

## A SIMPLE PROOF OF THE IMPROVED JOHNSON BOUND FOR BINARY CODES

LE THI NGOC GIAU AND PHAN THANH TOAN

ABSTRACT. In this paper, we give a simple proof of the improved Johnson bound for  $A(n, d)$ , the maximum number of codewords in a binary code of length  $n$  and minimum distance  $d$ , given by Mounits, Etzion and Litsyn.

### 1. The improved Johnson bound for $A(n, d)$

Let  $\mathbb{F} = \{0, 1\}$  and let  $n$  be a positive integer. The (*Hamming*) distance between two vectors  $u$  and  $v$  in  $\mathbb{F}^n$ , denoted by  $d(u, v)$ , is the number of coordinates where they differ. The (*Hamming*) weight of a vector  $u$  in  $\mathbb{F}^n$ , denoted by  $wt(u)$ , is the distance between it and the zero vector. The *minimum distance* of a subset of  $\mathbb{F}^n$  is the smallest distance between any two different vectors in that subset. An  $(n, d)$  code is a subset of  $\mathbb{F}^n$  having minimum distance  $\geq d$ . If  $\mathcal{C}$  is an  $(n, d)$  code, then an element of  $\mathcal{C}$  is called a *codeword* and the number of codewords in  $\mathcal{C}$  is called the *size* of  $\mathcal{C}$ , denoted by  $|\mathcal{C}|$ . The largest possible size of an  $(n, d)$  code is denoted by  $A(n, d)$ . An  $(n, d, w)$  *constant-weight code* is an  $(n, d)$  code such that every codeword has weight  $w$ . Denote by  $A(n, d, w)$  the largest possible size of an  $(n, d, w)$  constant-weight code.

The problem of determining the values of  $A(n, d)$  is one of the most fundamental problems in coding theory [15]. The exact value of  $A(n, d)$  is extremely difficult to find even for relatively small values of  $n$ . Hence, lower bounds and upper bounds for this function are usually considered. While lower bounds are obtained from explicit code constructions [3, 6, 10, 11, 13, 14, 16, 19, 21], upper bounds involve analytic methods [1, 8, 12, 17, 18, 20, 22]. The following equality is well-known (see for example [15]).

**Lemma 1.**  $A(n, d) = A(n + 1, d + 1)$  if  $d$  is odd.

---

Received March 29, 2018; Revised June 12, 2018; Accepted June 21, 2018.  
2010 *Mathematics Subject Classification.* 94B65.

*Key words and phrases.* binary code, Johnson bound, upper bound.

This research is funded by Foundation for Science and Technology Development of Ton Duc Thang University (FOSTECT), website: <http://fostect.tdtu.edu.vn>, under Grant FOSTECT.2017.BR.02.

One of the basis upper bounds on  $A(n, d)$ ,  $d = 2\delta + 1$ , is the sphere packing bound or the Hamming bound.

**Theorem 2** (Sphere packing bound).

$$(1) \quad A(n, 2\delta + 1) \leq \frac{2^n}{\sum_{i=0}^{\delta} \binom{n}{i}}.$$

The sphere packing bound follows from the fact that the spheres of radius  $\delta$  centered at the codewords of an  $(n, 2\delta + 1)$  code are disjoint, and each of which contains  $\sum_{i=0}^{\delta} \binom{n}{i}$  vectors in  $\mathbb{F}^n$ . Codes that attain the sphere packing bound are called *perfect codes*. So for a perfect code, the corresponding spheres cover the whole space  $\mathbb{F}^n$ . In [9], Johnson considered spheres of radius  $\delta + 1$  (the spheres may not be disjoint) and improved this sphere packing bound by showing:

**Theorem 3** (Johnson bound).

$$(2) \quad A(n, 2\delta + 1) \leq \frac{2^n}{\sum_{i=0}^{\delta} \binom{n}{i} + \frac{\binom{n}{\delta+1} - \binom{2\delta+1}{\delta} A(n, 2\delta+2, 2\delta+1)}{A(n, 2\delta+2, \delta+1)}}.$$

