Journal of Geometry
Vol. 54 (1995)

0047-2468/95/020030-14$1.50+0.20/0

ON NEAR-MDS CODES

Stefan Dodunekov† and Ivan Landgev ‡

1. INTRODUCTION

In the present paper we study a family of codes obtained by weakening the restrictions in the definition of classical Maximum-Distance-Separable (MDS) codes. This family of codes, which we call near-MDS (NMDS), contains remarkable representatives as the ternary Golay codes, the quaternary quadratic-residual $[11, 6, 5]$, the quaternary extended quadratic-residual $[12, 6, 6]$ code, as well as a large amount of algebraic geometric (AG) codes. Interesting connections of NMDS codes with arcs in finite projective planes, as well as with combinatorial designs, can be established.

The paper is written in the following way. In Section 2 we recall some necessary notions and results from coding theory. In Section 3 we introduce several definitions of an NMDS code and describe some of the basic properties of such codes. At the end of the section all binary NMDS codes are determined. The weight distribution of an NMDS code is calculated in Section 4. In Section 5 we investigate the maximum possible length of an NMDS code of fixed dimension and present some bounds on it. In the last Section 6 we study some geometric properties of NMDS codes, in particular, their connection with projective geometries. We improve on the bounds from Section 5 using some classical results on arcs in projective planes.

† Partially supported by the Bulgarian NSF under contract I-35/1991
‡ Supported by the Commission of the EC Contract CIPA 3511 CT 922960

2. PRELIMINARIES

At the beginning we recall some definitions and results about linear codes over an arbitrary finite field $\mathbf{F}_q = GF(q)$, $q = p^m$, p- a prime integer. For all notions and results not introduced here we refer to [9], [12], and [14].

Let $\mathcal{C}$ be a block code over $\mathbf{F}_q$. Denote by $\operatorname{supp}\mathcal{C}$ the set of coordinate positions, where not all codewords of $\mathcal{C}$ are zero and call it *the support* of $\mathcal{C}$. The support of a codeword is the set of its nonzero coordinate positions.

Let $\mathcal{C}$ be a linear $[n,k]$ code over $\mathbf{F}_q$, or an $[n,k]_q$ code for short [12,Ch.1]. The *r-th generalized Hamming weight* $d_r(\mathcal{C})$ [16] is defined to be the cardinality of the minimal support of an $[n,r]$ subcode of $\mathcal{C}$, $1 \le r \le k$, i.e.

$$d_r(\mathcal{C}) = \min\{\, |\operatorname{supp}\mathcal{D}| \,:\, \mathcal{D} \text{ is } [n,r]_q \text{ subcode of } \mathcal{C}\,\}. \tag{2.1}$$

Obviously, $d_1(\mathcal{C}) = d(\mathcal{C})$ is the minimum Hamming distance of $\mathcal{C}$.

For completeness we list below some results from [16] which we use throughout the paper.

LEMMA 2.1.[16] For every linear $[n,k]_q$ code $\mathcal{C}$

$$0 < d_1(\mathcal{C}) < d_2(\mathcal{C}) < \ldots < d_k(\mathcal{C}) \le n. \tag{2.2}$$

LEMMA 2.2.[16] Let $\mathbf{H}_\mathcal{C}$ be a parity check matrix of a linear code $\mathcal{C}$. Then $d_r(\mathcal{C}) = \delta$ if and only if

(a) any $\delta - 1$ columns of $\mathbf{H}_\mathcal{C}$ have rank greater or equal to $\delta - r$;

(b) there exist δ columns in $\mathbf{H}_\mathcal{C}$ of rank $\delta - r$.

LEMMA 2.3.[16] Let $\mathcal{C}$ be a linear $[n,k]_q$ code and let $\mathcal{C}^\perp$ be its dual. Then

$$\{\, d_r(\mathcal{C}) \,|\, r = 1,2,\ldots,k\,\} \cup \{\, n+1-d_r(\mathcal{C}^\perp) \,|\, r = 1,2,\ldots,n-k\,\} = \{\, 1,2,\ldots,n\,\}. \tag{2.3}$$

LEMMA 2.4.[16] (Generalized Singleton bound)

$$d_r(\mathcal{C}) \le n - k + r, \;\; r = 1,2,\ldots,k. \tag{2.4}$$

Let us recall also the Assmus-Mattson theorem [1], [2], or [12,Ch.6]. To emphasize that the minimum Hamming weight of an $[n,k]_q$ code $\mathcal{C}$ is equal to d we write its parameters as $[n,k,d]_q$.

