This is a previous version of the article published in Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas. 2020, 114:38. doi:10.1007/s13398-019-00744-y

Noname manuscript No.
(will be inserted by the editor)

Block Toeplitz matrices for burst-correcting convolutional codes

Joan-Josep Climent · Diego Napp · Verónica Requena

Received: date / Accepted: date

Abstract In this paper we study a problem in the area of coding theory. In particular, we focus on a class of error-correcting codes called convolutional codes. We characterize convolutional codes that can correct bursts of erasures with the lowest possible delay. This characterization is given in terms of a block Toeplitz matrix with entries in a finite field that is built upon a given generator matrix of the convolutional code. This result allows us to provide a concrete construction of a generator matrix of a convolutional code with entries being only zeros or ones that can recover bursts of erasures with low delay. This construction admits a very simple decoding algorithm and, therefore, simplifies the existing schemes proposed recently in the literature.

Keywords Linear algebra · (block) Toeplitz matrices · Error-correcting codes · Convolutional codes · Finite fields

Mathematics Subject Classification (2010) 94B10 · 11T71

1 Introduction

Coding theory has emerged out of the need for better communication and has rapidly developed as a mathematical theory in strong relationship with algebra and combinatorics. Error correction codes are used in practical applications constantly and have been the foundation of the revolutionary growth in digital communications and storage.

A very interesting class of error correcting codes is the class of convolutional codes [15]. Convolutional codes offer an approach to error control coding substantially different from block codes as they encode the entire data stream into

RACSAM˙2018˙Climent˙Napp˙Requena˙Revised.tex

All the authors are at
Departament de Matemàtiques,
Universitat d'Alacant, Spain
jcliment@ua.es, diego.napp@ua.es, vrequena@ua.es

a single codeword. Mathematically, convolutional codes are defined as $\mathbb{F}((D))$-subspaces of $\mathbb{F}((D))^n$, where $\mathbb{F}((D))$ is the field of Laurent polynomials with coefficients in a finite field $\mathbb{F}$. The design and construction of convolutional codes boil down to the construction of block Toeplitz matrices with entries in a finite field, typically, the binary field or fields with characteristic 2. The designed properties of this matrix will depend on the desired features of the associated convolutional code. For example, convolutional codes with optimal (column) distances require the construction of superregular matrices [8, 13, 25]. Currently, there are only two general construction of superregular matrices and both require huge finite fields [1, 2, 8]. Some computer search algorithms to seek superregular matrices have been recently proposed [6, 11, 12]. However, many problems in the area remain open, as for instance, the minimum field size needed for the existence of a superregular matrix of a given size [14], although some conjectures have been proposed in the literature, see for instance [8].

Recently, there has been a great interest in the theory of codes for streaming applications where a bitstream of data is transmitted sequentially in real-time under strict latency constraints [3–5, 7, 16, 17, 19]. This is due to the fact that in many multimedia applications, such as real-time video conference, the transmission must be performed sequentially and with minimal perceptible delay at the destination. As it has been shown in the literature, the problem of designing optimal error correcting codes that admit low-delay decoders have not been throughly treated before and have many unique differences from classical error correction designs. Classical error correcting codes require interleaving and long decoder delays which is not acceptable in many real-time multimedia communication applications.

Typically, streaming applications operate on packet networks and recent investigations have shown that packet losses occur in bursts of erasures rather than errors [17]. Moreover, it is known that the performance degradation due to burst losses is more relevant than random isolated losses. In the seminal work [19] the authors analysed this channel and established a fundamental trade-off bound between decoding delay and the burst length for a given rate. Moreover, they presented an explicit class of encoders, called Short codes, which achieve this bound with equality and have the shortest possible decoding delay required to correct bursts of a given length. Later in [5] this construction was simplified and a layer was added in order to deal also with isolated erasures. These codes were called Midas codes. These constructions were based on MDS Reed-Solomon block codes and m-MDS convolutional codes, respectively, and therefore the underlying field sizes are required to be relatively large.

In this work, we shall focus on convolutional codes over the burst erasure channel (formally defined below). We study and characterize the type of encoders that are optimal with respect to the rate, decoding delay and burst length. Moreover, we also present a new and novel class of encoders defined over the binary field and therefore it is also optimal with respect to the field size. As a consequence of this, the decoding is straightforward.

In contrast to previous contributions, we use the polynomial generator matrix approach to represent convolutional codes which can facilitate the analysis when considering the degree of the code or minimal input/state/output representations [9, 23].

2 Preliminaries

Let $\mathbb{F} = GF(q)$ be a finite field of size q, $\mathbb{F}((D))$ be the field of formal Laurent polynomials, $\mathbb{F}(D)$ be the field of rational polynomials and $\mathbb{F}[D]$ be the ring of polynomials all of them with coefficients in $\mathbb{F}$.