Since  $A(n, 2k, k) = \lfloor \frac{n}{k} \rfloor$ , we have:

**Corollary 4.**

$$(3) \quad A(n, 2\delta + 1) \leq \frac{2^n}{\sum_{i=0}^{\delta} \binom{n}{i} + \frac{\binom{n}{\delta+1} - \binom{2\delta+1}{\delta} A(n, 2\delta+2, 2\delta+1)}{\lfloor \frac{n}{\delta+1} \rfloor}}.$$

Codes that attain the Johnson bound are called *nearly perfect codes*. For more information on perfect codes and nearly perfect codes, see [4] and [15].

Besides other good bounds such as linear programming bound [5, 15] and semidefinite programming bound [7, 22], the Johnson bound is still one of the best known upper bounds on  $A(n, d)$ . In [17], Mounits, Etzion and Litsyn further improved the Johnson bound by showing the following.

**Theorem 5** (Improved Johnson bound).

$$(4) \quad A(n, 2\delta + 1) \leq \frac{2^n}{\sum_{i=0}^{\delta} \binom{n}{i} + \frac{\binom{n+1}{\delta+2} - \binom{2\delta+2}{\delta+2} A(n+1, 2\delta+2, 2\delta+2)}{A(n+1, 2\delta+2, \delta+2)}}.$$

They proved that this bound is always at least as good as the Johnson bound and that for each  $\delta \geq 1$ , there exist infinitely many values of  $n$  for which the bound is better than the Johnson bound.

## 2. A simple proof of the improved Johnson bound

In this section, we give a simple proof of the improved Johnson bound. In the proof of the improved Johnson bound, Mounits, Etzion and Litsyn considered  $A(n, 2\delta + 1)$  and since  $2\delta + 1$  is odd, an  $(n, 2\delta + 1)$  code can have both odd weight codewords and even weight codewords, which makes it more difficult to handle the codewords. In our proof the key difference is that we consider  $A(n + 1, 2\delta + 2)$  instead of  $A(n, 2\delta + 1)$ . For this, we restate the improved Johnson bound as below. The advantage of using  $A(n + 1, 2\delta + 2)$  is that every codeword of an  $(n + 1, 2\delta + 2)$  code can be assumed to have even weight since  $2\delta + 2$  is even. This makes the code simpler and hence so is the proof of the theorem (as showed below). Our proof is a modification of the proof of the Johnson bound in [15].

**Theorem 6** (Improved Johnson bound).

$$(5) \quad A(n + 1, 2\delta + 2) \leq \frac{2^n}{\sum_{i=0}^{\delta} \binom{n}{i} + \frac{\binom{n+1}{\delta+2} - \binom{2\delta+2}{\delta+2} A(n+1, 2\delta+2, 2\delta+2)}{A(n+1, 2\delta+2, \delta+2)}}.$$

*Proof.* (i) Let  $\mathcal{C}$  be an  $(n + 1, M, 2\delta + 2)$  code, i.e., an  $(n + 1, 2\delta + 2)$  code of size  $M$ , with  $M = A(n + 1, 2\delta + 2)$  such that  $\mathcal{C}$  contains the zero vector and each codeword in  $\mathcal{C}$  has even weight. First we consider the case when  $\delta$  is even. Let  $\mathbb{F}_{even}^{n+1}$  be the set of vectors in  $\mathbb{F}^{n+1}$  of even weight and let  $S_i$  be the set of vectors in  $\mathbb{F}_{even}^{n+1}$  at distance  $i$  from  $\mathcal{C}$ , i.e.,

$$(6) \quad S_i = \{u \in \mathbb{F}_{even}^{n+1} \mid d(u, v) \geq i \text{ for all } v \in \mathcal{C} \text{ and } d(u, v) = i \text{ for some } v \in \mathcal{C}\}.$$

Thus  $S_0 = \mathcal{C}$  and  $S_i$  is empty if  $i$  is odd (since the distance between two vectors of even weight is always even). We have

$$(7) \quad S_0 \cup S_2 \cup \dots \cup S_{2\delta} = \mathbb{F}_{even}^{n+1},$$

for if there were a vector of even weight at distance  $\geq 2\delta + 2$  from  $\mathcal{C}$ , then we could add it to  $\mathcal{C}$  and get a larger code.

