Generalized Hamming Weights of Nonlinear Codes and the Relation to the Z_4-Linear Representation

Ilan Reuven, *Student Member, IEEE,* and Yair Be'ery, *Member, IEEE*

***Abstract*—In this correspondence, we give a new definition of generalized Hamming weights of nonlinear codes and a new interpretation connected with it. These generalized weights are determined by the entropy/length profile of the code. We show that this definition characterizes the performance of nonlinear codes on the wire-tap channel of type II. The new definition is invariant under translates of the code, it satisfies the property of strict monotonicity and the generalized Singleton bound. We check the relations between the generalized weight hierarchies of Z_4-linear codes and their binary image under the Gray map. We also show that the binary image of a Z_4-linear code is a symmetric, not necessarily rectangular code. Moreover, if this binary image is a linear code then it admits a twisted squaring construction.**

***Index Terms*—Entropy/length profiles, generalized Hamming weights, linear codes over Z_4, nonlinear codes, twisted squaring construction.**

I. Introduction

Motivated by applications in keyless cryptography, Wei [26] defined the *generalized Hamming weight* (GHW) *hierarchy* of linear block codes. These weights characterize the performance of the code on the type II wire-tap channel. After Wei's study this subject received considerable interest, and the weight hierarchy of several classes of codes has been derived in numerous recent works (e.g., [6], [9], and [27]). Many other works introduced bounds on the generalized weights (e.g., [13] and [24]). Kasami *et al.* [15] showed that bounds on the state complexity profile of a trellis representation for a linear block code can be expressed in terms of the GHW. Forney [11] elaborated on this observation and called the GHW the *length/dimension profile* of the code C, $\boldsymbol{m}(C)$, whose component $m_j(C)$ is the minimum *effective length* of any subcode of C whose dimension is j, $0 \leq j \leq k$. The inverse profile, the *dimension/length profile* (DLP), $\boldsymbol{k}(C)$, which contains the same information about C, has been found useful for the study of trellis complexity of linear block codes. The jth component of the latter profile, $k_j(C)$, is the maximum dimension of any subcode of C whose effective length is not larger than j. A coordinate ordering of a linear block code that meets the DLP bound is called *efficient,* and the code for which such an efficient ordering do exists is said to satisfy the *two-way chain condition.*

Two proposals for generalization of the GHW notion to nonlinear codes have been suggested by Bassalygo [3] and by Cohen *et al.* [7]. Both definitions view the GHW as a certain minimum property of subcodes of C of a predefined cardinality. However, for nonlinear codes these subcodes do not represent the amount of information contained in the corresponding coordinates of the code, though both definitions coincide with the well-known definition when applied to linear codes. In this study, we suggest a new definition of the GHW for nonlinear codes. This definition seems to be a natural generalization of Wei's definition to nonlinear codes. The new definition has two main advantages over the previous ones. First, it fits in with Wei's framework. The new definition quantifies the minimum number of information bits (or q-ary symbols) contained in any, say t, symbols of a codeword. Second, the new definition generalizes the different dimension/length profiles to nonlinear codes by introducing the *entropy/length profile* (ELP). The relations between the different dimension/length profiles do exist also between the corresponding entropy/length profiles. Likewise, the new profiles can be used in a similar way as in [11] to investigate and to lower-bound various complexity measures of a trellis representation of nonlinear codes. The latter use of our new approach has been addressed in our previous work [20]. In the present work we hence concentrate on Wei's cryptographical viewpoint.

This study comprises two closely related parts. First, we present our new definition of generalized Hamming weights of nonlinear codes. We study the properties of the new notion and its relations to the other definitions. Likewise, we develop some bounds on these weights and illustrate their use in some examples. In the second part, we discuss the representation of binary nonlinear codes as Z_4-linear codes and the relations between this representation and its binary image. Hammons *et al.* [12] showed that some known classes of nonlinear codes such as Kerdock, Preparata, etc., can be constructed as binary images under the Gray map of linear codes over Z_4. The generalized Hamming weights of these codes over Z_4 have been considered in several recent studies [1], [2], [28]–[30]. These weights characterize the performance of the Z_4-linear codes over the wire-tap channel of type II. These studies have utilized the relation between the generalized weight hierarchy of dual codes. Recently, these codes have also been studied as binary nonlinear codes [21], [22]. It is easily verified that the GHW hierarchy of the codes over Z_4 can be used to lower-bound the ELP of the respective binary codes. We use the results of the above-referenced works to determine and to bound the ELP of the binary $[64, 2^{12}, 28]$ $\mathcal{K}(6)$ Kerdock, and $[64, 2^{52}, 6]$ $\mathcal{P}(6)$ Preparata codes, and to determine the first five generalized Hamming weights of the $\mathcal{G}(m)$ Goethals codes. Additionally, we show that the binary image under the Gray map of a linear code over Z_4 is a symmetric code (either when this image is linear or not). We show that when this binary image generates a linear code then it is equivalent to a twisted squaring construction code. Likewise, we give a counterexample to show that the binary image of a Z_4-linear code need not be a rectangular code.

The correspondence is organized as follows. Preliminaries and basic notations are presented in Section II. In Section III, we give our new definition for GHW of nonlinear codes and discuss the properties of this notion and the differences between our definition and other definitions ([3], [7]) of this notion. We also develop some bounds on the ELP which is the inverse of the GHW (according to our definition). In Section IV, we study the properties of the binary image of Z_4-linear codes and the relation between the generalized weight hierarchy of the code in these two representations. As an example we investigate the GHW of the binary $\mathcal{K}(6)$ Kerdock code, the $\mathcal{P}(6)$ Preparata code, and the $\mathcal{G}(m)$ Goethals codes. For the first two codes, we give the ELP for some indices and lower and upper bounds on this measure for the remaining indices.

II. Preliminaries and Notations

Let C be an $[n, k]$ linear code. The support of the subcode D, $D \subseteq C$, which is denoted by $\chi(D)$ is defined as follows:

$$\chi(D) \triangleq \{i: \exists (c_1, c_2, \cdots, c_n) = \boldsymbol{c} \in D, c_i \neq 0\}.$$

Manuscript received January 28, 1998; revised September 2, 1998.

The authors are with the Department of Electrical Engineering–Systems, Tel-Aviv University, Ramat-Aviv 69978, Tel-Aviv, Israel (e-mail: ybeery@eng.tau.ac.il).

Communicated by T. Kløve, Associate Editor for Coding Theory.
Publisher Item Identifier S 0018-9448(99)01387-5.

The rth generalized Hamming weight of a linear code C, $d_r(C)$, is the minimum support of any r-dimensional subcode of C, namely,

$$d_r(C) \triangleq \min_D\{|\chi(D)|\colon D \subseteq C, \dim(D) = r\}, \qquad 1 \le r \le k.$$

The set $\{d_1, d_2, \cdots, d_k\}$ is called the (generalized Hamming) weight hierarchy of C. In the appendix of [26], the connection between generalized Hamming weights and the security of the code when transmitted over the wire-tap channel of type II was explained. The suggested scheme used the $[n, n-k]$ linear code $C^{\perp}$: the encoder selects a coset according to the k information bits and transmits a codeword randomly chosen from that coset. The adversary has full knowledge of the code, but not of the random selection of a codeword in the coset. The adversary is allowed to tap s symbols from the transmitter. The drops of the adversary's *equivocation* curve for this scheme occur at the generalized weights of the $[n, k]$ code. That is, the adversary obtains r q-ary information symbols if and only if $s \ge d_r(C)$. The following lemma gives a new interpretation to the GHW hierarchy of a linear code. It refers to the adversary's equivocation when the codewords of the $[n, k]$ code, and not its dual, are transmitted over the channel. This interpretation of the GHW is useful to the study of nonlinear codes that do not have dual codes, and hence the above transmission scheme which is based on cosets of the code is not applicable to nonlinear codes.

