Des Codes Crypt (2007) 42:67–72
DOI 10.1007/s10623-006-9015-8

Improved lower bounds on sizes of single-error correcting codes

Simon Litsyn · Beniamin Mounits

Received: 16 January 2006 / Revised: 18 August 2006 /
Accepted: 25 August 2006 / Published online: 1 November 2006

Abstract A construction of codes of length $n = q + 1$ and minimum Hamming distance 3 over $GF(q)\setminus\{0\}$ is given. Substitution of the derived codes into a concatenation construction yields nonlinear binary single-error correcting codes with better than known parameters. In particular, new binary single-error correcting codes having more codewords than the best previously known in the range $n \le 512$ are obtained for the lengths 64–66, 128–133, 256–262, and 512.

Keywords $A(n,d)$ · MDS code · Weight distribution

AMS Classification 94B25

1 Introduction

Let $(n, M, d)_q$ denote a code of length n, minimum Hamming distance d and cardinality M over field $GF(q)$, whereas $[n, k, d]_q$ is a linear $(n, q^k, d)_q$ code . In a binary case we will omit the lower index and write (n, M, d). Let $A(n, d)$ denote the maximum number of codewords in a binary code of length n and minimum Hamming distance d. The quantity $A(n, d)$ is of basic interest in coding theory. Lower bounds on $A(n, d)$ are obtained by constructions. For a survey on the known lower bounds the reader is referred to [4].

Communicated by R. Hill.

S. Litsyn
Department of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel
e-mail: litsyn@eng.tau.ac.il

B. Mounits (✉)
CWI, Amsterdam, The Netherlands
e-mail: B.Mounits@cwi.nl

In this correspondence we consider lower bounds on $A(n,3)$. One of the most powerful tools in obtaining good lower bounds on $A(n,3)$ is the following method which consists of two steps:

- *Subalphabet subcode construction:* Suppose we have a nonbinary $(n,M,d)_Q$ code $\mathcal{C}$ over an alphabet $\mathcal{A} = \{\alpha_1,\alpha_2,\ldots,\alpha_Q\}$. Then, for $\mathcal{S} \subseteq \mathcal{A}$ we can construct $\left(n,M',d'\right)_{|\mathcal{S}|}$ subcode $\mathcal{C}'$ of $\mathcal{C}$ over subalphabet $\mathcal{S}$ of $\mathcal{A}$, i.e., $\mathcal{C}'$ consists of those codewords of $\mathcal{C}$ which have values from $\mathcal{S}$ in all the coordinates. It is clear that $M' \leq M$ and $d' \geq d$.
- *Concatenation construction:* Each coordinate value $\alpha_i \in \mathcal{S}$ is substituted by codewords of a binary code $\mathcal{C}_i$ with parameters (n_0,M_i,d), such that $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$ for $i \neq j$. Thus, the resulting binary code has length $n \cdot n_0$ and minimum distance d. The size of the code obtained depends on the values M_i, $1 \leq i \leq |\mathcal{S}|$.

To obtain good binary codes with minimum distance 3, one usually takes $|\mathcal{S}| = 2^m$ and codes $\mathcal{C}_i$ be cosets of the binary Hamming code of length $2^m - 1$. For a description of this method and some related constructions the reader is referred to [1–4,6].

In Sect. 2, we present a new construction which is a modification of the method given above. We construct a subalphabet subcode such that the alphabet sizes of the coordinates of the new code are not all equal $|\mathcal{S}|$. In Sect. 3, we apply the new construction and obtain improved lower bounds on $A(n,3)$.

The following notations will be used. The binary Hamming code of order s, $\mathcal{H}_1(s)$, is a $\left(2^s - 1, 2^{2^s-1-s}, 3\right)$ code. Given s, $\{\mathcal{H}_i(s) : 1 \leq i \leq 2^s\}$ denotes the collection of non-intersecting codes consisting of the binary Hamming code and its $2^s - 1$ cosets.

2 Construction

For $\mathcal{C}$ being an (n,M,d) code, let A_w be the number of codewords of weight w. The numbers $A_0, A_1, \ldots, A_n$ are called the *weight distribution* of C. Clearly $A_0 + A_1 + \cdots + A_n = M$.

