Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 428 (2008) 2585–2596
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On superregular matrices and MDP convolutional codes

Ryan Hutchinson [a,*], Roxana Smarandache [b], Jochen Trumpf [c,1]

[a] *Department of Mathematics and Computer Science, Bemidji State University, Bemidji, MN 56601-2699, USA*
[b] *Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182-7720, USA*
[c] *Department of Information Engineering, RSISE, The Australian National University, Canberra ACT 0200, Australia*

Received 19 September 2006; accepted 3 February 2008
Available online 1 April 2008
Submitted by U. Helmke

Abstract

Superregular matrices are a type of lower triangular Toeplitz matrix that arises in the context of constructing convolutional codes having a maximum distance profile. These matrices are characterized by the property that the only submatrices having a zero determinant are those whose determinants are trivially zero due to the lower triangular structure. In this paper, we discuss how superregular matrices may be used to construct codes having a maximum distance profile. We also present an upper bound on the minimum size a finite field must have in order that a superregular matrix of a given size can exist over that field. This, in turn, gives an upper bound on the smallest field size over which an MDP (n, k, δ) convolutional code can exist.

Keywords: Convolutional codes; Column distances; Maximum distance profile; Superregular matrices; Partial realization problem

1. Introduction

Convolutional codes are a class of error-correcting codes that have enjoyed wide use in practical applications due to the existence of efficient non-algebraic decoding algorithms. From a mathematical standpoint, however, the situation is still rather unsatisfying, as there are relatively few

* Corresponding author.
E-mail addresses: rhutchinson@bemidjistate.edu (R. Hutchinson), rsmarand@sciences.sdsu.edu (R. Smarandache), Jochen.Trumpf@anu.edu.au (J. Trumpf).

[1] Currently seconded to National ICT Australia Limited, which is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence Program.

0024-3795/$ - see front matter
doi:10.1016/j.laa.2008.02.011

algebraic constructions of convolutional codes having provably good distance properties or an accompanying algebraic decoding algorithm. Recent years have seen interesting developments in the algebraic theory of convolutional codes: the papers [5,7,8,12] extend the notion of cyclicity familiar from block code theory to convolutional codes; the papers [3,9] investigate weight enumerators and the existence of a MacWilliams Identity for convolutional codes; the paper [15] uses methods from systems theory to construct convolutional codes having a designed distance; and the papers [4,6,11,14,17] contain results concerning convolutional codes having certain maximal distance properties. Motivated by existence results proved in this last set of papers, we decided to investigate so-called *superregular matrices*. These matrices arise when one considers the problem of constructing convolutional codes having a maximum distance profile.

The remainder of this paper is structured as follows. The rest of this section contains a brief introduction to convolutional codes, the maximum distance profile property, and the superregularity property. In Section 2, we discuss how superregular matrices may be used to construct codes having a maximum distance profile. In Section 3, we present an upper bound on the minimum field size required for a superregular matrix of a given order to exist. In Section 4, we give some possible directions for future research in this area.

We first recall material from the theory of convolutional codes relevant to the presented work. Let $\mathbb{F}$ be a finite field. A *convolutional code $\mathscr{C}$ of rate k/n* is a rank-k direct summand of the polynomial module $\mathbb{F}[z]^n$. The elements of $\mathscr{C}$ are called *codewords*; when convenient, we will also view codewords as elements of $\mathbb{F}^n[z]$. As a summand of $\mathbb{F}[z]^n$, $\mathscr{C}$ may be viewed as the column space of a basic generator matrix $G(z) \in \mathbb{F}[z]^{n\times k}$; any other generator matrix of $\mathscr{C}$ may be obtained through multiplying $G(z)$ by some unimodular matrix $U(z) \in \mathbb{F}[z]^{k\times k}$. It is well-known (see, for example [2]) that we may assume $G(z)$ to be minimal, which means that its high-order coefficient matrix has full rank. The jth *column degree* of $G(z)$, δ_j, is the maximum degree possessed by an entry of the jth column of $G(z)$. The *degree* of $\mathscr{C}$ is the maximum degree of a polynomial determinant of a $k \times k$ submatrix of $G(z)$. A code of rate k/n and degree δ will be referred to as an (n, k, δ)-code.

We will be looking at convolutional codes from the point of view of linear systems theory, and we next describe briefly how this viewpoint is connected to the definition above. Throughout, 0 will be understood to be the zero matrix or vector of the appropriate size. Given matrices $A \in \mathbb{F}^{\delta\times\delta}$, $B \in \mathbb{F}^{\delta\times k}$, $C \in \mathbb{F}^{(n-k)\times\delta}$, and $D \in \mathbb{F}^{(n-k)\times k}$, with (A, B) a controllable pair and (A, C) an observable pair, one can use a sequence $\{u_t\}_{t\geqslant 0}$ of k-vectors over $\mathbb{F}$ to produce sequences $\{x_t\}_{t\geqslant 0}$ of δ-vectors and $\{y_t\}_{t\geqslant 0}$ of $(n-k)$-vectors via the equations

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t + Du_t, \\ x_0 &= 0. \end{aligned} \tag{1.1}$$

If there exists a $d \in \mathbb{N}_0$ such that $x_{d+1} = 0$ and $u_t = 0$ for $t \geqslant d+1$, we call $\{v_t\}_{t=0}^{d} = \left\{\binom{y_t}{u_t}\right\}_{t=0}^{d}$ a *finite-weight sequence for (A, B, C, D)* .

