the decoding complexity of punctured convolutional codes. The new technique also has the advantage of requiring a smaller path memory than what is needed when decoding punctured codes. This makes the PUM codes and the decoding technique presented here an attractive alternative for applications requiring codes with high rate.

### ACKNOWLEDGMENT

The authors wish to thank K. J. Hole for valuable suggestions concerning the presentation of this correspondence.

### REFERENCES

[1] L. N. Lee, "Short unit-memory byte-oriented binary convolutional codes having maximal free distance," *IEEE Trans. Inform. Theory*, vol. IT-22, no. 3, pp. 349–352, May 1976.
[2] G. S. Lauer, "Some optimal partial-unit-memory codes," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 2, pp. 240–243, Mar. 1979.
[3] J. Justesen, E. Paaske, and M. Ballan, "Quasicyclic unit memory convolutional codes," *IEEE Trans. Inform. Theory*, vol. 36, no. 3, pp. 540–547, May 1990.
[4] J. Justesen, "Bounded distance decoding of unit memory codes," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1616–1627, Sept. 1993.
[5] G. D. Forney, Jr., "Dimension/length profiles and trellis complexity of linear block codes," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1741–1752, Nov. 1994.
[6] R. J. McEliece, "On the BCJR trellis for linear block codes," submitted to *IEEE Trans. Inform. Theory*, 1995.
[7] T. Helleseth, T. Kløve, and Ø. Ytrehus, "Codes and the chain condition," in *Proc. Int. Workshop on Algebraic and Combinatorial Coding Theory* (Voneshda Voda, Bulgaria, 1992), pp. 88–91.
[8] ——, "Codes, weight hierarchies, and chains," in *Proc. 1992 Int. Symp. on Information Theory and its Applications, ISITA'92,* (Singapore).
[9] G. D. Forney, Jr., "Convolutional codes I: Algebraic structure," *IEEE Trans. Inform. Theory*, vol. IT-16, no. 6, pp. 720–738, Nov. 1970.
[10] R. Johannesson and Z. Wan, "A linear algebra approach to minimal convolutional encoders," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1219–1233, July 1993.
[11] A. Dholakia, *Introduction to Convolutional Codes With Applications.* Boston, MA: Kluwer, 1994.
[12] J. B. Cain, G. C. Clark Jr., and J. M. Geist, "Punctured convolutional codes of rate $(n-1)/n$ and simplified maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 1, pp. 97–100, Jan. 1979.
[13] K. J. Hole, "Rate $k/(k+1)$ minimal punctured convolutional encoders," *IEEE Trans. Inform. Theory*, vol. 37, no. 3, pp. 653–655, May 1991.
[14] G. D. Forney, Jr., "Cosets codes II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, vol. 34, no. 5, pp. 1152–1187, Sept. 1988.
[15] A. Vardy and Y. Be'ery, "Maximum-likelihood soft decision decoding of BCH codes," *IEEE Trans. Inform. Theory*, vol. 40, no. 2, pp. 546–554, Mar. 1994.
[16] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1983.
[17] V. K. Wei and K. Yang, "On the generalized Hamming weights of products codes", *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1709–1713, Sept. 1993.
[18] I. E. Bocharova and B. D. Kudryashov, "Rational rate convolutional codes for soft decision Viterbi decoding," Dept. Inform. Syste., St. Petersburg Acad. Airspace Instrum., Tech. Rep., St. Petersburg, Russia.

# On the Generalized Hamming Weights of Convolutional Codes

Joachim Rosenthal, *Senior Member, IEEE,*
and Eric V. York *Student Member, IEEE*

*Abstract*—Motivated by applications in cryptology, Wei introduced in 1991 the concept of a generalized Hamming weight for a linear block code. In this correspondence, we define generalized Hamming weights for the class of convolutional codes and we derive several of their basic properties. By restricting to convolutional codes having a generator matrix $G(D)$ with bounded Kronecker indices we are able to derive upper and lower bounds on the weight hierarchy.

*Index Terms*— Convolutional codes, generalized Hamming weights, weight hierarchy, length/dimension profile.

## I. INTRODUCTION

An important set of code parameters defined for a linear block code are the so-called generalized Hamming weights first introduced by Wei in [1]. By definition, the $r$th generalized Hamming weight $d_r(C)$ of a linear block code $C$ is equal to the smallest support of any $r$-dimensional subcode of $C$. In particular, $d_0(C) = 0$ and $d_1(C)$ is equal to the distance of $C$.

In this way, every $[n, k]$ linear block code has associated a whole weight hierarchy

$$0 = d_0(C) < d_1(C) < \cdots < d_k(C) \leq n. \qquad (1.1)$$

The determination of the weight hierarchy of codes is desirable for applications to trellis encoders and we refer to Forney [2] and Wei [4]. Forney [2] calls the generalized Hamming weights the length/dimension profile (LDP) of a code. As explained in detail in [2], there is a deep connection between LDP and the complexity of the minimal trellis diagram. In [2] Forney also points out that a study of LDP, i.e., generalized Hamming weights of convolutional codes and other trellis codes would be desirable and this motivates in part the investigation of this correspondence.

