A new rank metric for convolutional codes

P. Almeida · D. Napp

Abstract Let $\mathbb{F}[D]$ be the polynomial ring with entries in a finite field $\mathbb{F}$. Convolutional codes are submodules of $\mathbb{F}[D]^n$ that can be described by left prime polynomial matrices. In the last decade there has been a great interest in convolutional codes equipped with a rank metric, called sum rank metric, due to their wide range of applications in reliable linear network coding. However, this metric suits only for delay free networks. In this work we continue this thread of research and introduce a new metric that overcomes this restriction and therefore is suitable to handle more general networks. We study this metric and provide characterizations of the distance properties in terms of the polynomial matrix representations of the convolutional code. Convolutional codes that are optimal with respect to this new metric are investigated and concrete constructions are presented. These codes are the analogs of Maximum Distance Profile convolutional codes in the context of network coding. Moreover, we show that they can be built upon a class of superregular matrices, with entries in an extension field, that preserve their superregularity properties even after multiplication with some matrices with entries in the ground field.

1 Introduction

Within the area of coding theory, network coding has been a very active topic of research as it provides an effective tool to disseminate information (packets) over networks. Mathematically, we can consider the transmitted packets as columns of a matrix with entries in a finite field $\mathbb{F}_q$, and the linear combinations performed in the nodes of the network correspond to columns operations on this matrix. If no errors occur during the transmission over such a network, the $\mathbb{F}_q$-column space

P. Almeida
CIDMA - Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, Portugal
E-mail: palmeida@ua.pt

D. Napp
Department of Mathematics, University of Alicante, Spain
E-mail: diego.napp@ua.es

of the transmitted matrix remains invariant. To achieve a reliable communication over this channel, matrix codes are employed forming the so-called *rank metric codes* (see [31]). Rank metric codes such as Gabidulin codes are known to be able to protect packets in such a scenario. We call these codes *one-shot* codes, as they use the (network) channel only once. However, coding can also be performed over multiple uses of the network (*multi-shot*) in a sequential fashion, as it has been recently shown by several authors, see for instance [5, 18, 23, 27, 32].

The general idea of multi-shot network coding stems from the fact that creating dependencies among the sequence of transmitted codewords (subspaces) of different shots can improve the error-correction capabilities of the code. A first attempt to explain multi-shot network coding was presented in [27] and a type of concatenated n-shot codes ($n \geq 1$) was proposed based on a multilevel code (see also [24]). Apart from concatenated codes, another very natural way to spread redundancy across codewords is by means of convolutional codes (see [15, 22]). Adapting this class of codes to the context of networks gave rise to rank metric convolutional codes (see [5, 18, 19, 32]). The work in [32] was pioneer in this direction by presenting the first class of rank metric convolutional codes together with a decoding algorithm able to deal with errors, erasures and deviations. However, the results were only valid for unit memory convolutional codes and in [5, 18, 19] an interesting and more general class of rank metric convolutional codes was introduced to cope with network streaming applications. For a more general and theoretical framework for rank metric convolutional codes we refer to [23].

The first metric proposed in this context was the *active column sum rank distance* in [32] and it was defined using the state trellis graph of the convolutional code. The j-th active column sum distance just considers sequences that are constructed by exiting the zero state of the trellis at time instant 0 and not re-entering it for $1 \leq t \leq j - 1$. Thus, some sequences are not considered in the time interval $[0, j]$ and therefore this metric is not a sufficient metric to guarantee decoding within a time interval. Hence, a new distance, called *sum rank distance*, was introduced as a generalization of the active column rank distance and the rank distance used for one-shot network coding (see [19], [21] and [27]). This new distance has proven to be the proper notion in order to deal with networks that are delay-free. In delay-free networks it is assumed that the natural delay in the transmission (due, for instance, to the delay of the nodes) is so small that can be disregarded.

In this work, we continue this thread of research and consider convolutional codes tailor-made to handle networks with delays. We show that the previous metrics are not enough to characterize the error correction capability of the code in these networks (see Example 1) and consider a new (rank) metric, called column rank distance, that solves this problem. This distance is the rank analog of the so-called column distance of Hamming convolutional codes, see [10, 14, 15]. It can also be seen as the standard rank distance of a block code, but the code considered here is *nonlinear* as it is derived from truncated convolutional codes, see definition of column block code in (7). Hence, previous known results of rank metric linear block codes cannot be straightforward applied in this case. We will show that the column rank distance characterizes the error correcting capability of the convolutional code within a time interval, in more general network channels (see Theorem 4). Moreover such characterization leads to an efficient algorithm to recover the lost packages (see Example 2). We also present concrete constructions of convolutional codes

that are optimal with respect to the column rank distance using a particular type of superregular matrices that preserve superregularity after some linear operations.

In Section 2, we present fundamental results on classical one-shot network coding, multi-shot network coding and convolutional codes. With this basic knowledge we introduce, in Section 3, the metrics used so far for multi-shot network coding and introduce a new rank metric, presenting also some of its characterizations (see Theorem 3) and properties (see Theorem 4). Section 4 is devoted to convolutional codes that are optimal with respect to this metric, called Maximum Rank Profile, and we provide a characterization of them in terms of the sliding matrix of the code (see Theorem 6). We also present novel results on the optimization of the recovery of packets when this class of optimal codes are employed (see Theorem 7). In Section 5 we discuss and solve the problem of the existence and construction of these codes. We prove that this problem can be reduced to the construction of certain classes of matrices called superregular matrices (see Theorem 8) and present a concrete class of such matrices. The problem of deriving superregular matrices to build convolutional codes has become an active area of research, see for instance [3,6,12], and the results presented in this section extend the known results on this topic.

2 Preliminaries

In order to state more precisely our results, we introduce in this section the necessary material and notation on standard theory of (multi-shot) network coding and convolutional codes.

Let q be a prime power and let $\mathbb{F}_q$ denote the finite field with q elements and let $M > 1$ be an integer. It is well-known that there always exists a primitive element α of the extension field $\mathbb{F}_{q^M}$, such that $\mathbb{F}_{q^M}$ is isomorphic to $\mathbb{F}_q[\alpha]$. Moreover, $\mathbb{F}_{q^M}$ is isomorphic (as a vector space over $\mathbb{F}_q$) to the vector space $\mathbb{F}_q^M$. One then easily obtains the isomorphic description of matrices over the base field $\mathbb{F}_q$ as vectors over the extension field, $i.e.$, $\mathbb{F}_q^{M \times N} \sim \mathbb{F}_{q^M}^N$. Let $\phi : \mathbb{F}_{q^M}^N \to \mathbb{F}_q^{M \times N}$ be the isomorphism that converts vectors in $\mathbf{w} \in \mathbb{F}_{q^M}^N$ into matrices $\phi(\mathbf{w}) = W \in \mathbb{F}_q^{M \times N}$.

2.1 From One-shot to Multi-shot Network Coding

2.1.1 The network model: One-shot

Let $\mathbf{v} \in \mathbb{F}_{q^M}^n$ (or equivalently $V \in \mathbb{F}_q^{M \times n}$), called *channel packet*, $\mathbf{v}$ represents the n packets of length M to be sent through the network. We shall follow the approach of [19] and consider a rank-deficiency channel for one shot given by

$$\mathbf{x}^\top = A^\star \mathbf{v}^\top, \tag{1}$$

where $\mathbf{x} \in \mathbb{F}_{q^M}^n$ represents the received packets and $A^\star \in \mathbb{F}_q^{n \times n}$ is the *rank deficiency channel matrix*. The channel matrix $A^\star$ corresponds to the overall linear transformations applied by the network over the base field $\mathbb{F}_q$ and it is known by the receiver (as the combinations are carried over in the header bits of the packets). For perfect communications we have that $\text{rank}(A^\star) = n$, but failing and

deactivated links may cause a rank deficient channel matrix. We call $n - \mathrm{rank}(A^\star)$ the *rank deficiency* of the channel. In order to protect information in this setting, rank metric codes are used.

A *rank metric code* $\mathcal{C}$ is defined as a nonempty subset of $\mathbb{F}_q^{M \times n}$ equipped with the distance measure $d_{\mathrm{rank}}(V, W) = \mathrm{rank}(V - W)$, where $V, W \in \mathbb{F}_q^{M \times n}$ (see [16]).

The *rank distance* of a code $\mathcal{C} \subset \mathbb{F}_q^{M \times n}$ is defined as

$$d_{\mathrm{rank}}(\mathcal{C}) = \min_{V, W \in \mathcal{C}, \ V \neq W} d_{\mathrm{rank}}(V, W). \tag{2}$$

Although rank metric codes are in $\mathbb{F}_q^{M \times n}$, they are usually constructed as linear block codes of length n over the extension field $\mathbb{F}_{q^M}$ (see [16]). Then, with some obvious abuse of notation, we will sometimes use $d_{\mathrm{rank}}(\mathcal{C})$ when $\mathcal{C} \subset \mathbb{F}_{q^M}^n$. Consider linear codes over $\mathbb{F}_{q^M}$ and use k for their dimension. A linear (n, k) rank metric code over $\mathbb{F}_{q^M}$ satisfies the following analog of the Singleton bound:

$$d_{\mathrm{rank}}(\mathcal{C}) \leq n - k + 1.$$

A code that achieves this bound is called Maximum Rank Distance (MRD). Gabidulin codes, introduced in [9], are a well-known class of MRD codes, see also [8] and [26].

2.1.2 Multi-shot

In many situations the sender needs to transmit not only one single vector $\mathbf{v} \in \mathbb{F}_{q^M}^n$ of data but a sequence of vectors. Suppose that we have a stream of source packets $\mathbf{u}_0, \mathbf{u}_1, \cdots \in \mathbb{F}_{q^M}^k$ to be transmitted. Then, the idea of multi-shot network coding is that instead of encoding each $\mathbf{u}_t$ in $\mathbf{v}_t$ independently of other $\mathbf{u}_i$, one can generate each codeword $\mathbf{v}_t$ based on $\mathbf{u}_t$ and previous $\mathbf{u}_i$, $i \leq t$ to improve the error correcting capabilities of the code.

