

Concatenation of convolutional codes and rank metric codes for multi-shot network coding

D. Napp · R. Pinto · V. Sidorenko

Received: date / Accepted: date

Abstract In this paper we present a novel coding approach to deal with the transmission of information over a network. In particular we make use of the network several times (multi-shot) and impose correlation in the information symbols over time. We propose to encode the information via an inner and an outer code, namely, a Hamming metric convolutional code as an outer code and a rank metric code as an inner code. We show how this simple concatenation scheme can exploit the potential of both codes to produce a code that can correct a large number of error patterns.

Keywords Network coding · Multi-shot network coding · Concatenated Codes · Convolutional Codes · Rank Metric Codes

1 Introduction

The theory of random linear network coding developed so far is concerned to large extent with the so-called non-coherent *one-shot* network coding [13], where the encoder does not know the network structure. Hence the network is allowed to change very quickly, which is the case in many mobile applications.

D. Napp · R. Pinto ·

CIDMA - Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, Aveiro, Portugal.

This work was supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), within project UID/MAT/04106/2013.

E-mail: diego@ua.pt and raquel@ua.pt

V. Sidorenko

Institute for Communications Engineering, Technical University of Munich, Germany on leave from Institute for Information Transmission Problems, Moscow, Russian Academy of Sciences

E-mail: vladimir.sidorenko@tum.de

The problem is suitably modeled via the operator channel, which makes a very clear connection between network coding and classical information theory. The operator channel can be seen as a standard discrete memoryless channel. To achieve a reliable communication over this channel, matrix codes are employed forming a so-called rank metric code [28].

However, coding can also be performed over multiple uses of the network, whose internal structure may change at each shot, giving rise to the so-called *multi-shot network coding*. The potential of using multi-shot network coding was already observed in the seminal paper [13]. Recently it has been investigated by several authors [3, 16, 22, 31]. The general idea stems from the fact that creating dependencies among the transmitted codewords of different shots can improve the error-correction capabilities. A general construction was proposed for the first time in [22] of multi-shot codes for this channel based on the multilevel coding theory. Another way to impose correlation of codewords over time is by means of convolutional codes [8, 12, 20], see [4] for correlation in the 2D plane. In [3, 15, 16, 31] the use of convolutional codes for multi-shot network coding by means of rank metric convolutional codes was proposed. In [31] a particular class of unit memory rank metric convolutional codes was introduced together with a decoding algorithm able to deal with errors, erasures and deviations. In [3, 15, 16] (see also the references therein) an interesting and more general class of rank metric convolutional codes has been used to cope with network streaming applications such as video streaming. For a more general theoretical approach to rank metric convolutional codes see [19]. Finally we note that in the last years others papers have also appeared dealing with convolutional network coding using very different approaches [9, 24]. These codes do not transmit over the operator channel and therefore are not equipped with the rank metric.

In this paper, we aim to further explore the potential of multi-shot network coding. In particular we propose to use a rank metric code obtained by concatenation of a Hamming metric outer convolutional code and a rank metric inner block code which encodes each symbol of the outer code separately. This allows to decode at each shot, which is simpler than multi-shot decoding. If a symbol cannot be successfully decoded by the inner decoder then it will be declared an erasure. After this, the outer code corrects symbols that were not corrected by the inner code, hence the Hamming metric is the proper metric for the outer code, and well developed Hamming metric codes can be used here. We consider a convolutional outer code since it is less sensitive to distribution of wrong symbols (errors or erasures) in the received sequence in comparison with a block code which restricts the number of wrong symbols in each block. We show how these codes add complex dependencies to data streams in a quite simple way. An extension of the standard rank-metric over multiple shots, that is analogous to the *extended subspace distance* defined in [22], will provide the proper measure for the number of rank erasures (packet losses) that a code can tolerate.

Note that this concatenation scheme is non-standard as previous publications considered outer block codes. The reason of choosing this non-standard scheme is simple: we want to exploit the fact that convolutional codes perform very efficiently when dealing only with erasures. It was recently shown in [29, 5] that using the flexibility of selecting different sliding windows in convolutional codes allows to recover (using elementary linear algebra) patterns of erasures that cannot be decoded by an MDS block code of the same rate. Hence the rank metric code deals with the possible errors occurring in the transmission at each shot and delivers a correct symbol or an erasure to the convolutional code.

The paper is organized as follows. Section 2 provides the background for the development of the paper; the necessary concepts about convolutional codes and rank metric codes are introduced. Section 3 is devoted to present the concatenation scheme together with its main distance properties. Finally in Section 4 we study the performance of the proposed concatenated codes.

2 Preliminaries

This section contains the necessary mathematical background on convolutional codes and rank metric codes needed for the development of our results. Let \mathbb{F} be a finite field and $\mathbb{F}[D]$ be the ring of polynomials with coefficients in \mathbb{F} .

2.1 Convolutional codes

A *convolutional code* \mathcal{C} of rate k/n is an $\mathbb{F}[D]$ -submodule of $\mathbb{F}[D]^n$ of rank k . A full row rank matrix $G(D) \in \mathbb{F}[D]^{k \times n}$ with the property that

$$\mathcal{C} = \text{Im}_{\mathbb{F}[D]} G(D) = \{u(D)G(D) \mid u(D) \in \mathbb{F}^k[D]\},$$

is called a *generator matrix*. 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} .

A rate k/n code \mathcal{C} with degree δ is called an (n, k, δ) convolutional code [18]. An (n, k) block code is an (n, k, δ) convolutional code with $\delta = 0$.

We say that a generator matrix $G(D)$ is *basic* (see, e.g., [25, 27]) if it has a polynomial right inverse. If a code \mathcal{C} admits a basic generator matrix, then the code \mathcal{C} can be equivalently described using an $(n - k) \times n$ full row rank polynomial *parity-check matrix* $H(D)$, $H(D) = H_0 + H_1D + \dots + H_\gamma D^\gamma$ such that

$$\mathcal{C} = \ker H(D) = \{v(D) \in \mathbb{F}[D]^n \mid H(D)v(D)^\top = 0 \in \mathbb{F}[D]^{n-k}\}.$$

In the following, we will consider our codes generated by basic generator matrices, unless otherwise specified.

