Chapter 6 Codes with burst distance and periodical burst errors

The contents of this chapter are based on the paper:

• Das, P. K. and Haokip, L. Codes with burst distance and periodical burst errors. *Journal of Computational and Applied Mathematics*, 411:114240, 2022.

Chapter 6

Codes with burst distance and periodical burst errors

This chapter presents a study on the burst-b distance d_b of a linear code and provides periodical burst error detection and correction capability for a linear code equipped with minimum burst-b distance d_b . An (n,k) linear code with minimum burst-b distance d_b is written as (n,k,d_b) linear code. In [79], Villalba et al. extend the study on burst-b distance and derive an extended Reiger-Singleton bound for an (n,k,d_b) linear code as

$$b(d_b - 1) \le n - k. \tag{6.1}$$

Codes with $b(d_b - 1) = n - k$ are called Maximum Distance Separable (MDS) codes. A class of Maximum Distance Separable (MDS) codes is presented in [79] as below.

Lemma 6.1. [79] The code generated by the $b \times bd_b$ matrix G_b :

where 1's in each row are separated by exactly b-1 zeros, is a (bd_b, b, d_b) MDS code.

Observe that the code given in Lemma 6.1 is a cyclic code and every codeword has

same burst-b weight d_b . So, the code is a constant burst-b weight code. We denote the above (bd_b, b, d_b) code by C_b code.

This chapter investigates the periodical burst error detection and correction capability of the code C_b and its dual code along with a decoding procedure of the code C_b . The contents of this chapter are organised as follows. Section 6.1 gives bounds on d_b for a linear code and a constant burst-b weight linear code. Then, we give the connection between an (n, k, d_b) and $(n - b, k, d_b - 1)$ codes along with the cardinality of (n, k, d_b) code. We also provide a connection between linearly independent columns of the parity check matrix of any MDS (n, k, d_b) codes. Section 6.2 presents periodical burst error detection and correction capability of a linear (n, k, d_b) code. The same is investigated for the MDS code C_b . Finally, Section 6.3 gives a decoding procedure for the code C_b in the case of periodical burst error.

6.1 Codes with burst-b distance

In this section, we give bounds on d_b for a linear code and a constant burst-b weight linear code. Then, we give the existence of a linear code with minimum distance $d_b - 1$ from a linear code with minimum distance d_b . Further, an upper bound on the cardinality of a linear code with distance d_b is derived. The connection between linearly independent columns of the parity check matrix of any MDS code with burst-b distance is derived.

Theorem 6.2. Let C be an (n, k, d_b) linear code over GF(q). Then

$$d_b \le \frac{nq^{k-1}(q-1)}{q^k - 1}.$$

Proof. From Result 1.12, the total number of nonzero components in all the codewords of C is

$$nq^{k-1}(q-1).$$
 (6.2)

Again, the code has the minimum burst-b distance as d_b , so the number of nonzero components in each nonzero codeword is at least d_b . Therefore, the total number of

nonzero components in all the codewords of C is at least

$$(q^k-1)d_b$$
.

Hence

$$(q^k - 1)d_b \le nq^{k-1}(q - 1).$$

This proves the theorem.

Remark 6.3. This upper bound of d_b coincides with Result 1.12 when the distance is taken in Hamming sense for a q-ary (n, k, d) linear code.

Theorem 6.4. Let C be a constant burst-b weight (n, k, d_b) linear code over GF(q). Then

$$d_b \ge \frac{nq^{k-1}(q-1)}{(q^k-1)b}. (6.3)$$

Proof. Since every codeword of the code C has burst-b distance d_b , the number of nonzero components in each nonzero codeword is at most bd_b . Therefore, the total number of nonzero components in all the codewords of C can be at most

$$(q^k-1)bd_b.$$

But the total number of nonzero components in all the codewords of C is given by (6.2). Hence

$$(q^k - 1)bd_b \ge nq^{k-1}(q - 1).$$

This proves the theorem.

