Upper Bounds on the Weight Distribution Function for Some Classes of Linear Codes

Torleiv Kløve, *Fellow, IEEE*, and Jinquan Luo

***Abstract*—Upper bounds on the weight distribution function for codes of minimum distance at least 2 are given. Codes, where the bound is met with equality, are characterized. An improved upper bound on the weight distribution function for codes of minimum distance at least 3 is given. As an application, a sharp upper bound on the probability of undetected error for linear codes with full support is characterized.**

***Index Terms*—Full support, linear code, upper bound, weight distribution.**

I. Introduction

Let F_q denote the field of q elements and F_q^* the set of non-zero elements of F_q. Let $n \geq k \geq 1$. An $[n,k;q]$ code is a linear code of length n and dimension k over F_q. If the minimum distance of the code is d, it is also called an $[n,k,d;q]$ code.

For an $[n,k;q]$ code C, let $A_i(C)$ be the number of codewords of Hamming weight i. The sequence $A_0(C), A_1(C), \ldots, A_n(C)$ is known as the weight distribution of C, and

$$A_C(z) = \sum_{i=0}^{n} A_i(C)\, z^i$$

the weight distribution function of C.

The weight distribution function has several applications in coding theory. Important examples are bounding the error probability for maximum likelihood (ML) decoding when the code is used for error correction and expressing the probability of undetected error when the code is used for error detection.

The q-ary symmetric channel with error probability parameter p is a discrete memoryless channel, that is, a symbol $a \in F_q$ is modified into $b \in F_q$ independently of what happens to other symbols in the transmission. The probability that a is modified into $b \neq a$ is $p/(q-1)$.

Manuscript received November 02, 2011; revised February 25, 2012; accepted April 16, 2012. Date of publication May 1, 2012; date of current version July 10, 2012. This work was supported by the Norwegian Research Council under Grant 191104/V30. J. Luo was supported in part by the National Science Foundation (NSF) of China under Grant 60903036, in part by the NSF of Jiangsu Province under Grant 2009182, and in part by the open research fund of the National Mobile Communications Research Laboratory, Southeast University (No. 2010D12).

The authors are with the Department of Informatics, University of Bergen, N-5020 Bergen, Norway (e-mail: Torleiv.Klove@ii.uib.no; Jinquan.Luo@ii.uib.no).

Communicated by E. Arıkan, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2012.2197176

The word error decoding probability for ML decoding is upper bounded by

$$P_E \leq A_C(\gamma) - 1 \tag{1}$$

where

$$\gamma = 2\sqrt{\frac{p(1-p)}{q-1}} + \frac{(q-2)p}{q-1}$$

see, e.g., [1, Problem 7.10 and Th. 7.5]. The bound (1) is true also for many other channels, where the definition of γ depends on the channel; see, e.g., [2, Sec. IV].

For an $[n,k;q]$ code C, the probability of undetected error $P_{\rm ue}(C,p)$ is the probability that a codeword is changed to another codeword when transmitted over the q-ary symmetric channel.

It is known and easy to see (cf., [3, Th. 2.1]) that

$$P_{\rm ue}(C,p) = (1-p)^n \left\{ A_C\left(\frac{p}{(q-1)(1-p)}\right) - 1 \right\}. \tag{2}$$

The goal of this paper is to study upper bounds on $A_C(z)$ (for $0 \leq z \leq 1$) under some conditions on the code C. If

$$A_C(z) \leq f(z) \tag{3}$$

then

$$P_{\rm ue}(C,p) = (1-p)^n \left\{ f\left(\frac{p}{(q-1)(1-p)}\right) - 1 \right\} \tag{4}$$

for $p \in [0,(q-1)/q]$. We can also reverse the implication. We observe that

$$z = \frac{p}{(q-1)(1-p)}$$

runs through $[0,1]$ when p runs through $[0,(q-1)/q]$. Hence, (2) implies that if $P_{\rm ue}(C,p) \leq \phi(p)$ for all $p \in [0,(q-1)/q]$, then

$$A_C(z) \leq 1 + \big(1+(q-1)z\big)^n \phi\left(\frac{(q-1)z}{1+(q-1)z}\right). \tag{5}$$

An upper bound on A_i for an $[n,k,d;q]$ code was given by Levy [4]. The bound is also given in [3, Th. 1.22] in a slightly different formulation. For completeness, we quote Levy's bound, essentially using his formulations. Let $N_n(d,j,\nu)$ be the number of vectors of length n and weight $j+\nu$ within distance $\lfloor (d-1)/2 \rfloor$ of a fixed vector of weight j. An expression

for $N_n(d,j,\nu)$ was first given by MacWilliams [5]. Levy [4] gave essentially the following formulation of the expression:

$$N_n(d,j,\nu)=\sum_{i,\alpha}\binom{j}{i}\binom{j-i}{\alpha}\binom{n-j}{\nu+i}(q-2)^{\alpha}(q-1)^{\nu+i}$$

where the summation is over all $\alpha \geq 0$ and all i such that $\max(0,-\nu) \leq i \leq \big(\lfloor (d-1)/2 \rfloor - \nu - \alpha\big)/2$. Levy's upper bound for the weights of an $[n,k,d;q]$ code is as follows.

Theorem 1: If C is a code over F_q of length n and minimum distance d, then

$$A_j(C) \leq \min_{\substack{\nu \\ N_n(d,j,\nu)>0}} \frac{\binom{n}{j-\nu}(q-1)^{j-\nu}}{N_n(d,j,\nu)}.$$

We note that the bound does not depend on the dimension of the code.

Other early results [6]–[8] are upper bounds on the *average* probability of undetected error for all $[n,k;q]$ codes.

Two simple general bounds were given in [9] for the binary case, and generalized to the q-ary case in [3, Ths. 2.49 and 2.51]. We include the bounds (in terms of the weight distribution) and the proofs since a modified version will be used later to prove an improved bound.

Theorem 2: If C is an $[n,k,\delta;q]$ code and $\delta \geq d$, then

$$A_C(z) \leq 1+(q^k-1)z^d$$

for all $z \in [0,1]$.

Proof: For $z \in [0,1]$, we have $z^i \leq z^j$ for $i \geq j$. Since $w(\mathbf{x}) \geq d$ for all non-zero codewords, we get $z^{w(\mathbf{x})} \leq z^d$ and so

$$A_C(z)=\sum_{\mathbf{x}\in C} z^{w(\mathbf{x})} \leq 1+(q^k-1)z^d.$$

■

For $d<k$, we have the following improvement. Let

$$\begin{aligned} f_{q,k,d}(z) &= 1+\sum_{i=1}^{d}\binom{k}{i}(q-1)^i z^d+\sum_{i=d+1}^{k}\binom{k}{i}(q-1)^i z^i \\ &= (1+(q-1)z)^k+\sum_{i=1}^{d}\binom{k}{i}(q-1)^i\big(z^d-z^i\big). \end{aligned}$$

Theorem 3: If C is an $[n,k,\delta;q]$ code and $\delta \geq d$ where $d \leq k$, then $A_C(z) \leq f_{q,k,d}(z)$ for all $z \in [0,1]$.

