AN INFORMATION- AND CODING-THEORETIC STUDY OF BURSTY CHANNELS WITH APPLICATIONS TO COMPUTER MEMORIES

Thesis by

Khaled Ahmed Sabry Abdel-Ghaffar

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1986

(Submitted April 23, 1986)

ABSTRACT

This thesis is a study of two-dimensional bursty channels from the information-theoretic as well as the coding-theoretic points of view. An information-theoretic model of bursty channels is defined and analyzed using probabilistic arguments. Two-dimensional burst correcting codes are developed. Their combinatorial and algebraic structures are examined. Two-dimensional bursty channels are used to model computer memories. The results of this thesis give bounds on the storage capacities of computer memories if sophisticated codes are used.

ACKNOWLEDGEMENT

It is not only my duty, but also my pleasure, to thank those who have contributed to this thesis. I am indebted to Dr. Henk C. A. van Tilborg for his generous time and advice during his stay at Caltech. My sincere gratitude also extends to Dr. Andrew Odlyzko for his kind contributions. Drs. Li Fung Chang, Roch Guerin, Phil Merkey, and Kumar Swaminathan have been very patient in teaching me TEX. It is far beyond my ability to repay them. My appreciation goes also to Mr. Chi Chao for his constructive criticism. I would like also to thank Professor Charles Seitz and the Defense Advanced Research Projects Agency for their kind financial support.

My deepest thanks, however, go to my thesis advisor Professor Robert J. McEliece who gave me the opportunity to become one of his students. I have greatly profited from his knowledge, insight, and advice. His friendly and pleasant attitude had a great impact on me during my stay at Caltech.

CONTENTS

Abstract ii

Acknowledgement iii

Introduction

1. Overview 1

2. A Guide to the Thesis 3

Part One: Computer Memories and Bursty Channels:
An Information-Theoretic Study

Chapter I: Information Content per Chip

1. Random Access Memories 6

2. Soft Errors in Computer Memories 7

3. Minimum Area per Information Bit 10

Chapter II: Bursty Channels

1. Definitions 15

2. Channel Capacities 17

3. Asymptotic Values of Capacities 24

4. Smart Interleaving 28

5. Applications to Computer Memories

5.1. One-Dimensional Chips 34

5.2. Two-Dimensional Chips 37

References for Part One 39

Part Two: Two-Dimensional Burst Correcting Codes

Preliminaries 43

Chapter III: Burst Identification Codes

1. Definitions 47

2. One-Dimensional Burst Identification Codes 48

3. Two-Dimensional Burst Identification Codes 50

4. Some Specific Burst Identification Codes

4.1. $1 \times b$-Burst Identification Codes 58

4.2. 2×2-Burst Identification Codes 59

4.3. $b_1 \times b_2$-Burst Identification Codes of Redundancy $2b_1b_2 - 2$ 63

4.4. 3×2-Burst Identification Codes 70

Chapter IV: The Structure of Burst Correcting Codes

1. Definitions 76

2. Bounds on Two-Dimensional Burst Correcting Codes 77

3. Burst Locating Codes

3.1. Definitions 82

3.2. $\gamma\beta$-Codes 83

3.3. $\alpha\beta$-Codes 86

4. BIL-Codes 89

5. Fire-ish Codes 91

6. Cyclic Burst Correcting Codes of Minimum Redundancy

6.1. Definitions 94

6.2. Applying Weil's Estimates of Character Sums 100

6.3. The Existence of Cyclic Burst Correcting Codes of Minimum Redundancy 108

Appendix: Encoding and Decoding Two-Dimensional Burst Correcting Codes

1. Introduction 112

2. Encoding Burst Correcting Codes

2.1. Encoding Cyclic Burst Correcting Codes 114

2.2. Encoding BIL-Codes and Fire-ish Codes 116

3. Decoding Burst Correcting Codes

3.1. Decoding Cyclic Burst Correcting Codes 117

3.2. Decoding BIL-Codes 118

3.3. Decoding Fire-ish Codes 118

References for Part Two 120

to my parents

INTRODUCTION

1. Overview

This thesis is a study of bursty channels. These channels are common in many real communication problems where noise is time-dependent. In this work we are primarily concerned with two-dimensional bursty channels. Chapter I gives a model for such channels which motivates the rest of the thesis. As memory cells in VLSI chips are made smaller every year, the cells become sensitive to many sources of errors. One of the most important sources is alpha-particle radiation. In present day technology, an alpha-particle may affect only a single memory cell. However, as cell dimensions decrease, a single alpha-particle may cause a two-dimensional burst of errors. Although present day memories are quite immune from thermal and quantum effects, these sources of errors are expected to be fundamental in any theoretical estimation of ultimate limits on data storing in memory devices. If the dimensions of memory cells continue to shrink every year, using error correcting codes to combat errors will become a necessity. However, if coding is used, then only a fraction of the number of memory cells on the chip is used to store data. The remaining bits are used as parity checks in order to allow the chip some error correcting capability. Using information-theoretic techniques, an upper bound on the ultimate number of data bits that can be stored on a chip is derived, if present day technological trends continue.

In chapter II, we give an information-theoretic model of bursty channels. The channel capacities are derived, and their asymptotic values are obtained. A coding scheme which is asymptotically optimum is described. The asymptotic values of the channel capacities are used to obtain the ultimate fraction of area

that can be used to store data, if alpha-particles are the only source of errors. It may be surprising to note that this useful area, and not only the number of data bits, does increase by making the memory cells smaller, if the dimensions of the cells are already below a certain limit.

Chapters I and II compose the first part of the thesis which is dominated by information-theoretic arguments.

In the second part, two-dimensional burst correcting codes are investigated. This part starts with the essential preliminaries relevant to two-dimensional codes and bursts. In chapter III, the class of burst identification codes is defined and studied. Burst identification codes are codes that can identify the burst patterns but not necessarily their positions. We are primarily interested in the minimum number of redundant bits required to construct burst identification codes of arbitrarily large areas.

In chapter IV, we define two-dimensional burst correcting codes, and give a measure of their efficiency which we call excess redundancy. We also define and study two-dimensional burst locating codes. These codes can locate the positions of the bursts if their patterns are known. Burst locating codes are used extensively in constructing burst correcting codes. The first class of burst correcting codes developed in this chapter is the class of BIL-codes. These codes are constructed by combining burst identification and locating codes. The second class of codes developed in this chapter, is the class of Fire-ish codes, which is a generalization of Fire codes. These two classes of codes are generally better, with respect to excess redundancy, than any other class of two-dimensional burst correcting codes ever reported in the literature. However, the most important class of codes introduced in this chapter is the class of cyclic burst correcting codes of minimum redundancy. These codes, which are cyclic, have the smallest redundancies among all burst correcting codes of the same areas.

In the appendix, we briefly describe efficient encoding and decoding techniques for the codes developed in chapter IV.

2. A Guide to the Thesis

This thesis is composed of two parts which are held together with the same bond that holds together information theory and the theory of error correcting codes. The two parts are written independent of each other. However, chapter I is considered to be a motivation for the problems treated in both parts. Within the first part, chapter II, except section 5, is independent of the material of chapter I. The preliminaries of the second part are essential to chapters III and IV. Certain definitions and results of chapter III are used in sections 4 and 5 of chapter IV.

Each part has its own references. In the first part, the author makes free use of some basic defintions and results in information and probability theories. These definitions and results can be found in most textbooks. More specifically, the author is heavily influenced by [7],[13], [21] in information theory, and [5],[10] in probability theory. The general description of computer memories is primarily based on [15],[16].

Most of the arguments presented in the second part depends on very basic results in algebra, finite fields, and number theory. However, any result, which is not considered to be basic, is clearly stated and referenced. Moreover, proofs are given for most of these results wherever space and logical continuity permit. The author has primarily consulted [14] in algebra, [15],[18] in finite fields, and [11] in number theory. In coding theory, the author is a student of [3],[16],[17],[20].

The author considers Theorems 6 of chapter II, 8 of chapter III, and 34 of chapter IV to be the most important results of the thesis.

Finally, few remarks about notation. In this thesis, all logarithms have base

2. For real r, $\lfloor r \rfloor$ denotes the largest integer $\leq r$, and $\lceil r \rceil$ denotes the smallest integer $\geq r$. If q is a prime power, $\mathbf{F}_q$ denotes $GF(q)$, the finite field of order q. If $p(x, y) \in \mathbf{F}_q[x, y]$, then $\deg_x p(x, y)$ denotes the degree of $p(x, y)$ in x, i.e., the degree of $p(x, y)$ as a polynomial over $\mathbf{F}_q[y]$.

PART ONE

COMPUTER MEMORIES AND BURSTY CHANNELS:

AN INFORMATION-THEORETIC STUDY

CHAPTER I

INFORMATION CONTENT PER CHIP

This chapter is a motivation for the rest of the thesis. Section 1 contains a brief description of computer memories. Soft errors are described in section 2. In section 3, we will try to justify the use of error correcting codes in computer memories.

1. Random Access Memories

VLSI technology has made a great impact on semiconductor memories. Every year, semiconductor memories are made smaller, cheaper, and faster. One of the most important classes of computer semiconductor memories is the class of *random access memory*, known as RAM. In this type of memory, the content of the memory cells can be written or read in any desired sequence.

In dynamic RAMs, which are the most widely used class of computer memories, the content of each cell is stored on a small *storage capacitor*. Logical "0" or "1" are represented by the presence or absence of electric charge on the capacitor. Each cell is accessed by addressing a *word line*, which points to the row containing the cell, and a *bit line*, which points to its column. The information to be read or written on the cell is transmitted through the bit line. A *sense amplifier* is used to amplify the signal read from the cells.

Two important parameters related to the reliability of data storing in dynamic RAMs are the *critical charge* and the *switching energy*. The critical charge is the threshold used to decide if a given cell contains "0" or "1". The switching energy gives the minimum energy required to alter the content of a cell. For a dynamic RAM to be immune from noise, the critical charge and the switching

energy should be sufficiently large. However, in the current trend of increasing the information content per chip by reducing cell sizes, the critical charge and the switching energy are continuously reduced.

The most widely accepted set of rules for scaling RAM cells are due to Dennard et al. [3]. Although these rules are rarely followed literally, they give a simple model that can be used to give a reasonably clear view of the problems to be encountered with small cells. It should be mentioned that Dennard's scaling rules are not expected to be a good guideline as dimensions shrink below the submicron.

Suppose that a scaling factor n, where n is a positive integer, is applied to all physical dimensions, i.e., the lengths, widths, and heights of all devices are divided by n. Then, according to Dennard's rules, the critical charge is scaled down by $1/n^2$, while the switching energy is scaled down by $1/n^3$. This implies that as we progress towards smaller cells, the less reliable the cells become.

In the following section, we will examine some error mechanisms that take place as the switching energy and the critical charge decrease.

2. Soft Errors in Computer Memories

Errors in computer memories are traditionally divided into two classes: ***hard errors*** and ***soft errors***. Hard errors are associated with physical damage to the memory cells. Such damage is permanant and can affect single cells, columns, rows, or even the entire chip. Soft errors, on the other hand, do not cause any physical damage. Hence, the probability of error for each cell does not change after suffering from a soft error.

In the following, we will give a brief description of the sources of soft errors. The most important source which has been already noticed in the 64K RAM chips is alpha-particle radiation [11],[12]. Alpha-particles are emitted

from impurities in the chip package and hit the chip with a rate of .01-.1 alpha-particle/cm^2-hour. The energy spectrum of the emitted alpha-particles is in the range of 2–9 MeV. As an alpha-particle penetrates through a memory cell, it generates electron-hole pairs. The electrons move towards the storage capacitor, and may cause a change in the stored charge that exceeds the critical charge. This can cause an error in that particular cell. It turns out that an error may occur only if the content of the cell is "1". If the cell content is "0", the cell is not sensitive to the alpha-particle. However, it has been noticed that an alpha-particle may hit bit lines or sense amplifiers, and in such case a "0" can be read as "1" [23].

As the cells become smaller, a single alpha-particle may cause a two-dimensional burst of errors in the chip. The burst pattern is confined to a rectangle whose size depends on the energy of the particle as well as on its angle of incidence.

The errors caused by an alpha-particle are soft. The excess charge caused by the particle is completely removed in the following "write" action. It has been also noticed that cosmic rays can produce errors by essentially the same mechanism [24].

Two more sources of soft errors are attributed to quantum and thermal effects. These sources are fundamental in the sense that they are inherent in the basic operation of the cells. Quantum effects [1],[9] are related to the Heisenberg uncertainty principle. Thermal effects [9] are due to the random motion of electrons induced by thermal noise. In the following, we will derive an estimate for the error probability due to thermal effects as a function of the switching energy. The model we consider is the same as that given by Stein [19].* The equivalent electrical circuit is composed of a capacitor and a resistor. The mean

---

* *Unfortunately, the analysis in [19] is not mathematically rigorous, so we were inclined to rederive the results.*

thermal energy generated within the circuit is given by

$$E = \frac{1}{2}kT = \frac{1}{2}\Gamma\overline{u^2},$$

where k is Boltzmann's constant, T is the Kelvin's temperature, u is the noise voltage across the capacitor, $\overline{u^2}$ is its variance, and Γ is the capacitance.

The mean voltage across the capacitor is considered to be 0 or V according to whether the content of the cell is "0" or "1", respectively. The conditional probability density function of the voltage across the capacitor is given by

$$\Pr\{u|\text{“0”}\} = \frac{1}{\sqrt{2\pi kT/\Gamma}} \exp\left(-\frac{u^2\Gamma}{2kT}\right),$$
$$\Pr\{u|\text{“1”}\} = \frac{1}{\sqrt{2\pi kT/\Gamma}} \exp\left(-\frac{(u-V)^2\Gamma}{2kT}\right).$$

The bit error probability ϵ, assuming equal probabilities for storing "0" and "1", is given by

$$\epsilon = \int_{V/2}^{\infty} \Pr\{u|\text{“0”}\}du$$
$$= Q\left(\frac{V}{2\sqrt{kT/\Gamma}}\right),$$

where $Q(x) = \int_x^{\infty} \exp(-t^2/2)dt$. The switching energy is $E_{sw} = \Gamma V^2/2$. Hence,

$$\epsilon = Q\left(\sqrt{\frac{E_{sw}}{2kT}}\right). \tag{1}$$

Using the value of $E_{sw} = 10^{-12}$ Joules, typical for the 64K RAM, we get $\epsilon \approx 10^{-2.6\times10^7}$. This value is very low which reflects the fact that present day RAMs are quite immune from thermal effects. In fact, they are also immune from quantum effects as well [1],[9]. However, alpha-particles have a noticeable effect on these RAMs, which is expected to be intolerable in the near future if the trend towards higher densities increases without combating alpha-particles using new innovations in device technology. Thus, alpha-particles will impose a physical

limit to the persistent trend towards smaller and reliable memory cells, if no coding is used.

Using error correcting codes, the effect of soft errors can be greatly reduced, and hence, reliable memory cells whose sizes are below the apparent physical limits can be produced. This of course is done at the expense of using a certain number of memory cells on the chip as parity checks. To justify the use of error correcting codes, we need to argue that the *information content* per chip, i.e., the number of information bits on a chip, can be increased even if some of the cells are parity checks, as more cells are built per chip. This is the subject of the following section.

3. Minimum Area per Information Bit

We begin with an "abstract" chip of unit area which contains a single memory cell. We apply a scaling factor n, where n is a positive integer, to produce on the original chip n^2 cells. We assume that if error correction is present, it is performed at regular intervals of time, which may vary with n. We assume that the error probabilities of the memory cells are equal. For each memory cell, the probability of error is independent of the cell's content and all other cells. This bit error probability will be denoted by $\epsilon(n)$. With these assumptions, writing and reading bits from the chip is equivalent to transmitting them on a binary memoryless channel with error probability $\epsilon(n)$. Motivated by physical reasons, we assume that $\epsilon(n)$ is an increasing function of n, and $\epsilon(n) \to 1/2$ as $n \to \infty$. In the analysis that follows, $1 - 2\epsilon$ occurs more frequently than ϵ, so we define $\delta = 1 - 2\epsilon$. The channel capacity per cell is given by

$$\begin{aligned} C &= 1 + \epsilon \log \epsilon + (1 - \epsilon) \log(1 - \epsilon) \\ &= \frac{1}{2}[(1 - \delta) \log(1 - \delta) + (1 + \delta) \log(1 + \delta)]. \end{aligned}$$

We note that $C/\delta^2 \to \log e$ as $n \to \infty$.

If a code is used, so that k of the n^2 cells are used to store data, then the rate of the code is defined to be $R = k/n^2$. Shannon [17] has proved that there exist codes with arbitrarily small probability of error if $R < C$. On the other hand, if $R > C$, then no such codes exist. As n increases, we expect a larger probability of error, and so the rate of the code needed to have reliable chips should be reduced. We are interested in the minimum area per information bit (*MAPIB*) required for reliable storing of data, assuming the technology to produce a scaled chip for every positive integer n. This quantity is given by

$$MAPIB = \inf_n \frac{1}{n^2 C(n)}. \tag{2}$$

Since $C(n)/\delta^2(n) \to \log e$ as $n \to \infty$, a necessary and sufficient condition for *MAPIB* to be zero, is that $n\delta(n) \to \infty$ as $n \to \infty$. In that case, an infinite amount of information bits can be stored on a unit area. This is of course unrealistic.

Unfortunately, we cannot determine *MAPIB* from (2) since the bit error probability $\epsilon(n)$ is not known for all values of the scaling factor n. Another problem arises from the memoryless assumption, i.e., that the error probabilities for the cells are independent. This assumption will not hold as the cell dimensions shrink since a single alpha-particle may cause several cells to fail. This aspect will be considered in the next chapter. However, we simply note that neither the flux nor the energy spectrum of alpha-particles depend on scaling. This means that if a cell is small enough to be sensitive to alpha-particles of all energies in the energy spectrum, which ranges from 2–9 MeV, then its error probability depends primarily only on the probability that an alpha-particle hits the surrounding area of the cell, which does not depend on the scaling factor. Hence, alpha-particles impose no limit on the information content per chip by using error correcting codes. This shows the power of error correcting codes since, without them and with no major improvements in technology, alpha-particles undoubtedly impose a limit on the information content per chip.

On the other hand, thermal effects place a fundamental limit on information content per chip, even by using error correcting codes. This result was first reported by Swanson [20] for other types of memory devices. We will show that the result also holds for RAMs. From equation (1), we have for the unscaled chip, i.e., with $n = 1$

$$\delta(1) = 1 - 2\epsilon(1) = 1 - 2Q\left(\sqrt{\frac{E_{sw}}{2kT}}\right).$$

Using Dennard's scaling rules, the switching E_{sw} is scaled by a factor of $1/n^3$. Considering the area of a single cell in the 64K RAM to be unity, and using the value $E_{sw} = 10^{-12}$ Joules typical for the 64K RAM, we get $\delta(n) \approx 1 - 2Q(10^4 n^{-1.5})$. From (2), it follows that $MAPIB \approx 1.2 \times 10^{-5}$, and the capacity per cell C is about 0.5. The optimum value of n is about 400. Hence, even by using error correcting codes, it is impossible to produce a reliable chip with the same area as that of the 64K RAM (which is about 1mm^2, considering the area of the memory cells only) and whose information content exceeds 5.4 Gigabits. Of course, this estimate is very optimistic as we have ignored all other sources of error. Undoubtedly, at these very small dimensions new error mechanisms will be discovered, and more importantly, the whole model, including Dennard's scaling rules, will be invalid. Although this analysis shows that even by using error correcting codes, an ultimate limit on information content cannot be exceeded, the same analysis shows the power of error correcting codes. Indeed, for the optimal cell size, the bit error probability, without coding, is about 0.1, which is far from being tolerated if no coding is used. In other words, by using error correcting codes, reliable chips with very small cell dimensions and with high information densities can be obtained.

In the analysis we have pursued so far we did not consider the complexity of the encoder and the decoder. It was implicitly assumed that the area of the chip is entirely devoted to information and redundant bits. It is expected in

the future that the encoder and the decoder will be implemented on the chip itself. This will impose conditions on the codes to be used. In chapter IV, some two-dimensional burst error correcting codes, which may be helpful in combating alpha-particles, are developed. The first priority in considering these codes is their error correcting capabilities. In the appendix, efficient encoding and decoding algorithms for these codes will be described. However, we are in no position to argue that these algorithms are suitable to be implemented on chips. On the other hand, we expect in the future, that the chip architecture may be considerably modified to be compatible with the codes used.

In the rest of this section, we argue that in the "unrealistic" case of $n\delta(n) \to \infty$ as $n \to \infty$, an infinite number of information bits can still be stored on the chip even if the encoder and the decoder are placed on the chip. However, we will impose the unrealistic condition that the encoder and the decoder are completely immune from noise. Moreover, the area required for connections will be ignored.* Let $K(n)$ be the number of information bits on the scaled chip, and $N(n)$ be the codeword length. Hence, for reliable storing of data, we have $K(n)/N(n) < C(n)$. We assume that the encoder and the decoder are composed of cells that are scaled with the same scaling factor n. The encoder is a ROM (read only memory) that contains a list of all the 2^K codewords. An adder of length N is associated with each codeword. Each adder performs a componentwise addition modulo 2 of the retrieved word and the corresponding codeword, and gives the total number of 1's appearing in the sum, which is the Hamming distance. A codeword, with the closest distance to the retrieved word, is the decoder's estimate, which is a maximum likelihood decision. Hence, the number of equivalent memory cells of the encoder and the decoder N_{ed} is bounded by $cN \times 2^K$, where c is a constant which is independent of n. For the encoder and the decoder to be built on the chip, we should have $n^2 \geq N + N_{ed} \geq N + cN \times 2^K$. What we need to show is that if $n\delta(n) \to \infty$ as

* *The author encourages the reader not to be serious in reading this paragraph.*

$n \to \infty$, then we can make $K \to \infty$ such that the inequalities $K/N < C$ and $n^2 \geq N + cN \times 2^K$ are satisfied. Since $C/\delta^2 \to \log e$, then by having $N = n/\delta(n)$, the number of information bits per chip K can tend to infinity as $\log n\delta(n)$. It may be interesting to note that although the entire chip is almost devoted to the encoder and the decoder, and the tiny part that remains is itself almost entirely occupied by redundant bits, an infinite number of bits can still be stored on the chip.

CHAPTER II

BURSTY CHANNELS

In this chapter, we study a new class of channels with memory which we call bursty channels. In section 1, bursty channels are defined. Their capacities are derived in section 2. In section 3, asymptotic values of the capacities are obtained. Coding schemes that are asymptotically optimum are described in section 4. In section 5, we apply the model of bursty channels to computer memories suffering from alpha-particles.

Throughout this chapter, uppercase letters denote random variables. Sequences are written in boldface letters. The l-sequence $(v_1, v_2, \ldots, v_l)$ is denoted by $\mathbf{v}_l$. If $\mathcal{S}$ is a set, then the set $\mathcal{S}^l$ is the product of l copies of $\mathcal{S}$.

1. Definitions

Let $\mathcal{S}$ be a nonempty finite or countable set whose elements are called *states*. Let $\mathbf{S} = S_1, S_2, \ldots$, where S_i is $\mathcal{S}$-valued random variable, be a stochastic process such that the following conditions are satisfied:

(1) The stochastic process $\mathbf{S}$ is strictly stationary [4; chapter II], i.e., for any positive integer t, the multivariate distribution of $S_{h+1}, S_{h+2}, \ldots, S_{h+t}$ is independent of h, as long as $h \geq 0$.

(2) There exists a least nonnegative integer m, called *the duration of state memory*, such that any two finite sequences in $\mathbf{S}$ separated by at least m states are independently distributed, i.e., if t and t' are two positive integers, $s_1, \ldots, s_t, s'_1, \ldots, s'_{t'} \in \mathcal{S}$, $i \geq 0$, and $i' \geq i + t + m$, then,

$$\Pr\{S_{i+1} = s_1, \ldots, S_{i+t} = s_t, S_{i'+1} = s'_1, \ldots, S_{i'+t'} = s'_{t'}\} =$$

$$\Pr\{S_{i+1} = s_1, \ldots, S_{i+t} = s_t\} \Pr\{S_{i'+1} = s'_1, \ldots, S_{i'+t'} = s'_{t'}\}.$$

Let $\mathcal{X}$ and $\mathcal{Y}$ be two nonempty finite sets. For every state $s \in \mathcal{S}$, we associate a discrete memoryless channel, called the *component channel* defined by s, whose input and output alphabets are $\mathcal{X}$ and $\mathcal{Y}$, respectively, and whose channel probability matrix is $[p(y|x,s)]$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$.

The *bursty channel* is defined as follows. When a sequence $\mathbf{x} = x_1, x_2, \ldots$, where $x_i \in \mathcal{X}$, is transmitted, the ith component x_i is transmitted through the component channel defined by s_i, where s_i is the ith component of $\mathbf{S}$. The process $\mathbf{S} = S_1, S_2, \ldots$, is called the *state stochastic process* of the bursty channel. The l-sequence $(S_1, \ldots, S_l)$ is called the *state random l-sequence.*

As an example of a bursty channel, we consider the case where $\mathcal{S}$ is finite, and the components of $\mathbf{S}$ are independent and identically distributed random variables. In this case $m = 0$. This channel has been introduced by Shannon [18]. In the next section, we will make use of this simple channel to derive the capacity of the general bursty channel.

Block interference channels introduced by McEliece and Stark [14], are in certain aspects similar to bursty channels. The only difference is that in block interference channels, the state stochastic process $\mathbf{S} = S_1, S_2, \ldots$, is such that $S_{im+1} = S_{im+2} = \ldots = S_{(i+1)m}$ for every nonnegative integer i, where m is some positive integer. Moreover, S_{im+1} and S_{jm+1} are independent for $i \neq j$. Generally speaking, $\mathbf{S}$ is not stationary, and thus block interference channels may not be bursty channels. However, by considering the channel whose input is an m-sequence $(x_{im+1}, x_{im+2}, \ldots, x_{(i+1)m})$, where $i = 0, 1, \ldots$, then block interference channels can be viewed as bursty channels whose state stochastic process is a sequence of independent, identically distributed, random variables. In our treatment of bursty channels, we will use many techniques introduced in the study of block interference channels.

2. Channel Capacities

In this section, we derive the capacity of the bursty channel in the following three cases:*

(i) Neither the encoder nor the decoder knows the state which governs the transmission of each letter.

(ii) The decoder, but not the encoder, knows the state which governs the transmission of each letter.

(iii) Both the encoder and the decoder know the state which governs the transmission of each letter. We assume that the encoder knows all the states before encoding.

We note that in cases (i) and (ii) the input is independent of the states.

For each of the above three cases, a capacity is defined as follows. First, we define an (n, M, λ) code to be a code of M codewords, each of length n, and whose error probability does not exceed λ. Now, suppose that there exists a real number C such that if $\epsilon > 0$ and $0 < \lambda < 1$, then for all sufficiently large n, there exists an (n, M, λ) code with $M > 2^{n(C-\epsilon)}$, and any (n, M, λ) code satisfies $M < 2^{n(C+\epsilon)}$. If such C exists, it is called the *capacity*. We will argue in the following that in each of the three cases, the capacity exists. We denote by C, C_d, and C_{ed} the capacities corresponding to cases (i),(ii), and (iii), respectively.

For the example of bursty channels given in section 1, where S is finite and the components of $\mathbf{S}$ are independent and identically distributed random variables, the capacities C, C_d, and C_{ed} are known to exist. Let S, X, and Y denote single components of each of $\mathbf{S}$, the random input, and the random

---

* *Case (iv), in which the encoder, but not the decoder, knows the states, will not be considered here. In fact, we are mainly interested in case (i). Cases (ii) and (iii) are primarily used to gain a better understanding of case (i).*

output, respectively. Then, the capacities are given by [8],[18],[21; section 4.6],

$$C = \max_{p(x)} I(X;Y), \tag{1}$$

$$C_d = \max_{p(x)} I(X;Y|S), \tag{2}$$

$$C_{ed} = \max_{p(x|s)} I(X;Y|S), \tag{3}$$

where $I(X;Y)$ and $I(X;Y|S)$ denote the mutual information, and the conditional mutual information given S, between X and Y, respectively. The maximization is taken over all probability distributions, where $p(x)$ and $p(x|s)$ denote the distributions of x and of x given s, respectively. The expression of C_{ed} is an abbreviation for $\sum_{s\in\mathcal{S}} \max_{p(x)} I(X;Y|s)\Pr\{S = s\}$, The assumption that $\mathcal{S}$ is finite can be removed by using the results of [21; chapter 8].

Now, we return to the general bursty channel with state stochastic process $\mathbf{S}$, and duration of state memory m. We will consider only the capacity C to avoid repetition since the argument holds for C_d and C_{ed} as well. Let $n = k(l+m)+t$, where n, k, l, t are integers, $k, l \geq 1$, and $0 \leq t < l+m$. Let $\mathbf{S}_n = (S_1, \ldots, S_n)$ denote the first n components of $\mathbf{S}$, i.e., the state random n-sequence. Define the random k-sequence $\mathbf{R}_k = (R_1, \ldots, R_k)$, where $R_i = (S_{(i-1)(l+m)+1}, S_{(i-1)(l+m)+2}, \ldots, S_{(i-1)(l+m)+l})$ for $1 \leq i \leq k$. We define the channel Δ to be the channel whose input and output alphabets are $\mathcal{X}^l$ and $\mathcal{Y}^l$, respectively, and whose state random k-sequence is $\mathbf{R}_k$. Conditions (1) and (2) stated in section 1 ensure that the components of $\mathbf{R}_k$ are independent and identically distributed random variables. Hence, from (1), we have

$$C_\Delta(l) = \max_{p(x_l)} I(\mathbf{X}_l; \mathbf{Y}_l),$$

where $C_\Delta(l)$ denote the capacity of Δ. Here, $\mathbf{X}_l$ and $\mathbf{Y}_l$ denote the input and output random 1-sequences, respectively, of channel Δ. Let

$$C = \sup_l \frac{C_\Delta(l)}{l+m}.$$

In the following, we will argue that C is the capacity of the bursty channel corresponding to case (i). First, consider a discrete memoryless channel with zero capacity, and whose input and output alphabets are $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let s' denote the state of such channel. We also consider a discrete memoryless channel with capacity $\log|\mathcal{X}|$, whose input and output alphabets are $\mathcal{X}$ and $\bar{\mathcal{Y}}$, where $\bar{\mathcal{Y}}$ is a finite set which contains $\mathcal{Y}$ such that $|\bar{\mathcal{Y}}| \geq |\mathcal{X}|$. Let s'' denote the state of such channel. Note that these two discrete memoryless channels may actually be among the component channels of the bursty channel. From the random n-sequence $\mathbf{S}_n$, we define the random n-sequence $\mathbf{S}'_n = (S'_1, \ldots, S'_n)$ as

$$S'_i = \begin{cases} s', & \text{for } i \equiv l+1, l+2, \ldots, l+m \pmod{l+m}, \\ & \text{or } k(l+m) < i \leq n; \\ S_i, & \text{otherwise.} \end{cases}$$

The random n-sequence $\mathbf{S}''_n$ is defined similarly after replacing s' by s''. Let Δ' be the channel with input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and with state random n-sequence $\mathbf{S}'_n$. Similarly, let Δ'' be the channel with input and output alphabets $\mathcal{X}$ and $\bar{\mathcal{Y}}$, respectively, and with state random n-sequence $\mathbf{S}''_n$. Let $C_{\Delta'}(l)$ and $C_{\Delta''}(l)$ be the capacities of Δ' and Δ'', respectively. Clearly, these capacities are given by

$$C_{\Delta'}(l) = \frac{1}{l+m} C_{\Delta}(l), \tag{4}$$

and

$$C_{\Delta''}(l) = \frac{1}{l+m} C_{\Delta}(l) + \frac{m}{l+m} \log|\mathcal{X}|. \tag{5}$$

Suppose we are given ϵ and λ such that $\epsilon > 0$ and $0 < \lambda < 1$. Obviously, an (n, M, λ) code for channel Δ' is is an (n, M, λ) code for the bursty channel. Hence, it follows from the definition of the capacity of Δ' and (4), that there exists for sufficiently large n, an (n, M, λ) code for the bursty channel with

$$M > 2^{n(C_{\Delta'}(l) - \epsilon/2)} = 2^{n(C_{\Delta}(l)/(l+m) - \epsilon/2)} \geq 2^{n(C-\epsilon)}, \tag{6}$$

if l is chosen such that $C \leq C_{\Delta}(l)/(l+m) + \epsilon/2$. On the other hand, an (n, M, λ)

code for the bursty channel is an (n, M, λ) code for channel Δ''. Hence, it follows from the definition of the capacity of Δ'' and (5), that for sufficiently large n, any (n, M, λ) code for the bursty channel satisfies

$$M < 2^{n(C_{\Delta''}(l)+\epsilon/2)} = 2^{n(\frac{1}{l+m}C_{\Delta}(l)+\frac{m}{l+m}\log|\mathcal{X}|+\epsilon/2)}.$$

By choosing l large enough so that $m \log|\mathcal{X}|/(l+m) < \epsilon/2$, we get

$$M < 2^{n(\frac{1}{l+m}C_{\Delta}(l)+\epsilon)} \leq 2^{n(C+\epsilon)}. \tag{7}$$

From (6) and (7), it follows that the capacity of the bursty channel exists, and is equal to C. Furthermore, since $C_{\Delta'}(l) \leq C \leq C_{\Delta''}(l)$, we get by using (4) and (5),

$$\frac{1}{l+m}C_{\Delta}(l) \leq C \leq \frac{1}{l+m}C_{\Delta}(l) + \frac{m}{l+m}\log|\mathcal{X}|.$$

Hence,

$$C = \lim_{l\to\infty}\frac{C_{\Delta}(l)}{l} = \lim_{l\to\infty}\max_{p(\mathbf{x}_l)}\frac{I(\mathbf{X}_l;\mathbf{Y}_l)}{l}.$$

So far, we have treated case (i). The same argument holds for cases (ii) and (iii) by using (2) and (3). Hence, the next lemma is proved. Before stating it, recall the definition of $\mathbf{X}_l$ and $\mathbf{Y}_l$ as the input and output random 1-sequences, respectively, of channel Δ. The random variable $\mathbf{Y}_l$ depends on $\mathbf{X}_l$ and R_1, which is the state random variable of channel Δ. But $R_1 = (S_1, S_2, \ldots, S_l)$ is the state random l-sequence of the bursty channel. Hence, $\mathbf{X}_l$ and $\mathbf{Y}_l$ are the input and output random l-sequences of the bursty channel. Thus, we can state the following lemma without referring to channel Δ.

