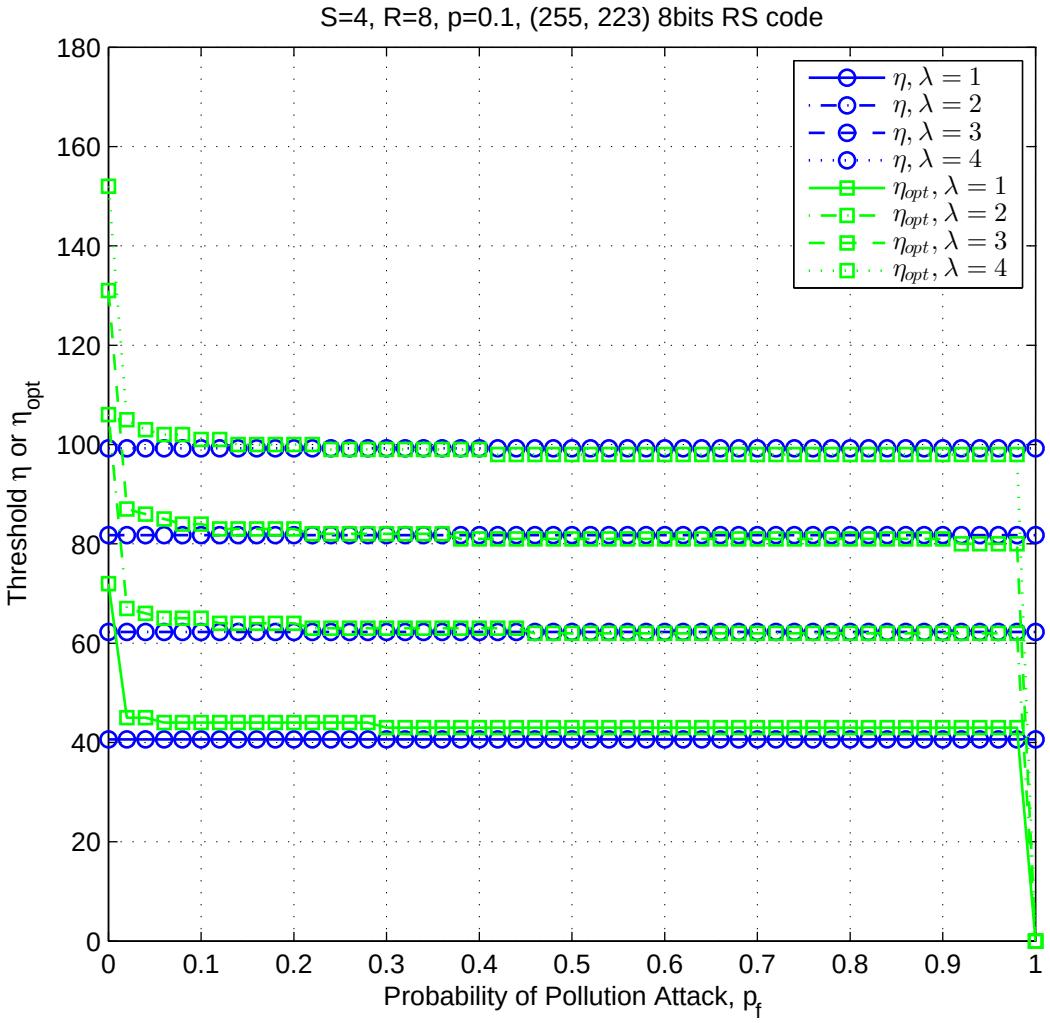

Figure 3.4 Threshold η or η_{opt} versus probability of pollution attack p_f; $S = 4, R = 8, p = 0.1, N = 255, K = 223$.

3.5.4 Scheme II

The scheme II selects coded packets having the lowest $W_H(\mathbf{z}_r)$'s. Therefore, as long as $\Lambda \leq R$, the destination can always choose Λ packets that are potentially unpolluted. Hence, the conditional probability that the destination selects λ coded packets given $\Lambda = \lambda$ is

$$P(\hat{\mathcal{F}}_\Lambda|\Lambda=\lambda) = \begin{cases} 1, & 0 < \lambda \leq R \\ 0, & \text{otherwise} \end{cases} \tag{3.65}$$

Let Γ be the number of polluted packets in which positions of nonzero symbols of $\mathbf{f}_r$ and $\mathbf{g}_r$ are completely overlapped. Then, the conditional probability that the selected coded packets are actually unpolluted given that they are selected and $\Lambda = \lambda$ is given by

$$P(\mathcal{F}_\Lambda|\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda) = \sum_{g=0}^{N}\sum_{r_f=0}^{R}\sum_{\gamma=0}^{r_f} P(\mathcal{F}_\Lambda,\Gamma=\gamma,R_f=r_f,G=g|\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda) \tag{3.66}$$

$$= \sum_{g=0}^{N}\sum_{r_f=0}^{R}\sum_{\gamma=0}^{r_f} \underbrace{P(\mathcal{F}_\Lambda|\Gamma=\gamma,R_f=r_f,G=g,\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda)}_{(3.68)}$$

$$\cdot \underbrace{P(\Gamma=\gamma,R_f=r_f|G=g,\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda)}_{(3.70)} \cdot \underbrace{P(G=g|\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda)}_{P(G=g|\Lambda=\lambda)=(3.47)}. \tag{3.67}$$

Since $W_H(\mathbf{z}_r|\mathbf{f}_r \neq \mathbf{0}) \geq W_H(\mathbf{z}_r|\mathbf{f}_r = \mathbf{0})$, the destination will select Λ packets randomly from $R - R_f + \Gamma$ coded packets where $R - R_f$ packets are unpolluted and Γ packets are polluted. The conditional probability that all selected packets are unpolluted given $\Gamma = \gamma, R_f = r_f, G = g, \hat{\mathcal{F}}_\Lambda, \Lambda = \lambda$ is given by

$$P(\mathcal{F}_\Lambda|\Gamma=\gamma,R_f=r_f,G=g,\hat{\mathcal{F}}_\Lambda,\Lambda=\lambda) = \begin{cases} \frac{\binom{\gamma}{0}\binom{R-r_f}{\lambda}}{\binom{R-r_f+\gamma}{\lambda}}, & \lambda \leq R - r_f \\ 0, & \text{otherwise} \end{cases} \tag{3.68}$$

The conditional probability of $\Gamma = \gamma$ and $R_f = r_f$ given $G = g$, $\hat{\mathcal{F}}_\Lambda$, $\Lambda = \lambda$ is given by

$$\begin{aligned} &P(\Gamma = \gamma, R_f = r_f | G = g, \hat{\mathcal{F}}_\Lambda, \Lambda = \lambda) \\ &= P(\Gamma = \gamma | R_f = r_f, G = g) \cdot P(R_f = r_f) \qquad (3.69) \\ &= \binom{r_f}{\gamma} P_{OL}(g)^\gamma (1 - P_{OL}(g))^{r_f - \gamma} \cdot P(R_f = r_f) \qquad (3.70) \end{aligned}$$

where

$$P_{OL}(g) = P(W_H(\mathbf{f}_r + \mathbf{g}_r) = g | \mathbf{f}_r \neq \mathbf{0}, G = g) \qquad (3.71)$$

$$= \begin{cases} \frac{\binom{g}{d_{min}}\binom{N-g}{0}}{\binom{N}{d_{min}}}, & g \geq d_{min}, \\ 0, & g < d_{min} \end{cases} \qquad (3.72)$$

is the probability that positions of nonzero symbols of $\mathbf{f}_r$ and $\mathbf{g}_r$ are overlapped. Therefore, it follows from (3.28) and (3.65)-(3.72) that the probability of decoding error is given by

$$P_E = 1 - P(\Lambda = 0) - \sum_{g=0}^{N} \sum_{r_f=0}^{R} \sum_{\gamma=0}^{r_f} \sum_{\lambda=1}^{\min(S, R-r_f)} \frac{\binom{R-r_f}{\lambda}}{\binom{R-r_f+\gamma}{\lambda}} \cdot \binom{r_f}{\gamma} P_{OL}(g)^\gamma (1 - P_{OL}(g))^{r_f - \gamma}$$

$$\cdot P(R_f = r_f) \cdot P(G = g | \Lambda = \lambda) \cdot P(\Lambda = \lambda). \qquad (3.73)$$

3.5.5 Asymptotic Analysis for Large N

If we assume that N and K get large enough with fixed rate $\frac{K}{N}$, then

$$\lim_{N,K \to \infty} W_H(\mathbf{e}_s) = Np \qquad (3.74)$$

by the law of large numbers [16]. Therefore, P_e in (3.30) goes to zero if $Np \leq t$ and one if $Np > t$. i.e.,

$$\lim_{N,K \to \infty} P_e = \begin{cases} 0, & \text{if } p \leq \frac{t}{N} \\ 1, & \text{if } p > \frac{t}{N} \end{cases} \qquad (3.75)$$

Therefore, $P(\Lambda = \lambda)$ in (3.29) approaches as follows.

$$\lim_{N,K\to\infty} P(\Lambda = \lambda) = \begin{cases} 1, & \text{if } p \leq \frac{t}{N} \text{ and } \lambda = 0 \\ 1, & \text{if } p > \frac{t}{N} \text{ and } \lambda = S \\ 0, & \text{otherwise} \end{cases} \tag{3.76}$$

This denotes that Λ is always zero when $p \leq \frac{t}{N}$ and S when $p > \frac{t}{N}$, respectively. From (3.27) and (3.76), we note that P_E goes to zero for $p \leq \frac{t}{N}$ regardless of the utilized scheme, because all S message packets are correctly received at the destination during phase 1. i.e.,

$$\lim_{N,K\to\infty} P_E = 0, \text{ if } p \leq \frac{t}{N} \tag{3.77}$$

Therefore, we find $\lim_{N,K\to\infty} P_E$ for $p > \frac{t}{N}$ for each scheme in the remaining part of this subsection.

3.5.5.1 Random Selection

From (3.34) and (3.75), we obtain

$$\lim_{N,K\to\infty} P_E = \begin{cases} 1, & \text{if } S > R \\ \underbrace{1-(1-p_f)^S}_{(a)}, & \text{if } S \leq R \end{cases} \tag{3.78}$$

where (a) denotes that probability that at least one of S coded packets are polluted.

3.5.5.2 Cryptographic Scheme

From (3.42) and (3.76), we obtain

$$\lim_{N,K\to\infty} P_E = \begin{cases} 1, & \text{if } S > R \\ \underbrace{1 - \sum_{r_f=0}^{R-S} \underbrace{P(R_f = r_f)}_{(3.41)}}_{(b)}, & \text{if } S \leq R \end{cases} \tag{3.79}$$

where (b) denotes the probability that the number of polluted coded packets are at most $R-S$.

3.5.5.3 Scheme I

From (3.74), we note that every row vector in $\dot{\mathbf{E}}$ has Np nonzero symbols which are randomly located. Hence, the number of nonzero columns in $\dot{\mathbf{E}}$ exists between Np and $\min(N, SNp)$. Therefore, from (3.47), the conditional probability of $G = g$ given $\Lambda = S$ is given by

$$\lim_{N,K\to\infty} P(G = g|\Lambda = S) = \lim_{N,K\to\infty} P(G_S = g) \tag{3.80}$$

$$= \sum_{g_1=0}^{N} \cdots \sum_{g_{S-1}=0}^{N} \lim_{N,K\to\infty} P(G_S = g, G_{S-1} = g_{S-1}, \cdots G_1 = g_1) \tag{3.81}$$

$$= \sum_{g_2=Np}^{\min(N,2Np)} \sum_{g_3=g_2}^{\min(N,Np+g_2)} \cdots \sum_{g_{S-1}=g_{S-2}}^{\min(N,Np+g_{S-2})} \left\{ \frac{\binom{Np}{2Np-g_2}\binom{N-Np}{g_2-Np}}{\binom{N}{Np}} \cdot \frac{\binom{g_2}{g_2-Np-g_3}\binom{N-g_2}{g_3-g_2}}{\binom{N}{Np}} \cdots \right.$$
$$\left. \cdots \frac{\binom{g_{S-1}}{g_{S-1}-Np-g}\binom{N-g_{S-1}}{g-g_{S-1}}}{\binom{N}{Np}} \right\} \tag{3.82}$$

where G_i denotes the nonzero columns of $\begin{bmatrix} \dot{\mathbf{e}}_1 \\ \vdots \\ \dot{\mathbf{e}}_i \end{bmatrix}$ and

$$\lim_{N,K\to\infty} P(G_S = g, G_{S-1} = g_{S-1}, \cdots G_1 = g_1)$$
$$= \lim_{N,K\to\infty} P(G_S = g|G_{S-1} = g_{S-1}, \cdots G_1 = g_1) \cdots$$
$$\cdots \lim_{N,K\to\infty} P(G_2 = g_2|G_1 = g_1) \lim_{N,K\to\infty} P(G_1 = g_1) \tag{3.83}$$

and

$$\lim_{N,K\to\infty} P(G_1 = g_1) = \begin{cases} 1, & \text{if } g_1 = Np \\ 0, & \text{otherwise} \end{cases} \tag{3.84}$$

denotes that every row vector of $\dot{\mathbf{E}}$ has Np nonzero symbols, and

$$\begin{aligned}
&\lim_{N,K\to\infty} P(G_i = g_i | G_{i-1} = g_{i-1}, \cdots, G_1 = g_1) \\
&= \lim_{N,K\to\infty} P(G_i = g_i | G_{i-1} = g_{i-1}) \qquad (3.85) \\
&= \begin{cases} \frac{\binom{g_{i-1}}{g_{i-1}+Np-g_i}\binom{N-g_{i-1}}{g_i-g_{i-1}}}{\binom{N}{Np}}, & \text{if } g_{i-1} \le g_i \le \min(N, Np+g_{i-1}) \\ 0, & \text{otherwise} \end{cases} \qquad (3.86)
\end{aligned}$$

denotes the conditional probability that $\begin{bmatrix} \dot{\mathbf{e}}_1 \\ \vdots \\ \dot{\mathbf{e}}_i \end{bmatrix}$ has g_i nonzero columns given that $\begin{bmatrix} \dot{\mathbf{e}}_1 \\ \vdots \\ \dot{\mathbf{e}}_{i-1} \end{bmatrix}$ has g_{i-1} nonzero columns.

Therefore, from above equations, the probability of decoding error of the scheme I in (3.61) is simplified to

$$\begin{aligned}
&\lim_{N,K\to\infty} P_E \\
&= \begin{cases} 1 - \sum_{g=Np}^{\min(N,SNp)} \sum_{c=0}^{R-S} \binom{c+S-1}{c} \underbrace{P(\hat{H}_1|G=g)}_{(3.52)}{}^{c} \underbrace{P(\hat{H}_0, H_0|G=g)}_{(3.63)}{}^{S} \underbrace{P(G=g|\Lambda=S)}_{(3.82)}, & \text{if } S \le R \\ 0, & \text{if } S > R \end{cases} \qquad (3.87)
\end{aligned}$$

3.5.5.4 Scheme II

From (3.76) and (3.73), the probability of decoding error of the scheme II is simplified to

$$\lim_{N,K\to\infty} P_E = 0 \qquad (3.88)$$

for $S > R$, and

$$\lim_{N,K\to\infty} P_E = 1 - \sum_{g=Np}^{\min(N,SNp)} \sum_{r_f=0}^{R-S} \sum_{\gamma=0}^{r_f} \frac{\binom{R-r_f}{S}}{\binom{R-r_f+\gamma}{S}} \binom{r_f}{\gamma} P_{OL}(g)^{\gamma}(1-P_{OL}(g))^{r_f-\gamma} \cdot P(R_f = r_f) \underbrace{P(G = g|\Lambda = S)}_{(3.82)} \tag{3.89}$$

for $S \leq R$.

3.5.6 Numerical Results

In the legend of each plot, 'ana' and 'sim' denotes analysis and simulation, respectively. Fig.3.5 shows the plot of the average probability of decoding error P_E versus the symbol error probability p for the case of $q = 256, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code. We can see that P_E of every scheme increases with increasing p. This is because larger p causes smaller probability that all message packets are correctly channel-decoded. Besides this reason, the proposed schemes have another reason. Since larger p statistically causes larger $W_H(\mathbf{g}_r)$ by (3.47)-(3.50), the probability of $W_H(\mathbf{f}_r + \mathbf{g}_r) = W_H(\mathbf{g}_r)$ (i.e., the probability that all nonzero symbols of $\mathbf{f}_r$ are completely overlapped with nonzero symbols of $\mathbf{g}_r$) increases. As a result, it gets harder that unpolluted coded packets are distinguished from polluted coded packets. We can see that P_E of the proposed scheme I with η is smaller than that of the random selection scheme, and is close to that of the proposed scheme I with η_{opt}, in the practical range of $p \leq 0.2$. This denotes that the polluted coded packets are detected well by η for the practically small p. We also notice that P_E of the proposed scheme II is much smaller than that of the proposed scheme I. This denotes that the proposed scheme II minimizes the case that polluted coded packets are mis-detected, by comparing all coded packets at the cost of full (R) delay. Fig.3.6 shows the plot of the average probability of decoding error P_E versus the symbol error probability p for the case of $q = 256, S = 5, R = 10, p_f = 0.3$, and $(N = 255, K = 223)$ 8bits (not shortened)

RS code. Only difference from Fig.3.5 is N and K. We can see the same trend with Fig.3.5 in this figure.

Fig.3.7 shows the plot of the average probability of decoding error P_E versus the probability of pollution attack p_f for the case of $q = 256, p = 0.1, S = 5, R = 10$, and $(N = 16, K = 14)$ 8bits shortened RS code. We can see that P_E of every scheme increases with increasing p_f. We also can see that P_E of the proposed scheme II is much lower than that of the proposed scheme I. As described earlier, this benefit of the proposed scheme II over the proposed scheme I comes from the cost that the proposed scheme II compares all R coded packets. Fig.3.8 shows the plot for the case of $q = 256, p = 0.1, S = 5, R = 10$, and $(N = 255, K = 223)$ 8bits RS code. In this figure, P_E of the proposed scheme II approaches that of the cryptographic scheme.

Fig.3.9 shows the plot of the average probability of decoding error P_E versus Hamming weight of false injection vectors $W_H(\mathbf{f}_r)$ for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code. We can see that P_E's of the random selection scheme and the cryptographic scheme are not functions of $W_H(\mathbf{f}_r)$ as given by (3.34) and (3.42), respectively. We also can see that P_E's of the proposed scheme I with η, with η_{opt}, and the proposed scheme II decay as $W_H(\mathbf{f}_r)$ increases. This is intuitively because larger $W_H(\mathbf{f}_r)$ results in the lower probability that a polluted coded packet is mis-detected as an unpolluted coded packet, by (3.59) and (3.72). In other words, to the proposed schemes, larger $W_H(\mathbf{f}_r)$ causes that polluted coded packets are more easily distinguished from unpolluted coded packets. We also notice that the proposed scheme I with η has P_E close to that of the proposed scheme I with η_{opt} and that the proposed scheme II has the much lower P_E than the proposed scheme I. Fig.3.10 is for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits (not shortened) RS code and has the same trend with Fig.3.9.

Fig.3.11 shows the plot of the average probability of decoding error P_E versus message length K for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$ and $(N = 16, K)$ 8bits

shortened RS code, while Fig.3.12 is for the case of $(N = 255, K)$ 8bits RS code. We can see that P_E's of all schemes increase with the increasing K. This is because error correction capability $t = \lfloor \frac{N-K}{2} \rfloor$ decreases as K increases. For proposed schemes, this is also because Hamming weight of false injection vector $W_H(\mathbf{f}_r) = d_{min}$ decays as K increases.

