

# Buffer-aided Link Selection with Network-coding in Multi-hop Networks

Zhao Tian, *Student Member, IEEE*, Yu Gong, Gaojie Chen, Zhi Chen, *Member, IEEE*  
and Jonathon Chambers, *Fellow, IEEE*

**Abstract**—This paper proposes a novel buffer-aided link selection scheme based on network-coding in the multiple hop relay network. Compared with existing approaches, the proposed scheme significantly increases the system throughput. This is achieved by applying data buffers at the relays to decrease the outage probability and using network-coding to increase the data rate. The closed-form expressions of both the average throughput and packet delay are successfully derived. The proposed scheme has not only significantly higher throughput than both the traditional and existing buffer-aided max-link scheme, but also smaller average packet delay than the max-link scheme, making it an attractive scheme in practice.

**Index Terms**—Multi-hop relay, link selection, buffer-aided relay, network coding

## I. INTRODUCTION

Relay network has been well investigated as an attractive scheme in wireless communications [1]. Current research mainly focuses on the 2-hop relay network that every data packet takes two hops to transmit from the source to destination through a relay node [2]–[4]. Relatively less has been studied for relay networks with more than two hops. The multi-hop relay network can be seen in many scenarios. A typical example is the device-to-device (D2D) communications in the cellular system, where some mobile users may directly communicate with each other (D2D communications) rather than through the base station (cellular communications) [5], [6]. Because the transmission powers for the D2D mobile users are usually strictly limited to avoid interfering the base station, multi-hop transmission can be required for D2D communications [7].

1) *Related work*: Conventionally the multi-hop links are consecutively selected for data transmission. Recent research shows that applying data buffers at the relays significantly improves the transmission performance. Due to the data buffers at the relays, when a data packet arrives at a node, it may not be immediately forwarded to the next node. Instead, other links with better signal-to-noise-ratio (SNR) may be selected for

data transmission. This so called adaptive link selection ([8], [9]) is particularly useful in the D2D cellular system, because the base station often has the knowledge of the channel state information (CSI) of the D2D links to coordinate the interference between the cellular and D2D communications. As a result, the base station can always select the best link for data transmission, rather than following the conventional hopping sequence.

Buffer-aided relay has attracted much attention recently. Beside the aforementioned adaptive link selection, buffer-aided relays have also been used in applications including relay selection [10]–[13], cognitive radio networks [14] and physical layer network security [15]. Of particular interest is the max-link relay selection scheme due to its excellent outage performance [10]. In the max-link relay selection, at any time slot, the link with the strongest channel SNR among all possible source-to-relay and relay-to-destination links is always selected for data transmission, leading to the diversity order of  $2N$  if the buffer size is large enough (where  $N$  is the number of relays). The max-link scheme can be straightforwardly used in the multi-hop link selection, simply by selecting the link with the highest SNR among all possible multi-hop links at any time slot.

Of particularly interest is the average throughput of the multi-hop relay network which is given by

$$\bar{\eta} = \bar{R} \cdot (1 - P_{out}), \quad (1)$$

where  $P_{out}$  is the outage probability of the system and  $\bar{R}$  is the average data rate (without considering the outage). It is known that the max-link scheme significantly reduces the outage probability. However, the max-link scheme still has the same data rate as the conventional scheme, because in both schemes only one link is selected for data transmission at any time slot. On the other hand, since the outage probability tends to be zero when the SNR goes to infinity, it is clear from (1) that the average throughput mainly depends on the average data rate at the high SNR range. This implies that the max-link selection scheme mainly improves the system throughput at the low SNR range.

On the other hand, it is well known that the physical layer network coding can be used to increase the data rate of the two-way relay network, where two source nodes exchange data packets through a single relay node [16]–[18]. To be specific, in the physical layer network coding scheme, the two sources can transmit packets to, or receive packets from, the relay node simultaneously. Thus the data rate can reach 1 packet per

Z. Tian and Y. Gong are with the Advanced Signal Processing Group, Loughborough University, Loughborough, Leicestershire, UK, Emails: {z.tian, y.gong}@lboro.ac.uk.

G. Chen is the corresponding author with Department of Engineering Science, University of Oxford, Oxford, UK, E-mails: gaojie.chen@eng.ox.ac.uk.

Z. Chen is with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China, E-mail: chenzhi@uestc.edu.cn.

J. A. Chambers is with the Communications, Sensors, Signal and Information Processing Group, Newcastle University, Newcastle Upon Tyne, UK, E-mails: jonathon.chambers@ncl.ac.uk.

time slot, rather than  $1/2$  in the conventional approach. This encourages us to apply the physical layer network coding in the multi-hop relay selection to increase the data rate. This can be achieved by simultaneously selecting two or more links for data transmission. As a result, the throughput of the multi-hop network at the high SNR range can be improved.

2) *Contribution*: In this paper, we propose a novel multi-hop link selection scheme which seamlessly integrates the max-link selection and physical layer network coding so that the average throughput is significantly improved at both low and high SNR ranges. The main contributions of this paper are listed as follows:

- *Proposing a novel buffer-aided network-coding link selection scheme for the multi-hop relay network*. The proposed scheme has significantly higher throughput than existing buffer-aided max-link scheme.
- *Describing a new analysis tool to obtain the average throughput of the proposed scheme*. Both the outage probability and average data rate are successfully derived to obtain the average throughput of the proposed scheme.
  - First, the outage probability analysis is based on the Markov chain of the buffer states, which is much more difficult than those in existing approaches (e.g. [10]) due to the complicated link selection rules. Particularly, we describe a trellis diagram to derive the transition probabilities between buffer states, based on which the outage probability is obtained.
  - Secondly, in the proposed multi-hop scheme, due to the simultaneous link transmission, the calculation of the average data rate is far from straightforward. In this paper, a trellis diagram is described to successfully obtain the average data rate. The analysis not only shows deep insight in understanding the multi-hop relay network, but also provides guidance in analyzing similar systems.
- *Deriving the closed-form expression of the average packet delay of the proposed scheme*. The average packet delay is an important issue in buffer-aided schemes. The analysis shows that the proposed scheme not only has larger throughput, but also shorter packet delay, than the max-link scheme, making it an attractive scheme in practice.

The remainder of the paper is organized as follows: Section II shows the system model of the  $N$ -hop relay network; Section III proposes the buffer-aided network-coding link selection scheme; Section IV and V analyze the outage probability and average data rate of the proposed scheme respectively; Section VI analyzes the average packet delay; Section VII shows simulation results to verify the proposed scheme; finally Section VIII concludes the paper.

## II. $N$ -HOP RELAY NETWORK

The system model of the  $N$ -hop relay network is shown in Fig. 1, where there are one source node ( $S$ ), one destination node ( $D$ ) and  $(N-1)$  number of relay nodes ( $R_1, \dots, R_{N-1}$ ). We assume that there are no direct links between two nodes separated by two hops or more, and all relays apply the

decode-and-forward (DF) protocol and operate in the half-duplex mode.



Fig. 1. The system model of the  $N$ -hop relay network.

For later use, the hopping links are consecutively named as  $link_1, link_2, \dots, link_N$  respectively, as is shown in Fig. 1. The channel coefficient and gain for  $link_i$  at time slot  $t$  is denoted as  $h_i(t)$  and  $\gamma_i(t) = |h_i(t)|^2$  respectively. We assume that all channel links are independent and identically distributed (i.i.d.) Rayleigh fading<sup>1</sup>, so that the channel gains  $\gamma_i(t)$  are exponentially distributed with the same average gain as  $\bar{\gamma} = E|h_i(t)|^2$  for all  $i = 1, \dots, N$ . We also assume without losing generality that transmission powers and all noise variances are normalized to unity.

The most straightforward way to transmit packets through the  $N$ -hop network is to let  $link_1, link_2, \dots, link_N$  be consecutively used for data transmission. This so-called ‘consecutive-hopping’ scheme is used as a baseline to compare with other schemes in the paper. If the transmission rates at all nodes are the same as  $r_t$ , the average data rate of the consecutive-hopping scheme is given by

$$\bar{R}^{(\text{con-hopping})} = \frac{1}{N} \cdot r_t. \quad (2)$$

In this paper, we assume the channels are quasi-static so that the coefficients remain unchanged during one hop interval but independently vary from one hop to another. We also assume that, when a link becomes outage, the packet will be re-sent by the transmission node corresponding to the link (rather than the source node  $S$ ). For Rayleigh fading channels, the probability that  $link_i$  becomes outage is given by

$$P_{out,i} = P(C_i < r_t) = 1 - e^{-\frac{\Delta}{\bar{\gamma}}} \quad (3)$$

where  $\Delta = 2^{r_t} - 1$ ,  $C_i = \log(1 + \gamma_i)$  which is the instantaneous capacity for  $link_i$ . Or the probability that a packet takes  $k$  time slots to successfully pass  $link_i$  is  $(P_{out,i})^{k-1}(1 - P_{out,i})$ . Thus the average number of time slots for a packet passing through  $link_i$  is given by

$$\begin{aligned} T_i &= \sum_{k=1}^{\infty} k \cdot (P_{out,i})^{k-1}(1 - P_{out,i}) \\ &= \frac{1}{1 - P_{out,i}} = \frac{1}{e^{-\frac{\Delta}{\bar{\gamma}}}} \end{aligned} \quad (4)$$

Because all channels are i.i.d., the average number of slots for a packet passing through the overall  $N$ -hop network is  $N \cdot T_i$ . Then the average throughput for the consecutive hopping scheme is obtained as

$$\eta^{(\text{con-hopping})} = \frac{r_t}{N \cdot T_i} = \frac{r_t}{N} \cdot e^{-\frac{\Delta}{\bar{\gamma}}}. \quad (5)$$

Comparing (1), (2) and (5), we can have the outage proba-

<sup>1</sup>While the analysis in this paper is based on the i.i.d. channel assumption, it can be generalized to the case that every link has different average channel gain.

bility as

$$P_{out}^{(\text{con-hopping})} = 1 - e^{-\frac{\Delta}{\gamma}}. \quad (6)$$

### III. BUFFER-AIDED LINK SELECTION BASED ON NETWORK-CODING

In this section, we will first apply the buffers at the relays to reduce the outage probability and use the physical layer network coding to increase the data rate. We then propose a novel link selection scheme for the multi-hop relay network by integrating the buffer-aided and network coding approaches.

