An Upper Bound on the Outage Probability of Random Linear Network Codes with Known Incidence Matrices

Birgit Schotsch[1], Michael Cyran[2], Johannes B. Huber[2], Robert F.H. Fischer[3], Peter Vary[1]
[1]Institute of Communication Systems and Data Processing, RWTH Aachen University, Aachen, Germany
[2]Lehrstuhl für Informationsübertragung, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[3]Institut für Nachrichtentechnik, Universität Ulm, Ulm, Germany
{schotsch|vary}@ind.rwth-aachen.de, {cyran|huber}@LNT.de, robert.fischer@uni-ulm.de

***Abstract*—Random linear network coding (RLNC) is a method to maximise the information flow in a network by forming random linear combinations over a finite field $\mathbb{F}_q$ of the received information packets at each intermediate node. The network between one source node and one destination node acts as a linear map $\mathbb{F}_q^n \to \mathbb{F}_q^N$, which is represented by the *network channel matrix*. The connectivity within the network is assumed to be given, i.e. it is considered to be fixed but arbitrary and thus, the incidence matrix of the network is said to be known. The optimal decoding method is equivalent to solving a consistent system of linear equations over the respective finite field, e.g. by means of Gaussian elimination. Therefore, decoding is only successful if the respective square or tall network channel matrix has full column rank. Since the incidence matrix of the network is given, there is one degree of randomness less compared to the usual notion of random matrices. By exploiting similarities of RLNC with Luby transform (LT) coding, a method to establish rateless erasure resilience, which is also based on random matrices over finite fields, we derive an upper bound on the outage probability for RLNC with known incidence matrices.**

I. Introduction and System Model

Network coding [1] is a promising strategy for information dissemination over packet-switched networks, which enables to achieve the network capacity. In contrast to routing, where intermediate nodes in a network just copy and retransmit their incoming messages, intermediate nodes in a network with network coding transmit *functions* of their incoming packets. The latter allows for a more efficient network usage. Particularly in linear network coding [2], intermediate nodes compute linear combinations of their incoming packets with factors from some finite field $\mathbb{F}_q$ of size q. Thus, the n packets (or symbols) sent by the source node $\mathcal{S}$ are linearly transformed en route to the destination node $\mathcal{D}$ and the communication network can be abstracted to a linear map $\mathbb{F}_q^n \to \mathbb{F}_q^N$ which is described by the so-called *network channel matrix*[1] $\mathbf{A} = [a_{i,j}] \in \mathbb{F}_q^{N\times n}$. The input-output relation of the considered scenario is referred to as the *multiplicative matrix channel* (MMC) [3]

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1}$$

where the n transmit symbols and N receive symbols form the transmit vector $\mathbf{x} = (x_0, \ldots, x_{n-1})^\mathsf{T} \in \mathbb{F}_q^n$ and the receive vector $\mathbf{y} = (y_0, \ldots, y_{N-1})^\mathsf{T} \in \mathbb{F}_q^N$, respectively. This is illustrated in Fig. 1 by means of a unicast linear network coding scenario which will be used as an example throughout this paper. The linear factors at each intermediate node are denoted *local encoding vector* or *local encoding kernel* [4]. In order to recover the data $\mathbf{x}$ from the receive vector $\mathbf{y}$, the above consistent system of linear equations (1) needs to be solved by means of Gaussian elimination or an equivalent yet computationally more efficient algorithm (cf. e.g. [5]–[9]). Moreover, to fully recover the original data $\mathbf{x}$, the network channel matrix needs to have full column rank, i.e. rank n.

Currently, there exist two methods to create proper encoding kernels: *deterministic network coding* and *random linear network coding* [10]. In deterministic network coding the local encoding kernels are carefully chosen (with the aid of specific offline algorithms, cf. e.g. [11]), so that the network channel matrix has full column rank, i.e. such that the destination node can recover the information sent by the source node. The main drawbacks of such approaches are a high computational complexity (the local kernels for all intermediate network nodes have to be computed) and that some form of cooperation between the nodes is needed. In this paper we assume that intermediate nodes are not able to cooperate.

Random linear network coding (RLNC) overcomes these issues. The conventional prerequisite is that the considered network is sufficiently dense and therefore can be modeled by a dense random network channel matrix A whose entries are i.i.d. and equiprobably sampled from $\mathbb{F}_q$. A so created random matrix is also refered to as standard random ensemble. Due to the randomness of this approach, however, it is not guaranteed that the matrix A has full column rank. The *outage probability*

$$P_{\text{out}} = \Pr\{\text{rank}(\mathsf{A}) < n\} \tag{2}$$

for RLNC in such usually examined networks, i.e. networks which are described by a *dense* network channel matrix or

This work has been supported in part by the UMIC Research Centre, RWTH Aachen University and by the German Research Foundation (DFG) within the framework COIN under grants HU 634/11-3 and FI 982/4-3.