Proof: Equivalent codes have the same weight distribution. There exists an equivalent code C' generated by a matrix $G=[I_k|Q]$, where I_k is the $k \times k$ identify matrix and Q is some $k \times (n-k)$ matrix, that is, $C'=\{\mathbf{x}G \mid \mathbf{x} \in F_q^k\}$. We have $\mathbf{x}G=(\mathbf{x}|\mathbf{x}Q)$ and so, for $\mathbf{x} \neq \mathbf{0}$, we have

$$w(\mathbf{x}G)=w(\mathbf{x})+w(\mathbf{x}Q) \geq \max(d,w(\mathbf{x})).$$

Hence

$$\begin{aligned} A_C(z) &= \sum_{\mathbf{x}\in F_q^k} z^{w(\mathbf{x}G)} = \sum_{i=0}^{k}\sum_{\substack{\mathbf{x}\in F_q^k \\ w(\mathbf{x})=i}} z^{w(\mathbf{x}G)} \\ &\leq 1+\sum_{i=1}^{d}\binom{k}{i}(q-1)^i z^d+\sum_{i=d+1}^{n}\binom{k}{i}(q-1)^i z^i. \end{aligned}$$

■

Since any code has minimum distance at least 1, putting $\delta=1$ in Theorem 3, we get the following trivial bound valid for all $[n,k;q]$ codes.

Corollary 1: If C is an $[n,k;q]$ code, then

$$A_C(z) \leq (1+(q-1)z)^k.$$

We note that the bounds in Theorems 2 and 3 depend on k, but not on n. For an $[n,k;q]$ code C, the dual code $C^{\perp}$ is defined by

$$C^{\perp}=\{\mathbf{x} \in F_q^n \mid \mathbf{c}\cdot\mathbf{x}=0 \quad \text{for all } \mathbf{c} \in C\}.$$

Here, $\cdot$ denotes the inner product, that is

$$(c_1,c_2,\ldots,c_n)\cdot(x_1,x_2,\ldots,x_n)=c_1x_1+c_2x_2+\cdots+c_nx_n.$$

The MacWilliams Theorem [5] states that if C is an $[n,k;q]$ code, then

$$A_{C^{\perp}}(z)=\frac{1}{q^k}(1+(q-1)z)^n A_C\left(\frac{1-z}{1+(q-1)z}\right). \quad (6)$$

This implies that if $A_C(z) \leq f(z)$ for all $z \in [0,1]$, then

$$A_{C^{\perp}}(z) \leq \frac{1}{q^k}(1+(q-1)z)^n f\left(\frac{1-z}{1+(q-1)z}\right)$$

for all $z \in [0,1]$. For example, Theorem 3 implies the following corollary.

Corollary 2: If C is an $[n,k;q]$ code with minimum distance $\delta \geq d$, then

$$\begin{aligned} q^k A_{C^{\perp}}(z) \leq{}& (1+(q-1)z)^n \\ &+\sum_{i=1}^{d}\binom{k}{i}(q-1)^i(1-z)^d(1+(q-1)z)^{n-d} \\ &+\sum_{i=d+1}^{n}\binom{k}{i}(q-1)^i(1-z)^i(1+(q-1)z)^{n-i} \end{aligned}$$

for all $z \in [0,1]$.

It is not simple to compare the bound on $A_C(z)$ in Theorem 3 and the bound obtained from Levy's bound in Theorem 1. As noted previously, the bound in Theorem 3 depends on k, but not on n. On the other hand, Levy's bounds depend on n, but not on k. Some numerical examples indicate that Theorem 3 is better

for small values of d whereas Levy's bound is better for larger d. In particular, for $d = 1$ and $d = 2$ we get $N_n(d, j, \nu) = 0$ for $\nu \neq 0$ and $N_n(d, j, 0) = 1$. Hence, Theorem 1 gives the trivial bounds $A_j \leq \binom{n}{j}(q-1)^j$.

The main goal of this paper is to give improvements of Theorem 3 under some conditions. In Section II, we give such bounds when $d \geq 2$ and in Section III, when $d \geq 3$. In Section V, we give bound for codes of full support and in Section VI, for codes of full support and minimum distance at least 2. Finally, we give some examples and a summary.

II. Upper Bound on $A_C(z)$ for Linear Codes C of Minimum Distance at Least 2

We now give an improvement of Theorem 3 when $\delta = 2$, that is, $d \geq 2$. We first give some lemmas that will be needed in the proof of the improved bound.

The first lemma is essentially the same as [3, Corollary 2.1]. We give the lemma and its proof.

Lemma 1: Let C be an $[n, k; q]$ code. If α_i, $i = 1, 2, \ldots, n$ are integers such that

$$\sum_{i=1}^{j} A_i(C) \leq \sum_{i=1}^{j} \alpha_i$$

for any $1 \leq j \leq n$, then, for all $z \in [0, 1]$, we have

$$A_C(z) \leq 1 + \sum_{i=1}^{n} \alpha_i z^i.$$

Moreover, we have equality for any $z \in (0, 1)$ if and only if $A_i(C) = \alpha_i$ for all i, $1 \leq i \leq n$.

Proof: Let $\sigma_0 = 0$ and

$$\sigma_j = \sum_{i=1}^{j} (\alpha_i - A_i) \quad \text{for } 1 \leq i \leq n.$$

Then

$$\begin{aligned}\left(1 + \sum_{i=1}^{n} \alpha_i z^i\right) - A_C(z) &= \sum_{i=1}^{n} (\alpha_i - A_i) z^i \\ &= \sum_{i=1}^{n} (\sigma_i - \sigma_{i-1}) z^i \\ &= \sum_{i=1}^{n} \sigma_i z^i - \sum_{i=0}^{n-1} \sigma_i z^{i+1} \\ &= \sigma_n z^n + \sum_{i=1}^{n-1} \sigma_i z^i (1 - z) \geq 0\end{aligned}$$

since $\sigma_i \geq 0$ by assumption, and $z^i(1 - z) > 0$. Moreover, if $\sigma_i > 0$ for any i, then $1 + \sum_{i=1}^{n} \alpha_i z^i - A_C(z) > 0$. ■

Let $D_{n,k,\mathbf{v}}$ be the code generated by

$$\begin{bmatrix} 1 & \mathbf{0}_{k-1} & \mathbf{v} \\ \mathbf{0}_{k-1}^T & I_{k-1} & O_{(k-1)\times(n-k)} \end{bmatrix}$$

where $\mathbf{0}_{k-1}$ is the all zero vector of length $k-1$, $\mathbf{0}_{k-1}^T$ is its transpose, $\mathbf{v} \in F_q^{n-k}$ is a vector of full support (that is, without zero in any position), and $O_{(k-1)\times(n-k)}$ is the $(k-1)\times(n-k)$ matrix of all zeros.

