Hierarchical Performance Analysis on Random Linear Network Coding

Dan Li, Xuan Guang, *Member, IEEE*, Zhiheng Zhou, *Member, IEEE*, Congduan Li, *Member, IEEE*, and Chee Wei Tan, *Senior Member, IEEE*

***Abstract*— Random linear network coding (RLNC) is a promising network coding solution when the network topology information is not fully available to all the nodes. However, in practice, nodes have partial knowledge of the network topology information. Motivated by this, we investigate the performance of RLNC and obtain different upper bounds on the failure probability of RLNC for the network constrained by different partial network topology information. These upper bounds not only improve the existing ones in the literature, but also show that the partial network topology information can bring benefits to the performance analysis of RLNC. On the other hand, it is observed that if more network topology information can be utilized, tighter upper bounds can be obtained, as expected. The upper bounds on two classical networks are compared for demonstration. To obtain a deeper understanding about the performance of RLNC, the asymptotic behavior of RLNC as the field size goes to infinity is also investigated.**

***Index Terms*— Random linear network coding, failure probability, hierarchical performance, upper bound, asymptotic behavior.**

I. Introduction

NETWORK coding as a cornerstone of network information theory [1] has attracted a substantial amount of research attentions in the communication and networking communities. It is well known that, in multicast, the capacity of information flow can be achieved by network coding. The essence of network coding is to encode the incoming messages at network nodes instead of merely performing a receive-and-forward operation. The seminal work by Ahlswede *et al.* [1] has shown that, in a directed acyclic network, the source node can multicast messages to all the sink nodes at capacity as the alphabet size goes to infinity. Li *et al.* [2] further showed that linear network coding (all encodings at network nodes are linear) with finite alphabet is sufficient for optimal multicast by means of a vector space approach. Independently, Koetter and Médard [3] developed an algebraic characterization of linear network coding by means of a matrix approach.

Naturally, network coding theory should be useful in practical networking systems. To implement network coding, Jaggi *et al.* [4] proposed a polynomial-time algorithm for a linear network code so that all coded messages received at each sink node are guaranteed to decode the source messages. This algorithm requires the knowledge of the global network topology information and thus has a high computational cost in exchanges of control packets. However, for many communication networks in practice, the network topology cannot be utilized because for instance, the network is huge in scale, or dynamic in structure. Hence, it is impossible to use such deterministic centralized network coding schemes that utilize global network topology information to construct network codes. To solve this problem, Ho *et al.* [5] proposed a localized randomized approach, called *random linear network coding (RLNC)*, where the strategy is as follows: each node (including the source node) transmits on each of its outgoing edges a linear combination of all the incoming messages, with randomly and uniformly-chosen coding coefficients from some finite field. Chou *et al.* [6] conceived that RLNC can be applied in practical networking systems by dividing a stream of source messages into generations and performing RLNC within each generation, and there are applications such as the peer-to-peer networks [7], [8] and multicast in wireless networks [9], [10].

However, RLNC does not consider the network topology or cooperative coding between the different nodes, it may not be able to achieve the best performance of network coding with respect to the successful decoding at all sinks. Therefore, the performance analysis of RLNC is not only of theoretical interest but also of practical importance. Toward characterizing the performance of RLNC, several failure probabilities of RLNC as the measures of the performance were introduced

Manuscript received April 26, 2017; revised September 16, 2017 and December 18, 2017; accepted December 18, 2017. Date of publication December 29, 2017; date of current version May 15, 2018. This work was supported by National Natural Science Foundation of China under Grant (No. 61771259), the Hong Kong Scholars Program (No. XJ2016028), the University Grants Committee of the Hong Kong SAR, China (Project No. AoE/E-02/08), the Vice-Chancellor's One-off Discretionary Fund of CUHK (Project Nos. VCF2014030 and VCF2015007), Research Grants Council of Hong Kong (No. 11207615), and Science, Technology and Innovation Commission of Shenzhen Municipality (Project No. JCYJ20160229165220746). The associate editor coordinating the review of this paper and approving it for publication was W. Saad. *(Corresponding author: Xuan Guang.)*

D. Li is with the Department of Mathematics, Tianjin University of Technology, Tianjin 300384, China (e-mail: lidaneileen@gmail.com).

X. Guang is with the School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, China, and also with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong SAR, China (e-mail: xguang@nankai.edu.cn).

Z. Zhou was with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong SAR, China. He is now with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: zhzhou0606@gmail.com).

C. Li and C. W. Tan are with the Department of Computer Science, City University of Hong Kong, Hong Kong SAR, China (e-mail: congduan.li@cityu.edu.hk; cheewtan@cityu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2017.2787991

and discussed [5], [11]–[14]. Ho *et al.* [5] showed that with RLNC, each sink node can decode the source messages successfully with probability close to one for a sufficiently large finite field. Following [5], Balli *et al.* [11] formally defined the failure probabilities of RLNC at the sink node and for the network (cf. (3) and (5), respectively) and gave upper bounds on them which improve the ones in [5]. Besides further improving the upper bounds on the failure probability of RLNC at the sink node, Li *et al.* [12] also gave a lower bound on this failure probability. Guang and Fu [13] introduced the *average failure probability of RLNC* (cf. (4)) and then gave several upper bounds on this average failure probability. The (asymptotical) tightness was also considered in [13]. In [11] and [14], RLNC for network error detection/correction was investigated by analyzing both the failure probabilities and the probability mass function of the minimum distance of RLNC.

In addition to the success guarantee of RLNC in the sense of probability, there is a trade-off between the failure probabilities of RLNC and the field size used in RLNC. The performance of RLNC also ties in well with considerations of its robustness and stability. Thus, we always desire to characterize the failure probabilities of RLNC more precisely, in particular, the failure probability of RLNC for the entire network, i.e., the probability that there exists decoding failure at at least one sink node, due to as the dual of the success probability of RLNC for the whole network, i.e., the probability that the decoding is successful for all the sink nodes.

In many communication problems, while the global network topology information is not utilized, some partial network topology information may still be available. Motivated by this fact, we put forward the problem whether we can improve the performance analysis of RLNC by taking advantage of the partial network topology information. In particular, we investigate the upper bounds on the failure probability of RLNC for the entire network when different partial network topology information are available. The obtained upper bounds improve the previous ones in the literature and thus show that the partial network topology information brings benefits to the performance analysis of RLNC.

In this paper, we consider the *single-source network coding model*, i.e., the message generated by the single source node is required to be transmitted to every sink node. For this model, the capacity is completely characterized by a max-flow min-cut bound theorem [1]. In the *multi-source network coding model*, the source nodes generate mutually independent messages and each one of them is multicast to a certain subset of the sink nodes. If the messages from all the source nodes are required to be multicast to the same set of the sink nodes, such a multi-source network coding model is equivalent to a single-source network coding model [1] and how to convert such a multi-source network coding model to a single-source network coding model was also given in [1]. However, for the general cases that different subsets of the sink nodes are required for the multiple source nodes, the characterization of the capacity region is difficult [1], [18]–[20]. For instance, when the network is acyclic, the capacity region can only be characterized implicitly in terms of achievable entropy functions [20] (necessarily involving non-Shannon type information inequalities), and computer-aided methodologies were used to solve these information theoretic problems [21], [22].

The main contributions of this paper are given as follows:

- Based on the different partial network topology information, we obtain different upper bounds on the failure probability of RLNC. These upper bounds improve the previous ones even under the same assumptions on the network topology information as in previous work. This shows that the partial network topology information can benefit the performance of RLNC.
- We observe that if more network topology information can be utilized, tighter upper bounds can be obtained. These different upper bounds are helpful to choose the appropriate parameters (e.g., finite field size) of RLNC with a requirement of a failure probability.
- To obtain a deeper understanding about the performance of RLNC, we study the limiting behavior of the failure probability as the field size goes to infinity. We give the worst case limiting behavior of RLNC that not only proves the asymptotic tightness of the upper bounds but also highlights the major factor affecting the failure probability of the network.

This paper is organized as follows. We first present RLNC and its performance measurements in Section II. In Section III, the hierarchical performance of RLNC is discussed. Specifically, we upper bound the failure probability over the entire network based on different partial network topology information. Numerical results are provided to compare the upper bounds obtained and to show the advantage of our methodologies. In Section IV, we investigate the limiting behavior of RLNC as the field size goes to infinity. We conclude in Section V with a summary of our results and a remark on future research.

II. Preliminaries

Let $G = (V, E)$ be a finite directed acyclic network with a single source node s and a set of sink nodes $T \subseteq V \setminus \{s\}$, where V and E are the sets of nodes and edges, respectively. Let $e \in E$ be a directed edge from node u to node v, where u is called the *tail* of e and v is called the *head* of e, denoted by $\mathrm{tail}(e)$ and $\mathrm{head}(e)$, respectively. For a node v, define $\mathrm{In}(v)$ as the set of incoming edges of v and $\mathrm{Out}(v)$ as the set of outgoing edges of v, i.e., $\mathrm{In}(v) = \{e \in E : \ \mathrm{head}(e) = v\}$ and $\mathrm{Out}(v) = \{e \in E : \ \mathrm{tail}(e) = v\}$. Assume $\mathrm{In}(s) = \emptyset$ and $\mathrm{Out}(t) = \emptyset$, $\forall\, t \in T$. Later, we will introduce the concept of the imaginary channels which are assumed to be the incoming channels of s providing source messages. An index taken from an alphabet can be transmitted on each edge $e \in E$ and parallel edges between two adjacent nodes are allowed. In other words, the capacity of each edge is taken to be 1.

Let ω be the information rate, i.e., the source node s generates ω source symbols per time step that are transmitted to every sink node. It is well known that with network coding, the communication is successful only if ω is not greater than the smallest minimum cut capacity between s and each sink node, i.e., $\omega \le \min_{t\in T} C_t$, where C_t is the minimum cut capacity between s and t. We assume that

s has ω *imaginary* incoming channels providing the source messages to s, denoted by $d_1, d_2, \cdots, d_\omega$, respectively, and let $\mathrm{In}(s) = \{d_1, d_2, \cdots, d_\omega\}$.

In linear network coding, the alphabet is a finite field $\mathcal{F}$ and the source symbols are ω elements in $\mathcal{F}$, say $X_1, X_2, \cdots, X_\omega$. We let $\mathbf{X} = \left[X_1, X_2, \cdots, X_\omega\right]$ and assume that X_i is transmitted to s through the imaginary incoming channel d_i, $1 \le i \le \omega$. Assign $k_{d,e}$ in $\mathcal{F}$ for every adjacent pair of channels (d, e) (the adjacent means $d \in \mathrm{In}(v)$ and $e \in \mathrm{Out}(v)$ for some node v), which is called the *local encoding coefficient* for (d, e). The set of all local encoding coefficients, i.e.,

$$\{k_{d,e} : \text{all adjacent pairs } (d, e) \text{ of channels on } G\},$$

is defined as a *linear network code* on G. In particular, if we independently and randomly assign $k_{d,e}$ from the field $\mathcal{F}$ for every adjacent pair of channels (d, e), the obtained code is a *random linear network code* (RLNC) on $\mathcal{F}$.

