

# **BOUNDS ON BLOCK CODES**

黃志文 簡永懿

- 1. Introduction**
- 2. Hamming bound**
- 3. Plotkin bound**
- 4. Griesmer bound**
- 5. Singleton bound**
- 6. Rieger bound for burst error**
- 7. Conclusion**
- 8. References**

# 1.Introduction:

- We exam the question of the relationship between the values of  $n$  and  $k$  and the amount of error correction that is possible.
  
- For a given  $n$  and  $d_{min}$ , it is customary to let  $A_r(n, d_{min})$  denote the largest possible size  $M=r^k$  for which there exists an  $r$ -ary  $(n, k, d_{min})$ -code. Thus,  
$$A_r(n, d_{min})=\max\{M/\text{there exists an } r\text{-ary } A_r(n, k, d_{min})\text{-code}\}$$
  
- Any  $(n, k, d_{min})$ -code  $C$  that has maxmum size  $M$  ( $M=r^k$ , and  $A_r(n, d_{min})=M$ ) is called an optimal code.
  
- Very little is currently known about the numbers  $A_r(n, d_{min})$ .

**Example 1.1:**  $A_2(4, 3) = 2$

**Proof:** Let  $C$  be a  $(4, k, 3)$ -code. Assume  $C$  contains the all-zero codeword  $\mathbf{0}$ . Since  $d(C) = 3$ , any other codeword  $\mathbf{c}$  in  $C$  must satisfy  $d(\mathbf{c}, \mathbf{0}) \geq 3$ , and so it leaves five codewords, namely

1110 1101 1011 0111 1111

But no pair of these has distance 3 apart, and so only one can be included in  $C$ . Hence,  $C$  can have at most two codewords, implying that  $A_2(4, 3) = 2$ .

- The approach is not go very far in determining values of  $A_2(n, d_{min})$ , and much more sophisticated method are needed.
- It some cases we do not know precise values, but there are a number of bounds.

## 2. Hamming bound:

- $C = \{c_1, c_2, \dots, c_M\}$  be the optimal  $(n, k, d_{min})$ -code, and set  $t = \frac{d_{min}-1}{2}$ . Since the packing spheres are disjoint we have  $\sum_{i=1}^M |S_r^n(c_i, t)| \leq |Z_r^n|$ , and since  $|S_r^n(c_i, t)| = V_r^n(t)$ , this is equivalent to

$$V_r^n(t) \cdot M \leq r^n,$$

since  $C$  is optimal,  $M = A_r(n, d_{min})$  and so

$$A_r(n, d_{min}) \leq \frac{r^n}{V_r^n(t)} = \frac{r^n}{V_r^n\left(\left\lfloor \frac{d_{min}-1}{2} \right\rfloor\right)}$$

This inequality is known as the sphere-packing bound for  $A_r(n, d_{min})$ .

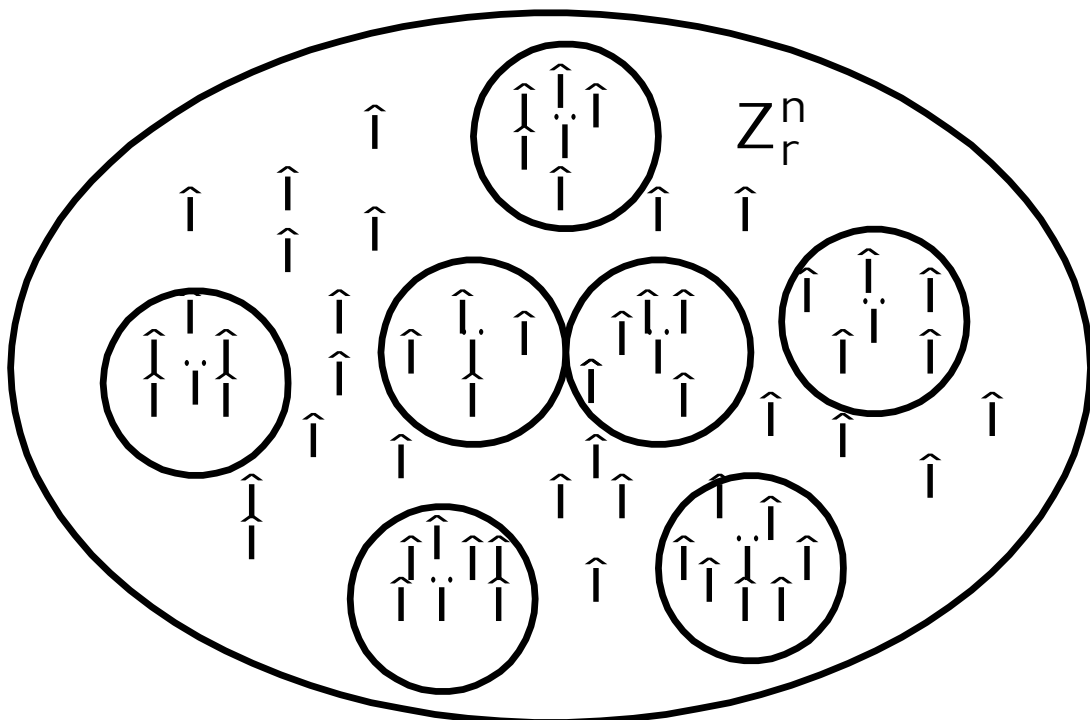
- For a binary  $t$ -error-correcting block code, the number of (received) words in a sphere of radius  $t$  is  $\sum_{i=0}^t \binom{n}{i}$ . Since there are  $2^k$  possible transmitted