Lemma 1: Let C be an $[n, k]$ linear code over GF(q) with $d_r(C) = t$, then $k - r$ information symbols can be reconstructed from any $n - t$ coordinates of a codeword of C.

Proof: Without loss of generality we assume that the first $n-t$ components of the codeword $\boldsymbol{c}$ are given. The generator matrix of C is equivalent to the following matrix:

$$\begin{bmatrix} A & B \\ 0 & C_t \end{bmatrix}$$

where the matrices C_t and B comprise t columns, $[0 \quad C_t]$ is a generator matrix of the largest subcode of C whose first $n - t$ components are all zeros. Clearly, the rank of A is identical to the total number of rows in it, i.e., its rank is at least $k-r$. We denote by l the rank of A. Thus there is a unique solution $(m_1, m_2, \cdots, m_l)$ to the equation set

$$(m_1, m_2, \cdots, m_l) \cdot A = (c_1, c_2, \cdots, c_{n-t})$$

for any $(n - t)$-tuple $(c_1, c_2, \cdots, c_{n-t})$ which comprises the first $(n - t)$ components of a codeword $\boldsymbol{c} \in C$. □

Consequently, we conclude that an adversary who taps any $n - t$ symbols of a codeword of C with $d_r(C) = t$ can gain at least $k - r$ information symbols. For example, a *maximum distance separable* code C has the worst immunity. This code meets the *generalized Singleton bound* $d_r(C) = n - k + r$. An adversary that listens to l components of the codewords gains the maximum possible information, namely, $\min(l, k)$ information symbols.

In the sequel we denote by $C[n, M]$ a block code that consists of M codewords of length n over a finite alphabet set of size q. When the minimum distance of the code, d, is of interest, the code will be denoted by the triple $[n, M, d]$. When we refer to an $[n, k]$ linear code, k will denote the dimension of the code. Likewise, we use parentheses to denote the pair (n, d), the length of the code, and its minimum distance and by $A(n, d)$ the largest number M of codewords in any binary (n, d) code. For nonlinear codes, it is more convenient to use a different definition of the support of a subcode. We will denote this measure by $\mathrm{supp}(D)$. We define $\mathrm{supp}(D)$ as follows:

$$\mathrm{supp}(D) \triangleq \{i\colon \exists (c_1, c_2, \cdots, c_n) = \boldsymbol{c} \in D \text{ and } (b_1, b_2, \cdots, b_n) = \boldsymbol{b} \in D \text{ with } c_i \ne b_i\}.$$

Let I be the index set for C, $I = \{1, 2, \cdots, n\}$ and let $P_J(\boldsymbol{c})$ denote the projection of a codeword $\boldsymbol{c} \in C$ onto $J \subseteq I$. The set of the projections of all codewords onto J is denoted by $P_J(C)$, and the complementary set of J in I will be denoted by $I - J$. For a fixed $i \in [1, n-1]$ we split a codeword $\boldsymbol{c} = (c_1, c_2, \cdots, c_n) \in C$ to its *past* portion or head $(c_1, c_2, \cdots, c_i)$ and its *future* portion or *tail* $(c_{i+1}, c_{i+2}, \cdots, c_n)$. The entropy of a random variable X is denoted by $H(X)$. All the logarithms here onwards are taken to base q, the cardinality of the alphabet set over which the code is defined.

III. Generalized Hamming Weights of Nonlinear Codes

We start with our generalization of the GHW concept to nonlinear block codes. Similarly to [20], we make the $[n, M]$ code C into a uniform probability space by assigning each codeword a probability of $1/M$. We denote by X_J a random $|J|$-tuple variable that takes on the values of the set $P_J(C)$ with probabilities that are induced by the uniform distribution of the codewords.

Definition 1 [20]: The *conditional entropy/length profile (conditional* ELP) of a code $C[n, M, d]$ over a finite alphabet set of size q is the sequence $\boldsymbol{h}(C) = \{h_i(C), 0 \le i \le n\}$ where

$$h_i(C) \triangleq \max_J\{H(X_J|X_{I-J})\colon |J| = i\}, \qquad 0 \le i \le n \tag{1}$$

with the convention $h_0(C) \triangleq 0$.

This definition extends the notion of DLP of [11] to nonlinear codes, and it coincides with it when applied to linear codes. Likewise, $0 \le h_{i+1}(C) - h_i(C) \le 1$, and $h_l(C) > 0$ if and only if $l \ge d$. The DLP of linear codes rises from 0 to the code's dimension k in k distinct unit steps. The increase of the ELP of nonlinear codes need not be in unit steps. Thus the total number of steps may exceed $\log_q M$.

Definition 2: The *generalized Hamming weight (distance)* of order r, $d_r^1(C)$, of the code $C[n, M, d]$ over a finite alphabet set of size q will be defined as the index of the rth nonzero value of the sequence $\boldsymbol{h}(C)$. That is, let L_r be the largest index for which $h_{L_r}(C) = k_r$, then $d_{r+1}^1(C) = L_r + 1$. In other words, the GHW hierarchy according to our definition is the set of values $\{l\colon h_l(C) \ne h_{l-1}(C)\}$.

The first value of the generalized Hamming weight hierarchy is the index of the first nonzero value of the sequence $\boldsymbol{h}(C)$, namely, the minimum distance of the code d. This definition reduces to Wei's definition when applied to linear codes. Lemma 1 can now be generalized to this definition: Let C be an $[n, M, d]$ block code over a finite alphabet set of size q with $d_r^1(C) = t$, then the entropy of any $n - t$ coordinates of C is at least $\log_q M - h_t(C)$ out of total $\log_q M$ information symbols.

A. Relations Between the Different Definitions

A different definition of the GHW of binary nonlinear codes was given by Cohen *et al.* [7].

$$d_r^2(C) \triangleq \min_{(\boldsymbol{t}, \boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_r) \subseteq C} \left| \bigcup_{j=1}^{r} \chi(\boldsymbol{c}_j + \boldsymbol{t}) \right| \tag{2}$$

such that the r vectors $\{(\boldsymbol{c}_1 + \boldsymbol{t}), (\boldsymbol{c}_2 + \boldsymbol{t}), \cdots, (\boldsymbol{c}_r + \boldsymbol{t})\}$ are linearly independent.

