Proc. of 35th Allerton Conference on Communication, Control and Computing, pp. 953–960. Allerton House, Monticello, Illinois, September 29–October 1, 1997.

Construction of Convolutional Codes using Methods from Linear Systems Theory *

Joachim Rosenthal
Department of Mathematics,
University of Notre Dame,
Notre Dame, Indiana 46556.
e-mail: Rosenthal.1@nd.edu

Roxana Smarandache
Department of Mathematics,
University of Notre Dame,
Notre Dame, Indiana 46556.
e-mail: rsmarand@nd.edu

September 26, 1997

Abstract

Using two different matrix pencil descriptions two new classes of convolutional codes with designed lower free distance are introduced.

Keywords: Convolutional codes, linear systems, generalized first order representations, BCH codes, Alternant codes, Grothendieck quot scheme.

1 Introduction

Until recently there did exist very few constructions of convolutional codes with a designed free distance. The constructions which did exist were mainly based on some quasi cyclic constructions of block codes as e.g. in [1, 2, 4, 6] or they made use of the parity check matrix $H(z)$ of the convolutional code as e.g. in [8, 15]. A survey of some of these results can be found in the monograph of Piret [7].

In [12, 14, 16] the first author together with York and Schumacher used some basic ideas in systems theory to derive some convolutional codes of Reed Solomon and BCH type. Different from the previous work in this area convolutional codes were constructed using a first order description of the encoder. It was explained in [14] that the algebraic structure of those codes can be used to algebraically decode those codes.

The paper is structured as follows: In Section 2 we review some of the results reported in [12, 14, 16] and we give conditions on the first order descriptions leading to 'desirable'

*Both authors were supported in part by NSF grant DMS-96-10389.

codes both from the point of view of the designed free distance and the possibility to decode those codes. In Section 3 we show how a small adaptation in the construction presented in [12] will lead to a class of convolutional codes of alternant type. Finally in Section 4 we show that some generalized first order representations well known in the systems literature can be used to derive some convolutional codes with excellent free distance.

2 First order representations of convolutional codes having the form (A, B, C, D) and what makes them desirable

In this paper, following [12, 16], we will define a convolutional code as a submodule $\mathcal{C} \subseteq \mathbb{F}^n[z]$. Since $\mathbb{F}[z]$ is a PID and since $\mathcal{C}$ is a submodule of the free module $\mathbb{F}^n[z]$ the code $\mathcal{C}$ is free as well and it has a well defined rank k. Let $\{g_1(z), \ldots, g_k(z)\} \subset \mathbb{F}^n[z]$ be a basis of the free module $\mathcal{C}$ and let $G(z)$ be the $n \times k$ polynomial matrix whose ith column consists of the polynomial vector $g_i(z); i = 1, \ldots, k$. The convolutional code $\mathcal{C}$ is then equivalently defined as:

$$\mathcal{C} = \left\{ v(z) \in \mathbb{F}^n[z] \mid v(z) = G(z)m(z),\ m(z) \in \mathbb{F}^k[z] \right\}.$$

As usual we will call $G(z)$ an *encoder* of the convolutional code $\mathcal{C}$ and we will say that $\mathcal{C}$ has *rate* k/n if the rank of the module $\mathcal{C}$ is k. If $\tilde{G}(z)$ is a second encoder then there exists a unimodular matrix $U(z) \in Gl_k(\mathbb{F}[z])$ such that $\tilde{G}(z) = G(z)U(z)$. The *free distance* of a convolutional code is defined as

$$d_f(\mathcal{C}) := \min\{\mathrm{wt}(v(z)) \mid v(z) \in \mathcal{C}, v(z) \neq 0\},$$

where wt denotes as usual the weight of a code word. On the side of the rate and the free distance of a convolutional code there is another important parameter called the *complexity*. The complexity δ of a convolutional code $\mathcal{C}$ is defined as the largest degree δ of the $k \times k$ full size minors of one and therefore any encoder $G(z)$. We can naturally identify convolutional codes of complexity zero with block codes. We define a convolutional code to be *observable* if one and therefore every encoder $G(z)$ is right prime. If $G(z)$ is an encoder of an observable convolutional code then $G(z)$ is necessarily a non-catastrophic encoder.