The multi-shot setting can be described as follows: the transmitter receives at each time instance t a source packet $\mathbf{u}_t \in \mathbb{F}_{q^M}^k$ (constituted by a set of k packets). A channel packet $\mathbf{v}_t \in \mathbb{F}_{q^M}^n$ (constituted by a set of n packets) is constructed using not only $\mathbf{u}_t$ but also previous source packets $\mathbf{u}_0, \ldots, \mathbf{u}_{t-1}$. Then, the channel packet $\mathbf{v}_t$ is sent through the network at each shot (time instance) t. The receiver observes not only a combination of the n packets sent at instant t but also of previous packets sent at time $i < t$. Hence, following the operator channel in (1) at each shot t the received packets $\mathbf{x}_t \in \mathbb{F}_{q^M}^n$ are linear combinations of the packets of $\mathbf{v}_t$ and, if there is delay in the transmission, also of combinations of the previous packets $\mathbf{v}_0, \ldots, \mathbf{v}_{t-1}$. Hence, for any $j \geq 0$, we have

$$\mathbf{x}_{[0,j]}^\top = A_{[0,j]}^\star \mathbf{v}_{[0,j]}^\top \tag{3}$$

where $\mathbf{x}_{[0,j]} = (\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_j)$, $\mathbf{v}_{[0,j]} = (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j) \in \mathbb{F}_{q^M}^{n(j+1)}$ and $A_{[0,j]}^\star$ is a block lower triangular *truncated channel matrix*

$$A_{[0,j]}^\star = \begin{bmatrix} A_{00}^\star & & & \\ A_{10}^\star & A_{11}^\star & & \\ \vdots & \vdots & \ddots & \\ A_{j0}^\star & A_{j1}^\star & \cdots & A_{jj}^\star \end{bmatrix}, \tag{4}$$

where $A^{\star}_{i\ell}$ are square (not necessarily nonsingular) matrices of order n, for $i \le \ell \le j$. The rank deficiency during the time interval $[0, j]$ is $n(j + 1) - \text{rank}(A^{\star}_{[0,j]})$ and measures the amount of information lost during this interval.

So far this channel model has not been proposed nor addressed in the literature in this generality and only the delay-free case has been considered, see [19] and references therein. In the delay-free case only combinations of packets of $\mathbf{v}_t$ arrive at time instance t and not of packets of $\mathbf{v}_i$, $i < t$ and therefore in this case the rank deficiency matrix $A^{\star}_{[0,j]}$ is a block diagonal matrix, *i.e.*, $A^{\star}_{i\ell} = 0$ for $i > \ell$. This assumption constitutes a restriction as there may exist a time-varying delay in the end-to-end transmission due to link delays. As we shall see, the extension to the above more general setting is not straightforward. In the next section, we introduce a metric that suits that framework.

Note that, in practice, the receiver can compute the rank deficiency as each packet carries a label identifying the shot (or generation) to which it corresponds. The possibility of using multi-shot coding in the context of network coding was already observed in the seminal papers [16] and [31]. In this setting, convolutional codes are a natural way to protect and process information in a sequential manner.

2.2 Convolutional codes

As opposed to block codes, convolutional codes treat the information as a stream of data. If we introduce a variable D, usually called the *delay operator*, to indicate the time instant in which each information arrived or each codeword was transmitted, then we can represent the sequence message $(\mathbf{v}_0, \mathbf{v}_1, \cdots)$ as a polynomial sequence $\mathbf{v}(D) = \mathbf{v}_0 + \mathbf{v}_1 D + \cdots$. Similarly, we can represent the information vector to be encoded as $\mathbf{u}(D)$ and the encoder (typically implemented by means of a linear finite-state shift register) as $G(D)$. Formally, we can define convolutional codes as follows (see [10]): A *convolutional code* $\mathcal{C}$ of rate k/n is an $\mathbb{F}_{q^M}[D]$-submodule of $\mathbb{F}_{q^M}[D]^n$ of rank k such that there exists a polynomial matrix $G(D) \in \mathbb{F}_{q^M}[D]^{k \times n}$, called generator matrix, that is *basic*, *i.e.*, there exists a polynomial matrix $X(D)$ such that $G(D)X(D) = I_k$ (that is, $G(z)$ has a polynomial right inverse), with the property that

$$\mathcal{C} = \text{im } G(D) = \left\{ \mathbf{v}(D) = \mathbf{u}(D)G(D) \ \mid \ \mathbf{u}(D) \in \mathbb{F}_{q^M}^k[D] \right\}.$$

Two generator matrices $G_1(D)$ and $G_2(D)$ generate the same convolutional code if there exists a unimodular matrix $U(D) \in \mathbb{F}_{q^M}^{k \times k}[D]$, that is, its determinant is a nonzero element of $\mathbb{F}_{q^M}$, such that $G_1(D) = U(D)G_2(D)$. The *degree* δ of a convolutional code $\mathcal{C}$ is the maximum of the degrees of the determinants of the $k \times k$ sub-matrices of one, and hence, any generator matrix of $\mathcal{C}$. The degree of $\mathcal{C}$ is the minimum size of a state space realization of $\mathcal{C}$ (the McMillan degree of the corresponding linear system, see [30]). For that, we feel that it is the single most important code parameter on the side of the transmission rate k/n. Note that a block code is a convolutional code with $\delta = 0$. For more general classes of convolutional codes see [4, 25]

In the sequel, we adopt the notation of McEliece [22, p. 1082] and denote a convolutional code of rate k/n and degree δ an (n, k, δ)-convolutional code.

If we write $\mathbf{v}(D) = \mathbf{v}_0 + \mathbf{v}_1 D + \ldots + \mathbf{v}_\ell D^\ell$, and we represent $G(D)$ as a matrix polynomial,

$$G(D) = G_0 + G_1 D + \cdots + G_\nu D^\nu,$$

where $G_\nu \neq 0$ and $G_i = 0$, for $i > \nu$, the *truncated sliding generator matrix* is

$$G_{trunc}(j) = \begin{bmatrix} G_0 & G_1 & \cdots & G_j \\ & G_0 & \cdots & G_{j-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{bmatrix} \in \mathbb{F}_{q^M}^{k(j+1) \times n(j+1)} \tag{5}$$

for any $j \geq 0$, and a truncated codeword can be represented as

$$\mathbf{v}_{[0,j]} = (\mathbf{v}_0, \mathbf{v}_1, \cdots \mathbf{v}_j) = (\mathbf{u}_0, \mathbf{u}_1, \cdots \mathbf{u}_j) \, G_{trunc}(j). \tag{6}$$

For each channel packet it holds that

$$\mathbf{v}_t = \sum_{i=0}^{\ell} \mathbf{u}_{t-i} G_i.$$

Define now the j-th *column block code* of $\mathcal{C}$ as

$$\begin{aligned} \mathcal{C}_j^c &= \left\{ (\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_j) G_{trunc}(j) \mid \mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_j \in \mathbb{F}_{q^M}^k, \mathbf{u}_0 \neq 0 \right\} \\ &= \left\{ (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j) \mid \sum_{h \in \mathbb{N}} \mathbf{v}_h D^h \in \mathcal{C}, \mathbf{v}_0 \neq 0 \right\} \subseteq \mathbb{F}_{q^M}^{(j+1)n}. \end{aligned} \tag{7}$$

Note that this block code is *nonlinear* (over $\mathbb{F}_{q^M}$) as the zero vector is not in $\mathcal{C}_j^c$. In this work we will be primarily interested in the correction capabilities of this nonlinear column block code rather than in the linear (over $\mathbb{F}_{q^M}[D]$) code $\mathcal{C}$.

The code $\mathcal{C}$ can be equivalently described using an $(n - k) \times n$ full rank polynomial *parity-check matrix* $H(D)$, defined by

$$\mathcal{C} = \ker H(D) = \left\{ \mathbf{v}(D) \in \mathbb{F}_{q^M}[D]^n \mid H(D)\mathbf{v}(D)^\top = 0 \in \mathbb{F}_{q^M}[D]^{n-k} \right\},$$

and the associated *truncated sliding parity-check matrix* of $H(D) = \sum_{i=0}^m H_i D^i$ is

$$H_{trunc}(j) = \begin{bmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{bmatrix} \tag{8}$$

with $H_j = 0$ when $j > m$, $j \in \mathbb{N}$.

3 Metrics for multi-shot network coding

The sum rank distance, the distance that has been widely considered for multi-shot network coding, was first introduced in [27] under the name of *extended rank distance*. We note that the sum rank distance has been also used in the context of block codes reducing their decoding complexity, see for instance [21].

Let $\mathbf{v} = (\mathbf{v}_0, \ldots, \mathbf{v}_t)$ and $\mathbf{w} = (\mathbf{w}_0, \ldots, \mathbf{w}_t)$ be two $(t+1)$-tuples of vectors in $\mathbb{F}_{q^M}^n$. The *sum rank distance* (SRD) between them is

$$d_{\mathrm{SR}}(\mathbf{v}, \mathbf{w}) = \sum_{i=0}^{t} \mathrm{rank}(\phi(\mathbf{v}_i) - \phi(\mathbf{w}_i)),$$

where ϕ is the isomorphism that converts vectors $\mathbf{v} \in \mathbb{F}_{q^M}^n$ into matrices $\phi(\mathbf{v}) = V \in \mathbb{F}_q^{M \times n}$.

For an (n, k, δ)-convolutional code $\mathcal{C}$ and $\mathbf{v}(D) = \mathbf{v}_0 + \mathbf{v}_1 D + \mathbf{v}_2 D^2 + \cdots \in \mathcal{C}$ we define its *free sum rank distance* (FSRD) as

$$d_{\mathrm{SR}}(\mathcal{C}) = \min \left\{ \sum_{i \geq 0} \mathrm{rank}(\phi(\mathbf{v}_i)) \mid \mathbf{v}(D) \in \mathcal{C} \text{ and } \mathbf{v}(D) \neq 0 \right\},$$

and, for $0 \leq j \leq t$, its *j-th Column Sum Rank Distance* (*j*-th CSRD) as

$$d_{\mathrm{SR}}^j(\mathcal{C}) = \min \left\{ \sum_{i=0}^{j} \mathrm{rank}(\phi(\mathbf{v}_i)) \mid \mathbf{v}(D) \in \mathcal{C} \text{ and } \mathbf{v}_0 \neq 0 \right\}.$$

In [32], concrete decoding algorithms for unit memory rank metric convolutional codes were presented using another distance, namely the active rank distance. However, in [19], it was shown that this metric fails to fully determine the error-correcting capabilities of rank metric convolutional codes with arbitrary memory and the *j*-th CSRD needs to be considered.

