Extended Low Rank Parity Check Codes and Their Efficient Decoding for Multisource Wireless Sensor Networks

Nicolas Aragon, Jean Pierre Cances, Philippe Gaborit, Imad El Quachchach

Extended Low Rank Parity Check Codes and Their Efficient Decoding for Multisource Wireless Sensor Networks

Nicolas ARAGON, Jean Pierre CANCES, Imad EL QACHCHACH and Philippe GABORIT
Xlim Institute of Technology, University of Limoges, France
Email: {nicolas.aragon, jean-pierre.cances, el-qachchach, philippe.gaborit}@xlim.fr

***Abstract*—In this paper, we consider a multisource network transmitting information through relays to a base station using Network Coding. We design a model for this scenario and use the rank metric to address the problem of packet errors (caused for example by a malicious user or a defective node). We introduce a new family of codes, the extended LRPC codes, that are very well suited to this model and extensively use the fact that the information comes from multiple sources to decode. They therefore improve the communication reliability compared to classical LRPC codes and Gabidulin codes. We provide a theoretical analysis of their decoding failure probability, both in a one source and multisource scenario, as well as simulation results confirming our analysis.**

I. Introduction

Network coding (NC) has been recently introduced to reduce the traffic in general networks. Plenty of works have investigated this idea in both wired and wireless networks. Indeed, NC is proved to be an appropriate solution increasing data throughput and reducing energy consumption for WSNs. NC was first introduced in the seminal paper [1] and since, it has been shown to significantly improve network efficiency by reducing the number of transmissions. Random linear network coding (RLNC) [2] is a class of network coding that uses a linear code generated randomly by every node of the network. It assumes that the data are vectors over a finite field and that each node of the network performs a random linear combination of all the received packets so far and forwards them to nearby nodes. Nevertheless, if packet error occurs, the erroneous packets are combined with unharmed ones causing the whole combination to be affected. This kind of errors can be illustrated in three use-cases. The first use-case is when a malicious user injects erroneous packets into the network to disrupt the overall system, such as the scenarios studied in [3] and [4]. The second use-case is depicted by the presence of a node failure within the network, see [5]. The third case is when we take into consideration the impact of background noise that is caused by propagation channel and electronic impairment (additive white Gaussian noise (AWGN) for example). In order to solve the problem of background noise, we propose to use convolutional codes. Each node uses a linear combination of the received packets and decodes them using convolutional decoder. The first and the second cases can be solved by using rank metric codes. It has been proven that rank codes are efficient against rank errors [6]. In particular, Gabidulin proposed a class of correcting codes named *Gabidulin codes* in order to apply them for correcting criss-cross errors. A class of rank metric codes has been proposed in [7], called Low Rank Parity Check (LRPC), that has approximately the same performance of Gabidulin codes. Koetter and Kschischang tested the performance of rank codes combined with RLNC schemes for intentional attacks [8].

In this paper, we investigate existing solution in [9] for multisource networks using error correcting codes and we propose a generalized solution. We also introduce a new family of codes, the extended LRPC codes, that are very well suited to this model.

The main contribution of this paper is a new family of LRPC codes, the extended LRPC codes, that features a probabilistic decoding algorithm whose decoding failure rate gets really low when using multiple sources. In particular, the error support can be naturally recovered from the first coordinate of the received word in such a way that the decoding capability is improved. We also derive theoretical expression of failure decoding probability at the destination for extended LRPC in multisource networks. Finally, we validate the theoretical results with simulations and we show that our proposition achieves good performance compared to existing ones. The simulations illustrate the advantages of using extended LRPC codes compared to classical LRPC and Gabidulin codes.

The remainder of this paper is organized as follows. In section II, notations and fundamental preliminaries of finite field and vector spaces are detailed. A detailed description of rank codes is provided in Section III. Section IV describes the system model and formulates the problem statement. The framework of the calculation of the failure decoding probability and the description of extended LRPC are expressed in Section V. In Section VI, we present the simulation results and the conclusions are drawn in Section VII.

II. Preliminaries

Let q be a power of prime number p and $\mathbf{u}$ be an element of $\mathbb{F}_{q^m} \setminus \mathbb{F}_q$. In this paper, all coefficients of a vector are in the finite field $\mathbb{F}_{q^m}$. Let $\mathbb{F}_q^{m\times N}$ denote the set of all $m \times N$ matrices over $\mathbb{F}_q$ such that $m \geq N$ and let $\mathbf{b} = \{b_1, b_2, \ldots, b_m\}$ be a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.