A major design problem in the area of convolutional codes is the construction of an observable code having largest free distance among all possible convolutional codes of a fixed rate and a fixed complexity.

It has been pointed out in [12] that after a possible permutation of the coordinates of $v(z)$, the code $\mathcal{C}$ can be represented through a familiar looking input/state/output description. For this let:

$$v(z) = v_0 z^\gamma + v_1 z^{\gamma-1} + \ldots + v_\gamma;\ v_t \in \mathbb{F}^n, t = 0, \ldots, \gamma.$$

If one partitions the vector v_t into $v_t = \binom{y_t}{u_t}$, where y_t has $n-k$ components and u_t has k components then the convolutional code is equivalently described by the familiar looking '(A, B, C, D)' representation

$$\begin{array}{rcl} x_{t+1} & = & Ax_t + Bu_t \\ y_t & = & Cx_t + Du_t, \quad x_0 = 0, \ x_{\gamma+1} = 0. \end{array} \tag{2.1}$$

If $\mathcal{C}$ has complexity δ then it is possible to choose the matrices (A, B, C, D) of size $\delta \times \delta$, $\delta \times k$, $(n-k) \times \delta$ and $(n-k) \times k$ respectively. A representation having these sizes is called a *minimal representation* and it is algebraically characterized through the condition that (A, B) forms a *controllable* matrix pair, i.e.

$$\text{rank } \left(B \ AB \ldots \ A^{\delta-1}B\right) = \delta. \tag{2.2}$$

As explained in [12] the convolutional code $\mathcal{C}$ defined through (2.1) represents an observable code if and only if the matrix pair (A, C) forms an *observable pair*, i.e.

$$\text{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{\delta-1} \end{pmatrix} = \delta. \tag{2.3}$$

The following theorem gives a general condition on matrices (A, B, C, D) which guarantees a designed distance d for a convolutional code $\mathcal{C}$:

Theorem 2.1 ([14, 16]) *Let d be a positive integer and let*

$$T := d \left\lceil \frac{\delta}{n-k} \right\rceil + 1. \tag{2.4}$$

Assume the matrices (A, B) form a controllable pair with the property that for any integer $\tau > 0$ the matrix

$$\left(A^{\tau}B \ \ A^{\tau+1}B \ \ \ldots \ \ A^{\tau+T-1}B\right) \tag{2.5}$$

defines the parity check matrix of a block code of distance d. If in addition (A, C) forms an observable pair then the convolutional code defined in (2.1) has designed free distance at least d.

The proof of this theorem was carried out in [12] in the situation where (2.5) represents a maximum distance separable code. In the general situation the idea is the same and details are given in [14, 16].

If for every integer τ the block code defined through (2.5) comes with an efficient (algebraic) decoding algorithm then it was shown in [14] that this algebraic decoding algorithm can be used to arrive at an algebraic decoding algorithm of the corresponding convolutional code. Both from the point of view of construction as well as from the point of view of decoding it is therefore desirable to choose matrices (A, B) where the block codes defined in (2.5) have some desirable distances and decoding algorithms.

3 A class of convolutional codes of alternant type

In this section we show how to choose matrices (A, B) such that the block codes appearing in (2.5) are of alternant type. First recall from [12] a Reed-Solomon type construction. For this assume that $d = \delta$, and that α is a primitive of the field $\mathbb{F}_q$ where the number of field elements satisfies $q > Tk$. Define:

$$A := \begin{pmatrix} \alpha^k & 0 & \cdots & 0 \\ 0 & \alpha^{2k} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha^{ck} \end{pmatrix}, \quad B := \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{k-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(k-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha^c & \alpha^{2c} & \cdots & \alpha^{c(k-1)} \end{pmatrix}. \tag{3.1}$$

With those choices it is guaranteed that for each $\tau > 0$ the matrices appearing in (2.5) are of Reed-Solomon type, in particular the parity check matrices appearing in (2.5) describe MDS codes.