Let $\text{wt}(v_i)$ is the number of the nonzero components of a vector $v_i \in \mathbb{F}^n$ and $\text{wt}(v(D))$ the Hamming weight of a polynomial vector $v(D) = \sum_{i \in \mathbb{N}} v_i D^i$ defined as $\text{wt}(v(D)) = \sum_{i \in \mathbb{N}} \text{wt}(v_i)$. An important distance measure for a convolutional code \mathcal{C} is its free distance or *Hamming distance* $d_{\text{H}}(\mathcal{C})$ defined as

$$d_{\text{H}}(\mathcal{C}) = \min \{ \text{wt}(v(D)) \mid v(D) \in \mathcal{C} \text{ and } v(D) \neq 0 \}.$$

In [23,26] it was shown that the free distance of an (n, k, δ) convolutional code is upper bounded by

$$d_{\text{H}}(\mathcal{C}) \leq (n - k) \left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1. \quad (1)$$

This bound is called the *generalized Singleton bound* since it generalizes in a natural way the Singleton bound for block codes. An (n, k, δ) convolutional code with its free distance equal to the generalized Singleton bound is called a *maximum distance separable* (MDS) code [26].

Let $v_{[0,j]}(D) = v_0 + v_1 D + \dots + v_j D^j$ be the j -th truncation of the codeword $v(D) = \sum_{i \in \mathbb{N}} v_i D^i \in \mathcal{C}$, $H(D) = H_0 + \dots + H_{\gamma} D^{\gamma}$ and

$$H_j^c = \begin{bmatrix} H_0 & & & & \\ H_1 & H_0 & & & \\ \vdots & \vdots & \ddots & & \\ H_j & H_{j-1} & \cdots & H_0 & \end{bmatrix} \in \mathbb{F}^{(j+1)(n-k) \times (j+1)n}, \quad (2)$$

where $H_j = 0$, for $j > \gamma$.

Another important distance measure for a convolutional code $\mathcal{C} = \ker H(D)$ is the j th *column distance* $d_{\text{H}}^j(\mathcal{C})$, given by the equivalent expressions

$$\begin{aligned} d_{\text{H}}^j(\mathcal{C}) &= \min \{ \text{wt}(v_{[0,j]}(D)) \mid v(D) \in \mathcal{C} \text{ and } v_0 \neq 0 \} \\ &= \min \left\{ \text{wt}(\hat{v}) \mid \hat{v} = (v_0, \dots, v_j)^{\top} \in \ker H_j^c \subset \mathbb{F}^{(j+1)n}, v_0 \neq 0 \right\} \end{aligned} \quad (3)$$

This notion is related to the free distance $d_{\text{H}}(\mathcal{C})$ in the following way

$$d_{\text{H}}(\mathcal{C}) = \lim_{j \rightarrow \infty} d_{\text{H}}^j(\mathcal{C}). \quad (4)$$

Lemma 1 [8] *Let \mathcal{C} be an (n, k, δ) convolutional code defined by a parity-check matrix $H(D) \in \mathbb{F}[D]^{(n-k) \times n}$ and $d \in \mathbb{N}$. Then the following properties are equivalent:*

- a) $d_{\text{H}}^j(\mathcal{C}) = d$;
- b) *none of the first n columns of H_j^c is contained in the span of any other $d - 2$ columns and one of the first n columns of H_j^c is in the span of some other $d - 1$ columns of that matrix.*

The j -th column distance is upper bounded as follows

$$d_H^j(\mathcal{C}) \leq (n-k)(j+1) + 1, \quad (5)$$

and the maximality of any of the j th column distances implies the maximality of all the previous ones, [11,8], *i.e.*,

$$d_H^j(\mathcal{C}) = (n-k)(j+1) + 1 \implies d_H^i(\mathcal{C}) = (n-k)(i+1) + 1, \quad \forall i \leq j.$$

Since no column distance can achieve a value greater than the generalized Singleton bound, there must exist an integer L for which the bound (5) could be attained for all $j \leq L$ and it is a strict inequality for $j > L$ [8]; this value is

$$L = \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor. \quad (6)$$

An (n, k, δ) convolutional code \mathcal{C} with every $d_H^j(\mathcal{C})$ maximal, for each $j \leq L$, is called a *maximum distance profile* (MDP) code [8,11,21]. Therefore, the column distances of MDP codes increase as rapidly as possible for as long as possible. We call $d_H^0(\mathcal{C}), d_H^1(\mathcal{C}), \dots, d_H^L(\mathcal{C})$ the *distance profile* of \mathcal{C} .

The following result characterizes the MDP convolutional codes in terms of the matrices defined in (2).

Theorem 1 [8] *Let \mathcal{C} be an (n, k, δ) convolutional code defined by a parity-check matrix $H(D) \in \mathbb{F}[D]^{(n-k) \times n}$. Then the following properties are equivalent:*

- a) $d_H^j(\mathcal{C}) = (n-k)(j+1) + 1$;
- b) every $(n-k)(j+1) \times (n-k)(j+1)$ full size minor of H_j^c formed from the columns with indices $1 \leq r_1 \leq \dots \leq r_{(j+1)(n-k)}$ where $r_{s(n-k)} \leq sn$ for $s = 1, \dots, j$, is nonzero.

2.2 Rank metric codes

A rank metric code \mathcal{C} is defined as any nonempty subset of $\mathbb{F}_q^{n \times m}$, the set of $n \times m$ matrices over \mathbb{F}_q . A natural metric for matrix codes is induced by the distance measure $d_{\text{rank}}(X, Y) = \text{rank}(X - Y)$, where $X, Y \in \mathbb{F}_q^{n \times m}$ [13]. In the context of the rank metric, a matrix code is called rank metric code. Rank metric codes in $\mathbb{F}_q^{n \times m}$ are usually constructed as block codes of length n over the extension field \mathbb{F}_{q^m} [13]. For a given basis of \mathbb{F}_{q^m} viewed as an m vector space over \mathbb{F}_q , any element of \mathbb{F}_{q^m} can be seen as a vector in \mathbb{F}_q^m . Analogously, any vector x of length n over \mathbb{F}_{q^m} can be regarded as an element X in $\mathbb{F}_q^{n \times m}$. We commit a harmless abuse of notation and define the rank of a vector $x \in \mathbb{F}_{q^m}^n$ as the rank of x as an $n \times m$ matrix over \mathbb{F}_q .

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

$$d_{\text{rank}}(\mathcal{C}) = \min_{X, Y \in \mathcal{C}, X \neq Y} d_{\text{rank}}(X, Y).$$

In this paper we will consider linear codes over \mathbb{F}_{q^m} and we use k for the dimension of the linear code over \mathbb{F}_{q^m} . To simplify presentation we will assume that $m \geq n$. In this case, for linear (n, k) rank metric codes over \mathbb{F}_{q^m} the following analog of the Singleton bound holds:

$$d_{\text{rank}}(\mathcal{C}) \leq n - k + 1. \quad (7)$$

A code that achieves this bound is called Maximum Rank Distance (MRD). Gabidulin codes are a well-known class of MRD codes [7], see also [6, 10].

