Multiple Access Networks Over Finite Fields - Optimality of Separation, Randomness and Linearity

by

Siddharth Ray

B.Tech, Electrical Engineering (2002)
Indian Institute of Technology Madras

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2003

Author . . . .
Department of Electrical Engineering and Computer Science
August 28, 2003

Certified by . . . .
Muriel Médard
Esther and Harold E. Edgerton Associate Professor
Thesis Supervisor

Certified by . . . .
Jinane Abounadi
Post Doctoral Lecturer
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Multiple Access Networks Over Finite Fields - Optimality of Separation, Randomness and Linearity

by

Siddharth Ray

Submitted to the Department of Electrical Engineering and Computer Science
on August 28, 2003, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

We consider a time-slotted multiple access noisy as well as noise-free channel in which the received, transmit and noise alphabets belong to a finite field. We show that source-channel separation holds when the additive noise is independent of inputs. However, for input-dependent noise, separation may not hold. For channels over the binary field, we derive the expression for the probability of source-channel separation failing. We compute this probability to be 1/4 when the noise parameters are picked independently and uniformly. For binary channels, we derive an upper bound of 0.0776 bit for the maximum loss in sum rate due to separate source-channel coding when separation fails. We prove that the bound is very tight by showing that it is accurate to the second decimal place.

We derive the capacity region and the maximum *code rate* for the noisy as well as noise-free channel where, *code rate* is defined as the ratio of the information symbols recovered at the receiver to the symbols sent by the transmitters in a slot duration. Code rate measures the overhead in transmitting in a slot under multiple access interference. We show for both noisy and noise-free channels that capacity grows logarithmically with the size of the field but the code rate is invariant with field size. For the noise-free channel, codes achieve maximum code rate if and only if they achieve capacity and add no redundancy to the shorter of the two information codewords.

For the noise-free multiple access channel, we consider the cases when both transmitters always transmit in a slot, as well as when each transmitter transmits in a bursty fashion according to a Bernoulli process. For the case when both transmitters always transmit, we propose a systematic code construction and show that it achieves the maximum code rate and capacity. We also propose a systematic random code construction and show that it achieves the maximum code rate and capacity with probability tending to 1 exponentially with codeword length and field size. This is a strong coding theorem for this channel. For the case when transmitters transmit according to a Bernoulli process, we propose a coding scheme to maximize the expected code rate. We show that maximum code rate is achieved by adding redundancy at the less bursty transmitter and not adding any redundancy at the more bursty transmitter.

For the noisy channel, we obtain the error exponents and hence, the expression for average probability of error when a random code is used for communicating over the channel.

Thesis Supervisor: Muriel Médard
Title: Esther and Harold E. Edgerton Associate Professor

Thesis Supervisor: Jinane Abounadi
Title: Post Doctoral Lecturer

Acknowledgments

I thank my advisor Professor Muriel Médard for her patience, guidance and support throughout my past year at MIT. It is a great privilege to work with such a highly talented person who is always very supportive and encouraging. Being her student has greatly helped in my development as a graduate student as well as a person.

I thank Dr. Jinane Abounadi who has been my co-thesis supervisor. Her valuable support and insights have helped me improve my research.

I also thank Professor Michelle Effros of California Institute of Technology for her helpful suggestions that have helped in bettering my understanding of my research area.

Finally, I thank my parents - Mrs. Chitra and Mr. Shankar Ray. For me, they have been representatives of God and whatever I am today, is because of their love, encouragement and sacrifices.

My thanks also go out to all my friends and well wishers all over the world for the good influence they have had on my life.

I acknowledge the following research grants that supported this thesis:

1. Hewlett-Packard, “Advanced Concepts in Wireless Networking”, Wireless Research Project Specification Award 008542-008.

2. NSF/University of Illinois, “ITR/SY: High Speed Wavelength-Agile Optical Network”, Subaward 02-194.

3. NSF Career Sponsor Award CCR - 0093349.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Over the past few years, there has been an interest in indoor wireless communication networks where nodes communicate with each other over short distances. The proximity of the nodes to each other makes multiple access interference from other nodes the chief source of interference rather than noise. A good design criterion for these networks is making the node receiver design invariant to the number of nodes in the network, allowing addition and removal of nodes without need for significant redesign.

Research on noise-free multiple access networks has generally assumed that the received alphabet size grows with the number of transmitters. An example of such a channel is the binary adder multiple access channel, where interference is additive and the received alphabet grows with the number of transmitters. For the case when all transmitters transmit, the capacity is derived and a coding technique is proposed in [1], which achieves capacity asymptotically with the number of users. Capacity achieving codes for this channel have also been proposed in [2, 3, 5]. Another example is the collision channel where a subset of the users transmit at the same time. Mathys in [6] has determined the capacity regions for asynchronous and slot synchronous users, and given constructive codes that approach all rates in these regions. Hughes in [4] has also considered design of codes for this channel. Caire *et.al* [7, 8] presented a novel class of signal space codes for bandwidth efficient transmission on a random access collision channel. All of these considered multiple access interference where the received alphabet size grows with the number of transmitters.

If the received alphabet size is allowed to grow with the number of transmitters, the receiver has more information about the transmissions than when the received alphabet is constrained to the same field as the transmitted alphabet. Rates are lower for a channel where the received alphabet is fixed as the receiver can observe only over the field in which transmissions occur. However, the receiver design becomes independent of the number of users and places the network in a consistent digital framework. In [14], transmission of information is considered for a modulo-2 multiple access channel and a code construction is proposed. However, it considers only the case where a proper subset of the users transmit. The work is confined to $\mathbb{F}_2$ and not finite field channels in general.

1.2 Thesis Outline

We consider a time-slotted additive multiple access network where two transmitters are independently transmitting to a single receiver. The transmitted and received alphabet sizes are the same. In general, the users may transmit only over a subset of the time, so interference may occur over any part of the slot. We will call this a multiple access *finite field adder channel* over $\mathbb{F}_{2^k}$ where, the transmitted and received elements belong to $\mathbb{F}_{2^k}$, for $1 \leq k$. The output of the channel is the sum of the inputs and noise with addition over $\mathbb{F}_{2^k}$. In this thesis, we consider the noise-free as well as the noisy multiple access finite field adder channel.

When we are transmitting two sources over a multiple access channel, we can do joint source-channel coding or separate source-channel coding. If we do joint source-channel coding, we use a single joint source-channel encoder at each source to encode it. On the other hand, when we do separate source-channel coding, we first use a source coder as described in [13] to compress the source. The source codewords are subsequently channel coded using techniques described in [21, 22] and sent over the channel. When the maximum achievable sum rate due to separate source-channel coding is the same as the maximum achievable sum rate due to joint source-channel coding, we say that separation holds for the multiple access channel and we can do compression and channel coding separately. On the other hand, if there is a loss in sum rate due to separate source-channel coding, we say that separation fails. For multiple access channels where the received alphabet is the integer sum of the inputs, it is shown by a simple example in [12] that source-channel separation does not hold. We investigate whether source channel separation holds for the multiple access finite field adder channel.

We know from [20] that for Slepian-Wolf source networks, the random coding error exponents are universally attainable by linear codes. Therefore, random linear source codes achieve performance arbitrarily close to the optimal one derived by Slepian and Wolf [13]. The codewords coming out of the Slepian-Wolf source coders are asymptotically independent. In this thesis, we show that source-channel separation holds when the noise is independent of the inputs. Moreover, we show that random linear channel codes achieve the maximum sum rate. This implies that we can combine the random linear source and channel codes into a random linear joint source-channel code. Thus, a random linear joint source-channel code is optimal for the multiple access finite field adder channel where noise is independent of inputs.

For the multiple access finite field adder channel we discuss issues of separation in Chapter 2. We show that if noise is input-dependent, then source-channel separation may not hold. Moreover, for channels over $\mathbb{F}_2$, we derive the expression for the probability of source-channel separation failing. We compute this probability to be $1/4$ when the noise parameters are picked independently and uniformly. For binary channels, we derive an upper bound of 0.0776 bit for the maximum loss in sum rate due to separate source-channel coding when separation fails. We prove that the bound is very tight by showing that it is accurate to the second decimal place.

In Chapter 3, we consider a time-slotted noise-free multiple access finite field adder channel. First, we develop a single-slot model for this channel. We use two metrics: capacity and *code rate*. Capacity represents the maximum achievable error-free rate.

A metric is needed to represent the overhead (in the form of redundancy) required for reliable communication. *Code rate* is defined as the ratio of the information symbols recovered at the receiver, to the symbols sent by the transmitters in a slot duration. Code rate tells us about this overhead. For the case when the transmit alphabet has a zero-energy symbol, i.e. no energy is required for transmitting 0 and all other symbols have equal energy, code rate represents the energy overhead required to transmit over a particular channel. In our work, we will assume that when transmitters do not transmit, they are transmitting 0. Note that code rate is a dimensionless quantity and has a maximum value of 1. It attains the value of 1 when there is no multiple access interference and no noise. The noise-free multiple access network becomes equivalent to a point-to-point channel without noise.

We derive the capacity region and maximum code rate and analyze the variation of these quantities with field size. A systematic code construction is presented which is shown to achieve maximum code rate and capacity. We prove that codes that achieve the maximum code rate also achieve capacity. Next, we consider systematic random codes and obtain conditions under which random codes achieve maximum code rate and capacity. We derive an expression for the probability of error when the codes are chosen randomly. The probability of error goes to 0 exponentially with code length and field size. We also look at the bursty case when transmitters transmit according to a Bernoulli process. We propose coding techniques to maximize code rate. For such bursty channels, we shown that, when the information codewords at the input to the channel encoders have the same size, maximum expected code rate is achieved by adding redundancy at the transmitter with a higher probability of transmission and not adding any redundancy at the transmitter with lower probability of transmission.

In Chapter 4, we consider a time-slotted noisy multiple access finite field adder channel. Noise is independent of the inputs and is additive over the same field as the input and output alphabet of the channel. First, we develop a model for communicating over this channel and then establish the capacity region and maximum code rate. Using the results of the noisy multiple access strong coding theorem developed by Liao in [22], we obtain the error exponents and hence the expression for average probability of error when a random code is used for communicating over this channel.

We present our conclusions and directions for future research in Chapter 5.

Chapter 2

Source-Channel Separation

In this chapter, we look at source-channel separation for two transmitter, single receiver, noisy and noise-free multiple access finite field adder channels. We prove that, when noise is input-dependent, source-channel separation may not hold. We derive the expression for the probability of separation failure for the channel over $\mathbb{F}_2$. Moreover, we compute an upper bound on the maximum loss in sum rate due to separation failure and show that the bound is accurate to the second decimal place. However, when noise is independent of input symbols, source-channel separation holds and thus, there is no loss in performing source and channel coding separately.

Since the noise-free case is a special case of the input-independent noise case (noise is 0 here), it follows that source-channel separation holds for the noise-free multiple access finite field adder channel.

2.1 Criteria for source-channel separation

Consider two sources generating symbols U and V, which have to be sent over the noisy multiple access finite field adder channel. Both these symbols are elements of a finite set that have to be sent using suitable encoding techniques. If we do joint source-channel coding, we will use a single joint source-channel coder at each source to encode U and V to codewords $\vec{X}_a$ and $\vec{X}_b$ respectively. The elements of $\vec{X}_a$ and $\vec{X}_b$ are from $\mathbb{F}_{2^k}$. $\vec{X}_a$ and $\vec{X}_b$ may be correlated if U and V are correlated. Codewords

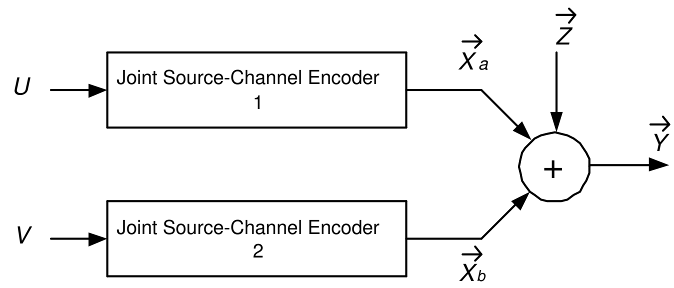

Figure 2-1: Joint source-channel coding.

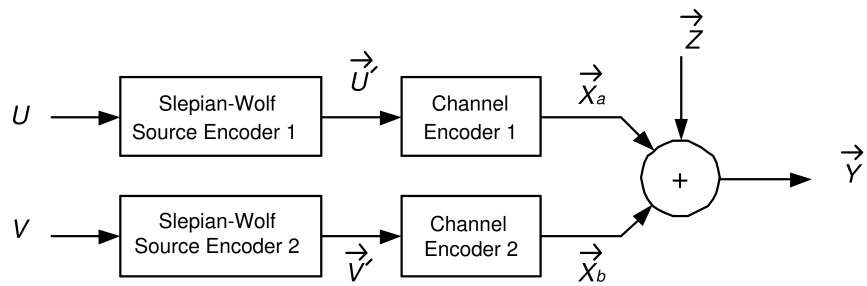

Figure 2-2: Separate source-channel coding.

$\vec{X}_a$ and $\vec{X}_b$ add over $\mathbb{F}_{2^k}$ with noise vector $\vec{Z}$ to yield $\vec{Y}$. Elements of $\vec{Z}$ and $\vec{Y}$ are also from $\mathbb{F}_{2^k}$. The joint source-channel coding is shown in Figure 2-1.

For separate source-channel coding we first do compression, which can be achieved by Slepian-Wolf source coding [13] of (U, V) to codeword pair $(\vec{U}', \vec{V}')$. Following source coding, $\vec{U}'$ and $\vec{V}'$ are asymptotically independent. $\vec{U}'$ and $\vec{V}'$ are subsequently channel coded to codewords $\vec{X}_a$ and $\vec{X}_b$ respectively. $\vec{X}_a$ and $\vec{X}_b$ have their elements in $\mathbb{F}_{2^k}$ and the codewords add over $\mathbb{F}_{2^k}$ with noise vector $\vec{Z}$ to give $\vec{Y}$. Elements of $\vec{Z}$ and $\vec{Y}$ are from $\mathbb{F}_{2^k}$. When we do separate source-channel coding, $\vec{X}_a$ and $\vec{X}_b$ are independent since $\vec{U}'$ and $\vec{V}'$ are independent. The scheme for separate source-channel coding is shown in Figure 2-2.

Joint source-channel coding gives the maximum possible sum rate for this channel. However, when separation holds, separate source-channel coding gives the same maximum sum rate as joint-source channel coding. When separation fails, we always incur a loss in sum rate by doing separate source-channel coding. Note that we assume that the joint source-channel encoder can generate the joint probability distribution needed to achieve the joint sum rate or in other words, it is able to match the source statistics to that needed by the channel. This may not be true for all source-channel pairs, which reduces the maximum joint sum rate possible, thereby reducing the loss in sum rate due to separation failure. However, by assuming that sources can be matched to the channel, we are looking at the maximum possible loss.

In this section, we establish criteria based on which we decide whether source-channel separation holds for a noisy multiple access finite field adder channel. Let R_a and R_b be the reliably transmitted rates of two transmitters a and b respectively, which are communicating with a single receiver over a noisy multiple access finite field adder channel. Using results from [21, 22], we know that the sum rate, R_{sum}, defined over the probability distribution of the input symbols X_a and X_b, follows the following equation:

$$R_{sum} = R_a + R_b \leq I(X_a, X_b; Y). \tag{2.1}$$

When we do separate source-channel coding, X_a and X_b are independent and we can transmit reliably at a maximum sum rate, $R_{sumSSCC}$, given by

$$R_{sumSSCC} = \max_{P_{X_a}(x_a)P_{X_b}(x_b)} I(X_a, X_b; Y) \tag{2.2}$$

where $P_{X_a}(x_a)$ and $P_{X_b}(x_b)$ are the probability mass functions of X_a and X_b respectively.

When we do joint source-channel coding, inputs X_a and X_b may not be independent of each other. We consider the joint probability mass function of X_a and X_b as $P_{X_aX_b}(x_a, x_b)$. By doing joint source-channel coding, we can transmit reliably at a maximum sum rate, $R_{sumJSCC}$, given by

$$R_{sumJSCC} = \max_{P_{X_aX_b}(x_a,x_b)} I(X_a, X_b; Y). \tag{2.3}$$

Criterion for separation not holding

If, for a channel,

$$R_{sumSSCC} < R_{sumJSCC}, \tag{2.4}$$

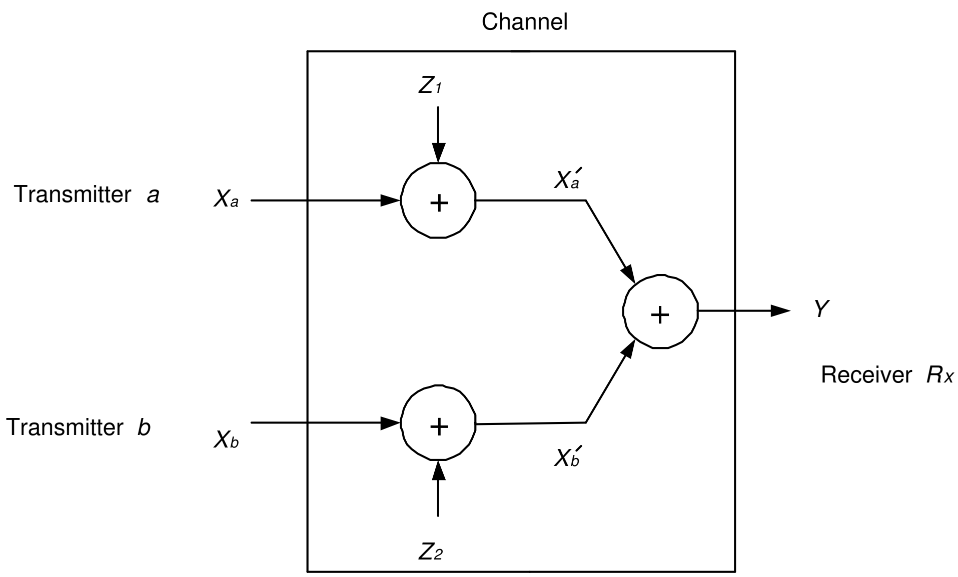

Figure 2-3: Noisy binary multiple access finite field adder channel with input-dependent noise. Transmitters a and b transmit binary {0,1} symbols, X_a and X_b respectively, to a single receiver, $R_\times$.

then we can achieve a higher sum rate by doing joint source-channel coding and separate source-channel coding does not give the highest possible sum rate. Therefore, if (2.4) holds for the channel, source-channel separation fails.

Criterion for separation holding

In order to prove that separation holds for a channel, it is sufficient to show that

$$R_{sumJSCC} = R_{sumSSCC}. \tag{2.5}$$

2.2 Source-channel separation when noise is input-dependent

We consider a noisy multiple access finite field adder channel over the binary field where two transmitters, a and b, transmit binary {0,1} symbols, X_a and X_b, respec-

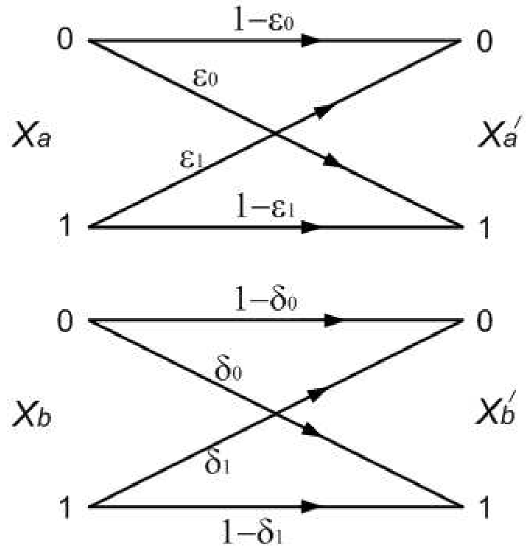

Figure 2-4: Input dependent noises Z_1 and Z_2 result in asymmetric transition probabilities between (X_a, X_a') and (X_b, X_b') respectively.

tively, to a single receiver, $R_\times$, as shown in Figure 2-3. The received symbol Y is also binary {0,1}. Binary noises (Z_1, Z_2) add on to each input. These noises depend on the input symbols being transmitted. When we consider input-dependent noise, we imply that the noise depends only on a single input. Thus, Z_1 depends on X_1 and Z_2 depends on X_2. The input dependence is represented by a binary asymmetric channel shown in Figure 2-4. Note that the asymmetry may be different for the two transmitters and the transition probabilities are parameterized by $\epsilon_0, \epsilon_1, \delta_0$ and δ_1. After addition of noises (Z_1, Z_2), the corrupted inputs are added together modulo-2. Therefore, the first process transforms X_a and X_b to X_a' and X_b' respectively, and the second process forms a mod-2 sum of X_a' and X_b'.

We show that separation may not hold for this channel when noise depends on the input symbols or in other words, the transition probability matrices of the part of the channel between (X_a, X_a') and (X_b, X_b') are asymmetric. Considering the binary case is sufficient to show that, for general fields, separation may not hold when noise is input-dependent. We show that source-channel separation does hold for the binary channel when noise is independent of the input symbols. In this case, the transition matrices for $(X_a, X_a^{'})$ and $(X_b, X_b^{'})$ are symmetric.