Lemma 1. *Let $\mathbf{X}_l$, $\mathbf{Y}_l$, and $\mathbf{S}_l$ be the input, output, and state random l-sequences, respectively, of the bursty channel. Then,*

$$C = \sup_l \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l)}{l+m} = \lim_{l\to\infty} \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l)}{l},$$

$$C_d = \sup_l \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)}{l+m} = \lim_{l\to\infty} \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)}{l},$$

$$C_{ed} = \sup_l \max_{p(\mathbf{x}_l|\mathbf{s}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)}{l+m} = \lim_{l\to\infty} \max_{p(\mathbf{x}_l|\mathbf{s}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)}{l}.$$

The next theorem implies that in cases (ii) and (iii), the statistical dependence between the components of $\mathbf{S}$ is immaterial given the first order distribution of the states.

Theorem 2. *Let l be a positive integer, $\mathbf{X}_l$, $\mathbf{Y}_l$, and $\mathbf{S}_l$ be the input, output, and state random l-sequences, respectively. Then,*

$$C_d = \max_{p(\mathbf{x}_l)} I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)/l,$$

$$C_{ed} = \max_{p(\mathbf{x}_l|\mathbf{s}_l)} I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l)/l.$$

Proof. For $\mathbf{s}_l = (s_1,\ldots,s_l) \in \mathcal{S}^l$ and $\Pr\{\mathbf{S}_l = \mathbf{s}_l\} > 0$, let $\mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l|\mathbf{s}_l}$ denote expectation on the joint sample space of $\mathbf{X}_l$ and $\mathbf{Y}_l$ given $\mathbf{S}_l = \mathbf{s}_l$. Since the component channels are memoryless, we have

$$p(\mathbf{y}_l|\mathbf{x}_l,\mathbf{s}_l) = \prod_{i=1}^{l} p(y_i|x_i,s_i).$$

From the definition of the mutual information, we have

$$I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{s}_l) = \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l|\mathbf{s}_l}\left[\log \frac{p(\mathbf{y}_l|\mathbf{x}_l,\mathbf{s}_l)}{p(\mathbf{y}_l|\mathbf{s}_l)}\right],$$

and

$$\begin{aligned}\sum_{i=1}^{l} I(X_i;Y_i|s_i) &= \sum_{i=1}^{l} \mathrm{E}_{X_i,Y_i|s_i}\left[\log \frac{p(y_i|x_i,s_i)}{p(y_i|s_i)}\right] \\ &= \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l|\mathbf{s}_l}\left[\log \frac{p(\mathbf{y}_l|\mathbf{x}_l,\mathbf{s}_l)}{\prod_{i=1}^{l} p(y_i|s_i)}\right].\end{aligned}$$

By using Jensen's inequality, we get

$$I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{s}_l) - \sum_{i=1}^{l} I(X_i;Y_i|s_i) = \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l|\mathbf{s}_l}\left[\log \frac{\prod_{i=1}^{l} p(y_i|s_i)}{p(\mathbf{y}_l|\mathbf{s}_l)}\right]$$
$$\leq \log \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l|\mathbf{s}_l}\left[\frac{\prod_{i=1}^{l} p(y_i|s_i)}{p(\mathbf{y}_l|\mathbf{s}_l)}\right] = 0,$$

with equality if, and only if, $p(\mathbf{y}_l|\mathbf{s}_l) = \prod_{i=1}^{l} p(y_i|s_i)$. The later equality is satisfied if $p(\mathbf{x}_l|\mathbf{s}_l) = \prod_{i=1}^{l} p(x_i|s_i)$. Hence, we have

$$\max_{p(\mathbf{x}_l|\mathbf{s}_l)} I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{s}_l) = \sum_{i=1}^{l} \max_{p(x_i|s_i)} I(X_i;Y_i|s_i).$$

Thus,

$$\max_{p(\mathbf{x}_l|\mathbf{s}_l)} I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{S}_l) = \sum_{\mathbf{s}_l\in\mathcal{S}^l} \max_{p(\mathbf{x}_l|\mathbf{s}_l)} I(\mathbf{X}_l;\mathbf{Y}_l|\mathbf{s}_l)\Pr\{\mathbf{S}_l = \mathbf{s}_l\}$$
$$= \sum_{\mathbf{s}_l\in\mathcal{S}^l}\sum_{i=1}^{l} \max_{p(x_i|s_i)} I(X_i;Y_i|s_i)\Pr\{\mathbf{S}_l = \mathbf{s}_l\}$$
$$= \sum_{i=1}^{l}\sum_{s\in\mathcal{S}} \max_{p(x_i|s)} I(X_i;Y_i|s)\ \Pr\{S_i = s\}$$
$$= l \max_{p(x_1|s_1)} I(X_1;Y_1|S_1),$$

since $\Pr\{S_i = s\}$ is independent of i by condition (1) of section 1. Now, it follows from Lemma 1 that Theorem 2 holds for C_{ed}. The same argument proves the theorem for C_d as well after taking into account that $\mathbf{X}_l$ and $\mathbf{S}_l$ are independent in case (ii). ∎

It is interesting to study the case when $C_d = C_{ed}$. First, we say that the component channels are *compatible* if there exists an input distribution that achieves capacity for all of the component channels whose states have positive probability. The next corollary then follows immediately from Theorem 2 by setting $l = 1$.

Corollary 3. *$C_{ed} = C_d$ if, and only if, the component channels are compatible.*

Of course, this corollary does not imply that the information provided to the

encoder about the states is useless if the component channels are compatible. In fact, by using this information the encoding and decoding complexity may be reduced considerably.

Theorem 2 offers an easy way to calculate C_{ed} and C_d by setting $l = 1$. However, for the capacity C, which is generally more interesting, Theorem 2 does not apply. In fact, the statistical dependence between the components of $\mathbf{S}$ plays a vital rule in determining C. This will be considered in the following section. We end this section with a theorem that implies a lower bound on C, which is tighter than that implied by Lemma 1.

Theorem 4. *Let $\mathbf{X}_l$ and $\mathbf{Y}_l$ be the input and output random l-sequences of the bursty channel, respectively. Then,*

$$C = \lim_{l\to\infty} \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l)}{l} = \sup_{l} \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l;\mathbf{Y}_l)}{l}.$$

Proof. The first equality follows from Lemma 1. To prove the second equality, it suffices to show that

$$\max_{p(\mathbf{x}_{lr})} \frac{I(\mathbf{X}_{lr};\mathbf{Y}_{lr})}{lr} \geq \max_{p(\mathbf{x}_r)} \frac{I(\mathbf{X}_r;\mathbf{Y}_r)}{r} \tag{8}$$

for all positive integers l and r. Without loss of generality, we give the proof for $r = 1$. Consider the components $X_1, \ldots, X_l$ of $\mathbf{X}_l$ to be independent. The following argument is from [13; chapter 1]. Let $\mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}$ denote expectation on the joint sample space of $\mathbf{X}_l$ and $\mathbf{Y}_l$. Then,

$$\begin{aligned} I(\mathbf{X}_l;\mathbf{Y}_l) &= \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}\left[\log \frac{p(\mathbf{x}_l|\mathbf{y}_l)}{p(\mathbf{x}_l)}\right] \\ &= \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}\left[\log \frac{p(\mathbf{x}_l|\mathbf{y}_l)}{\prod_{i=1}^{l} p(x_i)}\right], \end{aligned}$$

and

$$\sum_{i=1}^{l} I(X_i, Y_i) = \sum_{i=1}^{l} \mathrm{E}_{X_i,Y_i}\left[\log \frac{p(x_i|y_i)}{p(x_i)}\right]$$
$$= \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}\left[\log \frac{\prod_{i=1}^{l} p(x_i|y_i)}{\prod_{i=1}^{l} p(x_i)}\right].$$

By using Jensen's inequality, we get

$$\sum_{i=1}^{l} I(X_i; Y_i) - I(\mathbf{X}_l; \mathbf{Y}_l) = \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}\left[\log \frac{\prod_{i=1}^{l} p(x_i|y_i)}{p(\mathbf{x}_l|\mathbf{y}_l)}\right]$$
$$\leq \log \mathrm{E}_{\mathbf{X}_l,\mathbf{Y}_l}\left[\frac{\prod_{i=1}^{l} p(x_i|y_i)}{p(\mathbf{x}_l|\mathbf{y}_l)}\right] = 0.$$

Thus, we have

$$\max_{p(\mathbf{x}_l)} I(\mathbf{X}_l; \mathbf{Y}_l) \geq \sum_{i=1}^{l} \max_{p(x_i)} I(X_i; Y_i) = l \max_{p(x_1)} I(X_1; Y_1).$$

Thus, (8) holds for $r = 1$. ∎

3. Asymptotic Values of Capacities

Consider a bursty channel whose state stochastic process is $\mathbf{S} = S_1, S_2, \ldots$. We define the *run length random variable* to be the value of L such that $S_1 = S_2 = \ldots = S_L \neq S_{L+1}$, if such L exists. If no such L exists, we take $L = \infty$. We also define for each positive integer l, the *l-run random variable* U_l as

$$U_l = \begin{cases} 0, & \text{if } L \geq l; \\ 1, & \text{otherwise.} \end{cases}$$

Now, consider a sequence of bursty channels $\Gamma_1, \Gamma_2, \ldots$, defined on the same state set $\mathcal{S}$, input alphabet $\mathcal{X}$, and output alphabet $\mathcal{Y}$. However, the duration of state memory may not be the same for all channels. For channel Γ_n, let L_n, $U_{l,n}$, and $S_{l,n}$ be the run length random variable, the l-run random variable, and state random l-sequence, respectively. Every bursty channel satisfies conditions (1) and (2) of section 1 by definition. We assume that the sequence $\Gamma_1, \Gamma_2, \ldots$ satisfies also the following two conditions:

(3) The sequence of random variables $L_1, L_2, \ldots$ diverges to ∞ in probability.*

* *That is to say, for each l and $\epsilon > 0$, $\Pr\{L_n < l\} < \epsilon$ if n is sufficiently large [2; chapter 4].*

Equivalently, for each positive integer l, the sequence of random variables $U_{l,1}, U_{l,2}, \ldots$, converges to 0 in probability.

(4) The sequence of state random 1-sequences $S_{1,1}, S_{1,2}, \ldots$ converges in distribution to some random variable S_1.

Let $C(n), C_d(n)$, and $C_{ed}(n)$ be the capacities of channel Γ_n corresponding to cases (i),(ii), and (iii) of section 2, respectively. From Theorem 2, the next theorem follows because of condition (4).

Theorem 5. *Let X_1 and Y_1 be the input and output random 1-sequences, respectively. Then,*

$$\lim_{n\to\infty} C_d(n) = \max_{p(x_1)} I(X_1; Y_1|S_1),$$
$$\lim_{n\to\infty} C_{ed}(n) = \max_{p(x_1|s_1)} I(X_1; Y_1|S_1).$$

Now, we state and prove the most important result in this chapter. In the proof, as well as in the next section, some basic properties of mutual information and conditional mutual information will be used. These properties can be found in [7; chapter 2],[22].

Theorem 6.

$$\lim_{n\to\infty} C(n) = \lim_{n\to\infty} C_d(n).$$

Proof. Let n and l be positive integers to be determined later. Consider the bursty channel Γ_n. From Theorems 2 and 4, we have

$$C_d(n) - C(n) \le \frac{1}{l}\left[\max_{p(\mathbf{x}_l)} I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n}|\mathbf{S}_{l,n}) - \max_{p(\mathbf{x}_l)} I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n})\right],$$

where $\mathbf{X}_{l,n}$, $\mathbf{Y}_{l,n}$ and $\mathbf{S}_{l,n}$ are the input, output, and state random l-sequences, respectively, of channel Γ_n. Hence,

$$C_d(n) - C(n) \le \frac{1}{l}\max_{p(\mathbf{x}_l)}[I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n}|\mathbf{S}_{l,n}) - I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n})]. \qquad (9)$$

Since $\mathbf{X}_{l,n}$ and $\mathbf{S}_{l,n}$ are independent, $I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n}|\mathbf{S}_{l,n}) = I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n},\mathbf{S}_{l,n})$. We have

$$\begin{aligned} I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n},\mathbf{S}_{l,n}) - I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n}) &= I(\mathbf{X}_{l,n};\mathbf{S}_{l,n}|\mathbf{Y}_{l,n}) \\ &= I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}) - I(\mathbf{S}_{l,n};\mathbf{Y}_{l,n}) \\ &\leq I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}) \\ &= I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n},U_{l,n}) \\ &= I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}) + I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};U_{l,n}), \end{aligned}$$

where we have used the fact that $U_{l,n}$ is a function of $\mathbf{S}_{l,n}$. Since $U_{l,n}$ is a binary random variable, we have

$$I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n},\mathbf{S}_{l,n}) - I(\mathbf{X}_{l,n};\mathbf{Y}_{l,n}) \leq I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}) + 1. \tag{10}$$

But

$$\begin{aligned} I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}) = {} & I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}=0)\Pr\{U_{l,n}=0\} \\ & + I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}=1)\Pr\{U_{l,n}=1\}. \end{aligned}$$

Hence, we have

$$\begin{aligned} I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}) \leq {} & I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}=0) \\ & + l\log(|\mathcal{X}||\mathcal{Y}|)\Pr\{U_{l,n}=1\}. \end{aligned} \tag{11}$$

To bound $I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}=0)$, we use the same technique as that presented in [14]. Let T be the histogram describing the number of times each of the pairs $(x,y) \in \mathcal{X}\times\mathcal{Y}$ occurs among the l pairs $(X_1,Y_1),(X_2,Y_2),\ldots,(X_l,Y_l)$. Then T is a sufficient statistic for $\mathbf{S}_{l,n}$ given $U_{l,n}=0$, i.e., given T and $U_{l,n}=0$, the joint random variable $(\mathbf{X}_{l,n},\mathbf{Y}_{l,n})$ is independent of $\mathbf{S}_{l,n}$ [6; chapter 3]. Since T is a function of $\mathbf{X}_{l,n}$ and $\mathbf{Y}_{l,n}$, we have

$$\begin{aligned} I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|U_{l,n}=0) &= I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n},T;\mathbf{S}_{l,n}|U_{l,n}=0) \\ &= I(T;\mathbf{S}_{l,n}|U_{l,n}=0) + I(\mathbf{X}_{l,n},\mathbf{Y}_{l,n};\mathbf{S}_{l,n}|T,U_{l,n}=0) \\ &= I(T;\mathbf{S}_{l,n}|U_{l,n}=0) \end{aligned}$$

$$\leq H(T)$$

$$\leq |\mathcal{X}||\mathcal{Y}| \log(l+1), \tag{12}$$

since there are at most $(l+1)^{|\mathcal{X}||\mathcal{Y}|}$ such histograms. From (10), (11), and (12), we have

$$I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n}, \mathbf{S}_{l,n}) - I(\mathbf{X}_{l,n}; \mathbf{Y}_{l,n}) \leq |\mathcal{X}||\mathcal{Y}| \log(l+1) + l \log(|\mathcal{X}||\mathcal{Y}|)\Pr\{U_{l,n} = 1\} + 1.$$

From (9), we have

$$C_d(n) - C(n) \leq \frac{1}{l}\left[|\mathcal{X}||\mathcal{Y}| \log(l+1) + l \log(|\mathcal{X}||\mathcal{Y}|)\Pr\{U_{l,n} = 1\} + 1\right].$$

Now, we prove that $\lim[C_d(n) - C(n)] = 0$ as $n \to \infty$. Suppose $\delta > 0$ is given. Choose l large enough such that $[(|\mathcal{X}||\mathcal{Y}|) \log(l+1) + 1]/l < \delta/2$. From condition (3), it follows that $\log(|\mathcal{X}||\mathcal{Y}|)\Pr\{U_{l,n} = 1\} < \delta/2$ for all sufficiently large n. This proves that $C_d(n) - C(n) < \delta$ for such n. The theorem now follows since $\lim C_d(n)$ exists by Theorem 5. ∎

Theorem 6 can be heuristically explained as follows, in case of finite $\mathcal{S}$. As the state sequence form long runs, i.e., subsequences of elements of $\mathcal{S}$ of like kind, the decoder can infer the component channel governing the transmission of each letter with small probability of error. This can be done by an agreement between the encoder and the decoder to divide the word sent into blocks of some fixed length, and within each block a certain packet of shorter length that is known to the decoder is sent. The decoder assumes that all the letters transmitted in each block are governed by the same component channel, and guesses these channels from the output distributions of the packets. Hence, the information provided to the decoder about the states in case (ii) may be dispensible, if a certain probability of error in guessing the states is tolerated. McEliece and Stark [14] have noticed the same phenomenon in block interference channels where all the runs have the same length, and both the encoder and the decoder know where each run starts and ends. Thus, in block interference channels, by choosing the

blocks to have the same length as the runs, then the assumption made by the decoder that each block is governed by the same state holds. However, in bursty channels, the runs may start and end anywhere. Furthermore, even if the states form very long runs, the decoder will not be able, in general, to decide where each run starts or ends within an arbitrarily small probability of error.

In the following section, we will show that this argument does not only give a heuristic explanation for Theorem 6, but also offers a coding scheme to achieve the asymptotic value of $C(n)$.

4. Smart Interleaving

In this section we will confine ourselves to bursty channels with a finite state set $\mathcal{S}$. We are interested in case (i) of section 2 in which neither the encoder nor the decoder knows the state which governs the transmission of each letter. We assume, without loss of generality, that no two different states in $\mathcal{S}$ are associated with the same component channel.

Let t be a nonnegative integer. A *test packet* of length t is a fixed string of t letters in $\mathcal{X}$. Let $l > t$ be some positive integer. Now, suppose that the coding scheme is such that every codeword is divided into blocks of length l each, except possibly the last block which will be discarded. Each block starts with the test packet. The test packet and the number l are known to the decoder. The decoder, upon receiving a word, looks at every block of length l, and assumes that the states are the same in any particular block. Furthermore, the decoder infers the state of such block from the output corresponding to the test packet contained in the block. Having done that, the decoder decodes the received word assuming that it knows the state governing each letter sent.

The decisions made by the decoder from the test packets form an *estimated state stochastic process* $\hat{\mathbf{S}}$, whose components are in $\mathcal{S}$. The ith component $\hat{S}_i$ is the estimated state variable of the ith block. Let $m' = \lfloor m/l \rfloor$. Then $\hat{\mathbf{S}}$ satisfies

conditions (1) and (2) of section 1 if m is replaced by m'.

The coding scheme described so far can be looked at as follows. The encoder sends from every block the $l - t$ letters which are not used in the test packet. These letters form a "reduced" codeword. The decoder is supplied with the estimated state sequence, and decodes the received reduced word accordingly. Thus, the decoder is provided with a noisy version of the state sequence by depriving t/l of the length of the codeword from any information relevant to the message sent.

From section 3, there exists a coding scheme that achieves the capacity C if t is set to 0. The main problem in doing that is the coding complexity needed to utilize the statistical dependency of the components of the state sequence. Our aim is to find a good coding scheme for the reduced codewords that ignores competely this dependency. These later codes, if they exist, are in general easier to find and implement. The scheme which will be considered is to make the reduced codeword composed of $(l - t)(m' + 1)$ subcodewords. The ith subcodeword is composed of all letters whose coordinates are congruent to i mod $(l - t)(m' + 1)$. The states corresponding to such letters are independent since they are separated by at least m states in the original codeword, i.e., the codeword which contains the test packets. Furthermore, we require that the subcodewords are independent of each other such that the statistical dependency of the states of the subcodewords is completely ignored by both the encoder and the decoder. A coding scheme under these restrictions will be called *smart interleaving*, a term used in [14] in connection with block interleaving channels. The word "smart" is used to distinguish such scheme from "ordinary" interleaving in which no test packets are used, i.e., $t = 0$. Of course, in certain cases, as in the case of $m = 0$, smart interleaving with $t > 0$ in not the smartest thing to do and ordinary interleaving is better.

In general, the restrictions imposed on smart interleaving are too strict in

the sense that no coding scheme exists under these conditions that achieves the capacity C. In the next lemma, we will find the capacity C_{SI} of smart interleaving. First, note that the estimated state sequence is independent of the input, and that it is also independent of both the input and the output given the state sequence. Here, as well as in the following, the input, output, and state variables refer to the letters in the reduced codeword, i.e., the codeword that does not contain the test packets.

Lemma 7. *Let X_1, Y_1, and $\hat{S}_1$ be the input, output, and estimated state random 1-sequences, respectively. Then,*

$$C_{SI} = (1 - t/l) \max_{p(x_1)} I(X_1; Y_1, \hat{S}_1).$$

Proof. Clearly, C_{SI} is $(1 - t/l)$ times the capacity of the channel where $\hat{S}_1$ is known to the decoder, and the states are independent, identically distributed, and given by the random variable S_1, the state random 1-sequence. This channel may be considered to have input X_1 and output $(Y_1, \hat{S}_1)$. The capacity of this channel is $\max_{p(x_1)} I(X_1; Y_1, \hat{S}_1)$. ∎

In the following, we assume that the test packet and the estimate $\hat{S}$ are chosen to maximize C_{SI}.

One may consider the channel defined in the proof of Lemma 7 with input X_1 and output $(Y_1, \hat{S}_1)$ to be a cascade of a channel with input X_1 and output (Y_1, S_1), followed by a channel with input (Y_1, S_1) and output $(Y_1, \hat{S}_1)$ in which $\hat{S}_1$ depends only on S_1. This model is valid since $\hat{S}_1$ is independent of X_1 and Y_1 given S_1. The following lemma gives a bound on the loss of mutual information due to the second channel. In this lemma, H denotes the entropy function.

Lemma 8. *Let X_1, Y_1, S_1, and $\hat{S}_1$ be the input, output, state, and estimated state random 1-sequences, respectively. Then,*

$$I(X_1;Y_1,\hat{S}_1) \geq I(X_1;Y_1,S_1) - H(S_1|\hat{S}_1).$$

Proof.

$$\begin{aligned} I(X_1;Y_1,S_1,\hat{S}_1) &= I(X_1;Y_1,S_1) + I(X_1;\hat{S}_1|Y_1,S_1) \\ &= I(X_1;Y_1,\hat{S}_1) + I(X_1;S_1|Y_1,\hat{S}_1). \end{aligned}$$

But $I(X_1;\hat{S}_1|Y_1,S_1) = 0$ since $\hat{S}_1$ and X_1, Y_1 are independent given S_1. Hence,

$$\begin{aligned} I(X_1;Y_1,\hat{S}_1) &= I(X_1;Y_1,S_1) - I(X_1;S_1|Y_1,\hat{S}_1) \\ &\geq I(X_1;Y_1,S_1) - H(S_1|Y_1,\hat{S}_1) \\ &\geq I(X_1;Y_1,S_1) - H(S_1|\hat{S}_1). \end{aligned}$$

■

Let U_l be the l-run random variable defined in section 3. The next lemma shows that if $U_l = 0$, then $H(S_1|\hat{S}_1)$ can be made arbitrarily small by increasing t.

Lemma 9. *Let $\delta > 0$. Then for every sufficiently large t, there exists a test packet of length t such that $H(S_1|\hat{S}_1, U_l = 0) < \delta$ if $l > t$.*

Proof. Define the *distance* $d(s,s')$ between the states s and $s' \in \mathcal{S}$ by

$$d(s,s') = \max\{|p(y|x,s) - p(y|x,s')| : (x,y) \in \mathcal{X} \times \mathcal{Y}\}.$$

By the assumption made in the beginning of this section that no two different states share the same component channel, it follows that $d(s,s') > 0$ for $s \neq s'$. Define the *minimum distance* of the states as

$$d_{\min} = \min\{d(s,s') : s,s' \in \mathcal{S}, s \neq s'\}.$$

Consider a test packet of length $t \geq |\mathcal{X}|$, and let $c = \lfloor t/|\mathcal{X}| \rfloor$. In the following,

we will ignore the last $t - c|\mathcal{X}|$ letters in the test packet, and consider it to be composed of the first $c|\mathcal{X}|$ letters only. Let each letter in $\mathcal{X}$ appear c times in the test packet. Define the random variable $N(x,y)$, where $(x,y) \in \mathcal{X} \times \mathcal{Y}$, as the number of times the letter y appears in the output of the test packet corresponding to input x. Define the estimate $\hat{S}_1$ of S_1 as follows:

$$\hat{S}_1 = \begin{cases} s, & \text{if there exists a state } s \text{ such that} \\ & |N(x,y)/c - p(y|x,s)| < d_{\min}/2, \\ & \text{for all } (x,y) \in \mathcal{X} \times \mathcal{Y}; \\ \hat{s}, & \text{otherwise, where } \hat{s} \text{ is some fixed state in } \mathcal{S}. \end{cases}$$

From the definition of $d_{\min}$, it follows that this estimate is well defined, i.e., there is at most one state s for which $|N(x,y)/c - p(y|x,s)| < d_{\min}/2$ holds for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$. Let $\epsilon > 0$. From the law of large numbers, we have for sufficiently large t,

$$\Pr\{|N(x,y)/c - p(y|x,s)| < d_{\min}/2\} > 1 - \epsilon$$

for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$ if the block of length l that contains the test packet has the same state s. Hence, if t is sufficiently large,

$$\Pr\{\hat{S}_1 \neq S_1 | U_l = 0\} < \epsilon.$$

The lemma now follows from Fano's inequality [13; chapter 1]. ∎

Now, we consider the sequence of bursty channels $\Gamma_1, \Gamma_2, \ldots$ defined in section 3, where $\mathcal{S}$ is finite. This sequence satisfies conditions (3) and (4). The asymptotic value of $C(n)$ was derived in Theorem 6. Our main objective in this section is to prove that $C_{SI}(n)$, the capacity of Γ_n using smart interleaving, tends to the same limit as that of $C(n)$. We need a lemma before proving this result. The same notation will be used in the following after adding the subscript n to denote channel Γ_n.

Lemma 10. *For every $\delta > 0$ there exists $t_0(\delta)$ such that for every $t > t_0(\delta)$, $H(S_{1,n}|\hat{S}_{1,n}) < \delta$ for all sufficiently large n.*

Proof. For all $s, s' \in S$, we have

$$\Pr\{S_{1,n} = s|\hat{S}_{1,n} = s'\} = \Pr\{S_{1,n} = s|\hat{S}_{1,n} = s', U_{l,n} = 0\}\Pr\{U_{l,n} = 0\}$$
$$+ \Pr\{S_{1,n} = s|\hat{S}_{1,n} = s', U_{l,n} = 1\}\Pr\{U_{l,n} = 1\}$$

From condition (3) of section 3, $\Pr\{U_{l,n} = 0\} \to 1$ as $n \to \infty$. Hence,

$$\Pr\{S_{1,n} = s|\hat{S}_{1,n} = s'\} \to \Pr\{S_{1,n} = s|\hat{S}_{1,n} = s', U_{l,n} = 0\}.$$

Similarly,

$$\Pr\{S_{1,n} = s, \hat{S}_{1,n} = s'\} \to \Pr\{S_{1,n} = s, \hat{S}_{1,n} = s'|U_{l,n} = 0\},$$

and

$$\Pr\{S_{1,n} = s\} \to \Pr\{S_{1,n} = s|U_{l,n} = 0\}.$$

From the continuity of the entropy function it follows that

$$H(S_{1,n}|\hat{S}_{1,n}) \to H(S_{1,n}|\hat{S}_{1,n}, U_{l,n} = 0)$$

as $n \to \infty$. The lemma now follows from Lemma 9. ∎

The next theorem is the fundamental result of this section.

Theorem 11.

$$\lim_{n\to\infty} C_{SI}(n) = \lim_{n\to\infty} C(n) = \lim_{n\to\infty} C_d(n).$$

Proof. The last equality has been proved already in section 3. So, it suffices to show that $C_d(n) - C_{SI}(n) \to \infty$ as $n \to \infty$. By setting $l = 1$ in Theorem 2, it follows from Lemmas 7 and 8 that

$$C_d(n) - C_{SI}(n) \le \frac{t}{l} C_d(n) + (1 - \frac{t}{l}) H(S_{1,n}|\hat{S}_{1,n})$$
$$\le \frac{t}{l} \log|\mathcal{X}| + (1 - \frac{t}{l}) H(S_{1,n}|\hat{S}_{1,n}).$$

Suppose that $\delta > 0$ is given. Choose $t > t_0(\delta/2)$ of Lemma 10. Choose $l > t$ such that $t \log|\mathcal{X}|/l < \delta/2$. Then, from Lemma 10 it follows that $C_d(n) - C_{SI}(n) < \delta$

for all sufficiently large n. ∎

This theorem implies that smart interleaving is asymptotically optimum.

5. Applications to Computer Memories

In this section we apply the results of the previous sections to computer memories affected by alpha-particles. First, we consider the "one-dimensional chip", i.e., a chip whose width is one cell. Then, we consider "two-dimensional chips" and argue that every result obtained so far has a two-dimensional version.

In the following, we assume that the only source of errors is alpha-particles that hit memory cells. An error takes place when an alpha-particle hits the chip during an encoding/decoding cycle, and causes some cells to change their contents. This encoding/decoding cycle will be considered as a unit of time.

5.1. One-Dimensional Chips

Consider a chip whose width is 1 cell and whose length is infinite. The number of cells per unit length is n, where n is some positive integer. The cells on the chip are numbered $1, 2, \ldots$, consecutively. There are q types of alpha-particles, $\alpha_1, \alpha_2, \ldots, \alpha_q$. For each j, the number of alpha-particles of type α_j that fall on the ith cell per unit time is assumed to be a Poisson random variable $N_{i,j}$ of mean λ_j/n, for some positive number λ_j. We assume that the Poisson random variables $N_{i,j}$ for $i = 1, 2, \ldots$, and $j = 1, 2, \ldots, q$, are independent. It follows that the number of alpha-particles of type α_j falling on a unit length per unit time is a Poisson random variable with mean λ_j, which is independent of n.

Let $0 < w_1 < w_2 < \cdots < w_q$ be positive numbers, and define the burst length of α_j, where $1 \leq j \leq q$, as $b_j = \lfloor w_j n \rfloor$. Here, w_j is the effective range in which electron-hole pairs are generated due to an alpha-particle of type α_j. We assume that a cell is affected by an alpha-particle if, and only if, it lies completely

within its effective range. Hence, the burst length b_j gives the number of cells that are affected by an alpha-particle of type α_j. Note that alpha-particles of type α_q are considered to be the "strongest" in the sense that they affect more cells. This agrees with the physical evidence explained in chapter I that the effect of an alpha-particle depends on its energy and angle of incidence which vary considerably from particle to particle.

