A DECODING SCHEME FOR THE 4-ARY LEXICODES WITH $d_m = 4$

D.G. KIM

Liberal Arts and Science, Chungwoon University, South Korea

ABSTRACT. In this paper, we are interested in lexicographic codes which are greedily constructed codes. For an arbitrary length n, we shall find the basis of quaternary lexicographic codes, for short, lexicodes, with minimum distance $d_m = 4$. Also using a linear nim sum of some bases (such a vector is called the testing vector), its decoding algorithm will be found.

1. INTRODUCTION

In this paper, we shall introduce the surprising arithemetical operations which are used in the Game of Nim. Under these operations, the lexicodes are linear over some Galois field. Their definitions are derived from a greedy algorithm, that is, each codeword is chosen as the first word not prohibitively near to previous codewords.

The main aim of this paper is to find a decoding algorithm of the 4-ary lexicodes with minimum distance 4. Using the special vector, called the testing vector, we correct an error symbol of the received vector.

This paper is arranged as follows. The nim-operation is introduced in Section 2, and the lexicodes over the Galois field $GF(2^{2^a})$ are discussed in Section 3. In particular we get some bases of the 4-ary lexicodes with minimum distance 4, and give an algorithm to find the basis according to length n in Section 4. Finally, Section 5 gives a decoding algorithm for this code and its examples.

1991 *Mathematics Subject Classification.* 94Bxx.

Key words and phrases. Nim-operations, Hamming distance, Hamming weight, Basis, Decode.

This work was partially supported by Chungwoon University, 2003

2. Nim operation

First, we define two operations which are called the nim-addition $\oplus$ and nim-multiplication $\otimes$.

Definition 2.1. *Let x' be a variable that ranges over all elements strictly less than x and mex the least non-negative integer not of the form. Then we define the two operations:*
(1) $a \oplus b = mex\{a' \oplus b,\ a \oplus b'\}$
(2) $a \otimes b = mex\{(a' \otimes b) \oplus (a \otimes b') \oplus (a' \otimes b')\}$

Two operations, $\oplus$ and $\otimes$, convert the numbers $0, 1, 2, \cdots$ into a field of characteristic 2. Also, for $a \geq 0$, the numbers less than 2^{2^a} form a subfield and isomorphic to GF(2^{2^a}).

Theorem 2.2 ([2]). *The nim-operations turn the set of non-negative integers into a field of characteristic* 2.

Using the field laws, we shall fill out the first 4 by 4 corner of the addition and multiplication tables in nim. Consider the nim-addition of any two numbers from $0, 1, 2, 3$.

Theorem 2.3 ([1]). *We have $x \oplus 0 = 0 \oplus x = x$, for every number x.*

Since $\{0, 1, 2, 3\}$ is a field of characteristic 2, we have $x \oplus x = 0$ for all $x \in \{0, 1, 2, 3\}$. From Theorem 2.3, $1 \oplus 2$ can not be one of $0, 1, 2$ and so must be 3. Since $1 \oplus 3 \neq 0, 1, 3$, it must be 2. In the same way, we have $2 \oplus 3 = 1$. Therefore the sum of any two distinct numbers from $1, 2, 3$ is the third.

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

There is a nim-multiplication $\otimes$ which together with nim-addition $\oplus$ converts the integers into a field [2]. With nim-multiplication, we know that $0 \otimes x$ must be 0 which is the zero of the field. Also $1 \otimes x$ must be x. Since the elements other than 0, 1 satisfy $x^2 = x \oplus 1$ (here x^2 means $x \otimes x$) over GF(4), we have $2 \otimes 2 = 2 \oplus 1 = 3$ and $3 \otimes 3 = 3 \oplus 1 = 2$. Next $2 \otimes 3$ can not be one of $0, 2, 3$ and so must be 1.

$\otimes$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

The following is a rule enabling us to perform nim-additions. In its statement, the term 2-power means a power of 2, such as $1, 2, 4, 8, \cdots$, in the ordinary sense:

(1) If x is a 2-power and $y < x$, then $x \oplus y = x + y$.
(2) $x \oplus x = 0$ for any x.

For example, $15 \oplus 5 = (8 \oplus 4 \oplus 2 \oplus 1) \oplus (4 \oplus 1) = 8 \oplus 2 = 10$, since both 4's and 1's are cancelled.

For finite numbers, the nim-multiplication follows from the following rules, similar to those for nim-addition. In the following statement, the term *Fermat 2-power* means the number 2^{2^a}, such as $2, 4, 16, 256 \cdots$, in the ordinary sense:

(3) If x is a Fermat 2-power and $y < x$, then $x \otimes y = x \times y$.
(4) $x \otimes x = \frac{3}{2} \times x$ for any Fermat 2-power x.

For example $16 \otimes 2 = 32$, since $16 = 2^{2^2}$. By the equation (4), we have $2^2 = 2 \times \frac{3}{2} = 3$, $4^2 = 4 \times \frac{3}{2} = 6$, $16^2 = 16 \times \frac{3}{2} = 24, \cdots$.
Using the associative and distributive laws, $19 \otimes 11 = (16 \oplus 2 \oplus 1) \otimes (8 \oplus 2 \oplus 1) = (16 \otimes 8) \oplus (16 \otimes 2) \oplus (16 \otimes 1) \oplus (2 \otimes 8) \oplus (2 \otimes 2) \oplus (2 \otimes 1) \oplus (8 \oplus 2 \oplus 1) = 128 \oplus 32 \oplus 16 \oplus (2 \otimes 8) \oplus 2 \oplus 8 = 128 \oplus 32 \oplus 16 \oplus 4 \oplus 2 = 182$, since $2 \otimes 8 = 2 \otimes (4 \otimes 2) = 4 \otimes 2^2 = 4 \otimes 3 = 8 \oplus 4$.
Next, we compute the inverse value 15^{-1} satisfying $15 \otimes 15^{-1} = 1$. $15 \otimes 4 = (8 \oplus 4 \oplus 2 \oplus 1) \otimes 4 = (8 \otimes 4) \oplus (4 \otimes 4) \oplus (2 \otimes 4) \oplus (1 \otimes 4) = (2 \otimes 4 \otimes 4) \oplus 6 \oplus 8 \oplus 4 = (2 \otimes 6) \oplus (4 \oplus 2) \oplus 8 \oplus 4 = (2 \otimes (4 \oplus 2)) \oplus 2 \oplus 8 = 8 \oplus 3 \oplus 2 \oplus 8 = 3 \oplus 2 = 1$.
Hence $15^{-1} = 4$.

