

LEBANESE AMERICAN UNIVERSITY

Design and Analysis of Buffer-Aided Cooperative Networks
using Deterministic and Reinforcement Learning
Techniques

By

Sawsan El Zahr

A thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer
Engineering

School of Engineering

April 2022

THESIS APPROVAL FORM

Student Name: Sawsan El Zahr I.D. #: 201706764

Thesis Title: Design and Analysis of Buffer Aided Cooperative Networks using Deterministic and Reinforcement Learning Techniques.

Program: Master of Science in Computer Engineering.

Department: Electrical and Computer Engineering.

School: School of Engineering.


The undersigned certify that they have examined the final electronic copy of this thesis and approved it in Partial Fulfillment of the requirements for the degree of:

Master of Science in the major of Computer Engineering.

Thesis Advisor's Name: Chadi Abov-Rjeily

Signature:  Date: 29 / 04 / 2022
Day Month Year

Committee Member's Name: DANI TANNIR

Signature:  Date: 29 / 4 / 2022
Day Month Year

Committee Member's Name: Wessam FAWAZ

Signature:  Date: 29 / 04 / 2022
Day Month Year

THESIS COPYRIGHT RELEASE FORM

LEBANESE AMERICAN UNIVERSITY NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants the Lebanese American University (LAU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic formats and in any medium, including but not limited to audio or video. You agree that LAU may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that LAU may keep more than one copy of this submission for purposes of security, backup and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant LAU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN LAU, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. LAU will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Name: Sawsan El Zahr

Signature:  Date: 29 / 4 / 2022
Day Month Year

PLAGIARISM POLICY COMPLIANCE STATEMENT

I certify that:

1. I have read and understood LAU's Plagiarism Policy.
2. I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.
3. This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: Sawsan El Zahr

Signature: 

Date: 29 / 4 / 2022
Day Month Year

Acknowledgment

First and foremost, I'm thankful to God for offering me the opportunity to pursue this research work and the dedication to accomplish it.

Second, I acknowledge the help and support of many people to finalize this thesis. I take this opportunity to thank my supervisor, Professor Chadi Abou Rjeily, for his great guidance and help in formulating the research questions and methodology. Also thanks to the Graduate Studies and Research (GSR) office for securing the funding of this research degree and helping in developing my research skills.

Finally, I would like to thank my parents and friends for their continuous encouragement, support and love.

Design and Analysis of Buffer-Aided Cooperative Networks using Deterministic and Reinforcement Learning Techniques

Sawsan El Zahr

Abstract

Applications enabled by 5G technology resulted in an unforeseen increase in data traffic and data flows from any source to a destination are experiencing increased outages, delays, and drop-outs. Cooperative communication is one way to cope with this problem by integrating relays into systems to help a given source to communicate with its destination efficiently. Relays enable multiple shorter paths of better quality and equipping relays with buffers will further increase the degrees of freedom and allow for the mitigation of the fading effect. On the other hand, these benefits come at the expense of added complexity to the system. The problem of relay selection is a challenging task that can account for multiple parameters such as: the channel state information, the buffer states, and the position of relays. In this thesis, we address this problem in two ways: deterministic and learning-based techniques. Multi-hop systems with buffer-aided half-duplex relays are considered.

First, we propose a new relaying strategy that is dynamic and can achieve multiple levels of trade-off between the average packet delay and the outage probability. The system is analyzed in a Markov Chain framework and all theoretical results are checked for accuracy with simulation curves. Asymptotic analysis is the key approach to derive closed-form expressions solely dependent on adjustable parameters of the system. We could prove the superiority of this scheme compared with other benchmark schemes from the literature. Then, further relaying strategies are devised and compared. Additional performance levels are achieved to fit in different applications requirements. Next, for more complex setups, deterministic analysis becomes cumbersome, thus reinforcement learning techniques are used to efficiently boost the performance. A deep RL agent is trained with a joint reward until it converges to an optimum performance. We demonstrate the efficiency of this approach to further increase the throughput of cooperative systems under different interference and design constraints.

Keywords: Relaying, Cooperative Networks, Multi-Hop, Buffer, Data Queue, Performance Analysis, Markov Chain, Reinforcement Learning, Outage Probability, Queuing Delay, Diversity Order, Throughput, Optimization.

Contents

1	Introduction	1
1.1	RF Technology and 5G communications	1
1.2	Buffer-Free Cooperative Communications	2
1.2.1	Amplify-and-Forward versus Decode-and-Forward	2
1.2.2	Half-Duplex versus Full-Duplex Relays	3
1.2.3	Parallel Relaying	3
1.2.4	Serial Relaying	5
1.3	Buffer-Aided Cooperative Communications	5
1.3.1	HD DF parallel relaying	5
1.3.2	HD DF serial relaying	7
1.3.3	FD DF parallel relaying	8
1.3.4	FD DF serial relaying	9
1.4	RL Techniques for BA Systems	9
1.5	Contributions	11
1.6	List of Publications	11
2	Analysis Techniques	12
2.1	Reinforcement Learning (RL)	12
2.1.1	Markov Decision Process	12
2.1.2	RL Agent	12
2.1.3	Q-Learning	13
2.1.4	Deep Q-Learning	13
2.1.5	Handling Unfeasible Actions	14
2.2	Markov Chain Model	15
2.3	Asymptotic Analysis	15
3	Buffer State Based Relay Selection for Half-Duplex Buffer-Aided Serial Relaying Systems	17
3.1	Objectives	17
3.2	System Model and Relaying Strategy	17
3.2.1	Basic Parameters	17
3.2.2	Buffer State Based Relaying Strategy	19
3.3	Performance Analysis	22
3.3.1	Generalities	22
3.3.2	State Transition Matrix	23
3.3.3	Asymptotic Analysis	25
3.3.4	Conclusions about the design of the BA relaying scheme	30
3.4	Numerical Results	31

4	Relaying Strategies for Half-Duplex Buffer-Aided Serial Relaying Systems	37
4.1	Relaying Strategies	37
4.2	Performance Analysis	38
4.2.1	Generalities	38
4.2.2	Asymptotic analysis	39
4.2.3	Conclusions about the design of the BA relaying schemes .	42
4.3	Numerical Results	43
5	Optimized Relay Selection for Multi-Hop Cooperative Systems Using Deep Reinforcement Learning	46
5.1	Objectives	46
5.2	System Model	47
5.2.1	Serial Relaying	47
5.2.2	Parallel Relaying	49
5.3	Problem Formulation	50
5.4	Elements of the Reinforcement Learning Model	51
5.4.1	Environment	51
5.4.2	States	52
5.4.3	Actions	52
5.4.4	Reward Function	52
5.5	Deep Reinforcement Learning	53
5.5.1	Neural Network (NN)	53
5.5.2	Handling Unfeasible Actions	53
5.5.3	Agent	54
5.5.4	Algorithm	55
5.6	Simulation Results and Discussion	55
6	Conclusions and Future Work	63
A	Closed Subset for $1 < s < L$	70
B	Closed Subset for $s = 1$	76
C	Closed Subset for $s = L$	77
D	Transition Probabilities of Scheme 1	78
E	Derivation of the DO Expression	79

List of Figures

1	Half-Duplex vs Full-Duplex Relay	3
2	Parallel Relays	4
3	Serial Relays	5
4	Buffer-Aided Parallel Relays	6
5	Buffer-Aided Serial Relays	8
6	The agent-environment interaction in RL	13
7	Q-learning Approach	13
8	Deep Q-learning Approach	14
9	Closed Subset Approach	16
10	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.	33
11	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.	33
12	Asymptotic OP of the <i>max-link</i> scheme and the proposed scheme for different values of K	34
13	Asymptotic APD of the <i>max-link</i> scheme and the proposed scheme for different values of K	34
14	OP for $K = 5$ and $L = 5$	35
15	APD for $K = 5$ and $L = 5$	36
16	APD for different scenarios with $L=5$. Solid lines and dashed lines represent the analytical and simulation values, respectively.	44
17	OP for $K=5$ and $L=5$	45
18	APD for $K=5$ and $L=5$	45
19	Interference Constraint for Multiple Links Activation	48
20	Step 2 of Deep-RL Agent	54
21	Convergence of the throughput for $K = 4$ and $L = 10$ for all 4 serial setups at $SNR = 30$ dB.	57
22	Throughput for $K = 4$ and $L = 10$ for all 4 serial setups.	58
23	APD for $K = 4$ and $L = 10$ for all 4 serial setups.	59
24	Change in behavior scores for $K = 4$ and $L = 10$ for all 4 serial setups.	60
25	Convergence of the throughput for $K = 6$ and $L = 8$ for all 4 parallel setups at $SNR = 15$ dB.	61
26	Throughput for $K = 6$ and $L = 8$ for all 4 parallel setups.	61

27	APD for $K = 6$ and $L = 8$ for all 4 parallel setups.	62
28	Change in behavior scores for $K = 6$ and $L = 8$ for all 4 parallel setups.	62
29	Closed Subset for $1 < s < L$	74
30	Closed Subset for $s = 1$	76
31	Closed Subset for $s = L$	77

List of Tables

1	Closed subset of Scheme 3 for $\gamma > 0$	40
2	Closed subset of Scheme 3 for $\gamma < 0$	40
3	Summary of RL Symbols	55

List of Abbreviations and/or Symbols

DF	Decode-and-forward
AF	Amplify-and-forward
HD	Half-duplex
FD	Full-duplex
AWGN	Additive white gaussian noise
RF	Radio frequency
BA	Buffer aided
RL	Reinforcement learning
MC	Markov Chain
MDP	Markov decision process
UAV	Unmanned aerial vehicle
SWIPT	Simultaneous Wireless Information and Power Transfer
IoT	Internet of Things
D2D	Device to Device communications
IRI	Inter-relay interference
NN	Neural network
OP	Outage probability
APD	Average packet delay
DO	Diversity order
BER	Bit error rate
CSI	Channel state information
BSI	Buffer state information
K	Number of relays in a system
L	Buffer size of relays
η	Throughput
\bar{L}	Average queue length
$\bar{\gamma}$	Average signal-to-noise ratio
p	Outage probability of a particular link
q	Availability of a particular link
r_0	Target rate in bits per channel use (BPCU)
N_r	Number of rounds for RL
N_i	Number of iterations per round
N_c	Number of collected experiences per iteration
BS	Training batch size
γ	Discount factor
ϵ	Exploration rate
f	Decay factor

Chapter 1

Introduction

1.1 RF Technology and 5G communications

Wireless communication systems existed since 1979 when the first generation (1G) technology emerged. It started with analog signals and provided voice-only services. This first generation suffered from low security, bad quality and limited coverage. 2G technology appeared in 1991 and introduced digital signals with encryption capabilities. TDMA and CDMA were the main technologies and SMS services were supported. Later in 2001, 3G emerged with WCDMA to increase the data rate further. Multimedia messages and video conferencing were possible for the first time and smart phones became popular. The fourth generation 4G started in 2009 and used LTE technology. Advanced modulation/coding schemes were introduced and security of systems was leveraged which enabled high definition multimedia streaming and 3D gaming applications.

In fact, whenever communication systems evolve, new applications are enabled and consequently more bandwidth and performance requirements are needed. Technological advances nowadays resulted in additional dependence on communication systems: Internet of Things (IoT), Device-to-Device (D2D) communications, autonomous vehicles, environmental monitoring, smart grids, gaming, sports broadcasting and many more applications emerged and resulted in an unparalleled increase in data traffic. Thus, communication systems should develop to meet the necessary requirements of reliability, efficiency and reduced latency [1]. 5G systems appeared in 2019 and are in continuous development.

Communication techniques span five consecutive layers: Application layer, Transport layer, Network layer, Data Link layer and Physical layer. This thesis studies the communication techniques at the lowest layer which is the Physical layer. This layer is responsible for the transmission of bits over the communication channel. It deals with the encoding of data, the adaptation of the data rate, synchronization issues, topologies and transmission modes.

Specifically, among the 5G technologies, cooperative techniques are investi-

gated for being a promising solution to increase the efficiency and reliability of wireless systems [2]. Relays are integrated into networks to help a source or multiple sources to transmit packets of data to one or more destinations. In fact, relays can be placed either in parallel or in series to constitute multiple hops of communication. Dividing the end-to-end transmission distance allows for better quality of channels because of lower distances in addition to the ability to mitigate the fading effect by transmitting at different locations and time slots. Applications of such systems include unmanned aerial vehicle (UAV) relaying networks [3, 4] where multiple UAVs are employed as relay nodes to forward information packets. Another application is wireless backhauling of 5G small cells [5, 6] 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 packets reach the target SBS. Moreover, cooperation can be established for satellite communications where satellites can cooperate to serve ground terminals [7].

In this work, cooperative systems with Radio Frequency (RF) technology are considered. RF is a very old technology and its frequency range started from 1KHz to 1GHz but was recently extended to hundreds of GHz to the millimeter-wave bands. The obvious reason to increase the frequency band is to expand the spectrum and achieve higher bandwidth. However another reason is that the length of RF antennas is proportional to the wavelength and since smaller antennas are required when the frequency was further increased [8]. This technology has low transmit power requirements and is suitable for transmission over long distances because of its high spatial diversity gain [9]. There exist different channel models to represent the multipath fading of RF signals: Rayleigh, Rician and Nakagami. However, Rayleigh distribution is often assigned to RF links in the absence of a line-of-sight [10].

1.2 Buffer-Free Cooperative Communications

The first forms of implementing cooperative systems were buffer free where the relays do not possess any storing capabilities. In addition, different forms of relaying were implemented: amplify-and-forward (AF) vs decode-and-forward (DF), different modes of transmission: half-duplex (HD) vs full-duplex (FD) and different topologies: parallel vs serial relaying.

1.2.1 Amplify-and-Forward versus Decode-and-Forward

A relay in the AF mode amplifies the signal received and then retransmits it without decoding. However, in the DF mode, a relay decodes and remodulates

the signal before transmitting it to the destination. Although the AF mode is relatively simple, it has low performance since noise is amplified as well. In [11,12], it has been proven that for single-antenna multi-hop relaying, the DF protocol achieves higher capacity and lower outage probability than the AF protocol. This advantage comes at the drawback of added complexity because of the added processing capability for DF relays [13].

1.2.2 Half-Duplex versus Full-Duplex Relays

Fig. 1 shows a comparison between HD and FD modes. A HD relay receives a signal from the source in the first time slot and then transmits to the destination in the second time slot. Thus, no simultaneous reception and transmission is allowed. However, a FD relay can receive and transmit two different signals at the same time. In this mode, the two signals will interfere with each other and the nodes of the system should have extra processing capabilities to eliminate this interference.

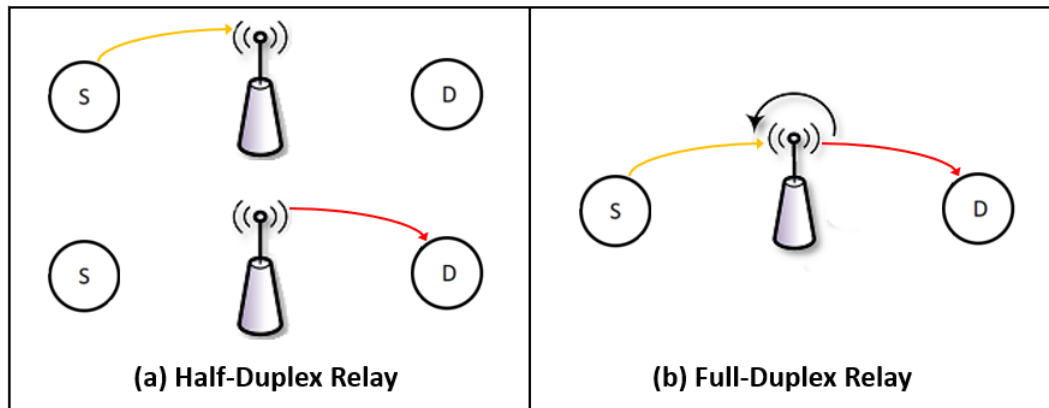


Figure 1: Half-Duplex vs Full-Duplex Relay

FD mode is more preferable only for systems with unconstrained bit-error-rate (BER) and with low transmit power and strong loop interference cancellation [14]. In [15], HD relays are used to build a virtual FD system where relays transmit packets one after another while the source is always sending packets to all relays. An inter-relay-interference (IRI) cancellation method is proposed by running a priori signal and then accordingly minimizing the power of the residual interference at every time slot.

1.2.3 Parallel Relaying

The first topology considered is parallel relaying and is illustrated in Fig. 2. K relays are placed between a source S and a destination D so that packets of information can be transmitted from S to any relay R_k and then R_k transmits

the packet to D with $k = 1, \dots, K$. With inter-relay cooperation enabled, a relay R_k will be able to transmit packets to an adjacent relay (i.e R_{k-1} or R_{k+1}) and hence the number of paths will be higher since $K - 1$ links are added.

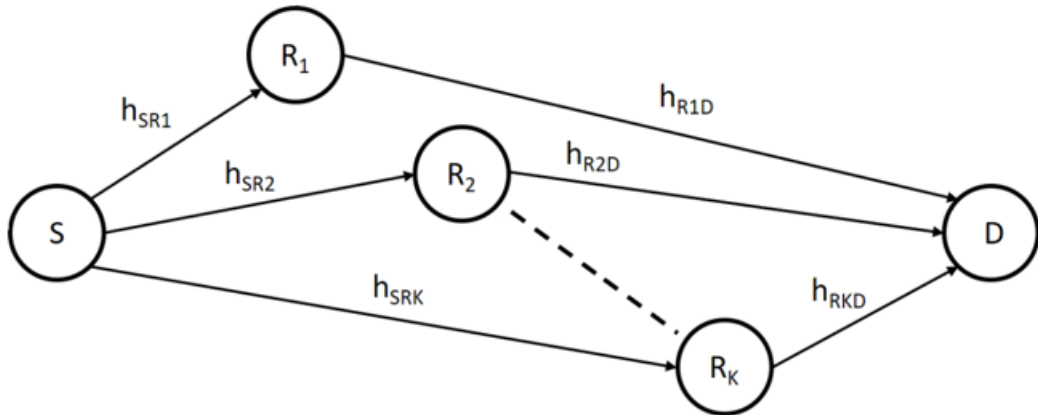


Figure 2: Parallel Relays

The very basic relaying strategy of buffer-free DF parallel setups is the max-min scheme where the best end-to-end link is selected [16]. In other words, the quality of all S-R and R-D links are checked and every path is judged by the minimum channel quality of either S- R_k or R_k -D link. Next, the relay allowing for the strongest path will receive a packet in the first time slot and then re-transmit the packet to D in the second time slot. Other feedback schemes were suggested in [17] and [18] in the context of DF and AF relays, respectively. A direct link between S and D was considered and relaying is used only if the destination node fails to correctly decode the signal from the source in case of the DF system or the SNR of the received signal is below a threshold in case of the AF system. In case relaying is required, D sends a feedback signal to all relays and then decides on the best node (among relays and the source) to transmit the packet. Finally, combination of the received signals is carried out at the destination node.