THEOREM 2.5.(Assmus and Mattson) Let $\mathcal{C}$ be a linear $[n,k,d]_q$ code. Suppose we can find an integer t, $0 < t < d$, such that there are at most $d-t$ non-zero weights σ_i with $0 < \sigma_i \leq n-t$ in $\mathcal{C}^\perp$, the dual of $\mathcal{C}$. Then the supports of the codewords of weight d in $\mathcal{C}$ form a t-design.

3. NEAR-MDS CODES

A linear $[n,k]_q$ code $\mathcal{C}$ is said to be *near-MDS* if

$$d_i(\mathcal{C}) = n-k+i, \quad \text{for } i = 2,3,\ldots,k; \tag{3.1}$$

$$d_1(\mathcal{C}) = n-k. \tag{3.2}$$

It follows easily from this definition that $\mathcal{C}$ is near-MDS iff $d_1(\mathcal{C}) = n-k$ and $d_2(\mathcal{C}) = n-k+2$. Note that the near-MDS codes are codes of genus at most 1 in the terminology of [15].

Lemma 2.2 yields the following useful result.

LEMMA 3.1. A linear $[n,k]_q$ code $\mathcal{C}$ is near-MDS if and only if a parity-check matrix of $\mathcal{C}$, say $\mathbf{H}_\mathcal{C}$, (and consequently everyone of its parity-check matrices) satisfies the conditions

(N1) any $n-k-1$ columns of $\mathbf{H}_\mathcal{C}$ are linearly independent;

(N2) there exist $n-k$ linearly dependent columns;

(N3) any $n-k+1$ columns of $\mathbf{H}_\mathcal{C}$ are of full rank.

LEMMA 3.2. If a linear $[n,k]_q$ code is near-MDS then so is its dual.

Proof. It follows from Lemma 2.3 that

$$\{n+1-d_r(\mathcal{C}^\perp) | \, r = 1,2,\ldots,n-k\} = \{1,2,\ldots,n-k-1,n-k+1\},$$

whence

$$d_{n-k}(\mathcal{C}^\perp) = n,\ d_{n-k-1}(\mathcal{C}^\perp) = n-1,\ldots,d_2(\mathcal{C}^\perp) = k+2,\ d_1(\mathcal{C}^\perp) = k,$$

i.e. $\mathcal{C}^\perp$ is an $[n,n-k]_q$ near-MDS code, as asserted. $\diamond$

COROLLARY 3.3. A linear $[n,k]_q$ code is near-MDS if and only if $d(\mathcal{C})+d(\mathcal{C}^{\perp})=n$.

Note that Corollary 3.3 provides an alternative definition for near-MDS codes.

Lemma 3.2 implies that the generator matrix $\mathbf{G}_{\mathcal{C}}$ of a linear $[n,k]_q$ near-MDS code $\mathcal{C}$ must satisfy some conditions similar to (N1)-(N3):

(N1′) any $k-1$ columns of $\mathbf{G}_{\mathcal{C}}$ are linearly independent;

(N2′) there exist k linearly dependent columns in $\mathbf{G}_{\mathcal{C}}$;

(N3′) any $k+1$ columns of $\mathbf{G}_{\mathcal{C}}$ are of full rank.

Let us note that not every $[n,k,n-k]_q$ code is necessarily a near-MDS code. The construction given below is a modification of a construction from [15,Ch.1.1] and yields $[n,k,n-k]_q$ codes which are not near-MDS.

We start with an $[n,k,n-k+1]_q$ MDS code $\mathcal{C}$ with parity check matrix $\mathbf{H}_{\mathcal{C}}$. Adjoin a row to $\mathbf{H}_{\mathcal{C}}$ which is not a linear combination of its rows and which is of weight less than $k-1$. Denote the matrix obtained by $\mathbf{H}_{\mathcal{C}_1}$ and consider it as a check matrix of an $[n_1,k_1,d_1]_q$ code $\mathcal{C}_1$. The code $\mathcal{C}_1$ has parameters

$$n_1 = n,\ \ k_1 = k-1,\ \ d_1 \geq n_1 - k_1.$$

Obviously, $d(\mathcal{C}_1^{\perp}) < k-1 = k_1$, whence $d_1 = n_1 - k_1$ and according to Corollary 3.3 $\mathcal{C}_1$ is an $[n_1,k_1,n_1-k_1]_q$ code which is not near-MDS.

However, as the next statement shows, if n is large enough, every $[n,k,n-k]_q$ code is near-MDS code.

THEOREM 3.4. If $n > k+q$ every $[n,k,n-k]_q$ code is near-MDS code.