(ii) Pick an arbitrary codeword  $P$  and move it to the origin. The codewords of weight  $2\delta + 2$  form a constant-weight code with distance  $\geq 2\delta + 2$ , i.e., the number of codewords of weight  $2\delta + 2$  is  $\leq A(n + 1, 2\delta + 2, 2\delta + 2)$ .

(iii) Let  $W_{\delta+2}$  be the set of vectors in  $\mathbb{F}^{n+1}$  of (even) weight  $\delta + 2$ . Any vector in  $W_{\delta+2}$  belongs to either  $S_\delta$  or  $S_{\delta+2}$ . Corresponding to each codeword  $Q$  of weight  $2\delta + 2$ , there are  $\binom{2\delta+2}{\delta+2}$  vectors of weight  $\delta + 2$  at distance  $\delta$  from  $Q$ . These vectors are in  $W_{\delta+2} \cap S_\delta$  and are all distinct. Therefore,

$$(8) \quad \begin{aligned} |W_{\delta+2} \cap S_{\delta+2}| &= |W_{\delta+2}| - |W_{\delta+2} \cap S_\delta| \\ &\geq \binom{n+1}{\delta+2} - \binom{2\delta+2}{\delta+2} A(n+1, 2\delta+2, 2\delta+2). \end{aligned}$$

(iv) A vector  $R$  in  $W_{\delta+2} \cap S_{\delta+2}$  is at distance  $\delta + 2$  from at most  $A(n + 1, 2\delta + 2, \delta + 2)$  codewords. For move the origin to  $R$  and consider how many

codewords can be at distance  $\delta + 2$  from  $R$  and have mutual distance  $2\delta + 2$ . The number of such codewords is  $\leq A(n + 1, 2\delta + 2, \delta + 2)$ .

(v) Now let  $P$  vary over all the codewords. For each  $i = 0, 2, \dots, \delta$ , we get

$$(9) \quad \begin{aligned} |S_i| &= A(n + 1, 2\delta + 2) \binom{n + 1}{i} \\ &= A(n + 1, 2\delta + 2) \left[ \binom{n}{i - 1} + \binom{n}{i} \right]. \end{aligned}$$

Also,

$$(10) \quad |S_{\delta+2}| \geq A(n + 1, 2\delta + 2) \frac{\binom{n+1}{\delta+2} - \binom{2\delta+2}{\delta+2} A(n + 1, 2\delta + 2, 2\delta + 2)}{A(n + 1, 2\delta + 2, \delta + 2)}.$$

The result then follows since

$$(11) \quad |S_0| + |S_2| + \dots + |S_\delta| + |S_{\delta+2}| \leq |\mathbb{F}_{even}^{n+1}| = 2^n.$$

The case when  $\delta$  is odd is proved similarly, where  $\mathbb{F}_{even}^{n+1}$  is replaced by  $\mathbb{F}_{odd}^{n+1}$ , the set of all vectors in  $\mathbb{F}^{n+1}$  of odd weight.  $\square$

### 3. Some examples

In this section, we give examples illustrating that the improved Johnson bound is one of the best upper bounds on  $A(n, d)$  and that it is really better than the Johnson bound. We first show known upper bounds on  $A(n, 4, 3)$  and  $A(n, 4, 4)$  (see [15] or [18]), which will be used in the examples.