Let $(x_1, \ldots x_n)$ n elements of $\mathbb{F}_{q^m}$. The $\mathbb{F}_q$-subspace generated by these elements is denoted $\langle x_1, \ldots, x_n\rangle$. If E and F are two subspaces of $\mathbb{F}_{q^m}^N$, then $\langle E.F\rangle$ denotes the subspace generated by the product of elements of E and F, ie $\langle E.F\rangle = \langle e_i f_j\rangle$ where the (e_i) (respectively the (f_j)) are a basis of E (respectively F). If X is a matrix in $\mathbb{F}_q^{m\times N}$, the row space of a matrix X is denoted by $\langle X\rangle$.

As it has been shown in [10], the number of t-dimensional subspace of an m-dimensional vector space over $\mathbb{F}_q$ is the Gaussian coefficient calculated by

$$\begin{bmatrix} m \\ t \end{bmatrix} \triangleq \prod_{i=0}^{t-1} \frac{q^m - q^i}{q^t - q^i}. \tag{1}$$

Hence, we can deduce from Equation (1) the number of matrices of rank t in the space $\mathbb{F}_q^{m\times N}$, which is

$$S(m, N, q, t) = \prod_{i=0}^{t-1} \frac{(q^m - q^i)(q^N - q^i)}{q^t - q^i}. \tag{2}$$

Let Y_1 and Y_2 be two $m \times N$ matrices over $\mathbb{F}_q$. The row space of a matrix Y_1 is denoted by $\langle Y_1\rangle$. It means that the the space $\langle Y_1\rangle$ is generated by the rows of the matrix Y_1. Then, we have

$$\left\langle \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \right\rangle = \langle Y_1\rangle + \langle Y_2\rangle. \tag{3}$$

Therefore

$$\begin{aligned} rank \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} &= dim\left(\langle Y_1\rangle + \langle Y_2\rangle\right) \\ &= rank(Y_1) + rank(Y_2) - dim(\langle Y_1\rangle \cap \langle Y_2\rangle). \end{aligned} \tag{4}$$

Let $\mathbf{u}$ be an element of $\mathbb{F}_{q^m} \setminus \mathbb{F}_q$ and E be a subspace of $\mathbb{F}_{q^m}$ of dimension r over $\mathbb{F}_q$. We suppose that $2r \ll m$ and we investigate the typical dimension of the subspace $E + \mathbf{u}E$. We rely on the following observation:

Proposition 1. *The probability that $E + \boldsymbol{u}E$ is of dimension $2r$ is given by*

$$\mathbb{P}(dim(E + \boldsymbol{u}E) = 2r) \approx 1 - \frac{q^{2r} - q^{r+1}}{q^m - q}.$$

Proof. Let us take a fixed r-dimensional subspace E in $\mathbb{F}_{q^m}$. Suppose that the dimension of $E + \mathbf{u}E$ is less than $2r$ for $\mathbf{u}$ randomly chosen in $\mathbb{F}_{q^m} \setminus \mathbb{F}_q$. It means that: $\exists (e_1, e_2) \in E^2$, that verifies $\mathbf{u}e_2 = e_1$. Now, we compute the number of possibilities of choosing $\mathbf{u}$ that verifies $\mathbf{u} = e_1 e_2^{-1}$, for $(e_1, e_2) \in E^2$. The number of possible values of (e_1, e_2) is at most q^{2r} and since $\mathbf{u}$ is not in $\mathbb{F}_q$ the case $(\alpha e, e)$ for $\alpha \in \mathbb{F}_q$ and $e \in E$ is not possible. Thus, the number of possibilities to choose $\mathbf{u}$ that verifies $\mathbf{u} = e_1 e_2^{-1}$, for $(e_1, e_2) \in E^2$ is $q^{2r} - q^{r+1}$. The number of possible values of $\mathbf{u}$ is $q^m - q$. □

Let A be a matrix in $\mathbb{F}_q{}^{2r\times N-k}$ and suppose that $2r \leq N - k$. By using (2), the probability that A is a full rank matrix is given by

$$\mathbb{P}(rank(A) = 2r) = \prod_{i=0}^{2r-1} (1 - q^{i-(N-k)}). \tag{5}$$

Let E be a subspace of dimension r over $\mathbb{F}_q$. Let s be a vector in $E + \mathbf{u}E$ of length $N - k$. We have the following proposition:

Proposition 2. *The probability that the subspace $\langle s\rangle$ is of dimension $2r$ over $\mathbb{F}_q$ is given by*

$$\begin{aligned} &\mathbb{P}(dim(\langle s\rangle) = 2r) \approx \\ &\left(1 - \frac{q^{2r} - q^{r+1}}{q^m - q}\right) \prod_{i=0}^{2r-1} (1 - q^{i-(N-k)}). \end{aligned}$$

Proof. Suppose that dimension of $E + \mathbf{u}E$ is $2r$ and let $\{E_1, E_2, ..., E_r, \mathbf{u}E_1, \mathbf{u}E_2, ..., \mathbf{u}E_r\}$ be a basis of $E + \mathbf{u}E$. All coefficients of the vector s are in $E + \mathbf{u}E$ by definition of s. The vector s can be written as follows:

$$s = (E_1, ..., E_r, \mathbf{u}E_1, ..., \mathbf{u}E_r) \times A,$$

where, A is a matrix in $\mathbb{F}_q{}^{2r\times N-k}$. Since the coefficients of s are random elements of $E+\mathbf{u}E$, the matrix A is also random. The probability that the set of all coefficients of s generates the whole space is the probability that A is a full rank matrix. From Equation (5), the probability that a random matrix A is full rank is $\prod_{i=0}^{2r-1} (1 - q^{i-(N-k)})$.

Now, the probability that $dim(E + \mathbf{u}E) = 2r$ is given in the Proposition 1. □

It is interesting to remark that in practice the probability $\mathbb{P}(dim(E + \mathbf{u}E) = dr)$ decreases much more faster to 0 when $dr \ll m$. Thus, the probability that $dim(\langle s \rangle) = dr$ given in the previous proposition can be approximated by:

$$\mathbb{P}(dim(\langle s \rangle) = dr) \approx \prod_{i=0}^{dr-1} (1 - q^{i-(N-k)}). \quad (6)$$

III. Rank metric

In this section, we present some concepts from rank metric coding theory. The reader is referred to [7] and [10] and references therein for further details. A brief overview of concepts relevant to this work can be found in [11]. Afterwards, we introduce Gabidulin codes and LRPC codes, and then we propose modified decoding algorithm of LRPC.

Let $\mathbf{v}$ be a vector of $\mathbb{F}_{q^m}^N$. For $i \in \{1, 2, \ldots, N\}$, we have $v_i = \sum_{j=1}^{m} v_{ij} b_j$ and $\mathbf{v}$ can be interpreted as a matrix $V = (v_{ij}) \in \mathbb{F}_q^{m \times N}$. We can define the rank weight of $\mathbf{v}$ over $\mathbb{F}_q$ as the rank of the associated matrix V, denoted $rank(\mathbf{v})$. The rank distance between two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ of $\mathbb{F}_{q^m}^N$ is defined by $d_r(\mathbf{v}, \mathbf{w}) = rank(\mathbf{v} - \mathbf{w})$. These definitions are independent of the choice of the basis $\{b_1, b_2, \ldots, b_m\}$.

We can now define the support of a vector. This definition differs from the Hamming metric:

Definition 3. *Let $\boldsymbol{v} \in \mathbb{F}_{q^m}^N$. The support of $\boldsymbol{v}$ is the $\mathbb{F}_q - subspace\ of\ \mathbb{F}_{q^m}$ generated by its coordinates:*

$$\mathrm{Supp}(\boldsymbol{v}) = \langle v_1, \ldots, v_N \rangle$$

Definition 4. *A rank code C of length N and dimension k over $\mathbb{F}_{q^m}$ is a subspace of dimension k of $\mathbb{F}_{q^m}^N$ equipped with the rank metric.*

Similar to the minimum Hamming distance for linear codes we define the minimum rank distance of a code C.

Definition 5. *The minimum rank distance of a code C is given by:*

$$d_r^m = min\{rank(\boldsymbol{v}) \mid \boldsymbol{v} \in C, \boldsymbol{v} \neq 0\}.$$

A. Gabidulin codes

Gabidulin codes are introduced in [10], the well-known class of Maximum Rank Distance (MRD) codes. They have been already used successfully in many applications such as cryptography [7], power-line communications [11] and network coding [8].

The Gabidulin code of length N, dimension k and support $g = (g_1, g_2, \ldots, g_N)$ is the set of words obtained by evaluating q-polynomials of q-degree at most $k-1$ at g_1, g_2 and g_N.

$$Gab(g, k, N) = \{(P(g_1), \ldots, P(g_N)) \mid deg_q(P) \leq k-1\}.$$

The decoding of Gabidulin codes can be done based on q-polynomials by using modified Berlekamp-Massey algorithm [12] or extended euclidean algorithm in the non-commutative ring of q-polynomial. They can decode errors of weight up to $\lfloor \frac{N-k}{2} \rfloor$ without probability of failure.