[1]We use the following notation: scalars are written in italic type (e.g. x). Boldfaced lower case letters denote column vectors (e.g. $\mathbf{x}$), while boldfaced capital letters denote matrices (e.g. $\mathbf{X}$). The corresponding random variables are set in sans serif font, e.g. x for random variables, x for random vectors and X for random matrices. In case of matrices, the product symbol denotes multiplication from the left, e.g. $\prod_{i=1}^{3} \mathbf{X}_i = \mathbf{X}_3\mathbf{X}_2\mathbf{X}_1$.

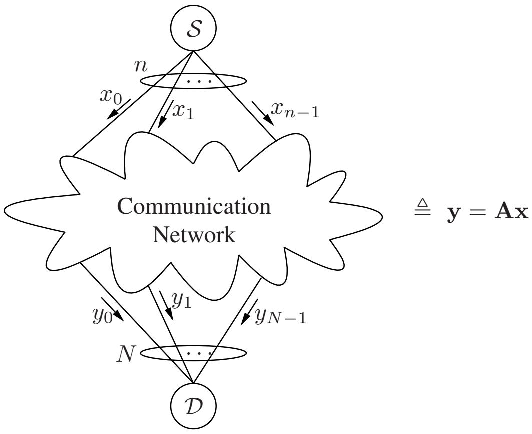

Fig. 1. Unicast linear network coding scenario.

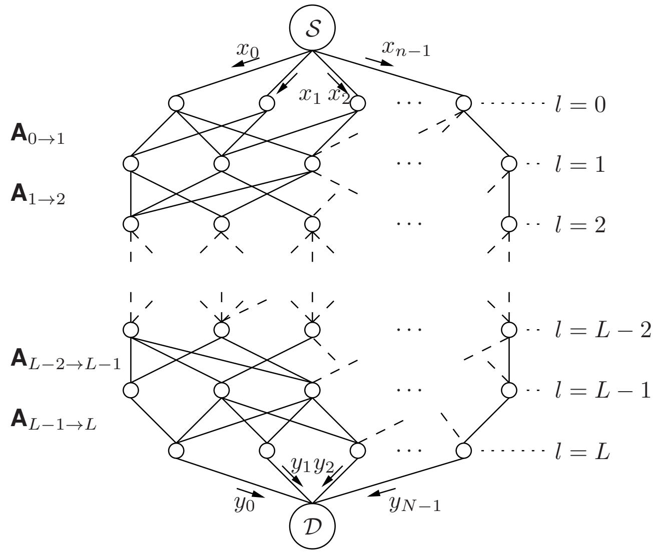

Fig. 2. Unicast scenario with random linear network coding, where the random network channel matrix A corresponds to the product of multiple interlayer random matrices $\mathsf{A}_{l\to l+1}$, i.e. the network is represented by an equivalent layered structure.

more specifically a matrix A whose entries are i.i.d. and equiprobably sampled from $\mathbb{F}_q$, is well known (cf. e.g. [12])

$$P_{\text{out}} = 1 - \prod_{i=N-n+1}^{N} (1 - q^{-i}). \tag{3}$$

However, the "dense network assumption" is not always valid and the structure of a network is usually not taken into account, though both structure and density have a significant influence on the outage probability.

A. Layered Networks

An arbitrary acyclic network can be transformed into an equivalent *layered* network by introducing additional redundant single-input, single-output nodes. The end-to-end *random* network channel matrix $\mathsf{A} \in \mathbb{F}_q^{n_L \times n_0}$, with $n_0 = n$ and $n_L = N$, can then be obtained as the product of all L interlayer *random* matrices $\mathsf{A}_{l\to l+1} \in \mathbb{F}_q^{n_{l+1}\times n_l}$ in a *layered* network as depicted in Fig. 2, i.e.

$$\mathsf{A} = \prod_{l=0}^{L-1} \mathsf{A}_{l\to l+1}, \tag{4}$$

where an interlayer matrix $\mathsf{A}_{l\to l+1}$ represents the linear transformation between layer l and layer $l+1$ of intermediate nodes, with $0 \le l \le L-1$. Accordingly, the total number of intermediate layers is $L+1$ and the number of intermediate nodes in layer l is n_l. The entries in the local encoding kernels of the interlayer matrices are chosen independently and uniformly at random from the set of non-zero Galois field elements $\mathbb{F}_q \setminus \{0\}$, since it is assumed that intermediate nodes act locally and are not able to cooperate with each other. This proposed layering technique does not change the outage probability of the overall network, but it enables the factorisation of the end-to-end network channel matrix A in (4) and thus marks a pivotal step towards a more structured, rigorous and general analysis of RLNC which is also valid for sparse networks.

