Rosenthal's Decoding Algorithm for Certain 1-Dimensional Convolutional Codes

Ángel Luis Muñoz Castañeda, José M. Muñoz-Porras, and Francisco J. Plaza-Martín

***Abstract*— The performance of Rosenthal's decoding algorithm is measured when it is applied to certain family of 1-dimensional MDS convolutional codes. It is shown that the performance of this algorithm depends on the algebraic structure of the convolutional code we are working with.**

***Index Terms*— Algebro-geometric codes, convolutional codes, maximum distance separable (MDS) codes, input/state/output realizations, algebraic decoding.**

I. Introduction

Convolutional codes were introduced by P. Elias [1] and were aimed at constructing codes whose correcting capability were better than that of block codes.

In the last decades, much work have been done in order to establish the mathematical formalism of convolutional coding theory, having special relevance the commutative algebra point of view, presenting a convolutional code as a submodule of a free $\mathbb{F}[z]$-module of finite rank, $\mathcal{C} \subset \mathbb{F}[z]^n$ (see the classical references [2], [3]), as well as the linear dynamical system point of view, presenting a convolutional code as a sequential machine whose states dynamics is governed by finitely many linear equations ([4]–[6] among many others). Likewise, the theory of convolutional codes can be established by using purely algebro-geometric techniques (see [7]), having especial relevance the theory of convolutional Goppa codes ([8]–[11])

Among the parameters that determine the properties of a convolutional code, say length, dimension etc, one of the most important is its free distance, since it is closely related to the error-correcting capability of the code [12]. Larger the free distance is, larger the error-correcting capability of the code is. A very important fact in this direction is the existence of an upper bound (generalized Singleton bound) for the free distance for fixed dimension and length [13]. Therefore, those convolutional codes that achieves the Singleton bound (MDS convolutional codes) are of much interest.

An important issue in coding theory is the construction of decoding schemes. If the transmission channel is assumed to be noise-free, a decoding scheme can be easily constructed. However, when the transmission channel is noisy the problem becomes harder. In this case, there are two essentially different kind of decoding schemes: probabilistic decoders [14], [15] and algebraic decoders [16]–[18]. Despite the former are efficient, they are computationally expensive, while the last ones make use of less physical resources.

In this paper, we look at the performance of Rosenthal's decoding algorithm (see [18]) on certain families of MDS convolutional codes (see [11]) compared with their free distances. Our strategy consists on computing explicitly the input/state/output (I/S/O) representation ([6]) of each member of the family of codes and, second, to determine conditions on the parameters that define the family in order for them to fulfill the assumptions given in Rosenthal's Theorem [18, Theorem 3.4].

II. A Family of 1-Dimensional MDS Convolutional Codes

A. The Family

Let p be a prime number and $q = p^r$ some power of p. Consider the family of convolutional codes parametrized by $\mathbb{F}_q$ generated by the family of convolutional encoders $G_{\{\lambda_i\}}(z) : \mathbb{F}_q[z] \to \mathbb{F}_q[z]^n$ given by

$$G_{\{\lambda_i\}}(z) := \begin{pmatrix} \lambda_0 + \lambda_1(a_1 z + b_1) + \ldots + \lambda_\delta (a_1 z + b_1)^\delta \\ \vdots \\ \lambda_0 + \lambda_1(a_n z + b_n) + \ldots + \lambda_\delta (a_n z + b_n)^\delta \end{pmatrix} \tag{1}$$

with $\lambda_i \in \mathbb{F}_q$ and $\delta \in \mathbb{N}$ being the complexity of the code. This is the convolutional Goppa encoder defined over the projective line by the divisors $G = \delta p_\infty$ and $D = \sum p_i$, with $p_i = a_i z + b_i$, and the restriction of the evaluation map to the subspace $< s := \lambda_0 + \lambda_1 t + \ldots + \lambda_\delta t^\delta > \subset L(G)$ (see [11, §IV] for more details). We assume that $\lambda_\delta, a_n \neq 0$. Thus, $G_{\{\lambda_i\}}(z)$ (from now on, just $G(z)$) is a minimal encoder. If we span the entries of $G(z)$ we get the following expressions for its rows,

$$(G(z))_l = \sum_{t=0}^{\delta} \lambda_t^{(l)} z^t, \ \lambda_t^{(l)} = \sum_{i=t}^{\delta} \lambda_i \begin{pmatrix} i \\ t \end{pmatrix} b_l^{i-t} a_l^t, \tag{2}$$

for $l = 1, \ldots, n$. Let $s_\lambda^{(t)}(u) = \sum_{i=t}^{\delta} \lambda_i \begin{pmatrix} i \\ t \end{pmatrix} u^{i-t}$ and suppose that $b_1, \ldots, b_n$ are equal and non-zero, and that $a_1, \ldots, a_n$ are pair-wise different and non-zero. Note that in this case $\lambda_t^{(l)} = a_l^t s_\lambda^{(t)}(b)$. Following [11, Lemma IV.1], the convolutional code generated by $G(z)$ is MDS if and only

Manuscript received August 27, 2018; revised April 29, 2019; accepted May 21, 2019. Date of publication June 6, 2019; date of current version November 20, 2019. This work was supported by the Research Project MTM2015-66760-P of MINECO, Spain.