B. Low Rank Parity Check codes

The LRPC code and its parity check matrix are described in the following definition.

Definition 6. *A Low Rank Parity Check code of low rank d, length N and dimension k and with a parity check matrix $\boldsymbol{H} = (h_{ij})$ over $\mathbb{F}_{q^m}$ such that the sub-vector space of $\mathbb{F}_{q^m}$, generated by the coefficients h_{ij} of the matrix $\boldsymbol{H}$, has dimension equals to d.*

Without loss of generality, in this article we are interested in the case $d = 2$. Let $M = (m_{ij})$ be a lower triangular matrix in $\mathbb{F}_q^{2(N-k) \times N}$ and let F be a subspace of $\mathbb{F}_{q^m}$ of dimension 2 generated by the basis $\{1, \mathbf{u}\}$. The matrix $\mathbf{H} = (h_{ij})$ is constructed such that $h_{ij} \in F$. Then, for $1 \leq i \leq N-k, 1 \leq j \leq N$, $h_{ij} = h_{ij1} + \mathbf{u} h_{ij2}$, where h_{ij1} and h_{ij2} are elements of $\mathbb{F}_q$. In order to reduce the complexity of decoding the LRPC codes, we set $h_{ij1} = m_{(2i-1),j}$ and $h_{ij2} = m_{2i,j}$, for $1 \leq i \leq N-k$ and $1 \leq j \leq N$.

Suppose that the error $(e_1, ..., e_N)$ is of weight r and e_i are elements of the error space E of dimension r generated by a basis $\{E_1, E_2, \cdots, E_r\}$. Then, all $e_i (1 \leq i \leq N)$ can be written as $e_i = \sum_{j=1}^{r} e_{ij} E_j$. Suppose that the dimension of the space $E + \mathbf{u}E$ is exactly $2r$ (see Proposition 1). It is then possible to express the system of equations $\mathbf{H}.e^T = s$ over $\mathbb{F}_{q^m}$ into system of equations over $\mathbb{F}_q$, by expressing the syndrome coordinates in the product basis $\{E_1, .., E_r, \mathbf{u}E_1, ..., \mathbf{u}E_r\}$, for $1 \leq i \leq N-k$, as follows:

$$s_i = \sum_{k=1}^{r} s_{i1k} E_k + \mathbf{u} \sum_{k=1}^{r} s_{i2k} E_k.$$

We have $\mathbf{A}_H^r.e'^T = s'$, where $e' = (e_{11}, ..., e_{1r}, e_{21}, ..., e_{nr})$ and $s' =$

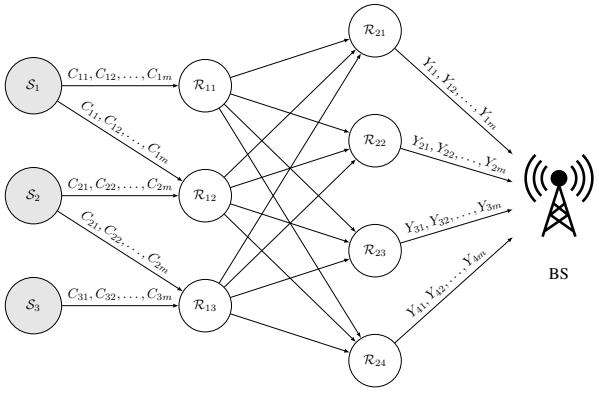

Figure 1: Example of a network composed of 3 sources, number of relay nodes and BS.

$(s_{111}, ..., s_{11r}, ..., s_{(n-k)2r})$. We have detailed the matrix A_H^r in a previous work (see [11]).

The decoding algorithm can fail if the support of s is of dimension strictly smaller than $2r$. Thus we have the following proposition:

Proposition 7. *An LRPC code of rank d, length N and dimension k can decode errors of weight up to $\left\lfloor \frac{N-k}{2} \right\rfloor$ with probability of $\approx 1 - q^{N-k+1-dr}$, where r is the rank of the error.*

Proof. According to Equation (6), we have

$$\begin{aligned} \mathbb{P}(dim(\langle s \rangle) = dr - 1) \approx & \prod_{i=0}^{dr-1} (1 - q^{i-(N-k)}) \\ \approx & \quad 1 - q^{-(N-k+1-dr)} \end{aligned}$$

□

IV. System Model and Problem Formulation

We consider a network comprising a base station BS, s source nodes $\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_s$ and a number of relay nodes. Each source node is attempting to transmit m packets to the BS through relay nodes, as illustrated in Figure 1.