codewords, there are  $2^k$  non-overlapping spheres, each have radius  $t$ .

■ The total number of (received) words enclose in the  $2^k$  spheres cannot exceed the  $2^n$  possible words.

■ Thus, a  $t$ -error-correcting block code must satisfy the following inequality:

$$2^n \geq 2^k \sum_{l=0}^t \binom{n}{l}.$$



- This called the Hamming bound or the sphere-packing bound for binary codes.
- A perfect code satisfies the Hamming bound with equality.

### 3. Plotkin bound:

- The Plotkin bound tends to set a tighter bound for low rate codes.
- For fixed values of  $n$  and  $d_{min}$ , we can derive another upper bound on the values of  $A_r(n, d_{min})$ .
- Let  $C = \{c_1, c_2, \dots, c_M\}$  be a binary  $(n, k, d_{min})$ -code, and consider the sum

$$S = \sum_{i < j} d(c_i, c_j)$$

this is the sum of the distances between

all pairs of the codewords in  $C$ . Since  $d(\mathbf{c}_i, \mathbf{c}_j) \leq d(C) = d_{\min}$ , for all codewords  $\mathbf{c}_i$  and  $\mathbf{c}_j$ , and since there are  $\binom{M}{2}$  pairs of codewords in  $C$ , we have

$$S = \sum_{i < j} d(\mathbf{c}_i, \mathbf{c}_j) \leq d_{\min} \binom{M}{2} \quad (4.9.1)$$

- Suppose the codewords in  $C$  have the form

$$\mathbf{c}_1 = \mathbf{c}_{11}\mathbf{c}_{12} \dots \mathbf{c}_{1n}$$

$$\mathbf{c}_2 = \mathbf{c}_{21}\mathbf{c}_{22} \dots \mathbf{c}_{2n}$$

.....

$$\mathbf{c}_M = \mathbf{c}_{M1}\mathbf{c}_{M2} \dots \mathbf{c}_{Mn}$$

consider the first positions  $\mathbf{c}_{11}, \mathbf{c}_{21}, \dots, \mathbf{c}_{M1}$  from each codeword. If  $p$  of these bits are 1, and  $M-p$  are 0, then these bits will contribute  $p(M-p)$  to  $S$ .

- If  $M$  is even, then  $p(M-p) \leq M^2/4$  and if

$M$  is odd, then  $p(M-p) \leq (M^2-1)/4$ .

- Hence the total contribution of all positions to the sum  $S$  is at most  $n(M^2/4)$  for  $M$  even and  $n((M^2-1)/4)$  for  $M$  odd. Thus,

$$S \hat{=} \begin{cases} \frac{nM^2}{4} & \text{for } M \text{ even} \\ \frac{n(M^2-1)}{4} & \text{for } M \text{ odd} \end{cases}$$

Putting this together with (4.9.1), we see that

$$d_{\min} \hat{=} \begin{cases} \frac{nM^2}{4} & \text{for } M \text{ even} \\ \frac{n(M^2-1)}{4} & \text{for } M \text{ odd} \end{cases}$$

After some algebraic simplification, we

got, for  $n < 2d_{\min}$ ,

$$M \hat{=} \begin{cases} \frac{2d_{\min}}{2d_{\min}-n} & \text{for } M \text{ even} \\ \frac{2d_{\min}}{2d_{\min}-1} \hat{-} n & \text{for } M \text{ odd} \end{cases}$$