Proof. Let $n > k+q$ and supose that $\mathcal{C}$ is an $[n,k,n-k]_q$ code with parity check matrix $\mathbf{H}_{\mathcal{C}}$ which is not near-MDS. Then $\mathbf{H}_{\mathcal{C}}$ satisfies (N1) and (N2), but does not satisfy (N3), i.e. there exist $n-k+1$ columns in $\mathbf{H}_{\mathcal{C}}$, say $\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_{n-k+1}$, which are of rank $n-k-1$ or less. It follows from (N1) that $\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_{n-k+1}$ are of rank exactly $n-k-1$. Assume for concreteness that the first $n-k-1$ of the $\mathbf{h}_i$'s are linearly independent. Then we can write

$$\mathbf{h}_{n-k} = \sum_{i=1}^{n-k-1} \alpha_i \mathbf{h}_i, \tag{3.3}$$

$$\mathbf{h}_{n-k+1} = \sum_{i=1}^{n-k-1} \beta_i \mathbf{h}_i, \tag{3.4}$$

where $\alpha_i, \beta_i \in \mathbf{F}_q^* = \mathbf{F}_q \setminus \{0\}$, $i = 1, 2, \ldots, n-k-1$. Now consider the set $\{\alpha_i \beta_i^{-1} | i = 1, 2, \ldots, n-k-1\}$. It contains elements from $\mathbf{F}_q^*$ and since $n-k-1 > q-1$ at least two of them coincide, say

$$\alpha_{i_1} \beta_{i_1}^{-1} = \alpha_{i_2} \beta_{i_2}^{-1} = \gamma \in \mathbf{F}_q^*.$$

Hence at least two of the coefficients in the righthand side of

$$\mathbf{h}_{n-k} - \gamma \mathbf{h}_{n-k+1} = \sum_{i=1}^{n-k-1} (\alpha_i - \gamma \beta_i) \mathbf{h}_i \tag{3.5}$$

are zero. Thus we obtain $n-k-1$ linearly dependent columns of $\mathbf{H}_{\mathcal{C}}$, a contradiction to (N1). $\diamond$

THEOREM 3.5. Let $\mathcal{C}$ be a $[n, k, n-k]_q$ code with $k \geq 2$. Then

(i) $n \leq 2q + k$;

(ii) $\mathcal{C}$ is generated by its codewords of weight $n-k$ and $n-k+1$; if $n > q+k$ $\mathcal{C}$ is generated by its minimum weight codewords.

Proof. (i) One gets from the Griesmer bound [8]

$$n \geq g(k, n-k) = \sum_{i=0}^{k-1} \left\lceil \frac{n-k}{q^i} \right\rceil \geq n - k + s + k - 2 = n + s - 2, \tag{3.6}$$

where $s = \left\lceil \frac{n-k}{q} \right\rceil$ ($\lceil x \rceil$ denotes the smallest integer $\geq x$). Hence $s = 1$ or 2, and $\left\lceil \frac{n-k}{q} \right\rceil \leq 2$ implies $n \leq 2q + k$.

(ii) From (i) $n \leq 1 + g(k, n-k)$ and hence the code $\mathcal{C}$ is generated by its codewords of weight not greater than $n-k+1$ (see [5]). If $n > q+k$ the code $\mathcal{C}$ meets the Griesmer bound and is generated by its minimum weight codewords [5]. $\diamond$

In contrast to the situation with MDS-codes there exist some non-trivial binary near-MDS codes. In the rest of this section we present the complete list of all binary near-MDS codes.

Let $\mathcal{C}$ be an $[n, k]_2$ near-MDS code and let $\mathbf{G}_{\mathcal{C}} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n]$ be its generator matrix. Here $\mathbf{g}_i \in \mathbf{F}_2^k$, $i = 1, 2, \ldots, n$ denote the columns of $\mathbf{G}_{\mathcal{C}}$. According to Lemma 3.2 $\mathcal{C}^{\perp}$ is an $[n, n-k]_2$ near-MDS code. Applying Theorem 3.5(i) we get immediately that $k \leq 4$, or

dually $k \geq n-4$. Without loss of generality we can assume that $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_k$ are linearly independent. For every $i \in \{1, 2, \ldots, n-k\}$ we have

$$\mathbf{g}_{k+i} = \sum_{j \in J_i} \mathbf{g}_j, \tag{3.7}$$

where $J_i \subset \{1, 2, \ldots, k\}$, $|J_i| = k-1$, or k. Suppose that $n > k+1$ and suppose there exist indices $j_1, j_2 \in \{1, 2, \ldots, n-k\}$ with $|J_{j_1}| = k$, $|J_{j_2}| = k-1$. Then