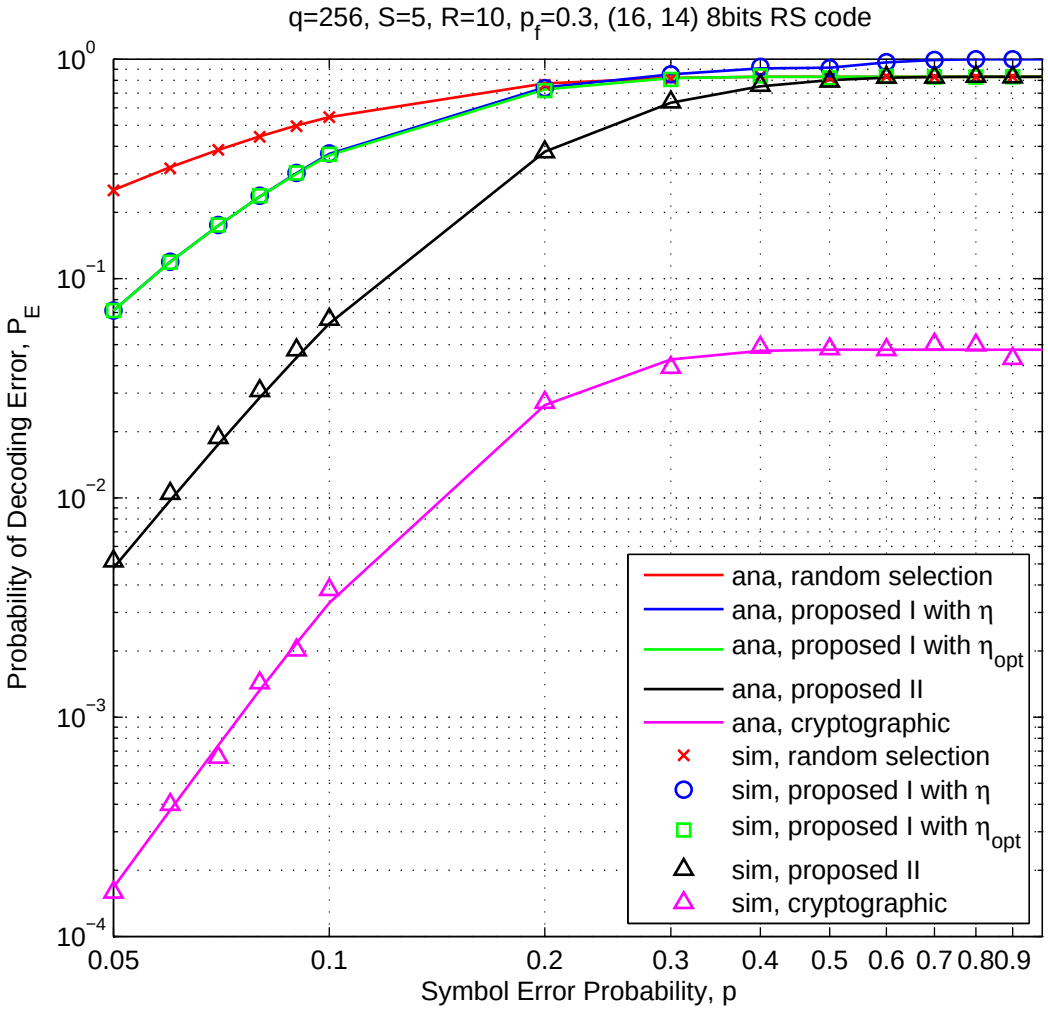

Figure 3.5 The average probability of decoding error P_E versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

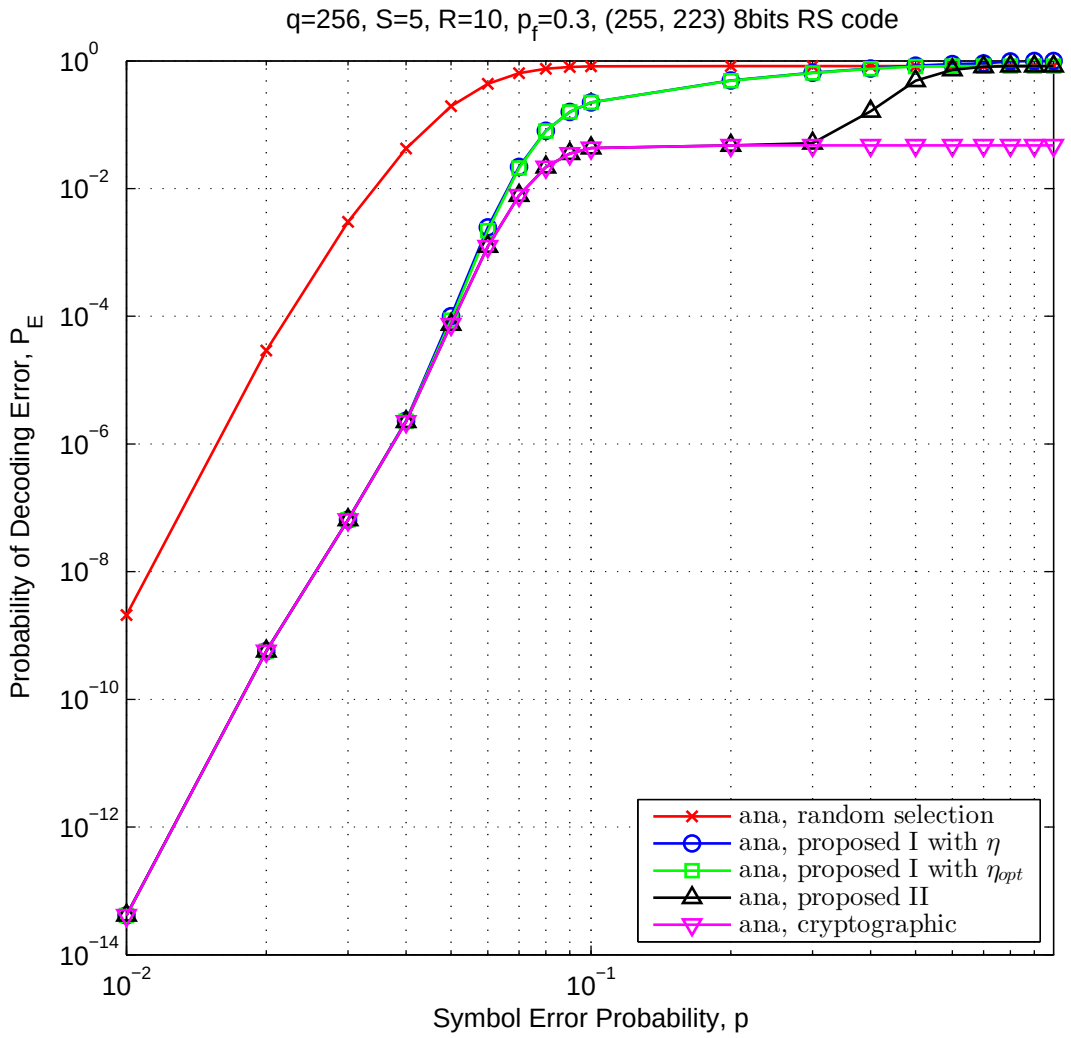

Figure 3.6 The average probability of decoding error P_E versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

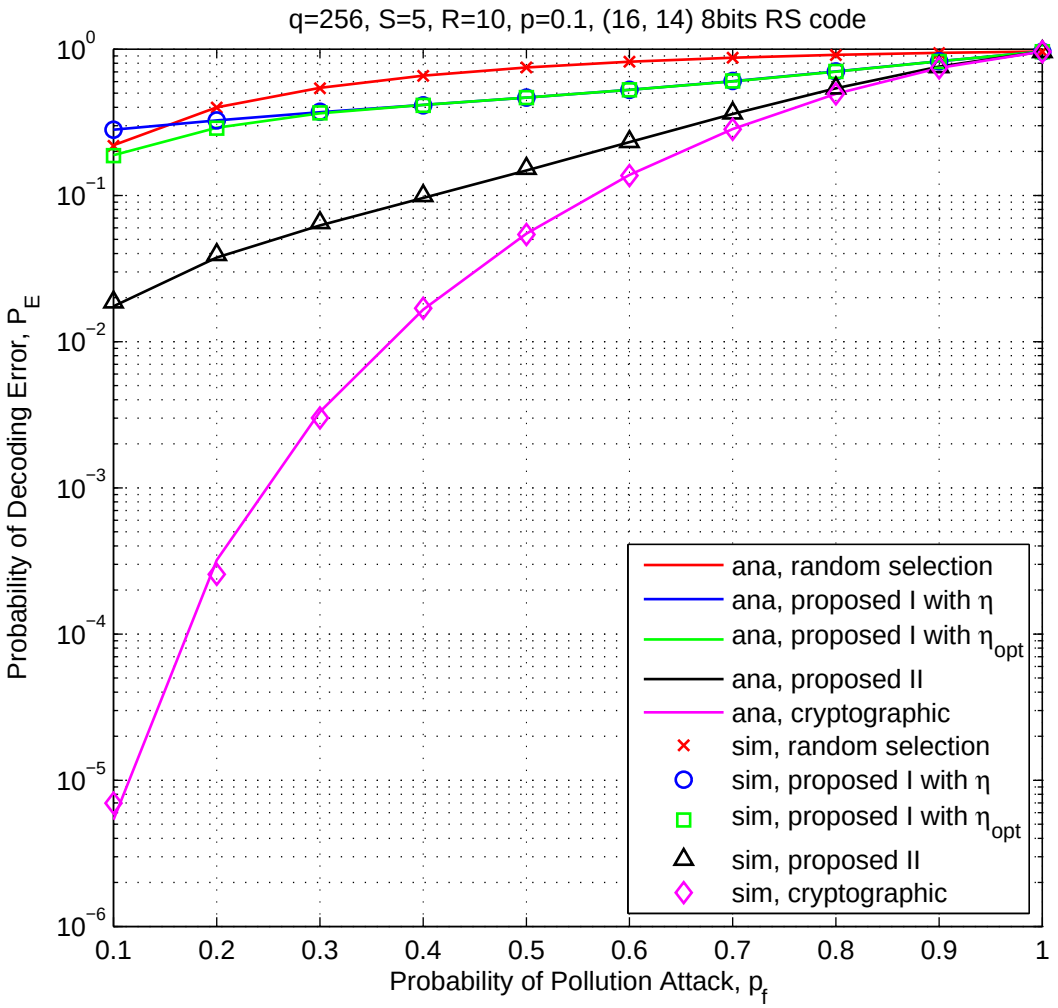

Figure 3.7 The average probability of decoding error P_E versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 16, K = 14$.

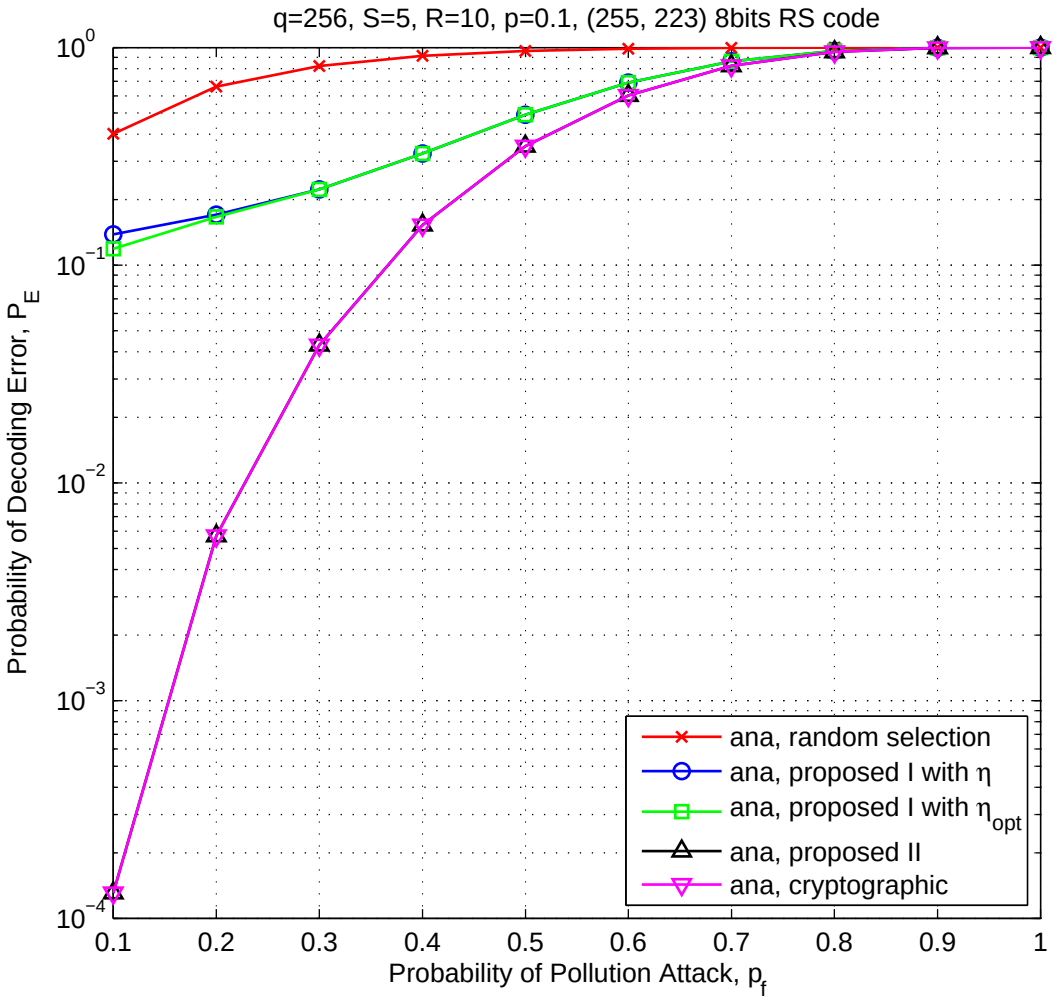

Figure 3.8 The average probability of decoding error P_E versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 255, K = 223$.

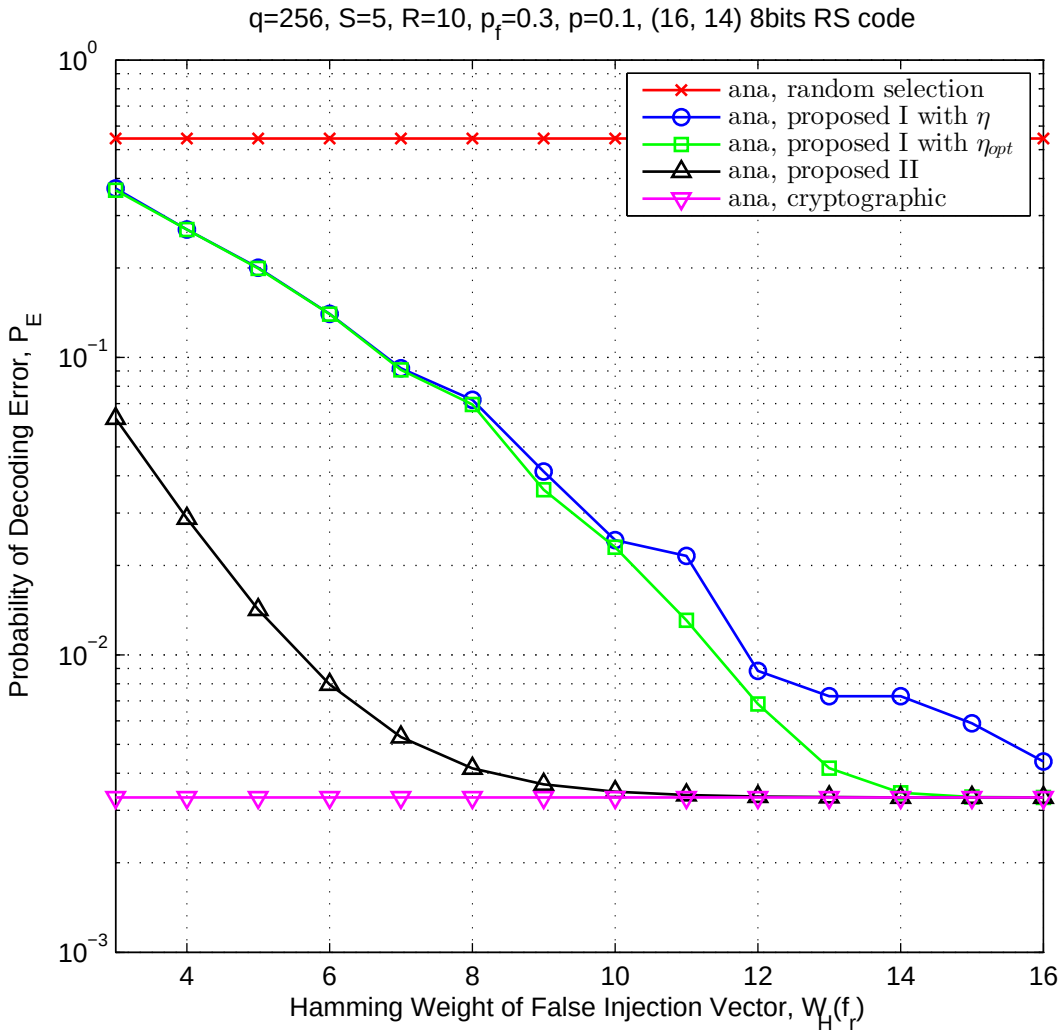

Figure 3.9 The average probability of decoding error P_E versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

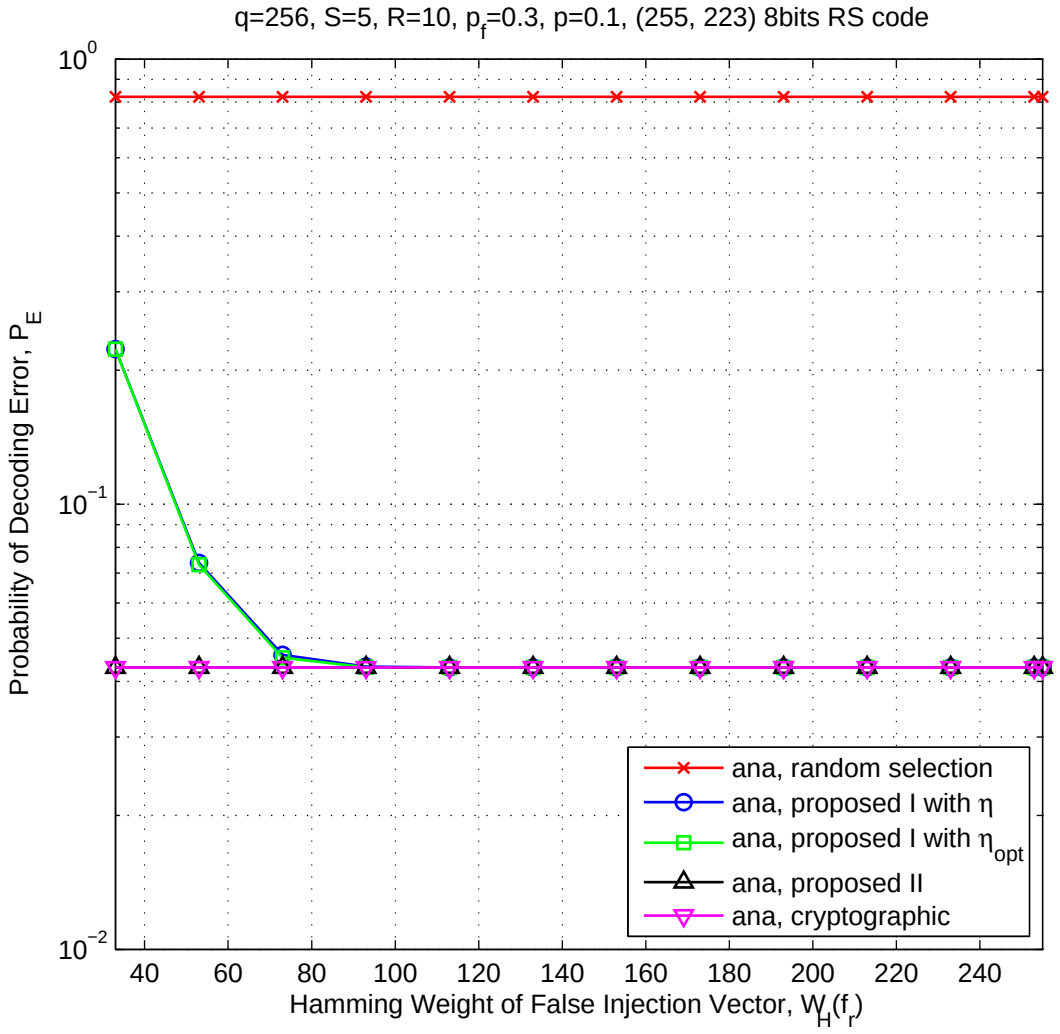

Figure 3.10 The average probability of decoding error P_E versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