#### A. Decrease the outage probability with buffers at the relays

The max-link relay selection scheme described in [10] can be straightforwardly applied in the multi-hop link selection. To be specific, in the buffer-aided link selection, every relay is equipped with a data buffer of the size  $L$ . We assume that the relay  $R_i$  has buffer  $Q_i$ , where  $i = 1, \dots, N-1$ . At any time slot, when a data packet arrives at a relay node, it is stored in the buffer. At the next time slot, unlike the traditional scheme, the stored data packet is not necessarily forwarded to the next node. Instead the link with the highest SNR among all of the “available” links is selected for data transmission. A link is considered available if the buffers of the corresponding transmitting and receiving nodes are not empty and full respectively. Thus in the max-link scheme, the link for data transmission is selected as

$$\text{link} = \arg \max_{\text{link}_i \in \mathcal{A}} \{\gamma_i\}, \quad (7)$$

where  $\mathcal{A}$  is the set containing all available links, and recall that  $\gamma_i$  is the instantaneous channel SNR for  $\text{link}_i$ . Without losing generality, we assume that the source  $S$  always have data to transmit and the buffer size is in the unit of “packet”. Because one packet is transmitted at one time slot at fixed rate, if an “available” link is selected, there must be a packet available for transmission and the buffer at the receiving node is ‘large’ enough to store the packet.

In the max-link scheme, because only one link is selected for data transmission at any time, the average data rate is still the same as that in the traditional scheme which is given by

$$\bar{R}^{(\text{max-link})} = \frac{1}{N} \cdot r_t, \quad (8)$$

Then the average throughput of the max-link scheme is given by

$$\bar{\eta}^{(\text{max-link})} = \frac{1}{N} \cdot \left(1 - P_{out}^{(\text{max-link})}\right) \cdot r_t, \quad (9)$$

where  $P_{out}^{(\text{max-link})}$  is the outage probability of the max-link scheme which can be obtained by following similar analysis as those in [10].

Because  $P_{out}^{(\text{max-link})} < P_{out}^{(\text{con-hopping})}$ , the throughput of the buffer-aided max-link scheme is higher than that of the traditional scheme. On the other hand, because the max-link and traditional schemes have the same data rate, and further noting that  $P_{out}^{(\text{max-link})} \rightarrow 0$  when  $\text{SNR} \rightarrow \infty$ , the two schemes have similar throughput when the SNR is high enough. This indicates that the buffer-aided link selection mainly improves the throughput at the low SNR range.

#### B. Increase the data rate with network coding

We suppose at one time slot, all odd numbered links transmit data at the same time, and at the next time slot all even numbered links transmit simultaneously. Thus a relay node may receive data from both the previous and next nodes. Without losing generality, at time  $t$ , we assume that node  $R_i$  receives data from its previous node  $a$  and next node  $b$  simultaneously. Then the received signal at relay  $R_i$  at time slot  $t$  is given by

$$y_i(t) = h_i(t) \cdot x_a + h_{b,i}(t) \cdot x_b + n_i(t), \quad (10)$$

where  $x_a$  and  $x_b$  are the data packets transmitted from nodes  $a$  and  $b$  respectively,  $h_{b,i}(t)$  is the channel coefficient for the  $b \rightarrow R_i$  link, and  $n_i(t)$  is the noise at node  $R_i$ .

It is clear from (10) that  $h_{b,i}(t) \cdot x_b$  forms the inter-relay interference. Because  $x_b$  is transmitted from  $R_i$  to node  $b$  previously, it can be stored at  $R_i$ . With the principle of physical layer network coding ([18]), the inter-relay interference can be completely removed from (10), so that the received signal at  $R_i$  becomes

$$y_i(t) = h_i(t) \cdot x_a + n_i(t). \quad (11)$$

Therefore, with the physical layer network coding, all odd (or even) numbered links can be used for data transmission simultaneously without causing any inter-relay interference. As an example, the network coding based transmission scheme for the 4-hop relay network is shown in Fig. 2.

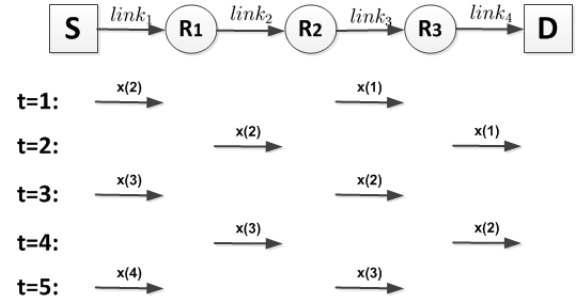


Fig. 2. Network-coding based 4-hop relay transmission.

As is illustrated in Fig. 2, on average it only takes two hops to transmit one data packet from  $S$  to  $D$ , no matter how many hops there are in the relay network. Thus the data rate for the network-coding based scheme is given by

$$\bar{R}^{(\text{net-coding})} = \frac{1}{2} \cdot r_t. \quad (12)$$

On the other hand, when the odd-numbered links are used for data transmission, the outage occurs when  $\min_{i \in \text{odd}} \{C_i\} < r_t$ . Similar to (4), the average number of time slots for a packet passing through odd-numbered links can be obtained as

$$T_{\text{odd}} = \frac{1}{1 - P_{out, \text{odd}}}, \quad (13)$$

where  $P_{out, \text{odd}} = P\left(\min_{i \in \text{odd}} \{C_i\} < r_t\right) = \left(1 - e^{-\frac{N_o \cdot \Delta}{\gamma}}\right)$ ,  $N_o = \lceil N/2 \rceil$  which is the number of odd-numbered links and  $\lceil \cdot \rceil$  rounds up the embraced value to the nearest integer.

Similarly, the average number of time slots for a packet passing through even-numbered links can be obtained as

$$T_{\text{even}} = \frac{1}{1 - P_{\text{out,even}}}, \quad (14)$$

where  $P_{\text{out,even}} = P\left(\min_{i \in \text{even}} \{C_i\} < r_t\right) = \left(1 - e^{-\frac{N_e \cdot \Delta}{\gamma}}\right)$ ,  $N_e = \lfloor N/2 \rfloor$  which is the number of odd-numbered links and  $\lfloor \cdot \rfloor$  rounds down the embraced value to the nearest integer. Thus the average throughput of the network-coding based scheme can be obtained as

$$\begin{aligned} \bar{\eta}^{(\text{net-coding})} &= \frac{r_t}{T_{\text{odd}} + T_{\text{even}}} \\ &= r_t \cdot \frac{(1 - P_{\text{out,odd}})(1 - P_{\text{out,even}})}{(1 - P_{\text{out,odd}}) + (1 - P_{\text{out,even}})} \quad (15) \\ &= r_t \cdot \frac{e^{-\frac{N_o \cdot \Delta}{\gamma}} e^{-\frac{N_e \cdot \Delta}{\gamma}}}{e^{-\frac{N_o \cdot \Delta}{\gamma}} + e^{-\frac{N_e \cdot \Delta}{\gamma}}} \end{aligned}$$

Comparing (1), (12) and (15), we can have the outage probability for the network-coding based scheme as

$$P_{\text{out}}^{(\text{net-coding})} = \frac{e^{-\frac{N_o \cdot \Delta}{\gamma}} + e^{-\frac{N_e \cdot \Delta}{\gamma}} - 2e^{-\frac{N_o \cdot \Delta}{\gamma}} e^{-\frac{N_e \cdot \Delta}{\gamma}}}{e^{-\frac{N_o \cdot \Delta}{\gamma}} + e^{-\frac{N_e \cdot \Delta}{\gamma}}}. \quad (16)$$

From (16), and noting that either  $N_o = N_e$  or  $N_o = N_e + 1$ ,  $P_{\text{out}}^{(\text{net-coding})}$  is bounded as

$$1 - e^{-\frac{N_e \cdot \Delta}{\gamma}} \leq P_{\text{out}}^{(\text{net-coding})} \leq 1 - e^{-\frac{N_o \cdot \Delta}{\gamma}} \quad (17)$$

Comparing (6) and (17) clearly shows that both upper and lower bounds of  $P_{\text{out}}^{(\text{net-coding})}$  are larger than  $P_{\text{out}}^{(\text{con-hopping})}$  so that

$$P_{\text{out}}^{(\text{net-coding})} \geq P_{\text{out}}^{(\text{con-hopping})} \quad (18)$$

Therefore, while the network-coding scheme has higher data rate than the consecutive-hopping scheme, its outage performance is however worse than the latter. To be more specific, because the outage probability  $P_{\text{out}} \rightarrow 0$  when the SNR  $\rightarrow \infty$ , the average throughput is mainly determined by the data rate when the SNR is large enough. Thus at the high SNR range, the average throughput of the  $N$ -hop network with the network-coding scheme is always about  $r_t/2$ . On the other hand, when the SNR  $\rightarrow -\infty$ , the outage probability  $P_{\text{out}} \rightarrow 1$  so that the throughput is more determined by the outage probability than by the data rate. This implies that, when the SNR is very small, the network-coding based scheme has lower throughput than the traditional scheme. Therefore, the network-coding scheme improves the throughput at the high SNR range.