- Finally this can be improved and extended by separating the cases where

$d_{min}$  is even and  $d_{min}$  is odd :

1. If  $d_{min}$  is even, for  $n < 2d_{min}$ ,  $k$

$$A_2(n, d_{min}) \hat{=} 2 \frac{d_{min}}{2d_{min} - n}$$

for  $n = 2d_{min}$ ,

$$A_2(2d_{min}, d_{min}) = 4d_{min}$$

2. If  $d_{min}$  is odd, for  $n < 2d_{min} + 1$ ,

$$A_2(n, d_{min}) \hat{=} 2 \left\lfloor \frac{d_{min} + 1}{2d_{min} + 1 - n} \right\rfloor$$

for  $n = 2d_{min} + 1$ ,

$$A_2(2d_{min} + 1, d_{min}) \hat{=} 4d_{min} + 4.$$

■ It is known as the Plotkin bound for fixed values of  $n$  and  $d_{min}$ .

■ For fixed values of  $n$  and  $k$ , for a  $q$ -ary code, the chance over the whole set of codewords of any position being nonzero is  $(q-1)/q$  and there are  $q^k$  codewords in the whole set. The number

of nonzero codewords is  $q^k - 1$  and so the average weight of a codeword is  $\frac{n \frac{q^k - 1}{q} q^k}{q^k - 1}$ .

- The  $d_{min}$  cannot be greater than this so

$$d_{min} \hat{=} \frac{n \frac{q^k - 1}{q} q^k}{q^k - 1}$$

For binary code this becomes

$$d_{min} \hat{=} \frac{n 2^{k-1}}{2^k - 1}$$

#### 4. Griesmer bound:

- The Griesmer bound is often tighter than the Plotkin bound, and its derivation leads to methods of constructing good codes.
- Let  $N(k, d_{min})$  represent the lowest possible value of length  $n$  for a linear code  $C$  of dimension  $k$ .

- Without lose of generality, the generate matrix can be taken to have a first row consisting of  $d_{min}$  ones followed by  $N(k,d_{min}) - d_{min}$  zeros.

$$\mathbf{G} = \begin{bmatrix} 111\dots1 & 000\dots0 \\ \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix}$$

- The  $\mathbf{G}_2$  generates a  $(N(k,d_{min})-d_{min}, k-1)$  code of minimum distance  $d_1$ , called the residue code.

- If  $\mathbf{u}$  is a residue code which when concatenated with a sequence  $\mathbf{v}$  of length  $d_{min}$  produce a codeword of C, then we can say  $d_1 + \text{weight}(\mathbf{v}) \geq d_{min}$ , However,  $\mathbf{u}$  concatenated with the complement of  $\mathbf{v}$  is also a codeword

$$d_1 + d_{min} - \text{weight}(\mathbf{v}) \geq d_{min}.$$

Therefore  $2d_1 \geq d_{min}$ , or  $d_1 \geq \left\lceil \frac{d_{min}}{2} \right\rceil$ .

- Since the code generated by  $G_2$  is of length  $N(k, d_{min}) - d_{min}$ , we say

$$N(k, d_{min}) = N(k-1, \lceil d_{min}/2 \rceil) + d_{min}.$$

- Apply the above result iteratively gives

$$N(k, d_{min}) = \sum_{i=0}^{k-1} \lceil \frac{d_{min}}{2^i} \rceil$$

This is the lowest possible value of length, so the general statement of the Griesmer bound for binary code is

$$n \geq \sum_{i=0}^{k-1} \lceil \frac{d_{min}}{2^i} \rceil$$

- For q-ary codes, the argument generalizes to give

$$n \geq \sum_{i=0}^{k-1} \lceil \frac{d_{min}}{q^i} \rceil$$

## 5. Singleton Bound

- For a linear  $(n, k, d_{\min})$  block code, exists an upper bound to minimum distance of  $d_{\min} \leq n - k + 1$ .

proof:

1. If we change one of the information symbols in a block code, the best we can hope for in terms of distance between codewords is that all the parity symbols will also change.
2. In this case the distance between the codewords will be  $n - k + 1$ .

Example:

1. The only binary codes that achieve this bound with equality are simple  $(n, 1)$  repetition code.
2. Reed Solomon Code:

$$d_{\min} = n - k + 1 = 2t + 1$$
$$n = q - 1 = 2^m - 1$$

## 6. Reiger Bound

■ If a linear  $(n, k)$  code  $C$  can correct all burst errors of length  $b$  or less, then

$$k \leq n - 2b$$

proof:

1. If  $2 \leq l \leq 2b$ , then any burst  $X$  of length  $l$  can be written as the difference  $x_1 \oplus x_2$  of two distinct bursts of length at most  $b$ .