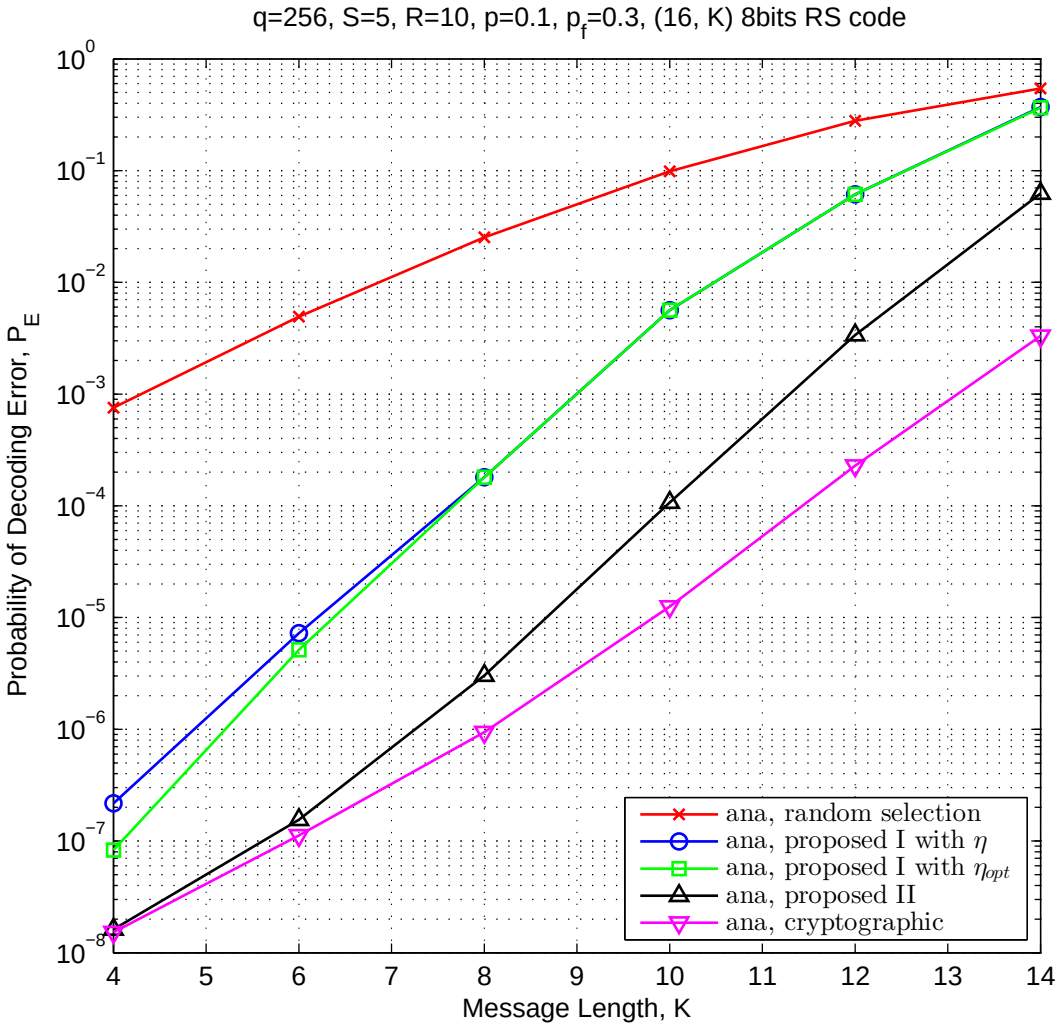

Figure 3.11 The average probability of decoding error P_E versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16$.

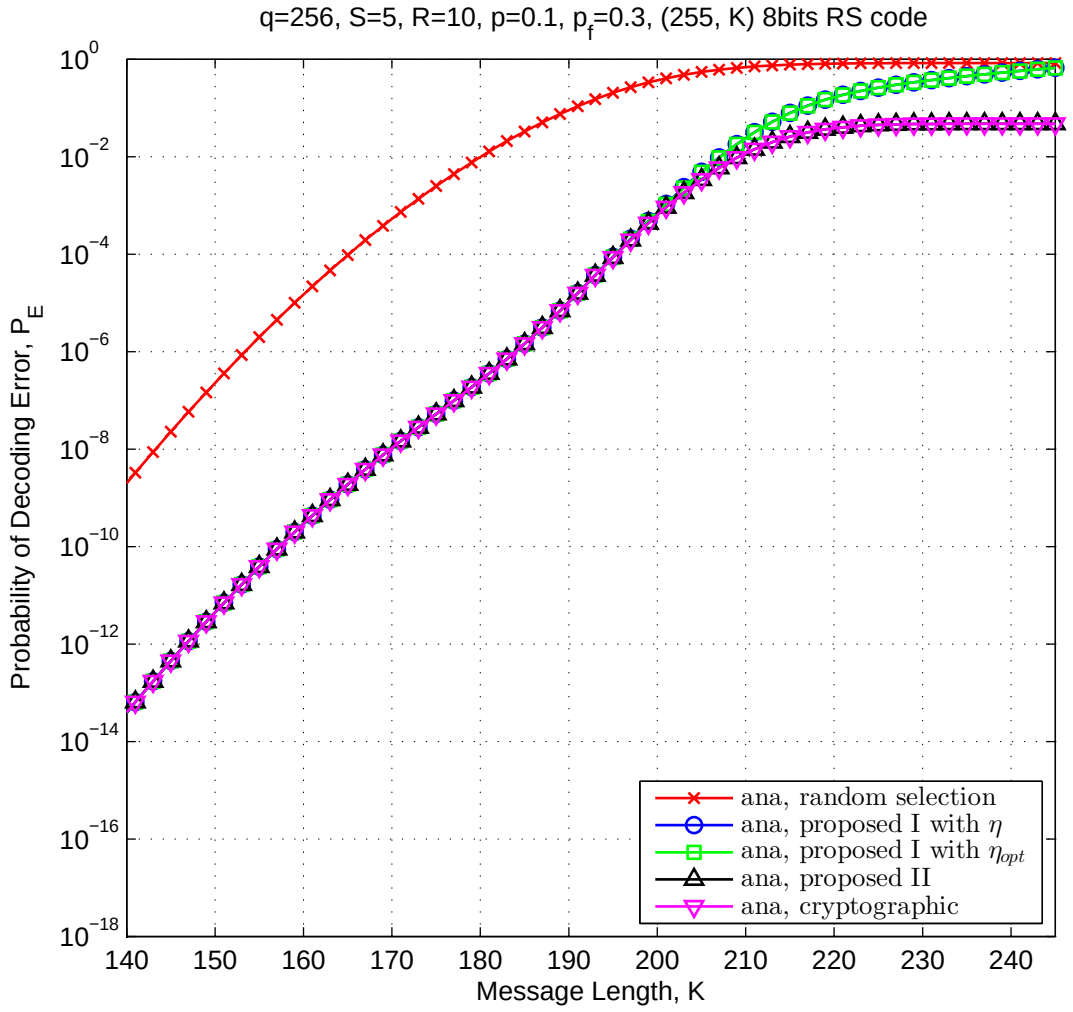

Figure 3.12 The average probability of decoding error P_E versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255$.

3.6 Average Delay

We define the delay Δ as the number of received coded packets before the decoding starts, regardless of whether the decoding is successful or not. Therefore, the average delay is given by

$$E[\Delta] = \sum_{\lambda=0}^{S} \sum_{\delta=0}^{R} \delta \cdot P(\Delta = \delta | \Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.90}$$

$$= \sum_{\lambda=1}^{S} \sum_{\delta=0}^{R} \delta \cdot P(\Delta = \delta | \Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.91}$$

where (3.91) follows from

$$P(\Delta = 0 | \Lambda = 0) = 1. \tag{3.92}$$

If $S < R$, then

$$E[\Delta] = \sum_{\lambda=1}^{S} \left\{ \sum_{\delta=\lambda}^{R-1} \delta \cdot P(\Delta = \delta | \Lambda = \lambda) + R \cdot \left(1 - \sum_{\delta=\lambda}^{R-1} P(\Delta = \delta | \Lambda = \lambda) \right) \right\} \cdot P(\Lambda = \lambda) \tag{3.93}$$

$$= \sum_{\lambda=1}^{S} \left\{ R - \sum_{\delta=\lambda}^{R-1} (R - \delta) \cdot P(\Delta = \delta | \Lambda = \lambda) \right\} \cdot P(\Lambda = \lambda) \tag{3.94}$$

$$= R(1 - P(\Lambda = 0)) - \sum_{\lambda=1}^{S} \sum_{\delta=\lambda}^{R-1} (R - \delta) \cdot P(\Delta = \delta | \Lambda = \lambda) \cdot P(\Lambda = \lambda) \tag{3.95}$$

where (3.93) follows from

$$P(\Delta = 0 | \Lambda = \lambda) = 0 \text{ for } \delta \leq \lambda - 1. \tag{3.96}$$

If $S \geq R$,

$$E[\Delta] = \sum_{\lambda=1}^{R-1}\sum_{\delta=0}^{R} \delta \cdot P(\Delta=\delta|\Lambda=\lambda)\cdot P(\Lambda=\lambda) + \sum_{\lambda=R}^{S}\sum_{\delta=0}^{R} \delta \cdot P(\Delta=\delta|\Lambda=\lambda)\cdot P(\Lambda=\lambda) \tag{3.97}$$

$$= \sum_{\lambda=1}^{R-1}\sum_{\delta=\lambda}^{R} \delta \cdot P(\Delta=\delta|\Lambda=\lambda)\cdot P(\Lambda=\lambda) + R\sum_{\lambda=R}^{S} P(\Lambda=\lambda) \tag{3.98}$$

$$= \sum_{\lambda=1}^{R-1}\left\{R - \sum_{\delta=\lambda}^{R-1}(R-\delta)P(\Delta=\delta|\Lambda=\lambda)\right\}\cdot P(\Lambda=\lambda) + R\sum_{\lambda=R}^{S} P(\Lambda=\lambda) \tag{3.99}$$

$$= R(1-P(\Lambda=0)) - \sum_{\lambda=1}^{R-1}\sum_{\delta=\lambda}^{R-1}(R-\delta)\cdot P(\Delta=\delta|\Lambda=\lambda)\cdot P(\Lambda=\lambda) \tag{3.100}$$

where (3.98) follows from

$$P(\Delta=\delta|\Lambda=\lambda) = 0 \text{ for } \delta \leq R-1, \lambda \geq R \tag{3.101}$$

and

$$P(\Delta=R|\Lambda=\lambda) = 1 \text{ for } \lambda \geq R \tag{3.102}$$

and (3.99) follows from

$$\sum_{\delta=\lambda}^{R} P(\Delta=\delta|\Lambda=\lambda) = 1. \tag{3.103}$$

Therefore, it follows from (3.95) and (3.100) that the average delay is given by

$$E[\Delta] = R(1-P(\Lambda=0)) - \sum_{\lambda=1}^{\min(S,R-1)}\sum_{\delta=\lambda}^{R-1}(R-\delta)\cdot P(\Delta=\delta|\Lambda=\lambda)\cdot P(\Lambda=\lambda) \tag{3.104}$$

where $P(\Lambda=\lambda)$ is given by (3.29).

3.6.1 Random Selection

Since the random selection scheme selects the first Λ coded packets, the conditional distribution of Δ given $\Lambda=\lambda$ is given by

$$P(\Delta=\delta|\Lambda=\lambda) = \begin{cases} 1, & \delta=\lambda \\ 0, & \text{otherwise} \end{cases} \tag{3.105}$$

if $\lambda \leq R$ and

$$P(\Delta = \delta | \Lambda = \lambda) = \begin{cases} 1, & \delta = R \\ 0, & \delta < R \end{cases} \tag{3.106}$$

if $\lambda > R$. Then, it follows from (3.104), (3.105), and (3.106) that the average delay is given by

$$E[\Delta] = R(1 - P(\Lambda = 0)) - \sum_{\lambda=1}^{\min(S,R-1)} (R - \lambda) \cdot P(\Lambda = \lambda) \tag{3.107}$$

$$= \begin{cases} \sum_{\lambda=1}^{S} \lambda \cdot P(\Lambda = \lambda), & \text{if } S < R \\ \sum_{\lambda=1}^{R-1} \lambda \cdot P(\Lambda = \lambda) + \sum_{\lambda=R}^{S} R \cdot P(\Lambda = \lambda), & \text{if } S \geq R \end{cases} \tag{3.108}$$

$$= \sum_{\lambda=1}^{S} \min(\lambda, R) \cdot P(\Lambda = \lambda). \tag{3.109}$$

3.6.2 Cryptographic Scheme

Since the cryptographic scheme is assumed to perfectly detect the presence of pollution attack, delay is equal to the number of received coded packets until Λth unpolluted coded packet is received. Therefore, the conditional probability of Δ given $\Lambda = \lambda$ is given by

$$P(\Delta = \delta | \Lambda = \lambda) = \begin{cases} 1, & \delta = \lambda = 0 \\ \binom{\delta - 1}{\delta - \lambda} \underbrace{P(H_1)}_{p_f}^{\delta - \lambda} \underbrace{P(H_0)}_{1-p_f}^{\lambda}, & 1 \leq \lambda \leq \delta \\ 0, & \text{otherwise} \end{cases} \tag{3.110}$$

if $\delta \leq R - 1$ and

$$P(\Delta = R | \Lambda = \lambda) = \begin{cases} 0, & \lambda = 0 \\ 1 - \sum_{d=\lambda}^{R-1} \binom{d-1}{d-\lambda} P(H_1)^{d-\lambda} P(H_0)^{\lambda}, & 1 \leq \lambda \leq R - 1 \\ 1, & \lambda \geq R \end{cases} \tag{3.111}$$

Therefore, it follows from (3.104), (3.110), and (3.111) that the average delay is given by

$$E[\Delta] = R(1 - P(\Lambda = 0)) - \sum_{\lambda=1}^{\min(S,R-1)} \sum_{\delta=\lambda}^{R-1} (R - \delta) \cdot \binom{\delta - 1}{\delta - \lambda} p_f^{\delta-\lambda} (1 - p_f)^{\lambda} \cdot P(\Lambda = \lambda). \tag{3.112}$$

3.6.3 Scheme I

The conditional distribution of Δ given $\Lambda = \lambda$ is given by

$$P(\Delta = \delta | \Lambda = \lambda) = \sum_{g=0}^{N} P(\Delta = \delta, G = g | \Lambda = \lambda) \tag{3.113}$$

$$= \sum_{g=0}^{N} \underbrace{P(\Delta = \delta | G = g, \Lambda = \lambda)}_{(3.115)} \cdot \underbrace{P(G = g | \Lambda = \lambda)}_{(3.47)} \tag{3.114}$$

where G is the number of nonzero columns in $\dot{\mathbf{E}}$ given by (3.45) and $P(G = g|\Lambda = \lambda)$ is given by (3.47). Similarly to (3.110), the conditional distribution of Δ given $G = g$ and $\Lambda = \lambda$ is given by

$$P(\Delta = \delta | G = g, \Lambda = \lambda)$$

$$= \begin{cases} 1, & \text{if } \lambda = \delta = 0 \\ \binom{\delta - 1}{\delta - \lambda} \underbrace{P(\hat{H}_1 | G = g)}_{(3.52)}^{\delta-\lambda} \underbrace{P(\hat{H}_0 | G = g)}_{(3.117)}^{\lambda}, & \text{if } 1 \le \lambda \le \delta \le R - 1 \\ 1 - \sum_{d=\lambda}^{R-1} \binom{d - 1}{d - \lambda} P(\hat{H}_1 | G = g)^{d-\lambda} P(\hat{H}_0 | G = g)^{\lambda}, & \text{if } 1 \le \lambda \le \delta = R \\ 1, & \text{if } \delta = R < \lambda \\ 0, & \text{otherwise} \end{cases} \tag{3.115}$$

where $P(\hat{H}_1|G = g)$ which denotes the conditional probability that a received coded packet is decided as polluted given $G = g$ is given by (3.52) and $P(\hat{H}_0|G = g)$ is similarly given by

$$P(\hat{H}_0 | G = g) = 1 - P(\hat{H}_1 | G = g) \tag{3.116}$$

$$= (1 - P_{FA}(g))(1 - p_f) + P_{MD}(g) p_f. \tag{3.117}$$

Therefore, by (3.113)-(3.117) and (3.104), the average delay is given by

$$E[\Delta] = R(1 - P(\Lambda = 0)) - \sum_{\lambda=1}^{\min(S,R-1)} \sum_{\delta=\lambda}^{R-1} \sum_{g=0}^{N} (R - \delta) \cdot \binom{\delta - 1}{\delta - \lambda} P(\hat{H}_1 | G = g)^{\delta - \lambda} P(\hat{H}_0 | G = g)^{\lambda} \cdot P(G = g | \Lambda = \lambda) \cdot P(\Lambda = \lambda). \quad (3.118)$$

3.6.4 Scheme II

Since the proposed scheme II requires that all R coded packets are compared as described in section 3.3, delay is always R when $\Lambda \geq 1$. Therefore, the average delay is given by

$$E[\Delta] = R(1 - P(\Lambda = 0)) \quad (3.119)$$

regardless of $S > R$ or $S \leq R$.

3.6.5 Asymptotic Analysis for Large N

In this subsection, we derive the average delay when N is large. When $p \leq \frac{t}{N}$, all S message packets are correctly channel-decoded during phase 1 due to (3.76). Hence,

$$\lim_{N,K \to \infty} E[\Delta] = 0, \text{ if } p \leq \frac{t}{N} \quad (3.120)$$

regardless of the utilized scheme during phase 2. Hence, we present the average delay for large N if $p > \frac{t}{N}$ in the remaining part of this subsection.

3.6.5.1 Random Selection

Following from (3.76) and (3.109) we obtain

$$\lim_{N,K \to \infty} E[\Delta] = \begin{cases} R, & \text{if } S > R \\ S, & \text{if } S \leq R \end{cases} \quad (3.121)$$

$$= \min(S, R). \quad (3.122)$$

3.6.5.2 Cryptographic Scheme

From (3.76) and (3.112), we obtain

$$\lim_{N,K\to\infty} E[\Delta] = \begin{cases} R, & \text{if } S \geq R \\ R - \sum_{\delta=S}^{R-1} (R-\delta) \underbrace{\binom{\delta-1}{\delta-S} p_f^{\delta-S}(1-p_f)^S}_{(a)}, & \text{if } S < R \end{cases} \tag{3.123}$$

where (a) is the probability that the δthly received coded packet is the Sth unpolluted coded packet.