We will be using the following notation throughout our discussion:

$$p_1 = Pr(X_a = 0), \tag{2.6}$$

$$p_2 = Pr(X_b = 0), \tag{2.7}$$

$$\mathcal{H}(q) = -q\log_2(q) - (1-q)\log_2(1-q). \tag{2.8}$$

2.2.1 Maximum sum rate using separate source-channel coding

We now compute $R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$, which is the maximum sum rate at which reliable communication is possible by separate source-channel coding. X_a and X_b are independent here. Since,

$$I(X_a, X_b; Y) = H(Y) - H(Y|X_a, X_b), \tag{2.9}$$

we have

$$R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) = \max_{p_1,p_2}[H(Y) - H(Y|X_a, X_b)]. \tag{2.10}$$

As X_a and X_b are independent,

$$\begin{aligned} Pr(X_a^{'} = 0) &= Pr(X_a^{'} = 0|X_a = 0)p_1 + Pr(X_a^{'} = 0|X_a = 1)(1-p_1) \\ &= \epsilon_1 + p_1(1 - \epsilon_0 - \epsilon_1), \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} Pr(X_b^{'} = 0) &= Pr(X_b^{'} = 0|X_b = 0)p_2 + Pr(X_b^{'} = 0|X_b = 1)(1-p_2) \\ &= \delta_1 + p_2(1 - \delta_0 - \delta_1). \end{aligned} \tag{2.12}$$

Computing $H(Y)$

The independence of X_a and X_b results in the independence of $X_a^{'}$ and $X_b^{'}$. Thus, we have

$$Pr(Y = 0) = Pr(X_a^{'} = 0; X_b^{'} = 0) + Pr(X_a^{'} = 1; X_b^{'} = 1)$$

$$= \quad 1 - Pr(X_a^{'} = 0) - Pr(X_b^{'} = 0) + 2Pr(X_a^{'} = 0)Pr(X_b^{'} = 0). \tag{2.13}$$

Substituting the values of $Pr(X_a^{'} = 0)$ and $Pr(X_b^{'} = 0)$ from (2.11) and (2.12) in (2.13), we obtain

$$\begin{aligned} Pr(Y = 0) &= 1 - \epsilon_1 - \delta_1 + 2\epsilon_1\delta_1 - p_1(1 - 2\delta_1)(1 - \epsilon_0 - \epsilon_1) \\ &\quad -p_2(1 - 2\epsilon_1)(1 - \delta_0 - \delta_1) + 2p_1p_2(1 - \epsilon_0 - \epsilon_1)(1 - \delta_0 - \delta_1). \end{aligned} \tag{2.14}$$

Therefore, $H(Y)$ is given by

$$\begin{aligned} H(Y) &= \mathcal{H}[\epsilon_1 + \delta_1 - 2\epsilon_1\delta_1 + p_1(1 - 2\delta_1)(1 - \epsilon_0 - \epsilon_1) + p_2(1 - 2\epsilon_1)(1 - \delta_0 - \delta_1) \\ &\quad -2p_1p_2(1 - \epsilon_0 - \epsilon_1)(1 - \delta_0 - \delta_1)]. \end{aligned} \tag{2.15}$$

Computing $H(Y|X_a, X_b)$

Using the fact that X_a and X_b are independent, we obtain that

$$\begin{aligned} H(Y|X_a, X_b) &= p_1p_2H(Y|X_a = 0, X_b = 0) + p_1(1 - p_2)H(Y|X_a = 0, X_b = 1) \\ &\quad +p_2(1 - p_1)H(Y|X_a = 1, X_b = 0) + (1 - p_1)(1 - p_2)H(Y|X_a = 1, X_b = 1). \end{aligned} \tag{2.16}$$

$X_a^{'}$ and $X_b^{'}$ are independent since X_a and X_b are independent. We may use this fact to compute the following conditional probabilities:

$$\begin{aligned} Pr(Y = 0|X_a = 0, X_b = 0) &= Pr(X_a^{'} = 0|X_a = 0)Pr(X_b^{'} = 0|X_b = 0) \\ &\quad +Pr(X_a^{'} = 1|X_a = 0)Pr(X_b^{'} = 1|X_b = 0) \\ &= 1 - \epsilon_0 - \delta_0 + 2\epsilon_0\delta_0. \end{aligned} \tag{2.17}$$

$$\begin{aligned} Pr(Y = 0|X_a = 0, X_b = 1) &= Pr(X_a^{'} = 0|X_a = 0)Pr(X_b^{'} = 0|X_b = 1) \\ &\quad +Pr(X_a^{'} = 1|X_a = 0)Pr(X_b^{'} = 1|X_b = 1) \\ &= \epsilon_0 + \delta_1 - 2\epsilon_0\delta_1. \end{aligned} \tag{2.18}$$

$$
\begin{aligned}
Pr(Y=0|X_a=1,X_b=0) &= Pr(X_a^{'}=0|X_a=1)Pr(X_b^{'}=0|X_b=0) \\
&\quad +Pr(X_a^{'}=1|X_a=1)Pr(X_b^{'}=1|X_b=0) \\
&= \epsilon_1+\delta_0-2\epsilon_1\delta_0. \qquad (2.19)
\end{aligned}
$$

$$
\begin{aligned}
Pr(Y=0|X_a=1,X_b=1) &= Pr(X_a^{'}=0|X_a=1)Pr(X_b^{'}=0|X_b=1) \\
&\quad +Pr(X_a^{'}=1|X_a=1)Pr(X_b^{'}=1|X_b=1) \\
&= 1-\epsilon_1-\delta_1+2\epsilon_1\delta_1. \qquad (2.20)
\end{aligned}
$$

Thus, $H(Y|X_a,X_b)$ is

$$
\begin{aligned}
H(Y|X_a,X_b) &= p_1p_2\mathcal{H}[\epsilon_0+\delta_0-2\epsilon_0\delta_0]+p_1(1-p_2)\mathcal{H}[\epsilon_0+\delta_1-2\epsilon_0\delta_1] \\
&\quad +p_2(1-p_1)\mathcal{H}[\epsilon_1+\delta_0-2\epsilon_1\delta_0]+(1-p_1)(1-p_2)\mathcal{H}[\epsilon_1+\delta_1-2\epsilon_1\delta_1].
\end{aligned} \qquad (2.21)
$$

Substituting (2.15) and (2.21) in (2.10), we obtain

$$
\begin{aligned}
R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1) &= \max_{p_1,p_2}\mathcal{H}[\epsilon_1+\delta_1-2\epsilon_1\delta_1+p_1(1-2\delta_1)(1-\epsilon_0-\epsilon_1) \\
&\quad +p_2(1-2\epsilon_1)(1-\delta_0-\delta_1)-2p_1p_2(1-\epsilon_0-\epsilon_1)(1-\delta_0-\delta_1)] \\
&\quad -p_1p_2\mathcal{H}[\epsilon_0+\delta_0-2\epsilon_0\delta_0]-p_1(1-p_2)\mathcal{H}[\epsilon_0+\delta_1-2\epsilon_0\delta_1] \\
&\quad -p_2(1-p_1)\mathcal{H}[\epsilon_1+\delta_0-2\epsilon_1\delta_0]-(1-p_1)(1-p_2)\mathcal{H}[\epsilon_1+\delta_1-2\epsilon_1\delta_1].
\end{aligned} \qquad (2.22)
$$

Rearranging the terms in (2.22), we obtain

$$
\begin{aligned}
R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1) &= \max_{p_1,p_2}\mathcal{H}[p_1p_2(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0)+p_1(1-p_2)(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
&\quad +p_2(1-p_1)(\epsilon_1+\delta_0-2\epsilon_1\delta_0)+(1-p_1)(1-p_2)(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1)] \\
&\quad -p_1p_2\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0)-p_1(1-p_2)\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
&\quad -p_2(1-p_1)\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)-(1-p_1)(1-p_2)\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1).
\end{aligned} \qquad (2.23)
$$

2.2.2 Maximum sum rate using joint source-channel coding

We now compute $R_{sumJSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$, which is the maximum sum rate at which reliable communication is possible by joint source-channel coding. We maximize

$X_a \backslash X_b$	0	1
0	P_{00}	P_{01}
1	P_{10}	P_{11}

Figure 2-5: Joint probability distribution of X_a and X_b.

$I(X_a, X_b; Y)$ over the joint probability distribution of (X_a, X_b). This joint distribution $P_{X_a X_b}(x_a, x_b)$ is specified by

$$P_{00} = P_{X_a X_b}(0, 0), \tag{2.24}$$

$$P_{01} = P_{X_a X_b}(0, 1), \tag{2.25}$$

$$P_{10} = P_{X_a X_b}(1, 0), \tag{2.26}$$

$$P_{11} = P_{X_a X_b}(1, 1), \tag{2.27}$$

and is illustrated in Figure 2-5. We have

$$\begin{aligned} R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) &= \max_{P_{X_a X_b}(x_a, x_b)} [H(Y) - H(Y|X_a, X_b)] && (2.28) \\ &= \max_{P_{00}, P_{01}, P_{10}, P_{11}} [H(Y) - H(Y|X_a, X_b)]. && (2.29) \end{aligned}$$

Computing $H(Y)$
We have that:

$$Pr(Y = 0) = Pr(X_a^{'} = 0, X_b^{'} = 0) + Pr(X_a^{'} = 1, X_b^{'} = 1) \tag{2.30}$$

We compute $Pr(X_a^{'} = 0, X_b^{'} = 0)$ and $Pr(X_a^{'} = 1, X_b^{'} = 1)$ as follows

$$\begin{aligned} Pr(X_a^{'} = 0, X_b^{'} = 0) &= Pr(X_a + Z_1 = 0, X_b + Z_2 = 0) \\ &= Pr(X_a = 0, Z_1 = 0, X_b = 0, Z_2 = 0) \\ &\quad + Pr(X_a = 1, Z_1 = 1, X_b = 1, Z_2 = 1) \\ &\quad + Pr(X_a = 0, Z_1 = 0, X_b = 1, Z_2 = 1) \\ &\quad + Pr(X_a = 1, Z_1 = 1, X_b = 0, Z_2 = 0) \end{aligned}$$

$$
\begin{aligned}
&= P_{00}(1-\epsilon_0)(1-\delta_0) + P_{01}\delta_1(1-\epsilon_0) \\
&\quad +P_{10}\epsilon_1(1-\delta_0) + P_{11}\epsilon_1\delta_1, \qquad (2.31)
\end{aligned}
$$

$$
\begin{aligned}
Pr(X_a^{'}=1, X_b^{'}=1) &= Pr(X_a+Z_1=1, X_b+Z_2=1) \\
&= Pr(X_a=0, Z_1=1, X_b=0, Z_2=1) \\
&\quad +Pr(X_a=1, Z_1=0, X_b=0, Z_2=1) \\
&\quad +Pr(X_a=0, Z_1=1, X_b=1, Z_2=0) \\
&\quad +Pr(X_a=1, Z_1=0, X_b=1, Z_2=0) \\
&= P_{00}\epsilon_0\delta_0 + P_{01}\epsilon_0(1-\delta_1) \\
&\quad +P_{10}\delta_0(1-\epsilon_1) + P_{11}(1-\epsilon_1)(1-\delta_1). \qquad (2.32)
\end{aligned}
$$

Substituting (2.31,2.32) in (2.30), we have

$$
\begin{aligned}
Pr(Y=0) \quad &= P_{00}(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0) + P_{10}(\delta_0+\epsilon_1-2\delta_0\epsilon_1) \\
&\quad +P_{01}(\delta_1+\epsilon_0-2\epsilon_0\delta_1) + P_{11}(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1). \qquad (2.33)
\end{aligned}
$$

Therefore we compute $H(Y)$ as

$$
\begin{aligned}
H(Y) \quad &= \mathcal{H}[P_{00}(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0) + P_{10}(\delta_0+\epsilon_1-2\delta_0\epsilon_1) \\
&\quad +P_{01}(\delta_1+\epsilon_0-2\epsilon_0\delta_1) + P_{11}(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1)]. \qquad (2.34)
\end{aligned}
$$

Computing $H(Y|X_a, X_b)$

We have that:

$$
\begin{aligned}
H(Y|X_a, X_b) \quad &= P_{00}H(Y|X_a=0, X_b=0) + P_{01}H(Y|X_a=0, X_b=1) \\
&\quad +P_{10}H(Y|X_a=1, X_b=0) + P_{11}H(Y|X_a=1, X_b=1). \qquad (2.35)
\end{aligned}
$$

We compute the conditional probabilities as

$$
\begin{aligned}
Pr(Y=0|X_a=0, X_b=0) &= Pr(X_a^{'}=0, X_b^{'}=0|X_a=0, X_b=0) \\
&\quad +Pr(X_a^{'}=1, X_b^{'}=1|X_a=0, X_b=0) \\
&= Pr(Z_1=0, Z_2=0|X_a=0, X_b=0) \\
&\quad +Pr(Z_1=1, Z_2=1|X_a=0, X_b=0) \\
&= 1-\epsilon_0-\delta_0+2\epsilon_0\delta_0, \qquad (2.36)
\end{aligned}
$$

$$
\begin{aligned}
Pr(Y=0|X_a=0,X_b=1) &= Pr(X_a^{'}=0,X_b^{'}=0|X_a=0,X_b=1) \\
&\quad +Pr(X_a^{'}=1,X_b^{'}=1|X_a=0,X_b=1) \\
&= Pr(Z_1=0,Z_2=1|X_a=0,X_b=1) \\
&\quad +Pr(Z_1=1,Z_2=0|X_a=0,X_b=1) \\
&= \epsilon_0+\delta_1-2\epsilon_0\delta_1, \qquad (2.37)
\end{aligned}
$$

$$
\begin{aligned}
Pr(Y=0|X_a=1,X_b=0) &= Pr(X_a^{'}=0,X_b^{'}=0|X_a=1,X_b=0) \\
&\quad +Pr(X_a^{'}=1,X_b^{'}=1|X_a=1,X_b=0) \\
&= Pr(Z_1=1,Z_2=0|X_a=1,X_b=0) \\
&\quad +Pr(Z_1=0,Z_2=1|X_a=1,X_b=0) \\
&= \epsilon_1+\delta_0-2\epsilon_0\delta_1, \qquad (2.38)
\end{aligned}
$$

$$
\begin{aligned}
Pr(Y=0|X_a=1,X_b=1) &= Pr(X_a^{'}=0,X_b^{'}=0|X_a=1,X_b=1) \\
&\quad +Pr(X_a^{'}=1,X_b^{'}=1|X_a=1,X_b=1) \\
&= Pr(Z_1=1,Z_2=1|X_a=1,X_b=1) \\
&\quad +Pr(Z_1=0,Z_2=0|X_a=1,X_b=1) \\
&= 1-\epsilon_1-\delta_1+2\epsilon_1\delta_1. \qquad (2.39)
\end{aligned}
$$

From (2.36), (2.37) , (2.38) and (2.39), we obtain

$$H(Y|X_a=0,X_b=0)=\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0), \qquad (2.40)$$

$$H(Y|X_a=0,X_b=1)=\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1), \qquad (2.41)$$

$$H(Y|X_a=1,X_b=0)=\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0), \qquad (2.42)$$

$$H(Y|X_a=1,X_b=1)=\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1). \qquad (2.43)$$

Substituting (2.40 - 2.43) in (2.35), we have

$$
\begin{aligned}
H(Y|X_a,X_b) &= P_{00}\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0)+P_{01}\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
&\quad +P_{10}\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)+P_{11}\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1). \qquad (2.44)
\end{aligned}
$$

Combining (2.34) and (2.44) we obtain

$$\begin{aligned} R_{sumJSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1) &= \max_{P_{00},P_{01},P_{10},P_{11}} \mathcal{H}[P_{00}(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0)+P_{01}(\delta_1+\epsilon_0-2\epsilon_0\delta_1) \\ &\quad +P_{10}(\delta_0+\epsilon_1-2\delta_0\epsilon_1)+P_{11}(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1)] \\ &\quad -P_{00}\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0)-P_{01}\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\ &\quad -P_{10}\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)-P_{11}\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1). \qquad (2.45) \end{aligned}$$

2.2.3 Failure of source-channel separation

We now give an example of a noisy binary multiple access finite field adder channel where noise is input-dependent. Since noise is input-dependent, $\epsilon_0 \neq \epsilon_1$ and $\delta_0 \neq \delta_1$. Let

$$\begin{aligned} \epsilon_0 &= 0.02, \\ \epsilon_1 &= 0.8, \\ \delta_0 &= 0.29, \\ \delta_1 &= 0.98. \qquad (2.46) \end{aligned}$$

We know from (2.23) that $R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$ is given by

$$R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1) = \max_{p_1,p_2} R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2), \qquad (2.47)$$

where $R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2)$ is defined as

$$\begin{aligned} R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2) &= \mathcal{H}[\epsilon_1+\delta_1-2\epsilon_1\delta_1+p_1(1-2\delta_1)(1-\epsilon_0-\epsilon_1) \\ &\quad +p_2(1-2\epsilon_1)(1-\delta_0-\delta_1)-2p_1p_2(1-\epsilon_0-\epsilon_1)(1-\delta_0-\delta_1)] \\ &\quad -p_1p_2\mathcal{H}[\epsilon_0+\delta_0-2\epsilon_0\delta_0]-p_1(1-p_2)\mathcal{H}[\epsilon_0+\delta_1-2\epsilon_0\delta_1] \\ &\quad -p_2(1-p_1)\mathcal{H}[\epsilon_1+\delta_0-2\epsilon_1\delta_0]-(1-p_1)(1-p_2)\mathcal{H}[\epsilon_1+\delta_1-2\epsilon_1\delta_1]. \end{aligned} \qquad (2.48)$$

$R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2)$ is defined for $(p_1,p_2) \in [0,1]\times[0,1]$. We set

$$\frac{\partial R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2)}{\partial p_1} = 0, \qquad (2.49)$$

$$\frac{\partial R_s(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2)}{\partial p_2} = 0. \qquad (2.50)$$

Equations (2.49,2.50) give a transcendental equation in p_1 and p_2 which is solved to get critical points of $R_s(\epsilon_0, \epsilon_1, \delta_0, \delta_1, p_1, p_2)$. Let these set of points form a set $\mathcal{C}$. The point where $R_s(\epsilon_0, \epsilon_1, \delta_0, \delta_1, p_1, p_2)$ attains a maximum may lie on the boundaries of the square $[0,1] \times [0,1]$. $R_s(\epsilon_0, \epsilon_1, \delta_0, \delta_1, p_1, p_2)$ is a concave function of one variable at each side of this square and thus, this single variable function has one maximum. We thus get 4 points, 1 at each boundary, at which the maximum value of $R_s(\epsilon_0, \epsilon_1, \delta_0, \delta_1, p_1, p_2)$ can occur. Let these points form a set $\mathcal{B}$. Now, $R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ is given by

$$R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) = \max_{(p_1,p_2) \in \mathcal{C} \bigcup \mathcal{B}} R_s(\epsilon_0, \epsilon_1, \delta_0, \delta_1, p_1, p_2). \tag{2.51}$$

Applying this to our example, we get

$$R_{sumSSCC}(0.02, 0.8, 0.29, 0.98) = 0.0975, \tag{2.52}$$

for $p_1 = 0.969$ and $p_2 = 0.43$.

Now, $R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ is a concave function of P_{00}, P_{01}, P_{10} and P_{11}. Hence, its maximum can be found by a gradient search technique. Using MATLAB, we compute

$$R_{sumJSCC}(0.02, 0.8, 0.29, 0.98) = 0.140, \tag{2.53}$$

where the maximum occurs at $P_{00} = 0$, $P_{01} = 0.6$, $P_{10} = 0.4$ and $P_{11} = 0$.

From (2.52) and (2.53), we have

$$R_{sumSSCC}(0.02, 0.8, 0.29, 0.98) < R_{sumJSCC}(0.02, 0.8, 0.29, 0.98). \tag{2.54}$$

Since (2.54) holds, we have satisfied the criterion established in section 2.1 for source-channel separation to not hold. Thus, separation may not hold when the transition matrices between $(X_a, X_a^{'})$ and $(X_b, X_b^{'})$ are asymmetric or, in other words, the noise is input-dependent.

Since separation may not hold when noise is input dependent for a binary noisy multiple access finite field adder channel, separation may not hold for any noisy multiple access finite field adder channel when noise is input-dependent. We have thus proved the following theorem:

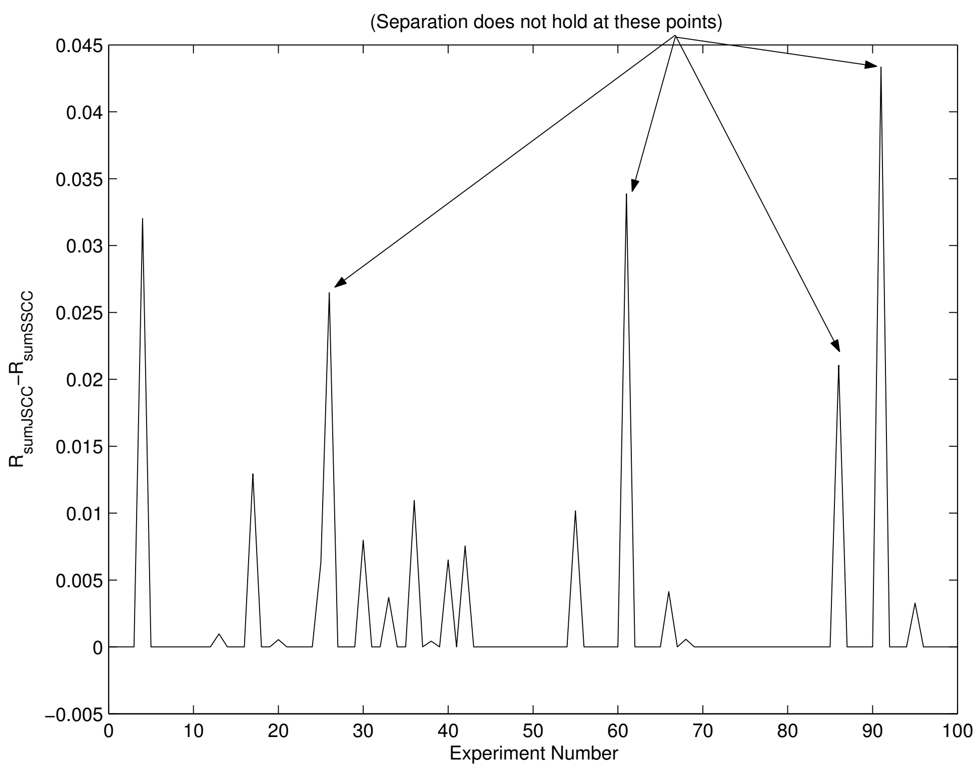

Figure 2-6: Separation does not hold for points above 0. Each experiment was carried out with a randomly picked set of $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$, where $\epsilon_0, \epsilon_1, \delta_0$ and δ_1 are uniformly distributed in $[0, 1]$.

Theorem 2.1 *Source-channel separation may not hold for a noisy two transmitter single receiver multiple access finite field adder channel where the input symbols, output symbol and noise are elements of* $\mathbb{F}_{2^k}$ *for* $1 \leq k$ *and interference occurs in* $\mathbb{F}_{2^k}$ *with noise being input-dependent.*

The example which we gave for source-channel separation failure is not an isolated one. To assess how often source-channel separation fails, we plotted the difference between $R_{sumJSCC}$ and $R_{sumSSCC}$ for 100 different sets of randomly chosen transition probabilities $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ using the same technique that we used for the example. Figure 2-6 shows how $R_{sumJSCC} - R_{sumSSCC}$ changes from set to set. Whenever we have $R_{sumJSCC} - R_{sumSSCC} > 0$, we satisfy the criterion established in section 2.1 for separation to fail. The plot shows several such points where separation does not hold and their difference. We can see from the plot that source-channel separation seems to fail frequently and that the magnitude of the difference is small. We confirm this fact in section 2.2.4 by deriving the expression for the probability of separation failing. This probability is significantly high. Moreover, in section 2.2.5, we derive an upper bound on the maximum loss in sum rate due to separation failing. This bound is accurate to the second decimal and is a small number. We show that although the probability that separation fails is high, the loss in sum rate due to separation failure is low.

2.2.4 Probability of separation failure

In this section, we derive the expression for probability not holding when the noise transition probabilities $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ are chosen at random. We will let $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ be continuous random variables whose probability density functions are L_2 functions. From (2.45) we see that the sum rate due to joint source-channel coding is given by

$$R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) = \max_{P_{00}, P_{01}, P_{10}, P_{11}} R_J(\epsilon_0, \epsilon_1, \delta_0, \delta_1, P_{00}, P_{01}, P_{10}, P_{11}) \qquad (2.55)$$

where

$$
\begin{aligned}
R_J(\epsilon_0,\epsilon_1,\delta_0,\delta_1,P_{00},P_{01},P_{10},P_{11}) \;\;=\;\; & \mathcal{H}[P_{00}(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0)+P_{01}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
& +P_{10}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)+P_{11}(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1)] \\
& -P_{00}\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0)-P_{01}\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
& -P_{10}\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)-P_{11}\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1).
\end{aligned} \tag{2.56}
$$

The sum rate obtained by separate source-channel coding is given by (2.23) as

$$
R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)=\max_{p_1,p_2} R_S(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2), \tag{2.57}
$$

where

$$
\begin{aligned}
R_S(\epsilon_0,\epsilon_1,\delta_0,\delta_1,p_1,p_2) \;\;=\;\; & \mathcal{H}[p_1p_2(1-\epsilon_0-\delta_0+2\epsilon_0\delta_0)+p_1(1-p_2)(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
& +p_2(1-p_1)(\epsilon_1+\delta_0-2\epsilon_1\delta_0)+(1-p_1)(1-p_2)(1-\epsilon_1-\delta_1+2\epsilon_1\delta_1)] \\
& -p_1p_2\mathcal{H}(\epsilon_0+\delta_0-2\epsilon_0\delta_0)-p_1(1-p_2)\mathcal{H}(\epsilon_0+\delta_1-2\epsilon_0\delta_1) \\
& -p_2(1-p_1)\mathcal{H}(\epsilon_1+\delta_0-2\epsilon_1\delta_0)-(1-p_1)(1-p_2)\mathcal{H}(\epsilon_1+\delta_1-2\epsilon_1\delta_1).
\end{aligned} \tag{2.58}
$$

The loss in sum rate is thus

$$
G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)=R_{sumJSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)-R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1). \tag{2.59}
$$

Note that since $R_{sumJSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)\geq R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$, $G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)\geq 0$. We want to find the maximum value of $G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$ for $0\leq\epsilon_0\leq 1$, $0\leq\epsilon_1\leq 1$, $0\leq\delta_0\leq 1$ and $0\leq\delta_1\leq 1$. Substituting in the equations for $R_{sumJSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$ and $R_{sumSSCC}(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$, we see that $G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$ obeys the following

$$
G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)=G(1-\epsilon_0,1-\epsilon_1,\delta_0,\delta_1), \tag{2.60}
$$

$$
G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)=G(\epsilon_0,\epsilon_1,1-\delta_0,1-\delta_1), \tag{2.61}
$$

$$
\Rightarrow G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)=G(1-\epsilon_0,1-\epsilon_1,1-\delta_0,1-\delta_1). \tag{2.62}
$$

Equations (2.60,2.61) divide the space spanned by $(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$ into 4 symmetric regions. Thus, in finding the maximum value of $G(\epsilon_0,\epsilon_1,\delta_0,\delta_1)$, we can confine our analysis to $0 \leq \epsilon_0 \leq 0.5$, $0 \leq \epsilon_1 \leq 1$, $0 \leq \delta_0 \leq 0.5$ and $0 \leq \delta_1 \leq 1$. We will call this space $\mathcal{D}$. Thus, if there are k points in $\mathcal{D}$ where $G(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ is maximum, in all there are $4k$ points in the entire space of $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ where $G(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ is maximum.