3 A concatenated code: Convolutional and Rank Metric codes

Let \mathcal{C}_I be a linear (n_I, k_I) rank metric code with (rank) distance $d_{\text{rank}}(\mathcal{C}_I)$ and generator matrix G_I . Let \mathcal{C}_o be a (n_o, k_o, δ) convolutional code over the field $\mathbb{F}_{q^{mk_I}}$ with (Hamming) distance $d_H(\mathcal{C}_o)$, column distance $d_H^j(\mathcal{C}_o)$ and a basic generator matrix $G_o(D)$. In this section we propose a concatenation scheme of \mathcal{C}_o as an outer code and \mathcal{C}_I as an inner code to obtain a concatenated code \mathcal{C} . We will also provide definitions according to the proposed scheme and finally the distance properties of \mathcal{C} will be investigated.

3.1 The concatenated code

Let $u(D) = u_0 + u_1D + u_2D^2 + \dots \in \mathbb{F}_{q^{mk_I}}[D]^{k_o}$ be the information (row) vector. Encode it through $G_o(D) \in \mathbb{F}_{q^{mk_I}}[D]^{k_o \times n_o}$ to obtain

$$v(D) = v_0 + v_1D + v_2D^2 + \dots = u(D)G_o(D) \in \mathcal{C}_o \subset \mathbb{F}_{q^{mk_I}}[D]^{n_o}.$$

We write

$$v_i = (v_i^0, v_i^1, \dots, v_i^{n_o-1}), \quad v_i^j \in \mathbb{F}_{q^{mk_I}}.$$

We identify $v_i^j \in \mathbb{F}_{q^{mk_I}}$ with a vector $\nu_i^j \in \mathbb{F}_{q^m}^{k_I}$ (for a given basis of $\mathbb{F}_{q^{mk_I}}$ over \mathbb{F}_{q^m}) and write

$$\nu_i = (\nu_i^0, \nu_i^1, \dots, \nu_i^{n_o-1}) \in (\mathbb{F}_{q^m}^{k_I})^{n_o}$$

and therefore

$$\nu(D) = \nu_0 + \nu_1D + \nu_2D^2 + \dots \in \mathbb{F}_{q^m}^{k_I}[D]^{n_o}.$$

Finally, the codewords $x(D)$ of the concatenated code \mathcal{C} are obtained through the matrix $G_I \in \mathbb{F}_{q^m}^{k_I \times n_I}$ in the following way:

$$x_i^j = \nu_i^j G_I \in \mathbb{F}_{q^m}^{n_I},$$

$$x_i = (x_i^0, x_i^1, \dots, x_i^{n_I-1}) \in (\mathbb{F}_{q^m}^{n_I})^{n_o}$$

and

$$x(D) = x_0 + x_1D + x_2D^2 + \dots \in \mathcal{C} \subset \mathbb{F}_{q^m}^{n_I}[D]^{n_o}.$$

Remark 1 It is standard to call the indeterminate D the delay operator [18]. Hence the vector v_i is being generated (or sent) at time instant i . Note however that in the proposed concatenated scheme D^i stands for transmission of one n_o -th block of the convolutional code, which is transmitted using n_o shots (using the network n_o times).

Remark 2 It is worth mentioning that different concatenations are also possible. Another very natural concatenation would be to let the inner code act directly on each coefficient $v_i \in (\mathbb{F}_{q^m}^{k_I})^{n_o}$ of the codewords $v(D) \in \mathcal{C}_o$ instead of on each v_i^j . The reasons to choose the above described concatenation scheme are twofold: first, field size needed to construct MDP convolutional codes is very large [1, 2, 8, 11] and therefore one can split $v_i \in (\mathbb{F}_{q^m}^{k_I})^{n_o}$ into components of smaller sizes. Secondly, the formulas for the distance properties are very simple for the proposed concatenation code (see theorems below).

3.2 Distances of the concatenated code

Let us now define the distance notions for this concatenated code \mathcal{C} .

Definition 1 Let $x(D)$, x_i and x_i^j be as defined above. Define

$$\text{sumrank}(x_i) := \sum_{j=0}^{n_o-1} \text{rank}(x_i^j),$$

and

$$\text{sumrank}(x(D)) := \sum_{i \geq 0} \text{sumrank}(x_i).$$

Then *sum rank distance* of \mathcal{C} is defined as

$$d_{SR}(\mathcal{C}) = \min_{0 \neq x(D) \in \mathcal{C}} \text{sumrank}(x(D)).$$

Let $x|_{[0,j]} = [x_0 \ x_1 \ \dots \ x_j]$ represents the j -th truncation of the codeword $x(D) = x_0 + x_1 D + x_2 D^2 + \dots \in \mathcal{C}$. Let

$$\text{sumrank}(x|_{[0,j]}) := \sum_{i=0}^j \text{sumrank}(x_i)$$

Then, the *column sum rank distance* of \mathcal{C} is defined as

$$d_{SR}^j(\mathcal{C}) = \min_{x(D) \in \mathcal{C} \text{ and } x_0 \neq 0} \text{sumrank}(x|_{[0,j]}).$$

According to the above described procedure one receives at each shot a vector $x_i^j \in \mathbb{F}_{q^m}^{n_I}$. Each coordinate of x_i^j represents a packet and therefore n_I packets are sent at each shot [13, 28]. For the sake of simplicity only packet losses are considered. However, it is very important to note that the inner code can deal also with errors and/or deviations. If the network is perfectly

functional then $\text{rank}(x_i^j) = n_I$ and no packet is lost. However the network may fail to transmit all the intended packets which results in a rank deficiency producing $\text{rank}(x_i^j) < n_I$ and we say the $n_I - \text{rank}(x_i^j)$ packets are lost. If there are too many packet losses for the inner code \mathcal{C}_I to recover, then the corresponding ν_i^j is also lost and we say that ν_i^j has been erased.

3.3 Distance properties

This concatenated code has the following distance properties (we assume throughout the paper that $m \geq n_I$).

The following result is a straightforward adaptation to the context of this paper of well-known results of concatenated codes, see for instance [14, Section 18.8].

