

# Strongly-MDS Convolutional Codes

Heide Gluesing-Luerssen, Joachim Rosenthal, *Senior Member, IEEE*, and Roxana Smarandache, *Member, IEEE*

**Abstract**—Maximum-distance separable (MDS) convolutional codes have the property that their free distance is maximal among all codes of the same rate and the same degree. In this paper, a class of MDS convolutional codes is introduced whose column distances reach the generalized Singleton bound at the earliest possible instant. Such codes are called strongly-MDS convolutional codes. They also have a maximum or near-maximum distance profile. The extended row distances of these codes will also be discussed briefly.

**Index Terms**—Column distances, convolutional codes, extended row distances, maximum-distance separable (MDS) codes, super-regular matrices, unit memory codes.

## I. INTRODUCTION

IN comparison to the literature on linear block codes, there exist only relatively few algebraic constructions of convolutional codes having some good designed distance. There are even fewer algebraic decoding algorithms which are capable of exploiting the algebraic structure of the code.

A large part of the literature on convolutional codes studies codes over the binary field. Codes are then typically presented by trellis and state diagrams and the decoding algorithm of choice is the Viterbi algorithm. The reader is referred to the standard books [1], [2] or the more recent articles [3]–[5] where also further references to the literature can be found.

In the early 1970s, there were some important constructions done for convolutional codes over large alphabets and we would like to mention the papers [6]–[9] and the monograph by Piret [10]. In [7], [8], Justesen and Hughes study the question on how large the free distance of a convolutional code over a large alphabet can be. In [11], the authors of the present paper derived a generalized Singleton bound and they define a convolutional code to be maximum-distance separable (MDS) if its free distance reaches this upper bound. Using a construction idea due to Justesen [6], they provided in [12] a first concrete construction of MDS convolutional codes for all rates and degrees.

Manuscript received March 20, 2003; revised October 14, 2005. This work was supported in part by the National Science Foundation under Grants DMS-0072383 and CCR-0205310. The material in this paper was presented at the 2002 IEEE International Symposium on Information Theory, Lausanne, Switzerland, June/July 2002 and at the International Symposium on the Mathematical Theory of Networks and Systems (MTNS), University of Notre Dame, Notre Dame, IN, August 2002.

H. Gluesing-Luerssen is with the Department of Mathematics, University of Groningen, 9700 AV Groningen, The Netherlands (e-mail: gluesing@math.rug.nl).

J. Rosenthal is with the Department of Mathematics, University of Zürich, CH-8057 Zürich, Switzerland (e-mail: rosen@ieee.org).

R. Smarandache is with the Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182-7720 USA (e-mail: rsmarand@sciences.sdsu.edu).

Communicated by J. Justesen, Associate Editor for Coding Techniques.

Digital Object Identifier 10.1109/TIT.2005.862100

In [13], the class of cyclic convolutional codes, as first introduced by Piret [14] and Roos [15], has been studied. It turned out that many of the constructed codes were MDS and/or were codes over large alphabets with excellent distance profile. More recently, Goppa convolutional codes have been introduced in [16] and many of these codes have excellent distances as well.

In this paper, we will consider the problem of constructing codes with rapid growth of their column distances. These distance parameters have been introduced by Costello in [17] and have been studied by many authors since. See, e.g., [1], [18], and the references therein.

Algorithmic searches for codes with large column distances have been carried out for instance in [19], [20]. We will now attack the question of convolutional codes with large column distances from the theoretical point of view. First, we will discuss how big these distances can possibly be, thereby introducing a new class of convolutional codes which we call strongly-MDS convolutional codes. These are codes with a column distance profile such that the free distance, i.e., the generalized Singleton bound, is reached at the earliest possible stage. The main part of the paper will be about existence and construction of such codes. At the end of the paper, we will also use our methods in order to briefly discuss the extended row distances of these codes in case they have unit memory. As opposed to the column distances, the extended row distances also grow beyond the free distance and thus contain some additional information about the performance of the code.

Convolutional codes over large alphabets are interesting both from a purely theoretical as well as from an applications point of view. On the theory side, the following questions arise naturally: How large can be the free distance of a convolutional code of some fixed rate and fixed degree? How to construct codes which achieve a maximal free distance? How good can be the column and row distance profile of these codes?

From a practical point of view, we can identify a convolutional code over a finite alphabet with a finite linear state machine (LFSM) having redundancy and which is capable of correcting processing errors. In a series of recent papers [21]–[23], Hadjicostis and Verghese showed how to error protect a given LFSM with a larger redundant LFSM capable of detecting and correcting state transition errors. The detection and correction of errors is done using nonconcurrent measurements of the state vectors of the redundant LFSM. The construction of the redundant system boils down to the construction of convolutional codes with good free distance over an alphabet which is in general nonbinary.

Let  $\mathbb{F}$  be any finite field and let  $\mathbb{F}[D]$  be the polynomial ring. For the definition of a convolutional code we take a module theoretic point of view [24], [25].

*Definition 1.1:* A convolutional code of rate  $k/n$  is a submodule  $\mathcal{C} \subseteq \mathbb{F}[D]^n$  of rank  $k$  such that there exists a  $k \times n$

polynomial encoder matrix  $G \in \mathbb{F}[D]^{k \times n}$  which is *basic*, i.e.,  $G$  has a polynomial right inverse, and which is *minimal*, i.e., the sum of the row degrees of  $G$  attains the minimal possible value, such that

$$\mathcal{C} := \{uG \mid u \in \mathbb{F}[D]^k\} \subseteq \mathbb{F}[D]^n.$$

We define the *degree* of  $\mathcal{C}$  as the sum of the row degrees of one, and hence any, minimal basic encoder.

In the sequel, we will adopt the notation used by McEliece [26, p. 1082] and call a convolutional code of rate  $k/n$  and degree  $\delta$  an  $(n, k, \delta)$  code. Every  $(n, k, \delta)$  code  $\mathcal{C}$  can be presented in terms of a parity-check matrix  $H \in \mathbb{F}[D]^{(n-k) \times n}$ , i. e.,

$$\mathcal{C} = \{v \in \mathbb{F}[D]^n \mid vH^T = 0\}.$$

It is clear that we can choose  $H$  to be basic and minimal and we will do so throughout the paper. Notice also that  $GH^T = 0$  for any generator matrix  $G$  of  $\mathcal{C}$ .

The *weight* of a vector  $v = \sum_{j=0}^L v_j D^j \in \mathbb{F}[D]^n$  is defined as

$$\text{wt}(v) := \sum_{j=0}^L \text{wt}(v_j) \in \mathbb{N}_0$$

where  $\text{wt}(v_j)$  denotes the Hamming weight of  $v_j \in \mathbb{F}^n$ . The *free distance* of the code  $\mathcal{C} \subset \mathbb{F}[D]^n$  is given as

$$d_{\text{free}} := \min\{\text{wt}(v) \mid v \in \mathcal{C}, v \neq 0\}.$$

Since we assume that  $G \in \mathbb{F}[D]^{k \times n}$  is minimal and basic, we also have

$$d_{\text{free}} = \min\{\text{wt}(v) \mid v = uG \text{ for some } u \in \mathbb{F}[D]^k \setminus \{0\}, u_0 \neq 0\}.$$

In Theorem 2.6, we will recall from the paper [11] an upper bound on  $d_{\text{free}}$  based on the parameters  $(n, k, \delta)$ . It generalizes the Singleton bound from block code theory and will play a central role in our paper.

The paper is structured as follows: In Section II, we review notions from convolutional coding theory such as column distances, the generalized Singleton bound and we introduce the important concepts for this paper, namely, the property of being strongly-MDS and having a maximum distance profile. In Section III, we show the existence of strongly-MDS codes of rate  $k/n$  with  $n-k \mid \delta$ . In order to do so, we introduce the interesting concept of a *superregular matrix* which could be of independent interest. In Section IV, we illustrate these concepts through a series of examples. In Section V, we investigate to what extent properties of MDS, strongly-MDS, and maximum distance profile carry over to the dual code. The main result of this section states that a code has a maximum distance profile if and only if its dual has this property. This allows us to show that for certain specific parameters, a code is strongly-MDS if and only if its dual is strongly-MDS. Finally, in Section VI, we will derive a lower bound for the extended row distances.

## II. STRONGLY-MDS CODES AND CODES WITH MAXIMUM DISTANCE PROFILE

Let  $\mathcal{C} \subseteq \mathbb{F}[D]^n$  be an  $(n, k, \delta)$  code with a basic minimal generator matrix

$$G = \sum_{j=0}^{\nu} G_j D^j \in \mathbb{F}[D]^{k \times n}, G_j \in \mathbb{F}^{k \times n}, G_{\nu} \neq 0 \quad (2.1)$$

and a basic parity-check matrix

$$H = \sum_{j=0}^{\mu} H_j D^j \in \mathbb{F}[D]^{(n-k) \times n}, H_j \in \mathbb{F}^{(n-k) \times n}, H_{\mu} \neq 0. \quad (2.2)$$

Notice that  $\nu$  is the memory of the code. For every  $j \in \mathbb{N}_0$ , the truncated sliding generator matrices  $G_j^c \in \mathbb{F}^{(j+1)k \times (j+1)n}$  and parity-check matrices  $H_j^c \in \mathbb{F}^{(j+1)(n-k) \times (j+1)n}$  are given by

$$G_j^c := \begin{bmatrix} G_0 & G_1 & \dots & G_j \\ & G_0 & \dots & G_{j-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{bmatrix} \quad (2.3)$$

$$H_j^c := \begin{bmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \dots & H_0 \end{bmatrix}$$

where we let  $G_j = 0$  (resp.,  $H_j = 0$ ) whenever  $j > \nu$  (resp.,  $j > \mu$ ); see also [1, p. 110]. The identity  $GH^T = 0$  and the full ranks of  $G_0$  and  $H_0$  immediately imply the full rank of the sliding matrices as well as the identities

$$\{uG_j^c \mid u \in \mathbb{F}^{(j+1)k}\} = \{v \in \mathbb{F}^{(j+1)n} \mid v(H_j^c)^T = 0\}$$

for all  $j \in \mathbb{N}_0$ . Using these equations and following [1, pp. 110], the  $j$ th *column distance* of  $\mathcal{C}$  is given as (2.4)–(2.5) at the bottom of the page. The column distances are invariants of the code, i.e., they do not depend on the choice of the generator matrix (see [1, Sec. 3.1]), and satisfy

$$d_0^c \leq d_1^c \leq d_2^c \dots \leq \lim_{j \rightarrow \infty} d_j^c = d_{\text{free}}. \quad (2.6)$$

The  $(\nu + 1)$ -tuple of numbers  $(d_0^c, d_1^c, d_2^c, \dots, d_{\nu}^c)$ , where  $\nu$  is the memory, is called the *column distance profile* of the code [1, p. 112].

Equation (2.5) immediately implies the following simple fact used several times in the paper.

*Proposition 2.1:* Let  $d \in \mathbb{N}$ . Then the following properties are equivalent:

- $d_j^c = d$ ;
- none of the first  $n$  columns of  $H_j^c$  is contained in the span of any other  $d - 2$  columns and one of the first  $n$  columns of  $H_j^c$  is in the span of some other  $d - 1$  columns of that matrix.

$$d_j^c = \min\{\text{wt}((u_0, \dots, u_j)G_j^c) \mid u_i \in \mathbb{F}^k, u_0 \neq 0\} \quad (2.4)$$

$$= \min\{\text{wt}(\hat{v}) \mid \hat{v} = (\hat{v}_0, \dots, \hat{v}_j) \in \mathbb{F}^{(j+1)n}, \hat{v}(H_j^c)^T = 0, \hat{v}_0 \neq 0\}. \quad (2.5)$$

We leave it to the reader to verify the equivalence of the statements. The following upper bound on the column distances is an immediate consequence of the previous result, along with the fact that  $H_j^c$  has full row rank.

*Proposition 2.2:* For every  $j \in \mathbb{N}_0$  we have

$$d_j^c \leq (n-k)(j+1) + 1.$$

This observation has already been made in the context of systematic convolutional codes in [8] and in [27]. The next proposition shows that maximality of  $d_j^c$  implies maximality of the preceding column distances.

*Corollary 2.3:* If  $d_j^c = (n-k)(j+1) + 1$  for some  $j \in \mathbb{N}_0$ , then  $d_i^c = (n-k)(i+1) + 1$  for all  $i \leq j$ .