Condition under which source-channel separation holds

Let us now establish the condition on the joint probability distribution for source-channel separation to hold. If (p_1, p_2) is able to achieve the maximum joint sum rate achieving distribution specified by $(P_{00}, P_{01}, P_{10}, P_{11})$, we have:

$$p_1 p_2 = P_{00}, \tag{2.63}$$

$$p_1(1 - p_2) = P_{01}, \tag{2.64}$$

$$p_2(1 - p_1) = P_{10}, \tag{2.65}$$

$$(1 - p_1)(1 - p_2) = P_{11}. \tag{2.66}$$

For (2.63,2.64,2.65,2.66) to hold, we need that

$$P_{11} P_{00} = P_{01} P_{10}. \tag{2.67}$$

Thus, whenever a capacity achieving joint probability distribution obeys (2.67), source-channel separation holds.

Maximizing of joint sum rate

Let us denote

$$\alpha_{00} = 1 - \epsilon_0 - \delta_0 + 2\epsilon_0\delta_0, \tag{2.68}$$

$$\alpha_{01} = \epsilon_0 + \delta_1 - 2\epsilon_0\delta_1, \tag{2.69}$$

$$\alpha_{10} = \epsilon_1 + \delta_0 - 2\epsilon_1\delta_0, \tag{2.70}$$

$$\alpha_{11} = 1 - \epsilon_1 - \delta_1 + 2\epsilon_1\delta_1. \tag{2.71}$$

Substituting in (2.55,2.56,2.57,2.58), the expressions for $R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ and $R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ become

$$R_{sumJSCC}(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}) = \max_{P_{00}, P_{01}, P_{10}, P_{11}} R_J(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}, P_{00}, P_{01}, P_{10}, P_{11})$$

$$= \max_{P_{00},P_{01},P_{10},P_{11}} \mathcal{H}[P_{00}\alpha_{00} + P_{01}\alpha_{01} + P_{10}\alpha_{10} + P_{11}\alpha_{11}] - P_{00}\mathcal{H}(\alpha_{00}) - P_{01}\mathcal{H}(\alpha_{01}) - P_{10}\mathcal{H}(\alpha_{10}) - P_{11}\mathcal{H}(\alpha_{11}), \tag{2.72}$$

$$\begin{aligned} R_{sumSSCC}(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}) &= \max_{p_1,p_2} R_S(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11},p_1,p_2) \\ &= \max_{p_1,p_2} \mathcal{H}[p_1p_2\alpha_{00} + p_1(1-p_2)\alpha_{01} \\ &\qquad + p_2(1-p_1)\alpha_{10} + (1-p_1)(1-p_2)\alpha_{11}] \\ &\quad - p_1p_2\mathcal{H}(\alpha_{00}) - p_1(1-p_2)\mathcal{H}(\alpha_{01}) \\ &\qquad - p_2(1-p_1)\mathcal{H}(\alpha_{10}) - (1-p_1)(1-p_2)\mathcal{H}(\alpha_{11}), \end{aligned} \tag{2.73}$$

and

$$G(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}) = R_{sumJSCC}(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}) - R_{sumSSCC}(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}). \tag{2.74}$$

Define

$$\alpha_{min} = \min\{\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}\}, \tag{2.75}$$

$$\alpha_{max} = \max\{\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}\}, \tag{2.76}$$

$$\alpha_1,\alpha_2 \in \{\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11}\} - \{\alpha_{min},\alpha_{max}\}. \tag{2.77}$$

Therefore, α_{min} and α_{max} are the smallest and largest α_{ij} respectively, where $i,j \in \{0,1\}$ and $\alpha_{min} \leq \alpha_1,\alpha_2 \leq \alpha_{max}$. We will now prove the following lemmas:

Lemma 2.1

$$\frac{\mathcal{H}(\alpha_{max}) - \mathcal{H}(\alpha)}{\alpha_{max} - \alpha} \leq \frac{\mathcal{H}(\alpha_{max}) - \mathcal{H}(\alpha_{min})}{\alpha_{max} - \alpha_{min}} \tag{2.78}$$

for $\alpha \in [\alpha_{min},\alpha_{max}]$.

Proof: Consider the function

$$t(\alpha) = \frac{\mathcal{H}(\alpha_{max}) - \mathcal{H}(\alpha)}{\alpha_{max} - \alpha}. \tag{2.79}$$

Differentiating with respect to α, we have

$$\frac{\partial t}{\partial \alpha} = \frac{(\alpha_{max}-\alpha)\log_2(\frac{\alpha}{1-\alpha}) + \mathcal{H}(\alpha_{max}) - \mathcal{H}(\alpha)}{(\alpha_{max}-\alpha)^2} \tag{2.80}$$

$$= \frac{-D(\alpha_{max}||\alpha)}{(\alpha_{max}-\alpha)^2} \tag{2.81}$$

$$\leq 0. \tag{2.82}$$

$D(\alpha_{max}||\alpha)$ is the Kullback Liebler distance (with logarithms taken base 2) between the probability distributions $(\alpha_{max}, 1-\alpha_{max})$ and $(\alpha, 1-\alpha)$. Since the first derivative is negative for $\alpha \in [\alpha_{min}, \alpha_{max}]$, $t(\alpha)$ is a monotonically decreasing function for $\alpha \in [\alpha_{min}, \alpha_{max}]$. Hence, it takes the maximum value when $\alpha = \alpha_{min}$. Thus, we have

$$t(\alpha) \leq t(\alpha_{min}) \qquad \alpha \in [\alpha_{min}, \alpha_{max}]. \tag{2.83}$$

The proof for the lemma is now complete. □

Lemma 2.2 *There exists some $p \in [0,1]$ such that $R_J(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}, p, 0, 0, 1-p) \geq R_J(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}, p'_{min}, p'_1, p'_2, p'_{max})$ where $(p'_{min}, p'_1, p'_2, p'_{max})$ specifies a joint probability distribution.*

Proof: Choose p such that

$$p\alpha_{min} + (1-p)\alpha_{max} = p'_{min}\alpha_{min} + p'_1\alpha_1 + p'_2\alpha_2 + p'_{max}\alpha_{max}. \tag{2.84}$$

This is valid, since $p \in [0,1]$. Now, the lemma holds if

$$p\mathcal{H}(\alpha_{min}) + (1-p)\mathcal{H}(\alpha_{max}) \leq p'_{min}\mathcal{H}(\alpha_{min}) + p'_1\mathcal{H}(\alpha_1) + p'_2\mathcal{H}(\alpha_2) + p'_{max}\mathcal{H}(\alpha_{max}). \tag{2.85}$$

Solving (2.84,2.85), we see that we require

$$0 \leq p'_1[\mathcal{H}(\alpha_1) - \mathcal{H}(\alpha_{max}) + \frac{\alpha_1 - \alpha_{max}}{\alpha_{max} - \alpha_{min}}\{\mathcal{H}(\alpha_{min}) - \mathcal{H}(\alpha_{max})\}]$$
$$+p'_2[\mathcal{H}(\alpha_2) - \mathcal{H}(\alpha_{max}) + \frac{\alpha_2 - \alpha_{max}}{\alpha_{max} - \alpha_{min}}\{\mathcal{H}(\alpha_{min}) - \mathcal{H}(\alpha_{max})\}]. \tag{2.86}$$

as a necessary and sufficient condition for the lemma to hold. Using Lemma 2.1, we see that

$$0 \le \mathcal{H}(\alpha_1) - \mathcal{H}(\alpha_{max}) + \frac{\alpha_1 - \alpha_{max}}{\alpha_{max} - \alpha_{min}}\{\mathcal{H}(\alpha_{min}) - \mathcal{H}(\alpha_{max})\}, \tag{2.87}$$

$$0 \le \mathcal{H}(\alpha_2) - \mathcal{H}(\alpha_{max}) + \frac{\alpha_2 - \alpha_{max}}{\alpha_{max} - \alpha_{min}}\{\mathcal{H}(\alpha_{min}) - \mathcal{H}(\alpha_{max})\} \tag{2.88}$$

since $\alpha_1, \alpha_2 \in [\alpha_{min}, \alpha_{max}]$. Thus, (2.86) holds, which completes the proof. □

We now prove a theorem that shows the probability distributions that maximize sum rate due to joint source-channel coding for a given channel. Let us define a function Ind that extracts the indices of its argument. For example

$$Ind(\alpha_{ij}) = (i, j). \tag{2.89}$$

Theorem 2.2 *$R_{sumJSCC}(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max})$ is maximized by the joint probability distribution $(p_{min}, 0, 0, p_{max})$ where p_{min} and p_{max} are defined as*

$$p_{min} = P_{Ind\{\min(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11})\}}, \tag{2.90}$$

$$p_{max} = P_{Ind\{\max(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11})\}}. \tag{2.91}$$

($p_{min} = P_{ij}$ where (i, j) are indices for which $\alpha_{min} = \alpha_{ij}$ and $p_{max} = P_{lm}$ where (l, m) are indices for which $\alpha_{max} = \alpha_{lm}$.) Therefore, $R_{sumJSCC}$ only depends on α_{min} and α_{max}.

Proof: For a joint distribution $(p'_{min}, p'_1, p'_2, p'_{max})$, we have from (2.72),

$$R_{sumJSCC}(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}) = \max_{p'_{min},p'_1,p'_2,p'_{max}} R_J(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}, p'_{min}, p'_1, p'_2, p'_{max}), \tag{2.92}$$

and $\alpha_{min} \le \alpha_1, \alpha_2 \le \alpha_{max}$. Using Lemma 2.2, we have

$$R_{sumJSCC}(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}) = \max_{p\in[0,1]} R_J(\alpha_{min}, \alpha_1, \alpha_2, \alpha_{max}, p, 0, 0, 1-p). \tag{2.93}$$

Let (2.93) be maximized at $p = q^*$. Note that q^* multiplies α_{min} and $(1-q^*)$ multiplies α_{max}. We define

$$p_{min} = q*, \tag{2.94}$$

$$p_{max} = 1 - q*. \tag{2.95}$$

Thus, the joint sum rate is maximized by the probability distribution $(p_{min}, 0, 0, p_{max})$. Since p_{min} multiplies α_{min} and p_{max} multiplies α_{max}, we can also denote p_{min} and p_{max} as $P_{k_1k_2}$ and $P_{r_1r_2}$ respectively where,

$$(k_1, k_2) = Ind\{\min(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})\}, \tag{2.96}$$

$$(r_1, r_2) = Ind\{\max(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})\}. \tag{2.97}$$

Therefore,

$$p_{min} = P_{Ind\{\min(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})\}}, \tag{2.98}$$

$$p_{max} = P_{Ind\{\max(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})\}}. \tag{2.99}$$

Thus, $R_{sumJSCC}$ is only dependent on $(\alpha_{min}, \alpha_{max})$ and is a function of only these two variables. The proof is now complete. □

For a channel where $\alpha_{ij} = \alpha_{i',j'}$ for $i, j, i', j' \in \{0, 1\}$, there is more than one joint distribution that achieves the maximum sum rate. Equation (2.67) must hold for at least one of these distributions for source-channel separation to hold. Hence, while optimizing, when we have more that one choice, we will choose p_{min} and p_{max} such that (2.67) holds. If (2.67) does not hold for any of the choices, then source-channel separation fails for the channel being considered.

Regions

We now further subdivide $\mathcal{D}$ into 4 disjoint regions $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ and $\mathcal{S}$. These regions are defined as:

$$\mathcal{P} : 0 \leq \epsilon_0, \epsilon_1, \delta_0, \delta_1 \leq 0.5 \tag{2.100}$$

$$\mathcal{Q} : 0 \leq \epsilon_0, \epsilon_1, \delta_0 \leq 0.5, 0.5 \leq \delta_1 \leq 1 \tag{2.101}$$

$$\mathcal{R} : 0 \leq \epsilon_0, \delta_0, \delta_1 \leq 0.5, 0.5 \leq \epsilon_1 \leq 1 \tag{2.102}$$

$$\mathcal{S} : 0 \leq \epsilon_0, \delta_0 \leq 0.5, 0.5 \leq \epsilon_1, \delta_1 \leq 1, \tag{2.103}$$

From Theorem 2.2, we see that in each of these regions the joint sum rate is maximized by the distribution $(p_{min}, 0, 0, p_{max})$ where p_{min} and p_{max} are defined as

$$p_{min} = P_{Ind\{\min(\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})\}}, \tag{2.104}$$

$$p_{max} = P_{Ind\{\max(\alpha_{00},\alpha_{01},\alpha_{10},\alpha_{11})\}}. \tag{2.105}$$

Region $\mathcal{P}$

The region $0 \leq \epsilon_0, \epsilon_1, \delta_0, \delta_1 \leq 0.5$ makes

$$0.5 \leq \alpha_{00} \leq 1, \tag{2.106}$$

$$0 \leq \alpha_{01} \leq 0.5, \tag{2.107}$$

$$0 \leq \alpha_{10} \leq 0.5, \tag{2.108}$$

$$0.5 \leq \alpha_{11} \leq 1. \tag{2.109}$$

Therefore,

$$\alpha_{00}, \alpha_{11} \geq \alpha_{01}, \alpha_{10}, \tag{2.110}$$

$$\Rightarrow p_{min} = p_{Ind\{\min(\alpha_{01},\alpha_{10})\}}, \tag{2.111}$$

$$p_{max} = p_{Ind\{\max(\alpha_{00},\alpha_{11})\}}. \tag{2.112}$$

$$\Rightarrow P_{00}P_{11} = P_{01}P_{10} = 0. \tag{2.113}$$

Thus, (2.113) satisfies (2.67) and source-channel separation holds for region $\mathcal{P}$.

Region $\mathcal{Q}$

The region $0 \leq \epsilon_0, \epsilon_1, \delta_0 \leq 0.5$, and $0.5 \leq \delta_1 \leq 1$ makes

$$0.5 \leq \alpha_{00} \leq 1, \tag{2.114}$$

$$0.5 \leq \alpha_{01} \leq 1, \tag{2.115}$$

$$0 \leq \alpha_{10} \leq 0.5, \tag{2.116}$$

$$0 \leq \alpha_{11} \leq 0.5, \tag{2.117}$$

and we will call this region $\mathcal{Q}$.

Therefore,

$$\alpha_{00}, \alpha_{01} \geq \alpha_{10}, \alpha_{11}. \tag{2.118}$$

If $\alpha_{00} \geq \alpha_{01}$

$$1 - \epsilon_0 - \delta_0 + 2\epsilon_0\delta_0 \geq \epsilon_0 + \delta_1 - 2\epsilon_0\delta_1, \tag{2.119}$$

$$\Rightarrow (1 - \delta_0 - \delta_1) \geq 0, \tag{2.120}$$

$$\Rightarrow (1 - 2\epsilon_1)(1 - \delta_0 - \delta_1) \geq 0, \tag{2.121}$$

$$\Rightarrow \alpha_{11} \geq \alpha_{10}, \tag{2.122}$$

$$\Rightarrow p_{min} = P_{10}, \tag{2.123}$$

$$p_{max} = P_{00}, \tag{2.124}$$

$$\Rightarrow P_{00}P_{11} = P_{01}P_{10} = 0. \tag{2.125}$$

If $\alpha_{00} \leq \alpha_{01}$,

$$1 - \epsilon_0 - \delta_0 + 2\epsilon_0\delta_0 \leq \epsilon_0 + \delta_1 - 2\epsilon_0\delta_1, \tag{2.126}$$

$$\Rightarrow (1 - \delta_0 - \delta_1) \leq 0, \tag{2.127}$$

$$\Rightarrow (1 - 2\epsilon_1)(1 - \delta_0 - \delta_1) \leq 0, \tag{2.128}$$

$$\Rightarrow \alpha_{11} \leq \alpha_{10}, \tag{2.129}$$

$$\Rightarrow p_{min} = P_{11}, \tag{2.130}$$

$$p_{max} = P_{01}, \tag{2.131}$$

$$\Rightarrow P_{00}P_{11} = P_{01}P_{10} = 0. \tag{2.132}$$

Thus, (2.125,2.132) satisfy (2.67)and source-channel separation holds for region $\mathcal{Q}$.

Region $\mathcal{R}$

The region $0 \leq \epsilon_0, \delta_0, \delta_1 \leq 0.5$, and $0.5 \leq \epsilon_1 \leq 1$ makes

$$0.5 \leq \alpha_{00} \leq 1, \tag{2.133}$$

$$0 \leq \alpha_{01} \leq 0.5, \tag{2.134}$$

$$0.5 \leq \alpha_{10} \leq 1, \tag{2.135}$$

$$0 \leq \alpha_{11} \leq 0.5, \tag{2.136}$$

and we will call this region $\mathcal{R}$.

Therefore,

$$\alpha_{00}, \alpha_{10} \geq \alpha_{01}, \alpha_{11}. \tag{2.137}$$

If $\alpha_{00} \geq \alpha_{10}$

$$
\begin{aligned}
1-\epsilon_0-\delta_0+2\epsilon_0\delta_0 &\geq \epsilon_1+\delta_0-2\epsilon_1\delta_0, && (2.138)\\
\Rightarrow (1-\epsilon_0-\epsilon_1) &\geq 0, && (2.139)\\
\Rightarrow (1-2\delta_1)(1-\epsilon_0-\epsilon_1) &\geq 0, && (2.140)\\
\Rightarrow \alpha_{11} &\geq \alpha_{01}, && (2.141)\\
\Rightarrow p_{min} &= P_{01}, && (2.142)\\
p_{max} &= P_{00}, && (2.143)\\
\Rightarrow P_{00}P_{11} &= P_{01}P_{10} = 0. && (2.144)
\end{aligned}
$$

If $\alpha_{00} \leq \alpha_{10}$

$$
\begin{aligned}
1-\epsilon_0-\delta_0+2\epsilon_0\delta_0 &\leq \epsilon_1+\delta_0-2\epsilon_1\delta_0, && (2.145)\\
\Rightarrow (1-\epsilon_0-\epsilon_1) &\leq 0, && (2.146)\\
\Rightarrow (1-2\delta_1)(1-\epsilon_0-\epsilon_1) &\leq 0, && (2.147)\\
\Rightarrow \alpha_{11} &\leq \alpha_{01}, && (2.148)\\
\Rightarrow p_{min} &= P_{11}, && (2.149)\\
p_{max} &= P_{10}, && (2.150)\\
\Rightarrow P_{00}P_{11} &= P_{01}P_{10} = 0. && (2.151)
\end{aligned}
$$

Thus, (2.144,2.151) satisfy (2.67) and source-channel separation holds for region $\mathcal{R}$.

Region $\mathcal{S}$

The region $0 \leq \epsilon_0, \delta_0 \leq 0.5$, and $0.5 \leq \epsilon_1, \delta_1 \leq 1$ makes

$$
\begin{aligned}
0.5 \leq \alpha_{00} \leq 1, && (2.152)\\
0.5 \leq \alpha_{01} \leq 1, && (2.153)\\
0.5 \leq \alpha_{10} \leq 1, && (2.154)\\
0.5 \leq \alpha_{11} \leq 1, && (2.155)
\end{aligned}
$$

and we will call this region $\mathcal{S}$. Note that (2.67) is satisfied iff an inequality in the expressions described henceforth for $\mathcal{S}$ is met with equality. However, such equalities translate to curves in the four dimension space spanned by $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$. Since $\epsilon_0, \epsilon_1, \delta_0$ and δ_1 are continuous random variables whose probability density functions are L_2 functions, the probability of being on the curve is 0. Thus we will not consider points on these boundary curves. We will show that except for these boundary curves, separation always fails in $\mathcal{S}$.

In $\mathcal{S}$, if $\alpha_{00} \geq \alpha_{01}, \alpha_{10}$

$$1 - \epsilon_0 - \epsilon_1 \geq 0, \tag{2.156}$$

$$1 - \delta_0 - \delta_1 \geq 0 \tag{2.157}$$

$$\Rightarrow \alpha_{11} \leq \alpha_{01}, \alpha_{10}, \tag{2.158}$$

$$\Rightarrow p_{min} = P_{11}, \tag{2.159}$$

$$p_{max} = P_{00}. \tag{2.160}$$

$$\Rightarrow P_{11}P_{00} \neq P_{01}P_{10}. \tag{2.161}$$

If $\alpha_{00} \leq \alpha_{01}, \alpha_{10}$

$$1 - \epsilon_0 - \epsilon_1 \leq 0, \tag{2.162}$$

$$1 - \delta_0 - \delta_1 \leq 0 \tag{2.163}$$

$$\Rightarrow \alpha_{11} \geq \alpha_{01}, \alpha_{10}, \tag{2.164}$$

$$\Rightarrow p_{min} = P_{00}, \tag{2.165}$$

$$p_{max} = P_{11}. \tag{2.166}$$

$$\Rightarrow P_{11}P_{00} \neq P_{01}P_{10}. \tag{2.167}$$

If $\alpha_{01} \geq \alpha_{00}, \alpha_{11}$

$$1 - \delta_0 - \delta_1 \leq 0, \tag{2.168}$$

$$1 - \epsilon_0 - \epsilon_1 \geq 0 \tag{2.169}$$

$$\Rightarrow \alpha_{10} \leq \alpha_{00}, \alpha_{11}, \tag{2.170}$$

$$\Rightarrow p_{min} = P_{10}, \tag{2.171}$$

$$p_{max} = P_{01}. \tag{2.172}$$

$$\Rightarrow P_{11}P_{00} \neq P_{01}P_{10}. \tag{2.173}$$

If $\alpha_{01} \leq \alpha_{00}, \alpha_{11}$

$$1 - \delta_0 - \delta_1 \geq 0, \tag{2.174}$$

$$1 - \epsilon_0 - \epsilon_1 \leq 0 \tag{2.175}$$

$$\Rightarrow \alpha_{10} \geq \alpha_{00}, \alpha_{11}, \tag{2.176}$$

$$\Rightarrow p_{min} = P_{01}, \tag{2.177}$$

$$p_{max} = P_{10} \tag{2.178}$$

$$\Rightarrow P_{11}P_{00} \neq P_{01}P_{10} \tag{2.179}$$

Thus in region $\mathcal{S}$, we see from (2.161,2.167,2.173,2.179) that (2.67) is never satisfied. Hence, we can say that source-channel separation fails in $\mathcal{S}$ with probability 1.

Probability that separation fails

We have seen that source-channel separation fails with probability 1 when we are in region $\mathcal{S}$. Due to the symmetry given by $(2.60, 2.61)$ the probability of separation failure $P_{failure}$ for $\epsilon_0, \epsilon_1, \delta_0, \delta_1 \in [0, 1]$ is

$$\begin{aligned} P_{failure} &= Pr(0 \leq \epsilon_0 \leq 0.5, 0.5 \leq \epsilon_1 \leq 1, 0 \leq \delta_0 \leq 0.5, 0.5 \leq \epsilon_1 \leq 1) \\ &\quad + Pr(0.5 \leq \epsilon_0 \leq 1, 0 \leq \epsilon_1 \leq 0.5, 0 \leq \delta_0 \leq 0.5, 0.5 \leq \epsilon_1 \leq 1) \\ &\quad + Pr(0 \leq \epsilon_0 \leq 0.5, 0.5 \leq \epsilon_1 \leq 1, 0.5 \leq \delta_0 \leq 1, 0 \leq \epsilon_1 \leq 0.5) \\ &\quad + Pr(0.5 \leq \epsilon_0 \leq 1, 0 \leq \epsilon_1 \leq 0.5, 0.5 \leq \delta_0 \leq 1, 0 \leq \epsilon_1 \leq 0.5). \end{aligned} \tag{2.180}$$

If $\epsilon_0, \epsilon_1, \delta_0$ and δ_1 are each chosen independently and uniformly from $[0, 1]$, we have

$$P_{failure} = 4.\frac{1}{16} \tag{2.181}$$

$$= \frac{1}{4}. \tag{2.182}$$

2.2.5 Maximum loss in sum rate due to separation failure

We now compute an upper bound on the maximum loss in sum rate by separate source-channel coding. Due to the symmetry given by $(2.60, 2.61)$, it is sufficient to consider $\mathcal{D}$. We have seen in section 2.2.4 that the only region in $\mathcal{D}$ where source-channel separation fails is $\mathcal{S}$. Separation holds on the boundary curves of $\mathcal{S}$ but these need not be considered since being on the curve is a 0 probability event. Thus, for investigating the maximum loss due to separate source-channel coding, it is sufficient if we look at $\mathcal{S}$ only.