### C. Buffer-aided network-coding link selection

In order to increase the average throughput over all SNR ranges, we propose a novel link selection scheme by integrating the buffer-aided max-link and network-coding approaches. This is achieved by adding simultaneous link transmission in the buffer-aided link selection rules.

Generalizing from the network-coding scheme, we understand that any links separated by two hops or more can be

simultaneously selected for data transmission. We denote  $N_s$  as the number of simultaneously transmitting links at one time slot. For the  $N$ -hop relay network, we have

$$1 \leq N_s \leq \lceil N/2 \rceil \quad (19)$$

For any  $N_s$ , there exist  $D(N_s)$  possible link selections, which is represented by the selection vector as

$$\mathbf{link}^{(N_s)} = [\text{link}^{(N_s)}(1), \dots, \text{link}^{(N_s)}(D(N_s))], \quad (20)$$

where  $\text{link}^{(N_s)}(i)$  is the  $i$ th link selection for  $N_s$  simultaneous link transmission. For later use, we denote  $\text{link}_{i_1+\dots+i_n}$  as the simultaneous transmission of  $\text{link}_{i_1}, \dots, \text{link}_{i_n}$ .

For example, in the 4-hop relay network, we have  $1 \leq N_s \leq 2$ , and

$$\begin{aligned} \mathbf{link}^{(N_s=1)} &= [\text{link}_1, \text{link}_2, \text{link}_3, \text{link}_4] \\ \mathbf{link}^{(N_s=2)} &= [\text{link}_{1+3}, \text{link}_{1+4}, \text{link}_{2+4}] \end{aligned} \quad (21)$$

The principle of the proposed scheme is to let as many links for simultaneous transmission as possible. To be specific, at time slot  $t$ , the link(s) for transmission is/are selected following the rules below:

*Step 1:* First, let  $N_s = \lceil N/2 \rceil$ , and find the selection vector  $\mathbf{link}^{(N_s)}$ , or list all possible link selections for  $N_s$  simultaneous link transmissions.

- If none of the link selections in  $\mathbf{link}^{(N_s)}$  is available, then go to *Step 2*.
- Otherwise, use the *max-min* to choose the best  $N_s$  simultaneous link transmission among all available links in  $\mathbf{link}^{(N_s)}$ , as

$$\text{link}_b^{(N_s)} = \arg \max_{\text{link}^{(N_s)}(i) \in \mathcal{A}} \left\{ \min_{\text{link}_i \in \text{link}^{(N_s)}(i)} \{\gamma_i\} \right\}, \quad (22)$$

where  $\mathcal{A}$  is the set containing all available links.

- Check whether the link selection  $\text{link}_b^{(N_s)}$  is in outage or not.
  - If  $\text{link}_b^{(N_s)}$  is in outage, then no  $N_s$  simultaneous link transmission is possible at time  $t$  and go to *Step 2*.
  - Otherwise select  $\text{link}_b^{(N_s)}$  for data transmission at time  $t$ .

*Step 2:* Let  $N_s \leftarrow (N_s - 1)$  and repeat *Step 1* until  $N_s = 1$ .

In order to better understand the proposed link selection rule, we consider the 4-hop relay network as an example. We suppose at time slot  $t$ , all links are available except  $\text{link}_3$ . Then the selection vectors for available links are obtained by removing all selections containing  $\text{link}_3$  in (21), so that we have

$$\begin{aligned} \mathbf{link}^{(N_s=1)} &= [\text{link}_1, \text{link}_2, \text{link}_4] \\ \mathbf{link}^{(N_s=2)} &= [\text{link}_{1+4}, \text{link}_{2+4}] \end{aligned} \quad (23)$$

Then the links are selected as following.

*Step 1:* Let  $N_s = 2$ , and find the best selection of 2 simultaneous link transmission as

$$\text{link}_b^{(N_s=2)} = \arg \max \{ \min\{\gamma_1, \gamma_4\}, \min\{\gamma_2, \gamma_4\} \} \quad (24)$$

We assume that solution from (24) is  $link_b^{(N_s=2)} = link_{1+4}$ . Then we check whether  $\min\{C_1, C_4\} < r_t$  or not

- If ‘no’,  $link_{1+4}$  is not in outage and is selected for data transmission at time slot  $t$ .
- Otherwise, go to Step 2.

Step 2: Let  $N_s = 1$ , and find the best selection of single link transmission as

$$link_b^{(N_s=1)} = \arg \max \{\gamma_1, \gamma_2, \gamma_4\} \quad (25)$$

We assume that solution from (25) is  $link_b^{(N_s=1)} = link_2$ . Then we check whether  $\min\{C_2\} < r_t$  or not

- If ‘no’, then choose  $link_2$  for data transmission.
- Otherwise, outage occurs.

The proposed buffer-aided network-coding scheme takes advantages of both the network-coding and max-link schemes. On the one hand, because higher link selection priority is given to simultaneous transmission, the average data rate is higher than that of the traditional scheme. Particularly, when  $SNR \rightarrow \infty$ , we have  $P_{out} \rightarrow 0$  so that the average throughput of the proposed scheme is  $r_t/2$ , which is the same as that for the network-coding scheme. On the other hand, in the proposed scheme, the outage occurs only if all available links are in outage. This is similar to the max-link scheme. Therefore, the outage performance of the proposed and max-link scheme are similar.

From (1), the average throughput of the proposed scheme is given by

$$\bar{\eta}^{(\text{buffer-code})} = \bar{R}^{(\text{buffer-code})} \cdot (1 - P_{out}^{(\text{buffer-code})}), \quad (26)$$

where  $P_{out}^{(\text{buffer-code})}$  and  $\bar{R}^{(\text{buffer-code})}$  are the outage probability and average data rate of the proposed scheme, which are given by (40) and (51) obtained in the following two sections respectively.

#### IV. OUTAGE PROBABILITY

At any time, the numbers of data packets in the relay buffers form a “state”. Because each buffer has size  $L$ , there are  $(L+1)^{N-1}$  states in total, where the  $i$ -th state vector is defined as

$$\mathbf{s}_i = [\Psi_i(Q_1), \Psi_i(Q_2), \dots, \Psi_i(Q_{N-1})], \quad i = 1, \dots, (L+1)^{N-1}, \quad (27)$$

where  $0 \leq \Psi_i(Q_k) \leq L$  for all  $k = 1, \dots, N-1$  which is the buffer length (or the number of data packets in the buffer) of  $Q_k$  at state  $\mathbf{s}_i$ . At every time, depending on which link(s) is/are selected for transmission, the state may move to several possible states at the next time, forming a Markov chain.

Considering all possible states, the outage probability of the buffer-aided network-coding scheme can be obtained as

$$P_{out}^{(\text{buffer-code})} = \sum_{i=1}^{(L+1)^{N-1}} \pi_i \cdot p_{out}^{s_i}, \quad (28)$$

where  $\pi_i$  and  $p_{out}^{s_i}$  are the stationary probability and outage probability for state  $\mathbf{s}_i$  respectively.

In the following two subsections, we derive  $p_{out}^{s_i}$  and  $\pi_i$  respectively.

##### A. $p_{out}^{s_i}$ : outage probability for state $\mathbf{s}_i$

According to the link selection rules of the proposed buffer-aided network-coding scheme, at state  $\mathbf{s}_i$ , outage occurs only if all available links are in outage. Recalling that a link is available when the buffers of the corresponding transmission and receiving nodes are not empty and full respectively, we define the available-link vector for the state  $\mathbf{s}_i$  in the  $N$ -hop network as

$$\mathbf{a}_i = [a_i(1), a_i(2), \dots, a_i(N)] \quad (29)$$

where  $a_i(n)$  can only be ‘1’ or ‘0’, indicating that the corresponding  $link_n$  is available or not available at state  $\mathbf{s}_i$  respectively. For instance, in the 4-hop example in Section III-C where the buffers are at the state that all links except  $link_3$  are available, we have  $\mathbf{a}_i = [1 \ 1 \ 0 \ 1]$ .

Because all channels are i.i.d., the outage probability for state  $\mathbf{s}_i$  is given by

$$p_{out}^{s_i} = (P(C_i < r_t))^{|a_i|^+}, \quad (30)$$

where  $P(C_i < r_t)$  is the probability that a single link becomes outage, and  $|a_i|^+$  is the total number of available links at state  $\mathbf{s}_i$  which is the number of ‘1’-s in  $\mathbf{a}_i$ .

