UNIT MEMORY CONVOLUTIONAL CODES WITH MAXIMUM DISTANCE

ROXANA SMARANDACHE*

Abstract. Unit memory codes and in particular, partial unit memory codes are reviewed. Conditions for the optimality of partial unit memory codes with degree $k-1$ are given, where optimal codes are the codes having the maximum free distance among all codes of the same parameters k, n and degree μ. A binary construction of unit memory codes with $\mu = k-1$ is discussed for the cases that satisfy the optimality conditions. This construction is generalized for codes over fields of characteristic $p > 2$.

Key words. Unit memory convolutional codes, MDS-convolutional codes.

1. Introduction. Maximum distance separable (MDS) convolutional codes, introduced first in [10], are convolutional codes characterized through the property that their free distance is maximum among all codes of the same parameters n, k and degree μ. The free distance of an MDS code represents a bound for the free distance of all the other codes with the same parameters. We call this bound the generalized Singleton bound, since in the case of block codes it obviously reduces to the Singleton bound. The existence of MDS convolutional codes was established in [10] by using elements of algebraic geometry and an input-state-output representation of convolutional codes. The proof was existential and could not provide a method of construction. A constructive proof was given later in [12]. The construction starts from a large Reed Solomon block code and therefore needs a fairly large finite field.

The question of finding a lower bound for the field size and MDS codes attaining this bound was now raised. Since unit memory codes have the simplest representation among the codes of nonzero memory we started our study by analyzing the conditions these codes need to satisfy in order to be maximum. We also took a new approach. We started with the binary field and came to study the binary partial unit memory codes with degree $\mu = k-1$, this being the only nontrivial case where binary MDS codes exist.

Binary partial unit memory codes were studied in the literature by Lauer [4] and Justesen [3] who showed that in some situations a unit memory code performs better than the codes having the same rate and degree but memory larger than 1. Some constructions given in [4] are the inspi-

*Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46556-5683, USA, *e-mail:* Smarandache.1@nd.edu. The author was supported by NSF grant DMS-96-10389 and by a fellowship from the Center of Applied Mathematics at the University of Notre Dame.

ration for the idea of this paper. In [3] quasi-cyclic unit memory codes are studied and some constructions and computer search results are presented. Furthermore some of the basic structural properties are discussed, such as noncatastrophicity, minimality conditions, distance measures, properties that we will use in this paper. Therefore we chose to use the same language as in these papers, only mentioning what this means in terms of the language of the MDS papers [10, 12]. For the development of this paper this is quite enough.

The paper consists of 6 sections, the first two being introductory and the last one an appendix section containing material that we will heavily use. In Section 3 we state some equivalent conditions for binary PUM codes to be optimal and in Section 4 we give a method of construction for this type of codes. Section 5 generalizes the binary construction to the case where the field has characteristic larger than 2. We add examples of both methods in Section 6.

2. Unit memory codes. Let $\mathbb{F}$ denote a finite field. A unit memory encoder is defined through the following encoding scheme:

$$v_t = u_t G_0 + u_{t-1} G_1 \tag{2.1}$$

where $u_t \in \mathbb{F}^k$ is the k information tuple at time $t, t = 0, 1, \ldots$ and $v_t \in \mathbb{F}^n$ is the n-tuple denoting the encoded vector at time t. By convention $u_t = 0$ for $t < 0$. The matrices G_0 and G_1 are defined over the field $\mathbb{F}$ and have size $k \times n$. We assume that G_0 has rank k.

Then a rate k/n *unit memory code* (UMC) is the set of all sequences generated by an encoder (G_0, G_1), with G_0 of rank k, satisfying the above encoding rule.

The code can also be defined through the compressed $k \times n$ matrix

$$G_0 + DG_1,$$

where D defines the time delay operator.

Following [4] we will say that two unit memory encoders: (G_0, G_1) and (G_0', G_1') are *equivalent* if there exists a nonsingular matrix T such that $G_0' = TG_0$ and $G_1' = TG_1$. Two equivalent encoders generate the same code. An encoder is called *catastrophic* if an information sequence with infinitely many nonzero information vectors produces an encoded sequence with finitely many nonzero encoded vectors. Thus two equivalent encoders are either both catastrophic or both noncatastrophic. We have the following criteria from [3]:

THEOREM 2.1. *[3] A UM encoder (G_0, G_1) is catastrophic if and only if there exists an $s \times k$ matrix P of rank $s, s > 0$ and a nonsingular $s \times s$ matrix Q such that*

$$QPG_0 = PG_1$$

Following the lines of [4] and [3] we define the *degree* μ of the encoder to be the rank of G_1. We consider the degree to be the third important parameter of a convolutional code $\mathcal{C}$ and in [10, 12] we define it in the general case of the convolutional codes of memory greater or equal to 1, as the maximal degree of the $k \times k$ full size minors of $G(D)$. (See [9] for details). If G_∞ denotes the high order coefficient matrix of a polynomial matrix $G(D)$, then we have that every code $\mathcal{C}$ of rate k/n has a $k \times n$ generator matrix $G(D)$ whose matrix G_∞ has rank k and whose row degrees are non-increasing or non-decreasing. The degree μ is in this case equal to the sum of the row degrees of the encoder $G(D)$ and we say that $G(D)$ is in *column proper form*. In the literature the degree μ is sometimes called the *total memory* of the code (see [5]) or *state-complexity*.