Consider a noisy multiple access finite field adder channel C_1 parameterized by $(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ such that $(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \in \mathcal{S}$. From the analysis done in section 2.2.4, $(\alpha_{min}, \alpha_{max}) \in \{(\alpha_{00}, \alpha_{11}), (\alpha_{11}, \alpha_{00}), (\alpha_{01}, \alpha_{10}), (\alpha_{10}, \alpha_{01})\}$. Therefore, this channel can be of two types. It is of type 1 if $(\alpha_{min}, \alpha_{max}) \in \{(\alpha_{00}, \alpha_{11}), (\alpha_{11}, \alpha_{00})\}$ and of type 2 if $(\alpha_{min}, \alpha_{max}) \in \{(\alpha_{01}, \alpha_{10}), (\alpha_{10}, \alpha_{01})\}$. Consider a type 2 channel $C_{1-type2}$ with noise transition probabilities $(\epsilon_0^*, \epsilon_1^*, \delta_0^*, \delta_1^*)$. This channel can be parameterized by $(\alpha_{00}^*, \alpha_{01}^*, \alpha_{10}^*, \alpha_{11}^*)$ such that

$$\alpha_{00}^* = 1 - \epsilon_0^* - \delta_0^* + 2\epsilon_0^*\delta_0^*, \tag{2.183}$$

$$\alpha_{01}^* = \epsilon_0^* + \delta_1^* - 2\epsilon_0^*\delta_1^*, \tag{2.184}$$

$$\alpha_{10}^* = \epsilon_1^* + \delta_0^* - 2\epsilon_1^*\delta_0^*, \tag{2.185}$$

$$\alpha_{11}^* = 1 - \epsilon_1^* - \delta_1^* + 2\epsilon_1^*\delta_1^*, \tag{2.186}$$

$$\alpha_{min}^* = \min\{\alpha_{00}^*, \alpha_{01}^*, \alpha_{10}^*, \alpha_{11}^*\}, \tag{2.187}$$

$$\alpha_{max}^* = \max\{\alpha_{00}^*, \alpha_{01}^*, \alpha_{10}^*, \alpha_{11}^*\}. \tag{2.188}$$

Now, if we have a channel with noise transition probabilities $(\epsilon_0^{**}, \epsilon_1^{**}, \delta_0^{**}, \delta_1^{**})$ and parameterized by $(\alpha_{00}^{**}, \alpha_{01}^{**}, \alpha_{10}^{**}, \alpha_{11}^{**})$ such that

$$\epsilon_0^{**} = \epsilon_0^*, \tag{2.189}$$

$$\epsilon_1^{**} = \epsilon_1^*, \tag{2.190}$$

$$\delta_0^{**} = 1 - \delta_1^*, \tag{2.191}$$

$$\delta_1^{**} = 1 - \delta_0^*, \tag{2.192}$$

$$\alpha_{min}^{**} = \min\{\alpha_{00}^{**}, \alpha_{01}^{**}, \alpha_{10}^{**}, \alpha_{11}^{**}\}, \tag{2.193}$$

$$\alpha_{max}^{**} = \max\{\alpha_{00}^{**}, \alpha_{01}^{**}, \alpha_{10}^{**}, \alpha_{11}^{**}\}. \tag{2.194}$$

For this channel we have $0 \leq \epsilon_0^{**} \leq 0.5$, $0.5 \leq \epsilon_1^{**} \leq 1$, $0 \leq \delta_0^{**} \leq 0.5$ and $0.5 \leq \delta_1^{**} \leq 1$ and thus $(\epsilon_0^{**}, \epsilon_1^{**}, \delta_0^{**}, \delta_1^{**}) \in \mathcal{S}$. From (2.183 - 2.194), we obtain

$$\alpha_{00}^{**} = \alpha_{01}^{*}, \tag{2.195}$$

$$\alpha_{01}^{**} = \alpha_{00}^{*}, \tag{2.196}$$

$$\alpha_{10}^{**} = \alpha_{11}^{*}, \tag{2.197}$$

$$\alpha_{11}^{**} = \alpha_{10}^{*}.$$

Also, $(\alpha_{min}^{**}, \alpha_{max}^{**}) \in \{(\alpha_{00}^{**}, \alpha_{11}^{**}), (\alpha_{11}^{**}, \alpha_{00}^{**})\}$ and this channel is of type 1. We will call it $C_{1-type1}$. Note that,

$$(\alpha_{min}^{*}, \alpha_{max}^{*}) = (\alpha_{min}^{**}, \alpha_{max}^{**}). \tag{2.199}$$

Thus,

$$R_{sumJSCC}(\epsilon_0^{**}, \epsilon_1^{**}, \delta_0^{**}, \delta_1^{**}) = R_{sumJSCC}(\epsilon_0^{*}, \epsilon_1^{*}, \delta_0^{*}, \delta_1^{*}). \tag{2.200}$$

Moreover, if a sum rate by separate source channel coding is achieved for $C_{1-type2}$ by a probability distribution (p_1, p_2), the same sum rate can be achieved for $C_{1-type1}$ by a probability distribution $(p_1, 1 - p_2)$. Thus

$$R_{sumSSCC}(\epsilon_0^{**}, \epsilon_1^{**}, \delta_0^{**}, \delta_1^{**}) = R_{sumSSCC}(\epsilon_0^{*}, \epsilon_1^{*}, \delta_0^{*}, \delta_1^{*}). \tag{2.201}$$

Therefore, we see from (2.200,2.201) that

$$G(\epsilon_0^{**}, \epsilon_1^{**}, \delta_0^{**}, \delta_1^{**}) = G(\epsilon_0^{*}, \epsilon_1^{*}, \delta_0^{*}, \delta_1^{*}). \tag{2.202}$$

Therefore, for every type 2 channel in $\mathcal{S}$, there exists a type 1 channel in $\mathcal{S}$ with the same loss in sum rate due to separate source-channel coding and vice versa. Hence, for finding the maximum loss in sum rate, we can confine our analysis to channel C_1 being a type 1 channel. Therefore, we consider C_1 to be such that $(\alpha_{min}, \alpha_{max}) \in \{(\alpha_{00}, \alpha_{11}), (\alpha_{11}, \alpha_{00})\}$.

Consider another noisy multiple access finite field adder channel C_2 parameterized by $(\epsilon^{'}, \delta^{'}, \epsilon^{'}, \delta^{'})$ where $(\epsilon^{'}, \delta^{'}, \epsilon^{'}, \delta^{'}) \in \mathcal{S}$. For this channel, we have

$$\alpha_{00}^{'} = 1 - 2\epsilon^{'} + 2\epsilon^{'2}, \tag{2.203}$$

$$\alpha'_{01} = \epsilon' + \delta' - 2\epsilon'\delta', \tag{2.204}$$

$$\alpha'_{10} = \epsilon' + \delta' - 2\epsilon'\delta', \tag{2.205}$$

$$\alpha'_{11} = 1 - 2\delta' + 2\delta'^2. \tag{2.206}$$

Since, $\alpha'_{01} = \alpha'_{10}$, we have for C_2: $(\alpha'_{min}, \alpha'_{max}) \in \{(\alpha'_{00}, \alpha'_{11}), (\alpha'_{11}, \alpha'_{00})\}$ and this channel is of type 1. Let us choose (ϵ', δ') such that

$$\alpha'_{00} = \alpha_{00}, \tag{2.207}$$

$$\alpha'_{11} = \alpha_{11}. \tag{2.208}$$

$$\Rightarrow (\alpha_{min}, \alpha_{max}) = (\alpha'_{min}, \alpha'_{max}). \tag{2.209}$$

Therefore, by Theorem 2.2,

$$R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) = R_{sumJSCC}(\epsilon', \delta', \epsilon', \delta'), \tag{2.210}$$

where $R_{sumJSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ was defined in (2.55,2.56). Since in $\mathcal{S}$, $\epsilon' \in [0, 0.5]$ and $\delta' \in [0.5, 1]$, we have from (2.207,2.208)

$$\epsilon' = \frac{1}{2}[1 - \sqrt{1 - 2(\epsilon_0 + \delta_0 - 2\epsilon_0\delta_0)}], \tag{2.211}$$

$$\delta' = \frac{1}{2}[1 + \sqrt{1 - 2(\epsilon_1 + \delta_1 - 2\epsilon_1\delta_1)}]. \tag{2.212}$$

If $\epsilon_0 \leq \delta_0$, we have

$$\epsilon_0(1 - 2\delta_0) \leq \delta_0(1 - 2\delta_0), \tag{2.213}$$

$$\Rightarrow \epsilon_0 + \delta_0 - 2\epsilon_0\delta_0 \leq 2\delta_0 - 2\delta_0^2, \tag{2.214}$$

$$\Rightarrow 1 - \sqrt{1 - 2(\epsilon_0 + \delta_0 - 2\epsilon_0\delta_0)} \leq 2\delta_0, \tag{2.215}$$

$$\Rightarrow \epsilon' \leq \delta_0. \tag{2.216}$$

Also,

$$\epsilon_0(1 - 2\epsilon_0) \leq \delta_0(1 - 2\epsilon_0), \tag{2.217}$$

$$\Rightarrow 2\epsilon_0 - 2\epsilon_0^2 \leq \epsilon_0 + \delta_0 - 2\epsilon_0\delta_0, \tag{2.218}$$

$$\Rightarrow 2\epsilon_0 \leq 1 - \sqrt{1 - 2(\epsilon_0 + \delta_0 - 2\epsilon_0\delta_0)}, \tag{2.219}$$

$$\Rightarrow \epsilon_0 \leq \epsilon', \tag{2.220}$$

and

$$0 \leq (\epsilon_0 - \delta_0)^2, \tag{2.221}$$

$$\Rightarrow 1 - 2\epsilon_0 - 2\delta_0 + 4\epsilon_0\delta_0 \leq 1 - 2\epsilon_0 - 2\delta_0 + 2\epsilon_0\delta_0 + \epsilon_0^2 + \delta_0^2, \tag{2.222}$$

$$\Rightarrow \sqrt{1 - 2(\epsilon_0 + \delta_0 - 2\epsilon_0\delta_0)} \leq 1 - \epsilon_0 - \delta_0, \tag{2.223}$$

$$\Rightarrow 1 - 2\epsilon^{'} \leq 1 - \epsilon_0 - \delta_0, \tag{2.224}$$

$$\Rightarrow \delta_0 - \epsilon^{'} \leq \epsilon^{'} - \epsilon_0. \tag{2.225}$$

Thus, from (2.216,2.220,2.225), we have

$$\epsilon_0 \leq \epsilon^{'} \leq \delta_0, \tag{2.226}$$

$$\delta_0 - \epsilon^{'} \leq \epsilon^{'} - \epsilon_0. \tag{2.227}$$

Similarly if $\delta_0 \leq \epsilon_0$, we obtain

$$\delta_0 \leq \epsilon^{'} \leq \epsilon_0, \tag{2.228}$$

$$\epsilon_0 - \epsilon^{'} \leq \epsilon^{'} - \delta_0. \tag{2.229}$$

We can thus say that

$$\min(\epsilon_0, \delta_0) \leq \epsilon^{'} \leq \max(\epsilon_0, \delta_0), \tag{2.230}$$

$$\max(\epsilon_0, \delta_0) - \epsilon^{'} \leq \epsilon^{'} - \min(\epsilon_0, \delta_0). \tag{2.231}$$

If $\epsilon_1 \leq \delta_1$, we have

$$\epsilon_1(2\delta_1 - 1) \leq \delta_1(2\delta_1 - 1), \tag{2.232}$$

$$\Rightarrow -2\epsilon_1 - 2\delta_1 + 4\epsilon_1\delta_1 \leq 4\delta_1^2 - 4\delta_1, \tag{2.233}$$

$$\Rightarrow \sqrt{1 - 2(\epsilon_0 + \delta_0 - 2\epsilon_0\delta_0)} \leq 2\delta_1 - 1, \tag{2.234}$$

$$\Rightarrow \delta^{'} \leq \delta_1. \tag{2.235}$$

Moreover,

$$\epsilon_1(2\epsilon_1 - 1) \leq \delta_1(2\epsilon_1 - 1), \tag{2.236}$$

$$\Rightarrow 4\epsilon_1^2 - 2\epsilon_1 \leq -2\epsilon_1 - 2\delta_1 + 4\epsilon_1\delta_1, \tag{2.237}$$

$$\Rightarrow 2\epsilon_1 \leq 1 - \sqrt{1 - 2(\epsilon_1 + \delta_1 - 2\epsilon_1\delta_1)}, \tag{2.238}$$

$$\Rightarrow \epsilon_1 \leq \delta^{'}, \tag{2.239}$$

and

$$0 \leq (\delta_1 - \epsilon_1)^2, \tag{2.240}$$

$$\Rightarrow 1 - 2\epsilon_1 - 2\delta_1 + 4\epsilon_1\delta_1 \leq 1 - 2\epsilon_1 - 2\delta_1 + 2\epsilon_1\delta_1 + \epsilon_1^2 + \delta_1^2, \tag{2.241}$$

$$\Rightarrow \sqrt{1 - 2(\epsilon_1 + \delta_1 - 2\epsilon_1\delta_1)} \leq \epsilon_1 + \delta_1 - 1, \tag{2.242}$$

$$\Rightarrow 2\delta^{'} - 1 \leq \epsilon_1 + \delta_1 - 1, \tag{2.243}$$

$$\Rightarrow \delta^{'} - \epsilon_1 \leq \delta_1 - \delta^{'}. \tag{2.244}$$

Thus, from (2.235,2.239,2.244), we have

$$\epsilon_1 \leq \delta^{'} \leq \delta_1, \tag{2.245}$$

$$\delta^{'} - \epsilon_1 \leq \delta_1 - \delta^{'}. \tag{2.246}$$

Similarly if $\delta_1 \leq \epsilon_1$, we obtain

$$\delta_1 \leq \delta^{'} \leq \epsilon_1, \tag{2.247}$$

$$\delta^{'} - \delta_1 \leq \epsilon_1 - \delta^{'}.$$

We can thus say that

$$\min(\epsilon_1, \delta_1) \leq \delta^{'} \leq \max(\epsilon_1, \delta_1), \tag{2.249}$$

$$\delta^{'} - \min(\epsilon_1, \delta_1) \leq \max(\epsilon_1, \delta_1) - \delta^{'}. \tag{2.250}$$

From (2.230,2.231,2.249,2.250) we see that the transition probabilities for C_2 are closer to 0.5 than those of C_1. This makes the noise variance of C_2 higher than that of C_1. Thus

$$R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \geq R_{sumSSCC}(\epsilon^{'}, \delta^{'}, \epsilon^{'}, \delta^{'}), \tag{2.251}$$

where $R_{sumSSCC}(\epsilon_0, \epsilon_1, \delta_0, \delta_1)$ was defined in (2.57,2.58). From (2.210,2.251), we obtain

$$G(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \leq G(\epsilon^{'}, \delta^{'}, \epsilon^{'}, \delta^{'}). \tag{2.252}$$

Thus, for finding the largest loss in sum rate it is sufficient to focus on channels parameterized by $(\epsilon, \delta, \epsilon, \delta)$ for $\epsilon \in [0, 0.5]$ and $\delta \in [0.5, 1]$. Equation (2.252) can thus be written as

$$G(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \leq G(\epsilon, \delta, \epsilon, \delta). \tag{2.253}$$

From (2.55,2.56) and recalling that these are type 1 channels, we have

$$\begin{aligned} R_{sumJSCC}(\epsilon,\delta,\epsilon,\delta) &= \max_{P_{00}\in[0,1]} \mathcal{H}[P_{00}(1-2\epsilon+2\epsilon^2)+(1-P_{00})(1-2\delta+2\delta^2)], \\ &\quad -P_{00}\mathcal{H}(1-2\epsilon+2\epsilon^2)-(1-P_{00})\mathcal{H}(1-2\delta+2\delta^2). \end{aligned} \tag{2.254}$$

The maximum occurs at

$$P_{00}^* = \frac{1-2\delta+2\delta^2}{2(\delta-\epsilon)(\delta+\epsilon-1)} - \frac{1}{2(\delta-\epsilon)(\delta+\epsilon-1)[1+\exp(\phi(\epsilon,\delta))]} \tag{2.255}$$

and

$$\begin{aligned} R_{sumJSCC}(\epsilon,\delta,\epsilon,\delta) &= \mathcal{H}[\frac{1}{1+\exp(\phi(\epsilon,\delta))}] - \frac{\phi(\epsilon,\delta)}{1+\exp(\phi(\epsilon,\delta))} \\ &\quad -\frac{(1-2\delta+2\delta^2)\mathcal{H}(2\epsilon(1-\epsilon))-(1-2\epsilon+2\epsilon^2)\mathcal{H}(2\delta(1-\delta))}{2(\delta-\epsilon)(\delta+\epsilon-1)}, \end{aligned} \tag{2.256}$$

where

$$\phi(\epsilon,\delta) = \frac{1}{2}[\frac{\mathcal{H}(2\delta(1-\delta))-\mathcal{H}(2\epsilon(1-\epsilon))}{(\delta-\epsilon)(\delta+\epsilon-1)}]. \tag{2.257}$$

The sum rate achievable by separate source-channel coding is given by (2.57,2.58). Since, the noise faced by both inputs is the same, the maximum occurs when

$$p_1 = p_2 = p, \tag{2.258}$$

where $p \in [0,1]$.

Let us define $R_D(\epsilon,\delta,\epsilon,\delta)$ as the sum rate achieved by minimizing the Euclidian distance between points $(P_{00}^*, 1-P_{00}^*)$ and $(p^2,(1-p)^2)$ in two dimensional space. Thus

$$p^* = \arg\min_p [(P_{00}^* - p^2)^2 + ((1-P_{00}^*)-(1-p)^2)^2], \tag{2.259}$$

and

$$\begin{aligned} R_D(\epsilon,\delta,\epsilon,\delta) &= \mathcal{H}[p^{*2}(1-2\epsilon+2\epsilon^2)+2p^*(1-p^*)(\epsilon+\delta-2\epsilon\delta)+(1-p^*)^2(1-2\delta+2\delta^2)] \\ &\quad -p^{*2}\mathcal{H}(1-2\epsilon+2\epsilon^2)-2p^*(1-p^*)\mathcal{H}(\epsilon+\delta-2\epsilon\delta)-(1-p^*)^2\mathcal{H}(1-2\delta+2\delta^2). \end{aligned} \tag{2.260}$$

Define $G_D(\epsilon, \delta, \epsilon, \delta)$ as

$$G_D(\epsilon, \delta, \epsilon, \delta) = R_J(\epsilon, \delta, \epsilon, \delta) - R_D(\epsilon, \delta, \epsilon, \delta). \tag{2.261}$$

The probability distribution that minimizes Euclidian distance cannot give a higher sum rate than the probability distribution that maximizes the sum rate. This gives,

$$R_D(\epsilon, \delta, \epsilon, \delta) \leq R_S(\epsilon, \delta, \epsilon, \delta). \tag{2.262}$$

Thus, we have

$$G(\epsilon, \delta, \epsilon, \delta) \leq G_D(\epsilon, \delta, \epsilon, \delta). \tag{2.263}$$

Combining (2.253,2.263) we get for $(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \in \mathcal{S}$ and $\epsilon \in [0, 0.5]$, $\delta \in [0.5, 1]$

$$G(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \leq G_D(\epsilon, \delta, \epsilon, \delta). \tag{2.264}$$

Note that computing $R_S(\epsilon, \delta, \epsilon, \delta)$ is difficult in general since the maximizing probability distribution is hard to evaluate. Hence, by using a probability distribution that minimizes Euclidian distance, we have a lower sum rate and an upper bound on the loss in sum rate by separate source-channel coding. Later we show with an example that the bound is very tight.

Minimizing Euclidian distance

Let

$$d^2 = (P^*_{00} - p^2)^2 + ((1 - P^*_{00}) - (1 - p)^2). \tag{2.265}$$

Setting

$$\frac{\partial d^2}{\partial p} = 0, \tag{2.266}$$

we obtain

$$2p^3 - 3p^2 + 2p - P^*_{00} = 0. \tag{2.267}$$

We have three roots for (2.267). Two are complex and should be omitted. The real root is given by

$$p^* = \frac{1}{2} - \frac{1}{2^{2/3}\sqrt[3]{108P^*_{00} - 54 + \sqrt{108 + (108P^*_{00} - 54)^2}}} + \frac{\sqrt[3]{108P^*_{00} - 54 + \sqrt{108 + (108P^*_{00} - 54)^2}}}{2^{1/3}6}. \tag{2.268}$$

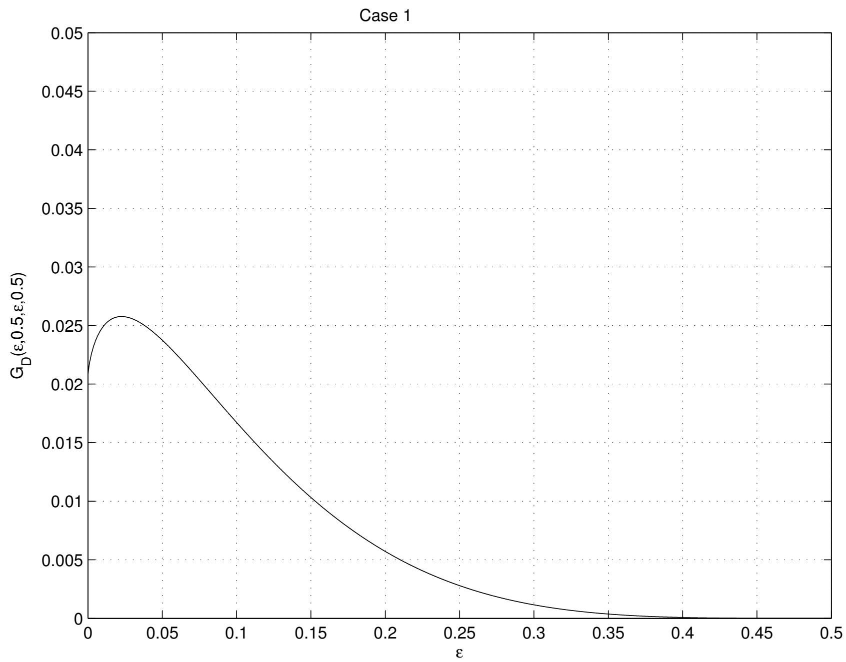

Figure 2-7: Case 1.

From (2.260,2.268), we have the explicit expression for $G_D(\epsilon, \delta, \epsilon, \delta)$.

For $\epsilon \in [0, 0.5]$ and $\delta \in [0.5, 1]$,

$$\frac{\partial G_D(\epsilon, \delta, \epsilon, \delta)}{\partial \epsilon} \neq 0, \tag{2.269}$$

$$\frac{\partial G_D(\epsilon, \delta, \epsilon, \delta)}{\partial \delta} \neq 0, \tag{2.270}$$

and thus the maximum values lie on the boundary of the square region we are looking at. Four cases arise:

Case 1: $\delta = 0.5$.