Inter-relay cooperation was tackled in [19] and its performance advantage was found to be only at low SNR condition and could achieve a lower outage probability and a higher throughput. The advantage of using inter-relay cooperation is further analyzed in [20], where the feedback from the destination is considered and packet transmission is done in three priority selections: (1) S-D link, (2) best quality R-D link, (3) an available R-R link where the sending relay correctly decoded the source's signal and the receiving relay successfully decoded the NACK from D. This receiving relay will then retransmit the packet to D. This scheme proved to achieve better outage probability and higher transmission rate compared with the conventional max-min scheme without a direct S-D link.

1.2.4 Serial Relaying

Relays can be placed serially to form multi-hop systems as shown in Fig. 3. With this topology, packets of information are transmitted from the source to the first relay R_1 and then from one relay to another until it reaches the destination node.

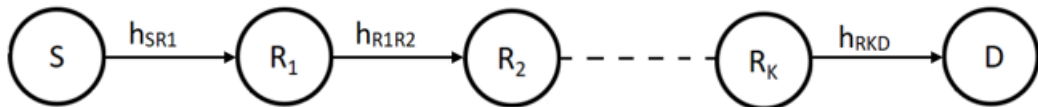


Figure 3: Serial Relays

Serial relaying has been extensively studied in the literature as a means of extending the coverage of wireless networks [21]. Similar to parallel relays, AF and DF modes can be implemented at the cooperating relays.

The most basic buffer free scheme consists of transmitting one packet from S through the relays to D , only when all links are available. Otherwise, an outage occurs. One weakness of this setup is its limited end-to-end outage performance that is dominated by the weakest of all hops; i.e. the hop with the highest outage probability [22]. Another scheme is implemented in [23] where the IRI is neglected and even and odd indexed relays transmit packets alternatively. Different modulation schemes were considered and the symbol error probability was derived for each.

1.3 Buffer-Aided Cooperative Communications

Relaying techniques evolved from buffer-free to buffer-aided (BA) relaying where the relays are equipped with buffers (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 can be exploited to mitigate channel fading and, hence, enhance the reliability of communications at the expense of introducing queuing delays [24, 25].

1.3.1 HD DF parallel relaying

The system model for BA parallel relaying is shown in Fig. 4. K relays are placed between S and D and are equipped with buffers of size L .

Systems with a single-relay were considered in [26–28]. In [26,27], the throughput was maximized while the case of buffer overflow was ignored. However in [28], finite buffers were considered and performance optimization revolved around the outage probability and the average delay.

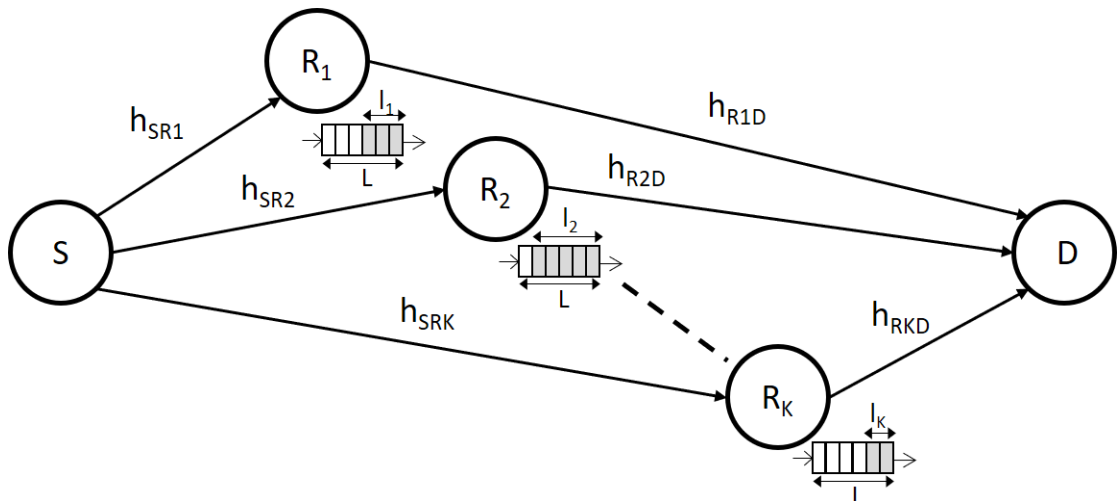


Figure 4: Buffer-Aided Parallel Relays

On the other hand, multiple relays were tackled in [29–41]. The authors in [29] suggested the max-link scheme that consists of selecting the strongest link among all available S-R and R-D links. This scheme is capable of achieving a full diversity order of $2K$ but only with infinite buffer sizes. Moreover, this scheme suffers from high APD values of $KL + 1$ that increase linearly with the buffer size. In [30], another scheme was suggested to reduce the average delay and improved on the max-link scheme by giving preference for the R-D links. It can achieve a smaller APD of 2, independent of the buffer size, but at the expense of decreasing the diversity order. The relaying strategy presented in [31] is similar to [29] but attempts to equalize the buffer lengths at the relays. This scheme outperformed the max-link scheme in terms of delay but has a limited diversity order equal to K as in buffer-free systems. In [32], the relay selection consisted of prioritizing the S-R and R-D links in odd and even time slots, respectively. This scheme is advantageous in the case of finite buffer sizes compared to the max-link protocol, where it achieves a slight increase in the diversity order.

Unlike [29–32] where the relay selection policy is based solely on the CSI, the buffer state information (BSI) is considered in [33–39]. A balancing BA scheme was analyzed in [33] for symmetrical networks and consisted of keeping the number of packets at each buffer the closest possible to $L/2$. Another priority-based max-link scheme was proposed in [34] and classifies relays in three priority classes; namely relays with full, empty and neither full nor empty buffers. In the case of quasi-symmetrical networks, this scheme also proved to achieve a diversity order equal to $2K$ for large values of L . The scheme proposed in [35] is based on a different relay classification. It assigns the relays as in the transmission mode or in the reception mode based on their current queue length and then,

decides on whether to transmit or receive based on the maximum and minimum number of stored packets for the transmission and reception modes, respectively. A full diversity order can be achieved for symmetrical networks with finite buffer sizes, along with a reduced asymptotic APD of $2K + 2$. In [36], the proposed scheme consisted of assigning a weight to each link that differentiates between S-R and R-D links and then, the link with the highest weight is selected. Similar to [35], the scheme in [36] is capable of achieving a diversity order of $2K$ in symmetrical networks for $L > 2$. In order to better address the cases where more than one link have the same weight, the scheme in [37] introduced a relay selection factor that includes the weight of the link and the link's quality as the first and second metrics, respectively. This resulted in the ability of the scheme to either prioritize the OP or the APD. Moreover, two delay-aware relaying strategies were suggested in [38] based on the availability of the links and buffer sizes. The first policy achieves an asymptotic APD of $4K - 1$, while the second one trade-offs the delay and the diversity order and reduces the delay to $2K + 1$ at the expense of decreasing the diversity order. The scheme in [39] introduced a threshold-based selection scheme where relays are divided into 4 priority classes based on the threshold of each relay. Multiple levels of tradeoffs between OP and APD could be achieved with a diversity order ranging from K to $2K$ and an average delay ranging from 2 to $2K + 2$. Moreover, a buffer size of 3 was proved to be sufficient to extract all the capabilities of the network.

While the BA relay selection schemes in [29–39] are deterministic, probabilistic schemes were considered in [40, 41]. In [40], the strongest available S-R and R-D links are selected first and then, a random selection is done between the two selected links. A full diversity order of $2K$ is achieved for quasi-symmetrical networks with infinite buffer sizes while allowing for different levels of tradeoff between OP and APD. In [41], after selecting the S-R and R-D links with the smallest and largest numbers of packets in the corresponding buffers, respectively, this scheme randomly selects between these links according to a probability distribution taking into account the delay constraints.

1.3.2 HD DF serial relaying

The system model for BA serial relaying is shown in Fig. 5. K relays are placed sequentially between S and D and are equipped with buffers of size L .

While the research on BA parallel relaying is extensive, BA serial relaying was less investigated in the literature [42–45]. Max-link selection was analyzed in [42–44] 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

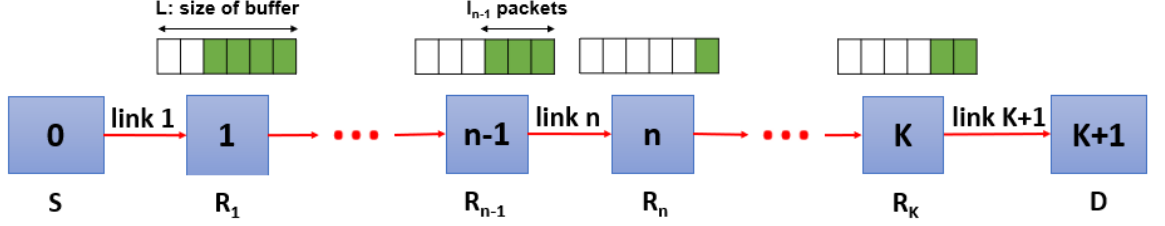


Figure 5: Buffer-Aided Serial Relays

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 [42], a lowerbound on the bit error rate was derived by loosening the aforementioned availability constraint and assuming that the buffers at the relays have an infinite size and that each relay always has packets to transmit. 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 [43], the BER and outage probability of the max-link scheme were analyzed. Results showed that the diversity order of $K + 1$ can be achieved exclusively with infinite buffer sizes while practical finite-size buffers can achieve only a fraction of this maximum diversity order. A Markov chain analysis was adopted to evaluate the outage and delay performance of the max-link scheme in [44]. Results are consistent with [43] 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.

HD BA DF serial relaying was considered in [45] with infinite buffer sizes. This work targeted the maximization of the average rate for a communication session that extends over an infinite number of fading blocks in a block fading environment. The implementation of the relaying protocol in [45] 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 [42–44], the relaying protocol in [45] is based solely on the CSI where the transmission modes are related to the rates that can be achieved over the different hops.

1.3.3 FD DF parallel relaying

Virtual FD parallel relaying is investigated in [46] and [47] in the sense that HD relays are used while the source and one of the relays are transmitting information packets simultaneously. The authors in [46] proposed joint relay selection and

beamforming schemes that account for the IRI problem. Relays were equipped with multiple antennas and the end-to-end rate was maximized using a weighted sum-rate maximization strategy. Results show that when increasing the number of relays and/or the number of antennas, the performance can asymptotically approach the upper-bound of the ideal FD relaying mode. On the other hand, the power minimization problem accounting for IRI was tackled in [47] with single antennas at the relays. This scheme has improved the system's performance in terms of the energy efficiency, the throughput and the average delay.

1.3.4 FD DF serial relaying

Virtual FD relaying was studied in [48]. The outage performance of the max-link scheme was improved assuming the absence of inter-relay-interference. 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 [48] 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 selection is based on the instantaneous SNR.

FD relays were considered in [49–51] 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 [49] where 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 [50] and [51] 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 for RF, self-interference (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 [50, 51] take into consideration that concurrent transmissions can take place along all hops in the absence of interference.

1.4 RL Techniques for BA Systems

Reinforcement learning techniques are recently integrated into cooperative networks for their ability to optimize complex problems that are hard to be solved in a theoretical framework. BA relaying systems are especially tackled since they can be modeled as Markov Decision Processes (MDP). Details about RL techniques are provided in Section 2.1.

Parallel BA relaying is considered in [52–57] in the context of RL algorithms. Joint throughput and delay optimization are tackled in [52–54] for different systems and under different constraints while the optimization in [55–57] addressed jointly the secrecy rate, the delay and the throughput of the system. Dealing with unfeasible actions is also a prime issue for RL techniques.

A deep Q-learning algorithm was proposed in [54] that allocates a negative reward for unfeasible actions and a positive reward for a packet that reaches the destination node within the delay constraint. A similar positive reward is used in [52] but unfeasible actions are assigned a zero Q-value in the target network while training the agent. This technique proved to further improve the throughput compared with the punishment-based algorithm in [54]. In [53], two algorithms were proposed and compared for hybrid OMA/NOMA BA relaying: asynchronous deep Q-Learning and asynchronous advantage actor-critic scheme. The former scheme demonstrated better performance and faster convergence for small action spaces whereas the latter showed better performance in the case of a large action space. Moreover, the RL agent was provided with a priori-information about unfeasible actions as well as inconvenient actions for the delay constraint. The agent uses this information to remove undesirable actions from the action space at each time slot so that the system could converge faster. The a-priori information could improve the performance for both aforementioned schemes.

On the other hand, the system in [56] comprises of a primary source, a secondary source, a primary destination, a secondary destination, K HD BA parallel relays and one eavesdropper. A deep Q-learning algorithm is proposed to maximize the throughput within a delay and secrecy rate constraints and allocates negative reward for unfeasible actions. The same system was considered in [57], and a a-priori information based Double Deep Q-learning (DDQN) algorithm was proposed to improve the convergence. In addition, the system in [55] makes use of an intelligent reflecting surface (IRS) with N reflecting elements. The problems of maximizing the throughput under the system’s constraints were divided into sub-tasks and a distributed multi-agent reinforcement learning scheme was proposed to solve the optimization problem.

Furthermore, RL is used in the context of simultaneous wireless information and power transfer (SWIPT) to optimize the energy consumption. Joint power allocation and throughput optimization for single relay systems is tackled in [58]. Deep Q-learning was used to find the optimal policy that maximizes the long-term reward. The reward function was based on the number of received and forwarded packets at the relay. Finally, a pretraining scheme was implemented to accelerate the convergence of the system by driving the DQN agent to choose some desirable actions at specific states. These preferences are assumed to hold

for any system transition function.

1.5 Contributions

1. HD DF serial relaying is not well investigated in the literature compared to parallel relaying. Previous schemes suffer from high average packet delay that increases with the buffer size. Also, large buffer sizes are needed to achieve the full diversity order. In fact, all relay selection schemes in [42–45] were based on the channel quality of links and the buffers' states were not considered. As such, in Chapter 3, we suggest a novel buffer state-based relaying scheme for HD BA DF multi-hop communications to further improve the performance. We analyze the proposed scheme in a theoretical framework using a MC formulation and derive closed-form analytical expressions for the APD, OP and diversity order.
2. In Chapter 4, we propose two more parameterized relaying techniques that achieve further performance improvements and levels of tradeoff between the average packet delay and the diversity order. We compare these schemes and deduce the adequate choices of the system parameters to optimize the performance.
3. Further setups of serial and parallel relaying accounting for interference constraints, inter-relay cooperation and variable transmission rates not investigated in the literature are considered in Chapter 5. We propose a deep RL algorithm to deduce the optimal relay selection that can achieve the maximum throughput and the minimum average packet delay. We compare the different setups and draw conclusions about the effect of the system's variables on the overall performance.

1.6 List of Publications

- "Buffer State Based Relay Selection for Half-Duplex Buffer-Aided Serial Relaying Systems," accepted for publication in IEEE Transactions on Communications.
- "Relaying Strategies for Half-Duplex Buffer-Aided Serial Relaying Systems," submitted/under revision in IEEE Communications Letters.
- "Optimized Relay Selection for Multi-Hop Cooperative Systems Using Deep Reinforcement Learning," under preparation.

Chapter 2

Analysis Techniques

2.1 Reinforcement Learning (RL)

2.1.1 Markov Decision Process

A Markov decision process (MDP) is a prime component for the reinforcement learning. A MDP is a stochastic model defined by a set of states, a set of actions and the one-step dynamics of the environment [59]. In other words, at time t , given a state s and assuming an action a is taken, the environment will output a reward r and move to the next state s' with a certain probability. The probability distribution defined in (2.1) for all possible states and actions of the system characterizes the MDP.

$$p(s', r|s, a) = \Pr(S_{t+1} = s', R_t = r|S_t = s, A_t = a) \quad (2.1)$$

2.1.2 RL Agent

RL is a branch of AI that is different from supervised and unsupervised learning. It is applied on Markov Decision Processes and if converges, results in an optimal policy that maximizes the reward function. The RL training approach is described in Fig. 6 where the agent at time slot t , given the current state S_t chooses the action A_t based on (2.2).

$$A_t = \arg \max_a Q(S_t, a) \quad (2.2)$$

where $Q(S_t, a)$ stands for the Q-value of selecting action a at the state S_t .

Next, the environment processes this action and outputs the new state S_{t+1} and a reward value R_t . This forms a new training experience expressed as (S_t, A_t, S_{t+1}, R_t) and accordingly the Q-values will be updated according to the Bellman equation (2.3).

$$Q(S_t, A_t) = R_t + \gamma * \max_a Q(S_{t+1}, a) \quad (2.3)$$

The parameter γ stands for the discount factor that determines the impact of future rewards on the learning process.

The Q-values are determined and processed by the agent using either a Q-table (Q-learning) or a neural network (Deep Q-learning) until a final function $S_t \rightarrow A_t$ converges.