Moreover, in [23, Theorem 4.1] it was derived a Singleton-type upper bound for the FSRD of $\mathcal{C}$ in a different setting. This can be adapted easily to our context to obtain:

$$d_{\mathrm{SR}}(\mathcal{C}) \leq (n - k) \left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1. \tag{9}$$

Note that this bound coincides with the so-called generalized Singleton bound of (Hamming) (n, k, δ)-convolutional codes (see [28] and [29]). Thus, the bound (9) could be also derived from the fact that the rank distance is upper bounded by the Hamming distance (see [9, 20]). Although some results were presented in [7], the problem of existence and construction of rank metric convolutional codes whose free rank distance achieves the bound (9) remains open.

The following result was obtained in [19] and shows that the *j*-th CSRD can be used to characterize the error-correcting capabilities of multi-shot codes in the context of delay-free networks.

Theorem 1 *[19, Theorem 2] Let $\mathcal{C} = \operatorname{im} G(D) \subset \mathbb{F}_{q^M}^n$ be a rank metric convolutional code, $\mathbf{v}(D) = \mathbf{u}(D)G(D) \in \mathcal{C}$, $\mathbf{u}_0 \neq 0$ and $A_{[0,j]}^*$ be the truncated block diagonal channel matrix. Then, $\mathbf{u}_0$ is recoverable at time instant j if*

$$d_{\mathrm{SR}}^j(\mathcal{C}) > n(j+1) - \operatorname{rank}(A_{[0,j]}^*). \tag{10}$$

The following example illustrates that, when the network has delays in the transmission of packets, *i.e.*, when $A_{[0,j]}^*$ is not necessarily block diagonal, the j-th CSRD fails to characterize the rank deficiency correcting capability of $\mathcal{C}$.

Example 1 Let $G(D) = G_0 + G_1 D \in \mathbb{F}_{2^6}[D]^{2\times 3}$ be a generator matrix of the convolutional code $\mathcal{C} \subset \mathbb{F}_{2^6}[D]^3$, α a primitive element of $\mathbb{F}_{2^6}$ such that $\alpha^6 = \alpha + 1$ and

$$G_0 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & \alpha & \alpha^2 \end{bmatrix}, \ G_1 = \begin{bmatrix} 0 & 1 & 0 \\ \alpha^3 & \alpha^4 & \alpha^5 \end{bmatrix}.$$

It is easy to see that $d_{\mathrm{SR}}^1(\mathcal{C}) = 2$ and that there exists a $\mathbf{v}(D) \in \mathcal{C}$ such that $\mathbf{v}_{[0,1]} = (1,0,0 \mid 0,1,0) \in \mathbb{F}_{2^6}^6$.

By Theorem 1 one can always recover $\mathbf{u}_0$ in delay-free networks as far as the j-th CSRD of $\mathcal{C}$ is larger than the rank deficiency in the window $[0,j]$, *i.e.*, if $n(j+1) - \operatorname{rank}(A_{[0,j]}^*) < d_{\mathrm{SR}}^j(\mathcal{C})$. Thus, in this example one can recover $\mathbf{u}_0$ if $6 - \operatorname{rank}(A_{[0,1]}^\star) \leq 1$ or equivalently, if $\operatorname{rank}(A_{[0,1]}^\star) \geq 5$. However, in presence of delays in the network this does not necessarily hold. Take the channel matrix with delays

$$A_{[0,1]}^\star = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{F}_2^{6\times 6}$$

that has rank equal to 5 and yields $A_{[0,1]}^\star v_{[0,1]}^\top = 0$, i.e., $\mathbf{v}_{[0,1]}$ is indistinguishable from the zero sequence and therefore cannot be corrected.

For the general case, *i.e.*, for not necessarily delay free networks, we introduce the *j-th Column Rank Distance* (*j*-th CRD) for convolutional codes as follows:

$$d_{\mathrm{CR}}^j(\mathcal{C}) = \min \left\{ \operatorname{rank}(\phi(\mathbf{v}_0) \mid \cdots \mid \phi(\mathbf{v}_j)) \mid \mathbf{v}(D) = \sum_{i=0}^{j} \mathbf{v}_i D^i \in \mathcal{C} \text{ and } \mathbf{v}_0 \neq 0 \right\}$$

$$= \min \left\{ \operatorname{rank}(\phi(\mathbf{v}_{[0,j]})) \mid \mathbf{v}_{[0,j]} = \mathbf{u}_{[0,j]} G_{trunc}(j) \text{ and } \mathbf{u}_0 \neq 0 \right\}$$

$$= \min \left\{ \operatorname{rank}(\phi(\mathbf{v}_{[0,j]})) \mid H_{trunc}(j) \mathbf{v}_{[0,j]}^\top = 0 \text{ and } \mathbf{v}_0 \neq 0 \right\}$$

$$= d_{\mathrm{rank}}(\mathcal{C}_j^c). \tag{11}$$

Note that this metric is the rank metric of the block code $\mathcal{C}_j^c$ but such a code is nonlinear and therefore existing results on linear rank metric codes cannot be directly applied to our case.

Clearly, $d_{\mathrm{CR}}^j(\mathcal{C}) \leq d_{\mathrm{SR}}(\mathcal{C})$, for any $j \geq 0$. Therefore, we have the following Singleton bound for column rank distance:

$$d_{\mathrm{CR}}^j(\mathcal{C}) \leq (n-k)\left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1\right) + \delta + 1. \tag{12}$$

Next, we show that the j-th CRD characterizes the recovery capability within an interval over a channel that admits delays in the transmission. We then investigate how to build convolutional codes with designed j-th column rank distance. For the sake of clarity we first revise the linear block case and the characterization of the rank distance in terms of the properties of the corresponding parity-check matrices.

Lemma 2 *Let $\mathcal{C} \subset \mathbb{F}_{q^M}^n$ be a linear block code and $H \in \mathbb{F}_{q^M}^{(n-k)\times n}$ a parity-check of $\mathcal{C}$. Then, $d_{\mathrm{rank}}(\mathcal{C}) = d$ (defined in (2)) if and only if both of the following conditions hold*

1. *for all nonsingular matrix $A \in \mathbb{F}_q^{n\times n}$, every set of $d-1$ columns of HA, are linearly independent;*
2. *for all nonsingular matrix $A \in \mathbb{F}_q^{n\times n}$, there exists d linearly dependent columns of HA.*

Proof The proof follows easily from [9, Theorem 1]. $\qquad \blacksquare$

Next we derive a counterpart result in the context of convolutional codes. First, for any $j \geq 0$, $A_{[0,j]}$ will represent a nonsingular block lower triangular matrix with entries in the base field $\mathbb{F}_q$, of the form

$$A_{[0,j]} = \begin{bmatrix} A_{00} & & & \\ A_{10} & A_{11} & & \\ \vdots & \vdots & \ddots & \\ A_{j0} & A_{j1} & \cdots & A_{jj} \end{bmatrix} \in \mathbb{F}_q^{n(j+1)\times n(j+1)}. \tag{13}$$

Theorem 3 *Let $\mathcal{C} = \ker H(D) \subset \mathbb{F}_{q^M}[D]^n$. Then, the following are equivalent*

1. $d_{\mathrm{CR}}^j(\mathcal{C}) = d_{\mathrm{rank}}(\mathcal{C}_j^c) = d$;
2. *none of the first n columns of $H_{trunc}(j)A_{[0,j]}$ is contained in the $\mathbb{F}_{q^M}$-span of any other $d-2$ columns of $H_{trunc}(j)A_{[0,j]}$ for all nonsingular block lower triangular $A_{[0,j]}$ as in (13) and, moreover, one of the first n columns of $H_{trunc}(j)A_{[0,j]}$ is in the $\mathbb{F}_{q^M}$-span of other $d-1$ columns of $H_{trunc}(j)A_{[0,j]}$, $A_{[0,j]}$ as in (13).*

Proof $(2. \Rightarrow 1.)$: Suppose that one of the first n columns of $H_{trunc}(j)A_{[0,j]}$ is in the $\mathbb{F}_{q^M}$-span of other $d-1$ columns of $H_{trunc}(j)A_{[0,j]}$ for a nonsingular lower triangular $A_{[0,j]} \in \mathbb{F}_q^{n(j+1)\times n(j+1)}$. Thus, $H_{trunc}(j)A_{[0,j]}$ has d columns linearly dependent, say $c_{i_1}, c_{i_2}, \ldots, c_{i_d}$, therefore there exists $\alpha_\ell \in \mathbb{F}_{q^M}$, $\ell \in S = \{i_1, i_2, \ldots, i_d\}$, such that

$$\sum_{\ell \in S} \alpha_\ell c_\ell = 0$$

and at least one element in S, say i_1, belongs to $\{1, \ldots, n\}$. Now take $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n(j+1)}) \in \mathbb{F}_{q^M}^{n(j+1)}$ with $\mathbf{w}_\ell = 0$ if $\ell \notin S$, $\mathbf{w}_\ell = \alpha_\ell$ if $\ell \in S$. As $H_{trunc}(j)A_{[0,j]}\mathbf{w}^\top = 0$, then $\mathbf{w}A_{[0,j]}^\top$ is a codeword. Since $\phi(\mathbf{w})$ has at most d nonzero columns and $A_{[0,j]}$ is nonsingular then $\phi(\mathbf{w}A_{[0,j]}^\top)$ has rank at most d. Thus $d_{\mathrm{CR}}^j(\mathcal{C}) \leq d$.

To show that $d_{\mathrm{CR}}^j(\mathcal{C}) \geq d$ we do it by contradiction. Assume $d_{\mathrm{CR}}^j(\mathcal{C}) < d$. Then, there exists $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j) \in \mathbb{F}_{q^M}^{n(j+1)}$ with $\mathbf{v}_0 \neq 0$, $\mathrm{rank}(\phi(\mathbf{v})) \leq d - 1$ and $H_{trunc}(j)\mathbf{v}^\top = 0$. Since $\mathbf{v}_0 \neq 0$, at least one of the first n coordinates of $\mathbf{v}$ is nonzero. Take a nonsingular block upper triangular matrix $B_{[0,j]} \in \mathbb{F}_q^{n(j+1) \times n(j+1)}$ such that $\phi(\mathbf{v})B_{[0,j]}$ has at most $d - 1$ nonzero columns. Then $\mathbf{v}B_{[0,j]}$ has at most $d - 1$ nonzero coordinates. Let $A_{[0,j]} = (B_{[0,j]}^\top)^{-1}$, then $A_{[0,j]}$ is as in (13) and

$$H_{trunc}(j)A_{[0,j]}(B_{[0,j]}^\top \mathbf{v}^\top) = 0,$$

this implies that one of the first n columns of $H_{trunc}(j)A_{[0,j]}$ is contained in the span of other $d - 2$ columns of $H_{trunc}(j)A_{[0,j]}$. By the first part of 2. we obtain a contradiction. Hence, $d_{\mathrm{CR}}^j(\mathcal{C}) = d$.