This definition has two prominent drawbacks. First, it is not a combinatorial definition for a combinatorial object (a nonlinear code). This definition depends on the field, ring, etc., over which the code is defined. The above definition applies only to binary codes, and different dependency relations should be checked in other cases. Likewise, this definition does not satisfy the property of strict monotonocity of the definition of GHW for linear codes. Another

definition, which is not subject to these deficiencies was given by Bassalygo [3]. This definition is related to the *cardinality/length profile* (CLP) notion of [17]. The CLP of a nonlinear code is defined as the sequence $\{\log_q M(1;C), \log_q M(2;C), \cdots, \log_q M(n;C)\}$ where

$$M(l;C) \triangleq \max_{D}\{|D|\colon D \subseteq C, |\mathrm{supp}\,(D)| \le l\}. \tag{3}$$

The CLP also reduces to the DLP when applied to linear codes. From (1) and (3) we have the obvious relation between the CLP and the ELP

$$\log_q M(l;C) \ge h_l(C). \tag{4}$$

Let L_r be the largest index for which $M(L_r;C) = k_r$, then if we define $d^3_{r+1}(C) = L_r + 1$ we achieve the hierarchy of GHW of [3]. Thus the GHW hierarchy of the latter reference is the set of values $\{l\colon M(l;C) \neq M(l-1;C)\}$. It is noteworthy to comment that $M(d_j;C)$ is denoted by M_j in [3]. Also, the index of the generalized weight according to our definition and that of [3] does not reflect the cardinality of the underlying subcode, unlike the definition of [7] and the definition of the GHW of linear codes. Both the ELP and the CLP satisfy $h_{d^1_j}(C) \le j$, $\log_q M(d^3_j;C) \le j$ where d^1_j and d^3_j stand for the respective definition of the jth generalized weight. The CLP also satisfies $M(d^3_j;C) \ge j+1$. This relation does not necessarily hold for the ELP. For both definitions of GHW of nonlinear codes $d^1_r(C)$ and $d^3_r(C)$ *strict* inequalities hold between successive values, $d^1_r(C) > d^1_{r-1}(C)$ and $d^3_r(C) > d^3_{r-1}(C)$. Likewise, using [20, Lemma 3] it is straightforward to show that our definition satisfies the *generalized Singleton bound.*

Lemma 2: The ELP of an $[n, M]$ code C satisfies the following inequality: $h_{n-\log_q M+r}(C) \ge r$.

Using (4) we also have $\log_q M(n-\log_q M+r;C) \ge r$. All three definitions coincide with the well-known definition of GHW when applied to linear codes. Likewise, under all three definitions of the GHW, a code and its translate have the same hierarchy of generalized weights. However, the definitions of [3] and [7] are *local* definitions that do not average the total amount of information contained in a partial set of the coordinates of the code. The total number of generalized weights $d^2_r(C)$ is equal to the dimension of the linear code generated by the linear span of the codewords of C. The total number of distinct generalized weights $d^1_r(C)$ and $d^3_r(C)$ is lower-bounded by $\log_q M$ and upper-bounded by $n-d+1$. In general, the total number of generalized weights according to the three different methods is different.

We illustrate the difference between the definitions by calculating the GHW of the Nordstrom–Robinson code according to the three definitions.

Example 1: The Nordstrom–Robinson code $\mathcal{N}_{16}$ is a $[16, 256, 6]$ nonlinear code. The profiles at the bottom of this page are the GHW hierarchy according to our definition and the corresponding ELP values (see the first list at the bottom of this page). The hierarchy of generalized weights of the code according to [7] is shown in the second listing at the bottom of this page. The next (final) lists on the bottom of this page give the sequence of GHW of the code according to the definition of [3] and the corresponding CLP values.

The generalized weight hierarchies $d^3_r(\mathcal{N}_6)$ and $d^1_r(\mathcal{N}_{16})$ may be deduced from [17] and [20]. In order to derive $d^2_r(\mathcal{N}_{16})$ we list the codewords as a union of eight cosets of the $[16, 5, 8]$ Reed–Muller code. We take the following eight coset representatives:

0000	0000	0000	0000	0001	0001	0111	1000
0000	0011	0101	0110	0001	0010	0001	1101
0000	0101	0110	0011	0001	0100	0100	1110
0000	0110	0011	1010	0001	0111	0010	0100

and use the standard bit order of the $[16, 5, 8]$ Reed–Muller code, i.e., we generate the code by the generator matrix

$$\mathrm{G}_{\mathrm{RM}(1,4)} = \begin{bmatrix} 1111111111111111 \\ 0000000011111111 \\ 0000111100001111 \\ 0011001100110011 \\ 0101010101010101 \end{bmatrix}.$$

We consider the following six codewords:

1010	1100	1001	0000	0110	1010	1100	0000
1110	1110	1000	0111	0010	1110	1101	1110
0010	1000	1000	1101	1110	1000	1101	1011

The first three codewords (left-hand side of the list) determine the second generalized weight according to [7], and all six codewords determine the next three generalized weights. The support of the eight coset representatives is 11, and the rank of their linear span is 6 and thus $d^2_6(\mathcal{N}_{16}) = 11$. The next generalized weight is determined by the coset representatives (actually, seven out of the eight representatives) and the second codeword in $G_{\mathrm{RM}\,(1,4)}$, and these codewords in addition to the third codeword of $G_{\mathrm{RM}\,(1,4)}$, determine $d^2_8(\mathcal{N}_{16}) = 13$. The next two generalized weights are determined by this set of codewords in addition to the fourth, and the fourth and fifth codewords of $G_{\mathrm{RM}\,(1,4)}$. The total number of generalized weights according to the definition of [7] is 11—the linear span of $\mathcal{N}_{16}$ is the $[16, 11]$ second-order Reed–Muller code.

The maximum cardinality of any subcode of $\mathcal{N}_{16}$ whose support is 6, 7, and 8 is 2. However, for $l = 8$, the codewords of $\mathcal{N}_{16}$ can be divided to 128 pairs such that each pair has eight common components at the same eight coordinates of the code, and for $l = 6$ and 7 one can find only 32 pairs with ten or nine common components, whereas the 192 remaining codewords have all distinct components at these coordinates. Thus the approach of [3] and [17] does not distinguish $l = 8$ as a generalized Hamming weight. Our approach, however, takes account of the fact that a further amount of information can be gained by intercepting nine components of the code instead of eight. Using Wei's terminology: the adversary's equivocation is different in these two cases.

r:	1	2	3	4	5	6	7	8	9	10
$d^1_r(\mathcal{N}_{16})$:	6	8	9	10	11	12	13	14	15	16
$h_{d^1_r}(\mathcal{N}_{16})$:	0.25	1	$\frac{3}{4}\log_2 3 = 1.19$	2.19	3	4	5	6	7	8

r:	1	2	3	4	5	6	7	8	9	10	11
$d^2_r(\mathcal{N}_{16})$:	6	9	10	10	10	11	12	13	14	15	16

r:	1	2	3	4	5	6	7	8	9
$d^3_r(\mathcal{N}_{16})$:	6	9	10	11	12	13	14	15	16
$\log_2 M(d^3_r(\mathcal{N}_{16});\mathcal{N}_{16})$:	1	$\log_2 3$	$\log_2 6$	3	4	5	6	7	8

B. Bounds on the Entropy/Length Profile

Some bounds on measures of trellis complexity and on the generalized weights (according to the definition of [3]) were derived by Muder [19] and Lafourcade and Vardy [17]. These bounds are usually expressed in terms of the maximum size of a code with a given minimum distance. These bounds do not consider the impact of asymmetrical structures in the examined code. Clearly, any upper bound on the CLP is also an upper bound on the ELP at the same level. However, the derivation of tighter bounds on trellis complexity of nonlinear codes entails the evaluation of the ELP or bounding this profile. We recall that this profile describes the inverse function to that determined by the GHW hierarchy. The derivation of bounds on the ELP rather than the GHW seems to be a reasonable approach due to the fact that the generalized weights were defined as the set of indices at which the ELP increases. The bounds of [3] were also derived not on the GHW hierarchy but on the inverse profile (CLP) according to the respective definition. The formulation of bounds on the ELP requires more delicate combinatorial considerations. In this section, we derive some bounds on the ELP using the basic parameters of the code $[n, M, d]$. We show that these bounds improve upon the CLP-based bounds.

