Buffer State Based Relay Selection for Half-Duplex Buffer-Aided Serial Relaying Systems

Sawsan El-Zahr and Chadi Abou-Rjeily, *Senior Member IEEE*

***Abstract*—This work investigates multi-hop serial relaying half-duplex (HD) networks comprising one source, one destination and K decode-and-forward (DF) relays. Buffer-aided (BA) relays are considered where finite size buffers are added to relays. For this setup, we propose a novel BA relaying strategy that selects the node that must transmit over one epoch in a block fading environment. Selecting a node depends on the buffer state information of all relays and on the availability of the hops. Moreover, the proposed relaying scheme is controlled by an adjustable parameter that allows the system to achieve multiple levels of tradeoff between the average packet delay (APD) and the outage probability (OP). A Markov chain model is adopted to evaluate the system's performance and simple closed-form expressions were derived for the APD and OP based on an asymptotic analysis. The performance analysis proves the ability of the suggested relaying scheme to achieve significant OP and APD gains with small buffer sizes. Two variants of the proposed relaying scheme are particularly appealing. The APD-prioritizing variant achieves the smallest APD of $2K$ without improving the diversity order. The second OP-prioritizing variant reaches the full diversity order of $K+1$ while increasing the asymptotic APD to $K(K+3)$.**

***Index Terms*—Serial relaying, multi-hop, Markov chain, diversity order, buffer, optimization, data queue, queuing delay, outage probability, performance analysis, cooperative networks.**

I. Introduction

In the literature, there is an increasing interest in studying cooperative relaying because of its promising ability to widen the coverage of wireless networks [1]. In scenarios where the direct communications between a source node (S) and destination node (D) are highly unreliable because of excessive path-loss, serial relaying techniques can be implemented to enhance the network connectivity. Consequently, signals from S to D propagate through a number of shorter hops/links that are less prone to signal degradation. In this context, multi-hop communication constitutes an attractive solution that has been recently investigated in several practical systems including Internet-of-Things (IoT), internet of vehicles, satellite, military systems and unmanned aerial vehicle (UAV) [2], [3]. Either regenerative decode-and-forward (DF) relaying or non-regenerative amplify-and-forward (AF) relaying can be implemented at the cooperating relays. For both relaying techniques, the end-to-end outage performance is mostly dependent on the weakest of all hops; i.e. the hop that suffers from the highest outage probability [4].

Relaying techniques evolved from buffer-free (BF) to buffer-aided (BA) relaying where relays are equipped with buffer (or data queues) that can temporarily store information packets until the links are available for transmission. This feature allowed for an additional degree of freedom that allows to mitigate channel fading and, hence, enhance the reliability of communications at the expense of introducing queuing delays [5], [6]. The existing literature focused primarily on the performance of parallel half-duplex (HD) DF BA relaying networks with the main problem of relay selection in these systems [7]–[11]. For such networks, one of the K neighboring relays is selected to relay the information from S to D involving a source-to-relay transmission phase and a subsequent relay-to-destination transmission phase. Enabling data buffering at the relays implies that the two transmission phases can take place over two non-consecutive time slots. As such, the relay selection protocol revolves around selecting a single relay to either receive (from S) or transmit (to D) over a certain time slot. The policy regulating the relay selection can be either based solely on the channel state information (CSI) as adopted in [7], [8] or on both the CSI and buffer state information (BSI) as in [9]–[11]. For the latter category of protocols, the numbers of packets stored in the relays' buffers are included in the relay selection process thus improving the system performance by avoiding the congestion and starvation of the buffers.

Relaying the information serially through K relays involves $(K+1)$-hop transmissions where the information packets from S are transmitted successively from one relay to another until they reach D. While the research on BA parallel relaying is extensive, BA serial relaying was less investigated in the literature [12]–[19]. Denote the relays by $\text{R}_1, \dots, \text{R}_K$, the source by R_0 and the destination by R_{K+1}. The serial relaying procedure revolves around selecting which link, among the $K+1$ links R_0-R_1, $\dots$, R_K-R_{K+1}, must be chosen for activation in each time slot where the transmission is limited to a single HD node in order to avoid interference. *Max-link* selection was analyzed in [12]–[14] where the link with the highest instantaneous signal-to-noise ratio (SNR) is selected. The selection is limited to the set of available links where the link R_{k-1}-R_k is available when the buffer at R_{k-1} is not empty, i.e. has at least one packet of information to be transmitted, and the buffer at R_k is not full, i.e. can accept at least one packet. In [12], it was possible to derive a lower-bound on the bit error rate (BER) by loosening the aforementioned availability constraint and assuming that the relays buffers have an infinite size and that each relay is always able to transmit a packet. The derived lower-bound has a diversity order of $K+1$ under independent and identically distributed Rayleigh fading with the same path-loss assumed along all hops. In [13], the BER and outage probability of the

The authors are with the Department of Electrical and Computer Engineering of the Lebanese American University (LAU), Byblos, Lebanon. (e-mails: sawsan.elzahr@lau.edu and chadi.abourjeily@lau.edu.lb).

max-link scheme were analyzed. Results proved that the full diversity order of $K+1$ is exclusively achievable with infinite buffer sizes while practical finite-size buffers can achieve only a fraction of this maximum diversity order. A Markov chain (MC) analysis was used to evaluate the outage and delay performance of the *max-link* scheme in [14]. Results are consistent with [13] where large buffer sizes are needed to extract the maximum diversity advantage from the underlying network. Moreover, results showed that the asymptotic average delay increases with the buffer size. As such, keeping the delay at acceptable levels incurs the implementation of buffers with small sizes at the relays which, in turn, reduces the diversity gain that can be reaped from the BA system. The outage performance of the *max-link* scheme was improved in [15] assuming the absence of inter-relay-interference (IRI). This strict assumption holds if perfect IRI cancellation techniques are implemented or if highly-directive antennas are deployed. Neglecting the IRI, the relaying protocol in [15] allows for the simultaneous transmissions along two hops that are selected from two groups out of the total number of three groups in which the hops are partitioned. As in the *max-link* scheme, the instantaneous SNR controls the selection of a node.

HD BA DF serial relaying was tackled in [16] with infinite buffer sizes. The authors in [16] targeted the maximization of the average rate for a communication session extended over an infinite number of fading blocks in a block fading environment. The implementation of the relaying protocol in [16] necessitates the availability of perfect instantaneous CSI at each transmitting node (source or relay) so that this node adapts its transmission rate to the underlying channel conditions where Gaussian codebooks are employed. As in [12]–[15], the relaying protocol in [16] is based solely on the CSI where the transmission modes are related to the rates that can be achieved over the different hops. While [12]–[16] considered HD relaying, full-duplex (FD) serial relaying was considered in [17]–[19] where the relays can transmit and receive at the same time and in the same frequency band. FD BA DF two-hop relaying was considered in [17] where, in a way that is analogous to [16], the relaying protocol revolved around maximizing the transmission rate over an infinite number of time slots based on the instantaneous SNRs along the two constituent hops. The buffer at the relay is assumed to be sufficiently large so that the incoming data can always be stored with no overflow. FD BA DF serial relaying was also studied in [18] and [19] in the context of millimeter-wave (mm-wave) and free-space optical (FSO) communications, respectively. While the self-interference (SI) impairment was taken into consideration in [16], SI and IRI can be neglected for mm-wave and FSO communications since the mm-wave and laser beams are highly directive. As such, the relaying protocols in [18], [19] take into consideration that concurrent transmissions can take place along all hops in the absence of interference.

In this paper, we suggest a novel BA relaying protocol for serial HD communications with any number of DF relays. The proposed strategy improves on the existing schemes [12]–[16] by including the buffer state in the link selection process. This approach results in the following advantages. (i): Achieving the maximum diversity order of $K+1$ with finite-size buffers whose storage capacity does not exceed five packets. (ii): Achieving acceptable levels of the queuing delays that do not increase with the buffer size. (iii): Achieving multiple levels of tradeoff between the average packet delay (APD) and outage probability (OP) by including a variable parameter in the link selection process. Unlike the *max-link* scheme [12]–[14] where increasing the buffer size improves the diversity order at the drawback of increasing the APD, such compromises on the buffer size are not needed for the proposed scheme that can reap the full capabilities of the multi-hop network with very small buffer sizes. As such, the proposed scheme relaxes the reported need to deploy very large buffer sizes at the relays of a serial system by relaxing the full dependence of the existing relaying protocols on the strengths of the hops as in [12]–[16]. We adopt a MC analysis to evaluate the performance of the proposed scheme and to study the effect of buffer size L, number of relays K and algorithm-controlling parameter s on the triad of APD, OP and diversity order. Since the total number of states of the MC increases exponentially with K, we also present an asymptotic analysis that focuses on the most probable states for large values of the SNR. The asymptotic analysis yields accurate closed-form expressions of the asymptotic APD and OP over Rayleigh block fading channels. From this asymptotic analysis, conclusions about the selection of the values of s and L are reached.

The contributions of this research paper are clarified in three-fold:

- We suggest a novel relaying scheme for HD BA DF multi-hop communications.
- We analyze the proposed scheme in a theoretical framework using a MC formulation.
- We derive closed-form analytical expressions for the APD, OP and diversity order.

This paper is structured as follows. The system model and the relaying strategy are presented in Section II. The performance analysis of the proposed relaying scheme is detailed in Section III including an asymptotic analysis and some observations about the system design. Section IV presents the numerical results and a demonstration of the main features of the system. Finally, Section V presents a conclusion of the paper.

II. System Model and Relaying Strategy

A. Basic Parameters

The system consists of a serial relaying network that involves $K+2$ nodes comprising K relay nodes denoted by R_1,. . ., R_K, a source node S and a destination node D. Because of possible long distances between S and D, the assumption of no direct link between S and D is valid and, consequently, a packet is transmitted from S to D in $K+1$ hops through the relays R_1 to R_K as depicted in Fig. 1. We denote S and D by R_0 and R_{K+1}, respectively, and we assume that each relay R_k can transmit a packet to the next relay R_{k+1} (if any). We assume that each node is equipped with only one antenna and that all nodes are HD which implies that simultaneous transmission and reception is impossible.

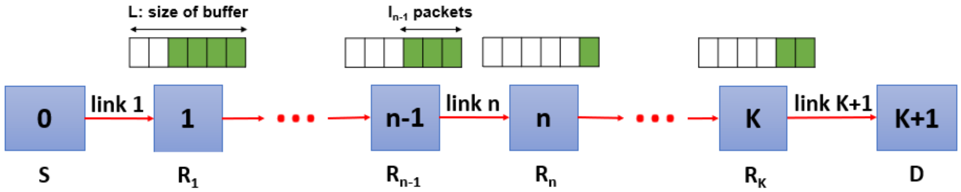

Fig. 1. System Model

The considered system model has the potential of broadening the network coverage and maintaining high spectral efficiency. The dissection of large communication distances into multi-hops allows to mitigate the fading effects and to reduce the outage when the transmitter's energy is limited. Practical examples include unmanned aerial vehicle (UAV) relaying networks where multiple UAVs are employed as relay nodes to forward information packets from a source to a destination [20]. Multi-hop relaying is critical for such applications in order to minimize the delays and improve the system reliability. Other applications include wireless backhauling of fifth generation (5G) small cells where a Macro cell Base Station (MBS) needs to send packets to a far Small cell Base Station (SBS). In this case, closer SBSs will forward packets sequentially in short range hops until these packets reach the target SBS [21].

