On Achievable Rates of Line Networks With Generalized Batched Network Coding

Jie Wang, Graduate Student Member, IEEE, Shenghao Yang, Member, IEEE, Yanyan Dong, and Yiheng Zhang

Abstract—To better understand the wireless network design with a large number of hops, we investigate a line network formed by general discrete memoryless channels (DMCs), which may not be identical. Our focus lies on Generalized Batched Network Coding (GBNC) that encompasses most existing schemes as special cases and achieves the min-cut upper bounds as the parameters batch size and inner block length tend to infinity. The inner blocklength of GBNC provides upper bounds on the required latency and buffer size at intermediate network nodes. By employing a “bottleneck status” technique, we derive new upper bounds on the achievable rates of GBNC. These bounds surpass the min-cut bound for large network lengths when the inner blocklength and batch size are small. For line networks of canonical channels, certain upper bounds hold even with relaxed inner blocklength constraints. Additionally, we employ a “channel reduction” technique to generalize the existing achievability results for line networks with identical DMCs to networks with non-identical DMCs. For line networks with packet erasure channels, we make refinement in both the upper bound and the coding scheme, and showcase their proximity through numerical evaluations.

Index Terms—Multi-hop network, line network, batched network code, capacity bound, buffer size, latency.

I. INTRODUCTION

We investigate multi-hop line topology networks formed by concatenating discrete memoryless channels (DMCs), which are fundamental channel models in communication systems. In this line network, the first node serves as the source node, the last node serves as the destination node, and the intermediate nodes establish connections between them. Multi-hop wireless communication networks find applications in diverse domains, including underwater acoustic networks [1], free space optical communication [2], deep space communication networks [3], field area networks [4], and terahertz communications [5].

In the absence of constraints on storage and latency at the intermediate nodes, the network capacity is determined by the min-cut from the source to the destination, achievable through the hop-by-hop implementation of capacity-achieving channel codes [6]. However, as the number of hops increases, the hop-by-hop coding approach introduces significant communication latency and storage requirements at the intermediate nodes, which are critical factors in multi-hop wireless networks [7], [8]. In their work [9], Niesen, Fragouli, and Tuninetti investigated the line network capacity by considering a fixed inner blocklength $N$ at the intermediate nodes. This blocklength has an impact on delay and buffer size. Assuming identical channels in the line network (referred to as $Q$), and when the zero-error capacity of $Q$ is non-zero, they demonstrated that using a constant $N$ allows achieving any constant rate below the zero-error capacity for any given number of hops $L$. Conversely, when the zero-error capacity of $Q$ is zero, a class of codes with a constant $N$ can achieve rates on the order of $\Omega(e^{-cL})$, where $c$ is a constant. Additionally, if $N$ is of the order of $\ln L$, it is possible to achieve any rate below the capacity of $Q$.

However, despite these achievability results, the min-cut remains the strongest upper bound for line networks. It is still uncertain whether the diminishing achievable rates observed with increasing network length are fundamental or if there exist more efficient coding strategies that can achieve higher rates. Furthermore, it is worth exploring the possibility of reducing the processing latency and buffer size requirements beyond the complexity of $O(N)$. With these inquiries in mind, we embark on a comprehensive investigation of line networks formed by DMCs.

Improving the general upper bound for multi-hop networks is an extremely challenging task, as suggested in the network information theory literature [10]. In this paper, our focus is on a specific class of codes called Generalized Batched Network Coding (GBNC). While batched network coding has been extensively studied for networks of packet erasure channels [11], [12], [13], [14], [15], [16], we extend batched network coding to accommodate general DMCs, which may not be identical. GBNC, introduced in §II of this paper, consists of an outer code and an inner code. The outer code...
encodes information messages into batches of coded symbols, while the inner code performs recoding operations within each batch. GBNC incorporates two key parameters: the batch size \( M \) and the inner blocklength \( N \). There are several reasons that make GBNC well-suited for our research objectives. Firstly, GBNC encompasses a wide range of codes as special cases. The coding scheme examined in [9] corresponds to GBNC with \( M = N \). Both decode-and-forward and retransmission schemes can be viewed as special inner codes for GBNC. Secondly, when both \( M \) and \( N \) can be arbitrarily large, GBNC has the capability to achieve the min-cut. Lastly, GBNC enables us to explicitly characterize latency and buffer size. Our formulation reveals that the recoding latency and buffer size at an intermediate node are upper-bounded by a linear order of \( N \).

In this paper, we derive both upper and lower bounds on the achievable rate of GBNC in terms of the parameters \( M, N \), and network length \( L \). Compared to our previous conference papers [17], [18], the main results presented in this paper are either improved or entirely new. Using a “bottleneck status” technique, we obtain new upper bounds on the achievable rate of GBNC for line networks consisting of channels with 0 zero-error capacity. We begin by proving the converses for a class of channels known as canonical channels, which are characterized by having an output symbol that occurs with a positive probability for all possible input symbols, and then extend the results to non-canonical channels (detailed in §III). We demonstrate through various cases that our upper bounds outperform the min-cut.

To gain a more explicit understanding, we conduct further analysis on how the upper and lower bounds scale with \( L \) for different scenarios of \( M \) and \( N \). Notably, when \( N = O(1) \), our upper bound reveals that the achievable rate must decay exponentially with \( L \), aligning with the achievable rates obtained in [9]. By utilizing a “channel reduction” technique (detailed in §IV-A and §IV-C), we extend the achievability results of [9] to line networks with non-identical DMCs. Additionally, when \( N = O(\ln L) \) and \( M = O(1) \), our upper bound indicates that the achievable rate is \( O(1/\ln L) \), which is a new scalability compared with the previous ones obtained in [9]. We demonstrate that rates of \( \Omega(1/\ln L) \) can be attained using \( M = O(1) \) and \( N = O(\ln L) \). In a general decode-and-forward approach, a buffer size of \( O(\ln L) \) is required. However, specific codes enable a reduced buffer size of \( O(\ln L) \) (refer to §IV-B). To exemplify this result, we consider a repetition coding scheme, which prompts us to explore simpler schemes for line networks with a large number of hops. A summarization of the scalability results can be found in Table I.

In the context of line networks with packet erasure channels, we make advancements in both the upper bound and the coding scheme. Through extensive numerical evaluations, we establish a close proximity between the upper bound and the achievable rates of the coding scheme (see §V). This finding serves as motivation for future research endeavors aimed at improving the upper bound and developing more efficient coding schemes tailored to specific channel characteristics.

### Table I

**Summary of the Achievable Rate Scalability for the Channels With () Zero-Error Capacity Using Batched Codes. Here, c and c' Have Constant Values That Do Not Change With L. The Upper/Lower Bound Marked With * Is Obtained in This Paper**

| batch size \( M \) | inner blk-length \( N \) | buffer size | upper bound                  |
|-------------------|--------------------------|-------------|-----------------------------|
| unbounded         | \( O(1) \)               | unbounded   | \( O(e^{-cL}) \)*           |
| \( O(1) \)        | \( O(\ln L) \)           | unbounded   | \( O(1/\ln L) \)*           |

| lower bound       |                                |             |                             |
|-------------------|---------------------------------|-------------|-----------------------------|
| \( O(1) \)        | \( O(1) \)                      | \( O(1) \)  | \( \Omega(e^{-cL}) \)*     |
| \( O(1) \)        | \( O(\ln L) \)                  | \( O(\ln \ln L) \) | \( \Omega(1/\ln L) \)*   |

### Table II

**Some Notations Used in the Paper, Listed in the Alphabetical Order**

| Notation | Explanation                  |
|----------|------------------------------|
| \( A \)  | Batch alphabet.              |
| \( C(Q) \) | Channel capacity of channel \( Q \). |
| \( C_0(Q) \) | Zero-error capacity of channel \( Q \). |
| \( C_0(L,M,N) \) | Maximum achievable rate of all recoding schemes with batch size \( M \) and inner blocklength \( N \). |
| \( E_{0,\ell} \) | Event that all \( N \) outputs of \( Q_\ell \) are equal to the same value regardless of channel input. |
| \( E_\ell \) | Event that there exists one link \( \ell \) such that \( E_{0,\ell} \) holds. |
| \( E_{\ell}^* \) | Smallest coding error exponent among all \( \ell \geq 1 \). |
| \( L \)  | Network length.              |
| \( M \)  | Batch size.                  |
| \( N \)  | Inner blocklength.           |
| \( Q_\ell \) | Discrete memoryless channel of link \( \ell \). |
| \( U_\ell | X_\ell \) | \( L \) uses of the \( \ell \)-th communication link. |
| \( W_L \) | End-to-end transition matrix of the batch channel from \( X \) to \( Y_L \). |
| \( X \in A^M \) | A generic batch. |
| \( X[k] \) | The \( k \)-th entry in \( X \). |
| \( Z_\ell \) | Channel status of \( Q_\ell \). |

Last, our results are extended to networks where certain channels have a positive zero-error capacity (see §VI).

Throughout this paper, we use log to denote the logarithm of base 2, and \( \ln \) to denote the natural logarithm of base \( e \). For random variables represented by uppercase letters (e.g., \( X \)), we use the corresponding lowercase letters (e.g., \( x \)) to represent their instances. We use \( P \) to denote the probability of events, and we may write \( P(X = x) \) as \( P(x) \) to simplify the notation. We use \( p_X \) to denote the probability mass function of the discrete random variable \( X \), where subscripts may be omitted. Most of the notations used throughout this manuscript are given in Table II for easy of reference. All omitted proofs can be found in the supplementary material online [19].