For each channel $e \in E$, we use U_e to denote the message transmitted on e. For the source node s, we assume without loss of generality that the ith source symbol is transmitted on the ith imaginary channel d_i, i.e., $U_{d_i} = X_i$. For other channels $e \in E$, the message U_e is calculated by the linear coding function

$$U_e = \sum_{d \in \mathrm{In}(\mathrm{tail}(e))} k_{d,e} U_d, \tag{1}$$

i.e., U_e is a linear combination of U_d for all $d \in \mathrm{In}(\mathrm{tail}(e))$ by applying the local encoding coefficients $k_{d,e}$, $d \in \mathrm{In}(\mathrm{tail}(e))$. Further, we see that, by applying local encodings recursively, U_e is a linear combination of the ω source symbols X_i, $1 \le i \le \omega$. That is, there is an ω-dimensional column vector f_e over $\mathcal{F}$ such that $U_e = \mathbf{X} \cdot f_e$ (also see [15], [16]). This column vector f_e is called the *global encoding kernel* of the channel e that can be determined by the local encoding kernels as follows:

$$f_e = \sum_{d \in \mathrm{In}(\mathrm{tail}(e))} k_{d,e} f_d \tag{2}$$

with boundary condition that the vectors f_{d_i}, $1 \le i \le \omega$, form the standard basis of the vector space $\mathcal{F}^\omega$. For each sink node $t \in T$, let the matrix $F_t \triangleq \left[f_e : \ e \in \mathrm{In}(t)\right]$ of size $\omega \times |\mathrm{In}(t)|$, called the *decoding matrix* of t, and $\mathbf{U}_t = \left[U_e : e \in \mathrm{In}(t)\right]$ be the received message vector of t. Both of them are available at the sink node t and we have the following *decoding equation* $\mathbf{U}_t = \mathbf{X} \cdot F_t$. Immediately, we see that t decodes $\mathbf{X}$ if and only if $\mathrm{Rank}(F_t) = \omega$.

We know that choosing the coefficients randomly can not guarantee the successful decoding at all sinks. To characterize the performance of RLNC, several failure probabilities are proposed in the literature as follows:

- The failure probability of RLNC at a sink node t ([5], [11], [12]):

$$\mathbf{P}_t(G) \triangleq \mathbf{Pr}(\mathrm{Rank}(F_t) < \omega); \tag{3}$$

- The average failure probability of RLNC ([13]):

$$\mathbf{P}_{av}(G) \triangleq \frac{\sum_{t \in T} \mathbf{P}_t(G)}{|T|}; \tag{4}$$

- The failure probability of RLNC for the network G ([5] [11]):

$$\mathbf{P}(G) \triangleq \mathbf{Pr}(\exists\ t \in T \text{ s.t. } \mathrm{Rank}(F_t) < \omega). \tag{5}$$

It is easy to see that the following inequality holds

$$\begin{aligned} \min_{t \in T} \mathbf{P}_t(G) &\le \mathbf{P}_{av}(G) \le \max_{t \in T} \mathbf{P}_t(G) \\ &\le \mathbf{P}(G) \le |T|\mathbf{P}_{av}(G) = \sum_{t \in T} \mathbf{P}_t(G). \end{aligned}$$

Note that $\mathbf{P}_t(G)$ (resp. $\min_{t \in T} \mathbf{P}_t(G)$ and $\max_{t \in T} \mathbf{P}_t(G)$) only characterizes the performance of RLNC for an individual sink node t. Although all sink nodes are involved for $\mathbf{P}_{av}(G)$, logically it only depends on $\mathbf{P}_t(G)$ for every individual sink node. On the contrast, $\mathbf{P}(G)$ is to characterize the performance of RLNC for the entire network and we will focus on $\mathbf{P}(G)$ throughout the paper. Frequent notations used in the paper are listed in Table I.

III. Hierarchical Performance Analysis of RLNC

A. Upper Bounds

Let $G = (V, E)$ be a finite directed acyclic network with a single source s and a set of ℓ sink nodes $T = \{t_1, t_2, \cdots, t_\ell\}$. In this subsection, we will upper bound $\mathbf{P}(G)$ to characterize the performance of RLNC.

We consider $\omega \le C_{t_i}$ for each $t_i \in T$, where C_{t_i} is the minimum cut capacity between s and t_i. The following Menger's theorem shows that the minimum cut capacity between node s to node t and the maximum number of channel-disjoint paths from s to t are really alternative ways to address the same issue.

Edge Version of Menger's Theorem ([23, Th. 6.7]): The maximum number of channel-disjoint paths from node s to node t equals the minimum cut capacity between node s and node t.

Thus, there exist ω channel-disjoint paths from s to t_i, denoted as $P_1^{(t_i)}, P_2^{(t_i)}, \ldots, P_\omega^{(t_i)}$, of which the set is denoted as $\mathcal{P}^{(t_i)}$, i.e.,

$$\mathcal{P}^{(t_i)} = \{P_1^{(t_i)}, P_2^{(t_i)}, \ldots, P_\omega^{(t_i)}\}.$$

Let r_i be the number of the intermediate nodes each of which is passed through by at least one path in $\mathcal{P}^{(t_i)}$ and R be the number of the intermediate nodes passed through by at least one path in $\cup_{t_i \in T} \mathcal{P}^{(t_i)} = \cup_{i=1}^{\ell} \mathcal{P}^{(t_i)}$. Then,

$$\max_{1 \le i \le \ell} r_i \le R \le \sum_{i=1}^{\ell} r_i.$$

Denote these R intermediate nodes by $i_1, i_2, \cdots, i_R$, and without loss of generality we assume that their ancestrally topological order is

$$s \triangleq i_0 \prec i_1 \prec i_2 \prec \cdots \prec i_R \prec \{t_1, t_2, \cdots, t_\ell\}.^1$$

For each sink node t_i, we will use a concept of cuts with respect to the set $\mathcal{P}^{(t_i)}$ of the paths from s to t_i, which is different from the standard concept of cuts in graph theory.

[1]Herein, we use i_0 to denote s for unification of notation.

TABLE I

NOTATION IN THE PAPER

Notation	Definition
$G = (V, E)$	a finite directed acyclic network with the set of the nodes V and the set of the channels E
s	the single source node
t or t_i	a sink node
T	the set of the sink nodes
$\mathrm{tail}(e)$	the tail node of the channel e
$\mathrm{head}(e)$	the head node of the channel e
$\mathrm{Out}(v)$	$= \{e \in E : \ \mathrm{tail}(e) = v\}$
$\mathrm{In}(v)$	$= \{e \in E : \ \mathrm{head}(e) = v\}$
ω	the information rate
C_t	the minimum cut capacity between s and t
$k_{d,e}$	the local encoding coefficient for the adjacent pair of the channels (d, e)
f_e	the global encoding kernel of the channel e
F_t	the decoding matrix of the sink node t
$\mathbf{P}(G)$	the failure probability of RLNC for the network G
$\mathcal{P}^{(t_i)}$	the set of ω channel-disjoint paths from the source node s to the sink node t_i
r_i	the number of the intermediate nodes each of which is passed through by at least one path in $\mathcal{P}^{(t_i)}$
R	the number of the intermediate nodes passed through by at least one path in $\cup_{t_i \in T}\mathcal{P}^{(t_i)}$
M_k	the set of the sink nodes t_i such that at least one path in $\mathcal{P}^{(t_i)}$ passes through the node i_k
N_k	the set of the sink nodes t_i such that not only at least one path in $\mathcal{P}^{(t_i)}$ passes through the node i_k but also i_k is the last non-sink node on the ω paths in $\mathcal{P}^{(t_i)}$
a	$= 1 - \prod_{h=1}^{\omega} (1 - 1/\lvert\mathcal{F}\rvert^h)$
$\mathbf{N}^*$	a RLNC problem $\{G, s, T, \omega\}$
$\Omega(\mathbf{N}^*)$	$= \limsup_{\lvert\mathcal{F}\rvert \to \infty} \lvert\mathcal{F}\rvert \mathbf{P}(G)$, the limiting behavior of $\mathbf{P}(G)$ as the field size goes to infinity
$\Lambda^{+\ *}_{(n,\ell)}$	$= \max_{\mathbf{N}^* \in \mathcal{M}^*_{n,\ell}} \Omega(\mathbf{N}^*)$, the worst case limiting behavior of RLNC

Here, each cut always contains exactly one edge of each path in $\mathcal{P}^{(t_i)}$. Specifically, the first cut is $\mathrm{CUT}_{t_i,0} = \mathrm{In}(s)$, i.e., the set of the ω imaginary incoming channels. Next, the cut $\mathrm{CUT}_{t_i,1}$, passing through $i_0 = s$, is the collection of the first edges (channels) corresponding to each of the ω paths in $\mathcal{P}^{(t_i)}$. We can obtain the cut $\mathrm{CUT}_{t_i,k+1}$ from $\mathrm{CUT}_{t_i,k}$ by passing through i_k as follows:

- if $\mathrm{In}(i_k) \cap \mathrm{CUT}_{t_i,k} \neq \emptyset$, then $\mathrm{CUT}_{t_i,k+1}$ is obtained by replacing each channel in $\mathrm{In}(i_k) \cap \mathrm{CUT}_{t_i,k}$ by its following channel in the corresponding path (the set of all updated channels is $\mathrm{Out}(i_k) \cap \mathcal{P}^{(t_i)}$), while the other channels remain the same as in $\mathrm{CUT}_{t_i,k}$;
- otherwise, $\mathrm{CUT}_{t_i,k+1}$ remains the same as $\mathrm{CUT}_{t_i,k}$.

Iteratively, we can determine all cuts $\mathrm{CUT}_{t_i,k}$ for $i = 1, 2, \cdots, \ell$ and $k = 0, 1, \cdots, R + 1$. Note that the concept of cuts here is similar to the idea of the algorithm for constructing a linear network code in [4], and also appeared in [11] and [13].