We can now adapt this construction slightly. Let h be any nonzero field element and define:

$$A := \begin{pmatrix} h^k\alpha^k & 0 & \cdots & 0 \\ 0 & h^k\alpha^{2k} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h^k\alpha^{ck} \end{pmatrix}, \quad B := \begin{pmatrix} 1 & h\alpha & h^2\alpha^2 & \cdots & h^{k-1}\alpha^{k-1} \\ 1 & h\alpha^2 & h^2\alpha^4 & \cdots & h^{k-1}\alpha^{2(k-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & h\alpha^c & h^2\alpha^{2c} & \cdots & h^{k-1}\alpha^{c(k-1)} \end{pmatrix}. \tag{3.2}$$

One immediately verifies that the block codes appearing in this case in (2.5) have now an 'alternant structure'. One can show that the subfield construction carried through in the situation when (A, B) are of the form (3.1) still can be carried through in this more general situation and the decoding algorithm presented in [14] still applies.

The distinct advantage to work with general matrices of type (3.2) is the fact that BCH block codes are known to be asymptotically bad whereas it is known that there are Alternant block codes which asymptotically are good.

We will address those issues in future work.

4 Construction of convolutional codes using generalized first order descriptions

In this section we explain that it is possible to construct some good convolutional codes using a purely module theoretic approach. In particular we will in this section not use an (A, B, C, D) description for the convolutional code.

It was first pointed out in [13] that a submodule $\mathcal{C} \subset \mathbb{F}^n[z]$ having rank k and complexity δ can also be represented in the following way:

Theorem 4.1 *Assume $\mathcal{C} \subseteq \mathbb{F}^n[z]$ is a rate $\frac{k}{n}$ convolutional code of complexity δ. Then there exist matrices P, Q having sizes $\delta \times (\delta + k)$, and a matrix R having size $n \times (\delta + k)$ (all defined over $\mathbb{F}$) such that the code $\mathcal{C}$ is described by*

$$\mathcal{C} = \{v(z) \in \mathbb{F}^n[z] \mid v(z) = R\zeta(z),\ (zP + Q)\zeta(z) = 0\}. \tag{4.1}$$

Moreover the following conditions are satisfied:

(1) P has full row rank;

(2) $\begin{bmatrix} P \\ R \end{bmatrix}$ has full column rank;

(3) $\begin{bmatrix} zP + Q \\ R \end{bmatrix}$ is right prime.

The representation given in Theorem 4.1 is dual to the '(K, L, M)' representation given in [12, Theorem 3.1.]. Both the (K, L, M) representation and above representation are well studied in the systems theory literature and we refer e.g. to [3], where also further references can be found.

Clearly if the triple of matrices (P, Q, R) have the sizes of Theorem 4.1 then (4.1) defines a submodule of $\mathbb{F}^n[z]$, i.e. a convolutional code. If (P, Q, R) is not a minimal triple then this convolutional code has in general neither rank k nor complexity δ. The next theorem will show that a convolutional code defined by (4.1) has rank k and has complexity δ as soon as the minimality conditions of Theorem 4.1 are satisfied.

It arises the question of how 'unique' the representation (4.1) is. The following Theorem answers this question. The theorem is well known in the systems literature (see e.g. [3, Theorem 4.34]) and it is the dual version of a theorem given in [12]. In systems theory the base field $\mathbb{F}$ is usually assumed to consist of the real numbers however a study of the proof in [3] verifies that the result is valid over any base field:

Theorem 4.2 *The matrices (P, Q, R) introduced in Theorem 4.1 are unique in the following way: if $(\tilde{P}, \tilde{Q}, \tilde{R})$ is a second triple of matrices describing the code $\mathcal{C}$ through*

$$\mathcal{C} = \{v(z) \in \mathbb{F}^n[z] \mid v(z) = \tilde{R}\zeta(z),\ (z\tilde{P} + \tilde{Q})\zeta(z) = 0\}.$$

and if (P, Q, R) and $(\tilde{P}, \tilde{Q}, \tilde{R})$ both satisfy the minimality conditions of Theorem 4.1 then there exist unique invertible matrices T and S such that

$$(\tilde{P}, \tilde{Q}, \tilde{R}) \ = \ (SPT^{-1}, SQT^{-1}, RT^{-1}). \tag{4.2}$$

Moreover every minimal triple which describes the convolutional code $\mathcal{C}$ is of the form (4.2). Finally the triple (P, Q, R) (and hence any triple of the form (4.2)) describes a rate k/n convolutional code of complexity δ.