The state of cell i is given by*

$$S_i = \sum_{j=1}^{q} \sum_{k=0}^{b_j-1} N_{i-k,j},$$

which is the number of alpha-particles that affect cell i. Hence, the set of states S is the set of nonnegative integers. Furthermore, S_i is a Poisson random variable with mean $\mu_n = \lambda_1 b_1/n + \cdots + \lambda_q b_q/n$. Clearly, the stochastic process $\mathbf{S} = S_1, S_2, \ldots$, satisfies conditions (1) and (2) of section 2, where the duration of state memory is $b_q - 1$. Hence, a bursty channel Γ_n can be defined. The input and output alphabets are $\{0, 1\}$. The effect of a single alpha-particle is modelled by transmitting the letters with state 1 through a binary memoryless channel Ω. The component channel defined by state $s > 0$ is considered as a cascade of s of the channels Ω. The component channel defined by state $s = 0$ is considered to be the binary noiseless memoryless channel.

Now, let $n = 1, 2, \ldots$, and define for each n the bursty channel Γ_n. The numbers $\lambda_1, \ldots, \lambda_q$ and $w_1, \ldots, w_q$ as well as the channel Ω are independent of n. Let $S_{i,n}$ be the state variable of cell i in channel Γ_n. Let $U_{l,n}$ be the l-run random variable of Γ_n, where l is a positive integer. We will argue that conditions (3) and (4) of section 3 are satisfied. Indeed $U_{l,n} = 1$ if, and only if, the states in the block of cells $1, 2, \ldots, l$ are not equal. Hence, $U_{l,n} = 1$ implies that at least one alpha-particle of type α_j, for some $1 \leq j \leq q$, has fallen on the

* *We will ignore the "edge effect", i.e., the case in which $i < b_q$. Of course this has no effect on the channel capacity as b_q is finite. So, in the following we assume that there are $b_q - 1$ cells numbered $2 - b_q, 3 - b_q, \ldots, 0$.*

block of cells $1, 2, \ldots, l$, or the block of cells $2 - b_j, 3 - b_j, \ldots, l + 1 - b_j$ in a unit time. Thus,

$$\Pr\{U_{l,n} = 1\} \leq 2 \sum_{j=1}^{q} [1 - e^{\lambda_j l/n}],$$

as $e^{\lambda_j l/n}$ is the probability that no alpha-particle of type α_j has fallen on a block of l cells in a unit time. Clearly, $U_{l,n} \to 0$ as $n \to \infty$ in probability, and condition (3) is satisfied. Since, as mentioned before, $S_{i,n}$ is a Poisson random variable with mean $\mu_n = \lambda_1 b_1/n + \cdots + \lambda_q b_q/n$, it follows that as $n \to \infty$, $S_{i,n}$ tends to a Poisson random variable with mean $\mu = \lambda_1 w_1 + \cdots + \lambda_q w_q$. Hence, condition (4) is also satisfied. From Theorems 5 and 6, we have

$$\lim_{n \to \infty} C(n) = \lim_{n \to \infty} C_d(n) = \max_{p(x)} \sum_{s=0}^{\infty} I(X; Y | S = s) \mu^s e^{-\mu}/s!,$$

where X and Y are the input and output random 1-sequences, respectively. On the other hand, from Theorem 5,

$$\lim_{n \to \infty} C_{ed}(n) = e^{-\mu} \sum_{s=0}^{\infty} C_s \mu^s/s!,$$

where C_s is the capacity of a cascade of s of the channels Ω. If Ω is a binary symmetric channel, or a channel with zero capacity, then the component channels are compatible as defined in section 2, and the limits of $C(n)$, $C_d(n)$, and $C_{ed}(n)$ become equal because of Corollary 3.

If Ω has zero capacity, then the set of component channels is composed of two channels only. In this case $C_0 = 1$ and $C_r = 0$ for $r > 0$. Since the two component channels are compatible, $C(n) \to C_{ed}(n) \to e^{-\mu}$ as $n \to \infty$. The set of states can be reduced to two states only, and smart interleaving can be used to achieve the asymptotic value of capacity.

As explained in chapter I, the output of a cell which is affected by an alpha-particle is 0, whether the input is 0 or 1. Hence, Ω is the binary asymmetric channel whose capacity is 0. In such case, $e^{-\mu}$ is the ultimate limit of the fraction of useful area that can be used to store data, as the memory cells continue to

shrink, if alpha-particles are the only source of errors. As a simple numerical example, consider the case in which we have one type of alpha-particles, i.e., $q = 1$. Let $w_1 = 1$ and $\lambda_1 = 0.1$. Hence $b_1 = n$ and $\mu_n = 0.1$ for all n. In case $n = 1$, the states of the cells are independent, and hence the capacity $C(1)$ can be calculated by using (1). We find that $C(1) \approx 0.77$, which is the fraction of the useful area when the cells are large so that an alpha-particle may affect only a single cell. Now, by decreasing the dimensions of the cells so that a single alpha-particle may affect n cells, the capacity $C(n) \to e^{-0.1} \approx 0.9$ as $n \to \infty$, i.e., 90% of the area becomes useful in storing data. This means that if the cells are affected by alpha-particles because they are too small, then we should make them even smaller and use coding!

5.2. Two-Dimensional Chips

The results of the previous sections all have two-dimensional versions. Let S be a finite or a countable set. Suppose $\mathbf{S} = (S_{i,j})$, where $S_{i,j} \in S$ for $i, j \geq 1$, is a two-dimensional stationary stochastic process such that there exists a least nonnegative integer m for which any two finite two-dimensional arrays in $\mathbf{S}$ separated by at least m states in either direction are independently distributed. These two conditions are the two-dimensional version of conditions (1) and (2) of section 1. The process $\mathbf{S}$ is called the two-dimensional stochastic process of the bursty channel, which can be defined exactly as in section 1. In this case, the input and output are two-dimensional arrays. Clearly, all the results obtained so far can be written in a two-dimensional form. For example, the capacity C of Theorem 4 becomes

$$C = \sup_l \max_{p(\mathbf{x}_l)} \frac{I(\mathbf{X}_l; \mathbf{Y}_l)}{l^2},$$

where $\mathbf{X}_l$ and $\mathbf{Y}_l$ are the input and output random $l \times l$-arrays, respectively.

We define a sequence of bursty channel $\Gamma_1, \Gamma_2, \ldots$, and for channel Γ_n, we define the $l \times l$-run binary random variable $U_{l,n}$ which is equal to 0 if the states $S_{i,j}$ are the same in the $l \times l$- block $1 \leq i, j \leq l$. We assume that $U_{l,n}$ tends to

0 in probability as $n \to \infty$. We also assume that condition (4) of section 3 is satisfied. The technique of smart interleaving can also be used if the state set is finite. In such case, the test packet is a $t \times t$ block for some positive integer t. The codeword is divided into $l \times l$ blocks for some $l > t$, and each block contains the test packet. The decoder estimates the states in every $l \times l$-block from the output of the test packet contained in the block. Two-dimensional chips affected by alpha-particles can be modelled as two-dimensional bursty channels, exactly as in the one-dimensional case.

References for Part One

[1] R. T. Bate, "Quantum-mechanical limitations on device performance," in *VLSI Electronics*, N. G. Einspruch, Ed., vol. 5, New York: Academic Press, 1982.

[2] K. L. Chung, *A Course in Probability Theory*, New York: Harcourt, Brace & World, 1968.

[3] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 256–268, October 1974.

[4] J. L. Doob, *Stochastic Processes*, New York: Wiley, 1960.

[5] W. Feller *An Introduction to Probability Theory and Its Applications*, vol. 1, New York: Wiley, 1968.

[6] T. S. Ferguson, *Mathematical Statistics*, New York: Academic Press, 1967.

[7] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

[8] C. Heegard and A. A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 731–739, September 1983.

[9] R. W. Keyes, "Physical limits in digital electronics," *Proc. IEEE*, vol. 63, pp. 740–767, May 1975.

[10] M. Loève, *Probability Theory*, Princeton, N.J.: Van Nostrand, 1963.

[11] T. C. May, "Soft errors in VLSI: present and future," *IEEE Trans. Compo-*

nents, Hybrids, and Manufacturing Technology, vol. CHMT-2, pp. 377–387, December 1979.

[12] T. C. May and M. H. Woods, "Alpha-particle-induced soft errors in dynamic memories," *IEEE Trans. Electron Devices*, vol. ED-26, pp. 2–9, January 1979.

[13] R. J. McEliece, *The Theory of Information and Coding*, Reading, MA: Addison-Wesley, 1977.

[14] R. J. McEliece and W. E. Stark, "Channels with block interference ," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 44–53, January 1984.

[15] C. Mead and L. Conway, *Introduction to VLSI Systems*, Reading, MA: Addison- Wesley, 1980.

[16] S. Muroga, *VLSI System Design*, New York: Wiley, 1982.

[17] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, July 1948.

[18] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Research and Development*, vol. 2, pp. 289–293, October 1958.

[19] K. Stein, "Noise-induced error rate as limiting factor for energy per operation in digital IC's," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 527–530, October 1977.

[20] J. A. Swanson, "Physical versus logical coupling in memory systems," *IBM J. Research and Development*, vol. 4, pp. 305–310, July 1960.

[21] J. Wolfowitz, *Coding Theorems of Information Theory*, Berlin: Springer-Verlag, 1978.

[22] A. Wyner, "A definition of conditional mutual information for arbitrary ensembles," *Inform. Contr.*, vol. 38, pp. 51–59, 1978.

[23] D. S. Yaney, J. T. Nelson, and L. L. Vanskika, "Alpha-particle tracks in silicon and their effect on dynamic MOS RAM reliability," *IEEE Trans. Electron Devices*, vol. ED-26, pp. 10–16, January 1979.

[24] J. F. Zieger and W. A. Lanford, "Effect of cosmic rays on computer memories," *Science*, vol. 206, pp. 776–788, 16 November 1979.

PART TWO

TWO-DIMENSIONAL BURST CORRECTING CODES

PRELIMINARIES

A *binary two-dimensional code* of area $n_1 \times n_2$,* where n_1 and n_2 are positive integers, is a set of $n_1 \times n_2$ binary arrays, whose elements are called *codewords.* The rows and columns of an $n_1 \times n_2$ array will be numbered $0, 1, \ldots, n_1 - 1$ and $0, 1, \ldots, n_2 - 1$, respectively. A *binary linear two-dimensional code* $\mathcal{C}$ of area $n_1 \times n_2$ is a subspace of the $n_1 n_2$-dimensional space of $n_1 \times n_2$ arrays over $\mathbf{F}_2$. Let k be the dimension of $\mathcal{C}$, and $[g_{i,j}^{(1)}], \ldots, [g_{i,j}^{(k)}]$, where $0 \le i < n_1, 0 \le j < n_2$, be a basis for $\mathcal{C}$. The $n_1 \times n_2$ matrix $G = [\mathbf{g}_{i,j}]$, where $\mathbf{g}_{i,j} = (g_{i,j}^{(1)}, \ldots, g_{i,j}^{(k)})$ is called a *generator matrix* of $\mathcal{C}$. The *dual code* of $\mathcal{C}$, denoted by $\mathcal{C}^{\perp}$, is the null space of $\mathcal{C}$. If the $n_1 \times n_2$ matrix $H = [\mathbf{h}_{i,j}]$ is a generator for $\mathcal{C}^{\perp}$, then H is called a *parity check matrix* of the code $\mathcal{C}$. The elements of H are elements in the r-dimensional vector space over $\mathbf{F}_2$, where $r = n_1 n_2 - k$ is called the *redundancy* of the code. The *syndrome* of a binary array $[a_{i,j}]$ of area $n_1 \times n_2$, with respect to the parity check matrix H of $\mathcal{C}$, is defined as $\sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} a_{i,j} \mathbf{h}_{i,j}$. Thus, a binary array of area $n_1 \times n_2$ is a codeword in $\mathcal{C}$ if, and only if, its syndrome is zero with respect to any given parity check martix.

The map $[a_{i,j}] \mapsto a(x,y) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} a_{i,j} x^i y^j$ defines an isomorphism between the $n_1 n_2$-dimensional vector space of $n_1 \times n_2$ arrays over $\mathbf{F}_2$ and the vector space of bivariate polynomials $\{p(x,y) \in \mathbf{F}_2[x,y] : \deg_x p(x,y) < n_1, \deg_y p(x,y) < n_2\}$. We will frequently identify each array with its image under this isomorphism.

A binary two-dimensional linear code $\mathcal{C}$ is said to be *cyclic* if $xc(x,y)$ and $yc(x,y)$, both mod $(x^{n_1}+1, y^{n_2}+1)$, are in $\mathcal{C}$ for each $c(x,y) \in \mathcal{C}$. Thus, a cyclic code of area $n_1 \times n_2$ is an ideal in the residue class ring $\mathbf{F}_2[x,y]/(x^{n_1}+1, y^{n_2}+1)$.

---

* *In the following, $n_1 \times n_2$ does not mean the product $n_1 n_2$, but rather the pair of integers (n_1, n_2). An array of n_1 rows and n_2 columns is said to have area $n_1 \times n_2$.*

The pairs of positive integers will be partially ordered by saying that $b_1 \times b_2$ is *less* than $n_1 \times n_2$ if $b_1 \leq n_1$, $b_2 \leq n_2$, and $b_1 \times b_2 \neq n_1 \times n_2$. A $b_1 \times b_2$-*burst*, where $b_1 \times b_2$ is less or equal to $n_1 \times n_2$, is a nonzero $n_1 \times n_2$ binary array whose nonzero components are confined to a rectangle of area $b_1 \times b_2$. Let $\{(i,j) : u_1 \leq i < u_1 + b_1', u_2 \leq j < u_2 + b_2'\}$, where $0 \leq u_1 \leq n_1 - b_1'$, $0 \leq u_2 \leq n_2 - b_2'$, be the smallest rectangle containing the nonzero components of the $b_1 \times b_2$-burst $B = [a_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$. Then, B is said to have area $b_1' \times b_2'$. For $b_1' \leq b_1'' \leq b_1$ and $b_2' \leq b_2'' \leq b_2$, we say that B has the *pattern* $[a_{i,j}]$, $u_1 \leq i < u_1 + b_1''$, $u_2 \leq j < u_2 + b_2''$, starting at *position* (u_1, u_2). In the following, it is more convenient to speak about "the" pattern of B by considering $[a_{i,j}]$, $u_1 \leq i < u_1 + b_1''$, $u_2 \leq j < u_2 + b_2''$, to represent the same pattern for all $b_1' \leq b_1'' \leq b_1$ and $b_2' \leq b_2'' \leq b_2$. By this convention, it is to be noted that the pattern and the starting position of any burst are unique. Thus, $[a_{i,j}]$ is a $b_1 \times b_2$-burst if, and only if, $a(x,y) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} a_{i,j} x^i y^j = x^{u_1} y^{u_2} b(x,y)$, for some $0 \leq u_1 \leq n_1 - b_1, 0 \leq u_2 \leq n_2 - b_2$, and $b(x,y) \in \mathcal{B}_{b_1 b_2}$, where

$$\mathcal{B}_{b_1 b_2} = \{p(x,y) \in \mathbf{F}_2[x,y] : \deg_x p(x,y) < b_1, \deg_y p(x,y) < b_2, \\ p(x,0) \neq 0, p(0,y) \neq 0\}.$$

In such case the burst pattern is given by the polynomial $b(x,y)$ and its starting position is (u_1, u_2). The array $[a_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, is called a $b_1 \times b_2$-*cyclic burst* if $a(x,y) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} a_{i,j} x^i y^j \equiv x^{u_1} y^{u_2} b(x,y) \pmod{x^{n_1}+1, y^{n_2}+1}$, for some $0 \leq u_1 < n_1, 0 \leq u_2 < n_2$, and $b(x,y) \in \mathcal{B}_{b_1 b_2}$. Thus, a $b_1 \times b_2$-burst is a $b_1 \times b_2$-cyclic burst, but the converse does not always hold. The starting position (u_1, u_2) of the cyclic burst and its pattern, which is given by $b(x,y)$, are not necessarily unique. This will be considered in the following lemma.

Lemma 1. *A necessary and sufficient condition for all $b_1 \times b_2$-cyclic bursts $[a_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, to have unique patterns and starting positions is $n_1 \geq 2b_1 - 1$ and $n_2 \geq 2b_2 - 1$.*

Proof. If $n_1 < 2b_1 - 1$, then the burst $[a_{i,j}]$, defined by $a_{i,j} = 1$ if, and only if, $(i,j) \in \{(0,0),(b_1-1,0)\}$ has starting positions $(0,0)$ and $(b_1-1,0)$ since

$$1 + x^{b_1-1} \equiv x^{b_1-1}(1 + x^{n_1-b_1+1}) \pmod{x^{n_1}+1, y^{n_2}+1},$$

and $\deg_x(1 + x^{n_1-b_1+1}) = n_1 - b_1 + 1 < b_1$. Thus $n_1 \geq 2b_1 - 1$ is a necessary condition. By similarity, $n_2 \geq 2b_2 - 1$ is also necessary. Conversely, suppose $n_1 \geq 2b_1 - 1$, $n_2 \geq 2b_2 - 1$, and

$$x^{u_1}y^{u_2}b'(x,y) \equiv x^{v_1}y^{v_2}b''(x,y) \pmod{x^{n_1}+1, y^{n_2}+1},$$

where $0 \leq u_1, v_1 < n_1$, $0 \leq u_2, v_2 < n_2$, $b'(x,y), b''(x,y) \in \mathcal{B}_{b_1 b_2}$. If $u_1 > v_1$, then

$$x^{n_1-v_1}y^{u_2}b'(x,y) \equiv x^{n_1-u_1}y^{v_2}b''(x,y) \pmod{x^{n_1}+1, y^{n_2}+1},$$

with $n_1 - v_1 < n_1 - u_1$. Hence we may assume, without loss of generality, that $u_1 \leq v_1$ and $u_2 \leq v_2$. Thus, we have

$$b'(x,y) + x^{v_1-u_1}y^{v_2-u_2}b''(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1},$$

Let $t_1 = \min\{v_1 - u_1, n_1 - v_1 + u_1\}$ and $t_2 = \min\{v_2 - u_2, n_2 - v_2 + u_2\}$. Then, $t_1 \leq n_1/2$ and $t_2 \leq n_2/2$. Thus, we have

$$f'(x,y) + x^{t_1}y^{t_2}f''(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1},$$

or

$$x^{t_1}f'(x,y) + y^{t_2}f''(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1},$$

where either $f'(x,y) = b'(x,y)$ and $f''(x,y) = b''(x,y)$, or $f'(x,y) = b''(x,y)$ and $f''(x,y) = b'(x,y)$. But both the polynomials $f'(x,y) + x^{t_1}y^{t_2}f''(x,y)$ and $x^{t_1}f'(x,y) + y^{t_2}f''(x,y)$ have degrees in x and y less than n_1 and n_2, respectively, as $n_1 \geq 2b_1 - 1$ and $n_2 \geq 2b_2 - 1$. Hence at least one of these polynomials is zero. Since $f'(x,0)$, $f'(0,y)$, $f''(x,0)$ and $f''(0,y)$ are nonzero, it follows that $t_1 = t_2 = 0$ and $f'(x,y) = f''(x,y)$. Thus, $u_1 = v_1$, $u_2 = v_2$, and $b'(x,y) = b''(x,y)$. ∎

The following lemma gives the number of distinct $b_1 \times b_2$-burst patterns, which we denote by $N(b_1, b_2)$.

Lemma 2. *The number of distinct $b_1 \times b_2$-burst patterns $N(b_1, b_2)$ is given by*

$$N(b_1, b_2) = 2^{b_1 b_2 - 1} + (2^{b_1 - 1} - 1)(2^{b_2 - 1} - 1) \times 2^{(b_1 - 1)(b_2 - 1)}.$$

Hence,

$$2^{b_1 b_2 - 1} \leq N(b_1, b_2) < 2^{b_1 b_2},$$

and the equality holds if, and only if, b_1 or b_2 is 1.

Proof. From the definitions, it follows that $N(b_1, b_2)$ is the total number of binary $b_1 \times b_2$-arrays with the property that their first row and column are nonzero. The first and second terms give the number of arrays satisfying this property with "1" and "0" at position (0,0), respectively. ∎

In this work, we will consider binary linear codes only. If $n_1 = 1$, we say that the code is a *one-dimensional code* of length n_2. In such case, the first dimension will be suppressed. Hence, from Lemma 2, it follows that the number of distinct b-burst patterns is given $N(b) = 2^{b-1}$.

CHAPTER III

BURST IDENTIFICATION CODES

In this chapter, we study burst identification codes which are defined in section 1. One-dimensional and two-dimensional burst identification codes are considered in sections 2 and 3, respectively. In section 4 we examine certain classes of two-dimensional burst identification codes.

1. Definitions

A two-dimensional linear code C is said to be a $b_1 \times b_2$-*burst identification code* if no codeword is a $b_1 \times b_2$-burst, or a sum of two $b_1 \times b_2$-bursts of different patterns. Equivalently, the code C is a $b_1 \times b_2$-burst identification code if, and only if, the syndromes of the $b_1 \times b_2$-bursts with respect to any given parity check matrix of C are nonzero and distinct for distinct burst patterns.

If a $b_1 \times b_2$-burst identification code is used over a channel that may add to any transmitted codeword a $b_1 \times b_2$-burst, then the receiver can determine the burst pattern added by the channel. It is important to note that the receiver may not be able to uniquely determine the burst position. Hence the transmitted codeword may not be uniquely determined. Thus, a $b_1 \times b_2$-burst correcting code, which is defined in the next chapter, is a $b_1 \times b_2$-burst identification code, but the converse does not always hold. In other words, a $b_1 \times b_2$-burst identification code may contain a codeword which is the sum of two $b_1 \times b_2$-bursts of the same pattern.

We define $r(b_1, b_2)$ to be the minimum redundancy required to construct a $b_1 \times b_2$-burst identification code of arbitrarily large area. Thus, if $r_{n \times n}$, for every positive integer n, denotes the redundancy, minimized over all $b_1 \times b_2$-burst

identification codes of area $n \times n$, or larger, then

$$r(b_1, b_2) = \lim_{n \to \infty} r_{n \times n}.$$

The validity of this definition, i.e., the existence of the limit will be shown later. It is to be noted that $r_{n \times n}$ is a nondecreasing function of n, and hence, it suffices to show that it is bounded.

Burst identification codes will be used in the next chapter to construct burst correcting codes. In the present chapter, we start studying burst identifcation codes by considering the one-dimensional case.

2. One-Dimensional Burst Identification Codes

A one-dimensional code is b-burst identification code if, and only if, the syndromes of the b-bursts are nonzero and distinct for distinct burst patterns. Since the number of different burst patterns is $N(b) = 2^{b-1}$, it follows that the minimum redundancy required to construct a b-burst identification code of arbitrarily large length is bounded by $r(b) \geq \lceil \log(1 + 2^{b-1}) \rceil$, which implies $r(b) \geq b$. It is obvious that $r(1) = 1$, which is achieved by a code whose parity check matrix is $H = [1, 1, \ldots, 1]$. The bound $r(b) \geq b$, also follows from the following lemma which is an immediate consequence of the definition.

Lemma 1. *Let H be a parity check matrix of a b-burst identification code. Then, every b consecutive elements of H are linearly independent over $\mathbf{F}_2$.*

The next lemma shows that a stronger bound holds for $b \geq 3$.

Lemma 2. *Let H be a parity check matrix of a b-burst identification code of length $n \geq 2b - 2$. Then, every $2b - 2$ consecutive elements of H are linearly independent over $\mathbf{F}_2$.*

Proof. If $b \leq 2$, the result follows from Lemma 1. So, we assume in the following that $b \geq 3$. Let $H = [\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_{n-1}]$ and suppose that $\sum_{i=0}^{u-1} a_i \mathbf{h}_{l+i} = \mathbf{0}$, where

$a_i \in \mathbf{F}_2$, $a_0 = a_{u-1} = 1$, $1 \leq u \leq 2b-2$, and $0 \leq l \leq n-u$. Lemma 1 implies that $u \geq b+1$. The burst pattern $(1, a_1+1, a_b, a_{b+1}, \ldots, a_{u-1})$ starting at position $l+b-2$ has syndrome $\mathbf{h}_{l+b-2}+(a_1+1)\mathbf{h}_{l+b-1}+\sum_{i=b}^{u-1} a_i\mathbf{h}_{l+i} = \sum_{i=0}^{b-3} a_i\mathbf{h}_{l+i}+(a_{b-2}+1)\mathbf{h}_{l+b-2}+(a_1+a_{b-1}+1)\mathbf{h}_{l+b-1}$. However, the burst pattern $(1, a_1, \ldots, a_{b-3}, a_{b-2}+1, a_1+a_{b-1}+1)$ starting at position l has the same syndrome, but obviously a different pattern. This contradicts the assumption that the code is a b-burst identification code. ∎

It follows from this lemma that $r(b) \geq 2b-2$. The following theorem gives a construction of b-burst identification codes, for $b \geq 2$, of arbitrarily large lengths and with redundancy $2b-2$. First, we define $\mathbf{e}_i = (e_{i,0}, e_{i,1}, \ldots, e_{i,2b-3})$, where $0 \leq i < 2b-2$, $e_{i,i} = 1$, and $e_{i,j} = 0$ for $i \neq j$, to be the ith canonical vector of length $2b-2$.

Theorem 3. *Let $b \geq 2$ and $\mathbf{e}_i$ be the ith canonical vector of length $2b-2$. Then, the code C of length n whose parity check matrix is given by $H = [\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_{n-1}]$, where $\mathbf{h}_i = \mathbf{e}_{i \bmod (2b-2)}$, is a b-burst identification code with redundancy $2b-2$.*

Proof. Consider a burst of length $b' \leq b$. Let $(a_0 = 1, a_1, \ldots, a_{b'-1} = 1)$ be its pattern and l its position, where $0 \leq l \leq n-b'$. The syndrome of this burst is

$$\mathbf{s} = \sum_{i=0}^{b'-1} a_i\mathbf{h}_{l+i} = \sum_{i=0}^{b'-1} a_i\mathbf{e}_{l+i \bmod (2b-2)}.$$

The vectors $\mathbf{e}_{l+i \bmod (2b-2)}$ for $0 \leq i < b'$ are distinct since $b' \leq b \leq 2b-2$. Hence, the weight of the vector $\mathbf{s}$, i.e., the number of its nonzero components, is equal to the weight of the burst pattern. This ends the proof for $b = 2$ since the burst patterns are either (1) or (1,1). Now, let $b \geq 3$. Then, $\mathbf{s} = (s_0, s_1, \ldots, s_{2b-3})$ is a cyclic shift of the $(2b-2)$-tuple $(a_0, a_1, \ldots, a_{b'-1}, 0, \ldots, 0)$. Hence, if $\mathbf{s}$ has a unique cyclic string of consecutive zeros of length $\geq b-2$, then the burst pattern $(a_0, a_1, \ldots, a_{b'-1})$ can be uniquely deduced from $\mathbf{s}$. If this is not the case, then $\mathbf{s}$ has two cyclic strings of $b-2$ zeros each, which occurs if, and only if, $b' = b$ and

the burst pattern is $(a_0 = 1, 0, 0, \ldots, 0, a_{b-1} = 1)$. Also in this case the burst pattern is uniquely determined from the syndrome. ∎

Combining Lemma 2 and Theorem 3, along with the fact that $r(1) = 1$, the following theorem follows.

Theorem 4. *$r(1) = 1$, and $r(b) = 2b - 2$ for $b \geq 2$.*

3. Two-Dimensional Burst Identification Codes

A two-dimensional code is a $b_1 \times b_2$-burst identification code if, and only if, the syndromes of the $b_1 \times b_2$-bursts are nonzero and distinct for distinct burst patterns. Since the number of different patterns is $N(b_1, b_2)$, as given in Lemma 2 in the preliminaries, it follows that the minimum redundancy required to construct a $b_1 \times b_2$-burst identification code of arbitrarily large area is bounded by $r(b_1, b_2) \geq \lceil \log(1 + N(b_1, b_2)) \rceil$, which implies $r(b_1, b_2) \geq b_1 b_2$. In this section, we will prove that $2b_1 b_2 - 2 \leq r(b_1, b_2) \leq 2b_1 b_2$. The following two lemmas are the two-dimensional versions of Lemmas 1 and 2.

Lemma 5. *Let $H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1, 0 \leq j < n_2$, be a parity check matrix of a $b_1 \times b_2$-burst identification code. Then, for every pair of integers (u_1, u_2), such that $0 \leq u_1 \leq n_1 - b_1$, $0 \leq u_2 \leq n_2 - b_2$, the vectors $\mathbf{h}_{i,j}$, for $u_1 \leq i < u_1 + b_1$, $u_2 \leq j < u_2 + b_2$ are linearly independent over $\mathbf{F}_2$.*

The proof of the previous lemma follows immediately from the definition.

Lemma 6. *Let $H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1, 0 \leq j < n_2$ be a parity check matrix of a $b_1 \times b_2$-burst identification code. Let $0 \leq u_1, v_1 \leq n_1 - b_1$, $0 \leq u_2, v_2 \leq n_2 - b_2$, $I_{u_1,u_2} = \{(i,j) : u_1 \leq i < u_1 + b_1,\ u_2 \leq j < u_2 + b_2\}$, and define I_{v_1,v_2} similarly. If $|I_{u_1,u_2} \bigcap I_{v_1,v_2}| \geq 2$, then the vectors $h_{i,j}$, $(i,j) \in I_{u_1,u_2} \bigcup I_{v_1,v_2}$ are linearly independent over $\mathbf{F}_2$.*

Proof. Without loss of generality, assume that $u_1 \leq v_1$ and $u_2 \leq v_2$. Let $J = I_{u_1,u_2} \bigcap I_{v_1,v_2} = \{(i,j) : v_1 \leq i < u_1 + b_1, v_2 \leq j < u_2 + b_2\}$, and suppose

$|J| \geq 2$. We may assume that $u_1 + b_1 - v_1 \geq 2$, otherwise interchange i and j. Now, suppose that $\sum_{(i,j)\in I_{u_1,u_2} \bigcup I_{v_1,v_2}} a_{i,j}\mathbf{h}_{i,j} = \mathbf{0}$, $a_{i,j} \in \mathbf{F}_2$, and not all are zero.

Let $c_{0,0} = 1$, x be an indeterminate, and $c_{i,j} = a_{v_1+i,v_2+j}$ for $(i,j) \in \{(i',j') : 0 \leq i' < b_1, 0 \leq j' < b_2\} - \{(0,0),(1,0)\}$. The array $[c_{i,j}]$, where $0 \leq i \leq b_1, 0 \leq j < b_2$ and $c_{1,0} = x$, defines the pattern of a $b_1 \times b_2$-burst starting at position (v_1, v_2). Let B_1 be this burst. The syndrome of B_1 is

$$\begin{aligned}\sum_{i=0}^{b_1-1}\sum_{j=0}^{b_2-1} c_{i,j}\mathbf{h}_{v_1+i,v_2+j} &= \mathbf{h}_{v_1,v_2} + x\mathbf{h}_{v_1+1,v_2} + \sum_{(i,j)\in I_{v_1,v_2}-\{(v_1,v_2),(v_1+1,v_2)\}} a_{i,j}\mathbf{h}_{i,j} \\ &= (1 + a_{v_1,v_2})\mathbf{h}_{v_1,v_2} + (x + a_{v_1+1,v_2})\mathbf{h}_{v_1+1,v_2} + \sum_{(i,j)\in I_{v_1,v_2}} a_{i,j}\mathbf{h}_{i,j} \\ &= \sum_{(i,j)\in I_{u_1,u_2}-J} a_{i,j}\mathbf{h}_{i,j} + (1 + a_{v_1,v_2})\mathbf{h}_{v_1,v_2} + (x + a_{v_1+1,v_2})\mathbf{h}_{v_1+1,v_2}.\end{aligned}$$

Define the array $[d_{i,j}]$, $0 \leq i < b_1$, $0 \leq j < b_2$ as

$$d_{i,j} = \begin{cases} a_{u_1+i,u_2+j}, & \text{if } (u_1+i, u_2+j) \in I_{u_1,u_2} - J; \\ 1 + a_{v_1,v_2}, & \text{if } (u_1+i, u_2+j) = (v_1, v_2); \\ x + a_{v_1+1,v_2}, & \text{if } (u_1+i, u_2+j) = (v_1+1, v_2); \\ 0, & \text{otherwise.} \end{cases}$$

Let B_2 be the burst whose pattern is defined by the array $[d_{i,j}]$, $0 \leq i < b_1$, $0 \leq j < b_2$ and whose starting position is (u_1', u_2'), where $u_1' = u_1 + t_1, u_2' = u_2 + t_2$, and t_1, t_2 are the maximum values for which the rectangle $\{(i,j) : t_1 \leq i < b_1,\ t_2 \leq j < b_2\}$ contains all the nonzero components of the array $[d_{i,j}]$ $0 \leq i < b_1$, $0 \leq j < b_2$. From Lemma 5, it follows that $(u_1', u_2') \in I_{u_1,u_2} - J$. The syndrome of B_2, which is $\sum_{i=0}^{b_1-1}\sum_{j=0}^{b_2-1} d_{i,j}\mathbf{h}_{u_1+i,u_2+j}$, is the same as the syndrome of B_1. Hence B_1 and B_2 should have the same pattern. This implies that $d_{u_1'+i,u_2'+j} = c_{i,j}$ for $0 \leq i < b_1 - u_1'$, $0 \leq j < b_2 - u_2'$. In particular it implies that $d_{u_1'+1,u_2'} = x$. But $(u_1', u_2') \in I_{u_1,u_2} - J$, implies $(u_1' + 1, u_2') \neq (v_1 + 1, v_2)$. Since d_{v_1+1,v_2} is the only element in the array $[d_{i,j}]$ $0 \leq i < b_1$, $0 \leq j < b_2$ that depends on x, it follows that $d_{u_1'+1,u_2'}$ can not depend on x. This contradiction proves the lemma. ∎

It follows from the previous lemma that $r(b_1, b_2) \geq 2b_1b_2 - 2$.