Á. L. Muñoz Castañeda is with the Departamento de Matemáticas, Universidad de León, 24071 León, Spain (e-mail: amunc@unileon.es).

J. M. Muñoz-Porras and F. J. Plaza-Martín are with the Departamento de Matemáticas, Universidad de Salamanca, 37008 Salamanca, Spain (e-mail: jmp@usal.es; fplaza@usal.es).

Communicated by V. Sidorenko, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2019.2921370

if $s_\lambda^{(t)}(b) \neq 0$ for all t. From now on, we will be assuming the following conditions on the parameters defining the family of convolutional codes,

Assumptions 1. 1) *$b_1, \ldots, b_n$ are equal and non-zero,*
2) *$a_1, \ldots, a_n$ are pair-wise different and non-zero,*
3) *$s_\lambda^{(t)}(b) \neq 0$ for all t, where $b := b_1 = \ldots = b_n$.*

B. I/S/O Representation

Given a (k, n) convolutional code $\mathcal{C} \subset \mathbb{F}_q[z]^n$ of complexity δ, there are matrices $K \in \mathbb{F}_q^{(\delta+n-k)\times\delta}$, $L \in \mathbb{F}_q^{(\delta+n-k)\times\delta}$, $M \in \mathbb{F}_q^{(\delta+n-k)\times n}$ representing the code

$$\mathcal{C} = \{v(z) \in \mathbb{F}_q[z]^n : \exists x(z) \in \mathbb{F}_q[z]^\delta \text{ s.t. } zKx(z) + Lx(z) + Mv(z) = 0\}, \quad (3)$$

and satisfying certain (minimal) properties (see [6, Theorem 5.1.1]). Such a triple is called a *minimal first order representation* of the convolutional code $\mathcal{C}$. Furthermore, there is always a permutation $\sigma \in \mathfrak{S}_n$ of the components of the code-words such that any minimal first order representation, (K, L, M), of the equivalent convolutional code $\sigma(\mathcal{C})$ admits a similarity transformation turning it into a minimal first order representation of the form (see [6, §5.2])

$$\left(\begin{pmatrix} -I \\ 0 \end{pmatrix}, \begin{pmatrix} A \\ C \end{pmatrix}, \begin{pmatrix} 0 & B \\ -I & D \end{pmatrix} \right). \quad (4)$$

The matrices $A \in \mathbb{F}_q^{\delta\times\delta}$, $B \in \mathbb{F}_q^{\delta\times k}$, $C \in \mathbb{F}_q^{(n-k)\times\delta}$, $D \in \mathbb{F}_q^{(n-k)\times k}$ determine a reachable linear dynamical system (called *input/state/output realization*, or just I/S/O to be shorter) that describes the encoding process for $\mathcal{C}$. In fact, if we denote $x(z) = x_0 z^\gamma + x_1 z^{\gamma-1} + \ldots + x_\gamma$, $u(z) = u_0 z^\gamma + u_1 z^{\gamma-1} + \ldots + u_\gamma$ and $y(z) = y_0 z^\gamma + y_1 z^{\gamma-1} + \ldots + y_\gamma$, then the equality given in (3) for K, L, M as in (4) turns into

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t + Du_t, \\ v_t &= \begin{pmatrix} y_t \\ u_t \end{pmatrix}, \ x_0 = 0, \ x_{\gamma+1} = 0 \end{aligned} \quad (5)$$

Furthermore, we can proceed in the other way around. Given δ, k, n and matrices A, B, C, D defining a linear dynamical system as in (5) we can construct K, L, M as in (4), which in turn describe a convolutional code as given in (3). This close relation between convolutional codes and linear dynamical systems given by I/S/O representations allows to reflect very important properties of codes in the linear dynamical systems side and viceversa (see [6] for a detailed exposition).

I/S/O representations have been proved to be a very useful tool to construct convolutional codes, but also to define decoding schemes for them.