In what follows, we assume a Rayleigh block fading channel and we indicate by h_k the channel coefficient of the k-th link between nodes R_{k-1} and R_k for $k = 1, \dots, K + 1$ as presented in Fig. 1. The channel coefficients are circularly symmetric complex Gaussian distributed random variables with zero mean and variances. We denote the variances by Ω_k for the k-th link. Finally, all links experience an additive white Gaussian noise (AWGN) that has zero mean and unit variance.

Each link is considered in outage when its corresponding channel capacity is lower than the targeted rate r_0 (in bits per channel use (BPCU)). Consequently, the outage probability along the k-th link is calculated as:

$$p_k = \Pr\left\{\frac{1}{K+1}\log_2(1+\bar{\gamma}|h_k|^2) \le r_0\right\} = 1 - e^{-\frac{2^{(K+1)r_0}-1}{\Omega_k\bar{\gamma}}}, \tag{1}$$

where $\bar{\gamma}$ denotes the average transmit signal-to-noise ratio (SNR). In (1), the division by $K + 1$ is introduced since the communication of a packet from S to D is performed in $K+1$ time slots.

It is assumed that each relay is equipped with one buffer of finite size L, which allows for temporarily storing the packets until better channel quality is available. We denote by $l_k \in \{0, \dots, L\}$ the actual amount of stored packets present in the buffer B_k at R_k for $k = 1, \dots, K$.

The unavailability probability of the k-th link is denoted by q_k. Three cases arise:

- Consider the first hop between S and R_1. The link 1 is judged unavailable if the buffer B_1 is full (cannot accommodate for an incoming packet) or the channel S-R_1 is in outage (that is with probability p_1).
- Consider the last hop between R_K and D. The link $K+1$ is unavailable if the buffer B_K is empty (no packet can be communicated to D) or the channel R_K-D is in outage (that is with probability p_{K+1}).
- Consider an intermediate hop between R_{k-1} and R_k. The link k, for $k = 2, \dots, K$, is unavailable if the buffer B_k is full or the buffer B_{K-1} is empty or the channel between R_{k-1} and R_k is in outage (that is with probability p_k).

Consequently, the unavailability probabilities $\{q_k\}_{k=1}^{K+1}$ can be expressed as:

$$q_k(l_1, \dots, l_K) = p_k + (1 - p_k)\times \begin{cases} \delta_{l_1=L}, & k = 1; \\ \delta_{l_{k-1}=0} + \delta_{l_k=L} - \delta_{l_{k-1}=0}\delta_{l_k=L}, & k = 2, \dots, K; , \\ \delta_{l_K=0}, & k = K + 1. \end{cases} \tag{2}$$

where δ_S is either equal to 1 if the statement S is true or equal to 0 if S is false.

B. Buffer State Based Relaying Strategy

To reap the maximum performance gains from the underlying serial system, the proposed relaying strategy will be based on the availabilities of the $K + 1$ links as well as the buffers' states captured by the vector $(l_1, \dots, l_K)$ representing the current number of packets stored in the K buffers $B_1, \dots,$ B_K. At each time interval, the relaying strategy determines the link $\hat{k}$ that must be activated as follows:

$$\hat{k} = \arg\max_{k\in\mathcal{L}_a}\{\Delta_k\}, \tag{3}$$

denoting that, in the corresponding time slot, $R_{\hat{k}-1}$ must transmit and $R_{\hat{k}}$ must receive. In (3), $\mathcal{L}_a \subset \{1, \dots, K + 1\}$ denotes the set of links that are available and Δ_k denotes the weight that is assigned to link k for $k = 1, \dots, K + 1$.

The relaying strategy that we propose in this work is based on defining the weights $\{\Delta_k\}_{k=1}^{K+1}$ as in (4).

$$\Delta_k = \begin{cases} s, & k = 1; \\ l_{k-1}, & k = 2, \dots, K + 1. \end{cases} . \tag{4}$$

The rationale behind (4) is as follows. In the goal of avoiding the excessive queuing of the packets at the relays' buffers which negatively impacts the queuing delay, the proposed strategy corresponds to the selection of the relay whose buffer is storing the highest number of packets at the transmitting node. Evidently, the selection is limited among the relays whose links with the subsequent relay (or D) is available since, otherwise, no packet can be successfully communicated along the link that must be activated. While the weight associated with each one of the K relays is determined from the number of packets stored in this relay's finite-size buffer, a distinct weight s is assigned to S (i.e. link 1). Note that the source is assumed to be equipped with an infinite size buffer and to be fully backlogged, i.e., it always has enough information packets to be transmitted. These assumptions are common in the open literature on BA relaying [7]–[19]. It is worth highlighting that any link that is not in outage will ensure the delivery of a packet from the transmitting node to the receiving node. In other words, if a link is stronger than another link while both links are not in outage, then there is no added value in activating the stronger link since, in both cases, the objective of successfully transmitting the dequeued packet is realized. As such, referring to (1), there is no need to include the explicit value of the link capacity in the link selection process as long as this capacity is above the threshold value.

The nonzero parameter s will be restricted to the set $\{1,\ldots,L\}$ to have comparable values with the buffer lengths $l_1,\ldots,l_K$. As will be highlighted later, the parameter s has a major impact on the achievable diversity orders and queuing delays and the subsequent performance analysis will suggest convenient options for selecting this parameter. Finally, in (3), in case multiple links share the same maximum weight, the highest order link (i.e. the closest link to D) will be selected. This selection will accelerate the arrival of packets at D and, thus, will contribute to reducing the average packet delay.

Given that each node in multi-hop networks can communicate only with the preceding and subsequent nodes, the signaling protocol in such networks differs from that implemented in parallel-relaying networks where S (or D) can broadcast signaling information to all relays. However, for the proposed relaying scheme, the decision on the link to be selected can be implemented in an advantageously simple sequential manner. In fact, instead of collecting all buffer state and channel state information and sharing it with a central node that makes a decision on the selected link, every relay can make an intermediate decision on whether this relay or the subsequent relay (if any) is better suited for transmission. The intermediate decision along with the corresponding recursive weight can be shared with the previous relay sequentially until the signaling information reaches S that makes the final decision on the selected link as follows:

- Starting from R_K, relay R_k performs the following tasks for $k = K,\ldots,1$. (i): It generates the recursive weight r_k and the index i_k of the link that is the best candidate so far. (ii): It shares the metrics (r_k, i_k) with the previous relay R_{k-1}.
- After K signaling time slots, S (relay R_0) receives (r_1, i_1) and generates the metrics (r_0, i_0). The integer i_0 will be equal to the index of the best link $\hat{k}$ in (3).
- The index i_0 then needs to be shared with the K relays over K additional signaling time slots. Starting from R_0, relay R_k shares the value of i_0 with R_{k+1} for $k = 0,\ldots,K-1$.
- Consequently, all nodes $R_0,\ldots,R_K$ have acquired the index of the best link and the corresponding node (if any) can initiate the data transmission.

For $k = K,\ldots,0$, r_k and i_k can be determined recursively as follows:

$$\begin{aligned} r_k &= \max\{\Delta_{k+1}s_{k+1}, r_{k+1}\} \\ i_k &= \begin{cases} k+1, & \Delta_{k+1}s_{k+1} > r_{k+1}; \\ i_{k+1}, & \text{otherwise.} \end{cases} \end{aligned} \quad , \tag{5}$$

where $r_{K+1} = 0$, $i_{K+1} = 0$, Δ_k is given in (4) and $s_k = 0$ (resp. $s_k = 1$) if link k is in outage (resp. not in outage). Note that, $\Delta_{k+1}s_{k+1} = 0$ if either $\Delta_{k+1} = 0$ (i.e. the buffer at R_k is empty) or $s_{k+1} = 0$ (i.e. the link $k+1$ over which R_k transmits is in outage). In both cases, link $k+1$ is unavailable and it cannot be selected as the best candidate link. Finally, $i_0 = 0$ means that all $K+1$ hops are unavailable and the network is in outage.

Comparing the signaling overhead with that of the parallel-relaying BA scheme in [11]:

- In the downlink (from relays to S), the proposed scheme requires the transmission of K messages of length $d_1 = \lceil \log_2((K+1)(L+1)) \rceil$ in K consecutive time slots since $\{(r_k, i_k)\}_{k=1}^{K} \subset \{0,\ldots,L\} \times \{0,2,\ldots,K+1\}$. Similarly, the scheme in [11] necessitates the transmission of K messages over K time slots but now the length of each message is $d_2 = \lceil \log_2(4(L+1)) \rceil$ where the factor four captures the joint availabilities of the links S-R_k and R_k-D. Note that, even with parallel-relaying, K distinct signaling time slots are needed since the K relays cannot transmit simultaneously in order to avoid interference.
- In the uplink (from S to relays), K messages of length $u_1 = \lceil \log_2(K+2) \rceil$ must be transmitted over K time slots to inform the relays which one of the $K+1$ hops must be activated (in addition to the option that all nodes must remain idle). For [11], one message of length $u_2 = \lceil \log_2(2K+1) \rceil$ can be broadcasted from S to inform all relays on the node to be selected (if any) and on whether the selected relay should transmit or receive.

As such, except for the incapability of broadcasting in any multi-hop network, the signaling overheads of the proposed scheme and [11] are comparable especially for practical systems comprising a limited number of relays K.

The proposed scheme is appealing from a signaling-overhead point of view for the following reasons. (1): The proposed scheme can be implemented with small buffer sizes which limits the portion of the signaling overhead pertaining to the buffer state information. In particular, we prove in the next section that there is no need to deploy buffers whose sizes exceed five. As such, an immaterial number of $d_1 = 6$ bits in the downlink can accommodate a network with up to nine relays. Therefore, the cost of collecting the buffer state information is not overwhelming. (2): The signaling overhead needs to indicate simply whether the links are in outage or not through the variable s_k in (5). Moreover, this variable can be further multiplied by the weight Δ_k since there is no need to report the number of stored packets if the corresponding link is in outage since this link will not be selected. As such, unlike the benchmark *max-link* scheme in [14], the proposed scheme does not include the actual values of the path gains $\{h_k\}_{k=1}^{K+1}$ in the decision-making process. In fact, feeding back the $K+1$ gains results in excessively long signaling messages if the real-valued path gains are to be quantized with a sufficiently high level of accuracy. In this context, it is worth highlighting that the signaling of the *max-link* scheme can be implemented sequentially as well. While the number of messages and the length of each message in the uplink remains the same as compared to our proposed scheme, the length of each one of the K messages in the downlink must increase from $d_1 = \lceil \log_2((K+1)(L+1)) \rceil$ to $\lceil \log_2((K+1)M) \rceil$ where M is the number of quantization levels that exceeds $L+1$ (whose maximum value is 6 with the proposed scheme) by several orders of magnitude.

III. Performance Analysis

A. Generalities

We will adopt a Markov Chain (MC) analysis to study the behavior of the BA serial relaying system where the features of interest are the outage probability (OP) and the average packet delay (APD). We define a state as the mixture of the current amount of stored packets in all buffers and will be denoted by $(l_1, l_2, \ldots, l_K)$. The number of states in this MC is $(L+1)^K$ in total since $l_k \in \{0, \ldots, L\}$ for $k = 1, \ldots, K$. In this work, we consider a finite buffer size L that yields a finite-state MC. This choice is motivated by the fact that infinite-size buffers are not practical since all storage devices have a finite capacity. Moreover, as will be discussed in Section III-D, we prove that the proposed BA relaying scheme is capable of extracting the full capabilities of the cooperative network with buffers having a finite size of five. As such, the use of infinite-size buffers and finite-size buffers with $L > 5$ is not justified since such options do not enhance the asymptotic performance gains.