3.6.5.3 Scheme I

Following from (3.76) and (3.118), the average delay of the scheme I is simplified as

$$\lim_{N,K\to\infty} E[\Delta]$$
$$= \begin{cases} R, & \text{if } S \geq R \\ R - \sum_{\delta=S}^{R-1} \sum_{g=Np}^{\min(N,SNp)} (R-\delta) \underbrace{\binom{\delta-1}{\delta-S} P(\hat{H}_1|G=g)^{\delta-S} P(\hat{H}_0|G=g)^S}_{(b)} \underbrace{P(G=g|\Lambda=S)}_{(3.82)}, & \text{if } S < R \end{cases} \tag{3.124}$$

where (b) denotes the probability that the δth coded packet is the Sthly selected coded packet.

3.6.5.4 Scheme II

Since delay of the proposed scheme II is always R for $\Lambda \geq 1$ as described in the section 3.6, we obtain

$$\lim_{N,K\to\infty} E[\Delta] = R. \tag{3.125}$$

3.6.6 Numerical Results

Fig.3.13 shows the plot of the average delay $E[\Delta]$ versus the symbol error probability p for the case of $q = 256, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code, while Fig.3.14 is for the case of $(N = 255, K = 223)$ 8bits (not shortened) RS code. We can see that P_E's of the random selection scheme, the proposed scheme II, and the cryptographic scheme increase as p increases. For the proposed scheme II, this is because $P(\Lambda = 0)$ decreases with increasing p, as shown by (3.119). For the random selection scheme and the cryptographic scheme, this is because Λ increases with increasing p. as shown in (3.109) and (3.112). We also see that $E[\Delta]$ of the proposed I with η_{opt} increases as p increases and turns into decreasing at a certain p and finally meets $E[\Delta]$ of the random selection scheme at large p. This is because of following reasons.

1. For small p, $E[\Delta]$ increases as p increases because Λ increases, such as the random selection scheme and the cryptographic scheme.

2. At a certain p, $E[\Delta]$ stops increasing and starts decreasing. This is because larger p causes the larger number of mis-detection.

3. At very large p, $E[\Delta]$ meets that of the random selection scheme. This is because so large p causes that $W_H(\mathbf{z}_r)$'s of the polluted coded packets and the unpolluted coded packets have the same value of N thus η_{opt} also becomes N. As a result, η_{opt} cannot distinguish polluted coded packets from unpolluted coded packets. In other words, every coded packet is detected as unpolluted, such as the random selection scheme.

Fig.3.15 shows the plot of the average delay $E[\Delta]$ versus the probability of pollution attack p_f for the case of $q = 256, p = 0.1, S = 5, R = 10$, and $(N = 16, K = 14)$ 8bits shortened RS code. We can see that $E[\Delta]$'s of the random selection and the proposed scheme II do not differ as p_f varies. This is because they are not functions of p_f as

shown in (3.109) and (3.119). On the other hand, we can see that $E[\Delta]$'s of the proposed scheme I with η, with η_{opt}, and the cryptographic scheme increases as p_f increases. We also notice that lines of proposed scheme I with η and η_{opt} are close to that of the cryptographic scheme. This denotes that they detect polluted coded packets with high probability close to one. Fig.3.16 is for the case of $(N = 255, K = 223)$ 8bits RS code and shows the same trend.

Fig.3.17 shows the plot of the average delay $E[\Delta]$ versus Hamming weight of false injection vectors $W_H(\mathbf{f}_r)$ for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code, while Fig.3.18 is for the case of $(N = 255, K = 223)$ 8bits RS code. In both figures, we can see that $E[\Delta]$'s of the random selection scheme, the cryptographic scheme, and the proposed scheme II are not functions of $W_H(\mathbf{f}_r)$ as shown by (3.109), (3.112), and (3.119). On the other hand, we can see that $E[\Delta]$'s of the proposed scheme I with η and with η_{opt} reduce to that of the cryptographic scheme as $W_H(\mathbf{f}_r)$ increases. This is intuitively because polluted coded packets are more easily distinguished with the larger $W_H(\mathbf{f}_r)$ thus the number of mis-detection decreases.

Fig.3.19 shows the plot of the average delay $E[\Delta]$ versus message length K for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$ and $(N = 16, K)$ 8bits shortened RS code, while Fig.3.20 shows the plot of the average delay $E[\Delta]$ versus message length K for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$ and $(N = 255, K)$ 8bits RS code. In both figures, we can see that $E[\Delta]$ increases as K increases.

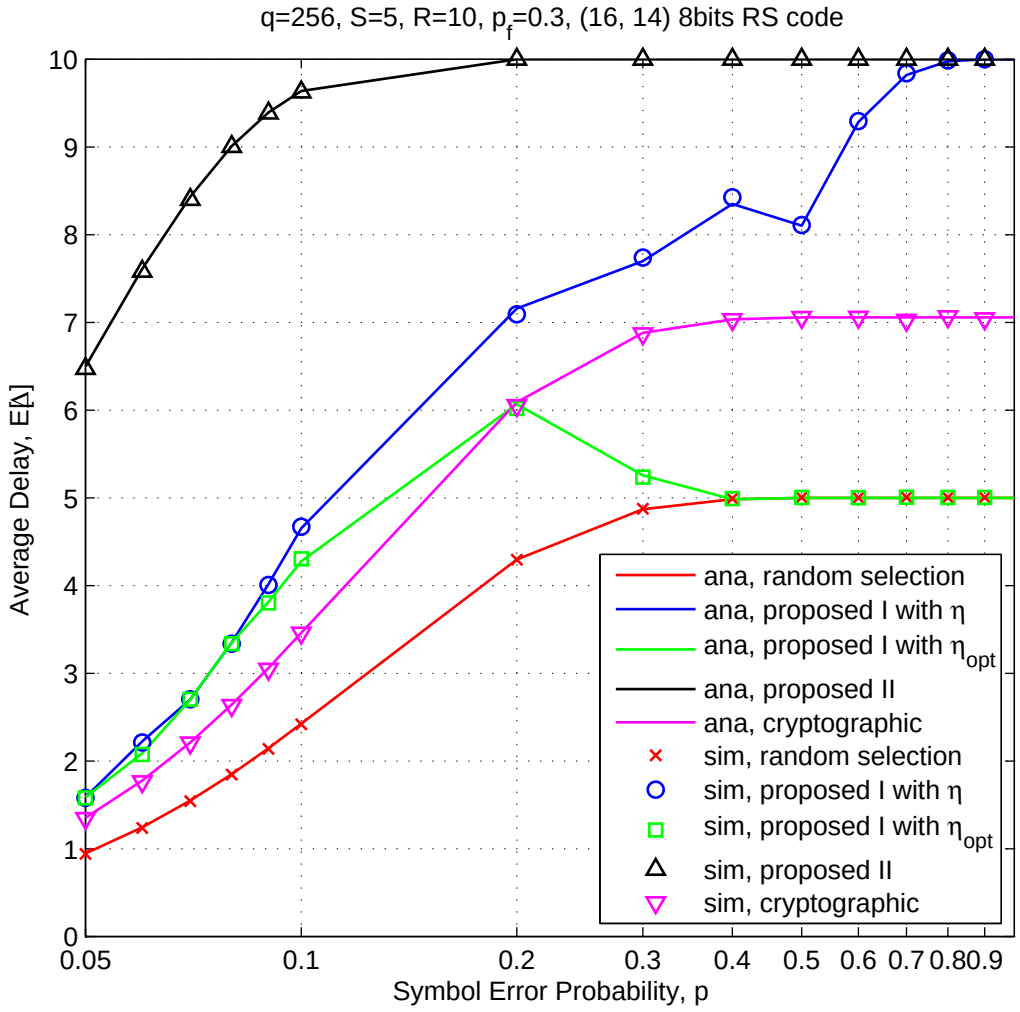

Figure 3.13 The average delay $E[\Delta]$ versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

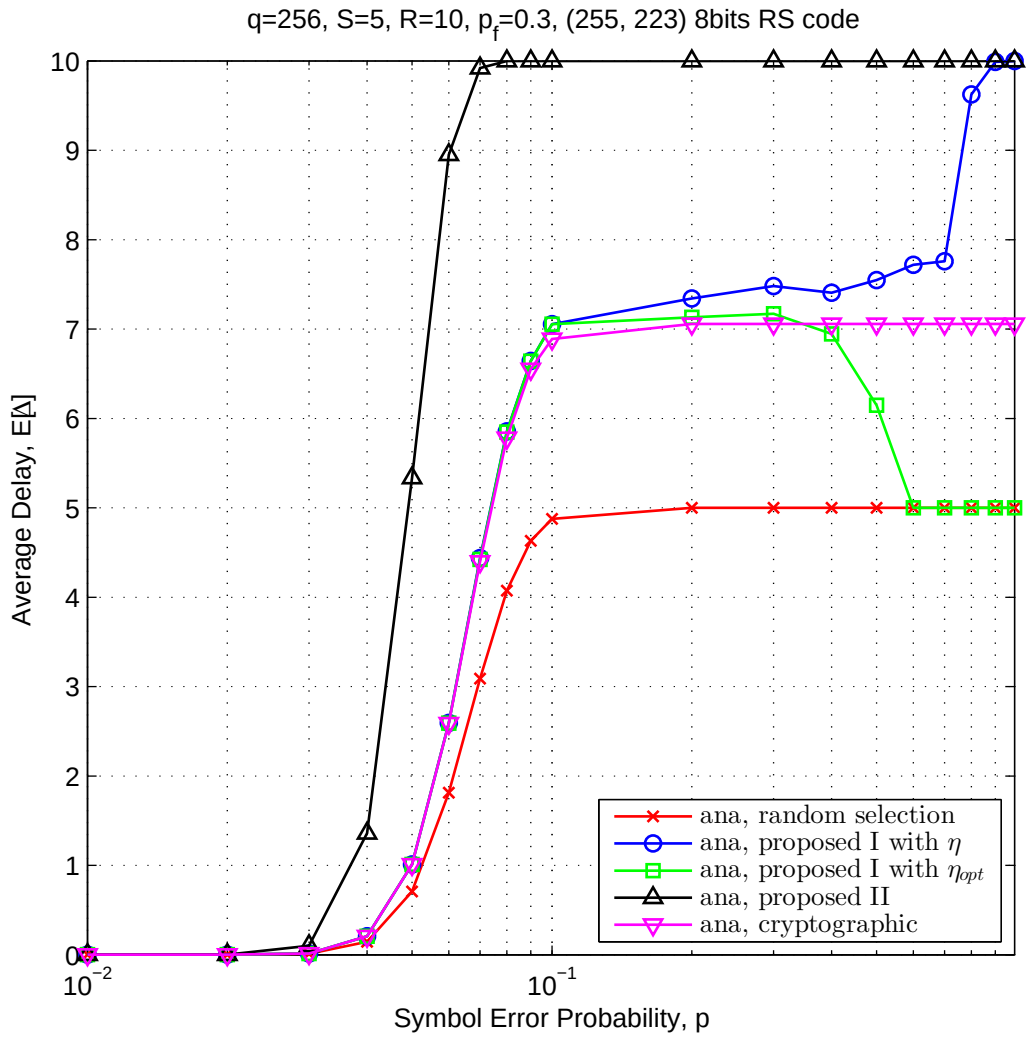

Figure 3.14 The average delay $E[\Delta]$ versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

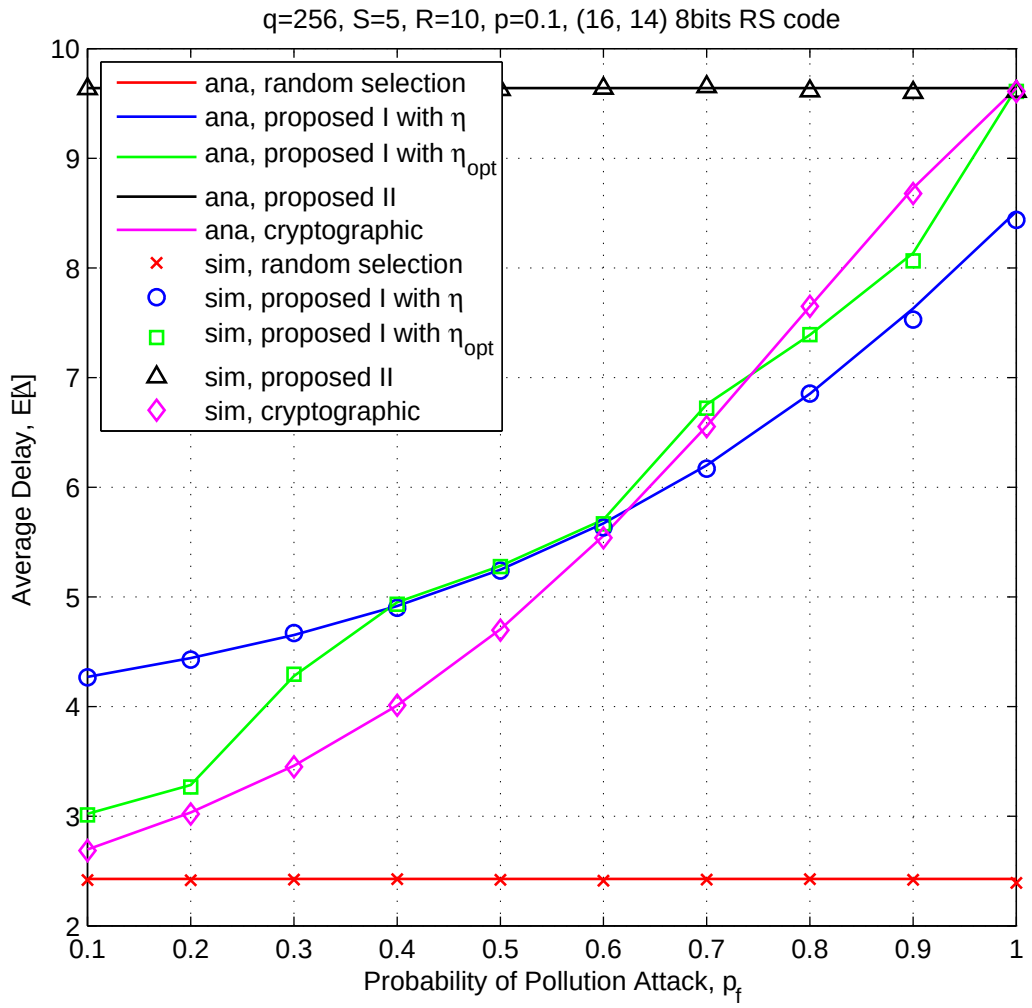

Figure 3.15 The average delay $E[\Delta]$ versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 16, K = 14$.

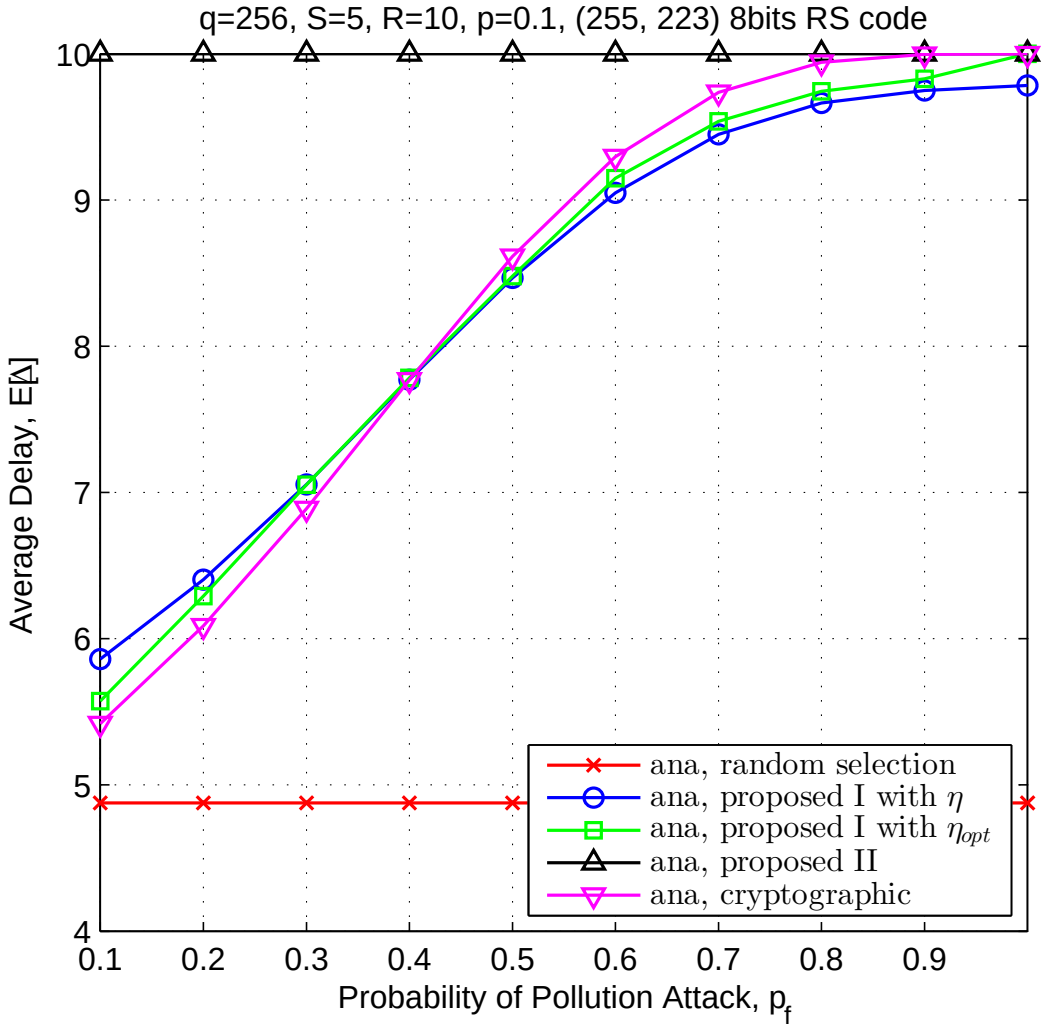

Figure 3.16 The average delay $E[\Delta]$ versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 255, K = 223$.