In order to apply Rosenthal's decoding algorithm, the I/S/O representation is required. The matrices (K, L, M) corresponding to a minimal first order representation of the convolutional code (1) are given by

$$K = \begin{pmatrix} -1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & 0 \\ -\lambda_1^{(1)} & -\lambda_2^{(1)} & \ldots & -\lambda_{\delta-1}^{(1)} & -\lambda_\delta^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ -\lambda_1^{(n)} & -\lambda_2^{(n)} & \ldots & -\lambda_{\delta-1}^{(n)} & -\lambda_\delta^{(n)} \end{pmatrix},$$

$$L = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -\lambda_0^{(1)} & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\lambda_0^{(n)} & 0 & 0 & \ldots & 0 \end{pmatrix}, M = \begin{pmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \\ 1 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 1 \end{pmatrix}$$

Observe that the full size minor given by the first block and the first $n-1$ columns of the last block

$$U = \left(\begin{array}{ccc|cccc} & -\mathrm{Id} & & & & 0 & \\ \hline -\lambda_1^{(1)} & \ldots & -\lambda_{\delta-1}^{(1)} & -\lambda_\delta^{(1)} & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\lambda_1^{(n-1)} & \ldots & -\lambda_{\delta-1}^{(n-1)} & -\lambda_\delta^{(n-1)} & 0 & \ldots & 1 \\ -\lambda_1^{(n)} & \ldots & -\lambda_{\delta-1}^{(n)} & -\lambda_\delta^{(n)} & 0 & \ldots & 0 \end{array} \right)$$

has determinant $(-1)^{n+\delta}\lambda_\delta^{(n)}$, which is non-zero by assumption. Therefore, it is invertible and the matrix $U^{-1}(K, L, M)$ is of the form given in (4). Then, the matrices A, B corresponding to an I/S/O representation are:

$$A = \begin{pmatrix} 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\lambda_0^{(n)}}{\lambda_\delta^{(n)}} & \frac{\lambda_1^{(n)}}{\lambda_\delta^{(n)}} & \frac{\lambda_2^{(n)}}{\lambda_\delta^{(n)}} & \ldots & \frac{\lambda_{\delta-1}^{(n)}}{\lambda_\delta^{(n)}} \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -\frac{1}{\lambda_\delta^{(n)}} \end{pmatrix}, \quad (6)$$

while the matrices $C \in$ and D are given by

$$C_{ij} = -\lambda_j^{(i)} + \frac{\lambda_\delta^{(i)}\lambda_j^{(n)}}{\lambda_\delta^{(n)}}, \ D_i = -\frac{\lambda_\delta^{(i)}}{\lambda_\delta^{(n)}} \quad (7)$$

Doing the computations for the code (1), it holds that

$$C_{ij} = a_i^\delta (a_i^{j-\delta-1} - a_n^{j-\delta-1}) s_\lambda^{(j-1)}(b).$$

III. APPLICATION OF ROSENTHAL'S DECODING ALGORITHM

Rosenthal's decoding algorithm [18] is based on the local description of trajectories theorem and depends on certain assumptions and certain parameters.

Assumptions 2. *In order to apply the decoding algorithm on a convolutional code of complexity δ described by an I/S/O realization (A, B, C, D), we have to assume that there are natural numbers $T > \Theta$ such that*

1) *A is invertible, the controllability matrix $\Phi_T(A, B)$ given by*

$$\Phi_T(A, B) = \left(B, AB, \ldots, A^\delta B, \ldots, A^{T-1}B \right) \quad (8)$$

has full row rank δ and its rows form a parity check matrix of a block code. Let d_1 be the minimum distance of this block code.

2) *The observability matrix $\Psi_\Theta(A, C)$ given by*

$$\Psi_\Theta(A, C)^t = \left(C^t, (CA)^t, \ldots, (CA^\delta)^t, \ldots, (CA^{T-1})^t\right) \tag{9}$$

has full column rank δ and its columns form a generator matrix of a block code. Let d_2 be the minimum distance of this block code.

The performance of the algorithm is determined in the following theorem.

Theorem III.1. [18, Theorem 3.4] *Let A, B, C, D be matrices satisfying the conditions of Assumption 2 and consider the received message $\left\{\begin{pmatrix} \widehat{y}_t \\ \widehat{u}_t \end{pmatrix}\right\}_{t\geq 0} = \left\{\begin{pmatrix} y_t + f_t \\ u_t + e_t \end{pmatrix}\right\}_{t\geq 0}$. Assume that for any $\tau \geq 0$ the error sequence $\begin{pmatrix} f_\tau \\ e_\tau \end{pmatrix}, \ldots, \begin{pmatrix} f_{\tau+T-1} \\ e_{\tau+T-1} \end{pmatrix}$ has weight at most $\lambda := \min\left(\left\lfloor \frac{d_1-1}{2} \right\rfloor, \left\lfloor \frac{T}{2\Theta} \right\rfloor\right)$. In this situation it is possible to uniquely compute the transmitted sequence $\left\{\begin{pmatrix} y_t \\ u_t \end{pmatrix}\right\}_{t\geq 0}$.*