We therefore have that for any PUM code generated by (G_0, G_1) there exists an encoder (TG_0, TG_1) with T nonsingular such that the first $k - \mu$ rows of TG_1 are zero. We say that this encoder is in *standard form*. We say that a standard form encoder (G_0, G_1) is *minimal* if among all encoders, G_1 has the smallest number of nonzero rows. We have from [3]:

THEOREM 2.2. *[3] A noncatastrophic UM encoder* (G_0, G_1) *of the form:*

$$\begin{bmatrix} G_0 & G_1 \end{bmatrix} = \begin{bmatrix} G_0' & 0 \\ G_0'' & G_1'' \end{bmatrix},$$

where G_0'', G_1'' *have* μ *rows, is minimal if and only if:*

$$\operatorname{rank} \begin{bmatrix} G_1'' \\ G_0' \end{bmatrix} = k.$$

Unit memory codes having $\mu < k$ are called *partial unit memory codes* (PUM) since the encoder requires only μ memory cells for storage.

There are several distance functions that are important when deciding on the decoding properties of a convolutional code. The free distance d_{free} of the code, defined as the minimum Hamming weight of the nonzero encoded sequences having minimum weight, seems to be the most important. We will define also the jth column distances d_j^c and the jth row distances d_j^r. We follow the approach of [1, 2].

The jth *order column distance* d_j^c is defined as the minimum of the weights of the truncated codewords $v_{[0,j]} := (v_0, v_1, \ldots, v_j)$ resulting from an information sequence $u_{[0,j]} := (u_0, u_1, \ldots, u_j)$ with $u_0 \neq 0$. The tuple $d^{\mathbf{p}} = [d_0^c, d_1^c]$ is called the *distance profile*. The limit $d_\infty^c = \lim_{j\to\infty} d_j^c$ exists and we have $d_0^c \leq d_1^c \leq \ldots \leq d_\infty^c$.

The jth *row distance* d_j^r is defined as the minimum of the weights of all the finite codewords $v_{[0,j+1]} := (v_0, v_1, \ldots, v_{j+1})$ resulting from an

information sequence
$u_{[0,j]} := (u_0, u_1, \ldots, u_j) \neq 0$. The limit $d_\infty^r = \lim_{j\to\infty} d_j^r$ exists and, if the encoder is noncatastrophic, we have (see [11, 2] for details):

$$d_0^c \leq d_1^c \leq \ldots \leq d_\infty^c = d_{free} = d_\infty^r \leq \ldots \leq d_1^r \leq d_0^r. \tag{2.2}$$

In terms of state space descriptions d_∞^r is equal to the minimal weight of a nonzero trajectory which starts from and returns to the all zero state. d_∞^c is equal to the minimal weight of a nonzero trajectory which starts from and not necessarily returns to the all zero state. Also it follows that for a non-catastrophic encoder the minimal weight codewords are generated by finite information sequences, so the free distance can be computed from the weights of finite encoded sequences.

We will discuss now the PUM codes with degree $\mu = k-1$ and we will search for conditions they need to satisfy in order that they are optimal among the codes with the same parameters, in the sense that they attain the maximum distance possible. We will work first over the binary field and then over larger finite fields.

We have the following obvious bound on the free distance of an PUM code with degree $\mu = k-1$:

THEOREM 2.3. *Let $\mathcal{C}$ be a rate k/n PUM convolutional code of degree $\mu < k$ generated by a minimal encoder (G_0, G_1) over $\mathbb{F}$:*

$$\left[\begin{array}{cc} G_0 & G_1 \end{array}\right] = \left[\begin{array}{cccccc} g_1 & \ldots & g_n & 0 & \ldots & 0 \\ & G_0' & & & G_1' & \end{array}\right] \tag{2.3}$$

Then the free distance

$$d_{free} \leq n - k + \mu + 1.$$

This bound is a particular case of the more general bound studied in [10] and [12]:

LEMMA 2.4. *Let $\mathcal{C}$ be a convolutional code of rate k/n and degree μ and let $G(D)$ be a polymnomial encoder in row proper form.*

Let ν denote the smallest value of the row degrees of $G(D)$. Let l be the number of row degrees having the value equal to ν. Then the free distance must satisfy:

$$d_{free} \leq n(\nu+1) - l + 1. \tag{2.4}$$

We called this bound the generalized Singleton Bound and we showed in [10] that there are codes attaining this bound, over sufficiently large

finite fields. We called them MDS convolutional codes. In [12] we were able to give a concrete construction of MDS codes, starting from some Reed Solomon block codes.

Therefore we know that there are PUM codes of degree $k-1$ attaining the maximum bound, there is n, over some finite field with enough elements. We will call them PUM-MDS codes with $\mu = k-1$. We will take a different approach from [12] in the construction of such codes. We will start with the field $\mathbb{F}_2$ and discuss the cases when maximum distance codes exist, and also we will give a construction in these specific cases. Then we will generalize the construction for fields $\mathbb{F}_p, p > 2$ obtaining constructions in some other cases not covered yet.

We conclude the section with a simple theorem that tells us how to obtain k'/n rate PUM codes of degree $\mu = k' - 1$ and maximum distance $d_{free} = n$ from k/n rate PUM-codes of degree $\mu = k - 1$ and $d_{free} = n$ maximum, where $k' < k$.