$$\mathbf{g}_{k+j_1} + \mathbf{g}_{k+j_2} + \mathbf{g}_\alpha = \mathbf{0},$$

where $\{\alpha\} = J_{j_1} \setminus J_{j_2}$ and $\mathbf{0}$ is the k-dimensional zero vector. Hence in this case $k \leq 3$. If for every $j \in \{1, 2, \ldots, n-k\}$, $|J_j| = k-1$, then

$$\mathbf{g}_{k+1} + \mathbf{g}_{k+2} + \mathbf{g}_\alpha + \mathbf{g}_\beta = \mathbf{0},$$

where $\{\alpha, \beta\} = \{1, 2, \ldots, k\} \setminus (J_1 \cap J_2)$. Therefore, $k \leq 4$. Below we list all $[n,k]_2$ near-MDS codes. Because of the duality (Lemma 3.2) it is sufficient to consider codes with $n \geq 2k$.

Case $k = 1$. The only code here is the trivial $[n, 1, n-1]$ code.

Case $k = 2$. From Theorem 3.5(i) $n \leq 6$. Let

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1\ldots1 & 0\ldots0 & 1\ldots1 \\ \underbrace{0\ldots0}_{l_1} & \underbrace{1\ldots1}_{l_2} & \underbrace{1\ldots1}_{l_3} \end{pmatrix}. \tag{3.8}$$

It follows from (N2′) that at least one of l_1, l_2, l_3 is greater or equal to 2. On the other hand (N3′) implies $l_1 \leq 2, l_2 \leq 2, l_3 \leq 2$. Now, if $n = 6$, then explicitly $l_1 = l_2 = l_3 = 2$ and the unique $[6,2]_2$ near-MDS code is generated by

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{3.9}$$

If $n = 5$, one of the l_i's must be equal to 1. Up to equivalence

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}. \tag{3.10}$$

generates the unique $[5,2,3]_2$ near-MDS code. Similarly, for $n = 4$ we get two non-equivalent $[4,2,2]_2$ near-MDS codes, generated by

$$\mathbf{G}_{\mathcal{C}'} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \tag{3.11}$$

and by

$$\mathbf{G}_{\mathcal{C}''} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{3.12}$$

Case $k = 3$. According to Theorem 3.5(i) $n \leq 7$, and (N1′) implies that any two columns of $\mathbf{G}_{\mathcal{C}}$ are linearly independent, i.e. there are no repeated columns. Therefore, if $n = 7$ the only near-MDS code is the simplex code:

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}. \tag{3.13}$$

If $n = 6$ there exists one (up to equivalence) $[6, 3]_2$ near-MDS code which has generator matrix

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{3.14}$$

Case $k = 4$. From Theorem 3.5(i) $n \leq 8$. On the other hand according to Lemma 3.2 it suffices to consider $n - k \geq 4$, i.e. the case $n = 8$. There exists one $[8, 4, 4]_2$ code - the extended Hamming code. One generator matrix of this code is

$$\mathbf{G}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{3.15}$$

4. THE WEIGHT DISTRIBUTION OF A NEAR-MDS CODE

Given a linear $[n, k]_q$ code $\mathcal{C}$ denote by A_i the number of codewords in $\mathcal{C}$ which are of Hamming weight i, $i = 0, 1, 2, \ldots, n$. The set $\{A_i \mid i = 0, 1, \ldots, n\}$ is called the *weight distribution* of $\mathcal{C}$.

Similarly to the MDS codes, the weight distribution of a near-MDS code can be completely determined. The only small price we have to pay in this case is that the numbers A_i, $i = n - k + 1, \ldots, n$ are linear functions of A_{n-k}.

THEOREM 4.1. Let $\mathcal{C}$ be an $[n,k]_q$ near-MDS code. Let further $\{A_i|i=0,1,\ldots,n\}$ be its weight distribution and $\{A'_i|i=0,1,\ldots,n\}$ be the weight distribution of $\mathcal{C}^\perp$. Then for every $s \in \{1,2,\ldots,k\}$

$$A_{n-k+s} = \binom{n}{k-s}\sum_{j=0}^{s-1}(-1)^j\binom{n-k+s}{j}(q^{s-j}-1) + (-1)^s\binom{k}{s}A_{n-k}, \tag{4.1}$$

$$A'_{k+s} = \binom{n}{k+s}\sum_{j=0}^{s-1}(-1)^j\binom{k+s}{j}(q^{s-j}-1) + (-1)^s\binom{n-k}{s}A'_k. \tag{4.2}$$

