Designs, Codes and Cryptography, , 1–20 ()
© Kluwer Academic Publishers, Boston. Manufactured in The Netherlands.

On Generalized Hamming Weights for Galois Ring Linear Codes

ALEXEI ASHIKHMIN alexei@c3serve.c3.lanl.gov
Los Alamos National Laboratory
Group CIC-3, MS P990, Los Alamos, NM 87545

Abstract. The definition of generalized Hamming weights (GHW) for linear codes over Galois rings is discussed. The properties of GHW for Galois ring linear codes are stated. Upper and existence bounds for GHW of $\mathbf{Z}_4-$linear codes and a lower bound for GHW of the Kerdock code over $\mathbf{Z}_4$ are derived. GHW of some $\mathbf{Z}_4-$ linear codes are determined.

Keywords: Galois ring linear codes, generalized Hamming weights

1. Introduction

For any code D, $\chi(D)$, the *support of* D, is the set of positions where not all the codewords of D are zero, and $w_s(D)$, the *support weight* of D, is the weight of $\chi(D)$. For an $[n,k]$ code C and any r, where $1 \leq r \leq k$, *the r-th Hamming weight* is defined [7],[14] by

$$d_r(C) = \min\{w_s(D) : D \text{ is an } [n,r] \text{ subcode of } C\}.$$

In [1],[4], and [5] different generalizations of GHW for nonlinear case were suggested. In the present paper we consider the case of Galois ring (GR) linear codes.

Outline of the paper is following.

In Section 2, we discuss the definition and state some properties of GHW of GR linear codes. In Section 3, we show that definition of GHW from [4] determines performance of group codes in the wire tap channel of type II. In Section 4, we find the number of different type subcodes in a $\mathbf{Z}_4-$linear code and state a connection between their support sizes. In Section 5, we use results of Section 4 to get bounds for GHW of $\mathbf{Z}_4-$linear codes. In Section 6, using results of sections 4 and 5, we obtain a lower bound for GHW of the Kerdock code over $\mathbf{Z}_4$ and complete weight hierarchys of some short $\mathbf{Z}_4-$ linear codes. We also show that though the minimum Hamming weight of a GR linear code can not exceed the minimum Hamming weight of an optimal linear codes, heigher weights of a GR linear code do can exceed corresponding weights of an optimal linear code of the same length and dimension.

2. Definition and Basic Properties

Let R be a Galois ring, i.e., a finite commutative ring with identity e, whose set of zero divisors has the form pR for a certain prime p. After Nechaev's paper [10], Galois rings have become easy to understand for a reader with a standard

background in finite fields. In particular, using the definition given, one can prove [10] that $|R| = q^m, m \geq 1$, where $q = p^s$ for some $s \geq 1$, and the characteristic of R (the order of e in the group $(R, +)$) equals p^m. Since fixing the numbers p^m and q^m identifies R up to isomorphism, it may be also denoted as $GR(q^m, p^m)$. All ideals of R form the following chain:

$$\mathcal{R}_0 = R \supset \mathcal{R}_1 = pR \supset \mathcal{R}_2 = p^2R \supset \ldots \supset \mathcal{R}_{m-1} = p^{m-1}R \supset \mathcal{R}_m = p^mR = 0,$$

and $|\mathcal{R}_i| = q^{m-i}$. A quotient ring $R/\mathcal{R} = R/pR$, is a Galois field of order q

$$R/pR = GF(q). \tag{1}$$

The ring R is constructed as a degree s Galois extension of $\mathbf{Z}_{p^m}$ in much the same way as one constructs the finite field from $\mathbf{Z}_p$. Note also that Galois rings encompass finite fields and residue rings as boundary cases. Namely, $GF(p^s) = GR(p^s, p)$ and $\mathbf{Z}_{p^m} = GR(p^m, p^m)$.

Let C be a GR linear code of length n. Let D be any GR linear subcode of C. The support of D is defined exactly as for codes over a Galois field, i.e., the support of D is the set of not-all-zero symbol positions of D.

In [1], the following definition was proposed for *the r-th generalized Hamming weight* of C

$$d_r(C) = \min\{w_s(D) : D \text{ is GR linear subcode of } C, \log_q |D| = r\}.$$

One can see that this definition generalizes the definition of GHW for linear codes over Galois fields.

L. A. Bassalygo proposed another definition for support and GHW [4]. Let A be any code of length n over arbitrary alphabet and let B be any subcode of A. Define the *support* of B as follows

$$X(B) = \{i : \exists \mathbf{a} = \{a_1, a_2, \ldots, a_n\}, \mathbf{b} = \{b_1, b_2, \ldots, b_n\} \in B \text{ such that } a_i \neq b_i\}.$$

In other words, the support of B is the set of positions where not all the codewords of B have the same symbols. Define the following function

$$\mathcal{F}_A(M) = \min\{|X(B)| : B \text{ is subcode of } A \text{ and } |B| = M\}, 1 \leq M \leq |A|.$$

We will consider only those points where the function $\mathcal{F}_A$ changes its values, that is, the following two sets $\{\delta_1, \delta_2, \ldots, \delta_t\}$ and $\{M_1, M_2, \ldots, M_t\}, t \geq 1$, where δ_i and M_i are defined as follows

$$\delta_i = \mathcal{F}_A(M) > \delta_{i-1} \text{ and } \mathcal{F}_A(M-1) = \delta_{i-1};\ \mathcal{F}_A(M_i) = \delta_i \text{ and } \mathcal{F}_A(M_i+1) > \delta_i,$$

and $\delta_0 = 0, M_0 = 1$. It is clear that the function $\mathcal{F}_A(M)$ is completely defined by the set of pairs $(\delta_1, M_1), (\delta_2, M_2), \ldots$ and so on. These pairs are called *generalized Hamming weights*.

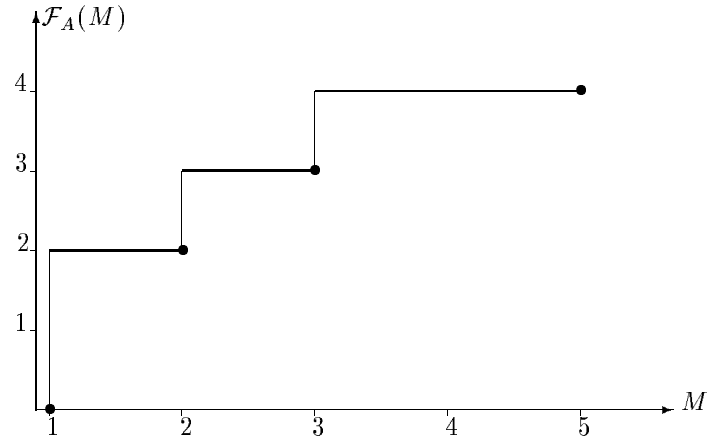
EXAMPLE: Consider the nonlinear code

$$A = \{(0,0,0,0),(0,0,1,1),(0,1,0,1),(1,1,0,0),(1,1,1,1)\}.$$

This code has the following values δ_i and M_i

$$\{\delta_1,\delta_2,\delta_3\} = \{2,3,4\} \text{ and } \{M_1,M_2,M_3\} = \{2,3,6\},$$

and the function $\mathcal{F}_A(M)$ can be depicted as follows (the function $\mathcal{F}_A(M)$ is defined only for integer values of M, and we draw it everywhere just for visualization) □



One can see that Bassalygo's definition works for both linear and nonlinear cases. In the case of a linear code over $GF(q)$, we have $M_i/M_{i-1} = q$ and therefore we can study only values δ_i. In the general case, we must consider pairs (δ_i, M_i).

In Corollary 1, we show that if C is a linear code over $GR(q^m,p^m)$, then $M_i/M_{i-1} = q^t, 1 \le t \le m$, i.e., the function $\mathcal{F}_C(M)$ can change its values only at $M = q^i$. Since $d_i(C) = \mathcal{F}_C(q^i)$, one can see that in the GR linear case there is one-to-one correspondence between the set $\{d_1,\ldots,d_{\log_q|C|}\}$ and the sets $\{\delta_1,\delta_2,\ldots,\delta_t\}$ and $\{M_1,M_2,\ldots,M_t\}$, i.e., knowledge of values d_i allows us to find values δ_j, M_j and vice versa.

Note that for GHW of nonlinear codes one more definition, nonequivalent to Bassalygo's one, was given in [5].

THEOREM 1 *Let C be a linear code of length n over $GR(q^m,p^m)$. Then*

$$1 \le d_1(C) \le d_2(C) \le \ldots \le d_{\log_q|C|}(C) \le n,$$

$$\textit{and if } d_r(C) = d_{r+1}(C) = \ldots = d_{r+m-1}(C) \textit{ then } d_{r+m-1}(C) < d_{r+m}(C).$$

Proof: Let D be a GR linear subcode of C such that $w_s(D) = d_r(C)$ and $\log_q|D| = r$. Let $j \in \chi(D)$ and $V_j = \{x_j : \exists(x_1,\ldots,x_j,\ldots,x_n) \in D\}$. From linearity of D it follows that $V_j = \mathcal{R}_{m-t}$ for some $1 \le t \le m$. Let $D_j = \{\mathbf{x} \in D : x_j = 0\}$.

Then $\log_q |D_j| = \log_q |D|/q^t$ and $d_{r-t}(C) \leq w_s(D_j) \leq w_s(D)-1 = d_r(C)-1$. Thus $d_{r-t}(C) < d_r(C)$ for some $t \leq m$.

To see that there exist a code A such that $d_r(A) = d_{r+1}(A) = \ldots = d_{r+m-1}(A)$, one can consider the code with generator matrix $[1, 1, \ldots, 1]$. ■

COROLLARY 1 *$M_i/M_{i-1} = q^t, 1 \leq t \leq m$.*

Note that if C is a linear code of length n over $GR(q^m, p^m)$ and $C^\perp$ is its dual code, then $|C||C^\perp| = q^{mn}$, i.e., $\log_q |C| + \log_q |C^\perp| = mn$. In the sequel, we suppose that $\{1, 2, \ldots, n\}$ is the set of positions of a code of length n.