Lemma 2: The weight distribution of $D_{n,k,\mathbf{v}}$ is

$$A_{D_{n,k,\mathbf{v}}}(z) = \big(1 + (q-1)z\big)^{k-1}\big(1 + (q-1)z^{n-k+1}\big).$$

Proof: The codewords in $D_{n,k,\mathbf{v}}$ are all vectors of the form $(\alpha|\mathbf{x}|\alpha\mathbf{v})$, where $\alpha \in F_q$ and $\mathbf{x} \in F_q^{k-1}$. We have

$$w(\alpha|\mathbf{x}|\alpha\mathbf{v}) = \begin{cases} w(\mathbf{x}), & \text{if } \alpha = 0 \\ w(\mathbf{x}) + n - k + 1, & \text{if } \alpha \neq 0. \end{cases}$$

Hence

$$\begin{aligned}A_{D_{n,k,\mathbf{v}}}(z) &= \sum_{\alpha \in F_q} \sum_{\mathbf{x} \in F_q^{k-1}} z^{w(\alpha|\mathbf{x}|\alpha\mathbf{v})} \\ &= \sum_{\mathbf{x} \in F_q^{k-1}} z^{w(\mathbf{x})} + \sum_{\alpha \in F_q^*} \sum_{\mathbf{x} \in F_q^{k-1}} z^{w(\mathbf{x})+n-k+1} \\ &= \big(1 + (q-1)z^{n-k+1}\big) \sum_{\mathbf{x} \in F_q^{k-1}} z^{w(\mathbf{x})} \\ &= \big(1 + (q-1)z^{n-k+1}\big)\big(1 + (q-1)z\big)^{k-1}.\end{aligned}$$

■

Let $E_{n,k,\mathbf{v}} = D_{n,n-k,\mathbf{v}}^{\perp}$. This code is generated by the matrix $[I_k|\mathbf{v}^T|O_{k\times(n-k-1)}]$.

Using (6), we see that

$$A_{E_{n,k,\mathbf{v}}}(z) = \frac{1}{q}\Big\{\big(1+(q-1)z\big)^{k+1} + (q-1)(1-z)^{k+1}\Big\}. \quad (7)$$

Lemma 3: Let $\mathbf{v}$ be a vector of full support. Then

a)

$$A_i(E_{n,k,\mathbf{v}}) = \frac{1}{q}\binom{k+1}{i}\left\{(q-1)^i + (q-1)(-1)^i\right\}.$$

b)

$$\begin{aligned}\sum_{i=2}^{j} A_i(E_{n,k,\mathbf{v}}) = &\sum_{i=1}^{j-1} \binom{k}{i}(q-1)^i \\ &+ \frac{1}{q}\binom{k}{j}\left\{(q-1)^j + (-1)^j(q-1)\right\}. \quad (8)\end{aligned}$$

Proof: We see that a) follows immediately from (7). From a), we get

$$\begin{aligned}\sum_{i=2}^{j} A_i(E_{n,k,\mathbf{v}}) = &\frac{1}{q}\sum_{i=2}^{j}\binom{k+1}{i}(q-1)^i \\ &+ \frac{q-1}{q}\sum_{i=2}^{j}\binom{k+1}{i}(-1)^i.\end{aligned}$$

Let

$$F(z) = \sum_{i=2}^{j} \binom{k+1}{i} z^i.$$

Then

$$\begin{aligned} F(z) &= \sum_{i=2}^{j} \binom{k}{i} z^i + \sum_{i=2}^{j} \binom{k}{i-1} z^i \\ &= \sum_{i=2}^{j} \binom{k}{i} z^i + \sum_{i=1}^{j-1} \binom{k}{i} z^{i+1} \\ &= (z+1) \sum_{i=1}^{j-1} \binom{k}{i} z^i + \binom{k}{j} z^j - kz. \end{aligned}$$

Hence

$$\begin{aligned} & q \sum_{i=2}^{j} A_i(E_{n,k,\mathbf{v}}) \\ &= F(q-1) + (q-1)F(-1) \\ &= q \sum_{i=1}^{j-1} \binom{k}{i} (q-1)^i + \binom{k}{j} (q-1)^j - (q-1)k \\ &\quad + (q-1)\binom{k}{j}(-1)^j + (q-1)k. \end{aligned}$$

Hence, b) follows. ■

Let

$$g_{q,k,2}(z) = \frac{1}{q}\left\{ (1+(q-1)z)^{k+1} + (q-1)(1-z)^{k+1} \right\}.$$

Theorem 4: If C is an $[n, k, d; q]$ code and $d \geq 2$, then

$$A_C(z) \leq g_{q,k,2}(z) \tag{9}$$

for all $z \in [0, 1]$, with equality if and only if $d = 2$ and C is equivalent to $E_{n,k,\mathbf{v}}$ for some vector $\mathbf{v} \in F_q^k$ of full support.

Proof: Without loss of generality, we may assume that C is generated by $G = [I_k|Q]$ where the rows of Q are $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ (and where $\mathbf{v}_i \neq \mathbf{0}$ for $1 \leq i \leq k$ since the minimum distance is at least 2). As noted in the proof of Theorem 3, for any $\mathbf{x} \in F_q^k$, the codeword $\mathbf{x}G = (\mathbf{x}|\mathbf{x}Q)$ has weight $w(\mathbf{x}G) = w(\mathbf{x}) + w(\mathbf{x}Q)$. Hence

$$\sum_{i=2}^{j} A_i(C) = S_1 + S_2 \tag{10}$$

where

$$\begin{aligned} S_1 &= |\{\mathbf{x} \mid \mathbf{x} \neq \mathbf{0}, w(\mathbf{x}) \leq j-1, w(\mathbf{x}Q) + w(\mathbf{x}) \leq j\}| \\ &\leq |\{\mathbf{x} \mid \mathbf{x} \neq \mathbf{0}, w(\mathbf{x}) \leq j-1\}| \\ &= \sum_{i=1}^{j-1} \binom{k}{i} (q-1)^i \end{aligned}$$

and

$$S_2 = |\{\mathbf{x} \mid w(\mathbf{x}) = j, \mathbf{x}Q = \mathbf{0}\}|.$$

To evaluate S_2, we first choose j positions out of k, the number of choices is $\binom{k}{j}$. Without loss of generality we can assume that $\mathbf{x} = (x_1, x_2, \ldots, x_k)$, where $x_1, x_2, \ldots, x_j$ are non-zero and $x_{j+1} = \ldots = x_k = 0$. Then, we have

$$\begin{cases} x_1, x_2, \ldots, x_j \neq 0 \\ x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_j\mathbf{v}_j = \mathbf{0}. \end{cases} \tag{11}$$

Let r be the rank of the matrix with rows $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_j$.

If $r = 1$, then for $1 \leq i \leq j$, $\mathbf{v}_i = t_i\mathbf{v}_j$ for some $t_i \in F_q^*$. Denote by n_j the number of solutions of (11). For arbitrary nonzero elements $x_1, x_2, \ldots, x_{j-1}$:

1) if $x_1t_1 + x_2t_2 + \cdots + x_{j-1}t_{j-1} = 0$, then $(x_1, x_2, \ldots, x_{j-1})$ contributes 1 to n_{j-1}.
2) if $x_1t_1 + x_2t_2 + \cdots + x_{j-1}t_{j-1} \neq 0$, then

$$x_j = -x_1t_1 - x_2t_2 - \cdots - x_{j-1}t_{j-1}$$

and $(x_1, x_2, \ldots, x_{j-1}, x_j)$ contributes 1 to n_j.