Theorem 2 *The sum rank distance of the concatenated code \mathcal{C} satisfies*

$$d_{SR}(\mathcal{C}) \geq d_H(\mathcal{C}_o)d_{\text{rank}}(\mathcal{C}_I).$$

Proof: The words of the concatenated code \mathcal{C} form an additive group, and hence the code distance $d_{SR}(\mathcal{C})$ is the minimum sum rank of a nonzero codeword $x(D)$ of \mathcal{C} . To get a nonzero codeword $x(D)$ we should take a nonzero codeword $v(D)$ of the outer code, which has a Hamming weight at least $d_H(\mathcal{C}_o)$, i.e., $v(D)$ has at least $d_H(\mathcal{C}_o)$ nonzero components v_i^j .

After inner encoding each $v_i^j \neq 0$ is encoded into x_i^j of rank at least $d_{\text{rank}}(\mathcal{C}_I)$ and statement of the theorem follows.

In the context of convolutional codes the notion of column distance plays a central role. In particular, in applications where the delay is important, such as streaming applications, large column distance are desired. The following result on column sum rank distance is straightforward. We include its short proof for completeness.

Theorem 3 *The column sum Rank distance of the concatenated code \mathcal{C} satisfies*

$$d_{SR}^j(\mathcal{C}) \geq d_H^j(\mathcal{C}_o)d_{\text{rank}}(\mathcal{C}_I).$$

Proof: By definition we have that for all $v(D) \in \mathcal{C}_o$, with $v_0 \neq 0$, $v|_{[0,j]}$ has at least $d_H^j(\mathcal{C}_o)$ components different from zero. Each component different from zero v_i^s , $i \in \{0, \dots, j\}$, $s \in \{0, \dots, n_o - 1\}$, gives rise to $x_i^s \in \mathcal{C}_I$ with $\text{rank}(x_i^s) \geq d_{\text{Rank}}(\mathcal{C}_I)$ and therefore the result follows. \square

To get an upper bound for the sum rank distance of the constructed concatenated code \mathcal{C} we can consider it as a slightly generalized finite-state (FS) code [23]. An (N, K, M) FS generator matrix is a Q^M state time-invariant

finite-state machine. The generator matrix starts from a fixed initial state. At each time instant t the generator matrix takes one of Q^K information messages as input, changes state and outputs a block V_t of N Q -ary code symbols. The set of all such code sequences (V_0, V_1, \dots) is called an FS code. The *Hamming* distance of Q -ary (N, K, M) FS code satisfies the upper bound [23]

$$d_H(\mathcal{C}) \leq (N - K) \left(\left\lfloor \frac{M}{K} \right\rfloor + 1 \right) + M + 1. \quad (8)$$

Our concatenated code is an (N, K, M) FS code over alphabet \mathbb{F}_{q^m} of size $Q = q^m$ where $N = n_o n_I$, $K = k_o k_I$, and $M = k_I \delta$. Since the sum rank distance of the concatenated code \mathcal{C} is upper bounded by the Hamming distance, i.e., $d_{SR}(\mathcal{C}) \leq d_H(\mathcal{C})$ we get the following result.

Theorem 4 *The sum rank distance of the concatenated code \mathcal{C} satisfies*

$$d_{SR}(\mathcal{C}) \leq (n_o n_I - k_o k_I) \left(\left\lfloor \frac{\delta}{k_o} \right\rfloor + 1 \right) + \delta k_I + 1. \quad (9)$$

Another way to prove this theorem is to observe that the concatenated code \mathcal{C} is an $(n_o n_I, k_o k_I, \delta k_I)$ convolutional code over \mathbb{F}_{q^m} . From this observation, equation (5) and inequality $d_{SR}(\mathcal{C}) \leq d_H(\mathcal{C})$ we also obtain

Theorem 5 *The column sum rank distance of the concatenated code \mathcal{C} satisfies*

$$d_{SR}^j(\mathcal{C}) \leq (n_o n_I - k_o k_I)(j + 1) + 1. \quad (10)$$

Let us compare achievable column sum rank distance $d_{SR}^j(\mathcal{C})$ of the concatenated code given by Theorem 3 with the upper bound (10). If we take an MDP convolutional code, which satisfies the upper bound (5) with equality as an outer code of rate $k/n = k_o/n_o$ and the inner rate k_I/n_I Gabidulin code that satisfy the Singleton bound (7) then from Theorem 3 we get the concatenated code with column sum rank distance

$$d_{SR}^j(\mathcal{C}) \geq d_H^j(\mathcal{C}_o) d_{rank}(\mathcal{C}_I) = ((n_o - k_o)(j + 1) + 1)(n_I - k_I + 1). \quad (11)$$

From (11) we see that the sum rank column distance of the concatenated code \mathcal{C} can reach the upper bound (10) if we take the trivial inner code $k_I = n_I = 1$. Otherwise, the sum rank distance of the concatenated code is less than the bound (10). But the advantage of the concatenated code is that it allows to make separately decoding of the network code and decoding of the outer classical code in Hamming metric. As a result decoding of a single complicated code is split to decoding of simpler component codes, simplifying the overall decoding complexity.

In [17] a construction of an (n, k, δ) convolutional code with optimal rank sum column distance profile satisfying $d_{SR}^j = (n - k)(j + 1) + 1$ was proposed. The codes in [17] are sum rank metric analogs of MDP codes [1, 8]. One disadvantage of these codes is the requirement of very large fields requiring double exponential size in the code degree δ .

4 Performance of the concatenated code

In this section the performance of the proposed concatenation scheme is investigated. As mentioned before and for the sake of simplicity we only consider rank deficiencies (packet losses). There exist decoding algorithms for rank metric codes to deal with errors, erasures and/or deviations (see [13,30]) that can be incorporated in this scheme. If the inner code fails to decode the received set of packets x_i^j then the corresponding v_i^j is declared an erasure. In case of packet losses only, the inner decoder never makes a mistake, i.e., it outputs either the correct symbol v_i^j or declares erasure at this position. Thus the convolutional code has to deal only with erasures. In this section we present conditions on the received data so that it can be completely decoded up to a given instant. Before these results can be stated we first need to derive some new key results on the decoding of convolutional codes over the erasure channel, that further develop the results of [29].