Theorem 3: Let C be an $[n, M, d]$ binary code. The value of the ELP at level i, $d \le i < \lceil 1.5d \rceil$ is bounded by

$$h_i(C) \le \frac{2A(n-i, \lceil d/2 \rceil - (i-d))}{M}. \tag{5}$$

Proof: Without loss of generality we assume that $h_i(C) = \max_J \{H(X_J|X_{I-J}): |J| = i\}$ is achieved for $J = [1, i]$. Clearly, there are at most two codewords with the same length-$(n-i)$ tail for indices i in the range $d \le i < \lceil 1.5d \rceil$. However, it may happen that not all length-$(n-i)$ tails have two distinct heads. Let us examine the cardinality of the largest subset D of C such that the tail of any codeword of D appears in two codewords of C(and D). For any four codewords which are partitioned to past and future portions at index i, $\{(\boldsymbol{x_1}, \boldsymbol{y_1}), (\boldsymbol{x_2}, \boldsymbol{y_1}), (\boldsymbol{x_3}, \boldsymbol{y_2}), (\boldsymbol{x_4}, \boldsymbol{y_2})\} \subseteq D$, we have

$$\text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_3}) + \text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_4}) \le 2i - \text{dist}\,(\boldsymbol{x_3}, \boldsymbol{x_4}) \le 2i - d.$$

Consequently, either $\text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_3})$ or $\text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_4})$ (or both) is smaller than $\lfloor \frac{1}{2}(2i-d) \rfloor$

$$\text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_3}) \le \lfloor d/2 \rfloor + (i-d) \quad \text{and/or}$$
$$\text{dist}\,(\boldsymbol{x_1}, \boldsymbol{x_4}) \le \lfloor d/2 \rfloor + (i-d).$$

Thus we have $\text{dist}\,(\boldsymbol{y_1}, \boldsymbol{y_2}) \ge \lceil d/2 \rceil - (i-d)$ and

$$|D| \le 2A(n-i, \lceil d/2 \rceil - (i-d))$$

which completes the proof. □

Example 2: For the Nordstrom–Robinson code $\mathcal{N}_{16}$ we have $\log_2 M(6; \mathcal{N}_{16}) = 1$. However, using Theorem 3 with the bound $A(10, 3) = 79$ [4], we have $h_6(\mathcal{N}_{16}) \le 79/128$. This value can be used to improve the bounds on trellis complexity of the code. In [20] it was found that the actual value of $h_6(\mathcal{N}_{16})$ is 0.25.

Example 3: The Kerdock code $\mathcal{K}(6)$ has the parameters $[64, 2^{12}, 28]$. Using the fact that the best known code with the parameters $(n, d) = (36, 14)$ comprises 2^9 codewords [18] we obtain $h_{28}(\mathcal{K}(6)) \le 0.25$ and $h_{29}(\mathcal{K}(6)) \le 0.25$ (postulating that cross sections of this code do not provide a better code than the linear code with parameters $[36, 9, 14]$). Both the constructions of Shany *et al.* [22] and Yang *et al.* [30] provide the trivial bound $h_{28}(\mathcal{K}(6)) > 0$, $h_{29}(\mathcal{K}(6)) > 0$.

Example 4: For n which is a multiple of 4, the (nonlinear) $\mathcal{B}_n$ and the $\mathcal{C}_n$ Hadamard codes have the parameters $[n-1, 2n, \frac{1}{2}n-1]$ and $[n, 2n, \frac{1}{2}n]$, respectively. It can be verified that for n values which are not divisible by 8 we have

$$h_{n/2-1}(\mathcal{B}_n) \le \frac{1}{2} + \frac{2}{n} \qquad h_{n/2}(\mathcal{C}_n) \le \frac{1}{2} + \frac{2}{n}.$$

Clearly, the bound of Theorem 3 is useful only when it is smaller than 1. The smallest length of a binary code with an even minimum distance d that comprises more than two codewords is $1.5d$ and $A(1.5d, d) = 4$. The next lemma bounds the ELP value at the index $1.5d$.

Lemma 4: Let C be an $[n, M, d]$ binary code where d is an even integer, and let F be the set of the length-$(n-1.5d)$ tails of codewords of C such that each tail in this set is comprised in four codewords of C. The minimum distance of the set F is lower-bounded by $\lceil \frac{1}{4}d \rceil$.

Proof: Suppose that $\{\boldsymbol{y_1}, \boldsymbol{y_2}\} \subseteq F$ and that

$$\{(\boldsymbol{x_{1i}}, \boldsymbol{y_1}), (\boldsymbol{x_{2i}}, \boldsymbol{y_2}), i = 1, 2, 3, 4\} \subseteq C.$$

We denote

$$d' = \min_{i=1,2,3,4;\ j=1,2,3,4} \text{dist}\,(\boldsymbol{x_{1i}}, \boldsymbol{x_{2j}}).$$

If $d' = 0$, then we have $\text{dist}\,(\boldsymbol{y_1}, \boldsymbol{y_2}) \ge d$. Otherwise,

$$\sum_{i=1}^{4}\sum_{j=1}^{4}[2 \cdot \text{dist}\,(\boldsymbol{x_{1i}}, \boldsymbol{x_{2j}}) + \text{dist}\,(\boldsymbol{x_{1i}}, \boldsymbol{x_{1j}}) + \text{dist}\,(\boldsymbol{x_{2i}}, \boldsymbol{x_{2j}})] \ge 24d + 32d'. \tag{6}$$

On the other hand, let L be the $8 \times 1.5d$ matrix whose rows are $\{\boldsymbol{x_{ij}}, i = 1, 2, j = 1, 2, 3, 4\}$. Suppose that the lth column of L comprises n_l 0's and $(8-n_l)$ 1's. Hence this column contributes $2n_l(8-n_l)$ to the sum on the left-hand side of (6), and the latter expression is maximized if $n_l = 4$. Thus

$$\sum_{i=1}^{4}\sum_{j=1}^{4}[2 \cdot \text{dist}\,(\boldsymbol{x_{1i}}, \boldsymbol{x_{2j}}) + \text{dist}\,(\boldsymbol{x_{1i}}, \boldsymbol{x_{1j}}) + \text{dist}\,(\boldsymbol{x_{2i}}, \boldsymbol{x_{2j}})] \le 1.5d \cdot 32 = 48d. \tag{7}$$

Combining (6) and (7) we have

$$d' \le \tfrac{3}{4}d.$$

Consequently,

$$\text{dist}\,(\boldsymbol{y_1}, \boldsymbol{y_2}) \ge \lceil \tfrac{1}{4}d \rceil, \qquad \forall \{\boldsymbol{y_1}, \boldsymbol{y_2}\} \subseteq F. \qquad \square$$

Thus the set F has a minimum distance of at least $\lceil \frac{1}{4}d \rceil$. The next theorem follows immediately.