2.1 Convolutional Codes

Definition 1 (Definition 2.3 of [20]) A **convolutional code** $\mathcal{C}$ of rate k/n is an $\mathbb{F}((D))$-subspace of $\mathbb{F}((D))^n$ of dimension k given by a rational **encoder matrix** $G(D) \in \mathbb{F}(D)^{k\times n}$, i.e.

$$\mathcal{C} = \mathrm{im}_{\mathbb{F}((D))} G(D) = \left\{ \boldsymbol{v}(D) = \boldsymbol{u}(D)G(D) \mid \boldsymbol{u}(D) \in \mathbb{F}^k((D)) \right\}.$$

Assume that $G(D) = \sum_{i=0}^{m} G_i D^i$ is polynomial. Then, m is called the **memory** of $G(D)$ and the **j-th associated sliding matrix** of $G(D)$ is

$$G_j^c = \begin{pmatrix} G_0 & G_1 & \cdots & G_j \\ & G_0 & \cdots & G_{j-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{pmatrix}, \text{ for } j \in \mathbb{N},$$

with $G_j = O$ when $j > m$, and analogously, the **infinite sliding matrix** of $G(D)$ is

$$G_\infty^c = \begin{pmatrix} G_0 & G_1 & \cdots & G_m & & & & \\ & G_0 & \cdots & G_{m-1} & G_m & & & \\ & & \ddots & \vdots & \vdots & \ddots & & \\ & & & G_0 & G_1 & \cdots & G_m & \\ & & & & \ddots & \ddots & & \ddots \end{pmatrix}.$$

The distance is an invariant of the code that serves as an indicator of the perfomance of the code [10, 21, 24]. In the context of convolutional codes there are two fundamental types of distances, the free distance and the column distance. Column distance is associated to the error-correcting capabilities of the convolutional code per time interval [22, 26]. Even though codes with optimal free distance and optimal column distance were used in [5] for streaming applications, these notions will not play an important role in the context of burst erasure channels. Instead, the notion of column span is more relevant for correction of bursts of erasures. In [19] this notion was introduced as the proper indicator of the error-burst-correcting capabilities of an encoder.

Definition 2 The **column span** of G_j^c is defined as

$$CS(j) = \min \left\{ \operatorname{span}(\boldsymbol{u}G_j^c) \mid \boldsymbol{u} = (\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_j),\ \boldsymbol{u}_0 \neq \boldsymbol{0} \right\}$$

where the span of a vector is $j - i + 1$, where i and j are the first and the last nonzero entries of such a vector, respectively.

Assume that a burst of maximum length L occurs within a window of length $W + 1$. From Lemma 1.1 of [7] it follows that, it can be corrected if and only if $CS(W) > L$.

2.2 Delay in Burst Erasure Channels

We follow previous approaches and regard the symbols $\boldsymbol{v}_i$ as packets and consider that losses occur on a packet level [5, 18, 26]. In the transmission of a stream of information at each time instant i we receive a symbol packet $\boldsymbol{v}_i \in \mathbb{F}^n$. Over an erasure channel the packets either arrive correctly or otherwise are regarded as an erasure. In burst erasure channels these erasures tend to occur in bursts. The goal of this work is to study how to construct polynomial encoders tailor-made to correct burst of erasures. Suppose that the information has been correctly received up to an instant i and a burst of length L is received at time instant i, i.e., one or more packets are lost from the sequence $(\boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \dots, \boldsymbol{v}_{i+L-1})$. Then, we say that the decoding delay is T if the encoder can reconstruct each source packet with a delay of T source packets, i.e., we can recover $\boldsymbol{u}_{i+j}$ (for $j \in \{0, 1, \dots, L-1\}$) once $\boldsymbol{v}_{i+L}, \boldsymbol{v}_{i+L+1}, \dots, \boldsymbol{v}_{i+j+T}$ are received. In [19] the following result on the trade-off between delay and redundancy was derived.

Theorem 1 (Theorem 1 of [19]) *If a rate R encoder enables correction of all burst of erasures of length L with decoding delay at most T, then,*

$$\frac{T}{L} \geq \max \left\{ 1, \frac{R}{1-R} \right\}. \tag{1}$$

A generalization of this result was later presented in [5] taking into account not only burst of erasures but isolated erasures as well. It is also worth mentioning the upperbounds given in [3, 7] on the maximum correctable burst length in terms of the encoder parameters n, k and m. In this work, we shall focus on low delay decoding under burst of erasures, and so consider only the inequality given in (1) without taking into consideration the memory of the encoder or isolated losses. A natural follow-up work will be to incorporate these parameters in the bound (1) and derive optimal codes with respect to this bound.

3 Systematic Encoders for Burst Erasure Correction with Low Delay

Convolutional codes with large column distance are very appealing for sequential decoding and have excellent error correction capabilities, but they require, in general, huge finite fields [2] and long delays. Even though these type of codes have been proposed for applications that consider erasure channels, see [26], they do not generally achieve the best trade-off between delay, redundancy, field size and burst correction.