Therefore, we have $n_{j-1} + n_j = (q-1)^{j-1}$. This recurrence relation and the first term $n_1 = 0$ imply that

$$n_j = \frac{1}{q}\left((q-1)^j + (-1)^j(q-1)\right). \tag{12}$$

If $r \geq 2$, then we may assume that the vectors

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$$

are linearly independent. For $j \geq r+1$, for any fixed non-zero elements $x_{r+1}, \ldots, x_j$, the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \cdots + x_r\mathbf{v}_r = -x_{r+1}\mathbf{v}_{r+1} - \cdots - x_j\mathbf{v}_j$$

has at most one solution. Therefore, the number of solutions of (11) is at most $(q-1)^{j-r}$ which is less than the expression for n_j given in (12), except when $r = 2$, $q = 2$, and j is odd, and

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_j = \mathbf{0}.$$

In this exceptional case, $n_j = 0 < 1 = (q-1)^{j-2}$ and at least one of $\mathbf{v}_i$ has Hamming weight at least 2 (since an odd number of binary vectors of weight 1 can not have sum $\mathbf{0}$). We may assume $w(\mathbf{v}_j) \geq 2$. Choose $\mathbf{x} = (1, 1, \ldots, 1, 0)$. Then, $w(\mathbf{x}) = j-1$ and

$$\mathbf{x}Q = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_{j-1} = \mathbf{v}_j.$$

Hence

$$w(\mathbf{x}G) = w(\mathbf{x}) + w(\mathbf{v}_j) \geq j - 1 + 2 = j + 1.$$

Therefore, in the exceptional case

$$S_1 < \sum_{i=1}^{j-1} \binom{k}{i} (q-1)^i.$$

In total, by (10), we obtain

$$\sum_{i=2}^{j} A_i(C) \le \sum_{i=1}^{j-1} \binom{k}{i}(q-1)^i + \frac{1}{q}\binom{k}{j}\left((q-1)^j + (-1)^j(q-1)\right) = \sum_{i=2}^{j} A_i(E_{n,k,\mathbf{v}}) \tag{13}$$

for $j \ge 2$ by (8).

By Lemma 1, we get that $A_C(z)$ takes the maximal value for any $z \in (0,1)$ if and only if C is (equivalent to) $E_{n,k,\mathbf{v}}$. ■

III. Upper Bound on $A_C(z)$ for Linear Codes C of Minimum Distance at Least 3

For $d \ge 3$, we get an improvement of Theorem 4. Let

$$g_{q,k,d}(z) = 1 + z^d \sum_{i=1}^{d-1} \binom{k}{i}(q-1)^i + \sum_{i=d}^{k} \binom{k}{i}\left((q-1)^i z + (q-1)^{i-d+1}(1-z)\right)z^i. \tag{14}$$

Theorem 5: If C is an $[n,k,d;q]$ code with $k \ge d \ge 3$, then $A_C(z) \le g_{q,k,d}(z)$ for all $z \in [0,1]$.

Proof: We use the notations and results in the proof of Theorem 4. Note that any collection of $d-1$ rows of Q are linearly independent. Indeed, if there exists $d-1$ linearly dependent rows, say, w.l.o.g, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{d-1}$, then there exists a non-zero vector $\mathbf{y} = (y_1, \ldots, y_{d-1}, 0, \ldots, 0)$ such that

$$y_1\mathbf{v}_1 + \cdots + y_{d-1}\mathbf{v}_{d-1} = \mathbf{0}$$

and then $w(\mathbf{y}Q) = w(\mathbf{y}) \le d-1$ which is a contradiction. Now, consider some vector $\mathbf{x} = (x_1, \ldots, x_k)$.

1) If $0 < w(\mathbf{x}) \le d-1$, then $w(\mathbf{x}G) \ge d$.
2) If $w(\mathbf{x}) = j \ge d$, then we may assume that $x_1, \ldots, x_j$ are non-zero and $x_{j+1} = \cdots = x_k = 0$. Then, the rank of $\mathbf{v}_1, \ldots, \mathbf{v}_j$ is at least $d-1$. Therefore, (11) has at most $(q-1)^{j-d+1}$ solutions. For the remaining nonzero $x_1, \ldots, x_j$ which is not a solution of (11), we get $\mathbf{x}Q = \mathbf{0}$ and $w(\mathbf{x}G) \ge j+1$.

Hence, we obtain

$$\begin{aligned} A_C(z) &\le 1 + z^d \sum_{i=1}^{d-1} \binom{k}{i}(q-1)^i + \sum_{i=d}^{k} \binom{k}{i}(q-1)^{i-d+1} z^i + \sum_{i=d}^{k} \binom{k}{i}\left((q-1)^i - (q-1)^{i-d+1}\right)z^{i+1} \\ &= 1 + z^d \sum_{i=1}^{d-1} \binom{k}{i}(q-1)^i + \sum_{i=d}^{k} \binom{k}{i}\left((q-1)^i z + (q-1)^{i-d+1}(1-z)\right)z^i. \end{aligned}$$

For $d \ge 3$, it appears to be quite complicated to get sharp upper bounds on $A_C(z)$.

For $d = 3$, Theorem 5 gives the upper bound

$$A_C(z) \le 1 + z^3\left\{k(q-1) + \binom{k}{2}(q-1)^2\right\} + \sum_{i=3}^{k} \binom{k}{i}(q-1)^i z^{i+1} + (1-z)\sum_{i=3}^{k} \binom{k}{i}(q-1)^{i-2} z^i. \tag{15}$$

If $[I_k|Q]$ generates an $[n,k,3;q]$ code C, then the rows of Q must have weight at least 2 and be non-proportional, that is, $\mathbf{x}$ and $\alpha\mathbf{x}$ can not both be rows of Q. Hence, the size σ of the support of Q must satisfy

$$(q-1) + (q-1)^2 + \cdots + (q-1)^{\sigma-1} \ge k.$$

In particular, $\sigma = 2$ is possible only if $k \le q-1$. Consider this case. The dual code $C^\perp$ is equivalent to the code generated by $G' = [I_{n-k}|Q^T]$, where Q^T has only two non-zero rows, an all-one row and $(a_1, a_2, \ldots, a_k)$ where the a_i are distinct and non-zero. The weight distribution of $C^\perp$ is easily seen to be

$$(1+(q-1)z)^{n-k-2} \cdot \left\{1 + (q-1)(k+2)z^{k+1} + (q-1)(q-1-k)z^{k+2}\right\}.$$

Using the MacWilliams theorem, we can get the weight distribution of C

$$\begin{aligned} q^{n-k}A_C(z) &= \big(1+(q-1)z\big)^n A_{C_\perp}\Big(\frac{1-z}{1+(q-1)z}\Big) \\ &= \big(1+(q-1)z\big)^n\Big(1+(q-1)\frac{1-z}{1+(q-1)z}\Big)^{n-k-2} \\ &\quad\cdot\Big\{1+(q-1)(k+2)\Big(\frac{1-z}{1+(q-1)z}\Big)^{k+1} + (q-1)(q-1-k)\Big(\frac{1-z}{1+(q-1)z}\Big)^{k+2}\Big\} \\ &= q^{n-k-2}\Big\{\big(1+(q-1)z\big)^{k+2} + (q-1)(k+2)\big(1+(q-1)z\big)(1-z)^{k+1} + (q-1)(q-1-k)(1-z)^{k+2}\Big\} \end{aligned}$$

and so

$$q^2 A_C(z) = \big(1+(q-1)z\big)^{k+2} + (q-1)(k+2)(1-z)^{k+1}\big(1+(q-1)z\big) + (q-1)(q-1-k)(1-z)^{k+2}.$$