## II. Line Networks and Generalized Batched Network Coding

In this section, we describe the line network model and introduce batched network coding.
A. Line Network Model

A line network of length $L$ consists of nodes labeled as $0, 1, \ldots, L$, with directed communication links from node $\ell - 1$ to node $\ell$. Each link is a discrete memoryless channel (DMC) with fixed finite input and output alphabets $Q_i$ and $Q_o$ respectively. The transition matrix for link $\ell$ is denoted as $Q_\ell$. The line network is formed by concatenating $Q_1, Q_2, \ldots, Q_L$. This study focuses on communication between the first node, referred to as the source node, and the last node, known as the destination node. The nodes numbered $1, 2, \ldots, L - 1$ are referred to as the intermediate nodes.

Let $C(Q)$ and $C_0(Q)$ denote the channel capacity and the zero-error capacity of a DMC with transition matrix $Q$ respectively. Without any constraints at the network nodes, the capacity of the network is given by $\min_{\ell=1}^L C(Q_\ell)$, which is also known as the min-cut. Achieving the min-cut involves using a capacity achieving code at each hop, where intermediate nodes decode the previous link’s code and encode the message using the next link’s code. This scheme is commonly referred to as decode-and-forward. However, as we will discuss later, decode-and-forward is not always the optimal solution when considering both latency and buffer size at the intermediate nodes. Next, we present a general coding scheme for the line network and examine the relationship between the coding parameters and latency as well as buffer size.

B. Generalized Batched Network Coding

A Generalized Batched Network Code (GBNC) comprises an outer code and an inner code. The outer code, executed at the source node, encodes a message from a finite set $\mathcal{A}$ to a finite set $\mathcal{X}$. The parameter $M$ is known as the batch size. The inner code operates on individual batches separately, employing recoding operations at nodes $0, 1, \ldots, L - 1$.

Let’s define the recoding process for a generic batch $X \in \mathcal{A}^M$. At the source node, the recoding transforms the original $M$ symbols of $X$ into $N$ recoded symbols $Y_1$ in $Q_1$, where $N$ is a positive integer referred to as the inner blocklength. The recoding at the source node is represented by the function $\phi_0 : \mathcal{A}^M \rightarrow \mathcal{Q}_1^N$, such that $U_1 = \phi_0(X)$.

At an intermediate node $\ell$, recoding is performed on the $N$ received symbols $Y_0, Y_1, \ldots, Y_{\ell-1}$ in $Q_\ell^N$ to generate $N$ recoded symbols $U_{\ell+1}$ in $Q_\ell^N$ for transmission on the outgoing link of node $\ell$. Due to the memoryless property of $Q_\ell$, the conditional probability of $Y_\ell = y$ given $U_\ell = u$ is

$$P(Y_\ell = y|U_\ell = u) = Q_\ell^{N}(y|u) \triangleq \prod_{i=1}^{N} Q_\ell(y[i]|u[i]),$$

where $y[k]$ $(1 \leq k \leq N)$ represents the $k$th entry in $y$. The recoding at node $\ell$ is represented by the function $\phi_\ell : \mathcal{Q}_\ell^N \rightarrow \mathcal{Q}_\ell^N$, such that $U_{\ell+1} = \phi_\ell(Y_\ell)$. In general, the number of recoded symbols transmitted by different nodes can vary [20], [21]. However, for simplicity, we assume they are all the same for the analysis.

At the destination node, all received symbols, which may belong to different batches, are jointly decoded. The inner code’s end-to-end operation, with the given recoding function $\phi_\ell$ at all nodes, can be viewed as a memoryless channel referred to as a batch channel, which takes $X$ as the input and produces $Y_L$ as the output. Fig. 1 illustrates the variables involved in the recoding process, forming the Markov chain:

$$X \rightarrow U_1 \rightarrow Y_1 \rightarrow \cdots \rightarrow U_L \rightarrow Y_L.$$  \hspace{1cm} (2)

The end-to-end transition matrix $W_L$ of the batch channel can be derived using $\phi_\ell$ and $Q_\ell$.

The outer code serves as a channel code for the batch channel $W_L$ to ensure end-to-end reliability. Given a recoding scheme $\{\phi_\ell\}$, the maximum achievable rate of the outer code is $\max_{p_X} I(X; Y_L)$ for $N$ channel uses, where $p_X$ represents the distribution of $X$. The objective of designing a recoding scheme, given parameters $M$ and $N$, is to maximize $\frac{1}{N} \max_{p_X} I(X; Y_L)$. Let $C_L(M, N)$ denote the maximum achievable rate among all recoding schemes with batch size $M$ and inner blocklength $N$, defined as:

$$C_L(M, N) = \max_{\{\phi_0, p_X\}} \frac{I(X; Y_L)}{N} = \max_{\{\phi_0, p_X\}} \frac{I(p_X; W_L)}{N}. \hspace{1cm} (3)$$

$C_L(M, N)$ is also referred to as the capacity of GBNCs with parameters $M$ and $N$. We can then maximize $C_L(M, N)$ while considering constraints on $M$ and $N$, which impact both the recoding latency and the buffer size.

Recoding functions $\{\phi_\ell\}$ can generally be random. However, the convexity of $I(p_X; W_L)$ for a fixed $p_X$ with respect to $W_L$ implies the existence of a deterministic recoding scheme that achieves $C_L(M, N)$. In particular, the coding scheme analyzed in [9] considers the case where $M = N$. A special inner code known as decode-and-forward will be discussed in §IV. GBNCs generalize the batched network codes studied for networks with packet erasure channels in literature (see discussion in §V).

C. Buffer Size and Latency at Intermediate Nodes

Let’s now delve into the buffer size requirement and latency at the intermediate nodes in GBNCs. In this discussion, we consider a sequential transmission model where symbols of a batch are transmitted consecutively. We will discuss the buffer size required for caching the received symbols for recoding at an intermediate node, as well as the latency between receiving the first symbol of a batch and transmitting the first symbol of the same batch. We will disregard the space and time costs associated with executing recoding $\phi_\ell$.

The key principle of GBNCs is the independent application of recoding to each batch. In the worst case scenario, an intermediate node begins transmitting the first recoded symbol of a batch only after receiving all $N$ symbols of that batch. Consequently, the latency of a batch at an intermediate node is upper bounded by $O(N)$. Since an intermediate node can only transmit symbols of a batch after receiving at least

[Fig. 1. A line network with the random variables involved in recoding.]
one symbol from that batch, the lower bound on the latency at an intermediate node is 1. The accumulated end-to-end recoding latency across all intermediate nodes falls within the range of $\Omega(L)$ to $O(NL)$.

Similarly, in the worst-case scenario, an intermediate node starts transmitting the first recoded symbol of a batch only after receiving all $N$ symbols of that batch. Additionally, these received symbols need to be cached for $N$ more channel uses. Therefore, an intermediate node needs to cache at most $2N$ symbols: $N$ symbols of the batch for transmitting and $N$ symbols of the same batch for receiving. This indicates that the buffer size required for caching symbols at an intermediate node is $O(N)$.

III. CONVERSE FOR LINE NETWORKS OF CHANNELS WITH 0 ZERO-ERROR CAPACITY

One known upper bound of $C_L(M, N)$ is the min-cut
\[
\min_{\ell=1}^{\infty} C(Q_\ell).
\]
However, this bound may not be sufficient for small values of $M$ and $N$. When $C_0(Q_\ell) = 0$ for all $\ell$, in this section, we introduce a technique called a “bottleneck status” to derive a potentially tighter bound on $C_L(M, N)$ when $M$ and $N$ are small.

The bottleneck status refers to an event $E_0$ that is associated with the channel $W_L$ and is independent of $X$. Let
\[
\begin{align*}
W_L^{(0)}(y | x) &= P(Y_L = y | X = x, E_0), \\
W_L^{(1)}(y | x) &= P(Y_L = y | X = x, \overline{E_0}).
\end{align*}
\]

The channel $W_L$ can be expressed as $W_L = W_L^{(0)}p_0 + W_L^{(1)}p_1$, where $p_0 = P(E_0)$, $p_1 = P(\overline{E_0})$. As mutual information $I(p_X, W_L)$ is convex w.r.t. $W_L$ for given $p_X$, we can establish the upper bound as follows:
\[
I(p_X, W_L) \leq p_0 I(p_X, W_L^{(0)}) + p_1 I(p_X, W_L^{(1)}).
\]

The crucial step is to design the event $E_0$ in order to obtain the desired upper bound.

**Definition 1:** For $0 < \varepsilon \leq 1$, we call a DMC $Q : Q_1 \rightarrow Q_0$ an $\varepsilon$-canonical channel if there exists $y^* \in Q_0$ such that for every $x \in Q_1$, $Q(y^* | x) \geq \varepsilon$.

For a canonical channel, there exists an output symbol $y^*$ that occurs with a positive probability for all inputs. The binary erasure channel (BEC) and binary symmetric channel (BSC) are both canonical channels, but a typewriter channel is non-canonical. Note that a canonical $Q$ has $C_0(Q) = 0$.

We first introduce our technique to design a bottleneck status for canonical channels, and then discuss the general channels.

A. Line Network of Canonical Channels

In this subsection, we study a line network consisting of $\varepsilon$-canonical channels $Q_\ell, \ell = 1, \ldots, L$. To design the bottleneck status $E_0$, we adopt a formulation of DMCs in [22, §7.1]. Define $Z = (Z[x], x \in Q_1)$, where $Z[x], x \in Q_1$ are independent random variables on $Q_1$ with the distribution $P(Z[x] = y) = Q(y|x)$. The relation between the input $X$ and output $Y$ of a DMC $Q$ can be modeled as
\[
Y = \alpha(X, Z = (Z[x], x \in Q_1)) \triangleq \sum_{x \in Q_1} 1\{X = x\} Z[x],
\]
where $1$ denotes the indicator function. Here $Z = (Z[x], x \in Q_1)$ is also called channel status variable, and $\alpha$ is called the channel function. We denote by $\alpha_\ell$ the channel function of $Q_\ell$.