Consider a systematic encoder $G(D) = [I_k \ \ \widehat{G}(D)]$, $\widehat{G}(D) = \sum_{j=0}^{m} \widehat{G}_j D^j$ and a submatrix of the sliding matrix $\widehat{G}_j^c$, with $j \geq L+T$

$$\widehat{G}_{L+T}^{trunc} = \begin{pmatrix} \widehat{G}_L & \widehat{G}_{L+1} & \cdots & \widehat{G}_T \\ \widehat{G}_{L-1} & \widehat{G}_L & \cdots & \widehat{G}_{T-1} \\ \vdots & \vdots & \cdots & \vdots \\ \widehat{G}_1 & \widehat{G}_2 & \cdots & \widehat{G}_{T-L+1} \end{pmatrix}, \tag{2}$$

of size $Lk \times (T-L+1)(n-k)$.

This matrix will play an important role in the construction of good burst correction convolutional codes with low delay. This fact will be evident in the next example. Before, we present a lemma which will need for our results. From now on, we denote by O the zero matrix of the appropiate size.

Lemma 1 *Assume that* $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, *where* A_1 *is a* $k \times t$ *matrix, with* $k \leq t$, $\boldsymbol{b}$ *a vector of length* t *and, consider the linear system*

$$(\boldsymbol{x}_0, \boldsymbol{x}_1)A = \boldsymbol{b} \tag{3}$$

where $\boldsymbol{x}_0$ *is a vector of length* k. *The following statements are equivalents*

1. $\boldsymbol{x}_0$ *is univocally determined by the system (3);*
2. *There exists an invertible matrix* P *of size* $t \times t$ *such that*

$$AP = \left(\begin{array}{c|c} I_k & O \\ \hline O & M \end{array}\right); \tag{4}$$

3. A_1 *has full row rank and,*

$$\text{rowspan}(I_k \ O) \cap \text{rowspan}(N \ M) = \{\mathbf{0}\}. \tag{5}$$

Proof (1 $\Leftrightarrow$ 2) This part follows from Gauss-Jordan elimination over the columns of A.

(2 $\Rightarrow$ 3) It is enough to observe that the invertible matrix P does not affect to the linearly independence of the rows of A. Thus, if one can transform A into a matrix of the form (4) the statement (3) readily follows.

(3 ⇒ 2) From the first part of condition 3, there exists an invertible $t \times t$ matrix Q such that

$$AQ = \begin{pmatrix} I_k & O \\ N & M \end{pmatrix}.$$

Now, from (5) it follows that

$$\operatorname{rowspan}(I_k \; O) \cap \operatorname{rowspan}(N \; M) = \{\mathbf{0}\}.$$

Assume that $\boldsymbol{v} \in \operatorname{rowspan}(N \; M)$ with $\boldsymbol{v} \neq \mathbf{0}$, then $\boldsymbol{v} \notin \operatorname{rowspan}(I_k \; O)$. We have that $\boldsymbol{v} = (\boldsymbol{u}M \; \boldsymbol{u}N)$ for some $\boldsymbol{u} \neq \mathbf{0}$. If $\boldsymbol{u}M = \mathbf{0}$, then $\boldsymbol{u}N = \mathbf{0}$; that is, the same rows that are linearly independent in M are also linearly independent in N; then, $\operatorname{rk}(N \; M) = \operatorname{rk}(M)$. In other words, there exists a $(t-k) \times k$ matrix R such that $N = MR$, so

$$\begin{pmatrix} I_k & O \\ N & M \end{pmatrix} \begin{pmatrix} I_k & O \\ -R & I_{t-k} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & M \end{pmatrix}.$$

Finally, the condition 2 follows taking

$$P = Q \begin{pmatrix} I_k & O \\ -R & I_{t-k} \end{pmatrix}.$$

□

Next, we present an illustrative example where we explain how to recover the lost packets in a burst of erasures.

Example 1 Assume that we want to correct a burst of erasures of maximum length $L = 3$ with a design decoding delay $T = 6$. To this end, we build a convolutional code with parameters $n = 6$ and $k = 4$. Note that with these parameters, we have that $R = \frac{2}{3}$ and the bound given in Theorem 1 is satisfied with equality. Suppose without lost of generality that the transmission starts at time instant zero. Then, we have that

$$G(D) = \left(I_k \; \widehat{G}(D)\right) = (I_k \; \widehat{G}_0) + (O \; \widehat{G}_1)D + (O \; \widehat{G}_2)D^2 + \cdots$$

and $\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D)$ with

$$\boldsymbol{u}(D) = \boldsymbol{u}_0 + \boldsymbol{u}_1 D + \boldsymbol{u}_2 D^2 + \cdots \quad \text{and} \quad \boldsymbol{v}(D) = \boldsymbol{v}_0 + \boldsymbol{v}_1 D + \boldsymbol{v}_2 D^2 + \cdots$$

and then, the received sequence is

$$\begin{aligned} \boldsymbol{v} = \Big(&\boldsymbol{u}_0 \; \boldsymbol{u}_0\widehat{G}_0, \boldsymbol{u}_1 \; \boldsymbol{u}_1\widehat{G}_0 + \boldsymbol{u}_0\widehat{G}_1, \\ &\boldsymbol{u}_2 \; \boldsymbol{u}_2\widehat{G}_0 + \boldsymbol{u}_1\widehat{G}_1 + \boldsymbol{u}_0\widehat{G}_2, \\ &\boldsymbol{u}_3 \; \boldsymbol{u}_3\widehat{G}_0 + \boldsymbol{u}_2\widehat{G}_1 + \boldsymbol{u}_1\widehat{G}_2 + \boldsymbol{u}_0\widehat{G}_3, \ldots\Big). \end{aligned}$$