Example 1: For $k = 3$, the upper bound is

$$1 + \big(3q^2 - 2q - 1\big)z^3 + \big(q^3 - 3q^2 + 2q\big)z^4$$

and

$$A_C(z) = 1 + 10\,(q-1)z^3 + 5\,(q-1)(q-3)z^4 + (q-1)(q^2-4q+6)z^5.$$

Hence, the upper bound is never sharp for $k = 3$. ■

IV. Comparing the Various Bounds

If $k \geq d \geq 3$ we now have three bounds on $A_C(z)$ for an $[n,k,d;q]$ code, $f_{q,k,d}(z)$ (see Theorem 3), $g_{q,k,2}(z)$ (see Theorem 4), and $g_{q,k,d}(z)$ (see Theorem 5).

We will now give a comparison of the various bounds.

Theorem 6: For $q \geq 2$ and $k \geq d \geq 2$, we have

$$f_{q,k,d-1}(z) - f_{q,k,d}(z) = z^{d-1}(1-z)\sum_{i=1}^{d-1}\binom{k}{i}(q-1)^i.$$

In particular, $f_{q,k,d-1}(z) > f_{q,k,d}(z)$ for all $z \in (0,1)$.

Proof:

$$\begin{aligned}
&f_{q,k,d-1}(z) - f_{q,k,d}(z)\\
&= 1 + \sum_{i=1}^{d-1}\binom{k}{i}(q-1)^i z^{d-1} + \sum_{i=d}^{k}\binom{k}{i}(q-1)^i z^i\\
&\quad - \Big\{1 + \sum_{i=1}^{d}\binom{k}{i}(q-1)^i z^d + \sum_{i=d+1}^{k}\binom{k}{i}(q-1)^i z^i\Big\}\\
&= \big(z^{d-1} - z^d\big)\sum_{i=1}^{d-1}\binom{k}{i}(q-1)^i.
\end{aligned}$$

■

Theorem 7: For $q \geq 2$ and $k \geq d \geq 4$, we have

$$\begin{aligned}
&g_{q,k,d-1}(z) - g_{q,k,d}(z)\\
&\quad = z^{d-1}(1-z)\sum_{i=1}^{d-2}\binom{k}{i}(q-1)^i\\
&\qquad + \binom{k}{d-1}(q-1)z^{d-1}(1-z)\\
&\qquad + (q-2)\sum_{i=d}^{k}\binom{k}{i}(q-1)^{i-d+1}z^i(1-z)\\
&\quad > 0
\end{aligned}$$

for all $z \in (0,1)$.

Proof:

$$\begin{aligned}
&g_{q,k,d-1}(z) - g_{q,k,d}(z)\\
&= 1 + z^{d-1}\sum_{i=1}^{d-2}\binom{k}{i}(q-1)^i + \binom{k}{d-1}(q-1)^{d-1}z^d\\
&\quad + \sum_{i=d}^{k}\binom{k}{i}(q-1)^i z^{i+1} + \binom{k}{d-1}(q-1)z^{d-1}(1-z)\\
&\quad + \sum_{i=d}^{k}\binom{k}{i}(q-1)^{i-d+2}z^i(1-z)\\
&\quad - 1 - z^d\sum_{i=1}^{d-2}\binom{k}{i}(q-1)^i - \binom{k}{d-1}(q-1)^{d-1}z^d\\
&\quad - \sum_{i=d}^{k}\binom{k}{i}(q-1)^i z^{i+1} - \sum_{i=d}^{k}\binom{k}{i}(q-1)^{i-d+1}z^i(1-z)\\
&= \big(z^{d-1} - z^d\big)\sum_{i=1}^{d-2}\binom{k}{i}(q-1)^i\\
&\quad + \binom{k}{d-1}(q-1)z^{d-1}(1-z)\\
&\quad + (q-2)\sum_{i=d}^{k}\binom{k}{i}(q-1)^{i-d+1}z^i(1-z).
\end{aligned}$$

■

Theorem 8: For $q \geq 2$ and $k \geq 2$, we have

$$\begin{aligned}
&f_{q,k,2}(z) - g_{q,k,2}(z) =\\
&\qquad \frac{q-1}{q}\sum_{j=2}^{k}\binom{k}{j}\big\{(q-1)^j - (-1)^j\big\}z^j(1-z).
\end{aligned}$$

In particular, $f_{2,2,2}(z) = g_{2,2,2}(z)$ for all $z \in [0,1]$, and $f_{q,k,2}(z) > g_{q,k,2}(z)$ for all $z \in (0,1)$ otherwise.

Proof:

$$\begin{aligned}
&f_{q,k,2}(z) - g_{q,k,2}(z)\\
&= \big(1 + (q-1)z\big)^k + k(q-1)(z^2 - z)\\
&\quad - \frac{1}{q}\big(1 + (q-1)z\big)^{k+1} - \frac{q-1}{q}(1-z)^{k+1}\\
&= \frac{1}{q}\big(1 + (q-1)z\big)^k\big\{q - 1 - (q-1)z\big\}\\
&\quad - k(q-1)z(1-z) - \frac{q-1}{q}(1-z)^{k+1}\\
&= \frac{q-1}{q}(1-z)\Big\{\big(1 + (q-1)z\big)^k - (1-z)^k - kqz\Big\}\\
&= \frac{q-1}{q}(1-z)\left\{\sum_{j=0}^{k}\binom{k}{j}\big\{(q-1)^j - (-1)^j\big\}z^j - kqz\right\}\\
&= \frac{q-1}{q}(1-z)\left\{\sum_{j=2}^{k}\binom{k}{j}\big\{(q-1)^j - (-1)^j\big\}z^j\right\}.
\end{aligned}$$

In particular, if $q > 2$, then $(q-1)^j - (-1)^j > 0$ for all $j \geq 2$. If $q = 2$, $(q-1)^j - (-1)^j > 0$ if j is odd. Hence, $f_{q,k,2}(z) > g_{q,k,2}(z)$, except when $k = q = 2$. ■

Theorem 9: For $q \geq 2$ and $k \geq d \geq 3$, we have

$$f_{q,k,d}(z) - g_{q,k,d}(z) = \sum_{i=d}^{k}\binom{k}{i}\big\{(q-1)^i - (q-1)^{i-d+1}\big\}z^i(1-z).$$

In particular, $f_{2,k,d}(z) = g_{2,k,d}(z)$ for all $z \in [0,1]$ and $f_{q,k,d}(z) > g_{q,k,d}(z)$ for all $z \in (0,1)$ when $q \geq 3$.