3. LEXICODES

Consider a lexicode over $GF(2^{2^a})$. A vector of this code is a sequence $\cdots x_3 x_2 x_1 = \mathbf{x}$, $x_i \in \{0, 1, \cdots, 2^{2^a} - 1\}$. For a convenience, we omit leading zeros (i.e., 012 = 12). The set of vectors is based on a lexicographic (i.e., dictionary) ordering of vectors, namely, the vector $\cdots x_3 x_2 x_1 = \mathbf{x}$ is smaller than the vector $\cdots y_3 y_2 y_1 = \mathbf{y}$, written $\mathbf{x} < \mathbf{y}$, if for some n we have $x_n < y_n$, but $x_N = y_N$ for all $N > n$. For example, 123 < 132, 312 < 1032.

Lexicodes are defined by saying a vector is in the code if it does not conflict with any earlier codewords. That is, the lexicode with minimum distance d_m is defined by saying that two vectors do not conflict if the Hamming distance between them is not less than d_m. The Hamming distance d between two vectors is simply the number of positions in which the vectors differ. Now we abbreviate the q-ary lexicodes with minimum distance d_m to $\mathcal{L}_{q,d_m}$.

Example. Applying the greedy algorithm, then the lexicode $\mathcal{L}_{4,3}$ contains the codewords, 0, 111, 222, 333, 1012, 1103, 1230, 1321, 2023, 2132, 2201, 2310, 3031, 3120, 3213, 3302.

In [3], it was shown that if $B = 2^a$, the lexicodes are closed under coordinatewise nim-addition over GF(B), and if $B = 2^{2^a}$, the lexicodes are closed under coordinatewise nim-multiplication by scalars k over GF(B). As a result we provide the following Lexicode Theorem.

Theorem 3.1 **([3]).** *If B is of the form 2^{2^a}, then the lexicode is a linear code over the Galois field GF(B).*

4. Basis

It is important that we obtain the basis for an arbitrary length n. This will give an information of decoding. So in this section, it is shown that according to the range of length, the basis has repeatedly the regular form in the first three symbols.

Now, any vectors will be shown in bold face.

Lemma 4.1 **([4]).** *Let* $\mathbf{e}_n$ *be the basis of length* n *of* $\mathcal{L}_{4,3}$. *Then* $111 = \mathbf{e}_3$, $1012 = \mathbf{e}_4$ *and* $10013 = \mathbf{e}_5$.

Lemma 4.2([4]). *There does not exist the basis of length* 6 *or* $17s + 5$ $(s \in \mathbb{N})$ *in* $\mathcal{L}_{4,3}$.

Notation : Let $\mathbf{E}_n$ be the basis of length n of $\mathcal{L}_{4,4}$ and $\alpha \in$ GF(4). An extra symbol of $\mathbf{E}_n$ is denoted by $(f_0)_n$.

Lemma 4.3. *We obtain* $1111 = \mathbf{E}_4$, $10123 = \mathbf{E}_5$ *and* $100132 = \mathbf{E}_6$.

Proof. From Lemma 4.1, we have obtained $\mathbf{e}_3$, $\mathbf{e}_4$ and $\mathbf{e}_5$. Using the basis of $\mathcal{L}_{4,3}$, we can make the basis of $\mathcal{L}_{4,4}$ by adding a symbol(called an extra symbol). This implies that an extra symbol would be added to the rightmost position of $\mathbf{e}_{n-1}$, for $4 \le n \le 6$. Here, the lexicographically earliest vector of distance 4 from a vector $\mathbf{0}$ must be 1111. So we get $1111 = \mathbf{E}_4$. From Lemma

4.1, 1012 = $\mathbf{e}_4$. So we may assume that $1012f_0 = \mathbf{E}_5$. If $(f_0)_5 = 0,1,2$, then the Hamming distance $d(\mathbf{E}_5, \alpha \otimes \mathbf{E}_4) = 3$. This contradicts the fact with the Hamming distance no less than 4. Thus we get $10123 = \mathbf{E}_5$. Also since $10013 = \mathbf{e}_5$, $\mathbf{E}_6$ is of the form $10013f_0$. If $(f_0)_6 = 0,1,3$, then $d(\mathbf{E}_6, \alpha \otimes \mathbf{E}_4) = 3$. Hence, we get $100132 = \mathbf{E}_6$. □

Lemma 4.4. *Let $s \in \mathbb{N}$. There does not exist the basis of length 7 or $17s+6$ in $\mathcal{L}_{4,4}$. Moreover, for $n > 7$, $\mathbf{E}_n$ has a symbol 1 in the 7th and the $17s+6$th positions.*

Proof. From Lemma 4.2, it was known that there does not exist $\mathbf{e}_n$ of length 6, $17s+5$ for all s. $\mathbf{E}_n$ can be obtained by adding an extra symbol to $\mathbf{e}_{n-1}$ of $\mathcal{L}_{4,3}$. For these reasons, neither $\mathbf{E}_7$ nor $\mathbf{E}_{17s+6}$ exist in $\mathcal{L}_{4,4}$. From [4] (see the proof of Theorem 2.2), the second result is clear. □

Notation : Let $[f_2f_1f_0]_n$ be the first three symbols of $\mathbf{E}_n$ and $[f_2f_1]_{n-1}$ the first two symbols of $\mathbf{e}_{n-1}$ of $\mathcal{L}_{4,3}$. Now we abbreviate the dimension k of $\mathcal{L}_{4,4}$ as $\mathcal{L}_{4,4}^k$.