The transition probability of going from the state $(l_1, \ldots, l_K)$ to the state $(l'_1, \ldots, l'_K)$ is denoted by $t_{(l_1,\ldots,l_K),(l'_1,\ldots,l'_K)}$. The transition matrix $\mathbf{T}$ of size $(L+1)^K \times (L+1)^K$ describes the evolution between the states. The (i,j)-th element of $\mathbf{T}$ is given by:

$$\mathbf{T}_{i,j} = t_{(l_1,\ldots,l_K),(l'_1,\ldots,l'_K)} \quad ; \quad i = \mathfrak{N}(l'_1, \ldots, l'_K) \ , \ j = \mathfrak{N}(l_1, \ldots, l_K), \tag{6}$$

where $j = \mathfrak{N}(l_1, \ldots, l_K)$ is the one-to-one function relating the integer $j \in \{1, \ldots, (L+1)^K\}$ and the state $(l_1, \ldots, l_K) \in \{0, \ldots, L\}^K$ and is expressed as: $j = 1 + \sum_{k=1}^{K} l_k (L+1)^{K-k}$.

We denote by $\pi_{l_1,\ldots,l_K}$ the steady-state probability of the system being in the state $(l_1, \ldots, l_K)$. These steady-state probabilities are calculated as follows [7]:

$$\pi = (\mathbf{T} + \mathbf{B} - \mathbf{I})^{-1} \mathbf{b}, \tag{7}$$

where π is a $(L+1)^K$-dimensional vector and its j-th element is equal to $\pi_{l_1,\ldots,l_K}$ with $j = \mathfrak{N}(l_1, \ldots, l_K)$. In (7), $\mathbf{B}$ and $\mathbf{I}$ are two matrices of size $(L+1)^K \times (L+1)^K$ that denote the all-one matrix and the identity matrix, respectively. $\mathbf{b}$ is a vector with $(L+1)^K$ elements all equal to 1.

The system is defined to be in outage only if none of its $K+1$ links can be activated, that is no packets can be transmitted along any of these links. Hence, for a given state $(l_1, \ldots, l_K)$ an outage occurs with the probability $\prod_{k=1}^{K+1} q_k(l_1, \ldots, l_K)$ following from (2). The steady state probabilities allow then the calculation of the outage probability as follows::

$$OP = \sum_{l_1=0}^{L} \cdots \sum_{l_K=0}^{L} \pi_{l_1,\ldots,l_K} \prod_{k=1}^{K+1} q_k(l_1, \ldots, l_K). \tag{8}$$

The queuing at the relays' buffers will imply a delay in the arrival of the packets to D. The average packet delay is formulated following from [14] and Little's law [22]:

$$APD = \frac{K + OP + (K+1)\bar{L}}{1 - OP}, \tag{9}$$

where the term $\bar{L}$ is denoting the average queue length of the buffers and is obtained as follows:

$$\bar{L} = \sum_{l_1=0}^{L} \cdots \sum_{l_K=0}^{L} \pi_{l_1,\ldots,l_K} \left[\sum_{k=1}^{K} l_k \right]. \tag{10}$$

It is worth highlighting that the presented MC analysis holds for multi-hop networks with any number $K \geq 1$ relays. For single-hop networks (i.e. $K = 0$), the MC framework is not needed since the network comprises only one link with no relays' buffers.

B. State Transition Matrix

In what follows, the unavailability probabilities in (2) will be written as q_k for simplicity. We will denote the state by $\mathbf{l} = (l_1, \ldots, l_K)$, the set of all relays by $\mathcal{A} = \{1, \ldots, K\}$ and $\mathbf{e}_k$ will denote the k-th row of the $K \times K$ identity matrix.

The self transition at any state $\mathbf{l}$ of the MC occurs only if all links are unavailable:

$$t_{\mathbf{l},\mathbf{l}} = \prod_{k=1}^{K+1} q_k. \tag{11}$$

To transit to another state, at least one link should be activated. Denote by a_k the probability of activating the link k for $k = 1, \ldots, K+1$. This probability is mapped to the transition probabilities as follows:

$$a_k = \begin{cases} t_{\mathbf{l},\mathbf{l}+e_k}, & k = 1; \\ t_{\mathbf{l},\mathbf{l}+e_k-e_{k-1}}, & k = 2, \ldots, K; \\ t_{\mathbf{l},\mathbf{l}-e_{k-1}}, & k = K+1. \end{cases} \tag{12}$$

A link k is selected to be activated if it is available and its weight Δ_k is the highest among all other available links:

$$a_k = (1 - q_k) \sum_{\mathcal{K} \subset \mathcal{A} \setminus \{k\}} \left[\prod_{i \in \mathcal{K}} (1 - q_i) \right] \left[\prod_{j \in \mathcal{A} \setminus \{k\} \cup \mathcal{K}} q_j \right] Q_{k,\mathcal{K}}, \tag{13}$$

where the set $\mathcal{K}$ comprises the indices of the links, other than the link k, that are available. In (13), $Q_{k,\mathcal{K}}$ designates the probability that Δ_k is *greater* than $\Delta_{k'}$ for all $k' \in \mathcal{K}$. We emphasize on the concept of *larger* Δ_k that considers the tie breaking rule following from the numbering of links based on their distances from S. As such, $Q_{k,\mathcal{K}} = \prod_{k' \in \mathcal{K}} Q_{k,k'}$ where $Q_{k,k'}$ denotes the probability that Δ_k is *greater* than $\Delta_{k'}$:

$$Q_{k,k'} = \delta_{k'<k} \delta_{\Delta_k \geq \Delta_{k'}} + \delta_{k'>k} \delta_{\Delta_k > \Delta_{k'}} \quad ; \quad k' \neq k, \tag{14}$$

$$a_k = (1-q_k)\left[\prod_{i=1, i\neq k}^{K+1} q_i + \sum_{k_1=1, k_1\neq k}^{K+1}(1-q_{k_1})\left[\prod_{j=1, j\neq k, j\neq k_1}^{K+1} q_j\right] Q_{k,k_1} + \sum_{k_1=1,k_1\neq k}^{K+1}\sum_{k_2=k_1+1,k_2\neq k}^{K+1}(1-q_{k_1})(1-q_{k_2})\left[\prod_{j=1,j\neq k,j\neq k_1,j\neq k_2}^{K+1} q_j\right] Q_{k,k_1}Q_{k,k_2} + \cdots\right]. \quad (15)$$

$$a_k = (1-q_k)\left[\prod_{i=1,i\neq k}^{K+1} q_i\right]\left[1+\sum_{k_1=1,k_1\neq k}^{K+1}\frac{(1-q_{k_1})}{q_{k_1}}Q_{k,k_1}\left[1+\sum_{k_2=k_1+1,k_2\neq k}^{K+1}\frac{(1-q_{k_2})}{q_{k_2}}Q_{k,k_2}\left[1+\ldots\left[1+\sum_{k_K=k_{K-1}+1,k_K\neq k}^{K+1}\frac{(1-q_{k_K})}{q_{k_K}}Q_{k,k_K}\right]\right]\right]\right], \quad (16)$$

since, for $\Delta_k = \Delta_{k'}$, it is preferred to activate the link that is farther from S that is having the higher index.

Equation (13) can be developed as (15) on the top of the page that further simplifies into (16) on the top of the page.

Equation (16) can be implemented recursively resulting in the following expression:

$$a_k = (1-q_k)\left[\prod_{i=1,i\neq k}^{K+1} q_i\right][1+f_r(\mathcal{A},k,0)], \quad (17)$$

where $f_r(\cdot,\cdot,\cdot)$ is the recursive function that can be derived using algorithm 1.

```
Function: f_r(Y, k, a)
Data: Y ⊂ A, k ∈ {1, ..., K+1} and
      a ∈ {0, ..., K+1};
Result: Sum;
initialization: Sum = 0;
if a + 1 > |Y| then
  | return 0
end
for m = a + 1 : |Y| do
  | k' = Y_m (m-th element of Y)
  | if k' ≠ k then
  |   | Sum = Sum + ((1 - q_k')/q_k') Q_{k,k'} [1 + f_r(Y, k, m)]
  | end
end
```

Algorithm 1: Recursive function $f_r(\mathcal{Y},k,a)$

C. Asymptotic Analysis

Using (17) to evaluate the transition probabilities in (12) then stacking these probabilities in the state transition matrix to determine the steady-state probabilities in (7) does not yield tractable expressions of the OP and APD especially when K and/or L are large. This observation follows from (i): the complexity of the recursive function in (17), (ii): the large number of states that can a state $\mathbf{l}$ transit to according to (12) and (iii): the need to invert a $(L+1)^K \times (L+1)^K$ matrix in (7) where there is an exponential increase of the number of states with the number of relays K. As such, we next resort to an asymptotic analysis that holds $\bar{\gamma} \gg 1$. This analysis yields tractable closed-form expressions of the APD and OP in the asymptotic regime and allows to draw useful conclusion about the system performance.

The $(L+1)^K$ states of the state space $\mathcal{S} \triangleq \{0,\ldots,L\}^K$ will no longer be considered in the steady-state probability calculations, instead, the asymptotic analysis will focus on a subset $\mathcal{S}_c$ of $\mathcal{S}$ where this subset comprises a much smaller number of states and where the MC is in $\mathcal{S}_c$ with a probability tending to one asymptotically. In other words, $\sum_{\mathbf{l}\in\mathcal{S}_c}\pi_{\mathbf{l}} \to 1$ while $\pi_{\mathbf{l}} \to 0\ \forall\ \mathbf{l} \notin \mathcal{S}_c$ for $\bar{\gamma} \gg 1$ where the steady-state probabilities satisfy (7). The set $\mathcal{S}_c$ is called the closed subset where the probability of exiting this set tends to zero asymptotically:

$$t_{\mathbf{l},\mathbf{l}'} \to 0\ \ \forall\ \ \mathbf{l}\in\mathcal{S}_c\ ,\ \mathbf{l}'\notin\mathcal{S}_c. \quad (18)$$

The performed asymptotic analysis shows that the closed subset $\mathcal{S}_c$ and the corresponding steady-state probabilities depend on the weight s of link 1 in (4). In particular, the cases $1 < s < L$, $s = 1$ and $s = L$ need to be considered separately. For the sake of notational simplicity, the following definitions of some states that depend on the parameter s are introduced:

$$\begin{aligned}
\mathbf{s}_1^{(1)} &= (s-1,\ldots,s-1),\\
\mathbf{s}_2^{(1)} &= (s+1,s-1,\ldots,s-1),\\
\mathbf{s}_3^{(1)} &= (s-1,\ldots,s-1,s-2),\\
\mathbf{s}_n^{(2)} &= (\underbrace{s-1,\ldots,s-1}_{n-1\ \text{times}},s,\underbrace{s-1,\ldots,s-1}_{K-n\ \text{times}}),\\
\mathbf{s}_n^{(3)} &= (s,\underbrace{s-1,\ldots,s-1}_{n-1\ \text{times}},s,\underbrace{s-1,\ldots,s-1}_{K-n-1\ \text{times}}),\\
\mathbf{s}_n^{(4)} &= (\underbrace{s-1,\ldots,s-1}_{n-1\ \text{times}},s,\underbrace{s-1,\ldots,s-1}_{K-n-1\ \text{times}},s-2),
\end{aligned} \quad (19)$$

with $n=1,\ldots,K$ for $\mathbf{s}_n^{(2)}$ and $n=1,\ldots,K-1$ for $(\mathbf{s}_n^{(3)},\mathbf{s}_n^{(4)})$.