Remark 6.5. For constant burst-b weight (n, k, d_b) linear code over GF(q), Theorem 6.2-6.4 give

$$\frac{nq^{k-1}(q-1)}{(q^k-1)b} \le d_b \le \frac{nq^{k-1}(q-1)}{q^k-1}.$$
(6.4)

Remark 6.6. For constant burst-b weight code C_b , k = b and q = 2. Then

$$\frac{nq^{k-1}(q-1)}{(q^k-1)b} = \frac{bd_b 2^{b-1}}{(2^b-1)b} \le d_b. \quad [\because \quad \frac{2^{b-1}}{(2^b-1)} \le 1]$$

Further, for b = 1

$$\frac{nq^{k-1}(q-1)}{(q^k-1)} = \frac{d_12^{1-1}}{(2^1-1)} = d_1,$$

and for $b \geq 2$

$$\frac{nq^{k-1}(q-1)}{(q^k-1)} = \frac{bd_b 2^{b-1}(2-1)}{(2^b-1)} = \frac{bd_b 2^{b-1}}{(2^b-1)} \ge d_b. \quad [\because \quad \frac{b2^{b-1}}{(2^b-1)} \ge 1]$$

Thus

$$\frac{nq^{k-1}(q-1)}{(q^k-1)} \ge d_b.$$

Therefore, the bound 6.4 is satisfied by the code C_b .

Theorem 6.7. If a q-ary (n, k, d_b) linear code C exists with $d_b \ge 2$, there also exists an $(n - b, k, d_b - 1)$ linear code C'.

Proof. There exist codewords x and y of C such that $d_b(x,y) = w_b(x-y) = d_b$. From the d_b sets in which nonzero components of x-y are confined, we choose a set of b consecutive components. Delete the b consecutive components of the set in each codeword of C, the resultant vectors will have burst-b weight at least $d_b - 1 (\geq 1)$ with one vector x-y having exactly burst-b weight $d_b - 1$, and they will form a subspace C' of C with q^k elements. This code C' is the required $(n-b,k,d_b-1)$ linear code C'.

Theorem 6.8. The cardinality M of a linear code of length n with minimum burst-b distance d_b is bounded above by

$$M \le \frac{q^n}{\sum_{i=0}^{(d_b-1)/2} \binom{n-ib+i}{i} (q-1)^i q^{i(b-1)}}.$$

Proof. By [23], the number of $(d_b-1)/2$ bursts of length up to b is given by

$$\frac{q^n}{\sum_{i=0}^{(d_b-1)/2} \binom{n-ib+i}{i} (q-1)^i q^{i(b-1)}}$$

By Result 1.18, the code can correct all $(d_b-1)/2$ bursts of length up to b. As every codeword disturbed by such errors must produce distinct words, we have

$$q^n \ge M \sum_{i=0}^{(d_b-1)/2} {n-ib+i \choose i} (q-1)^i q^{i(b-1)}$$

i.e.

$$M \le \frac{q^n}{\sum_{i=0}^{(d_b-1)/2} \binom{n-ib+i}{i} (q-1)^i q^{i(b-1)}}.$$

Theorem 6.9. Let C be an (n, k, d_b) code over GF(q) and n - k is a multiple of b. Let H be the parity check matrix of C. Then C is an MDS code if and only if every set of n - k columns that are formed from $\frac{n - k}{b}$ sets of b consecutive columns of H are linearly independent.

Proof. If C is an MDS code, then $b(d_b-1)=n-k$, which implies $d_b-1=\frac{n-k}{b}$, an integer. As every vector of burst-b weight d_b-1 cannot be a codeword, every linear combination of columns of H which consists of d_b-1 sets of b consecutive columns is nonzero. Also, the total number of columns in d_b-1 sets of b consecutive columns is n-k, so every set of n-k columns that are formed from $\frac{n-k}{b}$ sets of b consecutive columns of H is linearly independent. Now, if every set of n-k columns that are formed from $\frac{n-k}{b}$ sets of b consecutive columns are linearly independent, we cannot get a codeword whose nonzero components are confined to $\frac{n-k}{b}$ sets of b consecutive columns. The $\frac{n-k}{b}$ sets of b consecutive columns may be within the first n-k columns of H or not. In either case, the burst-b distance of the code C is at least $\frac{n-k}{b}+1$, i.e.

$$d_b \ge \frac{n-k}{b} + 1$$

$$\implies b(d_b - 1) \ge n - k.$$

Again from Equation (6.1), we have $b(d_b - 1) \le n - k$. So, $b(d_b - 1) = n - k$ and hence C is an MDS code.