As explained in [15], the set of finite-weight sequences for (A, B, C, D) corresponds with an (n, k, δ)-code $\mathscr{C}$. More specifically, $\{v_0, v_1, \ldots, v_{d-1}, v_d\}$ is a finite-weight sequence for (A, B, C, D) if and only if $v_d + v_{d-1}z + \cdots + v_1 z^{d-1} + v_0 z^d \in \mathscr{C}$. Let $G(z)$ be a minimal generator matrix for $\mathscr{C}$. We form the matrix $\overline{G(z)}$ by replacing the entry $g_{ij}(z)$ of $G(z)$ by $z^{\delta_j} g_{ij}(z^{-1})$. $\overline{G(z)}$ is also a minimal generator matrix of an (n, k, δ)-code $\overline{\mathscr{C}}$, and the codes $\mathscr{C}$ and $\overline{\mathscr{C}}$ are related by the fact that $v_d + v_{d-1}z + \cdots + v_1 z^{d-1} + v_0 z^d \in \mathscr{C}$ if and only if $v_0 +$

$v_1 z + \cdots + v_{d-1} z^{d-1} + v_d z^d \in \overline{\mathscr{C}}$ [10]. To summarize, then, $\{v_0, v_1, \ldots, v_{d-1}, v_d\}$ is a finite-weight sequence for (A, B, C, D) if and only if $v_0 + v_1 z + \cdots + v_{d-1} z^{d-1} + v_d z^d \in \overline{\mathscr{C}}$. In light of this, we will also refer to the set of finite-weight sequences for (A, B, C, D) as an (n, k, δ)-code, the sequences themselves as codewords, and $\overline{\mathscr{C}}$ as the code represented by (A, B, C, D).

In considering the potential performance of a code, one is often interested in the question of how many errors may be introduced to a codeword without jeopardizing the ability of a decoder to correct them. This leads to measures of distance for convolutional codes, which are defined via the Hamming weight. More precisely, if $v := \{v_t\}_{t=0}^d$ is a codeword, then the *weight* of v, $\mathrm{wt}(v)$, is given by

$$\mathrm{wt}(v) := \sum_{i=0}^{d} \mathrm{wt}(v_i).$$

The distance measure to which the results presented in this work are related is called the *column distance*. Column distances are relevant to the performance of sequential decoding algorithms (see, for example [13]) and are defined as follows:

Definition 1.1. Let $\mathscr{C}$ be a convolutional code. The jth *column distance* of $\mathscr{C}$, $d_j^c(\mathscr{C})$, is defined by

$$d_j^c(\mathscr{C}) := \min_{v \in \mathscr{C}} \left\{ \sum_{i=0}^{j} \mathrm{wt}(v_i) |\ v_0 \neq 0 \right\}$$

(where if $v = \{v_t\}_{t=0}^d$ and $j > d$, then $v_i = 0$ for $d < i \leqslant j$).

In [6], the following results concerning column distances are proved.

Proposition 1.2. *Let $\mathscr{C}$ be an (n, k, δ)-code, and set $L := \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor$. Then*

1. $d_j^c(\mathscr{C}) \leqslant (j+1)(n-k) + 1\ \forall j \in \mathbb{N}_0$.
2. *If* $d_j^c(\mathscr{C}) = (j+1)(n-k)+1$ *for some* j, *then* $d_i^c(\mathscr{C}) = (i+1)(n-k)+1\ \ \forall\ i \in \{0, 1, \ldots, j\}$.
3. *If* $d_j^c(\mathscr{C}) = (j+1)(n-k)+1$, *then* $j \leqslant L$.

An (n, k, δ)-code $\mathscr{C}$ is said to be *maximum distance profile* (*MDP*) if $d_L^c(\mathscr{C}) = (L+1)(n-k)+1$. If $\mathscr{C}$ is MDP, then it follows from statement 2 of Proposition 1.2 that $d_j^c(\mathscr{C}) = (j+1)(n-k)+1$ for all $j \in \{0, 1, \ldots, L\}$. In other words, the column distances of an MDP code are maximal for as long as possible.

We end this section with two definitions.

Definition 1.3. Consider a lower triangular block Toeplitz matrix

$$\mathscr{T} := \begin{bmatrix} T_1 & 0 & \cdots & 0 \\ T_2 & T_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ T_\gamma & \cdots & T_2 & T_1 \end{bmatrix} \in \mathbb{F}^{\gamma(n-k) \times \gamma k},$$

where each block has size $(n-k)\times k$. Let $m \in \{1,2,\ldots,\min\{\gamma(n-k),\gamma k\}\}$. Let $I := \{i_1,\ldots,i_m\}$, with $i_1 < \cdots < i_m$, be a set of row indices of $\mathscr{T}$ and $J := \{j_1,\ldots,j_m\}$, with $j_1 < \cdots < j_m$, a set of column indices of $\mathscr{T}$. We denote by $\mathscr{T}^{i_1,\ldots,i_m}_{j_1,\ldots,j_m}$ the $m\times m$ submatrix of $\mathscr{T}$ formed by intersecting the rows indexed by the members of I with the columns indexed by the members of J. $\mathscr{T}^{i_1,\ldots,i_m}_{j_1,\ldots,j_m}$ is said to be *proper* if, for each $\nu \in \{1,2,\ldots,m\}$, the inequality $j_\nu \leqslant \left\lceil \frac{i_\nu}{n-k} \right\rceil k$ holds.