1) Case 1: $1 < s < L$:

Proposition *1:* For $1 < s < L$, the closed subset comprises $3K+1$ states as follows:

$$\mathcal{S}_c = \{\mathbf{s}_n^{(1)} \ ; \ n = 1,2,3\} \cup \{\mathbf{s}_n^{(2)} \ ; \ n = 1,\dots,K\} \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} \ ; \ n = 1,\dots,K-1\}, \quad (20)$$

where the corresponding steady-state probabilities are given by:

$$\begin{cases} \pi_{\mathbf{s}_1^{(1)}} = \frac{1-\sum_{k=2}^{K+1} p_k}{K+1} \\ \pi_{\mathbf{s}_2^{(1)}} = \frac{p_2}{K+1} \\ \pi_{\mathbf{s}_3^{(1)}} = \frac{p_1}{K+1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1-\sum_{k=1}^{n} p_k}{K+1}, & \text{for } n = 1,\dots,K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=2}^{n+1} p_k}{K+1}, & \text{for } n = 1,\dots,K-1 \\ \pi_{\mathbf{s}_n^{(4)}} = \frac{p_1}{K+1}, & \text{for } n = 1,\dots,K-1. \end{cases} \quad (21)$$

Proof: The above proposition is proved in Appendix A. ■

2) Case 2: $s = 1$:

Proposition *2:* For $s = 1$, the closed subset comprises $2K+1$ elements:

$$\mathcal{S}_c = \{\mathbf{s}_n^{(1)} \ ; \ n = 1,2\} \cup \{\mathbf{s}_n^{(2)} \ ; \ n = 1,\dots,K\} \cup \{\mathbf{s}_n^{(3)} \ ; \ n = 1,\dots,K-1\}, \quad (22)$$

with the following steady-state probabilities:

$$\begin{cases} \pi_{\mathbf{s}_1^{(1)}} = \frac{1-\sum_{k=2}^{K+1} p_k}{K+1-Kp_1} \\ \pi_{\mathbf{s}_2^{(1)}} = \frac{p_2}{K+1-Kp_1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1-\sum_{k=1}^{n} p_k}{K+1-Kp_1}, & \text{for } n = 1,\dots,K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=2}^{n+1} p_k}{K+1-Kp_1}, & \text{for } n = 1,\dots,K-1. \end{cases} \quad (23)$$

Proof: The details of the proof are presented in Appendix B. ■

3) Case 3: $s = L$:

Proposition *3:* For $s = L$, the $3K$-element closed subset along with the steady state probabilities are presented below:

$$\mathcal{S}_c = \{\mathbf{s}_1^{(1)}, \mathbf{s}_3^{(1)}\} \cup \{\mathbf{s}_n^{(2)} \ ; \ n = 1,\dots,K\} \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} \ ; \ n = 1,\dots,K-1\}, \quad (24)$$

$$\begin{cases} \pi_{\mathbf{s}_1^{(1)}} = \frac{1-\sum_{k=3}^{K+1} p_k}{K+1} \\ \pi_{\mathbf{s}_3^{(1)}} = \frac{p_1}{K+1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1-\sum_{k=1}^{n} p_k}{K+1}, & \text{for } n = 1,\dots,K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=3}^{n+1} p_k}{K+1}, & \text{for } n = 1,\dots,K-1 \\ \pi_{\mathbf{s}_n^{(4)}} = \frac{p_1+p_2}{K+1}, & \text{for } n = 1,\dots,K-1. \end{cases} \quad (25)$$

Proof: The details of the proof are presented in Appendix C. ■

In (21), (23) and (25), the terms comprising the product of two or more outage probabilities among $\{p_k\}_{k=1}^{K}$ are ignored since these terms are small for large values of the SNR. It is obvious that the probabilities in (21) add up to one. The same holds for the probabilities in (23) and the probabilities in (25).

Assuming $L \geq 5$ and replacing (21), (23) and (25) in (8) implies the expressions of the asymptotic OP provided in (26) on the top of the next page.

It is worth highlighting that the asymptotically-dominant states in (19) comprise the buffer lengths $s-2$, $s-1$, s and $s+1$. Therefore, the cases $s = 2$, $s = 1$, $s = L$ and $s = L-1$ need to be considered separately in the OP derivations. In fact, for $s \in \{1, 2\}$ (resp. $s \in \{L-1, L\}$) some buffers are empty (resp. full) and, hence, cannot transmit (resp. receive) packets. In this context, only the case $2 < s < L-1$ implies that the unavailability probabilities in (2) satisfy $q_k = p_k$ for $k = 1,\dots,K+1$ for all states in (19) that determine the closed subset. As an illustration, for $s = 2$, the link $K+1$ is always unavailable for the states in $\mathcal{S}_u = \{\mathbf{s}_3^{(1)}, \mathbf{s}_1^{(4)}, \dots, \mathbf{s}_{K-1}^{(4)}\}$ in (20). Therefore, the product of the unavailability probabilities in (8) simplifies to $\prod_{k=1}^{K+1} q_k = \prod_{k=1}^{K} p_k$ for the states in $\mathcal{S}_u$ and to $\prod_{k=1}^{K+1} q_k = \prod_{k=1}^{K+1} p_k$ otherwise. As such, the asymptotic OP can be written as $OP_{\text{Asymp}} = (1 - \sum_{\mathbf{l}\in\mathcal{S}_u} \pi_{\mathbf{l}}) \prod_{k=1}^{K+1} p_k + (\sum_{\mathbf{l}\in\mathcal{S}_u} \pi_{\mathbf{l}}) \prod_{k=1}^{K} p_k$ which simplifies to the second expression in (26) after replacing the steady-state probabilities by their values from (21) and observing that $\sum_{\mathbf{l}\in\mathcal{S}_u} \pi_{\mathbf{l}} = \frac{p_1}{K+1} + (K-1)\frac{p_1}{K+1} = \frac{Kp_1}{K+1}$.

The asymptotic OP expressions in (26) yield the diversity order of the BA relaying system. The value of the diversity order can be extracted from the OP($\bar{\gamma}$) curve on a log-log scale as the negative slope of this curve. Asymptotically, the product of n terms among $\{p_1, \dots, p_{K+1}\}$ scales as $\bar{\gamma}^{-n}$ (given that each outage probability in (1) scales as $\bar{\gamma}^{-1}$) generating a diversity order of n. Consequently, the system's diversity order d is found to be:

$$d = \begin{cases} 1, & s = 1; \\ K+1, & 1 < s < L; , \\ K, & s = L. \end{cases} \quad (27)$$

implying that the choice $1 < s < L$ is the most appealing for maximizing the diversity order.

For the asymptotic APD derivations, the outage probabilities $\{p_k\}_{k=1}^{K+1}$ can be ignored in evaluating the steady-state probabilities in (21), (23) and (25). In fact, it was observed that this approach yields to a simple asymptotic APD expression that is highly accurate. Setting $p_1 = \ldots = p_{K+1} = 0$ in (21), (23) and (25) results in $\pi_{\mathbf{s}_1^{(1)}} = \pi_{\mathbf{s}_1^{(2)}} = \cdots = \pi_{\mathbf{s}_K^{(2)}} = \frac{1}{K+1}$ for all values of s while other steady-state probabilities can be ignored. Therefore, the average queue length in (10) is equal to $\bar{L} = \frac{1}{K+1}K(s-1) + \frac{K}{K+1}(K(s-1)+1)$ following from the definitions of the states $\mathbf{s}_1^{(1)}$ and $\mathbf{s}_n^{(2)}$ in (19). Replacing this value of $\bar{L}$ in (9) while ignoring OP that is very small asymptotically implies the below expression of the asymptotic APD that holds for all values of s:

$$APD_{\text{Asymp}} = 2K + (s-1)K(K+1). \quad (28)$$

implying that the choice $s = 1$ is the most appealing for minimizing the queuing delay.

$$OP_{\text{Asymp}} = \begin{cases} \frac{1-\sum_{k=2}^{K+1} p_k}{K+1-Kp_1} p_1 + \sum_{n=1}^{K} \frac{1-\sum_{k=1}^{n} p_k}{K+1-Kp_1} p_1 p_{n+1} + \frac{p_2}{K+1-Kp_1} p_1 p_2 + \sum_{n=2}^{K} \frac{\sum_{k=2}^{n+1} p_k}{K+1-Kp_1} p_1 p_2 p_{n+1}, & s=1 \\ \left[1 - \frac{Kp_1}{K+1}\right] \prod_{k=1}^{K+1} p_k + \frac{Kp_1}{K+1} \prod_{k=1}^{K} p_k, & s=2 \\ \prod_{i=1}^{K+1} p_i, & 2 < s < L-1 \\ \left[1 - \frac{p_2}{K+1}\right] \prod_{i=1}^{K+1} p_i + \frac{p_2}{K+1} \prod_{i=2}^{K+1} p_i, & s = L-1 \\ \frac{1 + p_1 - \sum_{k=3}^{K+1} p_k}{K+1} \prod_{i=1}^{K+1} p_i + \sum_{n=1}^{K} \frac{1 - \sum_{k=1}^{n} p_k}{K+1} \prod_{i=1, i\neq n}^{K+1} p_i & \\ \quad + \sum_{n=2}^{K} \frac{\sum_{k=3}^{n+1} p_k}{K+1} \prod_{i=2, i\neq n}^{K+1} p_i + \sum_{n=1}^{K-1} \frac{p_1 + p_2}{K+1} \prod_{i=1, i\neq n}^{K+1} p_i, & s = L. \end{cases} \quad (26)$$

Equation (28) demonstrates that the asymptotic APD increases as the number of relays K is increasing where the delay is accumulated as the information packets move from one relay's buffer to the buffer of the next relay. However, unlike the *max-link* scheme in [14], the asymptotic APD is independent of the buffer size L highlighting on the importance of including the buffer state information in the relaying strategy where the proposed relaying scheme revolves around avoiding the congestion of the relays' buffers.

The MC framework constitutes the broad mathematical tool to analyze queues [5], [7]–[11], [14], [15], [19]. The particularities of the underlying network and the implemented relaying strategy render the MC analysis different from one system to another. It is worth highlighting that the dynamics of the buffers in serial-relaying systems are more complicated compared to parallel-relaying systems as in [11]. In fact, for parallel-relaying, each packet is queued in one and only one relay buffer before being delivered to D. However, for serial-relaying, the packets move from one buffer to another and, hence, each packet will be sequentially queued in all relays' buffers before reaching D. Therefore, the transition probabilities derived in this paper differ substantially from those presented in [11]. The role of the source node S also differs substantially between [11] and the current work. A main challenge in the MC analysis performed in this paper resides in quantifying the role of S via a parameter s that was introduced in the link selection protocol in (4). As such, S has to compete with the relays for transmitting unlike [11]. As demonstrated in the presented performance analysis, the parameter s impacts the closed subset and, hence, three variants of the asymptotic analysis need to be carried out depending on the value of s. Unlike [11] where the closed subset contained only four states for any number of relays K, the asymptotic MC analysis presented in this paper is more challenging for the following reasons. (i): The number of states in the closed subset is not constant since it depends on the parameter s. (ii): The number of states in the closed subset is relatively large and increases with the number of relays K. As such, identifying the closed subset is much more difficult. Moreover, it is tougher to reach the asymptotic steady-state probabilities in equations (21), (23) and (25) (as compared to eq. (44) in [11]) since a larger number of balance equations involving a larger number of variables need to be solved. This also results in more complicated asymptotic OP expressions as can be observed by comparing (26) with eq. (45) in [11].