Denote by $\oplus$ the direct sum of two sets.

LEMMA 1 *Let $I \subseteq \{1, 2, \ldots, n\}$ and $C^\perp = S_I^\perp \oplus \{\mathbf{x} \in C^\perp : x_i = 0$ for all $i \in I\}$. Then $d_r(C) = \min\{|I| : m|I| - log_q|S_I^\perp| = r\}$.*

Proof: Let $S_I = \{\mathbf{y} : \chi(\mathbf{y}) \subseteq I, \mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{x} \in S_I^\perp\}$, where $\mathbf{x} \cdot \mathbf{y}$ denotes the inner product of $\mathbf{x}$ and $\mathbf{y}$. Then $\log_q |S_I| + \log_q |S_I^\perp| = m|I|$. Let I be such that $m|I| - \log_q |S_I^\perp| = r$. Then $\log_q |S_I| = r$. Since S_I is a linear subcode of C, we see that $d_r(C) \leq |S_I| \leq \min\{|I| : m|I| - log_q|S_I^\perp| = r\}$.

Let us establish inequality in the other direction. Let D be a linear subcode of C such that $w_s(D) = d_r(C)$ and $\log_q |D| = r$. Let $J = \chi(D)$. Then $\log_q |S_J^\perp| = m|J| - r$ and $d_r(C) = |J| \geq \min\{|I| : m|I| - \log_q |S_I^\perp| = r\}$. ■

THEOREM 2 *Let C be a GR linear code and let $C^\perp$ be its dual code. Then $\{d_r(C) : r = \overline{1, \log_q |C|}\} = \{\underbrace{1, \ldots, 1}_{m \text{ times}}, \underbrace{2, \ldots, 2}_{m \text{ times}}, \ldots, \underbrace{n, \ldots, n}_{m \text{ times}}\} \backslash \{n+1-d_r(C^\perp) : r = \overline{1, \log_q |C^\perp|}\}$.*

Proof: It suffices to prove that if

$$d_{s-l}(C^\perp) < d_{s-l+1}(C^\perp) = d_{s-l+2}(C^\perp) = \ldots = d_s(C^\perp) < d_{s+1}(C^\perp),$$

then there are no more than $m-l$ values $d_{t+1}(C) = d_{t+2}(C) = \ldots = d_{t+m-l}(C)$ such that $d_t(C) < d_{t+1}(C), d_{t+m-l}(C) < d_{t+m-l+1}(C)$, and $d_{t+1}(C) = n - d_s(C^\perp) + 1$ for some t.

At first we show that $d_t(C) \leq n - d_s(C^\perp)$, where $t = \log_q |C| + s - md_s(C^\perp)$. Let $D \subseteq C^\perp, w_s(D) = d_s(C^\perp)$ and $\log_q |D| = s$. Suppose $I = \{1, 2, \ldots, n\} \setminus \chi(D)$ and $C^\perp = C_I^\perp \oplus D$. Hence, $|C_I^\perp| = q^{\log_q |C^\perp| - s} = q^{mn - \log_q |C| - s}$. By Lemma 1, we obtain $d_t(C) \leq |I| = n - d_s(C^\perp)$ with $t = m|I| - mn + \log_q |C| + s = \log_q |C| + s - md_s(C^\perp)$.

Next we show that if $d_{t+\Delta}(C) = n - d_s(C^\perp) + 1$, then $\Delta \leq m - l$. Let D be a subcode of C such that $\log_q |D| = t + \Delta$ and $|D| = d_{t+\Delta}(C)$. Suppose $I = \{1, 2, \ldots, n\} \setminus \chi(D)$ and $C = C_I \oplus D$. Hence $|C_I| = q^{\log_q |C| - t - \Delta} = q^{md_s(C^\perp) - s - \Delta}$. By Lemma 1, we obtain $d_u(C^\perp) \leq n - d_{t+\Delta}(C) = d_s(C^\perp) - 1 = d_{s-l+1}(C^\perp) - 1$, where $u = m|I| - md_s(C^\perp) + s + \Delta = s - m + \Delta$. Hence, if $\Delta \geq m - l + 1$, then $u \geq s - l + 1$; a contradiction. ■

COROLLARY 2 *Let* $\mathcal{N} = \{n+1-\delta_i : \text{ for } i \text{ such that } M_i/M_{i-1} = q^m\}$. *Then*

$$\{\delta_i^\perp : 1 \le i \le s\} = \{1, 2, \ldots, n\} \setminus \mathcal{N},$$

where $s = n - \sum_{i=1}^{t} \left\lceil \frac{\log_q M_i/M_{i-1}}{m} \right\rceil$;

$$M_i^\perp = \begin{cases} M_{i-1}^\perp q^m, & \text{if there is not } \delta_j \text{ such that } \delta_i^\perp = n+1-\delta_j, \\ M_{i-1}^\perp q^m M_{j-1}/M_j, & \text{if } \delta_i^\perp = n+1-\delta_j. \end{cases}$$

Good nonlinear codes can be constructed by mapping GR linear codes into codes over $GF(q)$ [10],[6]. Therefore it is natural to ask whether GR linear codes themselves can have better characteristics than linear codes over Galois fields. The following theorem shows that GR linear codes can not be better (from the point of view of length, cardinality, and minimum Hamming distance) than optimal linear codes, i.e. than linear codes having the largest cardinality for a given length and distance.

THEOREM 3 *Let* C *be a GR linear code over* $GR(q^m, p^m)$ *of length* n *and minimum Hamming distance* d. *Then there exist a linear code* D *over* $GF(q^m)$ *of length* n *with minimum distance* d *such that* $|D| \ge |C|$.

Proof: We consider the case of a code C over $GR(q^2, p^2)$. The case of an arbitrary m can be proved by the same way. A generator matrix of C can be written in the form

$$G = \begin{bmatrix} G1 \\ pG_2 \end{bmatrix},$$

where G_1 and G_2 are $r_1 \times n$ and $r_2 \times n$ $GR(q^2, p^2)$–matrices, $|C| = q^{2r_1+r_2}$, $r_1 \ge 1, r_2 \ge 0$. Let $\widehat{C}$ be a subcode of C with the generator matrix

$$\begin{bmatrix} pG1 \\ pG_2 \end{bmatrix}.$$

Obviously, the code $\widehat{C}$ has the same length and the same minimum distance as C. Note that $\widehat{C}$ contains only elements from the ideal $pGR(q^2, p^2)$. Let us define exterior multiplication of an element $\alpha \in pGR(q^2, p^2)/p^2GR(q^2, p^2) = pGR(q^2, p^2)$ by an element $a \in GR(q^2, p^2)/pGR(q^2, p^2)$, $a = \bar{a} + pGR(q^2, p^2)$, as follows $\alpha a = \alpha\bar{a}$. Analogously, $a(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\bar{a}\alpha_1, \bar{a}\alpha_2, \ldots, \bar{a}\alpha_n)$. Now taking into account (1), it is easy to see that $\widehat{C}$ is a vector space over $GR(q^2, p^2)/pGR(q^2, p^2) = GF(q)$ of dimension $r_1 + r_2$. Hence $\widehat{C}$ is isomorphic to some $[n, r_1 + r_2]$ linear code $\widehat{D}$ over $GF(q)$. Let D be a code with a generator matrix of the code $\widehat{D}$ considered as a matrix over $GF(q^2)$. Then D is a linear code over $GF(q^2)$ of length n with minimum distance d and $|D| = q^{2r_1+2r_2} \ge q^{2r_1+r_2} = |C|$. ■

Using the Nordstrom-Robinson code over $\mathbf{Z}_4$ as an example, we shall show later that a GR linear code can have the same length, cardinality, and minimum distance as an optimal linear code, and the function $\mathcal{F}$ for such a GR linear code can exceed the function $\mathcal{F}$ for the optimal linear code.

3. An Application to The Wire-Tap Channel of Type II

One of motivations for the introduction of a notion of GHW was cryptographical one. Ozarow and Wyner [11] considered using linear codes on the wire-tap channel of Type II. One of their schemes uses an $[n, k]$ linear code, say C, over $GF(q)$. The code has q^{n-k} cosets, each representing a q-ary $(n-k)$-tuple. If the sender wants to transmit $n-k$ $q-$ary symbols of information to the receiver, he selects a random vector in the corresponding coset. The adversary has full knowledge of the code, but not of the random selection of a vector in a coset.

Wei showed [14] that in the linear case if the adversary is allowed to tap s symbols (of his choice) from the sender, he will obtain r $q-$ary symbols of information, if and only if $s \geq d_r(C^{\perp})$.

It is obvious that we can use a group code on the said channel. Let $\mathcal{C}$ be a group code of length n over an additive group $\Theta, |\Theta| = \theta$. The code $\mathcal{C}$ has $\theta^n/|\mathcal{C}|$ cosets. As in the linear case, the number of a coset is transmitted information and the sender transmits a random vector from the corresponding coset. Let $\mathcal{C}$ have values $\{\delta_1, \delta_2, \ldots, \delta_t\}, \{M_1, M_2, \ldots, M_t\}$ for some $t \geq 1$ and let the code $\mathcal{C}$ be used on the wire-tap channel of type II. We prove the following theorem.

THEOREM 4 *Let s symbols be taped from transmitter and $\delta_i \leq n - s < \delta_{i+1}$. Then the number of cosets of $\mathcal{C}$ that can (equiprobably) contain the transmitted vector (with known s taped symbols) is equal to or greater than $\frac{\theta^{n-s}}{M_i}$.*

Proof: Let I be a set of taped positions. By L denote a set of vectors that contain zeros on positions from I, i.e., $L = \{\mathbf{v} \in \Theta^n : v_i = 0 \text{ for all } i \in I\}$. Cosets of $\mathcal{C}$ that contain vectors from L form a subgroup, say Ψ, of the quotient group of cosets of $\mathcal{C}$. Cosets of $\mathcal{C}$ that contain vectors with given values on positions from I form a coset of Ψ. Hence w.l.o.g. we can assume that taped symbols are zeros. Denote by D a set of codewords of $\mathcal{C}$ that belong to L. It is clear that any coset of $\mathcal{C}$ that belongs to Ψ has exactly $|D|$ vectors from L and hence $|\Psi| = \theta^{n-s}/|D|$.