Generalized Hamming weights have also a very natural geometric interpretation and this was pointed out in [3]. For this, recall that a set of ordered points $\mathcal{P} := \{P_1, \cdots, P_n\}$ in a $k$-dimensional vector space $V$ is called an $[n, k]$ system if $\mathcal{P}$ is not contained in any hyperplane $H \subset V$. Two $[n, k]$ systems $\mathcal{P}$ and $\mathcal{P}'$ are called equivalent if there is an isomorphism on $V$ mapping $\mathcal{P}$ onto $\mathcal{P}'$. As explained in [17, Sec. 1.1.2] every block code $C \subset \mathbb{F}_q^n$ uniquely defines an equivalence class of $[n, k]$ systems. It can be shown (see [3]) that the $r$th generalized Hamming weight is then geometrically described through the formula

$$d_r(C) = |\mathcal{P}| - \max_{H_r \subset V} \{|H_r \cap \mathcal{P}|\}$$

where $H_r$ is an arbitrary hyperplane of co-dimension $r$. Hence the generalized Hamming weights correspond to how well the subspaces

of $C$ are in "general position." The weight hierarchy of convolutional codes provides us with similar geometrical information about their structure as well.

Since the appearance of Wei's original article, several authors (see, e.g., [2]–[4]) were studying the weight hierarchy of different classes of linear block codes. In this correspondence we will study the weight hierarchy of a convolutional code. After formally introducing this concept we will derive, in the next section, several of the basic properties. In particular, we will show that the generalized Hamming weights form an infinite strictly increasing sequence $d_r(C)$ of positive integers and, similar to the case of block codes, the free distance of the code is exactly $d_1(C)$. Knowledge of the weight hierarchy is desirable in the design of encoders and this motivates in part the investigation of this correspondence.

In Section III we give an overview of current existing upper bounds for $d_r(C)$ of a block code. Those bounds prepare for the main results of this correspondence, which are given in Section IV. In this section we will derive some upper and lower bounds for the weight hierarchy of different classes of convolutional codes. The bounds we derive depend on the rate and complexity of the code as opposed to depending on the rate and memory (see, e.g., [14]). More specific results are derived for the generalized Hamming weights of rate $1/n$ codes in which case the memory and the complexity are the same.

In the last section, several illustrative examples are provided. In those examples we compute the complete weight hierarchy of several classes of codes. In this way, we are able to show that some of the bounds derived in Section IV are tight for some classes of codes. We conclude the correspondence by providing several tables containing bounds for certain classes of convolutional codes.

## II. DEFINITIONS AND BASIC PROPERTIES

Let $\mathbb{F}_q$ be the Galois field of $q$ elements, $\mathbb{F}_q[D]$ be the polynomial ring over $\mathbb{F}_q$, and $\mathbb{F}_q(D)$ the ring of rational functions. In the following it will be convenient to view elements of $\mathbb{F}_q(D)$ as infinite (periodic) power series of the form

$$\sum_{i=0}^{\infty} x_i D^i, \quad x_i \in \mathbb{F}_q.$$

Let $C$ be a rate $k/n$ convolutional code represented through a noncatastrophic encoder

$$G(D) = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}. \quad (2.1)$$

Without loss of generality we will assume that the matrix $G(D)$ which is defined over $\mathbb{F}_q[D]$ is in row proper form; in other words, we will assume that the "high-order coefficient matrix" has full row rank. From this it follows that $G(D)$ is a minimal encoder, hence has (unique) ordered row (Kronecker) indices

$$\nu_1 \geq \cdots \geq \nu_k$$

where the indices $\nu_i$ are formally defined through

$$\nu_i = \max\{\deg(g_{ij}) \mid 1 \leq j \leq n\}, \qquad i = 1, \cdots, k.$$

We will denote the memory, complexity, and constraint length of a convolutional code by $m$, $c$, and $\eta$, respectively. In terms of the Kronecker indices we have: $m = \nu_1$

$$c = \sum_{i=1}^{k} \nu_i$$

and $\eta = n(\nu_1 + 1)$.

In an obvious way we can view $C$ as an (infinite-dimensional) linear $\mathbb{F}_q$ vector space. Let

$$\{u_1(D), \cdots, u_r(D)\}$$

be $r$ linearly independent vectors in $\mathbb{F}_q^k(D)$. Since $G(D)$ has by assumption linearly independent rows it follows that

$$\operatorname{span}_{\mathbb{F}_q}\{u_1(D)G(D), \cdots, u_r(D)G(D)\} \subset C \subset \mathbb{F}_q^n(D)$$

defines an $r$-dimensional subspace of $C$ and clearly every $r$-dimensional subspace $U \subset C$ is of this form.

*Definition 2.1:* Let $U \subset C$ be a linear subspace of $C$ and let

$$x(D) = (x_1(D), \cdots, x_i(D), \cdots, x_n(D)) \in U$$

be a codeword in $U$ whose $i$th component is of the form

$$x_i(D) = x_{i0} + x_{i1}D + \cdots + x_{is_i}D^{s_i} \in \mathbb{F}_q[D].$$

Then

$$\chi(U) := \{(i,j) \mid \exists\, (x_1(D), \cdots, x_n(D)) \in U, x_{ij} \neq 0\}$$

is called *the support* of $U$ and

$$d_r(C) := \min\{|\chi(U)| \mid U \subset C \text{ and } \dim U = r\}$$

is called the *$r$th generalized Hamming weight* of $C$.