THEOREM 2.5. *Let $\mathcal{C}$ be a PUM code of rate k/n generated by the minimal encoder (G_0, G_1) with $\mu = k - 1$. Let $(\bar{G}_0, \bar{G}_1) \in \mathbb{F}^{(k-1)\times 2n}$ be the matrix obtained from (G_0, G_1) by omitting any of the last $k-1$ rows of (G_0, G_1). If $\mathcal{C}$ has free distance n, then the same is true for the code $\bar{\mathcal{C}}$ generated by the encoder $(\bar{G}_0, \bar{G}_1)$.*

Proof. The theorem follows from the inclusion $\bar{\mathcal{C}} \subseteq \mathcal{C}$. □

3. Partial Unit Memory codes over $\mathbb{F}_2$. If $G_0,\ G_1$ generate a k/n PUM code of degree $\mu = k-1$ with maximum distance n over $\mathbb{F}_2$, then the matrices need to have the following form:

$$[\,G_0 \quad G_1\,] = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ & G_0' & & & G_1' & \end{bmatrix}, \text{ with } \operatorname{rank}(G_1') = k-1, \tag{3.1}$$

where G_0', G_1' need to satisfy some conditions that make the encoder G_0, G_1 noncatastrophic and minimal and the code generated by it through (2.1) optimal, i.e. MDS.

REMARK 3.1. It can be easily shown that if $2k - 1 \leq n$, the code is noncatastrophic provided that the matrix:

$$\begin{bmatrix} & G_1' & \\ 1 & \dots & 1 \\ & G_0' & \end{bmatrix} \text{ has full rank } 2k-1. \tag{3.2}$$

That assures the minimality as well.

For the next theorem we will need the following definition:

DEFINITION 3.1. A block code (k, n) is called *equidistant* if all nonzero codewords have the same weight d_{min}.

If a code is equidistant and G an arbitrary $k \times n$ encoder, then the entries of G have the property that all $\mathbb{F}$-linear combinations of its rows

have the same weight d_{min}. Such a matrix will be called an equidistant matrix.

Then we have the following theorem:

THEOREM 3.1. *Let (G_0, G_1) of the form (3.1) generate an UM-MDS code over $\mathbb{F}_2$. Then:*

1. n is even

2. G_0', G_1' generate equidistant $(k-1, n)$ block codes.

Proof. Let $u \in \mathbb{F}_2^{k-1},\ u \neq 0$ arbitrarily chosen. Let $x = \mathrm{wt}[uG_0'],\ y = \mathrm{wt}[uG_1']$. We need to prove that $x = y = n/2$. Let $u_1,\ u_{k+1} \in \mathbb{F}_2$.

Since $d_1^r = n$ we have that the weight of

$$(u_1, u, u_{k+1}) \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ & G_0' & & & G_1' & \\ & & & 1 & \dots & 1 \end{bmatrix}$$

is greater or equal to n. By giving different values to u_1, u_{k+1} we have:
$x + y \geq n, \quad n - x + y \geq n, \quad x + n - y \geq n, \quad n - x + n - y \geq n$
$\Rightarrow \quad x = y, \quad x + y = n$. Hence, we obtain that n is even and that

$$x = y = n/2,$$

which means that G_0', G_1' generate equidistant $(k-1, n)$ block codes. □

In the same way we proved that n is even we can prove that $2^{k-1} \mid n$. Hence $n = 2^{k-1}j$.

Actually we have the following straight forward lemma:

LEMMA 3.2. *The matrices G_0' and G_1' generate equidistant $(k, 2^{k-1})$ block codes if and only if the matrix*

$$\begin{bmatrix} G_0 & G_1 \end{bmatrix}$$

given through (3.1) is a generator matrix for a $(k+1, 2^k)$ equidistant block code.

Proof. Suppose G_0', G' are equidistant. If $u = (u_1, \dots, u_k) \in \mathbb{F}_2^k$ then uG_0 and uG_1 have the weight either n and respectively 0, if $(u_2, \dots, u_k) = 0$, or $n/2$, if not. Hence $\begin{bmatrix} G_0 & G_1 \end{bmatrix}$ is equidistant as well.

The other implication was just proved by the previous theorem 3.1. □

We therefore have a stronger statement:

THEOREM 3.3. *Suppose (G_0, G_1) of the form (3.1) generate a PUM code over $\mathbb{F}_2$. Suppose $2k - 1 \leq n$ and that condition (3.2) is satisfied (therefore the code is noncatastrophic). Then $\mathcal{C}$ is a noncatastrophic PUM-MDS convolutional code over $\mathbb{F}_2$ if and only if*

1. $n = 2^{k-1}j$.

2. G_0' and G_1' generate equidistant $(k-1, n)$ block codes.

In other words this statement gives us all the k/n MDS-PUM codes for $k \geq 4$ (so that $2k-1 \leq 2^{k-1}$).