Equivalently, using constant matrices one can write

$$(\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2, \cdots) = (\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2, \cdots) \left(\begin{array}{c|c|c|c} I_k \;\; \widehat{G}_0 & O \;\; \widehat{G}_1 & O \;\; \widehat{G}_2 & \cdots \\ \hline O & I_k \;\; \widehat{G}_0 & O \;\; \widehat{G}_1 & \cdots \\ \hline O & O & I_k \;\; \widehat{G}_0 & \cdots \\ \hline \vdots & \vdots & \vdots & \ddots \end{array}\right), \quad (6)$$

where we can expressed the vector $\boldsymbol{v}_i = (\boldsymbol{u}_i \; \widehat{\boldsymbol{v}}_i)$, with $\widehat{\boldsymbol{v}}_i = \sum_{j=0}^{i} \boldsymbol{u}_{i-j}\widehat{G}_j$. The identity matrices of the block diagonal of the matrix in (6) allow us to obtain the packets $\boldsymbol{u}_i$ directly without decoding, if the corresponding $\boldsymbol{v}_i$ has been received correctly.

Suppose that we have received a burst of erasures in the first three packets of our message

$$\boldsymbol{v} = (\bigstar, \bigstar, \bigstar, \boldsymbol{v}_3, \boldsymbol{v}_4, \boldsymbol{v}_5, \boldsymbol{v}_6, \boldsymbol{v}_7, \boldsymbol{v}_8),$$

that is, the burst of erasure starts at time instant 0. We need to recover $\boldsymbol{u}_0$ at time instant 6, because the decoding delay is $T = 6$; therefore, $\boldsymbol{u}_1$ will be recover at time instant 7, and $\boldsymbol{u}_2$ at time instant 8. Hence, to recover the packets $\boldsymbol{u}_0, \boldsymbol{u}_1$ and $\boldsymbol{u}_2$, we regard them as unknowns $\boldsymbol{x}_0, \boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively. Thus, it follows from the system (6) that to recover $\boldsymbol{u}_0$ we must solve the following linear system of equations

$$(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)\widehat{G}_9^{trunc} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) \begin{pmatrix} \widehat{G}_3 & \widehat{G}_4 & \widehat{G}_5 & \widehat{G}_6 \\ \widehat{G}_2 & \widehat{G}_3 & \widehat{G}_4 & \widehat{G}_5 \\ \widehat{G}_1 & \widehat{G}_2 & \widehat{G}_3 & \widehat{G}_4 \end{pmatrix} = (\boldsymbol{b}_0^{(0)}, \boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \boldsymbol{b}_3^{(0)}),$$

where,

$$\begin{aligned}\left(\boldsymbol{b}_0^{(0)}, \boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \boldsymbol{b}_3^{(0)}\right) = \Big(&\widehat{\boldsymbol{v}}_3 - \boldsymbol{u}_3\widehat{G}_0, \widehat{\boldsymbol{v}}_4 - \boldsymbol{u}_4\widehat{G}_0 - \boldsymbol{u}_3\widehat{G}_1, \\ &\widehat{\boldsymbol{v}}_5 - \boldsymbol{u}_5\widehat{G}_0 - \boldsymbol{u}_4\widehat{G}_1 - \boldsymbol{u}_3\widehat{G}_2, \\ &\widehat{\boldsymbol{v}}_6 - \boldsymbol{u}_6\widehat{G}_0 - \boldsymbol{u}_5\widehat{G}_1 - \boldsymbol{u}_4\widehat{G}_2 - \boldsymbol{u}_3\widehat{G}_3\Big).\end{aligned}$$

By Lemma 1, we will be able to obtain $\boldsymbol{x}_0$ univocally and, therefore recover $\boldsymbol{u}_0$, if the previous system satisfies the condition 3 of Lemma 1, that is, if the submatrix $\left(\widehat{G}_3 \; \widehat{G}_4 \; \widehat{G}_5 \; \widehat{G}_6\right)$ of $\widehat{G}_9^{trunc}$ has full row rank and

$$\text{rowspan}\left(\widehat{G}_3 \; \widehat{G}_4 \; \widehat{G}_5 \; \widehat{G}_6\right) \cap \text{rowspan}\begin{pmatrix} \widehat{G}_2 & \widehat{G}_3 & \widehat{G}_4 & \widehat{G}_5 \\ \widehat{G}_1 & \widehat{G}_2 & \widehat{G}_3 & \widehat{G}_4 \end{pmatrix} = \{\mathbf{0}\}.$$