Note that the generalized Hamming weights are well defined for any positive integer $r$ and not just for $r = 0, \cdots, k$ as it is the case for block codes. Also note that if $U$ is one-dimensional and $x \in U$ is any nonzero codeword then $|\chi(U)|$ is nothing else than the usual Hamming weight $w(x)$ of the codeword $x$. In particular, it follows in analogy to the block code case that $d_1(C)$ is equal to the free distance of $C$.

Since we assume that $G(D)$ is a minimal encoder, it is clear that in order to compute $d_r(C)$ it is sufficient to consider polynomial inputs of some bounded degree. Indeed, if $u(D) = (u_1(D), \cdots, u_k(D))$ is an infinite support input vector, then necessarily $u(D)G(D) = x(D)$ has infinite support by assumption. The following lemma provides a conservative bound on the degrees of the polynomial inputs $u(D) \in \mathbb{F}_q^k[D]$. Since we only need later in the correspondence that such a bound exists and for the sake of brevity, we omit the proof.

*Lemma 2.2:* Let $C$ be a convolutional code of rate $k/n$ and memory $\nu_1$. In order to compute $d_r(C)$ it is enough to consider subspaces of the form

$$U = \operatorname{span}\{u_1(D)G(D), \cdots, u_r(D)G(D)\}$$

where $u_i(D) \in \mathbb{F}_q^k[D]$ and the $\deg(u_i(D)) < \nu_1 nr(\nu_1 + r)$.

The following Lemma is a natural generalization of Wei's monotonicity theorem [1, Theorem 1] for block codes.

*Lemma 2.3:* The generalized Hamming weights of a convolutional code form a (strictly) increasing set of positive integers

$$0 = d_0(C) < d_1(C) < d_2(C) < \cdots.$$

*Proof:* Obviously, the sequence is weakly increasing. In order to show the strict inequality let $U \subset C$ have the property that $\dim U = r$ and $|\chi(U)| = d_r(C)$. Assume $(i, j)$ is in the support of $U$, i.e., there is a codeword

$$\left(\sum x_{1j}D^j, \cdots, \sum x_{nj}D^j\right) \in U, \quad x_{ij} \neq 0.$$

Let $V := \{c \in U \mid x_{ij} = 0\}$. But then one has $|\chi(V)| < |\chi(U)|$ and $\dim V = r - 1$. In other words, $d_{r-1}(C) < d_r(C)$. ∎

In [4, Sec. IV], the authors define the *chain condition* for block codes. This definition is easily generalized to convolutional codes as follows:

*Definition 2.4:* A convolutional code $C$ is said to satisfy the *chain condition* if there exists subcodes $C_r$ for $1 \leq r \leq \infty$ such that $\text{rank}(C_r) = r$, $|\chi(C_r)| = d_r(C)$, and $C_{r-1} \subset C_r$.

Similar to the block code case, we can describe the numbers $d_r(C)$ also algebraically. For this consider a parity check matrix $H(D) = H_0 + H_1 D + \cdots + H_m D^m$ of $C$ and let $H$ be the semi-infinite sliding-block matrix defined by the scalar matrices $H_0, H_1, \cdots, H_m$ (see, e.g., [12]). Then one has the immediate generalization of [1, Theorem 2]

*Theorem 2.5:* $C$ has generalized Hamming weight $d_r(C) = d$ if and only if $d$ is the smallest number with the property that there are $d$ columns of $H$ whose rank is $d - r$ or less.

### III. BOUNDS FOR THE WEIGHT HIERARCHY OF A BLOCK CODE

In this section we summarize the best general upper bounds known for the generalized Hamming weights of block codes. Those results will then be the basis in our investigation of the bounds of the generalized Hamming weights of convolutional codes.

The first bound was already given in [1] by Wei who called the bound the generalized Singleton bound.

*Lemma 3.1:* For an $[n, k]$ code $C$ one has

$$d_r(C) \leq n - k + r.$$

The next bound is the well-known Griesmer bound [5].

*Lemma 3.2:* For a linear block code over $\mathbb{F}_q$ with rate $k/n$ and distance $d$ we have

$$\sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil \leq n. \tag{3.1}$$

The following, given in [6, Theorem 5], is a generalization of the Griesmer bound to $d_r(C)$ :

*Theorem 3.3:* Let $C$ be a binary, linear $[n, k]$ code. Then for $1 \leq r \leq k$ one has:

$$d_r(C) + \sum_{i=1}^{k-r} \left\lceil \frac{d_r(C)}{(2^r - 1)2^i} \right\rceil \leq n. \tag{3.2}$$

When the distance of a block code is known, the Griesmer bound can also be used as a lower bound. The following result is referred to in [3, p. 276] as the *Griesmer–Wei Bound*:

*Lemma 3.4:* For a linear block code over $\mathbb{F}_q$ with rate $k/n$ and distance $d$ we have

$$d_r(C) \geq \sum_{i=0}^{r-1} \left\lceil \frac{d}{q^i} \right\rceil. \tag{3.3}$$

### IV. BOUNDS FOR THE WEIGHT HIERARCHY OF A CONVOLUTIONAL CODE

In this section we will derive a set of upper bounds on the generalized Hamming weights which have to be satisfied for all convolutional codes. In order to properly pose the problem it is of course necessary to restrict to certain classes of convolutional codes.