The best strategy for the adversary is to choose the set I such that the cardinality of D would be maximal. We claim that $|D| \leq M_i$. Indeed, suppose that there exists a subcode $B \subseteq \mathcal{C}$ such that $|B| > M_i$ and $w_s(B) \leq n - s < \delta_{i+1}$. But by the definition of GHW the support weight of any subcode of cardinality larger than M_i is equal to or greater than δ_{i+1}. A contradiction. Hence the transmitted vector equiprobably belongs to one of $|\Psi|$ cosets of $\mathcal{C}$, and $|\Psi| \geq \frac{\theta^{n-s}}{M_i}$. ■

4. Subcodes of a $\mathbf{Z}_4-$linear code

Let C be a $\mathbf{Z_4}-$linear code of length n and cardinality $4^{r_1}2^{r_2}$ with a generator matrix

$$G = \begin{bmatrix} G_1 \\ 2G_2 \end{bmatrix},$$

where G_1 and G_2 are $r_1 \times n$ and $r_2 \times n$ $\mathbf{Z}_4$–matrices. We will say that C is an $[n; r_1, r_2]$ code. By $\widehat{C}$ denote a subcode of C with the generator matrix $[2G_1]$ and by $\widetilde{C}$ denote a subcode of C with the generator matrix $\begin{bmatrix} 2G_1 \\ 2G_2 \end{bmatrix}$.

Remark. Throughout the rest of the paper we use symbols $\sim$ and $\wedge$ in this meaning.

Note that $\widehat{C}$ and $\widetilde{C}$ are $[n; 0, r_1]$ and $[n; 0, r_1 + r_2]$ codes respectively. Let D be an $[n; s_1, s_2], s_1 \leq r_1, s_1 + s_2 \leq r_1 + r_2$, subcode of C and let $\begin{bmatrix} \mathcal{G}_1 \\ 2\mathcal{G}_2 \end{bmatrix}$ be a generator matrix of D, where $\mathcal{G}_1$ and $\mathcal{G}_2$ are $s_1 \times n$ and $s_2 \times n$ $\mathbf{Z}_4$–matrices respectively. By $\nu(C, s_1, s_2)$ denote the set of subcodes $[n; s_1, s_2], s_1 \leq r_1, s_1 + s_2 \leq r_1 + r_2$, of C. In the sequel, we need the following propositions.

PROPOSITION 1

$$\mu(r_1, r_2, s_1, s_2) = |\nu(C, s_1, s_2)| = 2^{s_1(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1 + r_2 - s_1 \\ s_2 \end{bmatrix}.$$

Proof: Consider the number of ways of choosing s_1 $\mathbf{Z}_4$–linearly independent codewords. Note that linear independence over $\mathbf{Z}_4$ implies that none of these codewords consists only of zero divisors of $\mathbf{Z}_4$. The total number of codewords of C equals $2^{2r_1+r_2}$ and $2^{r_1+r_2}$ of them consist of zero divisors of $\mathbf{Z}_4$. Hence the first codeword, say $\mathbf{u}_1$, can be chosen by $2^{2r_1+r_2} - 2^{r_1+r_2}$ distinct ways. The second codeword $\mathbf{u}_2$ can not be equal to $a\mathbf{u}_1 + \mathbf{c}, a \in \mathbf{Z}_4, \mathbf{c} \in \widetilde{C}$. Indeed, if $\mathbf{u}_2 = a\mathbf{u}_1 + \mathbf{c}$, then either $2\mathbf{u}_2 = \mathbf{0}$ or $2\mathbf{u}_2 = 2a\mathbf{u}_1 + 2\mathbf{c} = 2\mathbf{u}_1$, i.e., codewords $\mathbf{u}_1$ and $\mathbf{u}_2$ are not $\mathbf{Z}_4$–linearly independent. Thus $\mathbf{u}_2$ can be chosen by $2^{2r_1+r_2} - 4 \cdot 2^{r_1+r_2-1}$ ways. Analogously, if we have codewords $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{t-1}$ then a codeword $\mathbf{u}_t$ can be chosen by $2^{2r_1+r_2} - 4^{t-1}2^{r_1+r_2-t+1}$ distinct ways. Thus the number of ways of choosing s_1 $\mathbf{Z}_4$–linearly independent codewords equals

$$(2^{2r_1+r_2} - 2^{r_1+r_2})(2^{2r_1+r_2} - 2^{r_1+r_2+1}) \ldots (2^{2r_1+r_2} - 2^{r_1+r_2+s_1-1}) \tag{2}$$

Now we have to choose s_2 $\mathbf{Z}_2$–linearly independent codewords that consist from zero divisors of $\mathbf{Z}_4$. These codewords should also be $\mathbf{Z}_2$–linearly independent from the s_1 codewords chosen earlier. The number of ways for choosing these codewords equals

$$(2^{r_1+r_2-s_1} - 1)(2^{r_1+r_2-s_1} - 2) \ldots (2^{r_1+r_2-s_1} - 2^{s_2-1}). \tag{3}$$

Using the same arguments, one can see that the number of distinct generator matrices of any $[n; s_1, s_2]$ code equals

$$2^{(s_1+s_2)s_1}(2^{s_1} - 1)(2^{s_1} - 2) \ldots (2^{s_1} - 2^{s_1-1})(2^{s_2} - 1)(2^{s_2} - 2) \ldots (2^{s_2} - 2^{s_2-1}). \tag{4}$$

To complete the proof, we multiply (2) by (3) and divide by (4). ■

In the linear case, any nonzero codeword of a linear code belongs to one and the same number of subcodes of a given dimension. The situation is different in the $\mathbf{Z}_4$–linear case. We should consider the following three cases.

PROPOSITION 2 *Let* $\mathbf{u} \in C, 2\mathbf{u} \neq \mathbf{0}$. *Then* $\mathbf{u}$ *belongs to*

$$\mathcal{A}_1 = 2^{(s_1-1)(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1-1 \\ s_1-1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1 \\ s_2 \end{bmatrix}$$

subcodes from $\nu(C, s_1, s_2)$.

Proof: Recall that the generator matrix of C has the form $\begin{bmatrix} G_1 \\ 2G_2 \end{bmatrix}$. To prove the proposition, we assume that one of rows of the matrix G_1 equals $\mathbf{u}$ and use arguments from the proof of Proposition 1. ■

PROPOSITION 3 *Let* $\mathbf{u} \in \widehat{C}, \mathbf{u} \neq \mathbf{0}$. *Then* $\mathbf{u}$ *belongs to*

$$\begin{aligned} \mathcal{A}_2 \;=\; & 2^{s_1(r_1+r_2-s_1-s_2)} \left(\begin{bmatrix} r_1-1 \\ s_1-1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1 \\ s_2 \end{bmatrix} \right. \\ & + \left. \frac{2^{s_1}(2^{r_1-s_1}-1)}{2^{r_1}-1} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1-1 \\ s_2-1 \end{bmatrix} \right) \end{aligned}$$

subcodes from $\nu(C, s_1, s_2)$.

Proof: Suppose a codeword $\mathbf{u} \in D$; then either $\mathbf{u} \in \widehat{D}$ or $\mathbf{u} \in \widetilde{D} \setminus \widehat{D}$.

Consider the case $\mathbf{u} \in \widehat{D}$. The subcode D contains $2^{s_1}-1$ nonzero codewords $\mathbf{w}$ such that $\mathbf{w} \in \widehat{D}$. The code C contains $2^{r_1}-1$ nonzero codewords $\mathbf{v}$ such that $\mathbf{v} \in \widehat{C}$. Hence the number of subcodes D such that $\mathbf{u} \in \widehat{D}$ equals the total number of $[n; s_1, s_2]$ subcodes of C multiplied by the factor $\frac{2^{s_1}-1}{2^{r_1}-1}$. Using Proposition 1, we get

$$2^{s_1(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1 \\ s_2 \end{bmatrix} \frac{2^{s_1}-1}{2^{r_1}-1}. \tag{5}$$

Consider the case $\mathbf{u} \in \widetilde{D} \setminus \widehat{D}$. From (5) it follows that the number of subcodes D such that $\mathbf{u} \notin \widehat{D}$ equals

$$2^{s_1(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1 \\ s_2 \end{bmatrix} \left(1 - \frac{2^{s_1}-1}{2^{r_1}-1}\right).$$

Assuming that one of rows of $\mathcal{G}_2$ equals $\mathbf{u}$ and using arguments like ones from Proposition 1, we get the number of subcodes D such that $\mathbf{u} \in \widetilde{D} \setminus \widehat{D}$. This number equals

$$2^{s_1(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1-1 \\ s_2-1 \end{bmatrix} \frac{2^{s_1}(2^{r_1-s_1}-1)}{2^{r_1}-1}. \tag{6}$$

Summation of (5) and (6) completes the proof. ■

PROPOSITION 4 *Let* $\mathbf{u} \in \widetilde{C} \setminus \widehat{C}$. *Then* $\mathbf{u}$ *belongs to*

$$\mathcal{A}_3 = 2^{s_1(r_1+r_2-s_1-s_2)} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1-1 \\ s_2-1 \end{bmatrix}$$

subcodes from $\nu(C, s_1, s_2)$.

Proof: Suppose that $\mathbf{u} \in D$. From the conditions of the proposition it follows that $\mathbf{u} \notin \widehat{D}$. Hence the matrix $\mathcal{G}_1$ can be formed by any s_1 $\mathbf{Z}_4$–linearly independent codewords of C. Assuming that one of rows of the matrix $\mathcal{G}_2$ equals $\mathbf{u}$ and using arguments from the proof of Proposition 1, we complete the proof. ■

By $wt_L(\mathbf{u})$ denote the Lee weight of a vector $\mathbf{u}$.

PROPOSITION 5 ([3], [13]) *For an* $[n; r_1, r_2]$ *code* C *we have*

$$\sum_{\mathbf{u} \in C} wt_L(\mathbf{u}) = 2^{2r_1+r_2} w_s(C).$$

Using an approach suggested in [8], we can obtain a connection between support weight of C and supports weights of its subcodes.