Figure 2-7 shows the plot of $G_D(\epsilon, 0.5, \epsilon, 0.5)$. To find the maximum value of $G_D(\epsilon, 0.5, \epsilon, 0.5)$ for $\epsilon \in [0, 0.5]$, we set

$$\frac{\partial G_D(\epsilon, 0.5, \epsilon, 0.5)}{\partial \epsilon} = 0. \tag{2.271}$$

We obtain a real root at $\epsilon = 0.0225$. Since this is the only critical point for $\epsilon \in [0, 0.5]$, it can be a maximum, minimum or point of inflexion. Using the same techniques as

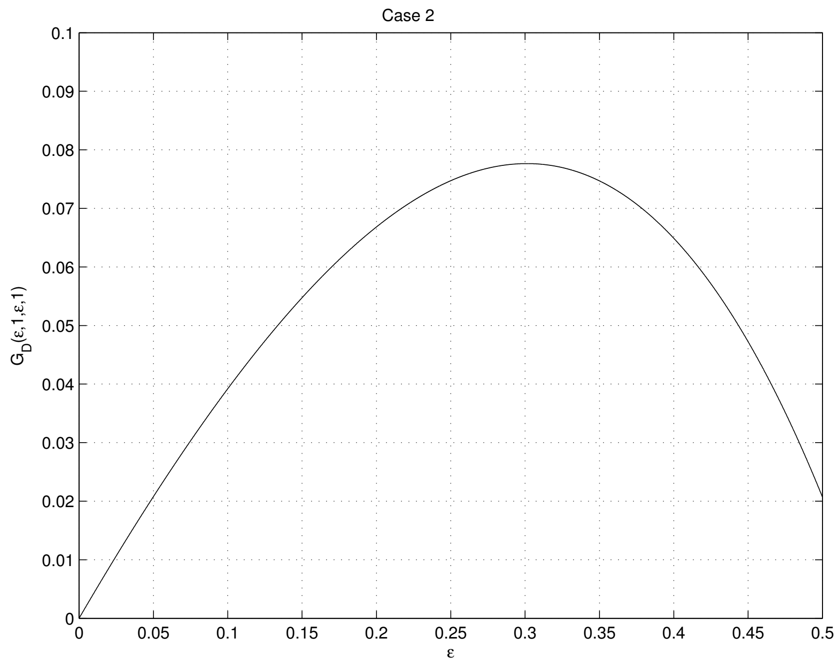

Figure 2-8: Case 2.

in section 2.2.3 for calculating the loss in sum rate due to separation failure for a specific channel, we compute

$$G_D(0, 0.5, 0, 0.5) = 0.0207, \tag{2.272}$$

$$G_D(0.0225, 0.5, 0.0225, 0.5) = 0.0258, \tag{2.273}$$

$$G_D(0.5, 0.5, 0.5, 0.5) = 0. \tag{2.274}$$

Since

$$G_D(0.0225, 0.5, 0.0225, 0.5) > G_D(0, 0.5, 0, 0.5), G_D(0.5, 0.5, 0.5, 0.5) \tag{2.275}$$

$G_D(0.0225, 0.5, 0.0225, 0.5)$ is the maximum. Thus,

$$G_D(\epsilon, 0.5, \epsilon, 0.5) \leq G_D(0.0225, 0.5, 0.0225, 0.5) = 0.0258. \tag{2.276}$$

Case 2: $\delta = 1$**.**

Figure 2-8 shows the plot of $G_D(\epsilon, 1, \epsilon, 1)$. To find the maximum value of $G_D(\epsilon, 1, \epsilon, 1)$ for $\epsilon \in [0, 0.5]$, we set

$$\frac{\partial G_D(\epsilon, 1, \epsilon, 1)}{\partial \epsilon} = 0. \tag{2.277}$$

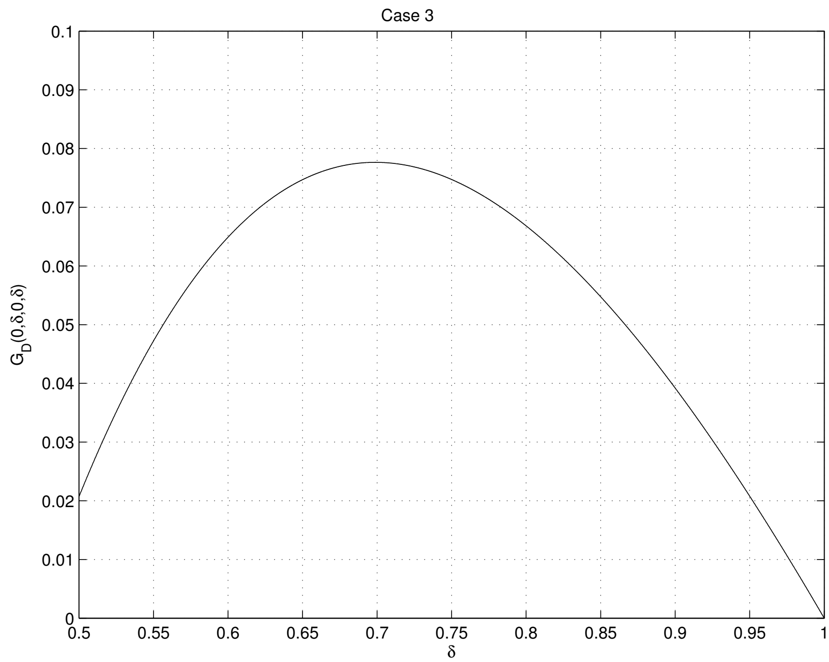

Figure 2-9: Case 3.

We obtain a real root at $\epsilon = 0.3012$. Since this is the only critical point for $\epsilon \in [0, 0.5]$, it can be a maximum, minimum or point of inflexion. Using the same techniques as in section 2.2.3 for calculating the loss in sum rate due to separation failure for a specific channel, we compute

$$G_D(0, 1, 0, 1) = 0, \tag{2.278}$$

$$G_D(0.3012, 1, 0.3012, 1) = 0.0776, \tag{2.279}$$

$$G_D(0.5, 1, 0.5, 1) = 0.0207. \tag{2.280}$$

Since

$$G_D(0.3012, 1, 0.3012, 1) > G_D(0, 1, 0, 1), G_D(0.5, 1, 0.5, 1), \tag{2.281}$$

$G_D(0.3012, 1, 0.3012, 1)$ is the maximum. Thus,

$$G_D(\epsilon, 1, \epsilon, 1) \leq G_D(0.3012, 1, 0.3012, 1) = 0.0776. \tag{2.282}$$

Case 3: $\epsilon = 0$.

Figure 2-9 shows the plot of $G_D(0, \delta, 0, \delta)$. To find the maximum value of $G_D(0, \delta, 0, \delta)$ for $\delta \in [0.5, 1]$, we set

$$\frac{\partial G_D(0, \delta, 0, \delta)}{\partial \delta} = 0. \tag{2.283}$$

We obtain a real root at $\delta = 0.6988$. Since this is the only critical point for $\delta \in [0.5, 1]$, it can be a maximum, minimum or point of inflexion. Using the same techniques as in section 2.2.3 for calculating the loss in sum rate due to separation failure for a specific channel, we compute

$$G_D(0, 0.5, 0, 0.5) = 0.0207, \tag{2.284}$$

$$G_D(0, 0.6988, 0, 0.6988) = 0.0776, \tag{2.285}$$

$$G_D(0, 1, 0, 1) = 0. \tag{2.286}$$

Since

$$G_D(0, 0.6988, 0, 0.6988) > G_D(0, 0.5, 0, 0.5), G_D(0, 1, 0, 1), \tag{2.287}$$

$G_D(0, 0.6988, 0, 0.6988)$ is the maximum. Thus,

$$G_D(0, \delta, 0, \delta) \leq G_D(0, 0.6988, 0, 0.6988) = 0.0776. \tag{2.288}$$

Case 4: $\epsilon = 0.5$.

Figure 2-10 shows the plot of $G_D(0.5, \delta, 0.5, \delta)$. To find the maximum value of $G_D(0.5, \delta, 0.5, \delta)$ for $\delta \in [0.5, 1]$, we set

$$\frac{\partial G_D(0.5, \delta, 0.5, \delta)}{\partial \delta} = 0. \tag{2.289}$$

We obtain a real root at $\delta = 0.9775$. Since this is the only critical point for $\delta \in [0.5, 1]$, it can be a maximum, minimum or point of inflexion. Using the same techniques as in section 2.2.3 for calculating the loss in sum rate due to separation failure for a specific channel, we compute

$$G_D(0.5, 0.5, 0.5, 0.5) = 0, \tag{2.290}$$

$$G_D(0.5, 0.9775, 0.5, 0.9775) = 0.0258, \tag{2.291}$$

$$G_D(0.5, 1, 0.5, 1) = 0.0207. \tag{2.292}$$

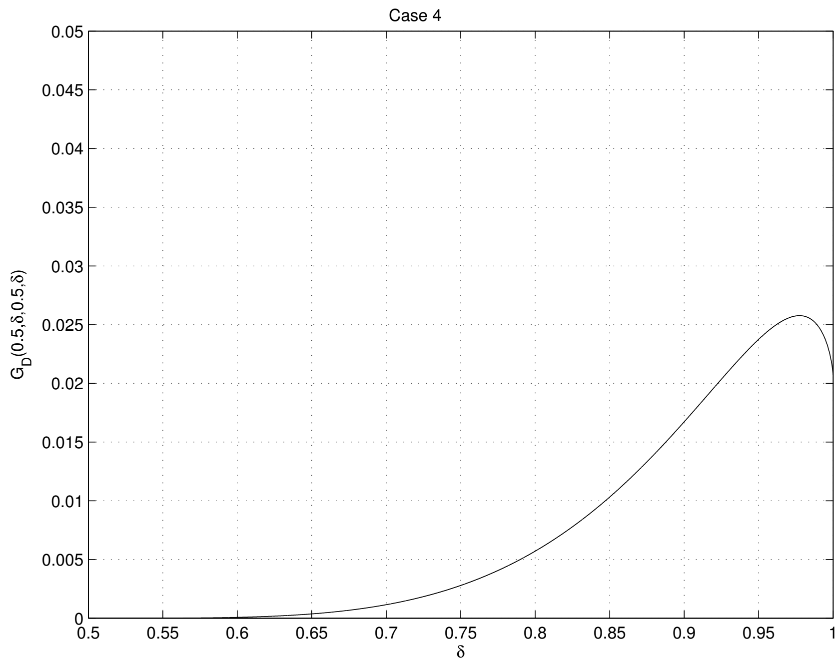

Figure 2-10: Case 4.

Since

$$G_D(0.5, 0.9775, 0.5, 0.9775) > G_D(0.5, 0.5, 0.5, 0.5), G_D(0.5, 1, 0.5, 1), \tag{2.293}$$

$G_D(0.5, 0.9775, 0.5, 0.9775)$ is the maximum. Thus,

$$G_D(0.5, \delta, 0.5, \delta) \leq G_D(0.5, 0.9775, 0.5, 0.9775) = 0.0258. \tag{2.294}$$

Since the maxima occur at the boundaries,

$$G_D(\epsilon, \delta, \epsilon, \delta) \leq \max\{G_D(\epsilon, 0.5, \epsilon, 0.5), G_D(\epsilon, 1, \epsilon, 1), G_D(0, \delta, 0, \delta), G_D(0.5, \delta, 0.5, \delta)\}. \tag{2.295}$$

Combining (2.276,2.282,2.288,2.294,2.295), we have for $\epsilon \in [0, 0.5]$, $\delta \in [0.5, 1]$

$$G_D(\epsilon, \delta, \epsilon, \delta) \leq 0.0776. \tag{2.296}$$

Thus, from (2.264,2.296) we have

$$G(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \leq 0.0776 \tag{2.297}$$

for $(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \in \mathcal{S}$.

Now, we know from section 2.2.4 that source-channel separation holds in regions $\mathcal{P}$, $\mathcal{Q}$ and $\mathcal{R}$. Moreover, from the symmetry in (2.60,2.61), we can say that for $\epsilon_0, \epsilon_1, \delta_0, \delta_1 \in [0, 1]$,

$$G(\epsilon_0, \epsilon_1, \delta_0, \delta_1) \leq 0.0776. \tag{2.298}$$

Tightness of the bound

We show that the bound on the maximum loss in sum rate due to separation failure is accurate to at least the second decimal place and is thus a very tight bound. For a channel specified by

$$\epsilon_0 = 0, \tag{2.299}$$

$$\delta_0 = 0.722, \tag{2.300}$$

$$\epsilon_1 = 0, \tag{2.301}$$

$$\delta_1 = 0.722, \tag{2.302}$$

we compute using techniques described in section 2.2.3

$$R_{sumJSCC}(0, 0.722, 0, 0.722) = 0.247, \tag{2.303}$$

$$R_{sumSSCC}(0, 0.722, 0, 0.722) = 0.1752, \tag{2.304}$$

$$\Rightarrow G(0, 0.722, 0, 0.722) = 0.0718. \tag{2.305}$$

From (2.305) we see that the bound given by (2.298) *is accurate to at least the second decimal position*. Thus, (2.298) is a very tight bound. Thus we see that *for noisy multiple access finite field adder channels over the binary field where noise is input-dependent, the maximum loss in sum rate by doing separate source-channel coding when separation fails is less that 0.0776 bit.*

Note that though there is a significantly high probability that source-channel separation does not hold for a channel over $\mathbb{F}_2$, the loss in sum rate due to separation failure is very small, especially, when the noise is low.

2.3 Source-channel separation when noise is independent of inputs

Let us first consider the binary noisy multiple access finite field adder channel shown in Figure 2-3 when noise does not depend upon the inputs. The transition matrices between $(X_a, X_a^{'})$ and $(X_b, X_b^{'})$ are now symmetric. Thus, we have

$$\epsilon_0 = \epsilon_1, \tag{2.306}$$

$$\delta_0 = \delta_1. \tag{2.307}$$

We define $\epsilon = \epsilon_1 = \epsilon_2$ and $\delta = \delta_0 = \delta_1$. From (2.23), we have $R_{sumSSCC}$ as

$$\begin{aligned} R_{sumSSCC}(\epsilon, \epsilon, \delta, \delta) &= \max_{p_1,p_2} \mathcal{H}[p_1 p_2(1-\epsilon-\delta+2\epsilon\delta) + p_1(1-p_2)(\epsilon+\delta-2\epsilon\delta) \\ &\quad + p_2(1-p_1)(\epsilon+\delta-2\epsilon\delta) + (1-p_1)(1-p_2)(1-\epsilon-\delta+2\epsilon\delta)] \\ &\quad - \mathcal{H}(\epsilon+\delta-2\epsilon\delta). \end{aligned} \tag{2.308}$$

Simplifying (2.308), we obtain

$$\begin{aligned} R_{sumSSCC}(\epsilon, \epsilon, \delta, \delta) &= \max_{p_1,p_2} \mathcal{H}[\epsilon+\delta-2\epsilon\delta + (1-2\delta)(1-2\epsilon)(p_1+p_2-2p_1p_2)] - \mathcal{H}[\epsilon+\delta-2\epsilon\delta]. \\ &= 1 - \mathcal{H}(\epsilon+\delta-2\epsilon\delta). \end{aligned} \tag{2.309}$$

We now compute $R_{sumJSCC}(\epsilon, \epsilon, \delta, \delta)$ from (2.45) as

$$\begin{aligned} R_{sumJSCC}(\epsilon, \epsilon, \delta, \delta) &= \max_{P_{00},P_{01},P_{10},P_{11}} \mathcal{H}[P_{00}(1-\epsilon-\delta+2\epsilon\delta) + P_{01}(\delta+\epsilon-2\epsilon\delta) \\ &\quad + P_{10}(\delta+\epsilon-2\delta\epsilon) + P_{11}(1-\epsilon-\delta+2\epsilon\delta] \\ &\quad - P_{00}\mathcal{H}(\epsilon+\delta-2\epsilon\delta) - P_{01}\mathcal{H}(\epsilon+\delta-2\epsilon\delta) \\ &\quad - P_{10}\mathcal{H}(\epsilon+\delta-2\epsilon\delta) - P_{11}\mathcal{H}(\epsilon+\delta-2\epsilon\delta). \end{aligned} \tag{2.310}$$

Simplifying, we get

$$\begin{aligned} R_{sumJSCC}(\epsilon, \epsilon, \delta, \delta) &= \max_{P_{00},P_{01},P_{10},P_{11}} \mathcal{H}[P_{00} + P_{11} + (P_{10} + P_{01} - P_{11} - P_{00})(\delta+\epsilon-\epsilon\delta)] \\ &\quad - \mathcal{H}(\epsilon+\delta-\epsilon\delta) \\ &= 1 - \mathcal{H}(\epsilon+\delta-\epsilon\delta). \end{aligned} \tag{2.311}$$

Expression (2.309) and (2.311) yield

$$R_{sumJSCC}(\epsilon, \epsilon, \delta, \delta) = R_{sumSSCC}(\epsilon, \epsilon, \delta, \delta), \tag{2.312}$$

for any $\epsilon, \delta \in [0, 1]$. Thus the criterion established in section 2.1 for separation to hold is satisfied and we have shown that separation holds when noise is not input dependent. We have thus proved the following theorem:

Theorem 2.3 *Source-channel separation holds for a two transmitter single receiver noisy multiple access finite field adder channel where the input symbols, output symbol and noise are elements of* $\mathbb{F}_2$ *and interference occurs in* $\mathbb{F}_2$ *with noise being independent of inputs.*

Theorem 2.3 establishes that separation holds for a noisy multiple access finite field adder channel over $\mathbb{F}_2$ when noise is independent of inputs. We will now show that separation holds for a channel over $\mathbb{F}_{2^k}$ for any $1 \leq k$, as long as noise is input-independent.

Let us look at the channel where the inputs X_a, X_b, noise Z and output Y are elements of $\mathbb{F}_{2^k}$. Noise is independent of inputs and $Y = X_a + X_b + Z$ with addition over $\mathbb{F}_{2^k}$. Consider the entropy of $Y = X_a + X_b + Z$. Since $X_a + X_b$ and Z are independent,

$$\begin{aligned} H(X_a + X_b) &= H(X_a + X_b|Z), \\ &= H(Y|Z), \\ &\leq H(Y). \end{aligned} \tag{2.313}$$

As addition is over a finite field, if X_a and X_b are independent and have a uniform distribution, then $X_a + X_b$ has a uniform distribution. When X_a and X_b are correlated, we can let $P_{X_aX_b}(x_a, x_b) = 2^{-2k}$ for all $(x_a, x_b) \in (\mathbb{F}_{2^k})^2$. This will make the probability distribution of $X_a + X_b$ uniform. Therefore, whether X_a and X_b are correlated or not, $X_a + X_b$ can always be made uniform and has a maximum entropy of k bits. This can be represented as

$$\max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(X_a + X_b) = \max_{P_{X_aX_b}(x_a,x_b)} H(X_a + X_b) = k. \tag{2.314}$$

From (2.313), we have

$$\max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(X_a + X_b) \leq \max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(Y), \tag{2.315}$$

$$\max_{P_{X_a X_b}(x_a,x_b)} H(X_a + X_b) \leq \max_{P_{X_a X_b}(x_a,x_b)} H(Y). \tag{2.316}$$

Since, $H(Y) \leq k$, we have

$$\max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(Y) = \max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(X_a + X_b) = k, \tag{2.317}$$

$$\max_{P_{X_a X_b}(x_a,x_b)} H(Y) = \max_{P_{X_a X_b}(x_a,x_b)} H(X_a + X_b) = k. \tag{2.318}$$

The maximum sum rate by separate source-channel coding is therefore

$$\begin{aligned} R_{sumSSCC} &= \max_{P_{X_a}(x_a)P_{X_b}(x_b)} I(X_a, X_b; Y), && (2.319) \\ &= \max_{P_{X_a}(x_a)P_{X_b}(x_b)} [H(Y) - H(Y|X_a, X_b)], && (2.320) \\ &= \max_{P_{X_a}(x_a)P_{X_b}(x_b)} [H(Y) - H(Z|X_a, X_b)], && (2.321) \\ &= \max_{P_{X_a}(x_a)P_{X_b}(x_b)} H(Y) - H(Z), && (2.322) \\ &= k - H(Z). && (2.323) \end{aligned}$$

Equation (2.322) holds since the noise is independent of the inputs and (2.323) follows from (2.317). When we do joint source-channel coding for the same channel we obtain

$$\begin{aligned} R_{sumJSCC} &= \max_{P_{X_a X_b}(x_a,x_b)} I(X_a, X_b; Y), && (2.324) \\ &= \max_{P_{X_a X_b}(x_a,x_b)} [H(Y) - H(Y|X_a, X_b)] && (2.325) \\ &= \max_{P_{X_a X_b}(x_a,x_b)} [H(Y) - H(Z|X_a, X_b)], && (2.326) \\ &= \max_{P_{X_a X_b}(x_a,x_b)} H(Y) - H(Z), && (2.327) \\ &= k - H(Z), && (2.328) \end{aligned}$$

where (2.327) holds since noise is independent of inputs and (2.328) follows from (2.318). Combining (2.323) and (2.328), we obtain

$$R_{sumSSCC} = R_{sumJSCC}. \tag{2.329}$$

Equation (2.329) satisfies the criterion established in section 2.1 for separation to hold. We have thus proved the following theorem which generalizes Theorem 2.3 to arbitrary field size:

Theorem 2.4 *Source-channel separation holds for a two transmitter single receiver noisy multiple access finite field adder channel where the input symbols, output symbol and noise are elements of* $\mathbb{F}_{2^k}$ *for* $1 \leq k$ *and interference occurs in* $\mathbb{F}_{2^k}$ *with noise being independent of inputs.*

Note that the noise-free multiple access finite field adder channel is a special case of the noisy finite field adder channel with input-independent noise.

Chapter 3

Time-Slotted Noise-Free Multiple Access Networks over Finite Fields

In this chapter, we will consider a time-slotted noise-free multiple access finite field adder channel. The transmitted elements X_a, X_b, and received element Y are from $\mathbb{F}_{2^k}$, for $1 \leq k$ and

$$Y = X_a + X_b. \tag{3.1}$$

Multiple access interference is additive over $\mathbb{F}_{2^k}$. First, we develop a single-slot model for this multiple access channel. We determine the capacity region and maximum code rate and study the dependance of these quantities on field size. We present a systematic code construction, which we show achieves maximum code rate and capacity. We show that codes that achieve the maximum code rate also achieve capacity. Next, we look at the performance of systematic random codes and obtain conditions under which random codes achieve maximum code rate and capacity. We provide explicit expressions for error probabilities, which are functions of code length, thus establishing the strong coding theorem for this channel. We also consider bursty transmissions, when transmitters transmit according to a Bernoulli process. We propose coding techniques to maximize code rate. We show that, when the information vectors at the input to the channel encoders have the same size, maximum expected code rate is achieved by adding redundancy at the transmitter with a higher proba-

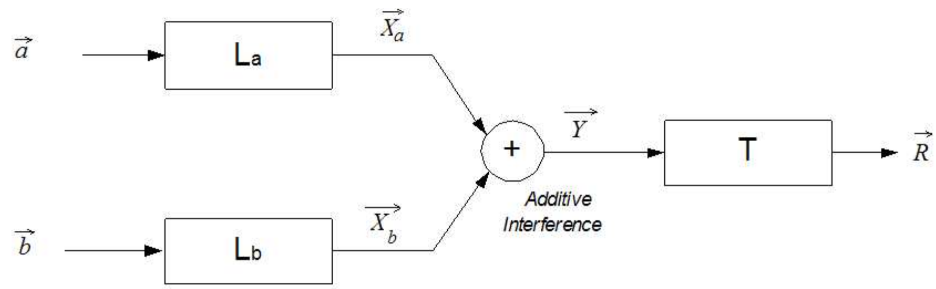

Figure 3-1: Single Slot Model for the Noise-Free Multiple Access Finite Field Adder Channel.

bility of transmission and not adding any redundancy at the transmitter with lower probability of transmission.

3.1 Single Slot Model

We consider a discrete time channel. The channel is time-slotted and we consider transmissions over the length of a slot or slot duration. Figure 3-1 shows a single slot model of the noise-free finite field multiple access channel. Information codewords $\vec{U}'$ and $\vec{V}'$, as described in Figure 2-2, coming out of the source coders at transmitters a and b in one slot duration will be represented as $\vec{a}$ and $\vec{b}$, respectively, in this chapter. $\vec{a}$ and $\vec{b}$ are thus vectors of sizes n_a and n_b respectively, which represent independent information. The elements of $\vec{a}$ and $\vec{b}$ are in $\mathbb{F}_{2^k}$, for $1 \leq k$. In Chapter 2, we showed in Theorem 2.4 that source-channel separation holds for a noise-free multiple access finite field adder channel. Thus, the scheme of separate source and channel coding is optimal. For this model, all operations, matrices and vectors are in $\mathbb{F}_{2^k}$. We will refer to $\vec{a}$ and $\vec{b}$ as transmit vectors and assume that $n_a \geq n_b$. (Otherwise, the arguments still hold with $\vec{a}$ and $\vec{b}$ interchanged.) L_a is a $(n_a + n_b) \times n_a$ size matrix and L_b is a $(n_a + n_b) \times n_b$ size matrix. These are the generator matrices for the channel codes at a and b respectively. $\vec{X}_a$ and $\vec{X}_b$ are the codewords that are sent over the channel and they interfere additively over $\mathbb{F}_{2^k}$. At the decoder, matrix T having a dimension of $(m_1 + m_2) \times (n_a + n_b)$ decodes the received vector to generate a subset of $\vec{a}$ and $\vec{b}$.

$\vec{R}$ is this decoded output containing m_1 elements of $\vec{a}$ and m_2 elements of $\vec{b}$.