Proof. Theorem 3.1 gives us the necessity implication. We still need to prove the sufficiency of the two conditions. From *2.* we have that $d_0^r = n$. Let $u_1, u_{k+1} \in \mathbb{F}_2$ and $u, v \in \mathbb{F}_2^{k-1}$, so that $(u_1, u) \neq 0$. The weight

$$\text{wt } (u_1, u, u_{k+1}, v) \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 & & & \\ & G_0' & & & G_1' & & & & \\ & & & 1 & \dots & 1 & 0 & \dots & 0 \\ & & & & G_0' & & & G_1' & \end{bmatrix} \geq$$

$$\geq \begin{cases} \text{wt } (u_1, u)G_0 + \text{wt } (v \cdot G_1') \geq n, \text{ if } v \neq 0 \\ \\ \text{wt } (u_1, u)G_0 + \text{wt } (u, u_{k+1}) \begin{bmatrix} & & G_1' & \\ 1 & \dots & 1 & \end{bmatrix} \geq n, \text{ if } v = 0 \end{cases},$$

because of condition (3.2). Hence $d_r^1 = n$. In the same way we show $d_i^r = n$. Also because of (3.2) we have:

$$d_0^c \geq n/2,\ d_1^c \geq n/2+1,\ d_2^c \geq n/2+2, \dots, d_{n/2}^c \geq n/2+n/2 = n \Rightarrow$$

$$\Rightarrow d_{free} = n.$$

Hence the code is MDS.

The noncatastrophicity is implied by the full rank condition on the $(2k-1, n)$ matrix. Due to this condition an infinite weight input can not produce a finite output. □

Therefore in order to construct rate $\frac{k}{2^{k-1}j}$ UPM codes with degree $\mu = k-1$ and maximum distance over $\mathbb{F}_2$, it is enough to construct rate $\frac{k}{2^{k-1}}$, $\mu = k-1$, $d_{free} = n$, MDS codes and concatenate them j times. From this, using the Theorem 2.5 we get PUM-MDS codes of rate $\frac{i}{2^{k-1}j}$, $1 \leq i \leq k$.

4. A binary construction of partial unit memory codes with maximum free distance. For the construction of PUM codes having maximal distance n over $\mathbb{F}_2$ we use an idea found in [4] but we will have a slightly different approach.

For that we need to introduce the following natural association:

REMARK 4.1. Through the following isomorphism of vector spaces:

$$\begin{array}{ccc} \mathbb{F}_2[X]/(X^{2^k}-1) & \longrightarrow & \mathbb{F}_2^{2^k} \\ a_0 + a_1 X + \ldots + a_{2^k-1} X^{2^k-1} & \longmapsto & (a_0, a_1, \ldots, a_{2^k-1}), \end{array} \tag{4.1}$$

any scalar encoded sequence in a PUM code $(v_0, v_1, v_2, \ldots)$, given through (2.1), where $v_i \in \mathbb{F}_2^{2^k}$ can be viewed as a polynomial encoded sequence:

$(v_0(X), v_1(X), v_2(X), \ldots)$, where all $v_i(X)$ are polynomials of degree at most $2^k - 1$.

Using the above isomorphism (4.1) we can also define an association between polynomial matrices $k \times 1$ and their coefficient matrices $k \times 2^k$:

$$A = \begin{bmatrix} a_{1,0} & a_{1,1} & \ldots & a_{1,2^k-1} \\ a_{2,0} & a_{2,1} & \ldots & a_{2,2^k-1} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k,0} & a_{k,1} & \ldots & a_{k,2^k-1} \end{bmatrix} \mapsto A(X) := \begin{bmatrix} a_{1,0} + a_{1,1}X + \ldots + a_{1,2^k-1}X^{2^k-1} \\ a_{2,0} + a_{2,1}X + \ldots + a_{2,2^k-1}X^{2^k-1} \\ \ldots \\ a_{k,0} + a_{k,1}X + \ldots + a_{k,2^k-1}X^{2^k-1} \end{bmatrix}. \tag{4.2}$$

With this association we have that

$$\mathrm{wt}\left[(u_1, \ldots, u_k)A\right] = \mathrm{wt}\left[(u_1, \ldots, u_k)A(X)\right], \forall (u_1, \ldots, u_k) \in \mathbb{F}^k.$$

It follows from Definition 3.2 and the above associations that an equidistant scalar matrix has the property that the associated polynomial matrix through (4.2) has all the polynomial entries of weight 2^{k-1} and any $\mathbb{F}_2$-linear combination of those polynomials gives another polynomial of the same weight. The weight of a polynomial is defined as the sum of the Hamming weights of all the coefficients.

Therefore instead of looking for $(k-1) \times (2^{k-1} - 1), k \geq 2$ equidistant scalar matrices G_0', G_1' we could instead look for $(k-1) \times 1$ polynomial matrices with the equivalent property. For this we will heavily use Lemmas 7.1 and 7.2 in the appendix. These lemmas will provide us such polynomial matrices. We have the following theorem:

THEOREM 4.1. *Let G_0', G_1' be $(k-1) \times (2^{k-1} - 1), k \geq 4$, scalar matrices associated to*

$$G_0'(X) := \begin{bmatrix} P_1(X) \\ P_2(X) \\ \ldots \\ P_{k-1}(X) \end{bmatrix}, \; G_1'(X) := \begin{bmatrix} Q_1(X) \\ Q_2(X) \\ \ldots \\ Q_{k-1}(X) \end{bmatrix}, \tag{4.3}$$

where all polynomials $P_i(X)$, $Q_j(X), i, j = \overline{1, k-1}$, have degree less or equal to $2^{k-1} - 2$. Then the rate $\frac{k}{2^{k-1}}$ PUM convolutional code generated by G_0, G_1 of the form in (3.1) is a noncatastrophic MDS code over $\mathbb{F}_2$ (i.e. it has maximal distance n) if and only if:

1. Any $\mathbb{F}_2$-linear combination of polynomials $P_1(X), \ldots, P_{k-1}(X)$ and any $\mathbb{F}_2$-linear combination of polynomials $Q_1(X), \ldots, Q_{k-1}(X)$ have weight 2^{k-1}.