Because  $C_i = \log(1 + \gamma_i)$  and the SNR  $\gamma_i$  is exponentially distributed, we have

$$P(C_i < r_t) = F_\gamma(\Delta) = \left(1 - e^{-\frac{\Delta}{\gamma}}\right), \quad (31)$$

where  $F_\gamma(\cdot)$  is the cumulative distribution function (CDF) of  $\gamma_i$  and  $\Delta = 2^{r_t} - 1$ . Substituting (31) into (30) gives

$$p_{out}^{s_i} = F_\gamma^{|\mathbf{a}_i|^+} = \left(1 - e^{-\frac{\Delta}{\gamma}}\right)^{|\mathbf{a}_i|^+}, \quad (32)$$

where  $\Delta$  is ignored in  $F_\gamma(\Delta)$  without causing any confusion.

##### B. $\pi_i$ : the stationary probability of state $\mathbf{s}_i$

In order to obtain the stationary probability  $\pi_i$  for every state, first we need to calculate the state transition matrix  $\mathbf{A}$  which is an  $(L+1)^{N-1}$  by  $(L+1)^{N-1}$  matrix, where the entry  $A_{j,i} = P(X_{t+1} = \mathbf{s}_j | X_t = \mathbf{s}_i)$  is the transition probability that the state moves from  $\mathbf{s}_i$  at time  $t$  to  $\mathbf{s}_j$  at time  $(t+1)$ .

We suppose that the buffer state is  $\mathbf{s}_i$  at time slot  $t$ . If the outage occurs, the buffer state remains at  $\mathbf{s}_i$  at the time slot  $(t+1)$ . Otherwise,  $\mathbf{s}_i$  may move to several possible states at  $(t+1)$ , which are denoted as  $\mathbf{s}_{j_1}, \dots, \mathbf{s}_{j_{Q_i}}$  respectively. The state transition from  $\mathbf{s}_i$  to  $\mathbf{s}_{j_q}$  ( $j_q \in \{j_1, \dots, j_{Q_i}\}$ ) is the result of one particular link selection which is represented by the selection vector defined as

$$\mathbf{sel}_i^{(j_q)} = [\mathbf{sel}_i^{(j_q)}(1), \dots, \mathbf{sel}_i^{(j_q)}(N)], \quad (33)$$

where  $\mathbf{sel}_i^{(j_q)}(n)$  can only take values of 1 or 0, indicating the corresponding  $link_n$  is selected or not respectively. For example, in the 4-hop network,  $\mathbf{sel}_i = [1 \ 0 \ 0 \ 1]$  represents the link selection of  $link_{1+4}$ . With these observations, we have

$$A_{j,i} = \begin{cases} p_{out}^{s_i}, & j = i \\ P(\mathbf{sel}_i^{(j)}), & j \in \{j_1, \dots, j_{Q_i}\} \\ 0, & \text{otherwise} \end{cases} \quad (34)$$

where  $P(\text{sel}_i^{(j)})$  is the probability to choose the link selection  $\text{sel}_i^{(j)}$  at state  $\mathbf{s}_i$ . While  $p_{out}^{s_i}$  is given by (32), below we calculate  $P(\text{sel}_i^{(j)})$ .

According to the proposed link selection rules, the link selection at state  $\mathbf{s}_i$  depends on the outage events at every available links, where the priority is given to as many simultaneously link transmission as possible. Only when a link is both available and not in outage, may it be used for data transmission. We define the *good-link* vector to indicate whether the links are ‘good’ or not for data transmission at state  $\mathbf{s}_i$  as

$$\mathbf{g}_i = [g_i(1), g_i(2), \dots, g_i(N)] \quad (35)$$

where  $g_i(n)$  can only take values of ‘1’, ‘-1’ or ‘0’,  $g_i(n) = 1$  indicates that the corresponding  $\text{link}_n$  is not only available but also not in outage,  $g_i(n) = -1$  indicates that  $\text{link}_n$  is available but in outage, and  $g_i(n) = 0$  indicates that  $\text{link}_n$  is not available.

Comparing (29) and (35) shows that, for every state  $\mathbf{s}_i$ , it corresponds to one available-link vector  $\mathbf{a}_i$ , which again corresponds a set of good-link vectors including all possible link outages of the available links. Because the state  $\mathbf{s}_i$  has  $|\mathbf{a}_i|^+$  available links, there are  $G_i = \binom{|\mathbf{a}_i|^+}{1} + \dots + \binom{|\mathbf{a}_i|^+}{|\mathbf{a}_i|^+-1}$  good-link vectors for  $\mathbf{s}_i$ , denoting as  $\mathbf{g}_i^{(1)}, \dots, \mathbf{g}_i^{(G_i)}$  respectively, where  $\binom{|\mathbf{a}_i|^+}{n}$  is the (combination) probability that  $n$  links become outage among all  $|\mathbf{a}_i|^+$  available links. The probability of the  $k$ -th good-link vector is obtained as

$$P(\mathbf{g}_i^{(k)}) = \bar{F}_\gamma^{|\mathbf{g}_i^{(k)}|^+} \cdot F_\gamma^{|\mathbf{g}_i^{(k)}|^-}, \quad k = 1, \dots, G_i \quad (36)$$

where  $|\mathbf{g}_i^{(k)}|^+$  and  $|\mathbf{g}_i^{(k)}|^-$  give the number of ‘1’-s and ‘-1’-s in  $\mathbf{g}_i^{(k)}$  respectively, and  $\bar{F}_\gamma = 1 - F_\gamma$  which is the probability that a single link is not in outage.

From the proposed link selection rules, for every good-link vector, it may lead to several possible link selections, depending on the channel gains at the current time slot. On the other hand, one link selection may also correspond to several good-link vectors. As a result, we can form a 2 stage trellis-like diagram for the state  $\mathbf{s}_i$ , as is illustrated in Fig. 4 for the 4-hop network. At the first stage, there are  $G_i$  nodes, where each node corresponds to one good-link vector  $\mathbf{g}_i$ . At the second stage, there are  $Q_i$  nodes, each corresponding to one link selection vector  $\text{sel}_i$ .

We assume that the  $k$ -th node at stage 1,  $\mathbf{g}_i^{(k)}$ , leads to  $N_k$  nodes at stage 2, denoting as  $\text{sel}_i^{(n_1)}, \dots, \text{sel}_i^{(n_{N_k})}$  respectively. Because the channels are i.i.d., the probabilities for the pathes from node  $\mathbf{g}_i^{(n)}$  to any of these  $N_k$  nodes at stage 2 are the same, or we have

$$P(\mathbf{g}_i^{(k)} \rightarrow \text{sel}_i^{(j)}) = \begin{cases} P(\mathbf{g}_i^{(k)}) \cdot \frac{1}{N_k}, & j \in \{n_1, \dots, n_{N_k}\} \\ 0, & \text{otherwise} \end{cases} \quad (37)$$

Then further from (34), the transition probability from  $\mathbf{s}_i$  to  $\mathbf{s}_j$  is the summation of the probabilities of all pathes that ends

at the node  $\text{sel}_i^{(j)}$ , which is given by

$$\begin{aligned} A_{j,i} &= P(\text{sel}_i^{(j)}) \\ &= \sum_{k=1}^{G_i} P(\mathbf{g}_i^{(k)} \rightarrow \text{sel}_i^{(j)}), \quad j \in \{j_1, \dots, j_{Q_i}\} \end{aligned} \quad (38)$$

Substituting (38) into (34), and applying it on all states, we can obtain the state transition matrix  $\mathbf{A}$ .

Because the transition matrix  $\mathbf{A}$  is column stochastic, irreducible and aperiodic<sup>2</sup>, the stationary state probability vector is obtained as (see [20] and [21])

$$\boldsymbol{\pi} = (\mathbf{A} - \mathbf{I} + \mathbf{B})^{-1} \mathbf{b}, \quad (39)$$

where  $\boldsymbol{\pi} = [\pi_1, \dots, \pi_{(L+1)^{N-1}}]^T$ ,  $\mathbf{b} = [\mathbf{1}, \dots, \mathbf{1}]^T$ ,  $\mathbf{I}$  is the identity matrix and  $\mathbf{B}_{n,l}$  is an  $n \times l$  all one matrix.

Finally, substituting (32) and (39) into (28) gives the outage probability of the overall system as

$$\begin{aligned} P_{out}^{(\text{buffer-code})} &= \sum_{i=1}^{(L+1)^{N-1}} \pi_i \cdot p_{out}^{s_i} = \text{diag}(\mathbf{A}) \cdot \boldsymbol{\pi} \\ &= \text{diag}(\mathbf{A}) \cdot (\mathbf{A} - \mathbf{I} + \mathbf{B})^{-1} \mathbf{b}, \end{aligned} \quad (40)$$

where  $\text{diag}(\mathbf{A})$  is the vector consisting of all diagonal elements of  $\mathbf{A}$ .

1) *Illustration - the 4-hop relay network*: In order to better understand the above analysis, we give an example of the 4-hop relay network with buffer size of  $L = 4$ . As an illustration, we consider the state transition for the state  $\mathbf{s}_i = [2 \ 0 \ 2]^T$ , or the buffer lengths at nodes  $R_1$ ,  $R_2$  and  $R_3$  are 2, 0 and 2 respectively. This is actually the same example in Section III-C where the selection rules are explained. As is shown in Fig. 3, there are 5 possible states that  $\mathbf{s}_i$  can move to at time  $(t+1)$ , denoting as  $\mathbf{s}_{j_1}, \dots, \mathbf{s}_{j_5}$  respectively, and each  $\mathbf{s}_{j_q}$  corresponds to one link selection.