Analogously, to recover $\boldsymbol{u}_1$ we must solve the system

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u}_3)\widehat{G}_9^{trunc} = \left(\boldsymbol{b}_1^{(1)}, \boldsymbol{b}_2^{(1)}, \boldsymbol{b}_3^{(1)}, \boldsymbol{b}_4^{(1)}\right)$$

where

$$(\boldsymbol{b}_1^{(1)}, \boldsymbol{b}_2^{(1)}, \boldsymbol{b}_3^{(1)}, \boldsymbol{b}_4^{(1)}) = \Big(\widehat{\boldsymbol{v}}_4 - \boldsymbol{u}_4\widehat{G}_0 - \boldsymbol{u}_0\widehat{G}_4, \widehat{\boldsymbol{v}}_5 - \boldsymbol{u}_5\widehat{G}_0 - \boldsymbol{u}_4\widehat{G}_1 - \boldsymbol{u}_0\widehat{G}_5,$$
$$\widehat{\boldsymbol{v}}_6 - \boldsymbol{u}_6\widehat{G}_0 - \boldsymbol{u}_5\widehat{G}_1 - \boldsymbol{u}_4\widehat{G}_2 - \boldsymbol{u}_0\widehat{G}_6,$$
$$\widehat{\boldsymbol{v}}_7 - \boldsymbol{u}_7\widehat{G}_0 - \boldsymbol{u}_6\widehat{G}_1 - \boldsymbol{u}_5\widehat{G}_2 - \boldsymbol{u}_4\widehat{G}_3 - \boldsymbol{u}_0\widehat{G}_7\Big),$$

can be computed from the correctly received data. We can rewrite the previous system simplifying it, since that the information vector $\boldsymbol{u}_3$ is known, so we only need to solve the following system

$$(\boldsymbol{x}_1, \boldsymbol{x}_2)\begin{pmatrix} \widehat{G}_3 & \widehat{G}_4 & \widehat{G}_5 & \widehat{G}_6 \\ \widehat{G}_2 & \widehat{G}_3 & \widehat{G}_4 & \widehat{G}_5 \end{pmatrix}$$
$$= \Big(\boldsymbol{b}_1^{(1)} - \boldsymbol{u}_3\widehat{G}_1, \boldsymbol{b}_2^{(1)} - \boldsymbol{u}_3\widehat{G}_2, \boldsymbol{b}_3^{(1)} - \boldsymbol{u}_3\widehat{G}_3, \boldsymbol{b}_4^{(1)} - \boldsymbol{u}_3\widehat{G}_4\Big)$$

obtaining the solution $\boldsymbol{x}_1 = \boldsymbol{u}_1$.

Finally, to recover $\boldsymbol{u}_2$, we have the system

$$(\boldsymbol{x}_2, \boldsymbol{u}_3, \boldsymbol{u}_4)\widehat{G}_9^{trunc} = \Big(\boldsymbol{b}_2^{(2)}, \boldsymbol{b}_3^{(2)}, \boldsymbol{b}_4^{(2)}, \boldsymbol{b}_5^{(2)}\Big)$$

where

$$(\boldsymbol{b}_2^{(2)}, \boldsymbol{b}_3^{(2)}, \boldsymbol{b}_4^{(2)}, \boldsymbol{b}_5^{(2)}) = \Big(\widehat{\boldsymbol{v}}_5 - \boldsymbol{u}_5\widehat{G}_0 - \boldsymbol{u}_0\widehat{G}_5 - \boldsymbol{u}_1\widehat{G}_4,$$
$$\widehat{\boldsymbol{v}}_6 - \boldsymbol{u}_6\widehat{G}_0 - \boldsymbol{u}_5\widehat{G}_1 - \boldsymbol{u}_0\widehat{G}_6 - \boldsymbol{u}_1\widehat{G}_5,$$
$$\widehat{\boldsymbol{v}}_7 - \boldsymbol{u}_7\widehat{G}_0 - \boldsymbol{u}_6\widehat{G}_1 - \boldsymbol{u}_5\widehat{G}_2 - \boldsymbol{u}_0\widehat{G}_7 - \boldsymbol{u}_1\widehat{G}_6,$$
$$\widehat{\boldsymbol{v}}_8 - \boldsymbol{u}_8\widehat{G}_0 - \boldsymbol{u}_7\widehat{G}_1 - \boldsymbol{u}_6\widehat{G}_2 - \boldsymbol{u}_5\widehat{G}_3 - \boldsymbol{u}_0\widehat{G}_8 - \boldsymbol{u}_1\widehat{G}_7\Big).$$

Reasoning as in the previous case, this system can be rewritten by

$$\boldsymbol{x}_2\left(\widehat{G}_3 \ \widehat{G}_4 \ \widehat{G}_5 \ \widehat{G}_6\right) = \Big(\boldsymbol{b}_2^{(2)} - \boldsymbol{u}_4\widehat{G}_1, \boldsymbol{b}_3^{(2)} - \boldsymbol{u}_4\widehat{G}_2, \boldsymbol{b}_4^{(2)} - \boldsymbol{u}_4\widehat{G}_3, \boldsymbol{b}_5^{(2)} - \boldsymbol{u}_4\widehat{G}_4\Big)$$

obtaining the solution $\boldsymbol{x}_2 = \boldsymbol{u}_2$.