2. For instance,

If  $n=11$  and  $b=3$ , then the burst  $X=000111001100$  of length 6 can be written as follow:

$$00011101100 =$$

$$00011100000 + 00000001100$$

3. Since  $C$  can correct  $x_1$  and  $x_2$ , these strings must be coset leader and therefore cannot be in the same coset of any standard array for  $C$ .

4. This implies that their difference, which is  $X$ , is not a codeword.

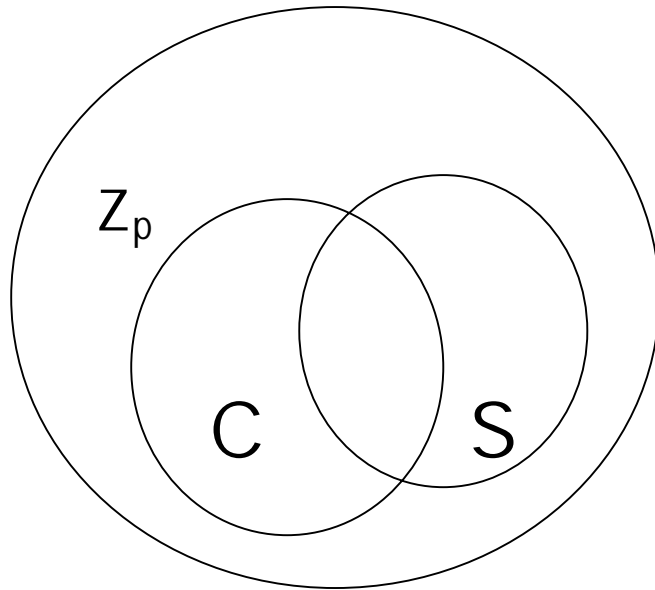
5. (with  $b$  replaced  $2b$ ), to get  $k \leq n - 2b$ .

Thm:

- Let  $C$  be a linear  $(n, k)$  code over  $Z_p$ .  
If  $C$  contains no bursts of length  $b$  or less, then  $k \leq n - b$ .

Proof:

1. We can find the set  $S$  of all strings in  $Z_p^n$  with 0s in the last  $n - b$  position.
2. Since the difference between any two strings in the same row of a standard array for  $C$  is a codeword.
3. If any two distinct strings in  $S$  lie in the same coset of a standard array, then their difference would be a nonzero burst of length at most  $b$  lying in  $C$ , which is not possible since  $C$  is assumed not to contain any such bursts.
4. Hence, the number of cosets of  $C$ , which is  $p^{n-k}$ , must be greater than or equals to the size of  $S$ , which is  $p^b$ .



$S_1 = (x \dots x 00 \dots 0)$  : 0s in the last  $n \rightarrow b$  position

$S_2 = (x \dots x 00 \dots 0)$

Table 1.

n	Actual value of $A(n,7)$	Plotkin Bound on $A(n,7)$	Sphere-packing Bound on $A(n,7)$
7	2	2	2
8	2	2	2
9	2	2	3
10	2	2	5
11	4	4	8
12	4	4	13
13	8	8	21
14	16	16	34
15	32	32	56

## **7. Conclusion:**

1. The relationship between the values of  $n$  and  $k$  and the amount of error correction.
2. It's found that there is not fixed relationship, but there are a number of upper bounds applying to minimum distance or error correction.
3. The Plotkin bound set a tighter bound for low rate codes, the Hamming bound being tighter for higher rates.

## 8.Reference:

1. Irving S.Reed & Xuemin Chen, “*Error-Control Coding for Data Network*”, Kluwer Academic Pulishers, 1999.
2. Peter Sweeney, “*Error Control Coding : An Introduction*”, Prentice Hall, 1991.
3. S. Aler, F.W. Gehring &P.R. Halmos, “*Introduction to Coding And Information Thoery*”, Springer 1996.
4. 林銀議, “錯誤保護及訂正碼”, 國立中央大學電機系, 2000.
5. S. H. Reiger, “*Correction of “Clustered” Error*”, IEEE Trans,on InformThoery,IT-6,pp16-21,1960.
6. R.C. Singleton, “*Maximum Distance Q-nary Codes* ”, IEEE Trans,on Inform Thoery, IT-10, pp116-118,1964.