Throughout $\mathcal{C}$ will denote the nonbinary MDS Hamming code having parameters $[n = q+1, k = q-1, 3]_q$ over $GF(q)$. Let $GF(q) = \{0, \alpha_1, \ldots, \alpha_i, \ldots, \alpha_{q-1}\}$. The weight distribution of MDS codes is known (see, e.g. [5, pp. 320–321]).

Theorem 1 *The number of codewords of weight w in an $[n,k,d = n-k+1]_q$ MDS code over $GF(q)$ is*

$$A_w = \binom{n}{w}(q-1)\sum_{j=0}^{w-d}(-1)^j\binom{w-1}{j}q^{w-d-j}. \tag{1}$$

We denote $\mathcal{C}_w = \{c \in \mathcal{C} : wt(c) = w\}$. Obviously $\mathcal{C} = \bigcup_{w=0}^{n} \mathcal{C}_w$ and $|\mathcal{C}_w| = A_w$.

Lemma 1 *For every coordinate and any $i \in \{1,\ldots,q-1\}$, the number of codewords of $\mathcal{C}_{q+1}$ having α_i in this coordinate is $\dfrac{A_{q+1}}{q-1}$.*

Proof Denote

$$B_\ell^j = \left\{u = (u_1, \ldots, u_k) \in GF(q)^k \ : \ uG = c \in \mathcal{C}_{q+1}, c_\ell = \alpha_j \in GF(q)\backslash\{0\}\right\},$$

where G is a generator matrix of $\mathcal{C}$ and c_ℓ is the ℓth coordinate of the codeword c. We wish to prove that for each ℓ, $1 \leq \ell \leq q+1$, and for any $i,j \in \{1, \ldots, q-1\}$, $|B_\ell^j| = |B_\ell^i|$. Note that $u \in B_\ell^j$ if and only if

$$\alpha_i \cdot \alpha_j^{-1} u = (\alpha_i \cdot \alpha_j^{-1} u_1, \ldots, \alpha_i \cdot \alpha_j^{-1} u_k) \in B_\ell^i \ ,$$

where α_j^{-1} denotes multiplicative inverse of α_j in $GF(q)$, which completes the proof. □

Lemma 2 *For every coordinate, the number of codewords of $\mathcal{C}_q$ having 0 in that coordinate is $A_q/(q+1)$.*

Proof Let

$$H = [h_1, h_2, \ldots, h_\ell, \ldots, h_n]$$

be a parity check matrix of $\mathcal{C}$ and

$$H'_\ell = [h_1, h_2, \ldots, h_{\ell-1}, h_{\ell+1} \ldots, h_n]$$

be a parity check matrix of the code $\mathcal{C}'$ which is obtained by deleting the ℓth column h_ℓ from H. Since $\mathcal{C}$ is MDS, it follows from [5, Corollary 3, p.319] that every $n-k = q+1-(q-1) = 2$ columns of H are linearly independent. Thus every two columns of H'_ℓ are also linearly independent, and by the same corollary [5, Corollary 3, p.319], the code $\mathcal{C}'$ having parameters $[n' = n-1 = q, k' = k-1 = q-2, d = n'-k'+1 = 3]_q$ is MDS. Obviously

$$\mathcal{C}'_q = \{(c_1, \ldots, c_{\ell-1}, c_{\ell+1}, \ldots, c_n) \ : (c_1, \ldots, c_{\ell-1}, 0, c_{\ell+1}, \ldots, c_n) \in \mathcal{C}_q\} \ .$$

Using (1) we have

$$A_q = |\mathcal{C}_q| = (q+1)(q-1)\sum_{j=0}^{q-d}(-1)^j \binom{q-1}{j} q^{q-d-j} \ ,$$