2. The polynomials $P_1(X), \ldots, P_{k-1}(X), Q_1(X), \ldots, Q_{k-1}(X)$ are linearly independent.

Proof. The linear independence of the polynomials is equivalent to the noncatastrophicity of the code, condition given by 3.2 and the fact that all polynomials have degree strictly less than $2^{k-1} - 1$. ∎

The following lemma will give an inductive construction of PUM codes with maximal distance n over $\mathbb{F}_2$:

THEOREM 4.2. *Let $P_1(X), \ldots, P_{k-1}(X)$ be polynomials of degree less or equal to $2^{k-1}-2$ and weight 2^{k-2}. Moreover, suppose that any linear combination of the $k-1$ polynomials has also weight 2^{k-2}. Then the following polynomials:*

$$P_1(X)(X^{2^{k-1}}+1), \quad \ldots, \quad P_{k-1}(X)(X^{2^{k-1}}+1), \quad (X+1)^{2^{k-1}-1} \tag{4.4}$$

form a set of k polynomials with the property that any linear combination of the polynomials has degree less than 2^k and weight 2^{k-1}.

The same weight property holds for the set of k polynomials :

$$P_1(X)(X^{2^{k-1}}+1), \quad \ldots, \quad P_{k-1}(X)(X^{2^{k-1}}+1), \quad [X(X+1)]^{2^{k-1}-1} \tag{4.5}$$

Moreover if $Q_1(X), \ldots, Q_{k-1}(X)$ form also a set of $k-1$ polynomials of degree less or equal to $2^{k-1}-2$ with the same property that any linear combination of the polynomials has weight 2^{k-2} and if the polynomials

$$P_1(X), \quad \ldots, \quad P_{k-1}(X), \quad Q_1(X), \quad Q_2(X), \quad \ldots, \quad Q_{k-1}(X) \tag{4.6}$$

$$\text{and } 1+X+X^2+\ldots+X^{2^{k-1}-2}$$

are $\mathbb{F}_2$-linearly independent, then the polynomials:

$$P_1(X)(X^{2^{k-1}}+1), \ldots, P_{k-1}(X)(X^{2^{k-1}}+1), (X+1)^{2^{k-1}-1}, \tag{4.7}$$

$$Q_1(X)(X^{2^{k-1}}+1), \ldots, Q_{k-1}(X)(X^{2^{k-1}}+1), [X(X+1)]^{2^{k-1}-1}$$

are $\mathbb{F}_2$-linearly independent.

Proof. Let $P(X) = u_1P_1(X) + u_2P_2(X) + \ldots + u_{k-1}P_{k-1}(X), u_i \in \mathbb{F}_2, \forall i = \overline{1, k-1}$ be a linear combination of $P_1(X), P_2(X), \ldots, P_{k-1}(X)$. A linear combination of the new k polynomials has the form:

$$u(X+1)^{2^{k-1}-1} + P(X)(X^{2^{k-1}}+1) = u(X+1)^{2^{k-1}-1} + P(X)(X+1)^{2^{k-1}} =$$

$$= (X+1)^{2^{k-1}-1}(u+P(X)(X+1)), \text{ with } u \in \mathbb{F}_2, \text{ or, in the second situation:}$$

$$u[X(X+1)]^{2^{k-1}-1} + P(X)(X^{2^{k-1}}+1) =$$

$$= (X+1)^{2^{k-1}-1}(uX^{2^{k-1}-1} + P(X)(X+1)).$$

If $u=0$ we obtain $P(X)(X+1)^{2^{k-1}}$ that has weight twice the weight of $P(X)$ as we stated before in the lemma 7.4. If $u=1$ we use the weight retaining property (7.3):

$$\text{wt}\left[(X+1)^{2^{k-1}-1}(u+P(X)(X+1))\right] \geq$$

$$\geq \text{wt}\left[(X+1)^{2^{k-1}-1}\right] \cdot \text{wt}\left[(u+P(X)(X+1))mod\ (X+1)\right] = 2^{k-1}.$$

The second case goes the same way.

For the second part let $Q(X) = v_1Q_1(X)+\ldots+v_{k-1}Q_{k-1}(X), v_i \in \mathbb{F}_2$, $\forall i = \overline{1,k-1}$ be a linear combination of $Q_1(X), Q_2(X), \ldots, Q_{k-1}(X)$. Let

$$(X+1)^{2^{k-1}-1}(u+P(X)(X+1))+(X+1)^{2^{k-1}-1}(vX^{2^{k-1}-1}+Q(X)(X+1)) =$$

$$= (X+1)^{2^{k-1}-1}(u+vX^{2^{k-1}-1}+(Q(X)+P(X))(X+1)) = 0,\ u,v \in \mathbb{F}_2,$$

be a linear combination of the new polynomials that is equal to zero. It implies $u = v$ and we obtain:

$$u(1+X^{2^{k-1}-1})+(Q(X)+P(X))(X+1) = 0 \Leftrightarrow$$

$$u(1+X+X^2+\ldots+X^{2^{k-1}-2})+Q(X)+P(X) = 0,$$

which leads to $u = u_1 = \ldots = u_{k-1} = v_1 = \ldots = v_{k-1} = 0$ because of (4.7). That gives the linear independence of the new polynomials. □

Basically, Theorem 4.2 says that if we have two equidistant matrices G_0' and G_1' of sizes $(k-1)\times(2^{k-1}-1),\ k \geq 4$ associated to the polynomial matrices $G_0'(X), G_1'(X)$ through (4.3), where the sets of polynomials $P_1(X), \ldots, P_{k-1}(X)$ and $Q_1(X), \ldots, Q_{k-1}(X)$ satisfy the conditions in Theorem 4.2, we can inductively construct equidistant matrices of size $j \times (2^j-1),\ j \geq k$.