*Proof:* It suffices to prove the assertion for  $i = j-1$ . In order to do so notice that

$$H_j^c = \left[ \begin{array}{c|c} H_{j-1}^c & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline H_j & H_0 \end{array} \right]$$

and assume that one of the first  $n$  columns of  $H_{j-1}^c$  is in the span of some other  $(n-k)j-1$  columns. Then  $\text{rank } H_0 = n-k$  implies that one of the first  $n$  columns of  $H_j^c$  is in the span of some other  $(n-k)j-1 + n-k = (n-k)(j+1) - 1$  columns of  $H_j^c$ . But this is a contradiction to the optimality of  $d_j^c$  by Proposition 2.1.  $\square$

The following characterizations of the  $j$ th column distance being maximal will be useful later on for constructing strongly-MDS codes and also when considering duality in Section V.

*Theorem 2.4:* Let  $G_j^c$  and  $H_j^c$  be as in (2.3). Then the following are equivalent:

- i)  $d_j^c = (n-k)(j+1) + 1$ ;
- ii) every  $(j+1)k \times (j+1)k$  full-size minor of  $G_j^c$  formed from the columns with indices  $1 \leq t_1 < \dots < t_{(j+1)k}$ , where  $t_{sk+1} > sn$  for  $s = 1, \dots, j$ , is nonzero;
- iii) every  $(j+1)(n-k) \times (j+1)(n-k)$  full-size minor of  $H_j^c$  formed from the columns with indices  $1 \leq r_1 < \dots < r_{(j+1)(n-k)}$ , where  $r_{s(n-k)} \leq sn$  for  $s = 1, \dots, j$ , is nonzero.

Notice that the index condition in part ii) simply says that for each  $s$ , the minors under consideration have at most  $sk$  columns out of the first  $sn$  columns of  $G_j^c$ . Clearly, all other full-size minors of  $G_j^c$  are singular. The proof of this purely matrix theoretical result will be given in Appendix A. The next corollary will provide a link of our results to the existing literature later in this section. It will also be helpful for the construction of codes with maximum column distances in the next section.

*Corollary 2.5:* Let  $G = [I, P]$  be a systematic matrix of memory  $\nu$  and let  $P = \sum_{i=0}^{\nu} P_i D^i \in \mathbb{F}[D]^{k \times (n-k)}$  where  $P_\nu \neq 0$ . Then the following are equivalent for  $j = 0, \dots, \nu$ :

- i)  $d_j^c = (n-k)(j+1) + 1$ ;

- ii) each  $i \times i$ -submatrix of

$$\hat{P} := \begin{bmatrix} P_0 & P_1 & \dots & P_j \\ & P_0 & \dots & P_{j-1} \\ & & \ddots & \vdots \\ & & & P_0 \end{bmatrix} \in \mathbb{F}^{(j+1)k \times (j+1)(n-k)}$$

that does not contain an  $s \times t$ -zero block for some  $s, t$  such that  $s+t \geq i+1$  is nonsingular. If one of these conditions is satisfied, then all entries of the matrices  $P_0, \dots, P_j$  are nonzero.

This result appeared already in [27]. However, since the paper is not easily accessible and no detailed proof is provided, we added a proof in Appendix A. We will come back to the relation of our work with [27] at the end of this section.

In the sequel, we will relate the upper bound for the column distances to an upper bound for the free distance of the code. The maximum possible value for the free distance of a convolutional code over any field has been established in [11]. Therein the following has been shown.

*Theorem 2.6:* The free distance of an  $(n, k, \delta)$  code satisfies

$$d_{\text{free}} \leq (n-k) \left( \left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1. \quad (2.7)$$

Notice that if  $\delta = 0$ , the number on the right-hand side of (2.7) reduces to the usual Singleton bound  $n-k+1$  for an  $(n, k)$  block code. Therefore, we call this number the *generalized Singleton bound* and codes whose distance attains this bound will be called *MDS codes*. It has been shown in [11] that for every set of parameters  $(n, k, \delta)$  and every prime number  $p$  there exists a suitably large finite field  $\mathbb{F}$  of characteristic  $p$  and an MDS code with parameters  $(n, k, \delta)$  over  $\mathbb{F}$ . The proof is based on techniques from algebraic geometry and is nonconstructive. In [12], a construction of MDS codes with parameters  $(n, k, \delta)$  was given for suitably large fields of characteristic coprime with  $n$ .

In the sequel, we will strengthen the MDS property by requiring that the generalized Singleton bound is attained by the earliest column distance possible. This will lead us to the notion of a strongly-MDS code.

*Proposition 2.7:* Suppose  $\mathcal{C}$  be an  $(n, k, \delta)$  MDS code with column distances  $d_j^c$ ,  $j \in \mathbb{N}_0$  and free distance  $d_{\text{free}}$ . Let  $M := \min\{j \in \mathbb{N}_0 \mid d_j^c = d_{\text{free}}\}$ . Then

$$M \geq \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lceil \frac{\delta}{n-k} \right\rceil.$$

*Proof:* From Proposition 2.2 we get

$$d_{\text{free}} = (n-k) \left( \left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1 = d_M^c \leq (n-k)(M+1) + 1. \quad (2.8)$$

This yields the assertion.  $\square$

The proof also shows that in the case  $j > \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lceil \frac{\delta}{n-k} \right\rceil$ , the column distance  $d_j^c$  never attains the upper bound  $(n-k)(j+1) + 1$  of Proposition 2.2, see also (2.6).

*Definition 2.8:* An  $(n, k, \delta)$  code with column distances  $d_j^c$ ,  $j \in \mathbb{N}_0$ , is called *strongly-MDS*, if

$$d_M^c = (n-k) \left( \left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1 \text{ for } M = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lceil \frac{\delta}{n-k} \right\rceil.$$

Because of (2.6) strongly-MDS codes are in particular MDS codes.

In the case where  $(n - k) \mid \delta$ , the strong-MDS property implies that  $d_M^c$  attains the upper bound  $(n - k)(M + 1) + 1$ , see Proposition 2.1. Hence, Corollary 2.3 shows that in this case *all* column distances attain their optimum value. In other words, the column distances are

$$(d_0^c, d_1^c, d_2^c, \dots) \\ = (n - k + 1, 2(n - k) + 1, \dots, M(n - k) + 1, S, S, \dots) \quad (2.9)$$

where  $S = (M + 1)(n - k) + 1$ . If  $(n - k) \nmid \delta$ , we always have  $d_M^c < (n - k)(M + 1) + 1$  as can be seen from (2.8). In this case, nothing can be concluded for the previous column distances. In order to also incorporate optimality of  $d_0^c, \dots, d_{M-1}^c$  in this general case we pose the following definition.

*Definition 2.9:* Let

$$L := \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n - k} \right\rfloor. \quad (2.10)$$

An  $(n, k, \delta)$  code with column distances  $d_j^c$ ,  $j \in \mathbb{N}_0$  is said to have a maximum distance profile if

$$d_j^c = (n - k)(j + 1) + 1, \quad \text{for } j = 1, \dots, L.$$

Using the notation of Definition 2.8 we have

$$L = \begin{cases} M, & \text{if } (n - k) \mid \delta \\ M - 1, & \text{otherwise.} \end{cases} \quad (2.11)$$

Obviously, a strongly-MDS code with maximum-distance profile satisfies (2.9) for  $S = (n - k)(\lfloor \frac{\delta}{k} \rfloor + 1) + \delta + 1$ . Furthermore, we obtain that if  $n - k$  divides  $\delta$  then an  $(n, k, \delta)$  code has maximum distance profile if and only if it is strongly-MDS. If  $n - k$  does not divide  $\delta$ , then neither property implies the other one as is verified by Examples in 2.12 later in the paper.

An immediate consequence of Corollary 2.3 is as follows.

*Remark 2.10:* An  $(n, k, \delta)$  code has a maximum distance profile if and only if the  $L$ th column distance satisfies

$$d_L^c = (n - k)(L + 1) + 1.$$

*Remark 2.11:* The concept is clearly related to the notion of optimum distance profile (ODP), see [1, p. 112]. For ODP it is required that the column distances are maximal up to the memory  $\nu$ . Hence, if  $\nu \leq L$  then a code with maximum distance profile is always ODP. In general, one expects a good code to have generic Forney indices, i.e., the indices attain only the two values  $\lceil \frac{\delta}{k} \rceil$  and  $\lfloor \frac{\delta}{k} \rfloor$ . McEliece [26, Corollary 4.3] calls such codes *compact codes*. It has been shown in [11] that an MDS code always has generic indices. Of course, if the indices are generic then  $\nu = \lceil \frac{\delta}{k} \rceil$  and thus,  $\nu \leq L + 1$ .

The notion of ODP seems also to be dependent on the base field which is usually assumed to be the binary field. A code with maximum distance profile, in general, does not exist over the binary field and it can only exist for sufficiently large base fields. This is similar to the situation of MDS block codes. Such codes are known to exist as soon as the field size of  $\mathbb{F}$  is larger than the block length  $n$ .

We close this section by relating our work to previous results in the literature. As indicated earlier, the papers [8], [27], [28]

deal with a notion closely related to the bound given in Proposition 2.2. In all these papers  $(n, k)$  convolutional codes with a systematic generator matrix of memory  $m$  are considered and such codes are called MDS if  $d_m^c = (n - k)(m + 1) + 1$ . In order to avoid confusion, we will for the rest of this section call codes that satisfy  $d_m^c = (n - k)(m + 1) + 1$ , where  $m$  is the memory,  $m$ -MDS codes. It is easy to see that the number  $(n - k)(m + 1) + 1$  is the maximum possible value for the free distance of a *systematic* convolutional code with these parameters. In the papers [27], [28] this property has been characterized by the equivalence we presented in Corollary 2.5. In [27], matrices  $\hat{P}$  with property ii) of the corollary are called strongly nonsingular. In the case  $k = n - k = 1$ , we will call such matrices superregular in the next section (see also Remark 3.7). In the same paper [27] it is claimed that the problem of constructing superregular matrices has been solved. Unfortunately this is not true. We will dwell on this in the next section. Therefore, to our knowledge, the problem of constructing strongly nonsingular matrices is still open.

Finally, we wish to verify that for general parameters the properties strongly-MDS, having a maximum distance profile, and  $m$ -MDS are not related to each other, meaning that neither of the properties implies the other. We list a suitable example for each case. The column distances have been computed by using straightforward computer routines.

*Example 2.12:*

- 1) The  $(7, 1, 2)$  code over  $\mathbb{F}_8$  given in Example 4.2, item 8) below is strongly-MDS in our sense, but not  $m$ -MDS and therefore does not have a maximum distance profile. In this case,  $L = 2$  and this is also the memory. It is worth being mentioned that over  $\mathbb{F}_8$  no  $(7, 1, 2)$  code with maximum distance profile exists and in particular no systematic  $m$ -MDS code with these parameters exists. This can be shown by some lengthy, but straightforward computations.

- 2) The  $(4, 1, 2)$  code over  $\mathbb{F}_{16}$  with generator matrix

$$\begin{bmatrix} z^2 + \alpha z + 1 & z^2 + \alpha z + \alpha^3 & z^2 + \alpha^6 z + \alpha^3 & \alpha z + \alpha^8 \end{bmatrix}, \\ \text{where } \alpha^4 + \alpha + 1 = 0$$

has a maximum distance profile (thus, by Corollary 2.3, is  $m$ -MDS), but is obviously not MDS in our sense and therefore in particular not strongly-MDS.

- 3) The code over  $\mathbb{F}_{16}$  with generator matrix

$$\begin{bmatrix} 1 & \alpha^{14} z + \alpha^2 & \alpha^3 z + \alpha^3 \\ \alpha z & \alpha^{11} z + \alpha^8 & \alpha^{10} z + \alpha^{10} \end{bmatrix}$$

is  $m$ -MDS but does not have a maximum distance profile.

### III. EXISTENCE OF STRONGLY-MDS CODES

During his investigation of algebraic decoding of convolutional codes, Allen conjectured in his dissertation [29] the existence of strongly-MDS convolutional codes in the situation when  $k = 1$  and  $n = 2$ . In this section, we will show the existence of strongly-MDS codes with parameters  $(n, k, \delta)$  such that  $n - k$  divides  $\delta$ . It follows from (2.11) and Remark 2.10 that these codes also have maximum distance profile. By Theorem 2.6, we have to find an  $(n, k, \delta)$  code where  $(n - k) \mid \delta$  such

that  $d_M^c = (n-k)(M+1) + 1$  for  $M = \lfloor \frac{\delta}{k} \rfloor + \frac{\delta}{n-k}$ . In order to do so, put  $m := \frac{\delta}{n-k}$  and let

$$H = [A, B], \text{ where } A = \sum_{i=0}^m A_i D^i \in \mathbb{F}[D]^{(n-k) \times (n-k)}$$

and

$$B = \sum_{i=0}^m B_i D^i \in \mathbb{F}[D]^{(n-k) \times k} \quad (3.1)$$

be a basic parity-check matrix of the desired code. Without loss of generality we may assume  $A_0 = I_{n-k}$ . The strongly-MDS property can now be expressed as follows.