Setting $n=2$ and $k=1$ gives a lower triangular Toeplitz matrix, and we have

Definition 1.4. A lower triangular Toeplitz matrix

$$\mathscr{T} := \begin{bmatrix} t_1 & 0 & \cdots & 0 \\ t_2 & t_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ t_\gamma & \cdots & t_2 & t_1 \end{bmatrix} \in \mathbb{F}^{\gamma\times\gamma}$$

is said to be *superregular* if every proper submatrix of $\mathscr{T}$ has a nonzero determinant.

One can see that, for any $\gamma \in \mathbb{N}$, there exists a finite field over which a superregular matrix of order γ exists (see Section 4). This leads to the problem of determining the smallest such field for a given γ. As we will see, a solution to this problem would give an upper bound on the size of the smallest finite field over which an MDP (n,k,δ)-code can exist. In Section 3, we take a first step toward solving this problem.

2. Using a superregular matrix to construct an MDP convolutional code

By iterating the equations of (1.1), we see that, if a sequence $\{v_t\}_{t=0}^{j} = \left\{\binom{y_t}{u_t}\right\}_{t=0}^{j}$ consists of the first $j+1$ vectors of a codeword, then it must satisfy the matrix equation

$$\left[\begin{array}{c|cccccc} & D & 0 & \cdots & \cdots & 0 \\ & CB & D & \ddots & & \vdots \\ -I_{(j+1)(n-k)} & CAB & CB & \ddots & \ddots & \vdots \\ & \vdots & \vdots & \ddots & \ddots & 0 \\ & CA^{j-1}B & CA^{j-2}B & \cdots & CB & D \end{array}\right] \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ \underline{y_j} \\ u_0 \\ u_1 \\ \vdots \\ u_j \end{bmatrix} = 0. \tag{2.2}$$

Consider the matrix

$$\mathscr{T}_L := \left[\begin{array}{c|cccccc} & D & 0 & \cdots & \cdots & 0 \\ & CB & D & \ddots & & \vdots \\ -I_{(L+1)(n-k)} & CAB & CB & \ddots & \ddots & \vdots \\ & \vdots & \vdots & \ddots & \ddots & 0 \\ & CA^{L-1}B & CA^{L-2}B & \cdots & CB & D \end{array}\right]. \tag{2.3}$$

If $\mathscr{C}$ is to be an MDP code, then the right-hand side of the matrix $\mathscr{T}_L$ must have the property that all of its proper submatrices have a nonzero determinant [11]. This property will be referred to as the *MDP property* for the matrix $\mathscr{T}_L$.

In [6], it is shown that, for parameters (n, k, δ) such that $(n-k)|\delta$, there exist MDP (n, k, δ)-codes over finite fields of arbitrary characteristic. The idea is that one can use a superregular matrix of order $(L+1)(n-1)$ to form a matrix having the shape of $\mathscr{T}_L$ and having the MDP property. One then uses this matrix to find a parity check matrix for an MDP (n, k, δ)-code. We should note that the work in [6] makes use of the polynomial module representation of a convolutional code instead of the (A, B, C, D)-representation considered in this paper.

As it turns out, one can do something similar if $(n-k)\nmid\delta$. The approach is derived from [10] and works for all parameters (n, k, δ). Let r be the remainder of δ on division by $n-k$. Let $\mathscr{T}$ be a superregular matrix of order $(L+1)(n-1)$ if $r=0$ and $(L+1)(n-1)+k+r-1$ if $r\neq 0$. The first "cutting out" step is done as in [6]: for $l \in \{0, \ldots, L\}$, set

$$F_l := \mathscr{T}_{1,2,\ldots,k}^{l(n-1)+k, l(n-1)+k+1, \ldots, (l+1)(n-1)}. \tag{2.4}$$

If $r \neq 0$, partially define the $(n-k)\times k$ matrix F_{L+1} through

$$F_{L+1} := \left[\begin{array}{c} \mathscr{T}_{1,\ldots,k}^{(L+1)(n-1)+k, \ldots, (L+1)(n-1)+k+r-1} \\ \hline * \end{array}\right],$$

where the bottom $n-k-r$ rows, denoted by $*$, are to be determined so that the sequence of matrices $\{F_1, F_2, \ldots, F_{L+1}\}$ has a minimal partial realization (A, B, C) with A of order δ (for (A, B, C) to be a partial realization means that $CA^{i-1}B = F_i$ for $i \in \{1, 2, \ldots, L+1\}$); that we can do this follows from [10][Theorem 4.3]. Setting $D := F_0$, we may then form a matrix $\mathscr{T}_L$ as in (2.3). By construction, this matrix has the MDP property. In other words, the (n, k, δ)-code represented by (A, B, C, D) is MDP.