Furthermore, we partition each $\mathrm{CUT}_{t_i,k}$ into two subsets $\mathrm{CUT}^{\mathrm{out}}_{t_i,k}$ and $\mathrm{CUT}^{\mathrm{in}}_{t_i,k}$, where

$$\mathrm{CUT}^{\mathrm{out}}_{t_i,k} = \{e : \ e \in \mathrm{CUT}_{t_i,k} \setminus \mathrm{In}(i_k)\}, \quad \text{and}$$
$$\mathrm{CUT}^{\mathrm{in}}_{t_i,k} = \{e : \ e \in \mathrm{CUT}_{t_i,k} \cap \mathrm{In}(i_k)\}.$$

For each i_k, $k = 0, 1, \cdots, R$, we respectively define two sets of sink nodes M_k and N_k as follows:

- M_k contains the sink nodes t_i such that at least one path in $\mathcal{P}^{(t_i)}$ passes through i_k, i.e.,
$$M_k = \{t_i : \ \mathrm{CUT}_{t_i,k} \neq \mathrm{CUT}_{t_i,k+1}\}; \tag{6}$$
- N_k contains the sink nodes t_i such that not only at least one path in $\mathcal{P}^{(t_i)}$ passes through i_k but also i_k is the last non-sink node (may be the source node i_0) on the ω paths in $\mathcal{P}^{(t_i)}$, i.e.,
$$N_k = \big\{t_i : \ \mathrm{CUT}_{t_i,k} \neq \mathrm{CUT}_{t_i,k+1} \text{ and } \mathrm{CUT}_{t_i,k+1} = \mathrm{CUT}_{t_i,R+1}\big\}. \tag{7}$$

Let $|M_k| = m_k$ and $|N_k| = n_k$. Then, we claim that

$$\sum_{k=0}^{R} n_k = \ell, \tag{8}$$
$$\sum_{k=0}^{R} m_k = (r_1 + 1) + (r_2 + 1) + \cdots + (r_\ell + 1) = \sum_{i=1}^{\ell} r_i + \ell. \tag{9}$$

We explain (8) and (9) as follows. The equality (8) holds since each sink node t_i has exactly one non-sink node to be the last non-sink node passed through by its ω paths in $\mathcal{P}^{(t_i)}$. The equality (9) follows from the assumption that for each sink node t_i, the number of the non-sink nodes passed through by the paths in $\mathcal{P}^{(t_i)}$ is $r_i + 1$ (r_i intermediate nodes and 1 source node). Consequently, we have

$$\sum_{k=0}^{R} (m_k - n_k) = \sum_{i=1}^{\ell} r_i. \tag{10}$$

In the following example, we use the butterfly network to illustrate the definitions introduced above.

Example 1: In the butterfly network G_1 depicted in Fig. 1, let $\omega = 2$ and $\mathrm{In}(s) = \{d_1, d_2\}$. For two sink nodes $t_1, t_2 \in T$, consider $\mathcal{P}^{(t_1)} = \{P^{(t_1)}_1, P^{(t_1)}_2\}$ and $\mathcal{P}^{(t_2)} = \{P^{(t_2)}_1, P^{(t_2)}_2\}$ with $P^{(t_1)}_1 = (e_1, e_3)$, $P^{(t_1)}_2 = (e_2, e_5, e_7, e_8)$, and $P^{(t_2)}_1 = (e_1, e_4, e_7, e_9)$, $P^{(t_2)}_2 = (e_2, e_6)$. By the above discussion,

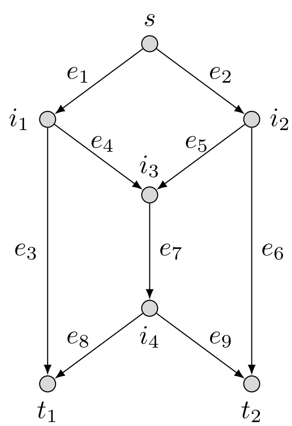


Fig. 1. The butterfly network G_1.

we have

$$\begin{aligned}
\mathrm{CUT}_{t_1,0} &= \{d_1, d_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,0} = \emptyset;\\
\mathrm{CUT}_{t_1,1} &= \{e_1, e_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,1} = \{e_2\};\\
\mathrm{CUT}_{t_1,2} &= \{e_3, e_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,2} = \{e_3\};\\
\mathrm{CUT}_{t_1,3} &= \{e_3, e_5\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,3} = \{e_3\};\\
\mathrm{CUT}_{t_1,4} &= \{e_3, e_7\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,4} = \{e_3\};\\
\mathrm{CUT}_{t_1,5} &= \{e_3, e_8\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_1,5} = \emptyset;
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{CUT}_{t_2,0} &= \{d_1, d_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,0} = \emptyset;\\
\mathrm{CUT}_{t_2,1} &= \{e_1, e_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,1} = \{e_2\};\\
\mathrm{CUT}_{t_2,2} &= \{e_4, e_2\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,2} = \{e_4\};\\
\mathrm{CUT}_{t_2,3} &= \{e_4, e_6\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,3} = \{e_6\};\\
\mathrm{CUT}_{t_2,4} &= \{e_7, e_6\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,4} = \{e_6\};\\
\mathrm{CUT}_{t_2,5} &= \{e_9, e_6\}, \quad \mathrm{CUT}^{\mathrm{out}}_{t_2,5} = \emptyset.
\end{aligned}$$

Consequently, we have

$$M_0 = M_1 = M_2 = M_3 = M_4 = \{t_1, t_2\} \quad \textit{and}$$
$$N_0 = N_1 = N_2 = N_3 = \emptyset, \quad N_4 = \{t_1, t_2\}.$$

The following theorem gives an upper bound on $\mathbf{P}(G)$.

Theorem 1: Consider RLNC scheme on the network G. Then,

$$\mathbf{P}(G) \le 1 - (1-a)^{\ell} \prod_{k=0}^{R-1} [1 - (m_k - n_k)a],$$

with $a = 1 - \prod_{h=1}^{\omega}(1 - 1/|\mathcal{F}|^h)$ and $|\mathcal{F}| > |T|$.[2]

Proof: We consider RLNC over a finite field $\mathcal{F}$ on the network G. For each sink node t_i, we define a series of $\omega \times \omega$ matrices as follows according to the corresponding cuts (as defined before)

$$F^{(k)}_{t_i} = [f_e : e \in \mathrm{CUT}_{t_i,k}], \quad k = 0, 1, \cdots, R+1.$$

[2]To simplify the notation, we always use a to stand for $1 - \prod_{h=1}^{\omega}(1 - 1/|\mathcal{F}|^h)$ in the rest of the paper. The constraint $|\mathcal{F}| > |T|$ is to guarantee the existence of an $\mathcal{F}$-valued linear network code on G [2], [4].

Denote the event "$\mathrm{Rank}(F^{(k)}_{t_i}) = \omega$" by $\Gamma_{t_i,k}$. In particular, $F^{(R+1)}_{t_i}$ is a submatrix of the decoding matrix $F_{t_i} = [f_e : e \in \mathrm{In}(t_i)]$, since $\mathrm{CUT}_{t_i,R+1} \subseteq \mathrm{In}(t_i)$. Thus, the event "$\exists\ t_i \in T$ s.t. $\mathrm{Rank}(F_{t_i}) < \omega$" implies the event "$\exists\ t_i \in T$ s.t. $\mathrm{Rank}(F^{(R+1)}_{t_i}) < \omega$". Together with

$$\begin{aligned}
\mathbf{P}(G) &= \mathbf{Pr}(\exists\ t \in T \text{ s.t. } \mathrm{Rank}(F_t) < \omega)\\
&= \mathbf{Pr}(\cup_{i=1}^{\ell} \mathrm{Rank}(F_{t_i}) < \omega),
\end{aligned}$$

we have

$$\begin{aligned}
\mathbf{P}(G) &\le \mathbf{Pr}(\cup_{i=1}^{\ell} \mathrm{Rank}(F^{(R+1)}_{t_i}) < \omega)\\
&= 1 - \mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,R+1}).
\end{aligned}$$

Note that

$$\begin{aligned}
&\mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,R+1})\\
&\ge \mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,R+1}, \cap_{i=1}^{\ell} \Gamma_{t_i,R}, \cdots, \cap_{i=1}^{\ell} \Gamma_{t_i,1}, \cap_{i=1}^{\ell} \Gamma_{t_i,0})\\
&= \mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,0}) \cdot \prod_{k=0}^{R} \mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \qquad (11)\\
&= \prod_{k=0}^{R} \mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}), \qquad (12)
\end{aligned}$$

where (11) holds since under the condition that no failure occurs up to the last node, the random encoding at the present node is independent of what happened before the last node, and (12) follows from

$$\mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,0}) = \mathbf{Pr}(\mathrm{Rank}([f_e : e \in \mathrm{In}(s)]) = \omega) = 1.$$

Furthermore, we have

$$\begin{aligned}
&\mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k})\\
&= \mathbf{Pr}(\cap_{t_i \in M_k} \Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \qquad (13)\\
&= \mathbf{Pr}\Big(\big(\cap_{t_i \in M_k \setminus N_k} \Gamma_{i,k+1}\big) \cap \big(\cap_{t_i \in N_k} \Gamma_{i,k+1}\big) \Big| \cap_{i=1}^{\ell} \Gamma_{t_i,k}\Big) \qquad (14)\\
&= \mathbf{Pr}\Big(\cap_{t_i \in M_k \setminus N_k} \Gamma_{i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}\Big)\\
&\quad \cdot \mathbf{Pr}\Big(\cap_{t_i \in N_k} \Gamma_{i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}\Big) \qquad (15)\\
&= \mathbf{Pr}\Big(\cap_{t_j \in M_k \setminus N_k} \Gamma_{j,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}\Big)\\
&\quad \cdot \prod_{t_i \in N_k} \mathbf{Pr}(\Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \qquad (16)\\
&= \mathbf{Pr}\Big(\cap_{t_j \in M_k \setminus N_k} \Gamma_{j,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}\Big)\\
&\quad \cdot \prod_{t_i \in N_k} \mathbf{Pr}(\Gamma_{t_i,k+1} | \Gamma_{t_i,k}). \qquad (17)
\end{aligned}$$

The equalities from (13) to (17) are explained below. (13) follows from $\mathrm{CUT}_{t_i,k+1} = \mathrm{CUT}_{t_i,k}$ for $t_i \notin M_k$. For a sink node $t_i \in N_k$, i.e., i_k is the last non-sink node passed through by the paths in $\mathcal{P}^{(t_i)}$, all "downstream" edges are only on the paths in $\mathcal{P}^{(t_i)}$. Hence, the event $\Gamma_{t_i,k+1}$ is conditionally independent to the event $\cap_{t_j \in M_k \setminus \{t_i\}} \Gamma_{t_j,k+1}$ under the condition $\cap_{i=1}^{\ell} \Gamma_{t_i,k}$ (implying (15) and (16)) and further conditionally independent under the condition $\Gamma_{t_i,k}$ (implying (17)). In addition, we reasonably set $\prod_{t_i \in N_k} \mathbf{Pr}(\Gamma_{t_i,k+1} | \Gamma_{t_i,k}) = 1$ when

$N_k = \emptyset$, and $\mathbf{Pr}(\cap_{t_j \in M_k \backslash N_k} \Gamma_{t_j,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) = 1$ when $M_k \backslash N_k = \emptyset$.