The paper is organised as follows: in Section II we explain how a dense overall random network channel matrix A may arise in RLNC even if the interlayer matrices are not dense and give reasons why one should be careful in analysing such an RLNC system in the conventional way by using (3). We further examine networks with two layers of intermediate nodes of which the network topology is assumed to be known, i.e. the incidence matrix M of the network graph is fixed but arbitrary:

$$\mathrm{M} = [m_{i,j}], \text{ where } m_{i,j} = \begin{cases} 1 & \text{if } \mathsf{a}_{i,j} \neq 0 \\ 0 & \text{if } \mathsf{a}_{i,j} = 0. \end{cases} \tag{5}$$

Finally, in Section III an upper bound on the outage probability of two-layer networks with known incidence matrices is derived, supported by a numerical evaluation as well as Monte Carlo simulations. For the derivation we adopt methods as applied for instance in [13], [14]. The obtained upper bound is particularly suited to analyse network topologies which result in a *sparse* network channel matrix.

II. Random Linear Network Coding – A Density Analysis –

The conventional approach to analyse RLNC is based on the assumption that the end-to-end network channel matrix A is a random matrix with i.i.d. entries uniformly chosen from $\mathbb{F}_q$, i.e. that A is a dense random matrix with $\Pr\{\mathsf{a}_{i,j} \neq 0\} = 1 - 1/q$. However, this assumption is only justified if the following conditions are met: the considered network is highly meshed, i.e. the interlayer matrices are not too sparse, and its equivalent layered representation consists of more than two layers of intermediate nodes, i.e. it is represented by more than one interlayer matrix.

The sparser the interlayer matrices, the more layers are required to obtain a dense end-to-end network channel matrix A. This can be explained by examining the probability that a fixed but arbitrary entry in the product of two interlayer matrices equals zero or equivalently that the scalar product of two random vectors is zero.

Lemma 1. *The probability that the scalar product of two independent random vectors* $\mathbf{a} \in \mathbb{F}_q^n$ *and* $\mathbf{b} \in \mathbb{F}_q^n$ *with given Hamming weight distributions is zero is equal to*

$$\Pr\left\{\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0\right\} = \sum_{w_a=0}^{n} \sum_{w_b=0}^{n} \Pr\left\{\|\mathbf{a}\|_{\mathrm{H}} = w_a\right\} \Pr\left\{\|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} \times \sum_{s=0}^{n} \frac{\binom{w_a}{s}\binom{n-w_a}{w_b-s}}{\binom{n}{w_b}} \cdot \frac{1}{q}\left(1-(1-q)^{1-s}\right), \quad (6)$$

where $\|\cdot\|_{\mathrm{H}}$ *denotes the Hamming weight of a vector.*

Proof: Since the two vectors $\mathbf{a} \in \mathbb{F}_q^n$ and $\mathbf{b} \in \mathbb{F}_q^n$ are independent, the probability that their scalar product is zero can be written as

$$\begin{aligned}\Pr&\left\{\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0\right\} \\ &= \sum_{w_a=0}^{n} \sum_{w_b=0}^{n} \Pr\left\{\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0 \,\middle|\, \|\mathbf{a}\|_{\mathrm{H}} = w_a,\, \|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} \\ &\times \Pr\left\{\|\mathbf{a}\|_{\mathrm{H}} = w_a\right\} \Pr\left\{\|\mathbf{b}\|_{\mathrm{H}} = w_b\right\}. \end{aligned} \quad (7)$$

Let $\mathbf{v} = (\mathsf{v}_0, \mathsf{v}_1, \ldots, \mathsf{v}_{n-1})^{\mathsf{T}}$ with $\mathsf{v}_j = \mathsf{a}_j \mathsf{b}_j$, where v_j, a_j, and b_j are the j^{th} elements of the vectors $\mathbf{v}$, $\mathbf{a}$, and $\mathbf{b}$, respectively, then $\mathbf{a}^{\mathsf{T}}\mathbf{b} = \sum_{j=0}^{n-1} \mathsf{v}_j$ and

$$\begin{aligned}\Pr&\left\{\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0 \,\middle|\, \|\mathbf{a}\|_{\mathrm{H}} = w_a,\, \|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} \\ &= \Pr\left\{\sum_{j=0}^{n-1} \mathsf{v}_j = 0 \,\middle|\, \|\mathbf{a}\|_{\mathrm{H}} = w_a,\, \|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} \\ &= \sum_{s=0}^{n} \Pr\left\{\|\mathbf{v}\|_{\mathrm{H}} = s \,\middle|\, \|\mathbf{a}\|_{\mathrm{H}} = w_a,\, \|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} \\ &\times \Pr\left\{\sum_{j=0}^{n-1} \mathsf{v}_j = 0 \,\middle|\, \|\mathbf{v}\|_{\mathrm{H}} = s\right\}. \end{aligned} \quad (8)$$

The probability of occurrence of exactly s non-zero elements in $\mathbf{v}$ is

$$\Pr\left\{\|\mathbf{v}\|_{\mathrm{H}} = s \,\middle|\, \|\mathbf{a}\|_{\mathrm{H}} = w_a,\, \|\mathbf{b}\|_{\mathrm{H}} = w_b\right\} = \frac{\binom{w_a}{s}\binom{n-w_a}{w_b-s}}{\binom{n}{w_b}}. \quad (9)$$