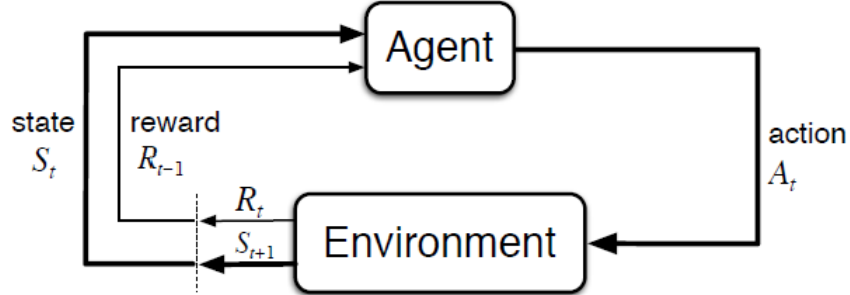


Figure 6: The agent-environment interaction in RL

2.1.3 Q-Learning

For M states and N actions, the Q-learning table will be of size $M \times N$ as illustrated in Fig. 7. For every state, the Q-value of every action is reported in this table and after every iteration these values are updated according to (2.3).

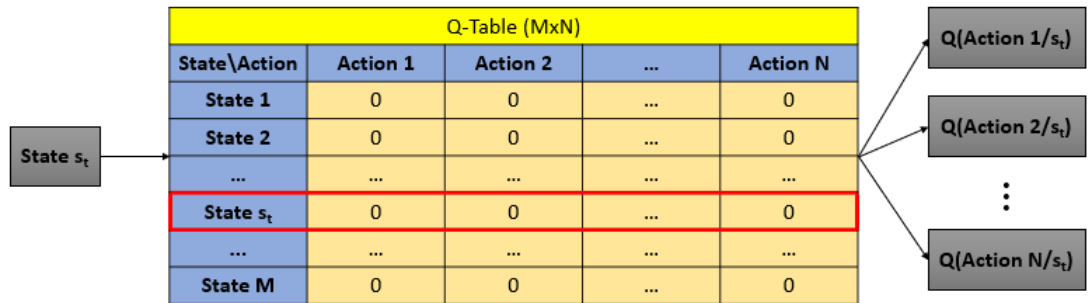


Figure 7: Q-learning Approach

2.1.4 Deep Q-Learning

For large sets of states and actions, the Q-table will become huge and the convergence will be slower. As an alternative, a neural network (NN) is trained to output the Q-values of actions at every state as shown in Fig. 8. In fact, two NN are required for the training, a prediction network and a target network. The

maximum future reward at S_{t+1} is computed from the target network to get the new Q-value in (2.3). According to this new Q-value and the old one (from the prediction network), a cost function is computed and the prediction NN will be updated.

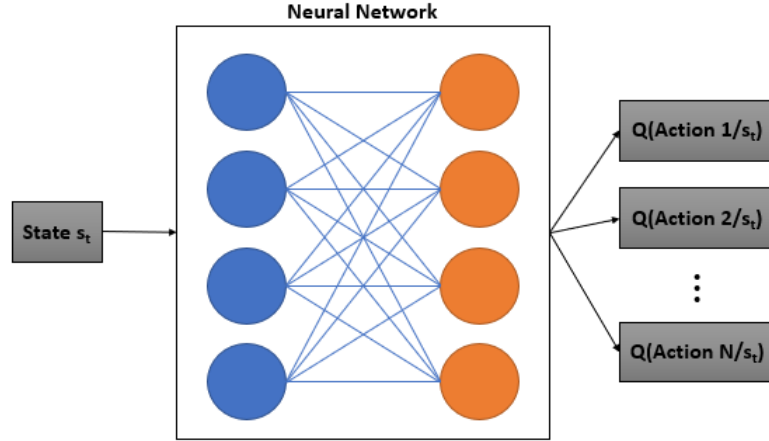


Figure 8: Deep Q-learning Approach

2.1.5 Handling Unfeasible Actions

At certain states, some actions might be unfeasible; i.e. the environment cannot process such actions. An example could be the action of activating a given link while in fact it is in outage. To handle these cases, different approaches were proposed in the literature:

- **Punishment:** Assigning a negative reward whenever an unfeasible action is chosen. This approach is good enough for Q-learning since it allocates a negative value at the corresponding entries of the Q-table. However, in DQN, the NN tries to deduce a smooth function that computes the Q-values, and unfeasible actions for some states can be the best actions for others, hence, this approach can be too harsh for DQN.
- **Decision Assisted DQN:** This approach changes the result of (2.3) to zero for unfeasible actions which is the actual final value of these Q-values. This approach can accelerate the convergence but, however, does not guarantee the selection of feasible actions especially for systems with a large set of unfeasible actions.
- **A-priori Information:** the agent can be given some a-priori information about unfeasible actions so that it can remove them from the action space at each time slot. This approach can accelerate the convergence and the elimination of unfeasible actions is guaranteed.

- **Pretraining:** The designer can preset some desirable actions at specific states and train the agent based on this specific dataset. Next, the agent will be trained according to the general approach described in Section 2.1.2. This approach can accelerate the convergence but, however, it can deviate the agent from converging to the best solution but instead, converge to what the designer thinks is a good solution.

2.2 Markov Chain Model

A Markov Chain (MC) model is a special case of an MDP where actions and rewards are eliminated; i.e. only one action is possible at every state and all rewards are set to zero. An MC model describes transition probabilities $p(s'/s)$ between the defined states of a system. This model is useful in wireless communication systems since probability distributions of different types of channels can be used. It is defined by a determined number of states and a transition matrix describing the transition probabilities between every two states. A variety of mathematical tools can be applied on these models to deduce the steady-state probabilities.

2.3 Asymptotic Analysis

This is an approximation approach of steady-state distributions of a MC model. It is based on the simplification of the transition probabilities so that some probabilities will tend to zero at certain conditions of the system (for example, at high SNR). A closed subset of the states will be formed, that is the probabilities of leaving this subset all tend to zero as described in Fig. 9. This allows to perform the analysis on a smaller number of states and to derive closed-form expressions of the system metrics.

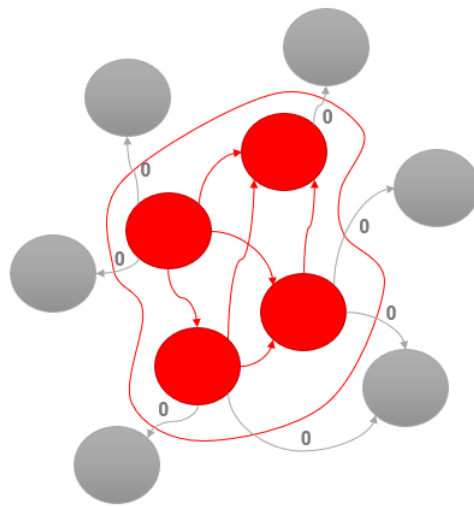


Figure 9: Closed Subset Approach

Chapter 3

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

3.1 Objectives

- 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.

3.2 System Model and Relaying Strategy

3.2.1 Basic Parameters

The system consists of a serial relaying network that involves $K + 2$ nodes comprising K relay nodes denoted by R_1, \dots, 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. 5. 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.

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. 5. 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:

$$p_k = \Pr \left\{ \frac{1}{K+1} \log_2(1 + \bar{\gamma}|h_k|^2) \leq r_0 \right\} = 1 - e^{-\frac{2^{(K+1)r_0}-1}{\Omega_k \bar{\gamma}}}, \quad (3.1)$$

where $\bar{\gamma}$ denotes the average transmit signal-to-noise ratio (SNR). In (3.1), 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:

$$q_k(l_1, \dots, l_K) = p_k + (1 - p_k) \times \begin{cases} \delta_{l_1=L}, & k = 1; \\ \delta_{l_{k-1}=0} + \delta_{l_k=L} - \delta_{l_{k-1}=0} \delta_{l_k=L}, & k = 2, \dots, K; \\ \delta_{l_K=0}, & k = K + 1. \end{cases} \quad (3.2)$$

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

3.2.2 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:

$$\hat{k} = \arg \max_{k \in \mathcal{L}_a} \{\Delta_k\}, \quad (3.3)$$

denoting that, in the corresponding time slot, $R_{\hat{k}-1}$ must transmit and $R_{\hat{k}}$ must receive. In (3.3), $\mathcal{L}_a \subset \{1, \dots, K + 1\}$ 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 (3.4).

$$\Delta_k = \begin{cases} s, & k = 1; \\ l_{k-1}, & k = 2, \dots, K + 1. \end{cases} \quad (3.4)$$

The rationale behind (3.4) 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 [29, 30, 35, 36, 39, 42–45, 48–51]. 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 (3.1), 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 $\{1, \dots, L\}$ to have com-

parable values with the buffer lengths l_1, \dots, 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 (3.3), 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 $k = K, \dots, 1$.
 - (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 (3.3).
- 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 $k = 0, \dots, K - 1$.
- Consequently, all nodes R_0, \dots, R_K have acquired the index of the best link and the corresponding node (if any) can initiate the data transmission.

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

$$\begin{aligned}
 r_k &= \max\{\Delta_{k+1}s_{k+1}, r_{k+1}\} \\
 i_k &= \begin{cases} k + 1, & \Delta_{k+1}s_{k+1} > r_{k+1}; \\ i_{k+1}, & \text{otherwise.} \end{cases}, \quad (3.5)
 \end{aligned}$$

where $r_{K+1} = 0$, $i_{K+1} = 0$, Δ_k is given in (3.4) 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 [39]:

- 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 $\{(r_k, i_k)\}_{k=1}^K \subset \{0, \dots, L\} \times \{0, 2, \dots, K+1\}$. Similarly, the scheme in [39] 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 [39], 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 [39] 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 (3.5). 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 [44], 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.

3.3 Performance Analysis

3.3.1 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, \dots, l_K) . The number of states in this MC is $(L + 1)^K$ in total since $l_k \in \{0, \dots, L\}$ for $k = 1, \dots, K$. 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 3.3.4, 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, \dots, l_K) to the state (l'_1, \dots, l'_K) is denoted by $t_{(l_1, \dots, l_K), (l'_1, \dots, l'_K)}$. 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:

$$\begin{aligned} \mathbf{T}_{i,j} &= t_{(l_1, \dots, l_K), (l'_1, \dots, l'_K)} \quad ; \\ i &= \mathfrak{N}(l'_1, \dots, l'_K), \quad j = \mathfrak{N}(l_1, \dots, l_K), \end{aligned} \tag{3.6}$$

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

We denote by π_{l_1, \dots, l_K} the steady-state probability of the system being in the state (l_1, \dots, l_K) . These steady-state probabilities are calculated as follows [29]:

$$\pi = (\mathbf{T} + \mathbf{B} - \mathbf{I})^{-1} \mathbf{b}, \quad (3.7)$$

where π is a $(L+1)^K$ -dimensional vector and its j -th element is equal to π_{l_1, \dots, l_K} with $j = \mathfrak{N}(l_1, \dots, l_K)$. In (3.7), \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, \dots, l_K) an outage occurs with the probability $\prod_{k=1}^{K+1} q_k(l_1, \dots, l_K)$ following from (3.2). The steady state probabilities allow then the calculation of the outage probability as follows::

$$OP = \sum_{l_1=0}^L \cdots \sum_{l_K=0}^L \pi_{l_1, \dots, l_K} \prod_{k=1}^{K+1} q_k(l_1, \dots, l_K). \quad (3.8)$$

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 [44] and Little's law [60]:

$$APD = \frac{K + OP + (K+1)\bar{L}}{1 - OP}, \quad (3.9)$$

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

$$\bar{L} = \sum_{l_1=0}^L \cdots \sum_{l_K=0}^L \pi_{l_1, \dots, l_K} \left[\sum_{k=1}^K l_k \right]. \quad (3.10)$$

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.

3.3.2 State Transition Matrix

In what follows, the unavailability probabilities in (3.2) will be written as q_k for simplicity. We will denote the state by $\mathbf{l} = (l_1, \dots, l_K)$, the set of all relays by $\mathcal{A} = \{1, \dots, K\}$ 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 un-

available:

$$t_{1,1} = \prod_{k=1}^{K+1} q_k. \quad (3.11)$$

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

$$a_k = \begin{cases} t_{1,1+e_k}, & k = 1; \\ t_{1,1+e_k-e_{k-1}}, & k = 2, \dots, K; \\ t_{1,1-e_{k-1}}, & k = K + 1. \end{cases} \quad (3.12)$$

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:

$$a_k = (1 - q_k) \sum_{\mathcal{K} \subset \mathcal{A} \setminus \{k\}} \left[\prod_{i \in \mathcal{K}} (1 - q_i) \right] \left[\prod_{j \in \mathcal{A} \setminus \{k\} \cup \mathcal{K}} q_j \right] Q_{k,\mathcal{K}}, \quad (3.13)$$

where the set \mathcal{K} comprises the indices of the links, other than the link k , that are available. In (3.13), $Q_{k,\mathcal{K}}$ 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 $\Delta_{k'}$:

$$Q_{k,k'} = \delta_{k' < k} \delta_{\Delta_k \geq \Delta_{k'}} + \delta_{k' > k} \delta_{\Delta_k > \Delta_{k'}} \quad ; \quad k' \neq k, \quad (3.14)$$

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 (3.13) can be developed as (3.15) that further simplifies into (3.16).

$$\begin{aligned} a_k = & (1 - q_k) \left[\prod_{i=1, i \neq k}^{K+1} q_i + \sum_{k_1=1, k_1 \neq k}^{K+1} (1 - q_{k_1}) \left[\prod_{j=1, j \neq k, j \neq k_1}^{K+1} q_j \right] Q_{k,k_1} \right. \\ & \left. + \sum_{k_1=1, k_1 \neq k}^{K+1} \sum_{k_2=k_1+1, k_2 \neq k}^{K+1} (1 - q_{k_1})(1 - q_{k_2}) \left[\prod_{j=1, j \neq k, j \neq k_1, j \neq k_2}^{K+1} q_j \right] Q_{k,k_1} Q_{k,k_2} + \dots \right]. \end{aligned} \quad (3.15)$$

$$a_k = (1 - q_k) \left[\prod_{i=1, i \neq k}^{K+1} q_i \right] \left[1 + \sum_{k_1=1, k_1 \neq k}^{K+1} \frac{(1 - q_{k_1})}{q_{k_1}} Q_{k, k_1} \left[1 + \sum_{k_2=k_1+1, k_2 \neq k}^{K+1} \frac{(1 - q_{k_2})}{q_{k_2}} Q_{k, k_2} \right. \right. \\ \left. \left. \left[1 + \dots \left[1 + \sum_{k_K=k_{K-1}+1, k_K \neq k}^{K+1} \frac{(1 - q_{k_K})}{q_{k_K}} Q_{k, k_K} \right] \right] \right] \right], \quad (3.16)$$

Equation (3.16) can be implemented recursively resulting in the following expression:

$$a_k = (1 - q_k) \left[\prod_{i=1, i \neq k}^{K+1} q_i \right] [1 + f_r(\mathcal{A}, k, 0)], \quad (3.17)$$

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

Function: $f_r(\mathcal{Y}, k, a)$
Data: $\mathcal{Y} \subset \mathcal{A}$, $k \in \{1, \dots, K + 1\}$ and $a \in \{0, \dots, K + 1\}$;
Result: *Sum*;
initialization: $Sum = 0$;
if $a + 1 > |\mathcal{Y}|$ **then**
 | return 0
end
for $m = a + 1 : |\mathcal{Y}|$ **do**
 | $k' = \mathcal{Y}_m$ (m -th element of \mathcal{Y})
 | **if** $k' \neq k$ **then**
 | | $Sum = Sum + \frac{(1 - q_{k'})}{q_{k'}} Q_{k, k'} [1 + f_r(\mathcal{Y}, k, m)]$
 | **end**
end

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

3.3.3 Asymptotic Analysis

Using (3.17) to evaluate the transition probabilities in (3.12) then stacking these probabilities in the state transition matrix to determine the steady-state probabilities in (3.7) 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 (3.17), (ii): the large number of states that can a state $\mathbf{1}$ transit to according to (3.12) and (iii): the need to invert a $(L + 1)^K \times (L + 1)^K$ matrix in (3.7) 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 $\mathcal{S} \triangleq \{0, \dots, L\}^K$ 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, $\sum_{\mathbf{l} \in \mathcal{S}_c} \pi_{\mathbf{l}} \rightarrow 1$ while $\pi_{\mathbf{l}} \rightarrow 0 \forall \mathbf{l} \notin \mathcal{S}_c$ for $\bar{\gamma} \gg 1$ where the steady-state probabilities satisfy (3.7). The set \mathcal{S}_c is called the closed subset where the probability of exiting this set tends to zero asymptotically:

$$t_{\mathbf{l}, \mathbf{l}'} \rightarrow 0 \quad \forall \mathbf{l} \in \mathcal{S}_c, \mathbf{l}' \notin \mathcal{S}_c. \quad (3.18)$$

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 (3.4). 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:

$$\begin{aligned} \mathbf{s}_1^{(1)} &= (s - 1, \dots, s - 1), \\ \mathbf{s}_2^{(1)} &= (s + 1, s - 1, \dots, s - 1), \\ \mathbf{s}_3^{(1)} &= (s - 1, \dots, s - 1, s - 2), \\ \mathbf{s}_n^{(2)} &= (\underbrace{s - 1, \dots, s - 1}_{n-1 \text{ times}}, \underbrace{s, s - 1, \dots, s - 1}_{K-n \text{ times}}), \\ \mathbf{s}_n^{(3)} &= (s, \underbrace{s - 1, \dots, s - 1}_{n-1 \text{ times}}, \underbrace{s, s - 1, \dots, s - 1}_{K-n-1 \text{ times}}), \\ \mathbf{s}_n^{(4)} &= (\underbrace{s - 1, \dots, s - 1}_{n-1 \text{ times}}, \underbrace{s, s - 1, \dots, s - 1}_{K-n-1 \text{ times}}, s - 2), \end{aligned} \quad (3.19)$$

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

3.3.3.1 Case 1

$1 < s < L$:

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

$$\begin{aligned} \mathcal{S}_c &= \{\mathbf{s}_n^{(1)} ; n = 1, 2, 3\} \cup \{\mathbf{s}_n^{(2)} ; n = 1, \dots, K\} \\ &\quad \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} ; n = 1, \dots, K - 1\}, \end{aligned} \quad (3.20)$$

where the corresponding steady-state probabilities are given by:

$$\left\{ \begin{array}{l} \pi_{\mathbf{s}_1^{(1)}} = \frac{1 - \sum_{k=2}^{K+1} p_k}{K+1} \\ \pi_{\mathbf{s}_2^{(1)}} = \frac{p_2}{K+1} \\ \pi_{\mathbf{s}_3^{(1)}} = \frac{p_1}{K+1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1 - \sum_{k=1}^n p_k}{K+1}, \quad \text{for } n = 1, \dots, K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=2}^{n+1} p_k}{K+1}, \quad \text{for } n = 1, \dots, K-1 \\ \pi_{\mathbf{s}_n^{(4)}} = \frac{p_1}{K+1}, \quad \text{for } n = 1, \dots, K-1. \end{array} \right. \quad (3.21)$$

Proof. The above proposition is proved in Appendix A. \square

3.3.3.2 Case 2

$s = 1$:

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

$$\mathcal{S}_c = \{\mathbf{s}_n^{(1)} ; n = 1, 2\} \cup \{\mathbf{s}_n^{(2)} ; n = 1, \dots, K\} \cup \{\mathbf{s}_n^{(3)} ; n = 1, \dots, K-1\}, \quad (3.22)$$

with the following steady-state probabilities:

$$\left\{ \begin{array}{l} \pi_{\mathbf{s}_1^{(1)}} = \frac{1 - \sum_{k=2}^{K+1} p_k}{K+1 - Kp_1} \\ \pi_{\mathbf{s}_2^{(1)}} = \frac{p_2}{K+1 - Kp_1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1 - \sum_{k=1}^n p_k}{K+1 - Kp_1}, \quad \text{for } n = 1, \dots, K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=2}^{n+1} p_k}{K+1 - Kp_1}, \quad \text{for } n = 1, \dots, K-1. \end{array} \right. \quad (3.23)$$

Proof. The details of the proof are presented in Appendix B. \square

3.3.3.3 Case 3

$s = L$:

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

$$\mathcal{S}_c = \{\mathbf{s}_1^{(1)}, \mathbf{s}_3^{(1)}\} \cup \{\mathbf{s}_n^{(2)} ; n = 1, \dots, K\} \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} ; n = 1, \dots, K-1\}, \quad (3.24)$$

$$\begin{cases} \pi_{\mathbf{s}_1^{(1)}} = \frac{1 - \sum_{k=3}^{K+1} p_k}{K+1} \\ \pi_{\mathbf{s}_3^{(1)}} = \frac{p_1}{K+1} \\ \pi_{\mathbf{s}_n^{(2)}} = \frac{1 - \sum_{k=1}^n p_k}{K+1}, \quad \text{for } n = 1, \dots, K \\ \pi_{\mathbf{s}_n^{(3)}} = \frac{\sum_{k=3}^{n+1} p_k}{K+1}, \quad \text{for } n = 1, \dots, K-1 \\ \pi_{\mathbf{s}_n^{(4)}} = \frac{p_1 + p_2}{K+1}, \quad \text{for } n = 1, \dots, K-1. \end{cases} \quad (3.25)$$

Proof. The details of the proof are presented in Appendix C. \square

In (3.21), (3.23) and (3.25), 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 (3.21) add up to one. The same holds for the probabilities in (3.23) and the probabilities in (3.25).

Assuming $L \geq 5$ and replacing (3.21), (3.23) and (3.25) in (3.8) implies the expressions of the asymptotic OP provided in (3.26).

$$OP_{\text{Asymp}} = \begin{cases} \frac{1 - \sum_{k=2}^{K+1} p_k}{K+1 - Kp_1} p_1 + \sum_{n=1}^K \frac{1 - \sum_{k=1}^n p_k}{K+1 - Kp_1} p_1 p_{n+1} \\ + \frac{p_2}{K+1 - Kp_1} p_1 p_2 + \sum_{n=2}^K \frac{\sum_{k=2}^{n+1} p_k}{K+1 - Kp_1} p_1 p_2 p_{n+1} \\ [1 - \frac{Kp_1}{K+1}] \prod_{k=1}^{K+1} p_k + \frac{Kp_1}{K+1} \prod_{k=1}^K p_k, & s = 1 \\ \prod_{i=1}^{K+1} p_i, & s = 2 \\ [1 - \frac{p_2}{K+1}] \prod_{i=1}^{K+1} p_i + \frac{p_2}{K+1} \prod_{i=2}^{K+1} p_i, & 2 < s < L-1 \\ \frac{1 + p_1 - \sum_{k=3}^{K+1} p_k}{K+1} \prod_{i=1}^{K+1} p_i + \sum_{n=1}^K \frac{1 - \sum_{k=1}^n p_k}{K+1} \prod_{i=1, i \neq n}^{K+1} p_i \\ + \sum_{n=2}^K \frac{\sum_{k=3}^{n+1} p_k}{K+1} \prod_{i=2, i \neq n}^{K+1} p_i + \sum_{n=1}^{K-1} \frac{p_1 + p_2}{K+1} \prod_{i=1, i \neq n}^{K+1} p_i, & s = L-1 \\ & s = L. \end{cases} \quad (3.26)$$

It is worth highlighting that the asymptotically-dominant states in (3.19) 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 (3.2) satisfy $q_k = p_k$ for $k = 1, \dots, K+1$ for all states in (3.19) 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 (3.20). Therefore, the product of the unavailability

probabilities in (3.8) 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 (3.26) after replacing the steady-state probabilities by their values from (3.21) and observing that $\sum_{\mathbf{l} \in \mathcal{S}_u} \pi_{\mathbf{l}} = \frac{p_1}{K+1} + (K-1) \frac{p_1}{K+1} = \frac{K p_1}{K+1}$.

The asymptotic OP expressions in (3.26) 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 (3.1) scales as $\bar{\gamma}^{-1}$) generating a diversity order of n . Consequently, the system's diversity order DO is found to be:

$$DO = \begin{cases} 1, & s = 1; \\ K + 1, & 1 < s < L; \\ K, & s = L. \end{cases} \quad (3.27)$$

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 (3.21), (3.23) and (3.25). In fact, it was observed that this approach yields to a simple asymptotic APD expression that is highly accurate. Setting $p_1 = \dots = p_{K+1} = 0$ in (3.21), (3.23) and (3.25) results in $\pi_{\mathbf{s}_1^{(1)}} = \pi_{\mathbf{s}_1^{(2)}} = \dots = \pi_{\mathbf{s}_K^{(2)}} = \frac{1}{K+1}$ for all values of s while other steady-state probabilities can be ignored. Therefore, the average queue length in (3.10) 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 (3.19). Replacing this value of \bar{L} in (3.9) while ignoring OP that is very small asymptotically implies the below expression of the asymptotic APD that holds for all values of s :

$$APD_{\text{Asymp}} = 2K + (s-1)K(K+1). \quad (3.28)$$

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

Equation (3.28) 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 [44], 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 [16, 29, 30, 35, 36, 39, 44, 48, 51]. 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 [39]. 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 chapter differ substantially from those presented in [39]. The role of the source node S also differs substantially between [39] and the current work. A main challenge in the MC analysis performed in this chapter resides in quantifying the role of S via a parameter s that was introduced in the link selection protocol in (3.4). As such, S has to compete with the relays for transmitting unlike [39]. 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 [39] where the closed subset contained only four states for any number of relays K , the asymptotic MC analysis presented in this thesis 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 (3.21), (3.23) and (3.25) (as compared to eq. (44) in [39]) 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 (3.26) with eq. (45) in [39].

3.3.4 Conclusions about the design of the BA relaying scheme

Following from (3.26)-(3.28), 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 (3.28), such large values of s penalize the APD while not presenting any advantage in terms of the diversity order as can be observed from (3.27).
- 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 (3.26). 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 3.3.3 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 3.3.3 with $s \in \{1, 2, 3\}$ holds for infinite-size buffers as well. In fact, the finite set of recurrent states in (3.19) will shape the asymptotic steady-state distribution of the MC even with an infinite number of states since all remaining states will be transient.

3.4 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 (3.1). In addition, we define the $(K + 1)$ -dimensional vector $\mathbf{\Omega}$ as $\mathbf{\Omega} = [\Omega_1, \dots, \Omega_{K+1}]$ capturing the strengths of the $K + 1$ hops.

Fig. 10 and Fig. 11 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 (3.26) and (3.28), 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 (3.8) and (3.9) 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 (3.27). This results in comparable OP performance where the corresponding OP curves are the steepest as can be observed from Fig. 10. However, from Fig. 11, increasing the value of s leads to higher APD values in coherence with (3.28). Among the above choices, the value $s = 3$ results in the smallest OP as predicted from (3.26); 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 3.3.4.

Fig. 12 and Fig. 13 make a comparison between the proposed scheme and the *max-link* scheme [44] 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 $\mathbf{\Omega} = [3, \dots, 3]$. 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 3.3.4. Results in Fig. 12 and Fig. 13 validate equations (3.26) and (3.28), 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 five. This observation does not hold for the *max-link* scheme where small buffer sizes present a small diversity

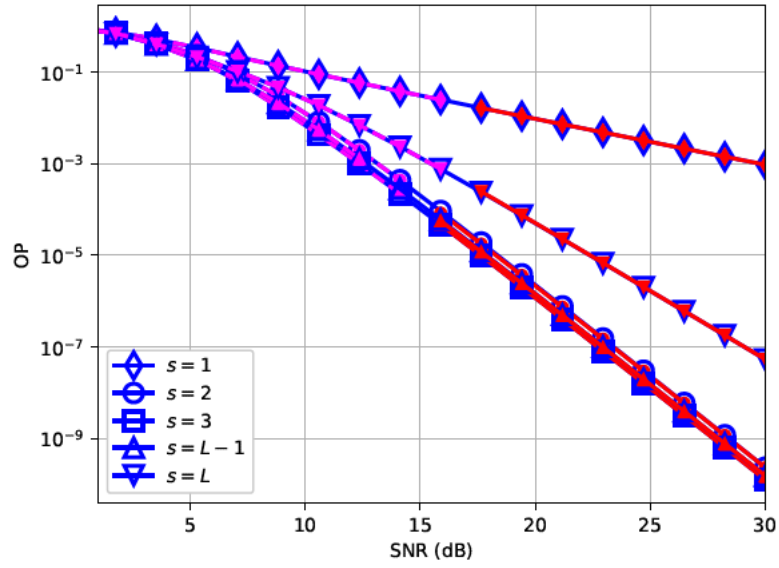


Figure 10: OP with $K = 3$ and $L = 8$. Dashed lines, solid lines with hollow markers and solid lines correspond to the simulation, theoretical and asymptotic values, respectively.

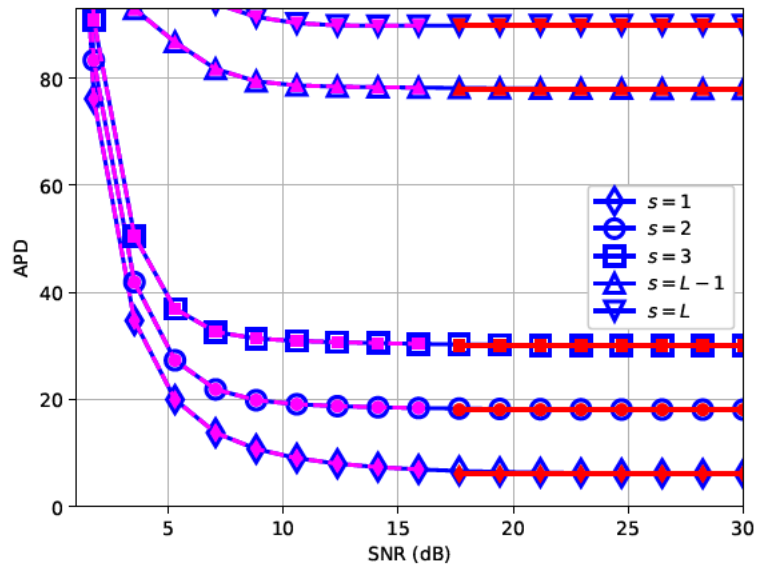


Figure 11: APD with $K = 3$ and $L = 8$. Dashed lines, solid lines with hollow markers and solid lines correspond to the simulation, theoretical and asymptotic values, respectively.

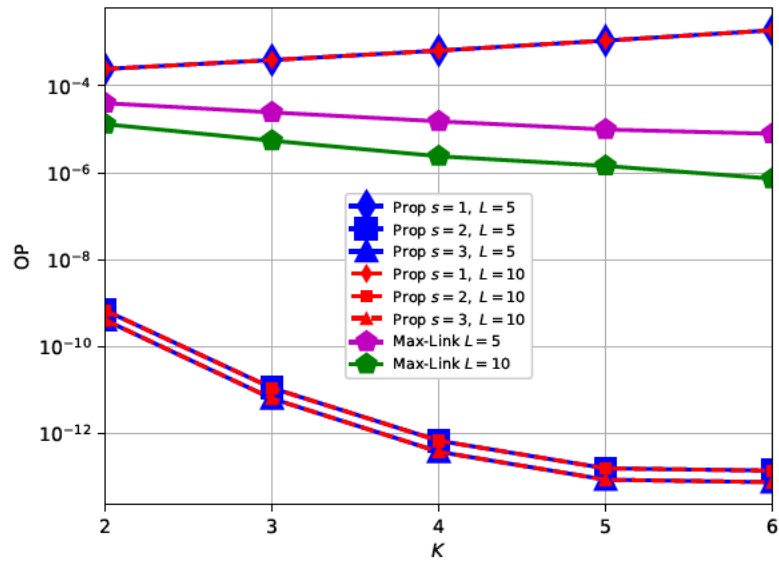


Figure 12: Asymptotic OP of the *max-link* scheme and the proposed scheme for different values of K .

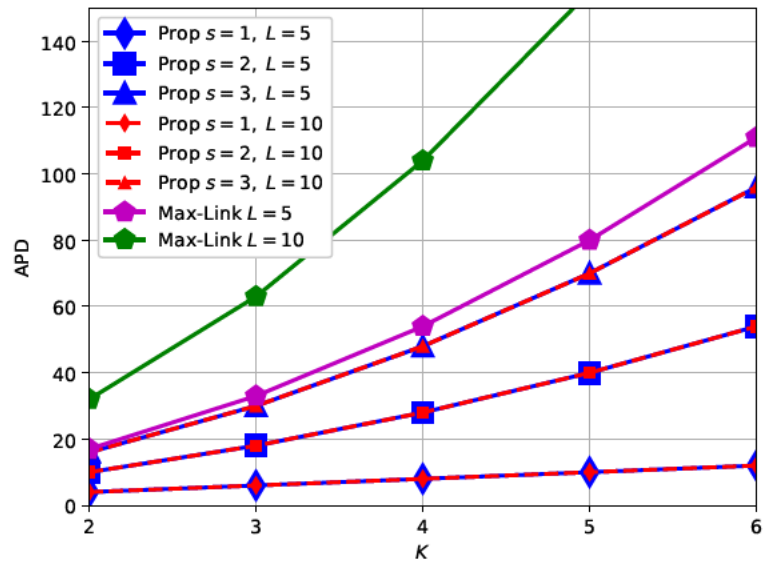


Figure 13: Asymptotic APD of the *max-link* scheme and the proposed scheme for different values of K .

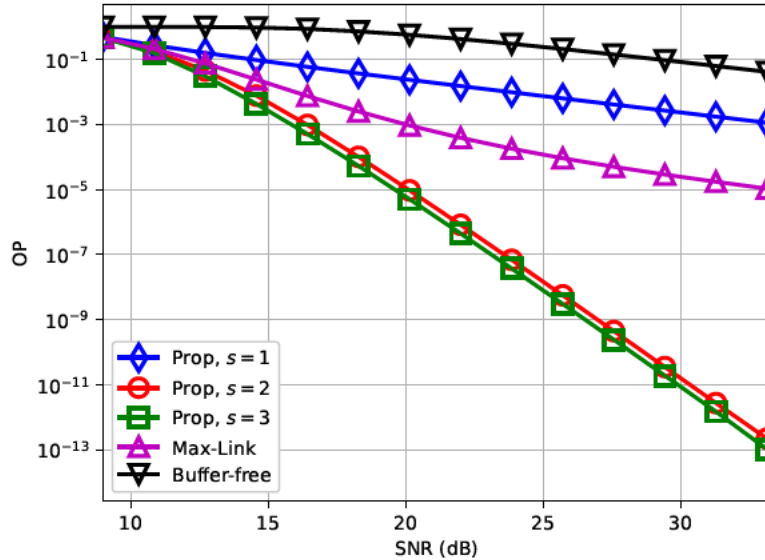


Figure 14: OP for $K = 5$ and $L = 5$.

order and full diversity is only reached with infinite buffer sizes [44]. 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 $L \rightarrow \infty$) 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. 12 and Fig. 13 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 [44] 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. 12 and Fig. 13, 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. 14 and Fig. 15 compare between the proposed scheme and the *max-link* scheme [44]. We added the performance of a buffer-free system whose OP is given

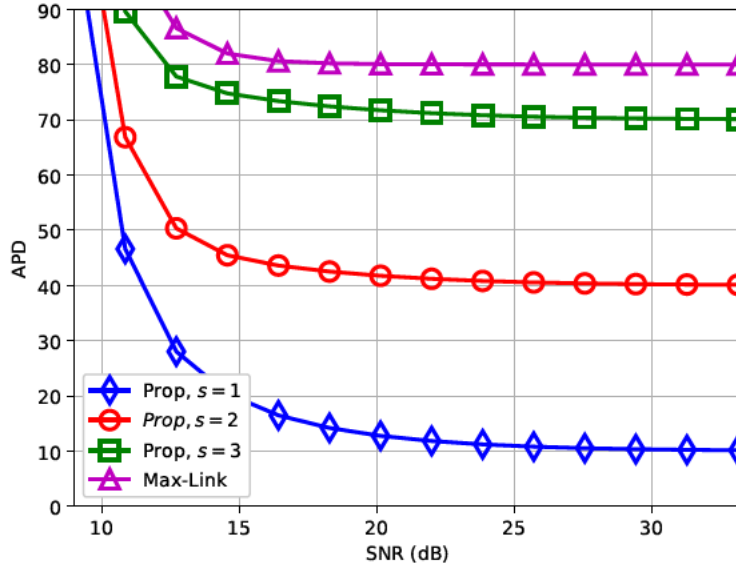


Figure 15: APD for $K = 5$ and $L = 5$.