We next consider an example to illustrate how one can find an MDP code using a superregular matrix. Consider the matrix

$$\mathscr{T} := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \omega & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \omega^9 & \omega & 1 & 0 & 0 & 0 & 0 & 0 \\ \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 & 0 & 0 \\ \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 & 0 \\ \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 \\ \omega & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 \\ 1 & \omega & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 \end{bmatrix}$$

defined over $\mathbb{F}_{64}$, where ω is a root of the primitive polynomial $x^6 + x + 1 \in \mathbb{F}_2[x]$ and thus a primitive field element. One may check that $\mathscr{T}$ is superregular. We set $(n, k, \delta) = (3, 2, 2)$ for this example, which means that $r = 0$, $L = 3$ and $(L+1)(n-1) = 8$. We may thus use $\mathscr{T}$ to find an MDP $(3, 2, 2)$-code. Using (2.4), we form the matrices

$$F_0 := [\omega \quad 1], \quad F_1 := [\omega^{33} \quad \omega^9], \quad F_2 := [\omega^9 \quad \omega^{33}], \quad F_3 := [1 \quad \omega].$$

We then compute a minimal partial realization

$$A := \begin{bmatrix} \omega^{62} & \omega^3 \\ \omega^{15} & \omega^{56} \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C := \begin{bmatrix} \omega^{33} & \omega^9 \end{bmatrix}$$

for the matrix sequence $\{F_1, F_2, F_3\}$ and set $D := F_0$. Since the matrix

$$\left[\begin{array}{cccc|cccccccc} 1 & 0 & 0 & 0 & \omega & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & \omega & \omega^9 & \omega^{33} & \omega^{33} & \omega^9 & \omega & 1 \end{array}\right]$$

has the MDP property, the code generated by (A, B, C, D) is an MDP $(3, 2, 2)$-code. A minimal generator matrix for this code is

$$G(z) = \begin{bmatrix} \omega^{36} + \omega^{54} z & \omega^{49} + \omega^6 z \\ \omega^{56} + \omega^{62} z & \omega^{48} + \omega^{47} z \\ \omega^{15} & z \end{bmatrix}$$

and one can verify that

$$H(z) = \begin{bmatrix} 1 + \omega^{57} z + \omega^{62} z^2 & \omega + \omega^{44} z + \omega^{54} z^2 & 1 + \omega^{17} z + \omega^{21} z^2 \end{bmatrix}$$

is a parity check matrix. We observe that $H(z)$ is the parity check matrix that one would have obtained by following [6, Appendix C].

3. An upper bound for the required field size

A fundamental question to consider in trying to better understand superregular matrices is that of how large a finite field must be so that a superregular matrix of a given order can exist over that field. For example, no 3×3 superregular matrix exists over the field $\mathbb{F}_2$. By definition, all entries in the lower triangular part of a superregular matrix must be nonzero, which leaves only the matrix

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

this matrix is clearly not superregular, since the lower left 2×2 submatrix has a zero determinant. In this section, we give an upper bound on the required field size by using a counting argument. We begin with two lemmas.

Lemma 3.1. *Let $i \in \mathbb{N}_0$ and $\gamma \in \mathbb{N}$. Define*

$$S_{i,\gamma} := \left\{ \{s_l\}_{l=0}^{i+1} \in \mathbb{N}_0^{i+2} | 0 = s_0 < s_1 < \cdots < s_{i+1} = \gamma,\ s_j + s_{i-j+1} \leqslant \gamma \right. \\ \left. \text{for } j \in \left\{ 0, 1, \ldots, \left\lceil \frac{i}{2} \right\rceil \right\} \right\}$$

and

$$T_{i,\gamma} := \left\{ \{t_l\}_{l=0}^{i+1} \in \mathbb{N}_0^{i+2} | 0 = t_0 < t_1 < \cdots < t_{i+1} = \gamma,\ \sum_{l=0}^{m} (-1)^l (t_{l+1} - t_l) \geqslant 0 \right. \\ \left. \text{for } m \in \{0, 1, \ldots, i\} \right\}.$$

Then $|S_{i,\gamma}| = |T_{i,\gamma}|$.

Proof. We construct injective functions $f : S_{i,\gamma} \to T_{i,\gamma}$ and $g : T_{i,\gamma} \to S_{i,\gamma}$. Throughout the proof, t_{-1} and s_{-1} are defined to be 0.

Let $\{s_l\}_{l=0}^{i+1} \in S_{i,\gamma}$. Starting with $j = 0$, form the sequence $\{t_l\}_{l=0}^{i+1}$ using the recursive formulae

$$t_{2j+1} := t_{2j} + s_{i-j+1} - s_{i-j}, \tag{3.5}$$

$$t_{2j} := t_{2j-1} + s_j - s_{j-1}. \tag{3.6}$$

It follows immediately that $\{t_l\}_{l=0}^{i+1}$ is a strictly increasing sequence and that $t_0 = 0$. Rewriting these formulae gives the identities

$$t_{l+1} - t_l = s_{i-\frac{l}{2}+1} - s_{i-\frac{l}{2}} \quad l \text{ even} \tag{3.7}$$

$$t_{l+1} - t_l = s_{\frac{l+1}{2}} - s_{\frac{l+1}{2}-1} \quad l \text{ odd}. \tag{3.8}$$