Before presenting a construction of a $b_1 \times b_2$-burst identification code with redundancy $2b_1b_2$, we have to say something about the notation that we shall use. Let $\mathcal{Q}$ be the set $\{0', 1', \ldots, (b_1b_2-1)'\} \cup \{0'', 1'', \ldots, (b_1b_2-1)''\}$. A vector $\mathbf{h} \in \mathbf{F}_2^{2b_1b_2}$ can be represented as $\mathbf{h} = (h_{0'}, h_{1'}, \ldots, h_{(b_1b_2-1)'}, h_{0''}, h_{1''}, \ldots, h_{(b_1b_2-1)''})$. We associate with every vector $\mathbf{h} \in \mathbf{F}_2^{2b_1b_2}$ its characteristic set $\mathcal{H} = \{q \in \mathcal{Q} : h_q = 1\}$. In particular, the parity check matrix $H = [\mathbf{h}_{i,j}]$ can be represented by the sets $\mathcal{H}_{i,j}$ instead of the vectors $\mathbf{h}_{i,j}$.

In the following construction of burst identification codes, we define a matrix $\hat{H} = [\hat{\mathbf{h}}_{i,j}]$, $0 \le i < 2b_1$, $0 \le j < 2b_2$. The parity check matrix of a code of area $n_1 \times n_2$, denoted by $H = [\mathbf{h}_{i,j}]$, $0 \le i < n_1, 0 \le j < n_2$, is then defined periodically by $\mathbf{h}_{i,j} = \hat{\mathbf{h}}_{i \bmod 2b_1, j \bmod 2b_2}$. The matrix $\hat{H}$ is called the *building block* of the code.

Theorem 7. *Let the elements of the $2b_1 \times 2b_2$ building block $\hat{\mathcal{H}}$ be defined as*

$$\begin{aligned}
\hat{\mathcal{H}}_{i,j} &= \{(ib_2 + j)'\}, \\
\hat{\mathcal{H}}_{i,j+b_2} &= \{(ib_2 + j)', (ib_2 + j)''\}, \\
\hat{\mathcal{H}}_{i+b_1,j} &= \{(ib_2 + j)', (((i+1) \bmod b_1)b_2 + j)''\}, \\
\hat{\mathcal{H}}_{i+b_1,j+b_2} &= \{(ib_2 + j)''\},
\end{aligned}$$

where $0 \le i < b_1$, $0 \le j < b_2$. Then, $\hat{\mathcal{H}}$ is the building of a $b_1 \times b_2$-burst identification code of redundancy $2b_1b_2$.

Before giving the proof, the following is an example of the construction of the building block for $b_1 = b_2 = 3$.

Example

With $b_1 = b_2 = 3$, the building block $\hat{\mathcal{X}}$ is

$$\begin{bmatrix} \{0'\} & \{1'\} & \{2'\} & \{0',0''\} & \{1',1''\} & \{2',2''\} \\ \{3'\} & \{4'\} & \{5'\} & \{3',3''\} & \{4',4''\} & \{5',5''\} \\ \{6'\} & \{7'\} & \{8'\} & \{6',6''\} & \{7',7''\} & \{8',8''\} \\ \{0',3''\} & \{1',4''\} & \{2',5''\} & \{0''\} & \{1''\} & \{2''\} \\ \{3',6''\} & \{4',7''\} & \{5',8''\} & \{3''\} & \{4''\} & \{5''\} \\ \{6',0''\} & \{7',1''\} & \{8',2''\} & \{6''\} & \{7''\} & \{8''\} \end{bmatrix}. \tag{1}$$

Proof. It is clear from the construction that the redundancy is $2b_1b_2$. Let B be a $b_1 \times b_2$-burst whose burst pattern is $[a_{i,j}]$, $0 \le i < b_1$, $0 \le j < b_2$, starting at position (u_1, u_2). Its syndrome is given by

$$\mathbf{s} = \sum_{i=0}^{b_1-1} \sum_{j=0}^{b_2-1} a_{i,j} \mathbf{h}_{u_1+i,u_2+j}.$$

Let $J = \{(u_1 + i, u_2 + j) : a_{i,j} = 1\}$ be the set of positions of the nonzero elements of B. The projection of the burst B on the building block of area $2b_1 \times 2b_2$ is a cyclic burst $\hat{B}$, where the position of its nonzero components are given by the set $\hat{J} = \{i \bmod 2b_1, j \bmod 2b_2 : (i,j) \in J\}$. From Lemma 1 in the preliminaries, it follows that the cyclic burst $\hat{B}$ has unique burst pattern and starting position. Hence, it suffices to show that the burst pattern of $\hat{B}$ can be uniquely determined from the syndrome $\mathbf{s}$. In fact, we will show that from $\mathbf{s}$ we can even uniquely determine $\hat{J}$, except for few cases in which $\hat{J}$ is determined up to a shift of b_1 and b_2, in the vertical and horizontal directions, respectively.

It follows from the definition of $\hat{\mathcal{X}}$ that $l \in \hat{\mathcal{X}}_{i,j} \cap \hat{\mathcal{X}}_{i',j'}$, where $(i,j) \neq (i',j')$, implies $|i - i'| \ge b_1$ or $|j - j'| \ge b_2$. So, if S denotes the characteristic set of the syndrome $\mathbf{s}$, then

$$S = \biguplus_{(i,j) \in \hat{J}} \mathcal{X}_{i,j}, \tag{2}$$

where $\biguplus$ denotes union of disjoint sets. It also follows from the definition of $\hat{\mathcal{X}}$ that if $\hat{\mathcal{X}}_{i,j}$ contains l' or l'', then $l \equiv j \pmod{b_2}$. Let $\hat{J}_l = \{(i,j) \in \hat{J} : j \equiv l$

$(\mathrm{mod}\ b_2)\}$, $S_l' = \{s' \in S : s \equiv l \pmod{b_2}\}$, $S_l'' = \{s'' \in S : s \equiv l \pmod{b_2}\}$, and $S_l = S_l' \uplus S_l''$, where $0 \leq l < b_2$. Hence, $\hat{J} = \bigcup_{0 \leq l < b_2} \hat{J}_l$ and $S = \uplus_{0 \leq l < b_2} S_l$, where $S_l = \uplus_{(i,j) \in \hat{J}_l} \hat{\mathcal{X}}_{i,j}$ for $0 \leq l < b_2$, which follows from (2). It is sufficient to show that from each individual S_l, where $0 \leq l < b_2$, we can determine uniquely the burst pattern defined by $\hat{J}_l$, which is a $b_1 \times 1$-burst. This is demonstrated only for $l = 0$ since the other values of l can be treated similarly.

Notice from the construction of $\hat{\mathcal{X}}$ that the number of elements contained in $\hat{\mathcal{X}}_{i,0}$ from the set $\{0', 1', \ldots, (b_1 b_2 - 1)'\}$ is at least equal to the number of elements contained in $\hat{\mathcal{X}}_{i,0}$ from the set $\{0'', 1'', \ldots, (b_1 b_2 - 1)''\}$. Of course, the same holds for disjoint unions of $\hat{\mathcal{X}}_{i,0}$. The converse holds for $\hat{\mathcal{X}}_{i,b_2}$. Equality occurs only in $\hat{\mathcal{X}}_{i,0}$ for $b_1 \leq i < 2b_1$, and in $\hat{\mathcal{X}}_{i,b_2}$ for $0 \leq i < b_1$.

Hence, we have the following set of rules for identifying the $b_1 \times 1$-burst pattern defined by $\hat{J}_0$ from $\hat{S}_0$. For other values of l, these rules are also applicable after obvious modifications.

Rule (1):
If $|S_0'| > |S_0''|$,
or $|S_0'| = |S_0''|$ and $S_0'' \neq \{(ib_2)'' : (ib_2)' \in S_0'\}$:
In this case, the elements of $\hat{J}_0$ are of the form $(i, 0)$. In fact, $\hat{J}_0 = \{(i,0) : (ib_2)' \in S_0',\ (((i+1) \bmod b_1)b_2)'' \notin S_0''\} \cup \{(i + b_1, 0) : (ib_2)' \in S_0',\ (((i+1) \bmod b_1)b_2)'' \in S_0''\}$.

Rule (2):
If $|S_0'| < |S_0''|$,
or $|S_0'| = |S_0''|$ and $S_0'' \neq \{(((i+1) \bmod b_1)b_2)'' : (ib_2)' \in S_0'\}$:
In this case, the elements of $\hat{J}_0$ are of the form (i, b_2). In fact, $\hat{J}_0 = \{(i, b_2) : (ib_2)' \in S_0',\ (ib_2)'' \in S''\} \cup \{(i + b_1, b_2) : (ib_2)' \notin S_0',\ (ib_2)'' \in S_0''\}$.

Rule (3):
If $S_0'' = \{(ib_2)'' : (ib_2)' \in S_0'\} = \{(((i+1) \bmod b_1)b_2)'' : (ib_2)' \in S_0'\}$:

In this case, either $S_0' = S_0'' = S_0 = \phi$, the null set, which implies $\hat{J}_0 = \phi$, or $S_0' = \{(ib_2)' : 0 \leq i < b_1\}$ and $S_0'' = \{(ib_2)'' : 0 \leq i < b_1\}$. The latter case implies that $\hat{J}_0 = \{(i,0) : b_1 \leq i < 2b_1\}$ or $\hat{J}_0 = \{(i,b_2) : 0 \leq i < b_1\}$. However, these two possibilities for $\hat{J}_0$ give the same pattern.

By applying this algorithm to S_l for $l = 1,2,\ldots,b_2 - 1$, this ambiguity will be resolved unless when rule (3) is applicable for all values of l. In the later situation, two possibilities for $\hat{J}$ can be deduced from the syndrome S. However, both give the same burst pattern. ∎

Example continued

Let $S = \{2',5',6',7',8',4'',5'',6'',7'',8''\}$ be a syndrome with respect to the building block of the 3 × 3-burst identification code given in (1). Then, $S_0' = \{6'\}$, $S_0'' = \{6''\}$. Rule (2) applies, and we find $\hat{J}_0 = \{(2,3)\}$. Similarly, $S_1' = \{7'\}$, $S_1'' = \{4'',7''\}$, and rule (2) yields $\hat{J}_1 = \{(2,4),(4,4)\}$. Finally, $S_2' = \{2',5',8'\}$, $S_2'' = \{5'',8''\}$, and rule (1) yields $\hat{J}_2 = \{(2,2),(3,2),(4,2)\}$. So, the cyclic burst $\hat{B}$ deduced from $\hat{J}_0, \hat{J}_1, \hat{J}_2$ is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and its pattern is given by

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Example continued

Let $S = \{0',1',3',6',0'',3'',4'',6'',8''\}$ be a syndrome with respect to the building block of the 3 × 3-burst identification code given in (1). Then, $S_0' = \{0',3',6'\}$, $S_0'' = \{0'',3'',6''\}$, and rule (3) yields $\hat{J}_0 = \{(3,0),(4,0),(5,0)\}$ or

$\{(0,3),(1,3),(2,3)\}$. We also have $S_1' = \{1'\}$, $S_1'' = \{4''\}$, and rule (1) yields $\hat{J}_1 = \{(3,1)\}$. Since the burst is assumed to be confined in a rectangle of area 3×3, then $\hat{J}_0 = \{(3,0)(4,0)(5,0)\}$. Finally, $S_2' = \phi$, $S_2'' = \{8''\}$, and rule(2) yields $\hat{J}_2 = \{(5,5)\}$. So, the cyclic burst $\hat{B}$ deduced from $\hat{J}_0, \hat{J}_1, \hat{J}_2$ is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and its pattern is given by

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

The next theorem, which is the most important result of this chapter, follows from Lemma 6 and Theorem 7.

Theorem 8. *Let $r(b_1, b_2)$ be the minimum redundancy required to construct a $b_1 \times b_2$- burst identification code of arbitrarily large area. Then,*

$$2b_1b_2 - 2 \leq r(b_1, b_2) \leq 2b_1b_2.$$

In the examples given, the burst projection $\hat{J}$ on the building block is uniquely determined from S. As mentioned in rule (3) of Theorem 7, this is not the case for all bursts. In the next chapter, when we study certain classes of burst correcting codes, we need to determine $\hat{J}$ uniquely from S for all bursts. The following lemma shows that this can be easily done by using b_2 more redundant bits.

Lemma 9. *Let the vectors of the $2b_1 \times 2b_2$ building block $\tilde{\mathcal{H}}$ be defined by their characteristic sets as*

$$\tilde{\mathcal{H}}_{i,j} = \begin{cases} \hat{\mathcal{H}}_{i,j} \cup \{(j - b_2)'''\}, & \text{for } i = 0,\ b_2 \leq j < 2b_2; \\ \hat{\mathcal{H}}_{i,j}, & \text{otherwise,} \end{cases}$$

where $\hat{\mathcal{H}}_{i,j}$ is given in Theorem 7. Then, $\tilde{\mathcal{H}}$ is a building block of a $b_1 \times b_2$-burst identification code of redundancy $2b_1b_2 + b_2$, which can uniquely determine the projection of any $b_1 \times b_2$-burst on the building block.

Example

With $b_1 = b_2 = 3$, the building block $\tilde{\mathcal{H}}$ is

$$\begin{bmatrix} \{0'\} & \{1'\} & \{2'\} & \{0',0'',0'''\} & \{1,'1'',1'''\} & \{2,'2'',2'''\} \\ \{3'\} & \{4'\} & \{5'\} & \{3',3''\} & \{4',4''\} & \{5',5''\} \\ \{6'\} & \{7'\} & \{8'\} & \{6',6''\} & \{7',7''\} & \{8',8''\} \\ \{0',3''\} & \{1',4''\} & \{2',5''\} & \{0''\} & \{1''\} & \{2''\} \\ \{3',6''\} & \{4',7''\} & \{5',8''\} & \{3''\} & \{4''\} & \{5''\} \\ \{6',0''\} & \{7',1''\} & \{8',2''\} & \{6''\} & \{7''\} & \{8''\} \end{bmatrix}. \quad (3)$$

Proof. The redundancy is $2b_1b_2 + b_2$. From the construction of $\hat{\mathcal{H}}$ and $\tilde{\mathcal{H}}$, it follows that the code defined by $\tilde{\mathcal{H}}$ is a subcode of the code defined by $\hat{\mathcal{H}}$. We use the same notation introduced in the proof of Theorem 7. The characteristic set S of the syndrome with respect to $\tilde{\mathcal{H}}$, is a disjoint union of the characteristic set of the syndrome with respect to $\hat{\mathcal{H}}$ and a subset of $\{0''', 1''', \ldots, (b_2 - 1)'''\}$. Hence, rules (1),(2) and (3) apply. However, rule (3) can be modified to be rule (3') by making use of the b_2 extra redundant bits as described in the following:

Rule (3'):

If $S_0' = S_0'' = \phi$, then $\hat{J}_0 = \phi$.

If $S_0' = \{(ib_2)' : 0 \le i < b_1\}$ and $S_0'' = \{(ib_2)'' : 0 \le i < b_2\}$, then

$$\hat{J}_0 = \begin{cases} \{(i,0) : b_1 \le i < 2b_1\}, & \text{if } 0''' \notin S; \\ \{(i,b_2) : 0 \le i < b_2\}, & \text{if } 0''' \in S. \end{cases}$$

This resolves the ambiguity in rule (3) of Theorem 7. ∎

Example continued

Let $S = \{0',1',3',4',6',7',0'',1'',3'',4'',6'',7'',0''',1'''\}$ be the syndrome with respect to the building block $\tilde{\mathcal{H}}$ given in (3). Then $S_0' = \{0',3',6'\}$, $S_0'' =$

$\{0'', 3'', 6''\}$, and rule (3') yields $\hat{J}_0 = \{(0,3),(1,3),(2,3)\}$. Similarly, $S_1' = \{1', 4', 7'\}$, $S_1'' = \{1'', 4'', 7''\}$, and rule (3') yields $\hat{J}_1 = \{(0,4),(1,4),(2,4)\}$. Finally, $S_2' = S_0'' = \phi$, and rule (3') yields $\hat{J}_2 = \phi$. So, the cyclic burst $\hat{B}$ deduced from $\hat{J}_0, \hat{J}_1, \hat{J}_2$ is

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and its pattern is given by

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Note that if the building block $\hat{X}$, given in (1), is used, then the syndrome becomes $S = \{0', 1', 3', 4', 6', 7', 0'', 1'', 3'', 4'', 6'', 7''\}$. In this case, $\hat{J}$ is not uniquely determined. Instead, we have the shown two possibilities that give, of course, the same pattern,

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

4. Some Specific Burst Identification Codes

In this section we will consider $b_1 \times b_2$-burst identification codes for some specific values of b_1 and b_2. We note that if H is a parity check matrix of a $b_1 \times b_2$-burst identification code, then the transpose of H is a parity check matrix of a $b_2 \times b_1$-burst identification code, and so $r(b_1, b_2) = r(b_2, b_1)$.

4.1. 1 × b-Burst Identification Codes

Obviously, $r(1,1) = 1$ which is achieved by a code whose parity check matrix is composed of 1's. So, in the following we take $b > 1$. The next theorem gives an

explicit construction for $1 \times b$-burst identification code with $2b - 2$ redundant bits.

Theorem 10. *Let $b > 1$ and let $\mathbf{e}_i$ be the ith canonical vector of length $2b - 2$. Then, the code of area $n_1 \times n_2$ whose parity check matrix is given by $H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, where $\mathbf{h}_{i,j} = \mathbf{e}_{j \bmod (2b-2)}$, is a $1 \times b$-burst identification code of redundancy $2b - 2$.*

Proof. The patterns of the $1 \times b$-bursts are the same as those of the one-dimensional b-bursts. In Theorem 3, a construction is given of a one-dimensional b-burst identification code. Hence, the code defined in Theorem 10, which is simply the code defined in Theorem 3 repeated n_1 times, is a $1 \times b$- burst identification code. The redundancy is obviously $2b - 2$. ∎

Combining this result with Theorem 8, we obtain the following theorem.

Theorem 11. *$r(1,1) = 1$, and $r(1,b) = r(b,1) = 2b - 2$ for $b > 1$.*

4.2. 2×2-Burst Identification Codes

From Theorem 8, we know that $6 \leq r(2,2) \leq 8$. Here, we will prove that $r(2,2) = 7$.

First, we will show that $r(2,2) > 6$. Suppose that H is a 3×4 submatrix of a parity check matrix of a 2×2-burst identification code with redundancy 6. By studying the structure of H, we will establish a contradiction. By Lemma 6, we may assume, without loss of generality, that H has the form

$$\begin{bmatrix} \mathbf{p} & \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{u} \\ \mathbf{q} & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{v} \\ \mathbf{t} & \mathbf{e}_4 & \mathbf{e}_5 & \mathbf{w} \end{bmatrix},$$

where $\mathbf{e}_i$ is the ith canonical vector of length 6, and $\mathbf{p}, \mathbf{q}, \mathbf{t}, \mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ are vectors of length 6. We shall write $\mathbf{p} = (p_0, p_1, \ldots, p_5)$, and the same notation holds for the other vectors.

Lemma 12. $q_1 = q_5 = v_0 = v_4 = 1$.

Proof. Suppose that $q_1 = 0$. By Lemma 6, applied to $\{(i,j) : i = 1,2\ j = 0,1,2\}$, it follows that $q_0 = 1$. The burst

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1+q_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has syndrome $(0,0,1,q_3,q_4,q_5)$, as does the burst

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & q_3 & 0 \\ 0 & q_4 & q_5 & 0 \end{bmatrix},$$

while, obviously, these two bursts have different patterns. This contradiction proves that $q_1 = 1$. For reasons of symmetry, $q_5 = v_0 = v_4 = 1$. ∎

Lemma 13. $\mathbf{p} = \mathbf{e}_5$, $\mathbf{t} = \mathbf{e}_1$, $\mathbf{u} = \mathbf{e}_4$, *and* $\mathbf{w} = \mathbf{e}_0$.

Proof. Let x be an indeterminate. Since $q_1 = 1$ by Lemma 12, it follows that the burst B_1, given by

$$\begin{bmatrix} 1 & p_0 + p_1 q_0 & 0 & 0 \\ p_1 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

has syndrome $(0,0,p_2+p_1q_2+x, p_3+p_1q_3, p_4+p_1q_4, p_5+p_1q_5)$, as does the burst B_2, given by

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & p_2+p_1q_2+x & p_3+p_1q_3 & 0 \\ 0 & p_4+p_1q_4 & p_5+p_1q_5 & 0 \end{bmatrix}.$$

Hence, these two bursts should have the same pattern. By taking $x = 1$, it follows that $p_2+p_1q_2 = 0$, $p_0+p_1q_0 = p_3+p_1q_3$, $p_1 = p_4+p_1q_4$, and $p_5+p_1q_5 = 1$. Hence

$$B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & x & p_0+p_1q_0 & 0 \\ 0 & p_1 & 1 & 0 \end{bmatrix}.$$

If we now take $x = 0$, and compare B_1 with B_2, it follows that $p_1 = p_0 = 0$. Substituting this in the previous equations, we get $p_2 = p_3 = p_4 = 0$ and $p_5 = 1$. Hence, $\mathbf{p} = \mathbf{e}_5$. By symmetry, we get $\mathbf{t} = \mathbf{e}_1$, $\mathbf{u} = \mathbf{e}_4$, and $\mathbf{w} = \mathbf{e}_0$. ∎

By Lemma 13, H has the form

$$\begin{bmatrix} \mathbf{e}_5 & \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{e}_4 \\ \mathbf{q} & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{v} \\ \mathbf{e}_1 & \mathbf{e}_4 & \mathbf{e}_5 & \mathbf{e}_0 \end{bmatrix}.$$

Furthermore, by Lemma 12 we have $q_1 = q_5 = 1$ and $v_0 = v_4 = 1$. The proof of the next lemma contradicts the assumption $r(2,2) = 6$.

Lemma 14. $r(2,2) > 6$.

Proof. By Lemma 6, applied to $\{(i,j) : i = 0,1,\ j = 0,1,2\}$, $\{(i,j) : i = 1,2,\ j = 0,1,2\}$, and $\{(i,j) : i = 0,1,2,\ j = 0,1\}$, we get $q_4 = q_0 = q_3 = 1$. Hence, $\mathbf{q} = (1,1,q_2,1,1,1)$. By symmetry, $\mathbf{v} = (1,1,1,v_3,1,1)$. However, the burst

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1+q_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has syndrome $(0,1,1,1,1,1)$, as does the burst

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1+v_3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

while, obviously, these bursts have different patterns. This contradiction proves the lemma. ∎

The following theorem gives a 2×2-burst identification code with redundancy 7. As in Theorem 7, the parity check matrix $H = [\mathbf{h}_{i,j}]$, $0 \le i < n_1$, $0 \le j < n_2$, is defined periodically by the 4×4 building block $\hat{H} = [\hat{\mathbf{h}}_{i,j}]$, $0 \le i < 4$, $0 \le j < 4$, as $\mathbf{h}_{i,j} = \hat{\mathbf{h}}_{i \bmod 4, j \bmod 4}$, where $\hat{\mathbf{h}}_{i,j} \in \mathbf{F}_2^7$.

Theorem 15. *Let*

$$\hat{H} = \begin{bmatrix} \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \mathbf{e}_4 & \mathbf{e}_5 & \mathbf{e}_6 & \mathbf{1} \\ \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_0 & \mathbf{e}_1 \\ \mathbf{e}_6 & \mathbf{1} & \mathbf{e}_4 & \mathbf{e}_5 \end{bmatrix},$$

where $\mathbf{e}_i$ is the ith canonical vector of length 7, and $\mathbf{1} = (1,1,1,1,1,1,1)$. Then $\hat{H}$ is the building block of a 2×2-burst idenification code of redundancy 7.

Proof. It is clear from the construction that the redundancy is 7. Since the building block has size 4×4, it suffices, as we have demonstrated in the proof of Theorem 7 by using Lemma 1 in the preliminaries, to show that the burst pattern of any 2×2-cyclic burst $\hat{B}$ on the building block can be uniquely determined from its syndrome.

Let $\hat{B}$ be a 2×2-cyclic burst, and let $\mathbf{s}$ be its syndrome, whose weight is denoted by $w(\mathbf{s})$. Let $\hat{J}$ be the set of positions of the nonzero components of $\hat{B}$. From the construction of $\hat{H}$, it follows that each vector $\mathbf{u} \in \{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_6, \mathbf{1}\}$ occurs twice in $\hat{H}$, namely at positions (i, j) and $(i+2 \bmod 4, j+2 \bmod 4)$, for some $0 \leq i, j < 4$. Since $\hat{B}$ is assumed to be of area 2×2, or less, it follows that no vector $\mathbf{u}$ can contribute twice to $\mathbf{s}$. The weight $w(\mathbf{s}) = 4$ if, and only if, the pattern of $\hat{B}$ is $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$. For all other burst patterns, $w(\mathbf{s}) > 4$ if, and only if, $\hat{J}$ contains $(1,3)$ or $(3,1)$. So, we may replace $\mathbf{1}$ by $\mathbf{e}_7$ in the burst identification algorithm, where we view now the $\mathbf{e}_i$'s as the canonical vectors of length 8. The weight of the burst is now equal to $w(\mathbf{s})$. Let $\mathbf{u}$ be one of the vectors that contributed to $\mathbf{s}$. The nonzero components of the burst $\hat{B}$ is contained in a 3×3 subarray, which corresponds to the 3×3 submatrix

$$\hat{H}(\mathbf{u}) = \begin{bmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 \\ \mathbf{u}_3 & \mathbf{u} & \mathbf{u}_4 \\ \mathbf{u}_2 & \mathbf{u}_5 & \mathbf{u}_0 \end{bmatrix},$$

where the vectors $\mathbf{u}_i$, $0 \leq i \leq 5$, and $\mathbf{u}$ are all different, and from the set $\{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_7\}$. Note that this 3×3 submatrix is the same for the two positions of $\mathbf{u}$ in $\hat{H}$. The burst pattern of $\hat{B}$ can now be easily determined from $\mathbf{s}$ and $\hat{H}(\mathbf{u})$. ∎

Combining Lemma 14 and Theorem 15, we arrive at the following theorem.

Theorem 16. $r(2,2) = 7$.

4.3. $b_1 \times b_2$-Burst Identification Codes of Redundancy $2b_1b_2 - 2$

From Theorem 8, we have $2b_1b_2 - 2 \leq r(b_1, b_2) \leq 2b_1b_2$. On the other hand, Theorem 11 gives $r(1, b) = r(b, 1) = 2b - 2$ if $b > 1$. In this section, we will state necessary conditions for a code with redundancy $2b_1b_2 - 2$, to be a $b_1 \times b_2$-burst identification code when $b_1, b_2 > 1$.

In Lemma 6, we proved that the parity check vectors of a $b_1 \times b_2$-burst identification code in two $b_1 \times b_2$-blocks, intersecting in two positions, are linearly independent. In the next lemma, we consider the linear relations that may exist among the parity check vectors in two $b_1 \times b_2$-blocks intersecting in one position.

Lemma 17. *Let $H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, be a parity check matrix of a $b_1 \times b_2$-burst identification code, where $b_1, b_2 > 1$. Let*

$$I = \{(i, j) : 0 \leq i < b_1,\ 0 \leq j < b_2\} \bigcup \{(i, j) : b_1 - 1 \leq i < 2b_1 - 1,\ b_2 - 1 \leq j < 2b_2 - 1\},$$

and $0 \leq u_1 \leq n_1 - 2b_1 + 1$, $0 \leq u_2 \leq n_2 - 2b_2 + 1$. Then, there exists a unique positive integer d, $d | \gcd(b_1 - 1, b_2 - 1)$, such that if

$$\sum_{(i,j) \in I} a_{u_1+i, u_2+j} \mathbf{h}_{u_1+i, u_2+j} = \mathbf{0}, \tag{4}$$

where $a_{i,j} \in \mathbf{F}_2$, and not all are zero, then $a_{u_1+i, u_2+j} = 1$ if, and only if, $(i, j) \in \{(kl_1, kl_2) : 0 \leq k \leq 2d, k \neq d\}$, where $l_1 = (b_1 - 1)/d$ and $l_2 = (b_2 - 1)/d$.

Proof. Without loss of generality, assume $u_1 = u_2 = 0$. Suppose that (4) holds. Let

$$J_1 = \{(i, j) : a_{i,j} = 1,\ 0 \leq i < b_1,\ 0 \leq j < b_2\} - \{(b_1 - 1, b_2 - 1)\}$$

and

$$J_2 = \{(i, j) : a_{i,j} = 1,\ b_1 - 1 \leq i < 2b_1 - 1,\ b_2 - 1 \leq j < 2b_2 - 1\} - \{(b_1 - 1, b_2 - 1)\}.$$

It follows from Lemma 6 that J_1 and J_2 are nonempty. From (4), we have

$$\sum_{(i,j)\in J_1} \mathbf{h}_{i,j} + x\mathbf{h}_{b_1-1,b_2-1} = \sum_{(i,j)\in J_2} \mathbf{h}_{i,j} + (a_{b_1-1,b_2-1} + x)\mathbf{h}_{b_1-1,b_2-1},$$

where x is an indeterminate. But the left hand side is the syndrome of a $b_1 \times b_2$-burst $B_1(x)$, whose nonzero components are confined to $\{(i,j) : 0 \le i < b_1,\ 0 \le j < b_2\}$, while the right hand side is a syndrome of a $b_1 \times b_2$-burst $B_2(x)$, whose nonzero components are confined to $\{(i,j) : b_1 - 1 \le i < 2b_1 - 1,\ b_2 - 1 \le j < 2b_2 - 1\}$. These bursts, having the same syndrome, have the same pattern. Suppose that $a_{b_1-1,b_2-1} = 1$. Then the number of nonzero components in $B_1(0)$ is $|J_1|$, and in $B_2(0)$ is $|J_2| + 1$. Hence, $|J_1| = |J_2| + 1$. However, $B_1(1)$ has $|J_1| + 1$ nonzero components, while $B_2(1)$ has $|J_2|$ nonzero components. This contradiction proves that $a_{b_1-1,b_2-1} = 0$.

Since $B_1(0)$ has the same pattern as $B_2(0)$, then there exists a pair of integers (l_1', l_2') such that

$$(i,j) \in J_1 \Leftrightarrow (i + l_1', j + l_2') \in J_2. \tag{5}$$

Similarly for $B_1(1)$ and $B_2(1)$, there exists a pair of integers (l_1'', l_2'') such that

$$(i,j) \in J_1 \bigcup \{(b_1 - 1, b_2 - 1)\} \Leftrightarrow (i + l_1'', j + l_2'') \in J_2 \bigcup \{(b_1 - 1, b_2 - 1)\}. \tag{6}$$

Applying Lemma 6, it follows that the parity check vectors whose positions are in the set $I - \{(0,j) : 0 \le j < b_2\}$ are linearly independent. Hence $(0, j_1) \in J_1$ for some $0 \le j_1 < b_2$. Using this statement in (6), it follows that $l_1'' = b_1 - 1$. Similarly, $l_2'' = b_2 - 1$. Substituting in (6), it follows that $(0,0) \in J_1$. Substituting in (5), it follows that $(l_1', l_2') \neq (b_1 - 1, b_2 - 1)$. Let $l_1 = l_1' - l_1''$ and $l_2 = l_2' - l_2''$. Then, $(l_1, l_2) \neq (0,0)$.