The last term in (8) corresponds to the number $N_0(s, q)$ of possibilities that s non-zero $\mathbb{F}_q$-elements add up to zero, taking the elements' order (i.e. the succession of the elements) into account, divided by the number $N(s, q)$ of all possibilities to draw s times with replacement from the set of the $q-1$ non-zero $\mathbb{F}_q$-elements again taking the order into account:

$$\Pr\left\{\sum_{j=0}^{n-1} \mathsf{v}_j = 0 \,\middle|\, \|\mathbf{v}\|_{\mathrm{H}} = s\right\} = \frac{N_0(s,\, q)}{N(s,\, q)}. \quad (10)$$

The problem of determining $N_0(s, q)$ is equivalent to finding the number of closed walks of length s in a complete graph of size q from some fixed but arbitrary vertex back to itself of which a closed form expression can be found, e.g. in [15]

$$N_0(s,\, q) = \frac{1}{q}\left[(q-1)^s + (q-1)(-1)^s\right]. \quad (11)$$

With $N(s, q) = (q-1)^s$ the expression in (10) results in

$$\Pr\left\{\sum_{j=0}^{n-1} \mathsf{v}_j = 0 \;\middle|\; \|\mathbf{v}\|_{\mathrm{H}} = s\right\} = \frac{1}{q}\left(1-(1-q)^{1-s}\right). \quad (12)$$

Finally, inserting (9) and (12) into (8), and the resulting expression into (7) concludes the assertion. ■

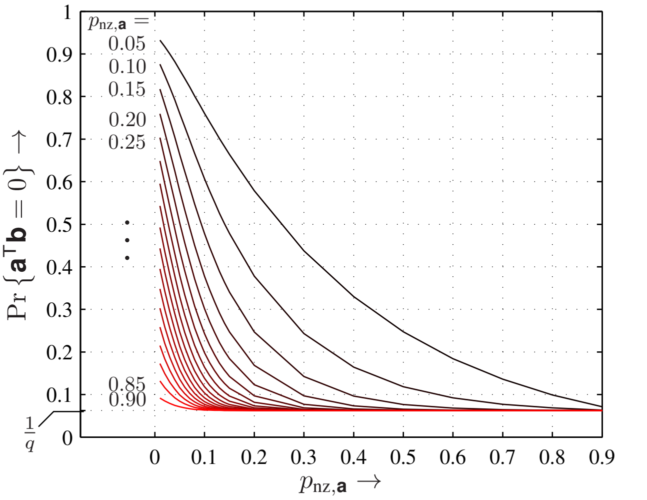

Fig. 3. Evaluation of (6) for two random vectors $\mathbf{a}$ and $\mathbf{b}$ over $\mathbb{F}_{16}$ of length $n = 50$ with expurgated binomial Hamming weight distributions.

Eq. (6) is evaluated in Fig. 3 for two random vectors $\mathbf{a}$ and $\mathbf{b}$ over $\mathbb{F}_{16}$ of length $n = 50$, whose Hamming weights w_a and w_b each follow an *expurgated* binomial distribution

$$\Pr\left\{\|\mathbf{a}\|_{\mathrm{H}} = w_a\right\} = \frac{1}{1-(1-p_{\mathrm{nz},\mathbf{a}})^n}\binom{n}{w_a} p_{\mathrm{nz},\mathbf{a}}^{w_a}\left(1-p_{\mathrm{nz},\mathbf{a}}\right)^{n-w_a}, \quad (13)$$

with $w_a = 1, \ldots, n$. The distribution of $\|\mathbf{b}\|_{\mathrm{H}}$ is defined analogously. Here, the term "expurgated" means that the probability of occurrence of an all-zero vector is forced to zero. The parameters $p_{\mathrm{nz},\mathbf{a}}$ and $p_{\mathrm{nz},\mathbf{a}}$ then denote the probabilities that a fixed but arbitrary entry in $\mathbf{a}$ or $\mathbf{b}$ is equal to a non-zero $\mathbb{F}_q$-element before expurgation. Despite the expurgation, these two parameters are approximately equal to the density of the vectors, i.e. the expected relative number of non-zero entries.

In Fig. 3 it can be observed that $\Pr\left\{\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0\right\}$ approaches $1/q$ already for medium densities of the two vectors $\mathbf{a}$ and $\mathbf{b}$. The results also imply that a matrix that is obtained as the product of two sparse matrices is much denser than either of them. Furthermore, one can conclude that multiplying *a few* (sparse) random matrices results in a product matrix whose density approaches $1 - 1/q$, while the multiplication of *many* (sparse) random matrices yields a product matrix whose density equals $1 - 1/q$ almost surely. While this result may not be surprising, it is now possible to quantify the density of the product matrix by means of Lemma 1.

Although the prerequisite for the conventional RLNC analysis of a dense overall matrix is oftentimes fulfilled, it does not immediately follow that the outage probability may be computed via the conventional approach in (3). Note that in a network with more than two layers, the overall outage probability is lower bounded by the product of the interlayer outage probabilities. This fact is usually overlooked in conventional RLNC analysis.