Recalling the definition of the set $S_{i,\gamma}$ and using the identities (3.7) and (3.8), one sees that, for $m \in \{0, 1, \ldots, i\}$

$$\begin{aligned}\sum_{l=0}^{m}(-1)^l(t_{l+1} - t_l) &= s_{i+1} + s_0 - \left(s_{\lceil \frac{m}{2} \rceil} + s_{i-\lfloor \frac{m}{2} \rfloor}\right)\\ &= \gamma - \left(s_{\lceil \frac{m}{2} \rceil} + s_{i-\lfloor \frac{m}{2} \rfloor}\right)\\ &\geqslant \gamma - \left(s_{\lceil \frac{m}{2} \rceil} + s_{i-\lceil \frac{m}{2} \rceil+1}\right)\\ &\geqslant 0.\end{aligned}$$

Using the identities (3.7) and (3.8) again and recalling that $t_0 = s_0 = 0$, one also sees that

$$t_{i+1} = t_{i+1} - t_0 = \sum_{l=0}^{i}(t_{l+1} - t_l) = s_{i+1} - s_0 = s_{i+1} = \gamma.$$

It follows that $\{t_l\}_{l=0}^{i+1} \in T_{i,\gamma}$. We may thus define a function $f : S_{i,\gamma} \to T_{i,\gamma}$ through $f(\{s_l\}_{l=0}^{i+1}) = \{t_l\}_{l=0}^{i+1}$, where $\{t_l\}_{l=0}^{i+1}$ is defined via (3.5) and (3.6). It follows immediately from these formulae that f is injective.

We next define a function $g : T_{i,\gamma} \to S_{i,\gamma}$ by rewriting (3.5) and (3.6). Let $\{t_l\}_{l=0}^{i+1} \in T_{i,\gamma}$. Form the sequence $\{s_l\}_{l=0}^{i+1}$ using first the recursive formula

$$s_j := s_{j-1} + t_{2j} - t_{2j-1} \tag{3.9}$$

for $j \in \left\{0, \ldots, \lceil \frac{i}{2} \rceil\right\}$ and then

$$s_{i-j+1} := s_{i-j} + t_{2j+1} - t_{2j} \tag{3.10}$$

for $j \in \left\{0, \ldots, \lfloor \frac{i}{2} \rfloor\right\}$ (starting with $j = \lfloor \frac{i}{2} \rfloor$). It is again immediate that $\{s_l\}_{l=0}^{i+1}$ is strictly increasing and that $s_0 = 0$. Noting that the identities (3.7) and (3.8) again apply, we see that

$$s_{i+1} = s_{i+1} - s_0 = \sum_{l=0}^{i}(t_{l+1} - t_l) = t_{i+1} - t_0 = t_{i+1} = \gamma.$$

As m ranges over the odd numbers of the set $\{1, \ldots, i\}$, $\lceil \frac{m}{2} \rceil$ ranges over the set $\left\{1, \ldots, \lceil \frac{i}{2} \rceil\right\}$. Since $s_0 = 0$ and $s_0 + s_{i+1} = \gamma$, we have

$$\begin{aligned}\gamma - \left(s_{\lceil \frac{m}{2} \rceil} + s_{i-\lfloor \frac{m}{2} \rfloor}\right) &= \gamma - \left(s_{\lceil \frac{m}{2} \rceil} + s_{i-\lceil \frac{m}{2} \rceil+1}\right) \\ &= s_{i+1} + s_0 - (s_{\lceil \frac{m}{2} \rceil} + s_{i-\lceil \frac{m}{2} \rceil+1}) \\ &= \sum_{l=0}^{m} (-1)^l (t_{l+1} - t_l) \\ &\geqslant 0\end{aligned}$$

for these odd values of m. We thus see that $s_j + s_{i-j+1} \leqslant \gamma$ for $j \in \left\{0, 1, \ldots, \left\lceil \frac{i}{2} \right\rceil\right\}$. It follows that $\{s_l\}_{l=0}^{i+1} \in S_{i,\gamma}$. Just as in the first part of the proof, we obtain a function $g : T_{i,\gamma} \to S_{i,\gamma}$ through $g(\{t_l\}_{l=0}^{i+1}) = \{s_l\}_{l=0}^{i+1}$, which is clearly injective. We conclude that $|S_{i,\gamma}| = |T_{i,\gamma}|$. □

Lemma 3.2. *Consider the lower triangular Toeplitz matrix of indeterminates*

$$\mathscr{X} := \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ x_2 & x_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ x_\gamma & \cdots & x_2 & x_1 \end{bmatrix} \in \mathbb{F}[x_1, x_2, \ldots, x_\gamma]^{\gamma \times \gamma}.$$

Let P denote the set of polynomials in which x_γ appears that arise as the determinant of at least one proper submatrix of $\mathscr{X}$. Then

$$|P| \leqslant \frac{1}{2}\left(\frac{1}{\gamma}\binom{2(\gamma-1)}{\gamma-1} + \binom{\gamma-1}{\left\lfloor \frac{\gamma-1}{2} \right\rfloor}\right).$$

Proof. Throughout the proof, we take $\{1, 2, \ldots, s-1\}$ to be the empty set if $s = 1$. Similarly, we take $\{2, \ldots, \gamma - 1\}$ to be the empty set if $\gamma = 1$ or 2.