PROPOSITION 6

$$\begin{aligned} \sum_{V \in \nu(C,s_1,s_2)} w_s(V) &= \frac{1}{2^{2s_1+s_2}} \left(2^{2r_1+r_2} \mathcal{A}_1 w_s(C) + 2^{r_1+r_2} (\mathcal{A}_3 - \mathcal{A}_1) w_s(\widetilde{C}) \right. \\ &+ \left. \left(2^{r_1} (\mathcal{A}_2 - \mathcal{A}_3) w_s(\widehat{C}) \right) \right), \end{aligned}$$

where $\mathcal{A}_1, \mathcal{A}_2$, *and* $\mathcal{A}_3$ *are numbers from Propositions 2,3, and 4.*

Proof: Using Proposition 5 and Propositions 2,3, and 4, we get (an abbreviation s.t. means such that)

$$\begin{aligned} &\sum_{V \in \nu(C,s_1,s_2)} w_s(V) = \sum_{V \in \nu(C,s_1,s_2)} \frac{1}{2^{2s_1+s_2}} \sum_{\mathbf{u} \in V \setminus \mathbf{0}} wt_L(\mathbf{u}) \\ &= \frac{1}{2^{2s_1+s_2}} \left[\sum_{\mathbf{u} \in C, 2\mathbf{u} \neq \mathbf{0}} wt_L(\mathbf{u}) \sum_{V \in \nu(C,s_1,s_2) \text{ s.t. } \mathbf{u} \in V} 1 \right. \\ &+ \sum_{\mathbf{u} \in \widehat{C} \setminus \mathbf{0}} wt_L(\mathbf{u}) \sum_{V \in \nu(C,s_1,s_2) \text{ s.t. } \mathbf{u} \in V} 1 \\ &+ \left. \sum_{\mathbf{u} \in \widetilde{C} \setminus \widehat{C}} wt_L(\mathbf{u}) \sum_{V \in \nu(C,s_1,s_2) \text{ s.t. } \mathbf{u} \in V} 1 \right] \\ &= \frac{1}{2^{2s_1+s_2}} \left[\mathcal{A}_1 \sum_{\mathbf{u} \in C, 2\mathbf{u} \neq \mathbf{0}} wt_L(\mathbf{u}) + \mathcal{A}_2 \sum_{\mathbf{u} \in \widehat{C} \setminus \mathbf{0}} wt_L(\mathbf{u}) + \mathcal{A}_3 \sum_{\mathbf{u} \in \widetilde{C} \setminus \widehat{C}} wt_L(\mathbf{u}) \right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2^{2s_1+s_2}}\left[\mathcal{A}_1\left(\sum_{\mathbf{u}\in C, 2\mathbf{u}\neq\mathbf{0}} wt_L(\mathbf{u}) + \sum_{\mathbf{u}\in\widehat{C}\setminus\mathbf{0}} wt_L(\mathbf{u}) + \sum_{\mathbf{u}\in\widetilde{C}\setminus\widehat{C}} wt_L(\mathbf{u})\right)\right.\\
&+ (\mathcal{A}_3-\mathcal{A}_1)\left(\sum_{\mathbf{u}\in\widehat{C}\setminus\mathbf{0}} wt_L(\mathbf{u}) + \sum_{\mathbf{u}\in\widetilde{C}\setminus\widehat{C}} wt_L(\mathbf{u})\right)\\
&\left.+ (\mathcal{A}_2-\mathcal{A}_1-(\mathcal{A}_3-\mathcal{A}_1))\sum_{\mathbf{u}\in\widehat{C}\setminus\mathbf{0}} wt_L(\mathbf{u})\right]\\
&= \frac{1}{2^{2s_1+s_2}}\left(2^{2r_1+r_2}\mathcal{A}_1 w_s(C) + 2^{r_1+r_2}(\mathcal{A}_3-\mathcal{A}_1)w_s(\widetilde{C}) + 2^{r_1}(\mathcal{A}_2-\mathcal{A}_3)w_s(\widehat{C})\right).
\end{aligned}
$$

■

It is known that if C is a $\mathbf{Z}_4-$linear code and C^* is its binary image under the Gray map, then minimum Lee distance of C equals minimum Hamming distance of C^* [6]. Such good nonlinear binary codes as Kerdock, Delsarte-Goethals, and Preparata codes can be constructed by Gray mapping $\mathbf{Z}_4-$linear codes [6]. Therefore it is natural to ask what is the minimal length, $N(r_1, r_2; d_L)$, of an $[n; r_1, r_2]$ $\mathbf{Z}_4-$linear code with minimum Lee distance d_L. We confine ourselves to the case of $[n; r, 0]$ codes and obtain an analog of the Griesmer bound for $N(r, 0; d_L)$.

THEOREM 5

$$N(r,0;d_L) \geq \sum_{i=0}^{r-1}\left\lceil \frac{3\cdot 2^{i(i-1)/2}}{4\cdot\prod_{j=0}^{i-1}(2^{i+1-j}+1)} d_L \right\rceil.$$

Proof: Let C be an $[n; r, 0]$ $\mathbf{Z}_4-$linear code with minimum Lee distance d_L. Note that in this case $\widehat{C} = \widetilde{C}$. Suppose $s_1 = r-1$ and $s_2 = 0$. Using Proposition 6, we get

$$2^{r-1}(2^r-1)N(r-1,0;d_L) \leq 2^r(2^{r-1}-1)w_s(C) + (2^{r-1}-1)w_s(\widehat{C}).$$

Obviously, $w_s(\widehat{C}) \leq w_s(C)$. Assuming $w_s(\widehat{C}) = w_s(C) = N(r, 0; d_L)$, we get

$$N(r,0;d_L) \geq \frac{2^{r-1}(2^r-1)}{(2^{r-1}-1)(2^r+1)} N(r-1,0;d_L), \quad \text{or} \tag{7}$$

$$N(r,0;d_L) \geq N(r-1,0;d_L) + \left\lceil \frac{N(r-1,0;d_L)}{(2^{r-1}-1)(2^r+1)} \right\rceil. \tag{8}$$

From (7) and (8) it follows that

$$N(r,0;d_L) \geq \sum_{i=0}^{r-1}\left\lceil \frac{2^{i(i-1)/2}}{\prod_{j=0}^{i-1}(2^{i+1-j}+1)} N(1,0;d_L) \right\rceil. \tag{9}$$

If $A = \{\mathbf{0}, \mathbf{a}, 2\mathbf{a}, 3\mathbf{a}\}$ is an $[n; r, 0]$ code with minimum distance d_L, then $w_s(A) \geq N(1, 0; d_L) = \lceil \frac{3}{4} d_L \rceil$. Otherwise $wt_L(2\mathbf{a}) < d_L$. Substituting the value for $N(1, 0; d_L)$ in (9) completes the proof. ■

Note that the Nordstrom-Robinson code over $\mathbf{Z}_4$ meets this bound.

5. Bounds on Higher Weights

We start with Plotkin type bound.

Let $d_{s_1,s_2}(C) = \min\{w_s(D) : D \text{ is an } [n; s_1, s_2] \text{ subcode of } C\}$.

THEOREM 6 *If C is an $[n; r_1, r_2]$ $\mathbf{Z}_4$–linear code then*

$$d_{2s}(C) \leq n\left(1 - \frac{(2^{r_1-s} - 1)}{(2^{r_1} - 1)2^s}\right),$$

$$d_{2s+1}(C) \leq n\left(1 - \frac{(2^{r_1+r_2-s-1} - 1)(2^{r_1-1} - 1)}{(2^{r_1} - 1)(2^{r_1+r_2-s} - 1)2^s}\right).$$

Proof: Calculate in two ways the sum, say S, of codeword Lee weights of all $[n; s_1, s_2]$ subcodes of C.

The total number of subcodes equals $\mu(r_1, r_2, s_1, s_2)$. According to Proposition 5 the sum of Lee weights of codewords of any $[n; s_1, s_2]$ subcode is greater than or equal to $2^{2s_1+s_2} d_{2s_1+s_2}(C)$. Hence

$$S \geq 2^{2s_1+s_2} \mu(r_1, r_2, s_1, s_2) d_{2s_1+s_2}(C).$$

Let A be a $2^{2r_1+r_2} \times n$ matrix whose rows are codewords of C. W.l.o.g. we can consider the first column of A. Any $a \in \mathbf{Z}_4$ occurs $2^{2r_1+r_2-2}$ times in the column. Let $\mathbf{v} = (a, v_2, \ldots, v_n) \in C$. If $a = 1$ or 3 then $\mathbf{v}$ belongs to $\mathcal{A}_1$ $[n; s_1, s_2]$ subcodes of C and there are $2 \cdot 2^{2r_1+r_2-2}$ such vectors. If $a = 2$ then there are three different cases. 1) $\mathbf{v} \notin \widetilde{C}$. Then $\mathbf{v}$ belongs to $\mathcal{A}_1$ $[n; s_1, s_2]$ subcodes of C and there are $2^{2r_1+r_2-2} - 2^{r_1+r_2-1}$ such vectors. 2) $\mathbf{v} \in \widehat{C}$. Then $\mathbf{v}$ belongs to $\mathcal{A}_2$ $[n; s_1, s_2]$ subcodes of C and there are 2^{r_1-1} such vectors. 3) $\mathbf{v} \in \widetilde{C} \setminus \widehat{C}$. Then $\mathbf{v}$ belongs to $\mathcal{A}_3$ $[n; s_1, s_2]$ subcofdes of C and there are $2^{r_1+r_2-1} - 2^{r_1-1}$ such vectors. Hence

$$S = n(\mathcal{A}_1(2 \cdot 2^{2r_1+r_2-2} + 2 \cdot (2^{2r_1+r_2-2} - 2^{r_1+r_2-1})) + \mathcal{A}_2 \cdot 2 \cdot 2^{r_1-1} + \mathcal{A}_3 \cdot 2 \cdot (2^{r_1+r_2-1} - 2^{r_1-1})).$$

After computations we get

$$d_{s_1,s_2}(C) \leq n\left(1 - \frac{(2^{r_1+r_2-s_1-s_2} - 1)(2^{r_1-s_1} - 1)}{(2^{r_1} - 1)(2^{r_1+r_2-s_1} - 1)2^{s_1}}\right).$$

Setting $s_1 = s, s_2 = 0$ and $s_1 = s - 1, s_2 = 1$ for d_{2s} and d_{2s+1} respectively, we get the assertion. ■

Let D be a $\mathbf{Z}_4$–linear code and C be its $[n; r_1, r_2]$ subcode such that $w_s(C) = d_{r_1,r_2}(D)$. Let γ_{r_1,r_2} and γ_{r_1} be upper bounds for support sizes of subcodes $\widetilde{C}$ and $\widehat{C}$.