Let $(i', j') \in J_1$, then from (5), we have $(i' + l_1', j' + l_2') \in J_2$, and from (6), $(i' + l_1, j' + l_2) \in J_1 \bigcup \{(b_1 - 1, b_2 - 1)\}$. If $(i' + l_1, j' + l_2) \neq (b_1 - 1, b_2 - 1)$, then by the same argument we have $(i' + 2l_1, j' + 2l_2) \in J_1 \bigcup \{(b_1 - 1, b_2 - 1)\}$. Thus, by iterating the same argument, it follows that there exists a positive integer $d_{i',j'}$

such that $i' + d_{i',j'}l_1 = b_1 - 1$ and $j' + d_{i',j'}l_2 = b_2 - 1$, and for all $0 \leq k < d_{i',j'}$, $(i' + kl_1, j' + kl_2) \in J_1$. In particular, since $(0,0) \in J_1$, then there exists $d_{0,0}$ such that $d_{0,0}l_1 = b_1 - 1$ and $d_{0,0}l_2 = b_2 - 1$. Now, we argue that

$$J_1 = \{(kl_1, kl_2) : 0 \leq k < d_{0,0}\}.$$

Indeed, the containment $\supseteq$ is already proved. To prove the containment $\subseteq$, note that if $(i,j) \in J_1$, then $i + d_{i,j}l_1 = b_1 - 1 = d_{0,0}l_1$ and $j + d_{i,j}l_2 = b_2 - 1 = d_{0,0}l_2$. Thus, $i = (d_{0,0} - d_{i,j})l_1$ and $j = (d_{0,0} - d_{i,j})l_2$. Hence, $(i,j) \in \{(kl_1, kl_2) : 0 \leq k < d_{0,0}\}$. This proves the containment $\subseteq$. Since $l_1' = l_1 + l_1'' = l_1 + b_1 - 1$ and $l_2' = l_2 + l_2'' = l_2 + b_2 - 1$, then it follows from (5) that

$$J_2 = \{((k+1)l_1 + b_1 - 1, (k+1)l_2 + b_2 - 1) : 0 \leq k < d_{0,0}\}.$$

Thus, $a_{i,j} = 1$ if, and only if,

$$(i,j) \in J_1 \bigcup J_2 = \{(kl_1, kl_2) : 0 \leq k \leq 2d,\ k \neq d\},$$

where $d = d_{0,0}$.

To prove that d is unique, suppose that

$$\sum_{(i,j) \in I} a'_{i,j}\mathbf{h}_{i,j} = \mathbf{0},$$

where $a'_{i,j} \in \mathbf{F}_2$, and not all are zero, such that $a'_{i,j} = 1$ if, and only if, $(i,j) \in \{(kt_1, kt_2) : 0 \leq k \leq 2m,\ k \neq m\}$, where $t_1 = (b_1 - 1)/m$ and $t_2 = (b_2 - 1)/m$, for some positive integer m. Then,

$$\sum_{(i,j) \in I} a''_{i,j}\mathbf{h}_{i,j} = \mathbf{0},$$

where $a''_{i,j} = a_{i,j} + a'_{i,j}$. Hence, $a''_{0,0} = 0$. If $m \neq d$, then not all the $a''_{i,j}$'s are zero. But, this contradicts the arguments given since $a''_{0,0} = 0$. ∎

If there exists a $b_1 \times b_2$-burst identification code with redundancy $2b_1b_2 - 2$,

where $b_1, b_2 > 1$, then (4) holds for every (u_1, u_2) such that $0 \le u_1 \le n_1 - 2b_1 + 1$ and $0 \le u_2 \le n_2 - 2b_2 + 1$. In the next theorem, we will prove that d of Lemma 4 does not depend on (u_1, u_2).

Theorem 18. *Suppose* $H = [\mathbf{h}_{i,j}]$, $0 \le i < n_1$, $0 \le j < n_2$ *is a parity check matrix of a* $b_1 \times b_2$*-burst identification code of redundancy* $2b_1b_2 - 2$, *where* $b_1, b_2 > 1$. *Let*

$$\bar{I} = \{(i,j) : 0 \le i < b_1,\ 0 \le j < b_2\} \bigcup \{(i,j) : b_1 - 1 \le i < 2b_1 - 1,\ b_2 - 1 \le j < 2b_2 - 1\}.$$

Then, there exists a unique positive integer $\bar{d} | \gcd(b_1 - 1, b_2 - 1)$ *such that if* $0 \le u_1 \le n_1 - 2b_1 + 1$, $0 \le u_2 \le n_2 - 2b_2 + 1$, *then*

$$\sum_{(i,j) \in \bar{I}} a_{u_1+i,u_2+j} \mathbf{h}_{u_1+i,u_2+j} = \mathbf{0},$$

for a unique nonempty set $\{(i,j) \in \bar{I} : a_{u_1+i,u_2+j} = 1\}$ *given by* $\{(k\bar{l}_1, k\bar{l}_2) : 0 \le k \le 2\bar{d},\ k \ne \bar{d}\}$, *where* $\bar{l}_1 = (b_1 - 1)/\bar{d}$, $\bar{l}_2 = (b_2 - 1)/\bar{d}$. *Similarly, let*

$$\underline{I} = \{(i,j) : 0 \le i < b_1,\ b_2 - 1 \le j < 2b_2 - 1\} \bigcup \{(i,j) : b_1 - 1 \le i < 2b_1 - 1,\ 0 \le j < b_2\}.$$

Then, there exists a unique positive integer $\underline{d} | \gcd(b_1 - 1, b_2 - 1)$ *such that if* $0 \le v_1 \le n_1 - 2b_1 + 1$, $0 \le v_2 \le n_2 - 2b_2 + 1$, *then*

$$\sum_{(i,j) \in \underline{I}} a_{v_1+i,v_2+j} \mathbf{h}_{v_1+i,v_2+j} = \mathbf{0},$$

for a unique nonempty set $\{(i,j) \in \underline{I} : a_{v_1+i,v_2+j} = 1\}$ *given by* $\{(k\underline{l}_1, (2\underline{d} - k)\underline{l}_2) : 0 \le k \le 2\underline{d},\ k \ne \underline{d}\}$, *where* $\underline{l}_1 = (b_1 - 1)/\underline{d}$, $\underline{l}_2 = (b_2 - 1)/\underline{d}$.

Proof. It suffices, by symmetry, to prove the theorem for $\bar{d}$. Since the redundancy is $2b_1b_2 - 2$, and $\bar{I}$ is of cardinality $2b_1b_2 - 1$, then if $0 \le u_1 \le n_1 - 2b_1 + 1$,

$0 \leq u_2 \leq n_2 - 2b_2 + 1$, there exists a nontrivial linear relation of the form

$$\sum_{(i,j)\in I} a_{u_1+i,u_2+j}\mathbf{h}_{u_1+i,u_2+j} = \mathbf{0}.$$

In Lemma 17, we proved that there exists a unique integer $d(u_1, u_2)$ such that $a_{u_1+i,u_2+j} = 1$ if, and only if, $(i,j) \in \{(kl_1(u_1,u_2), kl_2(u_1,u_2)) : 0 \leq k \leq 2d(u_1,u_2), k \neq d(u_1,u_2)\}$, where $l_1(u_1,u_2) = (b_1 - 1)/d(u_1,u_2)$ and $l_2(u_1,u_2) = (b_2 - 1)/d(u_1,u_2)$. It suffices, by induction and symmetry, to prove that $d(u_1,u_2) = d(u_1+1,u_2)$, where $0 \leq u_1 \leq n_1 - 2b_1$, $0 \leq u_2 \leq n_2 - 2b_2 + 1$. Let

$$J_1(u_1,u_2) = \{(kl_1(u_1,u_2), kl_2(u_1,u_2)) : 0 \leq k < d(u_1,u_2)\},$$

and

$$J_2(u_1,u_2) = \{(kl_1(u_1,u_2), kl_2(u_1,u_2)) : d(u_1,u_2) < k \leq 2d(u_1,u_2)\}.$$

By applying Lemma 17 to (u_1,u_2), we get

$$\sum_{(i,j)\in J_1(u_1,u_2)} \mathbf{h}_{u_1+i,u_2+j} = \sum_{(i,j)\in J_2(u_1,u_2)} \mathbf{h}_{u_1+i,u_2+j},$$

and by applying it to (u_1+1,u_2), we get

$$\sum_{(i,j)\in J_1(u_1+1,u_2)} \mathbf{h}_{u_1+1+i,u_2+j} = \sum_{(i,j)\in J_2(u_1+1,u_2)} \mathbf{h}_{u_1+1+i,u_2+j}.$$

Adding these equations, we get

$$\sum_{(i,j)\in I_1} \mathbf{h}_{u_1+i,u_2+j} = \sum_{(i,j)\in I_2} \mathbf{h}_{u_1+i,u_2+j}, \tag{7}$$

where

$$I_1 = \{(kl_1(u_1,u_2), kl_2(u_1,u_2)) : 0 \leq k < d(u_1,u_2)\}$$
$$\bigcup \{(kl_1(u_1+1,u_2)+1, kl_2(u_1+1,u_2)) : 0 \leq k < d(u_1+1,u_2)\}, \tag{8}$$

and

$$I_2 = \{(kl_1(u_1,u_2), kl_2(u_1,u_2)) : d(u_1,u_2) < k \le 2d(u_1,u_2)\}$$
$$\bigcup\{(kl_1(u_1+1,u_2)+1, kl_2(u_1+1,u_2)) : d(u_1+1,u_2) < k \le 2d(u_1+1,u_2)\}. \quad (9)$$

But

$$I_1 \subseteq \{(i,j) : 0 \le i < b_1,\ 0 \le j < b_2\},$$

and

$$I_2 \subseteq \{(i,j) : b_1 \le i < 2b_1,\ b_2 \le j < 2b_2\}.$$

Hence, from (7), there exist two $b_1 \times b_2$-bursts sharing the same syndrome, such that the positions of their nonzero components are given by I_1 and I_2. Hence, these two bursts have the same pattern, which implies that there exist a pair of positive integers t_1 and t_2 such that

$$(i,j) \in I_1 \Leftrightarrow (i+t_1, j+t_2) \in I_2.$$

From (8), it follows that $\{(0,0),(1,0)\} \subseteq I_1$ and $(i,j) \notin I_1$ for all $i < 0$. Hence, from (9), it follows that $t_1 = [d(u_1,u_2)+1]l_1(u_1,u_2)$ and $t_1 + 1 = [d(u_1+1,u_2)+1]l_1(u_1+1,u_2)+1$. But since $l_1(u_1,u_2)d(u_1,u_2) = l_1(u_1+1,u_2)d(u_1+1,u_2) = b_1 - 1$, then we have $d(u_1,u_2) = d(u_1+1,u_2)$. This, as argued before, suffices to prove the theorem. ∎

Suppose that there exists a $b_1 \times b_2$-burst identification code of area $n_1 \times n_2$ and redundancy $2b_1b_2 - 2$, where $0 < 2(b_1-1) < n_1$ and $0 < 2(b_2-1) < n_2$. Then, the integers $\bar{d}$ and $\underline{d}$ of Theorem 18 corresponding to the code, will be called the *diagonal parameters* of the code.

Corollary 19. *Suppose that* $H = [\mathbf{h}_{i,j}]$, $0 \le i < n_1$, $0 \le j < n_2$, *where* $0 < 4(b_1-1) < n_1$ *and* $0 < 4(b_2-1) < n_2$, *is a parity check matrix of a* $b_1 \times b_2$*-burst identification code of redundancy* $2b_1b_2 - 2$. *Then,* H *is uniquely determined from the submatrix* $H' = [\mathbf{h}_{i,j}]$, $0 \le i < 2(b_1-1)$, $0 \le j < 4(b_2-1)$, *and the diagonal parameters of the code.*

Proof. First, we show that $\mathbf{h}_{i,j}$, where $2(b_1-1) \le i < n_1$ and $0 \le j < 4(b_2-1)$, is uniquely determined from H'. This is proved by induction on i. Suppose that $\mathbf{h}_{i',j'}$ is uniquely determined for $i' < i$ and $0 \le j' < 4(b_2-1)$. Then, from Theorem 18, $\mathbf{h}_{i,j}$, where $2(b_2-1) \le j < 4(b_2-1)$, is uniquely determined from

$$\sum_{\substack{k=0 \\ k \neq d}}^{2d} \mathbf{h}_{i+k\bar{l}_1-2(b_1-1),j+k\bar{l}_2-2(b_2-1)} = \mathbf{0},$$

where $\bar{l}_1 = (b_1-1)/\bar{d}$ and $\bar{l}_2 = (b_2-1)/\bar{d}$. On the other hand, $\mathbf{h}_{i,j}$, where $0 \le j < 2(b_2-1)$, is uniquely determined from

$$\sum_{\substack{k=0 \\ k \neq \underline{d}}}^{2\underline{d}} \mathbf{h}_{i+k\underline{l}_1-2(b_1-1),j-k\underline{l}_2+2(b_2-1)} = \mathbf{0},$$

where $\underline{l}_1 = (b_1-1)/\underline{d}$ and $\underline{l}_2 = (b_2-1)/\underline{d}$. By interchanging i and j, the corollary follows. ∎

Corollary 19 implies that, for every $b_1 \times b_2$, there is a straightforward algorithm, though it may be very tedious computationally, to decide whether $r(b_1,b_2)$ equals $2b_1b_2-2$. Indeed, from this corollary it follows that the number of $b_1 \times b_2$-burst identification codes of redundancy $2b_1b_2-2$ is bounded by a function that depends only on b_1 and b_2, i.e., does not depend on the areas of the codes.

In the previous subsection, we proved that there are no 2×2-burst identification codes of large areas and redundancy 6. In the next subsection, we will show that there are no 3×2-burst identification codes of large areas and redundancy 10.

4.4. 3×2-Burst Identification Codes

From Theorem 8, it follows that $10 \leq r(3,2) \leq 12$. In the following, we will prove that $r(3,2) \neq 10$.

Suppose that H is a 7×5 submatrix of a parity check matrix of a 3×2-burst identification code with redundancy 10. By studying the structure of H, we will establish a contradiction. From Lemma 6 and Theorem 18 we may assume, without loss of generality, that

$$H = \begin{bmatrix} \mathbf{t} & \mathbf{w} & \mathbf{e}_6 & \mathbf{e}_7 & \mathbf{t} \\ \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{e}_8 & \mathbf{e}_9 & \mathbf{e}_0 \\ \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{p} & \mathbf{q} & \mathbf{e}_2 \\ \mathbf{e}_4 & \mathbf{e}_5 & \mathbf{u} & \mathbf{v} & \mathbf{e}_4 \\ \mathbf{e}_6 & \mathbf{e}_7 & \mathbf{t} & \mathbf{w} & \mathbf{e}_6 \\ \mathbf{e}_8 & \mathbf{e}_9 & \mathbf{e}_0 & \mathbf{e}_1 & \mathbf{e}_8 \\ \mathbf{p} & \mathbf{q} & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{p} \end{bmatrix}, \tag{10}$$

where $\mathbf{e}_i$ is the ith canonical vector of length 10, and $\mathbf{p}, \mathbf{q}, \mathbf{u}, \mathbf{v}, \mathbf{t}$, and $\mathbf{w}$ are vectors of length 10. We shall write $\mathbf{p} = (p_0, p_1, \ldots, p_9)$, and the same notation is used for the other vectors.

Lemma 20. *If $p_7 = 0$ or $q_6 = 0$, then $\mathbf{p} = (p_0, 1, 1, 0, 0, 0, 0, 0, p_8, 1)$ and $\mathbf{q} = (1, q_1, 0, 1, 0, 0, 0, 0, 1, q_9)$, where $p_0 = p_8$, $q_1 = q_9$, and $p_0 q_1 = p_8 q_9 = 0$.*

Proof. Suppose $p_7 = 0$. From Lemma 6 applied to $\{(i,j) : i = 0,1,2,\ j = 1,2\} \cup \{(i,j) : i = 1,2,3,\ j = 0,1\}$, it follows that $p_9 = 1$. The bursts

$$B_1 = \begin{bmatrix} 0 & 0 & p_6 & 0 & 0 \\ 0 & 0 & p_8 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ p_0 & p_1 & 0 & 0 & 0 \\ p_2 & p_3 & 0 & 0 & 0 \\ p_4 & p_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(p_0, p_1, p_2, p_3, p_4, p_5, 0, 0, 0, 0)$. Hence, they have the same pattern. By comparing their patterns, it follows that $p_5 = 0$. On the other hand, the bursts

$$B_3 = \begin{bmatrix} 0 & 0 & p_6 & 0 & 0 \\ 0 & p_1 & p_8 & 0 & 0 \\ 0 & p_3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & p_0 \\ 0 & 0 & 0 & 0 & p_2 \\ 0 & 0 & 0 & 0 & p_4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(p_0, 0, p_2, 0, p_4, 0, 0, 0, 0, 1)$. Hence, they have the same pattern. Thus, $p_4 = 0$. Substituting $p_4 = 0$ in B_1 and B_2, and comparing their patterns, it follows that $p_1 = p_2 = 1$, $p_3 = p_6 = 0$, and $p_0 = p_8$. Thus, $\mathbf{p} = (p_0, 1, 1, 0, 0, 0, 0, 0, p_8, 1)$, with $p_0 = p_8$.

Now, suppose that $q_6 = 1$. Then the bursts

$$B_5 = \begin{bmatrix} 0 & 0 & 1 & q_7 & 0 \\ 0 & 0 & q_8 + xp_8 & q_9 + x & 0 \\ 0 & 0 & x & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ q_0 + xp_0 & q_1 + x & 0 & 0 & 0 \\ q_2 + x & q_3 & 0 & 0 & 0 \\ q_4 & q_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where x is an indeterminate, share the syndrome $(q_0 + xp_0, q_1 + x, q_2 + x, q_3, q_4, q_5, 0, 0, 0, 0)$. Hence, they have the same pattern, which implies $x = q_4$, which is contrary to the definition of x. Hence, $q_6 = 0$. Due to the symmetry between $\mathbf{p}$ and $\mathbf{q}$, it follows that $\mathbf{q} = (1, q_1, 0, 1, 0, 0, 0, 0, 1, q_9)$, where $q_1 = q_9$. Finally, if $p_0 = q_1 = 1$, then $p_0 + q_0 = p_1 + q_1 = 0$, which contradicts Lemma 6 as applied to $\{(i, j) : i = 2, 3, 4, 5, 6,\ j = 0, 1\}$. ∎

Lemma 21. $p_7 = q_6 = 1$.

Proof. Suppose $p_7 = 0$ or $q_6 = 0$. Then, from Lemma 20, it follows that $\mathbf{p} = (p_0, 1, 1, 0, 0, 0, 0, 0, p_8, 1)$ and $\mathbf{q} = (1, q_1, 0, 1, 0, 0, 0, 0, 1, q_9)$, where $p_0 = p_8$, $q_1 = q_9$, and $p_0 q_1 = 0$.

Suppose $p_0 = 0$. Then $u_0 = 1$, since if $u_0 = 0$, then $\mathbf{u} = (u_1 + u_9)\mathbf{e}_1 + (u_2 + u_9)\mathbf{e}_2 + u_3\mathbf{e}_3 + u_4\mathbf{e}_4 + u_5\mathbf{e}_5 + u_6\mathbf{e}_6 + u_7\mathbf{e}_7 + u_8\mathbf{e}_8 + u_9\mathbf{p}$, which contradicts Lemma 6 as applied to $\{(i, j) : i = 1, 2, 3,\ j = 1, 2\} \cup \{(i, j) : i = 2, 3, 4,\ j = 0, 1\}$. The bursts

$$B_7 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 + u_8 & u_1 + u_9 & 0 \\ 0 & 0 & q_1 + u_1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ q_1+u_1+u_2 & 1+u_3 & 0 & 0 & 0 \\ u_4 & u_5 & 0 & 0 & 0 \\ u_6 & u_7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(0,0,q_1+u_1+u_2,1+u_3,u_4,u_5,u_6,u_7,0,0)$. Hence, they have the same pattern, which implies $u_7 = 0$. Since $u_7 = 0$, then the bursts

$$B_9 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & u_1+u_9 & 1+u_8 & 0 & 0 \\ 0 & 1+u_3 & q_1+u_9 & 0 & 0 \\ 0 & u_5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_{10} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & q_1+u_2+u_9 \\ 0 & 0 & 0 & 0 & u_4 \\ 0 & 0 & 0 & 0 & u_6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(1,q_1,q_1+u_2+u_9,1,u_4,0,u_6,0,1,q_9)$, and hence have the same pattern. If $u_1 \neq u_9$, then by comparing the patterns of B_7 and B_8, it follows that $u_5 = 1$, which leads to a contradiction by comparing the patterns of B_9 and B_{10}. Hence, $u_1 = u_9$. By comparing the patterns of B_9 and B_{10}, it follows that $u_6 = 0$. Then, by comparing the patterns of B_7 and B_8, where $u_6 = u_1 + u_9 = 0$, we get $u_2 = u_3 = u_5 = u_7 = 0, u_4 = u_8 = 1$. Thus, $u = (1,u_1,0,0,1,0,0,0,1,u_9)$, where $u_1 = u_9$. Hence, $\mathbf{u} = \mathbf{e}_0 + u_1\mathbf{e}_2 + \mathbf{e}_4 + \mathbf{e}_8 + u_1\mathbf{p}$, which contradicts Lemma 6 as applied to $\{(i,j) : i = 1,2,3,\ j = 0,1,2\}$. This contradiction yields $p_0 = 1$. But then, by Lemma 20, $q_1 = 0$, and a similar argument holds due to symmetry.

■

Lemma 22. $\mathbf{p} = (p_0, 1, p_2, p_3, 1, 0, p_6, 1, p_8, p_9)$, *where* $p_0 = p_6, p_2 = p_8, p_3 = p_9$, *and* $\mathbf{q} = (1, q_1, q_2, q_3, 0, 1, 1, q_7, q_8, q_9)$, *where* $q_1 = q_7, q_2 = q_8, q_3 = q_9$. *Moreover,* $p_0 q_1 = 0$ *and* $\mathbf{u} = u_2\mathbf{e}_2 + \mathbf{e}_3 + u_4\mathbf{e}_4 + u_5\mathbf{e}_5 + \mathbf{e}_6 + u_2\mathbf{e}_8 + \mathbf{e}_9 + u_4\mathbf{p} + u_5\mathbf{q}$.

Proof. By Lemma 21, we have $p_7 = q_6 = 1$. The bursts

$$B_{11} = \begin{bmatrix} 0 & 0 & p_6 & 1 & 0 \\ 0 & 0 & p_8 & p_9 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ p_0 & p_1 & 0 & 0 & 0 \\ p_2 & p_3 & 0 & 0 & 0 \\ p_4 & p_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(p_0, p_1, p_2, p_3, p_4, p_5, 0, 0, 0, 0)$. Hence, they have the same pattern, which implies $p_1 = p_4 = 1, p_5 = 0, p_0 = p_6, p_2 = p_8, p_3 = p_9$. Similarly, $q_0 = q_5 = 1, q_4 = 0, q_1 = q_7, q_2 = q_8, q_3 = q_9$.

By applying Lemma 6 to $\{(i,j) : i = 2,3,4,5,6,\ j = 0,1\}$, it follows that the set $\{\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_9, \mathbf{p}, \mathbf{q}\}$ forms a basis for $\mathbf{F}_2^{10}$. Thus, $p_0 q_1 = 0$. Hence, there exists an automorphism of $\mathbf{F}_2^{10}$ such that $\mathbf{e}_i \mapsto \mathbf{e}_{i-2}$ for $i = 2, 3, \ldots, 9$, $\mathbf{p} \mapsto \mathbf{e}_8$, and $\mathbf{q} \mapsto \mathbf{e}_9$. Since $\mathbf{p} = p_0\mathbf{e}_0 + \mathbf{e}_1 + p_2\mathbf{e}_2 + p_3\mathbf{e}_3 + \mathbf{e}_4 + p_0\mathbf{e}_6 + \mathbf{e}_7 + p_2\mathbf{e}_8 + p_3\mathbf{e}_9$, then by considering the image of H in (10) under this automorphism, it follows that $\mathbf{u} = u_2\mathbf{e}_2 + \mathbf{e}_3 + u_4\mathbf{e}_4 + u_5\mathbf{e}_5 + \mathbf{e}_6 + u_2\mathbf{e}_8 + \mathbf{e}_9 + u_4\mathbf{p} + u_5\mathbf{q}$. ∎

Theorem 23. $r(3,2) > 10$.

Proof. From Lemma 22, it follows that the bursts

$$B_{13} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & q_2 + u_2 + p_2 q_1 & 1 + q_3 + p_3 q_1 & 0 \\ 0 & 0 & q_1 + u_4 & 1 + u_5 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$B_{14} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ q_2 + u_2 + p_2 q_1 & 1 + q_3 + p_3 q_1 & 0 & 0 & 0 \\ q_1 + u_4 & 1 + u_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

share the syndrome $(1, 0, q_2 + u_2 + p_2 q_1, 1 + q_3 + p_3 q_1, q_1 + u_4, 1 + u_5, 0, 0, 0, 0)$, and hence have the same pattern. This implies $u_5 = q_3 + p_3 q_1 = 1$. This yields $\mathbf{u} + (q_1 + u_4)\mathbf{p} + (q_2 + u_2 + p_2 q_1)\mathbf{e}_3 = \mathbf{e}_0 + (q_2 + u_2 + p_2 q_1)\mathbf{e}_2 + (q_1 + u_4)\mathbf{e}_4$, which contradicts Lemma 6 as applied to $\{(i, j) : i = 1, 2, 3,\ j = 0, 1, 2\}$. This final contradiction proves that $r(3, 2) \neq 10$. ∎

CHAPTER IV

THE STRUCTURE OF BURST CORRECTING CODES

In this chapter, we study two-dimensional burst correcting codes, which are defined in section 1. In section 2, we give a number of bounds on the parameters of these codes. Burst locating codes, which are used in the construction of burst correcting codes, are examined in section 3. In sections 4, 5, and 6, three classes of burst correcting codes are developed.

1. Definitions

A two-dimensional linear code C is said to be a $b_1 \times b_2$-*burst correcting code* if no codeword, except the all zero codeword, is a $b_1 \times b_2$-burst or a sum of two $b_1 \times b_2$-bursts. Equivalently, the code C is a $b_1 \times b_2$-burst correcting code if, and only if, the syndromes of the $b_1 \times b_2$-bursts with respect to any given parity check matrix of C are nonzero and distinct.

This definition is useful because if a $b_1 \times b_2$-burst correcting code is used over a channel that may add to any transmitted codeword a $b_1 \times b_2$-burst, then the receiver can determine the burst added by the channel, and thus retrieve the transmitted codeword.

A two-dimensional code is said to be a $b_1 \times b_2$-*cyclic-burst correcting code* if no codeword, except the all zero codeword is a $b_1 \times b_2$-cyclic burst or a sum of two $b_1 \times b_2$-cyclic bursts. Note that a cyclic $b_1 \times b_2$-burst correcting code is a $b_1 \times b_2$-cyclic-burst correcting code, but the converse does not necessarily hold. Indeed, if C is a cyclic burst correcting code of area $n_1 \times n_2$, then $xc(x,y)$ and $yc(x,y)$, both $\bmod(x^{n_1}+1, y^{n_2}+1)$, are in C for each $c(x,y) \in C$. This is not necessarily true for cyclic-burst correcting codes.

We are mainly interested in $b_1 \times b_2$-burst correcting codes whose areas are much larger than $b_1 \times b_2$. Several classes of codes will be considered in this chapter. The following measure is used to estimate the redundancy required in each class. Consider an infinite class $\mathcal{S}$ of $b_1 \times b_2$-burst correcting codes, and suppose that for every positive integer n, the subset $\mathcal{S}(n)$ of codes in $\mathcal{S}$ whose areas are larger than $n \times n$ is nonempty. For each $C \in \mathcal{S}$, let $n_{1C} \times n_{2C}$ and r_C denote the area and redundancy of C, respectively. Then, we define the *excess redundancy** of the class $\mathcal{S}$ as

$$\underline{r}_{\mathcal{S}}(b_1, b_2) = \lim_{n \to \infty} \inf_{C \in \mathcal{S}(n)} \left(r_C - \log(n_{1C} n_{2C}) \right),$$

if such limit exists. It is to be noted that $\underline{r}_{\mathcal{S}}(b_1, b_2)$ exists if, and only if, $\inf_{C \in \mathcal{S}(n)} \left(r_C - \log(n_{1C} n_{2C}) \right)$ is bounded as a function of n since it is a nondecreasing function. If this function is unbounded, we take $\underline{r}_{\mathcal{S}}(b_1, b_2) = \infty$. The definition of excess redundancy may need some clarification. A $b_1 \times b_2$-burst correcting code of area $n_1 \times n_2$ and redundancy r must have distinct syndromes for all distinct 1×1-bursts. Since there are $n_1 n_2$ such bursts, it follows that r should be at least $\log(n_1 n_2)$. This explains the term "excess" used to describe $\underline{r}_{\mathcal{S}}(b_1, b_2)$. It follows from the definition of excess redundancy that if $\underline{r}_{\mathcal{S}}(b_1, b_2)$ is finite, then for every $\epsilon > 0$ and every positive integer n, there exists a $b_1 \times b_2$-burst correcting code in $\mathcal{S}$ of area $n_1 \times n_2$, for some $n_1 \times n_2$ greater than $n \times n$, whose redundancy is less than $\underline{r}_{\mathcal{S}}(b_1, b_2) + \log(n_1 n_2) + \epsilon$.

2. Bounds on Two-Dimensional Burst Correcting Codes

In this section, we will state and prove a number of bounds on two-dimensional burst correcting codes. The bounds stated in the following two theorems are extensions of similar bounds known for one-dimensional burst correcting codes.

The following bound is a Hamming-type volume bound, first stated by Fire

---

* *The concept of excess redundancy is a modified version of an earlier measure of efficiency of one-dimensional burst correcting codes developed by Fire [9].*

[9] in the one-dimensional case.

Theorem 1. *The redundancy r of a $b_1 \times b_2$-burst correction code of area $n_1 \times n_2$ satisfies*

$$2^r \geq 1 + (n_1 - b_1)(n_2 - b_2)N(b_1, b_2) + (n_1 - b_1)\sum_{b_2'=1}^{b_2} N(b_1, b_2') + (n_2 - b_2)\sum_{b_1'=1}^{b_1} N(b_1', b_2) + \sum_{b_1'=1}^{b_1}\sum_{b_2'=1}^{b_2} N(b_1', b_2'),$$

where $N(b_1, b_2)$ is the number of distinct patterns of $b_1 \times b_2$-bursts. If $n_1 \geq 2b_1 - 1$, $n_2 \geq 2b_2 - 1$, and the code is a $b_1 \times b_2$-cyclic-burst correcting code, then

$$2^r \geq 1 + n_1 n_2 N(b_1, b_2).$$

Proof. From the definition of a $b_1 \times b_2$-burst correcting code, it follows that the number of nonzero syndromes $2^r - 1$ should be at least equal to the number of $b_1 \times b_2$-bursts. $N(b_1, b_2)$ is the number of $b_1 \times b_2$-bursts starting at position (u_1, u_2) if $0 \leq u_1 < n_1 - b_1$ and $0 \leq u_2 < n_2 - b_2$. On the other hand, if $0 \leq u_1 < n_1 - b_1$ and $n_2 - b_2 \leq n_2 - b_2' = u_2 < n_2$, then the number of $b_1 \times b_2$-bursts starting at position (u_1, u_2) is $N(b_1, b_2')$. Similarly, the number of $b_1 \times b_2$-bursts starting at position (u_1, u_2) is $N(b_1', b_2)$ if $n_1 - b_1 \leq n_1 - b_1' = u_1 < n_1$ and $0 \leq u_2 < n_2 - b_2$. Finally, if $n_1 - b_1 \leq n_1 - b_1' = u_1 < n_1$ and $n_2 - b_2 \leq n_2 - b_2' = u_2 < n_2$, then the number of $b_1 \times b_2$-bursts starting at position (u_1, u_2) is $N(b_1', b_2')$. The first statement follows by counting the total number of $b_1 \times b_2$-bursts. Since, by Lemma 1 in the preliminaries, the number of $b_1 \times b_2$-cyclic bursts is $n_1 n_2 N(b_1, b_2)$ if $n_1 \geq 2b_1 - 1$ and $n_2 \geq 2b_2 - 1$, then the second statement also follows. ∎

The next lemma is analogous to the well known Varshamov-Gilbert bound, and was first stated for one-dimensional burst correcting codes by Campopiano [5],[20; chapter 4]. In the following, the bound is extended to the two-

dimensional case.*

Theorem 2. *There exists a $b_1 \times b_2$-burst correcting code of area $n_1 \times n_2$ and redundancy r if*

$$2^r \geq n_1 n_2 N^2(b_1, b_2).$$

Proof. We give an algorithm to construct a parity check matrix of a $b_1 \times b_2$-burst correction code of area $n_1 \times n_2$ and redundancy r if the inequality is satisfied. The algorithm consists of $n_1 n_2$ steps. In each step, an element $\mathbf{h}_{i,j} \in \mathbf{F}_2^r$ is obtained and placed in an $n_1 \times n_2$ array at position (i, j). By achieving the last step in the algorithm, the required parity check matrix is constructed. The elements are obtained in the order $\mathbf{h}_{n_1-1,n_2-1}, \mathbf{h}_{n_1-1,n_2-2}, \ldots, \mathbf{h}_{n_1-1,0}, \mathbf{h}_{n_1-2,n_2-1}, \ldots, h_{n_1-2,0}, \ldots, \mathbf{h}_{0,0}$. Start with $\mathbf{h}_{n_1-1,n_2-1} \neq \mathbf{0}$. Suppose now that we want to add an element $\mathbf{h}_{u_1,u_2}$ to the array that is partially filled. From the definition of $b_1 \times b_2$-burst correcting codes, it follows that $\mathbf{h}_{u_1,u_2}$ should not be a linear combination of any collection of elements already placed in the array whose positions are confined to a $b_1 \times b_2$-block containing the position (u_1, u_2) and any collection of elements already placed in the array whose positions are confined to a $b_1 \times b_2$-block. The number of distinct linear combinations of collections of elements already placed in the array whose positions are confined to a $b_1 \times b_2$-block, summed over all distinct $b_1 \times b_2$-blocks, is obviously bounded by $n_1 n_2 N(b_1, b_2)$. In the following, we will argue that the number of distinct linear combinations of collections of elements already placed in the array whose positions are confined to a $b_1 \times b_2$-block containing the position (u_1, u_2), is bounded by $N(b_1, b_2)$. Indeed, this is the number of $b_1 \times b_2$-bursts with "1" at position (u_1, u_2), and whose "1"s are confined to the positions of the elements already placed in the array. Consider two $b_1 \times b_2$-bursts B_1 and B_2

* *The bound of Campopiano is more refined than the one dimensional version of Theorem 2. It can be easily seen from the proof how the bound derived here can be improved. However, this improvement will not refine the upper bound on the excess redundancy derived in Theorem 3.*

satisfying these conditions, and let J_1 and J_2 denote the set of positions of their "1"s, respectively. It follows, from the order in which the elements are placed in the array, that the burst B_1 starts at position (u_1, j_1) for some $j_1 \leq u_2$, such that $(u, j) \notin J_1$ for $j_1 \leq j < u_2$. Similarly, the burst B_2 starts at position (u_1, j_2) for some $j_2 \leq u_2$ such that $(u_1, j) \notin J_2$ for $j_2 \leq j < u_2$. If B_1 and B_2 share the same pattern, then there exists a pair of integers (l_1, l_2) such that $(i, j) \in J_1$ if, and only if, $(i + l_1, j + l_2) \in J_2$. But from the characterization of the starting positions of B_1 and B_2, it follows that $l_1 = l_2 = 0$. Thus, $j_1 = j_2$, which implies $B_1 = B_2$. This proves the bound $N(b_1, b_2)$ for the number of distinct linear combinations of collections of elements already placed in the array whose positions are confined to a $b_1 \times b_2$-block containing the position (u_1, u_2). Hence, if $2^r \geq n_1 n_2 N^2(b_1, b_2)$, then we will succeed in finding the element $\mathbf{h}_{u_1,u_2}$. If we succeed in finding the element $\mathbf{h}_{0,0}$, then definitely we obtain a parity check matrix of a $b_1 \times b_2$-burst correcting code. ∎

The next theorem is an immediate consequence of Theorems 1 and 2.