In contrast to dense network channel matrices, where each element from $\mathbb{F}_q$ occurs with equal probability, sparse channel matrices are biased w.r.t. the zero element. We speak of a sparse matrix if $\Pr\{\mathsf{a}_{i,j}=0\} \gg 1/q$. A sparse network channel matrix arises if the network consists of only very few layers and if the interlayer matrices are sparse, too.

A. Two-Layer Networks with Known Incidence Matrices

In the extreme case of only two layers, i.e. only one interlayer matrix, a very accurate analysis of the outage probability is possible. Particularly if the incidence matrix $\mathbf{M}$ is known.

Two examples for the described case are depicted in Fig. 4. Intermediate nodes in the first and in the second layer are denoted $\mathcal{I}_{0,i}$ and $\mathcal{I}_{1,i}$, respectively. The resulting network channel matrices $\mathbf{A}_\mathrm{a} \in \mathbb{F}_q^{4\times 4}$ and $\mathbf{A}_\mathrm{b} \in \mathbb{F}_q^{6\times 4}$ are

$$\mathbf{A}_\mathrm{a} = \begin{pmatrix} \mathsf{a}_{0,0} & \mathsf{a}_{0,1} & 0 & 0 \\ \mathsf{a}_{1,0} & \mathsf{a}_{1,1} & \mathsf{a}_{1,2} & 0 \\ 0 & 0 & \mathsf{a}_{2,2} & \mathsf{a}_{2,3} \\ 0 & 0 & \mathsf{a}_{3,2} & \mathsf{a}_{3,3} \end{pmatrix}, \qquad (14)$$

$$\mathbf{A}_\mathrm{b} = \begin{pmatrix} \mathsf{a}_{0,0} & \mathsf{a}_{0,1} & 0 & 0 \\ \mathsf{a}_{1,0} & \mathsf{a}_{1,1} & 0 & 0 \\ \mathsf{a}_{2,0} & \mathsf{a}_{2,1} & \mathsf{a}_{2,2} & 0 \\ 0 & 0 & \mathsf{a}_{3,2} & \mathsf{a}_{3,3} \\ 0 & 0 & \mathsf{a}_{4,2} & \mathsf{a}_{4,3} \\ 0 & 0 & 0 & \mathsf{a}_{5,3} \end{pmatrix}. \qquad (15)$$

In this two-layer setting the matrix element $\mathsf{a}_{i,j}$ corresponds to the weight of the edge $(\mathcal{I}_{0,j} \to \mathcal{I}_{1,i})$. These weights are drawn uniformly at random from $\mathbb{F}_q \setminus \{0\}$.

III. Outage Probability of Random Matrices with Known Incidence Matrices

For such sparse network channel matrices as given above the outage probability is *not* equal to the expression in (3). Moreover, an analytic expression of the outage probability for such matrices is not known. Therefore, we derive an upper bound on the outage probability P_out which is valid for arbitrary random matrices over finite fields with a known incidence matrix and particularly for sparse matrices. In the derivation we adopt a method from [13], [14].

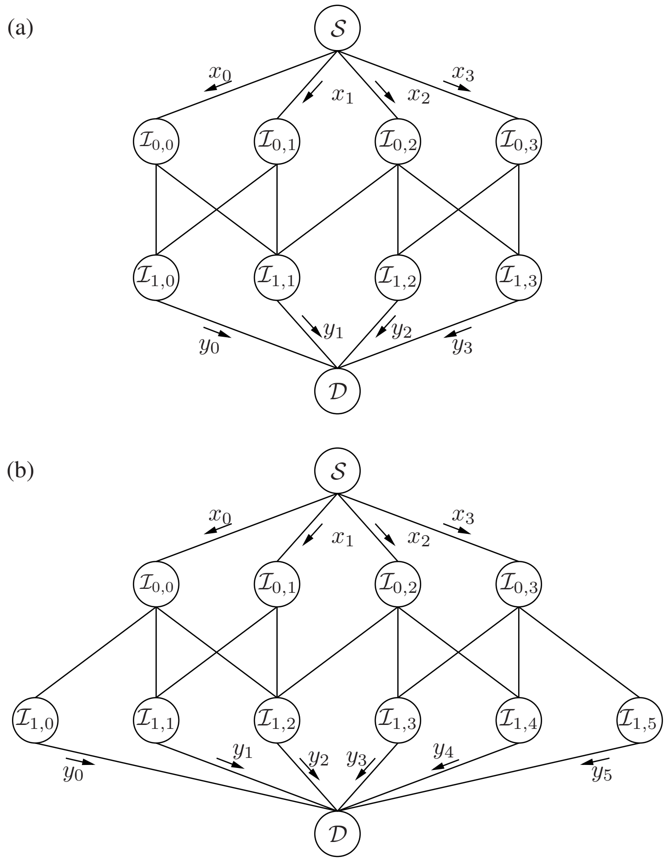

Fig. 4. Two examples of sparsely meshed networks with $(n,\,N) = (4,\,4)$ and $(n,\,N) = (4,\,6)$ whose network channel matrices are given in (14) and (15), respectively.