*Theorem 3.1:* Let  $H \in \mathbb{F}[D]^{n \times (n-k)}$  be as in (3.1), let  $A_0 = I_{n-k}$ , and define  $\mathcal{C} := \{v \in \mathbb{F}((D))^n \mid v\hat{H}^T = 0\}$  to be the code with parity-check matrix  $H$ . Furthermore, let

$$A^{-1}B = \sum_{i=0}^{\infty} P_i D^i \in \mathbb{F}((D))^{(n-k) \times k} \quad (3.2)$$

be the Laurent expansion of  $A^{-1}B$  over the field  $\mathbb{F}((D))$  of Laurent series and for  $M := \lfloor \frac{\delta}{k} \rfloor + \frac{\delta}{n-k}$  define

$$\hat{H} := [I_{(M+1)(n-k)} \quad \hat{P}] \quad \text{where} \quad \hat{P} := \begin{bmatrix} P_0 & & & \\ P_1 & P_0 & & \\ \vdots & \vdots & \ddots & \\ P_M & P_{M-1} & \cdots & P_0 \end{bmatrix}. \quad (3.3)$$

We call  $\hat{H}$  the  $M$ th systematic sliding parity-check matrix of  $\mathcal{C}$ . The following conditions are equivalent:

- $\mathcal{C}$  is strongly-MDS, i.e.,  $d_M^c = (n-k)(M+1) + 1$ ;
- none of the first  $k$  columns of  $\hat{P}$  is contained in the span of any other  $(M+1)(n-k) - 1$  columns of  $\hat{H}$ ;
- each  $j \times j$ -submatrix of  $\hat{P}$  that does not contain an  $s \times t$ -zero block where  $s+t \geq j+1$  is nonsingular.

Notice that b) automatically implies that none of the first  $n-k$  columns of  $\hat{H}$  is in the span of any other  $(M+1)(n-k) - 1$  columns.

*Proof:* After a column permutation the sliding parity-check matrix  $H_M^c$  of  $\mathcal{C}$  has the form

$$H^c := \left[ \begin{array}{ccc|ccc} I & & & B_0 & & \\ A_1 & I & & B_1 & B_0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ A_M & A_{M-1} & \cdots & I & B_M & B_{M-1} & \cdots & B_0 \end{array} \right].$$

It is straightforward to see that left multiplication of  $H^c$  by the inverse of the first block leads to the matrix  $\hat{H}$  of (3.3). After these transformations, Proposition 2.1 applied to the case  $j = M$  and  $d = (n-k)(M+1) + 1$  translates into the equivalence:  $\mathcal{C}$  is strongly-MDS if and only if neither any of the first  $k$  columns of  $H^c$  nor any of the first  $n-k$  columns of the second block of  $H^c$  is in the span of any other  $(n-k)(M+1) - 1$  columns of  $\hat{H}$ . But this, in turn, is equivalent to b) above. The equivalence of a) and c) is obtained from Corollary 2.5.  $\square$

In order to establish the existence of strongly-MDS codes we will proceed as follows. First, we will establish the existence of a systematic sliding parity-check matrix  $\hat{H}$  as in (3.3) with property c) of the preceding theorem. Thereafter, we will

show that there exists a basic polynomial matrix  $H = [A, B] \in \mathbb{F}[D]^{n \times (n-k)}$  of degree  $\delta$  such that

$$A^{-1}B = \sum_{i=0}^M P_i D^i + \text{higher powers.}$$

Theorem 3.1 then yields that the code with parity-check matrix  $H$  is a strongly-MDS  $(n, k, \delta)$  code.

As for the first step, let us have a look at the special case of  $(2, 1, \delta)$  codes. In this case,  $M = 2\delta$  and the systematic sliding parity-check matrix in (3.3) has the form

$$\hat{H} := \left[ \begin{array}{ccc|ccc} 1 & & & h_0 & 0 & \cdots & 0 \\ & 1 & & h_1 & h_0 & \ddots & \vdots \\ & & \ddots & \vdots & \ddots & \ddots & 0 \\ & & & 1 & h_{2\delta} & \cdots & h_1 & h_0 \end{array} \right] \\ =: [I_{2\delta+1}, T] \in \mathbb{F}^{(2\delta+1) \times (4\delta+2)}, \quad \text{where } h_j \in \mathbb{F}. \quad (3.4)$$

As we will see, the existence of matrices  $T$  of any given size and the structure above such that  $\hat{H}$  has the column property of Theorem 3.1, part b) will be the main tool for the existence of strongly-MDS codes even of length  $n > 2$ . Therefore, we will concentrate on these matrices first. The main point is to express the column condition on  $\hat{H}$  in terms of the minors of  $T$ .

*Definition 3.2:* Let  $R$  be a ring. For a matrix  $T \in R^{n \times k}$ , denote by  $T_{j_1, \dots, j_s}^{i_1, \dots, i_r} \in R^{r \times s}$  the  $r \times s$ -submatrix obtained from  $T$  by picking the rows with indices  $i_1, \dots, i_r$  and the columns with indices  $j_1, \dots, j_s$ .

In the sequel the following property will play a crucial role.

*Definition 3.3:* Let  $\mathbb{F}$  be field. A lower triangular matrix  $T \in \mathbb{F}^{n \times k}$  is said to be superregular,<sup>1</sup> if  $T_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  is nonsingular for all  $1 \leq r \leq \min\{k, n\}$  and all indices

$$1 \leq i_1 < \cdots < i_r \leq n, \quad 1 \leq j_1 < \cdots < j_r \leq k$$

which satisfy  $j_\nu \leq i_\nu$  for  $\nu = 1, \dots, r$ . We call the submatrices obtained by picking such indices the proper submatrices and their determinants the proper minors of  $T$ .

Obviously, a submatrix  $\hat{T}$  of  $T$  is proper if and only if no diagonal element of  $\hat{T}$  comes from strictly above the diagonal of  $T$ .

*Remark 3.4:* Observe that the proper submatrices are the only submatrices which can possibly be nonsingular. This can be seen as follows. If  $j_\nu > i_\nu$  for some  $\nu$ , then in the submatrix  $\hat{T} := T_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  the upper right block consisting of the first  $\nu$  rows and the last  $r - \nu + 1$  columns is identically zero. Hence, the first  $\nu$  rows of  $\hat{T}$  can have at most rank  $\nu - 1$ . In other words, the improper submatrices of  $T$  are trivially singular. For example, for  $T = (h_{ij})$  we have

$$T_{1,3,4}^{1,2,5} = \begin{bmatrix} h_{11} & 0 & 0 \\ h_{21} & 0 & 0 \\ h_{51} & h_{53} & h_{54} \end{bmatrix}.$$

Before we come to the existence of superregular matrices we will first present the following collection of characterizations of superregular matrices.

<sup>1</sup>We adopt this notion from [30], where it has been coined in a slightly different context.

*Theorem 3.5:* Let  $\mathbb{F}$  be a field and  $T$  be a lower triangular Toeplitz matrix, i.e.,

$$T = [T_1, \dots, T_l] = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{l-1} & \cdots & h_1 & h_0 \end{bmatrix} \in \mathbb{F}^{l \times l}. \quad (3.5)$$

Furthermore, put

$$\hat{H} := [I_l, T] = [e_1, \dots, e_l, T_1, \dots, T_l] \in \mathbb{F}^{l \times 2l}.$$

Then the following are equivalent:

- a)  $T$  is superregular, i.e., all proper submatrices in the sense of Definition 3.3 are nonsingular;
- b) all proper submatrices of  $T$  of the form  $T_{1, j_2, \dots, j_r}^{i_1, i_2, \dots, i_r}$ , where  $1 \leq i_1 < \dots < i_r \leq n$  and  $1 < j_2 < \dots < j_r \leq k$ , are nonsingular;
- c)  $\text{wt}(T_1 + \sum_{j=1}^s \beta_j T_{m_j}) \geq l - s$  for all  $1 \leq s \leq l - 1$ , all  $1 < m_1 < \dots < m_s \leq l$ , and all  $\beta_1, \dots, \beta_s \in \mathbb{F}$ ;
- d)  $T_1 \notin \text{span}\{T_{m_1}, \dots, T_{m_s}, e_{l_1}, \dots, e_{l_t}\}$  where  $1 < m_1 < \dots < m_s \leq l$  and  $1 \leq l_1 < \dots < l_t \leq l$  and  $s+t \leq l-1$ ;
- e) if  $v \in \mathbb{F}^{2l}$  satisfies  $v\hat{H}^T = 0$  and  $v_{l+1} \neq 0$ , then  $\text{wt}(v) \geq l + 1$ .
- f)  $e_1 \notin \text{span}\{T_{m_1}, \dots, T_{m_s}, e_{l_1}, \dots, e_{l_t}\}$  where  $1 \leq m_1 < \dots < m_s \leq l$  and  $1 < l_1 < \dots < l_t \leq l$  and  $s+t \leq l-1$ .
- g) If  $v \in \mathbb{F}^{2l}$  satisfies  $v\hat{H}^T = 0$  and  $v_1 \neq 0$ , then  $\text{wt}(v) \geq l + 1$ .

*Proof:*

a)  $\Leftrightarrow$  b) is obvious since in case of properness the Toeplitz structure implies

$$T_{j_1, \dots, j_r}^{i_1, \dots, i_r} = T_{j_1 - j_1 + 1, \dots, j_r - j_1 + 1}^{i_1 - j_1 + 1, \dots, i_r - j_1 + 1}.$$

b)  $\Rightarrow$  c): Let  $\hat{h} := T_1 + \sum_{j=1}^s \beta_j T_{m_j}$  and assume to the contrary  $\text{wt}(\hat{h}) < l - s$ . This implies that  $\hat{h}$  consists of at least  $s + 1$  zero entries, say at the positions  $i_1, \dots, i_{s+1}$ . Then

$$T_{1, m_1, \dots, m_s}^{i_1, \dots, i_{s+1}} \begin{pmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (3.6)$$

The superregularity yields  $m_\nu > i_{\nu+1}$  for some  $\nu \in \{1, \dots, s\}$ , which we can choose to be minimal with this property. Then the submatrix  $T_{m_\nu, \dots, m_s}^{i_1, \dots, i_{\nu+1}}$  is identically zero and therefore, we obtain from (3.6) the identity

$$T_{1, m_1, \dots, m_{\nu-1}}^{i_1, \dots, i_\nu} (1, \beta_1, \dots, \beta_{\nu-1})^T = 0,$$

a contradiction to superregularity since by minimality of  $\nu$  this coefficient matrix is nonsingular.

c)  $\Rightarrow$  b): Assume to the contrary that  $\det T_{1, m_1, \dots, m_s}^{i_1, \dots, i_{s+1}} = 0$  for some indices satisfying  $m_\nu \leq i_{\nu+1}$  for  $\nu = 1, \dots, s$ . We can assume  $s$  to be minimal with this property. Then there exists

$$(\beta_0, \beta_1, \dots, \beta_s) \in \mathbb{F}^{s+1} \setminus \{0\}$$

such that  $T_{1, m_1, \dots, m_s}^{i_1, \dots, i_{s+1}} (\beta_0, \dots, \beta_s)^T = 0$ . Minimality of  $s$  and the equivalence of a) and b) imply  $\beta_0 \neq 0$ . Hence, we can take  $\beta_0 = 1$  and (3.6) is satisfied. Thus,

$$\text{wt}(T_1 + \sum_{j=1}^s \beta_j T_{m_j}) \leq l - (s + 1),$$

a contradiction.

The properties (d) and (e) are simply reformulations of (c).

The equivalence (d)  $\Leftrightarrow$  (f) is clear from the structure of  $\hat{H}$  (a linear combination of  $T_1$  by the other columns of  $\hat{H}$  has to involve the column  $e_1$  and vice versa).

The property g) is a reformulation of f). □

The equivalence of e) and g) immediately implies the following.

*Corollary 3.6:* If  $T \in \mathbb{F}^{l \times l}$  is a superregular lower triangular Toeplitz matrix, then so is  $T^{-1}$ .

Using arguments as in the proof of Corollary 2.5 or by straightforward computations we obtain again the following.

*Remark 3.7:* Let  $T \in \mathbb{F}^{l \times l}$  be a lower triangular matrix with all elements on and below the diagonal being nonzero. Let  $\hat{T} := T_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  be a submatrix of  $T$ . Then  $\hat{T}$  is proper if and only if  $\hat{T}$  does not contain an  $s \times t$ -zero block where  $s + t \geq r + 1$ .