6.2 Burst-b distance and periodical burst errors

In this section, we present periodical burst detection and correction of an (n, k, d_b) linear code. Then we do the same investigation for code C_b and its dual code C_b^{\perp} .

Theorem 6.10. An (n, k, d_b) code can detect all vectors of $\psi_{(s,b),n,q}$ that start from the $(n+1-(d_b-1)(s+b))^{th}$ position.

Proof. Consider a vector of the error set $\psi_{(s,b),n,q}$ defined in Chapter 4 whose nonzero

components start from the j^{th} position (j = 1, 2, ..., n), so the errors are confined to the last n - j + 1 positions.

By Euclidean division algorithm, for integers n-i+1 and s+b, there exist integers λ_j and r_j such that

$$n - j + 1 = \lambda_j(s + b) + r_j$$
, where $0 \le r_j < s + b$. (6.5)

So, every vector of $\psi_{(s,b),n,q}$, where the error starts from the j^{th} position, has burst-b weight

$$\begin{cases} \left\lfloor \frac{n-j+1}{s+b} \right\rfloor \text{ if } r_j + p \leq b & (p: \text{the last nonzero position of the first nonzero set}) \\ \left\lceil \frac{n-j+1}{s+b} \right\rceil & \text{otherwise.} \end{cases}$$

An (n, k, d_b) code can detect error vectors of $\psi_{(s,b),n,q}$ if its burst-b weight is less than or equal to $d_b - 1$. So, if

$$d_b - 1 \ge \left\lceil \frac{n - j + 1}{s + b} \right\rceil$$
i.e., $(d_b - 1)(s + b) \ge n - j + 1$
i.e., $j \ge n + 1 - (d_b - 1)(s + b)$,

the code detects any error vector of $\psi_{(s,b),n,q}$ that starts from the j^{th} position. Therefore, if the starting position of the error pattern of $\psi_{(s,b),n,q}$ is $n+1-(d_b-1)(s+b)$, the code detects such errors.

Taking j = 1, we have $1 = n + 1 - (d_b - 1)(s + b)$, i.e., $d_b = \frac{n}{s + b} + 1$. Therefore, we have the following corollary.

Corollary 6.11. An (n, k, d_b) code detects all error vectors of $\psi_{(s,b),n,q}$ provided $d_b \ge \lceil \frac{n}{s+b} \rceil + 1$.

Theorem 6.12. An (n, k, d_b) code corrects all error vectors of $\psi_{(s,b),n,q}$ that start from the $\left(\left\lceil n+1-\frac{(d_b-1)}{2}(s+b)\right\rceil\right)^{th}$ position.

Proof. By Result 1.18, an (n, k, d_b) code can correct up to $(d_b - 1)/2$ bursts of length up to b each. Since every error vector of $\psi_{(s,b),n,q}$ that starts from j^{th} position (j = 1, 2, ..., n) has burst-b weight $\lceil \frac{n-j+1}{s+b} \rceil$ or less, an (n, k, d_b) code can correct an error vector of $\psi_{(s,b),n,q}$ if its burst-b weight is less than or equal to $(d_b - 1)/2$. Now

$$\frac{d_b - 1}{2} \ge \left\lceil \frac{n - j + 1}{s + b} \right\rceil$$

implies
$$\frac{(d_b - 1)}{2}(s + b) \ge n - j + 1$$

implies $j \ge n + 1 - \frac{(d_b - 1)}{2}(s + b),$

Therefore, if the starting position of an error vector of $\psi_{(s,b),n,q}$ is $\lceil n+1-\frac{(d_b-1)}{2}(s+b) \rceil$, the code corrects such errors.