Consider a GBNC with inner blocklength $N$ for the line network. With the alternative channel formulation (6), we can write for $\ell = 1, \ldots, L$, and $i = 1, \ldots, N$, $Y_\ell[i] = \alpha_\ell(U_\ell[i], Z_\ell[i])$. Here $Z_\ell[i] = (Z_\ell[i, x], x \in Q_1)$ is the channel status variable for the $i$th use of the channel $Q_\ell$, where
\[
P(Z_\ell[i, x] = y) = Q_\ell(y|x).
\]

Define $Z_\ell = (Z_\ell[i], i = 1, \ldots, N)$. For notation simplicity, we rewrite the channel relation as
\[
Y_\ell = \alpha_\ell(N) (U_\ell, Z_\ell).
\]

Given that $Q_\ell$ is $\varepsilon$-canonical, there exists an output denoted as $y^*_\ell$ satisfying
\[
Q_\ell(y^*_\ell | x) \geq \varepsilon \text{ for all } x \in Q_1.
\]

Let’s define
\[
E_{0, \ell} = \{Z_\ell[i, x] = y^*_\ell, i \in \{1, \ldots, N\}, x \in Q_1\}.
\]

Under the condition $E_{0, \ell}$, all $N$ outputs of $Q_\ell$ are equal to $y^*_\ell$ for any possible channel input, rendering the channel useless. We can quantify the probability of $E_{0, \ell}$ as follows:
\[
P(E_{0, \ell}) = \prod_{i \in \{1, \ldots, N\}, x \in Q_1} P(Z_\ell[i, x] = y^*_\ell) = \prod_{i \in \{1, \ldots, N\}, x \in Q_1} Q_\ell(y^*_\ell | x) \geq \varepsilon^N.
\]

where (11) follows from (7), and (13) follows from (9). Now we define the bottleneck status
\[
E_0 = \bigvee_{\ell=1}^L E_{0, \ell}.
\]

This event implies the existence of at least one link $\ell$ in the network that is deemed useless and hence the network is useless.

**Lemma 1:** When $Q_\ell$, $\ell = 1, \ldots, L$ are all $\varepsilon$-canonical channels, for $W_L^{(0)}$ defined in (4a) and $E_0$ defined in (14),
\[
I(p_X, W_L^{(0)}) = 0.
\]

**Lemma 2:** When $Q_\ell$, $\ell = 1, \ldots, L$ are all $\varepsilon$-canonical channels, for $W_L^{(1)}$ defined in (4b) and $E_0$ defined in (14), we have
\[
P(\overline{E_0}) \leq (1 - \varepsilon^{Q_1[N]}L)
\]

2) for any $\ell = 1, \ldots, L$
\[
I(p_X, W_L^{(1)}) \leq \max_{p_{U_\ell}} I(U_\ell; Y_\ell | E_{0, \ell}).
\]

In Lemma 2, $\max_{p_{U_\ell}} I(U_\ell; Y_\ell | E_{0, \ell})$ is the capacity of the channel $Q_\ell^N$ under the condition $E_{0, \ell}$. One upper bound is $\sum_{\ell=1}^L \max_{p_{U_\ell}} I(U_\ell; Y_\ell | E_{0, \ell}) \leq \log \min(|Q_1|, |Q_0|)$. In the following lemma, we give a better upper bound that converges $C(Q_\ell)$ when $N$ tends to infinity.

**Lemma 3:** Consider a channel $Q$ as defined in (6) by $(\alpha, Z)$. Fix an output $y^*$ such that $Q(y^* | x) = P(Z[x] = y)$...
y^*) \geq \epsilon \text{ for all input } x, \text{ where } \epsilon > 0. \text{ For } N \text{ uses of the channel, let } Z[i, x] \text{ be the channel variable of the } i\text{th uses associated with the input } x. \text{ Let } E_0 \text{ be the event that } \{Z[i, x] = y^*, i = 1, \ldots, N, x \in \mathcal{Q}_i\}. \text{ Let } W \text{ be the channel formed by } N \text{ uses of } Q \text{ under the condition of } E_0. \text{ Let }

\begin{equation}
D(Q, N) = (q^* + p_0) \log \frac{q^* - p_0}{e^{N} - p_0} + q^* \log \frac{e^{N}}{q^*} \tag{17}
\end{equation}

where \( p_0 = (\prod_x Q(y^*)|x)^N \), and \( q^* = \max_x Q^{\otimes N}(y^*|x) \).

Then

\[ \frac{1}{N} I(p, W) \leq C^*(Q, N) \triangleq \frac{1}{1 - p_0} \left( C(Q) + D(Q, N) \right). \tag{18} \]

Based on the relation (5), together with Lemmas 1, 2, and 3, we derive the following theorem.

**Theorem 4**: Consider a length-\( L \) line network of \( \varepsilon \)-canonical channels with finite input and output alphabets \( \mathcal{Q}_i \) and \( \mathcal{Q}_o \), respectively. The capacity of GBNCs with batch size \( M \) and inner blocklength \( N \) has the following upper bound:

\[ C_L(M, N) \leq (1 - \varepsilon|Q_i|)^L \min \left\{ C^*(Q_i, N), \right. \]

\[ \left. \times \log |Q_i|, \log |Q_o|, \frac{M \log |A|}{N} \right\} \tag{19} \]

Moreover,

1) when \( N = O(1) \), \( \max C_L(M, N) = O(1 - \varepsilon|Q_i|^L) \);
2) when \( M = O(1) \), \( \max C_L(M, N) = O(1/\ln L) \);
3) when \( M \) and \( N \) are arbitrary, \( \max C_L(M, N) = O(1) \).

**Proof**: Recall the capacity of GBNC in (3), where \( rCHI(p_X, W_L) \leq p_0 I(p_X, W_L^{(0)}) + p_1 I(p_X, W_L^{(1)}) \)

\[ = p_1 I(p_X, W_L^{(1)}) \leq (1 - \varepsilon|Q_i|^L) L I(p_X, W_L^{(1)}) \tag{22} \]

where (20) follows from (5), (21) is obtained by applying Lemma 1, and (22) follows from Lemma 2-1. The upper bound in (19) is proved by

\[ I(p_X, W_L^{(1)}) \leq H(X) \leq M \log |A| \tag{23} \]

\[ I(p_X, W_L^{(1)}) \leq I(U; Y, L | E_{0, L}) \leq N \log \min(|Q_i|, |Q_o|) \tag{24} \]

\[ I(p_X, W_L^{(1)}) \leq NC^*(Q_i, N), \tag{25} \]

where (24) follows from Lemma 2-2 and (25) holds due to Lemma 3.

The remainder part of the theorem is proved by analyzing the upper bound in (19) for different values of \( M \) and \( N \). In particular, Case 2) is obtained using the following Lemma 5.

**Lemma 5**: For fixed real number \( 0 < \epsilon < 1 \) and integer \( L > 1 \), the function \( F(N) = (1 - \varepsilon^L)/N \) of integer \( N \) is maximized when \( N = \Theta(\ln L) \), and the optimal value of \( F(N) \) is \( \Theta(\ln(1/\epsilon)/\ln L) \).

To illustrate the capacity upper bound in Theorem 4, we evaluate it for the network formed by BSCs in Fig. 2, and use the min-cut for baseline comparison. Fig. 2(a) depicts, for each hop length \( L \), the upper bound (19) when \( M = O(1) \). It reveals the exponential decay of the capacity with respect to \( L \), and the min-cut is in general a loose upper bound for sufficiently large \( L \). Fig. 2(b) shows the upper bound (19) when \( M = O(1) \), \( N = O(\ln L) \). In this case, the capacity decays slowly as \( L \) increases, and the min-cut is a loose upper bound as well.

**B. General Channels**

Consider a channel \( Q : \mathcal{Q}_i \rightarrow \mathcal{Q}_o \) with \( C_0(Q) = 0 \), modeled as in (6). Since \( Q \) may not be canonical, there may not exist an output symbol that occurs with a positive probability for all inputs. Furthermore, if \( Q \) is non-canonical, \( Q^{\otimes m} \) is also non-canonical for any positive integer \( m \). For instance, let’s define the channel \( Q_{3 \times 3} \) with \( \mathcal{Q}_i = \mathcal{Q}_o = \{0, 1, 2\} \) and

\[ Q_{3 \times 3}(0|0) = Q_{3 \times 3}(0|1) = Q_{3 \times 3}(1|0) = Q_{3 \times 3}(1|2) \]

\[ Q_{3 \times 3}(2|1) = Q_{3 \times 3}(2|2) = 1/2 \]. We can check that \( Q_{3 \times 3}^{\otimes m} = Q_{3 \times 3} \) is non-canonical. Consequently, the bottleneck status we observe for a canonical channel cannot be directly extended to non-canonical channels.

To investigate the converse of general channels, we employ a technique that involves concatenating multiple channels through recoding, resulting in a new channel that is canonical. Let’s use the example of \( Q_{3 \times 3} \) to illustrate this idea. We consider the concatenation of two copies of \( Q_{3 \times 3} \) using a \( 3 \times 3 \) deterministic transition matrix \( \Phi \), yielding the new channel \( W = Q_{3 \times 3} \Phi Q_{3 \times 3} \). In this setup, \( \Phi \) maps an output...
of the first channel as an input of the second channel. Refer to the illustration in Fig. 3. For the first channel, it is guaranteed that at one of the output in the set \{0, 1\} occurs with a positive probability for any input. Recoding \( \Phi \) can map the outputs 0 and 1 of the first channel to either the same input or two distinct inputs of the second channel. Due to the properties of \( Q_{3 \times 3} \), regardless of the specific mapping, there will always exist an output of \( W \) that occurs with a positive probability for any input of \( W \).