Theorem 3. *Let $\underline{r}(b_1, b_2)$ denote the excess redundancy of the class of all $b_1 \times b_2$-burst correcting codes. Then,*

$$\log N(b_1, b_2) \leq \underline{r}(b_1, b_2) \leq 2 \log N(b_1, b_2).$$

Our main aim is to develop two-dimensional burst correcting codes whose excess redundancy is small. Before doing that, it may be illuminating to consider one-dimensional codes, whose theory is better understood.

Hamming codes are 1-burst correcting codes whose excess redundancy is 0. For 2-burst correcting codes, Abramson codes [2] have excess redundancy 1. The excess redundancies of these two classes of codes satisfy the minimum bound of Theorem 3 with equality. For b-burst correcting codes, with $b \geq 3$, the

best known class until recently in terms of excess redundancy was Fire codes [9]. The excess redundancy of this class is $(2b-1)-\log(2b-1)$. However, it has been shown recently [1] that for every positive integer $b \geq 3$, there exists a class of cyclic b-burst correcting codes whose excess redundancy is $b-1$. This class satisfies the lower bound of Theorem 3 with equality.

We return to two-dimensional burst correcting codes. The first class of such codes ever reported in the literature is due to Elspas [8]. The codes in this class are products of cyclic codes. The excess redundancy of these codes is infinite for all values of b_1 and b_2. However, it should be noted that these codes have other error correcting capabilities, in addition to correcting two-dimensional bursts.

The $\gamma\beta$-codes developed by Nomura et al. [19], are cyclic 1×1-burst correcting codes whose excess redundancy is 0, which meets the lower bound of Theorem 3 with equality. The cyclic class of two-dimensional Fire codes [12] has excess redundancy $(2b_1-1)(2b_2-1)-\log(2b_1-1)(2b_2-1)$. Apart from the codes developed in this chapter, this excess redundancy is the best known value in case $b_1 \times b_2$ greater than 1×1.

Theorem 3 implies the existence of a class of codes whose excess redundancy is lower than that of Fire codes unless if $(b_1, b_2) = (1, b), (b, 1)$, or $(2, 2)$, where $b > 1$. However, it should be noted that there is no guarantee that these codes have the nice algebraic structure of Fire codes which make them easy to encode and decode. The most important result in this chapter, which is Theorem 34, states that there exists a class of cyclic $b_1 \times b_2$-burst correcting codes, for every pair (b_1, b_2) of positive integers, whose excess redundancy is $b_1 b_2$ if b_1 and b_2 are both larger than 1, and $b_1 b_2 - 1$, if otherwise. The encoding and decoding techniques for these codes, which are treated in the appendix, are easy, since the codes are cyclic.

3. Burst Locating Codes

3.1. Definitions

A two-dimensional linear code C is said to be a $b_1 \times b_2$-*burst locating code* if no codeword is a $b_1 \times b_2$-burst or a sum of two $b_1 \times b_2$-bursts of the same pattern. Equivalently, the code C is a $b_1 \times b_2$-burst locating code if, and only if, the syndromes of the $b_1 \times b_2$-bursts sharing the same pattern, with respect to any given parity check matrix of C are nonzero and distinct. A $b_1 \times b_2$-*cyclic-burst locating code* is defined similarly where the bursts involved in the definition are cyclic.

These codes are useful in practice. If a $b_1 \times b_2$-burst locating code is used over a channel that may add to any transmitted codeword a $b_1 \times b_2$-burst, then the receiver can determine the burst position if the burst pattern is known. It is important to note that the receiver may not be able to uniquely determine the burst position without first knowing its pattern. Thus, a $b_1 \times b_2$-burst correcting code is a $b_1 \times b_2$-burst locating code, but the converse does not always hold except if $b_1 = b_2 = 1$. In other words, a $b_1 \times b_2$-burst locating code may contain a codeword which is the sum of two $b_1 \times b_2$-bursts of different patterns.

The following is an immediate consequence of the definition.

Lemma 4. *If r is the redundancy of a burst locating code of* area $n_1 \times n_2$, *then*

$$r \geq \lceil \log(n_1 n_2 + 1) \rceil.$$

In the following, we will give two different constructions of $b_1 \times b_2$-burst locating codes which will show that this bound is tight.

3.2. $\gamma\beta$-Codes

In this subsection, we will present a class of cyclic burst locating codes which is already known in the literature.

Theorem 5. *Let m_1 and m_2 be positive integers. Let α be an element of order n in $\mathbf{F}_{2^{m_1 m_2}}$ whose minimal polynomial over $\mathbf{F}_2$ is of degree $m_1 m_2$. Let n_1, n_2, and η be positive integers such that the following conditions are satisfied:*

(1) $n_1 n_2 = n$.

(2) m_1 *is the multiplicative order of 2 modulo* n_1.

(3) $\gcd(n_1, n_2) = 1$.

(4) $\gcd(\eta, 2^{m_1 m_2} - 1) = 1$.

Let $\gamma = \alpha^{n_2}$ *and* $\beta = \alpha^{\eta n_1}$. *Then,*

(i) The orders of γ *and* β *are* n_1 *and* n_2, *respectively.*

(ii) The minimal polynomial of γ *over* $\mathbf{F}_2$ *is of degree* m_1, *and the minimal polynomial of* β *over* $\mathbf{F}_{2^{m_1}}$ *is of degree* m_2.

(iii) The elements $\gamma^i \beta^j$, *for* $0 \le i < m_1$, $0 \le j < m_2$, *are linearly independent over* $\mathbf{F}_2$.

(iv) $\gamma^i \beta^j = 1$ *if, and only if,* $n_1 | i$ *and* $n_2 | j$.

(v) The matrix $[\gamma^i \beta^j]$, $0 \le i < n_1$, $0 \le j < n_2$ *is a parity check matrix of a cyclic* $m_1 \times m_2$*-burst locating code of area* $n_1 \times n_2$ *and redundancy* $m_1 m_2$.

Proof. Part (i) immediately follows from conditions (1) and (4). From condition (2) it follows that the minimal polynomial of γ over $\mathbf{F}_2$ is of degree m_1. The degree of the minimal polynomial of β over $\mathbf{F}_{2^{m_1}}$ is the least positive integer d such that $2^{m_1 d} \equiv 1 \pmod{n_2}$. Conditions (1), (2), and (3) implies that for such d, we have $2^{m_1 d} \equiv 1 \pmod{n}$. Since the minimal polynomial of α over $\mathbf{F}_2$ is of degree $m_1 m_2$, it follows that $d = m_2$. This proves (ii).

Now, suppose that

$$\sum_{j=0}^{m_2-1} \sum_{i=0}^{m_1-1} a_{i,j} \gamma^i \beta^j = 0,$$

where $a_{i,j} \in \mathbf{F}_2$. The fact that the minimal polynomial of β over $\mathbf{F}_{2^{m_1}}$ is of degree m_2 implies $\sum_{i=0}^{m_1-1} a_{i,j}\gamma^i = 0$ for all $0 \leq j < m_2$, which implies $a_{i,j} = 0$ for all $0 \leq i < m_1$ and $0 \leq j < m_2$ as the minimal polynomial of γ over $\mathbf{F}_2$ is of degree m_1. This proves (iii).

To prove (iv), note that $n_1|i$ and $n_2|j$ implies $\gamma^i\beta^j = 1$ from (i). On the other hand, if $\gamma^i\beta^j = 1$, then $\gamma^{in_2} = \beta^{-jn_2} = 1$, which gives $n_1|in_2$. This implies $n_1|i$ by (3). Hence, $\beta^j = 1$, which gives $n_2|j$.

Next, we prove (v). Let $[c_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$ be an array over $\mathbf{F}_2$. The syndrome of this array is given by

$$s = \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} c_{i,j}\gamma^i\beta^j.$$

Thus, the array $[c_{i,j}]$ is a codeword if, and only if, $c(\gamma,\beta) = 0$, where $c(x,y) = \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} c_{i,j}x^iy^j$. Hence, the code is an ideal in $\mathbf{F}_2[x,y]/(x^{n_1}+1, y^{n_2}+1)$, and thus, is cyclic. Since the code is cyclic, then to show that it is $m_1 \times m_2$-burst locating code, it suffices to prove that if

$$c(x,y) \equiv b(x,y) + x^{u_1}y^{u_2}b(x,y) \pmod{x^{n_1}+1, y^{n_2}+1}$$

is a codeword, where $b(x,y) \in \mathcal{B}_{m_1m_2}{}^*$, then $c(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1}$. Suppose $c(x,y)$ is a codeword, then $b(\gamma,\beta) + \gamma^{u_1}\beta^{u_2}b(\gamma,\beta) = 0$. But from part (iii), it follows that $b(\gamma,\beta) \neq 0$, which implies $\gamma^{u_1}\beta^{u_2} = 1$. Part (iv) gives $n_1|u_1$ and $n_2|u_2$, which implies $c(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1}$.

Thus, the code is indeed a cyclic $m_1 \times m_2$-burst locating code of area $n_1 \times n_2$ and redundancy $\leq m_1m_2$. But from (iii), it follows that the redundancy is exactly m_1m_2. ∎

---

* *Recall from the preliminaries that*

$$\mathcal{B}_{m_1m_2} = \{p(x,y) \in \mathbf{F}_2[x,y] : \deg_x p(x,y) < m_1, \deg_y p(x,y) < m_2, p(x,0) \neq 0, p(0,y) \neq 0\}.$$

A code whose construction is as given in Theorem 5 will be called a *$\gamma\beta$-code* with parameters (m_1, m_2).

This class of codes can be traced to Gordon [10] who gave the construction of a subclass of the dual codes of the codes presented here as a two-dimensional generalization of M-sequences. However, Nomura et al. [19] have extensively generalized the work of Gordon. The $\gamma\beta$-codes presented here are the duals of the codes studied in [19], which are called $\gamma\beta$-array codes.

A $\gamma\beta$-code of area $n_1 \times n_2$ and redundancy $m_1 m_2$, where $n_1 n_2 = 2^{m_1 m_2} - 1$, is said to have maximal area. It is to be noted that a $\gamma\beta$-code of maximal area satisfies the bound of Lemma 4 with equality.

In the following, we will show that for all positive integers b_1 and b_2, there exists an infinite number of $b_1 \times b_2$-burst locating codes within the class of $\gamma\beta$-codes of maximal areas. The basic argument in the proof is due to Gordon [10]. First, we state without proof the following number-theoretic result attributed to T. S. Bang. A proof of this result can be found in [7].

Lemma 6. *If $1 < m \neq 6$ is a positive integer, then $2^m - 1$ has a prime factor that does not divide $2^l - 1$ for every positive integer $l < m$.*

Theorem 7. *For every pair of positive integers (m_1, m_2) such that $6 \neq m_1 \geq m_2$, there exists a $\gamma\beta$-code of maximal area with parameters (m_1, m_2).*

Proof. Let $6 \neq m_1 \geq m_2$. From Lemma 6, it follows that there exists a prime p such that m_1 is the multiplicative order of 2 modulo p, which implies $m_1 \leq p - 1$ by Fermat's Theorem. Let $p^a || 2^{m_1} - 1$ for some positive integer a.* If $p | (2^{m_1 m_2} - 1)/(2^{m_1} - 1)$, then

$$\underbrace{2^{(m_2-1)m_1} + 2^{(m_2-2)m_1} + \cdots + 2^{m_1} + 1}_{m_2 \text{ terms}} \equiv 0 \pmod{p}.$$

* *For positive integers l_1 and l_2, $l_1^a || l_2$ means that $l_1^a | l_2$ but $l_1^{a+1} \nmid l_2$.*

Since $2^{m_1} \equiv 1 \pmod{p}$, it follows that $p|m_2$. This implies $m_2 \geq p \geq m_1 + 1$, which contradicts $m_1 \geq m_2$. Hence, $p \not| (2^{m_1 m_2} - 1)/(2^{m_1} - 1)$, which implies $p \not| (2^{m_1 m_2} - 1)/p^a$. Let $n_1 = p^a$ and $n_2 = (2^{m_1 m_2} - 1)/p^a$. Then, conditions (1), (2) and (3) of Theorem 5 are satisfied, and hence there is a $\gamma\beta$-code with parameters (m_1, m_2). ∎

Corollary 8. *If b_1, b_2, and n are positive integers, then there exists a $b_1 \times b_2$-burst locating code which is a $\gamma\beta$-code of maximal area greater than $n \times n$ with parameters (m_1, m_2) for all sufficiently large m_1 and m_2.*

Proof. If $b_1 \times b_2$ is less or equal to $m_1 \times m_2$, then an $m_1 \times m_2$-burst locating code is a $b_1 \times b_2$-burst locating code. The corollary now follows from Theorem 7 and conditions (1) and (2) of Theorem 5. ∎

Some practical applications may require the areas of the burst locating codes to be squares or close to squares. In the construction given in Theorem 5, it follows that $n_1 \leq 2^{m_1} - 1$ and

$$n_2 = \frac{2^{m_1 m_2} - 1}{n_1} \geq \frac{2^{m_1 m_2} - 1}{2^{m_1} - 1} \geq 2^{m_1(m_2 - 1)}.$$

Thus, if n_1 and n_2 are required to be large and close in value, then m_2 is restricted to be less or equal to 2. But this may restrict the $\gamma\beta$-code to be a $b_1 \times b_2$-burst locating code with $b_2 = 1$ or 2 only.

In the following, we construct $b_1 \times b_2$-burst locating codes of square areas for all positive integers b_1 and b_2.

3.3. $\alpha\beta$-Codes

$\alpha\beta$-codes are cyclic burst locating codes that have square areas. The following theorem gives the structure of these codes.

Theorem 9. *Let b_1, b_2, t_1, t_2, and m be positive integers, $m > 1$. Let α and β be primitive elements in $\mathbf{F}_{2^m}$, not necessarily distinct. Suppose that the*

following conditions hold:

(1) The elements α^{it_1+j}, $0 \le i < b_1$, $0 \le j < b_2$, *are linearly independent over* $\mathbf{F}_2$.

(2) The elements β^{i+jt_2}, $0 \le i < b_1$, $0 \le j < b_2$, *are linearly independent over* $\mathbf{F}_2$.

(3) $\gcd(t_1t_2 - 1, 2^m - 1) = 1$.

Then the code with parity check matrix $[\mathbf{h}_{i,j}]$, $0 \le i, j < 2^m - 1$, *given by* $\mathbf{h}_{i,j} = (\alpha^{it_1+j}, \beta^{i+jt_2})$, *is a cyclic* $b_1 \times b_2$*-burst locating code of area* $2^m - 1 \times 2^m - 1$, *and redundancy* $2m$.

Proof. Let $n = 2^m - 1$. An array $[c_{i,j}]$, $0 \le i, j < n$, over $\mathbf{F}_2$ is a codeword if, and only if, its syndrome is zero, i.e., if, and only if,

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} c_{i,j}\alpha^{it_1+j} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} c_{i,j}\beta^{i+jt_2} = 0.$$

Let $c(x,y) = \sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1} c_{i,j}x^iy^j$, then it follows that $c(x,y)$ is a codeword if and only if $c(\alpha^{t_1}, \alpha) = c(\beta, \beta^{t_2}) = 0$. Thus, the code is an ideal in $\mathbf{F}_2[x,y]/(x^n + 1, y^n + 1)$, and hence is cyclic. To prove that the code is a $b_1 \times b_2$-burst locating code, it suffices to show that if

$$c(x,y) \equiv b(x,y) + x^{u_1}y^{u_2}b(x,y) \pmod{x^n + 1, y^n + 1},$$

is a codeword, where $b(x,y) \in \mathcal{B}_{b_1,b_2}$, then $c(x,y) \equiv 0 \pmod{x^n + 1, y^n + 1}$. Suppose $c(x,y)$ is a codeword, then

$$b(\alpha^{t_1}, \alpha) + \alpha^{u_1t_1+u_2}b(\alpha^{t_1}, \alpha) = b(\beta, \beta^{t_2}) + \beta^{u_1+u_2t_2}b(\beta, \beta^{t_2}) = 0.$$

From conditions (1) and (2), it follows that $b(\alpha^{t_1}, \alpha)$ and $b(\beta, \beta^{t_2})$ are nonzero, which implies

$$u_1t_1 + u_2 \equiv u_1 + u_2t_2 \equiv 0 \pmod{n}.$$

From condition (3) it follows that $u_1 \equiv u_2 \equiv 0 \pmod{n}$. Hence, $c(x,y) \equiv 0 \pmod{x^n + 1, y^n + 1}$. Thus, the code is a $b_1 \times b_2$-burst locating code of area

$2^m - 1 \times 2^m - 1$, and redundancy $2m$ at most. But by Lemma 4, it follows that the redundancy is exactly $2m$. ∎

A code whose construction is as given in Theorem 9 will be called an *$\alpha\beta$-code*.

For a given redundancy, $\alpha\beta$-codes are inferior to $\gamma\beta$-codes in the sense that the former have smaller areas.* However, as can be seen from Lemma 4, for the same redundancy, $\alpha\beta$-codes have the largest possible areas among all burst locating codes of square areas.

Conditions (1) and (2) of Theorem 9 may be tedious to check. In the following, we will give a systematic technique to satisfy these conditions if m is large with respect to b_1 and b_2.

Lemma 10. *Let b_1, b_2, t_1, t_2, and m be positive integers such that*

(i) $t_1 \geq b_2$ *and* $m \geq (b_1 - 1)t_1 + b_2$.

(ii) $t_2 \geq b_1$ *and* $m \geq (b_2 - 1)t_2 + b_1$.

Then, the following holds for any primitive elements α and β in $\mathbf{F}_{2^m}$:

(1) The elements α^{it_1+j}, $0 \leq i < b_1$, $0 \leq j < b_2$, are linearly independent over $\mathbf{F}_2$.

(2) The elements β^{it_2+j}, $0 \leq i < b_1$, $0 \leq j < b_2$, are linearly independent over $\mathbf{F}_2$.

Proof. It suffices, by symmetry, to prove that condition (i) implies (1). From (i), it follows that the numbers $it_1 + j$, where $0 \leq i < b_1$, $0 \leq j < b_2$, are distinct and lie between 0 and $m - 1$. Since the minimal polynomial of α has degree m, then condition (1) holds. ∎

The following is an immediate corollary of the previous lemma.

Corollary 11. *If b_1 and b_2 are positive integers, then there exists a $b_1 \times b_2$-burst locating code which is an $\alpha\beta$-code of area $2^m - 1 \times 2^m - 1$, for all sufficiently*

---

* *Here "area" does not mean the pair $n_1 \times n_2$, but rather the product $n_1 n_2$.*

large m.

4. BIL-Codes

Suppose that C_I is a $b_1 \times b_2$-burst identification code of area $n_1 \times n_2$. Then C_I has no nonzero codeword which is a $b_1 \times b_2$-burst or a sum of two $b_1 \times b_2$-bursts of different patterns. On the other hand, suppose that C_L is a $b_1 \times b_2$-burst locating code of the same area $n_1 \times n_2$. Then, C_L has no nonzero codeword which is a $b_1 \times b_2$-burst or a sum of two $b_1 \times b_2$-bursts of the same pattern. Hence, the subspace $C=C_I \cap C_L$ is a code of area $n_1 \times n_2$ which has no nonzero codeword which is a $b_1 \times b_2$-burst or a sum of two $b_1 \times b_2$-bursts. In other words, C is a $b_1 \times b_2$-burst correcting code. Let r_C denote the redundancy of C, and define r_{C_I}, r_{C_L}, and $r_{C_I+C_L}$ similarly. Then, we have

$$r_C = r_{C_I} + r_{C_L} - r_{C_I+C_L}.$$

From a practical point of view, the problem of constructing a burst identification code and a burst locating code, as two separate problems is often much easier than the problem of constructing a burst correcting code directly. On the other hand, any $b_1 \times b_2$-burst correcting code is a subspace of a $b_1 \times b_2$-burst identification code and a $b_1 \times b_2$-burst locating code, just by considering these two codes to be the same as the burst correcting code itself. Practically, if we want to have a simple construction technique for a burst correcting code, then we may only try to find a burst identification code C_I and a burst locating code C_L whose redundancies are as small as possible. In other words, in a simple construction technique, we may not deliberately consider minimizing $r_{C_I} + r_{C_L} - r_{C_I+C_L}$, but rather $r_{C_I} + r_{C_L}$. This is the basic motivation of the definition of the class of BIL-codes which follows.

Let C_I and C_L be $b_1 \times b_2$-burst identification and locating codes of the same area, respectively. Let $r_{C_I} + r_{C_L}$ denote the sum of their redundancies. Then, a

code of redundancy $r_{C_I} + r_{C_L}$, which is a subspace of $C_I \cap C_L$ is said to be a $b_1 \times b_2$-*BIL-code*. From the previous discussion, a $b_1 \times b_2$-BIL-code is a $b_1 \times b_2$-burst correcting code.

By Theorem 8, chapter III, the minimum redundancy $r(b_1, b_2)$ required to construct $b_1 \times b_2$-burst identification codes of arbitrarily large areas is bounded by

$$2b_1b_2 - 2 \leq r(b_1, b_2) \leq 2b_1b_2.$$

In Theorem 7, chapter III, we gave an explicit construction of $b_1 \times b_2$-burst identification codes of arbitrarily large areas whose redundancies are equal to $2b_1b_2$.

On the other hand, in subsection 3.2 we presented for every $m_2 \leq m_1 \neq 6$ such that $b_1 \leq m_1$ and $b_2 \leq m_2$, an explicit construction of a $b_1 \times b_2$-burst locating code which is a $\gamma\beta$-code of redundancy m_1m_2 and area $n_1 \times n_2$, where $n_1n_2 = 2^{m_1m_2} - 1$. Also, we have exhibited in subsection 3.3 an explicit construction of a $b_1 \times b_2$-burst locating code which is an $\alpha\beta$-code of redundancy $2m$ and area $2^m - 1 \times 2^m - 1$ if m is a sufficiently large integer. Using either construction, it follows that the excess redundancy $\underline{r}_{\text{BIL}}(b_1, b_2)$ of the class of $b_1 \times b_2$-BIL-codes is bounded by $r(b_1, b_2)$. On the other hand, Lemma 4 implies that $\underline{r}_{\text{BIL}}(b_1, b_2) = r(b_1, b_2)$. Thus, we have proved the following theorem.

Theorem 12. *The excess redundancy $\underline{r}_{\text{BIL}}(b_1, b_2)$ of the class of $b_1 \times b_2$-BIL-codes satisfies*

$$2b_1b_2 - 2 \leq \underline{r}_{\text{BIL}}(b_1, b_2) \leq 2b_1b_2.$$

In fact, $\underline{r}_{\text{BIL}}(b_1, b_2) = r(b_1, b_2)$ which is the minimum redundancy required to construct a $b_1 \times b_2$-burst identification code of arbitrarily large area.

5. Fire-ish Codes

One-dimensional Fire codes [9] form a well known class of one-dimensional burst correcting codes. The excess redundancy of this class of codes is given by

$$r_{\text{Fire}}(b) = (2b-1) - \log(2b-1),$$

which represents the smallest excess redundancy known among all b-burst correcting codes with $b \geq 3$ until very recently. For example, the excess redundancy of one-dimensional BIL-codes is given by Theorem 11 of chapter III as $r_{\text{BIL}}(b) = 2b-2$ for $b > 1$ and $r_{\text{BIL}}(1) = 1$. Hence, one-dimensional Fire codes are superior to BIL-codes in terms of excess redundancy for all values of $b \geq 2$.

Imai [12] has generalized one-dimensional Fire codes to two-dimensional cyclic burst correcting codes which are known as two-dimensional Fire codes. We will call this class of codes Fire codes for simplicity since one-dimensional Fire codes can be considered as a special case of two-dimensional Fire codes.

In this section, we will develop a class of two-dimensional burst correcting codes, called Fire-ish codes, which contains Fire codes as a subclass. It should be mentioned that Fire-ish codes, in contrast to Fire codes, may be noncyclic. Hence, the two-dimensional error trapping decoding technique developed by Imai [12] may not be applicable to Fire-ish codes in general. However, in the appendix, we will show that some Fire-ish codes that are superior to Fire codes can be decoded by essentially the same technique after some minor modifications. The construction of Fire-ish codes is explained in the following theorem.

Theorem 13. *Let* $H' = [\mathbf{h}'_{i,j}]$, $0 \leq i < n'_1$, $0 \leq j < n'_2$ *be a parity check matrix of a* $b_1 \times b_2$*-cyclic-burst locating code* C' *of area* $n'_1 \times n'_2$. *Let* $H'' = [\mathbf{h}''_{i,j}]$, $0 \leq i < n''_1$, $0 \leq j < n''_2$ *be a parity check matrix of a* $b_1 \times b_2$*-cyclic-burst correcting code* C'' *of area* $n''_1 \times n''_2$. *Suppose that* $n'_1 \times n'_2$ *and* $n''_1 \times n''_2$ *are greater or equal to* $2b_1 - 1 \times 2b_2 - 1$. *Then, the code* C *whose parity check matrix*

$H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, *where* $n_1 = \text{lcm}(n_1', n_1'')$, $n_2 = \text{lcm}(n_2', n_2'')$, *and* $\mathbf{h}_{i,j} = (\mathbf{h}'_{i \bmod n_1', j \bmod n_2'}, \mathbf{h}''_{i \bmod n_1'', j \bmod n_2''})$, *is a* $b_1 \times b_2$*-cyclic-burst correcting code of area* $n_1 \times n_2$.

Proof. Suppose that

$$c(x,y) \equiv x^{u_1'} y^{u_2'} b'(x,y) + x^{u_1''} y^{u_2''} b''(x,y) \pmod{x^{n_1}+1, y^{n_2}+1}$$

is a codeword in $\mathcal{C}$, where $0 \leq u_1', u_2'' < n_1$, $0 \leq u_2', u_2'' < n_2$, and $b'(x,y)$, $b''(x,y) \in \mathcal{B}_{b_1,b_2}$. Then, it follows from the construction of $\mathcal{C}$ that

$$x^{u_1' \bmod n_1''} y^{u_2' \bmod n_2''} b'(x,y) + x^{u_1'' \bmod n_1''} y^{u_2'' \bmod n_2''} b''(x,y) \pmod{x^{n_1''}+1, y^{n_2''}+1}$$

is a codeword in $\mathcal{C}''$. As $\mathcal{C}''$ is a $b_1 \times b_2$-cyclic-burst correcting code of area $n_1'' \times n_2''$ which is greater or equal to $2b_1 - 1 \times 2b_2 - 1$, it follows, by Lemma 1 in the preliminaries, that $b'(x,y) = b''(x,y)$, $n_1''|(u_1' - u_1'')$, and $n_2''|(u_2' - u_2'')$. Hence,

$$c(x,y) \equiv x^{u_1'} y^{u_2'} b'(x,y) + x^{u_1''} y^{u_2''} b'(x,y) \pmod{x^{n_1}+1, y^{n_2}+1}.$$

But, from the construction of $\mathcal{C}$, it follows that

$$x^{u_1' \bmod n_1'} y^{u_2' \bmod n_2'} b'(x,y) + x^{u_1'' \bmod n_1'} y^{u_2'' \bmod n_2'} b'(x,y) \pmod{x^{n_1'}+1, y^{n_2'}+1}$$

is a codeword in $\mathcal{C}'$. As $\mathcal{C}'$ is a $b_1 \times b_2$-cyclic burst locating code of area $n_1' \times n_2'$ which is greater or equal to $2b_1 - 1 \times 2b_2 - 1$, it follows that $n_1'|(u_1' - u_1'')$ and $n_2''|(u_1' - u_2'')$. Hence, $n_1|(u_1' - u_1'')$ and $n_2|(u_1'' - u_2'')$, which implies that $c(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1}$. ∎

If $r_{\mathcal{C}'}$, $r_{\mathcal{C}''}$, and $r_{\mathcal{C}}$ denote the redundancies of $\mathcal{C}'$, $\mathcal{C}''$, and $\mathcal{C}$, respectively, then $r_{\mathcal{C}} \leq r_{\mathcal{C}'} + r_{\mathcal{C}''}$.

Any $b_1 \times b_2$-cyclic-burst correcting code can be considered to have the construction of the previous theorem, just by taking $\mathcal{C}'$ and $\mathcal{C}''$ to be the same as the code itself. However, the problem of constructing a cyclic-burst correcting

code of small area and a cyclic-burst locating code, as two separate problems, is much easier than constructing a burst correcting code of large area. Practically, if we want a simple construction technique using Theorem 13, we may restrict ourselves to minimizing $r_{C'} + r_{C''}$, rather than r_C. With this motivation in mind, a subcode of redundancy $r_{C'} + r_{C''}$ of the code C whose construction is given in Theorem 13 is called a ***Fire-ish code.***

Note that there exists an infinite class of $\gamma\beta$-codes which are cyclic $b_1 \times b_2$-burst locating codes such that each code in this class has maximal area $n_1' \times n_2'$, for some n_1' and n_2' such that $n_1'n_2'$ is relatively prime to any given positive integer $n_1''n_2''$. This follows from Corollary 8 by properly choosing m_1m_2. Also, Corollary 11 implies that there exists an infinite number of $\alpha\beta$-codes, which are cyclic $b_1 \times b_2$-burst locating codes of area $2^m - 1 \times 2^m - 1$, for some $2^m - 1$ which is relatively prime to any given positive integer $n_1''n_2''$. Hence, $\gamma\beta$-codes and $\alpha\beta$-codes, along with any $b_1 \times b_2$-cyclic-burst correcting code C'' of redundancy $r_{C''}$ and area $n_1'' \times n_2''$ greater or equal to $2b_1 - 1 \times 2b_2 - 1$, can be used to construct an infinite class of $b_1 \times b_2$-burst correcting codes of arbitrarily large areas, whose excess redundancy is given by

$$\underline{r}_{\text{Fire-ish}(C'')}(b_1, b_2) = r_{C''} - \log n_1''n_2''. \tag{1}$$

By Lemma 4, there are no $b_1 \times b_2$-cyclic burst locating codes which can be used along with the code C'' to construct a class of Fire-ish codes with smaller excess redundancy.