Proof:

$$\begin{aligned}
&f_{q,k,d}(z) - g_{q,k,d}(z)\\
&= \sum_{i=d}^{k}\binom{k}{i}(q-1)^i z^i\\
&\quad - \sum_{i=d}^{k}\binom{k}{i}(q-1)^i z^{i+1} - \sum_{i=d}^{k}\binom{k}{i}(q-1)^{i-d+1}z^i(1-z)\\
&= \sum_{i=d}^{k}\binom{k}{i}z^i\big\{(q-1)^i - (q-1)^i z - (q-1)^{i-d+1}(1-z)\big\}\\
&= \sum_{i=d}^{k}\binom{k}{i}z^i(1-z)\big\{(q-1)^i - (q-1)^{i-d+1}\big\}.
\end{aligned}$$

In particular, if $q = 2$, then $(q-1)^i - (q-1)^{i-d+1} = 0$ for all i. If $q > 2$, then $(q-1)^i - (q-1)^{i-d+1} > 0$ for all $i \geq d$. ■

For small values of z, which are the most important for applications, Theorems 6–9 show that for $d \geq 3$, we have

$$f_{q,k,d-1}(z) \geq g_{q,k,d-1}(z) > f_{q,k,d}(z) > g_{q,k,d}(z).$$

For larger values of z, we may have $g_{q,k,d-1}(z) < f_{q,k,d}(z)$.

As noted previously, if $A_C(z) \le f(z)$ for all $z \in [0,1]$, then

$$A_{C^\perp}(z) \le h(z) = \frac{1}{q^k}\,(1+(q-1)z)^n\, f\left(\frac{1-z}{1+(q-1)z}\right)$$

for all $z \in [0,1]$. However, we note that if $f(z)$ is an upper bound on $A_C(z)$ for z close to 1, then $h(z)$ is an upper bound on $A_{C^\perp}(z)$ for small z. It is, therefore, of interest to compare the bounds for z close to 1. We see that if $f(1) = g(1)$ and $f'(1) > g'(1)$, then $f(z) < g(z)$ for z close to 1. We will use this observation to prove a couple of theorems.

Theorem 10: If $q \ge 3$ and

$$k \ge 1 + \frac{(d-2)\ln(q-1) + \ln 2}{\ln(q/2)} \tag{16}$$

or $q = 2$ and $k \ge c(d-1)$, where c is determined by

$$\frac{(c-2)^2}{4(c+1)} = \frac{\ln(d-1)}{d-1} \tag{17}$$

then $g_{q,k,2}(z) < f_{q,k,d}(z)$ for z close to 1.

Proof: We have $f_{q,k,d_1}(1) = g_{q,k,d_2}(1) = q^k$ for all q, k, d_1, d_2. We will show that $g'_{q,k,2}(1) > f'_{q,k,d}(1)$ under condition (16), respectively (17). We have

$$g'_{q,k,2}(z) = \frac{1}{q}\Big\{(k+1)(q-1)\,(1+(q-1)z)^k - (k+1)(q-1)(1-z)^k\Big\}$$

and so $g'_{q,k,2}(1) = (k+1)(q-1)q^{k-1}$. Further

$$f'_{q,k,d}(z) = k(q-1)(1+(q-1)z)^{k-1} + \sum_{i=1}^{d}\binom{k}{i}(q-1)^i\big(dz^{d-1} - iz^{i-1}\big).$$

We note that $(q-1)^i(d-i) \le (q-1)^{d-1}$ for $1 \le i \le d-1$ when $q \ge 3$. Hence

$$\begin{aligned} f'_{q,k,d}(1) &= k(q-1)q^{k-1} + \sum_{i=1}^{d}\binom{k}{i}(q-1)^i(d-i) \\ &\le k(q-1)q^{k-1} + (q-1)^{d-1}\sum_{i=1}^{d}\binom{k}{i} \\ &< k(q-1)q^{k-1} + (q-1)^{d-1}2^k. \end{aligned}$$

Therefore, for $q \ge 3$, we have

$$\begin{aligned} g'_{q,k,2}(1) - f'_{q,k,d}(1) &> (q-1)q^{k-1} - (q-1)^{d-1}2^k \\ &= (q-1)2^{k-1}\left\{\left(\frac{q}{2}\right)^{k-1} - 2(q-1)^{d-2}\right\} \\ &\ge 0 \end{aligned}$$

when (16) is satisfied. For $q = 2$, we have

$$\begin{aligned} f'_{2,k,d}(1) &= k\,2^{k-1} + \sum_{i=1}^{d}\binom{k}{i}(d-i) \\ &< k\,2^{k-1} + (d-1)\sum_{i=1}^{d}\binom{k}{i}. \end{aligned}$$

It is well known that if $k \ge 2m$ (see, e.g., [10, pp. 82–83]), then

$$\sum_{i=1}^{d}\binom{k}{i} < 2^{2m-1}\exp\Big(\frac{(m-d-1)^2}{2m-d-1}\Big). \tag{18}$$

For $k \ge 2m \ge c(d-1)$, the right-hand side of (18) is less than or equal to $2^{k-1}/(d-1)$ by the choice of c given in (17). Hence, $g'_{2,k,2}(1) - f'_{2,k,d}(1) > 0$. ∎

The bound $g_{q,k,2}(z)$ is not a special case of $g_{q,k,d}(z)$ given by (14). It turns out that $g_{q,k,2}(z)$ sometimes is better than $g_{q,k,3}(z)$. For z close to 1, we have the following situation.

Theorem 11: We have then $g_{2,k,2}(z) < g_{2,k,3}(z)$ for z close to 1 when $k \ge 6$. For $q \ge 3$ and $k \ge 1$, $g_{q,k,2}(z) > g_{q,k,3}(z)$ for z close to 1.

Proof: As shows in the proof of the previous theorem, $g'_{q,k,2}(1) = (k+1)(q-1)q^{k-1}$. Some calculations show that

$$\begin{aligned} g'_{q,k,3}(1) = q^{k-1}\Big(k(q-1) + q - \frac{q}{(q-1)^2}\Big) + \frac{1}{(q-1)^2} \\ - 1 + k(q-1) + \frac{k}{q-1} + \binom{k}{2}. \end{aligned}$$

Hence

$$\begin{aligned} g'_{q,k,3}(1) - g'_{q,k,2}(1) = q^{k-1}\Big(1 - \frac{q}{(q-1)^2}\Big) + \frac{1}{(q-1)^2} \\ + \big(k(q-1) - 1\big) + \frac{k}{q-1} + \binom{k}{2}. \end{aligned}$$

For $q \ge 3$, all the terms are positive. For $q = 2$, we have $1 - \frac{q}{(q-1)^2} = -1$. Hence, we see that $g'_{2,k,3}(1) - g'_{2,k,2}(1) < 0$ for k sufficiently large. Computation shows that $g'_{2,k,3}(1) - g'_{2,k,2}(1) < 0$ for $k \ge 6$. ∎

V. Upper Bounds on $A_C(z)$ for Linear Codes C of Full Support

For a code C of length n, the support $\chi(C)$ is the set of positions i such that $c_i \ne 0$ for some codeword $(c_1, c_2, \ldots, c_n) \in C$. The code has *full support* if $|\chi(C)| = n$, that is, for any position there is a codeword that is nonzero in this position. For example, the code generated by $[I_k|0_{k\times(n-k)}]$ has support of size k.