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 $\Omega = [4.5, 4, 3.75, 5, 4.5, 4.5]$. Results in Fig. 14 and Fig. 15 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. 15. 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. 14. From Fig. 15, 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.

Chapter 4

Relaying Strategies for Half-Duplex Buffer-Aided Serial Relaying Systems

This chapter is based on the analysis done in Chapter 3. The system model is the same as the scenario considered in Chapter 3 and the relaying strategy proposed in (3.3) is further extended with different weights. In fact, the scheme in Chapter 3 suffers from the limitation that only the two diversity order (DO) and average packet delay (APD) pairs $(DO, APD) = (1, 2K)$ and $(DO, APD) = (K+1, K^2+3K)$ can be achieved. Leveraging this limitation, this chapter suggests three relaying schemes that can achieve a broader range of tradeoff levels between DO and APD thus offering an improved flexibility for wider ranges of 5G applications depending on their delay tolerance. The proposed schemes differ by the weights they assign to the hops where these weights depend on the lengths of the buffers at the transmitting and/or receiving nodes. Different weights must be assigned to the first and last hops since the source always has packets to transmit while the destination can accommodate any incoming packet. As such, this chapter also targets the optimization of these weights for the sake of achieving any target DO level with the minimal delay. The objective of designing the relaying schemes with the best DO-APD tradeoffs was accomplished through a MC analysis that led to simple closed-form expressions of the DO and asymptotic APD.

4.1 Relaying Strategies

The proposed relaying strategies are based on assigning a weight Δ_k to link k for $k = 1, \dots, K+1$ and activating the link \hat{k} (between $R_{\hat{k}-1}$ and $R_{\hat{k}}$) with the largest weight:

$$\hat{k} = \arg \max_{k \in \mathcal{L}_a} \{\Delta_k\}, \quad (4.1)$$

where $\mathcal{L}_a \subset \{1, \dots, K+1\}$ denotes the set of links that are available. If more than one link share the highest weight, the link closer to D will be selected to

reduce the packet delay.

The proposed relaying schemes account for both the channel states through limiting the selection among the available links in \mathcal{L}_a and the buffer states through the weights $\{\Delta_k\}_{k=1}^{K+1}$ that depend on the buffer lengths $\{l_k\}_{k=1}^K$ as will be highlighted later. The considered schemes differ by their choices of the weights that include either the buffer length at the transmitting node or the buffer length at the receiving node or both.

Scheme 1: Privileging the transmission from congested buffers, the weights $\{\Delta_k\}_{k=1}^{K+1}$ are selected as follows:

$$\Delta_k = \begin{cases} s, & k = 1; \\ l_{k-1}, & k = 2, \dots, K + 1. \end{cases}, \quad (4.2)$$

where the parameter s is the weight associated with link 1. As such, (4.2) differentiates between the cases where either S or a relay is the transmitting node. Note that scheme 1 is the same as the scheme in Chapter 3 but is revisited here for completeness and since it fits the general presented mathematical framework.

Scheme 2 privileges the reception at under-filled buffers by assigning the following weights:

$$\Delta_k = \begin{cases} -l_k, & k = 1, \dots, K; \\ -d, & k = K + 1. \end{cases}, \quad (4.3)$$

where d is a parameter associated with link $K + 1$.

Scheme 3 is based on the buffer lengths l_{k-1} and l_k at the transmitting and receiving nodes, respectively, as follows:

$$\Delta_k = \begin{cases} \alpha - l_1, & k = 1; \\ l_{k-1} - l_k, & k = 2, \dots, K; \\ l_K - \beta, & k = K + 1. \end{cases}, \quad (4.4)$$

where α and β are parameters associated with the first and last hops, respectively.

4.2 Performance Analysis

4.2.1 Generalities

The Markov Chain (MC) framework constitutes the appropriate mathematical tool for analyzing queues [44]. The discrete-time discrete-value Markov chain that models the considered serial-relaying network comprises the $(L + 1)^K$ states $(l_1, l_2, \dots, l_K) \in \{0, \dots, L\}^K$. The transition probabilities between the states can be determined by applying algorithm 1 in Chapter 3 since the relaying scheme in Chapter 3 as well as the schemes considered in this chapter adopt the same strat-

egy of assigning weights to the links and selecting the link of maximum weight. As such, the derivations of the transition probabilities will not be repeated in this chapter since the weights in (4.2)-(4.4) can be readily replaced in (3.14). The transition probabilities can then be used to derive the steady-state probabilities using standard MC techniques (3.7). This formulation directly yields the derivation of the outage probability (OP) as in (3.8). Equation (3.8) allows for the derivation of the diversity order (DO) defined as the negative slope of OP in function of $\bar{\gamma}$ on a log-log scale as $\bar{\gamma} \rightarrow \infty$. The queuing at the relays' buffers will incur delays in the system. Following from [44], the APD is expressed as in (3.9).

4.2.2 Asymptotic analysis

An exact MC analysis is intractable since the number of states increases exponentially with the number of relays K . As such, in this section we resort to an asymptotic analysis that holds for large values of the SNR. This analysis yields closed-form expressions of the DO and asymptotic APD which allow us to study the impact of the parameters s , d , α and β on the performance of the proposed relaying schemes.

4.2.2.1 Closed Subset

The asymptotic analysis is based on the following observation. For the three considered schemes, we observe and prove the existence of a set of $K + 1$ states $\mathcal{S} \triangleq \{\mathbf{s}_i\}_{i=1}^{K+1} \subset \{0, \dots, L\}^K$ such that the transitions $\mathbf{s}_1 \rightarrow \mathbf{s}_2$, $\mathbf{s}_2 \rightarrow \mathbf{s}_3$, \dots , $\mathbf{s}_K \rightarrow \mathbf{s}_{K+1}$ and $\mathbf{s}_{K+1} \rightarrow \mathbf{s}_1$ occur with a probability that tends to one asymptotically. The implications of this observation are as follows. (i): At high SNR, the MC will be confined in the subset \mathcal{S} where the probability of exiting this subset tends to zero asymptotically. As such, instead of deriving the steady-state probabilities for all $(L + 1)^K$ states of the MC, it is sufficient to derive these probabilities only for the $K + 1$ elements of \mathcal{S} . (ii): Since elements of \mathcal{S} are connected to each other in a loop-like structure, then the asymptotic steady-state probability of each element of \mathcal{S} is equal to $\frac{1}{K+1}$.

We denote by $\mathbf{s}_i(k) \in \{0, \dots, L\}$ the k -th element of \mathbf{s}_i which stands for the number of packets stored in B_k at steady-state. The ordered sequences of states of the closed subset \mathcal{S} for schemes 1 and 2 are presented in (4.5) and (4.6), respectively:

$$\mathbf{s}_i(k) = \begin{cases} s - 1 + \delta_{k=i} & 1 \leq i \leq K; \\ s - 1 & i = K + 1. \end{cases}, \quad (4.5)$$

$$\mathbf{s}_i(k) = \begin{cases} d - \delta_{k=K-i+1} & 1 \leq i \leq K; \\ d & i = K + 1. \end{cases}. \quad (4.6)$$

Table 1: Closed subset of Scheme 3 for $\gamma > 0$

$\mathbf{s}_i(k)$	$1 \leq k < \gamma$	$k = \gamma$	$\gamma < k \leq K$
$i = 1$	$\alpha - k$	$\alpha - \gamma$	$\alpha - \gamma$
$1 < i \leq \gamma$	$\alpha - k - \delta_{k=\gamma-i+1}$	$\alpha - \gamma + 1$	$\alpha - \gamma$
$i = \gamma + 1$	$\alpha - k$	$\alpha - \gamma + 1$	$\alpha - \gamma$
$\gamma + 1 < i \leq K + 1$	$\alpha - k$	$\alpha - \gamma$	$\alpha - \gamma + \delta_{k=i-1}$

Table 2: Closed subset of Scheme 3 for $\gamma < 0$

$\mathbf{s}_i(k)$	$1 \leq k \leq K - \gamma $	$k = K - \gamma + 1$	$K - \gamma + 1 < k \leq K$
$1 \leq i \leq K - \gamma $	$\alpha - \delta_{k=K- \gamma -i+1}$	$\alpha + 1$	$\alpha + \gamma - (K - k + 1)$
$K - \gamma < i < K + 1$	α	$\alpha + \delta_{k=i}$	$\alpha + \gamma - (K - k + 1) + \delta_{k=i}$
$i = K + 1$	α	α	$\alpha + \gamma - (K - k + 1)$

For scheme 3, the closed subset depends on the parameter $\gamma \triangleq \alpha - \beta$. For $\gamma = 0$, elements of \mathcal{S} are provided in (4.7). For $\gamma > 0$ and $\gamma < 0$, the values of $\mathbf{s}_i(k)$ are reported in Table 1 and Table 2, respectively.

$$\mathbf{s}_i(k) = \begin{cases} \alpha - \delta_{k=K-i+1} & 1 \leq i \leq K; \\ \alpha & i = K + 1. \end{cases} \quad (4.7)$$

In Appendix D, we prove that the transition $\mathbf{s}_i \rightarrow \mathbf{s}_{i+1}$ occurs with probability 1 for scheme 1 with $1 \leq i \leq K - 1$. Similar proofs hold for all schemes and for all values of $i \in \{1, \dots, K + 1\}$ thus justifying the results in (4.5)-(4.7) and Tables 1-2. These proofs are omitted for the sake of brevity.

4.2.2.2 DO and Asymptotic APD

Limiting the analysis to elements of \mathcal{S} that are the most probable states in the MC, the asymptotic OP can be written as: $OP = \frac{1}{K+1} \sum_{i=1}^{K+1} \prod_{k=1}^{K+1} q_k(\mathbf{s}_i)$ following from (3.8). Since the unavailability probability q_k in (3.2) can be equal either to p_k or 1, then the DO can be determined from:

$$DO = \min_{i=1, \dots, K+1} \left\{ \sum_{k=1}^{K+1} \delta_{q_k(\mathbf{s}_i)=p_k} \right\}, \quad (4.8)$$

since (3.1) behaves asymptotically as $\bar{\gamma}^{-1}$.

Since the asymptotic OP is several orders of magnitude smaller than K and L , then (3.9) can be approximated by $APD = K + (K + 1)\bar{L}$ with $\bar{L} = \frac{1}{K+1} \sum_{i=1}^{K+1} \sum_{k=1}^K \mathbf{s}_i(k)$.

Therefore, the asymptotic APD and DO of the three proposed schemes are as follows.

For scheme 1 denoted by $Sc^{(1)}(s)$:

$$APD^{(1)}(s) = 2K + (s - 1)K(K + 1), \quad (4.9)$$

$$DO^{(1)}(s) = \begin{cases} 1, & s = 1; \\ K + 1, & 1 < s < L; \\ K, & s = L. \end{cases} \quad (4.10)$$

For scheme 2 denoted by $Sc^{(2)}(d)$:

$$APD^{(2)}(d) = dK(K + 1), \quad (4.11)$$

$$DO^{(2)}(d) = \begin{cases} 1, & d = L; \\ K, & d = 1; \\ K + 1, & 1 < d < L. \end{cases} \quad (4.12)$$

For scheme 3, denoted by $Sc^{(3)}(\alpha, \beta)$, the asymptotic APD is given by:

$$APD^{(3)}(\alpha, \beta) = \begin{cases} \alpha K(K + 1) & , \gamma = 0; \\ 2K + \beta K(K + 1) + \\ (\gamma - 1) \left(\frac{K + 1}{2} \gamma - 1 \right) & , \gamma > 0; \\ \alpha K(K + 1) + \frac{\gamma^2 - \gamma}{2} + \\ K \left(\frac{\gamma^2 + \gamma + 2}{2} \right) & , \gamma < 0. \end{cases} \quad (4.13)$$

The DO depends on whether $\gamma = 0$, $\gamma > 0$ or $\gamma < 0$:

$$DO^{(3)}(\alpha, \beta) = \begin{cases} K, & \alpha = 1; \\ K + 1, & 1 < \alpha < L; \\ 1, & \alpha = L. \end{cases} ; \gamma = 0, \quad (4.14)$$

$$DO^{(3)}(\alpha, \beta) = \begin{cases} K, & (\alpha, \gamma) = (L, 1); \\ K + 1, & \alpha > \gamma; \\ \gamma, & \alpha = \gamma; \end{cases} ; \gamma > 0, \quad (4.15)$$

$$DO^{(3)}(\alpha, \beta) = \begin{cases} K, & \alpha \in \{1, L + \gamma\}; \\ K + 1, & \text{elsewhere}; \end{cases} ; \gamma < 0. \quad (4.16)$$

The asymptotic APD expressions in (4.9), (4.11) and (4.13) follow directly from (4.5)-(4.7) and Tables 1-2. In Appendix E, we provide highlights on how

the DO of scheme 2 in (4.12) can be derived. Similar derivations hold for the two remaining schemes and are not provided here for conciseness.

4.2.3 Conclusions about the design of the BA relaying schemes

Similar to the buffer lengths, the parameters s , d , α and β are natural integers. As such, γ assumes integer values.

For scheme 3, (4.14) and (4.16) show that the diversity orders of K and $K + 1$ can be achieved by $\gamma = 0$ and $\gamma < 0$. However, from (4.13), it can be observed that for the same value of α the APD is always smaller in the former case since $\frac{\gamma^2 - \gamma}{2} > 0$ and $\frac{\gamma^2 + \gamma + 2}{2} > 0$ when γ is a nonzero negative integer. Therefore, the choice $\gamma < 0$ presents no advantage compared to the choice $\gamma = 0$ and the former option can be omitted.

The proposed schemes are capable of achieving different levels of tradeoff between DO and APD as highlighted below.

4.2.3.1 $DO = 1$

(4.10) and (4.15) show that $Sc^{(1)}(1)$ and $Sc^{(3)}(1, 0)$ are capable of achieving this DO with an APD value of $2K$ following from (4.9) and (4.13). Similarly, (4.12) and (4.14) show that $Sc^{(2)}(L)$ and $Sc^{(3)}(L, L)$ can achieve $DO = 1$ but with an increased APD of $LK(K + 1)$.

Therefore, $Sc^{(1)}(1)$ and $Sc^{(3)}(1, 0)$ are the best delay-prioritizing schemes that achieve the smallest possible delay of $2K$ at the expense of a reduced $DO = 1$.

4.2.3.2 $DO = K + 1$

$Sc^{(1)}(2)$, $Sc^{(2)}(2)$, $Sc^{(3)}(2, 2)$ and $Sc^{(3)}(2, 1)$ are all capable of achieving this maximum DO where the corresponding delays are $K^2 + 3K$, $2K^2 + 2K$, $2K^2 + 2K$ and $K^2 + 3K$, respectively. Note that, from (4.15), other values of α and γ (with $\alpha > \gamma$) can result in $DO = K + 1$ when $\gamma > 0$. However, the choice $\alpha = 2$ and $\gamma = 1$ (implying that $\beta = 1$) is the best option since the APD in (4.13) increases with α and γ when $\gamma > 0$.

Therefore, $Sc^{(1)}(2)$ and $Sc^{(3)}(2, 1)$ are the best outage-prioritizing schemes that achieve the highest possible DO of $K + 1$ with an APD value of $K^2 + 3K$.

4.2.3.3 $DO = K$

This DO can be achieved by $Sc^{(1)}(L)$, $Sc^{(2)}(1)$, $Sc^{(3)}(1, 1)$ and $Sc^{(3)}(L, L - 1)$ with the smallest delay of $APD = K^2 + K$ achieved by $Sc^{(2)}(1)$ and $Sc^{(3)}(1, 1)$.

4.2.3.4 $DO \in \{2, \dots, K - 1\}$

(4.15) shows that only scheme 3 with $\gamma > 0$ can achieve such diversity orders when $\alpha = \gamma = DO$. Therefore, $Sc^{(3)}(DO, 0)$ must be applied with $DO < K$. The corresponding APD is $2K + (DO - 1)(\frac{K+1}{2}DO - 1)$ that increases with DO . Note that this delay must not exceed $K^2 + K$ since, otherwise, $Sc^{(2)}(1)$ and $Sc^{(3)}(1, 1)$ will be better options since they achieve a smaller delay while increasing the DO to K . Therefore, $Sc^{(3)}(DO, 0)$ is particularly appealing with large values of K where the achievable APD that increases linearly with K will fall below $K^2 + K$.

Based on the above discussion and on the observation that scheme 1 and scheme 2 are easier to implement compared to scheme 3 (since the weights assume simpler expressions), the following conclusions can be drawn. (i): Scheme 1 is a good option capable of covering the extreme cases $(DO, APD) = (1, 2K)$ and $(DO, APD) = (K + 1, K^2 + 3K)$ giving the full priority to the APD and DO, respectively. (ii): Scheme 2 is a suitable choice for achieving $(DO, APD) = (K, K^2 + K)$ where, compared to the full-diversity case, DO is reduced by 1 with the advantage of reducing the APD by $2K$. (iii): Scheme 3 demonstrates the highest flexibility and can achieve all levels of tradeoffs between DO and APD.

4.3 Numerical Results

In (3.1), we fix $r_0 = 1$. We define the vector $\mathbf{\Omega} = [\Omega_1, \dots, \Omega_{K+1}]$ that captures the strengths of the $K+1$ hops.