For the purpose of constructing convolutional codes of a certain rate and a certain complexity it is therefore no limitation to work with minimal triples (P, Q, R) satisfying the minimality conditions of Theorem 4.1. Algorithms are available to go from this description to a polynomial encoder description.

There remains of course the problem of identifying those convolutional codes which are observable, i.e. in particular those codes which have non-catastrophic encoders. The following Lemma provides an answer to this question:

Lemma 4.3 *A matrix triple (P, Q, R) satisfying the minimality conditions of Theorem 4.1 defines an observable convolutional code if and only if the matrix pencil $[zP + Q]$ is left prime.*

Proof: Direct consequence of [11, Theorem 3.3]. □

In [14] it was shown that the Reed Solomon type convolutional codes achieve a free distance which is approximately $\frac{k}{n}$ times the best possible free distance found among all convolutional codes of rate $\frac{k}{n}$ and complexity δ. For low rate those code are therefore not as good. Justesen did construct in [2] rate $\frac{1}{n}$ convolutional codes of degree δ with optimal free distance. In the remainder of this section we will show how to construct (using the representations of Theorem 4.1) some low rate convolutional codes having a free distance superior to the free distance of the codes constructed in [12]. We start with a Lemma:

Lemma 4.4 *There exist matrices (P, Q) having sizes $\delta \times (\delta + k)$ such that the following is true:*

1. *If $\zeta(z) \in \mathbb{F}^{\delta+k}$ is a polynomial vector whose degree is strictly less than $\lfloor \frac{\delta}{k} \rfloor$ and if $(zP + Q)\zeta(z) = 0$ then necessarily $\zeta(z) = 0$.*

2. *$[zP + Q]$ is left prime.*

There are many ways of constructing such matrices (P, Q). A particular example could be $P = (I_\delta, 0)$ and $Q = (A, B)$ where (A, B) are either defined by (3.1) or by (3.2).

Now we can show that for certain low rates we can obtain better distance codes than the ones presented in [12, 14]:

Theorem 4.5 *Let (P, Q) be defined as in Lemma 4.4. Assume $n > \delta + k$ and let R be a matrix of size $n \times (\delta + k)$. If the columns of R form the generator matrix of a maximum distance separable code then the code defined by the triple (P, Q, R) has rate $\frac{k}{n}$, complexity δ and free distance at least $\left(\lfloor \frac{\delta}{k} \rfloor + 1\right)(n - \delta - k + 1)$.*

Proof: Let $v = v_0 + v_1 z + \cdots + v_\gamma z^\gamma$ be a code vector of degree γ and assume that $v_0 \neq 0$. Then, using the sliding matrix representation, we have

$$\begin{bmatrix} v_o \\ v_1 \\ \vdots \\ v_\gamma \end{bmatrix} = \begin{bmatrix} R & 0 & \cdots & 0 \\ 0 & R & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & R \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_\gamma \end{bmatrix}.$$

The first $\left(\lfloor \frac{\delta}{k} \rfloor + 1\right)$ components of the vector $\zeta(z)$ have to be nonzero by our assumption on (P, Q). Because the columns of R define a MDS code the claim follows. □

We conclude this section with some remarks about the algebraic geometric ingredient of the pencil descriptions considered in this section.

As it was explained in [10] a submodule of rank k and complexity δ in $\mathbb{F}^n[z]$ describes a quotient sheaf of rank k and degree δ. By a general theorem of Grothendieck it is possible to equip the set of all rank k submodules (quotient sheafs) of complexity (i.e. degree) δ with the structure of a scheme. Such a scheme is referred to as a *quot scheme* in the algebraic geometry literature. The quot scheme which parameterizes the rank k submodules of complexity δ turns out to be a smooth projective variety [9].