For example, if we take 1 ($k = 2$), multiply it with (X^2+1) and add the extra polynomial $1+X$, respectively $X(1+X)$ we obtain the 2×4 matrices:

$$G_0'(X) = \begin{bmatrix} 1+X \\ 1+X^2 \end{bmatrix},\ G_1'(X) = \begin{bmatrix} (1+X)X \\ (1+X^2) \end{bmatrix},$$

and the 3×8 matrices, after the next step:

$$G_0'(X) = \begin{bmatrix} (1+X)^3 \\ (1+X)^5 \\ (1+X)^6 \end{bmatrix},\ G_1'(X) = \begin{bmatrix} (1+X)^3X^3 \\ (1+X)^5X \\ (1+X)^6 \end{bmatrix}.$$

Of course this is not a good choice, since the polynomials obtained are not linearly independent, the code generated in this way being catastrophic. Therefore we have to change somehow these matrices in order to have the properties of Theorem 4.2. We will keep the matrix $G_0'(X)$ and change

the matrix $G_1'(X)$ by multiplying the entries with different powers of X modulo X^7+1. The following choice for $G_1'(X)$:

$$G_1'(X) = \begin{bmatrix} (1+X)^3 \cdot X^4 mod\ (X^7-1) \\ (1+X)^5 \cdot X^3 mod\ (X^7-1) \\ (1+X)^6 \cdot X^3 mod\ (X^7-1) \end{bmatrix}.$$

together with the $G_0'(X)$ constructed above will satisfy the condition of the theorem. Hence we could use the polynomial entries of $G_0'(X), G_1'(X)$ for the inductive construction of Theorem 4.2. We have the following construction theorem:

THEOREM 4.3. *Let* $P_1 = (1+X)^3,\ P_2 = (1+X)^5,\ P_3 = (1+X)^6$ *and* $Q_1 = (1+X)^3 \cdot X^4 mod\ (X^7-1),\ Q_2 = (1+X)^5 \cdot X^3 mod\ (X^7-1),\ Q_3 = (1+X)^6 \cdot X^3 mod\ (X^7-1)$.

Applying Theorem 4.2 inductively we obtain rate $\frac{k}{2^{k-1}}$ *noncatastrophic convolutional codes that have maximal free distance* 2^{k-1} *over* $\mathbb{F}_2$*, for all* $k \geq 4$.

REMARK 4.2. The rate $\frac{k}{2^{k-1}}$ code constructed above has the matrix G_0' associated to the following polynomial matrix:

$$G_0'(X) = \begin{bmatrix} (X+1)^{i_1} \\ (X+1)^{i_2} \\ \dots \\ (X+1)^{i_{k-1}} \end{bmatrix}$$

with $i_1, i_2, \dots, i_{k-1}$ nonnegative integer strictly less than 2^{k-1} of weight $k-2$, where we defined the weight of an integer in (7.1). We could apply (7.2) to show directly that the matrix G_0' generates an equidistant $(k-1, 2^{k-1})$ block code. We will use this direct approach rather than the inductive one, in the following section, for constructing MDS convolutional codes of rate k/n where n is odd. Of course we will have to use a larger field.

5. Constructions of partial unit memory codes with maximum free distance over $\mathbb{F}_p$. Let $\mathbb{F}_p$ be the field with p elements. Let $k \geq 1,\ n = p^{k-1}$.

THEOREM 5.1. *Let* G_0, G_1 *be the* $k \times n$ *scalar matrices associated to the following polynomial matrices:*

$$G_0(X) = \begin{bmatrix} (X+1)^{i_0} \\ (X+1)^{i_1} \\ \dots \\ (X+1)^{i_{k-1}} \end{bmatrix},\ G_1(X) = \begin{bmatrix} 0 \\ (X+1)^{j_1} \\ \dots \\ (X+1)^{j_{k-1}} \end{bmatrix}$$

with $i_0 = (p-1) + (p-1)p + \ldots + (p-1)p^{k-2} = p^{k-1} - 1 = n-1$,
$I := \{i_1, \ldots i_{k-1}\}$ *the set of all nonnegative integers with radix-p form (see*

Lemma 7.1) having one component equal to $p-2$ and the other $k-2$ components equal to $p-1$, and $J := \{j_1, \ldots, j_{k-1}\}$, the set of all nonnegative integers having one component equal to 0 and the other $k-2$ components equal to $p-1$. Both sets have $\binom{k-1}{k-2} = k-1$ elements. Then the convolutional code generated by G_0, G_1 over $\mathbb{F}_p$ is noncatastrophic and MDS.

Proof. We compute d_0^c and d_1^c.

By (7.1) we have: $\text{wt}[(X+1)^{i_l}] = \begin{cases} (p-1)p^{k-2}, & l \neq 0 \\ p^{k-1}, & l = 0 \end{cases}$ and $\text{wt}[(X+1)^{j_l}] = p^{k-2}$.