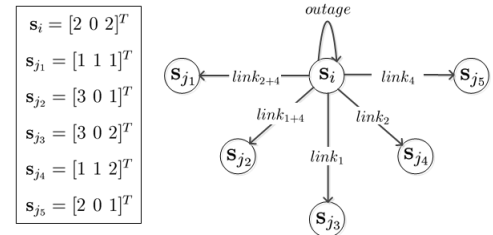


Fig. 3. State transition diagram for the state  $\mathbf{s}_i = [2 \ 0 \ 2]^T$  in the 4-hop relay network with buffer size of  $L = 4$ .

At state  $\mathbf{s}_i = [2 \ 0 \ 2]^T$ , all links except  $\text{link}_3$  are available, so that the available-link vector is given by

$$\mathbf{a}_i = [1 \ 1 \ 0 \ 1] \quad (41)$$

The trellis diagram for the transition probability of state  $\mathbf{s}_i$  is shown in Fig. 4, where there are 7 nodes (good-link vectors) at stage 1, and 5 nodes (link selection vectors) at stage 2. The probabilities for every link selection can be obtained from Fig.

<sup>2</sup>Column stochastic means all entries in any column sum up to one, irreducible means that it is possible to move from any state to any state, and aperiodic means that it is possible to return to the same state at any steps [19], [20]

4. For example, for  $link_{2+4}$  which is highlighted in red, we have

$$\begin{aligned}
 A_{j_1,i} &= P(s_i \rightarrow s_{j_1}) = P(\mathbf{sel}_i^{(j_1)}) \\
 &= P\left(\mathbf{g}_i^{(1)} = [1 \ 1 \ 0 \ 1]\right) \cdot \frac{1}{2} + P\left(\mathbf{g}_i^{(4)} = [-1 \ 1 \ 0 \ 1]\right) \cdot 1 \\
 &= \frac{1}{2}\bar{F}_\gamma^3 + F_\gamma\bar{F}_\gamma^2
 \end{aligned} \tag{42}$$

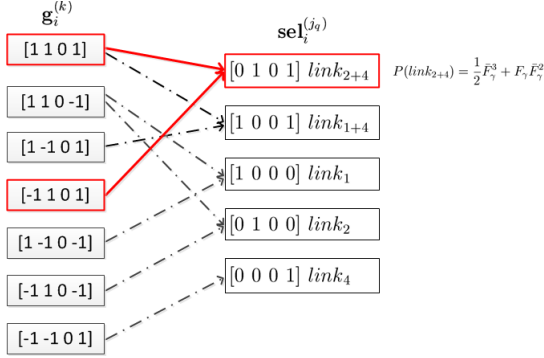


Fig. 4. Trellis diagram for the transition probability for the state  $s_i = [2 \ 0 \ 2]$  in the 4-hop relay network.

## V. AVERAGE DATA RATE

In this section, we first introduce the concept of “effective” hops and then use a trellis diagram to obtain the average data rate.

### A. Effective hops

In the proposed  $N$ -hop link selection scheme, although every data packet needs to go through the  $N$  hops consecutively to reach the destination, at some time slots, several packets may be simultaneously transmitted at different links. Thus by average, it takes fewer than  $N$  time slots to deliver one packet to the destination, or the number of ‘effective’ hops to transmit one packet is fewer than  $N$ . To be specific, at one time slot, if several packets are transmitted simultaneously, this time slot is only counted as one effective hop for one of the packets.

In order to better understand the influence of the simultaneous transmission on the effective hop number, we look at the example of the 4-hop network as is shown in Fig. 2. Specifically, for data packet  $x(2)$ , we have the following observations:

- At time slot  $t = 1$ , data packets  $x(2)$  and  $x(1)$  are simultaneously transmitted at  $link_1$  and  $link_3$  respectively. We assume that the time slot is counted as one effective hop only for the packet at the link with the lowest number. At  $t = 1$ , the lowest numbered link is  $link_1$ . Thus  $t = 1$  contributes one effective hop only for  $x(2)$  transmission, but not for  $x(1)$  transmission.
- Similarly,  $t = 2$  contributes one effective hop for  $x(2)$  transmission, but not for  $x(1)$ .

- At  $t = 3$ ,  $x(3)$  and  $x(2)$  are simultaneously transmitted at  $link_1$  and  $link_3$  respectively. Because the lowest numbered link is  $link_1$ ,  $t = 3$  contributes one effective hop only for  $x(3)$  transmission, but not for  $x(2)$ .
- Similarly  $t = 4$  is counted as one effective hopping time for  $x(3)$ , but not for  $x(2)$ .

Therefore, although  $x(2)$  goes through all 4 hops to reach the destination, only  $t = 1$  and  $t = 2$  are counted as its effective hopping times, or the number of effective hops for  $x(2)$  transmission is 2. This leads to the following rule as:

*Effective hopping rule: at any time slot ‘t’, if multiple data packets are transmitted simultaneously, the time slot is only counted as one effective hopping time for the packet transmitted at the lowest numbered link.*

In the proposed buffer-aided network coding scheme, because different simultaneous link transmissions may be selected at different time slots, different data packets have different numbers of effective hops and the average data rate is obtained as

$$R^{(\text{buffer-code})} = \frac{1}{\bar{n}} \cdot r_t, \tag{43}$$

where  $\bar{n}$  is the average number of effective hops to transmit one data packet.

### B. Trellis diagram to obtain $\bar{n}$

Below we use the trellis diagram to analyze average number of effective hops  $\bar{n}$ . In the proposed  $N$ -hop relay scheme, for any data packet, it must go through all links ( $link_1, \dots, link_N$ ) consecutively to reach the destination. We suppose that at time slot  $t$ , a packet needs to go through  $link_n$ . There exist several possible link selections to make this happen: either only  $link_n$  is selected, or  $link_n$  is selected simultaneously with other links. On the other hand, the link selections only depend on the buffer states at time slot  $t$ , but not on other packet transmissions. With this observation, we describe an  $N$ -stage trellis diagram to represent all possible link selections for one packet transmission, as is illustrated in Fig. 5 for the 4-hop network. Every stage contains a set of nodes, where every node corresponds to one possible link selection to pass through the corresponding link.

Trellis nodes at adjacent stages are inter-connected, forming ‘paths’ from stage 1 to  $N$ . The total number of paths is given by

$$N_p = N_t^{(1)} \times \dots \times N_t^{(N)} \tag{44}$$

where  $N_t^{(n)}$  is the number of trellis nodes at stage  $n$ . Every path corresponds to one combination of link selections for a packet passing through the network.

Supposing that the  $k$ th path consists of  $n_k$ -th trellis node at the  $n$ -th stage, the  $k$ -th path is represented as

$$path_k = \left\{ \mathbf{sel}^{(1_k)}, \dots, \mathbf{sel}^{(N_k)} \right\}, \quad k = 1, \dots, N_p, \tag{45}$$

where  $\mathbf{sel}^{(n_k)}$  is defined in (33) which is the  $n_k$ -th link selection for a packet passing through  $link_n$ .

In order to obtain the number of effective hops for the  $k$ -th path, we define a binary function  $\mathcal{H}(\mathbf{sel}^{(n_k)})$ . If  $\mathcal{H}(\mathbf{sel}^{(n_k)}) = 1$ , then the corresponding transmission at stage

$i$  contributes one effective hop; otherwise if  $\mathcal{H}(\text{sel}^{(n_k)}) = 0$ , no effective hop is contributed at this stage. From the *effective hopping rule*, we understand that a time slot is counted as one effective hopping time for a packet, only if there are no other packets are transmitting simultaneously at lower numbered links. Thus we have

$$\mathcal{H}(\text{sel}^{(n_k)}) = \begin{cases} 1, & \mathcal{L}(\text{sel}^{(n_k)}) < i \\ 0, & \mathcal{L}(\text{sel}^{(n_k)}) = i \end{cases} \quad (46)$$

where  $\mathcal{L}(\text{sel}^{(n_k)})$  gives the index of the first '1' in the selection vector  $\text{sel}^{(n_k)}$ .

Then number of effective hops for  $\text{path}_k$  is then given by

$$\mathcal{N}_e(\text{path}_k) = \sum_{n=1}^N \mathcal{H}(\text{sel}^{(n_k)}) \quad (47)$$

On the other hand, the probability to choose  $\text{path}_k$  is given by

$$P(\text{path}_k) = \prod_{n=1}^N P(\text{sel}^{(n_k)}), \quad (48)$$

where  $P(\text{sel}_i^{(n_k)})$  is the probability to select  $\text{sel}^{(n_k)}$  which is given by

$$P(\text{sel}^{(n_k)}) = \sum_{i=1}^{(L+1)^{N-1}} \pi_i \cdot P(\text{sel}_i^{(n_k)}) \quad (49)$$

where  $P(\text{sel}_i^{(n_k)})$  is the probability to select  $\text{sel}^{(n_k)}$  at state  $s_i$  which is given by (38).