The codes which we single out are all convolutional codes having a fixed rate $k/n$ and having a basic encoder (see, e.g., [7, Sec. 2.3]) with a fixed set of Kronecker indices $\nu = (\nu_1, \cdots, \nu_k)$. Clearly it is most natural to fix the rate. Moreover, the set of encoders having a fixed set of Kronecker indices is most natural too. Indeed, every encoder can be naturally identified with an associated Hermann–Martin map [8] from the projective line to a fixed Grassmann variety and the Kronecker indices correspond in this case exactly to the Grothendieck indices of the pull back of the tautological bundle. In system theory

the Kronecker indices correspond to the observability indices of the associated MA representation and the complexity

$$c = \sum_{i=1}^{k} \nu_i$$

is exactly the McMillan degree of the system. For readers interested in more details covering those interesting relations we refer to [9]–[11].

Because of the above mentioned reasons, we seek upper bounds on the weight hierarchy in the class of convolutional codes having fixed rate $k/n$ and having a basic encoder with a fixed set of Kronecker indices $\nu = (\nu_1, \cdots, \nu_k)$. Note that for $\nu_1 = 0$ (no memory) the problem is equivalent to estimating upper bounds of block codes as it was considered in the last section. In this way our problem can also be viewed as a natural generalization.

The basic strategy of how we will proceed to accomplish upper bounds is as follows (compare also with [7, Sec. 3.1]): Let $V \subset \mathbb{F}_q^n[D]$ be any finite-dimensional linear $\mathbb{F}_q$-subspace. Then

$$C_V := \{u(D)G(D) \mid u(D)G(D) \in V, u(D) \in \mathbb{F}_q^k[D]\}$$

and

$$C_V \subset C \subset \mathbb{F}_q^n[D].$$

$C_V$ defines a linear $[\mathcal{N}, \mathcal{K}]$ block code, where

$$\mathcal{N} = |\chi(C_V)| \leq |\chi(V)| \text{ and } \mathcal{K} = \dim C_V. \tag{4.1}$$

For every such linear block code $C_V$ we then necessarily have that

$$d_r(C) \leq d_r(C_V) \qquad \forall r. \tag{4.2}$$

Clearly, the bounds of $d_r(C_V)$ are expected to be tighter if the rate $\mathcal{K}/\mathcal{N}$ is large and because of this we will single out a set of subspaces $V$ which have a maximal dimension for a given support. Specifically we will consider for each integer $\gamma \geq 0$ the subspace

$$V\gamma := \{(x_1(D), \cdots, x_n(D)) \in \mathbb{F}_q^n[D] \mid$$
$$\deg x_i(D) \leq \gamma, i = 1, \cdots, n\}.$$

Note that $V_\gamma$ is in a natural way a $\mathbb{F}_q$ vector space of dimension $n\gamma + n$; indeed, one has natural vector space isomorphisms

$$V_\gamma \simeq \mathbb{F}_q^{\gamma+1} \otimes \mathbb{F}_q^n \simeq \mathbb{F}_q^{n\gamma+n}. \tag{4.3}$$

The following Lemma establishes the block size of the codes $C_{V_\gamma}$ which we will abbreviate with

$$C_\gamma := C_{V_\gamma}.$$

*Lemma 4.1:* Let $C$ be a rate $k/n$ convolutional code represented by a basic encoder having Kronecker indices $\nu_1, \cdots, \nu_k$. Let $\mathcal{N} = |\chi(C_\gamma)|$. Then for each $\gamma \geq 0$ the code $C_\gamma$ is a linear $[\mathcal{N}, \mathcal{K}]$ block code where

$$\mathcal{N} \leq n\gamma + n \tag{4.4}$$

$$\mathcal{K} = \sum_{i=1}^{k} \max(\gamma - \nu_i + 1, 0). \tag{4.5}$$

*Proof:* The fact that $C_\gamma$ is linear is obvious and the estimate (4.4) is a direct consequence of (4.1) and (4.3). Let $G(D)$ be a basic encoder with Kronecker indices $\nu_1, \cdots, \nu_k$ and let $u(D) = (u_1, u_2, \cdots, u_k)$. From the fact that the high-order coefficient matrix of $G(D)$ has full row rank it follows that

$$\deg u(D)G(D) = \max_{1 \leq i \leq k} \{\deg u_i + \nu_i\}.$$

In particular, it follows that every $(u_1, u_2, \cdots, u_k)$ having the property that $\deg u_i \leq \gamma - \nu_i$ results in a valid element of $C_\gamma$ and the map $u \mapsto uG$ induces a vector space isomorphism

$$C_\gamma \simeq \bigoplus_{i=1}^{k} \mathbb{F}_q^{\gamma - \nu_i + 1} \simeq \mathbb{F}_q^{\mathcal{K}}$$

where we made use of the convention $\mathbb{F}^\alpha = 0$ if $\alpha < 0$.    ■

The following examples illustrate the concepts introduced thus far.