4.1 Decoding of convolutional codes over the erasure channel

Let \mathcal{C}_o be an (n_o, k_o, δ) convolutional code, $d_H^T(\mathcal{C}_o)$ be its T -th column distance and let $H(D) = H_0 + H_1D + H_2D^2 + \dots + H_\gamma D^\gamma$ be a parity-check matrix of \mathcal{C}_o . Assume that we have been able to correctly decode up to an instant $t - 1$. Then for each received codeword $v(D) = v_0 + v_1D + v_2D^2 + \dots \in \mathcal{C} \subset \mathbb{F}_q[D]^{n_o}$ consider the system of linear equations

$$\left[\begin{array}{ccc|ccc} H_\gamma & \dots & \dots & H_1 & H_0 & \\ & H_\gamma & & & H_1 & H_0 \\ & & \ddots & & \vdots & \vdots & \ddots \\ & & & H_\gamma & \dots & \dots & H_0 \end{array} \right] \underbrace{\left[\begin{array}{ccc} H_0 & & \\ H_1 & H_0 & \\ \vdots & \vdots & \ddots \\ H_T & H_{T-1} & \dots & H_0 \end{array} \right]}_{H_T^c} \begin{bmatrix} v_{t-\gamma} \\ \vdots \\ v_{t-1} \\ v_t \\ \vdots \\ v_{t+T} \end{bmatrix} = 0 \quad (12)$$

where v_i , $t \leq i \leq t + T$ may contain some of its components erased and v_i , $t - \gamma \leq i < t$ are assumed to be correct. We can consider the columns of H_T^c that correspond to the coefficients of the erased elements to form a matrix denoted by \widehat{H}_T^c . The remaining columns will help us to compute the independent terms of a system, i.e., if \widetilde{H} is the matrix containing the remaining columns and \widetilde{v} the corresponding (known) coefficients in (12), then we obtain the non-homogeneous linear system with $(T + 1)(n - k)$ equations,

$$\widehat{H}_T^c Y = -\widetilde{H}\widetilde{v}, \quad (13)$$

where Y corresponds to the vector containing the erasures in $v_{[t, \dots, t+T]}$. Note that this system has always a solution since $v(D) \in \ker H(D)$. Thus, it will be possible to recover all the erasures in $v_{[t, \dots, t+T]}$ if and only if the system (13) has a *unique* solution, i.e., if and only if \widehat{H}_T^c has full column rank.

Lemma 2 *Let \mathcal{C}_o be an (n_o, k_o, δ) convolutional code and let $d_H^T(\mathcal{C}_o)$ be its T -th column distance, $T \leq L$. Assume that we have been able to correctly decode up to an instant $t - 1$. Let $E(t, t + T)$ be the number of erasures occurring in the time interval $[t, t + T]$. Then, we can recover v_t if*

$$E(t, t + T) \leq d_H^T(\mathcal{C}_o) - 1.$$

Proof: Let $H(D)$ and \widehat{H}_T^c as defined above. Let E_t be the erasures at time t . As $E(t, t + T) \leq d_H^T(\mathcal{C}_o) - 1$, by Lemma 1, it follows that none of the first E_t columns of \widehat{H}_T^c is contained in the span of any other $d_H^T(\mathcal{C}_o) - 2$ columns of \widehat{H}_T^c . This in particular implies that there exists an $i \in \{0, \dots, T\}$ such that \widehat{H}_i^c is full column rank and therefore we can recover v_t solving the corresponding system. \square

If instead of knowing the number of erasures within a window interval, we have information about the number of erasures at each time instant, a more accurate statement can be inferred.

Lemma 3 *Let $d_H^0(\mathcal{C}_o), d_H^1(\mathcal{C}_o), \dots, d_H^L(\mathcal{C}_o)$ be the distance profile of an (n_o, k_o, δ) convolutional code \mathcal{C}_o . Let E_i be the number of erasures at time instant i . Assume that we have been able to correctly decode up to an instant $t - 1$. Then, we can completely decode up to an instant $t + T$ where $T \leq L$ if*

$$\sum_{i=0}^s E_{T-i+t} \leq d_H^s(\mathcal{C}_o) - 1 \text{ for } s = 0, 1, \dots, T.$$

Proof: Take $s = T$ and apply Lemma 2 to recover v_t . Apply the same argument sequentially to recover $v_i, i = t + 1, \dots, t + T$. \square

It is important to note that the conditions of Lemma 3 may not be satisfied at time instant $s - 1$ but hold true at instant s . Therefore one may need to wait until these conditions are satisfied in order to proceed with the decoding. We shall develop an algorithm explaining this in detail at the end of the section.

Of course the best scenario is when the convolutional code is MDP as we can recover the maximum possible erasures per interval. Moreover, the conditions of the previous lemmas become sufficient when considering MDP convolutional codes as we state in the following theorem.

Theorem 6 *Let \mathcal{C}_o be an MDP (n_o, k_o, δ) convolutional code. Assume that we have been able to correctly decode up to an instant $t - 1$. Let E_i be the number of erasures at time instant i . Then, we can completely decode up to an instant $t + T$ where $T \leq L$ if and only if*

$$\sum_{i=0}^s E_{T-i+t} \leq (n_o - k_o)(s + 1) \text{ for } s = 0, 1, \dots, T. \quad (14)$$

Proof: Note that we can completely decode up to T if and only if the matrix \widehat{H}_c^T of system (13) is full column rank. By Theorem 1 and from the structure of matrix \widehat{H}_c^T , this happens if and only if the conditions in (14) are satisfied. \square

4.2 Decoding of the concatenated code \mathcal{C}

Let us now consider the concatenated code and analyze the conditions necessary to recover the missing packets. The error-correcting capabilities of these codes will depend on how the packet losses are distributed along (x_t, \dots, x_{t+T}) as we illustrate in the following example.

Example 1 Consider the concatenated code with an MDP $(2, 1, 25)$ convolutional code (and therefore $L = 50$) as the outer code, and a linear MRD $(3, 2)$ rank metric code as the inner code. Hence, at each time instant we send x_j^i in two shots x_j^0 and x_j^1 . At each shot we are able to recover the packet losses $n_I - \text{rank}(x_j^i)$ if $n_I - \text{rank}(x_j^i) \leq d_{\text{rank}}(\mathcal{C}_I) - 1$, see [28]. In this case we have that $n_I = 3$ and $d_{\text{rank}}(\mathcal{C}_I) = n_I - k_I + 1 = 2$ and therefore only one packet loss can be recovered at each shot. Next we present two extreme situations: in both cases 280 packets were lost in the time interval $[t, t + 69]$ however only in the second situation we can recover the missing data.

Assume that we have been able to correctly decode up to an instant $t - 1$ and that the data received after the time instant $t + 69$ is correct.

1) The received sequence $(x_t, x_{t+1}, \dots, x_{t+69})$ is such that

$$\text{rank}(x_j^0) = \text{rank}(x_j^1) = 1, \quad j = t, \dots, t + 69.$$

This means that two packets were lost in each received x_j^i and therefore 280 packets were lost in total. Note that the inner code can recover a received x_j^i if and only if $\text{rank}(x_j^i) \geq n_I - d_{\text{rank}}(\mathcal{C}_I) + 1$. Since this does not hold then all the corresponding v_j^0, v_j^1 , $i = t, \dots, t + 69$ will be considered erasures. In this extreme situation, it will not be possible to recover v_t .