Proof. We start with the equality

$$\sum_{i=0}^{n-\nu}\binom{n-i}{\nu}A_i = q^{k-\nu}\sum_{i=0}^{\nu}\binom{n-i}{\nu-i}A'_i \tag{4.3}$$

(see [12], Problem (6) on p.131, restated for arbitrary q). To prove (4.1) we use induction on s. For $s=1$ it can be obtained from (4.3) by setting $\nu = k-1$. Now suppose that (4.1) holds for every $s = 1,2,\ldots,\sigma$. Setting $\nu = k-\sigma+1$ in (4.3) one gets

$$A_{n-k+\sigma+1} = \binom{n}{k-\sigma+1}(q^{\sigma+1}-1) - \sum_{j=0}^{\sigma}\binom{k-j}{k-\sigma-1}A_{n-k+j},$$

whence after some tedious but straightforward calculations the desired expression for $A_{n-k+\sigma+1}$ can be obtained.

The formula (4.2) can be proven in the same fashion, setting $\nu = n-k, n-k-1,\ldots,2,1$ in (4.3). $\diamond$

Remark. Note that (4.3) yields $A_{n-k} = A'_k$ for $\nu = k$. Hence the weight distributions of $\mathcal{C}$ and $\mathcal{C}^\perp$ coincide for $n = 2k$, i.e. the code $\mathcal{C}$ is formally self-dual [15,Ch.1.1].

COROLLARY 4.2. For an $[n,k]_q$ near-MDS code $\mathcal{C}$

$$A_{n-k} \le \binom{n}{k-1}\frac{q-1}{k}, \tag{4.4}$$

with equality iff $A_{n-k+1} = 0$. By duality

$$A'_k \le \binom{n}{k+1}\frac{q-1}{n-k}, \tag{4.5}$$

with equality iff $A'_{k+1} = 0$.

Proof. The inequality (4.4) follows from

$$A_{n-k+1} = \binom{n}{k-1}(q-1) - kA_{n-k} \geq 0.$$

Similarly, (4.5) is derived from (4.2) and $A'_{k+1} \geq 0$. $\diamond$

Remark. Note that Theorem 1.1.16 from [15] yields the inequality $A_{n-k} \leq \binom{n}{k}(q-1)$, which is worse than (4.4).

COROLLARY 4.3. For an $[n,k]_q$ near-MDS code with $A_{n-k+1} = 0$ we have $k \leq \frac{n}{2}$.

Proof. It follows from Corollary 4.2 that

$$\binom{n}{k+1}\frac{q-1}{n-k} \geq A'_k = A_{n-k} = \binom{n}{k-1}\frac{q-1}{k},$$

whence $k \leq \frac{n}{2}$. $\diamond$

5. NEAR-MDS CODES OF MAXIMAL LENGTH

One of the most fascinating problems connected with MDS codes is the following: given k and q, find the largest value $m(k,q)$ of n, for which there exists an $[n,k,n-k+1]_q$ MDS code over $\mathbf{F}_q$ (see [12,Ch.11] and the references there). A general upper bound is

$$m(k,q) \leq q+k-1. \tag{5.1}$$

Define $m'(k,q)$ as the maximum possible length of a near-MDS code of fixed dimension k over a fixed field $\mathbf{F}_q$.

PROPOSITION 5.1. $m'(k,q) \leq 2q+k$.
In case of equality $A_{n-k+1} = 0$.

Proof. The first part follows from Theorem 3.5. We give an alternative proof below. One gets from (4.1)

$$A_{n-k+2} = \binom{n}{k-2}(q-1)(q+k-n-1) + \binom{k}{2}A_{n-k}.$$

The inequality (4.4) implies

$$\frac{q-1}{2}\binom{n}{k-2}(2q+k-n) = \binom{n}{k-2}(q-1)(q+k-n-1)+\binom{k}{2}\binom{n}{k-1}\frac{q-1}{k} \geq A_{n-k+2} \geq 0$$

whence (5.2) follows.

If $n = 2q + k$ the above inequality implies $A_{n-k} = \binom{n}{k-1}(q-1)/k$ and by Corollary 4.2 $A_{n-k+1} = 0$. $\diamond$

A near-MDS code meeting the bound from Proposition 5.1 will be called *extremal.*

PROPOSITION 5.2. It holds $m'(k,q) = k + 1$ for every $k > 2q$.