In the general case, for a burst of erasures of length L with decoding delay T, we can recover the packet $\boldsymbol{u}_i$, for $i = 0, 1, \ldots, L-1$, solving the system

$$(\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \ldots, x_{L-1}, \boldsymbol{u}_L, \boldsymbol{u}_{L+1}, \boldsymbol{u}_{L+i-1})\,\widehat{G}_{L+T}^{trunc} = \Big(\boldsymbol{b}_i^{(i)}, \boldsymbol{b}_{i+1}^{(i)}, \ldots, \boldsymbol{b}_{i+T-L}^{(i)}\Big),$$

where

$$\boldsymbol{b}_{i+j}^{(i)} = \boldsymbol{v}_{i+j+L} - \sum_{\ell=0}^{j} \boldsymbol{u}_{i+j+L-\ell}\widehat{G}_\ell - \sum_{\ell=0}^{i-1} \boldsymbol{u}_\ell\widehat{G}_{i+j+L-\ell},$$

for $j = 0, 1, \ldots T-L$, obtaining from Lemma 1 that the solution of this system is $\boldsymbol{x}_i = \boldsymbol{u}_i$. The second summation only exist when $i \geq 1$.

The following result characterizes convolutional encoders that admit decoding of burst of erasures of length up to L with delay T.

Theorem 2 *A systematic encoder $G(D) = [I_k \ \widehat{G}(D)] \in \mathbb{F}[D]^{k\times n}$ can recover a burst of erasures of length L with decoding delay T if the matrix $\widehat{G}^{trunc}_{L+T}$ given by (2) satisfies these conditions*

1. *the submatrix* $\left(\widehat{G}_L \ \widehat{G}_{L+1} \ \cdots \ \widehat{G}_T\right)$ *has full row rank,*
2. $\text{rowspan}\left(\widehat{G}_L \ \widehat{G}_{L+1} \ \cdots \ \widehat{G}_T\right) \cap \text{rowspan}\begin{pmatrix} \widehat{G}_{L-1} & \widehat{G}_L & \cdots & \widehat{G}_{T-1} \\ \vdots & \vdots & \dots & \vdots \\ \widehat{G}_1 & \widehat{G}_2 & \cdots & \widehat{G}_{T-L+1} \end{pmatrix} = \{\mathbf{0}\}.$

Proof Suppose without loss of generality that the transmission started at time zero and we have received the information packets $(\boldsymbol{v}_0, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{i-1})$ correctly up to an instant $i-1$ and from instant i to time instant $i+L-1$ we receive a burst of erasures. After the burst of erasures we receive the packets $(\boldsymbol{v}_{i+L}, \boldsymbol{v}_{i+L+1}, \dots)$ correctly again.

By the systematicy of $G(D)$ we straightforward recover $\boldsymbol{u}_j$, for $j \in \mathbb{N} \setminus \{i, i+1, \dots, i+L-1\}$.

Hence, it remains to recover the missing packets $(\boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \dots, \boldsymbol{u}_{i+L-1})$. If we regard this burst as unknowns to-be-determined and denote it as $\boldsymbol{x} = (\boldsymbol{x}_i, \boldsymbol{x}_{i+1}, \dots, \boldsymbol{x}_{i+L-1})$, then, taking into account the systematicity of $G(D)$ and the structure of G^c_∞, it follows that

$$\boldsymbol{x} \ \widehat{G}^{trunc}_{L+T} = \boldsymbol{b}$$

where $\boldsymbol{b}$ is a vector that can be computed from the correctly received data up to time $i+T$. From Lemma 1, the unknown components of $\boldsymbol{x} \in \mathbb{F}^{Lk}$ admit a unique solution if the corresponding row of $\widehat{G}^{trunc}_{L+T}$ is nonzero and independent of the other columns. Thus, assuming the conditions of the theorem it follows that $\boldsymbol{x}_i$ can be uniquely computed, or in other words, we can fully recover $\boldsymbol{u}_i$.

Next, due to the block-Toeplitz structure of G^c_∞ we can apply the same reasoning for $(\boldsymbol{x}_{i+1}, \boldsymbol{x}_{i+2}, \dots, \boldsymbol{x}_{i+L-1}, \boldsymbol{v}_{i+L})$, where $\boldsymbol{v}_{i+L}$ is known, to recover $\boldsymbol{x}_{i+1}$. Repeating the same arguments we can compute the remaining $\boldsymbol{x}_j$, $j = i+2, \dots, i+L-1$. This concludes the proof. □

The computation of each erased symbol (and therefore of the decoding) requires only linear algebra, as already noted in [25,26], and boils down to solving an associated linear system of equations as described in the proof of the theorem. Theorem 1, together with bound (1), characterizes convolutional encoders that admit low delay decoding over burst erasure channels. The encoders presented in [5,19] verify the conditions of Theorem 1 with a rate that achieves the bound (1) and therefore they admit optimal decoding delay. In the next section we present another instance of such encoders. As the proposed class of encoders is simple, and contains only zeros and ones, the decoding becomes straightforward.