Again taking j = 1, we have $1 = \left\lceil n + 1 - \frac{(d_b - 1)}{2}(s + b) \right\rceil$, i.e., $d_b \ge 2 \left\lceil \frac{n}{s + b} \right\rceil + 1$. This gives the following corollary:

Corollary 6.13. An (n, k, d_b) code can correct vectors of $\psi_{(s,b),n,q}$ provided $d_b \ge 2\left\lceil \frac{n}{s+b}\right\rceil + 1$.

Theorem 6.14. The code C_b $(d_b \ge 2)$ detects all

- (i) vectors of the error set $\psi_{(s,b),n,q}$ for any $s \geq b$, and
- (ii) vectors of the error set $\psi_{(s,b),n,q}$ that start from $(b+1)^{th}$ position for any s.

Proof. As $n = bd_b$ and $s \ge b$, then $\lceil \frac{n}{s+b} \rceil + 1 \le \lceil \frac{bd_b}{2b} \rceil + 1 = \lceil \frac{d_b}{2} \rceil + 1 \le d_b$ for $d_b \ge 2$. From Corollary 6.11, the code C_b detects all error vectors of $\psi_{(s,b),n,q}$ for any $s \ge b$. Further, if the error vector of $\psi_{(s,b),n,q}$ starts from the $(b+1)^{th}$ position for any s, then n becomes $(d_b-1)b$ and $\frac{b}{s+b} \le 1$, so

$$\left\lceil \frac{n}{s+b} \right\rceil + 1 = \left\lceil \frac{b(d_b - 1)}{s+b} \right\rceil + 1 \le d_b - 1 + 1 = d_b.$$

This proves the part (ii) by Corollary 6.11.

Theorem 6.15. The number of vectors of the error set $\psi_{(s,b),n,q}$ that go undetected by the code C_b is 2^b .

Proof. From Theorem 6.14, a vector of $\psi_{(s,b),n,q}$ may go undetected if it starts within the first b positions and s < b. Every codeword of C_b is a linear combination of rows of G_b , so every codeword can be written as

$$(a_0 + a_1X + a_2X^2 + \dots + a_jX^j) + X^b(a_0 + a_1X + a_2X^2 + \dots + a_jX^j) + \dots + X^{(d_b-1)b}(a_0 + a_1X + a_2X^2 + \dots + a_jX^j),$$

where $a_i \in GF(2)$ and $0 \le j \le b-1$.

Clearly, every nonzero codeword is a vector of $\psi_{(s,b),n,q}$ that starts within the first b

positions and s < b. Therefore, the number of vectors of $\psi_{(s,b),n,q}$ that go undetected by C_b is 2^b .

Theorem 6.16. The code C_b corrects all

(i) vectors of the error set $\psi_{(s,b),n,q}$ for any $s \geq 3b$, and

(ii) vectors of the set
$$\psi_{(s,b),n,q}$$
 that start from $\left(\left\lceil \frac{(d_b+1)b}{2}\right\rceil + 1\right)^{th}$ position for any s.

Proof. Since the difference between any two vectors of $\psi_{(s,b),n,q}$ for $s \geq 3b$ is a vector which contains at least b consecutive zeros, so the difference cannot be a codeword of C_b . Therefore, the code C_b corrects all error vectors of $\psi_{(s,b),n,q}$ for any $s \geq 3b$. This can be verified from Corollary 6.13 also:

As $n = bd_b$ and $s \ge 3b$, then $2\lceil \frac{n}{s+b} \rceil + 1 \le 2\lceil \frac{bd_b}{4b} \rceil + 1 = 2\lceil \frac{d_b}{4} \rceil + 1 \le d_b$ for $d_b \ge 3$.

Further, if an error vector of $\psi_{(s,b),n,q}$ starts from $\left(\left\lceil \frac{(d_b+1)b}{2}\right\rceil + 1\right)^{th}$ position, then n becomes $bd_b - \left\lceil \frac{(d_b+1)b}{2}\right\rceil \leq \frac{b(d_b-1)}{2}$. So

$$2\left\lceil \frac{n}{s+b}\right\rceil + 1 \le 2\left\lceil \frac{b(d_b-1)}{2(s+b)}\right\rceil + 1 \le d_b. \qquad (\because \frac{b}{s+b} \le 1)$$

This proves the part (ii).