Lemma 4.5. $[f_2f_1f_0]_{12} = [110]$.

Proof. We first prove that $[f_2f_1f_0]_{11}$ is [101]. Since $[f_2f_1]_{10} = [10]$, $[f_2f_1 f_0]_{11} = [10f_0]$. If $(f_0)_{11} = 0$, then the Hamming weight $wt(1\cdots 10f_0)$ is equal to 3. Assume that $(f_0)_{11} = 1$, i.e., $[101]_{11}$. For $4 \le n \le 6$, $d(\mathbf{E}_n, 1\cdots 10f_0) \ge 4$. In [3], we have obtained $[01f_0]_8$, $[02f_0]_9$ and $[03f_0]_{10}$. If we count the distinct symbols between $[101]_{11}$ and $[f_2f_1f_0]_n$ for $n = 8,9,10$, those numbers are at least 2 for any $(f_0)_n$. This means that $d(\mathbf{E}_{11}, \mathbf{E}_n) \ge 4$ for $n = 8,9,10$. Hence the vector with $(f_0)_{11} = 1$ is the lexicographically earliest vector of distance 4 from $\mathcal{L}_{4,4}^6$. We get $[101]_{11}$.

Now let us obtain $[11f_0]_{12}$(In [4], $[11]_{11}$). We have $(f_0)_{12} \ne 1$ because $(f_0)_{11} = 1$. Assume that $[110]_{12}$. For $n = 4,5,6,11$, $d(\mathbf{E}_n, 1\cdots 110) \ge 4$. For $\mathbf{E}_9$ with $[02f_0]_9$, we have $d(\mathbf{E}_9, 1\cdots 110) \ge 4$. For $\mathbf{E}_{10}$ with $[03f_0]_{10}$, then $d(\mathbf{E}_{10}, 1\cdots 110) \ge 4$. In the case of $[01f_0]_8$, we must have $(f_0)_8 \ne 0$. Otherwise, $wt(1\cdots 010) = 3$. Hence for any $(f_0)_8 \ne 0$, then $d(\mathbf{E}_8, 1\cdots 110) = 4$. Therefore we get $[110]_{12}$. □

Lemma 4.6. $[f_2f_1f_0]_{17} = [221]$.

Proof. From [4], $[f_2f_1]_{16} = [22]$. So we assume $[22f_0]_{17}$. If $(f_0)_{17} = 0$, then $2 \otimes [110]_{12} = [220]$. Since $d(1\cdots 22f_0, 2 \otimes \mathbf{E}_{12}) = 3$, it is a contradiction. If $(f_0)_{17} = 2$, then $2 \otimes [111]_4 = [222]$ and $d(1\cdots 22f_0, 2 \otimes \mathbf{E}_4) = 3$.

Suppose $(f_0)_{17} = 3$, i.e., $[223]_{17}$. We shall claim that $(f_0)_{15} = 3$. First, we have to obtain $[f_2f_1f_0]_{15}$. Since $[20]_{14}$, we assume $[20f_0]_{15}$. Here if $(f_0)_{15} =$

0, then $wt(1\cdots 20f_0) = 3$. If $(f_0)_{15} = 1$, then $d(1\cdots 20f_0, \mathbf{E}_{11}) = 3$ (see the proof of Lemma 4.5). In the case $(f_0)_{15} = 2$, then $d(1\cdots 20f_0, 2\otimes\mathbf{E}_{12}) = 3$ from Lemma 4.5. Hence, we get $(f_0)_{15} = 3$, i.e., $[203]_{15}$. Since $(f_0)_{15} = 3$, it contradicts to hypothesis. We get $[221]_{17}$. □

Theorems 4.7 and 4.11 show that the basis $\mathbf{E}_n$ of $\mathcal{L}_{4,4}$ has a regular form $[f_2f_1f_0]_n$ according to the range of length n.

Theorem 4.7. *For length n such that $8 \le n \le 22$, then $[f_2f_1f_0]_n$ takes over from* [011] *to* [331].

Proof. It is enough to find $[f_2f_1f_0]_8$ and $[f_2f_1f_0]_{22}$, because Table 1 of [4] gives $[f_2f_1]_n$ for $7 \le n \le 21$. Since $[01]_7$ and Lemma 4.5, we assume $[01f_0]_8$ for $(f_0)_8 \ne 0$. Thus when $(f_0)_8 = 1$, the vector with $[011]_8$ is the lexicographically earliest vector of distance 4 from $\mathcal{L}^3_{4,4}$. We get $[011]_8$.