Theorem 5: Let C be an $[n, M, d]$ binary code where d is an even integer. The value of the ELP at level $1.5d$ is upper-bounded by

$$h_{1.5d}(C) \le \min\left\{2, \log 3 + \frac{4A(n-1.5d, \lceil d/4 \rceil)}{M} \cdot \log\left(\tfrac{4}{3}\right)\right\}. \tag{8}$$

This bound is based on the observation that the code consists of at most $A(n-1.5d, \lceil d/4 \rceil)$ sets of four codewords with common elements on $(n-1.5d)$ positions while the remaining codewords may be partitioned to sets with at most three codewords with common bits at these $(n-1.5d)$ positions.

Example 5: When we apply Theorem 5 to the $[n, 2n, \frac{1}{2}n]\,\mathcal{C}_n$ Hadamard code we obtain

$$h_{3n/4}(\mathcal{C}_n) \le 2\left(\frac{1}{2}+\frac{4}{n}\right)+\left(\frac{1}{2}-\frac{4}{n}\right)\cdot \log_2 3, \quad \text{if } n \text{ is not divisible by } 16.$$

$$h_{3n/4}(\mathcal{C}_n) \le 2\left(\frac{1}{2}+\frac{2}{n}\right)+\left(\frac{1}{2}-\frac{2}{n}\right)\cdot \log_2 3, \quad \text{if } n \text{ is not divisible by } 8.$$

These results improve upon the trivial bound $h_{3n/4}(\mathcal{C}_n) \le 2$. Furthermore, one can elaborate on the above ideas and use some more subtle combinatorial considerations in order to achieve tighter bounds. Likewise, the prescribed bounds are not useful for codes with a relatively small minimum distance.

IV. Z_4-Linear Codes: GHW and the Relations to the Binary Image

Hammons *et al.* [12] have shown that some known families of nonlinear codes, such as Kerdock, Preparata, etc., can be constructed as binary images under the Gray map of linear codes over Z_4. The Kerdock and the Preparata codes are "formal duals" over the binary field in the sense that the weight distribution of one code is the MacWilliams transform of the weight distribution of the other. Moreover, these codes are dual over Z_4, and duality in the Z_4 domain implies that the binary images have dual weight distributions. The generalized Hamming weights of these codes over Z_4 have recently been considered in several works [1], [2], [28], [30]. These studies have utilized the relation between the generalized weight hierarchy of dual codes ([1], [2]). Recently, these codes have also been studied as binary nonlinear codes [21], [22]. In this section, we show that the binary image under the Gray map of a Z_4-linear code is a symmetric code. We also show that the binary image need not be a *rectangular* code [16]. Finally, we address the relation between the generalized weights of a code over Z_4 and those of its binary representation; in particular, we use the results of the above-referenced works to determine the ELP of the binary $[64, 2^{12}, 28]\mathcal{K}(6)$ and $[64, 2^{52}, 6]\,\mathcal{P}(6)$ codes for some indices and to derive bounds on the generalized weights for the remaining indices.

Let Z_4 be the ring of integers modulo 4. We say that two codes are *permutation-equivalent* if one can be obtained from the other by a coordinate permutation. A linear length-n code over Z_4 (quaternary code) is an additive subgroup of Z_4^n. Any quaternary linear code $\mathcal{C}$ is permutation-equivalent to a code with a generator matrix of the form [12]

$$\begin{bmatrix} I_{k_1} & A & B \\ 0 & 2I_{k_2} & 2D \end{bmatrix} \tag{9}$$

where A and D are binary matrices, B is a Z_4-matrix, and I_k denotes the $k \times k$ identity matrix. This code comprises $4^{k_1}2^{k_2}$ codewords. We define two maps β and γ from $Z_4 \to Z_2$ as follows:

c	$\beta(c)$	$\gamma(c)$
0	0	0
1	0	1
2	1	1
3	1	0

The Gray map ϕ: $Z_4 \to Z_2^2$ is given by

$$\phi(c) = (\beta(c), \gamma(c))$$

and the Gray map ϕ: $Z_4^n \to Z_2^{2n}$ is given by

$$\phi(\boldsymbol{c}) = (\beta(\boldsymbol{c}), \gamma(\boldsymbol{c})) \tag{10}$$

where

$$\beta(\boldsymbol{c}) = \beta((c_1, c_2, \cdots, c_n)) = (\beta(c_1), \beta(c_2), \cdots, \beta(c_n))$$

and, similarly,

$$\gamma(\boldsymbol{c}) = \gamma((c_1, c_2, \cdots, c_n)) = (\gamma(c_1), \gamma(c_2), \cdots, \gamma(c_n)).$$

The binary image of a quaternary code is the image $C = \phi(\mathcal{C})$ under the Gray map. A code of even length n is called *symmetric* if $(\boldsymbol{a}, \boldsymbol{b}) \in C$ implies $(\boldsymbol{b}, \boldsymbol{a}) \in C$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are $(n/2)$-tuples.

Lemma 6: The binary image of a Z_4-linear code is a symmetric code.

Proof: We denote by $\mathcal{C}$ the binary image of the Z_4-linear code C under the Gray map. Suppose that the binary codeword $(\boldsymbol{a}, \boldsymbol{b}) \in C$ is the image of a quaternary codeword $\boldsymbol{c} \in \mathcal{C}$, then the binary image of the quaternary codeword $3\boldsymbol{c} \in \mathcal{C}$ is $(\boldsymbol{b}, \boldsymbol{a})$. □

Let S be a finite set and let S/T denote a partition of S into $N = |S/T|$ disjoint subsets $T_1, T_2, \cdots, T_N$, where $T_i \subset S$ for $i = 1, 2, \cdots, N$. The Cartesian product set denoted by $T_i \times T_{j_i}$ comprises all pairs (t_i, t_{j_i}) such that $t_i \in T_i$ and $t_{j_i} \in T_{j_i}$. The *twisted squaring construction* (TSC) [10] is defined as the union U of N sets $T_i \times T_{j_i}$ such that both i and j_i run through $1, 2, \cdots, N$. We denote U by $\|S/T\|^2$.

Theorem 7: If the binary image under the Gray map of a Z_4-linear code is linear then there exists a coordinate ordering under which this binary image is a TSC code.

Proof: Consider the quaternary linear code $\mathcal{C}$ generated by

$$G = \begin{bmatrix} I_{k_1} & A & B \\ 0 & 2I_{k_2} & 2D \end{bmatrix}.$$

We further decompose the rows of G and carry out the following row operations. We find the maximum number k_{11} of Z_4-linearly independent codewords with binary entries only. Thus the cardinality of the Z_4-span of these codewords is $4^{k_{11}}$. We note that these codewords have the form $\boldsymbol{m}^T \cdot G$, where $\boldsymbol{m}$ is a binary vector. We denote these codewords by $[J_1 \quad J_2 \quad J_3]$, where the total number of columns in the submatrices J_1, J_2, and J_3 corresponds to the above partition of G. We append additional $k_{12} = k_1 - k_{11}$ quaternary (Z_4-linearly independent) rows to the new generator matrix. Obviously, J_1 is a full-rank matrix, and hence these additional k_{12} rows can be taken as k_{12} rows of $[I_{k_1} \quad A \quad B]$ from the original generator matrix such that the chosen rows are Z_4-linearly independent of the first k_{11} rows. We denote these rows by $[K_1 \quad K_2 \quad K_3]$. Clearly, K_1 and K_2 may be taken to be binary matrices.