Since we are concerned only with determinants in which x_γ appears, we restrict our attention to the submatrix $\mathscr{X}_1^\gamma$ and the submatrices $\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}$ with $s \in \{2, \ldots, \gamma - 1\}$. By doing a cofactor expansion along the first column of $\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}$, we see that x_γ appears in $\det\left(\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}\right)$ if and only if $\det\left(\mathscr{X}_{j_1,j_2,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1}}\right) \neq 0$. This is the case if and only if $\mathscr{X}_{j_1,j_2,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1}}$ is a proper submatrix of $\mathscr{X}$ [11], in other words if and only if $j_\nu \leqslant i_\nu$ for all $\nu \in \{1, 2, \ldots, s-1\}$. Thus, we are interested in the set of pairs of sequences $(\{2 \leqslant i_1 < \cdots < i_{s-1} \leqslant \gamma - 1\}, \{2 \leqslant j_1 < \cdots < j_{s-1} \leqslant \gamma - 1\})$, where $j_\nu \leqslant i_\nu$ for all $\nu \in \{1, 2, \ldots, s-1\}$ and $s \in \{2, \ldots, \gamma - 1\}$ Denote this set by N_γ.

We next observe that $j_\nu \leqslant i_\nu$ if and only if $\gamma - i_\nu + 1 \leqslant \gamma - j_\nu + 1$. We also observe that $\mathscr{X}_{1,\gamma-i_{s-1}+1,\ldots,\gamma-i_1+1}^{\gamma-j_{s-1}+1,\gamma-j_{s-2}+1,\ldots,\gamma}$ is the transpose of $\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}$ about the antidiagonal of $\mathscr{X}$. In fact, since $\mathscr{X}$ is symmetric about its antidiagonal, $\mathscr{X}_{1,\gamma-i_{s-1}+1,\ldots,\gamma-i_1+1}^{\gamma-j_{s-1}+1,\gamma-j_{s-2}+1,\ldots,\gamma}$ is the transpose of $\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}$ about the antidiagonal of $\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}$. Since transpose about the antidiagonal does not affect the determinant, it follows that $\det\left(\mathscr{X}_{1,j_1,\ldots,j_{s-1}}^{i_1,i_2,\ldots,i_{s-1},\gamma}\right) = \det\left(\mathscr{X}_{1,\gamma-i_{s-1}+1,\ldots,\gamma-i_1+1}^{\gamma-j_{s-1}+1,\gamma-j_{s-2}+1,\ldots,\gamma}\right)$. Let N'_γ denote the subset of N_γ consisting of those pairs of sequences satisfying $(\{2 \leqslant i_1 < \cdots < i_{s-1} \leqslant \gamma - 1\}, \{2 \leqslant j_1 < \cdots < j_{s-1} \leqslant \gamma - 1\}) = (\{2 \leqslant \gamma - j_{s-1} + 1 < \cdots < \gamma - j_1 + 1, \gamma - 1\}, \{2 \leqslant \gamma - i_{s-1} + 1 < \cdots < \gamma - i_1 + 1 \leqslant \gamma - 1\})$. From the preceding observations, it follows that $|P| \leqslant \frac{1}{2}(|N_\gamma| + |N'_\gamma|)$.

By observing that the members of N_γ are in bijective correspondence with the members of the set considered in problem (j^5) of [18][page 11] (with $n = \gamma - 1$), one sees that $|N_\gamma| = \frac{1}{\gamma}\binom{2(\gamma-1)}{\gamma-1}$, and it thus remains to compute $|N'_\gamma|$. To do this, we first observe that, if $(\{2 \leqslant i_1 < \cdots < i_{s-1} \leqslant \gamma - 1\}, \{2 \leqslant j_1 < \cdots < j_{s-1} \leqslant \gamma - 1\}) \in N'_\gamma$, then $\gamma - i_\nu = j_{s-\nu} - 1$ for all $\nu \in \{1, 2, \ldots, s-1\}$. It follows that this pair of sequences is completely determined by the $s-1$ integers $w_1 := j_1 - 1, w_2 := j_2 - 1, \ldots, w_{s-1} := j_{s-1} - 1$. Since $(\{2 \leqslant i_1 < \cdots < i_{s-1} \leqslant \gamma - 1\}, \{2 \leqslant j_1 < \cdots < j_{s-1} \leqslant \gamma - 1\}) \in N_\gamma$, we have $j_\nu = w_\nu + 1 \leqslant \gamma - w_{s-\nu} = i_\nu$ for all $\nu \in \left\{1, \ldots, \left\lceil \frac{s-1}{2} \right\rceil\right\}$. These inequalities may be rewritten as $w_\nu + w_{s-\nu} \leqslant \gamma - 1$. In other words, $\{0, w_1, \ldots, w_{s-1}, \gamma - 1\} \in S_{s-1,\gamma-1}$. Thus, each member of N'_γ may be associated with a unique member of $S_{s-1,\gamma-1}$. Similarly, each member of $S_{s-1,\gamma-1}$ may be associated with a unique member of N'_γ. We therefore have that $|N'_\gamma| = \sum_{y=0}^{\gamma-2} |S_{y,\gamma-1}|$. From Lemma 3.1, we know that $\sum_{y=0}^{\gamma-2} |S_{y,\gamma-1}| = \sum_{y=0}^{\gamma-2} |T_{y,\gamma-1}|$, and so it is sufficient to compute $\sum_{y=0}^{\gamma-2} |T_{y,\gamma-1}|$.