In a similar way, we can assume that $[33f_0]_{22}$. The distance between every pair of codewords should be compared. But it is enough to find the basis $\mathbf{E}_n$ ($n < 22$) such that $\alpha \otimes [f_2f_1f_0]_n = [33f_0]_{22}$. Then we have $3\otimes[111]_4 = [333]$, $3\otimes[110]_{12} = [330]$, $2\otimes[221]_{17} = [332]$ from Lemmas 4.5 and 4.6. Hence if $(f_0)_{22} = 3, 0, 2$, $d(1\cdots 33f_0, \alpha\otimes\mathbf{E}_n) = 3$ for $n = 4, 12, 17$. It is a contradiction. Hence we get $(f_0)_{22} = 1$, i.e., $[331]_{22}$. □

Definition 4.8. *We denote the Remainder of $\mathbf{E}_n$ by $\overline{E}_n$. This means $(n-3)$ coordinates excluded $[f_2f_1f_0]$ of $\mathbf{E}_n$. For two bases $\mathbf{E}_n$ and $\mathbf{E}_{n'}$, the distinct number between $\overline{E}_n$ and $\overline{E}_{n'}$ is denoted by $D(\overline{E}_n, \overline{E}_{n'})$.*

Lemma 4.9. *For length n such that $24 \le n \le 39$, then $[f_2f_1f_0]_n$ takes over from* [001] *to* [330].

Proof. It was mentioned in Table 2 of [4] that $[f_2f_1]_i$ takes over from [00] to [33], for $23 \le i \le 38$. It is only enough to find $[f_2f_1f_0]_{24}$ and $[f_2f_1f_0]_{39}$. We assume $[00f_0]_{24}$, $([00]_{23})$. Let $4 \le n \le 22$. We consider the Remainders $\overline{E}_n$ and $\overline{E}_{24}$. Since $\mathbf{E}_{24}$ has a symbol 1 in the 7th and 23rd positions from Lemma 4.4, then $D(\overline{E}_n, \overline{E}_{24}) \ge 3$. If $(f_0)_{24} = 0$, $wt(1\cdots 00f_0) = 3$. For all n such that $4 \le n \le 22$, there is no $[00f_0]_n$ with $f_0 \ne 0$. Hence for such n, $d(\mathbf{E}_n, 1\cdots 00f_0) \ge 4$ if $f_0 \ne 0$. So when $(f_0)_{24} = 1$, the vector with $[001]_{24}$ is the lexicographically earliest vector of distance 4 from $\mathcal{L}^{18}_{4,4}$. We get $[001]_{24}$.

We assume $[33f_0]_{39}$. Then $D(\overline{E}_n, \overline{E}_{39}) = 3$ for $8 \le n \le 22$. So we need to compare the distance between $[331]_{22}$ and $[33f_0]_{39}$. Clearly, $(f_0)_{39} \ne 1$. If we consider the lexicographic ordering of $[f_2f_1f_0]$, we should have $(f_0)_{39} = (f_0)_{22} \oplus 1$. Hence, we get $[330]_{39}$. □

Lemma 4.10. *For length n such that $41 \le n \le 56$, then $[f_2 f_1 f_0]_n$ takes over from* $[000]$ *to* $[331]$.

Proof. Table 2 of [4] gives $[f_2 f_1]_n$ for $40 \le n \le 55$. So it is enough to obtain $[f_2 f_1 f_0]_{41}$ and $[f_2 f_1 f_0]_{56}$. We assume $[00f_0]_{41}$. For any $(f_0)_{41}$ and $4 \le n \le 39$, we have $D(\overline{E}_{41}, \overline{E}_n)) \ge 4$. When $(f_0)_{41} = 0$, the vector with $[000]_{41}$ is the lexicographically earliest vector of distance 4 from $\mathcal{L}^{34}_{4,4}$. Hence we get $[000]_{41}$.

We can assume $[33f_0]_{56}$. Consider $[33f_0]_n$ for $4 \le n \le 55$. By Theorem 4.7 and Lemma 4.9, we have $[331]_{22}$ and $[330]_{39}$. For $24 \le n_2 \le 39$, $D(\overline{E}_{56}, \overline{E}_{n_2}) = 3$. Hence, $(f_0)_{56} \ne (f_0)_{39}$, i.e., $(f_0)_{56} \ne 0$. For $4 \le n_1 \le 22$, $D(\overline{E}_{56}, \overline{E}_{n_1}) \ge 4$. It may allow to have $(f_0)_{56} = (f_0)_{22}$, i.e., $(f_0)_{56} = 1$. Also the vector with $[331]_{56}$ is the lexicographically earliest vector of distance 4 from $\mathcal{L}^{49}_{4,4}$. Therefore we get $[331]_{56}$. □

Theorem 4.11. *Let $p \in \mathbb{N}$ such that $17p + 7 \le n_p \le 17p + 22$.*
(1) If p is odd, then $[f_2 f_1 f_0]_{n_p}$ takes over from $[001]$ *to* $[330]$.
(2) If p is even, then $[f_2 f_1 f_0]_{n_p}$ takes over from $[000]$ *to* $[331]$.

Proof. Let $q = p+1$ and $r = p+2$ such that $17i + 7 \le n_i \le 17i + 22$ for $i = q, r$. As before, $\overline{E}_{n_p}$, $\overline{E}_{n_q}$ and $\overline{E}_{n_r}$ refer to the Remainders of $\mathbf{E}_{n_p}$, $\mathbf{E}_{n_q}$ and $\mathbf{E}_{n_r}$, respectively. Then we have $D(\overline{E}_{n_p}, \overline{E}_{n_q}) = 3$, $D(\overline{E}_{n_q}, \overline{E}_{n_r}) = 3$ and $D(\overline{E}_{n_p}, \overline{E}_{n_r}) = 4$. Hence for n_p and n_q, we must have $(f_0)_{17p+a} \ne (f_0)_{17q+a}$ for $a = 7, 8, \cdots 22$. If we consider the lexicographic ordering, $(f_0)_{17q+a} = (f_0)_{17p+a} \oplus 1$. Similarly for n_q and n_r, we have $(f_0)_{17r+a} = (f_0)_{17q+a} \oplus 1$. On the other hand, $(f_0)_{17r+a} = (f_0)_{17p+a}$. Thus we have proved. □

From Lemma 4.4, $\mathbf{E}_n$ has a symbol 1 in the nth, the 7th and the $17s+6$th positions such that $7 < 17s + 6 < n$, $s \in \mathbb{N}$.