For the first term of (17), we obtain

$$\begin{aligned}
&\mathbf{Pr}(\cap_{t_j \in M_k \backslash N_k} \Gamma_{t_j,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \\
&= 1 - \mathbf{Pr}(\cup_{t_j \in M_k \backslash N_k} \Gamma_{j,k+1}^c | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \\
&\geq 1 - \sum_{t_j \in M_k \backslash N_k} \mathbf{Pr}(\Gamma_{t_j,k+1}^c | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \\
&= 1 - \sum_{t_j \in M_k \backslash N_k} \left[1 - \mathbf{Pr}(\Gamma_{t_j,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k})\right] \\
&= 1 - \sum_{t_j \in M_k \backslash N_k} \left[1 - \mathbf{Pr}(\Gamma_{t_j,k+1} | \Gamma_{t_j,k})\right] \\
&= 1 - \sum_{t_j \in M_k \backslash N_k} \left[1 - \prod_{h=1}^{\omega - |\mathrm{CUT}_{t_j,k}^{\mathrm{out}}|} \left(1 - \frac{1}{|\mathcal{F}|^h}\right)\right] \\
&\geq 1 - (m_k - n_k)\left[1 - \prod_{h=1}^{\omega} \left(1 - \frac{1}{|\mathcal{F}|^h}\right)\right] \\
&= 1 - (m_k - n_k)a. \qquad (18)
\end{aligned}$$

For the second term of (17), we obtain

$$\begin{aligned}
\prod_{t_i \in N_k} \mathbf{Pr}(\Gamma_{t_i,k+1} | \Gamma_{t_i,k}) &= \prod_{t_i \in N_k} \prod_{h=1}^{\omega - |\mathrm{CUT}_{t_i,k}^{\mathrm{out}}|} \left(1 - \frac{1}{|\mathcal{F}|^h}\right) \\
&\geq \prod_{t_i \in N_k} \prod_{h=1}^{\omega} \left(1 - \frac{1}{|\mathcal{F}|^h}\right) = (1-a)^{n_k}. \qquad (19)
\end{aligned}$$

It follows from the inequalities (13) to (19) that

$$\mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,k+1} | \cap_{i=1}^{\ell} \Gamma_{t_i,k}) \geq (1-a)^{n_k}[1 - (m_k - n_k)a]. \qquad (20)$$

Combining (11), (12), and (20), we obtain that

$$\begin{aligned}
\mathbf{Pr}(\cap_{i=1}^{\ell} \Gamma_{t_i,R+1}) &\geq \prod_{k=0}^{R} (1-a)^{n_k}[1 - (m_k - n_k)a] \qquad (21) \\
&= (1-a)^{\sum_{k=0}^{R} n_k} \prod_{k=0}^{R} [1 - (m_k - n_k)a] \qquad (22) \\
&= (1-a)^{\ell} \prod_{k=0}^{R} [1 - (m_k - n_k)a] \qquad (23) \\
&= (1-a)^{\ell} \prod_{k=0}^{R-1} [1 - (m_k - n_k)a], \qquad (24)
\end{aligned}$$

where (23) and (24) follow from $\sum_{k=0}^{R} n_k = \ell$ (see (8)) and $m_R = n_R$, respectively. The theorem is proved. ■

To compute the upper bound in Theorem 1, the quantities such as m_k, n_k, and R are necessary, which may be unavailable for some applications. Hence, we develop some other upper bounds so that fewer network topology information is required. It is conceivable that such new bounds may become looser.

Theorem 2: In the network G, let r_i be the number of intermediate nodes each of which is passed through by a path in a given set of ω channel-disjoint paths from s to t_i for $1 \leq i \leq \ell$. If $\sum_{i=1}^{\ell} r_i$ is upper bounded by a nonnegative integer n, then

$$\mathbf{P}(G) \leq 1 - (1-a)^{\ell}(1 - \ell a)^q (1 - ra),$$

where q and r are two nonnegative integers such that $n = \ell q + r$ and $0 \leq r \leq \ell - 1$.

In order to prove Theorem 2, we need the following lemma of which the proof is deferred to Appendix V.

Lemma 1: Let $h_0, h_1, \cdots, h_{\ell-1}$ be ℓ nonnegative integers such that

$$\ell h_0 + (\ell - 1)h_1 + (\ell - 2)h_2 + \cdots + 2h_{\ell-2} + h_{\ell-1} = \ell q + r, \qquad (25)$$

where q and r are two nonnegative integers with $0 \leq r \leq \ell - 1$. Then for any real number a with $0 \leq a < \frac{1}{\ell}$, the following inequality holds:

$$(1 - \ell a)^{h_0}(1 - (\ell - 1)a)^{h_1} \cdots (1-a)^{h_{\ell-1}} \geq (1 - \ell a)^q (1 - ra). \qquad (26)$$

Proof of Theorem 2: By Theorem 1, we can write

$$1 - \mathbf{P}(G) \geq (1-a)^{\ell} \prod_{k=0}^{R-1} [1 - (m_k - n_k)a].$$

Recall (10), i.e.,

$$\sum_{k=0}^{R-1} (m_k - n_k) = \sum_{k=0}^{R} (m_k - n_k) = \sum_{i=1}^{\ell} r_i,$$

and let $\sum_{i=1}^{\ell} r_i = \ell q' + r'$ with q' and r' being two nonnegative integers such that $0 \leq r' \leq \ell - 1$. Note that $0 \leq m_k - n_k \leq \ell$ for all $k = 0, 1, 2, \cdots, R-1$. Then we claim that

$$\begin{aligned}
&\big(1 - (m_{R-1} - n_{R-1})a\big)\big(1 - (m_{R-2} - n_{R-2})a\big) \\
&\quad \cdots \big(1 - (m_0 - n_0)a\big) \qquad (27) \\
&= (1 - \ell a)^{h_0}(1 - (\ell - 1)a)^{h_1} \cdots (1-a)^{h_{\ell-1}} \qquad (28) \\
&\geq (1 - \ell a)^{q'}(1 - r'a). \qquad (29)
\end{aligned}$$

The equality (28) and the inequality (29) are explained below. For (28), we use h_i, $\forall\ i = 0, 1, \cdots, \ell - 1$, to denote the number of such k's, $0 \leq k \leq R-1$, that $m_k - n_k = \ell - i$ ($h_i \geq 0$, $\forall\ i$). Hence, (27) can be written as (28). (29) follows from Lemma 1 by applying the fact that

$$1 - a = \prod_{h=1}^{\omega}(1 - \frac{1}{|\mathcal{F}|^h}) > 1 - \sum_{h=1}^{\infty} \frac{1}{|\mathcal{F}|^h} = 1 - \frac{1}{|\mathcal{F}| - 1},$$

which, together with $|\mathcal{F}| > |T|$, implies that

$$a < \frac{1}{|\mathcal{F}| - 1} \leq \frac{1}{|T|} = \frac{1}{\ell}.$$

Furthermore, note that

$$\ell q + r = n \geq \sum_{i=1}^{\ell} r_i = \ell q' + r'$$

and then we prove the following inequality by considering two cases below:

$$(1-\ell a)^q(1-ra) \leq (1-\ell a)^{q'}(1-r'a). \tag{30}$$

i) $q=q'$ and $r \geq r'$. It suffices to prove $1-ra \leq 1-r'a$, which is trivial since $r \geq r'$.
ii) $q > q'$. Eq. (30) follows from

$$(1-\ell a)^{q-q'}(1-ra) \leq 1-\ell a \leq 1-r'a.$$

Combining the above, the theorem is proved. ■

Remark 1: The above proof also implies that the smaller upper bound can be obtained if the smaller upper bound n on $\sum_{i=1}^{\ell} r_i$ can be used. In particular, when we can obtain the exact value of $\sum_{i=1}^{\ell} r_i$, regarded as its the smallest upper bound, we have the following corollary.

Corollary 3: In the network G, let r_i be the number of intermediate nodes each of which is passed through by a path in a given set of ω channel-disjoint paths from s to t_i for $1 \leq i \leq \ell$. Then,

$$\mathbf{P}(G) \leq 1-(1-a)^{\ell}(1-\ell a)^q(1-ra),$$

where $\sum_{i=1}^{\ell} r_i = \ell q + r$ with q and r being two nonnegative integers satisfying $0 \leq r \leq \ell-1$.

Furthermore, if we can choose a set $\mathcal{P}^{(t_i)}$ of ω channel-disjoint paths from s to t_i for each sink node t_i such that the sum of R_i, the number of intermediate nodes on $\mathcal{P}^{(t_i)}$, achieves the minimum, then

$$\mathbf{P}(G) \leq 1-(1-a)^{\ell}(1-\ell a)^Q(1-Ra),$$

where $\sum_{i=1}^{\ell} R_i = \ell Q + R$ with Q and R being two nonnegative integers such that $0 \leq R \leq \ell-1$.

Suppose fewer network topology information is available. To be specific, only the number of sink nodes $|T|$ and the number of intermediate nodes $|J|$ are available (the same assumption used in [11], [13]). Clearly, we have $r_i \leq |J|$, $\forall\ t_i \in T$ so that $\sum_{i=1}^{\ell} r_i \leq |T| \cdot |J|$. Thus, by replacing n in Theorem 2 by $|T| \cdot |J|$, we immediately obtain the following corollary which improves the corresponding upper bound in [11] ($\mathbf{P}_{av}(G)$ was considered in [13]).

Corollary 4: In the network G, T and J are the set of the sink nodes and the set of the intermediate nodes, respectively. Then,

$$\mathbf{P}(G) \leq 1-(1-a)^{|T|}(1-|T|a)^{|J|}.$$

Until now, we have obtained several upper bounds on $\mathbf{P}(G)$ when different partial network topology information are available, as shown in Theorems 1 and 2 and Corollaries 3 and 4. By comparing with them, it is observed that if more network topology information can be utilized, tighter upper bounds can be obtained, as stated in the following result.

Theorem 5: Let $G=(V,E)$ be a finite directed acyclic network, where s is the single source node, $T=\{t_1,t_2,\cdots,t_\ell\}$ is the set of the ℓ sink nodes, and J is the set of the intermediate nodes. For each t_i, let $\mathcal{P}^{(t_i)}$ be a set of ω ($\leq C_{t_i}$) channel-disjoint paths from s to t_i, r_i be the number of intermediate nodes each of which is passed through by at least one path in $\mathcal{P}^{(t_i)}$, R be the number of intermediate nodes passed through by at least one path in $\cup_{i=1}^{\ell}\mathcal{P}^{(t_i)}$, and n be a nonnegative integer such that $\sum_{i=1}^{\ell} r_i \leq n \leq \ell|J|$. Then,

$$\begin{aligned}\mathbf{P}(G) &\leq 1-(1-a)^{\ell}\prod_{k=0}^{R-1}[1-(m_k-n_k)a]\\ &\leq 1-(1-a)^{\ell}(1-\ell a)^q(1-ra)\\ &\leq 1-(1-a)^{\ell}(1-\ell a)^{q'}(1-r'a)\\ &\leq 1-(1-a)^{\ell}(1-\ell a)^{|J|},\end{aligned}$$

where q and r are two nonnegative integers such that $\sum_{i=1}^{\ell} r_i = \ell q + r$ and $0 \leq r \leq \ell-1$, and q' and r' are two nonnegative integers such that $n = \ell q' + r'$ and $0 \leq r' \leq \ell-1$.