Let m_a and m_b be the increase in the size of vectors $\vec{a}$ and $\vec{b}$, respectively, due to channel coding. We will denote l_a and l_b as the lengths of the vectors obtained by channel coding on $\vec{a}$ and $\vec{b}$ respectively. In general, $l_a \neq l_b$ and so both transmitters may not transmit for the entire slot duration. However, at least one transmitter will transmit for the whole slot duration. Therefore, the slot length is given by

$$S = \max(l_a, l_b) \textit{ symbols.} \tag{3.2}$$

Also,

$$l_a = n_a + m_a, \tag{3.3}$$

$$l_b = n_b + m_b. \tag{3.4}$$

3.2 Capacity Region and Maximum Code Rate

In this section, we derive the capacity region and maximum code rate over a very long slot or many slots. We first establish the capacity region and then derive the maximum code rate.

3.2.1 Capacity Region

We know from [21, 22] that the multiple access capacity region is the closure of the convex hull of all (R_a, R_b) satisfying

$$R_a \leq I(X_a; Y|X_b), \tag{3.5}$$

$$R_b \leq I(X_b; Y|X_a), \tag{3.6}$$

$$R_{sum} = R_a + R_b \leq I(X_a, X_b; Y). \tag{3.7}$$

The mutual information expressions can be simplified to

$$\begin{aligned} I(X_a; Y|X_b) &= H(Y|X_b) - H(Y|X_a, X_b) && (3.8) \\ &= H(X_a), && (3.9) \end{aligned}$$

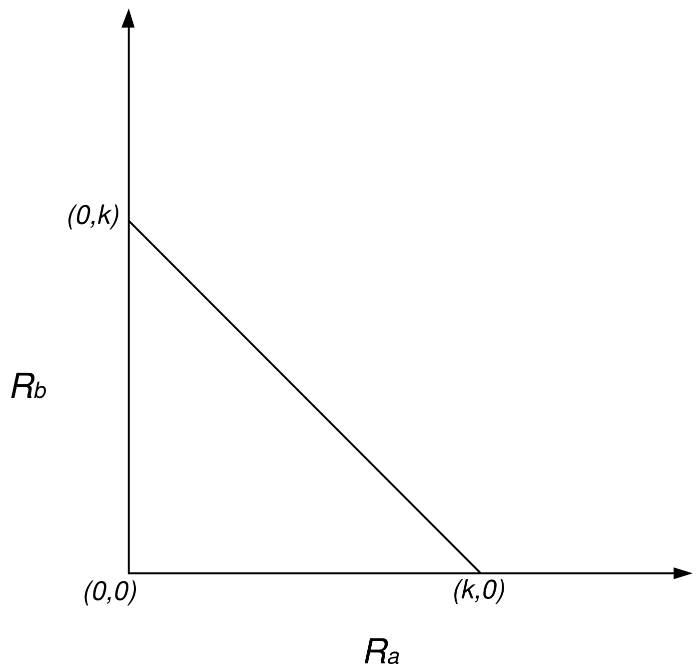

Figure 3-2: Capacity region of a noise-free multiple access finite field adder channel over a field of size 2^k.

$$\begin{aligned} I(X_b;Y|X_a) &= H(Y|X_a) - H(Y|X_a,X_b) && (3.10)\\ &= H(X_b), && (3.11)\\ I(X_a,X_b;Y) &= H(Y) - H(Y|X_a,X_b) && (3.12)\\ &= H(X_a+X_b). && (3.13)\end{aligned}$$

Uniform distribution of X_a and X_b maximizes R_a, R_b and R_{sum}. The multiple access capacity region is therefore the closure of the convex hull of all (R_a, R_b) satisfying

$$R_a \leq k, \tag{3.14}$$

$$R_b \leq k, \tag{3.15}$$

$$R_{sum} = R_a + R_b \leq k, \tag{3.16}$$

where the rates are in bits per channel use. The transmission rates will always be specified in bits per channel use. The capacity region is shown in Figure 3-2.

3.2.2 Maximum Code Rate

We now find the maximum code rate for this channel. Transmitters a and b transmit n_a and n_b information symbols per slot respectively and the codewords transmitted have a length of l_a and l_b symbols respectively. The transmission rates are

$$R_a = \frac{kn_a}{S}, \tag{3.17}$$

$$R_b = \frac{kn_b}{S}, \tag{3.18}$$

$$R_{sum} = \frac{k(n_a + n_b)}{S}. \tag{3.19}$$

From (3.16, 3.19), we obtain

$$n_a + n_b \leq S. \tag{3.20}$$

The code rate is a dimensionless quantity and is given by

$$\begin{aligned}
C_{rate} &= \frac{n_a + n_b}{l_a + l_b} && (3.21)\\
&= \frac{n_a + n_b}{\min(l_a, l_b) + \max(l_a, l_b)} && (3.22)\\
&= \frac{n_a + n_b}{\min(l_a, l_b) + S} && (3.23)\\
&\leq \frac{n_a + n_b}{n_b + n_a + n_b}. && (3.24)
\end{aligned}$$

Equation (3.23) is due to (3.2). Expression (3.24) follows from (3.20) and the fact that $n_b \leq \min(l_a, l_b)$. Thus, we have

$$C_{rate} \leq \frac{n_a + n_b}{n_a + 2n_b}. \tag{3.25}$$

We obtain an important lemma:

Lemma 3.1 *The capacity region of a two-transmitter noise-free finite field adder channel grows logarithmically with the size of the field but the code rate remains the same for all field sizes.*

Proof: From (3.14), (3.15) and (3.16), we see that the capacity region grows logarithmically with field size (k). However, (3.25) shows that the code rate is invariant to k. Thus we have proved the lemma. □

3.3 Code Construction

We now focus on construction of codes that achieve the capacity and maximum code rate for the noise-free multiple access finite field adder channel. We will call these codes that achieve both the maximum code rate and capacity as *optimal codes*. The

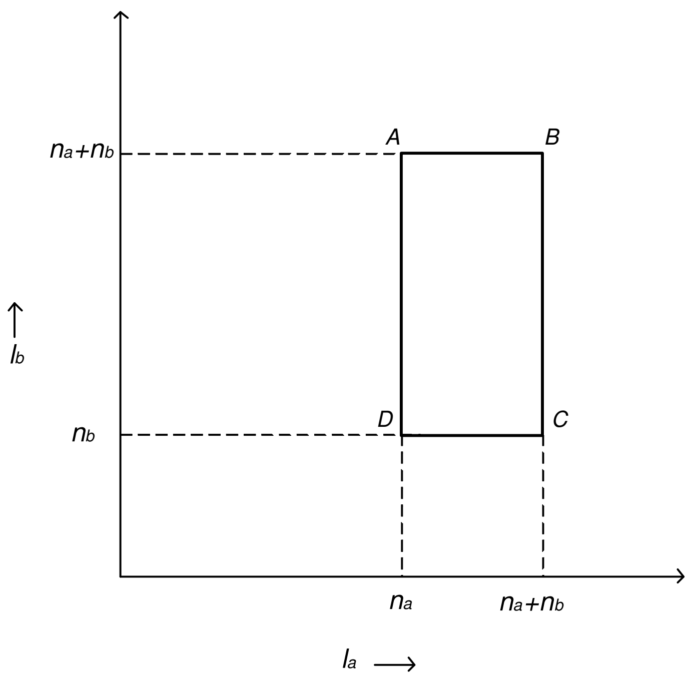

Figure 3-3: Region of Analysis.

code rate, transmission rates R_a, R_b, and sum rate R_{sum} for our model are given by

$$C_{rate} = \frac{m_1 + m_2}{l_a + l_b}, \tag{3.26}$$

$$R_a = \frac{km_1}{\max(l_a, l_b)}, \tag{3.27}$$

$$R_b = \frac{km_2}{\max(l_a, l_b)}, \tag{3.28}$$

$$R_{sum} = \frac{k(m_1 + m_2)}{\max(l_a, l_b)}. \tag{3.29}$$

We now construct a code that will be shown to be an optimal one. Note that there are many codes that can be optimal. We describe one such construction and prove its optimality.

By definition, $l_a \geq n_a$ and $l_b \geq n_b$. When $l_a = l_b = n_a + n_b$, all elements can be recovered after multiple access interference at the receiver. Thereby, we can safely reduce the region for finding optimal codes to the region $ABCD$ shown in Figure 3-3. All points outside $ABCD$ will have a lower code rate as the number of received elements remains same for increasing l_a and l_b. Hence, we confine our analysis of finding optimal codes to the region $ABCD$.

In this region $n_a \leq l_a \leq n_a + n_b$ and $n_b \leq l_b \leq n_a + n_b$ which makes $0 \leq m_a \leq n_b$, $0 \leq m_b \leq n_a$, $0 \leq m_1 \leq n_a$ and $0 \leq m_2 \leq n_b$. We also have the following relations:

$$\vec{X}_a = L_a \vec{a}, \tag{3.30}$$

$$\vec{X}_b = L_b \vec{b}, \tag{3.31}$$

$$\vec{Y} = L_a \vec{a} + L_b \vec{b}, \tag{3.32}$$

$$\vec{R} = T\vec{Y} = (TL_a)\vec{a} + (TL_b)\vec{b}. \tag{3.33}$$

Let $W_a = (TL_a)$ and $W_b = (TL_b)$. *We define a 1row as a row vector having only one non-zero element.* Since $\vec{R}$ contains m_1 elements of $\vec{a}$ and m_2 elements of $\vec{b}$, W_a should be a $(m_1 + m_2) \times n_a$ size matrix with m_1 1rows, W_b a $(m_1 + m_2) \times n_b$ size matrix with m_2 1rows and the 1row positions for these matrices should not overlap. Let $W = \left[\begin{array}{c|c} W_a & W_b \end{array} \right]$ and $L = \left[\begin{array}{c|c} L_a & L_b \end{array} \right]$. Thus, W should have $m_1 + m_2$ unique 1rows. Looking at the relations obtained from our model, we see that W is generated by receiver matrix T operating on L and the rows of W are linear combinations of the rows of L. By definition, W consists only of 1rows. For given L, we need to find the maximum number of 1rows in W that can be generated by linear combinations of the rows of L. This maximizes $m_1 + m_2$ for given $l_a + l_b$ which in turn maximizes the code rate and also specifies L_a, L_b and T. Therefore, codes that achieve the maximum code rate and capacity are found by jointly optimizing the encoder and decoder matrices. In our further discussion, $I_{k \times k}$ will represent a $k \times k$ identity matrix and $0_{p_1 \times p_2}$ a $p_1 \times p_2$ null matrix. We now prove the following lemma:

Lemma 3.2 *Let B_1 and B_2 be diagonal square matrices of size n_b with all diagonal elements non-zero. If s unique 1rows are inserted into a matrix J of the form*

$\left[\begin{array}{c} B_1 0_{n_b\times(n_a-n_b)} B_2 \end{array}\right]$ *to obtain matrix* C *such that the number of independent row vectors in* C *is* $s+n_b$ *and* k *1rows have their non-zero element in* $[n_b+1, n_a]$*, then the maximum number of 1rows possible by any linear combination of the rows of* C *is* $2s-k$*.*

Proof: The inserted 1rows that are non-zero in positions $[n_b+1, n_a]$ cannot give rise to any other 1row in C since, all the rows of J are 0 in that position. Any other inserted 1row can be combined with a row vector in J to give a unique 1row in C since all the row vectors of C are independent. Thus, the $s-k$ 1rows whose non-zero elements are not in the interval $[n_b+1, n_a]$, can generate $2(s-k)$ 1rows in C. These arguments are valid for $s \in [0, n_a]$ and $k \in [0, n_a - n_b]$. The total number of 1rows that can be generated by the s inserted 1rows is thus $2(s-k)+k = 2s-k$. □

Using this lemma, we know what we get by adding redundancy at the transmitters. Note that the proof depends only on B_1 and B_2's being diagonal with all diagonal elements being non-zero. There is no constraint on the values of the diagonal elements. Hence, we will set $B_1 = B_2 = I_{n_b\times n_b}$ and the non-zero element in a 1row to be 1.

Theorem 3.1 *Optimal codes are not contained in the region* $0 < m_b \leq n_a - n_b$*.*

Proof: Let $P = \left[\begin{array}{c|c} I_{n_a\times n_a} & \begin{array}{c} I_{n_b\times n_b} \\ 0_{(n_a-n_b)\times n_b} \end{array} \end{array}\right]$. We form matrix G by inserting 1rows to P in any row position. Adding redundancy of m_a generates m_a 1rows in G and reduces the number of 1rows the redundancy m_b can generate by $m_a + n_a - n_b$. This implies that

$$\begin{aligned} m_b &> m_a + n_a - n_b && (3.34)\\ &> \min(m_a) + n_a - n_b && (3.35)\\ &> n_a - n_b. && (3.36) \end{aligned}$$

Thus $m_b = 0$ (no redundancy added at transmitter b) or $m_b > n_a - n_b$, which completes the proof. □

Theorem 3.2 *Codes that achieve the maximum code rate and capacity are not contained on the line* $m_a = m_b - [n_a - n_b]$.

Proof: Let P be defined as before. Coding in this region results in the insertion of at least a row vector to P which is not a 1row. The inserted rows that are not 1rows contain a 1 in the first n_a positions and a 1 in the last n_b positions. The other elements are 0. The number of 1rows determine the size of the subset recoverable at the receiver. But in this case, the rows that are not 1rows increase redundancy but do not give us a larger subset. Thus, we should not insert any row that is not a 1row. This is not possible on the line $m_a = m_b - [n_a - n_b]$. Hence, this line does not contain optimal codes. □

3.3.1 Structure of Generator Matrices

We now develop the structure of the generator matrices for the encoders at the two transmitters. We will use *systematic codes* and show that they are optimal in terms of achieving maximum code rate and capacity. Using the results of Theorems 3.1 and 3.2, we get two cases:

Case 1: $m_b = 0$.

In this case, redundancy is added at transmitter a only. Let the redundancy added to $\vec{a}$ be m_a. This corresponds to appending m_a 1row vectors to P such that the 1 in each of these vectors lie in the first n_a positions and the resulting matrix consists of independent rows. Using Lemma 3.2, the maximum number of 1 rows that can be generated is $2(m_a + n_a - n_b) - (n_a - n_b) = 2m_a + n_a - n_b$. Now, $[n_a - n_b] + m_a$ of the 1rows generated will have their 1 in the first n_a positions and m_a 1rows will have their 1 in the last n_b position. Thus we have:

$$m_1 = [n_a - n_b] + m_a, \tag{3.37}$$

$$m_2 = m_a, \tag{3.38}$$

$$L_a = \begin{bmatrix} I_{n_a \times n_a} \\ M_{m_a \times n_a} \\ 0_{(n_b - m_a) \times n_a} \end{bmatrix}, \tag{3.39}$$

$$L_b = \begin{bmatrix} I_{n_b \times n_b} \\ 0_{n_a \times n_b} \end{bmatrix}, \tag{3.40}$$

where, $m_b = 0$, $0 \leq m_a \leq n_a$ and M is a matrix containing 1rows.

Case 2 : $n_a - n_b < m_b$.

Let $m_b = n_a - n_b + k$, where $0 < k \leq n_b$. In this case redundancies of m_a and m_b are added to $\vec{a}$ and $\vec{b}$ respectively.

When $m_a < k$, m_a 1rows are appended to P such that each 1row contains a 1 in the first n_a positions. Then, $k - m_a$ 1rows are appended to the matrix resulting from the previous step so that 1 is contained in one of the last n_b positions. The 1rows are appended such that all rows of the resulting matrix are independent. Using Lemma 3.2, we see that maximum number of 1rows that can be generated is given by $2m_b - [n_a - n_b]$. There are m_b 1rows with 1 in the first n_a positions and $m_b - [n_a - n_b]$ 1rows with 1 in the last n_b positions. Thus we have:

$$m_1 = m_b, \tag{3.41}$$

$$m_2 = m_b - [n_a - n_b], \tag{3.42}$$

$$L_a = \begin{bmatrix} I_{n_a \times n_a} \\ \Lambda_{m_a \times n_a} \\ 0_{(n_b - m_a) \times a} \end{bmatrix}, \tag{3.43}$$

$$L_b = \begin{bmatrix} I_{n_b \times n_b} \\ 0_{(n_a - n_b + m_a) \times n_b} \\ \Delta_{(m_b - m_a - [n_a - n_b]) \times n_b} \\ 0_{(n_a - m_b) \times n_b} \end{bmatrix}, \tag{3.44}$$

where, $[n_a - n_b] < m_b \leq n_a$ and $0 \leq m_a < m_b - [n_a - n_b]$. Λ and Δ are matrices containing unique 1rows.

When $m_a > k$, coding involves appending k 1rows to P such that each row contains the 1 in the last n_b positions. Then, $m_a - k$ 1rows are appended to the matrix resulting from the previous step so that a 1 is contained in the first n_a positions for each vector. 1rows are appended such that the resulting matrix consists of independent rows. Using Lemma 3.2, we see that the maximum number of 1rows that can be generated is given by $2m_a + [n_a - n_b]$. There are m_a 1rows with 1 in the last n_b positions and $m_a + [n_a - n_b]$ 1rows with 1 in the first n_a positions. Thus, we have:

$$m_1 = m_a + [n_a - n_b], \tag{3.45}$$

$$m_2 = m_a, \tag{3.46}$$

$$L_a = \begin{bmatrix} I_{n_a \times n_a} \\ 0_{(m_b - [n_a - n_b]) \times n_a} \\ S_{(m_a - m_b + [n_a - n_b]) \times n_a} \\ 0_{(n_b - m_a) \times n_a} \end{bmatrix}, \tag{3.47}$$

$$L_b = \begin{bmatrix} I_{n_b \times n_b} \\ 0_{(n_a - n_b) \times n_b} \\ K_{(m_b - [n_a - n_b]) \times n_b} \\ 0_{(n_b - m_b + n_a - n_b) \times n_b} \end{bmatrix}, \tag{3.48}$$

where, $[n_a - n_b] < m_b \leq n_a$ and $0 \leq m_a < m_b - [n_a - n_b]$. S and K are matrices containing 1rows.

3.3.2 Regions

The regions over which optimal codes exist can be now described and are shown in Figure 3-4.

Region 1: $n_b \leq l_a \leq n_a + n_b$ and $l_b = n_b$,

$$m_1 = m_a + [n_a - n_b], \tag{3.49}$$

$$m_2 = m_a, \tag{3.50}$$

$$C_{rate-R_1} = \frac{2m_a + [n_a - n_b]}{n_a + n_b + m_a}. \tag{3.51}$$

Region 2: $l_b < l_a \leq n_a + n_b$ and $n_a < l_b \leq n_a + n_b$,

$$m_1 = m_a + [n_a - n_b], \tag{3.52}$$

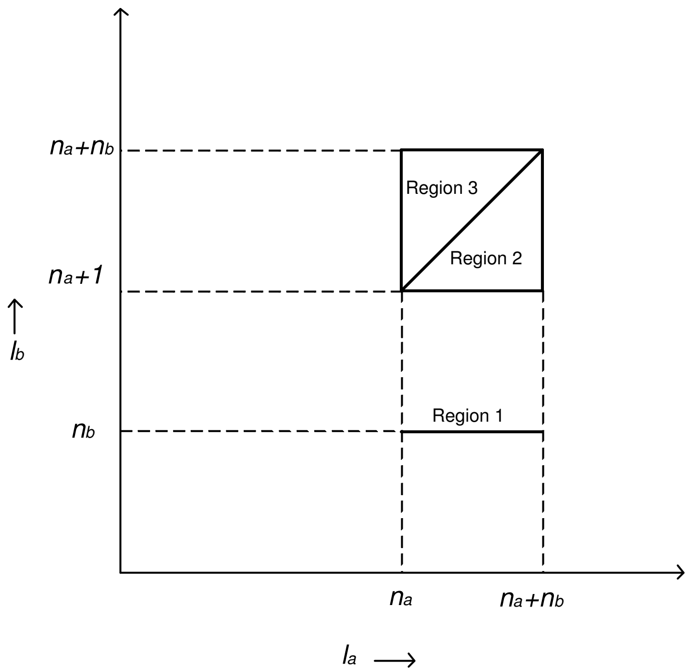

Figure 3-4: Gross (un-optimized) regions over which optimal codes exist.

$$m_2 = m_a, \tag{3.53}$$

$$C_{rate-R_2} = \frac{2m_a + [n_a - n_b]}{n_a + n_b + m_a + m_b}. \tag{3.54}$$

Region 3: $n_a \leq l_a < l_b$ and $n_a < l_b \leq n_a + n_b$,

$$m_1 = m_b, \tag{3.55}$$

$$m_2 = m_b - [n_a - n_b], \tag{3.56}$$

$$C_{rate-R_3} = \frac{2m_b - [n_a - n_b]}{n_a + n_b + m_a + m_b}. \tag{3.57}$$

3.3.3 Optimized Regions

Theorem 3.3 *To achieve the maximum code rate, it suffices to add redundancy to only one vector.*

Proof : We see from Figure 3-4 that in Region 1 and Region 2, m_1 and m_2 do not depend upon m_b. Thus, for higher code rate, m_b should be kept as low as possible.

We thus set $m_b = 0$ for Region 1 and $m_b = n_a - n_b + 1$ for Region 2. As $n_a \geq n_b$,

$$C_{rate-R_1} \geq C_{rate-R_2}. \tag{3.58}$$

Thus optimal codes cannot be in Region 2, as this region does not contain codes with higher code rate than Region 1. Hence, we do not consider this region in our further search for optimal codes. In Region 3, m_1 and m_2 do not depend on m_a. Therefore, it is best to keep m_a at its lowest, i.e. $m_a = 0$. We thus consider codes over Region 1 and Region 3, where we set $m_b = 0$ and $m_a = 0$, respectively. Thus, in order to achieve the optimal code rate, it suffices to add redundancy at only one transmitter.□

The optimized regions are shown in Figure 3-5.

Region A : $n_a \leq l_a \leq n_a + n_b$ and $l_b = n_b$

$$m_1 = m_a + [n_a - n_b], \tag{3.59}$$

$$m_2 = m_a \tag{3.60}$$

$$C_{rate-R_A} = \frac{2m_a + [n_a - n_b]}{n_a + n_b + m_a}. \tag{3.61}$$

Region B: $l_a = n_a$ and $n_a + 1 \leq l_b \leq n_a + n_b$

$$m_1 = m_b, \tag{3.62}$$

$$m_2 = m_b - [n_a - n_b], \tag{3.63}$$

$$C_{rate-R_B} = \frac{2m_b - [n_a - n_b]}{n_a + n_b + m_b}. \tag{3.64}$$

3.3.4 Achieving the capacity region and maximum code rate

From Theorem 3.3, we see that in order to achieve the maximum code rate, it suffices to add redundancy at only one transmitter. Let the redundancy be m. In Region A, $0 \leq m \leq n_b$ and

$$C_{rate-R_A} = \frac{2m + [n_a - n_b]}{n_a + n_b + m}. \tag{3.65}$$

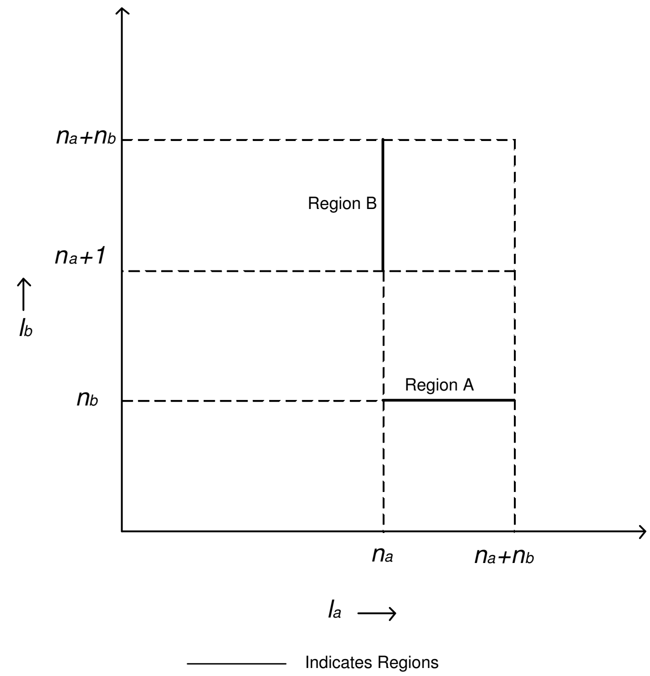

Figure 3-5: Optimized regions.