When applying Rosenthal's decoding algorithm to the family (1), we have to be sure that there are members in our family satisfying Assumptions 2:

i) Observe that $(-1)^\delta \lambda_\delta^n \det(A + z\mathrm{Id}) = s(p_n) = \lambda_0 + \lambda_1(a_n z + b) + \ldots + \lambda_\delta(a_n z + b)^\delta$. Therefore, $\det(A) = \frac{(-1)^\delta}{\lambda_\delta^{(n)}}(\lambda_0 + \lambda_1 b + \ldots + \lambda_\delta b^\delta)$, so $\det(A) \neq 0$ if and only if b is not a root of the polynomial $q_\lambda(t) = \lambda_0 + \lambda_1 t + \ldots + \lambda_\delta t^\delta$. We denote

$$W = \left\{ \begin{array}{c} (\lambda_0, \ldots, \lambda_\delta) \text{ such that} \\ \sum_{k=0}^\delta \lambda_k b^k \neq 0 \end{array} \right\} \subset \mathbb{A}^{\delta+1}.$$

Note that W is a dense open subset (respect to Zariski topology).

ii) Observe also that by [11, IV, Equation (13)] and Assumptions 1, there is an open subset

$$V = \bigcap_{1\leq i<j\leq n} \left\{ \begin{array}{c} (\lambda_0, \ldots, \lambda_\delta) \text{ such that} \\ \sum_{k=0}^\delta \lambda_k (b_j - b_i)^k (a_i - a_j)^{\delta-k} \neq 0 \end{array} \right\} = \left\{ \begin{array}{c} (\lambda_0, \ldots, \lambda_\delta) \text{ such that} \\ \lambda_0 \neq 0 \end{array} \right\} \subset \mathbb{A}^{\delta+1}$$

such that if $\lambda_0, \ldots, \lambda_\delta \in V$ then the convolutional code generated by the encoder $G(z)$ is observable. Therefore, the matrices (A, B, C, D) satisfy that $\mathrm{rank}(\Phi_\delta(A, B)) = \mathrm{rank}(\Phi_\delta(A, C)) = \delta$ (see [6, Lemma 5.3.5]), so any pair of integers $T > \Theta > \delta$ satisfy Assumptions 2.

By i) and ii), $V \cap W \neq \emptyset$ and parametrizes convolutional codes whose I/S/O realization (A, B, C, D) satisfy Assumptions 2.

In what follows we will give more explicit results regarding the parameters Θ, T, d_1, d_2.

A. The Controllability Matrix $\Phi_T(A, B)$

Recall that the controllability matrix $\Phi_T(A, B)$ is given by (8). Therefore, in order to compute the products $A^i B$, for A and B given in (6), the only important terms of A^i are those that belong to the last column. In fact, in case $T = \delta$, one can show that $\Phi_\delta(A, B)$ is of the form

$$\Phi_\delta(A, B) = -\frac{1}{\lambda_\delta^n} \begin{pmatrix} 0 & 0 & \ldots & 0 & (-1)^{\delta-1} \\ 0 & 0 & \ldots & (-1)^{\delta-2} & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & -1 & \ldots & * & * \\ 1 & * & \ldots & * & * \end{pmatrix} \tag{10}$$

Lemma III.2. *Let $T > \delta$. For $q \gg 0$ and generic values of the parameters $a_n, \lambda_0, \ldots, \lambda_\delta$ and b, the rows of $\Phi_T(A, B)$ form a parity check matrix of a block code whose minimum distance is $d_1 := d_{min} = \delta + 1$.*

Proof: From Equation (10) it follows that $\Phi_\delta(A, B)$ is invertible, so any $\delta + 1$ columns of $\Phi_T(A, B)$ are linearly dependent. Then, from [19, Theorem 4.12], it follows that $d_1 := d_{\min} \leq \delta + 1$. Let us check it is indeed an equality. For that it is enough to show that every minor of order δ of $\Phi_T(A, B)$ is non zero. Let us compute the minor associated to the columns in positions $i_1, \ldots, i_\delta$.

From Equation (1) and Equation (2), it follows that the characteristic polynomial of A, $P_A(z)$, is determined by

$$(-1)^\delta \lambda_\delta^{(n)} \det(A + z\,\mathrm{Id}) = s(p_n) = \sum_{i=0}^\delta \lambda_i (a_n z + b)^i$$

Let $\eta_1, \ldots, \eta_\delta$ be the roots of $P_A(z)$ in the algebraic closure of the base field $\mathbb{F}_q$. Since the parameters $a_n, \lambda_0, \ldots, \lambda_\delta, b$ are generic we can assume that there are no algebraic relations satisfied by the roots of $P_A(z)$. In particular the roots η_i are pair-wise different and A can be written as

$$A = \Gamma \cdot D \cdot \Gamma^{-1}$$

Γ being the Vandermonde matrix, whose (i, j) entry is given by η_j^{i-1}, and $D = \mathrm{diag}(-\eta_1, \ldots, -\eta_\delta)$. The i_jth column of $\Phi_T(A, B)$ can be written as $A^{i_j} B = \Gamma \cdot (D^{i_j} \cdot \Gamma^{-1} \cdot B)$ and, accordingly,