To this end, the source S_i segments data into m packets $u_{i1}, u_{i2}, ..., u_{im}$ of length k , then encodes them using a rank code and transmits the coded packets to the relay nodes. Let $C_{i1}, C_{i2}, ..., C_{im}$ denote the coded packets of node $\mathcal{S}_i$. Hence, $\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_s$ transmit $m \times s$ coded packets of length N to the relay nodes. Each relay node that receives the source packets employs RLNC to combine them and generates coded packets. Note that the coefficients are randomly chosen from $\mathbb{F}_q$, where q is the field size. Afterwards, relays send the generated packets to other relays until the coded packets are received by the destination BS. Let $Y_{11}, Y_{12}, ..., Y_{sm}$ denote the received packets which can be expressed in s block matrices of size $(m \times N)$.

We consider the application of Physical-layer Network Noding (PNC) between the relay nodes as shown in Figure 2. Each stage of the network behaves as independent network and differently of other stages. In this model, relays $\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_l$ send information to a node $\mathcal{N}$ in the next stage. We assume that all nodes are half-duplex. The first time slot corresponds to an uplink phase, in which nodes $\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_l$ transmit their coded packets simultaneously to the node $\mathcal{N}$. The node $\mathcal{N}$ then constructs a network coded packet based on the simultaneously received signals from $\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_l$. The second time slot corresponds to a downlink phase, in which $\mathcal{N}$ attempts to recover the original packet transmitted by $\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_l$ and sends it to next stage nodes.

In the following, we focus on improving the error decoding performance of convolutional code. As shown in Figure 2, nodes $\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_l$ adopt the same convolutional code with length N and k. In this paper, nodes use the same pseudo-random bit-interleaver instead of the conventional bit-interleaver to allocate the coded bits to different modulation levels. Without loss of generality, we focus on BPSK modulation. Our framework can be easily extended to higher order constellations. We assume that the power control and the synchronization at all nodes are perfect.

Consider transmission of l packets to the node $\mathcal{N}$. The received packet is:

$$y = (x_1 h_1 + n_1) + (x_2 h_2 + n_2) + \cdots + (x_l h_l + n_l), \quad (7)$$

where h_i is the channel coefficients of the channels between the node $\mathcal{N}_i$ and the node $\mathcal{N}$. It can be considered as an $N \times N$ diagonal matrix where diagonal coefficients have a Rayleigh distribution with parameter $\sigma = \sqrt{\frac{1}{2}}$. The parameter $n = n_1 + n_2 + \cdots + n_l$ represents the channel additive Gaussian noise (AWGN), where $n_1, n_2, ..., n_l$ are independent Gaussian variables with zero mean and variance $\sigma_1^2 = \sigma_2^2 = \cdots = \frac{N_0}{2}$; i.e. $n \sim \mathcal{N}(0, \frac{mN_0}{2})$.

In order to limit the impact of background noises that are caused by the nature of the wireless channel, we use a convolutional code. Each relay node verifies the integrity of the received packets. If the received packets is erroneous, the node uses convolutional decoder in order to recover the transmitted packet. However, if we combine a big number of packets the total variance of the noise increases significantly and then the convolutional decoder cannot recover the correct codeword.

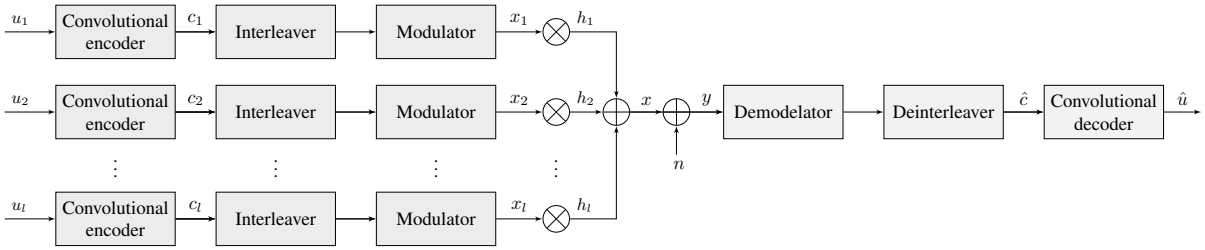

Figure 2: The system model for the inner code.