The proof of $(1. \Rightarrow 2.)$ follows the same reasoning.

We are now in a position to provide a necessary condition to recover rank deficiencies within a given time interval.

Theorem 4 *Let* $\mathcal{C} = \mathrm{im}\, G(D) = \ker H(D) \subset \mathbb{F}_{q^M}^n$ *be an* (n, k, δ)-*convolutional code,* $\mathbf{v}(D) = \mathbf{u}(D)G(D) \in \mathcal{C}$, $\mathbf{u}_0 \neq 0$ *and* $A_{[0,j]}^*$ *be the truncated channel matrix, where* $j \geq 0$. *Then,* $\mathbf{u}_0$ *is recoverable at time instant* j *if*

$$d_{\mathrm{CR}}^j(\mathcal{C}) > n(j + 1) - \mathrm{rank}(A_{[0,j]}^*). \tag{14}$$

Proof As $G(D)$ is basic, $G(0)$ is full row rank and therefore to recover $\mathbf{u}_0$ it is enough to decode $\mathbf{v}_0$. Obviously, if $\mathrm{rank}A_{00}^* = n$ there is no rank deficiency at time instant zero and $\mathbf{v}_0$ can be immediately recovered. Assume then that $\mathrm{rank}A_{00}^* < n$. The received vector $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_j) \in \mathbb{F}_{q^M}^{n(j+1)}$ and $A_{[0,j]}^*$ are known and we aim to find $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j) \in \mathbb{F}_{q^M}^{n(j+1)}$ such that $\mathbf{x}^\top = A_{[0,j]}^*(\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j)^\top$ and $H_{trunc}(j)(\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_j)^\top = 0$. We will show that $\mathbf{v}_0$ is uniquely determined.

By the Gauss elimination, there exists a nonsingular block lower triangular matrix $A_{[0,j]} \in \mathbb{F}_q^{n(j+1) \times n(j+1)}$ such that the nonzero columns of $A_{[0,j]}^* A_{[0,j]}$ are linearly independent. Clearly, the number of nonzero columns of $A_{[0,j]}^* A_{[0,j]}$ is equal to $\mathrm{rank}(A_{[0,j]}^*)$. Denote $\overline{\mathbf{v}}^\top = (\overline{\mathbf{v}}_0, \ldots, \overline{\mathbf{v}}_j)^\top = A_{[0,j]}^{-1}(\mathbf{v}_0, \ldots, \mathbf{v}_j)^\top$. First we find $\overline{\mathbf{v}}$ in order to obtain $\mathbf{v}_0$. From the equation

$$\mathbf{x}^\top = A_{[0,j]}^* A_{[0,j]} \overline{\mathbf{v}}^\top$$

$\mathrm{rank}(A_{[0,j]}^*)$ coordinates of $\overline{\mathbf{v}}$ can be uniquely determined. Notice that this system of equations always admits a nonzero solution. The remaining coordinates must satisfy the system of linear equations

$$(H_{trunc}(j)A_{[0,j]})\overline{\mathbf{v}}^\top = 0.$$

The unknowns in $\overline{\mathbf{v}}$ satisfy a system of linear equations $B\mathbf{y} = c$ where $\mathbf{y}$ represents the vector of t unknowns, with $t = n(j + 1) - \mathrm{rank}(A_{[0,j]}^*)$ and $c \in \mathbb{F}_{q^M}^{k(j+1)}$ is a constant vector (notice that some of the $k(j + 1)$ equations may be trivial). This system is always soluble since $H_{trunc}(j)\mathbf{v}^\top = 0$. Now, by (14), $t < d_{\mathrm{CR}}^j(\mathcal{C})$. Let y_i be one the unknowns of $\overline{\mathbf{v}}_0$. Since $t - 1 \leq d_{\mathrm{CR}}^j(\mathcal{C}) - 2$, the column corresponding to y_i of B is not a linear combination of the other $t - 1$ columns, by theorem 3.

Therefore y_i can be uniquely determined. Since this is true for any unknown of $\overline{\mathbf{v}}_0$, we can recover the whole vector $\overline{\mathbf{v}}_0$. Finally, as $A^{-1}_{[0,j]}$ is also a nonsingular and block lower triangular matrix, $\mathbf{v}_0$ is recovered as well.

Remark 1 It is important to remark that, as opposed to the proof of Theorem 1 provided in [19], the proof of Theorem 4 is constructive in the sense that it provides a decoding procedure to recover $\mathbf{u}_0$ when the conditions of the theorem are satisfied. The next example illustrates this procedure.

Example 2 Let $H(D) = H_0 + H_1 D \in \mathbb{F}_{2^6}[D]^{1\times 3}$ be a parity-check matrix of the convolutional code $\mathcal{C} \subset \mathbb{F}_{2^6}[D]^3$, α a primitive element of $\mathbb{F}_{2^6}$ such that $\alpha^6 = \alpha + 1$ and

$$H_0 = \begin{bmatrix} \alpha^5 & \alpha^3 & \alpha^2 \end{bmatrix}, \ H_1 = \begin{bmatrix} \alpha^4 & 1 & 1 \end{bmatrix}.$$

Thus,

$$H_{trunc}(1) = \begin{bmatrix} \alpha^5 \ \alpha^3 \ \alpha^2 & 0 \ \ 0 \ \ 0 \\ \alpha^4 \ 1 \ \ 1 & \alpha^5 \ \alpha^3 \ \alpha^2 \end{bmatrix},$$

and it is straightforward to verify that $d^1_{\mathrm{CR}}(\mathcal{C}) = 2$ and therefore by Theorem 4 each truncated codeword $\mathbf{v}_{[0,1]}$ can be recovered if the rank deficiency in the time interval $[0,1]$ is not larger than 1. Consider, for instance that the channel matrix is

$$A^{\star}_{[0,1]} = \begin{bmatrix} 1 \ 0 \ 0 & 0 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 0 \ 0 \\ 0 \ 0 \ 0 & 0 \ 0 \ 0 \\ \hline 0 \ 0 \ 0 & 1 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 1 \ 0 \\ 0 \ 0 \ 1 & 0 \ 0 \ 1 \end{bmatrix},$$

and the received vector is $\mathbf{x} = (0,1,0 \mid 0, 1+\alpha^3, \alpha)$. We aim at finding $\mathbf{v}_0 \in \mathbb{F}^3_{2^6}$ with $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1) \in \mathbb{F}^6_{2^6}$ such that $\mathbf{x} = A^{\star}_{[0,1]}\mathbf{v}$ and $H_{trunc}(1)\mathbf{v}^\top = 0$. Following the notation and the steps described in the proof of Theorem 4 one obtains

$$A^{*}_{[0,1]}A_{[0,1]} = \begin{bmatrix} 1 \ 0 \ 0 & 0 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 0 \ 0 \\ 0 \ 0 \ 0 & 0 \ 0 \ 0 \\ \hline 0 \ 0 \ 0 & 1 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 1 \ 0 \\ 0 \ 0 \ 1 & 0 \ 0 \ 1 \end{bmatrix} \begin{bmatrix} 1 \ 0 \ 0 & 0 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 0 \ 0 \\ 0 \ 0 \ 1 & 0 \ 0 \ 0 \\ \hline 0 \ 0 \ 0 & 1 \ 0 \ 0 \\ 0 \ 0 \ 0 & 0 \ 1 \ 0 \\ 0 \ 0 \ 1 & 0 \ 0 \ 1 \end{bmatrix} = \begin{bmatrix} 1 \ 0 \ 0 & 0 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 0 \ 0 \\ 0 \ 0 \ 0 & 0 \ 0 \ 0 \\ \hline 0 \ 0 \ 0 & 1 \ 0 \ 0 \\ 0 \ 1 \ 0 & 0 \ 1 \ 0 \\ 0 \ 0 \ 0 & 0 \ 0 \ 1 \end{bmatrix}.$$

Since the base field has characteristic 2, $A^{-1}_{[0,1]} = A_{[0,1]}$. Define $\overline{\mathbf{v}}^\top = A^{-1}_{[0,1]}\mathbf{v}^\top$. From equation

$$\mathbf{x}^\top = A^{*}_{[0,1]}A_{[0,1]}\overline{\mathbf{v}}^\top$$

we obtain

$$\overline{\mathbf{v}}_0 = (0,1,\mathbf{y}_2),$$
$$\overline{\mathbf{v}}_1 = (0,\alpha^3,\alpha),$$

where $\mathbf{y}_2$ can be computed solving $H_{trunc}(1)A_{[0,1]}\overline{\mathbf{v}}^\top = 0$, which yields

$$\begin{cases} \alpha^3 + \mathbf{y}_2\alpha^2 = 0 \\ 1 + (1+\alpha^2)\mathbf{y}_2 + \alpha^6 + \alpha^3 = 0 \end{cases}$$

and so $\mathbf{y}_2 = \alpha$. The recovered codeword is then $\mathbf{v}^\top = A_{[0,1]}\overline{\mathbf{v}}^\top = (0,1,\alpha \mid 0, \alpha^3, 0)$.

The next corollary readily follows from Theorem 4 and the fact that $\mathcal{C}$ is a linear time-invariant convolutional code.

Corollary 1 *Assume that we have been able to correctly decode up to an instant $T_0 - 1$. Let $T_f \geq T_0$ and $T = T_f - T_0$. Then, we can recover $\mathbf{u}_{T_0}$ at time instant T_f if*

$$d_{\mathrm{CR}}^T(\mathcal{C}) > n(T + 1) - \mathrm{rank}(A_{[T_0, T_f]}^*). \tag{15}$$

Corollary 1 illustrates how the notion of j-th CRD captures the error correcting capability of a rank metric convolutional code in a network with possible delays, within a limited time interval.