$$\det(A^{i_1} B|\ldots|A^{i_\delta} B) = (\lambda_\delta^{(n)})^\delta \det(D^{i_1}\Gamma^{-1} 1_\delta|\ldots|D^{i_\delta}\Gamma^{-1} 1_\delta)$$

where 1_δ denotes the column vector whose only non-zero entry is 1 in its δ-th position. Let $\gamma_1, \ldots, \gamma_\delta$ be the components of the column vector $\Gamma^{-1} 1_\delta$, then $D^{i_j}\Gamma^{-1} \cdot 1_\delta = (\eta_1^{i_j}\gamma_1, \ldots, \eta_\delta^{i_j}\gamma_\delta)^t$. Since Γ is Vandemonde, $\gamma_i \neq 0$ for each i, so $\det(A^{i_1} B|\ldots|A^{i_\delta} B) \neq 0$ if and only if

$$\begin{vmatrix} \eta_1^{i_1} & \eta_1^{i_2} & \ldots & \eta_1^{i_\delta} \\ \eta_2^{i_1} & \eta_2^{i_2} & \ldots & \eta_2^{i_\delta} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_\delta^{i_1} & \eta_\delta^{i_2} & \ldots & \eta_\delta^{i_\delta} \end{vmatrix} \neq 0$$

which is the generalized Vandermonde determinant associated to $\eta_1, \ldots, \eta_\delta$ and $i_1, \ldots, i_\delta$. Again, because of $a_n, \lambda_0, \ldots, \lambda_\delta, b$ are generic, we may assume that this condition is satisfied. $\square$

Remark III.3. *In the statement of the previous Lemma,* generic *means that there is a Zariski dense open subset $U \subseteq \mathbb{A}^{\delta+3}_{\mathbb{F}_q}$ such that if the parameters $(a_n, \lambda_0, \dots, \lambda_\delta, b)$ belong to U, then the conclusion of Lemma III.2 is true. Let us determine U explicitly. Let U_1 be the Zariski open subscheme of $\mathbb{A}^{\delta+3}_{\mathbb{F}_q}$ defined by $a_n \cdot \lambda_\delta \neq 0$. Note that $U_1 = \mathrm{Spec}\, R$ with $R := \mathbb{F}_q[a_n, a_n^{-1}, \lambda_0, \dots, \lambda_\delta, \lambda_\delta^{-1}, b]$. For any choice $\underline{i} \equiv 0 \leq i_1 < \dots < i_\delta \leq T$, let $U_{\underline{i}}$ be the Zariski open subscheme of U_1 where the generalized Vandermonde determinant associated to $\eta_1, \dots, \eta_\delta$ and $\underline{i}$ is non-zero. The desired open subscheme U is the intersection $\bigcap_{\underline{i}} U_{\underline{i}}$ is open since there are finitely many $\underline{i}$ as above. Finally, taking $q \gg 0$ this intersection is non-empty (thus, Zariski dense) and the conclusion of the lemma follows.*

B. The Observability Matrix $\Psi_\Theta(A, C)$

For this part, let us assume that $n > \delta$ and $\mathrm{rk}(C) = \delta$. This means that there is an invertible full size minor. Reordering the rows of $G(z)$ if it were necessary, we may assume that this full size minor is defined by the first δ rows of C, that is $C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ where C_1 is an invertible $\delta \times \delta$ matrix. Thus, we might consider the matrix $\Psi_\Theta(A, C) \cdot C_1^{-1}$, which satisfies that $\mathrm{Im}(\Psi_\Theta(A, C)) = \mathrm{Im}(\Psi_\Theta(A, C) \cdot C_1^{-1})$, so $\Psi_\Theta(A, C)$ and $\Psi_\Theta(A, C) \cdot C_1^{-1}$ can be seen as generator matrices of the same block code. Observe that

$$\Psi_\Theta(A, C) \cdot C_1^{-1} = \Big(\mathrm{Id}_{\delta\times\delta}, (C_2 C_1^{-1})^t, (CAC_1^{-1})^t, \dots, (CA^{\Theta-1}C_1^{-1})^t\Big)^t$$

Such a generator matrix admits a parity check matrix of the following form:

$$H := (H_1 | H_2) := \left(\begin{array}{c|c} -C_2C_1^{-1} & \\ -CAC_1^{-1} & \\ \vdots & \mathrm{Id}_{\alpha\times\alpha} \\ -CA^{\Theta-1}C_1^{-1} & \end{array}\right),$$

where $\alpha = (n-1)\Theta - \delta$, so $\mathrm{Im}(\Psi_\Theta(A, C)) = \mathrm{Ker}(H)$. Since we have assumed that A and C_1 are invertible, we know that $A^i C_1^{-1}$ is also invertible for every $i = 1, \dots, \Theta - 1$, and therefore

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} A^i C_1^{-1} = \begin{pmatrix} C_1 A^i C_1^{-1} \\ C_2 A^i C_1^{-1} \end{pmatrix}$$

is full column rank.