Remark. If we don't have an information on the structure of subcodes $\widetilde{D}$ and $\widehat{D}$ then we can say that $\gamma_{r_1,r_2} = \gamma_{r_1} = d_{r_1,r_2}(D)$. Sometimes, if we know weight distribution of support sizes of all subcodes of codes $\widetilde{D}$ and $\widehat{D}$, we can find exact values for support sizes of subcodes $\widetilde{C}$ and $\widehat{C}$. (see next chapter).

Now we can get a Griesmer type bound.

THEOREM 7 *If* $2^{r_1+r_2-s_1}(2^{r_1}-2^{s_1})+2^{s_1}-1 \geq 2^{r_1+r_2-s_1-s_2}(2^{r_1}-1)$ *then*

$$\begin{aligned} d_{r_1,r_2}(D) &\geq \frac{(2^{r_1}-1)}{(2^{s_1}-1)(2^{r_1-s_1})} d_{s_1,s_2}(D) - \left(\frac{(2^{r_1}-1)(2^{s_2}-1)2^{r_2-s_1-s_2}}{(2^{s_1}-1)(2^{r_1+r_2-s_1}-1)} - \frac{1}{2^{r_1}} \right) \gamma_{r_1,r_2} \\ &- \frac{2^{r_1+r_2-s_1}-2^{s_2}}{2^{s_1+s_2}(2^{r_1+r_2-s_1}-1)} \gamma_{r_1}, \end{aligned}$$

otherwise

$$\begin{aligned} d_{r_1,r_2}(D) &\geq \frac{(2^{r_1}-1)}{(2^{s_1}-1)(2^{r_1-s_1})} d_{s_1,s_2}(D) - \left(\frac{(2^{r_1}-1)(2^{s_2}-1)2^{r_2-s_1-s_2}}{(2^{s_1}-1)(2^{r_1+r_2-s_1}-1)} - \frac{1}{2^{r_1}} \right) d_{r_1,r_2}(D) \\ &- \frac{2^{r_1+r_2-s_1}-2^{s_2}}{2^{s_1+s_2}(2^{r_1+r_2-s_1}-1)} \gamma_{r_1}. \end{aligned}$$

Proof: From proposition 6 we have

$$\begin{aligned} & w_s(C) = d_{r_1,r_2}(D) \\ = & \frac{1}{2^{2r_1+r_2}\mathcal{A}_1} \left(2^{2s_1+s_2} \sum_{V \in \nu(C,s_1,s_2)} w_s(V) - 2^{r_1+r_2}(\mathcal{A}_3-\mathcal{A}_1) w_s(\widetilde{C}) \right. \\ - & \left. 2^{r_1}(\mathcal{A}_2-\mathcal{A}_3) w_s(\widehat{C}) \right) \\ \geq & \frac{1}{2^{2r_1+r_2}\mathcal{A}_1} \left(2^{2s_1+s_2} \mu(r_1,r_2,s_1,s_2) d_{s_1,s_2}(C) - 2^{r_1+r_2}(\mathcal{A}_3-\mathcal{A}_1) w_s(\widetilde{C}) \right. \\ - & \left. 2^{r_1}(\mathcal{A}_2-\mathcal{A}_3) w_s(\widehat{C}) \right). \end{aligned}$$

It is easy to check that $\mathcal{A}_2-\mathcal{A}_3$ is always nonnegative when $s_1 \leq r_1$ and $s_1+s_2 \leq r_1+r_2$; $\mathcal{A}_3-\mathcal{A}_1$ is nonnegative iff $2^{r_1+r_2-s_1}(2^{r_1}-2^{s_1})+2^{s_1}-1 \geq 2^{r_1+r_2-s_1-s_2}(2^{r_1}-1)$. Replacing $w_s(\widehat{C})$ by γ_{r_1} and $w_s(\widetilde{C})$ by $\gamma_{r_1,r_2}(C)$ in the case when $\mathcal{A}_3-\mathcal{A}_1 > 0$ and by $d_{r_1,r_2}(D)$ otherwise, we get the assertion. ■

As corollarys we have the following relations.
Although in general d_r can be equal to d_{r+1}, d_{s_1,s_2+1} is always greater than d_{s_1,s_2}.

COROLLARY 3

$$d_{s_1,s_2+1}(D) \geq \max \left\{ d_{s_1,s_2}(D) + \frac{1}{2^{s_1}(2^{s_2+1}-1)} \sum_{i=1}^{s_2+1} \left\lceil \frac{d_{0,s_1}(D)}{2^i(2^{s_1}-1)} \right\rceil, d_{0,s_1+s_2+1}(D) \right\},$$

Proof: Inequality $d_{s_1,s_2+1}(D) \geq d_{0,s_1+s_2+1}(D)$ is obvious. Assuming $r_1 = s_1, r_2 = s_2 + 1$, we get (like in Theorem 7)

$$d_{s_1,s_2+1}(D) \geq d_{s_1,s_2}(D) + \frac{w_s(\widetilde{D}) - w_s(\widehat{D})}{2^{s_1}(2^{s_2+1}-1)}.$$

For linear binary codes we have [8]

$$d_{r+s} \geq d_r + \sum_{i=1}^{s} \left\lceil \frac{d_r}{2^i(2^r-1)} \right\rceil.$$

Since $\widetilde{D}$ and $\widehat{D}$ are isomorphic to linear binary codes we have

$$w_s(\widetilde{D}) \geq w_s(\widehat{D}) + \sum_{i=1}^{s_2+1} \left\lceil \frac{w_s(\widehat{D})}{2^i(2^{s_1}-1)} \right\rceil.$$

Since $w_s(\widehat{D}) \geq d_{0,s_1}(D)$ the assertion follows. ■

COROLLARY 4

$$d_{s_1+1,s_2}(D) \geq \max\left\{ \frac{(2^{s_2+1}-1)(2^{s_1+1}-1)2^{s_1}}{(2^{s_2+1}-1)(2^{s_1+1}-1)2^{s_1}-1} d_{s_1,s_2}(D), d_{0,s_1+s_2+1}(D) \right\},$$

Proof: If $r_1 = s_1+1, r_2 = s_2$ then $2^{r_1+r_2-s_1}(2^{r_1}-2^{s_1})+2^{s_1}-1 \geq 2^{r_1+r_2-s_1-s_2}(2^{r_1}-1)$ and the assertion follows from assuming $\gamma_{r_1,r_2} = \gamma_{r_1} = d_{s_1,s_2}(D)$ and Theorem 7. ■

Note that as an estimation for value $d_{0,t}(D)$ we can use any lower bound for generalized Hamming weights of binary liner codes. For example using the Griesmer bound we get

$$d_{s_1,1}(D) \geq \max\left\{ d_{s_1,0}(D) + \frac{\sum_{i=0}^{s_1-1}\lceil d(D)/2^i \rceil}{2^{s_1+1}(2^{s_1}-1)}, \sum_{i=0}^{s_1} \left\lceil \frac{d(D)}{2^i} \right\rceil \right\}, \tag{10}$$

$$d_{s_1+1,0}(D) \geq \max\left\{ \frac{(2^{s_1+1}-1)2^{s_1}}{(2^{s_1+1}-1)2^{s_1}-1} d_{s_1,0}(D), \sum_{i=0}^{s_1} \left\lceil \frac{d(D)}{2^i} \right\rceil \right\}, \tag{11}$$

where $d(C)$ is minimum Hamming weight of C. Note that some codes have equalities in (10) and (11) (see Remark 5).

Our last bound is an existence one.

THEOREM 8 *For any $s_1 \leq r_1$ and $s_2, s_1+s_2 \leq r_1+r_2$ there exists an $[n; r_1, r_2]$ $\mathbf{Z}_4$−linear code with $d_{s_1,s_2} \geq d$ which satisfies*

$$\binom{n}{d} 2^{s_1(d+r_1+r_2-n-s_1-s_2)} \begin{bmatrix} d \\ s_1 \end{bmatrix} \begin{bmatrix} r_1 \\ s_1 \end{bmatrix} \begin{bmatrix} d-s_1 \\ s_2 \end{bmatrix} \begin{bmatrix} r_1+r_2-s_1 \\ s_2 \end{bmatrix} < \begin{bmatrix} n \\ s_1 \end{bmatrix} \begin{bmatrix} n-s_1 \\ s_2 \end{bmatrix}.$$

Proof: The number of $[n; r_1, r_2]$ codes that contain a given subcode $[n; s_1, s_2]$ equals

$$\mu(n, 0, r_1, r_2)\mu(r_1, r_2, s_1, s_2)/\mu(n, 0, s_1, s_2).$$

$\binom{n}{d}\mu(d, 0, s_1, s_2)$ is the number of $[n; s_1, s_2]$ codes with supports of size less than or equal to d. Hence if $\binom{n}{d}\mu(d, 0, s_1, s_2)\mu(n, 0, r_1, r_2)\mu(r_1, r_2, s_1, s_2)/\mu(n, 0, s_1, s_2)$ is less than the total number of $[n; r_1, r_2]$ codes, i.e. less than $\mu(n, 0, r_1, r_2)$, then at least one of $[n; r_1, r_2]$ codes has a subcode $[n; s_1, s_2]$ with support size greater than or equal to d. Substitution valies for the function μ gives us the assertion. ■

6. Determination of Weight Hierarchy

It is often that we know more about subcodes $\widetilde{C}$ and $\widehat{C}$, wich are isomorphic to linear binary codes. In those cases we can get better bounds.