Thus the code generated by G is equivalent to the code generated by the following generator matrix:

$$\tilde{G} = \begin{bmatrix} J_1 & J_2 & J_3 \\ K_1 & K_2 & K_3 \\ 0 & 2I_{k_2} & 2D \end{bmatrix}. \tag{11}$$

If $C = \phi(\mathcal{C})$ is linear, and $\mathcal{C}$ is defined by (11), then C has a generator matrix

$$G_C = \begin{bmatrix} J_1 & J_2 & J_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & J_1 & J_2 & J_3 \\ 0 & 0 & \beta(K_3) & K_1 & K_2 & \gamma(K_3) \\ K_1 & K_2 & \gamma(K_3) & 0 & 0 & \beta(K_3) \\ 0 & I_{k_2} & D & 0 & I_{k_2} & D \end{bmatrix}. \tag{12}$$

We denote

$$G_{S/T} \triangleq \begin{bmatrix} 0 & 0 & \beta(K_3) \\ K_1 & K_2 & \gamma(K_3) \\ 0 & I_{k_2} & D \end{bmatrix} \qquad G_T \triangleq [J_1 \quad J_2 \quad J_3] \tag{13}$$

to rewrite (12)

$$G_C = \begin{bmatrix} G_T & 0 \\ 0 & G_T \\ G_{S/T} & \tilde{G}_{S/T} \end{bmatrix} \tag{14}$$

where both $G_{S/T}$ and $\tilde{G}_{S/T}$ generate the same vector space $[S/T]$ which is a system of coset representatives of T in S, and G_T generates T. Thus the code C may be constructed as a union of $4^{k_{12}}2^{k_2}$ subsets $T_i \times T_{j_i}$, where $|T_i| = 2^{k_{11}}$ for all $i = 1, 2, \cdots, 4^{k_{12}}2^{k_2}$. □

Sidorenko *et al.* [23] have shown that the binary image of a Z_4-linear code under the Gray map is rectangular (e.g., [16]) at even indices when a symbol-by-symbol map is used. This map differs from the map of [12] which we adopt in our results. The following example shows that the map defined by (10) does not have this property.

Example 6: Consider the Z_4-linear code $\mathcal{C}$ generated by

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

This code comprises the codewords $\{1112, 1200, 1103\}$. The binary image of these codewords under the map defined by (10) is $\{0001\ 1111, 0100\ 1100, 0001\ 1100\}$. If we assume that the binary code is rectangular at the center level then it includes the codeword $0100\ 1111$. However, the corresponding quaternary codeword 1211 is not a codeword of $\mathcal{C}$.

Yet, the nonlinear binary image of a Z_4-linear code *may* be a rectangular code. For example, the Kerdock codes are rectangular under any coordinate permutation ([22], [23]).

We now revisit the issue of GHW and discuss the relations between the GHW of Z_4-linear codes and their binary image. The GHW hierarchy of Z_4-linear codes was defined by Ashikhmin [1]. These weights characterize the performance of the Z_4-linear codes over the wire-tap channel of type II. The rth generalized Hamming weight of a Z_4-linear code $\mathcal{C}$ is defined by

$$d_r(\mathcal{C}) \triangleq \min_D \{\, |\chi(D)| \colon D \text{ is a } Z_4\text{-linear subcode of } \mathcal{C},\ \log_4 |D| = r\},\ r = \{0.5, 1, 1.5, \cdots, \log_4 |C|\}. \tag{15}$$

Some binary codes can also be described as nonlinear codes over Z_4 [8]. The GHW hierarchy of the codes over Z_4 can be used to lower-bound the ELP of the codes over the binary field. That is, Let $C = \phi(\mathcal{C})$ be the binary image under the Gray map of (not necessarily linear) quaternary code $\mathcal{C}$. It is easily verified that

$$h_{2l}(C) \geq h_l(\mathcal{C}) = 2r, \qquad \text{for } d_r(\mathcal{C}) \leq l < d_{r+0.5}(\mathcal{C}) \tag{16}$$

$$h_{2l+1}(C) \geq h_l(\mathcal{C}) + 1, \qquad \text{when } h_{l+1}(\mathcal{C}) = h_l(\mathcal{C}) + 2 \tag{17}$$

where the logarithms in this inequality, as well as all the logarithms here onwards, are all taken to the base 2. These inequalities are justified by the fact that the support of the binary image of the set of quaternary codewords that determine each generalized weight upper-bounds the corresponding generalized weight of the binary code. Similar relations also exist between the CLP of both codes

$$\log M(2l; C) \geq \log M(l; \mathcal{C})$$
$$\log M(2l+1; C) \geq \log M(l; \mathcal{C}) + 1$$
$$\text{when } \log M(l+1; \mathcal{C}) = \log M(l; \mathcal{C}) + 2.$$

In the following theorems, we utilize these relations in conjunction with the results of [1], [2], [21], [22], [29], [30], and values of the function $A(n, d)$ or bounds on this function to calculate and bound the ELP of the binary Kerdock code $\mathcal{K}(6)$ and the Preparata code $\mathcal{P}(6)$. Using the same approach, we also determine the first five generalized weights of the $\mathcal{G}(m)$ Goethals codes.

Theorem 8: The entropy/length profile of the $[64, 2^{12}, 28]$ Kerdock code $\mathcal{K}(6)$ is determined/bounded as follows. When the value of the ELP is not known exactly we give two values, the first one is a lower bound and the second is an upper bound (see the bottom of this page).

Proof: All the lower bounds are deduced from the GHW hierarchy of the code over Z_4 [1] (which provides better bounds than those deduced from [22]). The upper bound for $i = 28, 29$ follows from Theorem 3. This bound relies on the parameters of the best known codes, and thus this entry in the table is denoted by an asterisk. The upper bounds and the exact values of the ELP for the indices 32–56 are derived from the Plotkin bound on $A(i, 28)$. The ELP for the indices $59 \leq i \leq 64$ is $i - 52$ since that the dual distance of the code d' is 6 and hence any set of $j \leq d' - 1$ columns of $\mathcal{K}(6)$ contains each j-tuple exactly $|\mathcal{K}(6)|/2^j$ times. Finally, the ELP value for the indices 57 and 58 is upper-bounded by the ELP value at $i = 59$. □

Theorem 9: The entropy/length profile of the $[64, 2^{52}, 6]$ standard Preparata code $\mathcal{P}(6)$ is determined/bounded as follows (see the bottom of this page).