Lemma III.4. *Fix $n, \Theta > \delta$. No set of $\Theta - 1$ columns of the matrix H is linearly dependent. Therefore, the minimum distance of the block code generated by the parity check matrix H is $d_2 := d_{min} \geq \Theta$.*

Proof: If the $\Theta - 1$ columns belong to the right hand side, H_2, of the matrix H, then they are clearly linearly independent. Assume, therefore, that there are some columns, let us say t, among the fixed $\Theta - 1$ columns that belong to the left hand side. The matrix composed with these $\Theta - 1$ columns is full rank if and only if there are t rows not containing a 1 in the right hand side such that the corresponding minor of size $(\Theta - 1) \times (\Theta - 1)$,

$$\left|\begin{array}{ccc|cccccc} \Psi^{s_1}_{i_1} & \dots & \Psi^{s_1}_{i_t} & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \Psi^{s_l}_{i_1} & \dots & \Psi^{s_l}_{i_t} & 0 & \dots & 1 & 0 & \dots & 0 \\ \Psi^{j_1}_{i_1} & \dots & \Psi^{j_1}_{i_t} & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \Psi^{j_{\Theta-t-l-1}}_{i_1} & \dots & \Psi^{j_{\Theta-t-l-1}}_{i_t} & 0 & \dots & 0 & 0 & \dots & 1 \\ \hline \Psi^{k_1}_{i_1} & \dots & \Psi^{k_1}_{i_t} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \Psi^{k_t}_{i_1} & \dots & \Psi^{k_t}_{i_t} & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right|,$$

is non zero which, in turn, is true if and only if the minor of size $t \times t$

$$\begin{vmatrix} \Psi^{k_1}_{i_1} & \dots & \Psi^{k_1}_{i_t} \\ \vdots & & \vdots \\ \Psi^{k_t}_{i_1} & \dots & \Psi^{k_t}_{i_t} \end{vmatrix}$$

is non zero. The number of rows not containing a 1 in the right hand side is equal to $\Theta(n-1) - \delta - (\Theta - 1 - t) = \Theta(n-1) - \delta - \Theta + 1 + t$, while the number of rows that remain when we remove one row in each block $CA^iC_1^{-1}$ of the left hand side is equal to $\Theta(n-1) - \delta - (\Theta - 1) = \Theta(n-1) - \delta - \Theta + 1$. Therefore, we can find all the rows of one of the blocks $CA^iC_1^{-1}$ in the left hand side. Since each of these blocks are full (column) rank, for any t columns there are t rows such that the corresponding minor is non-zero, so we are done. □

IV. Maximizing the Performance

From Theorem III.1 we know that this decoding algorithm allows to correct λ errors in a time-window of length T. Thus, in order to try to make use of the full error-correcting capacity of the code, we aim at maximizing the ratio λ/T.

Under Assumptions 2, and by Lemma III.2 and Lemma III.4, we have $T > \Theta > \delta$, $d_1 = \delta + 1$ and $d_2 \geq \Theta$. In particular, this implies that

$$\lambda = \min\left\{\left\lfloor\frac{\delta}{2}\right\rfloor, \left\lfloor\frac{T}{2\Theta}\right\rfloor\right\}.$$

Since δ is fixed, the ratio λ/T attains its maximal value when $T/2\Theta \geq \delta/2$ with T minimal. Since the minimal value of Θ is $\Theta_{min} = \delta + 1$, it follows that $T = \delta\Theta_{min}$. Hence, we have proved the following.

Theorem IV.1. *Consider the convolutional code* (1) *satisfying Assumptions 1 and those given in Lemma III.2. Then, using Rosenthal's algorithm as a decoding algorithm, the code is capable to correct $\left\lfloor\frac{\delta}{2}\right\rfloor$ errors in a time-window of length $T = \delta(\delta + 1)$.*

It is convenient to compare the previous result with the theoretical correcting capacity of the code. Recall that a convolutional code with free distance d_{free} can correct any error pattern of weight less than $\left\lfloor \frac{d_{free}}{2} \right\rfloor$ if the time-window is sufficiently large (see [20]). Thus, we can compare both bounds, the effective one versus the theoretical one, for the case of the code (1). We will assume that the code is MDS, Assumptions 1 are fulfilled and the time-window T is large enough. Under these hypothesis $d_{free} = n(\delta + 1)$ since the codes are MDS, so the error-correcting capability of the code allows to correct

$$\left\lfloor \frac{n(\delta+1)}{2} \right\rfloor$$

errors in a time-window of length T while the decoding algorithm allows to correct

$$\left\lfloor \frac{\delta}{2} \right\rfloor$$

errors in the same time-window, which shows that the efficiency of the algorithm decreases as the length n increases.

V. An Example

In order to illustrate different properties of this type of codes, let us produce an example of a family of convolutional codes satisfying Assumptions 1, Assumptions 2 and the conditions given in Theorem III.1 and Lemma III.2. Inspired by the proof of Lemma III.2, we proceed constructively as follows.