D. Conclusions about the design of the BA relaying scheme

Following from (26)-(28), we can reach the following conclusions pertaining to the values of the weight s and buffer size L.

- There is no interest in selecting $s > 3$. From (28), such large values of s penalize the APD while not presenting any advantage in terms of the diversity order as can be observed from (27).
- The values $\{1, 2, 3\}$ all constitute valid options for the parameter s thus allowing the proposed relaying scheme to achieve different levels of tradeoff between APD and OP.
- Setting s to 1 is the best choice if the most critical performance metric of a given application is the delay. Consequently, this will guarantee the minimal asymptotic APD value of $2K$ at the drawback of a minimal diversity order of 1. In this case, all values of $L \geq 3$ result in the same levels of the asymptotic APD and OP and, hence, there is no need to deploy buffer sizes exceeding 3 when $s = 1$.
- Setting s to $\{2, 3\}$ constitute the best choices in case the outage is set to be the most critical performance metric.
- Setting s to 2 permits to reach the maximum diversity order of $K + 1$ but with an asymptotic APD of $K(K + 3)$. In this case, the OP and APD performance does not improve by increasing L above 4 and, hence, setting $L = 4$ presents the best choice when s is fixed to 2.
- Setting s to 3 permits to achieve the maximum diversity order of $K + 1$ as well. However, comparing the choices $s = 3$ and $s = 2$, the former choice incurs an increase in the asymptotic APD value to $2K(K + 2)$ with the advantage of reducing the asymptotic OP following from the second and third expressions in (26). Therefore, increasing s from 2 to 3 maintains the same maximum diversity order with the disadvantage of increasing the delay by $K(K + 1)$ and the advantage of a coding

gain of $\frac{10}{K+1}\log_{10}\left(1+\frac{K}{K+1}\frac{\Omega_{K+1}}{\Omega_1}\right)$ decibels. Finally, for $s=3$, the buffer size of $L=5$ is sufficient to reap all the performance gains in the asymptotic regime. In fact, the derivations in Section III-C demonstrated that the probability of having more than five packets stored in any buffer tends to zero asymptotically. As such, there is no need to deploy buffers that can store more than five packets. Note that the delay-loss increases with K while the coding gain decreases with K rendering the choice $s=2$ more adequate to serial relaying systems with a large number of relays.

- Even though finite-size buffers were assumed in this work, the analysis presented in Section III-C with $s\in\{1,2,3\}$ holds for infinite-size buffers as well. In fact, the finite set of recurrent states in (19) will shape the asymptotic steady-state distribution of the MC even with an infinite number of states since all remaining states will be transient.

IV. Numerical Results

We next provide some numerical results supporting the theoretical expressions and conclusions derived in the previous sections. In what follows, r_0 is fixed to 1 BPCU in (1). In addition, we define the $(K+1)$-dimensional vector $\mathbf{\Omega}$ as $\mathbf{\Omega}=[\Omega_1,\ldots,\Omega_{K+1}]$ capturing the strengths of the $K+1$ hops.

Fig. 2 and Fig. 3 present the curves of OP and APD, respectively, for a network of three relays with $L=8$ and $\mathbf{\Omega}=[4,4.5,5,4.5]$. The results in these two figures demonstrate the accuracy of the asymptotic analysis and the validity of the formulated OP and APD asymptotic expressions in (26) and (28), respectively. In fact, for all values of s, the asymptotic and the exact OP and APD curves are perfectly matched for average-to-large values of SNR. Furthermore, the theoretical MC analysis that was performed is proved to be valid since the curves of theoretical OP and APD, from (8) and (9) respectively, are perfectly matched with their numerical counterparts that were generated by Monte Carlo simulations. Consequently, we can deduce the following observations. (i): The choice $s=1$ leads to the highest OP and lowest APD. In fact, assigning a small weight to link 1 privileges the transmissions from relays with non-empty buffers which positively contributes towards reducing the queuing delays. (ii): The choices $s=2$, $s=3$ and $s=L-1$ satisfy the condition $1<s<L$ and, hence, all achieve the maximum diversity order of $K+1$ following from (27). This results in comparable OP performance where the corresponding OP curves are the steepest as can be observed from Fig. 2. However, from Fig. 3, increasing the value of s leads to higher APD values in coherence with (28). Among the above choices, the value $s=3$ results in the smallest OP as predicted from (26); however, the coding gain with respect to the case of $s=2$ is small (around 0.65 dB) and does not justify the increase in the asymptotic APD from 18 to 30. (iii): Selecting $s=L$ results in a higher OP and a higher APD than the case of $1<s<L$ and, consequently, this choice does not present any advantage. As a conclusion, the above observations validate the findings reported in Section III-D.

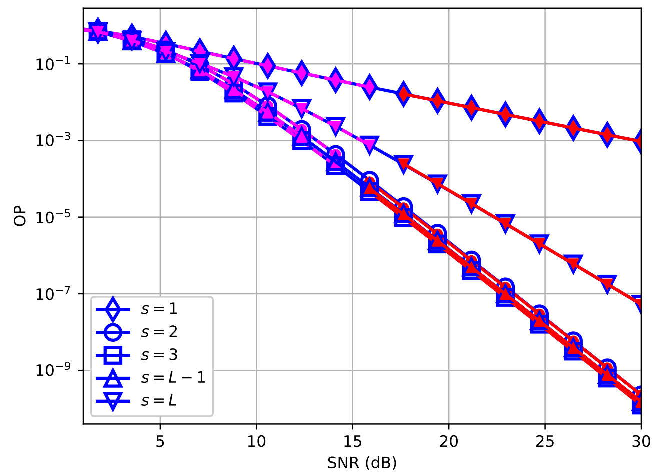

Fig. 2. OP with $K=3$ and $L=8$. Dashed lines, solid lines with hollow markers and solid lines correspond to the simulation, theoretical and asymptotic values, respectively.

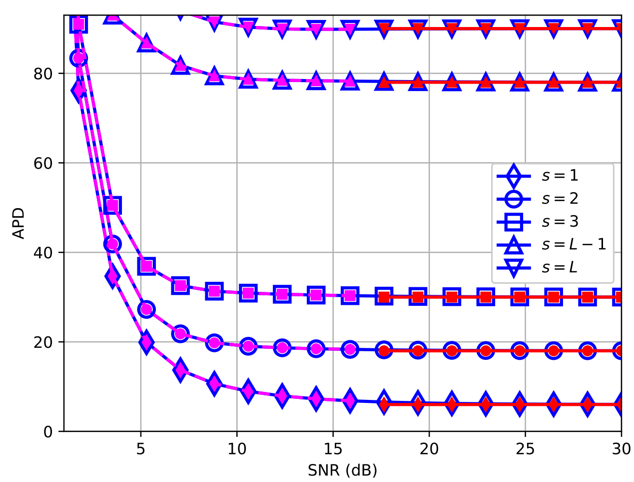

Fig. 3. APD with $K=3$ and $L=8$. Dashed lines, solid lines with hollow markers and solid lines correspond to the simulation, theoretical and asymptotic values, respectively.

Fig. 4 and Fig. 5 make a comparison between the proposed scheme and the *max-link* scheme [14] at a SNR of 35 dB with $L=5$ and $L=10$. These figures show the variations of the OP and APD, respectively, as function of the number of relays K for $\mathbf{\Omega}=[3,\ldots,3]$. For the proposed scheme, we consider the values of s in $\{1,2,3\}$ that constitute the valid values for this parameter following from Section III-D. Results in Fig. 4 and Fig. 5 validate equations (26) and (28), respectively, demonstrating that the asymptotic performance of the proposed scheme is independent of L as long as $L\geq 5$. In fact, the OP and APD curves of the proposed scheme pertaining to the cases $L=5$ and $L=10$ overlap for all values of s highlighting that the proposed scheme can be advantageously associated with a small buffer size of

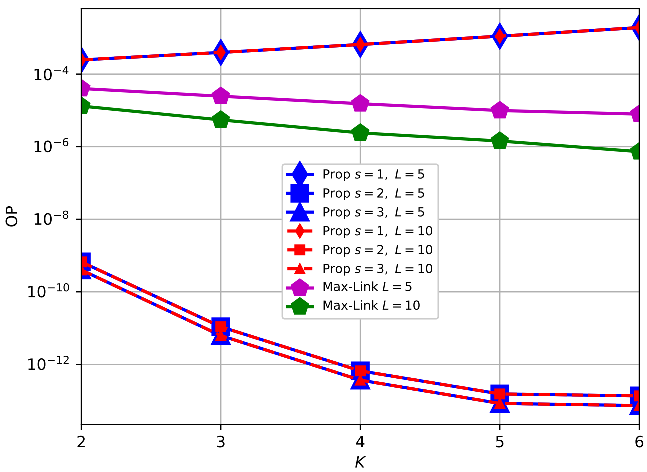

Fig. 4. Asymptotic OP of the *max-link* scheme and the proposed scheme for different values of K.

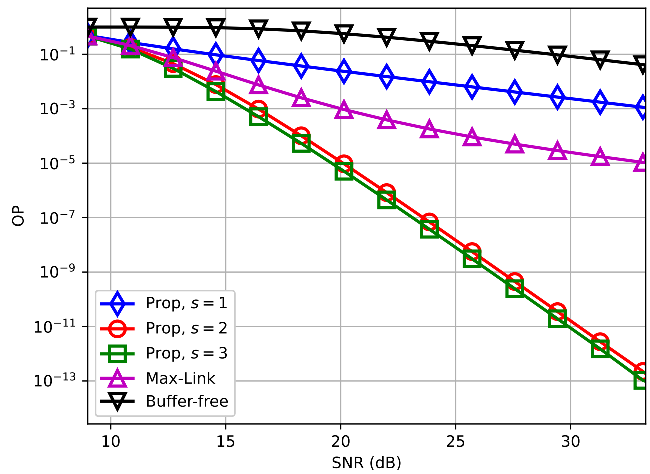

Fig. 6. OP for $K = 5$ and $L = 5$.

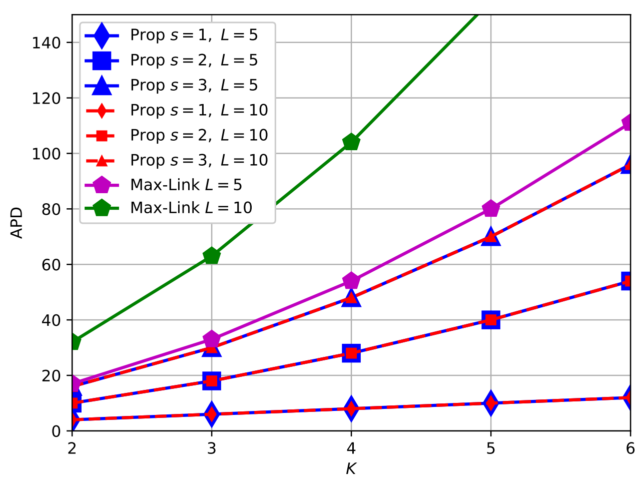

Fig. 5. Asymptotic APD of the *max-link* scheme and the proposed scheme for different values of K.