Table 1 gives $[f_2 f_1 f_0]_n$ such that $8 \le n \le 22$ or $17p + 7 \le n \le 17p + 22$ for $2 \mid p$, $p \in \mathbb{N}$.

$f_2 f_1 f_0$	000	011	023	032	101	110	122	133
n		8	9	10	11	12	13	14
n	41	42	43	44	45	46	47	48
$f_2 f_1 f_0$	203	212	221	230	302	313	320	331
n	15	16	17	18	19	20	21	22
n	49	50	51	52	53	54	55	56
				...				

Table 1

Table 2 gives $[f_2 f_1 f_0]_n$ such that $17p + 7 \le n \le 17p + 22$ for $2 \nmid p$, $p \in \mathbb{N}$.

$f_2 f_1 f_0$	001	010	022	033	100	111	123	132
n	24	25	26	27	28	29	30	31
n	58	59	60	61	62	63	64	65
$f_2 f_1 f_0$	202	213	220	231	303	312	321	330
n	32	33	34	35	36	37	38	39
n	66	67	68	69	70	71	72	73
				...				

Table 2

For length n such that $n \geq 8$ and $n \neq 17s+6$, we introduce the following algorithm to find $\mathbf{E}_n$.

ALGORITHM

Let $p, s \in \mathbb{N}$, and $17s + 6 < n$.

Step 1 : Let n be a length such that $8 \leq n \leq 22$.
The basis $\mathbf{E}_n$ has a symbol 1 in the 7th and the nth positions. Lexicographically, $[f_2 f_1 f_0]_n$ takes the $(n-7)$th ordered form from [011] to [331](see Table 1).

Step 2 : Let n be a length such that $17p + 7 \leq n \leq 17p + 22$ and $2 \mid p$.
Then $\mathbf{E}_n$ has a symbol 1 in the 7th, the $(17s + 6)$th and the nth positions. Lexicographically, $[f_2 f_1 f_0]_n$ takes the $(n - 17p - 6)$th ordered form from [000] to [331] (see Table 1).

Step 3 : Let n be a length such that $17p + 7 \leq n \leq 17p + 22$ and $2 \nmid p$.
Then $\mathbf{E}_n$ has a symbol 1 in the 7th, the $(17s + 6)$th and the nth positions. Lexicographically, $[f_2 f_1 f_0]_n$ takes the $(n - 17p - 6)$th ordered form from [001] to [330] (see Table 2).

The following Table 3 gives $\mathbf{E}_n$ such that $n \geq 8$, $n \neq 17s + 6$.

				1	1	0	0	0	0	1	1	$= \mathbf{E}_8$
			1	0	1	0	0	0	0	2	3	$= \mathbf{E}_9$
		1	0	0	1	0	0	0	0	3	2	$= \mathbf{E}_{10}$
	1	0	0	0	1	0	0	0	1	0	1	$= \mathbf{E}_{11}$
1	0	0	0	0	1	0	0	0	1	1	0	$= \mathbf{E}_{12}$
					$\vdots$							$\vdots$

				1	1	$\cdots$	1	0	0	0	0	0	1	$= \mathbf{E}_{24}$
			1	0	1	$\cdots$	1	0	0	0	0	1	0	$= \mathbf{E}_{25}$
		1	0	0	1	$\cdots$	1	0	0	0	0	2	2	$= \mathbf{E}_{26}$
	1	0	0	0	1	$\cdots$	1	0	0	0	0	3	3	$= \mathbf{E}_{27}$
1	0	0	0	0	1	$\cdots$	1	0	0	0	1	0	0	$= \mathbf{E}_{28}$
							$\vdots$							$\vdots$

				1	1	$\cdots$	1	$\cdots$	1	0	0	0	0	0	0	$= \mathbf{E}_{41}$
			1	0	1	$\cdots$	1	$\cdots$	1	0	0	0	0	1	1	$= \mathbf{E}_{42}$
		1	0	0	1	$\cdots$	1	$\cdots$	1	0	0	0	0	2	3	$= \mathbf{E}_{43}$
	1	0	0	0	1	$\cdots$	1	$\cdots$	1	0	0	0	0	3	2	$= \mathbf{E}_{44}$
1	0	0	0	0	1	$\cdots$	1	$\cdots$	1	0	0	0	1	0	1	$= \mathbf{E}_{45}$
									$\vdots$							$\vdots$

Table 3

Let us give some examples which are helpful to find the basis of $\mathcal{L}_{4,4}$.

EXAMPLES

(1) Consider a length $n = 14$. Then $\mathbf{E}_{14}$ has a symbol 1 in the 7th and the 14th positions by Step 1, i.e., $1000\ 0001000 f_2 f_1 f_0 = \mathbf{E}_{14}$. Since $8 \le n \le 22$, $[f_2 f_1 f_0]_{14}$ takes [133] which is the 7th ordering from [011]. Therefore, we get $1000\ 0001000133 = \mathbf{E}_{14}$.

(2) Let $n = 49$. When $s = 1, 2$, it satisfies $17s + 6 < 49$. So $\mathbf{E}_{49}$ has a symbol 1 in the 7th, the 23rd, the 40th and the 49th positions, by Step 2. Thus we get $100000000\ 1000000000\ 0000000100\ 0000000000\ 0001000 f_2 f_1 f_0 = \mathbf{E}_{49}$.