Theorem 6.17. The dual code C_b^{\perp} of C_b is also a $(bd_b, b(d_b - 1), 2)$ MDS code.

Proof. Since the code C_b is of order (bd_b, b) , the order of its dual code C_b^{\perp} is $(bd_b, b(d_b - 1))$ and its generator matrix H_b is given by

$$H_b = d_b - 1 \left\{ egin{bmatrix} I_b \ I_b \ dots & I_{b(d_b-1)} \ I_b \ I_b \ \end{pmatrix},$$

where I_b represents the identity matrix of order b.

Clearly, the minimum burst-b distance of C_b^{\perp} is $d_b' = 2$. So, for the dual code C_b^{\perp} , $n = bd_b$, $k = b(d_b - 1)$ and $d_b' = 2$. Now $b(d_b' - 1) = b = n - k$. Therefore, the dual code C_b^{\perp} is also an MDS code.

Theorem 6.18. If $d_b = 2$, the dual code C_b^{\perp} of C_b

- (i) detects all vectors of the error set $\psi_{(b,b),n,q}$ and
- (ii) corrects all vectors of the set $\psi_{(b,b),n,q}$ that start from the $(b+1)^{th}$ position.

Proof. From the previous theorem, we have that the length and the minimum burst-b distance of C_b^{\perp} are bd_b and $d_b' = 2$ respectively. Now taking s = b, we have

$$\left\lceil \frac{n}{s+b} \right\rceil + 1 = \left\lceil \frac{bd_b}{2b} \right\rceil + 1 = \left\lceil \frac{d_b}{2} \right\rceil + 1 \le d_b'.$$

So, by Corollary 6.11, the code C_b^{\perp} detects all vectors of $\psi_{(b,b),n,q}$ provided $d_b = 2$. Again

$$\lceil n+1 - \frac{(d_b'-1)}{2}(s+b) \rceil = \lceil 2b+1 - \frac{(2-1)}{2}(b+b) \rceil = b+1.$$

So, by Theorem 6.12, the dual code C_b^{\perp} corrects all vectors of $\psi_{(b,b),n,q}$ that start from the $(b+1)^{th}$ position.

6.3 Decoding of the MDS code C_b

In this section, we give a decoding method for periodical burst error by the code C_b . Suppose that v is a sent codeword of C_b and $w = (w_1, w_2, \ldots, w_{bd_b})$ is received after any vector of $\psi_{(s,b),n,q}$ (where $s \geq 3b$) which is $e = (e_1, e_2, \ldots, e_{bd_b})$. Then the sent codeword is v = w - e. In the following, we describe how to get the error from the received vector w. Now the syndrome of w is

$$S = wH_b^{\perp}$$

$$= eH_b^{\perp}$$

$$= \left(e_1 + e_{b+1}, e_2 + e_{b+2}, \dots, e_b + e_{2b}, e_1 + e_{2b+1}, e_2 + e_{2b+2}, \dots, e_b + e_{3b}, \dots, \dots, e_1 + e_{(d_b-1)b+1}, e_2 + e_{(d_b-1)b+2}, \dots, e_b + e_{d_bb}\right)$$

$$= \left(s_1, s_2, \dots, s_b, s_{b+1}, s_{b+2}, \dots, s_{2b}, \dots, \dots, s_{(d_b-2)b+1}, s_{(d_b-2)b+2}, \dots, s_{(d_b-1)b}\right)$$

If the periodical burst error starts from j^{th} position $(1 \le j \le b)$, all the three b-tuples $(s_1, s_2, \ldots, s_b), (s_{b+1}, s_{b+2}, \ldots, s_{2b}), (s_{2b+1}, s_{2b+2}, \ldots, s_{3b})$ of b consecutive components of S will be nonzero tuples and also $e_{j+b} = e_{j+b+1} = \cdots = e_{j+4b-1} = 0$. Taking the

majority of the three sets of b consecutive components gives us (e_1, e_2, \ldots, e_b) . Then subtracting (e_1, e_2, \ldots, e_b) from each set of b consecutive components of S gives the remaining $(d_b - 1)b$ components of e, i.e.