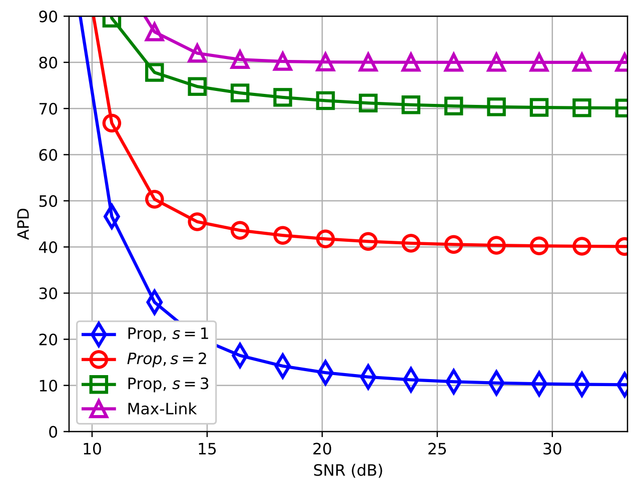

Fig. 7. APD for $K = 5$ and $L = 5$.

five. This observation does not hold for the *max-link* scheme where small buffer sizes present a small diversity order and full diversity is only reached with infinite buffer sizes [14]. Similarly, the *max-link* scheme admits an asymptotic APD of $K + L\sum_{i=0}^{K-1}(K - i)$ that depends on L. As such, for the *max-link* scheme, increasing the value of L for enhancing the diversity order will result in increasing the APD and a full diversity order (achievable with $L \to \infty$) will incur an infinitely large delay. However, for the proposed scheme, the full diversity order of $K+1$ can be achieved with a finite buffer size of five while keeping the APD bounded to $K(K + 3)$ and $2K(K + 2)$ for $s = 2$ and $s = 3$, respectively. Fig. 4 and Fig. 5 demonstrate that the proposed scheme outperforms the *max-link* scheme for all values of K. This superiority stems from including the buffer state information in the link selection algorithm unlike [14] where this selection depends solely on the values of the links' strengths. Moreover, the inclusion of the parameter s in the link selection process presents an additional degree of freedom that can optimize the performance. From Fig. 4 and Fig. 5, the proposed scheme with $s = 1$ results in a delay that is severally smaller than that of the *max-link* scheme at the expense of an increase in the OP since the choice $s = 1$ does not present any diversity gain. On the counterpart, the proposed scheme with $s \in \{2, 3\}$ outperforms the *max-link* scheme in both the OP and APD performance for all values of K and L. For example, for $s = 2$, $K = 4$ and $L = 10$, implementing the proposed scheme instigates sharp reductions in the APD from 104 to 28 and in the OP from around 10^{-6} to around 10^{-12} as compared to the *max-link* scheme.

Fig. 6 and Fig. 7 compare between the proposed scheme and the *max-link* scheme [14]. We added the performance of a buffer-free system whose OP is given by $1 - \prod_{k=1}^{K+1}(1 - p_k)$

where this system is not in outage only when all hops are not in outage. In these figures, we consider a 5-relay network with $L = 5$ and $\mathbf{\Omega} = [4.5, 4, 3.75, 5, 4.5, 4.5]$. Results in Fig. 6 and Fig. 7 demonstrate the capability of the proposed scheme to achieve a wide range of OP and APD levels, only by controlling the value of s. It can be seen that the best OP performance is achieved with $s = 3$ and the best APD performance is found for $s = 1$. The ability of the proposed relaying scheme in reaching the full diversity order with $s = 2$ and $s = 3$ manifests in the significant improvements in the OP where the performance gains exceed 10 dB at an OP of 10^{-5} as compared to the *max-link* scheme. The superiority in terms of the APD performance manifest clearly in Fig. 7. Finally, while both the buffer-free systems and the proposed scheme with $s = 1$ achieve the same diversity order of 1, the latter scheme results in OP improvements as can be observed from Fig. 6. From Fig. 7, these improvements are associated with an APD value that converges to $2K = 10$ as the SNR increases. For buffer-free systems on the other hand, a delay of $K + 1 = 6$ is incurred for an information packet to traverse the $K + 1$ hops between S and D in $K + 1$ time slots.

V. Conclusion

In this paper, we suggested a novel BA relaying strategy for serial-relaying HD networks for any number of relays. The priority of transmitting from a relay was captured by the current amount of packets stored in its buffer while the priority of transmitting from the source was quantified by a variable parameter s. Through an appropriate asymptotic analysis that limited the Markov chain evaluation to a tractable number of dominant states, we related the diversity order, outage probability and queuing delay to the parameter s. This analysis was culminated by suggesting convenient choices of the parameter s and the buffer size L. Advantageously, the delay does not increase with L and small buffer sizes not exceeding 5 are sufficient to achieve the ultimate performance gains of the network. These advantages constitute the main improvements of the suggested scheme with respect to the benchmark schemes in the literature.

Appendix A

For convenience, we introduce the following definition of the asymptotic order of a probability that can be written as the weighted sum of terms involving the product of 0 to $K + 1$ elements of $\{p_1, \dots, p_{K+1}\}$. For a probability $\mathfrak{p}$ that can be expressed as:

$$\mathfrak{p} = \sum_{i_1 \geq 0} \cdots \sum_{i_{K+1} \geq 0} c_{i_1,\dots,i_{K+1}} \prod_{n=1}^{K+1} p_n^{i_n}, \tag{29}$$

where $\{c_{i_1,\dots,i_{K+1}}\}$ are constants, then the asymptotic order of $\mathfrak{p}$ is defined as:

$$O(\mathfrak{p}) = \min_{i_1 \geq 0 \, \cdots \, i_{K+1} \geq 0} \left\{ \sum_{n=1}^{K+1} i_n \;\middle|\; c_{i_1,\dots,i_{K+1}} = 0 \;\; \forall \;\; \sum_{n=1}^{K+1} i_n < O(\mathfrak{p}) \right\}. \tag{30}$$

For example, $O(1/3 - p_1 + 2p_2 + \cdots) = 0$ and $O(p_1^2 + p_1p_2p_3 + \cdots) = 2$. From (30), it follows that $O(\mathfrak{p}\mathfrak{p}') = O(\mathfrak{p}) + O(\mathfrak{p}')$ and $O(\mathfrak{p} + \mathfrak{p}') = \min\{O(\mathfrak{p}), O(\mathfrak{p}')\}$.

The asymptotic analysis that we carry out in this appendix is an order-1 analysis where all probabilities whose asymptotic orders exceed 1 are ignored. In fact, for asymptotic SNR values, $p_k \ll 1$ for $k = 1, \dots, K + 1$ implying that the terms involving the product of two or more elements of $\{p_1, \dots, p_{K+1}\}$ can be ignored.

First, we demonstrate that the set $\mathcal{S}_c$ in (20) is closed by proving that the corresponding transition probabilities satisfy (18). A key element in the proof is to further partition the set $\mathcal{S}_c$ into two subsets $\mathcal{S}_{c,1}$ and $\mathcal{S}_{c,2}$ such that $\mathcal{S}_c = \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$ and (for $k \in \{1, \dots, K + 1\}$):

$$\forall \;\; \mathbf{l} \in \mathcal{S}_{c,1} : \qquad t_{\mathbf{l},\mathbf{l}'} = \begin{cases} 1 - p_k, & \mathbf{l}' \in \mathcal{S}_{c,1}; \\ p_k, & \mathbf{l}' \in \mathcal{S}_{c,2}. \end{cases} \tag{31}$$

$$\forall \;\; \mathbf{l} \in \mathcal{S}_{c,2} : \qquad t_{\mathbf{l},\mathbf{l}'} = \begin{cases} 1 - p_k, & \mathbf{l}' \in \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}; \\ p_k, & \mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases}, \tag{32}$$

where all transition probabilities whose asymptotic orders exceed one were ignored.

Proposition *4:* For the subsets $\mathcal{S}_{c,1}$ and $\mathcal{S}_{c,2}$ satisfying (31)-(32), the asymptotic orders of the steady-state probabilities of the corresponding states satisfy the relation in (33):

$$\begin{cases} O(\pi_{\mathbf{l}}) = 0, & \mathbf{l} \in \mathcal{S}_{c,1}; \\ O(\pi_{\mathbf{l}}) = 1, & \mathbf{l} \in \mathcal{S}_{c,2}; \\ O(\pi_{\mathbf{l}}) \geq 2, & \mathbf{l} \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases} . \tag{33}$$

Proof: We will prove that (33) satisfies the asymptotic orders of all balance equations. For any state $\mathbf{l}$, the balance equation at steady-state is generalized as follows:

$$\pi_{\mathbf{l}} = \sum_{\mathbf{l}' \in \mathcal{S}_{c,1}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} + \sum_{\mathbf{l}' \in \mathcal{S}_{c,2}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} + \sum_{\mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} \tag{34}$$

implying that:

$$O(\pi_{\mathbf{l}}) = \min \Bigg\{ \min_{\mathbf{l}' \in \mathcal{S}_{c,1}} \{ \underbrace{O(t_{\mathbf{l}',\mathbf{l}})}_{\triangleq o_1} + \underbrace{O(\pi_{\mathbf{l}'})}_{=0} \}, \min_{\mathbf{l}' \in \mathcal{S}_{c,2}} \{ \underbrace{O(t_{\mathbf{l}',\mathbf{l}})}_{\triangleq o_2} + \underbrace{O(\pi_{\mathbf{l}'})}_{=1} \}, \min_{\mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}} \{ \underbrace{O(t_{\mathbf{l}',\mathbf{l}})}_{\geq 0} + \underbrace{O(\pi_{\mathbf{l}'})}_{\geq 2} \} \Bigg\} \tag{35}$$

$$\triangleq \min \{ O_1(\pi_{\mathbf{l}}), O_2(\pi_{\mathbf{l}}), O_3(\pi_{\mathbf{l}}) \}, \tag{36}$$

where the asymptotic orders from (33) were replaced in (35). (i): For $\mathbf{l} \in \mathcal{S}_{c,1}$, $o_1 = 0$ and $o_2 = 0$ following from (31) and (32), respectively. Consequently, $O_1(\pi_{\mathbf{l}}) = 0$, $O_2(\pi_{\mathbf{l}}) = 1$ and $O_3(\pi_{\mathbf{l}}) \geq 2$ implying from (36) that $O(\pi_{\mathbf{l}}) = 0$ thus proving the first relation in (33). (ii): For $\mathbf{l} \in \mathcal{S}_{c,2}$, $o_1 = 1$ and $o_2 = 0$ following from (31) and (32), respectively. Consequently, $O_1(\pi_{\mathbf{l}}) = 1$, $O_2(\pi_{\mathbf{l}}) = 1$ and $O_3(\pi_{\mathbf{l}}) \geq 2$ implying from (36) that $O(\pi_{\mathbf{l}}) = 1$ thus proving the second relation in (33). (iii): For $\mathbf{l} \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$, $o_1 \geq 2$ and $o_2 = 1$ following from (31) and (32), respectively. Consequently, $O_1(\pi_{\mathbf{l}}) \geq 2$, $O_2(\pi_{\mathbf{l}}) = 2$ and $O_3(\pi_{\mathbf{l}}) \geq 2$ implying from (36) that $O(\pi_{\mathbf{l}}) = 2$ thus proving the third relation in (33). ■