$$|\mathcal{C}'_q| = (q-1)\sum_{j=0}^{q-d}(-1)^j \binom{q-1}{j} q^{q-d-j}$$

and therefore

$$|\mathcal{C}'_q| = \frac{A_q}{q+1}.$$

□

Lemma 3 *For a $[q+1, q-1, 3]_q$ MDS code $\mathcal{C}$, q odd,*

$$A_{q+1} = \frac{q-1}{q^2}\left((q-1)^q - q^2 + 1\right) \ . \tag{2}$$

Proof Using (1), we obtain

$$A_{q+1} = (q-1)\sum_{j=0}^{q-2}(-1)^j\tbinom{q}{j}q^{q-2-j} = \tfrac{q-1}{q^2}\sum_{j=0}^{q-2}(-1)^j\tbinom{q}{j}q^{q-j}$$
$$= \tfrac{q-1}{q^2}\left(\sum_{j=0}^{q}(-1)^j\tbinom{q}{j}q^{q-j} - \sum_{j=q-1}^{q}(-1)^j\tbinom{q}{j}q^{q-j}\right) = \tfrac{q-1}{q^2}\left((q-1)^q - q^2 + 1\right).$$

□

Lemma 4 *For a* $[q+1, q-1, 3]_q$ *MDS code* $\mathcal{C}$*,* q *odd,*

$$A_q = \frac{q+1}{q^2}\left((q-1)^q + (q-1)(q^2-q-1)\right). \tag{3}$$

Proof Using (1), we obtain

$$A_q = (q+1)(q-1)\sum_{j=0}^{q-3}(-1)^j\tbinom{q-1}{j}q^{q-3-j} = \tfrac{q^2-1}{q^2}\sum_{j=0}^{q-3}(-1)^j\tbinom{q-1}{j}q^{q-1-j}$$
$$= \tfrac{q^2-1}{q^2}\left(\sum_{j=0}^{q-1}(-1)^j\tbinom{q-1}{j}q^{q-1-j} - \sum_{j=q-2}^{q-1}(-1)^j\tbinom{q-1}{j}q^{q-1-j}\right)$$
$$= \tfrac{q^2-1}{q^2}\left((q-1)^{q-1} + q^2 - q - 1\right) = \tfrac{q+1}{q^2}\left((q-1)^q + (q-1)(q^2-q-1)\right).$$

□

Let $m \in \{1, \ldots, q-2\}$. We take $m \cdot \frac{A_{q+1}}{q-1}$ codewords of $\mathcal{C}_{q+1}$ having $\alpha_1, \ldots, \alpha_m$ in the ℓth coordinate, and $\frac{A_q}{q+1}$ codewords of the code $\mathcal{C}_q$ having 0 at the ℓth coordinate which we substitute by α_{m+1}. Therefore, we obtain a $\left(q+1, m\frac{A_{q+1}}{q-1} + \frac{A_q}{q+1}, 3\right)_{q-1}$ code over $GF(q)\setminus\{0\}$. Let us denote this code by $\mathcal{D}(m,q)$. If q is odd it is easy to evaluate, using (2) and (3), that $\mathcal{D}(m,q)$ has parameters

$$\left(q+1, \frac{(m+1)(q-1)^q + (q-1)(q^2-q-1) - m(q^2-1)}{q^2}, 3\right)_{q-1}.$$

Now, let us consider the case $m = 2^s - 1$ and $q = 2^t + 1$. We plug in the values for m and q and obtain that $\mathcal{D}(2^s-1, 2^t+1)$ is an $(n, M, 3)_{2^t}$ code, where

$$n = 2^t + 2, \qquad M = \frac{2^{t2^t+t+s} + 2^{3t} + 2^{2t+1} + 2^t - 2^{2t+s} - 2^{t+s+1}}{2^{2t} + 2^{t+1} + 1}.$$

We know that in the code $\mathcal{D}(2^s-1, 2^t+1)$, in the ℓth coordinate, only 2^s symbols can appear from the 2^t symbols of $GF(2^t+1)\setminus\{0\}$. Therefore, we can encode the codewords of $\mathcal{D}(2^s-1, 2^t+1)$ in the following way to obtain a binary code.