*Example 4.2:* Let $G(D) = (D^2 + D + 1, D + 1)$ be the generator matrix of the convolutional code $C$, then the rate of $C_\gamma$ is $(\gamma - 1)/(2\gamma + 1)$ for any $\gamma \geq 2$. An arbitrary element of $C_\gamma$ would be given by

$$y(D) = (a_\gamma D^\gamma + a_{\gamma-1} D^{\gamma-1} + \cdots a_1 D + a_0,$$
$$b_{\gamma-1} D^{\gamma-1} + b_{\gamma-2} D^{\gamma-2} + \cdots b_1 D + b_0).$$

The vector $y(D)$ can be naturally identified with the vector

$$(a_\gamma, a_{\gamma-1}, \cdots, a_1, a_0, b_{\gamma-1}, b_{\gamma-2}, \cdots, b_1, b_0)$$

viewed as element of the vector space $\mathbb{F}_2^{2\gamma+1}$.

*Example 4.3:* Let

$$G(D) = \begin{bmatrix} D^2 + 1 & D & D^2 \\ D + 1 & 1 & D + 1 \end{bmatrix}$$

be the generator matrix of the convolutional code $C$, then the rate of $C_\gamma$ is $(2\gamma - 1)/(3\gamma + 2)$ for $\gamma \geq 1$. An arbitrary element of $C_2$ would be given by

$$y(D) = (a_2 D^2 + a_1 D + a_0, \ b_1 D + b_0, \ c_2 D^2 + c_1 D + c_0).$$

The vector $y(D)$ can be identified with the vector

$$(a_2, a_1, a_0, b_1, b_0, c_2, c_1, c_0)$$

in $\mathbb{F}_2^8$. A generator matrix for the block code determined by $C_2$ is

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

We would like to remark at this point that for many codes $C$ the support $\mathcal{N} = |\chi(C_\gamma)|$ is strictly less than $n\gamma + n$ and only for a "generic code" in the class of rate $k/n$ codes with indices $\nu$ equality holds in (4.4). It therefore follows that if one restricts the class of convolutional codes further it is possible to achieve even sharper bounds than the bounds which we will derive shortly (in terms of the constraint length we say one only needs to consider the *effective constraint length* as pointed out by Costello in [13]).

Before we derive several upper bounds for the generalized Hamming weights $d_r(C)$ of a convolutional code $C$ we show through the next theorem that "optimal bounds" for the block codes $C_\gamma$ result in *optimal bounds* for the convolutional code $C$.

*Theorem 4.4:* Let $C$ be a rate $k/n$ convolutional code. Then the vector spaces

$$C_0 \subset C_1 \subset C_2 \subset \cdots$$

form a direct system of vector spaces with direct limit

$$\lim_{\gamma \to \infty} C_\gamma = C. \tag{4.6}$$

Moreover, for every integer $r \geq 1$ there exists a positive integer $n_o$ dependent only on $k$, $n$, $r$, $\nu$ having the property that

$$d_r(C_\gamma) = d_r(C) \tag{4.7}$$

for all $\gamma \geq n_o$.

*Proof:* The first part is a direct consequence of the definition of $C_\gamma$ and the definition of a direct limit. The second part follows from Lemma 2.2.    ■

The first upper bound which we will present is based on the generalized Griesmer bound as introduced in Theorem 3.3.

*Theorem 4.5:* Let $C$ be a binary rate $k/n$ convolutional code having a basic encoder with Kronecker indices $\nu = (\nu_1, \cdots, \nu_k)$. Let $\gamma$ be a positive integer and let

$$\mathcal{K} = \sum_{i=1}^{k} \max(\gamma - \nu_i + 1, 0).$$

Then the $r$th generalized Hamming weight of $C$ satisfies

$$d_r(C) + \sum_{i=1}^{\mathcal{K}-r} \left\lceil \frac{d_r(C)}{2^i(2^r - 1)} \right\rceil \leq n\gamma + n. \tag{4.8}$$

For convolutional codes defined over an arbitrary field $\mathbb{F}_q$ we have the well-known Griesmer bound:

*Theorem 4.6:* Let $C$ be a convolutional code over $\mathbb{F}_q$ with rate $k/n$, then for all $\gamma \geq 0$, $d_1(C)$ must satisfy

$$\sum_{i=0}^{\mathcal{K}-1} \left\lceil \frac{d_1(C)}{q^i} \right\rceil \leq n\gamma + n$$

where $\mathcal{K}$ is determined by (4.5).

Rate $1/n$ codes have been studied extensively, and there are several very effective techniques for constructing codes of this rate with *good* free distance (see, e.g., [18]–[20]). Next, we study the properties of the generalized Hamming weights for these types of codes.

Let $C$ be a convolutional code of rate $1/n$ and constraint length $\eta$. Then the fact that

$$d_r(C) \leq \eta + n(r - 1) \tag{4.9}$$

is obvious.