Consider the Kerdock code over $\mathbf{Z}_4$ [6],[10]. In our notation the Kerdock code $\mathcal{K}_m$ is an $[2^m; m+1, 0]$ code, m is odd. The minimum Lee distance of $\mathcal{K}_m$ is $2^m - 2^{(m-1)/2}$ and the minimum Hamming distance is 2^{m-1}. Moreover, in [6] compositions of all codewords and weight distribution of $\mathcal{K}_m$ were found. Suppose $l_t(\mathbf{u}) = |\{i : u_i = t\}|$. A nonzero codeword $\mathbf{u} \in \mathcal{K}_m$ such that $\chi(\mathbf{u}) \neq \{1, 2, \ldots, 2^m\}$ has one of the following compositions

$$l_0(\mathbf{u}) = 2^{m-1},\ l_2(\mathbf{u}) = 2^{m-1},\ l_1(\mathbf{u}) = 0,\ l_3(\mathbf{u}) = 0;\ \text{or}$$

$$l_0(\mathbf{u}) = 2^{m-2} + c_1 2^{(m-3)/2},\ l_1(\mathbf{u}) = 2^{m-2} + c_2 2^{(m-3)/2}$$
$$l_2(\mathbf{u}) = 2^{m-2} - c_1 2^{(m-3)/2},\ l_3(\mathbf{u}) = 2^{m-2} - c_2 2^{(m-3)/2},$$

where $c_1, c_2 = \pm 1$. Let A_i be the number of codewords of weight i. Then

$$\begin{array}{rclrcl} A_0 & = & 1, & A_{3\cdot 2^{m-2} \pm 2^{(m-3)/2}} & = & 2^{m+1}(2^m - 1), \\ A_{2^{m-1}} & = & 2^{m+1} - 2, & A_{2^m} & = & 2^{m+1} + 1. \end{array} \tag{12}$$

In [6], it was also shown that subcode $\widehat{\mathcal{K}}_m$ is isomorphic to the first order Reed-Muller code $RM(1, m)$ of length 2^m. To find a lower bound for $d_r(\mathcal{K}_m)$ we need the following lemma.

LEMMA 2 *The support size of any r dimensional subcode, say C, of $RM(1, m)$ equals to either $n = 2^m$ or $\frac{2^r-1}{2^{r-1}}2^{m-1}$.*

Proof: If C contains the all-one codeword then $w_s(C) = n$. Suppose that C does not contain the all-one codeword.

It is known [14](Theorem 5) that $d_r(RM(1, m)) = \frac{2^r-1}{2^{r-1}}2^{m-1}$ and any r rows (that are not all-one) in a standard generator matrix generate a subcode whose support size equals $d_r(RM(1, m))$. Any codeword of $RM(1, m)$ can be consedered as a boolean linear function of m variables, say $v_1, \ldots, v_m$, [9](Chapter 13). Any affine transformation of $v_1, \ldots, v_m$ belongs to the automorphism group of $RM(1, m)$ and hence it does not change the support size of C. Let $f_1(\mathbf{v}), \ldots, f_r(\mathbf{v}), \mathbf{v} =$

$(v_1, \ldots, v_m)$, be boolean functions corresponding to basis vectors of C. Assume that $f_1(\mathbf{v})$ has a term v_1. Then the affine transform $v_1 \leftrightarrow f_1(v_1, \ldots, v_m)$ transforms $f_1(v_1, \ldots v_m)$ to the form v_1 and $f_2(\mathbf{v})$ to some $f_2'(\mathbf{v})$. Since $f_2(\mathbf{v})$ is linearly independent of $f_1(\mathbf{v})$ and $f_1(\mathbf{v}) + f_2(\mathbf{v}) \neq \mathbf{1}$, $f_2'(\mathbf{v})$ has a term $v_j \neq v_1$, say v_2. The affine transform $v_2 \leftrightarrow f_2'(\mathbf{v})$ transforms $f_2'(\mathbf{v})$ to the form v_2. Continuing this procedure, we get $f_1(\mathbf{v}) = v_1, \ldots, f_r(\mathbf{v}) = v_r$, that is r rows from a standard generator matrix. ■

THEOREM 9

$$d_{2r}(\mathcal{K}_m) \geq \begin{cases} n(r,0) = \frac{2^r-1}{2^{2r-2}}\left(2^{r-1}(2^{m-1} - 2^{(m-3)/2}) + 2^{m-2}\right), & \\ & \text{if } r \leq (m+1)/2, \\ d_r(RM(1,m)) = \frac{2^r-1}{2^{r-1}}2^{m-1}, & \text{else.} \end{cases}$$

$$d_{2r+1}(\mathcal{K}_m) \geq \begin{cases} n(r,1) = \frac{2^r(2^r-1)}{2^{2r}-1}(3 \cdot 2^{m-2} - 2^{(m-3)/2}) + 2^{m-2}(2^{r+1} - 1 - 2^r(2^r-1)), & \\ & \text{if } r \leq (m-1)/2, \\ d_{r+1}(RM(1,m)) = \frac{2^{r+1}-1}{2^r}2^{m-1}, & \text{else.} \end{cases}$$

Proof: Let $s_1 = 1$ and $s_2 = 0$. Then we have (like in Theorem 7)

$$d_{r_1,r_2}(\mathcal{K}_m) \geq \frac{2^{r_1}-1}{(2^{r_1-1}-1)}d_{1,0}(\mathcal{K}_m) + \frac{w_s(C)}{2^{r_1}} - \frac{w_s(\widehat{C})}{2}, \tag{13}$$

where C is an $[n; r_1, r_2]$ subcode of $\mathcal{K}_m$ such that $w_s(C) = d_{r_1,r_2}(\mathcal{K}_m)$.

$\widehat{C}$ and $\widetilde{C}$ are $[n; 0, r_1]$ and $[n; 0, r_1 + r_2]$ subcodes of $\mathcal{K}_m$ respectively, and they are isomorphic to $[n, r_1]$ and $[n, r_1 + r_2]$ subcodes of $RM(1, m)$. Hence from the Lemma 2 we get

$$w_s(\widehat{C}) = \frac{2^{r_1}-1}{2^{r_1-1}}2^{m-1} \text{ and } w_s(\widetilde{C}) = \frac{2^{r_1+r_2}-1}{2^{r_1+r_2-1}}2^{m-1}.$$

From possible compositions of codewords of $\mathcal{K}_m$ it follows that if V is an $[n; 1, 0]$ subcode of $\mathcal{K}_m$ then $w_s(V) \geq 3 \cdot 2^{m-2} - 2^{(m-3)/2}$. Hence

$$d_{r_1,r_2}(\mathcal{K}_m) \geq \frac{2^{r_1}-1}{2^{r_1-1}-1}(3 \cdot 2^{m-2} - 2^{(m-3)/2}) + \frac{2^{r_1+r_2}-1}{2^{2r_1+r_2-1}}2^{m-1} - \frac{2^{r_1}-1}{2^{r_1}}2^{m-1}. \tag{14}$$

By $n(r_1, r_2)$ denote the right side of (14). It is evident that $d_r(\mathcal{K}_m) \geq \min\{n(r_1, r_2) : 2r_1 + r_2 = r\}$. It is easy to check that $n(r_1, r_2) \leq n(r_1 - 1, r_2 + 2)$. Hence, $d_{2r}(\mathcal{K}_m) \geq n(r, 0)$ and $d_{2r+1}(\mathcal{K}_m) \geq n(r, 1)$. On the other hand, since $w_s(C) \geq w_s(\widetilde{C})$, we have

$$\begin{aligned} d_r(\mathcal{K}_m) &\geq \min\{w_s(\widetilde{C}) : C \text{ is a } \mathbf{Z}_4 - \text{linear subcode of } \mathcal{K}_m, |C| = 2^{2r_1+r_2} = 2^r\} \\ &= \begin{cases} d_{r/2}(RM(1,m)), & \text{if } r \text{ is even,} \\ d_{(r+1)/2}(RM(1,m)), & \text{if } r \text{ is odd.} \end{cases} \end{aligned}$$

It is easy to check that $n(r, 0) > d_r(RM(1, m))$ at $r \leq (m + 1)/2$ and $n(r, 1) < d_r(RM(1, m))$ at $r \leq (m - 1)/2$. ■

It is obvious that for $\mathcal{K}_m$

$$\delta_2 = 3 \cdot 2^{m-2} - 2^{(m-3)/2} \text{ and } M_2 = 4. \tag{15}$$

Indeed, it is enough to take a subcode $C =< \mathbf{a} >= \{\mathbf{0}, \mathbf{a}, 2\mathbf{a}, 3\mathbf{a}\}$ generated by a codeword $\mathbf{a}$ with the composition $l_0(\mathbf{a}) = 2^{m-2} + 2^{(m-3)/2}, l_2(\mathbf{a}) = 2^{m-2} - 2^{(m-3)/2}, l_1(\mathbf{a}) + l_3(\mathbf{a}) = 2^{m-1}$ (note that in $\mathcal{K}_m$ there are also subcodes of cardinality 4 generated by a couple of codewords consisting of zero divisors of $\mathbf{Z}_4$), and check that support weight of C meets the lower bound and $|C| = 4$.