Proof. The existence of $[k+1, k]_q$ near-MDS codes is obvious for every prime power q. Suppose that $\mathcal{C}$ is an $[n,k]_q$ near-MDS code with $n > k + 1 > 2q + 1$. Let

$$\mathbf{G}_{\mathcal{C}} = [\mathbf{g}_1\, \mathbf{g}_2\, \ldots\, \mathbf{g}_n],\ \mathbf{g}_i \in \mathbf{F}_i^k \tag{5.2}$$

be its generator matrix and let the first k columns $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_k$ be linearly independent. Then

$$\mathbf{g}_{k+1} = \sum_{i \in J_1} \alpha_i \mathbf{g}_i,\ \alpha_i \in \mathbf{F}_q^*, \tag{5.3}$$

$$\mathbf{g}_{k+2} = \sum_{i \in J_2} \beta_i \mathbf{g}_i,\ \beta_i \in \mathbf{F}_q^*, \tag{5.4}$$

where J_1 and J_2 are subsets of $\{1, 2, \ldots, k\}$ of cardinality at least $k-1$. Since $|J_1 \cap J_2| \geq k-2 > 2(q-1)$ there exist indices i_1, i_2, i_3 such that $\alpha_{i_1}\beta_{i_1}^{-1} = \alpha_{i_2}\beta_{i_2}^{-1} = \alpha_{i_3}\beta_{i_3}^{-1} = \mu \in \mathbf{F}_q^*$. Then

$$\mathbf{g}_{k+1} - \mu \mathbf{g}_{k+2} = \sum_{i \in J} \varphi_i \mathbf{g}_i,\ \varphi_i \in \mathbf{F}_q^*, \tag{5.5}$$

wehre $J \subset \{1, 2, \ldots, n\}$, $|J| \leq k-3$, a contradiction to (N1'). $\diamond$

Extremal near-MDS codes with $k = 2$ exist for every q. To check this take two vectors from every one-dimensional subspace of the vector space of all pairs over $\mathbf{F}_q$. The matrix having these $2(q^2-1)/(q-1) = 2q+2$ vectors as columns is a generator matrix of an extremal near-MDS code.

PROPOSITION 5.3. The existence of an $[n,k]_q$ near-MDS code $\mathcal{C}$ implies

(i) the existence of an $[n-1, k-1]_q$ near-MDS code;

(ii) the existence of an $[n-1, k]_q$ near-MDS code.

Proof. (i) Delete a column from the parity check matrix of $\mathcal{C}$, preserving a set of $n-k$ linearly dependent columns.
(ii) Delete a column from the generator matrix of $\mathcal{C}$, preserving a set of $k+1$ columns which contains k linearly dependent columns.

COROLLARY 5.4. For every integer α, $0 \leq \alpha \leq k$, it holds $m'(k,q) \leq m'(k-\alpha,q)+\alpha$. In particular, the existence of an extremal near-MDS code of dimension k over $\mathbf{F}_q$ implies the existence of extremal near-MDS codes for every dimension $k' \leq k$.

Proof. Proposition 5.3(i) implies $m'(k-1,q) \geq m'(k,q)-1$. Applying this α times we get (5.3). If $m'(k,q) = 2q+k$ then $m'(k-\alpha,q) \geq 2q+k-\alpha$, and hence by (5.2) $m'(k-\alpha,q) = 2q+k-\alpha$. $\diamond$

Example. The ternary $[12,6,6]$ extended Golay code [7] is an extremal near-MDS code and yields extremal near-MDS codes over $\mathbf{F}_3$ for every $k \leq 6$.

It is known that one can construct algebraic geometric $[n,k,n-k]_q$ codes ($q = p^m$) of genus at most 1 (i.e. MDS, or near-MDS codes) for every n for which

$$n \leq \begin{cases} q+\lceil 2\sqrt{q} \rceil & \text{if } p \text{ divides } \lceil 2\sqrt{q} \rceil \text{ and } m \geq 3 \text{ is odd,} \\ q+\lceil 2\sqrt{q} \rceil + 1 & \text{otherwise,} \end{cases} \tag{5.6}$$

and arbitrary $k = 2,3,\ldots,n-2$ [15,Ch.3.2]. Therefore near-MDS codes of length greater than the one given by (5.6) are of special interest. One such example is the quaternary $[12,6,6]$ code [6], as well as the codes derived from it by Proposition 5.3.

Comparing the bound (5.1) with Proposition 5.1 we see that there exist near-MDS codes which are considerably longer than the longest possible MDS code with the same k and q. We close this section with the observation that near-MDS codes can sometimes produce t-designs.