Consider the base field $\mathbb{F}_{16}$ and let α be a generator of $\mathbb{F}_{16}^*$ such that $\alpha^4+\alpha^3+1 = 0$. Set $\delta = 3$ and $\eta_i = \alpha^i$ for $i = 1, 2, 3$. We will follow the notations of Lemma III.2. Consider 3×3 matrices defined by $D = \mathrm{diag}(\alpha, \alpha^2, \alpha^3)$ and

$$\Gamma = \begin{pmatrix} 1 & 1 & 1 \\ \alpha & \alpha^2 & \alpha^3 \\ \alpha^2 & \alpha^4 & \alpha^6 \end{pmatrix}, \ \Gamma^{-1} = \begin{pmatrix} \alpha^{12} & \alpha^6 & \alpha^7 \\ \alpha^7 & \alpha^{13} & \alpha^3 \\ \alpha^9 & \alpha^4 & \alpha^6 \end{pmatrix}$$

According to Theorem IV.1, set $T = \delta(\delta+1) = 12$. The parity check matrix $\Phi_T(A, B)$ acquires the following form

$$\Phi_T(A, B) = \Delta \cdot \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{11} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{22} \\ 1 & \alpha^3 & \alpha^6 & \dots & \alpha^{33} \end{pmatrix} \tag{11}$$

where Δ is a 3×3 invertible matrix. One checks that all maximal minors of the 3×12 rectangular matrix on the right hand side are non-zero.

Recalling that $A = \Gamma D \Gamma^{-1}$, setting $\lambda_\delta = 1$ and expanding the relation

$$(-1)^\delta \lambda_\delta^{(n)} \det(A + z\,\mathrm{Id}) = s(p_n) = \sum_{i=0}^{\delta} \lambda_i (a_n z + b)^i$$

we obtain the following system

$$\begin{aligned} \alpha^6 &= \lambda_0 + \lambda_1\, b + \lambda_2\, b^2 + b^3 \\ \alpha^{10} &= \lambda_1\, a_n + a_n b^2 \\ \alpha^8 &= \lambda_2\, a_n^2 + a_n^2\, b \\ 1 &= a_n^3 \end{aligned}$$

whose general solution for $a_n = 1$ is given by

$$\begin{aligned} \lambda_2 &= \alpha^8 + b \\ \lambda_1 &= \alpha^{10} + b^2 \\ \lambda_0 &= \alpha^6 + \alpha^{10} b + \alpha^8\, b^2 + b^3 . \end{aligned}$$

Summing up, the convolutional code is defined by the encoder $G(z) = (s(p_1), \dots, s(p_n))^t$ as in Equation (1) where

- $p_i := a_i z + b$ for $i = 1, \dots, n$ with $b \neq 0$, $a_n = 1$; and $a_1, \dots, a_n$ are pairwise different and non-zero;
- $s(t) = (\alpha^6+\alpha^{10}b+\alpha^8\, b^2+b^3)+(\alpha^{10}b^2)t+(\alpha^8+b)t^2+t^3$;

Then, we have that $s_\lambda^{(0)}(b) = \alpha^6, s_\lambda^{(1)}(b) = \alpha^{10}, s_\lambda^{(2)}(b) = \alpha^8, s_\lambda^{(3)}(b) = 1$, so all of them are non zero. By construction, the I/S/O representation of the convolutional code defined by the encoder $G(z)$ fulfills Assumptions 1, Assumptions 2 and the conditions given in Theorem III.1 and Lemma III.2. Furthermore, for decoding this convolutional code using Rosenthal's procedure, one uses the linear code with parity check matrix $\Phi_T(A, B)$ on received words. Having in mind Equation (11), it is clear that it is equivalent to consider the parity check matrix

$$\begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{11} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{22} \\ 1 & \alpha^3 & \alpha^6 & \dots & \alpha^{33} \end{pmatrix},$$

which is a shortening of a BCH(4,2) code. Morally, Rosenthal's algorithm and its complexity coincide with those of BCH codes.

VI. Conclusions

In this work, we have shown that the performance of Rosenthal's decoding algorithm applied to certain family of MDS convolutional codes is not far away from their theoretical error-correcting capability. However, it is also known that this algorithm can make full use of the theoretical error-correcting capability of some other convolutional codes (see [18, §6]). This means that the performance of the algorithm depends strongly on the algebraic properties of the I/S/O representation of the code to which we are applying the algorithm, so a deeper understanding of this dependency is necessary.

We must persevere in the study of convolutional code decoding algorithms based on block codes, as Rosenthal initiated, with the aim of achieving new algorithms whose corrective capacity is very close to the theoretical error-correcting capability of the code. In particular, the example that has been worked out support the hypothesis that the complexity of this decoding algorithm for convolutional codes might be similar to that of the decoding process of a BCH linear code.