The Kerdock code over $\mathbf{Z}_4$ can be also defined for the case of even m [6]. In this case

$$\begin{aligned} A_0 &= 1, & A_{3\cdot 2^{m-2} \pm 2^{m/2}} &= 2^m(2^m - 1), \\ A_{2^{m-1}} &= 2^{m+1} - 2, & A_{2^m} &= 2^{m+1} + 1. \end{aligned}$$

Similar to the case when m is odd we get

THEOREM 10

$$d_{2r}(\mathcal{K}_m) \geq \begin{cases} \frac{2^r-1}{2^{2r-2}} \left(2^{r-1}(2^{m-1} - 2^{m/2-1}) + 2^{m-2}\right), & \text{if } r \leq m/2, \\ \frac{2^r-1}{2^{r-1}} 2^{m-1}, & \text{else.} \end{cases}$$

$$d_{2r+1}(\mathcal{K}_m) \geq \begin{cases} \frac{2^r(2^r-1)}{2^{2r}-1}(3 \cdot 2^{m-2} - 2^{m/2}) + 2^{m-2}(2^{r+1} - 1 - 2^r(2^r - 1)), & \\ & r < m/2, \\ \frac{2^{r+1}-1}{2^r} 2^{m-1}, & \text{else.} \end{cases}$$

Remark. Theorems 9 and 10 were obtained for the first time in [2] [3]. Later and independently the following estimates were obtained for the Kerdock codes [13]:

$$d_{2r}(\mathcal{K}_m) \geq \begin{cases} \sum_{i=1}^{2r} 2^{m-i} - \sum_{i=1}^{r} 2^{(m-1)/2-i}, & \text{if } m \text{ is odd} \\ \sum_{i=1}^{2r} 2^{m-i} - \sum_{i=1}^{r} 2^{m/2-i}, & \text{if } m \text{ is even} \end{cases} \tag{16}$$

and

$$d_{2r+1}(\mathcal{K}_m) \geq d_r(\mathcal{K}_m) + 2^{m-2r-1}. \tag{17}$$

It is easy to check that estimates (16)and (17) coincides with estimates from Theorems 9 and 10 for $r < m/2$. For the $r \geq m/2$ estimates from the theorems are better than estimates (16) and (17).

Minimum distance of a code is only a small part of information contained in the spectrum of distance distribution of the code. In the same way, the number of subcodes of a given dimension with a given support weight gives much more information on a code than only minimum support weights of subcodes. For example we used distribution of support weights of the first order Reed-Muller code in the proof of Theorem 9 and 10. Papers [7],[12] are devoted to studying distributions of support weights of linear codes. We shall find weight distribution of $\mathbf{Z}_4-$linear subcodes of cardinality 2^3 (weight distribution of subcodes of cardinality 2^2 is evident). Let $A_i^3 = |\{C : C \text{ is a } \mathbf{Z}_4 - \text{linear subcode of } \mathcal{K}_m, |C| = 2^3, w_s(C) = i\}|$.

THEOREM 11

$$\begin{aligned} A^3_{7\cdot 2^{m-3}\pm 2^{(m-3)/2}} &= 2^{m-1}(2^m-1)(2^{m-1}-1), \\ A^3_{7\cdot 2^{m-3}} &= 2^m(2^m-1)(2^{m-1}-1)+\begin{bmatrix} m+1 \\ 3 \end{bmatrix}-\begin{bmatrix} m \\ 2 \end{bmatrix}, \\ A^3_{2^m} &= 3\cdot 2^{m-1}(2^m-1)+\begin{bmatrix} m \\ 2 \end{bmatrix}. \end{aligned}$$

Proof: We should consider subcodes $[2^m;1,1]$ and $[2^m;0,3]$.

Consider an $[2^m;1,1]$ subcode $C=<\mathbf{a},\mathbf{b}>=\{\mathbf{0},\mathbf{a},2\mathbf{a},3\mathbf{a},\mathbf{b},\mathbf{a}+\mathbf{b},2\mathbf{a}+\mathbf{b},3\mathbf{a}+\mathbf{b}\}$, generated by codewords $\mathbf{a},2\mathbf{a}\neq\mathbf{0},\chi(\mathbf{a})\neq\{1,2,\ldots,2^m\}$ and $\mathbf{b}$ with $l_0(\mathbf{b})=l_2(\mathbf{b})=2^{m-1}$.

Assuming $s_1=1,s_2=0,r_1=r_2=1$, and using Proposition 6, we get that $w_s(C)=7\cdot 2^{m-3}-2^{(m-3)/2}$ if $|\chi(\mathbf{a})|=|\chi(\mathbf{a}+\mathbf{b})|=3\cdot 2^{m-2}-2^{(m-3)/2}$, $w_s(C)=7\cdot 2^{m-3}$ if $|\chi(\mathbf{a})|=3\cdot 2^{m-2}-c2^{(m-3)/2}$ and $|\chi(\mathbf{a}+\mathbf{b})|=3\cdot 2^{m-2}+c2^{(m-3)/2}$, $c=\pm 1$, and $w_s(C)=7\cdot 2^{m-3}+2^{(m-3)/2}$ if $|\chi(\mathbf{a})|=|\chi(\mathbf{a}+\mathbf{b})|=3\cdot 2^{m-2}+2^{(m-3)/2}$.

Let S_1,S_2 and S_3 be sets of $[2^m;1,1]$ subcodes with support weights $7\cdot 2^{m-3}-2^{(m-3)/2}$, $7\cdot 2^{m-3}$, and $7\cdot 2^{m-3}+2^{(m-3)/2}$ respectively.

It is easy to see that if $w_s(<\mathbf{a},\mathbf{b}>)=7\cdot 2^{m-3}-2^{(m-3)/2}$, then

$$w_s(<\mathbf{a},\mathbf{b}+(2,2,\ldots,2)>)=7\cdot 2^{m-3}+2^{(m-3)/2}. \tag{18}$$

Hence, $|S_1|=|S_3|$.

As it was mentioned if $w_s(C)=7\cdot 2^{m-3}-2^{(m-3)/2}$ then any codeword $\mathbf{a}\in C, 2\mathbf{a}\neq\mathbf{0}$, has Hamming weight $3\cdot 2^{m-2}-2^{(m-3)/2}$. Hence the total number of codewords of weight $3\cdot 2^{m-2}-2^{(m-3)/2}$ in all the subcodes from S_1 equals

$$|S_1|\cdot 4. \tag{19}$$

In $\mathcal{K}_m$ there are $2^{m+1}-4$ codewords $\mathbf{b}, l_0(\mathbf{b})=l_2(\mathbf{b})=2^{m-1}$, that are $\mathbf{Z}_4-$linearly independent form a given $\mathbf{a}\in\mathcal{K}_m, 2\mathbf{a}\neq\mathbf{0}$. According to (18), half of these codewords $\mathbf{b}$ form subcodes $<\mathbf{a},\mathbf{b}>$ with $w_s(<\mathbf{a},\mathbf{b}>)=7\cdot 2^{m-3}-2^{(m-3)/2}$, and another half of them form subcodes $<\mathbf{a},\mathbf{b}>$ with $w_s(<\mathbf{a},\mathbf{b}>)=7\cdot 2^{m-3}+2^{(m-3)/2}$. So, a given codeword $\mathbf{a},2\mathbf{a}\neq\mathbf{0}$, belongs to 2^m-2 subcodes from S_1. Hence the the total number of codewords of weight $3\cdot 2^{m-2}-2^{(m-3)/2}$ in all the subcodes from S_1 equals the half number of codewords of this weight (codewords $\mathbf{a}$ and $3\mathbf{a}$ always belong to one and the same subcode) multiplied by 2^m-2. So, using (19) and (12), we get

$$|S_1|\cdot 4=2^{m+1}(2^m-1)(2^{m-1}-1) \text{ and } |S_1|=2^{m-1}(2^m-1)(2^{m-1}-1). \tag{20}$$

Any $[2^m;1,1]$ subcode of $\mathcal{K}_m$ with support $\{1,2,\ldots,2^m\}$ contains the codeword $(2,2,\ldots,2)$. Assuming $s_1=s_2=1,r_1=m+1,r_2=0$ and using Proposition 3, we get that the number of an $[2^m;1,1]$ subcodes with support $\{1,2,\ldots,2^m\}$ equals

$$\mathcal{A}_2=3\cdot 2^{m-1}(2^m-1). \tag{21}$$

According to Proposition 1 the total number of an $[2^m; 1, 1]$ subcodes equals $|\nu(\mathcal{K}_m; 1, 1)| = 2^{m-1}(2^{m+1} - 1)(2^m - 1)$. Hence the number of $[2^m; 1, 1]$ subcodes with support weight $7 \cdot 2^{m-3}$ equals $|S_2| = |\nu(\mathcal{K}_m; 1, 1)| - \mathcal{A}_2 - 2|S_1| = 2^m(2^m - 1)(2^{m-1} - 1)$.

Since $2\mathcal{K}_m$ is isomorphic to $RM(1, m)$, one can see that in $2\mathcal{K}_m$ there are $\begin{bmatrix} m \\ 2 \end{bmatrix}$ $[2^m; 0, 3]$ subcodes of $\mathcal{K}_m$ with support weight 2^m and $\begin{bmatrix} m+1 \\ 3 \end{bmatrix} - \begin{bmatrix} m \\ 2 \end{bmatrix}$ $[2^m; 0, 3]$ subcodes with support weight $7 \cdot 2^{m-3}$.

The numbers $|S_1|, |S_2| + \begin{bmatrix} m+1 \\ 3 \end{bmatrix} - \begin{bmatrix} m \\ 2 \end{bmatrix}, |S_3|$, and $\mathcal{A}_2 + \begin{bmatrix} m \\ 2 \end{bmatrix}$ give the numbers of $\mathbf{Z}_4$–linear subcodes of cardinality 8 with support weights $7 \cdot 2^{m-3} - 2^{(m-3)/2}, 7 \cdot 2^{m-3}, 7 \cdot 2^{m-3} + 2^{(m-3)/2}$, and 2^m respectively. ■

COROLLARY 5 *The values δ_3 and M_3 for the Kerdock code $\mathcal{K}_m$ are*

$$\delta_3 = 7 \cdot 2^{m-3} - 2^{(m-3)/2} \text{ and } M_3 = 8.$$

Using Theorems 2 and 7, we can find the weight hierarchy of some short $\mathbf{Z_4}$-linear codes.