Proof: The first inequality follows from Theorem 1, the second one follows from (27)-(29) in the proof of Theorem 2, the third one follows from Remark 1, and the last one follows from the argument before Corollary 4 and $\ell q' + r' = n \leq \ell|J|$. So the theorem is proved. ■

B. Numerical Results

In this subsection, we compare the different upper bounds on $\mathbf{P}(G)$, including the foregoing ones we obtained and the existing ones in [5] and [11], on two well-known networks, namely the butterfly network and the combination network.

We first reproduce the two existing upper bounds on $\mathbf{P}(G)$ in [5] and [11], respectively, as follows:

$$\mathbf{P}(G) \leq 1-\left(1-\frac{|T|}{|\mathcal{F}|}\right)^{|E|}, \tag{31}$$

$$\mathbf{P}(G) \leq 1-\left(1-\frac{|T|}{|\mathcal{F}|-1}\right)^{|J|+1}. \tag{32}$$

Clearly, (32) is an improvement over (31) because in most networks $|E| \gg |J|+1$ (see [11]). By a simple computation, our upper bound in Corollary 4 enhances (32). Together with Theorem 5, our upper bounds depending on partial network topological information improve the previous ones in [5] and [11].

1) Butterfly Network: The butterfly network G_1 is depicted in Fig. 1. Clearly, we see that $\ell=|T|=2$, $|J|=4$, and $|E|=9$. Let $\omega=2$. In G_1, the set $\mathcal{P}^{(t_i)}$ of 2 channel-disjoint paths of each sink node t_i, $i=1,2$, is unique (cf. Example 1), and so R, the number of the intermediate nodes passed through by at least one path in $\mathcal{P}^{(t_1)} \bigcup \mathcal{P}^{(t_2)}$, is 4. Furthermore, by Example 1, we see that

$$\begin{aligned} m_k &= |M_k| = 2, \quad 0 \leq k \leq 4, \quad \text{and}\\ n_k &= |N_k| = 0, \quad 0 \leq k \leq 3, \quad n_4 = |N_4| = 2.\end{aligned}$$

Based on the above discussions, the upper bounds in Theorem 1, and Corollaries 3 and 4 for G_1 are the same, i.e.,

$$\mathbf{P}(G_1) \leq 1-(1-a)^2(1-2a)^4. \tag{33}$$

In particular, for Corollary 3, we have $n = r_1 + r_2 = 8$ and so $q=4$ and $r=0$ from $n=\ell q + r$.

Furthermore, the exact value of $\mathbf{P}(G_1)$ has been given in [24] as

$$\mathbf{P}(G_1) = 1-\frac{(|\mathcal{F}|+1)(|\mathcal{F}|-1)^{10}}{|\mathcal{F}|^{11}}. \tag{34}$$

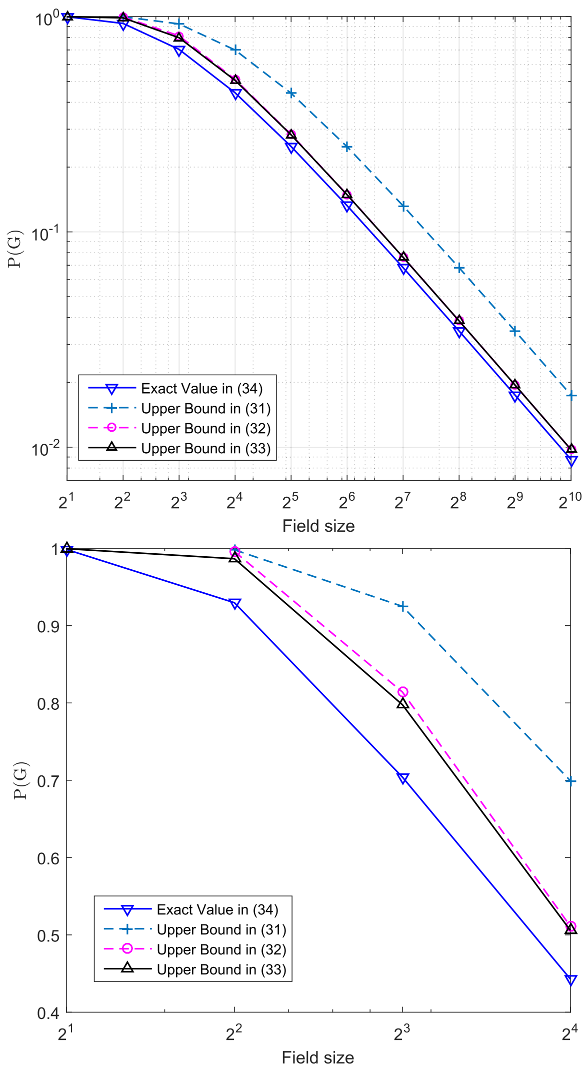


Fig. 2. The comparison for the butterfly network G_1.

In Fig. 2, we compare the 3 upper bounds in (31), (32), (33) and the exact value in (34). This comparison shows that our upper bounds are tighter than the existing ones, though the difference is not huge because the butterfly network is relatively small as it consists 7 nodes and 9 edges. For larger networks in scale, the advantage will be significant.

2) Combination Network: Combination network is a well-known type of network demonstrating that the gain of network coding over pure network routing can be unbounded. Here, we consider a $(2, 4)$ combination network $G_{(2,4)}$, i.e., $k = 2$ and $n = 4$, which is depicted in Fig. 3.

In $G_{(2,4)}$, we see that $|T| = 6$, $|J| = 4$, and $|E| = 16$. Let $\omega = 2$ and $\mathrm{In}(s) = \{d_1, d_2\}$. For the 6 sink nodes t_i, $1 \le i \le 6$, their sets $\mathcal{P}^{(t_i)}$ of 2 channel-disjoint paths are unique and given as follows:

$$\mathcal{P}^{(t_1)} = \{P_1^{(t_1)} = (e_1, e_5), \quad P_2^{(t_1)} = (e_2, e_8)\};$$
$$\mathcal{P}^{(t_2)} = \{P_1^{(t_2)} = (e_1, e_6), \quad P_2^{(t_2)} = (e_3, e_{11})\};$$
$$\mathcal{P}^{(t_3)} = \{P_1^{(t_3)} = (e_1, e_7), \quad P_2^{(t_3)} = (e_4, e_{14})\};$$
$$\mathcal{P}^{(t_4)} = \{P_1^{(t_4)} = (e_2, e_9), \quad P_2^{(t_4)} = (e_3, e_{12})\};$$

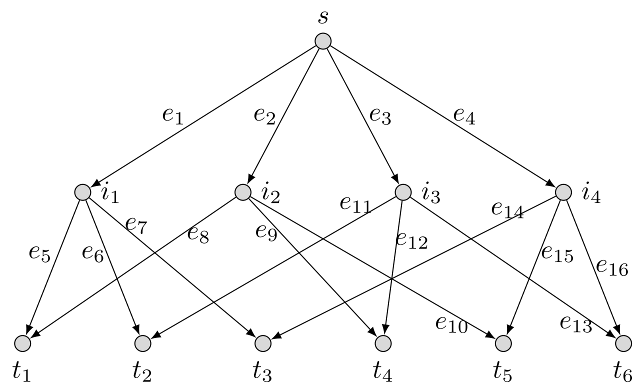


Fig. 3. The combination network $G_{(2,4)}$.

$$\mathcal{P}^{(t_5)} = \{P_1^{(t_5)} = (e_2, e_{10}), \quad P_2^{(t_5)} = (e_4, e_{15})\};$$
$$\mathcal{P}^{(t_6)} = \{P_1^{(t_6)} = (e_3, e_{13}), \quad P_2^{(t_6)} = (e_4, e_{16})\};$$

and note that the number r_i of the intermediate nodes each of which is passed through by at least one path in $\mathcal{P}^{(t_i)}$ is 2. We assume that the ancestrally topological order of the nodes is

$$s \triangleq i_0 \prec i_1 \prec i_2 \prec i_3 \prec i_4 \prec \{t_i : 1 \le i \le 6\}.$$

By (6), M_k is the set of the sink nodes t_i such that at least one path in $\mathcal{P}^{(t_i)}$ passes through i_k for $0 \le k \le 4$. More specifically, we have

$$M_0 = T, \quad M_1 = \{t_1, t_2, t_3\}, \quad M_2 = \{t_1, t_4, t_5\},$$
$$M_3 = \{t_2, t_4, t_6\}, \quad M_4 = \{t_3, t_5, t_6\}.$$

Let $|M_k| = m_k$ for $0 \le k \le 4$ and then $m_0 = 6$ and $m_1 = m_2 = m_3 = m_4 = 3$.

By (7), N_k is the set of the sink nodes t_i such that at least one path in $\mathcal{P}^{(t_i)}$ passes through i_k and i_k is the last non-sink node on the 2 paths in $\mathcal{P}^{(t_i)}$, $0 \le k \le 4$. More specifically, we have

$$N_0 = N_1 = \emptyset, \quad N_2 = \{t_1\}, \quad N_3 = \{t_2, t_4\}, \quad N_4 = \{t_3, t_5, t_6\}.$$

Let $|N_k| = n_k$ for $0 \le k \le 4$ and then $n_0 = n_1 = 0$, $n_2 = 1$, $n_3 = 2$, and $n_4 = 3$.

Therefore, the upper bounds in Theorem 1 and Corollaries 3 and 4 are given sequentially as follows:

$$\begin{aligned}\mathbf{P}(G_{(2,4)}) &\le 1 - (1-a)^6 \prod_{k=0}^{3} [1 - (m_k - n_k)a] \\ &= 1 - (1-a)^7(1-6a)(1-3a)(1-2a), \quad (35)\\ \mathbf{P}(G_{(2,4)}) &\le 1 - (1-a)^6(1-6a)^q(1-ra) \\ &= 1 - (1-a)^6(1-6a)^2, \quad (36)\\ \mathbf{P}(G_{(2,4)}) &\le 1 - (1-a)^{|T|}(1-|T|a)^{|J|} \\ &= 1 - (1-a)^6(1-6a)^4, \quad (37)\end{aligned}$$

where in (36), $q = 2$ and $r = 0$ follows from $\sum_{i=1}^{6} r_i = 6q + r$ and $r_i = 2$ for all $1 \le i \le 6$.