Then from (47) and (48), the average number of effective hops is obtained by averaging over all pathes in the trellis as

$$\bar{n} = \sum_{k=1}^{N_p} \mathcal{N}_e(\text{path}_k) \cdot P(\text{path}_k) \quad (50)$$

Substituting (50) into (43) gives the average data rate as

$$R^{(\text{buffer-code})} = \frac{1}{\sum_{k=1}^{N_p} \mathcal{N}_e(\text{path}_k) \cdot P(\text{path}_k)} \cdot r_t, \quad (51)$$

### C. An illustration of the hopping trellis diagram for the 4-hop relay network

Fig. 5 shows the hopping trellis diagram for the 4-hop relay network, where there are 4 stages (or columns) corresponding to a packet passing through  $\text{link}_1$  to  $\text{link}_4$  respectively. At stage 1, there are 3 nodes corresponding to 3 selection vectors, namely  $[1\ 0\ 0\ 0]$ ,  $[1\ 0\ 1\ 0]$  and  $[1\ 0\ 0\ 1]$  respectively. We note that the first element of all of the three vectors at stage 1 is 1. Therefore, for a packet to go through  $\text{link}_1$ , it must correspond to one of these selection vectors. Similarly, there are 2, 2 and 3 trellis nodes at stage 2, 3 and 4 respectively.

It is clear in Fig. 5 that there are  $3 \times 2 \times 2 \times 3 = 36$  paths for the 4-hop relay network, where each path corresponds to one combination of link selections for a packet passing through the network. For example, the  $k$ -th path, which is highlighted with red, is represented as

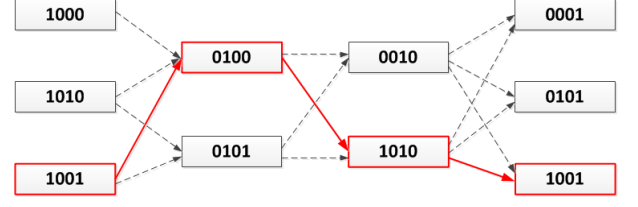


Fig. 5. Hopping trellis diagram for the 4-hop relay network.

$$\begin{aligned} & \{\text{sel}^{(1_k)}, \text{sel}^{(2_k)}, \text{sel}^{(3_k)}, \text{sel}^{(4_k)}\} \\ &= \{ [1\ 0\ 0\ 1], [0\ 1\ 0\ 0], [1\ 0\ 1\ 0], [1\ 0\ 0\ 1] \}, \end{aligned} \quad (52)$$

which corresponds to the combination of selections as  $\text{link}_{1+4}$ ,  $\text{link}_2$ ,  $\text{link}_{1+3}$  and  $\text{link}_{1+4}$  consecutively.

Table I lists the effective hops at every stage for the path in (52). Particularly, at stage 1,  $\text{sel}^{(1_k)} = [1\ 0\ 0\ 1]$  which is the link selection for the first hop for this path. It is clear that the index of the first '1' is  $\mathcal{L}(\text{sel}^{(1_k)}) = 1$  which is equal to the hop index (or the first hop). Thus  $\text{sel}^{(1_k)}$  contributes one effective hop for this path, or we have  $\mathcal{H}(\text{sel}^{(1_k)}) = 1$ . On the other hand, at stage 4, although  $\text{sel}^{(4_k)} = \text{sel}^{(1_k)} = [1\ 0\ 0\ 1]$ ,  $\text{sel}^{(4_k)}$  does not contribute one effective hop for this path. This is because that, for  $\text{sel}^{(4_k)}$ , the hop index is now 4 which is not equal to the index of the first '1' (which is still 1).

TABLE I  
EFFECTIVE HOPS FOR THE PATH IN (52)

$\text{sel}^{(n_k)}$	$[1\ 0\ 0\ 1]$	$[0\ 1\ 0\ 0]$	$[1\ 0\ 1\ 0]$	$[1\ 0\ 0\ 1]$
$n$	1	2	3	4
$\mathcal{L}(\text{sel}^{(n_k)})$	1	2	1	1
$\mathcal{H}(\text{sel}^{(n_k)})$	1	1	0	0

Then from (47), the number of effective hops for the path in (52) is given by  $\sum_{n=1}^N \mathcal{H}(\text{sel}^{(n_k)}) = 1 + 1 + 0 + 0 = 2$ . The probability to choose this packet is given by

$$P(\text{path}_k) = P([1\ 0\ 0\ 1]) \cdot P([0\ 1\ 0\ 0]) \cdot P([1\ 0\ 1\ 0]) \cdot P([1\ 0\ 0\ 1]) \quad (53)$$

## VI. AVERAGE PACKET DELAY

The delay of a packet in the  $N$ -hop network is defined as the duration between the time when the packet leaves the source and the time when it arrives the destination. In the non-buffer-aided schemes (e.g. the traditional or network-coding based scheme), when a packet reaches one node, it will be immediately forwarded to the next node at the following time slot, so that the delay for every packet is  $N$  time slots. On the other hand, in the buffer-aided scheme, because the data packets may queue at the relay nodes, the packet delay also includes the queuing time. We particularly note that the packet delay is different from the number of effective hops, where the latter does not take into account of the queueing times at the relays.



Because it takes one time slot to transmit a packet from the source to  $R_1$ , the average packet delay in the network is given by

$$\bar{D}^{(\text{buffer-code})} = 1 + \sum_{k=1}^{N-1} \bar{D}_k, \quad (54)$$

where  $\bar{D}_k$  is the average delay at relay  $R_k$ .

Using Little's law [22], the average delay at node  $i$  can be obtained as

$$\bar{D}_k = \frac{\bar{L}_k}{\bar{\eta}_k}, \quad k = 1, \dots, N \quad (55)$$

where  $\bar{L}_k$  and  $\bar{\eta}_k$  are the average queuing length and average throughput at node  $R_k$  respectively.

Because all nodes are connected in series, the average throughput at every node is the same, which is equal to the system average throughput as

$$\bar{\eta}_k = \bar{\eta}^{(\text{buffer-code})}, \quad k = 1, \dots, N \quad (56)$$

where  $\bar{\eta}^{(\text{buffer-code})}$  is given by (26).

On the other hand, the average queuing length at relay  $R_k$  is obtained by averaging the buffer lengths over all buffer states as

$$\bar{L}_k = \sum_{l=1}^{(L+1)^{N-1}} \pi_l \Psi_l(Q_k), \quad k = 1, \dots, N \quad (57)$$

where we recall that  $\Psi_l(Q_k)$  gives the number of packets (or the buffer length) of buffer  $Q_k$  at state  $s_l$ .

Substituting (56) and (57) into (55), and further into (54), gives the proposed average packet delay in the buffer-aided network-coding scheme as

$$\bar{D}^{(\text{buffer-code})} = 1 + \frac{\sum_{l=1}^{(L+1)^{N-1}} \sum_{k=1}^{N-1} \pi_l \Psi_l(Q_k)}{\bar{\eta}^{(\text{buffer-code})}}. \quad (58)$$

It is interesting to compare the average packet delays of the two buffer-aided schemes: the max-link and proposed schemes respectively. On the one hand, the proposed scheme has higher throughput than the max-link scheme, or  $\bar{\eta}^{(\text{buffer-code})} > \bar{\eta}^{(\text{max-link})}$ . On the other hand, because of the simultaneous data transmission in the proposed scheme, the data packets move more quickly through the system, resulting in shorter queuing lengths at the relays, than the max-link scheme. From the Little law (as is shown in (55)), the average packet delay of the proposed scheme is significantly smaller than that of the max-link scheme.

## VII. SIMULATION

In this section, numerical results are shown to verify the proposed scheme in this paper. In all simulations, the transmission powers and the noise powers are normalized to unity, the transmission rates are set as  $r_t = 1$ , all channels are i.i.d. Rayleigh fading, and the channel coefficients remains unchanged during one hopping time slot but vary independently from one time slot to another. Both the simulation and theoretical results are shown, where the simulation results are obtained by averaging over 100,000 independent runs. Other parameters including the buffer size and number of hops are set individually for every simulation.

### A. Average system throughput

Fig. 6 (a) and (b) shows outage probability and average data rate for consecutive-hopping, max-link, network-coding, and buffer-aided network-coding schemes in the 5-hop relay network respectively. First, for the proposed scheme, the simulations well match the theoretical results for both the outage probability and data rate, which verifies the analysis in this paper. It is interesting to observe in Fig. 6 (a) that the proposed and traditional max-link schemes have similar outage performance. This is not surprising because both proposed and max-link are buffer-aided schemes, where the outage occurs only when all of the available links are in outage. Fig. 6 (a) also shows that buffer-aided schemes (including both the proposed and traditional max-link schemes) have the best outage performance, while the network-coding scheme has even worse outage probability than the consecutive-hopping scheme. On the other hand, it is shown in Fig. 6 (b) that the network-coding scheme has the highest data rate (0.5 packet/time-slot), while both the consecutive-hopping and max-link schemes have the lowest data rate (0.2 packet/time-slot). It is interesting to observe that the proposed scheme has similar data rate as the consecutive-hopping scheme at low SNR range. But when the SNR is high enough, the data rate of the proposed scheme approaches to that of the network-coding scheme. Combining Fig. 6 (a) and (b), it is well expected that the proposed scheme must have the highest throughput among all schemes. This will be verified in the following simulation.