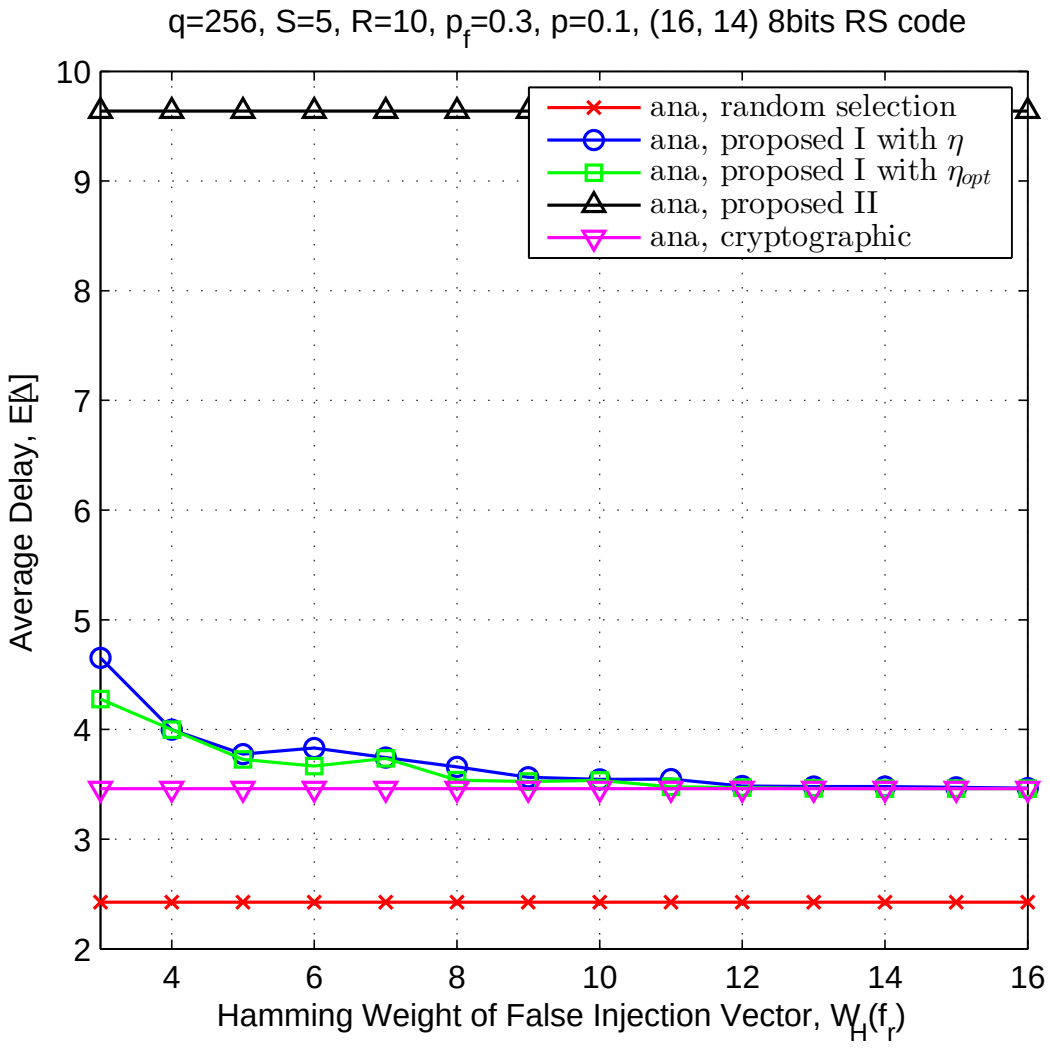

Figure 3.17 The average delay $E[\Delta]$ versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

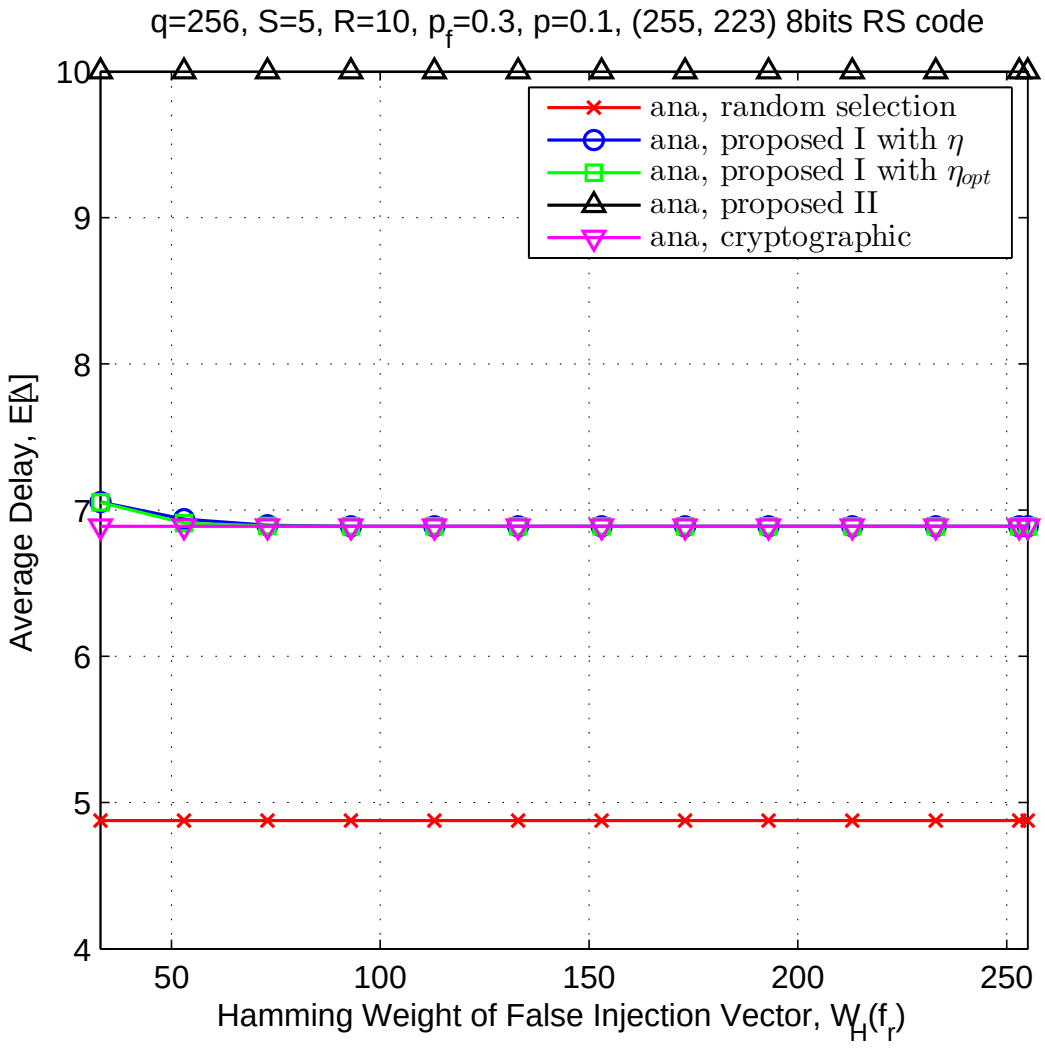

Figure 3.18 The average delay $E[\Delta]$ versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

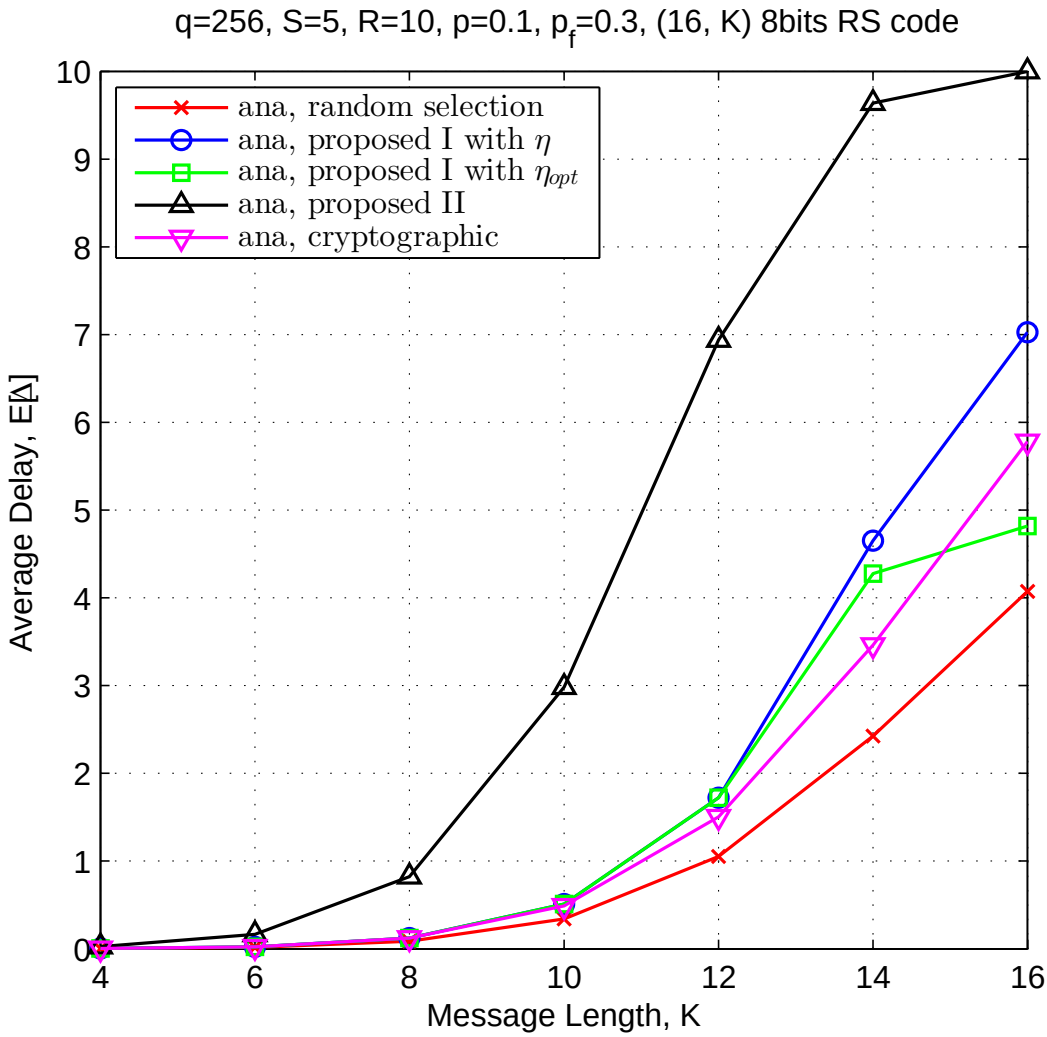

Figure 3.19 The average delay $E[\Delta]$ versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16$.

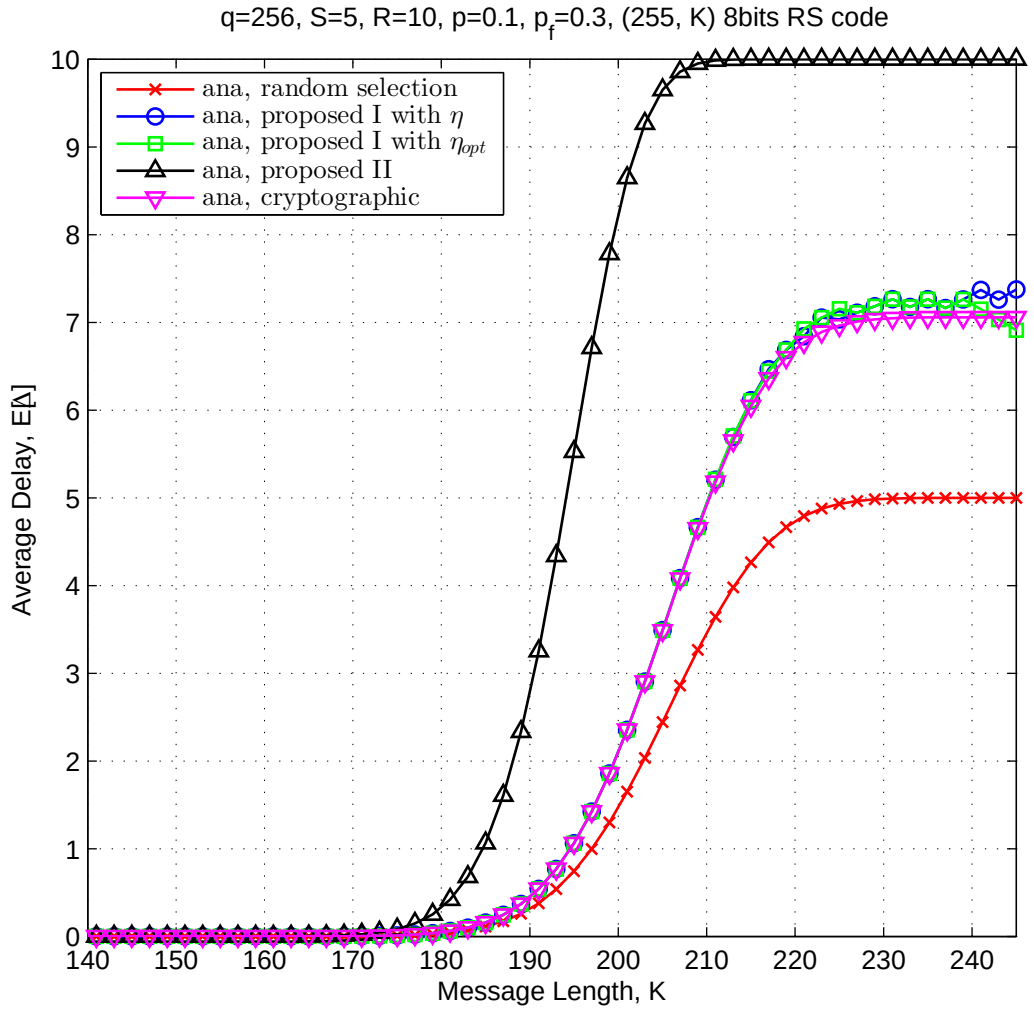

Figure 3.20 The average delay $E[\Delta]$ versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255$.

3.7 Average Throughput

We define the throughput as the number of successfully recovered message packets at the destination per packet transmission. Then, the average throughput is given by

$$W = P(\Lambda = 0) + \sum_{\lambda=1}^{S} \sum_{\delta=\lambda}^{R} \frac{(S-\lambda) + \lambda P_S(\delta, \lambda)}{S+\delta} P(\Delta = \delta | \Lambda = \lambda) P(\Lambda = \lambda) \tag{3.126}$$

for $S \leq R$ and

$$W = P(\Lambda = 0) + \sum_{\lambda=1}^{R} \sum_{\delta=\lambda}^{R} \frac{(S-\lambda) + \lambda P_S(\delta, \lambda)}{S+\delta} P(\Delta = \delta | \Lambda = \lambda) P(\Lambda = \lambda) + \sum_{\lambda=R+1}^{S} \frac{S-\lambda}{S+R} P(\Lambda = \lambda) \tag{3.127}$$

for $S > R$, where $P_S(\delta, \lambda)$ denotes the probability of decoding success given $\Delta = \delta$ and $\Lambda = \lambda$. We used $P(\Delta = \delta | \Lambda = \lambda)$'s derived for each scheme in (3.105), (3.106), (3.110), (3.111), and (3.113) in Section 3.6.

3.7.1 Random Selection

Since random selection scheme chooses the first received Λ coded packets,

$$P_S(\delta, \lambda) = \begin{cases} (1-p_f)^{\lambda}, & \text{if } \lambda = \delta \\ 0, & \text{otherwise} \end{cases} \tag{3.128}$$

Therefore, by (3.105), (3.106), and (3.128), average throughput is given by

$$W = \begin{cases} \sum_{\lambda=0}^{S} \frac{S-\lambda+\lambda(1-p_f)^{\lambda}}{S+\lambda} P(\Lambda = \lambda), & \text{if } S \leq R \\ \sum_{\lambda=0}^{R} \frac{S-\lambda+\lambda(1-p_f)^{\lambda}}{S+\lambda} P(\Lambda = \lambda) + \sum_{\lambda=R+1}^{S} \frac{S-\lambda}{S+R} P(\Lambda = \lambda), & \text{if } S > R \end{cases} \tag{3.129}$$

3.7.2 Cryptographic Scheme

Since the cryptographic scheme is assumed to perfectly detect polluted coded packets, decoding is successful when the λth unpolluted coded packet is received at the destination. Therefore, the probability of decoding success given $\Delta = \delta$ and $\Lambda = \lambda$ is represented

as

$$P_S(\delta,\lambda)=\begin{cases}1, & \text{if } \lambda=\delta=0\\ 1, & \text{if } 1\le\lambda\le\delta\\ 0, & \text{otherwise}\end{cases} \tag{3.130}$$

for $\delta < R$. If $\delta = R$, the joint probability that decoding is successful and $\Delta = R$ given $\Lambda = \lambda$ is given by

$$P_S(R,\lambda)P(\Delta=R|\Lambda=\lambda)=\binom{R-1}{R-\lambda}p_f^{R-\lambda}(1-p_f)^{\lambda} \tag{3.131}$$

which denotes the probability that the Rth coded packet is the λth unpolluted coded packet, given $\Lambda = \lambda$. And $P(\Delta = R|\Lambda = \lambda)$ is given by (3.111).

If $S < R$, following from (3.126), we obtain

$$W=P(\Lambda=0)+\sum_{\lambda=1}^{S}\left\{\sum_{\delta=\lambda}^{R-1}\frac{S}{S+\delta}\overbrace{P(\Delta=\delta|\Lambda=\lambda)}^{(3.110)}\right.$$
$$\left.+\frac{(S-\lambda)\overbrace{P(\Delta=R|\Lambda=\lambda)}^{(3.111)}+\lambda\overbrace{P_S(R,\lambda)P(\Delta=R|\Lambda=\lambda)}^{(3.131)}}{S+R}\right\}P(\Lambda=\lambda). \tag{3.132}$$

Similarly, if $S = R$,

$$W=P(\Lambda=0)+\sum_{\lambda=1}^{R-1}\left\{\sum_{\delta=\lambda}^{R-1}\frac{S}{S+\delta}P(\Delta=\delta|\Lambda=\lambda)\right.$$
$$\left.+\frac{(S-\lambda)P(\Delta=R|\Lambda=\lambda)+\lambda P_S(R,\lambda)P(\Delta=R|\Lambda=\lambda)}{S+R}\right\}P(\Lambda=\lambda)$$
$$+\frac{R(1-p_f)^R}{S+R}P(\Lambda=R). \tag{3.133}$$

Similarly, if $S > R$,

$$W=P(\Lambda=0)+\sum_{\lambda=1}^{R-1}\left\{\sum_{\delta=\lambda}^{R-1}\frac{S}{S+\delta}P(\Delta=\delta|\Lambda=\lambda)\right.$$
$$\left.+\frac{(S-\lambda)P(\Delta=R|\Lambda=\lambda)+\lambda P_S(R,\lambda)P(\Delta=R|\Lambda=\lambda)}{S+R}\right\}P(\Lambda=\lambda)$$
$$+\frac{(S-R)+R(1-p_f)^R}{S+R}P(\Lambda=R)+\sum_{\lambda=R+1}^{S}\frac{S-\lambda}{S+R}P(\Lambda=\lambda). \tag{3.134}$$

3.7.3 Scheme I

Following from (3.126), if $S \leq R$, average throughput is given by

$$W = P(\Lambda = 0) + \sum_{\lambda=1}^{S} \sum_{g=0}^{N} \sum_{\delta=\lambda}^{R} \frac{1}{S+\delta} \Bigg\{ (S-\lambda) \underbrace{P(\Delta = \delta | G = g, \Lambda = \lambda)}_{(3.115)} + \\ + \lambda \underbrace{P_S(\delta, g, \lambda) P(\Delta = \delta | G = g, \Lambda = \lambda)}_{(3.136)} \Bigg\} \underbrace{P(G = g | \Lambda = \lambda)}_{(3.47)} P(\Lambda = \lambda) \tag{3.135}$$

where $P_S(\delta, g, \lambda)$ denotes the conditional probability of decoding success given $\Delta = \delta$, $G = g$, $\Lambda = \lambda$ and

$$P_S(\delta, g, \lambda) P(\Delta = \delta | G = g, \Lambda = \lambda) \\ = \begin{cases} \binom{\delta - 1}{\delta - \lambda} \underbrace{P(\hat{H}_1 | G = g)}_{(3.52)}^{\delta - \lambda} \underbrace{P(\hat{H}_0, H_0 | G = g)}_{(3.63)}^{\lambda}, & \text{if } 1 \leq \lambda \leq \delta \\ 0, & \text{otherwise} \end{cases} \tag{3.136}$$

is the joint conditional probability of decoding success and $\Delta = \delta$ given $G = g$, $\Lambda = \lambda$, based on (3.51).