Also, packets generated by malicious nodes cannot be detected by the convolutional since the latter can use convolutional code too. In this case, relay node that receives the wrong packets combine them with the correct ones generating a wrong packet too. Let r denote the number of erroneous packets caused by the combination of a big number of received packets.

Suppose that r erroneous packets are injected into the network during the transmission of the $m \times s$ source packets. Since packets are randomly combined, errors may affect all the packets. Particularly, errors may affect all the packets of one source. At the BS, the packets of each source are put together in order to apply the rank decoder. By using a classical rank code, the decoding algorithm uses the information of m received packets so as to recover the source packets. For a particular source, if r is bigger than m, the rank error may be bigger than the decoding capability of the rank code. Thus, the BS cannot recover the source packets.

The main idea of this paper is to use the error information of all received packets in order to recover the error basis. Then, we use the error basis in the decoding algorithm to recover packets of each source.

V. Extended LRPC codes

A. Definition and decoding algorithm

Definition 8. *Extended LRPC codes*

An $[n+t,k]$ extended LRPC code of rank d over $\mathbb{F}_{q^m}$ is a code such that it has a parity check matrix $\boldsymbol{H}$ consisting of an $n \times (n-k)$ parity check matrix of an LRPC code, extended by an identity matrix of size t on the first coordinates :

$$H = \begin{pmatrix} I_t & 0 \\ 0 & H_{LRPC} \end{pmatrix}.$$

The probabilistic decoding algorithm of this family of codes is an adaptation of the decoding algorithm of the LRPC codes, to use the fact that the first syndrome coordinates are actually coordinates of the error. In the following we only consider extended LRPC codes of rank 2.

Algorithm 1: Decoding algorithm of the extended LRPC codes

Input: The parity check matrix $\boldsymbol{H}$, the syndrome $\boldsymbol{s}$

Output: The error vector $\boldsymbol{e}$ of rank r

$E' \leftarrow < s_1, \ldots, s_t >$

$S \leftarrow < E'.F > + < s_{t+1}, \ldots, s_{n-k+t} >$

$E \leftarrow F_1^{-1}.S \cap F_2^{-1}.S$, where $\{F_1, F_2\}$ is a basis of F

Try solving $\boldsymbol{H}.\boldsymbol{e}^t = \boldsymbol{s}$ with $\boldsymbol{e} \in E^{n+t}$

return $\boldsymbol{e}$

B. Probability of failure

In order to estimate the decoding failure rate of this algorithm, we need to study the probability that we do not recover the support E of the error. Since we can choose the parity check matrix $\boldsymbol{H}$ such that the system $\boldsymbol{H}.\boldsymbol{e}^t = \boldsymbol{s}$ is invertible, this can not be a source of failure.

Theorem 9. *An $[n+t,k]$ extended LRPC code of rank 2 can decode errors of rank r up to $\lfloor \frac{2t+k}{2} \rfloor$ with probability :*

$$\sum_{j=0}^{min(r-1,t)} \frac{S(t,r,q,j)}{q^{rt}} \times (1 - \frac{S(k+2j,2r,q,2r)}{q^{2r(k+2j)}})$$

Proof. The probability that the first t coordinates of the syndrome span a subspace of dimension j of E is equal to the number of matrices of size $t \times r$ of rank j over $\mathbb{F}_q$ divided by the total number of matrices of size $t \times r$ over $\mathbb{F}_q$: $\frac{S(t,r,q,j)}{q^{rt}}$. If the dimension is exactly r, then the algorithm will succeed. For each other potential dimension, we need to study the probability

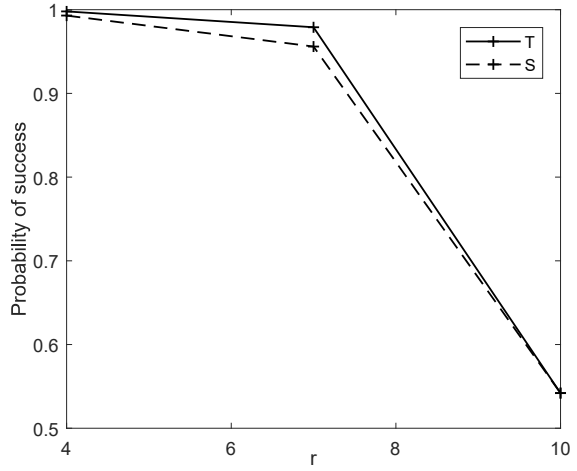

Figure 3: Simulation results for $d = 2, n = 24, k = 15$ and $t = 6$.

that $< E'.F > + < s_{t+1}, \ldots, s_{n-k+t} >$ span the whole product space $\langle E.F \rangle$.