Fig. 16 shows the variations of the APD as a function of the number of relays K for $L = 5$, $\bar{\gamma} = 30$ dB and $\Omega_k = 3, \forall k = 1, \dots, K+1$. Results show that the APD values in (4.9), (4.11) and (4.13) match the simulation results thus demonstrating the accuracy of the closed-form asymptotic expressions. Fig. 16 shows that the proposed schemes can achieve a wide range of tradeoffs between DO and APD for any number of relays. Compared with $Sc^{(2)}(1)$, $Sc^{(3)}(3, 0)$ compromises the DO for the sake of achieving reduced APD values for $K \geq 5$. For $K = 5$, the former scheme is better since it achieves a larger DO and a smaller APD. Fig. 16 highlights on the increased delays of the *max-link* scheme in [44] despite the fact that this scheme achieves full diversity only with infinite-size buffers.

Fig. 14 and Fig. 15 show the OP and APD, respectively, for $K = 5$, $L = 5$ and $\mathbf{\Omega} = [2, 3, 2.5, 2, 3, 3.5]$. Results highlight on the adjustability of the proposed schemes and on the impact of the control parameters s , d and (α, β) on the OP-APD tradeoffs. At $DO = 1$, $Sc^{(1)}(1)$ and $Sc^{(3)}(1, 0)$ manifest exactly the same OP and APD performances. At $DO = K + 1$, $Sc^{(1)}(2)$ achieves smaller OP levels compared with $Sc^{(3)}(2, 1)$ at the expense of higher APD levels for small-to-average values of the SNR. At large SNRs, both schemes achieve the same

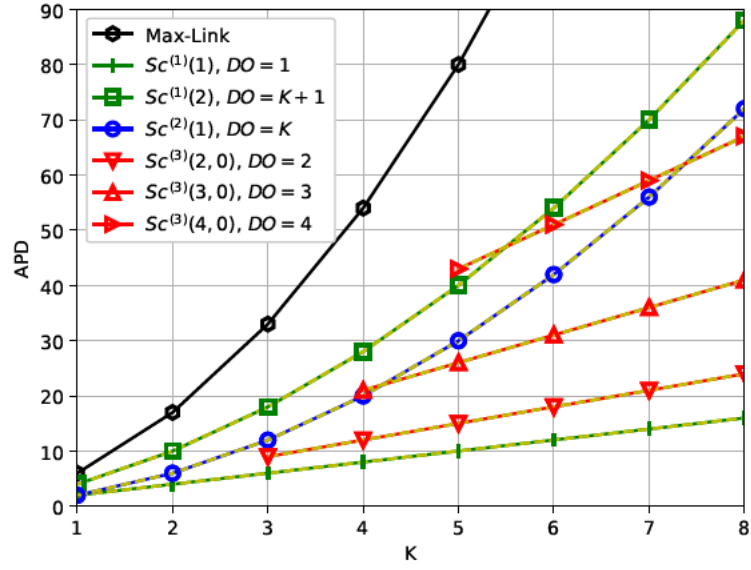


Figure 16: APD for different scenarios with $L=5$. Solid lines and dashed lines represent the analytical and simulation values, respectively.

asymptotic APD value of 40 as demonstrated in Section 4.2.3. At $DO = K$, $Sc^{(2)}(1)$ and $Sc^{(3)}(1,1)$ manifest comparable performances with a slight OP advantage and APD disadvantage of the former scheme. Finally, all eight variants of the proposed schemes in Fig. 14 and Fig. 15 result in significant APD reductions compared to [44].

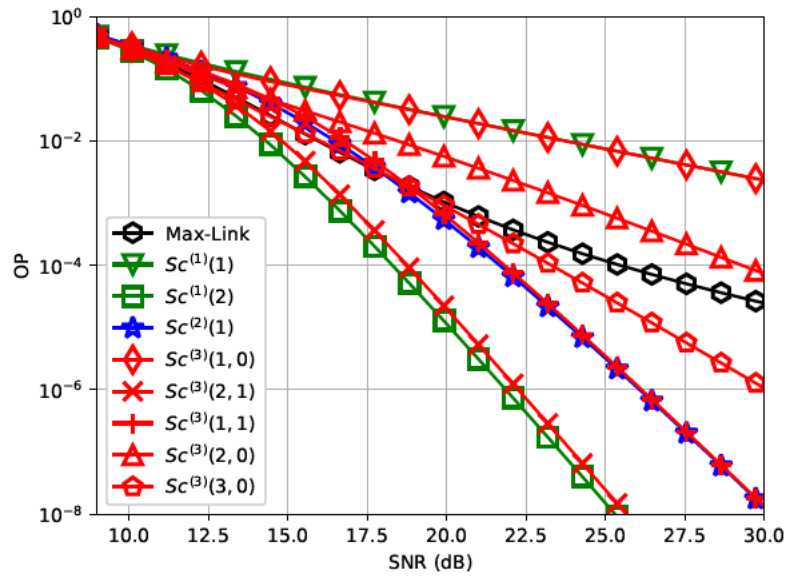


Figure 17: OP for K=5 and L=5.

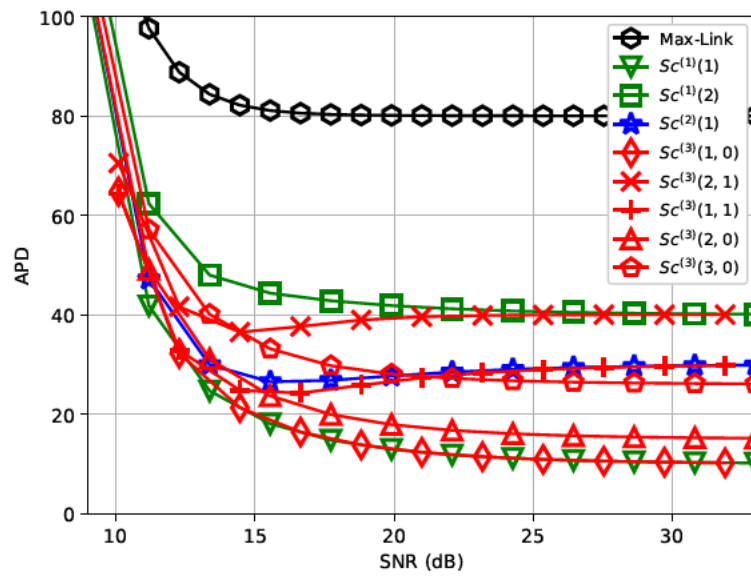


Figure 18: APD for K=5 and L=5.

Chapter 5

Optimized Relay Selection for Multi-Hop Cooperative Systems Using Deep Reinforcement Learning

The scenario considered in Chapters 3 and 4 presents a number of limitations: (1) the channel capacity is not used efficiently since for sufficiently large values of SNR, channels can allow for the transmission of multiple packets. (2) For large values of K , relays will be significantly distant so that multiple relays can transmit at the same time without interfering with each other. Thus, the interference constraint for HD relays becomes a limitation and might be loosened to gain some performance advantage.

To overcome these limitations, more complex setups for serial and parallel relaying are suggested in this chapter that allow for more degrees of freedom. The transmission of multiple packets and the activation of more than one link are introduced for serial setups whereas transmitting more than one packet and allowing for inter-relay cooperation are considered for parallel relaying setups.

Deterministic analysis of such networks is a challenging task because more states are involved and the transitions between these states are hard to be formulated within a MC model. Thus, RL techniques that are capable of analyzing large systems are considered in this chapter.

5.1 Objectives

- We formulate the problem of relay selection as a Markov Decision Process (MDP) for multiple setups of HD BA DF serial and parallel relaying.
- We optimize jointly the throughput and the average packet delay using a deep RL algorithm.

5.2 System Model

Different setups will be considered for serial and parallel relays. However, we assume the presence of K HD DF relays for all setups with no direct path linking S to D. Moreover, each relay is equipped with a buffer of size L . Rayleigh block fading channel is assumed as in previous chapters and all links are corrupted with AWGN.

5.2.1 Serial Relaying

5.2.1.1 Single Link Single Packet (SLSP)

This setup is equivalent to the scenario considered in Section 3.2. One link can be selected to be activated at every time slot and only one packet can be transmitted. The channel capacity C_k of link k ($k = 1, \dots, K + 1$) is given by (5.1) and outage happens if $C_k < r_0$.

$$C_k = \frac{1}{K + 1} \log_2(1 + \bar{\gamma}|h_k|^2) \quad (5.1)$$

The availability of a link is related to the buffer states: the buffer at the transmitting node should contain at least one packet and the buffer at the receiving node must not be full. The number of packets that can be transmitted is presented in (5.2).

$$n_{max} = \begin{cases} \min\{\delta_{C_k \geq r_0}, L - l_1\} & , \text{S-R}_1 \text{ link } (k = 1); \\ \min\{\delta_{C_k \geq r_0}, l_K\} & , \text{R}_K\text{-D link } (k = K + 1); \\ \min\{\delta_{C_k \geq r_0}, l_{k-1}, L - l_k\} & , \text{R}_{k-1}\text{-R}_k \text{ link for } k = 2, \dots, K. \end{cases} \quad (5.2)$$

where δ_S is 1 if the statement S is true and 0 otherwise.

From (5.2), the case of $l_{k-1} = 0$ means that the buffer at the sending relay is empty and hence, no packet can be transmitted over this link. Similarly, if $l_k = L$ then the number of remaining slots in the receiving relay is $L - l_k = 0$ and consequently, no additional packets can be accommodated at R_k . Finally, if $C_k < r_0$ then, $\delta_{C_k \geq r_0} = 0$ which means that the link k is in outage and cannot be activated.

5.2.1.2 Single Link Multiple Packets (SLMP)

The channel capacity is considered more carefully for this setup where n packets will be allowed to be transmitted through link k if the channel capacity C_k given in (5.1) satisfies (5.3). In other words, the number of transmitted packets will be

adapted to the channel quality where more than one packet can be transmitted along links with high SNR. However, $n' \leq n$ packets will be transmitted based on the channel capacity and the buffer states. The maximum possible number of packets follows (5.4).

$$C_k \geq n * r_0; \quad \text{with } n = \lfloor C_k/r_0 \rfloor \quad (5.3)$$

$$n_{max} = \begin{cases} \min\{\lfloor C_k/r_0 \rfloor, L - l_1\} & , \text{ S-R}_1 \text{ link } (k = 1); \\ \min\{\lfloor C_k/r_0 \rfloor, l_K\} & , \text{ R}_K\text{-D link } (k = K + 1); \\ \min\{\lfloor C_k/r_0 \rfloor, l_{k-1}, L - l_k\} & , \text{ R}_{k-1}\text{-R}_k \text{ link for } k = 2, \dots, K. \end{cases} \quad (5.4)$$

Moving from (5.2) to (5.4), the factor $\delta_{C_k \geq r_0}$ that limited the transmission to one packet is replaced by $\lfloor C_k/r_0 \rfloor$ to make use of the full capacity of the channel.

5.2.1.3 Multi Link Single Packet (MLSP)

The interference constraint for HD relays is loosened in this setup. It is assumed that when a relay R_k transmits a packet, it interferes only with the previous and next relays as illustrated in Fig. 19. This assumption holds in the scenario of long hops where the interference with distant nodes can be neglected. Accordingly, relay R_{k+2} cannot transmit simultaneously since in this case, it will interfere with R_{k+1} that is receiving a packet. R_{k+3} is the nearest relay that is allowed to transmit. Hence, for multiple link selection, every two nodes transmitting (source or relay) with indices k and k' should verify (5.5).

$$k' \geq k + 3; \quad k, k' \in \{0, 1, \dots, K \mid k < k'\} \quad (5.5)$$

In this setup, either no packet or a single packet can be transmitted along the selected links based on (5.2).

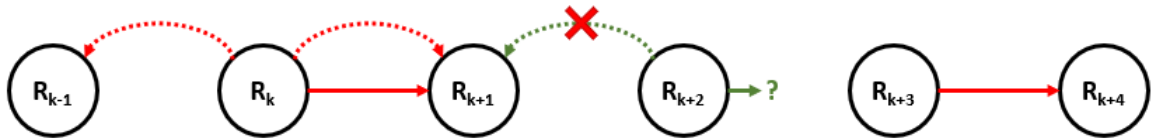


Figure 19: Interference Constraint for Multiple Links Activation

5.2.1.4 Multi Link Multiple Packets (MLMP)

This setup is a combination of SLMP and MLSP where multiple links can be selected to be activated and multiple packets per link can be transmitted. The

conditions in (5.4) and (5.5) should hold where up to n packets can be transmitted along the available links according to (5.3).

5.2.2 Parallel Relaying

5.2.2.1 Single Packet no Inter-Relay Cooperation (SPNI)

K parallel relays are placed between S and D and among $2K$ links, only one can be selected to carry a single packet. The channel capacity is different than the serial case, since every packet needs only two hops (one S-R and one R-D links) to reach D. Hence, the channel capacity C_k for parallel relays without inter-relay cooperation is given by (5.6).

$$C_k = \frac{1}{2} \log_2(1 + \bar{\gamma}|h_k|^2) \quad (5.6)$$

The link S- R_k (respectively R_k -D) is available if $C_k \geq r_0$ and the buffer at R_k is not full (respectively not empty). The maximum number of packets per link is given in (5.7).

$$n_{max} = \begin{cases} \min\{\delta_{C_k \geq r_0}, L - l_k\} & , \text{ S-}R_k \text{ link for } k = 1, \dots, K; \\ \min\{\delta_{C_k \geq r_0}, l_k\} & , \text{ }R_k\text{-D link for } k = 1, \dots, K. \end{cases} \quad (5.7)$$

where n_{max} can be either 0 or 1.

5.2.2.2 Multiple Packets no Inter-Relay Cooperation (MPNI)

Similarly to SLMP of the serial setup, n packets can be transmitted per link if the condition in (5.3) holds. Accounting for the buffer states, the maximum number of packets per link is given in (5.8).

$$n_{max} = \begin{cases} \min\{\lfloor C_k/r_0 \rfloor, L - l_k\} & , \text{ S-}R_k \text{ link for } k = 1, \dots, K; \\ \min\{\lfloor C_k/r_0 \rfloor, l_k\} & , \text{ }R_k\text{-D link for } k = 1, \dots, K. \end{cases} \quad (5.8)$$

5.2.2.3 Single Packet with Inter-Relay Cooperation (SPWI)

If inter-relay cooperation is allowed, packets have $K - 1$ possible additional links to reach D where R_k now have the option of forwarding the packet either to subsequent relay R_{k+1} (if any) or to the destination D. Consequently, a packet can pass by $K - 1 + 2 = K + 1$ links in a worst case scenario and the channel capacity again is different than the SPNI and MPNI cases but is similar to (5.1).

This implies that more links can be activated at the expense of a lower quality of the channel since the fraction of the total power allocated to each link will be smaller. The effect of the inter-relay cooperation will be discussed in Section 5.6. The maximum number of packets that can be transmitted follows (5.7) and (5.9).

$$n_{max} = \left\{ \min\{\delta_{C_k \geq r_0}, l_k, L - l_{k+1}\} \quad , \text{R}_k\text{-R}_{k+1} \text{ link for } k = 1, \dots, K - 1. \quad (5.9) \right.$$

5.2.2.4 Multiple Packets with Inter-Relay Cooperation (MPWI)

The difference between SPWI and MPWI is the ability to transmit multiple packets. At every time slot, the maximum number of packets that can be transmitted follows (5.8) and (5.10).

$$n_{max} = \left\{ \min\{\lfloor C_k/r_0 \rfloor, l_k, L - l_{k+1}\} \quad , \text{R}_k\text{-R}_{k+1} \text{ link for } k = 1, \dots, K - 1. \quad (5.10) \right.$$

5.3 Problem Formulation

The objectives of the RL agent are two-fold: (1) maximize the throughput of the system and (2) minimize the delay of packets arriving at D.

At every time slot t , we denote by the vector $a(t)$ the number of packets transmitted via every link. Thus, $a(t)$ has $K + 1$ elements for serial setups, $2K$ elements for parallel relaying setups without inter-relay cooperation such that elements $1 \leq k \leq K$ correspond to S-R $_k$ links and $K < k \leq 2K$ correspond to R $_k$ -D links. With inter-relay cooperation, $K - 1$ elements are added so that indices $2K < k \leq 3K - 1$ correspond to R $_k$ -R $_{k+1}$ links. To increase the throughput of the system, the number of packets passing through the last hop (R $_K$ -D) must be increased. This in turn will reduce the delay of packets since to reach the last hop, relays should forward packets faster. We denote by $\Psi(t)$ the set of indices of non-zero elements in $a(t)$.

The optimization problem over T time slots, can be formulated as:

$$\begin{cases} \max \frac{1}{T} \sum_{t=1}^T a(t)[K + 1] * \delta_{(a(t)[K+1] \leq n_{max}[K+1])} & \text{for serial setups} \\ \max \frac{1}{T} \sum_{t=1}^T \sum_{k=K+1}^{2K} a(t)[k] * \delta_{(a(t)[k] \leq n_{max}[k])} & \text{for parallel setups.} \end{cases} \quad (5.11)$$

where $a(t)[k]$ is the k -th element of $a(t)$. This implies that the objective of the RL agent is to increase the likelihood of selecting the last hop link, as well as to maximize the number of packets transmitted over this link under the constraint

of not exceeding n_{max} .

This optimization problem is subject to different constraints:

- For SLSP, SPNI and SPWI: (5.12) should hold.

$$\sum_k a(t)[k] \leq 1 \quad (5.12)$$

Hence, since $a(t)[k]$ is a positive integer then, at most one entry of the $a(t)$ vector will be equal to 1 and all other entries will be equal to 0. This is equivalent to allow at most one link to transmit one packet.

- For SLMP, MPNI and MPWI: (5.13) should hold.

$$|\Psi(t)| \leq 1 \quad (5.13)$$

(5.13) does not restrict the entries of $a(t)$ to be less than or equal to 1, but restricts the number of links that can be activated (i.e. have positive entries in $a(t)$). At most one link is allowed to be activated.

- For MLMP: the loosened interference constraint expressed in (5.14) should hold.

$$k' \geq k + 3, \quad \forall k, k' (k < k') \in \Psi(t) \quad (5.14)$$

- For MLSP: both (5.14) and (5.15) must hold.

$$a(t)[k] = 1, \quad \forall k \in \Psi(t) \quad (5.15)$$

In addition to the interference constraint, all non-zero entries in $\Psi(t)$ (i.e. all links selected for activation) should transmit only one packet.