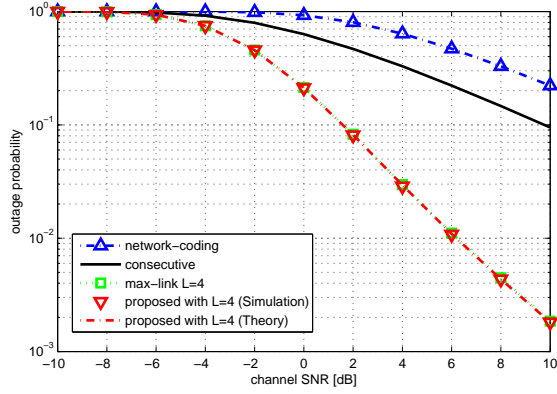


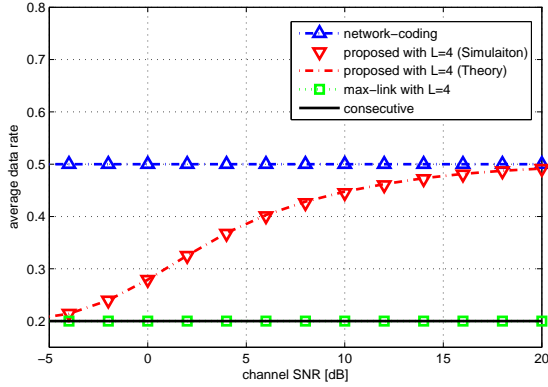
Fig. 7 (a) and (b) shows the average throughput for different schemes in the 5-hop and 3-hop relay networks respectively, where in both the max-link and proposed schemes, the buffer sizes for the 3-hop and 5-hop network are set as  $L = 3$  and  $L = 4$  respectively. In Fig. 7, the simulations also perfect matches the theoretical results for the proposed scheme. As is expected, in both 3-/5- hop networks, the network-coding scheme can achieve the maximum throughput of  $1/2$  at high SNRs (e.g. SNR  $> 20$  dB), but it has lower throughput than the consecutive-hopping scheme for small SNRs. The reason is shown in Fig. 6 that, compared with the consecutive-hopping scheme with data rate of  $R = 1/N$ , though the network-coding scheme increases the data rate to  $R = 1/2$ , it also increases the outage probability. For the max-link scheme, it significantly increases the throughput at low SNRs, but has the same throughput of  $1/N$  as the traditional scheme at high SNRs, where the reason is also shown in Fig. 6. This verifies the our expectation that the network-coding and max-link schemes improve the throughput at high SNRs and low SNRs respectively.

On the other hand, it is clearly shown in Fig. 7 that the proposed buffer-aided network-coding scheme takes advantage of both network-coding and max-link schemes, leading to significantly improvement in throughput at all SNR ranges. Particularly, when the SNR is large enough, the proposed scheme has the same maximum throughput of  $1/2$  as the network-coding scheme.

Fig. 8 compares the average throughput vs the buffer size  $L$  between the max-link and proposed schemes for the 3-hop



(a) Outage probability.



(b) Average data rate.

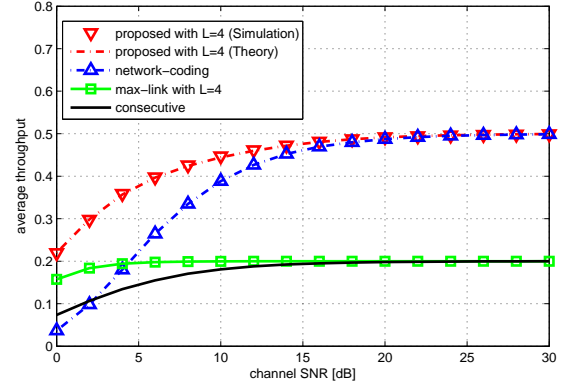
Fig. 6. Outage probability and average data rate for consecutive-hopping, max-link, network-coding, and buffer-aided network-coding schemes in the 5-hop relay network.

relay network. It is clearly shown that, for every buffer size  $L$ , the proposed scheme always has higher throughput than the max-link scheme, where the former can reach the data rate of  $1/2$  and the latter can only reach  $1/3$  when the SNR is very large. In both schemes, the average throughput becomes higher with larger buffer size, but the improvement becomes less significant when the buffer size is larger. For example, the throughput difference between those for  $L = 10$  and  $L = 5$  is trivial in both schemes.

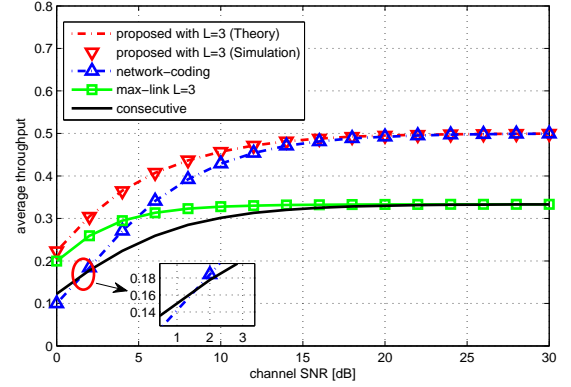
### B. Average packet delay

This simulation investigates the average packet delay. The unit of the delays is “time slot”, where one time slot is used for a packet transmitting from one node to the next. Table II compares the theoretical analysis (based on (58)) and simulation results of the proposed buffer-aided network-coding scheme for both 3-hop and 5-hop network, where the channel SNR is set as 15 dB. It is clearly shown that, in both networks, the theoretical analysis very well matches the simulation results. Together with the results in Fig. 8, we obtain that it is not necessary to have a very large buffer size  $L$  as otherwise it not only has little improvement in throughput but also unnecessarily increases the average packet delay.

Table III compares average packet delays between the max-link and proposed schemes for the 3-hop relay network. It is



(a) 5-hop.



(b) 3-hop.

Fig. 7. Throughput comparison among traditional, network-coding, max-link and buffer-aided network-coding schemes.

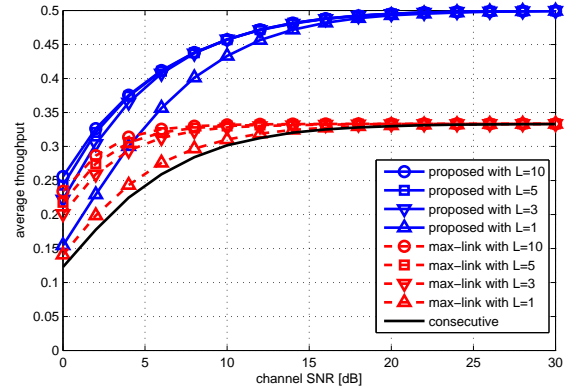


Fig. 8. Throughput vs buffer length  $L$ , for the max-link and buffer-aided network-coding schemes in the 3-hop relay network.

clearly shown that, while both schemes have larger average packet delays with larger buffer size  $L$ , the max-link scheme has approximately 50% larger average packet delay than the proposed scheme. This well matches our expectation in Section VI.

## VIII. CONCLUSION

In this paper, we proposed a novel buffer-aided network-coding link selection scheme for the  $N$ -hop relay network. The proposed scheme applied buffers at the relays to decrease

TABLE II  
AVERAGE PACKET DELAYS OF THE BUFFER-AIDED NETWORK-CODING SCHEME

	3-hop Average Delay		5-hop Average Delay	
Buffer size	Simulation	Theory	Simulation	Theory
L=1	3.09	3.09	5.28	5.30
L=3	7.15	7.18	13.82	13.62
L=5	11.28	11.30	21.78	21.95

TABLE III  
AVERAGE PACKET DELAYS COMPARISON BETWEEN THE MAX-LINK AND PROPOSED SCHEME IN THE 3-HOP NETWORK.

Channel SNR	3-hop Schemes	L=5	L=10	L=20
10 dB	Proposed	13.10	24.07	45.86
	Max-link	18.37	33.53	62.51
20 dB	Proposed	12.12	22.22	42.42
	Max-link	17.49	32.63	61.28

the outage, and used network-coding to increase the data rate. As a result, the throughput at all SNR ranges is increased. We described new analysis tools to analyze the outage probability and average data rate, based on which the average throughput of the proposed scheme was successfully obtained. We also analyzed the average packet delay. The analysis shows that, the proposed scheme not only has higher throughput, but also lower average packet delay, than the existing buffer-aided max-link scheme, making it an attractive approach in the multi-hop network.

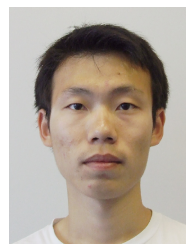
#### ACKNOWLEDGEMENT

The authors would like to thank the editors and anonymous reviewers for their constructive comments to improve the quality of this manuscript. Particularly the derivation of the average throughput for the traditional consecutive hopping scheme is based on the suggestion from one of the reviewers.