Now we discuss the general case. For a channel \( Q: Q_1 \rightarrow Q_o \), denote by \( \varepsilon_Q \) the maximum value such that for any \( x, x' \in Q_1 \), there exists \( y \in Q_o \) such that \( Q(y|x) \geq \varepsilon_Q \) and \( Q(y|x') \geq \varepsilon_Q \). In the case of \( Q_{1\times 2} \), we have \( \varepsilon_{Q_{1\times 2}} = 1/2 \). Note that \( \varepsilon_Q > 0 \) if and only if \( C_0(Q) = 0 \) (see [23]). Since \( C_0(Q_1) = 0 \), it is possible to observe the same output for any two channel inputs of \( Q_1 \). Exploiting this property, we can prove that for any subset \( S_1 \) of \( Q_1 \), there exists a subset \( S_0 \) of \( Q_o \) with a size less than half of \( S_1 \), such that for any input in \( S_1 \), it is possible to observe an output in \( S_0 \). This can be formally stated as the following lemma.

**Lemma 6:** Consider a DMC \( Q: Q_1 \rightarrow Q_o \) with \( \varepsilon_Q > 0 \) modelled by \((\alpha, Z)\). For any non-empty set \( S_1 \subseteq Q_1 \), there exist a subset \( Z \) of the range of \( Z \) and a subset \( S_0 \subseteq Q_o \) with \(|S_0| \leq \lceil |S_1|/2 \rceil \) such that \( \alpha(x, z) \in S_0 \) for any \( x \in S_1 \) and \( z \in Z \), and \( P(Z \in Z) \geq \varepsilon_Q |S_0| \).

Based on the aforementioned lemma, we can concatenate a sufficiently large number of consecutive channels in a line network to create a canonical channel. In order to establish the upper bound, we need to demonstrate that for a certain \( \varepsilon > 0 \) and any recoding schemes, a number of consecutive channels in the line network form an \( \varepsilon \)-canonical channel. The following lemma provides justification for this feasibility.

**Lemma 7:** Let \( K = \lceil N \log |Q_1| \rceil \). Consider a line network of \( K \) DMCs \( Q_1 \) with \( \varepsilon_{Q_1} \geq \varepsilon > 0 \). For any deterministic GBNC with the inner blocklength \( N \) and the recoding functions \( \{\phi_k\} \), let \( G = Q_1^{\otimes N} \phi_1 Q_2^{\otimes N} \cdots \phi_{K-1} Q_{K-1}^{\otimes N} \). Then \( G \) is \( \varepsilon(N(2|Q_1|+|K|)) \)-canonical.

**Proof:** Consider a deterministic GBNC as described in \( \S \). Channel \( Q_1^{\otimes N} \) can be modelled by the function \( \alpha_{N} \) with the channel status variable \( Z_t \) as in (8). As \( \varepsilon_{Q_1^{\otimes N}} \geq \varepsilon_{Q_1} \geq \varepsilon > 0 \), the condition of applying Lemma 6 on \( Q_1^{\otimes N} \) is satisfied.

Let \( S^{(1)}_1 = Q_1^{\otimes N} \). Applying Lemma 6 on \( Q_1^{\otimes N} \) w.r.t. \( S^{(1)}_1 \), there exists subsets \( Z^{(i)} \) of the range of \( Z_t \) and \( S^{(i)}_1 \subseteq Q_1^{\otimes N} \), with \(|S^{(i)}_0| \leq \lceil |S^{(i)}_1|/2 \rceil \) such that \( \alpha_{N}^{(i)}(x, z_1) \in S^{(i)}_0 \) for any \( x \in S^{(i)}_1 \) and \( z_1 \in Z^{(i)}_1 \), and \( P(Z_t \in Z^{(i)}) \geq \varepsilon (2|Q_1|) \).

For \( i = 2, 3, \ldots, K \), define recursively

\[
S^{(i)}_1 = \left\{ \phi_{i-1}(y) : y \in S^{(i-1)}_0 \right\},
\]

and \( S^{(i)}_0 \) and \( Z^{(i)} \) as in the proof of Lemma 6 w.r.t. \( Q_1^{\otimes N} \) and \( S^{(i)}_1 \) so that \( \alpha_{N}^{(i)}(x, z) \in S^{(i)}_0 \) for any \( x \in S^{(i)}_1 \) and \( z \in Z^{(i)}_1 \), and \( P(Z_t \in Z^{(i)}_1) \geq \varepsilon (2|Q_1|) \). According to the construction, \(|S^{(i)}_1| \leq |S^{(i-1)}_1|/2 \) and \(|S^{(i)}_0| \leq \lceil |S^{(i)}_1| \rceil/2 \). Hence \(|S^{(K)}_0| \leq \lceil |S^{(1)}|/2K \rceil = 1 \). Since the set \( S^{(K)}_0 \) is non-empty, we have \(|S^{(K)}_0| = 1 \), i.e., there exists an output of \( Q_1^{\otimes N} \) that occurs with a positive probability for all inputs of \( Q_1^{\otimes N} \).

Under the condition \( Z_t \in Z^{(i)}_1, i = 1, \ldots, K \), the output of \( G \) must be unique for all possible channel inputs. Note that

\[
P(Z_t \in Z^{(i)}_1, i = 1, \ldots, K) \geq \varepsilon (2|Q_1|) \cdot \frac{|G|}{|Q_1|}.
\]

The proof is completed.

Based on the aforementioned lemma, we are now ready to prove the upper bound for the general case. The main idea is to divide the line network into consecutive segments, each consisting of \( K \) consecutive channels. Lemma 7 guarantees that each segment can form a canonical channel. In contrast to the proof of Theorem 4, the key difference lies in the definition of the bottleneck status. In this case, we can utilize \( Z_t \) in the proof of Lemma 7 to define the bottleneck status. This demonstrates another way of applying the bottleneck status technique.

**Theorem 8:** Consider a length-\( L \) line network of channels \( \{Q_t\}_{t=1}^L \) with finite input and output alphabets and \( \varepsilon_{Q_t} \geq \varepsilon > 0 \) for all \( t \). When \( L > N \log |Q_t| \), the capacity of GBNCs with batch size \( M \) and inner blocklength \( N \) has the following upper bound:

\[
C_L(M, N) \leq (1 - \varepsilon(2|Q_1|+|K|))L/K \cdot \min \left\{ M/N \log |A|, \log |Q_1|, \log |Q_o| \right\},
\]

where \( K = \lceil N \log |Q_1| \rceil \). Moreover,

1) when \( N = O(1) \), \( \max C_L(M, N) = O((1 - \varepsilon)L) \) for certain \( \varepsilon \in (0, 1) \);
2) when \( M = O(1) \) and \( N = \Omega(\ln L) \), \( \max C_L(M, N) = O(1/\ln L) \);
3) when \( M \) and \( N \) are arbitrary, \( \max C_L(M, N) = O(1) \).

**Proof:** Let \( L' = \lceil L/K \rceil \). As \( L > N \log |Q_t| \), we have \( L' \geq 1 \). Consider a GBNC as described in \( \S \). Without loss of optimality, we assume a deterministic recoding scheme, i.e., \( \phi \) are deterministic. For \( i = 2, \ldots, L' \), define

\[
G_i = Q_1^{\otimes N}_1 \phi_{i-1}^{(i-1)+1} \phi_{i-1}^{(i-2)+1} \cdots \phi_{i-1}^{(1)} Q_{K-1}^{\otimes N}. 
\]

According to Lemma 7, we know that \( G_i \), \( i = 2, \ldots, L' \) are all \( \varepsilon(N(2|Q_1|+|K|)) \)-canonical and forms a length-\( L' \) network.

Let \( W_L' = \phi_1 G_1 \phi_2 G_2 \phi_3 G_3 \cdots G_L' \), which is the end-to-end transition matrix of a GBNC with inner blocklength 1 for the length-\( L' \) network of canonical channels \( G_i \). By the data processing inequality, \( I(p_{X_L} W_{L'}) \leq I(p_{X_L} W_{L}) \).

Fix an \( \ell \in \{1, 2, \ldots, L' \}. \) Considering the sets \( Z^{(i)}, i = 1, \ldots, K \) in the proof of Lemma 7 for \( G_{\ell} \), define

\[
E_{0, \ell} = \{ Z_t \in Z^{(i)}, i = 1, \ldots, K \}.
\]
Define the bottleneck status $E_0 = \land_{\ell=1}^{\ell'} E_{0,\ell}$. Let $p_1 = P(E_0)$ and $p_0 = 1 - p_1$. Using this bottleneck status $E_0$, we can define $W_{L'}^{(0)}$ and $W_{L'}^{(1)}$ as in (4). Similar as the proof of Theorem 4, we have

$$r\text{CH}(p_{X}, W_{L'}) \leq p_0 r\text{CH}(p_{X}, W_{L'}^{(0)}) + p_1 r\text{CH}(p_{X}, W_{L'}^{(1)}) \leq (1 - \varepsilon^{N(2\ell)^{N+K}}) L r\text{CH}(p_{X}, W_{L'}^{(1)})$$

(29)

$$= p_1 r\text{CH}(p_{X}, W_{L'}^{(1)})$$

(30)

$$\leq (1 - \varepsilon^{N(2\ell)^{N+K}}) L r\text{CH}(p_{X}, W_{L'}^{(1)})$$

(31)

where (31) follows from (27). The proof is completed by bounding $r\text{CH}(p_{X}, W_{L'}^{(1)})$ using the alphabet size.