The numerical result for the comparison of the upper bounds on $\mathbf{P}(G_{(2,4)})$ is given in Fig. 4. This characterizes not only the usefulness of our upper bounds over the existing ones, but

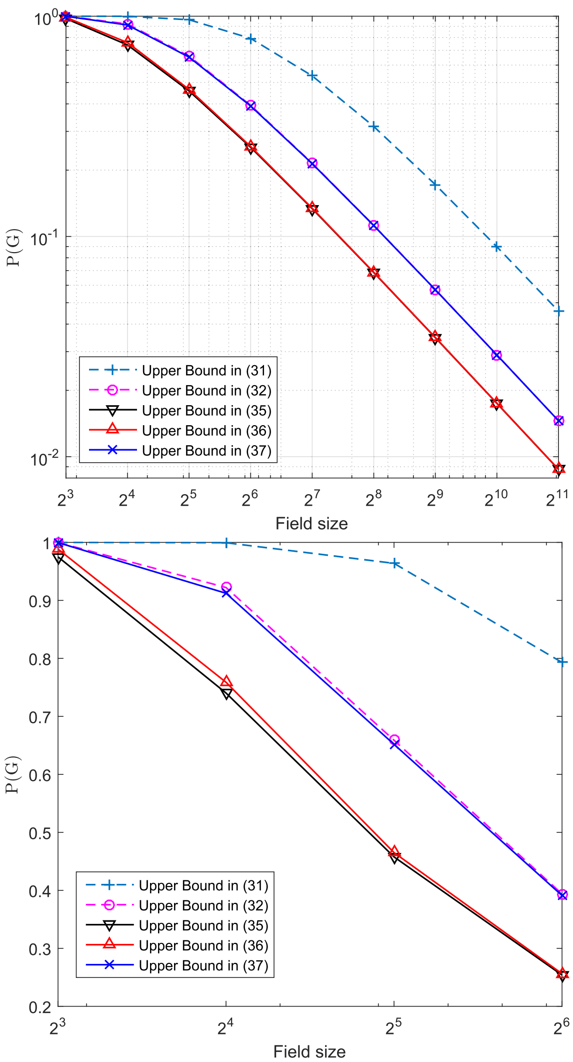


Fig. 4. The comparison for the combination network $G_{(2,4)}$.

also the usefulness of our upper bounds when more network topology information is available.

IV. Asymptotic Behavior of RLNC

In this section, we will investigate the asymptotic behavior of the failure probability $\mathbf{P}(G)$ as the field size goes to infinity. This makes it possible for us to ignore some complicated minor terms during the derivation so that we can obtain a deeper understanding about the behavior of failure probability $\mathbf{P}(G)$, for instance, the major term affecting the value of $\mathbf{P}(G)$. In addition, the main result in this section (Theorem 6) also indicates the asymptotic tightness of the upper bounds obtained in the last section.

We use $\mathbf{N}^*$ to denote a RLNC problem, which is fully characterized by the network G, the source node s, the set of the sink nodes T, and the information rate ω. Hence, we write $\mathbf{N}^*$ as a quadruple $\{G, s, T, \omega\}$. For such a RLNC problem $\mathbf{N}^*$, we define the *limiting behavior* of $\mathbf{P}(G)$ as the field size goes to infinity as follows:

$$\Omega(\mathbf{N}^*) = \limsup_{|\mathcal{F}|\to\infty} |\mathcal{F}| \cdot \mathbf{P}(G). \tag{38}$$

Next, we consider all the RLNC problems $\mathbf{N}^*$ with the following constraints:

1) $|T| = \ell$, i.e., the number of the sink nodes is ℓ;
2) $\sum_{i=1}^{\ell} r_i \leq n$, i.e., for each sink node t_i, $1 \leq i \leq \ell$, there exists a set $\mathcal{P}^{(t_i)}$ of ω channel-disjoint paths from s to t_i in which r_i is the number of the intermediate nodes passed through by at least one path in $\mathcal{P}^{(t_i)}$ such that the sum of r_i, $1 \leq i \leq \ell$, does not exceed a fixed nonnegative integer n.

The set of all such RLNC problems is denoted by $\mathcal{M}^*_{n,\ell}$, and we further define

$${\Lambda^+_{(n,\ell)}}^* = \max_{\mathbf{N}^* \in \mathcal{M}^*_{n,\ell}} \Omega(\mathbf{N}^*), \tag{39}$$

which is the worst case limiting behavior of $\Omega(\mathbf{N}^*)$ over all $\mathbf{N}^*$ in $\mathcal{M}^*_{n,\ell}$. The following theorem gives the worst case limiting behavior ${\Lambda^+_{(n,\ell)}}^*$.

Theorem 6: For the worst case limiting behavior ${\Lambda^+_{(n,\ell)}}^*$,

$${\Lambda^+_{(n,\ell)}}^* = \ell + n.$$

Proof: We first prove that $\Omega(\mathbf{N}^*) \leq \ell + n$ for any $\mathbf{N}^* \in \mathcal{M}^*_{n,\ell}$, which immediately implies that ${\Lambda^+_{(n,\ell)}}^* \leq \ell + n$ by (39). Let $\mathbf{N}^* = \{G, s, T, \omega\}$ be a RLNC problem in $\mathcal{M}^*_{n,\ell}$. With the constraints 1) and 2), we obtain by Theorem 2 that

$$\mathbf{P}(G) \leq 1 - (1-a)^{\ell}(1-\ell a)^{q}(1-ra), \tag{40}$$

where q and r are two nonnegative integers such that $n = \ell q + r$ and $0 \leq r \leq \ell - 1$.

Recall that $a = 1 - \prod_{h=1}^{\omega}(1 - 1/|\mathcal{F}|^h)$, and then we have

$$a = 1 - \prod_{h=1}^{\omega}\left(1 - \frac{1}{|\mathcal{F}|^h}\right) \geq 1 - \left(1 - \frac{1}{|\mathcal{F}|}\right) = \frac{1}{|\mathcal{F}|}$$

and

$$a = 1 - \prod_{h=1}^{\omega}\left(1 - \frac{1}{|\mathcal{F}|^h}\right) \leq \sum_{h=1}^{\omega} \frac{1}{|\mathcal{F}|^h} < \sum_{h=1}^{\infty} \frac{1}{|\mathcal{F}|^h} = \frac{1}{|\mathcal{F}|-1}.$$

For $\mathcal{F}$ sufficiently large, we further have

$$0 < \frac{1}{|\mathcal{F}|} \leq a \leq ra \leq \ell a < \frac{\ell}{|\mathcal{F}|-1} \leq 1.$$

This implies that

$$\begin{aligned}(1-a)^{\ell}(1-\ell a)^{q}(1-ra) &\geq (1-\ell a)(1-\ell q a)(1-ra)\\ &\geq 1-(\ell+\ell q+r)a = 1-(\ell+n)a.\end{aligned}$$

By (40), we obtain

$$\mathbf{P}(G) \leq (\ell+n)a \leq \frac{\ell+n}{|\mathcal{F}|-1}.$$

Immediately, we obtain

$$\limsup_{|\mathcal{F}|\to\infty} |\mathcal{F}|\mathbf{P}(G) \leq (\ell+n) \lim_{|\mathcal{F}|\to\infty} \frac{|\mathcal{F}|}{|\mathcal{F}|-1} = \ell + n.$$

Thus, ${\Lambda^+_{(n,\ell)}}^* \leq \ell + n$.

It remains to prove that there exists a RLNC problem $\mathbf{N}^* \in \mathcal{M}^*_{n,\ell}$ such that

$$\limsup_{|\mathcal{F}|\to\infty} |\mathcal{F}|\mathbf{P}(G) = \ell + n.$$

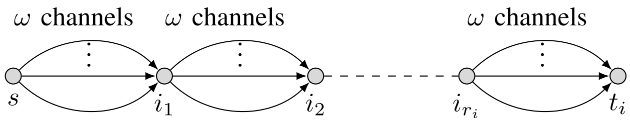


Fig. 5. The plait network G_i'.

Now, consider the network G' with a single source node s and ℓ sink nodes $t_1, t_2, \cdots, t_\ell$ as follows: for each sink node t_i, let G_i' be a plait network[3] with r_i intermediate nodes, depicted in Fig. 5, such that $\sum_{i=1}^{\ell} r_i = n$; and all the networks G_i', $1 \leq i \leq \ell$, share the common source node s.

We first calculate $\mathbf{P}(G_i')$ for the plaint network G_i', $1 \leq i \leq \ell$, as follows:

$$\begin{aligned}
1 - \mathbf{P}(G_i') &= 1 - \mathbf{Pr}(\mathrm{Rank}(F_{t_i}) = \omega) \\
&= 1 - \mathbf{Pr}\big(\cap_{k=0}^{r_i} \mathrm{Rank}(\big[f_e : e \in \mathrm{Out}(i_k)\big]) = \omega \big) \\
&= 1 - \mathbf{Pr}\big(\mathrm{Rank}(\big[f_e : e \in \mathrm{Out}(s)\big]) = \omega \big| \\
&\qquad\qquad \mathrm{Rank}(\big[f_e : e \in \mathrm{In}(s)\big]) = \omega\big) \\
&\quad \cdot \prod_{k=1}^{r_i} \mathbf{Pr}\Big(\mathrm{Rank}(\big[f_e : e \in \mathrm{Out}(i_k)\big]) = \omega \big| \\
&\qquad\qquad \mathrm{Rank}(\big[f_e : e \in \mathrm{Out}(i_{k-1})\big]) = \omega\Big) \\
&= 1 - (1-a)^{r_i + 1}.
\end{aligned}$$

Thus, for $\mathbf{P}(G')$, we have

$$\begin{aligned}
\mathbf{P}(G') &= 1 - \mathbf{Pr}(\cap_{i=1}^{\ell} \mathrm{Rank}(F_{t_i}) = \omega) \\
&= 1 - \prod_{i=1}^{\ell} \mathbf{Pr}(\mathrm{Rank}(F_{t_i}) = \omega) \\
&= 1 - \prod_{i=1}^{\ell} (1-a)^{r_i + 1} \\
&= 1 - (1-a)^{\sum_{i=1}^{\ell} r_i + \ell} = 1 - (1-a)^{n+\ell} \\
&\geq 1 - \Big(1 - \frac{1}{|\mathcal{F}|}\Big)^{n+\ell} \\
&= \frac{\ell + n}{|\mathcal{F}|} + \mathcal{O}(\frac{1}{|\mathcal{F}|^2}),
\end{aligned}$$

where the inequality follows from $a \geq 1/|\mathcal{F}|$. Therefore, $\lim_{|\mathcal{F}| \to \infty} |\mathcal{F}| \cdot \mathbf{P}(G') \geq \ell + n$, implying ${\Lambda_{(n,\ell)}^{+}}^{*} \geq \ell + n$. Combining the above, the theorem is proved. ■

V. Conclusion

In this paper, depending on the different partial network topology information available, we have given the hierarchical performance analysis of RLNC by upper bounding the failure probability of RLNC for the entire network. Our results show that the partial network topology information can bring benefits to the performance analysis of RLNC and tighter upper bounds can be obtained if more network topology information can be utilized, as expected. To obtain a deeper understanding about the performance of RLNC, we also investigated the asymptotic behavior of RLNC as the field size goes to infinity.