4 Our Construction

Suppose that in the channel only burst of erasures of length L occur. We first consider the case $k > n-k$, say, $(n-k)\lambda + \gamma = k$ for some integer λ and $\gamma < n-k$. Let $G(D) = (I_k \;\; \widehat{G}(D)) \in \mathbb{F}^{k\times n}$, $\widehat{G}(D) = \sum_{j\geq 0} \widehat{G}_j D^j$ be a systematic encoder given by

$$\widehat{G}_{iL} = \begin{pmatrix} O_{(i-1)(n-k)\times(n-k)} \\ I_{n-k} \\ O_{k-i(n-k)\times(n-k)} \end{pmatrix}, \text{ for } i = 1, 2, \ldots, \lambda.$$

If $(n-k) \nmid k$, i.e., $\gamma \neq 0$, then, we also define

$$\widehat{G}_{(\lambda+1)L} = \begin{pmatrix} O_{(k-\gamma)\times\gamma} & O_{(k-\gamma)\times(n-k)} \\ I_\gamma & O_{\gamma\times(n-k)} \end{pmatrix}.$$

The remaining coefficients of $\widehat{G}(D)$ are null matrices.

Suppose that a burst of erasure of length L occurs at time j. Then, one can verify that at time instant $j + iL$, we recover $n-k$ coordinates of $\boldsymbol{u}_j$ for $i = 1, 2, \ldots, \lambda$, and wait until time $j+(\lambda+1)L$ to retrieve the remaining part of $\boldsymbol{u}_j$, if required. Then, the delay to recover $\boldsymbol{u}_j$ is $T = \left\lceil \frac{k}{n-k} \right\rceil$. Furthermore, due to the block-Toeplitz structure of the sliding matrix it follows that T is also the delay for decoding all the remaining erasures of $\boldsymbol{v}_s$, $s = j+1, \ldots, j+L-1$. Assume now for simplicity that $\gamma = 0$ to show that the bound in (1) is achieved with equality. First note that $\frac{R}{1-R} = \lambda$ for the selected parameters $(n-k)\lambda = k$ and $R = k/n$. Now, it is easy to verify that $T = \lambda L$ and therefore $\frac{T}{L} = \lambda = \frac{R}{1-R}$.

Thus, the proposed construction admits an optimal delay decoding when only bursts of erasures of length up to L occur. Note that this construction requires only binary entries, whereas previous contributions (see for instance [18,19]) require larger finite fields. As a consequience, the decoding of our construction is computationally more efficient. The case $k \leq n-k$ readily follows by considering

$$\widehat{G}_L = \begin{pmatrix} O \; I_k \end{pmatrix}$$

and the remaining coefficients of $G(D)$, G_j, $j \notin \{0, L\}$ null matrices.

Example 2 Consider the parameters given in Example 1. The construction described above yields the following matrices

$$\widehat{G}_0 = \widehat{G}_1 = \widehat{G}_2 = \widehat{G}_4 = \widehat{G}_5 = O_{4\times 2}, \widehat{G}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \widehat{G}_6 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So, we have

$$\widehat{G}_9^{trunc} = \left(\begin{array}{cc|cc|cc|cc} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ \hline 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{array}\right).$$

Assume that we have lost the first three packets $\boldsymbol{v}_0, \boldsymbol{v}_1$ and $\boldsymbol{v}_2$, in other words, we need to recover the unknowns $\boldsymbol{u}_0, \boldsymbol{u}_1$ and $\boldsymbol{u}_2$ from message. We denote these missing information packets by $\boldsymbol{x}_0 = (x_{00}, x_{01}, x_{02}, x_{03})$, $\boldsymbol{x}_1 = (x_{10}, x_{11}, x_{12}, x_{13})$ and $\boldsymbol{x}_2 = (x_{20}, x_{21}, x_{22}, x_{23})$; to recover $\boldsymbol{u}_0$ we have to solve the system

$$(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)\widehat{G}_9^{trunc} = \left(\widehat{\boldsymbol{v}}_3, \widehat{\boldsymbol{v}}_4, \widehat{\boldsymbol{v}}_5, \widehat{\boldsymbol{v}}_6 - \boldsymbol{u}_3\widehat{G}_3\right),$$

where $(\widehat{\boldsymbol{v}}_3, \widehat{\boldsymbol{v}}_4, \widehat{\boldsymbol{v}}_5, \widehat{\boldsymbol{v}}_6)$ is known. In our example, we have that

$$\begin{aligned}(u_{00}, u_{01}, u_{02}, u_{03}, u_{10}, u_{11}, u_{12}, u_{13}, u_{20}, u_{21}, u_{22}, u_{23})\, \widehat{G}_9^{trunc} \\ = (u_{00}\ u_{01}, u_{10}\ u_{11}, u_{20}\ u_{21}, u_{02}\ u_{03}).\end{aligned}$$

From the known information of $\widehat{\boldsymbol{v}}_3$ and $\widehat{\boldsymbol{v}}_6$, we recover directly the vector $\boldsymbol{u}_0$ completely, at time instant 6 because of the form of our submatrices $\widehat{G}_3$ and $\widehat{G}_6$. To recover the rest of the packets $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ we repeat the arguments given in Example 1 but using our construction.

Acknowledgements

This work was partially supported by Spanish grants AICO/2017/128 of the Generalitat Valenciana and VIGROB-287 of the Univesitat d' Alacant.