When $p = 2$, it satisfies $17p + 7 \le 49 \le 17p + 22$. So $[f_2 f_1 f_0]_{49}$ takes [203] which is the 9th ordering from [000]. Hence, we get $100000000\ 1000000000\ 0000000100\ 0000000000\ 0001000203 = \mathbf{E}_{49}$.

(3) Let $n = 59$. When $s = 1, 2, 3$ it satisfies $17s + 6 < 59$. So $\mathbf{E}_{59}$ has a symbol 1 in the 7th, the 23rd, the 40th, the 57th and the 59th positions, by Step 3. So $101000000\ 0000000000\ 1000000000\ 0000000100\ 0000000000\ 0001000 f_2 f_1 f_0 = \mathbf{E}_{59}$.

When $p = 3$, it satisfies $17p + 7 \le 59 \le 17p + 22$. So $[f_2 f_1 f_0]_{59}$ takes [010] which is the 2nd ordering from [001]. Hence, we get $101000000\ 0000000000\ 1000000000\ 0000000100\ 0000000000\ 0001000010 = \mathbf{E}_{59}$.

5. Decoding

In this section, we describe a decoding algorithm of $\mathcal{L}_{4,4}$ using the testing vector, and give its examples.

Definition 5.1. *Given a received vector $r_{n-1}\cdots r_2r_1r_0 = \mathbf{r}$ over $GF(4)$, The testing vector* $\mathbf{t}$ *of $\mathcal{L}_{4,4}$ is defined by $\bigoplus_{k=4}^{n} (r_{k-1}\otimes\mathbf{E}_k)$ where $k \neq 7, 17s+6,\ s \in \mathbb{N}$.*

Theorem 5.2. *Let* $\mathbf{r}$, $\mathbf{t}$ *be the received and the testing vector, respectively. $d(\mathbf{r},\mathbf{t}) = 1$ if and only if one of r_{i-1}'s is not correct for $i = 1,2,3,7,17s+6,\ s \in \mathbb{N}$.*

Proof. ($\Rightarrow$) Given a received $r_{n-1}\cdots r_2r_1r_0 = \mathbf{r}$, an error-corrupted vector, suppose r_{i-1} is correct for all $i = 1,2,3,7,17s+6$. That is, let r_{l-1} $(l \neq i)$ be an error symbol. If $4 \leq l \leq 6$, all symbols of $[f_2f_1f_0]_l$ are nonzero. Thus, $d(\mathbf{r},\mathbf{t}) \geq 4$. If $l \geq 8$ and $l \neq 34s+7$, then $[f_2f_1f_0]_l$ has of nonzero symbol no less than 1. In addition, $\mathbf{E}_l$ has a symbol 1 in the 7th and the $17s+6$th positions. Thus, $d(\mathbf{r},\mathbf{t}) \geq 2$. Finally if $l = 34s+7$, we have $[f_2f_1f_0]_l = [000]$. But $\mathbf{E}_l$ has a symbol 1 in the positions no less than 3. Hence, $d(\mathbf{r},\mathbf{t}) \geq 3$. Therefore in any case, $d(\mathbf{r},\mathbf{t}) \neq 1$.
($\Leftarrow$) Since there does not exist $\mathbf{E}_i$ corresponding to r_{i-1} in $\mathbf{t}$, so $\mathbf{t}$ is not affected by an error symbol r_{i-1}. Hence $d(\mathbf{r},\mathbf{t}) = 1$. □

Corollary 5.3. *As for Theorem 5.2, let us have* $\mathbf{r}$ *and* $\mathbf{t}$. *If $d(\mathbf{r},\mathbf{t}) > 1$, then one of r_{i-1}'s is not correct for $i \geq 4$ and $i \neq 7, 17s+6,\ s \in \mathbb{N}$.*

In the following remark, we explain how to find a decoding algorithm in more detail.

Remark 5.4. Given $r_{n-1}\cdots r_2r_1r_0 = \mathbf{r}$, then we obtain $t_{n-1}\cdots t_2\,t_1t_0 = \mathbf{t}$ from Definition 5.1. Let $c_{n-1}\cdots c_1c_0 = \mathbf{c}$ be a desired codeword.

(A) If $d(\mathbf{r},\mathbf{t}) = 1$, clearly $\mathbf{t}$ is obtained by sum of the bases which do not depend on error symbol. Therefore, we have the desired codeword $\mathbf{c} = \mathbf{t}$.

If $d(\mathbf{r},\mathbf{t}) > 1$, from Corollary 5.3 there is r_{k-1} such that $r_{k-1} \neq c_{k-1}$ for $k \geq 4$ and $k \neq 7, 17s+6$. We consider the following four cases.

(B) In the case of $c_{k-1} = 0$ and $r_{k-1} \neq 0$ $(k \leq n)$, then $r_{k-1}\otimes\mathbf{E}_k$ must be deleted in $\mathbf{t}$. Hence, we obtain $\mathbf{c} = \mathbf{t} \oplus (r_{k-1}\otimes\mathbf{E}_k)$. On the other hand, we replace r_{k-1} by 0 in $\mathbf{r}$.

Here, there is $r_{k-1}\otimes\mathbf{E}_k$ with $[d_2d_1d_0]$ in $\mathbf{t}$ such that $[t_2t_1t_0] \oplus [d_2\,d_1d_0] = [r_2r_1r_0]$.

Let r'_{k-1} be a nonzero correct symbol, i.e., $r'_{k-1} = c_{k-1}$.

(C) In the case of $c_{k-1} \neq 0$ and $r_{k-1} = 0$ $(k \leq n)$, then $r'_{k-1} \otimes \mathbf{E}_k$ must be added to $\mathbf{t}$. Hence, we obtain $\mathbf{c} = \mathbf{t} \oplus (r'_{k-1} \otimes \mathbf{E}_k)$.