Now we will turn to the existence of superregular matrices. As indicated already earlier, there exists literature seemingly closely related to this problem, but unfortunately not solving it. Indeed, in [31], [32] (see also [33, p. 322]) triangular configurations are constructed for which all square submatrices inside the configuration are nonsingular. An example of such a configuration over  $\mathbb{F}_8$  is given by the matrix at the bottom of the page, where  $\alpha^3 + \alpha + 1 = 0$ . In [31, Theorem 3] and [32, p. 107] it has been shown that all square submatrices inside this triangular configuration are nonsingular. However, the triangular matrix shown at the top of the following page is not superregular, since, for instance,  $\det T_{1,2,3}^{2,3,4} = 0$ . The same applies to the triangular configurations given in [31, Theorem 5]. As this example shows, the main obstacle for constructing superregular matrices are those proper submatrices that are partly located in the zero triangle of the matrix. This produces a type of irregularity making it hard to come up with an algebraic construction of such matrices, even though examples in Example 3.10, item 1) indicate that such construction should be possible. However, existence of superregular matrices, even with Toeplitz structure, is guaranteed by the following lemma.

$$\begin{matrix} 1 & & & & & & \\ 1 & 1 & & & & & \\ 1 & (1-\alpha)^{-1} & (1-\alpha^2)^{-1} & \dots & \dots & (1-\alpha^5)^{-1} & (1-\alpha^6)^{-1} \\ 1 & (1-\alpha^2)^{-1} & \dots & \dots & (1-\alpha^5)^{-1} & (1-\alpha^6)^{-1} & \\ \vdots & \vdots & & & & & \\ 1 & (1-\alpha^6)^{-1} & & & & & \end{matrix}$$

$$T := \begin{bmatrix} (1-\alpha^6)^{-1} & 0 & 0 & \dots & 0 \\ (1-\alpha^5)^{-1} & (1-\alpha^6)^{-1} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ (1-\alpha^2)^{-1} & (1-\alpha^3)^{-1} & \dots & (1-\alpha^6)^{-1} & 0 \\ (1-\alpha)^{-1} & (1-\alpha^2)^{-1} & \dots & (1-\alpha^5)^{-1} & (1-\alpha^6)^{-1} \end{bmatrix} \in \mathbb{F}_8^{6 \times 6}$$

*Lemma 3.8:* Let  $\mathbb{F}$  be a field and  $X_1, \dots, X_l$  be independent indeterminates over  $\mathbb{F}$ . Define the matrix

$$A := \begin{bmatrix} X_1 & 0 & \dots & 0 \\ X_2 & X_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ X_l & \dots & X_2 & X_1 \end{bmatrix} \in \mathbb{F}(X_1, \dots, X_l)^{l \times l}.$$

Then  $A$  is superregular.

*Proof:* We proceed by contradiction. Assume there exists a singular proper submatrix

$$\hat{A} := A_{j_1, \dots, j_r}^{i_1, \dots, i_r}.$$

We can take the size  $r$  to be minimal. Then certainly  $r > 1$ . By properness, we know that  $j_\nu \leq i_\nu$  for  $\nu = 1, \dots, r$ .

Notice that for  $\mu \leq \nu$  the entry of  $A$  at the position  $(\nu, \mu)$  is given by  $A_\mu^\nu = X_{\nu-\mu+1}$ . Hence, the indeterminate with the largest index appearing in  $\hat{A}$  is  $X_{i_r-j_1+1}$ . It appears only once in the matrix and that is in the lower left corner. Thus, its coefficient in  $\det \hat{A}$  is  $\pm \det \tilde{A}$ , where

$$\tilde{A} := A_{j_2, \dots, j_r}^{i_1, \dots, i_{r-1}}.$$

Singularity of  $\hat{A}$  now implies  $\det \tilde{A} = 0$ . By minimality of  $r$ , this yields that  $\tilde{A}$  is an improper submatrix of  $A$ , i.e., there exists an index  $\tau \in \{2, \dots, r\}$  such that  $j_\tau > i_{\tau-1}$ . Picking  $\tau$  minimal we get  $i_1 < \dots < i_{\tau-1} < j_\tau < \dots < j_r$  and therefore the first  $\tau - 1$  rows of  $\hat{A}$  have the form

$$\begin{bmatrix} * & \dots & * & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \dots & * & 0 & \dots & 0 \end{bmatrix}$$

where the block of possibly nonzero elements consists of  $\tau - 1$  columns. Hence,  $\hat{A}$  is a blocktriangular matrix and we have

$$0 = \det \hat{A} = \det A_{j_1, \dots, j_{\tau-1}}^{i_1, \dots, i_{\tau-1}} \det A_{j_\tau, \dots, j_r}^{i_\tau, \dots, i_r}.$$

Since both factors are proper minors we get a contradiction to the minimality of the size  $r$ .  $\square$

The following consequence is standard.

*Theorem 3.9:* For every  $l \in \mathbb{N}$  and every prime number  $p$  there exists a finite field  $\mathbb{F}$  of characteristic  $p$  and a superregular matrix  $T \in \mathbb{F}^{l \times l}$  having Toeplitz structure.

*Proof:* Consider the prime field  $\mathbb{F}_p$  and the matrix of the previous lemma with entries in  $\mathbb{F}_p(X_1, \dots, X_l)$ . All its proper minors are nonzero polynomials in  $\mathbb{F}_p[X_1, \dots, X_l]$ . Over an algebraic closure  $\bar{\mathbb{F}}_p$ , a point  $a := (a_1, \dots, a_l) \in \bar{\mathbb{F}}_p^l$  can be found such that none of the minors vanishes at  $a$ . Hence, the Toeplitz matrix  $T$  having  $(a_1, \dots, a_l)^T$  as its first column is superregular. Since each  $a_i$  is algebraic over  $\mathbb{F}_p$ , the matrix  $T$  has its entries in a finite field extension  $\mathbb{F}$  of  $\mathbb{F}_p$ .  $\square$

In particular, for every size  $l \in \mathbb{N}$  there exist superregular Toeplitz matrices over a field of characteristic 2. Unfortunately,

the theorem above is nonconstructive and it is not at all clear what the minimum field of characteristic 2 is to allow a superregular Toeplitz matrix of given size  $l \times l$ . We present some examples.

*Example 3.10:*

1) Using a computer algebra program one checks that the following matrices are superregular. The first examples are all over prime fields  $\mathbb{F}_p$ .

$$\begin{aligned} & \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{2 \times 2}, & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \in \mathbb{F}_3^{3 \times 3} \\ & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{bmatrix} \in \mathbb{F}_5^{4 \times 4}, & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 6 & 1 & 2 & 1 & 0 \\ 4 & 6 & 1 & 2 & 1 \end{bmatrix} \in \mathbb{F}_7^{5 \times 5} \\ & & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 2 & 1 & 0 \\ 4 & 3 & 1 & 1 & 2 & 1 \end{bmatrix} \in \mathbb{F}_{11}^{6 \times 6} \\ & & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 1 & 0 & 0 & 0 & 0 & 0 \\ 13 & 7 & 1 & 0 & 0 & 0 & 0 \\ 2 & 13 & 7 & 1 & 0 & 0 & 0 \\ 1 & 2 & 13 & 7 & 1 & 0 & 0 \\ 4 & 1 & 2 & 13 & 7 & 1 & 0 \\ 14 & 4 & 1 & 2 & 13 & 7 & 1 \end{bmatrix} \in \mathbb{F}_{17}^{7 \times 7}. \end{aligned}$$

The following examples represent superregular matrices over finite fields of characteristic 2. For this assume that  $\alpha, \beta$ , and  $\gamma$  satisfy

$$\alpha^2 + \alpha + 1 = 0, \quad \beta^3 + \beta + 1 = 0, \quad \text{and} \quad \gamma^4 + \gamma + 1 = 0.$$

Then the following matrices are superregular over  $\mathbb{F}_4, \mathbb{F}_8$ , and  $\mathbb{F}_{16}$ , respectively:

$$\begin{aligned} & \begin{bmatrix} 1 \\ \alpha & 1 \\ 1 & \alpha & 1 \end{bmatrix} \in \mathbb{F}_{2^2}^{3 \times 3}, & \begin{bmatrix} 1 & & & & \\ \beta & 1 & & & \\ \beta^3 & \beta & 1 & & \\ \beta & \beta^3 & \beta & 1 & \\ 1 & \beta & \beta^3 & \beta & 1 \end{bmatrix} \in \mathbb{F}_{2^3}^{5 \times 5} \\ & & \begin{bmatrix} 1 & & & & & \\ \gamma & 1 & & & & \\ \gamma^5 & \gamma & 1 & & & \\ \gamma^5 & \gamma^5 & \gamma & 1 & & \\ \gamma & \gamma^5 & \gamma^5 & \gamma & 1 & \\ 1 & \gamma & \gamma^5 & \gamma^5 & \gamma & 1 \end{bmatrix} \in \mathbb{F}_{2^4}^{6 \times 6}. \end{aligned}$$

Assume  $\epsilon, \omega$  satisfy

$$\epsilon^5 + \epsilon^2 + 1 = 0 \quad \text{and} \quad \omega^6 + \omega + 1 = 0.$$

Then the following matrices are superregular over  $\mathbb{F}_{32}$  and  $\mathbb{F}_{64}$ , respectively:

$$\begin{bmatrix} 1 & & & & & & & \\ \epsilon & 1 & & & & & & \\ \epsilon^6 & \epsilon & 1 & & & & & \\ \epsilon^9 & \epsilon^6 & \epsilon & 1 & & & & \\ \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 & & & \\ \epsilon & \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 & & \\ 1 & \epsilon & \epsilon^6 & \epsilon^9 & \epsilon^6 & \epsilon & 1 & \end{bmatrix} \in \mathbb{F}_{2^5}^{7 \times 7}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \omega & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \omega^9 & \omega & 1 & 0 & 0 & 0 & 0 & 0 \\ \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 & 0 & 0 \\ \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 & 0 \\ \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 \\ \omega & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 \\ 1 & \omega & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 \end{bmatrix} \in \mathbb{F}_{2^6}^{8 \times 8}.$$

Notice that the matrices above have even more symmetry than required. One can easily show that there is no superregular  $4 \times 4$  matrix over  $\mathbb{F}_4$  of general Toeplitz structure. However, the above suggests a question whether one can find for every  $l \geq 5$  a superregular  $l \times l$ -Toeplitz matrix over  $\mathbb{F}_{2^{l-2}}$ .

2) In Appendix B, we prove that for every  $n \in \mathbb{N}$  the proper minors of the Toeplitz matrix

$$T_n := \begin{bmatrix} \binom{n-1}{0} & & & & \\ \binom{n-1}{1} & \binom{n-1}{0} & & & \\ \vdots & \ddots & \ddots & & \\ \binom{n-1}{n-1} & \cdots & \binom{n-1}{1} & \binom{n-1}{0} \end{bmatrix} \in \mathbb{Z}^{n \times n}$$

are all positive. Hence, for each  $n \in \mathbb{N}$  there exists a smallest prime number  $p_n$  such that  $T_n$  is superregular over the prime field  $\mathbb{F}_{p_n}$ . One can check that

$$p_2 = 2, p_3 = 5, p_4 = 7, p_5 = 11, p_6 = 23, p_7 = 43.$$

Now we can establish the existence of strongly-MDS codes in the following sense.

*Theorem 3.11:* For every  $n, k, \delta \in \mathbb{N}$  such that  $n-k$  divides  $\delta$  and for every prime number  $p$  there exists a strongly-MDS code with parameters  $(n, k, \delta)$  over a suitably large field of characteristic  $p$ .

The proof of this theorem is rather long and technical and therefore it is given in Appendix C.

*Remark 3.12:* It would be of course interesting to find good bounds on the size of the field where an  $(n, k, \delta)$  strongly-MDS code exists. Using the fact that an  $n \times n$  matrix, whose entries have magnitude at most  $m$ , can have a determinant of at most  $m^n n^{n/2}$  it is possible to bound the largest minor of the matrix  $T_n$  from Example 3.10, item 2) above. This, in turn, provides then a very rough bound for a prime field where the existence of strongly-MDS codes is guaranteed. In his upcoming dissertation, R. Hutchinson will provide sharper bounds for the smallest field size where the existence of strongly-MDS codes are guaranteed. Unfortunately, examples show that these bounds are still far away from being optimal.

There is of course the natural question if strongly-MDS convolutional codes and codes with maximum distance profile exist for all parameters  $(n, k, \delta)$ . This section showed that such codes exist whenever  $n-k$  divides  $\delta$ . In [34], it has been shown that codes with a maximum distance profile exist for all parameters  $(n, k, \delta)$  over sufficiently large fields. For other small values of  $(n, k, \delta)$  we have found strongly-MDS convolutional codes and codes with maximum distance profile by making computer searches. In the next section, we present a series of examples of such codes. Based on this wealth of data we conjecture the following.