PROPOSITION 5.5. Let $\mathcal{C}$ be an $[n,k]_q$ near-MDS code. Suppose there exists an integer $s \geq 1$ such that $A_{n-k+s} = 0$. Then the words of weight k in $\mathcal{C}^\perp$ form a $(k-s)$-design . In particular, the words of minimal weight in the dual of an extremal near-MDS code form a Steiner system $S(k-1;k,2q+k)$.

Proof. Use the Assmus-Mattson theorem (Theorem 2.5) with $t = k-s$. For the second part note that in the design obtained from the dual of an extremal near-MDS code each

$(k-1)$-tuple is contained in

$$\lambda = \frac{A'_k}{q-1}\frac{\binom{k}{k-1}}{\binom{2q+k}{k-1}} = 1$$

block. $\diamond$

Steiner systems $S(t;k,v)$ with large t are extremely rare. Therefore, one might expect that so are extremal near-MDS codes. Later on we shall prove that the $[12,6,6]$ ternary extended Golay code and the codes obtained from it by Proposition 5.3(i) are the only extremal near-MDS codes with $q \geq 3$, $k \geq 3$.

6. NEAR-MDS CODES AND PROJECTIVE GEOMETRIES

Let $\mathcal{C}$ be a near-MDS code with $k \geq 3$ and let $\mathbf{G}_\mathcal{C} = [\mathbf{g}_1\, \mathbf{g}_2\, \ldots\, \mathbf{g}_n], \mathbf{g}_i \in \mathbf{F}_q^k$ be its generator matrix. The columns of $\mathbf{G}_\mathcal{C}$ can be looked at as different (because of $k \geq 3$) points in the projective geometry $\mathbf{PG}(k-1,q)$. In other words a near-MDS code of dimension $k \geq 3$ is always projective (cf. [4]). The existence of an $[n,k]$ near-MDS code is equivalent to the existence of a set $\mathcal{S}$ of points in $\mathbf{PG}(k-1,q)$ having the properties

(N1″) every $k-1$ points from $\mathcal{S}$ generate a hyperplane in $\mathbf{PG}(k-1,q)$;

(N2″) there exist k points in $\mathcal{S}$ lying on a hyperplane;

(N3″) every $k+1$ points from $\mathcal{S}$ generate $\mathbf{PG}(k-1,q)$.

These properties become very simple for $k=3$. In such case the existence of a near-MDS code is equivalent to the existence of a set $\mathcal{S}$ of points in the projective plane $\mathbf{PG}(2,q)$ having the properties

- there exist three collinear points in $\mathcal{S}$;
- no four points from $\mathcal{S}$ lie on a line.

A (κ,ν)-*arc* in the projective plane $\mathbf{PG}(2,q)$ is a set $\mathcal{S}$ of κ points such that each line meets $\mathcal{S}$ in at most ν points and there exists a line meeting it in exactly ν points. For every (κ,ν)-arc in $\mathbf{PG}(2,q)$ we have

$$\kappa \leq (\nu-1)q + \nu, \tag{6.1}$$

(cf. [11], p.322). It is clear that near-MDS codes of dimension 3 over $\mathbf{F}_q$ are equivalent to $(n,3)$-arcs in $\mathbf{PG}(2,q)$. The inequality (6.1) coincides with (5.2) for $k=\nu=3$.

Let τ_i, $i = 0, 1, \ldots, \nu$ denote the number of lines meeting $\mathcal{S}$ in exactly i poits. The numbers τ_i determine the weight distribution of the corresponding near-MDS code by

$$(q-1)\tau_i = A_{n-i},\ i = 0, 1, 2, 3. \tag{6.2}$$

Example. Consider a Desarguesian configuration $\mathcal{D}$ in $\mathbf{PG}(2, 7)$. The set of the intersecting points of the 10 lines of $\mathcal{D}$ which are not in $\mathcal{D}$ is a (15,3)-arc with

$$\tau_0 = 12,\ \tau_1 = 0,\ \tau_2 = 15,\ \tau_3 = 30. \tag{6.3}$$

Therefore, there exists a $[15, 3, 12]_7$ near-MDS code $\mathcal{C}$ with $A_{14} = 0$. Proposition 5.5 implies that the words of weight 3 in $\mathcal{C}^{\perp}$ (which is a $[15, 12, 3]_7$ code) form a 1-design.

The bound (5.2) can be strengthened using a well-known result of Thas [13] (see also [11, p.335]).