In practical applications, one usually uses codes with full support. We expect to find a sharper upper bound on $P_{\rm ue}(C,p)$ for codes of full support. In this section we find the best possible upper bound on $A_C(z)$ for linear codes C of full support. This is given in Theorem 12. First, we need a lemma.

Lemma 4: An $[n,k;q]$ code C has full support if and only if $C^\perp$ has minimum distance at least 2.

Proof: The result follows from the observation that if i is not in the support, then the unit vector $\mathbf{e}_i$ is contained in $C^\perp$ and vice versa. ∎

The weight distribution of $D_{n,k,\mathbf{v}}$ was given by Lemma 2. Let

$$h_{q,k,1}(z) = \big(1+(q-1)z\big)^{k-1}(1+(q-1)z^{n-k+1}).$$

From Lemma 4 and Theorem 9, we immediately get the following theorem.

Theorem 12: If C is an $[n,k;q]$ code of full support, then $A_C(z) \le h_{q,k,1}(z)$ for all $z \in [0,1]$, with equality if and only if C is equivalent to $D_{n,k,\mathbf{v}}$ for some vector $\mathbf{v}$ of full support.

This bound is tighter than the bound $f_{q,k,1}(z)$ in Corollary 1. The improvement of Theorem 12 over Corollary 1, for $z \in [0,1]$, is

$$f_{q,k,1}(z) - h_{q,k,1}(z) = (q-1)z(1-z^{n-k})\big(1+(q-1)z\big)^{k-1}.$$

VI. On Codes of Full Support and Minimum Distance Two

As shown previously in Theorem 12, a worst case $[n,k;q]$ code of full support is equivalent to a code with generator matrix $D_{n,k,\mathbf{v}}$, where $\mathbf{v}$ has full support. Also, a worst case $[n,k,2;q]$ code is equivalent to a code with generator matrix $E_{n,k,\mathbf{v}}$, where $\mathbf{v}$ has full support.

A natural question is then: what is a worst case $[n,k,2;q]$ code of full support. The answer seems to be the code generated by the matrix

$$G_1 = \begin{bmatrix} \mathbf{0} & \mathbf{v} \\ G_2 & O_{(k-1)\times(n-k)} \end{bmatrix}$$

where $\mathbf{v}$ is a vector of length $n-k$ and full support, and G_2 is the generator matrix for a worst case $[k,k-1,2;q]$ code. If C is the code generated by G_1 and D is the code generated by G_2, we get, by Theorem 4,

$$\begin{aligned} A_C(z) = h_{q,k,2}(z) &= (1+(q-1)z^{n-k})A_D(z) \\ &= \frac{1}{q}(1+(q-1)z^{n-k}) \\ &\quad \cdot \Big\{(1+(q-1)z)^k + (q-1)(1-z)^k\Big\}. \end{aligned}$$

We conjecture that $A_C(z) \le h_{q,k,2}(z)$ for all $[n,k,2;q]$ codes C of full support, but we do not have a general proof. A proof for $k=2$ goes as follows. Consider a generator matrix G for an $[n,2,2;q]$ code C of full support. Since equivalent codes have the same weight distribution, we may assume that the first non-zero element in each column is 1. Then the possible columns are $(0,1)$ and $(1,a)$, $a \in F_q$. If $(0,1)$ appears y times and $(1,a)$ appears x_a times, then

$$y + \sum_{a \in F_q} x_a = n. \tag{19}$$

Moreover

$$y \ge 1 \text{ and } x_0 \ge 1. \tag{20}$$

The weight distribution of C is

$$A_C(z) = 1 + (q-1)z^{n-y} + (q-1)\sum_{a \in F_q} z^{n-x_a}.$$

Because of the symmetry of this expression, we may assume that

$$x_0 \ge x_a \quad \text{for all } a \in F_q \text{ and } x_0 \ge y. \tag{21}$$

Finally, since the minimum distance of the code is 2, we have

$$x_0 \le n-2. \tag{22}$$

Proposition 1: For $n \ge 4$, a worst case $[n,2,2;q]$ code of full support is obtained for $y=2$ and $x_0 = n-2$, for which we get

$$\begin{aligned} A_C(z) &= 1 + (q-1)z^2 + (q-1)z^{n-2} + (q-1)^2 z^n \\ &= (1+(q-1)z^2)(1+(q-1)z^{n-2}) = h_{q,2,2}(z). \end{aligned}$$

Proof: Let C be a worst case code with column count y and x_a, $a \in F_q$ satisfying (19)–(22). If $y > 2$, we get a new code C' by decreasing the number of $(0,1)$ columns by $y-2$ and increasing the number of $(1,0)$ columns by the same amount (that is $y \to 2$ and $x_0 \to x_0 + y - 2$). Then

$$\begin{aligned} A_{C'}(z) - A(z) &= z^{n-2} - z^{n-y} + z^{n-x_0-y+2} - z^{n-x_0} \\ &= z^{n-x_0-y+2}(1-z^{x_0-2})(1-z^{y-2}) > 0 \end{aligned}$$

for all $z \in (0,1)$ since $x_0 \ge y > 2$. Hence, C is not a worst case code, a contradiction. Hence, $y=2$. Similarly, we show that $x_a = 0$ for all $a \ne 0$ and $x_0 = n-2$. Hence, a generator matrix is

$$G = \begin{bmatrix} 0 & 0 & 1 & 1 & \dots & 1 \\ 1 & 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

■

VII. Examples

Example 2: As a very simple example, we consider the case $k=2$. For each code we give the weight distribution and a pair $V = (\mathbf{v}_1, \mathbf{v}_2)$ such that

$$\begin{bmatrix} 1 & 0 & \mathbf{v}_1 \\ 0 & 1 & \mathbf{v}_2 \end{bmatrix}$$

is a generator matrix for the code.

1) For all $[n,2;q]$ codes

$$\begin{aligned} A_C(z) \le A_{C_{n,1}}(z) &= f_{q,2,1}(z) \\ &= 1 + 2(q-1)z + (q-1)^2 z^2 \end{aligned}$$

where

$$V = V_{n,1} = ((0000\dots0),(0000\dots0)).$$

2) For all $[n,2,2;q]$ codes

$$\begin{aligned} A_C(z) \le A_{C_{n,2}}(z) &= g_{q,2,2}(z) \\ &= 1 + 3(q-1)z^2 + (q-1)(q-2)z^3 \end{aligned}$$

where

$$V = V_{n,2} = ((1000\dots0),(1000\dots0)).$$

We note that $f_{q,2,2}(z) = 1 + (q^2-1)z^2$.