4 MRD convolutional codes: A matrix characterization

A natural follow-up question is to know bounds on the j-th CRD for a given set of parameters (n, k, δ). An upper-bound on these distances is presented next.

Corollary 2 *Let $\mathcal{C}$ be an (n, k, δ) code. Then, for any $j \geq 0$,*

$$d_{\mathrm{CR}}^j(\mathcal{C}) \leq (j + 1)(n - k) + 1. \tag{16}$$

Proof Note that if $A_{[0,j]}$ has full row rank then A_{ii} is nonsingular for $i = 0, \ldots, j$ and therefore $H_0 A_{ii}$ has also full row rank since H_0 is full row rank (as $H(D)$ is basic). Thus, $H_{trunc}(j) A_{[0,j]}$ has full row rank which implies, in particular, that any of the columns of $H_{trunc}(j) A_{[0,j]}$ is linear combination of at most $(n-k)(j+1)$ columns of $H_{trunc}(j) A_{[0,j]}$ in particular, one of the first n columns. Then, the result follows from the first part of the proof $(2. \Rightarrow 1.)$ of Theorem 3.

Note that Singleton bound on classical rank metric codes still holds for non-linear codes and therefore Corollary 2 would also follow directly from this fact.

Since no column distance can achieve a value greater than the Singleton-type upper bound in (12), there must exist an integer L for which the bound (16) could be attained for all $j \leq L$ and it is a strict upper bound for $j > L$. It is a matter of straightforward computations to verify that this value is (see [13] for more details):

$$L = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n - k} \right\rfloor. \tag{17}$$

An (n, k, δ)-convolutional code $\mathcal{C}$ with every $d_{\mathrm{CR}}^j(\mathcal{C})$ maximal, *i.e.*, $d_{\mathrm{CR}}^j(\mathcal{C}) = (n - k)(j + 1) + 1$, for each $j \leq L$, is called a *Maximum Rank Profile* (MRP) code. The j-th column rank distances of MRP codes increase as rapidly as possible for as long as possible. Fast growth of the column distances is an important property for codes to be used with sequential decoding since they have the potential to correct a maximal number of errors per time interval.

Another interesting property of MRP codes is that Theorem 4 can be strengthened when these codes are used. Suppose that all the packets have been recovered up to an instant $T_0 - 1$. In order to recover $\mathbf{u}_{T_0}$ as soon as possible, the condition of Corollary 1 is checked sequentially at $T_0, T_0 + 1, \ldots$ up to an instant, say T_f,

when the condition is finally satisfied. Then, by Corollary 1, $\mathbf{u}_{T_0}$ can be recovered. However, if MRP codes are employed in the transmission not only $\mathbf{u}_{T_0}$ but the complete vector $(\mathbf{u}_{T_0}, \ldots, \mathbf{u}_{T_f})$ can be actually recovered at time instant T_f. This will be properly stated and proved next.

First we show that, as it happens with the column (Hamming) distance, the maximality of the j-th column rank distance implies maximality of the previous i-th column rank distances, $i = 0, 1, \ldots, j - 1$.

Lemma 5 *If $d_{CR}^j(\mathcal{C}) = (j+1)(n-k) + 1$, then $d_{CR}^i(\mathcal{C}) = (i+1)(n-k) + 1$ for $i = 0, 1, \ldots, j - 1$.*

Proof It is proved by contradiction. It is enough to prove it for $i = j - 1$. Then, assume that $d_{\mathrm{CR}}^{j-1}(\mathcal{C}) \leq j(n-k)$, that is, there exists an $A_{[0,j-1]}$ as in (13) such that one of the first n columns of $H_{trunc}(j-1)A_{[0,j-1]}$ is in the span of some other $(n-k)j - 1$ columns. Note that for all $A_{[0,j]}$ it holds that

$$H_{trunc}(j)A_{[0,j]} = \left[\begin{array}{c|c} H_{trunc}(j-1)A_{[0,j-1]} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \widetilde{H}_j \; \widetilde{H}_{j-1} \; \cdots \; \widetilde{H}_1 & H_0 A_{jj} \end{array} \right],$$

where $\widetilde{H}_s = \sum_{t=0}^s H_{s-t} A_{t+j-s \; j-s}$. As A_{jj} is full row rank, then so it is $H_0 A_{jj}$ and therefore one of the first n columns of $H_{trunc}(j)A_{[0,j]}$ is in the span of some other $(n-k)j - 1 + n - k = (n-k)(j+1) - 1$ columns of $H_{trunc}(j)A_{[0,j]}$, which, by Theorem 3, implies that $d_{\mathrm{CR}}^j(\mathcal{C}) \leq (j+1)(n-k)$.

A characterization of MRP codes in terms of the full size minors of $G_{trunc}(j)A_{[0,j]}^\top$ is given next.

Theorem 6 *Let $\mathcal{C}$ be an (n, k, δ)-convolutional code with generator matrix $G(D)$. The following statements are equivalent:*

1. *$d_{\mathrm{CR}}^j(\mathcal{C}) = d_{\mathrm{rank}}(\mathcal{C}_j^c) = (j+1)(n-k) + 1$;*
2. *every $(j+1)k \times (j+1)k$ full-size minor of $G_{trunc}(j)A_{[0,j]}^\top$ formed by the columns with indices $1 \leq c_1 < \cdots < c_{(j+1)k}$ where $c_{sk+1} > sn$ for $s = 1, \ldots, j$, is nonzero, where $A_{[0,j]}$ as in (13).*

In particular, $\mathcal{C}$ is an MRP convolutional code if $d_{\mathrm{CR}}^L(\mathcal{C}) = d_{\mathrm{rank}}(\mathcal{C}_L^c) = (L+1)(n-k) + 1$, with L as defined in (17).

Proof (1⇒2): Suppose that 2. does not hold and $A_{[0,j]}$ as in (13) is such that $G_{trunc}(j)A_{[0,j]}^\top$ contains a singular square submatrix $\widetilde{G}$ of size $(j+1)k$ formed by column of $G_{trunc}(j)A_{[0,j]}^\top$ with indices $c_1 < c_2 < \cdots < c_{(j+1)k}$, $c_{sk+1} > sn$ for $s = 1, \ldots, j$. Take $(\mathbf{u}_0, \ldots, \mathbf{u}_j)$ such that $(\mathbf{u}_0, \ldots, \mathbf{u}_j)\widetilde{G} = 0$. Then,

$$\mathrm{rank}(\phi((\mathbf{u}_0, \ldots, \mathbf{u}_j)G_{trunc}(j)A_{[0,j]}^\top)) < (n-k)(j+1) + 1.$$

If $\mathbf{u}_0 \neq 0$ then $d_{\mathrm{CR}}^j(\mathcal{C}) < (j+1)(n-k) + 1$ contradicting 1.

Assume then that $\mathbf{u}_0 = 0$ and consider $i > 0$ such that $\mathbf{u}_i \neq 0$ and $\mathbf{u}_\ell = 0$ for $0 \leq \ell < i \leq j$. The last $n(j-i)$ rows and columns of $A_{[0,j]}^\top$ is again a nonsingular upper block triangular matrix. Then, due to the Toeplitz structure of $G_{trunc}(j)$ and that $c_{sk+1} > sn$ for $s = 1, 2, \ldots, j$, it holds that $(\mathbf{u}_i, \ldots, \mathbf{u}_j)G_{trunc}(j-i)A_{[i,j]}^\top$ has weight (and therefore rank) strictly less than $(j-i+1)(n-k)+1$. This implies that $d_{\mathrm{CR}}^{j-i}(\mathcal{C}) \leq (j-i+1)(n-k)$ and by Lemma 5 $d_{\mathrm{CR}}^j(\mathcal{C}) \leq (j+1)(n-k)$.

$(2 \Rightarrow 1)$: Again it is done by contradiction. Suppose that $d_{SR}^j(\mathcal{C}) \leq (j+1)(n-k)$ and let ℓ be the first instant such that the upper bound (16) is not achieved. Hence, there exists a $(\mathbf{u}_0, \ldots, \mathbf{u}_\ell)$, $\mathbf{u}_0 \neq 0$ and an $\bar{A}_{[0,\ell]}$ as in (13) such that $(\mathbf{v}_0, \ldots, \mathbf{v}_\ell) = (\mathbf{u}_0, \ldots, \mathbf{u}_\ell)G_{trunc}(\ell)\bar{A}_{[0,\ell]}^\top$ with

$$\mathrm{rank}(\phi((\mathbf{v}_0, \ldots, \mathbf{v}_\ell))) \leq (n-k)(\ell+1).$$

This means that there exists $\widehat{A}_{[0,\ell]}$ as in (13) such that $(\mathbf{v}_0, \ldots, \mathbf{v}_\ell)\widehat{A}_{[0,\ell]}^\top$ has at least $k(\ell+1)$ zeros.

Let $A_{[0,\ell]} = \widehat{A}_{[0,\ell]}\bar{A}_{[0,\ell]}$ and build a square matrix L by selecting the columns of $G_{trunc}(j)A_{[0,j]}^\top$ where $(\mathbf{u}_0, \ldots, \mathbf{u}_\ell)G_{trunc}(j)A_{[0,j]}^\top$ is zero and denote these indices by $c_1, \ldots, c_{(\ell+1)k}$.

At this point we observe that as $d_{\mathrm{CR}}^i(\mathcal{C}) = (i+1)(n-k)+1$, for $i = 1, 2, \ldots, \ell-1$, it follows that $(\mathbf{u}_0, \ldots, \mathbf{u}_i)G_{trunc}(i)A_{[0,i]}^\top$ has at most $k(i+1)-1$ zeros for $i = 0, 1, \ldots, \ell-1$. This, in particular, implies that $c_{sk+1} > sn$ for $s = 1, 2, \ldots, \ell$. Extend the square matrix L by taking the matrix formed by the last $k(j-\ell)$ columns of $G_{trunc}(j)A_{[0,i]}^\top$, which we denote by J, and build a $k(j+1) \times k(j+1)$ submatrix of $G_{trunc}(j)A_{[0,j]}^\top$ of the form

$$\left[\begin{array}{c|c} L \\ \hline 0 \end{array}\,\middle|\, J \right].$$

We call the indices of the selected columns $c_1, \ldots, c_{(j+1)k}$. Clearly, the indices satisfy $c_{sk+1} > sn$, for $s = 1, 2, \ldots, j$ and has determinant equal to zero as L is singular by construction. This means that **2.** does not hold which concludes the proof. The last statement readily follows from Lemma 5.