A. Derivation of an Upper Bound on the Outage Probability

Network channel matrices $\mathbf{A} \in \mathbb{F}_q^{N\times n}$ of sparsely meshed, two-layer networks are random matrices with entries $\mathsf{a}_{i,j}$. At known positions some entries are a priori set to zero, while all others are chosen uniformly at random from $\mathbb{F}_q \setminus \{0\}$, cf. (14) and (15). In other words, the *incidence matrices* $\mathbf{M} \in \mathbb{F}_2^{N\times n}$, cf. (5), are known.

Theorem 1. *An upper bound $\overline{P}_{out}$ on the outage probability of a network channel matrix $\mathbf{A} \in \mathbb{F}_q^{N\times n}$ with known incidence matrix $\mathbf{M}$ is given by (29) on the last page.*

Proof: The outage probability P_out is the probability of $\mathbf{A}$ not having full column rank n

$$P_\mathrm{out} = \Pr\{\mathrm{rank}(\mathbf{A}) < n\}, \qquad (16)$$

which is equivalent to the probability that the cardinality (i.e. the finite number of elements, as spaces over $\mathbb{F}_q$ contain a finite number of elements) of the non-trivial kernel is non-zero

$$P_\mathrm{out} = \Pr\{|\ker(\mathbf{A}) \setminus \{\mathbf{0}\}| \geq 1\}. \qquad (17)$$

This can further be upper bounded as follows

$$P_{\text{out}} = \sum_{j\geq 1} \Pr\left\{|\ker(\mathbf{A}) \setminus \{\mathbf{0}\}| = j\right\} \tag{18}$$

$$\leq \sum_{j\geq 0} j \cdot \Pr\left\{|\ker(\mathbf{A}) \setminus \{\mathbf{0}\}| = j\right\}. \tag{19}$$

Since (19) corresponds to the expected cardinality of the non-trivial kernel of $\mathbf{A}$, the probability P_{out} can be upper bounded accordingly

$$P_{\text{out}} \leq \mathrm{E}\left\{|\ker(\mathbf{A}) \setminus \{\mathbf{0}\}|\right\}. \tag{20}$$

However, this bound can be tightened by a factor of $q-1$ by exploiting the fact that if some $\mathbf{x} \in \ker(\mathbf{A}) \setminus \{\mathbf{0}\}$, then also $c\mathbf{x} \in \ker(\mathbf{A}) \setminus \{\mathbf{0}\}$, $\forall c \in \mathbb{F}_q \setminus \{0\}$. So, in order to bound the outage probability from above, it is sufficient to count just one of the $q-1$ scaled versions of $\mathbf{x}$

$$P_{\text{out}} \leq \overline{P}_{\text{out}} := \frac{1}{q-1} \cdot \mathrm{E}\left\{|\ker(\mathbf{A}) \setminus \{\mathbf{0}\}|\right\} \tag{21}$$

and w.l.o.g. this is accomplished by counting only those vectors $\mathbf{x}$ that have been normalised w.r.t. their first non-zero entry, i.e. vectors $\mathbf{x}$ whose first non-zero entry is $x_i = 1$

$$\overline{P}_{\text{out}} = \frac{1}{q-1} \sum_{\mathbf{x}\in\mathbb{F}_q^n\setminus\{\mathbf{0}\}} \Pr\left\{\mathbf{A}\mathbf{x} = \mathbf{0}\right\} \tag{22}$$

$$= \sum_{\substack{\mathbf{x}\in\mathbb{F}_q^n,\\ x_i=1}} \Pr\left\{\mathbf{A}\mathbf{x} = \mathbf{0}\right\}. \tag{23}$$

To yield $\mathbf{A}\mathbf{x} = \mathbf{0}$, each row $\mathbf{a}_i^{\mathsf{T}}$ of $\mathbf{A}$, with $\mathbf{a}_i \in \mathbb{F}_q^n$ and $0 \leq i < N-1$, must individually fulfil $\mathbf{a}_i^{\mathsf{T}}\mathbf{x} = 0$

$$\overline{P}_{\text{out}} = \frac{1}{q-1} \sum_{\mathbf{x}\in\mathbb{F}_q^n\setminus\{0\}} \prod_{i=0}^{N-1} \Pr\left\{\mathbf{a}_i^{\mathsf{T}}\mathbf{x} = 0\right\}. \tag{24}$$

Before continuing with the determination of $\Pr\left\{\mathbf{a}_i^{\mathsf{T}}\mathbf{x} = 0\right\}$, some quantities need to be defined:

- d_i denotes the row weight of $\mathbf{a}_i^{\mathsf{T}}$, i.e. the number of non-zero positions in row i of $\mathbf{A}$.
- $d_{i_1,\ldots,i_l}$ denotes the number of intersecting non-zero positions in rows $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_l}$.
- Let $\mathbf{v}_i = \left(\mathsf{v}_{i,0}, \mathsf{v}_{i,1}, \ldots, \mathsf{v}_{i,n-1}\right)^{\mathsf{T}}$ with $\mathsf{v}_{i,j} = \mathsf{a}_{i,j}x_j$.
- l_i and $l_{i_1,\ldots,i_l}$ denote the row weights of $\mathbf{v}_i$ and the number of intersecting non-zero positions in $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}$, respectively.
- d_Σ and l_Σ denote the union of non-zero positions in all rows $\mathbf{a}_i$ and $\mathbf{v}_i$, respectively. Applying the inclusion-exclusion principle (cf. e.g. [15]), these quantities can be computed as follows

$$d_\Sigma = d_0 + \cdots + d_{N-1} - d_{0,1} - \cdots - d_{N-2,N-1} + \cdots - \cdots \pm d_{0,\ldots,N-1}\ , \tag{25}$$

$$l_\Sigma = l_0 + \cdots + l_{N-1} - l_{0,1} - \cdots - l_{N-2,N-1} + \cdots - \cdots \pm l_{0,\ldots,N-1} \tag{26}$$

- $d_{\Sigma,i}$ denotes the number of non-zero positions that occur in row $\mathbf{a}_i^{\mathsf{T}}$, but not in any other row

$$\begin{aligned} d_{\Sigma,i} = d_i & \\ & - d_{0,i} - \cdots - d_{i-1,i} - d_{i,i+1} - \cdots - d_{i,N-1} \\ & + d_{0,1,i} + \cdots + d_{0,i-1,i} + d_{0,i,i+1} + \cdots + d_{0,i,N-1} \\ & + d_{1,2,i} + \cdots + d_{1,i-1,i} + d_{1,i,i+1} + \cdots + d_{1,i,N-1} \\ & \vdots \\ & + d_{i,N-2,N-1} \\ & \vdots \\ & \pm d_{0,\ldots,i,\ldots,N-1}. \end{aligned} \tag{27}$$

- $d_{\Sigma,i_1,\ldots,i_l}$ denotes the number of joint non-zero positions that occur only in rows $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_l}$, but not in any other row

$$\begin{aligned} d_{\Sigma,i_1,\ldots,i_l} &= d_{i_1,\ldots,i_l} \\ & - d_{0,i_1,\ldots,i_l} - \cdots - d_{i_1,\ldots,i_l,N-1} \\ & + d_{0,1,i_1,\ldots,i_l} + \cdots + d_{0,i_1,\ldots,i_l,N-1} \\ & + d_{1,2,i_1,\ldots,i_l} + \cdots + d_{1,i_1,\ldots,i_l,N-1} \\ & \vdots \\ & + d_{i_1,\ldots,i_l,N-2,N-1} \\ & \vdots \\ & \pm d_{0,\ldots,i_1,\ldots,i_l,\ldots,N-1} \end{aligned} \tag{28}$$

- $l_{\Sigma,i}$ and $l_{\Sigma,i_1,\ldots,i_l}$ can be computed accordingly.

Combining these quantities with (24) we obtain (29). The first row in (29) contains the summation over the cardinalities of all possible (joint) support sets of $\mathbf{v}_i$ as well as the summation over the weight of $\mathbf{x}$. The term $(q-1)^{w-1}$ counts the number of different realisations of $\mathbf{x}$ with weight w, excluding non-trivial multiplicities, given that $\mathbf{x}$ has a fixed but arbitrary support set. This term together with the binomial coefficients gives the total number of choices of $\mathbf{x}$ (excluding non-trivial multiplicities) which have suitable support sets. The last term of the first row can be derived by the same reasoning as (12) in Section II. ■

B. Simulation Results

In Fig. 5 the simulated outage probabilities of the random network channel matrices $\mathbf{A}_{\text{a}}$ and $\mathbf{A}_{\text{b}}$, corresponding to networks (a) and (b) in Fig. 4, are depicted as cross markers in blue and red, respectively. The corresponding upper bounds on the outage probabilities according to (29) are depicted as solid lines. For comparison, the outage probabilities of dense random matrices of sizes (4×4) and (6×4) are determined according to (3) and are plotted as dashed lines.

From Fig. 5 it is obvious that for sparse network channel matrices, the outage probabilities can deviate significantly from the outage probabilities of dense matrices in (3). A good estimate on the actual outage probability is the introduced upper bound according to (29). It can be seen that this upper