Similarly, if $S > R$, we obtain

$$W = P(\Lambda = 0) + \sum_{\lambda=1}^{R} \sum_{g=0}^{N} \sum_{\delta=\lambda}^{R} \frac{1}{S+\delta} \Bigg\{ (S-\lambda) \underbrace{P(\Delta = \delta | G = g, \Lambda = \lambda)}_{(3.115)} + \\ + \lambda \underbrace{P_S(\delta, g, \lambda) P(\Delta = \delta | G = g, \Lambda = \lambda)}_{(3.136)} \Bigg\} \underbrace{P(G = g | \Lambda = \lambda)}_{(3.47)} P(\Lambda = \lambda) \\ + \sum_{\lambda=R+1}^{S} \frac{S-\lambda}{S+R} P(\Lambda = \lambda). \tag{3.137}$$

3.7.4 Scheme II

The scheme II always has delay of R in phase 2. i.e.,

$$P(\Delta=\delta|\Lambda=\lambda)=\begin{cases}1, \text{ if } \lambda=0 \text{ and } \delta=0\\ 1, \text{ if } \lambda\geq 1 \text{ and } \delta=R\\ 0, \text{ otherwise}\end{cases} \tag{3.138}$$

By (3.126) and (3.138), If $S\leq R$, the average throughput is given by

$$W=P(\Lambda=0)+\sum_{r_f=0}^{R}\sum_{\lambda=1}^{S}Z(\lambda,r_f)\underbrace{P(R_f=r_f)}_{(3.41)}P(\Lambda=\lambda) \tag{3.139}$$

where

$$Z(\lambda,r_f)=\frac{S-\lambda}{S+R} \tag{3.140}$$

if $\lambda>\min(S,R-r_f)$, and

$$\begin{aligned}Z(\lambda,r_f)=\sum_{g=0}^{N}\sum_{\gamma=0}^{r_f}\frac{1}{S+R}\left\{S-\lambda+\lambda\frac{\binom{R-r_f}{\lambda}}{\binom{R-r_f+\gamma}{\lambda}}\right\}\\ \cdot\binom{r_f}{\gamma}\underbrace{P_{OL}(g)}_{(3.72)}{}^{\gamma}(1-P_{OL}(g))^{r_f-\gamma}\underbrace{P(G=g|\Lambda=\lambda)}_{(3.47)}\end{aligned} \tag{3.141}$$

if $\lambda\leq\min(S,R-r_f)$, based on (3.73).

3.7.5 Asymptotic Analysis for Large N

In this subsection, we provide the average throughput when N is large. If $p\leq\frac{t}{N}$, all S message packets are correctly channel-decoded in phase 1 by (3.76) thus

$$\lim_{N,K\to\infty}W=1, \text{ if } p\leq\frac{t}{N} \tag{3.142}$$

regardless of the used scheme during phase 2. Therefore, we derive the average throughput for large N when $p>\frac{t}{N}$, in the remainder of this subsection.

3.7.5.1 Random Selection

Following from (3.76) and (3.129) we obtain

$$\lim_{N,K\to\infty} W = \begin{cases} 0, & \text{if } S > R \\ \frac{(1-p_f)^S}{2}, & \text{if } S \leq R \end{cases} \tag{3.143}$$

3.7.5.2 Cryptographic Scheme

Following from (3.76) and (3.132)-(3.134) we obtain

$$\lim_{N,K\to\infty} W = \begin{cases} 0, & \text{if } S > R \\ \sum_{\delta=S}^{R} \frac{S}{S+\delta} \underbrace{\binom{\delta-1}{\delta-S} p_f^{\delta-S}(1-p_f)^S}_{(a)}, & \text{if } S \leq R \end{cases} \tag{3.144}$$

where (a) denotes the probability that the δth coded packet is the Sth unpolluted coded packet.

3.7.5.3 Scheme I

From (3.76) and (3.135)-(3.137), the average throughput of the scheme I is simplified to

$$\lim_{N,K\to\infty} W = \begin{cases} 0, & \text{if } S > R \\ \sum_{g=Np}^{\min(N,SNp)} \sum_{\delta=S}^{R} \frac{S}{S+\delta} \underbrace{P_S(\delta,g,S)P(\Delta=\delta|G=g,\Lambda=S)}_{(3.136)} \underbrace{P(G=g|\Lambda=S)}_{(3.82)}, & \text{if } S \leq R \end{cases} \tag{3.145}$$

3.7.5.4 Scheme II

From (3.76) and (3.139)-(3.141), the average throughput of the scheme II is simplified as

$$\lim_{N,K\to\infty} W = \sum_{r_f=0}^{R} Z(S,r_f) \underbrace{P(R_f=r_f)}_{(3.41)} \tag{3.146}$$

where

$$Z(S, r_f)$$
$$= \begin{cases} 0, & \text{if } S > R - r_f \\ \sum_{g=Np}^{\min(N,SNp)} \sum_{\gamma=0}^{r_f} \frac{S}{S+R} \frac{\binom{R-r_f}{S}}{\binom{R-r_f+\gamma}{S}} \binom{r_f}{\gamma} \underbrace{P_{OL}(g)}_{(3.72)}^{\gamma} (1 - P_{OL}(g))^{r_f-\gamma} \underbrace{P(G = g|\Lambda = S)}_{(3.82)}, & \text{if } S \leq R - r_f \end{cases} \tag{3.147}$$

3.7.6 Numerical Results

Fig. 3.21 shows the plot of the average throughput W versus the symbol error probability p for the case of $q = 256, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code, while Fig. 3.22 is for $(N = 255, K = 223)$ 8bits RS code. In both figures, we can see that throughput of the scheme I is higher than that of the scheme II. This is because the scheme II requires that the destination receives all R coded packets before reconstructing message packets, while the scheme I starts reconstructing upon receiving Λ coded packets which are linearly independent and decided to be unpolluted. Therefore, the scheme I has shorter delay than the scheme II, so that the scheme I has better throughput than the scheme II. We also can see that throughput of all schemes decrease as p increases. This is intuitively because larger p causes larger probability of decoding error P_E and/or larger average delay $E[\Delta]$, as we saw in Fig. 3.5, 3.6, 3.13, and 3.14.

Fig. 3.23 shows the plot of the average throughput W versus the probability of pollution attack p_f for the case of $q = 256, p = 0.1, S = 5, R = 10$, and $(N = 16, K = 14)$ 8bits shortened RS code, while Fig. 3.24 shows the plot for $(N = 255, K = 223)$ shortened RS code. We can see that the throughput decrease with increasing p_f. This is because the probability of decoding error P_E and average delay $E[\delta]$ decreases with increasing p_f, as shown in Fig. 3.7, 3.8, 3.15, and 3.16. We also can see that the scheme I has better throughput than the scheme II for the smaller p_f, due to the smaller $E[\Delta]$

of the scheme I. We also can see that throughput of each scheme is generally higher in Fig. 3.23 than Fig. 3.24. This is intuitively because $p = 0.1$ causes that (255,223) code has more Λ in phase 1 thus more $E[\Delta]$ in phase 2 than (16,14) RS code.

Fig. 3.25 shows the plot of the average throughput W versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$ for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$, and $(N = 16, K = 14)$ 8bits shortened RS code, while Fig. 3.26 is for $(N = 255, K = 223)$ RS code. Although we assume that $W_H(\mathbf{f}_r) = d_{min}$ through this chapter, we exceptionally assume that $W_H(\mathbf{f}_r)$ can differ from d_{min} here. In both figures, we can see that the scheme I outperforms the scheme II and approaches the cryptographic scheme, as $W_H(\mathbf{f}_r)$ increases. This is intuitively because larger $W_H(\mathbf{f}_r)$ causes that polluted coded packets are more easily distinguished from unpolluted coded packets, thus the scheme I can early start reconstructing message packets with high detection accuracy, while the scheme II still has disadvantage that all R coded packets are required. We also can see that the throughput of the cryptographic scheme and random selection is not a function of $W_H(\mathbf{f}_r)$, thus has the same value of the throughput for all $W_H(\mathbf{f}_r)$'s.

Fig. 3.27 shows the plot of the average throughput W versus message length K for the case of $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3$, and $N = 16$, while Fig. 3.28 is for $N = 255$. In both of figures, we note that throughput of all schemes decrease as K increases. This is because larger K causes smaller error correction capability t, thus more Λ, larger $E[\Delta]$, finally results in smaller throughput W. For the scheme I and scheme II, this is also because larger K causes smaller d_{min} which is equal to $W_H(\mathbf{f}_r)$. We can see that the scheme I outperforms the scheme II and approaches the cryptographic scheme in the practical range of K (i.e., not very large K). This is because if K is not very large, the scheme I's early recovering property benefits its throughput performance while the scheme II is suffering from the requirement to receive all R coded packets. If K is very large, d_{min} gets smaller, so it becomes harder to detect polluted coded packets. As a result, the early recovering property does not benefit the scheme I's throughput.

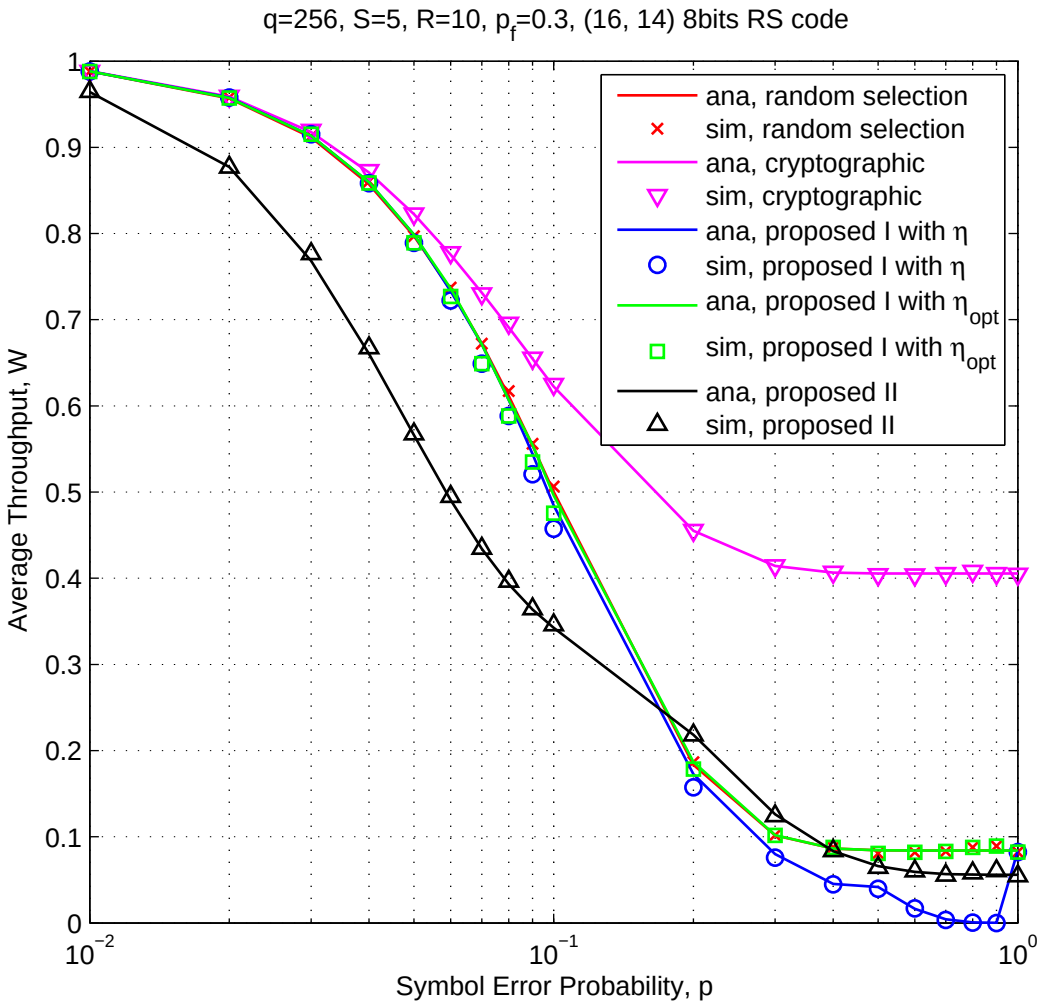

Figure 3.21 The average throughput W versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

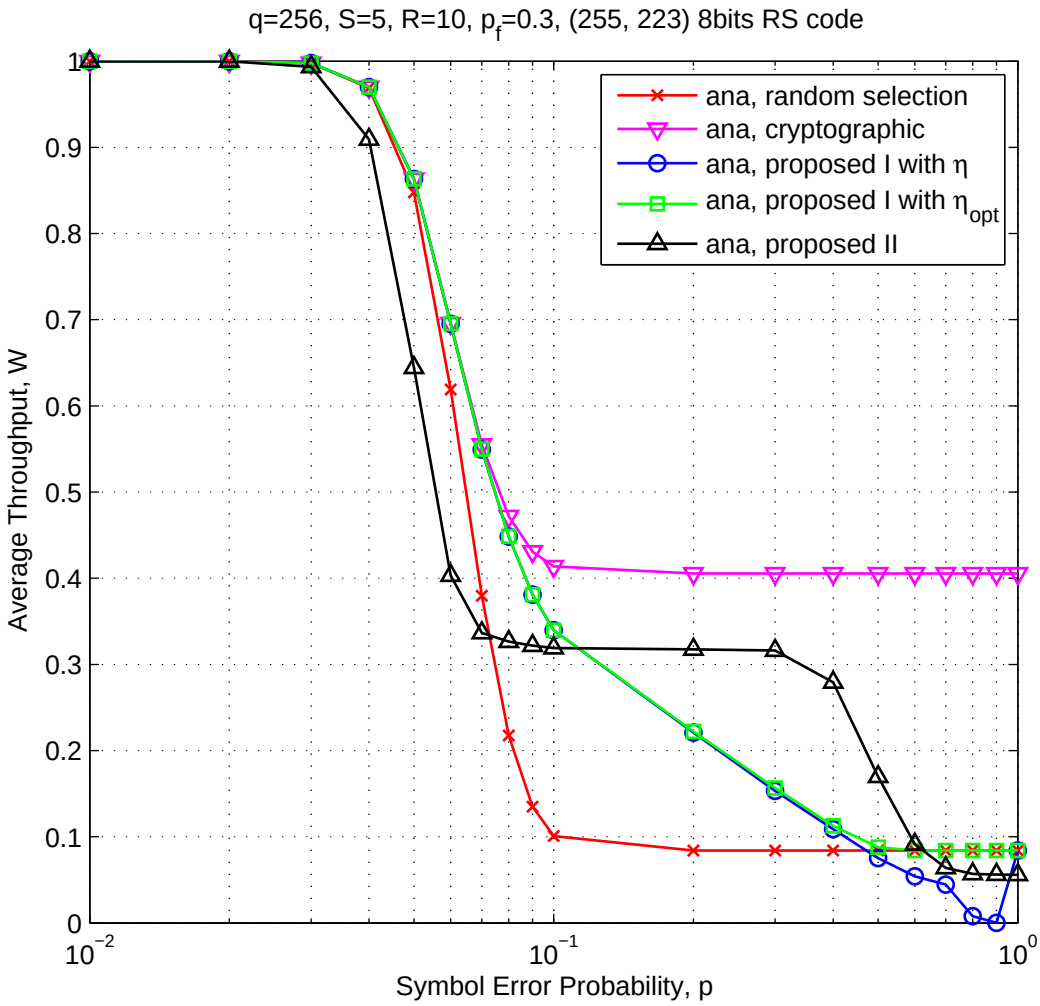

Figure 3.22 The average throughput W versus the symbol error probability p; $q = 256, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

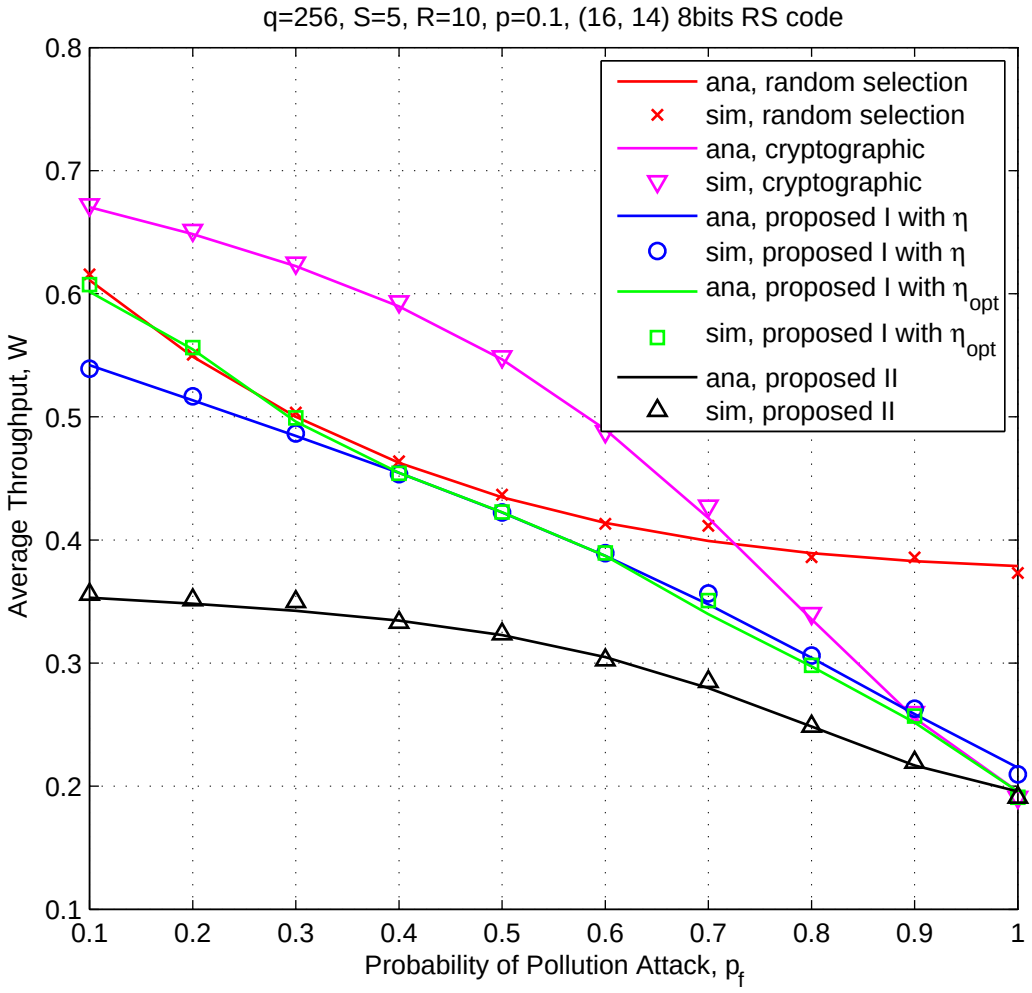

Figure 3.23 The average throughput W versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 16, K = 14$.