5.4 Elements of the Reinforcement Learning Model

5.4.1 Environment

The environment is the entity responsible for saving the current state and processing the action chosen by the agent. It necessarily implements the functions:

- *reset()*: to re-initiate the system at the end of every round.
- *step()*: to process the action.
- *observe()*: to return the resulting state.
- *evaluate()*: to return the reward corresponding to the processed action.

- *check_unfeasible()*: to inform the agent of the feasibility of a potential action.

5.4.2 States

Both the CSI and buffer state information (BSI) are included in the state. The state vector S is divided into two parts. In order to unify the notation for serial and parallel setups, assume that K is the number of relays and X is the number of links in the system, thus, S has $K + X$ entries: the first K elements correspond to l_k the number of packets stored in the buffer of every relay. The next X elements represent n_{max} for every link in the system. n_{max} is computed based on (5.2), (5.4), (5.7)-(5.10).

5.4.3 Actions

The number of actions is specific for every setup. An action $a(t)$ is defined by the number of packets to be transmitted via every link. For serial setups, $a(t)$ has $K + 1$ elements corresponding to $K + 1$ links whereas for parallel setups, the vector $a(t)$ has either $2K$ entries if no inter-relay cooperation is allowed, or $3K - 1$ entries otherwise. $a(t)$ will follow the preset constraints for every setup (5.12)-(5.15).

5.4.4 Reward Function

For serial systems, setting a reward based solely on the goal of transmitting a packet to D will make the convergence, if any, very slow. Another joint reward function is suggested based on the number of packets transmitted in a time slot along with their relative positions from D.

Let $A = (i_1, i_2, \dots, i_{K+1})$ be the action chosen at a state s_t and i_k is the number of packets transmitted at link k . The reward function is defined as in (5.16):

$$r_t = \begin{cases} -10 & , \text{ in case of outage} \\ \alpha * \sum_k i_k + \beta * \sum_k (\frac{k}{K+1} * \delta_{i_k > 0}) & , \text{ otherwise} \end{cases} \quad (5.16)$$

where α and β are two tuning parameters. This implies that the higher the number of packets and the closer they are to D, the higher the reward is.

The same concept will be applied to parallel relaying setups. A higher reward should be given to the R-D links whereas R-R links do not add any advantage in terms of the packets' position with respect to D. For parallel setups, only one link k can be selected with i_k standing for the number of packets chosen to be

transmitted. Thus, the reward for parallel setups is as in (5.17).

$$r_t = \begin{cases} -10 & , \text{ in case of outage} \\ 0 & , \text{ R-R link is selected} \\ \alpha * i_k + \beta * \frac{m_k}{2} & , \text{ otherwise} \end{cases} \quad (5.17)$$

where $m_k = 1$ for S-R links ($1 \leq k \leq K$) and $m_k = 2$ for R-D links ($K < k \leq 2K$).

5.5 Deep Reinforcement Learning

The number of states and actions associated with the considered complex setups is relatively high and Q-learning tables with such a huge number of entries cannot converge. Hence, deep Q-learning is the convenient choice in this case.

5.5.1 Neural Network (NN)

The neural network has the state $S(t)$ as input and outputs the Q-values of all possible actions. The implemented NN has three layers:

1. Dense layer with 64-neurons and a leaky-Relu activation function ($\alpha = 0.3$).
2. Dense layer with 64-neurons and a leaky-Relu activation function ($\alpha = 0.3$).
3. Dense layer with $(K + X)$ -neurons and a Relu activation function.

5.5.2 Handling Unfeasible Actions

A-priori information will be used in this work to enforce the agent to choose feasible actions at all times. This is based on letting the agent use the previous knowledge about feasible actions so that when training and testing, the agent selects only feasible actions. This will allow the agent to train the NN on a smaller set of actions and hence will have a faster convergence.

Unlike the work in [53] where information about unfeasible actions as well as inconvenient actions assumed by the authors was provided to the agent, in this work, only unfeasible actions are eliminated and the agent is free to learn convenient actions to improve the performance.

Let ϕ_{s_t} be the set of feasible actions for a given state s_t . The selection of an action with a-priori information is done based on (5.18).

$$a_t = \begin{cases} \text{Random selection from } \phi_{s_t} & \text{with probability } \epsilon \\ \max_{a \in \phi_{s_t}} (Q_{\text{predicted}}(s_t, a)) & \text{with probability } 1 - \epsilon \end{cases} \quad (5.18)$$

where ϵ is the exploration rate. The agent starts with a high value for ϵ to better explore the action space, and then ϵ is decreased to allow the agent to move to the exploitation phase and refine the selection to the best action.

5.5.3 Agent

The deep RL agent consists of a prediction network and a target network, both having the same structure. The function of this agent can be divided into three steps as explained below.

1. Generate N_c experiences: At every iteration, the agent forwards the environment in N_c actions according to (5.18) and stores all (s_t, a_t, s_{t+1}, r_t) experiences in the set ψ .
2. Update the prediction network: Recall that the output of the prediction network is the Q-values of all actions. This network is updated according to the difference between the predicted Q-values and the target Q-values that form a loss function. The prediction network is updated using the Adam algorithm with a learning rate of 0.01.

For every experience in ψ , the predicted Q-value is calculated from the prediction network as $Q(s_t, a_t)$. On the other hand, the target Q-value is calculated from the target network as $r_t + \gamma * (\max_a Q_{tar}(s_{t+1}, a))$ which is based on the current reward r_t and the maximum future reward. This step is illustrated in Fig. 20.

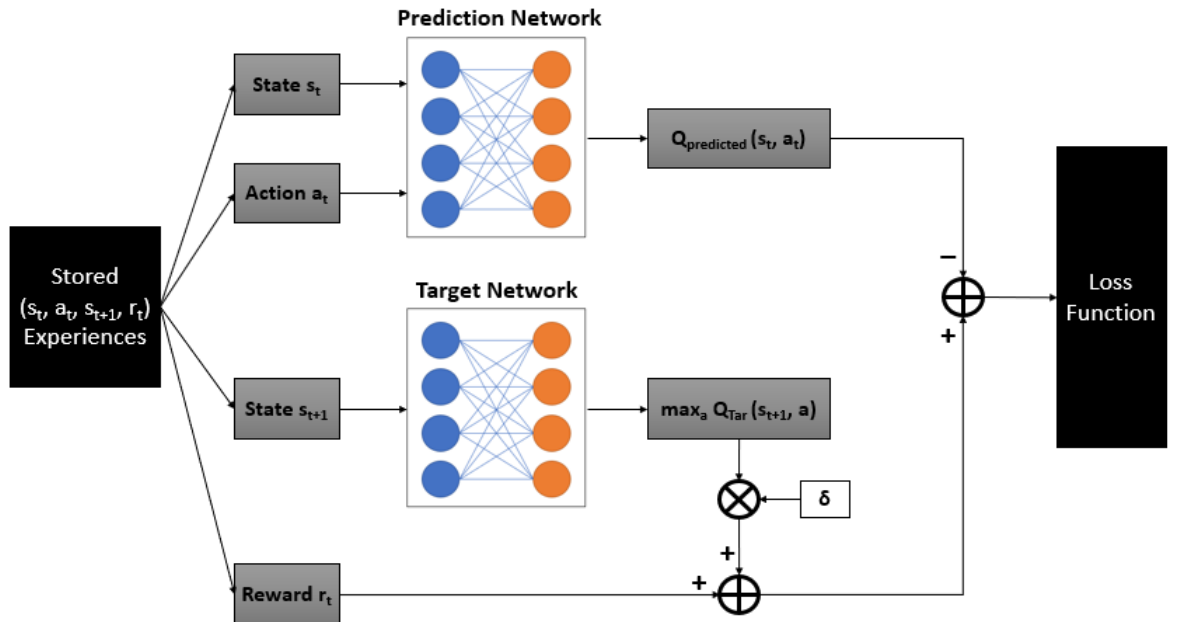


Figure 20: Step 2 of Deep-RL Agent

3. Update the target network: simply by copying the weights of the prediction network to the target network. This step is done every round (every N_i iterations) until the system converges (the weights of the prediction and target networks become approximately the same).

5.5.4 Algorithm

The learning process is summarized in Algorithm 2. A brief explanation of the terms is shown in Table 3.

Table 3: Summary of RL Symbols

N_r	Number of rounds for RL
N_i	Number of iterations per round
N_c	Number of collected experiences per iteration
BS	Training batch size
γ	Discount factor
ϵ	Exploration rate
f	Decay factor
ψ	Set of saved experiences per iteration

5.6 Simulation Results and Discussion

The main performance metrics considered are the throughput and the average packet delay (APD). The throughput of a system is defined as the average number of packets arriving at D and is measured in packets per time slot. The APD is measured by averaging the delay of all packets that arrived to D and its unit is normalized per time slot.

In what follows, we fix $N_r = 25$, $N_i = 50$, $N_c = 200$, $BS = 32$, $\gamma = 0.8$, $f = 0.999$ and $\epsilon_{min} = 0.1$. A network of 4 relays placed in series is considered in Fig. 21-Fig. 24 with $\Omega = [12, 12, 12, 12, 12]$. All deployed buffers have a size $L = 10$.

Fig. 21 shows the throughput of the system for all 4 setups at an SNR of 30 dB. It can be seen that all setups have a fast convergence. The setup MLMP took a longer time to converge (approximately after 400 iterations) because the number of actions considered is the largest (multiple links and multiple packets are allowed to be transmitted) and hence the NN is relatively bigger than other setups. Clearly, SLSP has the lower throughput since the number of packets that can be transmitted in one time slot is limited whereas setups allowing for multiple packets per link have the highest throughputs.

Fig. 22, Fig. 23 and Fig. 24 show the variations of the throughput, the average packet delay and the percentage of change in behavior with respect to the SNR,

Initialization: $\psi = \emptyset$, build prediction and target networks

for $n = 1 : Nr$ **do**

Reset the environment's variables

for $i = 1 : N_i$ **do**

Update ϵ to move from exploration to exploitation phase:

$$\epsilon = \max(f^{nN_i+i}, \epsilon_{min}) \quad (5.19)$$

for $j = 1 : N_c$ **do**

Get a_t based on (5.18)

Perform a_t and get s_{t+1} and r_t

Store the experience (s_t, a_t, s_{t+1}, r_t) in ψ

end

Randomly choose BS experiences from ψ

for $j = 1 : BS$ (for every experience) **do**

From target network get maximum future reward:

$$T(j) = \max_a Q_{tar}(s_{t+1}, a)$$

Get target weight: $Tar(j) = r_t + \gamma * T(j)$

From prediction network get: $Pr(j) = Q_{predicted}(s_t, a_t)$

end

Get the Loss Function L_i (based on MSE):

$$L_i = \sum_{j=1}^{BS} (Tar(j) - Pr(j))^2 \quad (5.20)$$

Update the Prediction Network based on (5.20)

Clear $\psi = \emptyset$

end

Update Target Network: Copy the weights of the prediction network into the target network

end

Algorithm 2: Deep RL with a-priori information for all setups of serial and parallel relay selection

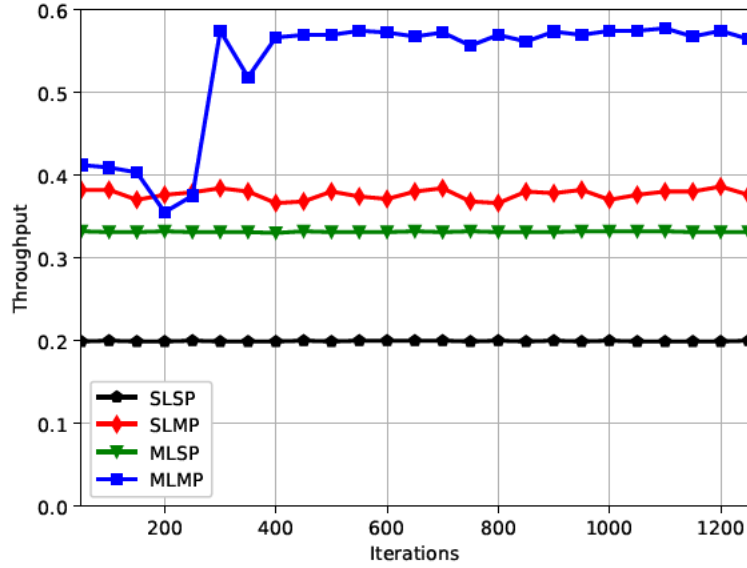


Figure 21: Convergence of the throughput for $K = 4$ and $L = 10$ for all 4 serial setups at $SNR = 30$ dB.

respectively. The change in behavior is the ratio of actions taken that allow for multiple packets and/or multiple links for a given setup (i.e. improvement on the standard SLSP setup). Different observations can be noted:

- The throughputs of SLSP and MLSP converge to a limit at high SNR whereas when multiple packets can be transmitted per link, the throughput keeps increasing. Ideally, the number of packets per link is limited to L but this happens at very high SNRs only when the capacity of links can accommodate for the transmission of L packets and this seem to take place at excessively large and unrealistic values of the SNR.
- Setups allowing for multiple links transmission have a higher throughput at all values of SNR whereas the transmission of multiple packets does not yield any advantage until the channel quality gets better and this takes place around an SNR of 20 dB. This is reflected by the increase of throughput of MLMP with respect to MLSP and SLMP with respect to SLSP starting at $SNR = 15$ dB and becomes significant at $SNR = 20$ dB.
- In terms of delay, the sudden increase in throughput at 20 dB is reflected by a slight increase in the delay of MLMP and SLMP and this is because more packets are being transmitted and buffers are more congested. However, once the channel gets better when the SNR increases, packets can be transmitted faster and the delay will decrease again.

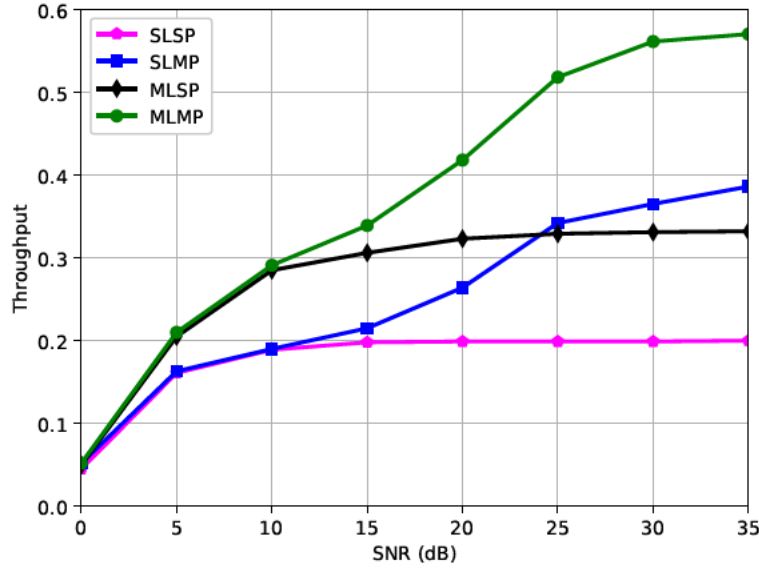


Figure 22: Throughput for $K = 4$ and $L = 10$ for all 4 serial setups.

- The average delays of SLSP and MLSP are close for mid-to-low SNR, however, asymptotically, each setup tends to a different limit. MLSP has a lower APD since for example, R_1 does not have to wait for R_4 to transmit (having a priority following from the reward function), but can transmit simultaneously and consequently, the number of accumulated packets at B_1 will be reduced and hence, the overall APD will be lower.
- SLMP has an asymptotic delay similar to SLSP since even if multiple packets are transmitted per link but relays cannot transmit simultaneously and should wait for one another. Similarly, MLMP has similar asymptotic delay to MLSP.
- Finally, Fig. 24 shows that MLMP and MLSP are benefiting from simultaneous transmission over different links starting low SNR, however, SLMP is not able to allow multiple packets to be transmitted until SNR is 15 or 20 dB. At high SNR, agents with MLMP and SLMP are not selecting actions with one packet anymore but are near 100% benefiting from simultaneous link activation and high capacity of links.

Parallel relays are considered in Fig. 25- Fig. 28 with $K = 6$, $L = 8$, $\Omega_{SR} = [4, 3.5, 3, 2.5, 2, 1.5]$, $\Omega_{RD} = [1, 1.5, 2, 2.5, 3, 3.5]$ and $\Omega_{RR} = [16, 16, 16, 16, 16]$.

Fig. 25 shows the throughput of the parallel-relaying system for all 4 setups at an SNR of 15 dB. All setups converged fast with MPNI having the highest throughput. A throughput of 0.5 was achieved by the simplest model of SPNI.

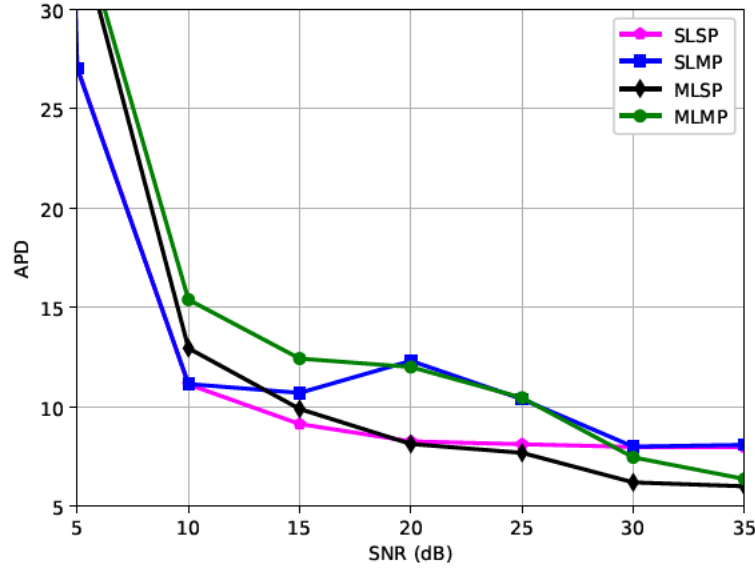


Figure 23: APD for $K = 4$ and $L = 10$ for all 4 serial setups.