In Region B, $(n_a - n_b + 1) \leq m \leq n_b$ and

$$C_{rate-R_B} = \frac{2m - [n_a - n_b]}{n_a + n_b + m}. \tag{3.66}$$

Case 1 : $n_a > n_b$. When $0 \leq m \leq n_a - n_b$ Region B is excluded and Region A provides the only solution. For all other m, $C_{rate-R_A} > C_{rate-R_B}$. Thus, Region A always provides a higher code rate than Region B. From the code rate equations derived earlier, we see that the maximum code rate is obtained when m is largest, i.e. $m = n_b$. Therefore, $(m_a, m_b) = (n_b, 0)$ is the optimal point. This corresponds to $(l_a, l_b) = (n_a + n_b, n_b)$. Thus, for obtaining the maximum code rate, we add redundancy to only the larger transmit vector and the size of the redundancy is the size of the smaller transmit vector. The code rate is

$$C_{rate} = \frac{n_a + n_b}{n_a + 2n_b}. \tag{3.67}$$

Case 2: $n_a = n_b = n$. Here, for given m, both regions give the same code rate and we can add redundancy to any of the two vectors. A symmetry exists about the line $l_a = l_b$ and there are two optimal points. Code rate is maximum when m is

maximum, i.e. $m = n$. These points are $(m_a, m_b) \in \{(0, n), (n, 0)\}$ corresponding to $(l_a, l_b) \in \{(2n, n), (n, 2n)\}$. In this case, coding results in the size of redundancy being equal to the transmit vector size and the code rate is $2/3$.
The transmission rates of the code are

$$R_a = \frac{kn_a}{n_a + n_b}, \tag{3.68}$$

$$R_b = \frac{kn_b}{n_a + n_b}, \tag{3.69}$$

$$R_{sum} = R_a + R_b = k. \tag{3.70}$$

We see from (3.67,3.70) that this code achieves the maximum code rate and capacity for this channel and is thus an optimal code. Moreover, this code obeys the property (that will be proved in Theorem 3.4) of any maximum code rate achieving code, i.e. no redundancy is added to the smaller transmit vector.

We now prove the following theorem:

Theorem 3.4 *For a noise-free multiple access finite field adder channel, codes achieve the maximum code rate if and only if they are capacity achieving and no redundancy is added to the smaller transmit vector.*

Proof: We first prove the forward part. Let a code be capacity approaching without redundancy being added to the smaller transmit vector. From the code construction described earlier, we have shown that such codes exist. We therefore have the following relations:

$$S = n_a + n_b, \tag{3.71}$$

$$n_b = \min(l_a, l_b), \tag{3.72}$$

$$C_{rate} = \frac{n_a + n_b}{\min(l_a, l_b) + S}. \tag{3.73}$$

$$\Rightarrow C_{rate} = \frac{n_a + n_b}{n_a + 2n_b}. \tag{3.74}$$

Therefore, capacity approaching codes with no redundancy added to the smaller transmit vector achieve the maximum code rate. This completes the forward part of the proof.

We prove the reverse part now. A code that achieves the maximum code rate must meet the inequality in (3.24) with equality. From the code construction described in the beginning of this section, we know that codes that achieve the maximum code rate exist. Therefore, maximum code rate achieving codes must satisfy

$$S = n_a + n_b, \tag{3.75}$$

$$\min(l_a, l_b) = n_b. \tag{3.76}$$

Hence, maximum code rate achieving codes achieve capacity and do not add redundancy to the smaller transmit vector. The reverse part of the proof is now complete and we have proved the theorem. □

3.4 Random Coding

In the previous section, we described a systematic optimal code construction, i.e., it was capacity and maximum code rate achieving. In this section, we focus on the performance of systematic random codes and show their asymptotic optimality.

Theorem 3.5 *For a two transmitter noise-free multiple access finite field adder channel over $\mathbb{F}_{2^k}$, as the codeword lengths or field size tends to infinity, a random code becomes optimal with probability tending to 1 exponentially with codeword length and field size.*

We will prove the theorem in two ways.

Proof 1: Let us code randomly over the larger transmit vector, $\vec{a}$, with a systematic random code to generate redundancy, $\vec{g}$, of n_b elements. We split $\vec{a}$ into two vectors, $\vec{a_1}$ and $\vec{a_2}$, where $\vec{a_1}$ is a vector representing the first n_b elements of $\vec{a}$ and $\vec{a_2}$ is a vector representing the last $n_a - n_b$ elements of $\vec{a}$. Thus, $\vec{a} = \begin{bmatrix} \vec{a_1} & \vec{a_2} \end{bmatrix}$. The vectors (codewords) coming out of the encoders are

$$\vec{X_a} = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \vec{g} \end{bmatrix}, \tag{3.77}$$

$$\vec{X_b} = \begin{bmatrix} \vec{b} \end{bmatrix}. \tag{3.78}$$

Now, consider any vector of length n containing elements from $\mathbb{F}_{2^{\mathrm{k}}}$ as a single element belonging to $\mathbb{F}_{2^{\mathrm{kn}}}$. Thus, for any x_1, x_3 and $y_1 \in \mathbb{F}_{2^{\mathrm{kn_b}}}$ and $x_2 \in \mathbb{F}_{2^{\mathrm{k(n_a-n_b)}}}$, we can consider $x_1 \equiv \vec{a_1}$, $x_2 \equiv \vec{a_2}$, $x_3 \equiv \vec{g}$, $y_1 \equiv \vec{b}$ and x_3 is generated from x_1 as

$$x_3 = \alpha x_1 \tag{3.79}$$

where, α is randomly (uniformly) picked from $\mathbb{F}_{2^{\mathrm{kn_b}}}$. *This multiplication by a random field element represents the random coding.* After multiple access interference, we get $\vec{Y}$ of length $n_a + n_b$ such that $\vec{Y} = \vec{X_a} + \vec{X_b}$. $\vec{Y} \equiv \begin{bmatrix} x_1 + y_1 & x_2 & x_3 \end{bmatrix}$, where addition is over $\mathbb{F}_{2^{\mathrm{kn_b}}}$. $x_1 + y_1$ corresponds to the first n_b, x_2 the next $n_a - n_b$ and x_3 to the last n_b elements of $\vec{Y}$. We need to get x_1, x_2 and y_1 after decoding. We get x_2 as it does not suffer multiple access interference. For $m \in \mathbb{F}_{2^{\mathrm{kn_b}}}$, denote $m = x_1 + y_1$, with addition in $\mathbb{F}_{2^{\mathrm{kn_b}}}$ and m represents the first n_b elements of $\vec{Y}$.

We know from the property of finite fields that each non-zero element in the field has an unique inverse. Thus, for any $\alpha \in \mathbb{F}_{2^{\mathrm{kn_b}}}$, $\alpha \neq 0$, $\exists \alpha^* \in \mathbb{F}_{2^{\mathrm{kn_b}}}$ such that

$$\alpha^* \alpha = 1. \tag{3.80}$$

x_3 corresponds to the last n_b elements of $\vec{Y}$. At the decoder, we get x_1 and y_1 from a two step decoding process. In the first step we get y_1:

$$y_1 = m + \alpha^* x_3. \tag{3.81}$$

Using y_1, we get x_1 in the next step :

$$x_1 = m + y_1, \tag{3.82}$$

where all operations are in $\mathbb{F}_{2^{\mathrm{kn_b}}}$. Thus, we recover x_1, x_2 and y_1 and thereby completely recover $\vec{a}$ and $\vec{b}$. The transmission rates and code rates are given as:

$$R_a = \frac{kn_a}{n_a + n_b}, \tag{3.83}$$

$$R_b = \frac{kn_b}{n_a + n_b}, \tag{3.84}$$

$$R_{sum} = k, \tag{3.85}$$

$$C_{rate} = \frac{n_a + n_b}{n_a + 2n_b}. \tag{3.86}$$

Thus, we achieve the capacity and maximum code rates given by $(3.16, 3.25)$ as long as (3.80) holds. Equation (3.80) holds iff $\alpha \neq 0$. The probability P_R, that a random code is an optimal code, is given by

$$P_R = 1 - P(\alpha = 0). \tag{3.87}$$

Since α is chosen randomly (uniformly) from $\mathbb{F}_{2^{\mathrm{kn_b}}}$,

$$P(\alpha = 0) = \frac{1}{2^{kn_b}} \quad \Rightarrow P_R = 1 - \frac{1}{2^{kn_b}}. \tag{3.88}$$

Since $n_a \geq n_b$, by letting the code lengths n_a, $n_b \to \infty$, or by letting the field size $k \to \infty$, we obtain

$$\lim_{n_a, n_b \to \infty} P_R = 1, \tag{3.89}$$

$$\lim_{k \to \infty} P_R = 1. \tag{3.90}$$

The proof is now complete. □

We have an *alternate proof* to the theorem which is based on the probability of a matrix, composed of randomly chosen elements from a finite field, being full rank.

Proof 2: Let us use a systematic random code on the larger information codeword $\vec{a}$, and not code on the smaller information codeword $\vec{b}$. We will be able to achieve capacity and maximum code rate given by (3.16,3.25), as long as we completely recover n_a elements of $\vec{a}$ and n_b elements of $\vec{b}$. We set $n_b = \beta n_a$, where, $\beta \in [0, 1)$ and is fixed. Random coding will give an optimal code if $W = \left[\begin{array}{c|c} TL_a & TL_b \end{array} \right]$ has $n_a + n_b$ 1rows, i.e., has a rank of $n_a + n_b$. We know from our model that

$$W = LT, \tag{3.91}$$

$$\Rightarrow Rank(W) = \min\{Rank(L), Rank(T)\}. \tag{3.92}$$

Since, $Rank(W) = n_a + n_b$, for the random code to be optimal we require

$$Rank(L) = n_a + n_b, \tag{3.93}$$

$$Rank(T) = n_a + n_b. \tag{3.94}$$

Thus, L and T need to be full rank matrices. Random systematic coding makes L take the form $L = \begin{bmatrix} I_{n_a \times n_a} & I_{n_b \times n_b} \\ M & 0_{n_a \times n_b} \end{bmatrix}$ where, M is a $n_b \times n_a$ sized random matrix containing elements from $\mathbb{F}_{2^k}$. Clearly, L will be full rank if the rows of M are all independent. Once L has full rank, it can always be transformed into W using a $n_a + n_b$ sized full rank square matrix T. Thus random coding becomes optimal iff the rows of M are independent. In other words, if and only if L is full rank, does a random code become optimal.

Let us look at the $n_b \times (n_a + n_b)$ sized matrix M. We choose the elements of this matrix uniformly from $\mathbb{F}_{2^k}$. In order for the rows to be independent, the number of ways, N_j, to choose the j^{th} row in M is

$$N_j = 2^{kn_a} - 2^{k(j-1)} \quad j \in \{1, 2, ..., n_b - 1, n_b\}. \tag{3.95}$$

Therefore, the total number of full rank matrices, N_{total} is

$$\begin{aligned} N_{total} &= \prod_{j=1}^{n_b} N_j & (3.96) \\ &= 2^{kn_a n_b} \prod_{j=1}^{n_b} [1 - 2^{-k(n_a - j + 1)}]. & (3.97) \end{aligned}$$

Since the total number of random matrices is $2^{kn_a n_b}$, the probability of M having all independent rows, P_{ind}, is

$$\begin{aligned} P_{ind} &= \prod_{j=1}^{n_b} [1 - 2^{-k(n_a - j + 1)}] & (3.98) \\ &= exp\{\sum_{j=1}^{n_b} ln(1 - 2^{-k(n_a - j + 1)})\} & (3.99) \\ &\geq exp\{-\sum_{j=1}^{n_b} (2^{-k(n_a - j + 1)} + 2^{-2k(n_a - j + 1)})\}, & (3.100) \end{aligned}$$

where the inequality in (3.100) follows from the fact that, for $x \in [0, 1/2]$, $-(x + x^2) \leq \ln(1 - x)$. The probability of random coding being optimal, P_R, is the same as the probability of the rows of M being independent, P_{ind}. Simplifying (3.100), we obtain

$$exp\{-[\frac{2^{-k(1-\beta)n_a} - 2^{-kn_a}}{2^k - 1} + \frac{2^{-2k(1-\beta)n_a} - 2^{-2kn_a}}{2^{2k} - 1}]\} \leq P_R \leq 1. \tag{3.101}$$

In (3.101), if we let the code lengths $n_a, n_b \to \infty$, or if we let the field size $k \to \infty$, we have

$$\lim_{n_a, n_b \to \infty} P_R = 1, \tag{3.102}$$

$$\lim_{k \to \infty} P_R = 1. \tag{3.103}$$

The proof is now complete. □

Since the probability of a random code being optimal goes to 1 exponentially with k, n_a and n_b, a random code becomes optimal with moderate codeword lengths or field sizes (over which elements are defined). This is a strong coding theorem for this channel. A similar result can also be obtained by using the Schwartz-Zippel Theorem in [9]. Using this theorem we can consider a square matrix to be equivalent to a multi-variate polynomial whose elements are defined over a finite field. Now, if the determinant of the matrix is 0, the polynomial will be 0. The theorem in [9] proves that as the size of the field goes to infinity, the probability that the polynomial is 0 tends to 0 asymptotically with the size of the field.

Deterministic code constructions are difficult in general since they involve solving equations over finite fields. As the error probability for systematic random codes tends to 1 in an exponential manner, systematic random codes become optimal with moderate codeword lengths, making code construction easy. Therefore, when we choose our code book randomly, the number of tries to get a code book for the optimal code is very small.

3.5 Multiple access for bursty transmitters

In our discussion in the previous sections, it is assumed that each transmitter has a codeword to transmit in a slot. We now look at the case when the channel encoders may not always have an input information codeword to encode. Each transmitter transmits in a slot according to a Bernoulli process. The lower the probability of transmission, the burstier the transmitter. We will like to know as to what coding technique to use in order to obtain the maximum code rate over a large number of transmissions. It should be expected that bursty transmissions will reduce multiple access interference and increase the code rate. Moreover, we should be able to obtain the code rate of 1, (code rate of a point-to-point noise-free channel) in the limit that one transmitter stops transmitting. In this section, we illustrate the coding technique to be used for bursty multiple access and also show that the limits that we expect actually hold.

Let the probability that transmitters a and b have a codeword to transmit in a slot be p_a and p_b respectively. We define the sizes of $\vec{a}$ and $\vec{b}$ as n_a and n_b again. Consider the *expected code rate*. The mean sizes of $\vec{a}$ and $\vec{b}$ are $p_a n_a$ and $p_b n_b$ respectively. We will therefore define the mean sizes $n_1^{'}, n_2^{'}$ where $n_2^{'} \leq n_1^{'}$ as

$$n_1^{'} = \max(p_a n_a, p_b n_b), \tag{3.104}$$

$$n_2^{'} = \min(p_a n_a, p_b n_b). \tag{3.105}$$

The maximum expected code rate is thus

$$E(C_{rate}) = \frac{n_1^{'} + n_2^{'}}{n_1 + 2n_2^{'}}, \tag{3.106}$$

where (from Theorem 3.4) no redundancy is added to the transmit vector with smaller mean size. Therefore, two cases arise.

Case 1: We add redundancy only at a and not at b if the mean size of $\vec{a}$ is greater than the mean size of $\vec{b}$. This implies that

$$n_1^{'} = p_a n_a, \tag{3.107}$$

$$n_2^{'} = p_b n_b, \tag{3.108}$$

$$E(C_{rate}) = \frac{p_a n_a + p_b n_b}{p_a n_a + 2 p_b n_b}. \tag{3.109}$$

Case 2: We add redundancy only at b and not on a if the mean size of $\vec{a}$ is smaller than the mean size of $\vec{b}$, which implies that

$$n_1^{'} = p_b n_b, \tag{3.110}$$

$$n_2^{'} = p_a n_a, \tag{3.111}$$

$$E(C_{rate}) = \frac{p_a n_a + p_b n_b}{2 p_a n_a + p_b n_b}. \tag{3.112}$$

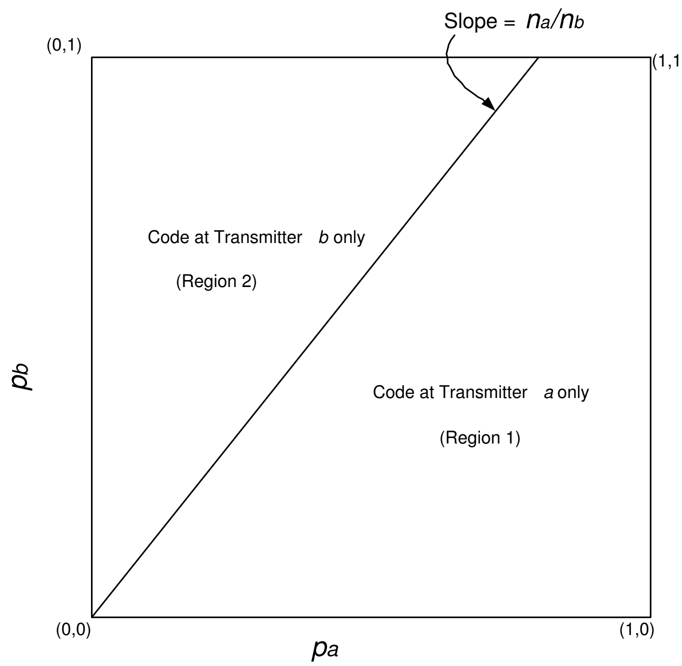

Figure 3-6: Coding Regions for Bursty Multiple Access.

When $n_1' = n_2'$, we can use either technique. Figure 3-6 shows the regions of the two dimensional space of (p_a, p_b) where the cases apply. Region 1 corresponds to the first case and Region 2 to the second.
Let us denote $\alpha = \frac{p_a}{p_b}$ and $\beta = \frac{n_a}{n_b}$. We have $\alpha \in [0, \infty)$ and $\beta \geq 1$. Thus, the maximum expected code rate expression can be written as

$$E(C_{rate}(\alpha, \beta)) = \frac{1 + \alpha\beta}{1 + \alpha\beta + \min(1, \alpha\beta)}. \tag{3.113}$$

For $\alpha \in [0, \frac{1}{\beta}]$, the mean size of $\vec{a}$ is less than or equal to the size of $\vec{b}$ and we add redundancy only at b. For $\alpha \in [\frac{1}{\beta}, \infty)$ the mean size of $\vec{a}$ is larger than or equal to the mean size of $\vec{b}$, and we add redundancy only at a. Note that for $\alpha = \frac{1}{\beta}$, we may add redundancy at a or b and still obtain the same expected code rate. The expected code rate is minimum and has a value of $2/3$ when $p_a n_a = p_b n_b$. Figure 3-7 shows how the expected code rate changes with α. We now look at the limit when transmitter a stops transmitting, i.e $\alpha \to 0$ and when transmitter b stops transmitting, i.e. $\alpha \to \infty$.

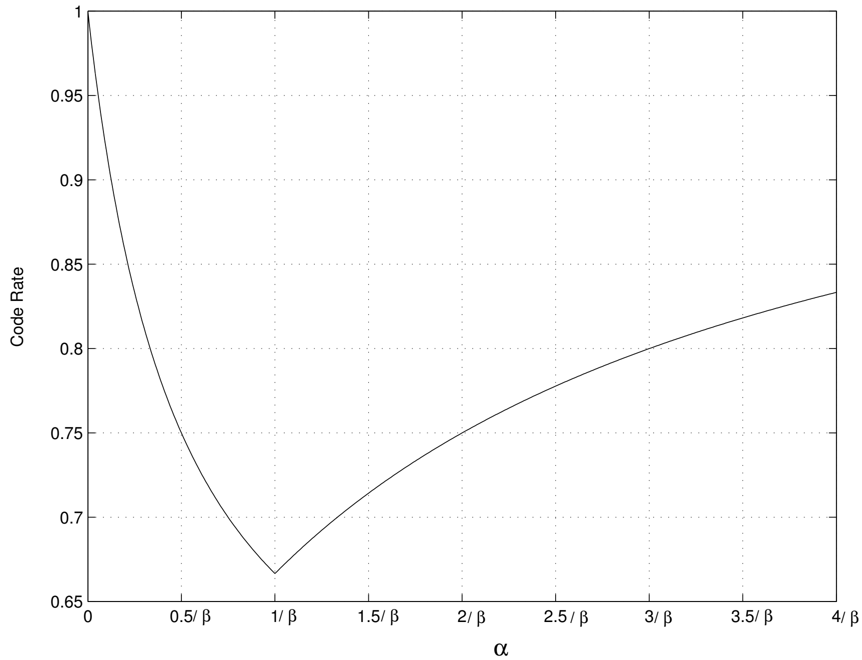

Figure 3-7: Variation of code rate with α.

Evaluating the value of expected code rate as α tends to 0 or ∞, we get

$$\lim_{\alpha \to 0} C_{rate}(\alpha, \beta) = 1, \tag{3.114}$$

$$\lim_{\alpha \to \infty} C_{rate}(\alpha, \beta) = 1. \tag{3.115}$$

These limits are what we had expected since, in both cases, one transmitter transmits in a slot with probability 1 and the other does not transmit at all. There is no multiple access interference and the average code rate becomes the code rate of a point-to-point noiseless channel, i.e 1.

When $\beta = 1$, i.e $n_a = n_b$, we see that if $p_b \leq p_a$, we add redundancy only at a and when $p_b > p_a$ we add redundancy only at b. This gives rise to the following lemma:

Lemma 3.3 *When the information codewords at the input to the channel encoders have the same size, maximum expected code rate is achieved by adding redundancy at the less bursty transmitter not adding any redundancy at the more bursty transmitter.*

□

This lemma gives us a coding technique when transmissions are probabilistic.

Chapter 4

Time-Slotted Noisy Multiple Access Networks over Finite Fields

In this chapter, we will consider a time-slotted noisy multiple access finite field adder channel where the noise is independent of the inputs. The transmitted elements X_a, X_b, noise element Z and received element Y are from $\mathbb{F}_{2^k}$, for $1 \leq k$ and

$$Y = X_a + X_b + Z. \tag{4.1}$$

Noise and multiple access interference are additive over $\mathbb{F}_{2^k}$. First, we develop a model for communicating over this channel. Then, we establish the capacity region and maximum code rate and study their dependence on field size. Using the results of the noisy multiple access strong coding theorem developed by Liao in [22], we obtain the error exponents and hence the expression for average probability of error when a random code is used for communicating over this channel.

4.1 Single Slot Model

We consider a discrete time channel. The channel is time-slotted and we consider transmissions over the length of a slot or slot duration. Figure 4-1 shows a single slot model of the noise-free finite field multiple access channel. Information codewords $\vec{U}'$ and $\vec{V}'$, as described in Figure 2-2, coming out of the source coders at transmitters

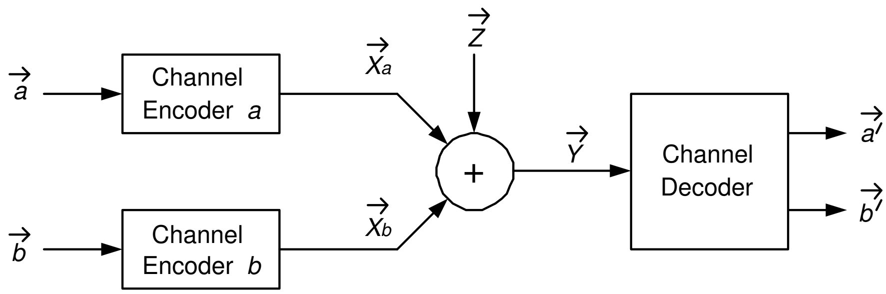

Figure 4-1: Single slot model for the noisy multiple access finite field adder channel.

a and b in one slot duration will be represented as $\vec{a}$ and $\vec{b}$, respectively, in this chapter. $\vec{a}$ and $\vec{b}$ are thus vectors of sizes n_a and n_b respectively, which represent independent information. The elements of $\vec{a}$ and $\vec{b}$ are in $\mathbb{F}_{2^k}$, for $1 \leq k$. Noise $\vec{Z}$ is independent of the inputs. In Chapter 2, we showed in Theorem 2.4 that source-channel separation holds for a noisy multiple access finite field adder channel if the noise is independent of inputs. Thus, the scheme of separate source and channel coding is optimal. For this model, all operations, matrices and vectors are in $\mathbb{F}_{2^k}$. We will refer to $\vec{a}$ and $\vec{b}$ as transmit vectors and assume that $n_a \geq n_b$. (Otherwise, the arguments still hold with $\vec{a}$ and $\vec{b}$ interchanged.) Channel encoders at a and b encode the information codewords into codewords $\vec{X_a}$ and $\vec{X_b}$ respectively. The values $\vec{X_a}$ and $\vec{X_b}$ are subsequently transmitted over the channel which adds with noise vector $\vec{Z}$ to yield $\vec{Y}$. At the receiver, the multiple access channel decoder acts on $\vec{Y}$ to produce estimates $\vec{a'}$ and $\vec{b'}$ of $\vec{a}$ and $\vec{b}$ respectfully.