**Remark 1:** Theorem 4 provides stronger results for line networks of canonical channels compared to Theorem 8. The upper bound given in (19) is strictly better than the one in (28). For general channels, it is possible to further improve Theorem 8 by enhancing Lemma 3. However, when directly applying Lemma 3 to canonical channels $G_{\ell}$ in the proof of Theorem 8, the resulting $D(G_{\ell}, 1)$ depends on the specific GBNC employed. In order to prove an upper bound that holds independently of the chosen GBNC, it would be necessary to establish a GBNC-independent upper bound on $D(G_{\ell}, 1)$. This matter is not discussed in the current paper.

**IV. Achievable Rates Using Decode-and-Forward**

In this section, we discuss the lower bounds of the achievable rates of line networks. We will first study the achievable rates when $N = O(\ln L)$ using two recoding schemes: decode-and-forward and repetition, which can achieve different scalability of the buffer size. When $N = O(1)$, for a line network of identical channels, a rate that exponentially decays with $L$ can be achieved as proved in [9]. We will extend their results for line networks where channels may not be identical.

A. Decode-and-Forward Recoding

We discuss a class of GBNC recoding called decode-and-forward. When there is a trivial outer code, decode-and-forward has been extensively studied and widely applied in the existing communication systems [10]. We first describe decode-and-forward recoding in the GBNC framework, and then discuss the achievable rates.

Following the notations in §II-B, we consider a GBNC with batch size $M$. Let $(f_{\ell}, g_{\ell})$ be a channel code for $Q_{\ell}$ where $f_{\ell}: A^{M} \rightarrow Q_{\ell}^{N}$ and $g_{\ell}: Q_{\ell}^{N} \rightarrow A^{M}$ are the encoding and decoding functions, respectively. Consider the transmission of a generic batch $X$. The source node transmits $U_{1} = f_{1}(X)$. Each intermediate node $\ell$ first receives $Y_{\ell}$ and then transmits $U_{\ell+1} = f_{\ell+1}(g_{\ell}(Y_{\ell}))$. In other words, the encoding function $f_{\ell}$ behaves as follows:

- For $i = 1, \ldots, N$, the node $\ell$ just keeps the received symbols in the buffer. Therefore, the buffer size is $\Theta(N)$.
- After receiving the $N$ symbols of $Y_{\ell}$, the node $\ell$ generates $f_{\ell+1}(g_{\ell}(Y_{\ell}))$. If the decoding is correct at nodes $1, \ldots, \ell$, then $g_{\ell}(Y_{\ell}) = X$ and $U_{\ell+1} = f_{\ell+1}(X)$.

Let $\varepsilon_{\ell}$ denote the maximum decoding error probability of $(f_{\ell}, g_{\ell})$ for $Q_{\ell}$. Due to the fact that if the decoding is correct at all the nodes $1, \ldots, L$, it holds that $g_{L}(Y_{L}) = X$, we have

$$P(g_{L}(Y_{L}) \neq X) \leq 1 - \prod_{\ell=1}^{L} (1 - \varepsilon^{N\varepsilon_{\ell}(r)})$$

(32)

where $\varepsilon_{\ell}$ is the random coding error exponent for $Q_{\ell}$.

Using a similar argument as in the proof of [9, Theorem 5.3], the GBNC achieves rate $r (1 - e^{-N\varepsilon_{\ell}(r)})^L - 1/N$. Next, we discuss the scalability of the rate for different scalings of $M$ and $N$. 1) Suppose $M = \Theta(N)$, i.e., $r_1 \leq r \leq r_2$ for some $0 < r_1 < r_2 < C'$, and $N = O(\ln L)$. In this case,

$$\max C_{L}(M, N) \geq r_1 \left(1 - e^{-N\varepsilon_{\ell}(r_2)}\right)^L - \frac{1}{N}. \quad (37)$$

Using a similar argument as in the proof of [9, Theorem V.3], the GBNC achieves rate $r (1 - e^{-N\varepsilon_{\ell}(r_2)})^L - 1/N$. Next, we discuss the scalability of the rate for different scalings of $M$ and $N$. 1) Suppose $M = \Theta(N)$, i.e., $r_1 \leq r \leq r_2$ for some $0 < r_1 < r_2 < C'$, and $N = O(\ln L)$. In this case,

$$\max C_{L}(M, N) \geq r_1 \left(1 - e^{-N\varepsilon_{\ell}(r_2)}\right)^L - \frac{1}{N}. \quad (37)$$

Since $N = O(\ln L)$, it holds that $(1 - e^{-N\varepsilon_{\ell}(r_2)})^L = \Theta(1)$ and $1/N = o(1)$. Consequently, the lower bound in (37) is $\Theta(1)$. 
We first discuss the case $M = 1$. For any $\ell$, let $Q_\ell^f$ be the maximal subset of $Q$ such that for any $x \neq x' \in Q_\ell^f$, $Q_\ell(x) \neq Q_\ell(x')$. For $\ell = 1, \ldots, L$, assume $|Q_\ell^f| \geq |A| \geq 2$, and let $u_\ell$ be a one-to-one mapping from $A$ to $Q_\ell^f$. For a generic batch $x \in A$ with $M = 1$, node $\ell - 1$ transmits $u_\ell(x)$ for $N$ times, i.e.,

$$f_\ell(x) = (u_\ell(x), \ldots, u_\ell(x)).$$

(39)

Suppose $y_\ell = y_\ell^o$, i.e., node $\ell$ receives $y_\ell$ for the transmission $f_\ell(x)$. The decoding function $g_\ell$ is defined based on the maximum likelihood (ML) criterion:

$$g_\ell(y_\ell) = \arg \max_{x \in A} \prod_{i=1}^N Q_\ell(y_\ell[i] | u_\ell(x)),$$

(40)

where a tie is broken arbitrarily. Let

$$L_\ell(x, y_\ell) = \sum_{i=1}^N \ln Q_\ell(y_\ell[i] | u_\ell(x)) = \sum_{y \in Q_\ell^f} N(y) \ln Q_\ell(y | u_\ell(x)),$$

(41)

where $N(y \mid y_\ell)$ denote the number of times that $y$ appears in $y_\ell$. Then the ML decoding problem can be equivalently written as $g_\ell(y_\ell) = \arg \max_{x \in A} L_\ell(x; y_\ell)$.

To perform the ML decoding, node $\ell$ needs to count the frequencies of symbols $y$ for any $y \in Q_\ell^f$ among $N$ received symbol. As a result, a buffer of size $O(\log(N))$ is required. Additionally, the computation cost of the repetition recoding is $O(N)$ per batch. The following lemma bounds the maximum decoding error probability $\epsilon_\ell$ of $(f_\ell, g_\ell)$ for $Q_\ell$.

**Lemma 10:** For any $\ell = 1, \ldots, L$, under the condition $|Q_\ell^f| \geq |A| \geq 2$, using the repetition encoding $f_\ell$ and the ML decoding $g_\ell$ in (39) and (40), respectively, the maximum decoding error probability $\epsilon_\ell$ for $Q_\ell$ satisfies $\epsilon_\ell \leq (|A| - 1) \exp(-NE_\ell)$, where $E_\ell > 0$ is a constant depends only on the channel $Q_\ell$.

Consider a sequence of DMCs $Q_\ell, \ell = 1, 2, \ldots$ with $C(Q_\ell) > 0, \ell \geq 1$. Let

$$E^* = \inf \{E_\ell, \ell \geq 1\}.$$

(42)

We choose the alphabet $A$ such that $|A| \in [2, S^*]$, where

$$S^* = \inf\{|Q_\ell^f| : \ell \geq 1\}.$$

(43)

Note that when $C(Q_\ell) > 0, |Q_\ell^f| \geq 2$. Hence $S^* \geq 2$. Considering the repetition coding,

$$P(g_L(Y_L) = X) \geq \left(1 - e^{-NE^*}\right)^L.$$

Applying an argument in [9, Theorem V.3], we obtain the following theorem.

**Theorem 11:** For the line network of length $L$, the GBNC with repetition recoding scheme, batch size $M = 1$, inner blocklength $N$, and batch alphabet $A$ achieves rate

$$C_L(1, N) \geq \frac{1}{N} \left\{ \log |A| - H \left(\left(1 - e^{-NE^*}\right)^L\right) \right. \left. - \left(1 - \left(1 - e^{-NE^*}\right)^L\right) \log(|A| - 1) \right\},$$

(44)
where $H(\cdot)$ denotes the binary entropy function. When $N = O(\ln L)$, $\max C_L(1, N) = \Omega(1/\ln L)$.

We plot the rate of repetition coding using BSC with crossover error probability $\epsilon \in \{0.05, 0.1, 0.15, 0.2\}$ and $|A| = 2$ with respect to the hop length $L$ in Fig. 5. In Fig. 5(a), for each hop length $L$, we plot the optimal value of $N$ maximizing the lower bound (44), which is denoted as $N^*_L$. This illustration highlights the observed trend of $N^*_L$ increasing roughly in the order of $\ln L$. In Fig. 5(b), we plot the lower bound (44) for each hop length $L$, showcasing an approximate decrease rate in the order of $1/\ln L$.