Acknowledgment

The authors would like to thank the reviewers for their careful revision and their comments, that helped to improve the final version of this paper.

At the beginning of this year, José María Muñoz Porras, second author of this article, passed away. We would like to remember and recognize how much we learned from him, both from the human point of view and from the mathematical point of view, and make this article a small tribute to his memory. As a brilliant mathematician and a very warm person, we will always remember him.

REFERENCES

[1] P. Elias, "Coding for noisy channels," *IRE Nat. Conv. Rec.*, vol. 3, no. 4, pp. 37–46, 1955.

[2] G. Forney, "Convolutional codes I: Algebraic structure," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 6, pp. 720–738, Nov. 1970.

[3] R. J. McEliece, "The algebraic theory of convolutional codes," in *Handbook of Coding Theory*, vol. 1, V. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998, pp. 1065–1138.

[4] J. Rosenthal, "Connections between linear systems and convolutional codes," in *Codes, Systems, and Graphical Models*, B. Marcus and J. Rosenthal, Eds. New York, NY, USA: Springer, 2001, pp. 39–66.

[5] J. Rosenthal, J. M. Schumacher, and E. V. York, "On behaviors and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1881–1891, Nov. 1996.

[6] E. V. York, "Algebraic description and construction of error correcting codes: A linear systems point of view," Ph.D. dissertation, Dept. Math., Univ. Notredam, Notre Dame, IN, USA, 1997.

[7] V. Lomadze, "Convolutional codes and coherent sheaves," *Appl. Algebra Eng., Commun. Comput.*, vol. 12, no. 4, pp. 273–326, 2000.

[8] J. A. D. Pérez, J. M. M. Porras, and G. S. Sotelo, "Convolutional codes of Goppa type," *Appl. Algebra Eng., Commun. Comput.*, vol. 15, pp. 51–61, Jun. 2004.

[9] J. I. I. Curto, A. L. M. Castañeda, J. M. M. Porras, and G. S. Sotelo, "Every convolutional code is a Goppa code," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6628–6641, Oct. 2013.

[10] J. M. M. Porras, J. A. D. Perez, J. I. I. Curto, and G. S. Sotelo, "Convolutional Goppa codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 340–344, Jan. 2006.

[11] F. J. Plaza-Martín, J. I. Iglesias-Curto, and G. Serrano-Sotelo, "On the construction of 1-dimensional MDS convolutional Goppa codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4615–4625, Jul. 2013.

[12] J. Justesen and T. Hoholdt, *A Course in Error-Correcting Codes*. Zürich, Switzerland: European Mathematical Society, 2004.

[13] J. Rosenthal and R. Smarandache, "Maximum distance separable convolutional codes," *Appl. Algebra Eng., Commun. Comput.*, vol. 10, no. 1, pp. 15–32, 1999.

[14] G. D. Forney, Jr., "Convolutional codes II. Maximum-likelihood decoding," *Inf. Control*, vol. 25, pp. 222–266, Jul. 1974.

[15] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 260–269, Apr. 1967.

[16] M. I. G. Planas, E. M. Souidi, and L. E. Um, "Convolutional codes under linear systems point of view. Analysis of output-controllability," *WEAS Trans. Math.*, vol. 11, no. 4, pp. 324–333, 2012.

[17] H. Gluesing-Luerssen, U. Helmke, and J. I. I. Curto, "Algebraic decoding for doubly cyclic convolutional codes," *Adv. Math. Commun.*, vol. 4, no. 1, pp. 260–269, 2010.

[18] J. Rosenthal, "An algebraic decoding algorithm for convolutional codes," in *Dynamical Systems, Control, Coding, Computer Vision*, G. Picci and D. S. Gilliam, Eds. Basel, Switzerland: Birkhäuser Basel, 1999, pp. 343–360.

[19] O. Pretzel, *Error-Correcting Codes and Finite Fields*. Oxford, U.K.: Oxford Univ. Press, 1992.

[20] D. J. Costello, "A construction technique for random-error-correcting convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 5, pp. 631–636, Sep. 1969.

Ángel Luis Muñoz Castañeda received his Ph.D. in Mathematics from the Freie Universität Berlin. He is currently assistant professor at the University of León and belongs to the Research Institute of Applied Sciences in Cybersecurity of the same university. His research interests concern moduli problems in algebraic geometry and applications of algebraic geometry in coding and systems theory.

José M. Muñoz-Porras was born in Salamanca, Spain, in 1956. He received the Ph.D. in Mathematics from University of Salamanca in 1984. He was Professor of Algebra at University of Salamanca. His research interests were algebraic geometry and applications to physics, arithmetic and coding theory.

Francisco J. Plaza-Martín was born in Jan, Spain, in 1970. He received the Ph. D. in Mathematics from University of Salamanca in 1997 where he is currently Professor of Geometry and Topology. His research topics are algebraic geometry and its applications to physics and information theory.