**Theorem 7.**

$$(12) \quad A(n, 4, 3) = \begin{cases} \left\lfloor \left\lfloor \frac{n}{3} \left\lfloor \frac{n-1}{2} \right\rfloor \right\rfloor \right. & \text{if } n \not\equiv 5 \pmod{6}, \\ \left\lfloor \left\lfloor \frac{n}{3} \left\lfloor \frac{n-1}{2} \right\rfloor \right\rfloor - 1 \right. & \text{if } n \equiv 5 \pmod{6}. \end{cases}$$

**Theorem 8.**

$$(13) \quad A(n, 4, 4) = \begin{cases} \frac{n(n-1)(n-3)}{24} & \text{if } n \equiv 1 \text{ or } 3 \pmod{6}, \\ \frac{n(n-1)(n-2)}{24} & \text{if } n \equiv 2 \text{ or } 4 \pmod{6}, \\ \frac{n(n^2-3n-6)}{24} & \text{if } n \equiv 0 \pmod{6}, \end{cases}$$

$$(14) \quad A(n, 4, 4) \leq \begin{cases} \frac{n^3 - 4n^2 + n - 6}{24} & \text{if } n \equiv 5 \pmod{12}, \\ \frac{n^3 - 4n^2 + n - 18}{24} & \text{if } n \equiv 11 \pmod{12}. \end{cases}$$

**Example 9.** Consider  $n = 22$  and  $\delta = 1$ . By Theorem 7, we have

$$(15) \quad A(22, 4, 3) = 73.$$

The Johnson bound gives

$$\begin{aligned} A(22, 3) &\leq \frac{2^{22}}{\sum_{i=0}^1 \binom{22}{i} + \frac{\binom{22}{2} - \binom{3}{1} A(22, 4, 3)}{\lfloor \frac{22}{2} \rfloor}} \\ &= \frac{2^{22}}{1 + 22 + \frac{231 - 3 \cdot 73}{11}} \\ &= \frac{46137344}{265} \\ (16) \quad &< 174104. \end{aligned}$$

Hence,

$$(17) \quad A(23, 4) = A(22, 3) \leq 174103.$$

On the other hand, by Theorems 7 and 8, we have

$$(18) \quad A(23, 4, 3) = 83$$

and

$$(19) \quad A(23, 4, 4) \leq 419.$$

In fact,  $A(23, 4, 4) = 419$  (see [2]) but this equality is not necessary in evaluating the upper bound. The improved Johnson bound gives

$$\begin{aligned} A(23, 4) &\leq \frac{2^{22}}{\sum_{i=0}^1 \binom{22}{i} + \frac{\binom{23}{3} - \binom{4}{3} A(23, 4, 4)}{A(23, 4, 3)}} \\ &= \frac{2^{22}}{1 + 22 + \frac{1771 - 4 \cdot 419}{83}} \\ &= \frac{87031808}{501} \\ (20) \quad &< 173717. \end{aligned}$$

Therefore,

$$(21) \quad A(23, 4) \leq 173716 < 174103.$$

The best known upper bound of  $A(23, 4)$  is

$$(22) \quad A(23, 4) \leq 172361,$$

which is from [18].

**Example 10.** Consider  $n = 23$  and  $\delta = 1$ . By Theorems 7 and 8, we have

$$(23) \quad A(n + 1, 2\delta + 2, \delta + 2) = A(24, 4, 3) = 88$$

and

$$(24) \quad A(n+1, 2\delta+2, 2\delta+2) = A(24, 4, 4) = 498.$$

The improved Johnson bound gives

$$(25) \quad \begin{aligned} A(24, 4) &\leq \frac{2^{23}}{\sum_{i=0}^1 \binom{23}{i} + \frac{\binom{24}{3} - \binom{4}{3} A(24, 4, 4)}{A(24, 4, 3)}} \\ &= \frac{2^{23}}{1 + 23 + \frac{2024 - 4 \cdot 498}{88}} \\ &= \frac{23068672}{67} \\ &< 344309. \end{aligned}$$

Therefore,

$$(26) \quad A(24, 4) \leq 344308.$$

The upper bound  $A(24, 4) \leq 344308$  is the best known upper bound for  $A(24, 4)$  up to now.

**Acknowledgement.** The authors thank the reviewer for his/her comments and suggestions to improve the presentation of the paper.