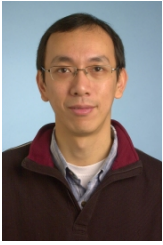
#### REFERENCES

- [1] W. J. Huang, Y. W. Hong, and C. C. J. Kuo, "Lifetime Maximization for Amplify-and-Forward Cooperative Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1800–1805, May 2008.
- [2] A. Papadogiannis, H. Bang, D. Gesbert, and E. Hardouin, "Efficient Selective Feedback Design for Multicell Cooperative Networks," *IEEE Trans. Veh. Technology*, vol. 60, no. 1, pp. 196–205, Jan. 2011.
- [3] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative Communications in Wireless Networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, 2004.
- [4] J. N. Laneman, D. N. C. Tes, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: efficient Protocols and Outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [5] P. Pahlavan, M. Hündebell, M. V. Pedersen, D. E. Lucani, H. Charaf, F. H. P. Fitzek, H. Bagheri, and M. Katz, "Novel Concepts for Device-to-Device Communication Using Network Coding," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 32–39, Apr. 2014.
- [6] C. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource Sharing Optimization for Device-to-Device Communication Underlying Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Jun. 2011.
- [7] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator Controlled Device-to-Device Communications in LTE-advanced Networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 96–104, Jun. 2012.
- [8] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-Aided Relaying with Adaptive Link Selection," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1530–1542, Aug. 2013.

- [9] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-Aided Relaying with Adaptive Link Selection-Fixed and Mixed Rate Transmission," *IEEE Trans. Inform. Theory*, vol. 59, no. 5, pp. 2816–2840, May. 2013.
- [10] I. Krikidis, T. Charalambous, and J. S. Thompson, "Buffer-Aided Relay Selection for Cooperative Diversity Systems without Delay Constraints," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1957–1967, May 2012.
- [11] A. Ikhlef, D. S. Michalopoulos, and R. Schober, "Buffers Improve the Performance of Relay Selection," in *Proc. 2011 IEEE Global Commun. Conf. Houston, Texas, USA*, pp. 1–6, Dec. 2011.
- [12] Z. Tian, G. Chen, Z. Chen, Y. Gong, and J. A. Chambers, "Buffer-aided Max-link Relay Selection in Amplify-and-Forward Cooperative Networks," *IEEE Trans. Veh. Technology*, vol. 64, no. 2, pp. 553–565, May 2014.
- [13] B. Xia, Y. Fan, J. Thompson, and H. V. Poor, "Buffering in a Three-Node Relay Network," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4492–4496, Nov. 2008.
- [14] G. Chen, Z. Tian, Y. Gong, and J. A. Chambers, "Decode-and-Forward Buffer-Aided Relay Selection in Cognitive Relay Networks," *IEEE Trans. Veh. Technology*, vol. 63, no. 9, pp. 4723–4728, Mar. 2014.
- [15] G. Chen, Z. Tian, Y. Gong, Z. Chen, and J. A. Chambers, "Max-Ratio Relay Selection in Secure Buffer-Aided Cooperative Wireless Networks," *IEEE Trans. Inform. Forensics and Security*, vol. 9, no. 4, pp. 719–729, Apr. 2014.
- [16] H. Yang, W. Meng, B. Li, and G. Wang, "Physical Layer Implementation of Network Coding in Two-way Relay Networks," in *Proc. 2012 IEEE Int. Conf. Commun., Ottawa, Canada*, pp. 671–675, Jun. 2012.
- [17] W. Liang, Y. Chen, K. Xu, H. Tian, and Y. Xu, "Joint LDPC and Physical-layer Network Coding for Two-way Relay Channels with Different Carrier Frequency Offsets," in *Proc. 2012 International Conference on Wireless Communications & Signal Processing, Huangshan, China*, pp. 1–4, Oct. 2012.
- [18] R. H. Y. Louie, Y. Li, and B. Vucetic, "Practical Physical Layer Network Coding for Two-way Relay Channels: Performance Analysis and Comparison," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 764–777, Feb. 2010.
- [19] C. M. Grinstead and J. L. Snell, "Introduction to Probability: Second Revised Edition," *American Mathematical Society*, 1991.
- [20] J. R. Norris, "Markov Chains," *Cambridge University Press*, 1998.
- [21] A. Berman and R. J. Plemmons, "Nonnegative Matrices in the Mathematical Sciences," *Society of industrial and applied mathematics*, 1994.
- [22] J. D. C. Little and S. C. Graves, "Little's Law," *International Series in Operations Research & Management Science*, vol. 115, pp. 81–100, 2008.

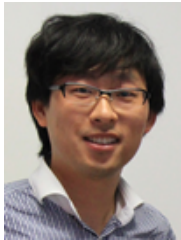


**Zhao Tian** (S'12-M'15) received the B. Eng and Ph.D degrees in School of Electronic, Electrical and Systems Engineering from Loughborough University, UK, in 2012 and 2015, respectively. He received full postgraduate scholarship from Engineering and Physical Sciences Research Council (EPSRC) when he was working toward the Ph.D degree. He is currently a Knowledge Transfer Project Associate in Computer Science of Aston University, Birmingham and IGI Ltd., Birmingham. His current research interests include the general field of wireless communications with emphasis on buffer-aided relaying, machine learning and data mining.



**Yu Gong** is with School of Electronic, Electrical and Systems Engineering, Loughborough University, UK, in July 2012. Dr Gong obtained his BEng and MEng in electronic engineering in 1992 and 1995 respectively, both at the University of Electronics and Science Technology of China. In 2002, he received his PhD in communications from the National University of Singapore. After PhD graduation, he took several research positions in Institute of Informcomm Research in Singapore and Queens University of Belfast in the UK respectively. From 2006 and

2012, Dr Gong had been an academic member in the School of Systems Engineering, University of Reading, UK. His research interests are in the area of signal processing and communications including wireless communications, cooperative networks, non-linear and non-stationary system identification and adaptive filters.



**Gaojie Chen** (S'09-M'12) received the B. Eng. and B. Ec. in Electrical Information Engineering and International Economics and Trade from the Northwest University, Shaanxi, China, in 2006, and the M.Sc (Distinction) and Ph.D degrees from Loughborough University, Loughborough, UK, in 2008 and 2012, respectively, all in Electrical and Electronic Engineering. From 2008 to 2009 he worked, as a software engineering in DTmobile, Beijing, China, and from 2012 to 2013 as a Research Associate in the School of Electronic, Electrical and Systems

Engineering at the Loughborough University, Loughborough, UK. Then he was a Research Fellow with the 5GIC, the Faculty of Engineering and Physical Sciences, University of Surrey, U.K., from 2014 to 2015. He is currently a Research Associate with the Department of Engineering Science, University of Oxford, U.K. His current research interests include information theory, wireless communications, cooperative communications, cognitive radio, secrecy communication and random geometric networks.



**Zhi Chen** (M'04) received B. Eng, M. Eng., and Ph.D. degree in Electrical Engineering from University of Electronic Science and Technology of China (UESTC), in 1997, 2000, 2006, respectively, all in electrical engineering. After his Ph.D. graduation, he joined the National Key Lab of Science and Technology on Communications, UESTC, where he was promoted to Professor in August 2013.. He was a visiting scholar at University of California, Riverside during 2010-2011. His current research

interests include wireless communication and signal processing, specifically relay and cooperative communications, interference coordination and cancellation. Dr. Chen has served as a reviewer for various international journals and conferences, including IEEE Transactions on Signal Processing, IEEE Transactions on Vehicular Technology, and many more.



**Jonathon A. Chambers** (S'83-M'90-SM'98-F'11) received the Ph.D. and D.Sc. degrees in signal processing from the Imperial College of Science, Technology and Medicine (Imperial College London), London, U.K., in 1990 and 2014, respectively.

From 1991 to 1994, he was a Research Scientist with the Schlumberger Cambridge Research Centre, Cambridge, U.K. In 1994, he returned to Imperial College London as a Lecturer in signal processing and was promoted to Reader (Associate Professor) in 1998. From 2001 to 2004, he was the Director of

the Centre for Digital Signal Processing and a Professor of signal processing with the Division of Engineering, Kings College London. From 2004 to 2007, he was a Cardiff Professorial Research Fellow with the School of Engineering, Cardiff University, Cardiff, U.K. Between 2007-2014, he led the Advanced Signal Processing Group, within the School of Electronic, Electrical and Systems Engineering and is now a Visiting Professor. In 2015, he joined the School of Electrical and Electronic Engineering, Newcastle University, where he is a Professor of signal and information processing and heads the ComS<sup>2</sup>IP group. He is also a Guest Professor at Harbin Engineering University, China. He is a co-author of the books *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (New York, NY, USA: Wiley, 2001) and *EEG Signal Processing* (New York, NY, USA: Wiley, 2007). He has advised more than 60 researchers through to Ph.D. graduation and published more than 400 conference proceedings and journal articles, many of which are in IEEE journals. His research interests include adaptive and blind signal processing and their applications.

Dr. Chambers is a Fellow of the Royal Academy of Engineering, U.K., and the Institution of Electrical Engineers. He was the Technical Program Chair of the 15th International Conference on Digital Signal Processing and the 2009 IEEE Workshop on Statistical Signal Processing, both held in Cardiff, U.K., and a Technical Program Co-chair for the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, Prague, Czech Republic. He received the first QinetiQ Visiting Fellowship in 2007 for his outstanding contributions to adaptive signal processing and his contributions to QinetiQ, as a result of his successful industrial collaboration with the international defence systems company QinetiQ. He has served on the IEEE Signal Processing Theory and Methods Technical Committee for six years and the IEEE Signal Processing Society Awards Board for three years. He is currently a member of the IEEE Signal Processing Conference Board and the European Signal Processing Society Best Paper Awards Selection Panel. He has also served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING for three terms over the periods 1997-1999, 2004-2007, and since 2011 as a Senior Area Editor.