2) The received sequence $(x_t, x_{t+1}, \dots, x_{t+69})$ is such that

$$\text{rank}(x_j^0) = \text{rank}(x_j^1) = 0, \quad j = t, \dots, t + 19, t + 50, \dots, t + 69,$$

$$\text{rank}(x_j^0) = \text{rank}(x_j^1) = 2, \quad j = t + 20, \dots, t + 39,$$

and

$$\text{rank}(x_j^0) = \text{rank}(x_j^1) = 3, \quad j = t + 40, \dots, t + 49.$$

The total number of packet losses is 280, as in 1). By the same reasoning as in 1), it will be not possible to correct x_j^0 and x_j^1 for $j = t, \dots, t + 19$ and for $j = t + 50, \dots, t + 69$ and therefore the corresponding v_j^i will be declared erased symbols. The x_j^0 and x_j^1 , $j = t + 40, \dots, t + 49$ are correct and the inner

code can correct x_j^0 and x_j^1 , for $j = t + 20, \dots, t + 39$. Thus, the distribution of the erasures along the time interval $[t, t + 69]$ is the following:

$$\cdots v_{t-1}^0 v_{t-1}^1 | \overbrace{\star \star \cdots \star \star}^{2 \times 20} \overbrace{v_{t+20}^0 v_{t+20}^1 \cdots v_{t+49}^0 v_{t+49}^1}^{2 \times 30} \overbrace{\star \star \cdots \star \star}^{2 \times 20} | v_{t+70}^0 v_{t+70}^1 \cdots$$

In this case it is possible to completely recover all information sequence by carefully selecting the appropriate intervals (sliding windows) to decode. Consider first $T = 39$ and the window $[t, t + 39]$. The number of erasures E_i , $i = t, \dots, t + 39$ satisfy condition (14) and therefore by Theorem 6, it is possible to completely decode up to the instant $t + 39$. Now we can shift the sliding window and consider the window $[t + 50, t + 89]$ in order to completely decode the remaining erasures.

Note that in the situation of example 1 part 2), if instead of considering the $(2, 1, 25)$ MDP convolutional code as the outer code we consider a $[140, 70]$ MDS block code, it would not be possible to recover the symbols in the window $[t, t + 69]$. In fact, this block code can correct up to 70 erasures in a window of 140 symbols and therefore it would not be able to decode the 80 erasures appearing in the window $[t, t + 69]$. Thus, although both codes have the same erasure recovering capability of 50%, the flexibility in selecting the sliding window when decoding a convolutional code improves its decoding efficiency. More details can be found in [29].

Next theorem presents conditions on the pattern of packet losses so that one can completely recover the transmitted sequence.

Theorem 7 *Let \mathcal{C}_o , \mathcal{C}_I and \mathcal{C} be the codes defined in the previous section. Let $d_H^0(\mathcal{C}_o), d_H^1(\mathcal{C}_o), \dots, d_H^L(\mathcal{C}_o)$ be the distance profile of \mathcal{C}_o . Let P_s be the number of packet losses in x_s . Assume that we have been able to correctly decode up to an instant $t - 1$. Then, we can completely decode up to an instant $t + T$ where $T \leq L$ if*

$$\sum_{i=0}^s P_{T-i+t} \leq d_H^s(\mathcal{C}_o) d_{\text{rank}}(\mathcal{C}_I) - 1 \quad \text{for } s = 0, 1, \dots, T.$$

Proof: First note that each $v_t^j \in \mathbb{F}_{q^m}^{k_I}$ can be recovered from $x_t^j \in \mathbb{F}_{q^m}^{n_I}$ if $n_I - \text{rank}(x_t^j) \leq d_{\text{rank}}(\mathcal{C}_I) - 1$, see [28], i.e., there must be at least $d_{\text{rank}}(\mathcal{C}_I)$ lost packets at shot j and time instant t in order to lose v_t^j (or equivalently v_t^j).

Hence if the total number of lost packets in $x_i = [x_i^0 \cdots x_i^{n_o-1}]$ at time instant i is P we have that $\lfloor \frac{P}{d_{\text{rank}}(\mathcal{C}_I)} \rfloor$ is the maximum number of v_i^j in $[v_i^0 \cdots v_i^{n_o-1}]$ that are erased at time instant i .

Thus $\sum_{i=0}^s \lfloor \frac{P_{T-i+t}}{d_{\text{rank}}(\mathcal{C}_I)} \rfloor$ is the maximum number of v_i^j , and therefore of v_i^j , $T - s + t \leq i \leq T + t$, $0 \leq j < n_o$ that are erased in the sliding window $[v_{T-s+t} \cdots v_{T+t}]$. Since for $s = 0, 1, \dots, T$, it holds that

$$\sum_{i=0}^s \left\lfloor \frac{P_{T-i+t}}{d_{\text{rank}}(\mathcal{C}_I)} \right\rfloor \leq \left\lfloor \frac{\sum_{i=0}^s P_{T-i+t}}{d_{\text{rank}}(\mathcal{C}_I)} \right\rfloor \leq \left\lfloor d_H^s(\mathcal{C}) - \frac{1}{d_{\text{rank}}(\mathcal{C}_I)} \right\rfloor = d_H^s(\mathcal{C}_o) - 1$$

it follows, by Lemma 3, that v_t can be recovered. Now take the window $[\nu_{t+1} \cdots \nu_{t+T}]$ and follow the same arguments to decode v_{t+1} . Repeat this procedure up to recovering v_{t+T} . \square

The next theorem considers a concatenated code in which the outer code is MDP and presents a necessary condition to ensure that the recovery is impossible within a given maximal delay.