Even through we have proved that the partial network topology information brings benefits to the analysis of the performance of RLNC, we still considered using the (complete) random linear network coding, i.e., all the local encoding coefficients are independently random variables uniformly distributed over some finite field. In other words, we do not apply the available partial network topology information to modify the design of network codes. Thus, an interesting problem for further research is whether the available partial network topology information can lead to better network codes in RLNC, e.g., possibly using *partial* random linear network coding instead of a complete random linear network coding, so as to improve the performance of network coding in essence.

Appendix
Proof of Lemma 1

Before proving Lemma 1, we first give the following two lemmas.

Lemma 2: Let k and ℓ be two arbitrary nonnegative integers with $k \leq \ell$. Then for any real number a with $0 \leq a \leq \frac{1}{\ell}$, the following inequality holds:

$$(1 - ka)^{\ell} \geq (1 - \ell a)^k. \tag{41}$$

Proof: For $a = \frac{1}{\ell}$, the lemma is trivial. For $a \in [0, \frac{1}{\ell})$, note that $1 \geq 1 - ka \geq 1 - \ell a > 0$. We hence rewrite the inequality (41) as

$$\ell \ln^{(1-ka)} \geq k \ln^{(1-\ell a)}. \tag{42}$$

Now, we prove the inequality (42). Let $f(x) = \ell \ln^{(1-kx)} - k \ln^{(1-\ell x)}$, $x \in [0, \frac{1}{\ell})$, be a function. We calculate the differential of $f(x)$ and obtain that

$$f'(x) = k\ell \left(\frac{1}{1 - \ell x} - \frac{1}{1 - kx} \right) \geq 0, \quad \forall x \in [0, \frac{1}{\ell}). \tag{43}$$

This implies that $f(x)$ is a monotone increasing function on the interval $[0, \frac{1}{\ell})$. Together with $f(0) = 0$, we obtain that $f(x) \geq 0$, $\forall x \in [0, \frac{1}{\ell})$. This lemma is proved. ■

Lemma 3[4]: Let $h_0, h_1, \cdots, h_{\ell-1}$ be ℓ nonnegative integers such that

$$\ell h_0 + (\ell - 1)h_1 + (\ell - 2)h_2 + \cdots + 2h_{\ell-2} + h_{\ell-1} = \ell q \tag{44}$$

for a nonnegative integer q. Then for any real number a ($0 \leq a < \frac{1}{\ell}$), the following inequality holds:

$$(1 - \ell a)^{h_0}(1 - (\ell - 1)a)^{h_1} \cdots (1 - a)^{h_{\ell-1}} \geq (1 - \ell a)^q. \tag{45}$$

Proof: The inequality (45) can be proved as follows:

$$\begin{aligned}
&(1 - \ell a)^{h_0}(1 - (\ell - 1)a)^{h_1} \cdots (1 - a)^{h_{\ell-1}} && (46) \\
&= \prod_{i=0}^{\ell-1} \Big[\big(1 - (\ell - i)a\big)^{\ell}\Big]^{\frac{h_i}{\ell}} && (47) \\
&\geq \prod_{i=0}^{\ell-1} \Big[(1 - \ell a)^{\ell - i}\Big]^{\frac{h_i}{\ell}} && (48)
\end{aligned}$$

[3]The plait network was first introduced in [13] to characterize the tightness of their upper bound on the average failure probability of RLNC.

[4]Lemma 3 is a special case of Lemma 1 for $r = 0$.

$$= (1-\ell a)^{\frac{1}{\ell}\sum_{i=0}^{\ell-1}(\ell-i)h_i} \tag{49}$$
$$= (1-\ell a)^{\frac{1}{\ell}\cdot \ell q} \tag{50}$$
$$= (1-\ell a)^{q}, \tag{51}$$

where (48) and (50) follow from Lemma 2 and the equality (44), respectively. ■

We now proceed to prove Lemma 1. We will prove the lemma by induction on ℓ. First, note that the lemma is clear for $\ell = 1$. Next, we assume that the inequality (26) in Lemma 1 holds for all positive integers $1, 2, \cdots, \ell-1$. Now, we consider the positive integer ℓ. If $r = 0$, we complete the proof by Lemma 3. Otherwise, for $r > 0$ we consider the following two cases.

Case 1: $h_{\ell-r} \geq 1$.

In this case, the term $(1-ra)$ appears in both the left and the right hand sides of the inequality (26). Specifically,

$$\begin{aligned}&(1-\ell a)^{h_0}\cdots(1-ra)^{h_{\ell-r}}\cdots(1-a)^{h_{\ell-1}}\\ &\quad= \big[(1-\ell a)^{h_0}\cdots(1-ra)^{h_{\ell-r}-1}\cdots(1-a)^{h_{\ell-1}}\big](1-ra)\\ &\quad\geq (1-\ell a)^q(1-ra),\end{aligned} \tag{52}$$

where the inequality (52) follows from Lemma 3 and the equality

$$\ell h_0 + \cdots + (r+1)h_{\ell-r-1} + r(h_{\ell-r}-1) + (r-1)h_{\ell-r+1}\cdots + h_{\ell-1} = \ell q.$$

Case 2: $h_{\ell-r} = 0$. We further consider the following two subcases.

Case 2.1: $(r-1)h_{\ell-r+1}\cdots + h_{\ell-1} \geq r$.

Let $(r-1)h_{\ell-r+1}\cdots + h_{\ell-1} = rp + b$, where p and b are two nonnegative integers satisfying $p \geq 1$ and $0 \leq b \leq r-1$. By the induction hypothesis, we obtain

$$(1-(r-1)a)^{h_{\ell-r+1}}\cdots(1-a)^{h_{\ell-1}} \geq (1-ra)^p(1-ba), \tag{53}$$

implying that

$$\begin{aligned}&(1-\ell a)^{h_0}(1-(\ell-1)a)^{h_1}\cdots(1-a)^{h_{\ell-1}}\\ &= \big[(1-\ell a)^{h_0}\cdots(1-(r+1)a)^{h_{\ell-r-1}}\big]\\ &\quad\cdot\big[(1-(r-1)a)^{h_{\ell-r+1}}\cdots(1-a)^{h_{\ell-1}}\big] &&(54)\\ &\geq \big[(1-\ell a)^{h_0}\cdots(1-(r+1)a)^{h_{\ell-r-1}}\big]\big[(1-ra)^p(1-ba)\big] &&(55)\\ &= \big[(1-\ell a)^{h_0}\cdots(1-(r+1)a)^{h_{\ell-r-1}}(1-ra)^{p-1}(1-ba)\big]\\ &\quad\cdot(1-ra) &&(56)\\ &\geq (1-\ell a)^q(1-ra), &&(57)\end{aligned}$$

where the equality (54) holds since $h_{\ell-r} = 0$, the inequality (55) follows from (53) and the inequality (57) follows from Lemma 3 and the equality

$$\ell h_0 + (\ell-1)h_1 + \cdots + (r+1)h_{\ell-r-1} + r(p-1) + b = (\ell q + r) - r = \ell q.$$

Case 2.2: $(r-1)h_{\ell-r+1}\cdots + h_{\ell-1} \triangleq r' < r$.

By the induction hypothesis, we first obtain

$$(1-(r-1)a)^{h_{\ell-r+1}}\cdots(1-a)^{h_{\ell-1}} \geq 1 - r'a. \tag{58}$$

Let $\upsilon = r - r'$ $(0 < \upsilon < r)$. Then, it suffices to prove that

$$\begin{aligned}&(1-\ell a)^{h_0}(1-(\ell-1)a)^{h_1}\cdots(1-(r+1)a)^{h_{\ell-r-1}}\\ &\quad\geq (1-\ell a)^q(1-\upsilon a),\end{aligned} \tag{59}$$

because the inequalities (58) and (59) imply that

$$\begin{aligned}&(1-\ell a)^{h_0}(1-(\ell-1)a)^{h_1}\cdots(1-a)^{h_{\ell-1}}\\ &\quad= \big[(1-\ell a)^{h_0}\cdots(1-(r+1)a)^{h_{\ell-r-1}}\big]\\ &\qquad\cdot\big[(1-(r-1)a)^{h_{\ell-r+1}}\cdots(1-a)^{h_{\ell-1}}\big]\\ &\quad\geq (1-\ell a)^q(1-\upsilon a)(1-r'a)\\ &\quad\geq (1-\ell a)^q(1-ra).\end{aligned}$$

What remains is to prove (59). First, we say

$$\ell h_0 + (\ell-1)h_1 + \cdots + (r+1)h_{\ell-r-1} \tag{60}$$

as the *"1st order" coefficient* of the left hand side (LHS) of (59), since (60) in fact is the coefficient of a in the expansion of the LHS of (59) when considering it as a polynomial of a.

Assume that h_n is the last one not equal to zero in $h_0, h_1, \cdots, h_{\ell-r-1}$ $(n \leq \ell-r-1)$. Then we claim that

$$n(h_n-1) + (n-1)h_{n-1} + \cdots + 2h_2 + h_1 \geq \ell - n - \upsilon. \tag{61}$$

Suppose the contrary that

$$0 \leq h_1 + 2h_2 + \cdots + n(h_n-1) \triangleq \ell - n - \upsilon' < \ell - n - \upsilon, \tag{62}$$

i.e., $\upsilon < \upsilon' \leq \ell - n$. Then we have

$$\begin{aligned}\ell q + \upsilon &= \ell h_0 + (\ell-1)h_1 + \cdots + (r+1)h_{\ell-r-1} &&(63)\\ &= \ell h_0 + (\ell-1)h_1 + \cdots + (\ell-n)h_n &&(64)\\ &= \ell h_0 + (\ell-1)h_1 + \cdots + (\ell-n)(h_n-1) + (\ell-n) &&(65)\\ &= \ell h_0 + \ell h_1 + \cdots + \ell(h_n-1) + (\ell-n)\\ &\quad - \big[h_1 + 2h_2 + \cdots + n(h_n-1)\big] &&(66)\\ &= \ell\big[h_0 + h_1 + \cdots + h_{n-1} + (h_n-1)\big] + \upsilon', &&(67)\end{aligned}$$

where the last equality follows from (62). Thus, we have $q = h_0 + h_1 + \cdots + h_{n-1} + (h_n - 1)$ and $\upsilon = \upsilon'$, contradicting the assumption $\upsilon < \upsilon'$.