In the ℓ-th coordinate: If $s = 1$, then $\alpha_1 \to 0$ and $\alpha_2 \to 1$. If $s \geq 2$, then

$\alpha_1 \to$ all the codewords of $\mathcal{H}_1(s)$,
$\alpha_2 \to$ all the codewords of $\mathcal{H}_2(s)$,
⋮
$\alpha_i \to$ all the codewords of $\mathcal{H}_i(s)$,

$$\vdots$$
$$\alpha_{2^s} \to \text{all the codewords of } \mathcal{H}_{2^s}(s).$$

In the rest of coordinates: We encode using the following rules

$$\alpha_1 \to \text{all the codewords of } \mathcal{H}_1(t),$$
$$\alpha_2 \to \text{all the codewords of } \mathcal{H}_2(t),$$
$$\vdots$$
$$\alpha_i \to \text{all the codewords of } \mathcal{H}_i(t),$$
$$\vdots$$
$$\alpha_{2^t} \to \text{all the codewords of } \mathcal{H}_{2^t}(t).$$

By this encoding the code $\mathcal{D}(2^s-1, 2^t+1)$ transforms into a binary code, that will be denoted by $\mathcal{B}(s,t)$, having parameters $(n, M, 3)$, where

$$n = 2^{2t} + 2^s - 2, \qquad M = \frac{2^{t2^t+t+s} + 2^{3t} + 2^{2t+1} + 2^t - 2^{2t+s} - 2^{t+s+1}}{2^{2t} + 2^{t+1} + 1} 2^{2^{2t}+2^s-t2^t-t-s-2}.$$

Therefore, we have proved the following theorem.

Theorem 2 *Let q be a prime power of the form $q = 2^t + 1$, and m be an integer of the form $m = 2^s - 1$, $s \le t$. There exists a binary $(n, M, 3)$ code $\mathcal{B}(s,t)$, where*

$$n = 2^{2t} + 2^s - 2, \qquad M = \frac{2^{t2^t+t+s} + 2^{3t} + 2^{2t+1} + 2^t - 2^{2t+s} - 2^{t+s+1}}{2^{2t} + 2^{t+1} + 1} 2^{2^{2t}+2^s-t2^t-t-s-2}.$$

3 Improved lower bounds on $A(n,3)$ for $n \le 512$

Here we apply the construction from the previous section to improve on the best known values of $A(n,3)$ for $n \le 512$. The following table presents these improvements. Codes that are obtained by shortening and having the same redundancy do not appear in the table.

Length	$\mathcal{B}(s,t)$	$\lvert\mathcal{B}(s,t)\rvert$	Previous bound
64	$\mathcal{B}(1,3)$	1657012×2^{37}	1657009×2^{37}
66	$\mathcal{B}(2,3)$	1657010×2^{39}	1657009×2^{39}
256	$\mathcal{B}(1,4)$	$1021273028302258920 \times 2^{188}$	$1021273028302258913 \times 2^{188}$
258	$\mathcal{B}(2,4)$	$1021273028302258916 \times 2^{190}$	$1021273028302258913 \times 2^{190}$
262	$\mathcal{B}(3,4)$	$1021273028302258914 \times 2^{194}$	$1021273028302258913 \times 2^{194}$

By using the $(u, u+v)$ construction [5, p.76] on the new codes of lengths 64–66 and 256, we obtain codes of lengths 128–133 and 512 in the range $n \le 512$, that improve on the best known values.

Length	New bound	Previous bound
128	1657012×2^{100}	1657009×2^{100}
133	1657010×2^{105}	1657009×2^{105}
512	$1021273028302258920 \times 2^{443}$	$1021273028302258913 \times 2^{443}$

Acknowledgments The authors wish to thank Tuvi Etzion for many helpful discussions and two anonymous referees for their constructive comments. This research was supported in part by grant ISF551-03.

References

1. Hämäläinen HO (1988) Two new binary codes with minimum distance three. IEEE Trans Inform Theory 34(4):875
2. Kabatyanskii GA, Panchenko VI (1988) Packings and coverings of the Hamming space by sphers of radius one. Problems Inform Trans 24(4):3–16
3. Kalbfleisch JG, Stanton RG, Horton JD (1971) On covering sets and error-correcting codes. J Comb Theory 11:233–250
4. Litsyn S (1998) An updated table of the best binary codes known. In: Pless VS, Huffman WC (eds) Handbook of Coding Theory vol. 1. Elsevier, Amsterdam, pp 463–498
5. MacWilliams FJ, Sloane NJA (1977) The theory of error-correcting codes. North-Holland, Amsterdam
6. Östergård PRJ, Kaikkonen MK (1996) New single-error-correcting codes. IEEE Trans Inform Theory 42(4):pp 1261–1262