If we write $< E'.F > + < s_{t+1}, \ldots, s_{n-k+t} >$ as a $k + 2j \times 2r$ matrix over $\mathbb{F}_q$, then the probability that these vectors do not span the whole space $< E'.F >$ is $1 - \frac{S(k+2j,2r,q,2r)}{q^{2r(k+2j)}}$, hence the result.

□

We use the expression for the failure decoding probability given in Theorem 9 and compare the resulting values with the simulation results. Figure 3 depicts simulated (S) and theoritical (T) expression of successful decoding probability for $d = 2, n = 24, k = 15, s = 1$ and $t = 6$ as a function of the number of erroneous packets. It can be observed that the system performance is close to the formula given in Theorem 9.

C. Multisource case

Theorem 10. *Using syndromes from N sources, the extended LRPC codes can decode errors of rank r up to $r \leqslant \lfloor \frac{2Nt+Nk}{2} \rfloor$ with probability :*

$$\sum_{j=0}^{min(r-1,Nt)} \frac{S(Nt,r,q,j)}{q^{rNt}} \times (1 - \frac{S(Nk+2j,2r,q,2r)}{q^{2r(Nk+2j)}}).$$

Proof. The proof is similar to the proof of theorem 9, except that we get Nt elements of the vector space E, and Nk elements of $\langle E.F \rangle$ in the syndrome coordinates.

□

VI. NUMERICAL RESULTS

In this section, we investigate the performance of the proposed model via simulation and compare the results of the proposed extended LRPC code with the the classical LRPC code. First, we test the behavior of the two codes in the absence of AWGN noise and then, we evaluate the impact of background noise on both codes.

A. The comparison between extended LRPC and classical LRPC in the absence of AWGN

We set the number of source packets to 80 and the number of source nodes to 1, 2 and 3 respectively. The source coded packets have the same length n. The relevant dimensions of the parity-check matrix are $n = 17$, $k = 10$, $d = 2$ and $t = 3$. We use a binary phase shift keying (BPSK).

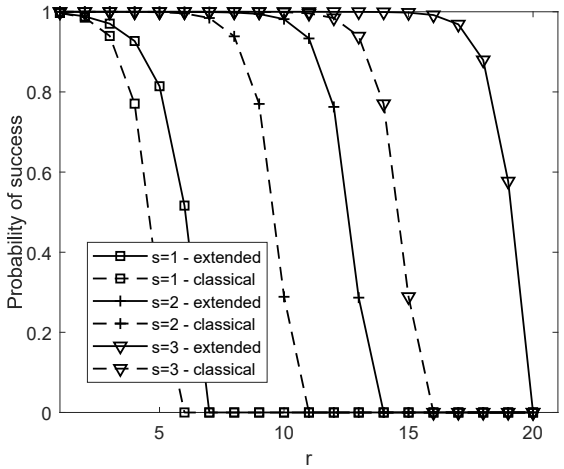

Figure 4: Probability of successful decoding for $n = 17$, $k = 10, t = 3$ using 1, 2 and 3 sources.

Figure 4 illustrates the probability of successful decoding as a function of the number of erroneous packets injected into the network for different numbers of sources. It can be observed that extended LRPC has a good behavior compared to the classical LRPC. By increasing the number of sources, the gap between the two graphs becomes increasingly important. This is because extended LRPC code has $s \times t/2$ additional information of the error support that uses in the decoding process.

B. The comparison between extended LRPC and classical LRPC in the presence of AWGN

In the second experiment we compare the performance of the extended LRPC and classical LRPC in the presence of additive white Gaussian noise. We fix the number of erroneous packets injected into the network to 4.

We use extended LRPC and classical LRPC as *outer codes*. Then, the coded packets are coded again using convolutional code at the source nodes, and transmitted to the next relays. At the intermediate levels, we use the classical RLNC. For convolutional encoder, with a standard $rate = \frac{1}{2}$ and $K = 7$, we use an interleaver to improve the error correction.