By (61), we can see that for some $0 \leq i \leq n$, there exist integers $0 \leq t \leq n-i$ and $h'_{n-i} < h_{n-i}$ such that

$$n(h_n-1) + (n-1)h_{n-1} + \cdots + (n-i)h'_{n-i} + t = \ell - n - \upsilon. \tag{68}$$

Furthermore, it is not difficult to obtain that for any nonnegative integers u_1, u_2 and δ with $0 \leq \delta \leq u_1 \leq u_2$,

$$(1-u_1 a)(1-u_2 a) \geq (1-(u_1-\delta)a)(1-(u_2+\delta)a) \tag{69}$$

and note that the two "1st order" coefficients of the LHS and the RHS of (69) are identical.

Then, we obtain

$$
\begin{aligned}
&(1-(\ell-n)a)^{h_n}(1-(\ell-n+1)a)^{h_{n-1}} \\
&\quad\cdots(1-(\ell-n+i)a)^{h'_{n-i}+1} &(70)\\
&=(1-(\ell-n)a)(1-(\ell-n)a)^{h_n-1}(1-(\ell-n+1)a)^{h_{n-1}} \\
&\quad\cdots(1-(\ell-n+i)a)^{h'_{n-i}+1} &(71)\\
&\ge\Big[1-\big[\ell-n-\big(n(h_n-1)+(n-1)h_{n-1} \\
&\quad+\cdots+(n-i)h'_{n-i}+t\big)\big]a\Big]\cdot\Big[(1-\ell a)^{h_n-1}(1-\ell a)^{h_{n-1}} \\
&\quad\cdots(1-\ell a)^{h'_{n-i}}(1-(\ell-n+i+t)a)\Big] &(72)\\
&=(1-va)\cdot\Big[(1-\ell a)^{h_n-1}(1-\ell a)^{h_{n-1}} \\
&\quad\cdots(1-\ell a)^{h'_{n-i}}(1-(\ell-n+i+t)a)\Big] &(73)\\
&=(1-va)\cdot\Big[(1-\ell a)^{(h_n-1)+h_{n-1}+\cdots+h'_{n-i}} \\
&\quad\cdot(1-(\ell-n+i+t)a)\Big], &(74)
\end{aligned}
$$

where (72) follows by using (69) repeatedly, and (74) follows from (68).

Note that the two "1st order" coefficients of the LHS and the RHS of (72) are identical. Therefore, we further obtain

$$
\begin{aligned}
&(1-\ell a)^{h_0}(1-(\ell-1)a)^{h_1}\cdots(1-(r+1)a)^{h_{\ell-r-1}} &(75)\\
&=(1-\ell a)^{h_0}(1-(\ell-1)a)^{h_1}\cdots(1-(\ell-n)a)^{h_n} &(76)\\
&=\Big[(1-(\ell-n)a)^{h_n}(1-(\ell-n+1)a)^{h_{n-1}}\cdots \\
&\quad(1-(\ell-n+i)a)^{h'_{n-i}+1}\Big]\cdot\Big[(1-(\ell-n+i)a)^{h_{n-i}-h'_{n-i}-1} \\
&\quad\cdot(1-(\ell-n+i+1)a)^{h_{n-i-1}}\cdots(1-\ell a)^{h_0}\Big] &(77)\\
&\ge\Big[(1-va)(1-\ell a)^{(h_n-1)+h_{n-1}+\cdots+h'_{n-i}} \\
&\quad\cdot(1-(\ell-n+i+t)a)\Big]\cdot\Big[(1-(\ell-n+i)a)^{h_{n-i}-h'_{n-i}-1} \\
&\quad\cdot(1-(\ell-n+i+1)a)^{h_{n-i-1}}\cdots(1-\ell a)^{h_0}\Big] &(78)\\
&=(1-va)\Big[(1-\ell a)^{(h_n-1)+h_{n-1}+\cdots+h'_{n-i}} \\
&\quad\cdot(1-(\ell-n+i+t)a)\cdot(1-(\ell-n+i)a)^{h_{n-i}-h'_{n-i}-1} \\
&\quad\cdot(1-(\ell-n+i+1)a)^{h_{n-i-1}}\cdots(1-\ell a)^{h_0}\Big]. &(79)
\end{aligned}
$$

where (78) follows from (70) to (74). Note that still the two "1st order" coefficients of the LHS and the RHS of the inequality (78) are identical. Hence, the "1st order" coefficient of (79) is

$$\ell h_0+(\ell-1)h_1+\cdots+(\ell-n)h_n-v=\ell q.$$

Together with Lemma 3, we obtain from (79) that

$$
\begin{aligned}
&(1-\ell a)^{(h_n-1)+h_{n-1}+\cdots+h'_{n-i}}(1-(\ell-n+i+t)a) \\
&\cdot(1-(\ell-n+i)a)^{h_{n-i}-h'_{n-i}-1}(1-(\ell-n+i+1)a)^{h_{n-i-1}} \\
&\cdot s(1-\ell a)^{h_0}\ge(1-\ell a)^q. &(80)
\end{aligned}
$$

Combining (79) and (80), the inequality (59) is proved.

Acknowledgment

The authors would like to thank the editor and the anonymous reviewers for their valuable suggestions and comments that helped to greatly improve the paper.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

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

[3] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[4] S. Jaggi *et al.*, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973–1982, Jun. 2005.

[5] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[6] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Commun. Control Comput.*, Oct. 2003, pp. 40–49. [Online]. Available: http://sites.google.com/site/saloot2/practical-network-coding.pdf

[7] B. Li and D. Niu, "Random network coding in peer-to-peer networks: From theory to practice," *Proc. IEEE*, vol. 99, no. 3, pp. 513–523, Mar. 2011.

[8] M. Wang and B. Li, "R^2: Random push with random network coding in live peer-to-peer streaming," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 9, pp. 1655–1666, Sep. 2007.

[9] T. Tran, H. Li, W. Lin, L. Liu, and S. Khan, "Adaptive scheduling for multicasting hard deadline constrained prioritized data via network coding," in *Proc. IEEE GLOBECOM*, Dec. 2012, pp. 5843–5848.

[10] B. Li, H. Li, and R. Zhang, "Adaptive random network coding for multicasting hard-deadline-constrained prioritized data," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8739–8744, Oct. 2016.

[11] H. Balli, X. Yan, and Z. Zhang, "On randomized linear network codes and their error correction capabilities," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3148–3160, Jul. 2009.

[12] D. Li, X. Guang, and F.-W. Fu, "The failure probabilities of random linear network coding at sink nodes," *IEICE Trans. Fundam., Electron., Commun. Comput. Sci.*, vol. E99-A, no. 6, pp. 1255–1259, Jun. 2016.

[13] X. Guang and F.-W. Fu, "The average failure probabilities of random linear network coding" *IEICE Trans. Fundam., Electron., Commun. Comput. Sci.*, vol. E94-A, no. 10, pp. 1991–2001, Oct. 2011.

[14] X. Guang, F.-W. Fu, and Z. Zhang, "Construction of network error correction codes in packet networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 1030–1047, Feb. 2013.

[15] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, "Network coding theory," *Found. Trends Commun. Inf. Theory*, vol. 2, nos. 4–5, pp. 241–381, 2005.

[16] R. W. Yeung, *Information Theory and Network Coding*. New York, NY, USA: Springer, 2008.

[17] S. W. Ho, C. W. Tan, and R. W. Yeung, "Proving and disproving information inequalities," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 2814–2818.

[18] X. Yan, J. Yang, and Z. Zhang, "An outer bound for multisource multisink network coding with minimum cost consideration," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2373–2385, Jun. 2006.

[19] R. Dougherty, C. Freiling, and K. Zeger, "Networks, matroids and non-Shannon information inequalities," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 1949–1969, Jun. 2007.

[20] X. Yan, R. W. Yeung, and Z. Zhang, "An implicit characterization of the achievable rate region for acyclic multisource multisink network coding," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5625–5639, Sep. 2012.

[21] S. W. Ho, C. W. Tan, and R. W. Yeung, "Proving and disproving information inequalities," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 2814–2818.

[22] C. Li, S. Weber, and J. M. Walsh, "On multi-source networks: Enumeration, rate region computation, and hierarchy," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7283–7303, Nov. 2017, doi: 10.1109/TIT.2017.2745620.

[23] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

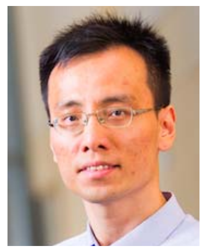
[24] X. Guang and F.-W. Fu, "On random linear network coding for butterfly network," *Chin. J. Electron.*, vol. 20, no. 2, pp. 283–286, Apr. 2011.

Dan Li received the B.S. degree in mathematics from Tianjin University, Tianjin, China, in 2005, and the M.S. and Ph.D. degrees in mathematics from Nankai University, Tianjin, in 2008 and 2012, respectively. She is currently with the Department of Mathematics, Tianjin University of Technology, Tianjin. Her main research interests include functional analysis, operator algebra, and applied mathematics. Her current research is focused on network coding theory.

Xuan Guang (M'12) received the Ph.D. degree in mathematics from Nankai University, Tianjin, China, in 2012. He was a Visiting Researcher and a Post-doctoral Research Fellow with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA. He is currently an Associate Professor with Nankai University. He is also with the Institute of Network Coding, The Chinese University of Hong Kong by the Hong Kong Scholars Program. His research interests include network coding theory, network computation theory, and network information theory.

Zhiheng Zhou (M'17) received the B.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2007 and 2014, respectively. From 2015 to 2017, he was a Post-Doctoral Research Fellow with the Institute of Network Coding, The Chinese University of Hong Kong. He is currently a Lecturer with UESTC. His major research interests include network coding, reliability study, and network optimization.

Congduan Li (S'11–M'15) received the B.S. degree from the University of Science and Technology Beijing, China, in 2008, the M.S. degree from Northern Arizona University, Flagstaff, AZ, USA, in 2011, and the Ph.D. degree from Drexel University, Philadelphia, PA, USA, in 2015, all in electrical engineering. From 2015 to 2016, he was a Post-Doctoral Research Fellow with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong. He is currently a Post-Doctoral Research Fellow with the Department of Computer Science, City University of Hong Kong, Hong Kong.



Chee Wei Tan (M'08–SM'12) received the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, USA. He was a Post-Doctoral Scholar with the California Institute of Technology, Pasadena, CA, USA. He was a Visiting Faculty with Qualcomm Research and Development, San Diego, CA, USA, in 2011. He is currently an Associate Professor with the City University of Hong Kong. His research interests are in networks and graph analytics, statistical inference, cyber-security, and convex optimization theory and its applications. He received the 2008 Princeton University Wu Prize for Excellence and the 2015 Chapter of the Year Award. He was the Chair of the IEEE Information Theory Society Hong Kong Chapter. He was twice selected to participate at the U.S. National Academy of Engineering China-America Frontiers of Engineering Symposium in 2013 and 2015. He currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE/ACM TRANSACTIONS ON NETWORKING.