THEOREM 12 *The Nordstrom-Robinson code ($\mathcal{NR}$) over $\mathbf{Z_4}$ has weight hierarchy*

$$\{\delta_1, \ldots, \delta_5\} = \{4, 5, 6, 7, 8\} \text{ and } \{M_1, \ldots, M_5\} = \{2, 2^2, 2^4, 2^6, 2^8\}.$$

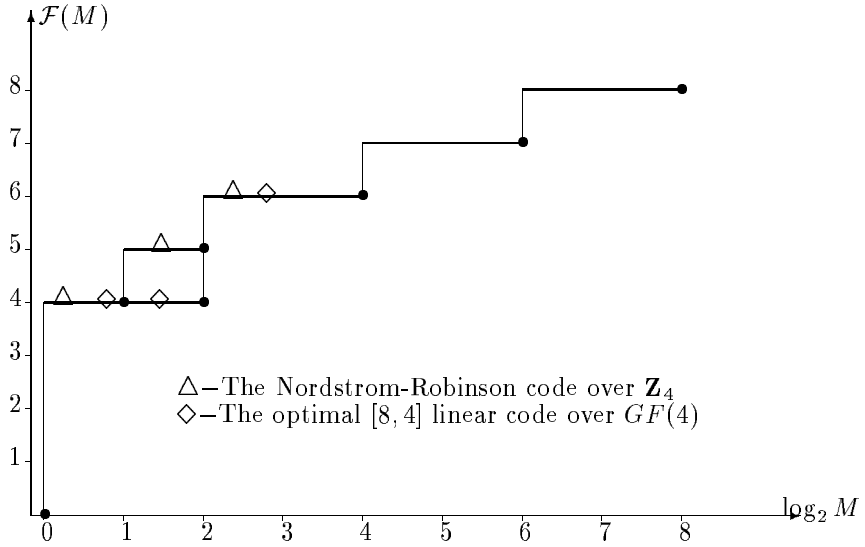
Proof: The Nordstrom-Robinson code is a self-dual code with $d_1(\mathcal{NR}) = 4$. According to (15), $d_2(\mathcal{NR}) = 5$. From this fact and Theorem 2 it follows that $\{d_1(\mathcal{NR}), d_2(\mathcal{NR}), \ldots, d_8(\mathcal{NR})\} = \{4, 5, 6, 6, 7, 7, 8, 8\}$. The claim of the theorem follows from this hierarchy. ■

Remark. The optimal linear code, say B, over $GF(4)$ of length 8 and dimension 4 has $d_1(B) = 4$. The best possible function $\mathcal{F}_B(M)$ for such a code and the function $\mathcal{F}_{\mathcal{NR}}(M)$ have the following form (see the picture). One can see that $\mathcal{F}_B(3) = \mathcal{F}_B(4) = 4 < \mathcal{F}_{\mathcal{NR}}(3) = \mathcal{F}_{\mathcal{NR}}(4) = 5$. Thus the Nordstrom-Robinson over $\mathbf{Z}_4$ code is better, in some sense, than the optimal linear code over $GF(4)$ of the same length and dimension. From cryptographical point of view it means that if we use the code B on the wire tap channel the adversary must tap 4 symbols from transmitter to get 2 bits of information, whereas in the case of using the Nordstrom-Robinson code the adversary must tap at least 5 symbols to get 2 bits of information. Whether there exist other GR linear codes (or codes over another ring or additive group) that are better (from the point of view of GHW) than optimal linear codes over $GF(q)$ is an open question.

THEOREM 13 *The Kerdock code of length 16 over $\mathbf{Z_4}$ has weight hierarchy*

$$\{\delta_1, \ldots, \delta_6\} = \{8, 10, 12, 14, 15, 16\}; \{M_1, \ldots, M_6\} = \{2, 2^2, 2^4, 2^6, 2^8, 2^{10}\}.$$

Proof: From (15) and Corollary 5 it follows that $\{d_1(\mathcal{K}_4), d_2(\mathcal{K}_4), d_3(\mathcal{K}_4)\} = \{8, 10, 12\}$. Let $\mathcal{P}_4 = \mathcal{K}_5^{\perp}$ be the Preparata code over $\mathbf{Z}_4$ of length 16. Using the



MacWilliams identities for $\mathbf{Z}_4$−linear codes [6], we see that in $\mathcal{P}_5$ there exists a codeword $\mathbf{a}$ such that $2\mathbf{a} \neq \mathbf{0}$ and $|\chi(\mathbf{a})| = 4$. Using this fact and taking into account that minimum Hamming weight of $\mathcal{P}_4$ equals 4, we get $d_1(\mathcal{P}_4) = 4$ and $d_2(\mathcal{P}_4) = 4$. Now using Theorem 2, we get the hierarchy of $\mathcal{K}_4$. ■

Remark. Note that if we start with known values $d_1(\mathcal{K}_m)$ and $d_2(\mathcal{K}_m), m = 3, 4$ then in (10) and (11) we have equality for all $d_{s_1,0}$ and $d_{s_1,1}$.

Using Theorems 2 and 13, we can also find weight hierarchy of the Preparata code $\mathcal{P}_4$

$$\{d_1(\mathcal{P}_4), \ldots, d_{22}(\mathcal{P}_5)\} = \{1, 1, 2, 2, \ldots, 32, 32\} \setminus \{1, 1, 2, 2, 3, 3, 5, 5, 7, 9\}.$$

THEOREM 14 *The Kerdock code of length* 32 *over* $\mathbf{Z_4}$ *has weight hierarchy*

$$\{\delta_1, \ldots, \delta_9\} = \{16, 22, 26, 27, 28, 29, 30, 31, 32\};$$

$$\{M_1, \ldots, M_9\} = \{2, 2^2, 2^3, 2^4, 2^5, 2^6, 2^8, 2^{10}, 2^{12}\}.$$

Proof: From (15) and Theorem 7 it follows that $\{d_1(\mathcal{K}_5), d_2(\mathcal{K}_5), d_3(\mathcal{K}_5)\} = \{16, 22, 26\}$. Let $\mathcal{P}_5 = \mathcal{K}_5^{\perp}$ be the Preparata code over $\mathbf{Z}_4$ of length 32. Using the MacWilliams identities for $\mathbf{Z}_4$−linear codes [6], we see that in $\mathcal{P}_5$ there exists a codeword $\mathbf{a}$ such that $2\mathbf{a} \neq \mathbf{0}$ and $|\chi(\mathbf{a})| = 5$. Using this fact and taking into account that minimum Hamming weight of $\mathcal{P}_5$ equals 4, we get $d_1(\mathcal{P}_5) = 4$ and $d_2(\mathcal{P}_5) = 5$. Now using Theorem 2, we get $\{d_5(\mathcal{K}_5), \ldots, d_{12}(\mathcal{K}_5)\} = \{28, 29, 30, 30, 31, 31, 32, 32\}$. From Theorem 6 it follows that $d_4(\mathcal{K}_5) = 27$ or 28. Suppose $d_4(\mathcal{K}_5) = 28$; then $d_2(\mathcal{P}_5) > 5$. A contradiction. Thus $d_4(\mathcal{K}_5) = 27$. The claim of the theorem follows from the hierarchy of the values $d_i(\mathcal{K}_5)$. ■

Using Theorems 9 and 2, we can also find weight hierarchy of the Preparata code $\mathcal{P}_5$

$$\{d_1(\mathcal{P}_5), \ldots, d_{52}(\mathcal{P}_5)\} = \{1, 1, 2, 2, \ldots, 32, 32\} \setminus \{1, 1, 2, 2, 3, 3, 4, 5, 7, 9, 17\}.$$

References

1. A. Ashikhmin,"Generalized Hamming Weights for $\mathbf{Z_4}$-Linear Codes," in: Proc. *the 1994 IEEE International Symposium on Information Theory*, Trondheim, Norway, June 27 - July 1, 1994, p. 306,
2. A. Ashikhmin, "On Generalized Hamming Weights for Galois Ring Linear Codes," in *Proc. Seventh Joint Swedish-Russian Int. Workshop on Information Theory* June 17-22, 1995, St.-Petersburg, Russia, pp. 20–24.
3. A. Ashikhmin, "On Generalized Hamming Weights for Galois Ring Linear Codes," *Proc. Mediterranean Workshop on Coding and Information Integrity*, February, 1996.
4. L. A. Bassalygo, "Supports of a Code," in Proc. Int. Symp. Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, *Lecture Notes in Computer Science,* 948, (1995) Springer Verlag, pp.1–3.
5. G. Cohen, S. Litsyn, and G. Zemor, "Upper Bounds on Generalized Distances," *IEEE Trans. Inf. Theory.*, IT–40, no.6, (1994) pp. 2090–2092.
6. A. R. Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The $\mathbf{Z_4}$-Linearity of Kerdock, Preparata, Goethals and Related Codes," *IEEE Trans. Inform. Theory*, IT–40, (1994) pp.301–319.
7. T. Helleseth, T. Kløve, and J.Mykkelveit, "The Weight Distribution of Irreducible Cyclic Codes with Block Lengths $n_1((q^l - 1)/N)$," *Discrete Math.,* vol. 18, (1977) pp. 179–211.
8. T. Helleseth, T. Kløve. and Ø. Ytrehus, "Generalization of the Griesmer Bound," *Lecture Notes in Computer Science*, vol. 829, (1993) pp. 41–52.
9. F. J. K. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977.
10. A. A. Nechaev, "Kerdock Code in a Cyclic Form," *Diskret. Mat.,* vol. 1, no.4, pp.123–139, 1989 (in Russian). English translation in *Discrete Math. Appl.*, vol. 1, (1991) pp. 365–384.
11. L. H. Ozarow and A. D. Wyner, "Wire-Tap Channel II," *AT&T Bell Labs. Tech. J.*, vol. 63, (1984) pp. 2135–2157.
12. J. Simonis, "The Effective Length of Subcodes," *Applicable Algebra in Engineering, Communication and Computing*, no.5, (1994) pp. 371–377.
13. K. Yang, T. Hellethseth, P. V. Kumar, and A. G. Shanbhag, "On the Weight Hierarchy of the Kerdock Codes Over $\mathbf{Z}_4$," *IEEE Trans. Inform. Theory*, IT–42, (1996) pp.1587–1593.
14. V. K. Wei, "Generalized Hamming Weights for Linear Codes," *IEEE Trans. Inform. Theory*, IT–37, no.5, (1991) pp. 1412–1418.