We are now able to prove a stronger version of Theorem 4 and Corollary 1 if an MRP code is employed. In this case, when condition (14) is satisfied at time T_f then we can recover the whole vector $(\mathbf{u}_{T_0}, \ldots, \mathbf{u}_{T_f})$. This property is particularly useful when erased packets must be recovered within tight delay constrains as it happens in many streaming applications.

Theorem 7 *Let $\mathcal{C} = \mathrm{im}\, G(D)$ be an MRP code. Assume that we have been able to recover all packets up to an instant T_0 and that T_f is the first instant such that the condition (15) is satisfied. Then, the whole vector $(\mathbf{u}_{T_0}, \ldots, \mathbf{u}_{T_f})$ can be recovered.*

Proof Denote $T = T_f - T_0$. Let $A_{[T_0, T_0+j]}^\star$ be the truncated channel matrix, $0 \leq j \leq T$. Trying to recover $\mathbf{u}_{T_0}$ unsuccessfully up to $T_f - 1$ using an MRP code is equivalent to

$$\mathrm{rank}(A_{[T_0, T_0+j]}^\star) \leq n(j+1) - d_{\mathrm{CR}}^j(\mathcal{C}) = k(j+1) - 1$$

for $0 \leq j < T - 1$. Since condition (15) is satisfied at T_f,

$$\mathrm{rank}(A_{[T_0, T_f]}^\star) \geq k(T+1).$$

By Corollary 1 $\mathbf{u}_{T_0}$ (and obviously also $\mathbf{v}_{T_0}$) can be recovered. Now, shift one time instant forward and consider the time window $[T_0 + 1, T_f]$ for decoding. As $\text{rank}(A^\star_{[T_0, T_0]}) \leq k - 1$, it follows from the lower block triangular configuration of $A^\star_{[T_0, T_f]}$ that

$$\text{rank}(A^*_{[T_0+1, T_f]}) \geq kT,$$

and thus again Corollary 1 guarantees that $\mathbf{u}_{T_0+1}$ can be computed. The same procedure is applied sequentially to retrieve the remaining $\mathbf{u}_{T_0+j}$, $j = 2, 3, \ldots, T$.

5 Superregular matrices to build MRP codes

In this section we introduce a class of matrices that will be essential for the construction of convolutional codes that possess optimal rank distance properties: Superregular matrices. We will first explain the relation between superregular matrices and MRP. We then present a concrete class of superregular matrices that can be used to produce MRP codes. This, in particular, proves that MRP codes exist and the bounds given above are optimal for this new rank metric.

5.1 Block Toeplitz Superregular matrices

Superregular matrices are important in coding theory as they can be used to construct codes with optimal Hamming distance. Roughly speaking, this is due to the fact that a full row rank superregular matrix has the following property: Take any one of its rows with Hamming weight, say d. Then, any combination of this row with t other rows yields a vector of Hamming weight $\geq d - t$. In this paper we will show that a particular class of superregular matrices can be also used to build convolutional codes with optimal rank distance. Next, we formally introduce the notion of superregular matrix.

Let $F = [\mu_{ij}]$ be a square matrix of order m over $\mathbb{F}_{q^M}$ and S_m the symmetric group of order m. Recall that the determinant of F is given by

$$|F| = \sum_{\sigma \in S_m} \text{sgn}(\sigma) \mu_{1\sigma(1)} \cdots \mu_{m\sigma(m)}, \tag{18}$$

where the sign of the permutation σ, $\text{sgn}(\sigma)$, is 1 (resp. -1) according as if σ can be written as product of an even (resp. odd) number of transpositions. A *trivial term* of the determinant is a term of (18), $\mu_{1\sigma(1)} \cdots \mu_{m\sigma(m)}$, equal to zero. If F is a square submatrix of a matrix B, with entries in $\mathbb{F}_{q^M}$, and all the terms of the determinant of F are trivial we say that $|F|$ is a *trivial minor* of B. We say that B is *superregular* if all its non-trivial minors are different from zero.

Notice that any full size minor of $G_{trunc}(j) A^\top_{[0,j]}$ formed by more than sk columns in the first sn columns, with $s \leq j$, is trivially zero. Indeed, if all the entries of $G_{trunc}(j)$ are nonzero, then, it holds by Theorem 6 that the maximum j-th column rank distance is achieved if and only if all the nontrivial full size minors of $G_{trunc}(j) A^\top_{[0,j]}$ are nonzero. This leads immediately to a sufficient condition to obtain MRP codes.

Theorem 8 *Let* $\mathcal{C} = \mathrm{im}\, G(D) \subset \mathbb{F}_{q^M}^n$ *be an* (n, k, δ)-*convolutional code with* L *as in* *(17) and the entries of the coefficients* G_i *of* $G(D)$ *all nonzero. If* $G_{trunc}(L)A_{[0,L]}^\top$ *is* *superregular for* $A_{[0,L]}^\top$ *as in (13), with* $j = L$, *then* $\mathcal{C}$ *is MRP.*

Proof If the entries of the coefficients G_i of $G(D)$ are all nonzero, then the conditions in Theorem 6 on the column indices amounts to consider the nontrivial full size minors of $G_{trunc}(j)A_{[0,j]}^\top$. By Lemma 5 it is enough to consider $j = L$. Finally, the nontrivial full size minors of $G_{trunc}(L)A_{[0,j]}^\top$ are nonzero as $G_{trunc}(L)A_{[0,j]}^\top$ is assumed to be superregular. $\qquad\square$

The main goal of this section is to build a $G(D)$ that satisfies conditions of Theorem 8. To this end we will first recall a known $G(D)$ having the property that $G_{trunc}(L)$ is superregular and then show that $G_{trunc}(L)A_{[0,j]}^\top$ remains superregular. This last proof is not straightforward and requires the recall and modification of several previous results.

For the purposes of this paper we propose below a type of superregular matrices. This class of matrices were first introduced in the context of (Hamming) convolutional codes with $q = 2$ in [2] and will bring about a new class of convolutional codes with optimal rank distances. Recall that the extension field $\mathbb{F}_{q^M}$ is isomorphic to the M-dimensional vector space $\mathbb{F}_q^M$ over the base field $\mathbb{F}_q$. A basis of this vector space, say $\alpha_1, \alpha_2, \ldots, \alpha_M \in \mathbb{F}_{q^M}$, is said to be *normal* when there exists $\alpha \in \mathbb{F}_{q^M}$ such that $\alpha_i = \alpha^{[i]} = \alpha^{q^i}$, for all $1 \leq i \leq M$. If so, α is called a *normal element* of $\mathbb{F}_{q^M}$ and there is always at least one element that is both normal and primitive (see [11]). Hence, every element φ of $\mathbb{F}_{q^M}$ can be written as a linear combination in $\mathbb{F}_q$ of the elements of the basis, *i.e.*,

$$\varphi = \sum_{i=1}^M \varphi_i \alpha_i = \sum_{i=1}^M \varphi_i \alpha^{[i]}. \tag{19}$$

We shall make extensive use of linearized polynomials which are polynomials with its monomials terms having a Frobenius power. Hence, if $\varphi_i \mathbb{F}_q$ are the coordinates of the element φ of $\mathbb{F}_{q^M}$, as in 19, then the linearized polynomial $f_\varphi(x) = \sum_{i=1}^M \varphi_i x^{[i]}$ satisfies $f_\varphi(\alpha) = \varphi$.

The following result follows straightforward from the *Freshman's Rule*.

Lemma 9 *Let* u *be a positive integer and* $f(x) = \sum_{i=1}^M a_i x^{[i]}$ *be a linearized polynomial over* $\mathbb{F}_{q^M}$, *then*

$$f^{[u]}(x) = \sum_{i=1}^M a_i x^{[i+u]}.$$

Now, let α be a primitive and normal element of a finite field $\mathbb{F}_{q^M}$ with q^M elements and consider $G(D) = \sum_{i=0}^\nu G_i D^i \in \mathbb{F}_{q^M}^{k \times n}(D)$ where each G_i, for $0 \leq i \leq \nu$, is the $k \times n$ matrix

$$\begin{bmatrix} \alpha^{[ni]} & \alpha^{[ni+1]} & \cdots & \alpha^{[n(i+1)-1]} \\ \alpha^{[ni+1]} & \alpha^{[ni+2]} & \cdots & \alpha^{[n(i+1)]} \\ \alpha^{[ni+2]} & \alpha^{[ni+3]} & \cdots & \alpha^{[n(i+1)+1]} \\ \vdots & \vdots & \vdots & \\ \alpha^{[ni+1+k-1]} & \alpha^{[ni]} & \cdots & \alpha^{[n(i+1)+k-2]} \end{bmatrix}, \tag{20}$$

and where again we use the notation $\alpha^{[j]} = \alpha^{q^j}$ to denote the j-th Frobenius power of $\alpha \in \mathbb{F}_{q^M}$.

Note that the entries of (20) are the values of linearized monomials evaluated at α but the determinant is not necessarily the value of a linearized polynomial evaluated at α. Nevertheless, the next result states that if M is sufficiently large then $G_{trunc}(j)$ (defined in (5)) is a superregular matrix for all $j \leq \lfloor \frac{\delta}{k} \rfloor + \lfloor \frac{\delta}{n-k} \rfloor$.

Theorem 10 *Let* $L = \lfloor \frac{\delta}{k} \rfloor + \lfloor \frac{\delta}{n-k} \rfloor, n \in \mathbb{N}$, α *be a primitive and normal element of a finite field* $\mathbb{F}_{q^M}$ *of characteristic* p *and consider* $G_{trunc}(L) \in \mathbb{F}_{q^M}^{(L+1)k \times (L+1)n}$ *as in (5), with* G_i *submatrices of (20). If* $M \geq q^{n(L+2)-1}$ *then the matrix* $G_{trunc}(L)$ *is superregular (over* $\mathbb{F}_{q^M}$).

Proof By observing that permutation of columns does not affect superregularity and using [2, Theorem 3.2] the theorem immediately follows for a prime q. The extension of that theorem obtained by R. Mahmood in [17] for any $q = p^r$, where p is prime and r is a positive integer, concludes the proof.