$$\overline{P}_{\text{out}} = \sum_{l_0=0}^{d_0} \cdots \sum_{l_{N-1}=0}^{d_{N-1}} \sum_{l_{0,1}=0}^{d_{0,1}} \sum_{l_{0,2}=0}^{d_{0,2}} \cdots \sum_{l_{N-2,N-1}=0}^{d_{N-2,N-1}} \cdots \sum_{l_{0,\ldots,N-1}=0}^{d_{0,\ldots,N-1}} \sum_{w=\min(1,l_\Sigma)}^{n} (q-1)^{w-1} \left(\prod_{i=0}^{N-1} \frac{1}{q} \left[1-(1-q)^{1-l}\right] \right)$$
$$\times \left(\prod_{i=0}^{N-1} \binom{d_{\Sigma,i}}{l_{\Sigma,i}} \right) \left(\prod_{0\le i_1<i_2\le N-1} \binom{d_{\Sigma,i_1,i_2}}{d_{\Sigma,i_1,i_2}} \right) \left(\prod_{0\le i_1<i_2<i_3\le N-1} \binom{d_{\Sigma,i_1,i_2,i_3}}{d_{\Sigma,i_1,i_2,i_3}} \right) \cdots \binom{d_{\Sigma,0,\ldots,N-1}}{l_{\Sigma,0,\ldots,N-1}} \binom{n-d_\Sigma}{w-l_\Sigma} \quad (29)$$

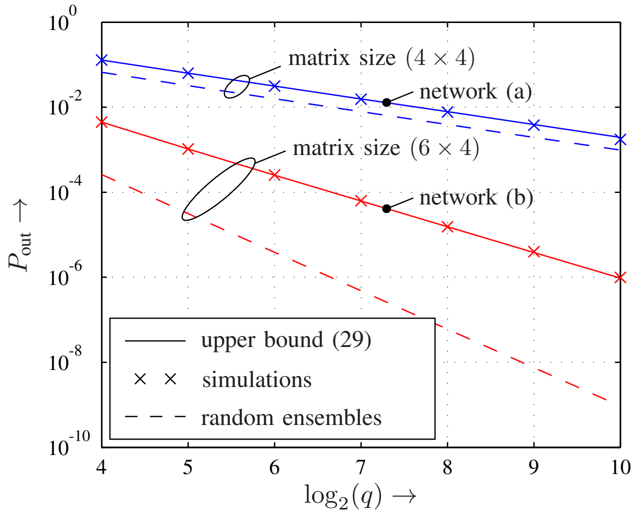

Fig. 5. For the two exemplary networks (a) and (b) of size (4×4) and (6×4), respectively, the simulated outage probabilities (cross markers) and upper bounds on the outage probability according to (29) (solid lines) are depicted. For comparison, the outage probabilities according to (3) of the standard random ensembles of the same sizes are plotted as well.

bound is very tight and nearly coincides with the simulation points.

IV. Conclusion

In this paper we have questioned and discussed the validity of the conventional random linear network coding approach, in which the network is modelled as a dense random matrix over a finite field and where usually the well known expression (3) is used to assess the outage probability. We have introduced a layering technique to transform arbitrary acyclic networks into layered networks which allows a more structured, rigorous and general analysis of RLNC. We have shown how multiple layers of sparsely or even densely meshed networks result in a dense overall network channel matrix and have pointed out that in such a case the outage probability should not be computed via the conventional approach.

Particularly for two-layer networks with known incidence matrices we have derived an upper bound on the outage probability which proves particularly useful for sparse networks, whose outage probability usually deviates significantly from the one of the standard random ensemble. The extension of the upper bound to networks with more than two layers of intermediate nodes is currently under investigation.

References

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[2] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[3] D. Silva, F. Kschischang, and R. Koetter, "Communication over finite-field matrix channels," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1296–1305, Mar. 2010.

[4] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, *Network Coding Theory*. now Publishers, 2005.

[5] D. H. Wiedemann, "Solving Sparse Linear Equations over Finite Fields," *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 54–62, 1986.

[6] H. Pishro-Nik and F. Fekri, "On Decoding of Low-Density Parity-Check Codes over the Binary Erasure Channel," *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 439–454, 2004.

[7] A. Shokrollahi, S. Lassen, and R. Karp, "Systems and Processes for Decoding Chain Reaction Codes Through Inactivation," U.S. Patent 6,856,263, Feb. 2005.

[8] G. Liva, B. Matuz, E. Paolini, and M. Chiani, "Pivoting Algorithms for Maximum Likelihood Decoding of LDPC Codes over Erasure Channels," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2009, pp. 1–6.

[9] E. Paolini, G. Liva, B. Matuz, and M. Chiani, "Maximum Likelihood Erasure Decoding of LDPC Codes: Pivoting Algorithms and Code Design," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3209–3220, 2012.

[10] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On Randomized Network Coding," in *Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, Oct. 2003.

[11] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, and L. M. G. M. Tolhuizen, "Polynomial Time Algorithms for Multicast Network Code Construction," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1973–1982, June 2005.

[12] R. Lidl, H. Niederreiter, and P. M. Cohn, *Finite Fields*. Cambridge University Press, 1997.

[13] B. Schotsch, G. Garrammone, and P. Vary, "Analysis of LT Codes over Finite Fields under Optimal Erasure Decoding," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1826–1829, Sep. 2013.

[14] B. E. Schotsch, "Rateless coding in the finite length regime," Dissertation, IND, RWTH Aachen University, 2014.

[15] R. P. Stanley, *Enumerative Combinatorics*, 2nd ed. Cambridge University Press, 2011, vol. 1.