Let m_a and m_b be the increase in the size of $\vec{a}$ and $\vec{b}$, respectively, due to coding. Let l_a and l_b denote the lengths of the vectors obtained by coding on $\vec{a}$ and $\vec{b}$ respectively. In general $l_a \neq l_b$ and so both transmitters may not transmit for the entire slot duration. However, at least one transmitter will transmit for the whole slot duration. Therefore, the slot length is given by

$$S' = \max(l_a, l_b) \text{ } symbols. \tag{4.2}$$

We have

$$l_a = n_a + m_a, \tag{4.3}$$

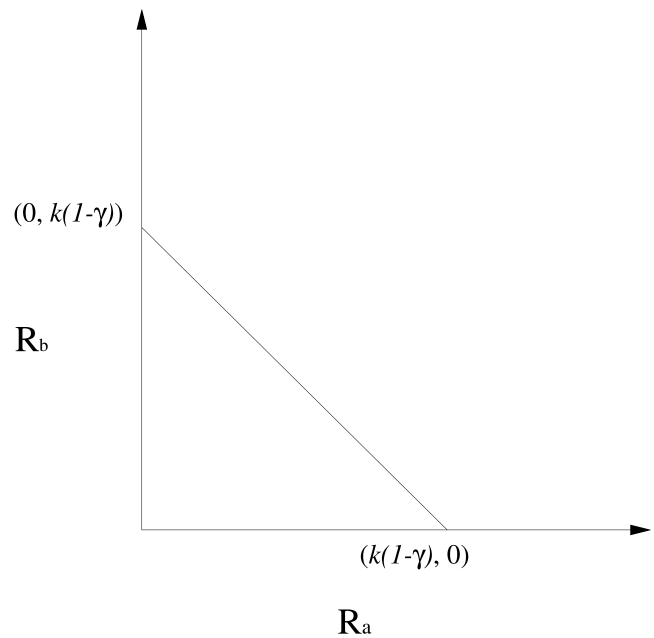

Figure 4-2: Capacity region of a noisy multiple access finite field adder channel over a field of size 2^k.

$$l_b = n_b + m_b. \tag{4.4}$$

4.2 Capacity Region and Maximum Code Rate

In this section, we derive the capacity region and maximum code rate over a very long slot or many slots. First, we establish the capacity region and then the maximum code rate.

4.2.1 Capacity Region

From [21, 22], we know that the multiple access capacity region is the closure of the convex hull of all (R_a, R_b) satisfying

$$R_a \leq I(X_a; Y|X_b), \tag{4.5}$$

$$R_b \leq I(X_b; Y|X_a), \tag{4.6}$$

$$R_{sum} = R_a + R_b \leq I(X_a, X_b; Y). \tag{4.7}$$

Simplifying the mutual information expressions, we get

$$
\begin{aligned}
I(X_a; Y|X_b) &= H(Y|X_b) - H(Y|X_a, X_b) && (4.8)\\
&= H(X_a + Z) - H(Z), && (4.9)\\
I(X_b; Y|X_a) &= H(Y|X_a) - H(Y|X_a, X_b) && (4.10)\\
&= H(X_b + Z) - H(Z), && (4.11)\\
I(X_a, X_b; Y) &= H(Y) - H(Y|X_a, X_b) && (4.12)\\
&= H(X_a + X_b + Z) - H(Z). && (4.13)
\end{aligned}
$$

Uniform distribution of X_a and X_b maximizes R_a, R_b and R_{sum}. Let us denote

$$H(Z) = k\gamma, \tag{4.14}$$

where $\gamma \in [0, 1]$. The multiple access capacity region is therefore the convex hull of all (R_a, R_b) satisfying

$$
\begin{aligned}
R_a &\leq k(1 - \gamma), && (4.15)\\
R_b &\leq k(1 - \gamma), && (4.16)\\
R_{sum} &\leq k(1 - \gamma), && (4.17)
\end{aligned}
$$

where, the rates are in bits per channel use. The capacity region is shown in Figure 4-2.

4.2.2 Maximum Code Rate

We now find the maximum code rate for this channel. Transmitters a and b transmit n_a and n_b information symbols per slot, respectively, and the codewords transmitted have a length of l_a and l_b symbols, respectively. The transmission rates are

$$
\begin{aligned}
R_a &= \frac{kn_a}{S'}, && (4.18)\\
R_b &= \frac{kn_b}{S'}, && (4.19)\\
R_{sum} &= \frac{k(n_a + n_b)}{S'}. && (4.20)
\end{aligned}
$$

From (4.17, 4.20), we obtain

$$\frac{n_a + n_b}{1-\gamma} \leq S'. \tag{4.21}$$

The code rate is a dimensionless quantity and is given by

$$\begin{aligned}
C_{rate} &= \frac{n_a + n_b}{l_a + l_b} && (4.22)\\
&= \frac{n_a + n_b}{\min(l_a, l_b) + \max(l_a, l_b)} && (4.23)\\
&= \frac{n_a + n_b}{\min(l_a, l_b) + S'} && (4.24)\\
&\leq \frac{n_a + n_b}{n_b + \frac{n_a + n_b}{1-\gamma}} && (4.25)\\
&= \frac{(1-\gamma)(n_a + n_b)}{n_a + (2-\gamma)n_b}. && (4.26)
\end{aligned}$$

Equation (4.24) is due to (4.2) and (4.25) follows from (4.21) and the fact that $n_b \leq \min(l_a, l_b)$. Thus, we have

$$C_{rate} \leq \frac{(1-\gamma)(n_a + n_b)}{n_a + (2-\gamma)n_b}. \tag{4.27}$$

We now prove an important lemma:

Lemma 4.1 *The capacity region of a two-transmitter noisy finite field adder channel grows logarithmically with the size of the field but the code rate remains the same for all field sizes.*

Proof: From (4.15), (4.16) and (4.17), we see that the capacity region grows logarithmically with field size (k). However, (4.27) shows that the code rate is invariant to k. Thus, we have proved the lemma. □

Note that this lemma for the noisy channel is similar to Lemma 3.1 for the noise-free channel.

4.3 Error Exponents

Liao in [22] developed a strong coding theorem for noisy multiple access channels. In this section, we present the derivation of the multiple access coding theorem as done in [11]. We then compute the error exponents and evaluate the average probability of error by random coding for the noisy multiple access finite field adder channel. Consider block coding with a given block length N using M_1 codewords $\{\vec{x_1}, \vec{x_2}, ..., \vec{x_{M_1}}\}$ for transmitter 1, and M_2 codewords $\{\vec{w_1}, \vec{w_2}, ..., \vec{w_{M_2}}\}$ for transmitter 2; each codeword is a sequence of N channel inputs. For convenience we refer to a code with these parameters as an (N, M_1, M_2) code. The rates of the two sources are defined as

$$R_1 = \frac{\ln M_1}{N}, \tag{4.28}$$

$$R_2 = \frac{\ln M_2}{N}. \tag{4.29}$$

Each N units of time, source 1 generates an integer m_1 uniformly distributed from 1 to M_1, and source 2 independently generates an integer m_2 uniformly distributed from 1 to M_2. The transmitters send $\vec{x_{m_1}}$ and $\vec{w_{m_2}}$, respectively, and the corresponding channel output $\vec{y}$ enters the decoder and is mapped into a decoded "message" $m_1^{'}, m_2^{'}$. If both $m_1^{'} = m_1$ and $m_2^{'} = m_2$, the decoding is correct and otherwise a decoding error occurs. The probability of decoding error P_e is minimized for each $\vec{y}$ by a maximum likelihood decoder, choosing $(m_1^{'}, m_2^{'})$ as integers $m_1^* \in [1, M_1]$, $m_2^* \in [1, M_2]$ that maximize $P(\vec{y}|\vec{x_{m_1^*}}\vec{w_{m_2^*}})$. If the maximum is non-unique, any maximizing (m_1^*, m_2^*) can be chosen with no effect on P_e. Both sets of codewords, $\{\vec{x_1}, \vec{x_2}, ..., \vec{x_{M_1}}\}$ and $\{\vec{w_1}, \vec{w_2}, ..., \vec{w_{M_2}}\}$ are known to the decoder, but, of course, the source outputs m_1, m_2 are unknown.

Let $Q_1(x)$ and $Q_2(w)$ be probability assignments on input alphabets X and W. The input alphabets X, W and output alphabet y are elements from $\mathbb{F}_{2^k}$. Consider an ensemble of (N, M_1, M_2) codes where each codeword $\vec{x_{m_1}}$ for $m_1 \in [1, M_1]$, is independently selected according to the probability assignment

$$Q_1(\vec{x}) = \prod_{n=1}^{N} Q_1(x_n), \qquad \vec{x} = (x_1, x_2, ..., x_N) \tag{4.30}$$

and each code word $\vec{w_{m_2}}$ for $m_2 \in [1, M_2]$ is independently chosen according to

$$Q_2(\vec{w}) = \prod_{n=1}^{N} Q_2(w_n), \qquad \vec{w} = (w_1, w_2, ..., w_N). \tag{4.31}$$

For each code in the ensemble, the decoder uses maximum likelihood decoding, and we want to upper bound the expected value $\overline{P_e}$ of P_e for this ensemble. Define an error event to be of type 1 if the decoded pair $(m_1^{'}, m_2^{'})$ and the original source pair (m_1, m_2) satisfy $m_1^{'} \neq m_1, m_2^{'} = m_2$. An error event is of type 2 if $m_1^{'} = m_1, m_2^{'} \neq m_2$, and is of type 3 if $m_1^{'} \neq m_1, m_2^{'} \neq m_2$. Let P_{ei} for $i \in \{1, 2, 3\}$ be the probability, over the ensemble, of a type i error event. We have

$$\overline{P_e} = P_{e1} + P_{e2} + P_{e3}. \tag{4.32}$$

Consider P_{e3} first. Note that, when (m_1, m_2) enters the encoder, there are $M_1 - 1$ choices for $m_1^{'}$ and $M_2 - 1$ choices for $m_2^{'}$, or $(M_1 - 1)(M_2 - 1)$ pairs, that yield a type 3 error. For each such pair $(m_1^{'}, m_2^{'})$, the codeword pair $(\vec{x_{m_1^{'}}}, \vec{w_{m_2^{'}}})$ is statistically independent of $(\vec{x_{m_1}}, \vec{w_{m_2}})$ over the ensemble of codes. Thus, regarding $(\vec{x}, \vec{w})$ as a combined input to a single input channel with input alphabet $X \times W$, we can directly apply the coding theorem which asserts that, for all $\rho \in [0, 1]$,

$$P_{e3} \leq [(M_1 - 1)(M_2 - 1)]^{\rho} \sum_{\vec{y}} [\sum_{\vec{x}, \vec{w}} Q_1(\vec{x}) Q_2(\vec{w}) P(\vec{y}|\vec{x}\vec{w})^{\frac{1}{1+\rho}}]^{1+\rho}. \tag{4.33}$$

Using the product form of Q_1, Q_2, P and definition of rates, we have

$$P_{e3} \leq \exp{[\rho N(R_1 + R_2)]}[\sum_{y} [\sum_{x,w} Q_1(x) Q_2(w) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}]^N. \tag{4.34}$$

Next consider P_{e1}, the probability that $m_1^{'} \neq m_1$ and $m_2^{'} = m_2$. We first condition this probability on a particular message m_2 entering the second encoder, and a choice of code with a particular $\vec{w_{m_2}}$ transmitted at the second input. Given $\vec{w_{m_2}}$, we can view the channel as a single input channel with input $\vec{x_{m_1}}$ and with transition probabilities $P(\vec{y}|\vec{x_{m_1}}\vec{w_{m_2}})$.

A maximum likelihood detector for that single input channel will make an error (or be ambiguous) if

$$P(\vec{y}|\vec{x_{m_1}^{*}}\vec{w_{m_2}}) \geq P(\vec{y}|\vec{x_{m_1}}\vec{w_{m_2}}), \tag{4.35}$$

for at least one $m_1^{*} \neq m_1$. Since this event must occur whenever a type 1 error occurs, the probability of a type 1 error, conditional on $\vec{w_{m_2}}$ being sent, is upper bounded by the probability of error or ambiguity on the above single input channel. Using the coding theorem again for this single input channel, we have, for any $\rho \in [0,1]$,

$$P[type\ 1\ error|\vec{w_{m_2}}] \leq [M_1 - 1]^{\rho} \sum_{\vec{y}} [\sum_{\vec{x}} Q_1(\vec{x}) P(\vec{y}|\vec{x}\vec{w})^{\frac{1}{1+\rho}}]^{1+\rho}. \tag{4.36}$$

Using the product form of Q_1, Q_2, P and taking the expected value over $\vec{w_{m_2}}$, we have

$$P_{e1} \leq \exp[\rho N R_1] [\sum_{y,w} Q_2(w) [\sum_x Q_1(x) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}]^N. \tag{4.37}$$

Applying the same argument to type 2 errors for $\rho \in [0,1]$,

$$P_{e2} \leq \exp[\rho N R_2] [\sum_{y,x} Q_1(x) [\sum_w Q_2(w) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}]^N. \tag{4.38}$$

Let us define $E_{01}(\rho, Q)$, $E_{02}(\rho, Q)$ and $E_{03}(\rho, Q)$ as

$$E_{01}(\rho, Q) = -\ln \sum_{y,w} Q_2(w) [\sum_x Q_1(x) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}, \tag{4.39}$$

$$E_{02}(\rho, Q) = -\ln \sum_{y,x} Q_1(x) [\sum_w Q_2(w) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}, \tag{4.40}$$

$$E_{03}(\rho, Q) = -\ln \sum_{y} [\sum_{x,w} Q_1(x) Q_2(w) P(y|xw)^{\frac{1}{1+\rho}}]^{1+\rho}. \tag{4.41}$$

Substituting (4.39,4.40,4.41) in (4.34,4.37,4.38) we obtain

$$P_{e1} \leq \exp[-N[-\rho R_1 + E_{01}(\rho, Q)]], \tag{4.42}$$

$$P_{e2} \leq \exp[-N[-\rho R_2 + E_{02}(\rho, Q)]], \tag{4.43}$$

$$P_{e3} \leq \exp[-N[-\rho(R_1 + R_2) + E_{03}(\rho, Q)]]. \tag{4.44}$$

We now compute the error exponents and evaluate the average probability of error. The capacity achieving distribution maximizes $E_{01}(\rho, Q)$, $E_{02}(\rho, Q)$ and $E_{03}(\rho, Q)$. Since the capacity achieving distribution for this channel is uniform distribution of X and W, we have

$$\max_Q E_{01}(\rho, Q) = k(2+\rho)\ln 2 - \ln \sum_{y,w} [\sum_x P(y|x,w)^{\frac{1}{1+\rho}}]^{1+\rho}, \tag{4.45}$$

$$\max_Q E_{02}(\rho, Q) = k(2+\rho)\ln 2 - \ln \sum_{y,x} [\sum_w P(y|x,w)^{\frac{1}{1+\rho}}]^{1+\rho}, \tag{4.46}$$

$$\max_Q E_{03}(\rho, Q) = 2k(1+\rho)\ln 2 - \ln \sum_{y} [\sum_{x,w} P(y|x,w)^{\frac{1}{1+\rho}}]^{1+\rho}. \tag{4.47}$$

Owing to the symmetry of the channel, we have

$$\sum_x P(y|x,w)^{\frac{1}{1+\rho}} = \sum_w P(y|x,w)^{\frac{1}{1+\rho}} = \sum_{i=0}^{2^k-1} P(Z=i)^{\frac{1}{1+\rho}}. \tag{4.48}$$

We will denote

$$\mathcal{F}(\rho) = \sum_{i=0}^{2^k-1} P(Z=i)^{\frac{1}{1+\rho}}. \tag{4.49}$$

From (4.42 - 4.49), we have

$$P_{e1} \leq \exp[-N[\max_{\rho\in[0,1]}(k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho R_1)]], \tag{4.50}$$

$$P_{e2} \leq \exp[-N[\max_{\rho\in[0,1]}(k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho R_2)]], \tag{4.51}$$

$$P_{e3} \leq \exp[-N[\max_{\rho\in[0,1]}(k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho(R_1+R_2))]]. \tag{4.52}$$

Therefore, the error exponents $E_1(R_1)$, $E_2(R_2)$ and $E_3(R_1, R_2)$ for type 1, type 2 and type 3 errors respectively, are

$$E_1(R_1) = \max_{\rho\in[0,1]}[k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho R_1], \tag{4.53}$$

$$E_2(R_2) = \max_{\rho\in[0,1]}[k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho R_2], \tag{4.54}$$

$$E_3(R_1,R_2) = \max_{\rho\in[0,1]}[k\rho\ln 2 - (1+\rho)\ln\mathcal{F}(\rho) - \rho(R_1+R_2)], \tag{4.55}$$

and the average probability of error is

$$\overline{P_e} \leq \exp[-NE_1(R_1)] + \exp[-NE_2(R_2)] + \exp[-NE_3(R_1,R_2)]. \tag{4.56}$$

All the error exponents are positive only when the rate pairs are inside the capacity region. Thus, for rate pairs inside the capacity region, the probability of error goes to 0 exponentially with codeword length and reliable communication is possible. However, the converse to the multiple access coding theorem described in [22] shows that for all rate pairs outside the capacity region, the probability of decoding error cannot be made arbitrarily small no matter what encoding or decoding procedures are used. Reliable communication is thus not possible for all rate pairs outside the capacity region.

Chapter 5

Conclusion

We have considered the noisy as well as the noise-free multiple access finite field adder channel. It is shown that source-channel separation holds when noise is independent of inputs but may not when noise is input-dependent. For channels over the binary field, we derive the expression for the probability of source-channel separation failing. We compute this probability to be 1/4 when the noise parameters are picked independently and uniformly. For binary channels, we derive an upper bound of 0.0776 bit for the maximum loss in sum rate due to separate source-channel coding when separation fails. We prove that the bound is very tight by showing that it is accurate to the second decimal place. Thus we see that though there is a significantly high probability that source-channel separation does not hold for a channel over $\mathbb{F}_2$, the loss in sum rate is very small, especially, when the noise is low.

Source-channel separation does not hold for multiple access channels where the alphabet size grows with the number of transmitters. However, such systems are designed by separating source and channel coding which can be inefficient as the maximum sum rate achievable by separate source-channel coding can be significantly lower than the maximum sum rate possible by joint source-channel coding. In this thesis we show that this problem does not exist for multiple access channels when the received alphabet size is fixed. Since source-channel separation holds when noise is independent of inputs, we do not lose optimality be separating the source and channel coding. Moreover, the loss is very small even when the noise is input-dependent.

For the noise-free finite field adder channel, a single slot model is developed and the capacity and maximum code rate are derived. It is shown that the capacity of this channel grows logarithmically with the field size but the code rate remains the same for all field sizes. Moreover, we show that codes achieve the maximum code rate if and only if they achieve capacity and added no redundancy to the smaller transmit vector.

For the case when both transmitters transmit in a slot, we propose a systematic code construction that achieves the maximum code rate and capacity. We consider the performance of systematic random codes where we derive an expression for the probability of error when the codes are chosen randomly. The probability of error goes to 0 exponentially with code length and field size. Thus we have a strong coding theorem for this channel. Deterministic code constructions are difficult in general since they involve solving equations over finite fields. Since, the error probability tends to 1 in an exponential manner, systematic random codes become optimal with moderate codeword lengths and field sizes. Thus code construction becomes easy since we have to just choose the code book randomly and the number of tries to get an optimal one is very small. For the case when transmitters transmit according to a per-slot Bernoulli process, a coding scheme is proposed to maximize the expected code rate. It is shown that, when the information codewords at the input to the channel encoders have the same size, maximum code rate is achieved by adding redundancy at the less bursty transmitter and not adding any redundancy at the more bursty transmitter.

We also consider a time-slotted noisy multiple access finite field adder channel where noise is independent of the inputs and additive over the same field as the inputs and output. For this noisy channel, we develop a model and establish the capacity region and maximum code rate. We show, as in the noise-free multiple access finite field adder channel, that the capacity grows logarithmically with the field size but the code rate remains the same for all field sizes. Using the results of the noisy multiple access strong coding theorem developed by Liao in [22], we obtain error exponents and hence the expression for average probability of error when a random code is used for communicating over this channel.

In this thesis, we have looked at source-channel separation when noise is dependent on only one input. Future research can look at the problem where the noise depends on both inputs. The question of separation, maximum loss in sum rate by doing separate source-channel coding and probability of separation failure can be looked at. Moreover, whenever we compute loss in sum rate be separate source-channel coding, we assume that sources can be perfectly matched to the channel by a joint source-channel coder. This may not hold in general and research should look at when such a match is or is not possible. Unmatched sources make the loss due to separate source-channel coding lower than if they were perfectly matched. For this reason, the probability of separation failing may also be lower.

Bibliography

[1] S. Chang and E.J.Weldon,"Coding for T-user multiple access channels", *IEEE Transactions on Information Theory, Vol.IT-25, No.6, Nov 1979.*

[2] J. Cheng and Y. Watanabe,"T-user code with arbitrary code length for multiple access adder channel", *IEICE Transactions Fundamentals, Vol.E82-A, No.10, October 1999.*

[3] G.H. Khachatrian and S.S. Martirossian,"Code construction for the T-user noiseless adder channel", *IEEE Transaction on Information Theory, Vol.44, No.5, Sep 1998.*

[4] B. Hughes,"Capacity and coding for T active users out of M on the collision channel.", *ISIT 1993.*

[5] B.L. Hughes and A.B. Cooper, "Nearly optimal multiuser codes for the binary adder channel", *IEEE Transactions on Information Theory, Vol.-42, No.2, March 1996.*

[6] P. Mathys, "A class of codes for T active users out of N multiple access communication system", *IEEE Transactions on Information Theory, Vol. 36, No. 6, November 1990.*

[7] G.Caire, E.Leonardi, E. Viterbo,"Improving performance on wireless networks using collision resistant modulations.", *Globecom '98.*

[8] G. Caire, E. Leonardi and E. Viterbo,"Modulation and coding for the Gaussian collision channel", *IEEE Transactions on Information Theory, Vol. 46, No. 6, September 2000.*

[9] R. Motwani and P. Raghavan, "Randomized Algorithms", *Cambridge University Press, 2000.*

[10] R.G. Gallager, "Information Theory and Reliable Communication", *Wiley, 1968.*

[11] R.G. Gallager, "A Perspective on Multiaccess Channels", *IEEE Transactions on Information Theory, Vol. 31, No. 2, March 1985.*

[12] T.M. Cover and J.A. Thomas, "Elements of Information Theory", *Wiley, 1991.*

[13] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources", *IEEE Transactions on Information Theory, IT-19: 471-480, 1973.*

[14] G. Poltyrev and J. Snyders, "Linear codes for the sum mod-2 multiple access channel with restricted access.", *IEEE Transactions on Information Theory, Vol. 41, No.3, May 1995.*

[15] G. Thomas,"Capacity of the wireless packet collision without feedback", *IEEE Transactions on Information Theory, Vol. 46, No.3, May 2000.*

[16] S. Laufer and J. Snyders,"Short coding for the collision channel without feedback.", *ISIT 1994.*

[17] T. J. Ketseoglou,"Certain generalizations on the collision channel without feedback.", *ISIT 1993.*

[18] I. Lim and J. Heo,"Performance analysis of CDMA ALOHA/FEC scheme in the centralized packet radio networks.", *ICICS 1997, Singapore.*

[19] A. J. Grant and C. Schlegel,"Collision-type multiple user communications", *IEEE Transactions on Information Theory, Vol. 43, No. 5, September 1997.*

[20] I. Csiszár, "Linear codes for sources and source networks: Error exponents, Universal Coding", *IEEE Transactions on Information Theory, 28:585-592, July 1982.*

[21] R. Ahlswede, "Multiway communication channels", *ISIT 1971.*

[22] H. Liao, "Multiple access channels", *Ph.D. Dissertation, Department of Electrical Engineering, University of Hawaii, Honolulu, 1972.*