3) For all $[n,2;q]$ codes of full support

$$\begin{aligned} A_C(z) \le A_{C_{n,3}}(z) &= h_{q,2,1}(z) \\ &= 1 + (q-1)z + (q-1)z^{n-1} + (q-1)^2 z^n \end{aligned}$$

where

$$V = V_{n,3} = ((0000\ldots0),(1111\ldots1)).$$

4) For all $[n,2,2;q]$ codes of full support

$$\begin{aligned} A_C(z) \le\ & A_{C_{n,4}}(z) = h_{q,2,2}(z) \\ & = 1 + (q-1)z^2 + (q-1)z^{n-2} + (q-1)^2 z^n \end{aligned}$$

where

$$V = V_{n,4} = ((1000\ldots0),(0111\ldots1)).$$

Clearly

$$h_{q,2,2}(z) \le g_{q,2,2}(z) \le f_{q,2,1}(z)$$

and

$$h_{q,2,2}(z) \le h_{q,2,1}(z) \le f_{q,2,1}(z)$$

for all $z \in [0,1]$. It is easy to see that $g_{q,2,2}(z) < h_{q,2,1}(z)$ for values of z close to 0 and $g_{q,2,2}(z) > h_{q,2,1}(z)$ for values of z close to 1. A little calculus actually shows that there exists a $z_n \in (0,1)$ such that $g_{q,2,2}(z) < h_{q,2,1}(z)$ for $z \in (0,z_n)$ and $g_{q,2,2}(z) > h_{q,2,1}(z)$ for $z \in (z_n,1)$.

Proposition 2: We have $z_n \sim \omega + \theta\,\omega^{n-2}$, when $n \to \infty$, where

$$\omega = \frac{2}{3+\sqrt{4q+1}} \text{ and } \theta = \frac{1+(q-1)\omega}{3+2\,(q-2)\omega}.$$

Proof: We sketch that proof. By definition

$$\begin{aligned} & 1 + 3\,(q-1)z_n^2 + (q-1)(q-2)z_n^3 \\ & \quad = 1 + (q-1)z_n + (q-1)z_n^{n-1} + (q-1)^2 z_n^n \end{aligned}$$

and so

$$3\,z_n + (q-2)z_n^2 = 1 + z_n^{n-2} + (q-1)z_n^{n-1}. \tag{23}$$

Let $\omega = \lim_{n\to\infty} z_n$. Since $0 < z_n < 1$, we get

$$\lim_{n\to\infty} z_n^{n-2} = \lim_{n\to\infty} z_n^{n-1} = 0.$$

Hence, (23) implies that

$$3\,\omega + (q-2)\omega^2 = 1. \tag{24}$$

For $q=2$, we get $\omega = 1/3$. For $q > 2$, the equation

$$3\,x + (q-2)x^2 = 1$$

has two roots:

$$x_1 = \frac{-3+\sqrt{4q+1}}{2(q-2)} = \frac{2}{3+\sqrt{4q+1}}$$
$$x_2 = \frac{-3-\sqrt{4q+1}}{2(q-2)}.$$

Since $\omega > 0$ and $x_2 < 0$, we must have $\omega = x_1 = \frac{2}{3+\sqrt{4q+1}}$. We note that this gives the correct value, 1/3, also for $q=2$.

We can conclude that $z_n \approx \omega$. Let $z_n = \omega + y_n$. Substituting this in (23) and simplifying, we get

$$y_n\big(3+2(q-2)\omega+(q-2)y_n\big) = (\omega+y_n)^{n-2} + (q-1)(\omega+y_n)^{n-1}$$

and so

$$\frac{y_n}{\omega^{n-2}} = \Big(\frac{\omega+y_n}{\omega}\Big)^{n-2} \frac{1+(q-1)(\omega+y_n)}{3+2(q-2)\omega+(q-2)y_n}.$$

Hence

$$\frac{y_n}{\omega^{n-2}} \to \frac{1+(q-1)\omega}{3+2(q-2)\omega} = \theta \text{ when } n \to \infty.$$

■

VIII. Summary

We have first given an upper bound on the weight distribution function for codes of minimum distance at least 2. We have shown that the bound is best possible for codes of minimum distance equal to 2, and we characterized the codes meeting the bound.

Next, we have given an improved bound for codes of minimum distance at least 3.

Next, we noted that a code has full support if and only if the dual code has minimum distance 2, and we used this fact to determine a best possible upper bound on the weight distribution function for linear codes of full support. We also characterized the codes meeting this bound.

We discussed the weight distribution function for linear codes that both have full support and minimum distance 2. We have shown a best possible upper bound for such codes of dimension $k=2$ and we conjecture a bound for general k.

References

[1] R. J. McEliece, *The Theory of Information and Coding*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[2] H. Jin and R. J. McEliece, "Coding theorems for turbo code ensembles," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1451–1461, Jun. 2002.

[3] T. Kløve, *Codes for Error Detection*. Singapore: World Scientific, 2007.

[4] J. E. Levy, "A weight distribution bound for linear codes," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 3, pp. 487–490, May 1968.

[5] J. MacWilliams, "A theorem on the distribution of weights in a systematic code," *Bell Syst. Tech. J.*, vol. 42, pp. 79–94, Jan. 1963.

[6] V. I. Korzhik, "Bounds on undetected error probability and optimum group codes in a channel with feedback," *Telecommun. Radio Eng.*, vol. 20, no. 1, pt. 2, pp. 87–92, Jan. 1965.

[7] V. I. Levenshtein, "Bounds on the probability of undetected error," *Prob. Inform. Transmission*, vol. 13, no. 1, pp. 1–12, 1978.

[8] T. Kasami, T. Kløve, and S. Lin, "Linear block codes for error detection," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 1, pp. 131–136, Jan. 1983.

[9] T. Kløve and V. I. Korzhik, *Error Detecting Codes, General Theory and Their Application in Feedback Communication Systems*. Boston, MA: Kluwer, 1995.

[10] L. Lovasz, J. Pelikan, and K. L. Vesztergombi, *Discrete Mathematics: Elementary and Beyond*. New York: Springer-Verlag, 2003.

Torleiv Kløve (M'89–SM'91–F'03) was born in Voss, Norway, 1943. He received the Cand. Mag., Cand. Real., and Dr. Philos. degrees from the University of Bergen, Norway, in 1966, 1967, and 1971, respectively.

He has been with the University of Bergen since 1971, first as Senior Lecturer in Mathematics, and, since 1982, Professor of Informatics. During the academic years 1975–1976, 1981–1982, and 1990–1991 he spent sabbaticals at the University of Hawaii at Manoa. During the academic year 2001–2002, he was a Visiting Professor at The Chinese University of Hong Kong, and during the fall semester 2002, he was Visiting Professor at Hong Kong University of Science and Technology. During the academic year 2008–2009 he spent a sabbatical at the University of California, Santa Cruz, and the academic year 2011–2012 he spent a sabbatical at the Università Politecnica delle Marche, Ancona, Italy.

His research interests include coding theory, number theory, and combinatorics. He is co-author of the book *Error Detecting Codes* (Kluwer 1995) and the author of the book *Codes for Error Detection* (World Scientific 2007).

He was Chairman of the 1996 IEEE Information Theory Workshop, Longyearbyen, Norway, Associate Editor for Coding Theory for IEEE TRANSACTIONS ON INFORMATION THEORY (1996–1999), and member of the Board of Governors of the IEEE Information Theory Society (2001–2003).

Jinquan Luo was born in February 1980, Anhui, China. He received the B.S. degree from Zhejiang University, Hangzhou, China, in July 2001 and the Ph.D. degree from Tsinghua University, Beijing, China, in January 2007, both in mathematics.

He joined Yangzhou University, China from 2007 and later become a research fellow at Nanyang Technological University, Singapore from 2009. Currently, he servers as postdoctor at the Department of Informatics, University of Bergen, Norway. His major research interests are coding theory, cryptology, and number theory.