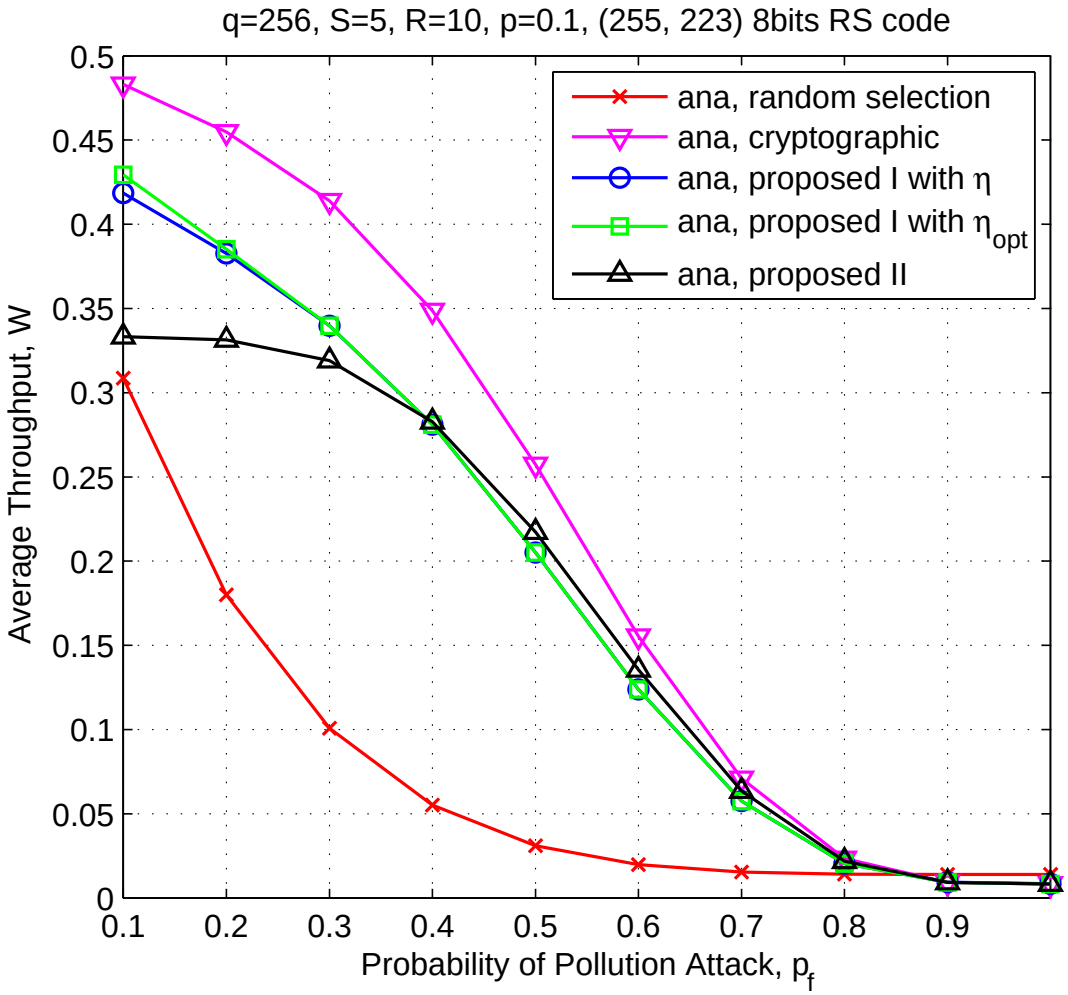

Figure 3.24 The average throughput W versus the probability of pollution attack p_f; $q = 256, p = 0.1, S = 5, R = 10, N = 255, K = 223$.

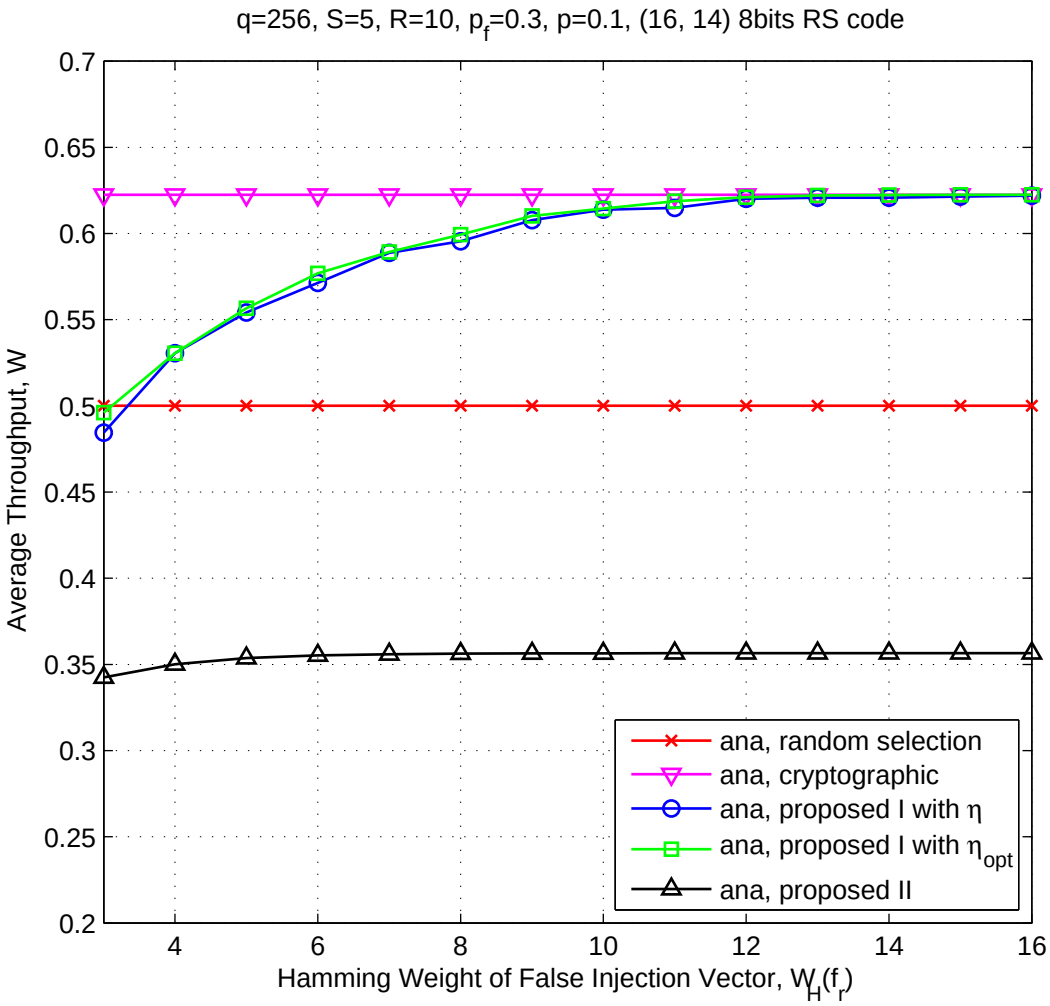

Figure 3.25 The average throughput W versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16, K = 14$.

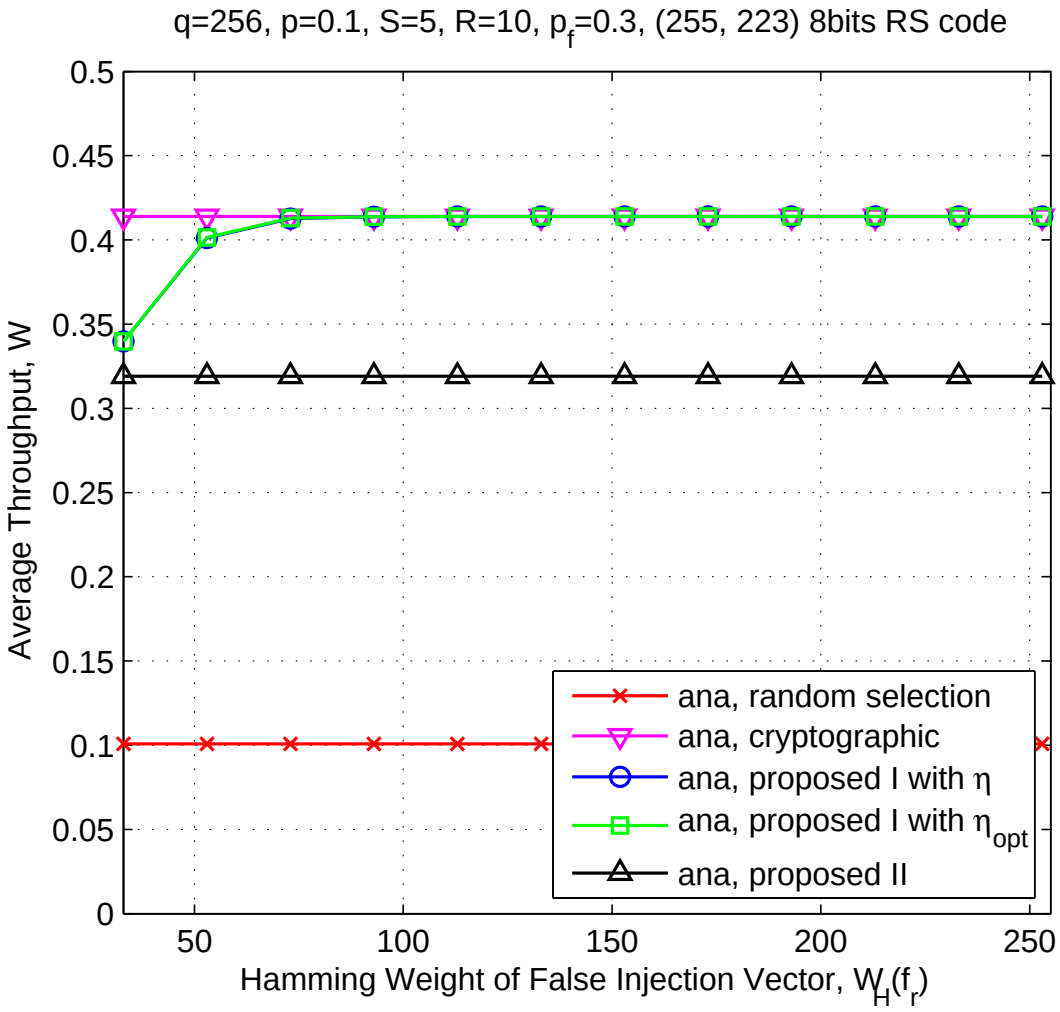

Figure 3.26 The average throughput W versus Hamming weight of false injection vector $W_H(\mathbf{f}_r)$; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255, K = 223$.

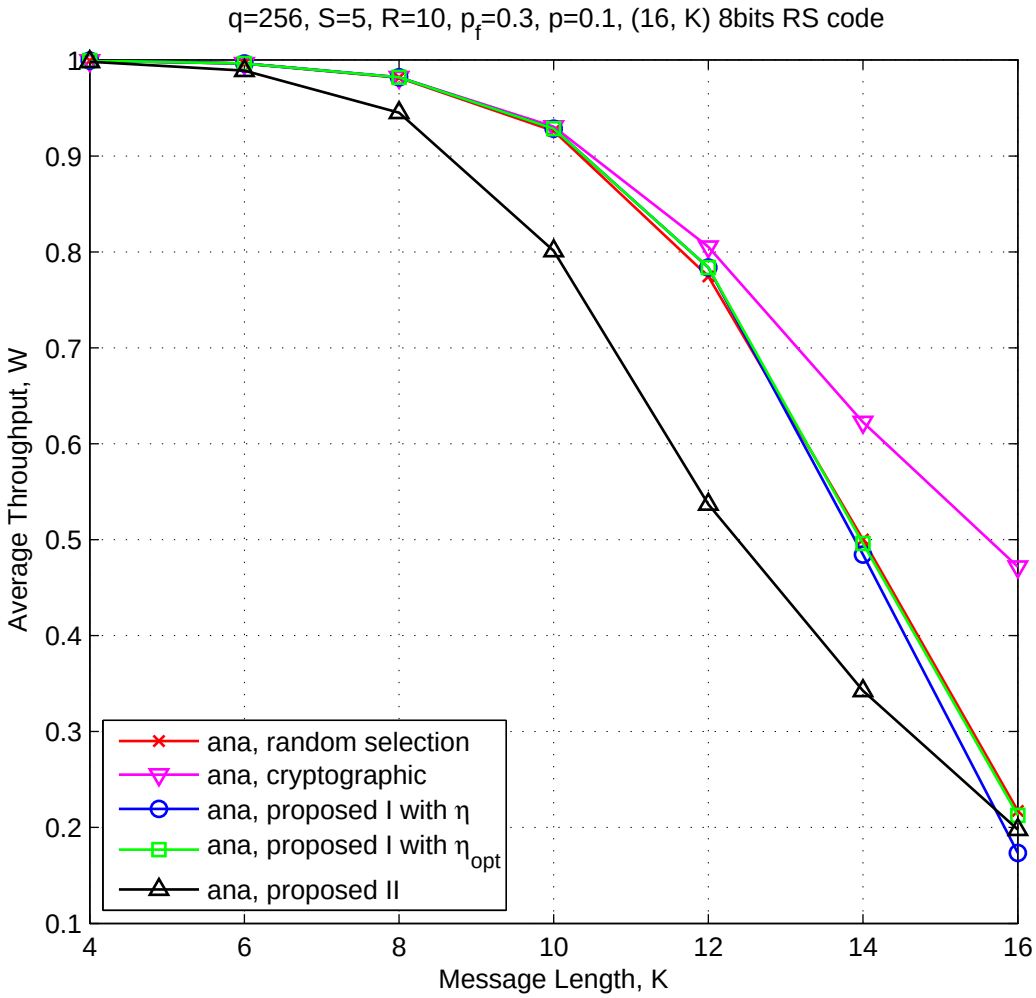

Figure 3.27 The average throughput W versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 16$.

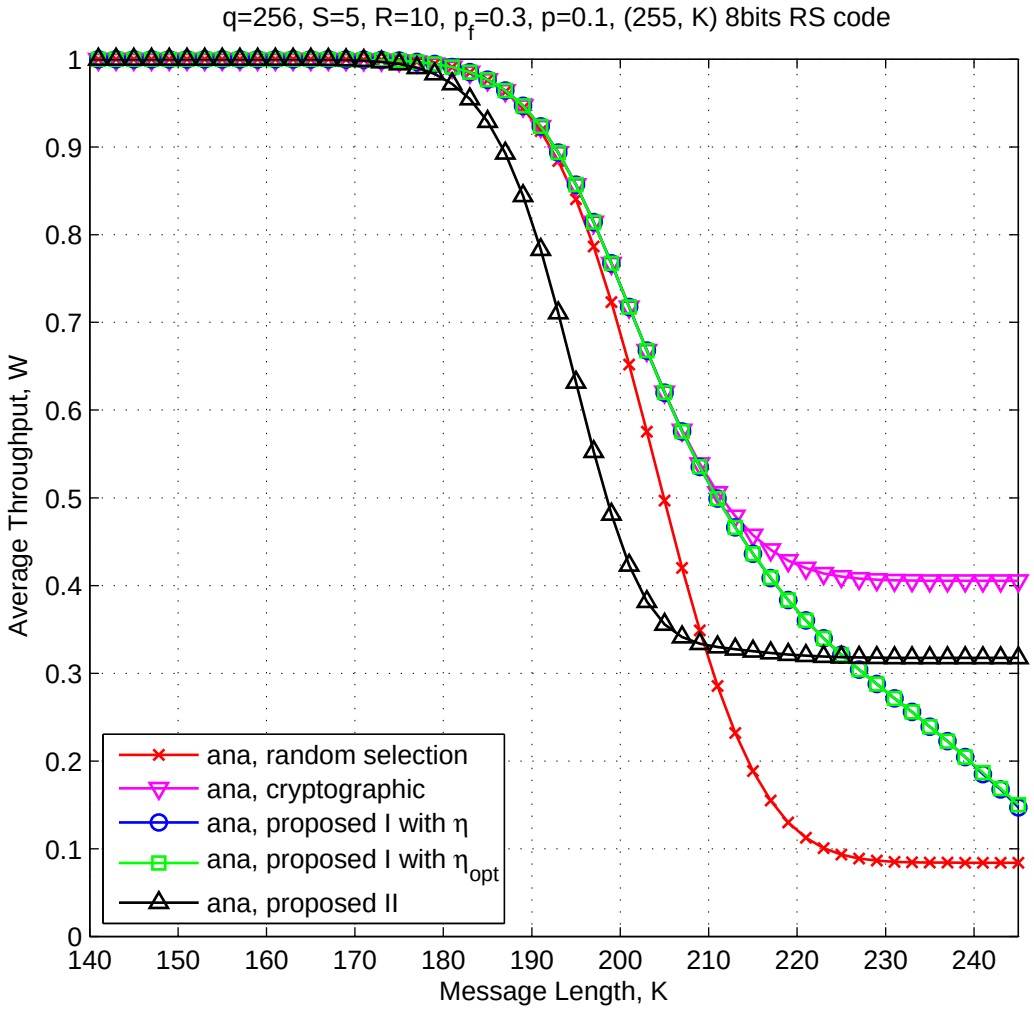

Figure 3.28 The average throughput W versus message length K; $q = 256, p = 0.1, S = 5, R = 10, p_f = 0.3, N = 255$.

3.8 Conclusion

In this chapter, we proposed two physical-layer approaches to exclude polluted packets from decoding at the destination by utilizing noisy information overheard from sources to the destination, on the two-hop wireless networks with multiple sources, single relay, and single destination. We provide the analytical derivation and numerical results for the probability of decoding error, the average delay, and the average throughput of the proposed schemes, the random selection scheme, and the cryptographic scheme. In the practical level of source-to-destination wireless channel error, the probability of decoding error of the proposed schemes are close to that of the cryptographic scheme which perfectly detects polluted packets and much better than that of the random selection scheme which does not detect pollution attacks. While the proposed scheme II has the lower probability of decoding error than the proposed scheme I at the cost of longer delay, the proposed scheme I has the higher average throughput.

CHAPTER 4. CONCLUSIONS AND FUTURE WORKS

In this dissertation, we investigated two problems in order to improve the integrity of random network coded information under pollution attack, on wireless two-hop relay networks.

The first problem is about how the probability of symbol error and the throughput are influenced by the finite field size or the number of combined packets when received data might be polluted by malicious adversaries. Chapter 2 presented that there is an optimized field size to minimize the probability of symbol error or to maximize the throughput. It is also shown that the optimized finite field size to minimize the probability of symbol error decreases with the decreasing trustworthiness of node. Based on this result, if the trustworthiness of node is small, it is proposed to apply a smaller finite field size. It is also shown that the probability of correct decoding exponentially decreases as the number of combined packets increases and that the decaying rate is faster when the trustworthiness of node is smaller. This work can be further extended as follows.

- If channel coding is applied for each packet transmission, the throughput and the probability of symbol error can be improved because some channel error can be corrected by the error correction capability of channel coding. This might cause the different value of the optimized field size to minimize the probability of symbol error or to maximize the throughput. Therefore, analysis with channel coding would be an interesting future work.

- This chapter assumes that the finite field size and QAM constellation size are the

same as q. If we let M be the constellation size of QAM, assuming $q > M$ or $q < M$ could result in the difference for the probability of symbol error, the throughput, and the optimized q.

The second problem is about how to reduce the damage from pollution attack by detecting the polluted packets at the destination node. Chapter 3 suggested two approaches which detect and discard polluted packets at the destination node by exploiting overheard noisy information. The main advantage of proposed schemes is that cryptography-based signature is not required. For the practical (i.e., not very large) values of source-to-destination symbol error probability p, the probability of decoding error of the proposed schemes approaches that of the cryptographic scheme which is assumed to perfectly detect all polluted packets. Analysis of the average throughput shows that the proposed scheme I has higher throughput than the proposed scheme II in the practical range of p because the scheme I can start reconstructing message packets earlier than the scheme II which requires all coded packets in phase 2. Future research directions for this work are as follows.

- Let a group be a set of multiple received coded packets at the destination. Let us say that a group is polluted if at least one polluted coded packet is included in the group. Then, the destination detects the presence of pollution attack for each group (not for each coded packet), by calculating

$$W_H\left(\sum_{i\in\mathbf{I}}\mathbf{z}_i\right) \tag{4.1}$$

  where $\mathbf{I}$ denotes the set of coded packet indices in the group and $\mathbf{z}_i$ denotes a sufficient statistic of the packet by (3.15). If a group is detected as unpolluted, all coded packets in the group are regarded as unpolluted without detecting each packet in one by one manner. If the group is detected as polluted, there could be several further options as follows.

- All coded packets in the group are regarded as polluted, thus all of them are discarded.
- Detect each packet in the group one by one.
- Divide the group into multiple subgroups and detect if each subgroup is polluted or not.