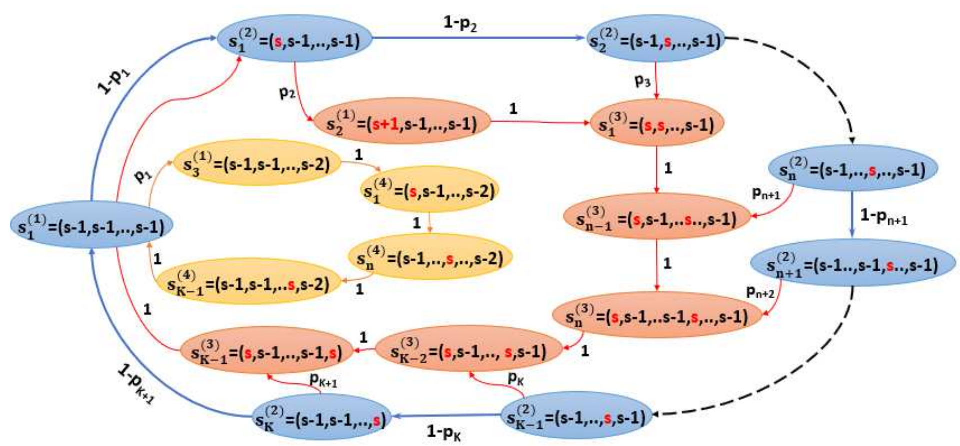

Fig. 8. Closed Subset for $1 < s < L$.

Lemma *1:* For the subsets $\mathcal{S}_{c,1}$ and $\mathcal{S}_{c,2}$ satisfying (31)-(32), the set $\mathcal{S}_c = \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$ is closed asymptotically.

Proof: For $\mathbf{l} \notin \mathcal{S}_c = \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$, $O(\pi_{\mathbf{l}}) \geq 2$ from (33) implying that the steady-state probability $\pi_{\mathbf{l}}$ can be ignored when carrying out the order-1 asymptotic analysis. Therefore, the MC is always in one of the states of $\mathcal{S}_c$ asymptotically implying that this set is closed. ■

Lemma *2:* For the order-1 asymptotic analysis, the transition probabilities in (32) can be approximated by:

$$\forall \ \mathbf{l} \in \mathcal{S}_{c,2}: \qquad t_{\mathbf{l},\mathbf{l}'} \approx \begin{cases} 1, & \mathbf{l}' \in \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}; \\ 0, & \mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases} \quad . \tag{37}$$

Proof: Since the transitions from $\mathcal{S}_{c,2}$ to states outside $\mathcal{S}_c$ are ignored since the set $\mathcal{S}_c$ is asymptotically closed following from lemma 1, then p_k can be approximated by 0 in (32). In fact, the inclusion of the probability p_k in the transitions appearing in (32) will only yield to terms whose asymptotic orders exceed two and, hence, can be ignored. ■

Proposition *5:* The following subsets of the set $\mathcal{S}_c$ in (20) satisfy the conditions in (31)-(32):

$$\begin{aligned} &\mathcal{S}_{c,1} = \{\mathbf{s}_1^{(1)}\} \cup \{\mathbf{s}_n^{(2)} \ ; \ n = 1, \dots, K\} \\ &\mathcal{S}_{c,2} = \{\mathbf{s}_n^{(1)} \ ; \ n = 2, 3\} \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} \ ; \ n = 1, \dots, K-1\}. \end{aligned} \tag{38}$$

Proof: We first prove the condition in (31). Ignoring the terms involving the product of two or more elements of $\{p_k\}_{k=1}^{K+1}$ asymptotically, the non-zero transition probabilities from elements of $\mathcal{S}_{c,1}$ in (38) are:

$$\begin{aligned} &t_{\mathbf{s}_1^{(1)},\mathbf{s}_1^{(2)}} = 1 - p_1 \quad ; \quad t_{\mathbf{s}_K^{(2)},\mathbf{s}_1^{(1)}} = 1 - p_{K+1} \\ &t_{\mathbf{s}_n^{(2)},\mathbf{s}_{n+1}^{(2)}} = 1 - p_{n+1} \ , \ n = 1, \dots, K-1, \end{aligned} \tag{39}$$

and:

$$\begin{aligned} &t_{\mathbf{s}_1^{(1)},\mathbf{s}_3^{(1)}} = p_1 \quad ; \quad t_{\mathbf{s}_1^{(2)},\mathbf{s}_2^{(1)}} = p_2 \\ &t_{\mathbf{s}_n^{(2)},\mathbf{s}_{n-1}^{(3)}} = p_{n+1} \ , \ n = 2, \dots, K, \end{aligned} \tag{40}$$

where the proof of (39) and (40) follows directly from the relaying strategy in (3)-(4). We will next provide the proof for the states $\mathbf{s}_1^{(1)}$ and $\mathbf{s}_n^{(2)}$ for $n = 2, \dots, K-1$. The proof for other states in $\mathcal{S}_{c,1}$ follows in a similar manner and, hence, will be omitted for the sake of brevity. (i): For $\mathbf{s}_1^{(1)} = (s-1, \dots, s-1)$, $\Delta_1 = s$ and $\Delta_2 = \dots = \Delta_{K+1} = s-1$ implying that preference is given for transmission from S. In this case, if link 1 is not in outage (with probability $1-p_1$), there will a transmission of a packet from S to R_1 which implies an increase in the number of packets stored in the buffer of R_1 by 1 thus moving to the state $\mathbf{s}_1^{(2)} = (s, s-1, \dots, s-1)$. If link 1 is in outage (with probability p_1) and since $\Delta_2 = \dots = \Delta_{K+1}$, the priority will be given for the transmission along the link with the highest index $K+1$ according to the tie breaking rule adopted in the relaying protocol. Therefore, with probability $p_1(1-p_{K+1}) \approx p_1$, a packet will be transmitted from R_K to D implying that the MC will move to the state $\mathbf{s}_3^{(1)} = (s-1, \dots, s-1, s-2)$. If link $K+1$ is in outage, the subsequent transition probabilities will involve the multiplicative term $p_1 p_{K+1}$ implying that such terms can be neglected in the order-1 asymptotic analysis. (ii): For $\mathbf{s}_n^{(2)}$ (with $n = 2, \dots, K-1$), $\Delta_1 = \Delta_{n+1} = s$ while $\Delta_2 = \dots = \Delta_n = \Delta_{n+2} = \dots = \Delta_{K+1} = s-1$. As such, priority will be given for transmission along link $n+1$ followed by link 1 if the link $n+1$ is in outage. Therefore, with probability $1-p_{n+1}$, a packet will be transmitted from R_n to R_{n+1} implying that the number of packets stored in R_n will be reduced by 1 while the number of packets stored in R_{n+1} will rise by 1, thus incurring a transition to the state $\mathbf{s}_{n+1}^{(2)}$. On the other hand, with probability $p_{n+1}(1-p_1) \approx p_{n+1}$, a packet will be transmitted from S to R_1 thus incurring a transition to the state $\mathbf{s}_{n-1}^{(3)}$. Other transition probabilities will involve the term $p_{n+1}p_1$ and, hence, can be ignored for large values of the SNR. As a conclusion, the transition probabilities in (39)-(40) satisfy the condition in (31).

Next, we prove the condition in (32). For elements of $\mathcal{S}_{c,2}$ in (38), the transitions that are confined in $\mathcal{S}_c = \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$ occur with the following probabilities:

$$\begin{aligned} &t_{\mathbf{s}_2^{(1)},\mathbf{s}_1^{(3)}} = 1 - p_2 \ ; \ t_{\mathbf{s}_3^{(1)},\mathbf{s}_1^{(4)}} = 1 - p_1 \\ &t_{\mathbf{s}_{K-1}^{(3)},\mathbf{s}_1^{(2)}} = 1 - p_{K+1} \ ; \ t_{\mathbf{s}_{K-1}^{(4)},\mathbf{s}_1^{(1)}} = 1 - p_K \\ &t_{\mathbf{s}_n^{(3)},\mathbf{s}_{n+1}^{(3)}} = 1 - p_{n+2} \ , \ n = 1, \dots, K-2 \\ &t_{\mathbf{s}_n^{(4)},\mathbf{s}_{n+1}^{(4)}} = 1 - p_{n+1} \ , \ n = 1, \dots, K-2, \end{aligned} \tag{41}$$

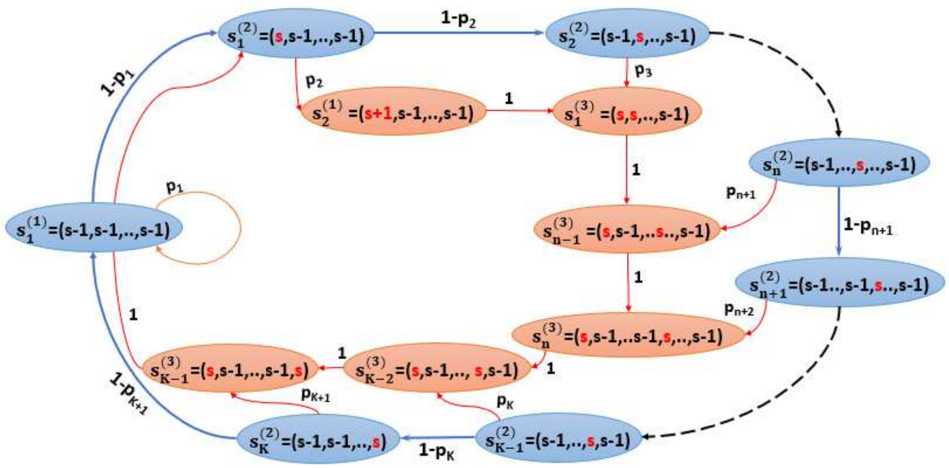

Fig. 9. Closed Subset for $s = 1$.

while the transitions leading to states outside $\mathcal{S}_c$ occur with the following probabilities:

$$\begin{aligned}
&t_{\mathbf{s}_2^{(1)},\mathbf{s}_2^{(1)}+\mathbf{e}_1} = p_2 \;\; ; \;\; t_{\mathbf{s}_3^{(1)},\mathbf{s}_3^{(1)}-\mathbf{e}_{K-1}+\mathbf{e}_K} = p_1 \\
&t_{\mathbf{s}_{K-1}^{(3)},\mathbf{s}_{K-1}^{(3)}-\mathbf{e}_1+\mathbf{e}_2} = p_{K+1} \;\; ; \;\; t_{\mathbf{s}_{K-1}^{(4)},\mathbf{s}_{K-1}^{(4)}+\mathbf{e}_1} = p_K \\
&t_{\mathbf{s}_n^{(3)},\mathbf{s}_n^{(3)}-\mathbf{e}_1+\mathbf{e}_2} = p_{n+2} \;,\; n = 1,\ldots,K-2 \\
&t_{\mathbf{s}_n^{(4)},\mathbf{s}_n^{(4)}+\mathbf{e}_1} = p_{n+1} \;,\; n = 1,\ldots,K-2,
\end{aligned} \tag{42}$$

where (41)-(42) follow directly from the relaying strategy in (3)-(4). As an illustration, we will provide the proof for the state $\mathbf{s}_n^{(3)}$ with $n = 1,\ldots,K-2$ and the proof for other states of $\mathcal{S}_{c,2}$ will follow in a similar manner. For the state $\mathbf{s}_n^{(3)}$ (with $n = 1,\ldots,K-2$), $\Delta_1 = \Delta_2 = \Delta_{n+2} = s$ while $\Delta_k = s-1$ for $k \neq 1,2,n+2$ implying that the highest priority is to transmit along the link $n+2$ followed by the link 2 (in case the link $n+2$ is in outage). Therefore, with probability $1-p_{n+2}$, R_{n+1} transmits and R_{n+2} receives resulting in the transition $\mathbf{s}_n^{(3)} \to \mathbf{s}_n^{(3)} - \mathbf{e}_{n+1} + \mathbf{e}_{n+2} = \mathbf{s}_{n+1}^{(3)}$. Ignoring the outage of more than one link asymptotically, with probability p_{n+2}, R_1 transmits and R_2 receives resulting in the transition $\mathbf{s}_n^{(3)} \to \mathbf{s}_n^{(3)} - \mathbf{e}_1 + \mathbf{e}_2 \notin \mathcal{S}_c$. As a conclusion, the transition probabilities in (41)-(42) satisfy the condition in (32). ■

Therefore, the union of the sets in (38) is closed asymptotically following from lemma 1 and the transition probabilities in (41) and (42) can be approximated by 1 and 0, respectively, following from lemma 2. This results in the simplified closed state diagram illustrated in Fig. 8.