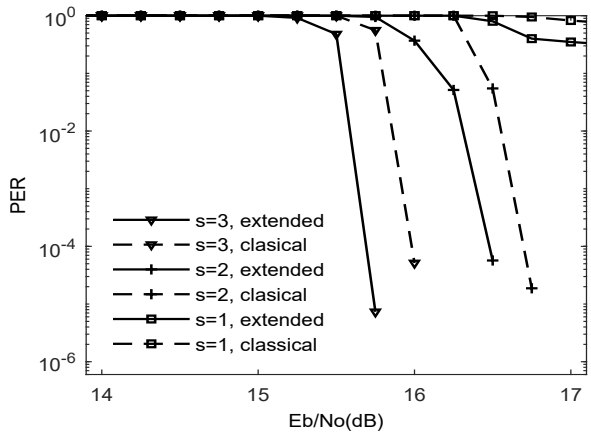

Figure 5: Packet Error Rate as a function of SNR for $r = 4$ using 1, 2 and 3 sources.

We can observe, in Figure 5, that the extended LRPC is about $0.4dB$ better than the classical LRPC. The use of a rank code does not have a beneficial contribution regards to the channel errors. This is because of the property of the white noise, each symbol has a big probability to generate a rank error and therefore reducing the error-correction capability. This is the reason of using a convolutional code to reduce the channel errors impact. It is obvious that the performance of both rank codes deteriorate for $s = 1$ this is because the decoding failure probability of extended LRPC and classical LRPC are affected by the rank error in the case when $s = 1$.

VII. Conclusion

In this paper, we proposed a new family of LRPC codes, extended LRPC codes, which are particularly well suited for use in multisource network using RLNC. We propose a new decoding algorithm that takes into account the fact that the information comes from multiples sources, which is not possible when using Gabidulin codes, and reduces the decoding failure rate over the classical LRPC codes.

The considered scenario takes into account not only errors caused by the nature of the wireless channel, but also errors introduced by a malicious users or due to node failures. In fact, we use extended LRPC as an outer code and we use the convolutional code as an inner code to deal with the wireless channel errors. We have derived analytically the exact expression for the decoding probability of extended LRPC codes. Numerical results have shown that both the simulation and the theoretical expression for the decoding probability of extended LRPC codes are very tight and accurately predict the decoding probability. Our analysis has also exposed the clear benefits of the extended LRPC in terms of recovery accuracy compared to both the classical LRPC codes and the Gabidulin codes.

References

[1] R. Ahlswede, N. Cai, S. yen Robert Li, and R. W. Yeung, "Network information flow," *IEEE TRANSACTIONS ON INFORMATION THEORY*, vol. 46, no. 4, pp. 1204–1216, 2000. *1*

[2] T. Ho, R. Koetter, M. Medard, D. R. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," 2003. *1*

[3] A. Fiandrotti, R. Gaeta, and M. Grangetto, "Simple countermeasures to mitigate the effect of pollution attack in network coding-based peer-to-peer live streaming," *IEEE Transactions on Multimedia*, vol. 17, pp. 562–573, April 2015. *1*

[4] T. Ho, B. Leong, R. Koetter, M. Medard, M. Effros, and D. R. Karger, "Byzantine modification detection in multicast networks with random network coding," *IEEE Transactions on Information Theory*, vol. 54, pp. 2798–2803, June 2008. *1*

[5] T. Ho, R. Koetter, M. Medard, D. R. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, pp. 442+, IEEE, June 2003. *1*

[6] S. Plass, G. Richter, and A. H. Vinck, "Coding schemes for crisscross error patterns," *Wireless Personal Communications*, vol. 47, no. 1, pp. 39–49, 2008. *1*

[7] P. Gaborit, G. Murat, O. Ruatta, and G. Zémor, "Low rank parity check codes and their application to cryptography," in *Proc. WCC*, pp. 168–180, 2013. *1, 3*

[8] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Transactions on Information Theory*, vol. 54, pp. 3579–3591, Aug 2008. *1, 3*

[9] I. El Qachchach, O. Habachi, J.-P. Cances, and V. Meghdadi, "Efficient multi-source network coding using low rank parity check code," in *Wireless Communications and Networking Conference (WCNC), 2018 IEEE*, pp. 1–6, IEEE, 2018. *1*

[10] E. M. Gabidulin, "Theory of codes with maximum rank distance.," *Problems of Information Transmission (English translation of Problemy Peredachi Informatsii)*, vol. 21, no. 1, 1985. *2, 3*

[11] A. K. Yazbek, I. EL Qachchach, J.-P. Cances, and V. Meghdadi, "Low rank parity check codes and their application in power line communications smart grid networks," *International Journal of Communication Systems*, 2017. *3, 4*

[12] G. Richter and S. Plass, "Fast decoding of rank-codes with rank errors and column erasures," in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, pp. 398–398, IEEE, 2004. *3*