References

1. Almeida, P., Napp, D., Pinto, R.: A new class of superregular matrices and MDP convolutional codes. Linear Algebra and its Applications **439**(7), 2145–2157 (2013)
2. Almeida, P., Napp, D., Pinto, R.: Superregular matrices and applications to convolutional codes. Linear Algebra and its Applications **499**, 1–25 (2016)
3. Arai, M., Yamaguci, A., Fukumpto, S., Iwasaki, K.: Method to recover lost internet packets using (n, k, m) convolutional code. Electronics and Communications in Japan **88**(7), Part 3 (2005)
4. Badr, A., Khisti, A., Tan, W.T., Apostolopoulos, J.: Robust streaming erasure codes based on deterministic channel approximations. In: 2013 IEEE International Symposium on Information Theory, pp. 1002–1006 (2013)

5. Badr, A., Khisti, A., Tan, W.T., Apostolopoulos, J.: Layered constructions for low-delay streaming codes. IEEE Trans. Inform. Theory **63**(1), 111–141 (2017)
6. Barbero, A., Ytrehus, Ø.: Rate $(n-1)/n$ systematic memory maximum distance separable convolutional codes. IEEE Trans. Inform. Theory **64**(4), 3018–3030 (2018)
7. Deng, H., Kuijper, M., Evans, J.: Burst erasure correction capabilities of $(n, n-1)$ convolutional codes. In: 2009 IEEE International Conference on Communications, pp. 1–5 (2009)
8. Gluesing-Luerssen, H., Rosenthal, J., Smarandache, R.: Strongly MDS convolutional codes. IEEE Trans. Inform. Theory **52**(2), 584–598 (2006)
9. Gómez-Torrecillas, J., Lobillo, F.J., Navarro, G.: A Sugiyama-like decoding algorithm for convolutional codes. IEEE Trans. Inform. Theory **63**(10), 6216–6226 (2017)
10. Guardia, G.G.L.: On negacyclic MDS-convolutional codes. Linear Algebra and its Applications **448**(Supplement C), 85 – 96 (2014)
11. Hansen, J., Østergaard, J., Kudahl, J., Madsen, J.: On the construction of jointly superregular lower triangular toeplitz matrices. International Symposium on Information Theory (ISIT) (2016)
12. Hansen, J., Østergaard, J., Kudahl, J., Madsen, J.H.: Superregular lower triangular Toeplitz matrices for low delay wireless streaming. IEEE Transactions on Communications **65**(9), 4027–4038 (2017)
13. Hutchinson, R.: The existence of strongly MDS convolutional codes. SIAM Journal on Control and Optimization **47**(6), 2812–2826 (2008)
14. Hutchinson, R., Smarandache, R., Trumpf, J.: On superregular matrices and MDP convolutional codes. Linear Algebra and its Applications **428**, 2585–2596 (2008)
15. Johannesson, R., Zigangirov, K.S.: Fundamentals of Convolutional Coding. IEEE Press, New York (2015)
16. Kuijper, M., Bossert, M.: On (partial) unit memory codes based on reed-solomon codes for streaming. In: 2016 IEEE International Symposium on Information Theory (ISIT), pp. 920–924 (2016)
17. Mahmood, R., Badr, A., Khisti, A.: Streaming-codes for multicast over burst erasure channels. IEEE Trans. Inform. Theory **61**(8), 4181–4208 (2015)
18. Mahmood, R., Badr, A., Khisti, A.: Convolutional codes with maximum column sum rank for network streaming. IEEE Trans. Inform. Theory **62**(6), 3039–3052 (2016)
19. Martinian, E., Sundberg, C.E.W.: Burst erasure correction codes with low decoding delay. IEEE Transactions on Information Theory **50**(10), 2494–2502 (2004)
20. McEliece, R.J.: The algebraic theory of convolutional codes. In: Handbook of Coding Theory, vol. 1, pp. 1065–1138. Elsevier Science Publishers (1998)
21. Muñoz Porras, J.M., Domínguez Pérez, J.A., Iglesias Curto, J.I., Serrano Sotelo, G.: Convolutionall Goppa codes. IEEE Trans. Inform. Theory **52**(1), 340–344 (2006)
22. Napp, D., Smarandache, R.: Constructing strongly MDS convolutional codes with maximum distance profile. Advances in Mathematics of Communications **10**(2), 275–290 (2016)
23. Rosenthal, J.: An algebraic decoding algorithm for convolutional codes. In: G. Picci, D. Gilliam (eds.) Dynamical Systems, Control, Coding, Computer Vision: New Trends, Interfaces, and Interplay, pp. 343–360. Birkäuser, Boston-Basel-Berlin (1999)
24. Rosenthal, J., Smarandache, R.: Maximum distance separable convolutional codes. Appl. Algebra Engrg. Comm. Comput. **10**(1), 15–32 (1999)
25. Tomás, V.: Complete-mdp convolutional codes over the erasure channel. Ph.D. thesis, Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante, Alicante, España (2010)
26. Tomás, V., Rosenthal, J., Smarandache, R.: Decoding of convolutional codes over the erasure channel. IEEE Trans. Inform. Theory **58**(1), 90–108 (2012)