*Conjecture 3.13:* For all  $n > k > 0$  and for all  $\delta \geq 0$  there exists an  $(n, k, \delta)$  code over a sufficiently large field which is both strongly-MDS and has a maximum distance profile.

#### IV. EXAMPLES

In this section, we will present some examples of strongly-MDS codes with small parameters. The first set of examples is constructed according to the proof of Theorem 3.11 by utilizing the superregular matrices in Example 3.10.

*Example 4.1:* Recall the first part of the proof of Theorem 3.11.

- 1) We can construct strongly-MDS  $(2, 1, \delta)$  codes once a  $\tau \times \tau$  superregular matrix, where  $\tau = 2\delta + 1$ , is available. Thus, the  $5 \times 5$  and  $7 \times 7$  matrices given in Example 3.10, item 1) lead to the strongly-MDS  $(2, 1, 2)$  code over  $\mathbb{F}_8$  (where  $\beta^3 + \beta + 1 = 0$ ) with a parity-check matrix

$$H = [a, b] \\ = [1 + \beta^2 D + \beta^5 D^2, 1 + \beta^4 D + \beta^5 D^2] \in \mathbb{F}_8[D]^2$$

and to the strongly-MDS  $(2, 1, 3)$  code over  $\mathbb{F}_{32}$  (where  $\epsilon^5 + \epsilon^2 + 1 = 0$ ) with a parity-check matrix

$$H = [a, b] \\ = [1 + \epsilon^{18} D + \epsilon^{11} D^2 + \epsilon^{29} D^3, 1 + D + \epsilon^{27} D^2 + \epsilon^{18} D^3] \\ \in \mathbb{F}_{32}[D]^2.$$

Indeed, one checks that

$$\frac{1 + \beta^4 D + \beta^5 D^2}{1 + \beta^2 D + \beta^5 D^2} \\ = 1 + \beta D + \beta^3 D^2 + \beta D^3 + D^4 + \text{higher powers}$$

and

$$\frac{1 + D + \epsilon^{27} D^2 + \epsilon^{18} D^3}{1 + \epsilon^{18} D + \epsilon^{11} D^2 + \epsilon^{29} D^3} \\ = 1 + \epsilon D + \epsilon^6 D^2 + \epsilon^9 D^3 + \epsilon^6 D^4 + \epsilon D^5 + D^6 \\ + \text{higher powers.}$$

Hence, the free distance of the two codes above is 6 (resp., 8), and this is also the forth (resp., sixth) column distance.

- 2) Using the  $8 \times 8$ -superregular matrix of Example 3.10, item 1), one can construct a strongly-MDS  $(3, 2, 2)$  code over  $\mathbb{F}_{64}$ . Hence, the code has free distance equal to its third column distance, and this value is 5. Using the construction of the proof of Theorem 3.11 and going through



some tedious calculations in the field  $\mathbb{F}_{64}$  (where  $\omega^6 + \omega + 1 = 0$ ) one finally arrives at the parity-check matrix

$$H = [1 + \omega^{57}D + \omega^{62}D^2, \omega + \omega^{44}D + \omega^{54}D^2, \\ 1 + \omega^{17}D + \omega^{21}D^2] \in \mathbb{F}_{64}^3.$$

- 3) A strongly-MDS  $(4, 3, 1)$  code has free distance 3 and this is identical with the first column distance. It can be obtained from a  $6 \times 6$ -superregular matrix using the construction of the proof of Theorem 3.11. Indeed, the matrix

$$\hat{H} = \left[ \begin{array}{cc|ccc} 1 & 0 & \gamma^5 & \gamma & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & \gamma & \gamma^5 & \gamma^5 & \gamma & 1 \end{array} \right]$$

has been obtained from the superregular Toeplitz matrix of Example 3.10, item 1) and thus it satisfies property b) of Theorem 3.1. Hence, a parity-check matrix of a strongly-MDS  $(4, 3, 1)$  code over  $\mathbb{F}_{16}$  (where  $\gamma^4 + \gamma + 1 = 0$ ) is given by

$$H = [1, \gamma^5 + D, \gamma + \gamma D, 1 + \gamma^5 D] \in \mathbb{F}_{16}[D]^4.$$

- 4) Of course, not every MDS code is strongly-MDS. For instance, the code with parity-check matrix

$$H = [10 + 3D + 2D^2, 4 + 2D + D^2] \in \mathbb{F}_{11}[D]^2$$

is an MDS code, but not strongly-MDS. In this example, the MDS property follows from the fact that this code is the result of the construction of MDS codes as presented in [12]. However, a  $(2, 1, 1)$  code is strongly-MDS if and only if it is an MDS code. This can be checked directly by using Theorem 3.1 and the fact that for the (basic) parity-check matrix  $[a_0 + a_1D, b_0 + b_1D]$  of an MDS code all coefficients as well as  $a_0b_1 - a_1b_0$  are nonzero.

The next series of examples has been found by completely different methods. They are all cyclic convolutional codes in the sense of [13], [35], [14], [15]. In those papers, convolutional codes having some additional algebraic structure are being investigated. This additional structure is a generalization of cyclicity of block codes but is a far more complex notion for convolutional codes. In particular, cyclicity of convolutional codes does *not* mean invariance under the cyclic shift in  $\mathbb{F}[D]^n$ . We will not go into the details but rather refer to [13], [35]. However, in order to understand and test the following examples there is no need in understanding the concept of cyclicity for convolutional codes since in the following we provide all information needed to specify the codes. We present the generator matrices and also provide all column distances; they have been computed with a computer algebra program. All matrices given below are minimal basic. We would like to mention that just like for cyclic block codes, the length of the code and the characteristic of the field have to be coprime. Therefore, only codes with odd length are given below.

One should note that most of the following codes exist over comparatively smaller alphabets than the examples of Example 4.1. However, we do not know any general construction

for strongly-MDS cyclic convolutional codes yet. But the abundance of (small) examples suggests that such a construction might be possible and might lead to smaller alphabets for given parameters than the construction of the last section. We will leave this as an open question for future research.

*Example 4.2:*

- 1) A strongly-MDS  $(3, 1, 1)$  code over  $\mathbb{F}_4$

$$G = [\alpha + \alpha D, \alpha^2 + \alpha D, 1 + \alpha D].$$

The column distances are  $d_0^c = 3$ ,  $d_1^c = 5$ ,  $d_j^c = 6$  for  $j \geq 2$ .

- 2) A strongly-MDS  $(3, 1, 2)$  code over  $\mathbb{F}_{16}$  (where  $\beta^4 + \beta + 1 = 0$ )

$$G = [\beta + \beta D + D^2, \beta^6 + \beta D + \beta^{10}D^2, \beta^{11} + \beta D + \beta^5D^2].$$

The column distances are  $d_0^c = 3$ ,  $d_1^c = 5$ ,  $d_2^c = 7$ ,  $d_j^c = 9$  for  $j \geq 3$ .

- 3) A strongly-MDS  $(3, 2, 2)$  code over  $\mathbb{F}_{16}$

$$G = \begin{bmatrix} \beta^5 + \beta^4D & \beta^3 + \beta^8D & \beta^9 + \beta^2D \\ \beta^9 + \beta^{12}D & \beta^5 + \beta^{14}D & \beta^3 + \beta^3D \end{bmatrix}.$$

The column distances are  $d_0^c = 2$ ,  $d_1^c = 3$ ,  $d_2^c = 4$ ,  $d_j^c = 5$  for  $j \geq 3$ .

- 4) A strongly-MDS  $(5, 1, 1)$  code over  $\mathbb{F}_{16}$

$$G = [\beta + \beta D, \beta^{13} + \beta^{10}D, \beta^{10} + \beta^4D, \beta^7 + \beta^{13}D, \beta^4 + \beta^7D].$$

The column distances are  $d_0^c = 5$ ,  $d_1^c = 9$ ,  $d_j^c = 10$  for  $j \geq 2$ .

- 5) A strongly-MDS  $(5, 1, 2)$  code over  $\mathbb{F}_{16}$

$$G = [\beta + \beta^4D + \beta D^2, \beta^7 + \beta D + \beta^{10}D^2, \beta^{13} + \beta^{13}D + \beta^4D^2, \\ \beta^4 + \beta^{10}D + \beta^{13}D^2, \beta^{10} + \beta^7D + \beta^7D^2].$$

The column distances are  $d_0^c = 5$ ,  $d_1^c = 9$ ,  $d_2^c = 13$ ,  $d_j^c = 15$  for  $j \geq 3$ .

- 6) A strongly-MDS  $(5, 2, 2)$  code over  $\mathbb{F}_{16}$  as shown at the bottom of the page. The column distances are  $d_0^c = 4$ ,  $d_1^c = 7$ ,  $d_j^c = 9$  for  $j \geq 2$ .

- 7) A strongly-MDS  $(7, 1, 1)$  code over  $\mathbb{F}_8$  (where  $\gamma^3 + \gamma + 1 = 0$ )

$$G = [\gamma + \gamma D, \gamma^3 + D, \gamma^5 + \gamma^6D, 1 + \gamma^5D, \\ \gamma^2 + \gamma^4D, \gamma^4 + \gamma^3D, \gamma^6 + \gamma^2D].$$

The column distances are  $d_0^c = 7$ ,  $d_1^c = 13$ ,  $d_j^c = 14$  for  $j \geq 2$ .

- 8) A strongly-MDS  $(7, 1, 2)$  code over  $\mathbb{F}_8$

$$G = [\gamma^2 + \gamma D + D^2, \gamma^5 + \gamma^3D + \gamma^6D^2, \\ \gamma + \gamma^5D + \gamma^5D^2, \gamma^4 + D + \gamma^4D^2, \\ 1 + \gamma^2D + \gamma^3D^2, \gamma^3 + \gamma^4D + \gamma^2D^2, \\ \gamma^6 + \gamma^6D + \gamma D^2].$$

The column distances are  $d_0^c = 7$ ,  $d_1^c = 13$ ,  $d_2^c = 18$ ,  $d_j^c = 21$  for  $j \geq 3$ .

$$G = \begin{bmatrix} \beta + \beta D & \beta^{13} + \beta^{10}D & \beta^{10} + \beta^4D & \beta^7 + \beta^{13}D & \beta^4 + \beta^7D \\ 1 + \beta^5D & \beta^3 + \beta^{11}D & \beta^6 + \beta^2D & \beta^9 + \beta^8D & \beta^{12} + \beta^{14}D \end{bmatrix}.$$

- 9) It is worth noting that there does not exist even an MDS  $(7, 2, 2)$  code over  $\mathbb{F}_8$ , since the generalized Singleton bound in this case is 13, but due to the Griesmer bound (see [1, p. 133] for the binary case) the parameters of an  $(n, k, \delta)$  code over  $\mathbb{F}_q$  with memory  $m$  and distance  $d$  satisfy

$$\sum_{l=0}^{k(m+i)-\delta-1} \left\lceil \frac{d}{q^l} \right\rceil \leq n(m+i), \quad \text{for all } i \in \mathbb{N}_0.$$

Hence, a  $(7, 2, 2)$  code over  $\mathbb{F}_8$  with memory 1 has at most distance 12. The inequality applied to  $i = 1$  shows that the field size has to be at least 13 in order to allow the existence of an MDS  $(7, 2, 2)$  code.

One should notice that the codes in Example 4.2, items 1)–7) are not only strongly-MDS but also have *all* column distances being optimal in the sense that they reach the upper bound given in Proposition 2.2. In particular, they also have a maximum distance profile in the sense of Definition 2.9. For the  $(7, 1, 2)$  code in item 8), only the second column distance is not optimal, but rather one less than the upper bound, which is 19 in this case. The implications of this have been discussed already in Example 2.12, item 1).

## V. THE DUAL OF A STRONGLY MDS-CODE

In this section, we will present some results concerning the dual code of a strongly-MDS code. The main result shows that a convolutional code has a maximum distance profile if and only if its dual has this property. This then implies for certain parameters that a code is strongly-MDS if and only if its dual has this property. These results are very appealing as they generalize the situation for block codes.