Let $u = (u_0, \ldots, u_{k-1}) \in \mathbb{F}_p^k, u \neq 0$. Then:

$$\text{wt}[uG_0] = \text{wt}[uG_0(X)] = \text{wt}\left[\sum_{l=0}^{k-1} u_l(X+1)^{i_l}\right] \geq$$

$$\geq \text{wt}\left[(X+1)^{i_{min}}\right] \geq (p-1)p^{k-2},$$

by (7.2). We denoted by i_{min} the smallest of all integers $i_l, l \in \{0, \ldots, k-1\}$ with, $u_l \neq 0$. Therefore $d_0^c \geq (p-1)p^{k-2}$ and since there is a row of this weight we have:

$$d_0^c = (p-1)p^{k-2}.$$

For d_1^c we do the same. Let $u = (u_0, \ldots, u_{2k-1}), \in \mathbb{F}_p^{2k-1}, u \neq 0$. If $(u_1, \ldots, u_{2k-1}) = 0$, we obtain the codeword associated to $u_0(X+1)^{i_0}$ which has weight p^{k-1} by the choice of i_0. If $(u_1, \ldots, u_{2k-1}) \neq 0$ then the weight

$$\text{wt } u \begin{bmatrix} G_1 \\ G_0 \end{bmatrix} = \text{wt } u \begin{bmatrix} G_1(X) \\ G_0(X) \end{bmatrix} =$$

$$\text{wt}\left[\sum_{s=1}^{k-1} u_l(X+1)^{j_s} + \sum_{l=1}^{k-1} u_{l+k-1}(X+1)^{i_l}\right] \geq p^{k-2},$$

by (7.2), since all the powers $i_l,\ l = \overline{0, k-1}$ differ from $j_s,\ s == \overline{1, k-1}$. Then

$$d_1^c \geq (p-1)p^{k-2} + p^{k-2} = p^{k-1} = n.$$

Therefore $d_1^c = d_{free} = n$ and the code is noncatastrophic and MDS. □

REMARK 5.1. Since the main fact we used was that the sum

$$\text{wt}[(X+1)^{i_l}] + \text{wt}[(X+1)^{j_s}] \geq p^{k-1},$$

for any $l = \overline{0, k-1},\ s = \overline{1, k-1}$, we could use instead in the construction the sets I and J, with I and J formed by all nonnegative integers having the radix-p form with one component equal to $p-i$, respectively $i-2$, and the

rest $k-2$ components equal to $p-1$, for all i such that $p-i > i-2$, i.e. for all $i = \overline{2, \lfloor \frac{p+2}{2} \rfloor}$. The weights $\mathrm{wt}[(X+1)^{i_l}] = \begin{cases} (p-i+1)p^{k-2}, & l \neq 0 \\ p^{k-1}, & l = 0 \end{cases}$ and $\mathrm{wt}[(X+1)^{j_l}] = (i-1)p^{k-2}$ have also the sum greater than p^{k-1}. Also the sets I and J formed like this have both $k-1$ elements as it is needed.

6. Examples. We will give here two concrete examples to show how Theorems 4.3 and 5.1 are applied, the rate 4/8 and 3/9. After that we will discuss also the cases $k = 2, k = 3$ that have not been covered by the binary theorem. We will use here the polynomial matrix representation $G(D) = G_0 + DG_1$.

EXAMPLE 1. We already showed in the previous section how to choose the matrices G_0, G_1 of sizes 4×8 over $\mathbb{F}_2$. In conformity with Theorem reffinal we have that the polynomial matrix $G(D) = G_0 + DG_1$, given by:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1+D & 1 & 1 & 1 & D & D & D & 0 \\ 1+D & 1+D & 0 & D & 1+D & 1 & 0 & 0 \\ 1+D & 0 & 1+D & D & 1 & D & 1 & 0 \end{bmatrix}$$

generates a rate 4/8 PUM convolutional code of degree $\mu = 3$ and maximum distance 8.

EXAMPLE 2. Let G_0, G_1 be the 3×9 matrices over $\mathbb{F}_3$ associated to the following polynomial matrices:

$$\begin{bmatrix} (X+1)^8 \\ (X+1)^5 \\ (X+1)^7 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ (X+1)^2 \\ (X+1)^6 \end{bmatrix}.$$

The convolutional code generated by

$$G(D) = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1+D & -1 & 1 & 1-D & -1 & 1 & D & 0 & 0 \\ 1+D & 1-D & D & -1 & -1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

is noncatastrophic, has degree $\mu = 2$ and maximum distance 9.

EXAMPLE 3. We will construct rate $2/n$ PUM convolutional codes that are MDS and noncatastrophic, for all $n \geq 3$.

1. In the case n even we can do the construction over the binary field. Let:

$$G(D) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \underbrace{1+D \quad 1 \quad 0 \quad D}_{j \text{ times}} \end{bmatrix}, G(D) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \underbrace{1+D \quad 1 \quad 0 \quad D}_{j \text{ times}} & 1+D & 0 \end{bmatrix},$$

for $n = 4j$, respectively $n = 4j + 2$. Then the $2/n$ code generated by $G(D)$ is noncatastrophic and has distance n over $\mathbb{F}_2$. The code has the column distances: $d_0^c = n/2, d_1^c = n = d_{free}$ in both cases.