We are now going to obtain a result that extends this theorem and states that the superregularity of the matrices $G_{trunc}(j)$ built from block matrices of the form (20), remains invariant under right multiplications of $A_{[0,j]}^{\top}$.

In [19], the authors proved the superregularity of $F = T A_{[0,m]}$ with

$$T = \begin{bmatrix} & & T_0 \\ & T_0 & T_1 \\ \ddots & \vdots & \vdots \\ T_0 & \cdots & T_{m-1} & T_m \end{bmatrix} \quad \text{and} \quad A_{[0,m]} = \begin{bmatrix} A_0 & & & \\ & A_1 & & \\ & & \ddots & \\ & & & A_m \end{bmatrix}, \qquad (21)$$

where the matrices T_i are given by

$$\begin{bmatrix} \alpha^{[ni]} & \alpha^{[ni+1]} & \cdots & \alpha^{[n(i+1)-1]} \\ \alpha^{[ni+1]} & \alpha^{[ni+2]} & \cdots & \alpha^{[n(i+1)]} \\ \alpha^{[ni+2]} & \alpha^{[ni+3]} & \cdots & \alpha^{[n(i+1)+1]} \\ \vdots & \vdots & \vdots & \\ \alpha^{[n(i+1)-1]} & \alpha^{[n(i+1)]} & \cdots & \alpha^{[n(i+2)-2]} \end{bmatrix}, \qquad (22)$$

and A_i are invertible matrices, for $0 \leq i \leq m$. To obtain their proof, they first showed the following properties:

1. Let $F_{ij} = T_i A_j$, for $0 \leq i \leq j \leq m$. Then

$$F = \begin{bmatrix} & & & F_{0m} \\ & & F_{0(m-1)} & F_{1m} \\ \ddots & \vdots & \vdots \\ F_{00} & \cdots & F_{(m-1)(m-1)} & F_{mm} \end{bmatrix}$$

and

$$F_{ij} = \begin{bmatrix} f_{0j}^{[ni]} & f_{1j}^{[ni]} & \cdots & f_{n-1\,j}^{[ni]} \\ f_{0j}^{[ni+1]} & f_{1j}^{[ni+1]} & \cdots & f_{n-1\,j}^{[ni+1]} \\ \vdots & \vdots & \vdots & \\ f_{0j}^{[n(i+1)-1]} & f_{1j}^{[n(i+1)-1]} & \cdots & f_{n-1\,j}^{[n(i+1)-1]} \end{bmatrix}, \qquad (23)$$

where for each $0 \le r \le j \le n - 1$, f_{rj} is the value of a linearized polynomial evaluated at α and

(a) $(f_{0j}, f_{1j}, \ldots, f_{n-1\,j})$ are linearly independent over $\mathbb{F}_q$;

(b) If $s' < s$, the entries of a row s of F_{ij} are q-powers of the corresponding entries in row s', by the Freshman's Rule.

(c) If $i' < i$ the entries of a row s of F_{ij} are q-powers of the corresponding entries in row s of $F_{i'j}$, by the Freshman's Rule.

2. The q-degrees of the polynomials associated to the elements of F_{ij} strictly increase downwards on any fixed column, for any $0 \le i \le j \le m$.

3. If $i < i'$ and $s \ge i$ then the q-degree of the polynomial associated to any entry of the row r of F_{is} is smaller than the q-degree of of the polynomial associated to any entry of the row r of $F_{i'(s+i'-i)}$, for any $0 \le r \le n - 1$ (notice that the block matrices F_{is} and $F_{i'(s+i'-i)}$ are in the same block row of F).

4. If $i < i' \le j$ then the q-degree of the polynomial associated to any entry of the column s of F_{ij} is smaller than the q-degree of the polynomial associated to any entry of the column s of $F_{i'j}$, for any $0 \le s \le n - 1$.

5. If D is a square submatrix of F there exists an invertible matrix M such that

(a) the q-degrees of the polynomials associated with the entries of any row of DM are strictly increasing,

(b) the q-degrees of the polynomials associated with the entries of any column of DM are strictly increasing.

Next they used the proof of Theorem 3.2 of [2], where it is shown, when q is a prime number, that if a matrix B satisfies properties 5 (a) and 5 (b) above then B is invertible. This result is also valid for any $q = p^r$, with p prime, as it was shown in [17].

Therefore, DM is invertible and since M is invertible also D is invertible. Therefore, F is superregular.

Based on these properties the next theorem was derived in [19].

Theorem 11 *[19, Theorem 5] For any $0 \le t \le m$ let $A_t \in \mathbb{F}_q^{n \times n}$ be non-singular matrices and let $A_{[0,m]} = diag(A_0, \ldots, A_m)$. Let $T \in \mathbb{F}_{q^M}^{n(m+1) \times n(m+1)}$ be of the form*

$$\begin{bmatrix} & & & T_0 \\ & \cdot^{\cdot^{\cdot}} & & \vdots \\ T_0 & \cdots & T_{m-1} \\ T_0 & T_1 & \cdots & T_m \end{bmatrix}.$$

where T_i, for each $0 \le i \le m$, is given by (20).

If $M > q^{n(m+2)-1}$ and α is a primitive and normal element of $\mathbb{F}_{q^M}$, then $F = TA_{[0,m]}$ is superregular.

In our case instead of having a block diagonal matrix we have a block lower triangle matrix $A_{[0,L]}$, but the corresponding versions of all the above properties are still satisfied.

Theorem 12 *Let $\mathcal{C}$ be an (n, k, δ)-convolutional code with generator matrix $G(D) = \sum_{i=0}^{\nu} G_i D^i$ where each G_i is a $k \times n$ matrix given by (20). Let $M \ge q^{n(L+2)-1}$ and let*

α be a primitive and normal element of F_{q^M}. Furthermore, let $A_{[0,L]}$ be nonsingular as described in (13), with $j = L$. Then

$$S = G_{trunc}(L)A_{[0,L]}^\top \text{ is a superregular matrix,}$$

and therefore

$$\mathcal{C} \text{ is an MRP convolutional code.}$$

Proof We have,

$$S = G_{trunc}(L)A_{[0,L]}^\top = \begin{bmatrix} G_0 & G_1 & \cdots & G_L \\ & G_0 & \cdots & G_{L-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{bmatrix} \begin{bmatrix} A_{00} & A_{01} & \cdots & A_{0L} \\ & A_{11} & \cdots & A_{1L} \\ & & \ddots & \vdots \\ & & & A_{LL} \end{bmatrix}$$

$$= \begin{bmatrix} S_{00} & S_{01} & \cdots & S_{0L} \\ & S_{11} & \cdots & S_{1L} \\ & & \ddots & \vdots \\ & & & S_{LL} \end{bmatrix},$$

where

$$S_{ij} = \sum_{\ell=0}^{j-i} G_\ell A_{i+\ell\, j}.$$

Let P be the permutation matrix corresponding to the permutation σ defined by

$$\sigma = \begin{pmatrix} 1 & \cdots & k & k+1 & \cdots 2k & \cdots kL+1 & \cdots & k(L+1) \\ kL+1 & \cdots & k(L+1) & k(L-1)+1 & \cdots kL & \cdots 1 & \cdots & k \end{pmatrix}.$$

Then S is superregular if and only if PS is superregular. Define

$$F = PS = \begin{bmatrix} & & & F_{0L} \\ & & F_{0(L-1)} & F_{1L} \\ & \ddots & \vdots & \vdots \\ F_{00} & \cdots & F_{(L-1)(L-1)} & F_{LL} \end{bmatrix}$$

where $F_{ij} = S_{j-i\, j}$.

Each element of F_{ij} is the value of a linearized polynomial evaluated at α, moreover, if the first row of F_{ij} is $(f_{ij0}, \ldots, f_{ij(n-1)})$ then, for $s \leq n$, the s-th row is $(f_{ij0}^{[s]}, \ldots, f_{ij(n-1)}^{[s]})$. So we obtain similar properties to 1. b) above. The corresponding propriety 1. c) is not valid, though. In our case, if $i' < i$ the entries of a row s of F_{ij} are not q-powers of the corresponding entries in row s of $F_{i'j}$ since our F_{ij} instead of being just the product of two matrices, are the sum of products of matrices, so for $i \neq i'$, the linearized polynomials from which we construct the entries of F_{ij} may be completely different from the ones we use to construct the entries of $F_{i'j}$. That is also the reason why the exponents of the linearized

polynomials used to obtain the entries of F_{ij} don't depend on i, contrary to what happened above in (23). Hence, we have

$$
F_{ij} = \begin{bmatrix}
f_{ij0}^{[0]} & f_{ij1}^{[0]} & \cdots & f_{ij(n-1)}^{[0]} \\
f_{ij0}^{[1]} & f_{ij1}^{[1]} & \cdots & f_{ij(n-1)}^{[1]} \\
f_{ij0}^{[2]} & f_{ij1}^{[2]} & \cdots & f_{ij(n-1)}^{[2]} \\
\vdots & \vdots & \vdots & \\
f_{ij0}^{[n-1]} & f_{ij1}^{[n-1]} & \cdots & f_{ij(n-1)}^{[n-1]}
\end{bmatrix}.
\tag{24}
$$

As α is normal over $\mathbb{F}_q$ and each A_{ii} is invertible, for $1 \le i \le n$, then it holds that $(f_{ij0}, f_{ij1}, \ldots, f_{ij(n-1)})$ are linearly independent over $\mathbb{F}_q$, which corresponds to property 1. a) above. Clearly, The q-degrees of the polynomials associated to the elements of F_{ij} strictly increase downwards on any fixed column, for any $0 \le i \le j \le L$, so we get property 2.. It is also easy to see that the proprieties corresponding to the properties 3. 4. and 5. above are also satisfied. Hence F is superregular. The last statement is Theorem 8.

The following example illustrates the proprieties mentioned above.