$$\begin{pmatrix} e_{b+1}, e_{b+2}, \dots, e_{2b}, e_{2b+1}, e_{2b+2}, \dots, e_{3b}, \dots, e_{(d_b-1)b+1}, e_{(d_b-1)b+2}, \dots, e_{d_bb} \end{pmatrix}$$

$$= \begin{pmatrix} \underbrace{s_1, s_2, \dots, s_b}_{b}, \underbrace{s_{b+1}, s_{b+2}, \dots, s_{2b}}_{b}, \dots, \underbrace{s_{(d_b-2)b+1}, s_{(d_b-2)b+2}, \dots, s_{(d_b-1)b}}_{b} \end{pmatrix}$$

$$- \begin{pmatrix} \underbrace{e_1, e_2, \dots, e_b}_{b}, \underbrace{e_1, e_2, \dots, e_b}_{b}, \dots, \underbrace{e_1, e_2, \dots, e_b}_{b} \end{pmatrix}$$

$$= \begin{pmatrix} \underbrace{s_1 - e_1, s_2 - e_2, \dots, s_b - e_b}_{b}, \underbrace{s_{b+1} - e_1, s_{b+2} - e_2, \dots, s_{2b} - e_b}_{b}, \dots, \underbrace{s_{(d_b-2)b+1} - e_1, s_{(d_b-2)b+2} - e_2, \dots, s_{(d_b-1)b} - e_b}_{b} \end{pmatrix} .$$

$$\underbrace{s_{(d_b-2)b+1} - e_1, s_{(d_b-2)b+2} - e_2, \dots, s_{(d_b-1)b} - e_b}_{b} \end{pmatrix} .$$

If the periodical burst error starts from j^{th} position $(j \ge b + 1)$, then

$$\begin{split} S &= w H_b^{\perp} \\ &= e H_b^{\perp} \\ &= \left(e_{b+1}, e_{b+2}, \dots, e_{2b}, e_{2b+1}, e_{2b+2}, \dots, e_{3b}, \dots, \dots, e_{(d_b-1)b+1}, e_{(d_b-1)b+2}, \dots, e_{d_bb} \right) \\ &= \left(s_1, s_2, \dots, s_b, s_{b+1}, s_{b+2}, \dots, s_{2b}, \dots, \dots, s_{(d_b-2)b+1}, s_{(d_b-2)b+2}, \dots, s_{(d_b-1)b} \right). \end{split}$$

Thus, if the periodical burst starts from j^{th} position (j > b), at least one b-tuple of $\{(s_1, s_2, \ldots, s_b), (s_{b+1}, s_{b+2}, \ldots, s_{2b}), (s_{2b+1}, s_{2b+2}, \ldots, s_{3b})\}$ of b consecutive components of S will be a zero tuple. Then, the error vector e will be $e = (\underbrace{00 \ldots 0}_{b} s_1 s_2 \ldots s_{b(d_b-1)})$.

Example 6.19. For b = 4 and $d_4 = 6$, the parity check matrix of C_4 is given by

 $H_4 =$

 20×24

Therefore, the error vector is e = (0110100000000001010000) and the sent

codeword is

$$v = w - e = (101101011101110111001101) - (011010000000000001010000)$$

= (11011101110111011101).

(ii) Let the received vector be w = (1101010011011101110110001) after an error vector of $\psi_{(3b=12,3),24,2}$. Now the syndrome of w is $wH_4^T = (10010000000000001100)$. As the second and third tuples of b components are all zero, the error starts after b=4 positions.

Therefore the error vector is e = (0000100100000000001100) and the sent codeword is

$$v = w - e = (1101010011011101110110001) - (000010010000000000001100)$$

= 110111011101110111011.