Fire codes are the subclass of Fire-ish codes whose code C' is a $\gamma\beta$-code, and whose code C'' is the code of area $2b_1 - 1 \times 2b_2 - 1$, which contains the zero codeword only. Obviously, this is the most simple construction for the code C''. From (1), it follows that the excess redundancy of Fire codes is given by

$$\underline{r}_{\text{Fire}}(b_1, b_2) = (2b_1 - 1)(2b_2 - 1) - \log(2b_1 - 1)(2b_2 - 1).$$

The main problem in constructing good Fire-ish codes is to find a code C'' with small $r_{C''} - \log(n_1'' n_2'')$. In Lemma 9 of chapter III, we have actually constructed a $b_1 \times b_2$-cyclic-burst correcting code of area $2b_1 \times 2b_2$ and redundancy $2b_1 b_2 + b_2$. The parity check matrix of this code is the building block $\tilde{H}$ defined in the lemma. Let $\tilde{C}$ denote this code. Using $\tilde{C}$ to construct Fire-ish codes, we get a subclass of Fire-ish codes whose excess redundancy is given by

$$\underline{r}_{\text{Fire-ish}(\tilde{C})}(b_1, b_2) = 2b_1 b_2 + \min\{b_1, b_2\} - \log(4b_1 b_2),$$

where we have used $\min\{b_1, b_2\}$ instead of b_2 because of symmetry. Note that $\underline{r}_{\text{Fire-ish}(\tilde{C})}(b_1, b_2)$ is less than $\underline{r}_{\text{Fire}}(b_1, b_2)$ unless if b_1 or b_2 is 1, or $(b_1, b_2) = (2, 2)$.

6. Cyclic Burst Correcting Codes of Minimum Redundancy

6.1. Definitions

From Theorem 1, it follows that the redundancy r and the area $n_1 \times n_2$ of a cyclic $b_1 \times b_2$-burst correcting code satisfy

$$r \geq \left\lceil \log\left(1 + n_1 n_2 N(b_1, b_2)\right) \right\rceil.$$

A cyclic $b_1 \times b_2$-burst correcting code of area $n_1 \times n_2$, where $n_1 n_2 = 2^m - 1$ and redundancy $r = m + \lceil \log N(b_1, b_2) \rceil$ for some positive integer m, is said to be a *cyclic $b_1 \times b_2$-burst correcting code of minimum redundancy*. In this section, we will prove that for all positive integers b_1 and b_2, there exist cyclic $b_1 \times b_2$-burst correcting codes of minimum redundancy of arbitrarily large areas. The excess redundancy of this class of codes is $\lceil \log N(b_1, b_2) \rceil$. The reason for calling this class of codes minimum redundancy codes is explained in the following theorem.

Theorem 14. *Let C be a $b_1 \times b_2$-burst correcting code of redundancy r and area $n_1 \times n_2$, where $n_1 n_2 = 2^m - 1$ for some positive integer m. Then*

$$r \geq m + \lceil \log N(b_1, n_2) \rceil,$$

if n_1 and n_2 are sufficiently large. If C is a $b_1 \times b_2$-cyclic-burst correcting code, then the above inequality holds for every $n_1 \times n_2$ greater or equal to $2b_1 - 1 \times 2b_2 - 1$.

Proof. The first statement follows from the definition of excess redundancy. Indeed, if $r \leq m + \lceil \log N(b_1, b_2) \rceil - 1$ for an infinite sequence of areas $n_1 \times n_2$, where n_1 and n_2 are increasing, then there exists a class of $b_1 \times b_2$-burst correcting codes whose excess redundancy is $\lceil \log N(b_1, b_2) \rceil - 1$, which contradicts Theorem 3. Now, we prove the second statement. From Theorem 1, we have

$$2^r \geq 1 + (2^m - 1)N(b_1, b_2).$$

So, it suffices to show that

$$2^{m+K-1} < 1 + (2^m - 1)N(b_1, b_2), \tag{2}$$

where $K = \lceil \log N(b_1, b_2) \rceil$. If b_1 or b_2 is 1, then from Theorem 2 in the preliminaries, it follows that (2) holds. So, assume that b_1 and $b_2 > 1$, and (2) does not hold. Then,

$$\begin{aligned} N(b_1, b_2) &\leq \frac{2^{m+K-1} - 1}{2^m - 1} \\ &= 2^{K-1} + \frac{2^{K-1} - 1}{2^m - 1} \\ &\leq 2^{K-1} + \frac{2^{K-1} - 1}{b_1 b_2} \end{aligned}$$

Using Theorem 2 in the preliminaries, which implies $K = b_1 b_2$, we get

$$(2^{b_1 - 1} - 1)(2^{b_2 - 1} - 1)2^{(b_1 - 1)(b_2 - 1)} \leq \frac{2^{b_1 b_2 - 1} - 1}{b_1 b_2}.$$

This inequality does not hold if $b_1, b_2 \geq 2$. ∎

Theorem 14 implies that there is no $b_1 \times b_2$-burst correcting code with smaller redundancy than that of a cyclic $b_1 \times b_2$-burst correcting code of minimum redundancy and the same area, if the area is sufficiently large. They may exist, however, a $b_1 \times b_2$-burst correcting code which has the same redundancy, but

larger area than that of a cyclic $b_1 \times b_2$-burst correcting code of minimum redundancy. This is the reason why the excess redundancy of cyclic $b_1 \times b_2$-burst correcting codes of minimum redundancy differs from the lower bound of Theorem 3.

Before giving a scheme to construct cyclic $b_1 \times b_2$-burst correcting codes of minimum redundancy, we need the following definitions. Define the set of polynomials $\mathcal{F}_{b_1,b_2}$ as

$$\mathcal{F}_{b_1,b_2} = \{p(x,y) \in \mathcal{B}_{b_1,b_2} : p(x,y) \text{ irreducible over } \mathbf{F}_2\}.$$

Let γ and β be elements in $\mathbf{F}_q$, where q is a power of 2. Let ν be a primitive element in $\mathbf{F}_q$, and $p(x,y)$ be a nonzero polynomial over $\mathbf{F}_q$. We define the index $\text{ind}_\nu(p(\gamma,\beta))$ of $p(\gamma,\beta)$ as

$$p(\gamma,\beta) = \nu^{\text{ind}_\nu(p(\gamma,\beta))},$$

where $\text{ind}_\nu(p(\gamma,\beta))$ is reduced modulo $q-1$. It follows that if α is a primitive element in $\mathbf{F}_q$, then

$$\text{ind}_\alpha(p(\gamma,\beta)) \equiv \eta \text{ind}_\nu(p(\gamma,\beta)) \pmod{q-1},$$

for some η relatively prime to $q-1$. It is to be noted that $\mathbf{F}_2[x,y]$ is a unique factorization domain [4; chapter I],[14; chapter V], and hence every $b(x,y) \in \mathcal{B}_{b_1,b_2}$ can be written uniquely as a product $\prod_{i=1}^{k} f_i(x,y)$ of irreducibles in $\mathcal{F}_{b_1,b_2}$. This implies that

$$\text{ind}_\nu b(\gamma,\beta) \equiv \sum_{i=1}^{k} \text{ind}_\nu(f_i(\gamma,\beta)) \pmod{q-1}.$$

A cyclic code $\mathcal{C}$ of redundancy $\lceil \log N(b_1,b_2) \rceil$ such that if $b'(x,y), b''(x,y) \in \mathcal{B}_{b_1,b_2}$, and $b'(x,y) + b''(x,y) \in \mathcal{C}$, then $b'(x,y) + b''(x,y) = 0$, is said to be a $b_1 \times b_2$-*code*. Thus, a $b_1 \times b_2$-code has redundancy $b_1 b_2 - 1$ if b_1 or b_2 is 1, and

$b_1 b_2$, otherwise.

Theorem 15. *For every pair of positive integers (b_1, b_2), there exists a $b_1 \times b_2$-code of area $n_1' \times n_2'$ for some odd, and relatively prime n_1' and n_2'.*

Proof. If $b_2 = 1$, let $g(x) \in \mathbf{F}_2[x]$ be a squarefree polynomial of degree $b_1 - 1$ which is not divisible by x. Let C be the one-dimensional cyclic code generated by $g(x)$. Suppose $b'(x,0) + b''(x,0) \in C$, where $b'(x,0), b''(x,0) \in \mathcal{B}_{b_1,1}$. Then,

$$b'(x,0) + b''(x,0) \equiv 0 \pmod{g(x)}.$$

As $\deg_x b'(x,0), \deg_x b''(x,0) < b_1$ and $\deg_x g(x) = b_1 - 1$, it follows that $b'(x,0) + b''(x,0) = g(x)$. But from the definition of $\mathcal{B}_{b_1,1}$, we have $b'(0,0) = b''(0,0) = 1$, which contradicts $g(0) = 1$ if $b'(x,0) \neq b''(x,0)$. Hence, C is a $b_1 \times 1$-code. For reasons of symmetry, there exists a $1 \times b_2$-code for every positive integer b_2.

Now we consider the case $b_1, b_2 > 1$. In this case, a $\gamma\beta$-code with parameters (b_1, b_2), as described in Theorem 5, is a $b_1 \times b_2$-code which satisfies the required conditions. From Theorem 7, it suffices, by symmetry, to prove that there exist $\gamma\beta$-codes with parameters $(b_1, b_2) = (6,2)$, $(6,3)$, $(6,4)$, $(6,5)$, and $(6,6)$. In case $b_1 = 6$ and $b_2 = 2, 4$, or 5, let α be a primitive element in $\mathbf{F}_{2^{6b_2}}$, $n_1' = 63$, and $n_2' = (2^{6b_2} - 1)/63$. In case $b_1 = 6$ and $b_2 = 3$ or 6, let α be an element in $\mathbf{F}_{2^{6b_2}}$ of order $(2^{6b_2} - 1)/3$, $n_1' = 63$, and $n_2' = (2^{6b_2} - 1)/189$. In all cases, the minimal polynomial of α over $\mathbf{F}_2$ is of degree $b_1 b_2$, and conditions (1),(2), and (3) of Theorem 5 are satisfied, and thus there exist $b_1 \times b_2$-codes for $b_1 = 6$ and b_2 =2,3,4,5, and 6. ∎

The following theorem gives a technique to construct cyclic $b_1 \times b_2$-burst correcting codes of minimum redundancy if certain conditions hold.

Theorem 16. *Let $m_1 \geq b_1$ and $m_2 \geq b_2$ be positive integers. Let $C_{\gamma\beta}$ be a $\gamma\beta$-code of maximal area $n_1 \times n_2$ and parameters (m_1, m_2). Let $[\gamma^i \beta^j]$, $0 \leq i < n_1$, $0 \leq j < n_2$, be its parity check matrix. Let n_1' and n_2' be positive*

integers. Suppose that the following conditions hold:

(1) $\gcd(n_1', n_2') = 1$.

(2) $n_1' | n_1$ *and* $n_2' | n_2$.

(3) $n_1' n_2' | \mathrm{ind}_\nu(f(\gamma, \beta))$ *for all* $f \in \mathcal{F}_{b_1, b_2}$, *where* ν *is some primitive element in* $\mathbf{F}_q$, *and* $q = 2^{m_1 m_2}$.

Suppose that $[\mathbf{h}_{i,j}]$, $0 \le i < n_1'$, $0 \le j < n_2'$ *is a parity check matrix of a* $b_1 \times b_2$*-code of area* $n_1' \times n_2'$. *Then,* $H = [(\gamma^i \beta^j, \mathbf{h}_{i \bmod n_1', j \bmod n_2'})]$, $0 \le i < n_1$, $0 \le j < n_2$, *is a parity check matrix of a cyclic* $b_1 \times b_2$*-burst correcting code of minimum redundancy.*

Proof. Let $\mathcal{C}$ be the code whose parity check matrix is H. Clearly, $\mathcal{C}$ is cyclic. Hence, it suffices to show that if

$$c(x, y) \equiv b'(x, y) + x^{u_1} y^{u_2} b''(x, y) \pmod{x^{n_1} + 1, y^{n_2} + 1}$$

is a codeword in $\mathcal{C}$, where $b'(x, y), b''(x, y) \in \mathcal{B}_{b_1, b_2}$, then $c(x, y) \equiv 0 \pmod{x^{n_1} + 1, y^{n_2} + 1}$. Suppose $c(x, y)$ is a codeword in $\mathcal{C}$, then

$$b'(\gamma, \beta) + \gamma^{u_1} \beta^{u_2} b''(\gamma, \beta) = 0. \tag{3}$$

This implies that

$$\alpha^{\mathrm{ind}_\alpha(b'(\gamma,\beta))} + \alpha^{\mathrm{ind}_\alpha(b''(\gamma,\beta)) + u_1 n_2 + u_2 n_1 \eta} = 0,$$

where α is a primitive element in $\mathbf{F}_q$ such that $\gamma = \alpha^{n_2}$ and $\beta = \alpha^{n_1 \eta}$ for some η relatively prime to $q - 1 = n_1 n_2$. Thus, from condition (2) we have

$$\mathrm{ind}_\alpha(b'(\gamma, \beta)) - \mathrm{ind}_\alpha(b''(\gamma, \beta)) \equiv u_1 n_2 + u_2 n_1 \eta \pmod{n_1' n_2'}.$$

But $n_1' n_2' | \mathrm{ind}_\alpha(b'(\gamma, \beta))$ and $n_1' n_2' | \mathrm{ind}_\alpha(b''(\gamma, \beta))$ from condition (3). Hence,

$$u_1 n_2 + u_2 n_1 \eta \equiv 0 \pmod{n_1' n_2'}.$$

From conditions (1) and (2), it follows that $n_1' | u_1$. Similarly, $n_2' | u_2$ since

$\gcd(\eta, n_1 n_2) = 1$. But from the construction of the code C, it follows that

$$b'(x,y) + b''(x,y) \equiv b'(x,y) + x^{u_1 \bmod n_1'} y^{u_2 \bmod n_2'} b''(x,y) \pmod{x^{n_1'}+1, y^{n_2'}+1}$$

is a codeword in the $b_1 \times b_2$-code. Hence, $b'(x,y) = b''(x,y)$. From (3), we have $\gamma^{u_1}\beta^{u_2} = 1$. But then Theorem 5 implies $n_1|u_1$ and $n_2|u_2$. Hence, $c(x,y) \equiv 0 \pmod{x^{n_1}+1, y^{n_2}+1}$. Thus the code C is indeed a $b_1 \times b_2$-burst correcting code. Its redundancy is obviously $m_1 m_2 + \lceil \log N(b_1, b_2) \rceil$, and hence C is of minimum redundancy. ∎

Our main concern now is to use Theorem 16 to show that for all positive integers b_1 and b_2 there exist cyclic $b_1 \times b_2$-burst correcting codes of arbitrarily large areas. The main problem is that this theorem is burdened by so many conditions that need to be satisfied. From the construction of $\gamma\beta$-codes given in Theorem 5, the following lemma lists all the conditions needed in Theorem 16.

Lemma 17. *Let b_1 and b_2 be positive integers. Let n_1' and n_2' be two positive odd integers which are relatively prime. Suppose that $m_1 \geq b_1$ and $m_2 \geq b_2$ be some integers such that the following conditions are satisfied for some positive integers n_1 and n_2:*

(1) $n_1 n_2 = 2^{m_1 m_2} - 1$.

(2) m_1 *is the multiplicative order of 2 modulo* n_1.

(3) $\gcd(n_1, n_2) = 1$.

(4) $n_1'|n_1$.

(5) $n_2'|n_2$.

(6) $n_1' n_2' | \mathrm{ind}_\nu(f(\gamma,\beta))$ *for all* $f \in \mathcal{F}_{b_1,b_2}$, *where* $\gamma, \beta \in \mathbf{F}_q$ *are of orders* n_1 *and* n_2, *respectively,* $q = 2^{m_1 m_2}$, *and* ν *is some primitive element in* $\mathbf{F}_q$.

Then, there exists a cyclic $b_1 \times b_2$-burst correcting code of minimum redundancy whose area is $n_1 \times n_2$.

In the following subsection, we will be concerned with condition (6).

6.2. Applying Weil's Estimates of Character Sums

In this section, we will make use of Weil's estimates of character sums with polynomial arguments to prove the existence of cyclic burst correcting codes of minimum redundancy. Many preliminary results are needed before making use of Weil's estimates. We start by giving a brief survey of characters of the multiplicative groups of finite fields.

Let G be a finite abelian group of order n. A character χ is a homomorphism from G into the multiplicative group of $\mathbf{C}$, the field of complex numbers. Hence, for $g \in G$, $\chi(g)$ is an nth root of unity. If G is cyclic, $G = \langle g \rangle$, then

$$\chi^{(l)}(g^k) = e^{2\pi ilk/n},$$

defines a character $\chi^{(l)}$ of G, where $i = \sqrt{-1}$ and $0 \leq l, k < n$ are integers. Moreover, if χ is a character of G, then $\chi = \chi^{(l)}$ for some $0 \leq l < n$. By identifying $\chi^{(l)}$ with l, it follows that the characters of the cyclic group G forms a cyclic group of order n. The identity of the character group is $\chi^{(0)}$, which is called the identity character. A character χ of order j will be denoted by χ_j. Thus, χ_1 is the identity character.

In the following, let $q = 2^m$, where m is some positive integer. Let G be the multiplicative group of $\mathbf{F}_q$, denoted by $\mathbf{F}_q^*$. This group is cyclic. The characters of $\mathbf{F}_q^*$ are called multiplicative characters. If χ is a multiplicative character, we define $\chi(0) = 0$. As usual, let μ, ϕ and d denote the Möbious, Euler, and divisor functions, respectively. The following lemma is obvious.

Lemma 18. *Let $h > 1$ be a positive integer, z be an indeterminate, and $\xi \in \mathbf{C}$ be a primitive h th root of unity. Then*

$$\prod_{j=1}^{h-1} (1 - z\xi^j) = \sum_{j=0}^{h-1} z^j.$$

The following lemma appears in [6].

Lemma 19. *Let e_1 and e_2 be positive integers such that $e_1e_2 = q-1$. Let*

$$\psi(\omega) = \frac{1}{e_2}\sum_{k|e_1}\frac{\mu(k)}{k}\sum_{\chi^{ke_2}=\chi_1}\chi(\omega),$$

where $\omega \in \mathbf{F}_q^*$. Then

$$\psi(\omega) = \begin{cases} 1, & \text{if } \omega \text{ is of order } e_1; \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Let n be the order of ω. From the characterization of characters of finite cyclic groups, it follows that any character χ of $\mathbf{F}_q^*$ is $\chi^{(l)}$ for some $0 \le l < q-1$. For such character χ, we have $\chi(\omega) = e^{2\pi il/n}$, and if $\chi^{ke_2} = \chi_1$, then $q-1|lke_2$ which implies $e_1|lk$. Hence, if $k|e_1$, we get

$$\begin{aligned}\frac{1}{ke_2}\sum_{\chi^{ke_2}=\chi_1}\chi(\omega) &= \frac{1}{ke_2}\sum_{t=0}^{ke_2-1} e^{2\pi te_1/nk} \\ &= \begin{cases} 1, & \text{if } k|(e_1/n); \\ 0, & \text{otherwise.} \end{cases}\end{aligned}$$

Thus, we have

$$\psi(\omega) = \sum_{k|(e_1/n)}\mu(k) = \begin{cases} 1, & \text{if } e_1 = n; \\ 0, & \text{otherwise.} \end{cases}$$

■

In the following, $\mathcal{F} = \{f_t(x) \in \mathbf{F}_q[x] : t = 1, 2, \ldots, M\}$ is a set of pairwise relatively prime polynomials of positive degree and $f_t(0) \neq 0$ for $1 \le t \le M$.

Lemma 20. *Let $h > 1$ be an integer which divides $q-1$, and ν be a primitive element in $\mathbf{F}_q$. Let*

$$\theta(\omega) = \psi(\omega)\prod_{t=1}^{M}\sum_{j=0}^{h-1}\chi_h^j\left(f_t(\omega)\right).$$

where $\omega \in \mathbf{F}_q^*$ and $\psi(\omega)$ is as defined in Lemma 19. Then

$$\theta(\omega) = \begin{cases} h^M, & \text{if } \omega \text{ is order } e_1 \text{ and } h|\mathrm{ind}_\nu(f_t(\omega)) \text{ for all } 1 \le t \le M; \\ 0, & \text{otherwise.} \end{cases}$$

Proof. From Lemma 19, it follows that $\theta(\omega) = 0$ if the order of ω is not e_1. So, let ω be of order e_1, and apply Lemma 18 to the summation with $z = \chi_h\big(f_i(\omega)\big)$. ∎

In the following, Weil's estimates of character sums with polynomial arguments are used. We state these estimates in the following lemma [15; chapter 5],[21; chapter II].

Lemma 21. *Weil's Estimates: Let χ be a multiplicative character of order $j > 1$, and let $f \in \mathbf{F}_q[x]$ be a polynomial which is not a jth power of a polynomial. Let s be the number of distinct roots of f in its splitting field over $\mathbf{F}_q$. Then we have*

$$\Big| \sum_{\omega \in \mathbf{F}_q} \chi\big(f(\omega)\big) \Big| \leq (s-1)q^{1/2}.$$

In the following, we need to estimate character sums over $\mathbf{F}_q^*$ rather than $\mathbf{F}_q$. Write $f(x) = x^L g(x)$, for nonnegative integer L such that $g(0) \neq 0$. Suppose that $g(x)$ is a polynomial of positive degree that is not a jth power of a polynomial. Let s' denote the number of distinct roots of $g(x)$ in its splitting field over $\mathbf{F}_q$. If $L = 0$, then

$$\Big| \sum_{\omega \in \mathbf{F}_q^*} \chi\big((f(\omega)\big) \Big| \leq \Big| \sum_{\omega \in \mathbf{F}_q} \chi\big((f(\omega)\big) \Big| + 1 \leq (s'-1)q^{1/2} + 1 \leq s'q^{1/2}.$$

On the other hand, if $L \neq 0$, then $\chi\big(f(0)\big) = 0$, and we get

$$\Big| \sum_{\omega \in \mathbf{F}_q^*} \chi\big((f(\omega)\big) \Big| = \Big| \sum_{\omega \in \mathbf{F}_q} \chi\big((f(\omega)\big) \Big| \leq s'q^{1/2},$$

since $f(x)$ has in this case $s' + 1$ distinct roots. Thus, the next lemma follows.

Lemma 22. *Let χ be a multiplicative character of order $j > 1$, and let $f \in \mathbf{F}_q[x]$ be a polynomial which is not a jth power of a polynomial. Let s' be*

the number of distinct nonzero roots of f in its splitting field over $\mathbf{F}_q$. Then

$$\Big| \sum_{\omega \in \mathbf{F}_q^*} \chi\big(f(\omega)\big) \Big| \leq s' q^{1/2}.$$

Lemma 23.

$$\Big| \sum_{\omega \in \mathbf{F}_q^*} \theta(\omega) - \phi(e_1) \Big| \leq A(h, \mathcal{F}) q^{1/2} d(e_1),$$

where $\theta(\omega)$ is as defined in Lemma 20, and $A(h, \mathcal{F}) = (h-1)h^{M-1} \sum_{t=1}^{M} \deg f_t$.

Proof. From Lemma 20, we have $\theta(\omega) = \psi(\omega) + R(\omega)$, where

$$R(\omega) = \psi(\omega) \underbrace{\sum_{i_1=0}^{h-1} \cdots \sum_{i_M=0}^{h-1}}_{(i_1,\ldots,i_M) \neq (0,\ldots,0)} \chi_h^{i_1}\big(f_1(\omega)\big) \cdots \chi_h^{i_M}\big(f_M(\omega)\big). \tag{4}$$

Summing over all $\omega \in \mathbf{F}_q^*$, and using Lemma 19 along with the fact that there are exactly $\phi(e_1)$ elements in $\mathbf{F}_q^*$ of order e_1, we get

$$\sum_{\omega \in \mathbf{F}_q^*} \theta(\omega) = \phi(e_1) + \sum_{\omega \in \mathbf{F}_q^*} R(\omega).$$

Hence, the proof of the lemma depends on showing that

$$\Big| \sum_{\omega \in \mathbf{F}_q^*} R(\omega) \Big| \leq (h-1)h^{M-1} \sum_{t=1}^{M} (\deg f_t)\, q^{1/2} d(e_1). \tag{5}$$

From Lemma 19, we have by considering a typical term in the sum of the right hand side of (4),

$$\psi(\omega) \chi_h^{i_1}\big(f_1(\omega)\big) \cdots \chi_h^{i_M}\big(f_M(\omega)\big) = \frac{1}{e_2} \sum_{k \mid e_1} \frac{\mu(k)}{k} \sum_{\chi^{ke_2} = \chi_1} \chi(\omega) \chi_h^{i_1}\big(f_1(\omega)\big) \cdots \chi_h^{i_M}\big(f_M(\omega)\big). \tag{6}$$

In the inner sum $\chi = \chi_j$ for some $j \mid ke_2$. Hence,

$$\chi_j(\omega) \chi_h^{i_1}\big(f_1(\omega)\big) \cdots \chi_h^{i_M}\big(f_M(\omega)\big) = \chi_{q-1}\big(\omega^L g(\omega)\big),$$

where $L = (q-1)/j$ and

$$g(x) = \prod_{t=1}^{M} \big(f_t(x)\big)^{(q-1)i_t/h}.$$

The polynomial $g(x)$ is a polynomial of positive degree which is not a $(q-1)$-st power of a polynomial since $(i_1,\dots,i_M) \neq (0,\dots,0)$, $0 \leq i_t \leq h-1$ for $1 \leq t \leq M$ and the polynomials $f_1(x),\dots,f_M(x)$ are pairwise relatively prime. Moreover, $g(0) \neq 0$. The number of distinct roots of $g(x)$ in its splitting field is at most $\sum_{t=1}^{M} u(i_t) \deg f_t$, where $u(i)$ is defined over the nonnegative integers as $u(0) = 0$ and $u(i) = 1$ for $i \geq 1$. Hence from Lemma 22, we get

$$\Big|\sum_{\omega\in\mathbf{F}_q^*} \chi_j(\omega)\,\chi_h^{i_1}\big(f_1(\omega)\big)\dots\chi_h^{i_M}\big(f_M(\omega)\big)\Big| \leq \sum_{t=1}^{M} u(i_t)(\deg f_t)q^{1/2}.$$

Using (6), and noting that there are exactly ke_2 characters χ such that $\chi^{ke_2} = \chi_1$, we get

$$\Big|\sum_{\omega\in\mathbf{F}_q^*} \psi(\omega)\,\chi_h^{i_1}\big(f_1(\omega)\big)\dots\chi_h^{i_M}\big(f_M(\omega)\big)\Big| \leq \sum_{t=1}^{M} u(i_t)(\deg f_t)\,q^{1/2}d(e_1).$$

Using (4), we get

$$\begin{aligned}\Big|\sum_{\omega\in\mathbf{F}_q^*} R(\omega)\Big| &\leq \underbrace{\sum_{i_1=0}^{h-1}\dots\sum_{i_M=0}^{h-1}}_{(i_1,\dots,i_M)\neq(0,\dots,0)} \sum_{t=1}^{M} u(i_t)(\deg f_t)\,q^{1/2}d(e_1)\\ &= \sum_{t=1}^{M} h^{M-1}\sum_{i_t=1}^{h-1} u(i_t)(\deg f_t)\,q^{1/2}d(e_1).\end{aligned}$$

Since $u(i) = 1$ for $i > 1$, (5) is proved. ∎

The following lemma is a direct consequence of Lemma 23.

Lemma 24. *Let q be a power of 2, and let h and e_1 be positive divisors of $q-1$. Let $\mathcal{F} = \{f_1(x),\dots,f_M(x)\}$ be a set of pairwise relatively prime polynomials of positive degree in $\mathbf{F}_q[x]$ such that $f_t(0) \neq 0$ for $1 \leq t \leq M$. Let $A = h^M \sum_{t=0}^{M} \deg f_t(x)$, and suppose that $\phi(e_1) > Aq^{1/2}d(e_1)$. Then, there exists an element $\omega \in \mathbf{F}_q$ of order e_1 such that $h|\mathrm{ind}_\nu\big(f_t(\omega)\big)$ for $0 \leq t \leq M$, where ν*

is primitive in $\mathbf{F}_q$.

In the following, we will work with a set of irreducible polynomials in $\mathbf{F}_q[x, y]$ with positive degree in y. We want to ensure that given an element $\gamma \in \mathbf{F}_q$, then for any two distinct polynomials $f_1(x, y)$ and $f_2(x, y)$ in this set, the polynomials $f_1(\gamma, y)$ and $f_2(\gamma, y)$ are relatively prime.

Let $f_1(x, y)$, $f_2(x, y) \in \mathbf{F}_q[x, y]$ be of positive degree in y. Hence, we have

$$f_1(x, y) = a_{1,d_1}(x)y^{d_1} + a_{1,d_1-1}(x)y^{d_1-1} + \cdots + a_{1,0}(x)$$
$$f_2(x, y) = a_{2,d_2}(x)y^{d_2} + a_{2,d_2-1}(x)y^{d_2-1} + \cdots + a_{2,0}(x),$$

where $d_1 \geq 1$ and $d_2 \geq 1$ are the degrees of $f_1(x, y)$ and $f_2(x, y)$ over $\mathbf{F}_q[x]$, respectively.

The resultant $R(f_1, f_2)(x)$ of $f_1(x, y)$ and $f_2(x, y)$ over $\mathbf{F}_q[x]$ is the determinant of the $(d_1 + d_2) \times (d_1 + d_2)$ matrix

$$\begin{bmatrix}
a_{1,d_1}(x) & a_{1,d_1-1}(x) & \cdots & a_{1,0}(x) & 0 & \cdots & 0 \\
0 & a_{1,d_1}(x) & a_{1,d_1-1}(x) & \cdots & a_{1,0}(x) & \cdots & 0 \\
\vdots & \ddots & \ddots & & & \ddots & \vdots \\
0 & \cdots & a_{1,d_1}(x) & a_{1,d_1-1}(x) & & \cdots & a_{1,0}(x) \\
a_{2,d_2}(x) & a_{2,d_2-1}(x) & \cdots & a_{2,0}(x) & 0 & \cdots & 0 \\
0 & a_{2,d_2}(x) & a_{2,d_2-1}(x) & \cdots & a_{2,0}(x) & \cdots & 0 \\
\vdots & \ddots & \ddots & & & \ddots & \vdots \\
0 & \cdots & a_{2,d_2}(x) & a_{2,d_2-1}(x) & & \cdots & a_{2,0}(x)
\end{bmatrix}.$$

Lemmas 25 and 26 appear in [4; chapter I].

Lemma 25. *There exist polynomials* $p_1(x, y)$, $p_2(x, y) \in \mathbf{F}_q[x, y]$ *such that* $\deg_y p_1(x, y) < d_2$, $\deg_y p_2(x, y) < d_1$, *and*

$$R(f_1, f_2)(x) = p_1(x, y)f_1(x, y) + p_2(x, y)f_2(x, y).$$

Proof. Let $\mathbf{c}_i$ denote the ith column in the matrix of $R(f_1, f_2)(x)$. Replace the

last column $\mathbf{c}_{d_1+d_2}$ by

$$\mathbf{c} = \mathbf{c}_1 y^{d_1+d_2-1} + \mathbf{c}_2 y^{d_1+d_2-2} + \cdots + \mathbf{c}_{d_1+d_2},$$

then evaluate the determinant from the last column. ∎

Lemma 26. *Let $g(x,y) = \gcd(f_1(x,y), f_2(x,y))$. Then, $\deg_y g(x,y) \geq 1$ if, and only if, $R(f_1,f_2)(x) = 0$.*

Proof. From Lemma 25, and the fact $R(f_1,f_2)(x) \in \mathbf{F}_q[x]$, it follows that $R(f_1,f_2)(x) \neq 0$ implies $\deg_y g(x,y) = 0$. On the other hand, if $R(f_1,f_2)(x) = 0$, then

$$p_1(x,y)f_1(x,y) = p_2(x,y)f_2(x,y)$$

for some polynomials $p_1(x,y)$ and $p_2(x,y)$ as stated in Lemma 25. We have

$$\begin{aligned} p_2(x,y)g(x,y) &= \gcd(p_2(x,y)f_1(x,y), p_2(x,y)f_2(x,y)) \\ &= \gcd(p_2(x,y)f_1(x,y), p_1(x,y)f_1(x,y)), \end{aligned}$$

which implies $f_1(x,y) | p_2(x,y)g(x,y)$. As $\deg_y p_2(x,y) < \deg_y f_1(x,y)$, it follows that $\deg_y g(x,y) \geq 1$. ∎

Lemma 27. *Let $f_1(x,y)$ and $f_2(x,y)$ be relatively prime polynomials in $\mathbf{F}_2[x,y]$ of positive degree in y. Let $\gamma \in \mathbf{F}_q$, where q is a power of 2. Suppose that the degree of the minimal polynomial of γ over $\mathbf{F}_2$ is greater than $\deg_x f_1 \deg_y f_2 + \deg_x f_2 \deg_y f_1$. Then, $f_1(\gamma,y)$ and $f_2(\gamma,y)$ are relatively prime.*

Proof. From Lemma 26, we have $R(f_1,f_2)(x) \neq 0$ as $f_1(x,y)$ and $f_2(x,y)$ are relatively prime. But $R(f_1,f_2)(x) \in \mathbf{F}_2[x]$ is of degree not greater than $\deg_x f_1 \deg_y f_2 + \deg_x f_2 \deg_y f_1$, which follows from the definition of $R(f_1,f_2)(x)$. Hence, $R(f_1,f_2)(\gamma) \neq 0$ as the minimal polynomial of γ over $\mathbf{F}_2$ is of degree greater than the degree of $R(f_1,f_2)(x)$. By Lemma 26, it follows that $\gcd(f_1(\gamma,y), f_2(\gamma,y)) = 1$. ∎

Now we return to the cyclic burst correcting codes of minimum redundancy whose construction is given in Theorem 16.