From Fig. 8, the $3K+1$ balance equations in the closed subset $\mathcal{S}_c$ are given by:

$$\pi_{\mathbf{s}_1^{(1)}} = (1-p_{K+1})\pi_{\mathbf{s}_K^{(2)}} + \pi_{\mathbf{s}_{K-1}^{(4)}} \tag{43}$$

$$\pi_{\mathbf{s}_1^{(2)}} = (1-p_1)\pi_{\mathbf{s}_1^{(1)}} + \pi_{\mathbf{s}_{K-1}^{(3)}} \tag{44}$$

$$\pi_{\mathbf{s}_n^{(2)}} = (1-p_n)\pi_{\mathbf{s}_{n-1}^{(2)}} \quad , \quad n = 2,\ldots,K \tag{45}$$

$$\pi_{\mathbf{s}_2^{(1)}} = p_2\pi_{\mathbf{s}_1^{(2)}} \tag{46}$$

$$\pi_{\mathbf{s}_1^{(3)}} = p_3\pi_{\mathbf{s}_2^{(2)}} + \pi_{\mathbf{s}_2^{(1)}} \tag{47}$$

$$\pi_{\mathbf{s}_n^{(3)}} = p_{n+2}\pi_{\mathbf{s}_{n+1}^{(2)}} + \pi_{\mathbf{s}_{n-1}^{(3)}} \quad , \quad n = 2,\ldots,K-1 \tag{48}$$

$$\pi_{\mathbf{s}_3^{(1)}} = p_1\pi_{\mathbf{s}_1^{(1)}} \tag{49}$$

$$\pi_{\mathbf{s}_1^{(4)}} = \pi_{\mathbf{s}_3^{(1)}} \tag{50}$$

$$\pi_{\mathbf{s}_n^{(4)}} = \pi_{\mathbf{s}_{n-1}^{(4)}} \quad , \quad n = 2,\ldots,K-1. \tag{51}$$

Solving the relation $\sum_{\mathbf{l}\in\mathcal{S}_c} \pi_{\mathbf{l}} = 1$ along with the above balance equations, generates the expressions of the steady-state probabilities presented in (21).

APPENDIX B

The case of $s = 1$ differs from the case $1 < s < L$ presented in Appendix A by the elimination of the states $\mathbf{s}_3^{(1)}$ and $\mathbf{s}_n^{(4)}$ for $n = 1,\ldots,K-1$ since these states do not exist for $s = 1$ (since $s-2$ becomes negative in this case). Removing these states from (20) results in the closed subset provided in (22). The closed subset is now as presented in Fig. 9 that is obtained by removing the above mentioned K states from the state diagram in Fig. 8.

Consequently, in the balance equations (43)-(51), equations (49)-(51) must be removed while equation (43) must be replaced by:

$$\pi_{\mathbf{s}_1^{(1)}} = (1-p_{K+1})\pi_{\mathbf{s}_K^{(2)}} + p_1\pi_{\mathbf{s}_1^{(1)}}. \tag{52}$$

Solving (44)-(48) and (52) along with the relation $\sum_{\mathbf{l}\in\mathcal{S}_c} \pi_{\mathbf{l}} = 1$ generates the steady-state probabilities presented in (23).

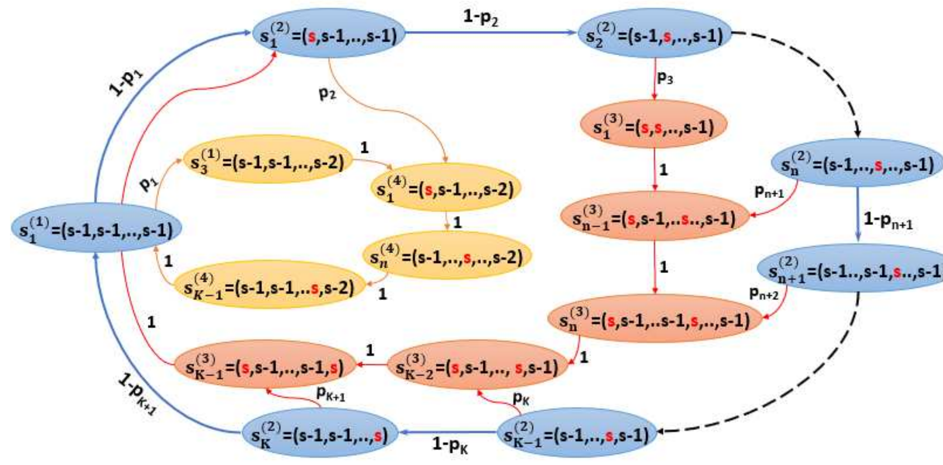

Fig. 10. Closed Subset for $s = L$.

Appendix C

The case of $s = L$ differs from the case $1 < s < L$ presented in Appendix A by the elimination of the state $\mathbf{s}_2^{(1)} = (s+1, s-1, \ldots, s-1)$ since $s+1$ exceeds the buffer size L for $s = L$. As such, the closed subset in (20) reduces to the one given in (24). The reduced state diagram for $s = L$ is illustrated in Fig. 10 where a transition can occur from $\mathbf{s}_1^{(2)}$ to $\mathbf{s}_1^{(4)}$ with probability p_2. In fact, in the state $\mathbf{s}_1^{(2)} = (s, s-1, \ldots, s-1)$, the highest priority is given for transmission along link 2 followed by the transmission along link $K+1$ since the buffer at R_1 is full for $s = L$ and the link 1 is unavailable. Consequently, equation (46) must be removed while equations (47) and (50) must be replaced by:

$$\pi_{\mathbf{s}_1^{(3)}} = p_3 \pi_{\mathbf{s}_2^{(2)}} \tag{53}$$

$$\pi_{\mathbf{s}_1^{(4)}} = p_2 \pi_{\mathbf{s}_1^{(2)}} + \pi_{\mathbf{s}_3^{(1)}}. \tag{54}$$

Solving equations (43)-(45), (48)-(49), (51) and (53)-(54) along with the relation $\sum_{\mathbf{l} \in \mathcal{S}_c} \pi_{\mathbf{l}} = 1$ generates the steady-state probabilities presented in (25).

References

[1] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1820–1830, Oct. 2004.

[2] L. Bai, L. Zhu, X. Zhang, W. Zhang, and Q. Yu, "Multi-satellite relay transmission in 5G: Concepts, techniques, and challenges," *IEEE Network*, vol. 32, no. 5, pp. 38–44, Sep. 2018.

[3] L. Yang, J. Chen, M. O. Hasna, and H.-C. Yang, "Outage performance of UAV-assisted relaying systems with RF energy harvesting," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2471–2474, Dec. 2018.

[4] M. O. Hasna and M.-S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1126–1131, Nov. 2003.

[5] N. Nomikos, T. Charalambous, I. Krikidis, D. N. Skoutas, D. Vouyioukas, M. Johansson, and C. Skianis, "A survey on buffer-aided relay selection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1073–1097, second quarter 2015.

[6] M. El-Rajab, C. Abou-Rjeily, and R. Kfouri, "Buffer-aided relaying: a survey on relay selection policies," *IET Communications*, vol. 14, no. 21, pp. 3715–3734, Jan. 2021.

[7] 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.

[8] Z. Tian, Y. Gong, G. Chen, and J. A. Chambers, "Buffer-aided relay selection with reduced packet delay in cooperative networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2567–2575, Mar. 2016.

[9] S. Luo and K. C. Teh, "Buffer state based relay selection for buffer-aided cooperative relaying systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5430–5439, Oct. 2015.

[10] P. Xu, Z. Ding, I. Krikidis, and X. Dai, "Achieving optimal diversity gain in buffer-aided relay networks with small buffer size," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8788–8794, Oct. 2015.

[11] S. El-Zahr and C. Abou-Rjeily, "Threshold based relay selection for buffer-aided cooperative relaying systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6210–6223, Sep. 2021.

[12] L.-L. Yang, C. Dong, and L. Hanzo, "Multihop diversity - a precious source of fading mitigation in multihop wireless networks," in *IEEE Global Telecommun. Conf. (GLOBECOM 2011)*, 2011, pp. 1–5.

[13] C. Dong, L.-L. Yang, and L. Hanzo, "Performance analysis of multihop-diversity-aided multihop links," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2504–2516, July 2012.

[14] B. Manoj, R. K. Mallik, and M. R. Bhatnagar, "Buffer-aided multi-hop DF cooperative networks: A state-clustering based approach," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 4997–5010, 2016.

[15] A. K. Shukla, B. Manoj, and M. R. Bhatnagar, "Virtual full-duplex relaying in a buffer-aided multi-hop cooperative network," in *Int. Conf. on Signal Processing and Commun. (SPCOM)*. IEEE, July 2020, pp. 1–5.

[16] V. Jamali, N. Zlatanov, H. Shoukry, and R. Schober, "Achievable rate of the half-duplex multi-hop buffer-aided relay channel with block fading," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6240–6256, Nov. 2015.

[17] M. M. Razlighi and N. Zlatanov, "Buffer-aided relaying for the two-hop full-duplex relay channel with self-interference," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 477–491, Jan. 2018.

[18] G. Yang, M. Haenggi, and M. Xiao, "Traffic allocation for low-latency multi-hop networks with buffers," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3999–4013, May 2018.

[19] C. Abou-Rjeily and W. Fawaz, "Buffer-aided serial relaying for FSO communications: asymptotic analysis and impact of relay placement," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8299–8313, Dec. 2018.

[20] B. Li, S. Zhao, R. Miao, and R. Zhang, "A survey on unmanned aerial vehicle relaying networks," *IET Communications*, vol. 15, pp. 1262–1272, Feb. 2021.

[21] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[22] J. D. Little and S. C. Graves, "Little's law," in *Building intuition*. Springer, 2008, pp. 81–100.