If the degree $\delta = 0$ the Grothendieck quot scheme is exactly the Grassmann variety $\mathrm{Grass}(k, \mathbb{F}^n)$ consisting of all k dimensional subspaces of the vector space $\mathbb{F}^n$. This variety parameterizes the set of all linear block codes of rate $\frac{k}{n}$ defined over the field $\mathbb{F}$. For an arbitrary complexity δ the Grothendieck quot scheme parameterizes in a natural way all rate $\frac{k}{n}$ convolutional codes of complexity δ.

Linear systems described by matrix triples (P, Q, R) have been studied widely in the systems literature and probably the most comprehensive account is given in the monograph of Kuijper [3]. It was pointed out by Lomadze [5] that a matrix pencil of the form $\left[\begin{smallmatrix} zP+Q \\ R \end{smallmatrix}\right]$ represents exactly the linear free resolution of the associated quotient sheaf and in this way such matrix pencils appear naturally in the algebraic geometry literature as well. Finally we would like to note that we can view (4.2) as a group action of the reductive group $Gl_\delta \times Gl_{\delta+k}$ on the vector space consisting of all matrix triples (P, Q, R) of a fixed size. The uniqueness Theorem 4.2 expresses the fact that the group orbits in (4.2) correspond to the submodules of $\mathbb{F}^n[z]$, i.e. the convolutional codes.

Actually much more is true: The geometric quotient in the sense of GIT (=geometric invariant theory) induced by the group action (4.2) is exactly the Grothendick quot scheme. The minimality conditions provided in Theorem 4.1 guarantee that the associated orbit is a 'stable orbit' in the sense of GIT. This is true for an arbitrary base field and this statement is a geometric formulation of the uniqueness Theorem 4.2. The reader who is interested in more details is referred to [10].

References

[1] J. Justesen. New convolutional code constructions and a class of asymptotically good time-varying codes. *IEEE Trans. Inform. Theory*, IT-19(2):220–225, 1973.

[2] J. Justesen. An algebraic construction of rate $1/\nu$ convolutional codes. *IEEE Trans. Inform. Theory*, IT-21(1):577–580, 1975.

[3] M. Kuijper. *First-Order Representations of Linear Systems*. Birkhäuser, Boston, 1994.

[4] Y. Levy and D.J. Costello Jr. An algebraic approach to constructing convolutional codes from quasi-cyclic codes. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 14:189–198, 1993.

[5] V. Lomadze. Finite-dimensional time-invariant linear dynamical systems: Algebraic theory. *Acta Appl. Math*, 19:149–201, 1990.

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

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

[8] Ph. Piret. A convolutional equivalent to Reed-Solomon codes. *Philips J. Res.*, 43(3-4):441–458, 1988.

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

[10] M. S. Ravi and J. Rosenthal. A general realization theory for higher order linear differential equations. *Systems & Control Letters*, 25(5):351–360, 1995.

[11] M. S. Ravi, J. Rosenthal, and J. M. Schumacher. Homogeneous behaviors. *Math. Contr., Sign., and Syst.*, 10:61–75, 1997.

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

[13] J. Rosenthal and E.V. York. Linear systems defined over a finite field, dynamic programming and convolutional codes. In *Proc. of the IFAC Conference on System Structure and Control*, pages 466–471, Nantes, France, July 1995.

[14] J. Rosenthal and E.V. York. BCH convolutional codes. Preprint, October 1997.

[15] R.M. Tanner. Convolutional codes from quasi-cyclic codes: A link between the theories of block and convolutional codes. Computer Research Laboratory, Technical Report, USC-CRL-87-21, November 1987.

[16] E.V. York. *Algebraic Description and Construction of Error Correcting Codes, a Systems Theory Point of View.* PhD thesis, University of Notre Dame, 1997. Available at http://www.nd.edu/~rosen/preprints.html.