Proof: All the lower bounds are deduced from the construction of [21]. The upper bounds for $10 \leq i \leq 16$ follow from the known values of $A(i, 6)$. The upper bounds for the indices $17 \leq i \leq 36$ are

i	1–27	28–29	30–31	32–41	42–43	44–46	47–51	52–53
$h_i[\mathcal{K}(6)]$	0	$^*[0, 0.25]$	$[0, 1]$	1	$[1, 2]$	2	$\left[2, 1 + \log \left\lfloor \frac{28}{56-i} \right\rfloor\right]$	$\left[3, 1 + \log \left\lfloor \frac{28}{56-i} \right\rfloor\right]$

i	54–55	56	57	58	59–64
$h_i[\mathcal{K}(6)]$	$\left[4, 1 + \log \left\lfloor \frac{28}{56-i} \right\rfloor\right]$	$[5, \log 112]$	$[5, 7]$	$[6, 7]$	$i - 52$

i	1–5	6–8	9	10	11	12	13–14	15–16
$h_i[\mathcal{P}(6)]$	0	1	2	$[2, \log 6]$	$[3, \log 12]$	$[3, \log 24]$	$[i-9, i-8]$	$[i-10, i-8]$

i	17–20	21	22–24	25–32	33–34	35–36	37–64
$h_i[\mathcal{P}(6)]$	$^*[i-10, i-9]$	$^*[10, \log 2560]$	$^*[i-11, i-10]$	$^*(i-11)$	$^*[I-12, i-11]$	$^*(i-12)$	$i-12$

derived from the table of the best binary codes *known* ([18], and [5] for $i = 34$), and hence these entries are denoted by asterisks. The upper bound for the indices $37 \leq i \leq 64$ is due to the fact that the dual distance of this code is $d' = 28$. Actually the use of bounds on the maximum size of codes with a given length and minimum distance increases the upper bound by at most 1 for the indices 17–24, 32–34, 36, and by no more than 2 for the indices 25–31 and 35. □

The dual of the Z_4-linear Kerdock code $\mathcal{K}(m)$ is a code whose binary image is a $[2^m, 2^{2^m-2m}, 6]$ nonlinear code. This code has the same weight distribution as the standard Preparata code but it differs from the standard Preparata code. Thus the GHW of the Z_4-linear Preparata code cannot be used to bound the ELP of the standard binary Preparata code. Yet, the GHW hierarchy of the quaternary Preparata code as given in [28] provides looser bounds on the ELP relative to the bounds deduced from the construction of [21]. The profile of Theorem 8 can be used to evaluate lower bounds on the trellis complexity of the Kerdock code in either one of the two representations of this code. Theorem 9 can be used to bound the trellis complexity of the standard binary Preparata code.

For $m = 2t + 2 \geq 6$, the $\mathcal{G}(m)$ Goethals code is a nonlinear $[2^m, 2^{2^m-3m+1}, 8]$ code. The Z_4-linear Goethals codes also differ from the standard binary Goethals codes though they have the same weight distribution. Using values of $A(n, d)$, it can be shown that both the standard Goethals codes and the binary image of the Z_4-linear Goethals codes achieve the maximum possible value of the ELP for the first 16 indices, and this fact in turn gives the first five generalized weights of these codes. Hence the first five generalized weights of these codes as derived in [29] give the exact value of the ELP of the corresponding binary images (for the respective indices).

Theorem 10: Both the standard Goethals codes and the binary image of the Z_4-linear Goethals codes have the following entropy/length profile at indices $1 \leq i \leq 16$:

i	1–7	8–11	12–13	14	15	16
$h_i[\mathcal{G}(m)]$	0	1	2	3	4	5

Thus the first five generalized weights of these codes are $\{8, 12, 14, 15, 16\}$.

Proof: The Plotkin bound implies that the ELP of the Goethals codes for the indices 1–16 cannot be larger than indicated in the above table. Conversely, The first five generalized weights of the Goethals codes over Z_4 are $\{4, 6, 7, 8, 8\}$ [29]. Using (16) and (17), it is clear that the binary image of these codes meet the above ELP profile. Likewise, in [25] it is proved that the first five generalized weights of triple-error-correcting primitive binary BCH codes of length $2^m - 1$ are $\{7, 11, 13, 14, 15\}$. Using these values along with the construction of [21], it follows that the standard binary Goethals codes also meet the profile of the above table. The first five generalized weights of these codes are derived from this profile. □

Finally, we mention that there exists another definition of generalized weights, generalized Lee weights, for (not necessarily linear) codes over Z_4 [14]. This definition follows a similar approach as the one by Bassalygo [3] and extends it to Z_4 codes by checking Lee weights instead of Hamming weights. These generalized Lee weights of a Z_4 code coincide with the generalized Hamming weights of its binary image under the definition of [3].

ACKNOWLEDGMENT

The authors wish to thank G. Cohen for directing them to [3] and [7]. The discussions with him while the second author visited ENST, Paris, were the origin of the new definition of the GHW. The authors also wish to thank Y. Shany for helpful discussions.

REFERENCES

[1] A. E. Ashikhmin, "Generalized Hamming weights for Z_4-linear codes," in *Proc. IEEE Int. Symp. Information Theory* (Trondheim, Norway, June 27–July 1), 1994, p. 306.
[2] ——, "On generalized Hamming weights for Galois ring linear codes," *Designs, Codes Cryptogr.*, vol. 14, pp. 107–126, 1998.
[3] L. A. Bassalygo, "Supports of a code," in *Proc. Int. Symp. Applied Algebra, Algebraic Algorithms and Error-Correction Codes, AAECC-11, Lecture Notes on Computer Science*, no. 948, G. Cohen, M. Giusti, and T. Mora Eds. Paris, France: Springer-Verlag, 1995, pp. 1–3.
[4] M. R. Best, "Binary codes with a minimum distance of four," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 738–742, Nov. 1980.
[5] A. E. Brouwer and T. Verhoeff, "An updated table of minimum-distance bounds for binary linear codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 662–677, Mar. 1993.
[6] J. Cheng and C.-C. Chao, "On generalized Hamming weights of binary primitive BCH codes with minimum distance one less than a power of two," *IEEE Trans. Inform. Theory*, vol. 43, pp. 294–299, Jan. 1997.
[7] G. Cohen, S. Litsyn, and G. Zemor, "Upper bounds on generalized distances," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2090–2092, Nov. 1994.
[8] J. H. Conway and N. J. A. Sloane, "Quaternary constructions for the binary single-error correcting codes of Julin, Best, and others," *Designs, Codes Cryptogr.*, vol. 4, pp. 31–42, 1994.
[9] G. L. Feng, K. K. Tzeng, and V. K. Wei, "On the generalized Hamming weights of several classes of cyclic codes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1125–1130, May 1992.
[10] G. D. Forney, Jr., "Coset codes—Part II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1152–1187, Sept. 1988.
[11] ——, "Dimension/length profiles and trellis complexity of linear block codes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1741–1752, Nov. 1994.
[12] A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The Z_4-linearity of Kerdock, Preparata, Goethals, and related codes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 301–319, Mar. 1994.
[13] T. Helleseth, T. Kløve, and Ø. Ytrehus, "Generalized Hamming weights of linear codes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1133–1140, May 1992.
[14] B. Ho, "Generalized Lee weights for codes over Z_4," in *Proc. 1997 Int. Symp. Information Theory*, Ulm, Germany, June 29–July 4, 1997, p. 203.
[15] T. Kasami, T. Takata, T. Fujiwara, and S. Lin, "On the optimum bit orders with respect to the state complexity of trellis diagrams for binary linear codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 242–245, Jan. 1993.
[16] F. R. Kschischang, "The trellis structure of maximal fixed-cost codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1828–1838, Nov. 1996.
[17] A. Lafourcade and A. Vardy, "Lower bounds on trellis complexity of block codes," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1938–1954, Nov. 1995.
[18] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* New York: North-Holland, 1977.
[19] D. J. Muder, "Minimal trellises for block codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1049–1053, Sept. 1988.
[20] I. Reuven and Y. Be'ery, "Entropy/length profiles, bounds on the minimal covering of bipartite graphs, and trellis complexity of nonlinear codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 580–598, Mar. 1998.
[21] Y. Shany and Y. Be'ery, "The Preparata and Goethals codes: Trellis complexity and twisted squaring constructions," *IEEE Trans. Inform. Theory*, to be published.
[22] Y. Shany, I. Reuven, and Y. Be'ery, "On the trellis representation of the Delsarte–Goethals codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1547–1554, July 1998.
[23] V. Sidorenko, I. Martin, and B. Honary, "On separability of nonlinear block codes," *IEEE Trans. Inform. Theory*, to be published.
[24] M. A. Tsfasman and S. G. Vlăduţ, "Geometric approach to higher weights," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1564–1588, Nov. 1995.
[25] G. van der Geer and M. van der Vlugt, "Generalized Hamming weights of BCH(3) revisited," *IEEE Trans. Inform. Theory*, vol. 41, pp. 300–301, Jan. 1995.
[26] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1412–1418, Sept. 1991.
[27] V. K. Wei and K. Yang, "On the generalized Hamming weights of product codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1709–1713, Sept. 1993.
[28] K. Yang and T. Helleseth, "On the weight hierarchy of Preparata codes

over Z_4," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1832–1842, Nov. 1997.