Suppose $\{t_l\}_{l=0}^{s} \in T_{s-1,\gamma-1}$. To this sequence, we may associate a nonnegative planar walk of length $\gamma - 1$ with $s+1$ vertices. The walk begins at the origin, and steps are given by the vectors $(1, 1)$ and $(1, -1)$. The vertices of the walk are the origin, the endpoint of the walk, and the points where the direction of the walk changes from up to down or from down to up. We make the association by letting t_i be the x-coordinate of the ith vertex. The nonnegativity of the walk is guaranteed by the condition defining membership in $T_{s-1,\gamma-1}$. Conversely, the x-coordinates of the $s+1$ vertices in a nonnegative planar walk of length $\gamma - 1$ may be used to form a sequence in $T_{s-1,\gamma-1}$. Therefore, this association gives a bijective correspondence between sequences in $\cup_{y=0}^{\gamma-2} T_{y,\gamma-1}$ and nonnegative planar walks of length $\gamma - 1$. It is a fact (see, for example [1]) that the number of nonnegative planar walks of length $\gamma - 1$ is given by $\left(\binom{\gamma-1}{\lfloor \frac{\gamma-1}{2} \rfloor}\right)$. This means that $\sum_{y=0}^{\gamma-2} |T_{y,\gamma-1}| = \left(\binom{\gamma-1}{\lfloor \frac{\gamma-1}{2} \rfloor}\right)$. Consequently, $|N'_\gamma| = \left(\binom{\gamma-1}{\lfloor \frac{\gamma-1}{2} \rfloor}\right)$. □

Set $B_\gamma := \frac{1}{2}\left(\frac{1}{\gamma}\binom{2(\gamma-1)}{\gamma-1} + \left(\binom{\gamma-1}{\lfloor \frac{\gamma-1}{2} \rfloor}\right)\right)$. We then have the following theorem.

Theorem 3.3. *Let $\mathbb{F}$ be a finite field such that $|\mathbb{F}| > B_\gamma$. Then, there exists a superregular matrix of order γ over $\mathbb{F}$.*

Proof. The proof is by induction. The claim is clearly true if $\gamma = 1$. Suppose that the claim holds for $\gamma = k$, and let $\mathbb{F}$ be a finite field such that $|\mathbb{F}| > B_{k+1}$. Since $B_{k+1} \geqslant B_k$, we may assume that a superregular matrix

$$\mathscr{T}_k := \begin{bmatrix} t_1 & 0 & \cdots & 0 \\ t_2 & t_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ t_k & \cdots & t_2 & t_1 \end{bmatrix}$$

of order k exists over $\mathbb{F}$. We want to see that we may substitute an element of $\mathbb{F}$ for x in the matrix

$$\mathscr{T}_{k+1} := \begin{bmatrix} t_1 & 0 & \cdots & \cdots & 0 \\ t_2 & t_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ t_k & & \ddots & \ddots & 0 \\ x & t_k & \cdots & t_2 & t_1 \end{bmatrix},$$

so that it is also superregular. This will be the case if and only if we may substitute an element of $\mathbb{F}$ for x so that all of the determinants of submatrices in which x appears are nonzero (the determinants of all other proper submatrices are already nonzero, since $\mathscr{T}_k$ is assumed to be superregular). By Lemma 3.2, there are at most B_{k+1} such determinants. Noting that these determinants are all linear in x, it follows that, since $|\mathbb{F}| > B_{k+1}$, we may find such an element of $\mathbb{F}$. This completes the proof. □

Recalling how a superregular matrix can be used to construct an MDP (n, k, δ)-code, we have

Corollary 3.4. *Let r be the remainder of δ on division by $n-k$. Let $\mathbb{F}$ be a finite field satisfying $|\mathbb{F}| > B_{(L+1)(n-1)}$ or $|\mathbb{F}| > B_{(L+1)(n-1)+k+r-1}$ as $r = 0$ or $r \neq 0$, respectively. Then, an MDP (n, k, δ)-code exists over $\mathbb{F}$.*

For small values of γ, we were able to determine the exact minimum field size required for the existence of a superregular matrix of order γ. The results may be seen in Table 1. Note that we do not claim that each field having a size larger than the minimum required size admits the existence of a superregular matrix, though computer searches have led us to believe that this is in fact the case. Looking at Table 1, it is apparent that the upper bound $B_\gamma + 1$ grows much more quickly than the minimum required field size.

It is still an open problem as to how this bound may be refined. Based on examples in [6], we performed computer searches that led us to make the following conjecture; if true, it would offer a significant improvement to the bound given above:

Conjecture 3.5. For $\gamma \geqslant 5$, there exists a superregular matrix of order γ over the field $\mathbb{F}_{2^{\gamma-2}}$.