APPENDIX A. PROOF OF (2.24) AND (2.25)

We prove (2.24) and (2.25). Let

$$w_s = \sum_{i=1}^{S} \tilde{c}_{s,i} u_{[i]}. \tag{A.1}$$

It follows from (2.15) and (2.20) that

$$P(\hat{x}_s = x_s | \mathbf{u} = \mathbf{0}) = P(v_s + w_s = 0 | \mathbf{u} = \mathbf{0}) \tag{A.2}$$

$$= P(v_s = 0 | \mathbf{u} = \mathbf{0}) \tag{A.3}$$

$$= P(v_s = 0) \tag{A.4}$$

$$= P(f_s = 0) P(e_{s,[s]} = 0) + \frac{(1 - P(f_s = 0))(1 - P(e_{s,[s]} = 0))}{q - 1} \tag{A.5}$$

where the last equality follows from the assumption that nonzero values of f_s are equiprobable. This proves (2.24).

Similarly,

$$P(\hat{x}_s = x_s | \mathbf{u} \neq \mathbf{0})$$

$$= P(v_s + w_s = 0 | \mathbf{u} \neq \mathbf{0}) \tag{A.6}$$

$$= P(v_s = 0 | \mathbf{u} \neq \mathbf{0}) P(w_s = 0 | \mathbf{u} \neq \mathbf{0}) + \frac{P(v_s \neq 0 | \mathbf{u} \neq \mathbf{0}) P(w_s \neq 0 | \mathbf{u} \neq \mathbf{0})}{q - 1} \tag{A.7}$$

$$= P(v_s = 0) P(w_s = 0 | \mathbf{u} \neq \mathbf{0}) + \frac{P(v_s \neq 0) P(w_s \neq 0 | \mathbf{u} \neq \mathbf{0})}{q - 1} \tag{A.8}$$

Let

$$\tilde{\mathbf{c}}_s = \begin{bmatrix} \tilde{c}_{s,1} \\ \vdots \\ \tilde{c}_{s,S} \end{bmatrix}. \tag{A.9}$$

Then, the set of $\tilde{\mathbf{c}}_s$'s that satisfy

$$w_s = \mathbf{u} \cdot \tilde{\mathbf{c}}_s = 0 \tag{A.10}$$

is the null space of $\mathbf{u}$. Since $\mathbf{u}$ is not a zero vector, the rank of $\mathbf{u}$ is one. By the rank theorem [14], the dimension of the null space of $\mathbf{u}$ is equal to $S - \text{rank}(\mathbf{u}) = S - 1$. Hence, the number of $\tilde{\mathbf{c}}_s$'s that cause $\mathbf{u} \cdot \tilde{\mathbf{c}}_s = 0$ is q^{S-1}. Therefore,

$$\begin{aligned}
&P(\mathbf{u} \cdot \tilde{\mathbf{c}}_s = 0 | \mathbf{u} \neq \mathbf{0}) \\
&= \frac{\text{The number of nonzero } \tilde{\mathbf{c}}_s\text{'s that satisfy } \mathbf{u} \cdot \tilde{\mathbf{c}}_s = 0 \text{ given that } \mathbf{u} \neq \mathbf{0}}{\text{The number of nonzero } \tilde{\mathbf{c}}_s\text{'s}} && \text{(A.11)} \\
&= \frac{q^{S-1} - 1}{q^S - 1} && \text{(A.12)} \\
&\approx \frac{1}{q}. && \text{(A.13)}
\end{aligned}$$

Therefore, it follows from (A.8) and (A.12) we obtain

$$P(\hat{x}_s = x_s | \mathbf{u} \neq \mathbf{0}) = \frac{q^{S-1} - P(v_s = 0)}{q^S - 1}. \tag{A.14}$$

This proves (2.25).

APPENDIX B. PROOF OF (2.49)

It follows from (2.24) and (2.28) that if $\alpha = 1$, i.e., $p_f + p_e = \frac{qp_fp_e}{q-1}$, then $P_C = 1$. Hence,

$$\lim_{S,N\to\infty} W = \frac{\beta \log_2 q}{1+\beta}. \tag{B.1}$$

If $\alpha < 1$, i.e., $p_f + p_e > \frac{qp_fp_e}{q-1}$, then $\lim_{S,N\to\infty} P_C = \frac{1}{q}$. Hence,

$$\lim_{S,N\to\infty} W = \frac{\beta \log_2 q}{(1+\beta)q}. \tag{B.2}$$

This completes the proof of (2.49).

APPENDIX C. PROOF OF (3.14)

In this Appendix we prove that (3.14) is equivalent to $\mathbf{f}_{[1]} = \mathbf{0}, \cdots, \mathbf{f}_{[\Lambda]} = \mathbf{0}$. First of all, (3.14) is equivalent to

$$\underbrace{\begin{bmatrix} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(S-\Lambda)} \\ \hline \mathbf{c}_{[1]} \\ \vdots \\ \mathbf{c}_{[\Lambda]} \end{bmatrix}^{-1}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \hline \mathbf{f}_{[1]} \\ \vdots \\ \mathbf{f}_{[\Lambda]} \end{bmatrix}}_{\mathbf{F}} = \begin{bmatrix} \mathbf{0} \\ \\ \vdots \\ \\ \mathbf{0} \end{bmatrix} \tag{C.1}$$

by (3.10). If we let $\mathbf{j}_n$ be the nth column vector of $\mathbf{F}$, the set of $\mathbf{j}_n$ satisfying

$$\mathbf{C}^{-1}\mathbf{j}_n = \left.\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}\right\} S \text{ zeros} \tag{C.2}$$

is null space of $\mathbf{C}^{-1}$ where the right hand side of (C.2) is a column vector of S zeros. Since $\mathbf{C}^{-1}$ is an invertible matrix, its rank is S. By the rank theorem [14], dimension of null space of $\mathbf{C}^{-1}$ is $S - S = 0$. Therefore, the number of $\mathbf{j}_n$ satisfying (C.2) is $q^0 = 1$. Since a zero vector is always included in null space, the only $\mathbf{j}_n$ satisfying (C.2) is a zero vector. Since this holds for any $n \in \{1, \cdots, N\}$, all N column vectors of $\mathbf{F}$ are zero vectors. This is equivalent to $\mathbf{f}_{[1]} = \mathbf{0}, \cdots, \mathbf{f}_{[\Lambda]} = \mathbf{0}$.

APPENDIX D. PROOF OF SUFFICIENT STATISTIC $\mathbf{z}_r$ IN (3.15)

In this Appendix we prove that $\mathbf{z}_r$ is a sufficient statistic, by showing that

$$I(\mathbf{f}_r; \mathbf{r}_1, \cdots, \mathbf{r}_S, \mathbf{p}_r) = I(\mathbf{f}_r; \mathbf{z}_r). \tag{D.1}$$

When $S = 2$,

$$I(\mathbf{f}_r; \mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r) = H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r) - H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r | \mathbf{f}_r) \tag{D.2}$$

where

$$H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r) = H(\mathbf{x}_1 + \mathbf{e}_1, \mathbf{x}_2 + \mathbf{e}_2, c_{r,1}\mathbf{x}_1 + c_{r,2}\mathbf{x}_2 + \mathbf{f}_r) \tag{D.3}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, c_{r,1}\mathbf{x}_1 + c_{r,2}\mathbf{x}_2 + \mathbf{f}_r | \underbrace{\mathbf{x}_1 + \mathbf{e}_1}_{:=\mathbf{u}_1}) \tag{D.4}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, c_{r,1}(\mathbf{u}_1 - \mathbf{e}_1) + c_{r,2}\mathbf{x}_2 + \mathbf{f}_r | \mathbf{u}_1) \tag{D.5}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, -c_{r,1}\mathbf{e}_1 + c_{r,2}\mathbf{x}_2 + \mathbf{f}_r) \tag{D.6}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2) + H(\underbrace{-c_{r,1}\mathbf{e}_1 + c_{r,2}\mathbf{x}_2 + \mathbf{f}_r}_{=-c_{r,1}\mathbf{e}_1 + c_{r,2}(\mathbf{u}_2 - \mathbf{e}_2) + \mathbf{f}_r} | \underbrace{\mathbf{x}_2 + \mathbf{e}_2}_{:=\mathbf{u}_2}) \tag{D.7}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2) + H(-c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2 + \mathbf{f}_r) \tag{D.8}$$

and

$$H(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r|\mathbf{f}_r) = H(\mathbf{x}_1 + \mathbf{e}_1, \mathbf{x}_2 + \mathbf{e}_2, c_{r,1}\mathbf{x}_1 + c_{r,2}\mathbf{x}_2) \tag{D.9}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, \underbrace{c_{r,1}\mathbf{x}_1 + c_{r,2}\mathbf{x}_2}_{=c_{r,1}(\mathbf{u}_1 - \mathbf{e}_1) + c_{r,2}\mathbf{x}_2} | \underbrace{\mathbf{x}_1 + \mathbf{e}_1}_{:=\mathbf{u}_1}) \tag{D.10}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, -c_{r,1}\mathbf{e}_1 + c_{r,2}\mathbf{x}_2) \tag{D.11}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2) + H(\underbrace{-c_{r,1}\mathbf{e}_1 + c_{r,2}\mathbf{x}_2}_{=-c_{r,1}\mathbf{e}_1 + c_{r,2}(\mathbf{u}_2 - \mathbf{e}_2)} | \underbrace{\mathbf{x}_2 + \mathbf{e}_2}_{:=\mathbf{u}_2}) \tag{D.12}$$

$$= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2) + H(-c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2). \tag{D.13}$$

Therefore, by substituting (D.8) and (D.13) into (D.2),

$$I(\mathbf{f}_r; \mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_r) = H(-c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2 + \mathbf{f}_r) - \underbrace{H(-c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2)}_{=H(-c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2 + \mathbf{f}_r|\mathbf{f}_r)} \tag{D.14}$$

$$= I(\mathbf{f}_r; -c_{r,1}\mathbf{e}_1 - c_{r,2}\mathbf{e}_2 + \mathbf{f}_r) \tag{D.15}$$

$$= I(\mathbf{f}_r; \mathbf{z}_r). \tag{D.16}$$

Therefore, (D.1) holds for $S = 2$.

This can be generalized to $S > 2$ sources. In general,

$$I(\mathbf{f}_r; \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r) = H(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r) - H(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r|\mathbf{f}_r) \tag{D.17}$$

where

$$
\begin{aligned}
& H(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r) \\
&= H(\mathbf{x}_1 + \mathbf{e}_1, \cdots, \mathbf{x}_S + \mathbf{e}_S, \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r) && \text{(D.18)} \\
&= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, \cdots, \mathbf{x}_S + \mathbf{e}_S, \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r | \underbrace{\mathbf{x}_1 + \mathbf{e}_1}_{:=\mathbf{u}_1}) && \text{(D.19)} \\
&= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, \cdots, \mathbf{x}_S + \mathbf{e}_S, -c_{r,1}\mathbf{e}_1 + \sum_{s=2}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r) && \text{(D.20)} \\
&\ \ \vdots \\
&= \sum_{s=1}^{S} H(\mathbf{x}_s + \mathbf{e}_s) + H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s + \mathbf{f}_r) && \text{(D.21)}
\end{aligned}
$$

and

$$
\begin{aligned}
& H(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r | \mathbf{f}_r) \\
&= H(\mathbf{x}_1 + \mathbf{e}_1, \cdots, \mathbf{x}_S + \mathbf{e}_S, \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s + \mathbf{f}_r | \mathbf{f}_r) && \text{(D.22)} \\
&= H(\mathbf{x}_1 + \mathbf{e}_1, \cdots, \mathbf{x}_S + \mathbf{e}_S, \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s) && \text{(D.23)} \\
&= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, \cdots, \mathbf{x}_S + \mathbf{e}_S, \sum_{s=1}^{S} c_{r,s}\mathbf{x}_s | \underbrace{\mathbf{x}_1 + \mathbf{e}_1}_{:=\mathbf{u}_1}) && \text{(D.24)} \\
&= H(\mathbf{x}_1 + \mathbf{e}_1) + H(\mathbf{x}_2 + \mathbf{e}_2, \cdots, \mathbf{x}_S + \mathbf{e}_S, -c_{r,1}\mathbf{e}_1 + \sum_{s=2}^{S} c_{r,s}\mathbf{x}_s) && \text{(D.25)} \\
&\ \ \vdots \\
&= \sum_{s=1}^{S} H(\mathbf{x}_s + \mathbf{e}_s) + H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s). && \text{(D.26)}
\end{aligned}
$$

By substituting (D.21) and (D.26) into (D.17), we obtain

$$I(\mathbf{f}_r; \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_S, \mathbf{p}_r) = H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s + \mathbf{f}_r) - H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s) \tag{D.27}$$

$$= H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s + \mathbf{f}_r) - H(-\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s + \mathbf{f}_r|\mathbf{f}_r) \tag{D.28}$$

$$= I(\mathbf{f}_r; -\sum_{s=1}^{S} c_{r,s}\mathbf{e}_s + \mathbf{f}_r) \tag{D.29}$$

$$= I(\mathbf{f}_r; \mathbf{z}_r). \tag{D.30}$$

Therefore, (D.1) holds.

BIBLIOGRAPHY

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204-1216, 2000.

[2] T. Ho, R. Koetter, M. Médard, D. R. Karger and M. Effros, "The benefits of coding over routing in a randomized setting," in *Proc. of International Symposium on Information Theory*, 2003.

[3] D. S. Lun, M. Médard, and M. Effros. "On coding for reliable communication over packet networks," in *Proc. of Allerton Conference on Communication, Control, and Computing*, 2004.

[4] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. of Allerton Conference on Communication, Control, and Computing*, October 2003.

[5] L. Lima, M. Médard, and J. Barros, "Random linear network coding: A free cypher?" in *Proc. of the IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[6] M. Kim, L. Lima, F. Zhao, J. Barros, M. Médard, R. Koetter, T. Kalkert, and K. J. Han, "On counteracting byzantine attacks in network coded peer-to-peer networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 5, pp. 692-702, Jun. 2010.

[7] M. Kim, M. Médard, and J. Barros, "Algebraic watchdog: mitigating misbehavior in wireless network coding," *IEEE Journal on Selected Areas in Communications (JSAC)*, December 2011.

[8] T. Ho, B. Leong, R. Koetter, M. Médard, M. Effros, and D. R. Karger, "Byzantine modification detection in multicast networks with random network coding," *IEEE Trans. Information Theory*, vol. 54, no. 6, pp. 2798-2803, June 2008.

[9] S. Kim and S. W. Kim "Recycling polluted packet at the physical layer in wireless network coding," *IEEE Communications Letters*, Jun. 2013.

[10] K. Han, T. Ho, R. Koetter, M. Médard, and F. Zhao, "On network coding for security," in *Proc. of Milcom*, 2007.

[11] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, and M. Médard, "Resilient network coding in the presence of byzantine adversaries," in *Proc. of IEEE INFOCOM*, 2007.

[12] B. Shrader and A. Ephremides, "On packet lengths and overhead for random linear coding over the erasure channel," in *Proc. Int. Wireless Communications and Mobile Computing Conf.*, Honolulu, HI, Aug. 2007.

[13] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the Society for Industrial and Applied Mathematics,* 8:300-304, 1960.

[14] D. C. Lay, *Linear Algebra and Its Applications*, Addison Wesley, 2nd ed., 1997.

[15] J. Kurose and K. Ross. *Computer Networking: A Top-Down Approach Featuring the Internet*, Addison Wesley, 3rd ed., 2003.

[16] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, Oxford, U.K.: Oxford Univ. Press, 2006.

[17] S. Lin and D. J. Costello, *Error Control Coding.* Upper Saddle River, NJ: Prentice-Hall, 2004.

[18] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products,* 6th ed. Academic Press, 2000.

[19] S. W. Kim, "Securing wireless network coding against pollution attack at the physical layer," in *Proc. of IEEE Milcom*, Orlando, Oct. 2012.

[20] D. H. Yoon and S. W. Kim, "Field size of random network coding in untrustworthy networks, in *Proc. of IEEE International Conference on Communications*, June 2013.

[21] D. H. Yoon and S. W. Kim, "Trustworthy decoding of random network coded packets under pollution attack: physical-layer approach," in *Proc. of IEEE International Symposium on Network Coding*, June 2013.

[22] A. Goldsmith, *Wireless Communications,* Cambridge University Press, p.190, 2005.

[23] S. Wicker, *Error Control Systems for Digital Communications and Storage,* Prentice-Hall, Englewood Cliffs, NJ, 1995.

[24] C. Cheng and T. Jiang, "An efficient homomorphic mac with small key size for authentication in network coding," *IEEE Transactions on Computers,* 2013. Volume:62, Issue: 10, 2096-2100.

[25] C. Cheng, T. Jiang, and Q. Zhang, "TESLA-based homomorphic mac for authentication in P2P system for live streaming with network coding," *IEEE Journal on Selected Areas in Communications (JSAC),* September 2013.

[26] C. Cheng and T. Jiang, "A novel homomorphic MAC scheme for authentication in network coding," *IEEE Communications Letters*, November 2011.

[27] S. Agrawal and D. Boneh, "Homomorphic MACs: MAC-based integrity for network coding," in *Proc. of Applied Cryptography Netw. Security,* 2009, pp. 292-305.

[28] T. T. Tran, H. Li, G. Ru, R. J. Kerczewski, L. Liu, and S. U. Khan, "Secure wireless multicast for delay-sensitive data via network coding," *IEEE Transactions on Wireless Communications,* Vol. 12, No. 7, pp. 3372-3387, July 2013.

[29] A. Le and A. Markopoulou, "Cooperative Defense Against Pollution Attacks in Network Coding Using SpaceMac," *IEEE Journal on Selected Areas in Communications (JSAC),* vol. 30, no. 2, Feb. 2012.

[30] C. Gkantsidis and P. Rodriguez, "Cooperative security for network coding file distribution," *in Proc. of IEEE INFOCOM*, 2006, pp. 113.

[31] Z. Yu, Y. Wei, B. Ramkumar, and Y. Guan, "An efficient signature based scheme for securing network coding against pollution attacks," *in Proc. of IEEE INFOCOM*, 2008, pp. 14091417.

[32] Z. Yu, Y. Wei, B. Ramkumar, and Y. Guan, "An efficient scheme for securing XOR network coding against pollution attacks," *in Proc. of IEEE INFOCOM*, 2009, pp. 406414.

[33] F. Oggier and H. Fathi, "An authentication code against pollution attacks in network coding," *IEEE/ACM Transactions on Networking*, Issue 99, March 2011. CoRR abs/0909.3146, 2009.

[34] Y. Jiang, Y. Fan, X. Shen, and C. Lin, "A self-adaptive probabilistic packet filtering scheme against entropy attacks in network coding," *Computer Networks*, vol. 53, no. 18, pp. 3089-3101, Dec. 2009.

[35] Y. Li, H. Yao, M. Chen, S. Jaggi, and A. Rosen, "RIPPLE authentication for network coding," *in Proc. of IEEE INFOCOM*, Mar. 2010.

[36] E. Kehdi and B. Li, "Null keys: limiting malicious attacks via null space properties of network coding," *in Proc. of IEEE INFOCOM*, Apr. 2009.

[37] A. Newell, R. Curtmola, and C. Nita-Rotaru, "Entropy attacks and countermeasures in wireless network coding," *in Proc. of 2012 ACM Conference on Security and Privacy in Wireless and Mobile Network*.

[38] J. Dong , R. Curtmola , R. Sethi, and C. Nita-Rotaru "Toward secure network coding in wireless networks: threats and challenges," *in Proc. IEEE NPSEC*, pp.33-38 2008