The repetition coding scheme discussed previously has a limitation that $|A| \leq |Q|$. We can extend the scheme by multiple uses of $Q_\ell$. For an integer $m$, let $Q_i^{m, \ell}$ be the maximum subset of $Q_i^m$ such that for any $x \neq x' \in Q_i^{m, \ell}$, $Q_\ell^{\otimes m}(x) \neq Q_\ell^{\otimes m}(x')$. Define

$$S^{m, \ell} = \inf \{|Q_i^{m, \ell}| : \ell \geq 1\}. \quad (45)$$

Fix $m, M$ and a finite alphabet $A$ such that $|A|^M \in [2, S^{m, \ell}]$. Consequently, we can view the line network of channels $Q_1, \ldots, Q_L$ as one of $Q_1^{2m}, \ldots, Q_L^{2m}$. For the latter, we can apply the repetition recoding with batch size $1$, inner blocklength $\tilde{N}$ and the batch alphabet $A^{\tilde{N}}$, which for the original line network of $Q_1, \ldots, Q_L$ is a GBNC with batch size $M$, inner blocklength $m\tilde{N}$ and the batch alphabet $A$. Based on Theorem 11, such a coding scheme achieves rate

$$\frac{1}{N} \left\{ \log |A|^M - H \left( \left( 1 - e^{-N E^{\epsilon}} \right)^L \right) - \left( 1 - \left( 1 - e^{-N E^{\epsilon}} \right)^L \right) \log(|A|^M - 1) \right\}. \quad (46)$$

While the repetition code may appear straightforward, it serves as an illustrative example of how to reduce the buffer size at the intermediate node. Using convolutional codes with Viterbi decoding, due to their analogous encoding and decoding nature, can achieve the same order of the buffer size. However, the corresponding achievable rate is challenging to analyze.

### C. Channel Reduction

When all the links in the line network are identical DMCs, it has been shown in [9] that an exponentially decreasing rate can be achieved using $N = O(1)$, which corresponds to the first case in Theorem 8. Here we discuss how to generalize this scalability result to line networks where the DMCs $Q_\ell$ are not necessarily identical. Our approach is to perform recoding so that the line network is reduced to one with identical channels.

We introduce the reduction of an $m \times n$ stochastic matrix $Q$ with $C(Q) > 0$. Let $r = \text{rank}(Q)$. Note that $C(Q) > 0$ if and only if $r \geq 2$. Let $s$ be an integer such that $2 \leq s \leq r$. We would like to reduce $Q$ by multiplying an $s \times m$ matrix $R$ and an $n \times s$ matrix $S$ before and after $Q$, respectively, so that $RQS$ becomes an $s \times s$ matrix $U_s(q)$ with $(U_s(q))_{i,j} = q$ if $i = j$ and otherwise $(U_s(q))_{i,j} = \frac{1 - q}{s - 1}$, where $q$ is a parameter in the range $(1/s, 1]$. When $1/s < q \leq 1$, among all the $s \times s$ stochastic matrices with trace $s\theta$, $U_s(q)$ is the one that has the least mutual information for the uniform input distribution (ref. [9, Theorem V.3]). The reduction described above, if exists, is called uniform reduction.

We give an example of uniform reduction with $s = 2$. Choose $R$ so that $RQ$ is an $s$-row matrix formed by $s$ linearly independent rows of $Q$. Let $a_{ij}$ be the $(i, j)$ entry of $RQ$, where $i = 1, 2$ and $1 \leq j \leq n$. Define an $n \times 2$ stochastic matrix $W = (w_{ij})$ as

$$w_{ij} = \begin{cases} \frac{a_{1j}}{a_{11} + a_{21}} & \text{if } a_{11} + a_{21} > 0, \\ 1 & \text{otherwise}, \end{cases} \quad (47)$$

and $w_{i2} = 1 - w_{i1}$, where $1 \leq i \leq n$. With the above $R$ and $W$, we see that $RQW = U_2(q)$, where $q = \sum_{k:a_{1k}+a_{2k}>0} \frac{a_{1k}}{a_{11} + a_{21}}$. The following lemma states a range of $q$ such that the reduction to $U_2(q)$ is feasible.

**Lemma 12:** For a stochastic matrix $Q$ such that $C(Q) > \epsilon$ for some $\epsilon > 0$, there exists a constant $B > 1/2$ depending only on $\epsilon$ such that $Q$ has a uniform reduction to $U_2(q)$ for all $1/2 < q \leq B$.

Fix any $\epsilon > 0$. Consider the line network formed by $Q_1, \ldots, Q_L$, where $C(Q) > \epsilon$ and hence $\text{rank}(Q_\ell) \geq 2$. We discuss a GBNC with $|A| = 2$ and $M = N = 1$.

By Lemma 12, there exists $q > 1/2$ such that for any $\ell$, there exists stochastic matrices $R_\ell$ and $S_\ell$ such that...
Define the recoding at the source node as $R_1$, and for $\ell = 1, \ldots, L - 1$, define the recoding at node $\ell$ as $S_{R_\ell+1}$. At the destination node, process all the received batches by $R_L$. The overall operation of a batch from the source node to the destination node is $W_1^L \triangleq (U_2(g))^L$. Applying the argument in [9, Theorem III.5], we get

$$
\log \left( \frac{1}{2^\rho - 1} \right) \leq \lim \inf_{L \to \infty} - \frac{1}{L} \log C(W'_L) \quad (48)
$$

$$
\leq \lim \sup_{L \to \infty} - \frac{1}{L} \log C(W'_L) \quad (49)
$$

$$
\leq 2 \log \left( \frac{1}{2^\rho - 1} \right). \quad (50)
$$

where $\frac{1}{2^\rho - 1}$ is the second largest eigenvalue of $U_2(g)$. Therefore, a channel code for the transition matrix $W_L$ as the outer code can achieve the rate $Q(e^{-CL})$ as $L \to \infty$, where the constant $c$ is between $\log \left( \frac{1}{2^\rho - 1} \right)$ and $2 \log \left( \frac{1}{2^\rho - 1} \right)$. The above discussion is summarized as the following theorem:

**Theorem 13:** Consider a sequence of DMCs $Q_\ell, \ell = 1, 2, \ldots$ with inf{$C(Q_\ell), \ell \geq 1$} > 0. For the line network of length $L$, the $i$th link is $Q_i$, the GBNC with $M = O(1)$ and $N = O(1)$ achieves rate $C_L(M, N) \geq c^I \cdot e^{-CL}$, where $c$ is a constant between $\log \left( \frac{1}{2^\rho - 1} \right)$ and $2 \log \left( \frac{1}{2^\rho - 1} \right)$, and $c^I > 0$ is a constant.

The technique used in the proof of Theorem 13 can be generalized for $M, N \geq 1$. We first show that for an $m \times n$ stochastic matrix $Q$ with rank($Q$) $\geq 2$, for any $2 \leq s \leq r$, the uniform reduction to $U_s(g)$ exists if $g$ is sufficiently close to 1/s. For an integer $2 \leq s \leq r$, let

$$
\kappa_s(Q) = \max_{n \times s \text{ stochastic matrix } R, s \times m \text{ stochastic matrix } W} \min \text{inv}(RQW) \quad (51)
$$

where mininv$(RQW)$ is the minimum value of $(RQW)^{-1}$ when $RQW$ is invertible and is $\infty$ otherwise. We give an example of $R$ and $W$ such that $RQW$ is invertible. Choose $R$ so that $RQ$ is an s-row matrix formed by s linearly independent rows of Q. Let $a_{ij}$ be the $(i, j)$ entry of $RQ$, where $1 \leq i \leq s$ and $1 \leq j \leq n$. To simplify the discussion, we assume all the columns of $R$ are non-zero. Define $W = D(RQ)^T$, where $D$ is an $n \times n$ diagonal matrix with the $(i, i)$ entry $1/\sum_{j} a_{ji}$. With the above $R$ and $W$, we see that $RQW$ is positive definite and hence invertible. Let $\rho_s(Q) = \min_{\kappa_s(Q)}-1/\min \text{inv}(\kappa_s(Q), 0)^{-1}$. We see that $\rho_s(Q) > 1/s$. The following lemma states a range of $g$ such that the reduction to $U_s(g)$ is feasible.

**Lemma 14:** Consider an $m \times n$ stochastic matrix $Q$ with rank $r \geq 2$. For any $2 \leq s \leq r$ and $1/s < g \leq \rho_s(Q)$, there exist an $s \times m$ stochastic matrix $R$ and an $n \times s$ stochastic matrix $S$ such that $RQS = U_s(g)$.

**Remark 2:** Lemma 12 is stronger than Lemma 14 for the case $s = 2$ as the former gives a uniform bound on $B$ that does not depend on $Q$ as long as $C(Q) > \epsilon$.

Consider a line network formed by $Q_1, \ldots, Q_L$, where $C(Q_L) > 0$ and hence rank($Q_L$) $\geq 2$. Let $r = \min_{L} \text{rank}(Q_L)$. Assuming $r \geq |A|$, we first discuss a recoding scheme with $M = N = 1$. Let $g = \min_{L} \rho_r(Q_L)$. By Lemma 14, there exists stochastic matrices $R_\ell$ and $S_\ell$ such that $R_\ell Q_\ell S_\ell = U_r(g)$. The following argument is similar as that of the proof of Theorem 13. Now we consider recoding with $M, N = O(1)$. Fix $M, N = O(1)$ and a finite alphabet $A$ such that $r^{MN} \geq |A|^M$. Regarding the line network $L$ as one formed by $Q_1^L, \ldots, Q_L^L$, we can apply the above GBNC with batch size 1, inner blocklength 1 and the batch alphabet $A^M$, which for the original line network $L$ of $Q_1, \ldots, Q_L$ is a GBNC with batch size $M$, inner blocklength $N$ and the batch alphabet $A$.