*Theorem 4.7:* Let $C$ be a convolutional code of rate $1/n$ generated by $G(D)$. Then

$$d_r(C) + n \leq d_{r+1}(C). \tag{4.10}$$

*Proof:* Suppose that $V \subset C$ has dimension $r + 1$ and support $d_{r+1}(C)$. Then, by Lemma 2.2, there exist $r + 1$ linearly independent vectors $u_s(D)$ such that

$$V = \operatorname{span}\{u_1(D)G(D), \cdots, u_{r+1}(D)G(D)\}.$$

Since the $u_s(D)$'s are linearly independent, by row reducing we can obtain the polynomials

$$\{\tilde{u}_1(D), \tilde{u}_2(D), \cdots, \tilde{u}_{r+1}(D)\}$$

such that

$$V = \operatorname{span}\{\tilde{u}_1(D)G(D), \cdots, \tilde{u}_{r+1}(D)G(D)\}$$

and

$$\deg(\tilde{u}_1(D)) < \deg(\tilde{u}_2(D)) < \cdots < \deg(\tilde{u}_{r+1}(D)).$$

The $r$-dimensional subspace

$$\tilde{V} = \mathrm{span}\{\tilde{u}_1(D)G(D), \cdots, \tilde{u}_r(D)G(D)\}$$

has

$$d_r(C) \le |\chi(\tilde{V})| \le d_{r+1}(C) - n. \qquad \blacksquare$$

*Corollary 4.8:* Let $C$ be a convolutional code of rate $1/n$ and memory $m$. Then if

$$d_r(C) = \eta + n(r - 1)$$

for some $r \ge 1$ we have

$$d_j(C) = \eta + n(j - 1), \forall j > r.$$

## V. TABLES AND EXAMPLES

In this section we will give several examples illustrating the concepts defined throughout this correspondence. We also present tables containing the bounds for $d_r(C)$ for some low-rate codes with particular Kronecker indices.

*Example 5.1:* Consider the class of convolutional codes over $\mathbb{F}_2$ with rate $k/n = \frac{1}{2}$ and memory $m = \nu_1 = 16$. Using (4.6) and considering the elements in $C_{19}$, we obtain

$$\sum_{j=0}^{3} \left\lceil \frac{d_1(C)}{2^j} \right\rceil \le 40.$$

This implies that $d_1(C) \le 20$. In [12, p. 330] it is shown that there exists a rate $1/2$ code having memory $m = 16$ and distance $d_1(C) = 20$. The bound is therefore tight in this particular example.

*Example 5.2:* If $C$ is generated by $G(D) = (D^2 + D + 1, D^2 + 1)$ then one has $d_0(C) = 0$, $d_1(C) = 5$. By Lemma 3.4 we must have $d_2(C) \ge 8$, hence by Corollary 4.8 we must have $d_r(C) = 2(r-1) + 6$, $\forall i > 1$. Furthermore, let $c_i(D) = D^{i-1}G(D)$, and set

$$D_r = \mathrm{span}\{c_1(D), c_2(D), \cdots, c_r(D)\}.$$

Then one can easily verify that $|\chi(D_r)| = d_r(C)$, hence $C$ has optimal generalized Hamming weight and satisfies the chain condition for convolutional codes.

*Example 5.3:* Consider the class of convolutional codes over $\mathbb{F}_2$ with rate $2/3$, and $\nu = (2, 3)$. Using (4.6) and considering the elements in $C_3$, we obtain

$$\sum_{j=0}^{2} \left\lceil \frac{d_1(C)}{2^j} \right\rceil \le 12.$$

This implies that $d_1(C) \le 6$. In [12, p. 330] it is shown that the rate $2/3$, memory $m = 3$ code $C$ which is generated by

$$G = \begin{bmatrix} D^2 + D^3 & 1 & 1 + D + D^2 + D^3 \\ 1 + D & D + D^2 & 1 + D + D^2 \end{bmatrix}$$

has a free distance $d_1(C)$ of 6. The bound $d_1(C) \le 6$ is therefore tight for the class of rate $2/3$, codes having Kronecker indices $\nu = (2, 3)$. Note: If one were to consider the class of rate $2/3$, codes having Kronecker indices $\nu = (3, 3)$ then the Griesmer bound

gives $d_1(C) \le 8$, hence consideration of the Kronecker indices does give a refinement on existing bounds.

Next consider $c_1 = (0, 1 + D)G(D)$, $c_2 = (1 + D, 0)G(D)$, and $c_3 = (1, D^2)G(D)$. Then one can verify that

$$|\chi(\mathrm{span}\{c_1, c_2\})| = 9$$

and

$$|\chi(\mathrm{span}\{c_1, c_2, c_3\})| = 11.$$

It therefore follows that $d_2(C) \le 9$ and $d_3(C) \le 11$. By Lemma 3.4, $d_2(C) \ge 9$ and $d_3(C) \ge 11$ hence $d_2(C) = 9$ and $d_3(C) = 11$.