Here, there is no $r_{i-1} \otimes \mathbf{E}_i$ with $[d_2\ d_1 d_0]$ in $\mathbf{t}$ such that $[t_2 t_1 t_0] \oplus [d_2\ d_1 d_0] = [r_2 r_1 r_0]$. But there is $r'_{k-1} \otimes \mathbf{E}_k$ with $[d_2 d_1 d_0]$ for $k \leq n$. As a result, we have a nonzero r'_{k-1}. Therefore, we replace 0 by r'_{k-1} in $\mathbf{r}$.

(D) In the case of $c_{k-1}, r_{k-1} \neq 0$ and $r_{k-1} \neq c_{k-1}$ $(k \leq n)$, then $r_{k-1} \otimes \mathbf{E}_k$ must be deleted in $\mathbf{t}$, and $r'_{k-1} \otimes \mathbf{E}_k$ must be added to $\mathbf{t}$. Hence, we obtain $\mathbf{c} = \mathbf{t} \oplus (r_{k-1} \otimes \mathbf{E}_k) \oplus (r'_{k-1} \otimes \mathbf{E}_k)$.

In order to find r'_{k-1}, we should know that $r_{k-1} \neq c_{k-1}$. There is no $r_{i-1} \otimes \mathbf{E}_i$ with $[d_2 d_1 d_0]$ in $\mathbf{t}$ such that $[t_2 t_1 t_0] \oplus [d_2\ d_1 d_0] = [r_2 r_1 r_0]$. And if there is $\alpha \in \mathrm{GF}(4)$ such that $\alpha \otimes (r_{k-1} \otimes [f_2 f_1 f_0]_k) = [d_2 d_1 d_0]$, then $r_{k-1} \neq c_{k-1}$, i.e., r_{k-1} is not correct. From the above equation $\mathbf{c} = \mathbf{t} \oplus (r_{k-1} \otimes \mathbf{E}_k) \oplus (r'_{k-1} \otimes \mathbf{E}_k)$, we get r'_{k-1} such that $(r_{k-1} \oplus r'_{k-1}) \otimes [f_2 f_1 f_0]_k = [r_2 r_1 r_0] \oplus [t_2 t_1 t_0]$, because $[c_2 c_1 c_0] = [r_2 r_1 r_0]$. Therefore, we replace r_{k-1} by r'_{k-1} in $\mathbf{r}$.

(E) In the case of $c_{k-1} \neq 0$ and $r_{k-1} = 0$ $(k > n)$, then $r'_{k-1} \otimes \mathbf{E}_k$ must be added to $\mathbf{t}$. Hence, we obtain $\mathbf{c} = \mathbf{t} \oplus (r'_{k-1} \otimes \mathbf{E}_k)$.

Here, there is no vector $(r_{i-1} \otimes \mathbf{E}_i)$ with $[d_2 d_1 d_0]$ in $\mathbf{t}$. And there is no vector $\alpha \otimes (r_{i-1} \otimes \mathbf{E}_i)$ with $[d_2 d_1 d_0]$ for all $i \leq n$. We should find r'_{k-1} for $k > n$. If $n \leq 7$, we obtain $r'_{k-1} \otimes \mathbf{E}_k$ such that $r'_{k-1} \otimes [f_2 f_1 f_0]_k = [d_2 d_1 d_0]$ for $8 \leq k \leq 22$. If $7 < n \leq 23$, we obtain $r'_{k-1} \otimes \mathbf{E}_k$ with $[d_2 d_1 d_0]$ for $24 \leq k \leq 39$. If $n > 23$ and $17p+6 < n \leq 17p+23$, we obtain $r'_{k-1} \otimes \mathbf{E}_k$ with $[d_2 d_1 d_0]$ for $17p + 24 \leq k \leq 17p + 39$. Therefore, we replace 0 by r'_{k-1} in $\mathbf{r}$.

DECODING ALGORITHM

Step 1 : Suppose $d(\mathbf{r}, \mathbf{t}) = 1$.
Then $\mathbf{c} = \mathbf{t}$. Otherwise, i.e., $d(\mathbf{r}, \mathbf{t}) > 1$, we go to Step 2.

Step 2 : Suppose $d(\mathbf{r}, \mathbf{t}) > 1$.
If there is $r_{k-1} \otimes \mathbf{E}_k$ with $[d_2 d_1 d_0]$ in $\mathbf{t}$, then $\mathbf{c} = \mathbf{t} \oplus (r_{k-1} \otimes \mathbf{E}_k)$. Otherwise, we go to Step 3.

Step 3 : Suppose there is no $r_{i-1} \otimes \mathbf{E}_i$ with $[d_2 d_1 d_0]$ in $\mathbf{t}$. (Here, $r_{k-1} = 0$)
If there is $r'_{k-1} \neq 0$ such that $r'_{k-1} \otimes [f_2 f_1 f_0]_k = [d_2 d_1 d_0]$ for $k \leq n$, then $\mathbf{c} = \mathbf{t} \oplus (r'_{k-1} \otimes \mathbf{E}_k)$. Otherwise, we go to Step 4.

Step 4 : Suppose $r_{k-1} \neq 0$ and there is no r_{k-1} such that $r_{k-1} \otimes [f_2 f_1 f_0]_k = [d_2 d_1 d_0]$ for $k \leq n$.

If there is $\alpha \otimes (r_{k-1} \otimes \mathbf{E}_k)$ with $[d_2 d_1 d_0]$ for $k \leq n$, we can get r'_{k-1} $(\neq r_{k-1})$ such that $(r_{k-1} \oplus r'_{k-1}) \otimes [f_2 f_1 f_0]_k = [r_2 r_1 r_0] \oplus [t_2 t_1 t_0]$. Then $\mathbf{c} = \mathbf{t} \oplus (r_{k-1} \otimes \mathbf{E}_k) \oplus (r'_{k-1} \otimes \mathbf{E}_k)$. Otherwise, we go to Step 5.