Recall that if

$$\mathcal{C} = \{uG \mid u \in \mathbb{F}[D]^k\} = \{v \in \mathbb{F}[D]^n \mid vH^\top = 0\} \subseteq \mathbb{F}[D]^n$$

is an  $(n, k, \delta)$  code with generator matrix  $G \in \mathbb{F}[D]^{k \times n}$  and parity-check matrix  $H \in \mathbb{F}[D]^{(n-k) \times n}$ , then the dual code, defined as

$$\mathcal{C}^\perp = \{w \in \mathbb{F}[D]^n \mid wv^\top = 0 \text{ for all } v \in \mathcal{C}\}$$

is given by

$$\mathcal{C}^\perp = \{uH \mid u \in \mathbb{F}[D]^{n-k}\} = \{w \in \mathbb{F}[D]^n \mid wG^\top = 0\}$$

and thus an  $(n, n-k, \delta)$  code. In contrast to the block code situation, almost nothing is known about the relation between the distances of a code and its dual. In particular, it has been shown in [36] that no MacWilliams identity relating the weight distributions of  $\mathcal{C}$  and  $\mathcal{C}^\perp$  exists. In block code theory, a very simple relation between the distances of a code and its dual is given in the case of MDS codes. In fact, if  $\mathcal{C}$  is an MDS  $(n, k)$  block code, then the dual  $\mathcal{C}^\perp$  is an MDS  $(n, n-k)$  code, see [33, Ch. 11, Par. 2] and very specific knowledge on the weight enumerator and its dual is known [33, Ch. 11]. Therefore, it is quite natural to investigate whether the dual of an MDS (or strongly-MDS) convolutional code is MDS (or strongly-MDS), too. Unfortunately, this is in general not the case.

*Example 5.1:* In general, the dual of a strongly-MDS code is not even an MDS code. This can be seen from the dual of the code given in Example 4.1, item 3). The dual has generator matrix

$$G = [1, \gamma^5 + D, \gamma + \gamma D, 1 + \gamma^5 D] \in \mathbb{F}_{16}[D]^4$$

which obviously has weight less than the generalized Singleton bound 8 (see Theorem 2.6).

As we will show next, the property of maximum distance profile carries over under dualization. In addition, for specific code parameters the strongly-MDS property carries over to the dual code as well. To this end, recall from Definition 2.8 that an  $(n, k, \delta)$  code is strongly-MDS if the  $M$ th column distance attains the generalized Singleton bound where

$$M = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lceil \frac{\delta}{n-k} \right\rceil.$$

Thus, the dual code  $\mathcal{C}^\perp$  is MDS if the  $\hat{M}$ th column distance attains the generalized Singleton bound where

$$\hat{M} = \left\lfloor \frac{\delta}{n-k} \right\rfloor + \left\lceil \frac{\delta}{k} \right\rceil.$$

Obviously, these two numbers differ by one when  $k$  divides  $\delta$  but  $n-k$  does not or *vice versa*. What remains equal for both the code and its dual is the quantity

$$L = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor$$

used in Definition 2.9, where we introduced the concept of maximum distance profile.

Theorem 2.4 provides us with the following nice duality result.

*Theorem 5.2:* An  $(n, k, \delta)$  code  $\mathcal{C} \subseteq \mathbb{F}[D]^n$  has a maximum distance profile if and only if the dual code  $\mathcal{C}^\perp \subseteq \mathbb{F}[D]^n$  has this property.

*Proof:* Let  $\mathcal{C}$  have generator matrix  $G$  and parity-check matrix  $H$  as given in (2.1) and (2.2). Assume  $\mathcal{C}$  has a maximum distance profile. By Theorem 2.4, every  $(L+1)(n-k) \times (L+1)(n-k)$  full-size minor formed from the columns of  $H_L^\perp$  with indices  $1 \leq r_1 < \dots < r_{(L+1)(n-k)}$ , where  $r_{s(n-k)} \leq sn$  for  $s = 1, \dots, L$ , is nonzero.

Consider now the dual code  $\mathcal{C}^\perp$  which is defined as the row-space of the  $(n-k) \times n$  matrix  $H$ . It follows from (2.4) that the  $L$ th column distance of the dual code  $\mathcal{C}^\perp$  is given by

$$\hat{d}_L^\perp = \min \{ \text{wt}((u_L, \dots, u_0)H_L^\perp) \mid u_i \in \mathbb{F}^{n-k}, u_0 \neq 0 \}.$$

Taking the reversed ordering into account, we obtain again from Theorem 2.4 that the dual code  $\mathcal{C}^\perp$  has maximum distance profile as well.  $\square$

*Corollary 5.3:* When both  $k$  and  $n-k$  divide  $\delta$  then an  $(n, k, \delta)$  code  $\mathcal{C} \subseteq \mathbb{F}[D]^n$  is strongly-MDS if and only if  $\mathcal{C}^\perp \subseteq \mathbb{F}[D]^n$  has this property.

*Proof:* From  $k \mid \delta$  and  $(n-k) \mid \delta$ , it follows that  $L = M$  and

$$d_M^\perp = (n-k) \left( \frac{\delta}{k} + 1 \right) + \delta + 1$$

which is the generalized Singleton bound of the code  $\mathcal{C}$  and

$$\hat{d}_M^c = k \left( \frac{\delta}{n-k} + 1 \right) + \delta + 1$$

which is the generalized Singleton bound of the dual code  $\mathcal{C}^\perp$ .  $\square$

The preceding result gives us another class of strongly-MDS codes by dualizing Theorem 3.11.

*Corollary 5.4:* For every  $n, \delta \in \mathbb{N}_0$  such that both  $k$  and  $n - k$  divide  $\delta$  and every prime number  $p$  there exists a strongly-MDS  $(n, k, \delta)$  code over some suitably large field of characteristic  $p$ .

*Example 5.5:*

- a) Corollary 5.3 tells us that the duals of the  $(2, 1, \delta)$  codes given in Example 4.1, item 1) are strongly-MDS. But this is obviously so, since they are—up to ordering—identical to the given codes.

- b) Dualizing the code of Example 4.1, item 2) gives us a strongly-MDS  $(3, 1, 2)$  code with generator matrix

$$G = [1 + \omega^{57}D + \omega^{62}D^2, \omega + \omega^{44}D + \omega^{54}D^2, 1 + \omega^{17}D + \omega^{21}D^2] \in \mathbb{F}_{64}^3.$$

- c) Dualizing the codes given in Example 4.2, items 2) and 3), we obtain another two strongly-MDS codes with generator matrices

$$H_1 = \begin{bmatrix} 1 & \beta D + \beta^9 & \beta^6 D + \beta^8 \\ \beta^{14} D & \beta^7 D + \beta^6 & \beta^8 D + \beta \end{bmatrix} \in \mathbb{F}_{16}^{2 \times 3}$$

and

$$H_2 = [D^2 + D + \beta^2, \beta^{10} D^2 + D + \beta^7, \beta^5 D^2 + D + \beta^{12}] \in \mathbb{F}_{16}^3.$$

It is known that these codes are also cyclic convolutional codes in the sense of [13], see [13, Theorem 7.5].

Finally, we would like to mention that even in the case where  $k \mid \delta$  and  $(n - k) \mid \delta$ , the dual of an MDS code is not MDS in general. An example is given by the following code.

*Example 5.6:* The  $(3, 1, 2)$  code  $\mathcal{C} \subseteq \mathbb{F}[D]^3$ , where  $\mathbb{F} = \mathbb{F}_{16}$ , with generator matrix

$$G = [1 + \beta D + \beta^4 D^2, \beta^{10} + \beta^2 D + \beta^4 D^2, \beta^8 + \beta^5 D + D^2]$$

and parity-check matrix

$$H = \begin{bmatrix} 1 & \beta^{14} D + \beta^2 & \beta^3 D + \beta^3 \\ \beta D & \beta^{11} D + \beta^8 & \beta^{10} D + \beta^{10} \end{bmatrix}$$

is an MDS code, but not strongly-MDS. It satisfies  $d_3^c = 8$  and  $d_4^c = 9$ . The dual code generated by  $H$  is not MDS. Its distance is 4.

## VI. ESTIMATES FOR THE EXTENDED ROW DISTANCES

In this section, we will use the information about the column distances in order to present a lower bound for the extended row distances of a strongly-MDS code with unit memory.

For this, let  $G = G_0 + G_1 z \in \mathbb{F}[z]^{k \times n}$  be the generator matrix of a unit memory code of degree  $k$ , thus,  $G_1$  has full row rank, and let the code be strongly-MDS and have a maximum

distance profile. Hence,  $d_j^c = (n - k)(j + 1) + 1$  for all  $j = 0, \dots, M - 1$  where  $M := 1 + \lceil \frac{k}{n-k} \rceil$  and we have

$$d_M^c = 2n - k + 1 = d_{\text{free}}(\text{im } G).$$

Denote by  $\hat{d}_j^r$  the  $j$ th extended row distances of the code, thus,  $\hat{d}_j^r$  is the minimum weight of all codewords  $v = \sum_{i=0}^{j-1} u_i z^i G$  of degree  $j$  where  $u_i \neq 0$  for all  $i = 0, \dots, j - 1$ . Define

$$A_M^c = \begin{bmatrix} G_1 & & & \\ G_0 & G_1 & & \\ & G_0 & \ddots & \\ & & \ddots & G_1 \\ & & & G_0 \end{bmatrix} \in \mathbb{F}^{(M+1)k \times Mn}.$$

*Lemma 6.1:* Let  $u_0 \neq 0$ . Then

$$\text{wt}((u_0, u_1, \dots, u_M) A_M^c) \geq n - k + 1.$$

*Proof:* Let  $(u_0, u_1, \dots, u_M) A_M^c = (v_1, \dots, v_M)$ . Then  $(u_0, u_1, \dots, u_M) G_M^c = (v_0, v_1, \dots, v_M)$  for some  $v_0 \in \mathbb{F}^n$ . Since  $u_0 \neq 0$ , we have

$$\text{wt}(v_0, v_1, \dots, v_M) \geq d_M^c = 2n - k + 1.$$

Estimating the weight of  $v_0$  by  $n$  we obtain the desired result.  $\square$

*Theorem 6.2:* Let  $j \geq M$  and write  $j = aM + t$  where  $a \in \mathbb{N}$  and  $0 \leq t < M$ . Then we have

$$\hat{d}_j^r \geq \frac{n - k + 1}{M} j + (n - k)(M - 1) + \max\{0, t(n - k) + 1 - n\}.$$

Hence, the extended row distances are bounded from below by a linear function with slope  $\frac{n-k+1}{M}$ .

Note that the constant part of this linear function is always positive.

*Proof:* Write the matrix at the top of the following page. We have to estimate  $\text{wt}((u_0, \dots, u_j) G_j^r)$  where all  $u_i \neq 0$ . Thus, we may use the lower bound  $d_{M-1}^c = M(n-k)+1$  for the first block, the lemma for the next  $a-1$  blocks,  $\max\{0, d_{t-1}^c - n\}$  and  $n-k+1$  for the last two blocks, respectively. Hence,

$$\begin{aligned} \hat{d}_j^r &\geq M(n - k) + 1 + (a - 1)(n - k + 1) \\ &\quad + \max\{0, t(n - k) + 1 - n\} + n - k + 1 \\ &= (n - k)(M + a) + a + 1 + \max\{0, t(n - k) + 1 - n\} \\ &= (n - k) \left( M + \left\lfloor \frac{j}{M} \right\rfloor \right) + \left\lfloor \frac{j}{M} \right\rfloor + 1 + \max\{0, t(n - k) + 1 - n\} \\ &\geq (n - k) \left( \frac{j}{M} + M - 1 \right) + \frac{j}{M} + \max\{0, t(n - k) + 1 - n\} \\ &= \frac{n - k + 1}{M} j + (n - k)(M - 1) + \max\{0, t(n - k) + 1 - n\}. \quad \square \end{aligned}$$

*Remark 6.3:* We computed the weight distribution for some of our codes (those with number of states not bigger than 64), and in all cases we even obtain  $\hat{d}_j^r = (n - k)(j + 1) + 2$ . This is in general a much better slope than the estimate of the preceding theorem.

## VII. CONCLUSION

In this paper, we introduced two new classes of convolutional codes called strongly-MDS convolutional codes and codes

$$G_j^r = \left( \begin{array}{c|c|c|c|c}
 \begin{array}{ccc} G_0 & G_1 & \\ & G_0 & \ddots \\ & & \ddots & G_1 \\ & & & G_0 \end{array} & & & & \\
 \hline
 & \begin{array}{ccc} G_1 & & \\ G_0 & G_1 & \\ & G_0 & \ddots \\ & & \ddots & G_1 \\ & & & G_0 \end{array} & G_1 & & \\
 \hline
 & & & \ddots & & \\
 \hline
 & & & & G_1 & \\
 \hline
 & & & \begin{array}{ccc} G_0 & G_1 & \\ & G_0 & \ddots \\ & & \ddots & G_1 \\ & & & G_0 \end{array} & & G_1 \\
 \hline
 \end{array} \right) \in \mathbb{F}^{jk \times (j+1)n}.$$

$\underbrace{\hspace{10em}}_{M \text{ blocks}} \quad \underbrace{\hspace{10em}}_{M \text{ blocks}} \quad \dots \quad \underbrace{\hspace{10em}}_{t \text{ blocks}} \quad \underbrace{\hspace{10em}}_{1 \text{ block}}$   
 $a \text{ times}$

having maximum distance profile. Strongly-MDS convolutional codes have the property that the generalized Singleton bound is attained at the earliest possible column distance. Codes with maximum distance profile have a maximal possible increase of the column distances.