[29] ____, "On the weight hierarchy of Goethals codes over Z_4," *IEEE Trans. Inform. Theory*, vol. 44, pp. 304–307, Jan. 1998.

[30] K. Yang, T. Helleseth, P. V. Kumar, and A. G. Shanbhag, "On the weight hierarchy of Kerdock codes over Z_4," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1587–1593, Sept. 1996.

On the Rectangularity of Nonlinear Block Codes

Vladimir Sidorenko, Ian Martin, and
Bahram Honary, *Senior Member, IEEE*

***Abstract*— We give simple sufficient conditions for a code to be rectangular and show that large families of well-known nonlinear codes are rectangular. These include Hadamard, Levenshtein, Delsarte–Goethals, Kerdock, and Nordstrom–Robinson codes. Being rectangular, each of these codes has a unique minimal trellis that can be used for soft-decision maximum-likelihood decoding.**

***Index Terms*— Codes, Delsarte–Goethals, Hadamard, Kerdock, Levenshtein, nonlinear, Nordstrom–Robinson, rectangular, trellis.**

I. Introduction

Modern practice requires from coding theory powerful codes and optimum soft-decision maximum-likelihood (ML) decoding. Usually, linear codes are used, since the theory of linear codes is well developed and linearity simplifies coding and decoding procedures. However, several famous families of nonlinear codes have more codewords than any comparable linear code presently known. This fact explains interest in nonlinear codes despite coding and decoding problems connected with the nonlinearity of the codes.

It is traditional to implement ML decoding using the Viterbi algorithm [24] applied to a code trellis. Let each edge in a trellis be labeled by a single code symbol. In this case, the Viterbi decoding algorithm requires $|E|$ additions and $|E|-|V|+1$ comparisons [15], where $|E|$ is the number of edges and $|V|$ is the number of vertices in a code trellis. For a given code many trellises can be constructed. To minimize the Viterbi decoding complexity we have to use a code trellis that has the minimum number of edges $|E|$ and minimum cycle rank $|E|-|V|+1$. However, historically *a minimal trellis* of a code C (with fixed order of codeword coordinates) is defined as one having the minimum number of vertices $|V|$ [4].

For linear codes (and group codes) it was shown that the minimal trellis is unique [5], [16] and it has the minimum number of edges $|E|$ [15] and the minimum cycle rank [20], [23]. Thus, for linear codes the minimal trellis minimizes the Viterbi decoding complexity. As for constructions of minimal trellises, it was proved [3], [9], [15], [27] that the classical methods of trellis design suggested by Bahl–Cocke–Jelinek–Raviv [1], Wolf [26], and Massey [14] produce the minimal trellis of a linear code. Recently, new methods of structured minimal trellis design were proposed in [10], [15], [21].

For nonlinear codes a minimal trellis is, generally speaking, not unique. Moreover, determination of a minimal trellis for a nonlinear code in general appears to be computationally infeasible [10]. However, there exists a class of codes having unique minimal trellis. These codes are called *rectangular* [11] or *separable* [20]. A code C of length n is *rectangular* if for each position t, $0 < t < n$, it can be separated into disjoint subcodes in such a way that concatenation of the t-length head of one codeword $\mathbf{c}_1$ with the $(n-t)$-length tail of another codeword $\mathbf{c}_2$ belongs to C if and only if $\mathbf{c}_1$ and $\mathbf{c}_2$ are in the same subcode. A rectangular code has a unique minimal trellis that has minimal number of vertices $|V|$ and edges $|E|$ and has minimum cycle rank $|E|-|V|+1$ [20], [23]. As a result, the Viterbi decoding complexity of a rectangular code is minimum when using the minimal code trellis.

Using the partition of a rectangular code it is straightforward (at least theoretically) to obtain the minimal trellis of the code [11], [20]. Thus, if a nonlinear code is rectangular, then complexity of ML decoding may be reduced with the aid of the minimal trellis. For nonlinear codes, encoding is also a problem. Using the minimal code trellis, the encoding problem for rectangular code can be solved as well [11].

All group codes (including linear codes) are rectangular [5], [11], [20]. In this paper, we investigate rectangularity of some known nonlinear codes. We say that a code is *permutation-rectangular* if it is rectangular and all equivalent codes obtained by permutations of codeword coordinates are also rectangular.

In Section III we obtain some general results concerning rectangularity of codes. We show that codes with $2d > n$ are permutation-rectangular, where d is the Hamming distance and n is code length. We prove that some transformations of a rectangular code (including an extension or translation) yields a rectangular code. We show that if $3d > n$ then a binary self-complementary code is permutation-rectangular.

In Section IV, we show that the Hadamard and the Levenshtein codes are permutation rectangular. It is tempting to conjecture that all known codes (or at least codes described in [13]) are rectangular. However, the example of a conference matrix code shows that this is not the case. A conference matrix code C_9, in the standard coordinate ordering, is not rectangular but can be made rectangular by coordinate permutation.

In Section V, we investigate rectangularity of binary images under the Gray map of Z_4 linear codes. We show that such a code can be nonrectangular. We prove that the Delsarte-Goethals codes, the Kerdock codes, and the Nordstrom-Robinson code are permutation rectangular.

II. Definitions

A block code C is a set of n words $\mathbf{c} = (c_1, \cdots, c_n)$ of length n over an alphabet $Q = \{0, 1, \cdots, q-1\}$. Denote by (n, M, d) a code with length n, which has M codewords and the minimum code distance d

$$d = d_{\min} = \min_{\mathbf{c}_1, \mathbf{c}_2 \in C, \mathbf{c}_1 \neq \mathbf{c}_2} D[\mathbf{c}_1, \mathbf{c}_2]$$

where $D[.,.]$ stands for the Hamming distance.

First, we give four equivalent definitions for a t rectangular code. We will later use the definition that gives us the simplest proof or explanation. Let t be an integer $t \in \{1, \cdots, n-1\} = [1, n-1]$. Split

Manuscript received October 20, 1996; revised March 5, 1998. This work was supported in part by the Royal Society, U.K.

V. Sidorenko is with the Institute for Problems of Information Transmission, Russian Academy of Science, Moscow GSP-4, Russia.

I. Martin and B. Honary are with the Communications Research Centre, Lancaster University, Lancaster LA1 4YR, U.K.

Communicated by A. Vardy, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(99)01409-1.