*Example 5.4:* Let $\tilde{C}$ be the $1/4$ code generated by $\tilde{G} = (D + 1, 1, D, 1)$. Let $G$ be the matrix from Example 5.3, then the $3/7$ code $C'$ generated by

$$G' = \begin{bmatrix} \tilde{G} & 0 \\ 0 & G \end{bmatrix}$$

has $d_1(C') = 5$ and by Example 5.3 and Lemma 4.7 we have $d_2(C') = 9$ $d_3(C') = 11$ and any three-dimensional subspace $V$ that contains the vector $(D + 1, 1, D, 1)$ must have $d_3(V) \ge 13$, hence $C'$ does not satisfy the chain condition for convolutional codes.

*Example 5.5:* Consider the rate $1/4$ code $C$ given by

$$G(D) = (1 + D^2 + D^3, 1 + D + D^3, 1 + D + D^3, 1 + D + D^2 + D^3).$$

By consulting [12, p. 330] we see that $d_1(C) = 13$. By Lemma 3.4 and (4.9) we have $20 \ge d_2(C) \ge 20$ hence by Corollary 4.8 we have

$$d_r(C) = 4(r - 1) + 16, \qquad \forall r > 1$$

By considering the vectors $D^i G(D)$ one can show that the code $C$ satisfies the chain condition as well.

The following tables were obtained using (4.8).

| rate 1/2 | | | | | |
|---|---|---|---|---|---|
| $\nu_1$ | $d_1(C) \le$ | $d_2(C) \le$ | $d_3(C) \le$ | $d_4(C) \le$ | $d_5(C) \le$ |
| 1 | 4 | 6 | 8 | 10 | 12 |
| 2 | 5 | 8 | 10 | 12 | 14 |
| 3 | 6 | 10 | 12 | 14 | 16 |
| 4 | 8 | 12 | 14 | 16 | 18 |
| 5 | 8 | 13 | 16 | 18 | 20 |

| rate 1/3 | | | | | |
|---|---|---|---|---|---|
| $\nu_1$ | $d_1(C) \le$ | $d_2(C) \le$ | $d_3(C) \le$ | $d_4(C) \le$ | $d_5(C) \le$ |
| 1 | 6 | 9 | 12 | 15 | 18 |
| 2 | 8 | 12 | 15 | 18 | 21 |
| 3 | 10 | 15 | 18 | 21 | 24 |
| 4 | 12 | 18 | 21 | 24 | 27 |
| 5 | 13 | 20 | 24 | 27 | 30 |

| rate 2/4 | | | | | |
|---|---|---|---|---|---|
| $(\nu_1, \nu_2)$ | $d_1(C) \le$ | $d_2(C) \le$ | $d_3(C) \le$ | $d_4(C) \le$ | $d_5(C) \le$ |
| (1, 0) | 4 | 6 | 8 | 11 | 12 |
| (1, 1) | 5 | 8 | 11 | 12 | 15 |
| (2, 0) | 6 | 9 | 11 | 12 | 15 |
| (2, 1) | 6 | 10 | 12 | 15 | 16 |
| (3, 0) | 4 | 8 | 12 | 15 | 16 |
| (2, 2) | 8 | 12 | 14 | 16 | 19 |
| (3, 1) | 8 | 12 | 14 | 16 | 19 |
| (4, 0) | 4 | 8 | 12 | 16 | 19 |
| (3, 2) | 8 | 13 | 16 | 19 | 20 |
| (4, 1) | 8 | 12 | 16 | 19 | 20 |
| (5, 0) | 4 | 8 | 12 | 16 | 20 |

| rate 2/3 | | | | | |
|---|---|---|---|---|---|
| $(\nu_1, \nu_2)$ | $d_1(C) \leq$ | $d_2(C) \leq$ | $d_3(C) \leq$ | $d_4(C) \leq$ | $d_5(C) \leq$ |
| (1, 0) | 3 | 5 | 6 | 8 | 9 |
| (1, 1) | 4 | 6 | 8 | 9 | 11 |
| (2, 0) | 3 | 6 | 8 | 9 | 11 |
| (2, 1) | 4 | 7 | 9 | 11 | 12 |
| (3, 0) | 3 | 6 | 9 | 11 | 12 |
| (2, 2) | 6 | 9 | 11 | 12 | 14 |
| (3, 1) | 6 | 9 | 11 | 12 | 14 |
| (4, 0) | 3 | 6 | 9 | 12 | 14 |
| (3, 2) | 6 | 10 | 12 | 14 | 15 |
| (4, 1) | 6 | 9 | 12 | 14 | 15 |
| (5, 0) | 3 | 6 | 9 | 12 | 15 |

## REFERENCES

[1] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 5, pp. 1412–1418, Sept. 1991.

[2] G. D. Forney, Jr., "Dimension/length profiles and trellis complexity of linear block codes," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1741–1752, Nov. 1994.

[3] J. W. P. Hirschfeld, M. A. Tsfasman, and S. G. Vlăduţ, "The weight hierarchy of higher dimensional Hermitian codes," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 275–278, Jan. 1994.

[4] V. K. Wei and K. Yang, "On the generalized Hamming weights of product codes," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1709–1713, Sept. 1993.