From an applications point of view, strongly-MDS convolutional codes are particularly suited in situations where codes over large alphabets are required and in situations where algebraic decoding is desirable. Hadjicostis [21], [22] has recently demonstrated that convolutional codes over large alphabets are very desirable in areas of process control via linear finite-state machines, where large numbers of nonconcurrent errors should be detected and corrected. It seems that strongly-MDS convolutional codes have potential for such applications.

#### APPENDIX A

We will need the following lemma.

*Lemma A.1:* Let  $A \in \mathbb{F}^{k \times n}$  and  $B \in \mathbb{F}^{n \times (n-k)}$  such that

$$AB = 0 \text{ and } \text{rank } A = k, \quad \text{rank } B = n - k.$$

Then the following are equivalent:

- a) the  $k \times k$ -submatrix of  $A$  consisting of the columns with indices  $1 \leq t_1 < \dots < t_k \leq n$  is singular;
- b) the  $(n-k) \times (n-k)$ -submatrix of  $B$  obtained by taking the rows with indices in  $\{1, \dots, n\} \setminus \{t_1, \dots, t_k\}$  is singular;

*Proof:* Without loss of generality, assume  $(t_1, \dots, t_k) = (1, \dots, k)$  and partition  $A = (A_1 \ A_2)$ , where  $A_1$  is the  $k \times k$  submatrix under consideration. If  $A_1$  is invertible then

$$\ker A = \text{colspan}_{\mathbb{F}} \begin{pmatrix} A_1^{-1} A_2 \\ -I_{n-k} \end{pmatrix} = \text{colspan}_{\mathbb{F}}(B).$$

This shows that the bottom  $(n-k) \times (n-k)$ -submatrix of  $B$  is invertible.  $\square$

#### *Proof of Theorem 2.4:*

i)  $\Rightarrow$  ii): Assume there are indices  $1 \leq t_1 < \dots < t_{(j+1)k}$  satisfying  $t_{sk+1} > sn$  for  $s = 1, \dots, j$  whose corresponding minor is zero. It follows that there is a vector  $u = (u_0, \dots, u_j)$  such that  $uG_j^c$  has zero coordinates at positions  $t_1, \dots, t_{(j+1)k}$ . Let  $\ell := \min\{i \mid u_i \neq 0\}$ . Consider the vector

$$(u_\ell, \dots, u_j)G_{j-\ell}^c \in \mathbb{F}^{(j-\ell+1)n}.$$

The weight of this vector is at most  $(j-\ell+1)(n-k)$  as there are at least  $(j-\ell+1)k$  zero coordinates. From (2.4), it follows that  $d_{j-\ell}^c \leq (j-\ell+1)(n-k)$  and by Corollary 2.3 we also have  $d_j^c \leq (j+1)(n-k)$ , contradicting i).

ii)  $\Rightarrow$  i): Assume that  $d_j^c \leq (n-k)(j+1)$ . Let

$$m := \min\{i \mid d_i^c \leq (n-k)(i+1)\}.$$

It follows that there is a vector  $u = (u_0, \dots, u_m)$ ,  $u_0 \neq 0$  such that  $uG_m^c$  has at least  $k(m+1)$  zeros. As a submatrix inside  $G_j^c$  we select the columns corresponding to the first  $k(m+1)$  positions where  $uG_m^c$  has a zero and we augment it by the last  $k(j-m)$  columns of  $G_j^c$ . We call the indices of the selected columns  $t_1, \dots, t_{(j+1)k}$ . This gives a  $(j+1)k \times (j+1)k$  full-size minor and we claim that this minor is zero and that the indices  $t_1, \dots, t_{(j+1)k}$  satisfy  $t_{sk+1} > sn$  for  $s = 1, \dots, j$ . In order to prove the latter, note that  $d_i^c = (n-k)(i+1) + 1$  for  $i = 0, \dots, m-1$ . It therefore follows that  $(u_0, \dots, u_i)G_i^c$  has at most  $k(i+1) - 1$  zeros for  $i = 0, \dots, m-1$ . In particular,  $t_{sk+1} > sn$  for  $s = 1, \dots, m$ . Clearly, it is also true for  $s = m+1, \dots, j$ . It remains to be shown that the minor is zero. For this, note that the selected matrix has the form  $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ , where  $A$  is an  $(m+1)k \times (m+1)k$  submatrix of  $G_m^c$  which is singular by construction. The full size minor is therefore zero.

As for the equivalence of ii) and iii) recall that  $G_j^c(H_j^c)^T = 0$  and that both matrices have full rank. The minor in  $H_j^c$  complementary to the minor of  $G_j^c$  with the indices as in ii) has indices as given in iii). Therefore, Lemma A.1 completes the proof.  $\square$

*Proof of Corollary 2.5:* In this special case, where  $G$  is systematic, the truncated sliding generator matrix has the form

$$G_j^c = \begin{bmatrix} I & P_0 & 0 & P_1 & \cdots & \cdots & 0 & P_{j-1} & 0 & P_j \\ & I & P_0 & & & & 0 & P_{j-2} & 0 & P_{j-1} \\ & & & \ddots & \ddots & & \vdots & \vdots & \vdots & \vdots \\ & & & & I & P_0 & 0 & P_1 \\ & & & & & I & P_0 & 0 & P_1 \end{bmatrix}.$$

First of all, using Theorem 2.4, item ii) it is easy to see that both conditions in the corollary imply that the matrices  $P_i$  do not contain any zero entries. Therefore, we may assume that all entries of  $P_0, \dots, P_j$  are nonzero. Second, notice that the  $(j+1)k \times (j+1)k$  submatrices  $M$  of  $G_j^c$  are in one to one relation to the square submatrices  $\hat{M}$  of  $\hat{P}$ . Precisely, let  $M$  be obtained from  $G_j^c$  by picking, say,  $a_i$  columns from the block columns containing the identity matrix and, say,  $b_i$  columns from the block column starting with  $P_i$ . Then

$$\sum_{i=0}^j (a_i + b_i) = (j+1)k \quad (\text{A1})$$

and  $M$  satisfies the index condition in Theorem 2.4, item ii) if and only if

$$\sum_{i=0}^t (a_i + b_i) \leq (t+1)k, \quad \text{for all } t = 0, \dots, j. \quad (\text{A2})$$

The submatrix  $M$  contains a  $\hat{b} \times \hat{b}$ -submatrix  $\hat{M}$  of  $\hat{P}$ , where  $\hat{b} := \sum_{i=0}^j b_i$ , and obviously, the matrix  $M$  is nonsingular if and only if  $\hat{M}$  is. Therefore, it remains to prove that  $M$  satisfies the index condition in Theorem 2.4, item ii) if and only if  $\hat{M}$  does not contain an  $s \times t$ -zero block where  $s + t \geq \hat{b} + 1$ . Since all entries of the matrices  $P_i$  are nonzero, the largest zero blocks contained in the submatrix  $\hat{M}$  are of sizes

$$\left( (j-t+1)k - \sum_{i=t}^j a_i \right) \times \sum_{i=0}^{t-1} b_i, \quad \text{for } t = 1, \dots, j.$$

Using (A1) it is easy to see that

$$(j-t+1)k - \sum_{i=t}^j a_i + \sum_{i=0}^{t-1} b_i \leq \hat{b} \iff \sum_{i=0}^{t-1} (a_i + b_i) \leq tk.$$

But this is just the condition in (A2) and, by virtue of Theorem 2.4, that proves the equivalence of i) and ii).  $\square$

## APPENDIX B

We will prove that the proper minors of the matrix  $T_n$  given in Example 3.10, part 2) are all positive. In order to do so consider the matrix

$$X = \begin{bmatrix} 1 & & & & & & & & & \\ 1 & 1 & & & & & & & & \\ & 1 & 1 & & & & & & & \\ & & \ddots & \ddots & & & & & & \\ & & & & 1 & 1 & & & & \\ & & & & & 1 & 1 & & & \\ & & & & & & 1 & 1 & & \end{bmatrix} \in \mathbb{Z}^{n \times n}$$

and notice that for all  $k \in \{1, \dots, n-1\}$  we have

$$X^k = \begin{bmatrix} 1 & & & & & & & & & \\ \binom{k}{1} & 1 & & & & & & & & \\ \binom{k}{2} & \binom{k}{1} & 1 & & & & & & & \\ \vdots & \ddots & \ddots & \ddots & & & & & & \\ \vdots & & \ddots & \ddots & \ddots & & & & & \\ \vdots & & & & \ddots & \ddots & & & & \\ 1 & \cdots & \cdots & \binom{k}{2} & \binom{k}{1} & 1 & & & & \\ & 1 & \cdots & \cdots & \binom{k}{2} & \binom{k}{1} & 1 & & & \\ & & \ddots & & & & & \ddots & & \\ & & & 1 & \cdots & \cdots & \binom{k}{2} & \binom{k}{1} & 1 & \end{bmatrix}. \quad (\text{B1})$$

In particular,  $X^{n-1} = T_n$ . Therefore, the positivity of the proper minors is a consequence of the following theorem.

**Theorem B.1:** Let  $k \in \{1, \dots, n-1\}$  and

$$1 \leq i_1 < \dots < i_r \leq n, \quad 1 \leq j_1 < \dots < j_r \leq n$$

and define  $\hat{X} := (X^k)_{j_1, \dots, j_r}^{i_1, \dots, i_r}$ . Then  $\det \hat{X} \geq 0$  and

$$\det \hat{X} > 0 \iff j_l \in \{i_l, i_l - 1, \dots, i_l - k\} \text{ for all } l = 1, \dots, r.$$

*Proof:*

1) We first show that

$$j_l \notin \{i_l, i_l - 1, \dots, i_l - k\}, \quad \text{for some } l \implies \det \hat{X} = 0. \quad (\text{B2})$$

To this end notice that

$$X_{ij} = 0, \quad \text{for } j > i \text{ or } j < i - k$$

and thus,

$$\hat{X}_{ef} = X_{i_e j_f} = 0, \quad \text{for } j_f > i_e \text{ or } j_f < i_e - k.$$

Assume now  $j_l > i_l$  for some  $l$ . Then for all  $e \leq l$  and  $f \geq l$  we have  $j_f \geq j_l > i_l \geq i_e$  and thus,  $\hat{X}_{ef} = 0$ . Hence, the first  $l$  rows of  $\hat{X}$  have at most rank  $l-1$  and thus,  $\det \hat{X} = 0$ . Similarly, if  $j_l < i_l - k$  for some  $l$ , then we have  $\hat{X}_{ef} = 0$  for all  $e \geq l$  and  $f \leq l$  and the first  $l$  columns of  $\hat{X}$  have at most rank  $l-1$ .

2) It remains to prove the implication " $\Leftarrow$ " of the equivalence given in the theorem.