Theorem 8 *Let \mathcal{C}_o , \mathcal{C}_I and \mathcal{C} be the codes defined in the previous section such that \mathcal{C}_o is MDP and $T \leq L$. Assume that we have been able to correctly decode up to an instant $t - 1$. Let P_s be the number of packet losses in x_s . Then, it will not be possible to completely decode (x_t, \dots, x_{T+t}) if*

$$\sum_{i=0}^T P_{t+i} \geq (n_o - k_o)(T + 1)n_I + k_o(T + 1)(d_{\text{rank}}(\mathcal{C}_I) - 1) + 1.$$

Proof: By Theorem 6 the maximal amount of erasures that can occur in the time interval $[t, t+T]$ in order to be possible to completely decode (v_t, \dots, v_{t+T}) is $(n_o - k_o)(T + 1)$. In this case, there will be $(T + 1)n_o - (n_o - k_o)(T + 1) = k_o(T + 1)$ elements in (x_t, \dots, x_{T+t}) that were decoded correctly and therefore will not produce erasures. Since the maximal number of packet losses that a received element x_i^j can have in order to be corrected is $d_{\text{rank}}(\mathcal{C}_I) - 1$, we have that the maximal number of packet losses in (x_t, \dots, x_{t+T}) that will produce this number of erasures is $(n_o - k_o)(T + 1)n_I + k_o(T + 1)(d_{\text{rank}}(\mathcal{C}_I) - 1)$ and the results follows immediately. \square

Next, we present a decoding algorithm for the concatenated code constituted by an (n_o, k_o, δ) convolutional code over a field $\mathbb{F}_{q^{m k_I}}$, \mathcal{C}_o , as the outer code, and an (n_I, k_I) rank metric code as the inner code, \mathcal{C}_I . The algorithm fixes a delay T . At a certain instant t , it considers that the received sequence was correctly decoded up to an instant $t - 1$ and tries to decode v_t , by successively considering the received sequence in the interval $[t, t+i]$, for $i = 0, \dots, T$, until v_t is decoded (then it returns v_t and declares success - *Success = TRUE*) or declares failure (*Success = FALSE*) if it is not able to recover v_t in the time interval $[t, t + T]$. If we succeed we look for the next vector v_i with erasures and proceed in the same way.

The algorithm uses the function $Correct(x_t, \dots, x_{t+\ell})$ that tries to decode v_t using $(x_t, \dots, x_{t+\ell})$. This function is also presented after the algorithm. It returns v_t and *Correctable = TRUE* if it succeeds or *Correctable = FALSE*, otherwise. It also resorts to the subfunction $innercorrect(x_i^j)$ that decodes x_i^j into ν_i^j , using a decoding algorithm of \mathcal{C}_I , and then computes the corresponding v_i^j . $\widehat{H}_\ell^c[col]$ represents the submatrix of \widehat{H}_ℓ^c constituted by the columns with indices in col . The vector $v[col]$ is defined in a similar way.

Next we give an upper bound on the decoding failure probability in terms of the probability of obtaining a symbol erasure at each transmission shot. Let

Algorithm

Result: $(v_t, \text{Success})$.

```

for  $\ell = 0, \dots, T$  do
   $(v_t, \text{Correctable}) = \text{Correct}(x_t, \dots, x_{t+\ell})$ 
  if  $\text{Correctable} = \text{TRUE}$  then
    | Return  $v_t$  and  $\text{Success} = \text{TRUE}$ 
  end
end
Return  $\text{Success} = \text{FALSE}$ 

Correct  $(x_t, \dots, x_{t+\ell})$ :
Data:  $x_i = (x_i^0, \dots, x_i^{n_o-1})$ ,  $i = t, \dots, t + \ell$ 
Result:  $(v_t, \text{Correctable})$ .
 $col = \emptyset$ ;
for  $i = 0, \dots, \ell$  do
  for  $j = 1, \dots, n_o$  do
    | if  $n_I - \text{rank}(x_i^j) \geq d_{\text{rank}}(C_I)$  then
      | |  $v_i^j = \text{erasure}$ 
      | |  $col = col \cup \{in_o + j\}$ 
    | else
      | |  $v_i^j = \text{innercorrect}(x_i^j)$ 
    | end
  end
   $v_i = (v_i^1, \dots, v_i^{n_o})$ 
end
if  $col \cap \{1, \dots, n_o\} = \emptyset$  then
  | Return  $v_t$  and  $\text{Correctable} = \text{TRUE}$ 
else
  | if  $\widehat{H}_\ell^c[col]$  is full column rank then
  | | solve the system (13) in the unknowns  $v[col]$ 
  | | Return  $v_t$  and  $\text{Correctable} = \text{TRUE}$ 
  | else
  | | Return  $\text{Correctable} = \text{FALSE}$ 
  | end
end

```

\mathcal{C}_o , \mathcal{C}_I and \mathcal{C} be the outer code, the inner code and the concatenated codes, respectively, defined as in the previous section. If at a certain transmission shot, the number ℓ of lost packets is more than $d_{\text{rank}}(\mathcal{C}_I) - 1$, the inner decoder gives to the outer code the symbol of erasure. Assume that we transmit via a network n_I linearly independent packets, and each packet can be lost with probability p independently of other packets. Then the probability p_o of symbol erasure of the outer code is

$$p_o = \sum_{\ell=d_I}^{n_I} \binom{n_I}{\ell} p^\ell (1-p)^{n_I-\ell}. \quad (15)$$

After decoding sufficient information via the inner code, we decode using the outer convolutional code. Assume that we have been able to correctly decode up to an instant $t - 1$ and that there are erasures in v_t . According to Lemma 2, we can recover v_t using the received information in a window of

size $T + 1$, $(v_t, v_{t+1}, \dots, v_{t+T})$, if the number of erasures in the window is less than $\hat{d} = d_H^T(\mathcal{C}_o)$. Hence, the probability of failure P_f is upper bounded by the probability of having at least \hat{d} erasures in the window $[t, t + T]$ (given by (16)), where at least one erasure should be in the block v_t (see correction given by (17)). Since every symbol of the outer code can be independently erased with probability p_o , we obtain

$$P_f(p) \leq \sum_{\ell=\hat{d}}^{n_o(T+1)} \binom{n_o(T+1)}{\ell} p_o^\ell (1-p_o)^{n_o(T+1)-\ell} \quad (16)$$

$$- \sum_{\ell=\hat{d}}^{n_o T} \binom{n_o T}{\ell} p_o^\ell (1-p_o)^{n_o(T+1)-\ell}. \quad (17)$$

Example 2 For the case of an $(3, 2, \delta)$ (with $\delta \geq 1$) outer MDP convolutional code, a linear $(3, 2)$ MRD inner code and a window parameter $T = 2$, we get the function of failure probability $P_f(p)$ shown in Fig. 1.