[5] J. H. Griesmer, "A bound for error-correcting codes," *IBM J. Res. Devel.*, vol. 4, pp. 532–542, 1960.

[6] T. Helleseth, T. Kløve, and Ø. Ytrehus, "Generalized Hamming weights of linear codes," *IEEE Trans. Inform. Theory*, vol. 38, no. 3, pp. 1133–1140, May 1992.

[7] P. Piret, *Convolutional Codes, an Algebraic Approach.* Cambridge, MA: MIT Press, 1988.

[8] C. F. Martin and R. Hermann, "Applications of algebraic geometry to system theory: The McMillan degree and Kronecker indices as topological and holomorphic invariants," *SIAM J. Contr. Optim.*, vol. 16, pp. 743–755, 1978.

[9] G. D. Forney, Jr., "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr. Optim.*, vol. 13, no. 3, pp. 493–520, 1975.

[10] M. S. Ravi and J. Rosenthal, "A smooth compactification of the space of transfer functions with fixed McMillan degree," *Acta Appl. Math.*, vol. 34, pp. 329–352, 1994.

[11] J. Rosenthal, M. Sain, and X. Wang, "Topological considerations for autoregressive systems with fixed Kronecker indices," *Circuits Syst. Sig. Processing*, vol. 13, nos. 2–3, pp. 295–308, 1994.

[12] S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[13] D. Costello, "Free distance bounds for convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 3, pp. 356–365, May 1974.

[14] R. Johannesson and K. Zigangirov, "Distances and distance bounds for convolutional codes—An overview," in *Topics in Coding Theory* (In honor of L. H. Zetterberg) *Lecture Notes in Control and Information Sciences*, no. 128. New York: Springer-Verlag, 1989, pp. 109–136.

[15] F. J. MacWilliams and N. J. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North Holland, 1977.

[16] S. Roman, *Coding and Information Theory, Graduate Texts in Mathematics.* New York/Berlin: Springer-Verlag, 1992.

[17] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes, Mathematics and Its Applications.* Amsterdam, The Netherlands: Kluwer, 1991.

[18] J. L. Massey, D. J. Costello, and J. Justesen, "Polynomial weights and code constructions," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 1, pp. 101–110, Jan. 1973.

[19] W. J. Reitsema "A lower bound on the free distance of convolutional codes related to cyclic codes," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 5, pp. 638–369, Sept. 1973.

[20] J. Justesen, "An algebraic construction of rate $1/\nu$ convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 1, pp. 577–580, Jan. 1975.

# A Forbidden Rate Region for Generalized Cross Constellations

E. A. Gelblum and A. R. Calderbank, *Member, IEEE*

*Abstract*— An analysis of the Generalized Cross Constellation (GCC) is presented and a new perspective on its coding algorithm is described. We show how the GCC can be used to address generic sets of symbol points in any multidimensional space through an example based on the matched spectral null coding used in magnetic recording devices. We also prove that there is a forbidden rate region of fractional coding rates that are practically unrealizable using the GCC construction. We introduce the idea of a constellation tree and show how its decomposition can be used to design GCC's matching desired parameters. Following this analysis, an algorithm to design the optimal rate GCC from a restriction on the maximum size of its constellation signal set is given, and a formula for determining the size of the GCC achieving a desired coding rate is derived. We finish with an upper bound on the size of the constellation expansion ratio.

*Index Terms*— Data transmission, QAM signaling, generalized cross constellations

## I. INTRODUCTION

The principal use of two-dimensional QAM constellations is the transmission of voiceband modem data [1]–[3]. Traditionally, two-dimensional constellations consisting of $2^n$ channel symbols for some integer $n$, are addressed symbol-by-symbol via a low-complexity lookup table with $2^n$ entries. The ability to shape these constellations, however, is restricted in only two dimensions, and the cardinality of the channel symbol sets are necessarily limited to an integer power of two. In constrast to the two-dimensional constellation, multidimensional constellations, formed by the concatenation of $N$ two-dimensional QAM channel symbols, are able to provide higher shaping gain than the two-dimensional variety at the expense, however, of higher complexity. This added complexity is introduced by the $2N$-dimensional codebook whose mere size makes traditional lookup-table addressing methods unwieldy. The generalized cross constellation (GCC), first mentioned in [1], and subsequently described in [4], is a multidimensional constellation that both exhibits low addressing complexity and allows the transmission of a nonintegral number of bits per channel symbol.

The $2N$-dimensional GCC selects a block of $N$ two-dimensional points from among a family of simply defined constituent subconstellations in a two-step process. First, it chooses a constrained sequence of the subconstellations, and second it selects an individual channel symbol from each one in the sequence. This construction simplifies the addressing procedure by reducing the multidimensional addressing problem to a series of $N$ two-dimensional subconstellation mappings [5]. In addition, since the probability with which each subconstellation is selected is purposely disproportionate to its relative size, the GCC makes it possible to use the lower power signal sets with increased frequency and thereby also reduces average transmitted signal power. While this addressing technique can be applied to channel constellations of any type, generalized cross constellations have hitherto found application in QAM modems. Ungerboeck-like trellis coding of these $N$-dimensional signal points is also possible and is further described in [5].