We begin with proving the statement for  $k = 1$ , i.e., for the matrix  $X$ . In order to do so, we proceed by induction on  $r$ . For  $r = 1$  we have to consider the submatrices  $X_{i_1}^{i_1}$  and  $X_{i_1-1}^{i_1}$ . They all trivially have determinant 1. Now let  $r > 1$ . We suppose the statement is true for all  $(r-1) \times (r-1)$  proper submatrices with the according restriction on the indices and we have to show that the assertion is also true for  $\hat{X} = X_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  where  $j_l \in \{i_l, i_l - 1\}$  for all  $l$ . Notice that the first column of  $\hat{X}$  has either one or two nonzero entries and they are equal to 1. If the first column of  $\hat{X}$  has one 1 only, then it is on the first row. Applying cofactor expansion along that column we obtain

$$\det \hat{X} = 1 \cdot \det X_{j_2, \dots, j_r}^{i_2, \dots, i_r}. \quad (\text{B3})$$

The  $(r-1) \times (r-1)$ -submatrix satisfies  $j_l \in \{i_l, i_l - 1\}$  for all  $l = 2, \dots, r$  and hence by induction has positive determinant. This proves  $\det \hat{X} > 0$  in this case. If the first column of  $X_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  has two entries equal to 1, then they are necessarily

on the first two rows, thus,  $i_2 = i_1 + 1$  and  $j_1 = i_1$ . Since  $j_2 \in \{i_2, i_2 - 1\} = \{i_1 + 1, i_1\}$  and  $j_2 > j_1$ , we can only have  $j_2 = i_1 + 1$ . Then the first row will have only one nonzero entry equal to 1 on the first position, and applying cofactor expansion along that row, we obtain again (B3) and thus  $\det \hat{X} > 0$ . We now proceed by induction on  $k$  in order to prove the desired result for  $X^k$  where  $k > 1$ . Assume  $X^{k-1}$  has the stated property. Using  $X^k = X \cdot X^{k-1}$  and the Cauchy–Binet formula for minors we obtain

$$\det \hat{X} = \sum_{\substack{1 \leq s_1 < \dots < s_r \leq n, \\ s_l \in \{i_l, i_l - 1\} \cap \{j_l, j_l + 1, \dots, j_l + k - 1\}}} \det X_{s_1, \dots, s_r}^{i_1, \dots, i_r} \cdot \det (X^{k-1})_{j_1, \dots, j_r}^{s_1, \dots, s_r}.$$

Due to part 1) of the proof, the sum indeed expands only over the given indices. By induction, all nonsingular submatrices of both matrices  $X$  and  $X^{k-1}$  have positive determinant, hence if there are any nonzero terms in the sum, it is necessarily positive. Therefore, the only thing left to be proven is that there is a nonzero term in the above sum. But all products of the form  $\det X_{i_1, \dots, i_r}^{i_1, \dots, i_r} \cdot \det (X^{k-1})_{j_1, \dots, j_r}^{i_1, \dots, i_r}$  with

$$j_l \in \{i_l, i_l - 1, i_l - 2, \dots, i_l - (k - 1)\}$$

for all  $l$  are nonzero. Thus,  $\det \hat{X} > 0$  and the proof is complete.  $\square$

## APPENDIX C

### *Proof of Theorem 3.11:*

**Step 1:** We will show the existence of a matrix  $\hat{P}$  as in (3.3) satisfying part c) of Theorem 3.1. This can be accomplished as follows. Define  $\tau = (M+1)(n-1)$  and pick a superregular Toeplitz matrix

$$T = \begin{bmatrix} t_0 & 0 & \dots & 0 \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ t_{\tau-1} & \dots & t_1 & t_0 \end{bmatrix} \in \mathbb{F}^{\tau \times \tau}.$$

Theorem 3.9 guarantees the existence of such a matrix over a suitably large field of characteristic  $p$ . For  $l = 0, \dots, M$  define the matrices

$$P_l = T_{1, \dots, k}^{l(n-1)+k, l(n-1)+k+1, \dots, (l+1)(n-1)} \in \mathbb{F}^{(n-k) \times k}.$$

Then, due to the Toeplitz structure of  $T$ , we have for all  $r = 0, \dots, M$

$$P_l = T_{r(n-1)+1, \dots, r(n-1)+k}^{(l+r)(n-1)+k, \dots, (l+r+1)(n-1)} \text{ for } l = 0, \dots, M - r$$

and therefore the matrix  $\hat{P}$  in (3.3) is obtained from  $T$  by picking the rows with indices

$$k, k+1, \dots, n-1, n-1+k, \dots, 2(n-1), \dots, M(n-1)+k, \dots, (M+1)(n-1)$$

and the columns with indices

$$1, \dots, k, n, \dots, n-1+k, 2(n-1)+1, \dots, 2(n-1)+k, \dots, M(n-1)+1, \dots, M(n-1)+k.$$

But then it is obvious that the matrix  $\hat{P}$  inherits from  $T$  the property that all  $j \times j$ -submatrices not containing an  $s \times r$ -zero block where  $s+r \geq j+1$  are nonsingular, cf. Remark 3.7. This provides us with a matrix  $\hat{H}$  as in (3.3) satisfying the equivalent conditions of Theorem 3.1.

**Step 2:** We now establish the existence of a polynomial matrix  $H \in \mathbb{F}[D]^{(n-k) \times n}$  having  $\hat{H}$  in (3.3) as  $M$ th systematic sliding parity-check matrix. For this, define  $m := \frac{\delta}{n-k}$  and start with  $H$  as in (3.1). Without loss of generality we may assume  $A_0 = I_{n-k}$ . Then (3.2) tells us that we need matrices  $A$  and  $B$  satisfying

$$B = A \left( \sum_{i=0}^M P_i D^i + \text{higher terms} \right). \quad (\text{C1})$$

Comparing the coefficients of  $D^{m+1}, \dots, D^M$  we obtain the matrix equation

$$[A_m \quad \dots \quad A_1] \begin{bmatrix} P_{M-m} & \dots & P_1 \\ P_{M-m+1} & \dots & P_2 \\ \vdots & & \vdots \\ P_{M-1} & \dots & P_m \end{bmatrix} = -[P_M \quad \dots \quad P_{m+1}]. \quad (\text{C2})$$

Denote the matrix occurring on the left-hand side by  $\mathcal{P}$ . In order to see that the matrix equation is solvable we will show that  $\mathcal{P}$  has full column rank. Notice that  $\mathcal{P} \in \mathbb{F}^{m(n-k) \times (M-m)k}$ . Using  $M = \lfloor \frac{\delta}{k} \rfloor + \frac{\delta}{n-k}$  and  $m = \frac{\delta}{n-k}$  one can easily see that  $(M-m)k \leq m(n-k)$ . But then the full column rank of  $\mathcal{P}$  follows from part c) of Theorem 3.1. Thus, we can find matrices  $A_1, \dots, A_m$  satisfying (C2). Comparing now the powers of  $D^0, \dots, D^m$  in (C1) we obtain  $B_0, \dots, B_m$ . Then the equation is fully satisfied by setting  $P_{M+1}, P_{M+2}, \dots$  suitably.

**Step 3:** It remains to see that the code

$$\mathcal{C} := \{v \in \mathbb{F}[D]^n \mid vH^T = 0\}$$

has degree  $\delta$ . But this follows directly from the construction. Indeed, recall that  $M = \lfloor \frac{\delta}{k} \rfloor + \frac{\delta}{n-k}$ . Therefore, using Theorem 3.1, part a), we know that the  $M$ th column distance of  $\mathcal{C}$  satisfies

$$\begin{aligned} d_M^c &= (n-k)(M+1) + 1 \\ &= (n-k) \left( \lfloor \frac{\delta}{k} \rfloor + 1 \right) + \delta + 1. \end{aligned}$$

Now the generalized Singleton bound in Theorem 2.6 shows that the code  $\mathcal{C}$  cannot have a degree smaller than  $\delta$ . Thus,  $\mathcal{C}$  is a strongly-MDS  $(n, k, \delta)$  code,  $n-k \mid \delta$  and the proof is complete.  $\square$

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments and for bringing the papers [27], [28] to their attention.

## REFERENCES

- [1] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*. Piscataway, NJ: IEEE Press, 1999.
- [2] S. Lin and D. J. Costello Jr., *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [3] A. R. Calderbank, G. D. Forney Jr., and A. Vardy, "Minimal tail-biting trellises: The Golay code and more," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1435–1455, Jul. 1999.
- [4] J. Justesen, "Upper bounds on the number of errors corrected by a convolutional code," *IEEE Trans. Inf. Theory*, vol. 50, no. 2, pp. 350–353, Feb. 2004.
- [5] P. Ståhl, J. B. Anderson, and R. Johannesson, "Optimal and near-optimal encoders for short and moderate-length tail-biting trellises," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2562–2571, Nov. 1999.
- [6] J. Justesen, "New convolutional code constructions and a class of asymptotically good time-varying codes," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 2, pp. 220–225, Mar. 1973.
- [7] —, "An algebraic construction of rate  $1/\nu$  convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 1, pp. 577–580, Jan. 1975.
- [8] J. Justesen and L. R. Hughes, "On maximum-distance-separable convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, p. 288, Mar. 1974.
- [9] J. L. Massey, D. J. Costello Jr., and J. Justesen, "Polynomial weights and code constructions," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 1, pp. 101–110, Jan. 1973.
- [10] P. Piret, *Convolutional Codes, an Algebraic Approach*. Cambridge, MA: MIT Press, 1988.
- [11] J. Rosenthal and R. Smarandache, "Maximum distance separable convolutional codes," *Appl. Algebra Engrg. Comm. Comput.*, vol. 10, pp. 15–32, 1999.
- [12] R. Smarandache, H. Gluesing-Luerssen, and J. Rosenthal, "Constructions for MDS-convolutional codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 2045–2049, Jul. 2001.
- [13] H. Gluesing-Luerssen and W. Schmale, "On cyclic convolutional codes," *Acta Appl. Math.*, vol. 82, pp. 183–237, 2004.
- [14] P. Piret, "Structure and constructions of cyclic convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 2, pp. 147–155, Mar. 1976.
- [15] C. Roos, "On the structure of convolutional and cyclic convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 6, pp. 676–683, Nov. 1979.
- [16] J. A. Dominguez Perez, J. M. Muñoz Porras, and G. Serrano Sotelo, "Convolutional codes of Goppa type," *Appl. Algebra Engrg. Comm. Comput.*, vol. 15, pp. 51–61, 2004.
- [17] D. J. Costello Jr., "A construction technique for random-error-correcting convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 5, pp. 631–636, Sep. 1969.
- [18] C. Thommesen and J. Justesen, "Bounds on distances and error exponents of unit memory codes," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 5, pp. 637–649, Sep. 1983.
- [19] J. Justesen, E. Paaske, and M. Ballan, "Quasicyclic unit memory convolutional codes," *IEEE Trans. Inf. Theory*, vol. 36, no. 3, pp. 540–547, May 1990.
- [20] J. E. Porath and T. Aulin, "Algorithmic construction of trellis codes," *IEEE Trans. Commun.*, vol. 41, no. 5, pp. 649–654, May 1993.
- [21] C. N. Hadjicostis, "Nonconcurrent error detection and correction in fault-tolerant discrete-time LTI dynamic systems," *IEEE Trans. Circuits Syst. I*, vol. 50, no. 1, pp. 45–55, Jan. 2003.
- [22] —, "Finite-state machine embeddings for nonconcurrent error detection and identification encoded dynamics for fault tolerance in linear finite-state machines," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 142–153, Feb. 2005.
- [23] C. N. Hadjicostis and G. C. Verghese, "Encoded dynamics for fault tolerance in linear finite-state machines," *IEEE Trans. Autom. Control*, vol. 47, no. 1, pp. 189–192, Jan. 2002.
- [24] J. Rosenthal, "Connections between linear systems and convolutional codes," in *Codes, Systems and Graphical Models*, B. Marcus and J. Rosenthal, Eds. New York: Springer-Verlag, 2001, vol. 123, IMA, pp. 39–66.
- [25] J. Rosenthal, J. M. Schumacher, and E. V. York, "On behaviors and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1881–1891, Sep. 1996.
- [26] R. J. McEliece, "The algebraic theory of convolutional codes," in *Handbook of Coding Theory*, V. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998, vol. 1, pp. 1065–1138.
- [27] E. M. Gabidulin, "Convolutional codes over large alphabets," in *Proc. Int. Workshop on Algebraic Combinatorial and Coding Theory*, Varna, Bulgaria, 1988, pp. 80–84.
- [28] E. M. Gabidulin and D. K. Zigangirov, "Further results on convolutional codes over large alphabets," in *Proc. IEEE Int. Workshop on Information Theory*, Moscow, Russia, 1994, pp. 39–40.
- [29] B. M. Allen, "Linear Systems Analysis and Decoding of Convolutional Codes," Ph.D. dissertation, Univ. Notre Dame, Notre Dame, IN, 1999.
- [30] R. M. Roth and A. Lempel, "On MDS codes via Cauchy matrices," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1314–1319, Nov. 1989.
- [31] R. M. Roth and G. Seroussi, "On generator matrices of MDS codes," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 826–831, Nov. 1985.
- [32] A. K. Aydinian, "On matrices with nondegenerate square submatrices," *Probl. Transm. Inform.*, vol. 22, pp. 104–108, 1986.
- [33] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North Holland, 1977.
- [34] R. Hutchinson, J. Rosenthal, and R. Smarandache, "Convolutional codes with maximum distance profile," *Syst. Control Lett.*, vol. 54, pp. 53–63, 2005.
- [35] H. Gluesing-Luerssen, W. Schmale, and M. Striha, "Some small cyclic convolutional codes," in *Proc. 15th Int. Symp. Mathematical Theory of Networks and Systems*, D. Gilliam and J. Rosenthal, Eds., Notre Dame, IN, Aug. 2002.
- [36] J. B. Shearer and R. J. McEliece, "There is no MacWilliams identity for convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 6, pp. 775–776, Nov. 1977.