### References

- [1] E. Agrell, A. Vardy, and K. Zeger, *A table of upper bounds for binary codes*, IEEE Trans. Inform. Theory **47** (2001), no. 7, 3004–3006.
- [2] J. Bao and L. Ji, *The completion determination of optimal (3, 4)-packings*, Des. Codes Cryptogr. **77** (2015), no. 1, 217–229.
- [3] M. R. Best, *Binary codes with a minimum distance of four*, IEEE Trans. Inform. Theory **26** (1980), no. 6, 738–742.
- [4] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*, North-Holland Mathematical Library, **54**, North-Holland Publishing Co., Amsterdam, 1997.
- [5] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. No. **10** (1973), vi+97 pp.
- [6] K. Elssel and K.-H. Zimmermann, *Two new nonlinear binary codes*, IEEE Trans. Inform. Theory **51** (2005), no. 3, 1189–1190.
- [7] D. C. Gijswijt, H. D. Mittelmann, and A. Schrijver, *Semidefinite code bounds based on quadruple distances*, IEEE Trans. Inform. Theory **58** (2012), no. 5, 2697–2705.
- [8] W. Haas, *On the failing cases of the Johnson bound for error-correcting codes*, Electron. J. Combin. **15** (2008), no. 1, Research paper 55, 13 pp.
- [9] S. M. Johnson, *A new upper bound for error-correcting codes*, IRE Trans. **IT-8** (1962), 203–207.
- [10] M. K. Kaikkonen, *A new four-error-correcting code of length 20*, IEEE Trans. Inform. Theory **35** (1989), no. 6, 1344.
- [11] ———, *Codes from affine permutation groups*, Des. Codes Cryptogr. **15** (1998), no. 2, 183–186.
- [12] H. K. Kim and P. T. Toan, *Improved semidefinite programming bound on sizes of codes*, IEEE Trans. Inform. Theory **59** (2013), no. 11, 7337–7345.

- [13] Y. Klein, S. Litsyn, and A. Vardy, *Two new bounds on the size of binary codes with a minimum distance of three*, Des. Codes Cryptogr. **6** (1995), no. 3, 219–227.
- [14] A. Laaksonen and P.R.J. Östergård, *Constructing error-correcting binary codes using transitive permutation groups*, Discrete Appl. Math. **233** (2017), 65–70.
- [15] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes. I*, North-Holland Publishing Co., Amsterdam, 1977.
- [16] M. Milshtein, *A new binary code of length 16 and minimum distance 3*, Inform. Process. Lett. **115** (2015), no. 12, 975–976.
- [17] B. Mounits, T. Etzion, and S. Litsyn, *Improved upper bounds on sizes of codes*, IEEE Trans. Inform. Theory **48** (2002), no. 4, 880–886.
- [18] ———, *New upper bounds on codes via association schemes and linear programming*, Adv. Math. Commun. **1** (2007), no. 2, 173–195.
- [19] P. R. J. Östergård, *Two new four-error-correcting binary codes*, Des. Codes Cryptogr. **36** (2005), no. 3, 327–329.
- [20] ———, *On optimal binary codes with unbalanced coordinates*, Appl. Algebra Engrg. Comm. Comput. **24** (2013), no. 3-4, 197–200.
- [21] P. R. J. Östergård, T. Baicheva, and E. Kolev, *Optimal binary one-error-correcting codes of length 10 have 72 codewords*, IEEE Trans. Inform. Theory **45** (1999), no. 4, 1229–1231.
- [22] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), no. 8, 2859–2866.

LE THI NGOC GIAU  
FACULTY OF MATHEMATICS AND STATISTICS  
TON DUC THANG UNIVERSITY  
HO CHI MINH CITY, VIETNAM  
Email address: lethingocgiau@tdt.edu.vn

PHAN THANH TOAN  
FRACTIONAL CALCULUS, OPTIMIZATION AND ALGEBRA RESEARCH GROUP  
FACULTY OF MATHEMATICS AND STATISTICS  
TON DUC THANG UNIVERSITY  
HO CHI MINH CITY, VIETNAM  
Email address: phanthanhtoan@tdt.edu.vn