Step 5 : Suppose there is no vector $\alpha \otimes (r_{i-1} \otimes \mathbf{E}_i)$ with $[d_2 d_1 d_0]$ for all $i \leq n$. (Here, $r_{k-1} = 0, k > n$)

We can get $r'_{k-1} \neq 0$ such that $r'_{k-1} \otimes [f_2 f_1 f_0]_k = [d_2 d_1 d_0]$ from Remark(E). Then $\mathbf{c} = \mathbf{t} \oplus (r'_{k-1} \otimes \mathbf{E}_k)$.

EXAMPLES

(1) Given a received vector 2200 0000000000 1003300312 = $\mathbf{r}$, we get $(2 \otimes \mathbf{E}_{24}) \oplus (1 \otimes \mathbf{E}_{10}) \oplus (3 \otimes \mathbf{E}_6)$ = 2200 0000000000 1003300311 = $\mathbf{t}$. Since $d(\mathbf{r}, \mathbf{t}) = 1$, we have the desired codeword $\mathbf{c} = \mathbf{t}$.

(2) Given 2200 0000000002 1003300311 = $\mathbf{r}$, we get $(2 \otimes \mathbf{E}_{24}) \oplus (2 \otimes \mathbf{E}_{11}) \oplus (1 \otimes \mathbf{E}_{10}) \oplus (3 \otimes \mathbf{E}_6)$ = 2200 0000000002 1001300113 = $\mathbf{t}$. From [113] $\oplus$ $[d_2 d_1 d_0]$ = [311], $[d_2 d_1 d_0]$ = [202]. Since $d(\mathbf{r}, \mathbf{t}) > 1$ and there is $2 \otimes \mathbf{E}_{11}$ with [202] in $\mathbf{t}$, therefore $\mathbf{t} \oplus (2 \otimes \mathbf{E}_{11})$ = 2200 0000000000 1003300311 = $\mathbf{c}$.

(3) Given 2200 0000000000 3001300120 = $\mathbf{r}$, we get $(2 \otimes \mathbf{E}_{24}) \oplus (3 \otimes \mathbf{E}_{10}) \oplus (3 \otimes \mathbf{E}_6)$ = 2200 0000000000 3001300302 = $\mathbf{t}$. Using [302] $\oplus$ $[d_2 d_1 d_0]$ = [120], we have $[d_2 d_1 d_0]$ = [222]. Then $d(\mathbf{r}, \mathbf{t}) > 1$ and there is no $(r_{i-1} \otimes \mathbf{E}_i)$ with [222] in $\mathbf{t}$. But there is $2 \otimes \mathbf{E}_4$ with [222]. Therefore, $\mathbf{t} \oplus (2 \otimes \mathbf{E}_4)$ = 2200 0000000000 3001302120 = $\mathbf{c}$.

(4) Given 2200 0000000000 3003300311 = $\mathbf{r}$, we get $(2 \otimes \mathbf{E}_{24}) \oplus (3 \otimes \mathbf{E}_{10}) \oplus (3 \otimes \mathbf{E}_6)$ = 2200 0000000000 3001300302 = $\mathbf{t}$. Hence, we have $[d_2 d_1 d_0]$ = [013] such that [302] $\oplus$ $[d_2 d_1 d_0]$ = [311]. So $d(\mathbf{r}, \mathbf{t}) > 1$ and there is no $r_{i-1} \otimes \mathbf{E}_i$ with [013] in $\mathbf{t}$. When $\alpha = 3$ and $r_9 = 3$, it satisfies $\alpha \otimes (r_9 \otimes [032]_{10}) = [013]$. Thus $r_9 \neq c_9$. Now we can get $r'_9 = 1$ satisfying $(r_9 \oplus r'_9) \otimes [032]_{10} = [311] \oplus [302]$. Therefore, $\mathbf{t} \oplus ((3 \oplus 1) \otimes \mathbf{E}_{10})$ = 2200 0000000000 1003300311 = $\mathbf{c}$.

(5) Given 0200 0000000000 1003300311 = $\mathbf{r}$, we get $(1 \otimes \mathbf{E}_{10}) \oplus (3 \otimes \mathbf{E}_6)$ = 1001300313 = $\mathbf{t}$, and $[d_2 d_1 d_0]$=[002]. Then $d(\mathbf{r}, \mathbf{t}) > 1$ and there is no $r_{i-1} \otimes \mathbf{E}_i$ with [002] in $\mathbf{t}$. Also, there is no vector $\alpha \otimes (r_{i-1} \otimes \mathbf{E}_i)$ with [002] for $i \leq 23$. Since $n = 23$, we can obtain $2 \otimes \mathbf{E}_{24}$ such that $2 \otimes [001]_{24}$ = [002]. Therefore, $\mathbf{t} \oplus (2 \otimes \mathbf{E}_{24})$ = 2200 0000000000 1003300311 = $\mathbf{c}$.

References

[1] J.H. Conway, Integral Lexicographic Codes, Discrete Mathematics 83(1990), pp. 219-235.

[2] J.H. Conway, On Numbers and Games, Academic Press, New York (1976).
[3] J.H. Conway and N. J. A. Sloane, Lexicographic Codes: Error-Correcting Codes from Game Theory, IEEE Trans. Inform. Theory IT-32(3)(1986), pp. 337-348.
[4] D.G. Kim and H.K. Kim, A decoding scheme for the 4-ary lexicodes with d=3, IJMMS 29:5(2002), pp. 265-270.