**V. LINE NETWORKS OF PACKET ERASURE CHANNELS**

For line networks of **packet erasure channels**, GBNC is also called batched network coding (BNC). In this section, we discuss line networks with identical packet erasure channels, for which, we demonstrate stronger converse and achievability results than the general ones.

Fix the alphabet $A$ with $|A| \geq 2$. Suppose that the input alphabet $Q_0$, and the output alphabet $Q_\infty$ are both $A \cup \{\epsilon\}$ where $\epsilon \notin A$ is called the erasure. For example, we may use a sequence of bits to represent a packet so that $A = \{0, 1\}^T$, i.e., each packet is a sequence of $T$ bits. Henceforth, a symbol in $A$ is also called a packet in this section. A packet erasure channel with erasure probability $\epsilon$ ($0 < \epsilon < 1$) has the transition matrix $Q_\text{era}$: for each $x \in A$, $Q_\text{era}(y|x) = 1 - \epsilon$ if $y = x$ and $Q_\text{era}(y|x) = \epsilon$ if $y = \epsilon$. The input $\epsilon$ can be used to model the input when the channel is not used for transmission and we define $Q_\text{era}(\epsilon|\epsilon) = 1$. When the input $\epsilon$ is not used for encoding information, erasure codes can achieve a rate of $1 - \epsilon$ symbols (in $A$) per use. It is also clear that $C_0(Q_\text{era}) = 0$.

**A. Upper Bound**

We obtain a refined upper bound by using a simpler channel function for packet erasure channels: The relation between the input $X$ and output $Y$ of a packet erasure channel can be written as a function

$$
Y = \alpha_\text{era}(X, Z) = \begin{cases} X & \text{if } Z \neq \epsilon, \\
\epsilon & \text{if } Z = \epsilon,
\end{cases} \quad (52)
$$

where $Z$ is a discrete random variable independent of $X$ with $P(Z = \epsilon) = \epsilon$. In other words, $Z$ indicates whether the channel output is the erasure or not.

For a line network of length $L$ with a GBNC of inner blocklength $N$, the bottleneck status can be defined as

$$
E_0 = \{ \forall_{i=1}^L (Z_\ell[i] = \epsilon, i = 1, \ldots, N) \}, \quad (53)
$$

where $Z_\ell[i]$ is the channel variable of the $i$th use of $Q_\ell$. With this bottleneck status, $p_1 = P(E_0) = (1 - \epsilon^L)^N$. Following a similar procedure as in the proof of Theorem 4, we have

$$
C_L(M, N) \leq \frac{(1 - \epsilon^L)^N}{N} \min \{ M \log |A|, N \log |Q_0| \}, \quad (54)
$$

which is a tighter upper bound than (19).
B. Achievability by Random Linear Recoding

We now introduce a class of inner codes with batch size \( M = O(1) \), which provides the achievability counterpart for the cases 1) and 2) in Theorem 4. Let \( \mathbb{F}_q \) be the finite field of \( q \) symbols, and let \( T > 0 \) be an integer. Suppose \( \mathcal{A} = \mathbb{F}_q^T \), i.e., each packet is a sequence of \( T \) symbols from the finite field \( \mathbb{F}_q \). The outer code generates batches that consist of \( M \) packets in \( \mathcal{A} \), and can be represented as a \( T \times M \) matrix over \( \mathbb{F}_q \). In each packet generated by the outer code, the first \( M \) symbols in \( \mathbb{F}_q \) are called the coefficient vector. A batch \( X \) has the first \( M \) rows of the matrix formed by the coefficient vector. Denote \( F \) as a packet in the batches. In other words, when a packet is erased, an intermediate node assumes 0 is received.

The inner code is formed by random linear recoding, which have been studied in random linear network coding (RLNC). A random linear combination of vectors in \( \mathcal{A} \) has the linear combination coefficients chosen uniformly at random from \( \mathbb{F}_q \). The inner code includes the following operations:
- The source node generates \( N \) packets for a batch using random linear combinations of the \( M \) packets of the batch generated by the outer code.
- Each intermediate node generates \( N \) packets for a batch using random linear combinations of all packets of the received packets of the batch.

Note that for each batch, only the packets with linearly independent coefficient vectors are needed for random linear recoding. Therefore, the buffer size used to store batch content is \( O(MT \log q) \) bits. Also, the computational cost of the above recoding scheme for each intermediate node is \( O(N^2T \log q) \) per batch.

At each node, the rank of the coefficient matrix of a batch (i.e., the first \( M \) rows of the matrix formed by the generated/received packets of the batch) is also called the rank of the batch. At each node, the ranks of all the batches follow an identical and independent distribution. Denote by \( \pi_\ell \) the rank distribution of a batch at node \( \ell \). As all the batches at the source node have rank \( M \), we know that \( \pi_0 = (0, 0, \ldots, 0, 1) \). Moreover, the rank distributions \( \pi_0, \pi_1, \ldots, \pi_L \) form a Markov chain so that for \( \ell = 1, \ldots, L \), it holds that

\[
\pi_\ell = \pi_{\ell-1} P
\]

where \( P \) is the transition matrix characterized in [25, Lemma 4.2].

The maximum achievable rate of this class of BNC is

\[
(1 - \frac{MT}{T}) E[|\pi_L|] N \log |\mathcal{A}|. \tag{56}
\]

In Fig. 6, we compare numerically the upper bound and the achievable rates of BNC by evaluating (54) and (56), respectively. Throughout the experiment, we specify parameters \( \epsilon = 0.2, q = 256 \) and \( T = 1024 \) following the same setup as in [14, Fig. 10], which are decided based on the following considerations: Firstly, in many practical wireless communication systems, a packet loss rate of around 10 to 20 percent is commonly observed. Secondly, a finite field of size 256 is frequently utilized in real-world implementations. Lastly, a packet of 1024 bytes is a typical choice in internet-based communication scenarios. Note that each packet has 8192 bits and the min-cut is 65536 bits per use.

First, we consider fixed \( M = N = 2, 3, 4 \), and plot the calculation for \( L \) up to 1000 in Fig. 6(a). We see from the figure that for a fixed \( N \), the achievable rates of BNC and the upper bound in (54) share the same exponential decreasing trend.

Second, we consider fixed \( M = 2, 4, 8, 16, 32 \). For each value of \( M \), we find the optimal value of \( N \), denoted by \( N^*_L \), that maximizes \( \text{BNC}_{L}(M, N) \). We see from Fig. 6(b) that \( N^*_L \) demonstrates a low increasing rate with \( L \). We further illustrate \( \text{BNC}_{L}(M, N^*_L) \) and \( \text{PEC}_{UB}(M, N^*_L) \) for each value of \( M \) in Fig. 6(c).

The following theorem justifies the scalability of \( \text{BNC}_{L}(M, N) \) when \( L \) is large, where the \( M = 1 \) case was proved in [25].

**Theorem 15:** Consider a line network of \( L \) packet erasure channels with erasure probability \( \epsilon \). For GBNCs of batch size \( M < T \) and inner blocklength \( N \) using random linear recoding,

\[
\text{BNC}_{L}(M, N) = \Theta \left( \frac{(1 - (1 - \epsilon + (1 - \epsilon)/q)^N)^L}{N} \right). \tag{57}
\]

When \( q \) is relatively large, \( \text{BNC}_{L}(M, N) \) has nearly the same scalability as \( \text{PEC}_{UB}(M, N) \), as illustrated by Fig. 6(c). Consider two cases of \( N \) for the scalability of \( \text{BNC}_{L}(M, N) \): When \( N \) is a fixed number, \( \text{BNC}_{L}(M, N) \) decreases exponentially with \( L \). When \( M \) is a fixed number and \( N \) is unconstrained, based on the optimization theory (see, e.g., [17, Lemma 1]) we know that \( \max_{N} \text{BNC}_{L}(M, N) = \Theta(1/\ln L) \), and the maximum is achieved by \( N = \Theta(\ln L) \).

VI. LINE NETWORKS WITH CHANNELS OF POSITIVE ZERO-ERROR CAPACITY

Last, we discuss how to extend our study so far to line networks of channels that have positive capacity but may also have positive zero-error capacity. Denote by \( \mathcal{L} \) a line network of length \( L \) formed by channels \( Q_1, \ldots, Q_L \), where it is not necessary that \( C_0(Q_\ell) = 0 \). For a GBNC on \( \mathcal{L} \), the end-to-end transition matrix of a batch is denoted by \( W_{\mathcal{L}} \). Denote the maximum achievable rate of all recoding schemes with batch size \( M \) and inner blocklength \( N \) for \( \mathcal{L} \) as \( C_{\mathcal{L}}(M, N) \). Let \( L_0 \) be the number of channels in \( \mathcal{L} \) with 0 zero-error capacity, i.e., \( L_0 = |\{1 \leq \ell \leq L : C_0(Q_\ell) = 0\}| \). In the following, we argue that \( C_{\mathcal{L}}(M, N) \) scales like a line network of length \( L_0 \) formed by channels with 0 zero-error capacity.

Let \( \{l_1, \ldots, l_{L_0}\} = \{1 \leq \ell \leq L : C_0(Q_\ell) = 0\} \) where \( l_1 < l_2 \cdots < l_{L_0} \). Denote by \( \mathcal{L}' \) the line network formed by the concatenation of \( Q_{l_1}, \ldots, Q_{l_{L_0}} \). For any given GBNC on \( \mathcal{L} \), we can find proper recoding operations for the GBNC on \( \mathcal{L}' \) so that \( W_{\mathcal{L}} = W_{\mathcal{L}'} \), and hence \( C_{\mathcal{L}}(M, N) \leq C_{\mathcal{L}'}(M, N) \). For network \( \mathcal{L}' \), [III] provides the upper bounds on the achievable
rates as functions of length $L_0$ under certain coding parameter sets, which are also upper bounds for network $L$.