4. For future research: finding a construction

At this point, little is understood about how to construct superregular matrices. The problem of finding constructions appears to be a very hard one. One must find a way of guaranteeing that

Table 1
Comparison of minimum required field size and $B_\gamma + 1$

γ	Minimum required field size	$B_\gamma + 1$
3	3	3
4	5	5
5	7	11
6	11	27
7	17	77
8	31	233
9	59	751
10	$\leqslant$127	2495

all proper submatrices with any number of zeroes above the diagonal have a nonzero determinant and do so with additional constraints coming from the Toeplitz structure. In [16], a method is given for constructing, for any prime number p, a triangular Toeplitz array of order p over $\mathbb{F}_p$ having the property that all full square submatrices (submatrices with no zero entries) have a nonzero determinant. Unfortunately, there seems to be no way to extend this construction to the much more general situation considered in this paper.

In [6], the following result is proved:

Theorem 4.1. *For all $\gamma \in \mathbb{N}$, the matrix*

$$X_\gamma := \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \binom{\gamma-1}{1} & 1 & \ddots & & & \vdots \\ \binom{\gamma-1}{2} & \binom{\gamma-1}{1} & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \binom{\gamma-1}{\gamma-2} & \binom{\gamma-1}{\gamma-3} & \cdots & \binom{\gamma-1}{1} & 1 & 0 \\ 1 & \binom{\gamma-1}{\gamma-2} & \cdots & \cdots & \binom{\gamma-1}{1} & 1 \end{bmatrix} \in \mathbb{Z}^{\gamma\times\gamma}$$

has the property that the determinants of all of its proper submatrices are positive.

Thus, for a sufficiently large prime number p, taking the entries of this matrix modulo p results in a superregular matrix. This gives a construction insofar as one knows that, modulo a large enough prime number, the matrix X_γ is superregular. It is not clear, however, how one may give a good bound as to how large p must be for a given γ, and this must also be left for future research.

5. Conclusions

In this paper, we investigated superregular matrices in connection with convolutional codes. In particular, we discussed how superregular matrices may be used to construct codes having a maximum distance profile. The main result of this paper is an upper bound on the minimum size a finite field must have in order that a superregular matrix of a given order can exist over it.

Acknowledgements

The authors would like to thank the referees for their careful reading and comments.

References

[1] W. Feller, An Introduction to Probability Theory and its Applications, third ed., Wiley, New York, 1950.
[2] G. Forney Jr., Minimal bases of rational vector spaces with applications to multi-variable linear systems, SIAM J. Control 13 (3) (1975) 493–520.
[3] H. Gluesing-Luerssen, On the weight distribution of convolutional codes, Linear Algebra Appl. 408 (2005) 298–326.
[4] H. Gluesing-Luerssen, B. Langfeld, A class of one-dimensional MDS convolutional codes, J. Algebra Appl. 5 (2006) 505–520.

[5] H. Gluesing-Luerssen, B. Langfeld, On the algebraic parameters of convolutional codes with cyclic structures, J. Algebra Appl. 5 (2006) 53–76.
[6] H. Gluesing-Luerssen, J. Rosenthal, R. Smarandache, Strongly MDS convolutional codes, IEEE Trans. Inform. Theory 52 (2) (2006) 584–598.
[7] H. Gluesing-Luerssen, W. Schmale, On cyclic convolutional codes, Acta Appl. Math. 82 (2004) 183–237.
[8] H. Gluesing-Luerssen, W. Schmale, On doubly-cyclic convolutional codes, Appl. Algebra Eng. Commun. Comput. 17 (2006) 151–170.
[9] H. Gluesing-Luerssen, G. Schneider, On the Mac Williams identity for convolutional codes, 2006. <http://www.math.rug.nl/∼gluesing/Publ.html>.
[10] R. Hutchinson, The existence of strongly – MDS convolutional codes, 2006. Available from: arXiv math.OC/0801.0184.
[11] R. Hutchinson, J. Rosenthal, R. Smarandache, Maximum distance profile convolutional codes, Systems Control Lett. 54 (1) (2005) 53–63.
[12] B. Langfeld, Minimal cyclic convolutional codes. Master's Thesis, University of Oldenburg, 2003. <http://www-m9.ma.tum.de/dm/homepages/langfeld/thesis.pdf>.
[13] S. Lin, D. Costello Jr., Error Control Coding: Fundamentals and Applications, Prentice-Hall, Englewood Cliffs, NJ, 2004.
[14] J. Rosenthal, R. Smarandache, Maximum distance separable convolutional codes, Appl. Algebra Eng. Commun. Comput. 10 (1) (1999) 15–32.
[15] J. Rosenthal, E. York, BCH convolutional codes, IEEE Trans. Inform. Theory 45 (6) (1999) 1833–1844.
[16] R. Roth, G. Seroussi, On generator matrices of MDS codes, IEEE Trans. Inform. Theory 31 (6) (1985) 826–831.
[17] R. Smarandache, H. Gluesing-Luerssen, J. Rosenthal, Constructions of MDS – convolutional codes, IEEE Trans. Inform. Theory 47 (5) (2001) 2045–2049.
[18] R. Stanley, Catalan Addendum to Enumerative Combinatorics, vol. 2, <http://www-math.mit.edu/∼rstan/ec/catadd.pdf>.