Example 3 Let $k = 3, n = 4, q = 2$ and $\delta = 1$. Let $M = 2048$ and α a primitive and normal element of $\mathbb{F}_{2^M}$. Then $L = 1$ and

$$
G_{trunc}(1) = \begin{bmatrix}
\alpha^{[0]} & \alpha^{[1]} & \alpha^{[2]} & \alpha^{[3]} & \alpha^{[4]} & \alpha^{[5]} & \alpha^{[6]} & \alpha^{[7]} \\
\alpha^{[1]} & \alpha^{[2]} & \alpha^{[3]} & \alpha^{[4]} & \alpha^{[5]} & \alpha^{[6]} & \alpha^{[7]} & \alpha^{[8]} \\
\alpha^{[2]} & \alpha^{[3]} & \alpha^{[4]} & \alpha^{[5]} & \alpha^{[6]} & \alpha^{[7]} & \alpha^{[8]} & \alpha^{[9]} \\
0 & 0 & 0 & 0 & \alpha^{[0]} & \alpha^{[1]} & \alpha^{[2]} & \alpha^{[3]} \\
0 & 0 & 0 & 0 & \alpha^{[1]} & \alpha^{[2]} & \alpha^{[3]} & \alpha^{[4]} \\
0 & 0 & 0 & 0 & \alpha^{[2]} & \alpha^{[3]} & \alpha^{[4]} & \alpha^{[5]}
\end{bmatrix}.
$$

Take

$$
A_{[0,1]} = \begin{bmatrix}
0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}.
$$

Then, taking

$$
P = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}.
$$

we obtain $F = PS = PG_{trunc}(1)A_{[0,1]}^{\top} =$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \alpha^{[0]} + \alpha^{[1]} & \alpha^{[1]} + \alpha^{[2]} + \alpha^{[3]} & \alpha^{[1]} & \alpha^{[0]} + \alpha^{[2]} \\ 0 & 0 & 0 & 0 & \alpha^{[1]} + \alpha^{[2]} & \alpha^{[2]} + \alpha^{[3]} + \alpha^{[4]} & \alpha^{[2]} & \alpha^{[1]} + \alpha^{[3]} \\ 0 & 0 & 0 & 0 & \alpha^{[2]} + \alpha^{[3]} & \alpha^{[3]} + \alpha^{[4]} + \alpha^{[5]} & \alpha^{[2]} & \alpha^{[2]} + \alpha^{[4]} \\ \alpha^{[1]} + \alpha^{[2]} & \alpha^{[0]} & \alpha^{[0]} + \alpha^{[2]} & \alpha^{[3]} & \alpha^{[0]} + \alpha^{[2]} + \alpha^{[4]} + \alpha^{[5]} & \alpha^{[1]} + \alpha^{[2]} + \alpha^{[5]} + \alpha^{[6]} + \alpha^{[7]} & \alpha^{[0]} + \alpha^{[1]} + \alpha^{[5]} & \alpha^{[3]} + \alpha^{[4]} + \alpha^{[6]} \\ \alpha^{[2]} + \alpha^{[3]} & \alpha^{[1]} & \alpha^{[1]} + \alpha^{[3]} & \alpha^{[4]} & \alpha^{[1]} + \alpha^{[3]} + \alpha^{[5]} + \alpha^{[6]} & \alpha^{[2]} + \alpha^{[3]} + \alpha^{[6]} + \alpha^{[7]} + \alpha^{[8]} & \alpha^{[1]} + \alpha^{[2]} + \alpha^{[6]} & \alpha^{[4]} + \alpha^{[5]} + \alpha^{[7]} \\ \alpha^{[3]} + \alpha^{[4]} & \alpha^{[2]} + \alpha^{[4]} & \alpha^{[5]} & \alpha^{[2]} + \alpha^{[4]} + \alpha^{[6]} + \alpha^{[7]} & \alpha^{[3]} + \alpha^{[4]} + \alpha^{[7]} + \alpha^{[8]} + \alpha^{[9]} & \alpha^{[2]} + \alpha^{[3]} + \alpha^{[7]} & \alpha^{[5]} + \alpha^{[6]} + \alpha^{[8]} \end{bmatrix}.$$

Now, the above proprieties can easily be checked.

6 Conclusions

We have studied rank metric convolutional codes and propose a novel metric suitable for networks with delays. We have fully characterized and built optimal rank convolutional codes with respect to this metric and therefore extended previous research in the area of multi-shot network coding. The results were investigated for a given rate and degree allowing the field size unrestricted. Although we need an enormous field size in the Theorem 12, these type of constructions can be used to obtain superregular matrices in fields much smaller, but in that case these matrices have to be checked individually for superregularity. This approach was already explored in [19] and [1] and is one avenue of research we are interested to investigate. Another important issue that remains open is to provide not only sufficient (as given in Theorem 8) but also necessary conditions for a given convolutional code to be MRP in terms of superregular matrices. We conjecture that this is possible and would require the superregularity of a smaller matrix than $G_{trunc}(L)$ and $H_{trunc}(L)$ which would allow to build MRP over much smaller fields. These issues require further research.

References

1. P. Almeida, U. Martínez-Peñas, and D. Napp. Systematic maximum sum rank codes. *Finite Fields and Their Applications*, 65:101677, August 2020.
2. P. Almeida, D. Napp, and R. Pinto. A new class of superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 439(7):2145–2157, 2013.
3. P. Almeida, D. Napp, and R. Pinto. Superregular matrices and applications to convolutional codes. *Linear Algebra and its Applications*, 499:1–25, 2016.
4. P. Almeida, D. Napp, and R. Pinto. MDS 2D convolutional codes with optimal 1D horizontal projections. *Designs, Codes and Cryptography*, 86(2):285–302, 2018.
5. A. Badr, A. Khisti, Wai-Tian. Tan, and J. Apostolopoulos. Layered constructions for low-delay streaming codes. *IEEE Trans. Inform. Theory*, 63(1):111–141, 2017.

6. Á. Barbero and Ø. Ytrehus. Rate $(n-1)/n$ systematic memory maximum distance separable convolutional codes. *IEEE Trans. Inform. Theory*, 64(4):3018–3030, 2018.
7. J. Rosenthal D. Napp, R. Pinto and P. Vettori. MRD rank metric convolutional codes. *IEEE International Symposium on Information Theory (ISIT) 2017*, pages 2766–2770, 2017.
8. Ph. Delsarte. Bilinear forms over a finite field, with applications to coding theory. *Journal of Combinatorial Theory, Series A*, 25(3):226–241, 1978.
9. È.M. Gabidulin. Theory of codes with maximum rank distance. *Probl. Inf. Transm.*, 21:1–12, 1985.
10. H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache. Strongly MDS convolutional codes. *IEEE Trans. Inform. Theory*, 52(2):584–598, 2006.
11. Jr. H. W. Lenstra and R. J. Schoof. Primitive normal bases for finite fields. *Mathematics of Computation*, 48(177):217–231, 1987.
12. J. Hansen, J. Østergaard, J. Kudahl, and J. H. Madsen. Superregular lower triangular Toeplitz matrices for low delay wireless streaming. *IEEE Transactions on Communications*, 65(9):4027–4038, 2017.
13. R. Hutchinson, J. Rosenthal, and R. Smarandache. Convolutional codes with maximum distance profile. *Systems & Control Letters*, 51(1):53–63, 2005.
14. R. Hutchinson, R. Smarandache, and J. Trumpf. On superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 428:2585–2596, 2008.
15. R. Johannesson and K. Sh. Zigangirov. *Fundamentals of Convolutional Coding*. IEEE Press, New York, 2015.
16. R. Kötter and F.R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Trans. Inform. Theory*, 54(8):3579–3591, Aug 2008.
17. R. Mahmood. Rank metric convolutional codes with applications in network streaming. Master of applied science, Graduate Department of Electrical and Computer Engineering, University of Toronto, 2015.
18. R. Mahmood, A. Badr, and A. Khisti. Streaming-codes for multicast over burst erasure channels. *IEEE Trans. Inform. Theory*, 61(8):4181–4208, 2015.
19. R. Mahmood, A. Badr, and A. Khisti. Convolutional codes with maximum column sum rank for network streaming. *IEEE Trans. Inform. Theory*, 62(6):3039–3052, 2016.
20. U. Martínez-Peñas. On the similarities between generalized rank and hamming weights and their applications to network coding. *IEEE Transactions on Information Theory*, 62(7):4081–4095, July 2016.
21. U. Martínez-Peñas. Theory of supports for linear codes endowed with the sum-rank metric. *Designs, Codes and Cryptography*, 87(10):2295 –2320, 2019.
22. R. J. McEliece. The algebraic theory of convolutional codes. In *Handbook of Coding Theory*, volume 1, pages 1065–1138. Elsevier Science Publishers, 1998.
23. D. Napp, R. Pinto, J. Rosenthal, and P. Vettori. Rank metric convolutional codes. *Proceedings of the 22nd International Symposium on Mathematical Theory of Network and Systems (MTNS), Minnesota, USA*, 2016.
24. D. Napp, R. Pinto, and V. R. Sidorenko. Concatenation of convolutional codes and rank metric codes for multi-shot network coding. *Designs, Codes and Cryptography*, 86(2):303–318, 2018.
25. D. Napp, R. Pinto, and M. Toste. On MDS convolutional codes over $\mathbb{Z}_{p^r}$. *Designs, Codes and Cryptography*, 83:101–114, 2017.
26. A. Neri, A. Horlemann-Trautmann, T. Randrianarisoa, and J. Rosenthal. On the genericity of maximum rank distance and Gabidulin codes. *Designs, Codes and Cryptography*, 86:341 –363, 2018.
27. R.W. Nóbrega and B.F. Uchoa-Filho. Multishot codes for network coding using rank-metric codes. In *Wireless Network Coding Conference (WiNC), 2010 IEEE*, pages 1–6, June 2010.
28. F. Pollara, R. J. McEliece, and K. Abdel-Ghaffar. Finite-state codes. *IEEE Transactions on Information Theory*, 34(5):1083 –1089, 1988.
29. J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Appl. Algebra Engrg. Comm. Comput.*, 10(1):15–32, 1999.
30. J. Rosenthal and E. V. York. BCH convolutional codes. *IEEE Trans. Automat. Control*, 45(6):1833–1844, 1999.
31. D. Silva, R. Kötter, and F.R. Kschischang. A rank-metric approach to error control in random network coding. *IEEE Trans. Inform. Theory*, 54(9):3951–3967, 2008.
32. A. Wachter-Zeh, M. Stinner, and V. Sidorenko. Convolutional codes in rank metric with application to random network coding. *IEEE Trans. Inform. Theory*, 61(6):3199–3213, June 2015.