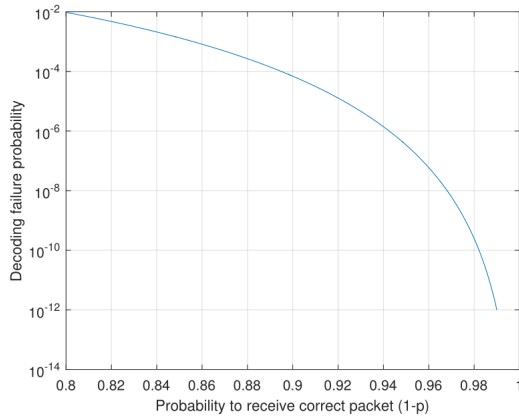


Fig. 1 Failure probability (Example 2)

Remark 3 Another interesting variation of the proposed decoding procedure is to adopt a more conservative approach when decoding with the inner code. Hence, instead of decoding packet losses up to the maximum rank distance $d_{\text{rank}}(\mathcal{C}_I) - 1$ one can decode up to $d_{\text{rank}}(\mathcal{C}_I) - 1 - \varepsilon$ so that the probability of failure when decoding (wrong codeword is found by the decoder) is reduced.

5 Acknowledgement

We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments.

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. Badr, A., Khisti, A., Tan, W.T., Apostolopoulos, J.: Layered constructions for low-delay streaming codes. *IEEE Trans. Inform. Theory* (2013 accepted). DOI <http://arxiv.org/pdf/1308.3827v1.pdf>
4. Climent, J., Napp, D., Perea, C., Pinto, R.: Maximum distance separable 2D convolutional codes. *IEEE Trans. Inform. Theory* **62**(2), 669–680 (2016)
5. Climent, J., Napp, D., Pinto, R., Simões, R.: Decoding of 2D convolutional codes over the erasure channel. *Advances in Mathematics of Communications* **10**(1), 179–193 (2016)
6. Delsarte, P.: Bilinear forms over a finite field, with applications to coding theory. *Journal of Combinatorial Theory, Series A* **25**(3), 226–241 (1978)
7. Gabidulin, É.: Theory of codes with maximum rank distance. *Probl. Inf. Transm.* **21**, 1–12 (1985)
8. Gluesing-Luerssen, H., Rosenthal, J., Smarandache, R.: Strongly MDS convolutional codes. *IEEE Trans. Inform. Theory* **52**(2), 584–598 (2006)
9. Guo, W., Shi, X., Cai, N., Medard, M.: Localized dimension growth: A convolutional random network coding approach to managing memory and decoding delay. *IEEE Transactions on Communications* **61**(9), 3894–3905 (2013)
10. Horlemann-Trautmann, A., Marshall, K.: New criteria for MRD and Gabidulin codes and some rank-metric code constructions. *arXiv: 1507.08641*
11. Hutchinson, R., Rosenthal, J., Smarandache, R., Trunpf, J.: Convolutional codes with maximum distance profile. *Systems & Control Letters*
12. Johannesson, R., Zigangirov, K.S.: *Fundamentals of Convolutional Coding*. IEEE Press, New York (1999)
13. Kötter, R., Kschischang, F.: Coding for errors and erasures in random network coding. *IEEE Trans. Inform. Theory* **54**(8), 3579–3591 (2008)
14. MacWilliams, F.J., Sloane, N.J.: *The Theory of Error-Correcting Codes*. North Holland, Amsterdam (1977)
15. Mahmood, R.: Rank metric convolutional codes with applications in network streaming. Master of applied science (2015). URL <https://tspace.library.utoronto.ca/handle/1807/70480>
16. Mahmood, R., Badr, A., Khisti, A.: Streaming-codes for multicast over burst erasure channels. *IEEE Trans. Inform. Theory* **61**(8), 4181–4208 (2015)
17. Mahmood, R., Badr, A., Khisti, A.: Convolutional codes with maximum column sum rank for network streaming. *IEEE Trans. Inform. Theory* **to appear** (2016). DOI <http://arxiv.org/pdf/1506.03792.pdf>
18. McEliece, R.J.: The algebraic theory of convolutional codes. In: V. Pless, W. Huffman (eds.) *Handbook of Coding Theory*, vol. 1, pp. 1065–1138. Elsevier Science Publishers, Amsterdam, The Netherlands (1998)
19. Napp, D., Pinto, R., Rosenthal, J., Vettori, P.: Rank metric convolutional codes. *Proceedings of the 22nd International Symposium on Mathematical Theory of Network and Systems (MTNS), Minnesota, USA* (2016)
20. Napp, D., Pinto, R., Toste, T.: On MDS convolutional codes over \mathbb{Z}_p^r . *Designs, Codes and Cryptography* pp. 1–14 (2016). DOI 10.1007/s10623-016-0204-9

21. Napp, D., Smarandache, R.: Constructing strongly MDS convolutional codes with maximum distance profile. *Advances in Mathematics of Communications* **10**(2), 275–290 (2016)
22. Nóbrega, R., Uchoa-Filho, B.: Multishot codes for network coding using rank-metric codes. In: *Wireless Network Coding Conference (WiNC)*, 2010 IEEE, pp. 1–6 (2010)
23. Pollara, F., McEliece, R.J., Abdel-Ghaffar, K.: Finite-state codes. *IEEE Trans. Inform. Theory* **34**(5), 1083–1089 (1988)
24. Prasad, K., Rajan, B.S.: Network error correction for unit-delay, memory-free networks using convolutional codes. *IEEE International Conference on Communications (ICC)*, 2010 pp. 1–6 (2010)
25. Rosenthal, J.: Connections between linear systems and convolutional codes. In: B. Marcus, J. Rosenthal (eds.) *Codes, Systems and Graphical Models*, IMA Vol. 123, pp. 39–66. Springer-Verlag (2001)
26. Rosenthal, J., Smarandache, R.: Maximum distance separable convolutional codes. *Appl. Algebra Engrg. Comm. Comput.* **10**(1), 15–32 (1999)
27. Rosenthal, J., York, E.V.: BCH convolutional codes. *IEEE Trans. Automat. Control* **45**(6), 1833–1844 (1999)
28. Silva, D., Kötter, R., Kschischang, F.: A rank-metric approach to error control in random network coding. *IEEE Trans. Inform. Theory* **54**(9), 3951–3967 (2008)
29. Tomas, V., Rosenthal, J., Smarandache, R.: Decoding of convolutional codes over the erasure channel. *IEEE Trans. Inform. Theory* **58**(1), 90–108 (2012)
30. Wachter-Zeh, A., Afanassiev, V., Sidorenko, V.R.: Fast decoding of Gabidulin codes. *Des. Codes Cryptogr.* **66**, 57–73 (2013)
31. Wachter-Zeh, A., Stinner, M., Sidorenko, V.: Convolutional codes in rank metric with application to random network coding. *IEEE Trans. Inform. Theory* **61**(6), 3199–3213 (2015)