2. The cases where n is odd requires more field elements. It turns out that a field with 3 elements is enough for a construction. Therefore, over $\mathbb{F}_3$ we obtain:

$$G(D)=\left[\begin{array}{cccccc} 1 & 1 & 1 & 1 & & 1 \\ \underbrace{1+D \quad 1 \quad 0 \quad D}_{j \text{ times}} & & & & & D+2 \end{array}\right], G(D)=\left[\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & & 1 & 1 \\ \underbrace{1+D \quad 1 \quad 0 \quad D}_{j \text{ times}} & & & & 2+D & & 1+D & 0 \end{array}\right],$$

for $n = 4j + 1$, respectively $n = 4j + 3$. The column distances are in both cases:
$d_0^c = \lfloor n/2 \rfloor + 1,\ d_1^c = n = d_{free}$.

EXAMPLE 4. The construction Theorem 3.3 can not be applied in the case of $k = 3$. It turns out that any choice of binary matrices $G_0,\ G_1$ we take gives a catastrophic encoder. Therefore there is no noncatastrophic PUM convolutional code of rate $3/4$, degree 2, having distance 4. The smallest field we can construct such a $3/4$, code with degree 2, distance 4 is $\mathbb{F}_3$. Taking

$$G_0 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},\ G_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

we obtain an MDS code but over a field of characteristic $p \neq 2$. The column distances are $d_0^c = 2,\ d_1^c = 2,\ d_2^c = 3,\ d_3^c = 3,\ d_4^c = 4 = d_{free}$.

7. Appendix. We state here some results that we need along the paper. For more details see [6].

LEMMA 7.1. *[6] Let $c \in \mathbb{F}, c \neq 0$ and let $i \geq 1$ with radix-p form $[i_0, i_1, \ldots, i_{m-1}]$, i.e. $i = i_0 + i_1 p + \ldots + i_{m-1} p^{m-1}$. Then:*

$$\text{wt}[(X + c)^i] = \prod_{j=0}^{m-1} (i_j + 1). \tag{7.1}$$

In particular, for $p = 2$,

$$\text{wt}[(X + 1)^i] = 2^{\text{wt}(i)}, \tag{7.2}$$

where $\text{wt}(i)$ is the number of 1's in $\{i_0, i_1, \ldots, i_{m-1}\}$.

LEMMA 7.2. *[6] Let I be any nonempty finite set of nonnegative integers with least integer i_{min} and let*

$$P(X) = \sum_{i \in I} b_i (X - c)^i,$$

where $c, b_i \in \mathbb{F}$, *all nonzero. Then:*

$$\text{wt}[P(X)] \geq \text{wt}[(X+c)^{i_{min}}]. \tag{7.3}$$

LEMMA 7.3. *[6] For any polynomial* $P(X)$ *over* $\mathbb{F}$, *any* $c \in \mathbb{F}, c \neq 0$, *and any nonnegative integers* n *and* N,

$$\text{wt}\left[P(X)(X^n+c)^N\right] \geq \text{wt}\left[(X+c)^N\right] \text{wt}\left[P(X) mod\ (X^n-c)\right]. \tag{7.4}$$

The following lemma gives a very obvious result that we need for the constructions. It could be also seen as a corollary to Lemma 7.3:

LEMMA 7.4. *If* $P(X)$ *is a polynomial over* $\mathbb{F}_2$ *of degree less or equal to* 2^k-1, *then the weight* $\text{wt}\left[P(X)(X^{2^k}+1)\right] = 2wt[P(X)]$.

REFERENCES

[1] R. Johannesson and K. Zigangirov. Distances and distance bounds for convolutional codes – an overview. In *Topics in Coding Theory. In honour of L. H. Zetterberg.*, Lecture Notes in Control and Information Sciences # 128, pages 109–136. Springer Verlag, 1989.

[2] R. Johannesson and K. Sh. Zigangirov. *Fundamentals of Convolutional Coding.* IEEE Press, New York, 1999.

[3] J. Justesen, E. Paaske, and M. Ballan. Quasi-cyclic unit memory convolutional codes. *IEEE Trans. Inform. Theory*, IT-36(3):540–547, 1990.

[4] G.S. Lauer. Some optimal partial-unit-memory codes. *IEEE Trans. Inform. Theory*, 25:240–243, 1979.

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

[6] J. L. Massey, D. J. Costello, and J. Justesen. Polynomial weights and code constructions. *IEEE Trans. Inform. Theory*, IT-19(1):101–110, 1973.

[7] R. J. McEliece. The algebraic theory of convolutional codes. In V. Pless and W.C. Huffman, editors, *Handbook of Coding Theory*, volume 1, pages 1065–1138. Elsevier Science Publishers, Amsterdam, The Netherlands, 1998.

[8] Ph. Piret. *Convolutional Codes, an Algebraic Approach.* MIT Press, Cambridge, MA, 1988.

[9] J. Rosenthal, J. M. Schumacher, and E. V. York. On behaviors and convolutional codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1881–1891, 1996.

[10] J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Appl. Algebra Engrg. Comm. Comput.*, 10(1):15–32, 1999.

[11] J. Rosenthal and E. V. York. BCH convolutional codes. *IEEE Trans. Inform. Theory*, 45(6):1833–1844, 1999.

[12] R. Smarandache, H. Gluesing-Luerssen, and J. Rosenthal. Constructions of MDS-convolutional codes. Submitted to IEEE Trans. Inform. Theory, August 1999.