PROPOSITION 6.1. For every $(\kappa, 3)$-arc in $PG(2, q)$ with $q > 3$ we have $\kappa \leq 2q + 1$. $\diamond$

PROPOSITION 6.2. If $q > 3$ then

$$m'(k, q) \leq 2q + k - 2. \tag{6.4}$$

The only ternary extremal near-MDS codes are the extended Golay code and the codes with parameters $[12 - i, 6 - i, 6]_3$, $1 \leq i \leq 5$.

Proof. From every $[n, k]_q$ near-MDS code we can construct (see Proposition 5.3(i)) an $[n-k+3, 3]_q$ near-MDS code and hence an $(n-k+3, 3)$-arc in $PG(2, q)$. Now Proposition 6.1 yields

$$n - k + 3 \leq 2q + 1, \tag{6.5}$$

which proves (6.4). The rest follows from Proposition 5.2. It has to be noted that all the codes with parameters $[12 - i, 6 - i, 6]_3$, $i = 0, 1, \ldots, 5$, are unique. $\diamond$

The maximal cardinalities for 3-arcs are known for $q = 4, 5, 7, 8, 9$ [3],[10]. (The cases $q = 2$, and $q = 3$ are trivial and have been already discussed.) This implies that $m'(3, 4) = 9$, $m'(3, 5) = 11$, $m'(3, 7) = 15$, $m'(3, 8) = 15$, $m'(3, 9) = 17$. It follows from [3] that $21 \leq m'(3, 11) \leq 23$, $23 \leq m'(3, 13) \leq 27$. It is conjectured that for every $(\kappa, 3)$-arc in

$PG(2,q)$, where $q \geq 8$, $\kappa \leq 2q-1$. This is only known to be true for $q = 8, 9$. Of course, every restriction on the number of points in a maximal 3-arc would imply an improvement on (6.4).

Acknowledgements. The paper was written during a visit of the authors at the department of Electrical Engineering, Linköping University.

REFERENCES

[1] E.F.Assmus,Jr., H.F.Mattson, New 5-designs, *J.Combin. Theory* **6**(1969), 122–151.

[2] E.F.Assmus,Jr., H.F.Mattson, Coding and Combinatorics, *SIAM Review* **16**(1974), 349–388.

[3] S.M.Ball, On Sets of Points in Finite Planes, manuscript.

[4] R.Calderbank, W.M.Kantor, The Geometry of Two-weight Codes, *Bull. London Math. Soc.* **18**(1986), 97–122.

[5] S.M.Dodunekov, A Note on the Weight Structure of Generator Matrices of Linear Codes, *Problemi Peredachi Informacii* **26**(1990), 101– 104, in Russian.

[6] I.I.Dumer, V.A.Zinoviev, Some New Maximal Codes over GF(4), *Problemi Peredachi Informacii* **14**(1978), 24–34 in Russian.

[7] M.J.E.Golay, Notes on Digital Coding, *Proc. IEEE* **37**(1949), 657.

[8] J.H.Griesmer, A Bound for Error-correcting Codes, *IBM J. Res. Develop.* **4**(1960), 532–542.

[9] W.Heise, P.Quattrocchi, *Informations- und Codierungstheorie*, Springer, Berlin - Heidelberg, 1988.

[10] R.Hill, J.Mason, On (k,n)-arcs and the Falsity of Lunelli-Sce Conjecture, *London Math. Soc. Lecture Note Series* **49**(1981), 153–168.

[11] J.W.P.Hirschfeld, *Projective Geometries over Finite Fields*, Clarendon Press, Oxford, 1979.

[12] F.J.MacWilliams, N.J.A.Sloane, *The Theory of Error-correcting Codes*, North Holland, Amsterdam, 1977.

[13] J.Thas, Some Results Concerning $((q+1)(n-1),n)$-arcs, *J.Combin. Theory Ser.A* **19**(1975), 228–232.

[14] V.D.Tonchev, *Combinatorial Configurations*, Longman, Wiley, New York, 1988.

[15] M.A.Tsfasman, S.G.Vladut, *Algebraic-geometric Codes*, Kluwer Academic Publishers, Dordrecht-Boston-London, 1991.

[16] V.K.Wei, Generalized Hamming Weights for Linear Codes, *IEEE Trans. Inform. Theory* **IT-37**(1991), 1412–1418.

S.M.Dodunekov
Institute of Mathematics
Bulgarian Academy of Sciences
8. Acad. G.Bonchev Str.
1113 Sofia, Bulgaria

I.N.Landgev
Institute of Mathematics
Bulgarian Academy of Sciences
8 Acad. G.Bonchev Str.
1113 Sofia, Bulgaria

Eingegangen am 14. Februar 1994