We derive a lower bound of achievable rates of $L$ using the uniform reduction approach introduced in §IV-C. Suppose $C = \inf \{C(Q_i) : i \geq 1\} > 0$. By Lemma 12, there exists a constant $B \in (1/2, 1)$ depending only on $C$ such that there exist stochastic matrices $R_i$ and $S_i$ with $R_i Q_i S_i = U_2(B)$ for all $i$. For $Q_i$ with $C(Q_i) > 0$, we can find $R_i$ and $S_i$ so that $R_i Q_i S_i$ equals the identity matrix $I_2$. The existence of $R_i$ and $S_i$ is guaranteed by the following lemma.

**Lemma 16:** For an $m \times n$ stochastic matrix $Q$ with $C_0(Q) > 0$, there exists a $2 \times m$ stochastic matrix $R$ and a $n \times 2$ stochastic matrix $S$ such that $SR = I_2$, the $2 \times 2$ identity matrix.

**Proof:** For a DMC $Q$, two channel inputs $x_1$ and $x_2$ are said to be adjacent if there exists an output $y$ such that $Q(y|x_1)Q(y|x_2) > 0$. Denote by $M_0(Q)$ the largest number of inputs in which adjacent pairs do not exist. For a DMC with $C_0 > 0$, we have that $M_0(Q) \geq 2$ and then $C_0(Q) \geq 1$, since otherwise it is easy to verify $M_0(Q^{2Q}) \leq 1$ for any $n$ which leads to $C_0(Q) = 0$.

When the channel $Q$ satisfies $C_0(Q) > 0$, we have $M_0(Q) \geq 2$. Define $R$ as a two-row deterministic stochastic matrix that selects two rows of $Q$ that correspond to two non-adjacent inputs. Denote by $a_{ij}$ the $(i, j)$ entry of $RQ$. We have $a_{1j}a_{2j} = 0$ for all $j = 1, \ldots, n$. Let $S$ be defined same as the matrix $W$ in defined in (47).

Denote by $L'$ the line network formed by the concatenation of $L_0$ identical channels $U_2(B)$. Hence, we obtain $C_L(M, N) \geq C_{L'}(M, N)$, where the later can be lower bounded by the techniques in §IV. In particular, the error exponent condition in Theorem 11 can be verified by checking the proof of Lemma 10 for the special case of BSCs.

**VII. CONCLUDING REMARKS**

This paper examines the achievable rates of generalized batched network codes (GBNCs) in line networks with general discrete memoryless channels (DMCs). The findings suggest that capacity-achieving codes for DMCs may not be the only consideration for the inner code. Simple codes like repetition and convolutional codes can achieve the same rate order while requiring lower buffer sizes. Additionally, reliable hop-by-hop communication is not always optimal when buffer size and latency constraints are present.

Feedback is useful in certain communication scenarios. Hop-by-hop feedback does not increase the network capacity (the min-cut). However, exploring its potential benefits is an intriguing area of research in the context of GBNC. Hop-by-hop feedback within batches does not increase the upper bound since it does not increase the capacity of a DMC. However, when hop-by-hop feedback crosses batches, it introduces memory in the batched channel, which may increase the capacity. Additionally, feedback can also simplify coding schemes.

Future research directions also include investigating better upper bounds and recoding schemes for line networks with special channels like BSCs, generalizing the analysis to channels with infinite alphabets and continuous channels, and exploring whether the upper bound holds for more general codes beyond GBNCs would be valuable.

**REFERENCES**

[1] W. Zhang, M. Stojanovic, and U. Mitra, “Analysis of a linear multihop underwater acoustic network,” *IEEE J. Ocean. Eng.*, vol. 35, no. 4, pp. 961–970, Oct. 2010.
[2] X. Tang, Z. Wang, Z. Xu, and Z. Ghassenmooly, “Multihop free-space optical communications over turbulence channels with pointing errors using heterodyne detection,” J. Lightw. Technol., vol. 32, no. 15, pp. 2597–2604, Aug. 15, 2014.

[3] Z. Huakai, D. Guangliang, and L. Haitao, “Simplified BATS codes for deep space multihop networks,” in Proc. IEEE Inf. Theory, Netw., Electron. Autom. Control Conf., May 2016, pp. 311–314.

[4] H. Hariadna, K. Mizutani, J. Fujiwara, K. Mochizuki, K. Obata, and R. Okumura, “IEEE 802.15.4g based Wi-SUN communication systems,” IEICE Trans. Commun., vol. 7, pp. 1032–1043, Jul. 2017.

[5] P. Bhattacharjee and S. M. Zafaruddin, “On the performance of multihop THz wireless system over mixed channel fading with shadowing and antenna misalignment,” IEEE Trans. Commun., vol. 70, no. 11, pp. 7703–7763, Nov. 2022.

[6] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[7] A. M. Bedewy, Y. Sun, and N. B. Shroff, “The age of information in multihop networks,” IEEE/ACM Trans. Netw., vol. 27, no. 3, pp. 1248–1257, Jun. 2019.

[8] S. Farazi, A. G. Klein, and D. R. Brown, “Fundamental bounds on the age of information in multi-hop global status update networks,” J. Commun. Netw., vol. 21, no. 3, pp. 268–279, Jun. 2019.

[9] U. Niesen, C. Fragouli, and D. Tannetti, “On capacity of line networks,” IEEE Trans. Inf. Theory, vol. 53, no. 11, pp. 4039–4058, Nov. 2007.

[10] A. E. Gammal and Y.-H. Kim, Network Information Theory. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[11] D. Silva, W. Zeng, and F. R. Kschischang, “Sparse network coding with overlapping classes,” in Proc. Workshop Netw. Coding, Theory, Appl., Jun. 2009, pp. 74–79.

[12] A. Heidarzadeh and A. H. Banihashemi, “Overlapped chunked network coding,” in Proc. IEEE/ACM Trans. Netw. Inf. Theory Workshop Inf. Theory (ITW, Cairo), Jan. 2010, pp. 1–5.

[13] Y. Li, E. Soljanin, and P. Spasojevic, “Effects of the generation size and overlap on throughput and complexity in randomized linear network coding,” IEEE Trans. Inf. Theory, vol. 57, no. 2, pp. 1111–1123, Feb. 2011.

[14] S. Yang and R. W. Yeung, “Batched sparse codes,” IEEE Trans. Inf. Theory, vol. 60, no. 9, pp. 5322–5346, Sep. 2014.

[15] B. Tang, S. Yang, B. Ye, Y. Yin, and S. Lu, “Expander chunked codes,” EURASIP J. Adv. Signal Process., vol. 2015, no. 1, pp. 1–13, Dec. 2015.

[16] B. Tang and S. Yang, “An LDPC approach for chunked network codes,” IEEE/ACM Trans. Netw., vol. 26, no. 1, pp. 605–617, Feb. 2018.

[17] S. Yang, J. Wang, Y. Dong, and Y. Zhang, “On the capacity scalability of line networks with buffer size constraints,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jul. 2019, pp. 1507–1511.

[18] J. Wang, S. Yang, Y. Dong, and Y. Zhang, “On achievable rates of line networks with generalized batched network coding,” 2021, arXiv:2105.07669.

[19] S. Yang and R. W. Yeung, BATS Codes: Theory and Practice (Synthesis Lectures on Communication Networks). San Rafael, CA, USA: Morgan & Claypool, 2017.

[20] S. Yang, J. Meng, and E.-H. Yang, “Coding for linear operator channels over finite fields,” in Proc. IEEE Int. Symp. Inf. Theory, Jun. 2010, pp. 2413–2417.

[21] Z. Zhou, C. Li, S. Yang, and X. Guang, “Practical inner codes for BATS codes in multi-hop wireless networks,” IEEE Trans. Veh. Technol., vol. 68, no. 3, pp. 2751–2762, Mar. 2019.

Jie Wang (Graduate Student Member, IEEE) received the B.S. degree from The Chinese University of Hong Kong, Shenzhen, in 2020. He is currently pursuing the Ph.D. degree with the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. His research interests include optimization, statistics, and information theory. He received the Best Theoretical Paper Award in 2023 INFORMS Workshop on Data Mining and Data Analytics, the ISyE Robert Goodell Brown Research Excellence Prize in 2022, and the Best Poster Award in 2022 INFORMS Annual Meeting.

Shenghao Yang (Member, IEEE) received the B.S. degree from Nankai University in 2001, the M.S. degree from Peking University in 2004, and the Ph.D. degree in information engineering from The Chinese University of Hong Kong in 2008. He was a Visiting Student with the Department of Informatics, University of Bergen, Norway, in Spring 2017. He was a Post-Doctoral Fellow with the University of Waterloo from 2008 to 2009 and the Institute of Network Coding, The Chinese University of Hong Kong, from 2010 to 2012. He was with the Institute for Interdisciplinary Information Sciences, Tsinghua University, from 2012 to 2015. He is currently an Associate Professor with The Chinese University of Hong Kong, Shenzhen. His research interests include network coding, information theory, coding theory, and quantum information.

Yanyan Dong received the B.S. degree from Jilin University in 2017 and the Ph.D. degree from The Chinese University of Hong Kong, Shenzhen, in 2022. She was a Research Fellow with the National University of Singapore from December 2022 to December 2023. Her research interests include information theory, coding theory, network coding, and machine learning.

Yiheng Zhang received the B.S. degree from The Chinese University of Hong Kong, Shenzhen, in 2021. He is currently pursuing the master's degree with Carnegie Mellon University. His current research interests include mathematical theorem proving with machine learning.