Lemma 28. *Let $m_1 \geq 2b_1b_2$ and $m_2 \geq b_2$ be positive integers. Define $q = 2^{m_1m_2}$ and $q_1 = 2^{m_1}$. Let n_1 and n_2 satisfy the following conditions:*

(1) $n_1n_2 = q - 1$.

(2) m_1 *is the multiplicative order of 2 modulo* n_1.

(3) $\gcd(n_1, n_2) = 1$.

Let n_1' and n_2' be positive integers such that $n_1'|n_1$ and $n_2'|(q-1)/(q_1-1)$. Let $A_1 = b_1|\mathcal{F}_{b_11}|(n_1')^{|\mathcal{F}_{b_11}|}$ and $A_2 = b_2|\mathcal{F}_{b_1b_2}|(n_1'n_2')^{|\mathcal{F}_{b_1b_2}|}$. Suppose that $\phi(n_1) > A_1q_1^{1/2}d(n_1)$, and $\phi(n_2) > A_2q^{1/2}d(n_2)$. Then, there exist elements $\gamma, \beta \in \mathbf{F}_q$ of orders n_1 and n_2, respectively, such that $n_1'n_2'|\mathrm{ind}_\nu\big(f(\gamma,\beta)\big)$ for all $f \in \mathcal{F}_{b_1b_2}$, where ν is primitive in $\mathbf{F}_q$.

Proof. The set $\mathcal{F}_{b_11}$ is a set of irreducible polynomials over $\mathbf{F}_2$ whose degrees are less than b_1. Using Lemma 24, after substituting q_1 for q, n_1' for h, n_1 for e_1, and $\mathcal{F}_{b_11}$ for $\mathcal{F}$, it follows that there exist an element $\gamma \in \mathbf{F}_{q_1}$ of order n_1 such that $n_1'|\mathrm{ind}_{\nu_1}\big(f(\gamma)\big)$ for all $f \in \mathcal{F}_{b_11}$, where ν_1 is primitive in $\mathbf{F}_{q_1}$. This implies that $\nu_1 = \nu^{\frac{q-1}{q_1-1}\eta}$, where η is relatively prime to $q-1$. Hence,

$$\mathrm{ind}_\nu\big(f(\gamma)\big) = \frac{q-1}{q_1-1}\eta\mathrm{ind}_{\nu_1}\big(f(\gamma)\big) \pmod{q-1},$$

which implies $n_1'n_2'|\mathrm{ind}_\nu\big(f(\gamma)\big)$ for all $f \in \mathcal{F}_{b_11}$ since $n_2'|(q-1)/(q_1-1)$.

From condition (2), m_1 is the degree of the minimal polynomial of γ over $\mathbf{F}_2$. Since $m_1 \geq 2b_1b_2$, then from Lemma 27, it follows that

$$\mathcal{F} = \{f(\gamma, y) : f(x, y) \in \mathcal{F}_{b_1b_2} - \mathcal{F}_{b_11}\}$$

is a set of pairwise relatively prime polynomials. Lemma 24, after substituting $n_1'n_2'$ for h and n_2 for e_1, implies the existence of an element $\beta \in \mathbf{F}_q$ of order n_2 such that $n_1'n_2'|\mathrm{ind}_\nu\big(f(\gamma,\beta)\big)$ for all $f(x,y) \in \mathcal{F}_{b_1b_2} - \mathcal{F}_{b_11}$. This concludes the proof of the lemma. ∎

In this subsection, we have succeeded in finding sufficient conditions that

guarantee the satisfaction of condition (6) of Lemma 17.

6.3. The Existence of Cyclic Burst Correcting Codes of Minimum Redundancy

In this final subsection, we will prove that for all positive integers b_1 and b_2, there exist cyclic $b_1 \times b_2$-burst correcting codes of minimum redundancy with arbitrarily large areas. The strategy is to prove that there exists a sequence of areas $n_1 \times n_2$ such that both n_1 and n_2 are increasing, and for each $n_1 \times n_2$ in the sequence, the conditions of Lemma 17 are satisfied. First, we use Lemma 28 to replace condition (6) of Lemma 17.

Lemma 29. *Let b_1 and b_2 be positive integers. Suppose that $n_1' \times n_2'$ is the area of a $b_1 \times b_2$-code where n_1' and n_2' are odd and relatively prime. Let $m_1 \geq b_1$, $m_2 \geq b_2$, n_1, and n_2 be positive integers. Suppose that the following conditions are satisfied:*

(1) $n_1 n_2 = 2^{m_1 m_2} - 1$.

(2) m_1 *is the multiplicative order of 2 modulo* n_1.

(3) $\gcd(n_1, n_2) = 1$.

(4) $n_1' | n_1$.

(5) $n_2' | (2^{m_1 m_2} - 1)/(2^{m_1} - 1)$.

(6) $m_1 \geq 2b_1 b_2$.

(7) $\phi(n_1) > A_1 2^{m_1/2} d(n_1)$, *where* $A_1 = b_1 |\mathcal{F}_{b_1 1}| (n_1')^{|\mathcal{F}_{b_1 1}|}$.

(8) $\phi(n_2) > A_2 2^{m_1 m_2/2} d(n_2)$, *where* $A_2 = b_2 |\mathcal{F}_{b_1 b_2}| (n_1' n_2')^{|\mathcal{F}_{b_1 b_2}|}$.

Then, there exists a cyclic $b_1 \times b_2$-burst correcting code of minimum redundancy whose area is $n_1 \times n_2$.

If we find a sequence of pairs of integers (m_1, m_2) that satisfy the conditions of Lemma 29 such that both m_1 and m_2 are increasing, then the sequence of areas $n_1 \times n_2$ is such that both n_1 and n_2 are increasing. This is a direct consequence of conditions (1) and (2).

For a proof of the following lemma, see [11; chapter XVIII].

Lemma 30. *Let $\epsilon > 0$. Then, $d(n) < n^{\epsilon}$ and $\phi(n) > n^{1-\epsilon}$ for all sufficiently large n.*

In the following, we will refer to the conditions of Lemma 29 by their numbers.

Lemma 31. *If $m_2 \geq 3$ and conditions (1) and (2) hold, then condition (8) holds for all sufficiently large n_2.*

Proof. From conditions (1) and (2), we have

$$n_2 = \frac{2^{m_1 m_2} - 1}{n_1} \geq \frac{2^{m_1 m_2} - 1}{2^{m_1} - 1} \geq 2^{m_1(m_2-1)}.$$

If $m_2 \geq 3$, then

$$2^{m_1 m_2/2} \leq 2^{3m_1(m_2-1)/4} \leq n_2^{3/4}.$$

The lemma now follows from Lemma 30. ∎

Lemma 32. *There exists a sequence of pairs of integers (m_1, n_1) such that the sequence of m_1 is increasing, and for each (m_1, n_1), the following conditions hold:*

(2) m_1 *is the multiplicative order of 2 modulo* n_1.

(4) $n_1' | n_1$.

(7) $\phi(n_1) > A_1 2^{m_1/2} d(n_1)$.

(9) $n_1' n_2' | 2^{m_1} - 1$.

(10) $\gcd(n_1, n_2') = 1$.

(11) For every prime p, if $p \nmid n_2'$, then $p \nmid (2^{m_1} - 1)/n_1$.

Proof. In the following, we will write $\exp_h(2)$, for odd integer h, to denote the multiplicative order of 2 modulo h. Let $n_2' = {p_1}^{a_1} \ldots {p_r}^{a_r}$ be the prime factorization of n_2' where the p_i's are distinct and the a_i's are positive. From Lemma 6, it follows that there exists a positive integer m_0 such that for every integer $m > m_0$, there exists a prime $q_m \notin \{p_1, \ldots, p_r\}$ such that $m = \exp_{q_m}(2)$. Let $m > m_0$

and $\exp_{n_1'n_2'}(2)|m$. Define $n_1 = \prod_p p^{c_p}$, where p is a prime, $p \notin \{p_1, \ldots, p_r\}$, and $p^{c_p}||2^m - 1$. Then, $p \not| (2^{m_1} - 1)/n_1$ if $p \not| n_2'$. We also have $\gcd(n_1, n_2') = 1$. Since $\gcd(n_1', n_2') = 1$, it follows that $n_1'|n_1$. As $q_m \notin \{p_1, \ldots, p_r\}$, it follows that m is the multiplicative order of 2 modulo n_1. For each p_i, $1 \leq i \leq r$, let d_i be the least positive integer such that $\exp_{p_i^{d_i}}(2) \not| \exp_{n_1'n_2'}(2)$. Now choose $m > m_0$, $\exp_{n_1'n_2'}(2)|m$, and $\exp_{p_i^{d_i}}(2) \not| m$ for $1 \leq i \leq r$. Certainly, there exists an infinite number of choices for such m. Let S be the infinite set of all such m's. Hence, for $m \in S$, $p_i^{d_i} \not| 2^m - 1$ for each $1 \leq i \leq r$. Let $t = p_1^{d_1} \ldots p_r^{d_r}$, which is independent of the choice of m. Then, from the definition of n_1, it follows that $n_1 > (2^m - 1)/t$. From Lemma 30, it follows that if m is sufficiently large, then $\phi(n_1) > A_1 2^{m/2} d(n_1)$. Now, the lemma follows by choosing m_1 sufficiently large in the set S. ∎

Lemma 33. *There exists a sequence of pairs of integers (m_1, m_2) such that both m_1 and m_2 are increasing and for each (m_1, m_2) the conditions of Lemma 29 are satisfied for some n_1 and n_2.*

Proof. Take (m_1, n_1) in the sequence of Lemma 32 such that $m_1 \geq 2b_1b_2$. Choose m_2 such that $n_2'|m_2$ but if $p|2^{m_1} - 1$ and $p \not| n_2'$, then $p \not| m_2$. Certainly, there exists an infinite number of choices for such m_2. Let $n_2 = (2^{m_1m_2} - 1)/n_1$. So condition (1) is satisfied. Since

$$\frac{2^{m_1m_2} - 1}{2^{m_1} - 1} = 2^{m_1(m_2-1)} + 2^{m_1(m_2-2)} + \cdots + 2^{m_1} + 1,$$

and $2^{m_1} \equiv 1 \pmod{n_2'}$ by condition (9) of Lemma 32, it follows that

$$\frac{2^{m_1m_2} - 1}{2^{m_1} - 1} \equiv \underbrace{1 + \cdots + 1}_{m_2 \text{ terms}} \equiv m_2 \pmod{n_2'}.$$

Since $n_2'|m_2$, condition (5) is satisfied. If m_2 is chosen sufficiently large, then from Lemma 31, it remains to prove that condition (3) holds. Suppose $p|n_1$ for some prime p. Then $p|2^{m_1} - 1$. Condition (10) of Lemma 32 implies $p \not| n_2'$. By

condition (11), $p \not| (2^{m_1}-1)/n_1$. By the choice of m_2, we have $p \not| m_2$. But

$$\frac{2^{m_1 m_2}-1}{2^{m_1}-1} = 2^{m_1(m_2-1)} + 2^{m_1(m_2-2)} + \cdots + 2^{m_1} + 1 \equiv m_2 \pmod{p},$$

which implies $p \not| (2^{m_1 m_2}-1)/(2^{m_1}-1)$. Hence,

$$p \not| \left(\frac{2^{m_1 m_2}-1}{2^{m_1}-1}\right)\left(\frac{2^{m_1}-1}{n_1}\right) = n_2.$$

■

Now we combine Theorem 15 and Lemmas 29 and 33 to obtain the most important result of this chapter.

Theorem 34. *For every pair (b_1, b_2) of positive integers, there exists a class of cyclic $b_1 \times b_2$-burst correcting codes of arbitrarily large areas whose excess redundancy is $\lceil \log N(b_1, b_2) \rceil$.*

APPENDIX

ENCODING AND DECODING TWO-DIMENSIONAL BURST CORRECTING CODES

1. Introduction

In this appendix, we give an encoding and decoding technique for the burst correcting codes developed in chapter IV. Recall that a BIL-code, as explained in section 4, chapter IV, is constructed from a burst identification code and a burst locating code. A Fire-ish code, as explained in section 5, chapter IV, is constructed from a cyclic-burst correcting code and a cyclic-burst locating code. In this appendix, we will assume that the burst locating codes in the BIL-codes and the Fire-ish codes are cyclic. As explained in chapter IV, the classes of $\gamma\beta$-codes and $\alpha\beta$-codes are cyclic burst locating codes which can be used to construct BIL-codes and Fire-ish codes whose excess redundancy is minimum. It should be noted that we do not assume that the BIL-codes and the Fire-ish codes considered in this appendix are necessarily cyclic. The encoding and decoding techniques presented in this appendix are also applicable to cyclic burst correcting codes of minimum redundancy described in section 6, chapter IV.

The techniques presented here are generalizations of encoding and decoding techniques of one-dimensional cyclic burst correcting codes. Both one-dimensional encoding and decoding techniques where generalized by Imai in [12],[13] to two-dimensional cyclic codes of odd areas. In the following, we do not assume that the areas are necessarily odd.

In this introduction, we will consider cyclic burst locating codes. These codes may be those used in the construction of BIL-codes or Fire-ish codes.

They may also be the cyclic burst correcting codes of minimum redundancy.

Let $H = [\mathbf{h}_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$ be the parity check matrix of a cyclic $b_1 \times b_2$-burst locating code $\mathcal{C}$ of redundancy r. Let $\Omega = \{(i,j) : 0 \leq i < n_1, 0 \leq j < n_2\}$. From the definition of a $b_1 \times b_2$-burst locating code, it follows that no codeword is a $b_1 \times b_2$-burst. Hence, there exists a set of parity check positions Π of cardinality r such that $\{(i,j) : 0 \leq i < b_1, 0 \leq j < b_2\} \in \Pi$.*

The set $\{\mathbf{h}_{k,l} : (k,l) \in \Pi\}$ forms a basis for $\mathbf{F}_2^r$. Hence, any element $\mathbf{h}_{i,j}$ in H can be expressed as

$$\mathbf{h}_{i,j} = \sum_{(k,l)\in\Pi} h_{k,l}^{(i,j)} \mathbf{h}_{k,l},$$

where $h_{k,l}^{(i,j)} \in \mathbf{F}_2$ and $(i,j) \in \Omega$. Equivalently, $\mathbf{h}_{i,j}$ can be represented as a polynomial $h_{i,j}(x,y)$ given by

$$h_{i,j}(x,y) = \sum_{(k,l)\in\Pi} h_{k,l}^{(i,j)} x^k y^l.$$

Obviously, $h_{i,j}(x,y) = x^i y^j$ for $(i,j) \in \Pi$. It follows, from the definition of the parity check matrix, that $c(x,y) \in \mathbf{F}_2[x,y]$ is a codeword in $\mathcal{C}$ if, and only if,

$$\sum_{(i,j)\in\Omega} c_{i,j} h_{i,j}(x,y) = 0.$$

For a subset $\Phi \in \Omega$, define $\mathcal{P}[\Phi]$ to be the set of all polynomials $p(x,y) = \sum_{(i,j)\in\Phi} p_{i,j} x^i y^j$ over $\mathbf{F}_2$. For a polynomial $p(x,y) \in \mathbf{F}_2[x,y]$ let $\overline{p(x,y)}$ denote the polynomial in $\mathcal{P}[\Omega]$ such that

$$\overline{p(x,y)} \equiv p(x,y) \pmod{x^{n_1}+1, y^{n_2}+1}.$$

---

* *We do not require that Π satisfies other conditions. This is different from Imai's approach in which Π is restricted to be of some specific form which does not guarantee that $\{(i,j) : 0 \leq i < b_1, 0 \leq j < b_2\} \in \Pi$ in general.*

The operator $T_x : \mathcal{P}[\Pi] \longrightarrow \mathcal{P}[\Pi]$ is defined as

$$T_x p(x,y) = \overline{xp(x,y)} + \sum_{(i,j)\in\Pi_{\partial x}} p_{i,j} g_{i+1,j}(x,y),$$

where $g_{i,j}(x,y) = x^i y^j + h_{i,j}(x,y)$, $\Pi_{\partial x} = \{(i,j) \in \Pi : ((i+1) \bmod n_1, j) \in \Omega - \Pi\}$. The operator $T_y : \mathcal{P}[\Pi] \longrightarrow \mathcal{P}[\Pi]$ is defined as

$$T_y p(x,y) = \overline{yp(x,y)} + \sum_{(i,j)\in\Pi_{\partial y}} p_{i,j} g_{i,j+1}(x,y),$$

where $\Pi_{\partial y} = \{(i,j) \in \Pi : (i, (j+1) \bmod n_2) \in \Omega - \Pi\}$.

In the following, we will list some results from [12]. These results are direct consequences of the definitions, and the fact that $g_{i,j}(x,y)$, where $(i,j) \in \Omega$, is a codeword in $\mathcal{C}$. However, for a complete proof of these results, see [12].*

Lemma 1. $T_x^{n_1} 1 = T_y^{n_2} 1 = 1$.

Lemma 2. *If* $p(x,y) \in \mathcal{P}[\Pi]$, *then* $p(T_x, T_y) = p(x,y)$.

Lemma 3. *Let* $p_1(x,y) = \overline{p_2(x,y)}$, *where* $p_2(x,y) \in \mathbf{F}_2[x,y]$. *Then,* $p_1(T_x, T_y) = p_2(T_x, T_y)$.

Theorem 4. *Let* $c(x,y) \in \mathcal{P}[\Omega]$. *Then,* $c(x,y)$ *is a codeword in* $\mathcal{C}$ *if, and only if* $c(T_x, T_y) = 0$.

2. Encoding Burst Correcting Codes

2.1. Encoding Cyclic Burst Correcting Codes

The encoding technique described here is due to Imai [13].

Let $m_{i,j}$, where $(i,j) \in \Omega - \Pi$, be the information bits. Let $m(x,y) = \sum_{(i,j)\in\Omega} m_{i,j} x^i y^j$, where $m_{i,j} = 0$ for $(i,j) \in \Pi$. To encode $m(x,y)$, we first

* *Although in [12], it is assumed that* Π *satisfies certain conditions, the proofs are equally valid for any* Π *as defined in this appendix.*

calculate $m(T_x, T_y)$, and then let

$$c(x, y) = m(x, y) + m(T_x, T_y).$$

By Theorem 4, $c(x, y)$ is a codeword.

So, the main problem is to give an efficient technique to calculate $m(T_x, T_y)$. We use a two-dimensional feedback shift register in which $|\Pi|$ storage devices are arranged in the form of the parity check positions given by Π. We represent the contents of the shift register by $\sigma(x, y) = \sum_{(i,j)\in\Pi} \sigma_{i,j} x^i y^j$, where $\sigma_{i,j} \in \mathbb{F}_2$ is the content of the storage device at position (i, j). The shift register can be shifted in the x-direction and the y-direction. The feedback connections are set such that if the register is shifted in the x-direction, the contents become $T_x\sigma(x, y)$, and if it is shifted in the y-direction, the contents become $T_y\sigma(x, y)$. The register has one input line which is connected to the input of the storage device at position $(0, 0)$ only when the register is shifted in the x-direction.

The storage device is set initially to zero. The coefficients of $m(x, y)$ are fed into the register in the order

$$\begin{array}{cccc} m_{n_1-1,n_2-1} & m_{n_1-2,n_2-1} & \cdots & m_{0,n_2-1} \\ m_{n_1-1,n_2-2} & m_{n_1-2,n_2-2} & \cdots & m_{0,n_2-2} \\ \vdots & \vdots & & \vdots \\ m_{n_1-1,0} & m_{n_1-2,0} & \cdots & m_{0,0}. \end{array}$$

Each row is shifted into the register by n_1 successive shifts in the x-direction. Then, the register is shifted once in the y-direction with no input. Thus, n_1 successive shifts in the x-direction are followed by one shift in the y-direction with no input until $m_{0,0}$ enters the register. Hence, after entering $m_{i,j}$, the register is shifted $n_1 j + i$ times in the x-direction, and j times in the y-direction until $m_{0,0}$ enters. Therefore, after entering $m_{0,0}$, the contents become

$$\sum_{(i,j)\in\Omega} T_x^{n_1 j+i} T_y^j m_{i,j} = \sum_{(i,j)\in\Omega} T_x^i T_y^j m_{i,j} = m(T_x, T_y),$$

where Lemma 1 was used.

The encoding technique presented here can be applied to encode cyclic burst locating codes with minimum redundancy.

2.2. Encoding BIL-Codes and Fire-ish Codes

Consider a $b_1 \times b_2$-burst correcting code which may be a BIL-code or a Fire-ish code. The parity check matrix in both codes can be written as $(\mathbf{h}'_{i,j}, \mathbf{h}''_{i,j})$, $0 \leq i < n_1$, $0 \leq j < n_2$, where $[\mathbf{h}'_{i,j}]$ is the parity check matrix, or copies of the parity check matrix of a cyclic burst locating code, and $[\mathbf{h}''_{i,j}]$ is the parity check matrix of a burst identification code or copies of the parity check matrix of a cyclic-burst correcting code. In both cases, the code whose parity check matrix is $[\mathbf{h}''_{i,j}]$, $0 \leq i < n_1$, $0 \leq j < n_2$, will be called the burst code.

Let Π be check positions for the BIL-code or the Fire-ish code such that $(i,j) \in \Pi$ for $0 \leq i < b_1$, $0 \leq j < b_2$. The information bits are represented by the polynomial $m(x,y) = \sum_{(i,j)\in\Omega} m_{i,j} x^i y^j$, where $m_{i,j} = 0$ for $(i,j) \in \Pi$. Let $m(x,y)$ be encoded for the cyclic code using the two-dimensional shift register described in subsection 2.1. Let $c'(x,y)$ be the codeword corresponding to $m(x,y)$ in the cyclic code. Naturelly, $c'(x,y)$ may not be a codeword in the burst code. Let $\mathbf{s}_B$ denote its syndrome with respect to the burst code. Now, we will obtain a codeword $c''(x,y) \in \mathcal{P}[\Pi]$ in the cyclic code whose syndrome with respect to the burst code is $\mathbf{s}_B$. The polynomial $c''(x,y) = \sum_{(i,j)\in\Pi} c''_{i,j} x^i y^j$ is determined by

$$\sum_{(i,j)\in\Pi} c''_{i,j} \mathbf{h}'_{i,j} = 0,$$
$$\sum_{(i,j)\in\Pi} c''_{i,j} \mathbf{h}''_{i,j} = \mathbf{s}_B.$$

From the definition of Π, it follows that the elements $(\mathbf{h}'_{i,j}, \mathbf{h}''_{i,j})$, where $(i,j) \in \Pi$, are independent. Hence, these equations can be solved to obtain $c''(x,y)$.

Now, $c(x,y) = c'(x,y) + c''(x,y)$ is the codeword in the burst correcting

code corresponding to the the information bits represented by the polynomial $m(x, y)$.

3. Decoding Burst Correcting Codes

Consider a $b_1 \times b_2$-burst correcting code. Let $v(x, y) = c(x, y) + e(x, y)$ be the received word where $c(x, y)$ is a codeword and $e(x, y)$ is a $b_1 \times b_2$-cyclic burst of error added to the codeword $c(x, y)$. Then, we have $e(x, y) = \overline{x^{u_1} y^{u_2} b(x, y)}$, for some $(u_1, u_2) \in \Omega$, where $b(x, y)$ is the burst pattern.

3.1. Decoding Cyclic Burst Correcting Codes

The following is a two-dimensional version of the error trapping algorithm which is well known for one-dimensional cyclic burst correcting codes. The two-dimensional version described here is due to Imai [12].

Since $c(T_x, T_y) = 0$ by Theorem 4, $v(T_x, T_y) = e(T_x, T_y)$. Now, $b(x, y) \in P[\Pi]$, and hence Lemmas 2 and 3 imply

$$\begin{aligned} v(T_x, T_y) = e(T_x, T_y) &= T_x^{u_1} T_y^{u_2} b(T_x, T_y) \\ &= T_x^{u_1} T_y^{u_2} b(x, y). \end{aligned} \tag{1}$$

We calculate $v(T_x, T_y)$ by using the two-dimensional shift register described in section 2. After that the register is shifted with no input in the same order, i.e., n_1 shifts in the x-direction followed by a shift in the y-direction. Let $\sigma_{i,j}(x, y)$ be the contents of the shift register when it is shifted $n_1 j + i$ times in the x-direction and j-times in the y-direction after calculating $v(T_x, T_y)$, i.e., with

$\sigma_{0,0} = v(T_x, T_y)$. Then,

$$\begin{aligned}\sigma_{i,j}(x,y) &= T_x^{n_1 j+i} T_y^j \sigma_{0,0}(x,y) \\ &= T_x^i T_y^j \sigma_{0,0}(T_x, T_y) \\ &= T_x^i T_y^j v(T_x, T_y) \\ &= T_x^{i+u_1} T_y^{j+u_2} b(x,y),\end{aligned}$$

where Lemma 1 and equation (1) were used. From Theorem 4, it follows that

$$\sigma_{i,j}(x,y) + \overline{x^{i+u_1} y^{j+u_2} b(x,y)} \tag{2}$$

is a codeword. We continue shifting the shift register until $\sigma_{i,j}(x,y)$ is a burst pattern of a $b_1 \times b_2$-burst, i.e., $\sigma_{i,j}(x,y) \in \mathcal{B}_{b_1,b_2}$. In such case, (2) is the sum of two $b_1 \times b_2$-cyclic bursts, and hence must be zero. This determines $e(x,y) = \overline{x^{u_1} y^{u_2} b(x,y)}$.

3.2. Decoding BIL-Codes

BIL-codes are, in general, not cyclic-burst correcting codes. So, we assume that $e(x,y)$ is a $b_1 \times b_2$-burst, i.e., $e(x,y) = x^{u_1} y^{u_2} b(x,y)$, where $b(x,y) \in \mathcal{B}_{b_1,b_2}$ and $(u_1, u_2) \in \Omega$ such that $e(x,y) \in \mathcal{P}[\Omega]$. We use the burst identification code to determine the burst pattern $b(x,y)$. Then we use the shift register as described in subsection 3.1. applied to the cyclic burst locating code until $\sigma_{i,j}(x,y) = b(x,y)$ for some (i,j). In such case (2) is the sum of two $b_1 \times b_2$-cyclic bursts of the same pattern and is a codeword in a cyclic burst locating code. Hence, this codeword is zero, and thus $e(x,y) = x^{u_1} y^{u_2} b(x,y)$ is determined.

3.3. Decoding Fire-ish Codes

Consider a Fire-ish code of area $n_1 \times n_2$ which is a $b_1 \times b_2$-cyclic-burst correcting code. The parity check matrix of this code has the form $[(\mathbf{h}'_{i \bmod n'_1, j \bmod n'_2}, \mathbf{h}''_{i \bmod n''_1, j \bmod n''_2})]$, $0 \le i < n_1$, $0 \le j < n_2$, where $n_1 = \mathrm{lcm}(n'_1, n''_1)$, $n_2 = \mathrm{lcm}(n'_2, n''_2)$, $[\mathbf{h}'_{i,j}]$, $0 \le i < n'_1$, $0 \le j < n'_2$, is the parity check matrix of

a cyclic $b_1 \times b_2$-burst locating code of area $n_1' \times n_2'$, and $[\mathbf{h}_{i,j}'']$, $0 \leq i < n_1''$, $0 \leq j < n_2''$, is the parity check matrix of a $b_1 \times b_2$-cyclic-burst correcting code of area $n_1'' \times n_2''$.

To decode a Fire-ish code, we use the cyclic-burst correcting code to determine $u_1 \bmod n_1'$, $u_2 \bmod n_2'$, and the burst pattern $b(x, y)$. Then we use the shift register applied to the cyclic burst locating code until $\sigma_{i,j}(x, y) = b(x, y)$ for some $(i, j) \in \Omega$. In such case, (2) is the sum of two $b_1 \times b_2$-cyclic bursts of the same pattern which is a codeword in a cyclic $b_1 \times b_2$-burst locating code of area $n_1' \times n_2'$. Hence, this codeword is zero, and thus $u_1 \bmod n_1'$ and $u_2 \bmod n_2'$ can be determined. It follows that u_1 and u_2 can be determined modulo $\text{lcm}(n_1', n_1'') = n_1$ and $\text{lcm}(n_2', n_2'') = n_2$, respectively. This determines $e(x, y) = \overline{x^{u_1} y^{u_2} b(x, y)}$.

It is to be noted that in many cases, encoding and decoding noncyclic BIL-codes and Fire-ish codes are not computationally difficult. The reason is that the noncyclic codes used in these classes of codes as given in the previous chapters have small areas or periodic whose periods are in the order of $b_1 \times b_2$.

References for Part Two

[1] K. A. S. Abdel-Ghaffar, R. J. McEliece, A. M. Odlyzko, H. C. A. van Tilborg, "On the existence of optimum cyclic burst correcting codes," *IEEE Trans. Inform. Theory*, to appear.

[2] N. M. Abramson, "A class of systematic codes for non-independent errors," *IRE Trans. Inform. Theory*, vol. IT-5, pp. 150–157, December 1959.

[3] E. L. Berlekamp, *Algebraic Coding Theory*, Laguna Hills, CA: Aegean Park Press, 1984.

[4] G. A. Bliss, *Algebraic Functions*, New York: American Mathematical Society, 1933.

[5] C. N. Campopiano, "Bounds on burst-error-correcting codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 257–259, April 1962.

[6] L. Carlitz, "Primitive roots in a finite field," *Trans. American Mathematical Society*, vol. 73, pp. 373–382, 1952.

[7] L. E. Dickson, "On the cyclotomic function," *Amer. Math. Monthly*, vol. 12, pp. 86–89, 1905.

[8] B. Elspas, "Notes on multidimensional burst-error correction," presented at the IEEE Int. Symp. on Inform. Theory, San Remo, Italy, September 1967.

[9] P. Fire, "A class of multiple-error-correcting binary codes for non-independent errors," Sylvania Reconnaisance Sys. Lab., Mountain View, Calif., Sylvania Rept. RSL-E-2, March 1959.

[10] B. Gordon, "On the existence of perfect maps," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 486–487, October 1966.

[11] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford: Oxford University Press, 1983.

[12] H. Imai, "Two-dimensional Fire codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 796–806, November 1973.

[13] H. Imai, "A theory of two-dimensional cyclic codes," *Inform. Contr.*, vol. 34, pp 1–21, 1977.

[14] S. Lang, *Algebra*, Menlo Park, CA: Addison-Wesley, 1984.

[15] R. Lidl and H. Niederreiter, *Finite Fields*, Reading, MA: Addison-Wesley, 1983.

[16] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland, 1981.

[17] R. J. McEliece, *The Theory of Information and Coding*, Reading, MA: Addison-Wesley, 1983.

[18] R. J. McEliece, *Finite Fields*, Hingham, MA: Klewer, 1986.

[19] T. Nomura, H. Miyakawa, H. Imai, and A. Fukuda, "A theory of two-dimensional linear recurring arrays," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 775–785, November 1972.

[20] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, Camdridge, MA: M.I.T. Press, 1981.

[21] W. M. Schmidt, *Equations over Finite Fields*, Lecture Notes in Math., vol. 539, Berlin: Springer-Verlag, 1976.