However, both setups with inter-relay cooperation achieve the same throughput level that is much lower than 0.5. MPWI does not present any advantage over SPWI.

Fig. 26, Fig. 27 and Fig. 28 show the variations of the throughput, the average packet delay and the percentage of change in behavior for parallel setups with respect to the SNR, respectively. Different observations can be noted:

- MPNI setup can achieve the best throughput that is three times higher than the standard 0.5 value of SPNI at high SNR values. However, the transmission of multiple packets results in an increase in the APD that is significant for the average to low SNR range. This difference is eliminated asymptotically as MPNI converges to an APD of 2 which is the same as SPNI.
- Setups allowing for inter-relay cooperation have a reduced channel quality and consequently for low SNR, no packet can be transmitted. As SNR increases, the channel gets better and the throughput could converge to 0.5 for $\text{SNR} > 20$ dB as highlighted in Fig. 27.
- Fig. 28 shows that inter-relay cooperation is used for average SNR only and then as SNR increases, it is no longer needed and the behavior of SPWI and MPWI converges to the same value as SPNI.
- MPWI is the same as SPWI since the reduction in the channel quality results in $\lfloor C_k/r_0 \rfloor$ never exceeding 1. Thus only one packet can be trans-

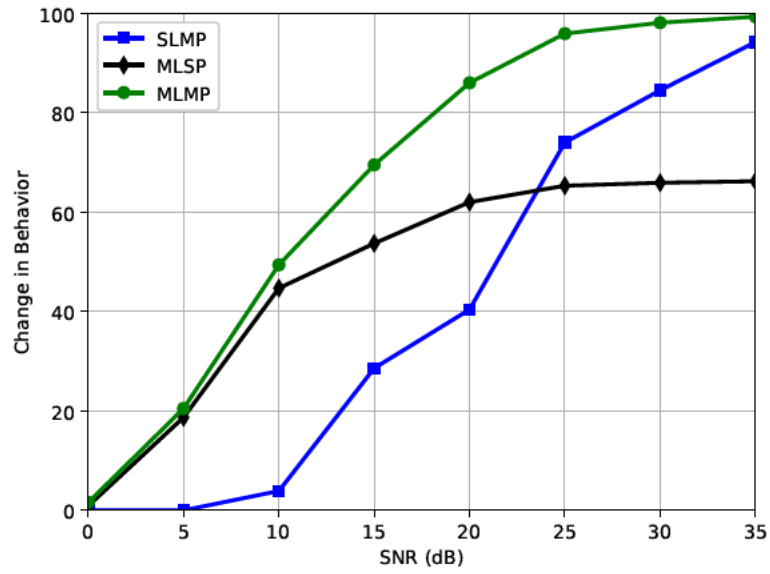


Figure 24: Change in behavior scores for $K = 4$ and $L = 10$ for all 4 serial setups.

mitted. Higher values of SNR might show some advantage for MPWI over SPWI but such SNR range seems to be unrealistic.

- In terms of delay, the RL agent proved to be efficient since all setups converged to the minimum APD that is 2.

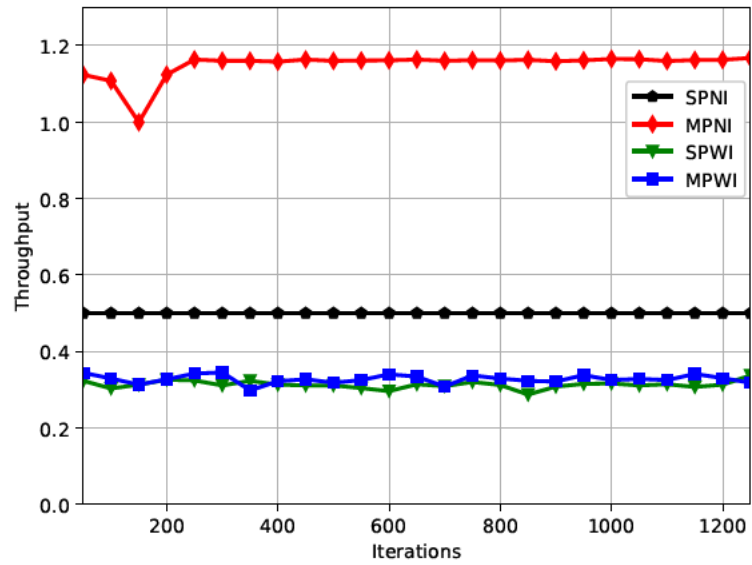


Figure 25: Convergence of the throughput for $K = 6$ and $L = 8$ for all 4 parallel setups at $SNR = 15$ dB.

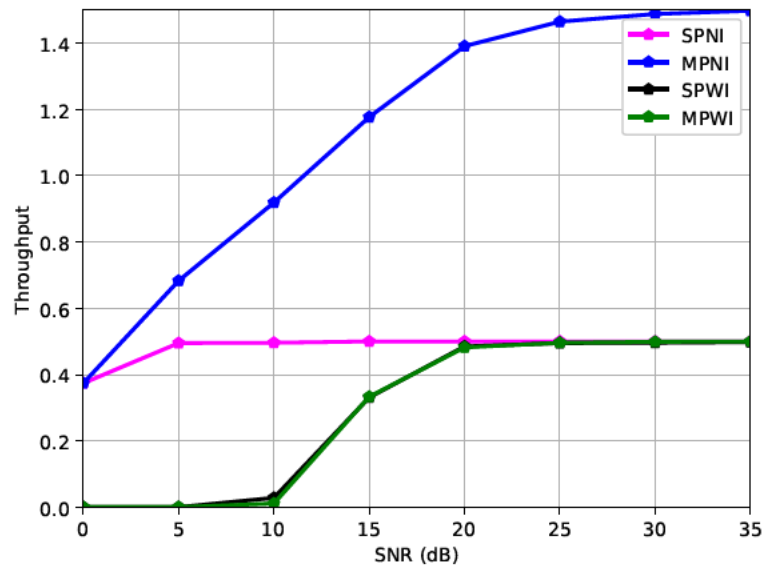


Figure 26: Throughput for $K = 6$ and $L = 8$ for all 4 parallel setups.

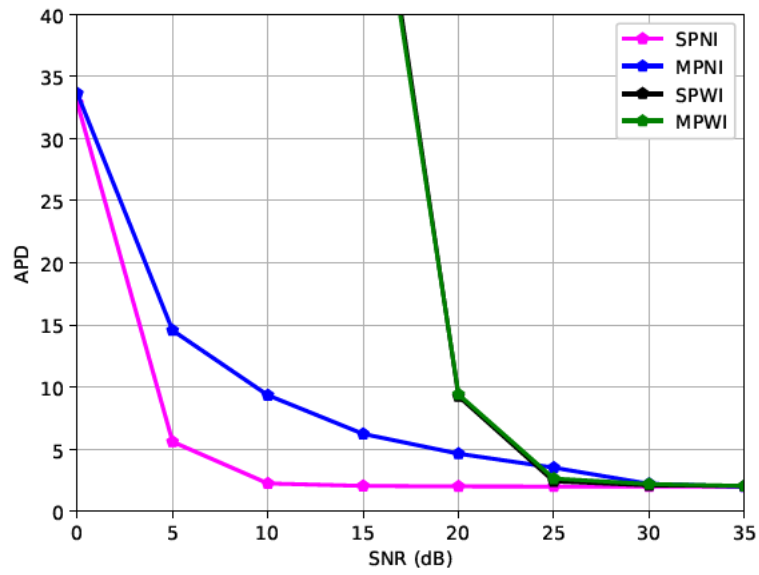


Figure 27: APD for $K = 6$ and $L = 8$ for all 4 parallel setups.

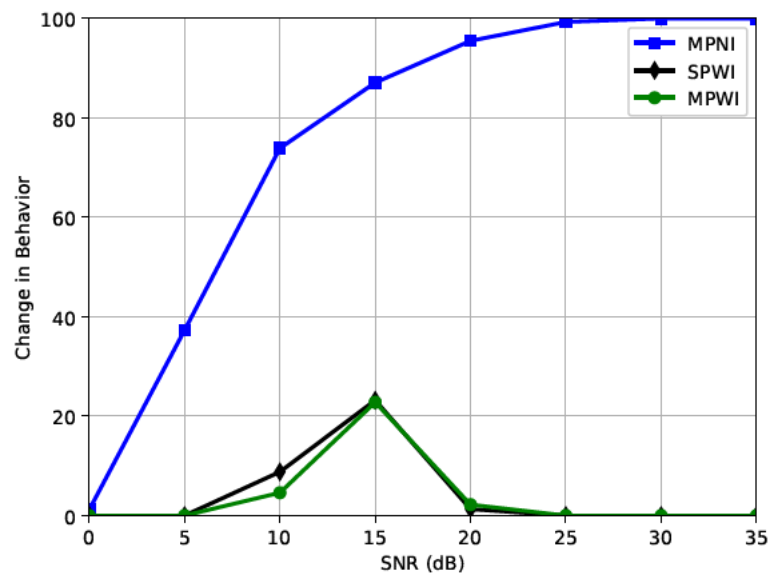


Figure 28: Change in behavior scores for $K = 6$ and $L = 8$ for all 4 parallel setups.

Chapter 6

Conclusions and Future Work

In this thesis, we investigated the problem of relay selection for BA cooperative relaying. Deterministic and learning-based techniques were the main tools for analyzing the proposed schemes and both serial and parallel positioning of HD relays were considered.

First, we proposed three novel deterministic schemes for HD multi-hop systems with an arbitrary number of relays. Based on a MC analysis and closed subset formulation, we identified the most probable states to derive closed-form expressions of the OP and APD. These expressions were essential for optimizing the control parameters of the tunable relaying schemes allowing to achieve different levels of tradeoff between outage and delay.

However, moving to more complex setups, deterministic analysis becomes cumbersome and modeling the transitions between states in a MC model becomes challenging. Hence, a RL algorithm was proposed to improve the performance of different serial and parallel relaying setups. A new design of the reward function was suggested and led to a fast convergence of the system. Results show that allowing for multiple packets transmission per link was advantageous for all setups at average to high SNR values, that is when the channel capacity is high. Moreover, loosening the interference constraint for serial setups resulted in a higher throughput even at low values of the SNR. For parallel relaying on the other hand, inter-relay cooperation, by adding more links to the system, decreases the fraction of the total power allocated to each link and hence achieves a low throughput and a very high delay.

Finally, we note that the RL model tackled in this thesis consisted of discrete states and actions to solve the relay selection problem. A future work might address the power allocation problem that is based on continuous state and action spaces and need more advanced algorithms of RL. Power efficient systems are required especially for applications of wireless sensor networks.

References

- [1] M. Agiwal, A. Roy, and N. Saxena, “Next generation 5g wireless networks: A comprehensive survey,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.
- [2] J. S. Banerjee, A. Chakraborty, and A. Chattopadhyay, “A decision model for selecting best reliable relay queue for cooperative relaying in cooperative cognitive radio networks: the extent analysis based fuzzy ahp solution,” *Wireless Networks*, vol. 27, no. 4, pp. 2909–2930, 2021.
- [3] L. Yang, J. Chen, M. O. Hasna, and H.-C. Yang, “Outage performance of UAV-assisted relaying systems with RF energy harvesting,” vol. 22, no. 12, pp. 2471–2474, Dec. 2018.
- [4] B. Li, S. Zhao, R. Miao, and R. Zhang, “A survey on unmanned aerial vehicle relaying networks,” *IET Communications*, 2021.
- [5] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, “Wireless backhauling of 5g small cells: Challenges and solution approaches,” *IEEE Wireless Communications*, vol. 22, no. 5, pp. 22–31, 2015.
- [6] G. Shen, J. Liu, D. Wang, J. Wang, and S. Jin, “Multi-hop relay for next-generation wireless access networks,” *Bell Labs Technical Journal*, vol. 13, no. 4, pp. 175–193, 2009.
- [7] 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.
- [8] R. A. Alsemmeiri, S. T. Bakhsh, and H. Alsemmeiri, “Free space optics vs radio frequency wireless communication,” *Int. J. Inf. Technol. Comput. Sci*, vol. 8, no. 9, pp. 1–8, 2016.
- [9] K. Prabu, R. Rajendran, and D. S. Kumar, “Spectrum analysis of radio over free space optical communications systems through different channel models,” *Optik*, vol. 126, no. 11-12, pp. 1142–1145, 2015.

- [10] S. S. Ikki and M. H. Ahmed, "Performance analysis of incremental-relaying cooperative-diversity networks over rayleigh fading channels," *IET communications*, vol. 5, no. 3, pp. 337–349, 2011.
- [11] G. Farhadi and N. C. Beaulieu, "On the ergodic capacity of multi-hop wireless relaying systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2286–2291, 2009.
- [12] S. Loyka and G. Levin, "On outage probability and diversity-multiplexing tradeoff in mimo relay channels," *IEEE transactions on communications*, vol. 59, no. 6, pp. 1731–1741, 2011.
- [13] G. Levin and S. Loyka, "Amplify-and-forward versus decode-and-forward relaying: Which is better?" in *22th International Zurich seminar on communications (IZS)*. Eidgenössische Technische Hochschule Zürich, 2012.
- [14] Y. Gu, H. Chen, Y. Li, and B. Vucetic, "Ultra-reliable short-packet communications: Half-duplex or full-duplex relaying?" *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 348–351, 2017.
- [15] E. Yang, X. Wen, R. Chen, and J. Wei, "Virtual full-duplex relaying protocol with network interference cancellation," in *2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*. IEEE, 2020, pp. 1–2.
- [16] 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, 2nd Quarter 2015.
- [17] R. Tannious and A. Nosratinia, "Spectrally-efficient relay selection with limited feedback," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1419–1428, 2008.
- [18] S. S. Ikki, M. Uysal, and M. H. Ahmed, "Performance analysis of incremental-best-relay amplify-and-forward technique," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*. IEEE, 2009, pp. 1–6.
- [19] S. A. Fares, F. Adachi, and E. Kudoh, "A novel cooperative relaying network scheme with inter-relay data exchange," *IEICE transactions on communications*, vol. 92, no. 5, pp. 1786–1795, 2009.

- [20] H. Wang, G. Xu, and M. Li, “A selection cooperation scheme with feedback and effective utilization of inter-relay links,” *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 428–431, 2016.
- [21] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, “Multihop diversity in wireless relaying channels,” vol. 52, no. 10, pp. 1820–1830, Oct. 2004.
- [22] M. O. Hasna and M.-S. Alouini, “End-to-end performance of transmission systems with relays over Rayleigh-fading channels,” vol. 2, no. 6, pp. 1126–1131, Nov. 2003.
- [23] Y. G. Kim and N. C. Beaulieu, “Relay advantage criterion for multihop decode-and-forward relaying systems,” *IEEE transactions on wireless communications*, vol. 13, no. 4, pp. 1988–1999, 2014.
- [24] R. Kumar and A. Hossain, “Survey on half-and full-duplex relay based cooperative communications and its potential challenges and open issues using markov chains,” *IET Communications*, vol. 13, no. 11, pp. 1537–1550, May 2019.
- [25] M. El-Rajab, C. Abou-Rjeily, and R. Kfourri, “Buffer-aided relaying: a survey on relay selection policies,” *IET Communications*, vol. 14, no. 21, pp. 3715–3734, Jan. 2021.
- [26] N. Zlatanov, R. Schober, and P. Popovski, “Buffer-aided relaying with adaptive link selection,” vol. 31, no. 8, pp. 1530–1542, Aug. 2013.
- [27] N. Zlatanov and R. Schober, “Buffer-aided relaying with adaptive link selection—fixed and mixed rate transmission,” vol. 59, no. 5, pp. 2816–2840, May 2013.
- [28] C. Abou-Rjeily, “Toward a better comprehension of decode-and-forward buffer-aided relaying: Case study of a single relay,” *IEEE Communications Letters*, vol. 24, no. 5, pp. 1005–1009, May 2020.
- [29] I. Krikidis, T. Charalambous, and J. S. Thompson, “Buffer-aided relay selection for cooperative diversity systems without delay constraints,” vol. 11, no. 5, pp. 1957–1967, May 2012.
- [30] Z. Tian, Y. Gong, G. Chen, and J. A. Chambers, “Buffer-aided relay selection with reduced packet delay in cooperative networks,” vol. 66, no. 3, pp. 2567–2575, Mar. 2016.

- [31] S.-L. Lin and K.-H. Liu, "Relay selection for cooperative relaying networks with small buffers," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6562–6572, 2015.
- [32] M. Oiwa, C. Tosa, and S. Sugiura, "Theoretical analysis of hybrid buffer-aided cooperative protocol based on max–max and max–link relay selections," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 9236–9246, 2016.
- [33] A. A. M. Siddig and M. F. M. Salleh, "Balancing buffer-aided relay selection for cooperative relaying systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8276–8290, 2017.
- [34] B. Manoj, R. K. Mallik, and M. R. Bhatnagar, "Performance analysis of buffer-aided priority-based max-link relay selection in df cooperative networks," *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 2826–2839, 2018.
- [35] S. Luo and K. C. Teh, "Buffer state based relay selection for buffer-aided cooperative relaying systems," vol. 14, no. 10, pp. 5430–5439, Oct. 2015.
- [36] P. Xu, Z. Ding, I. Krikidis, and X. Dai, "Achieving optimal diversity gain in buffer-aided relay networks with small buffer size," vol. 65, no. 10, pp. 8788–8794, Oct. 2015.
- [37] W. Raza, N. Javaid, H. Nasir, N. Alrajeh, and N. Guizani, "Buffer-aided relay selection with equal-weight links in cooperative wireless networks," *IEEE Communications Letters*, vol. 22, no. 1, pp. 133–136, 2017.
- [38] N. Nomikos, D. Poulimeneas, T. Charalambous, I. Krikidis, D. Vouyioukas, and M. Johansson, "Delay-and diversity-aware buffer-aided relay selection policies in cooperative networks," *IEEE Access*, vol. 6, pp. 73 531–73 547, 2018.
- [39] S. El-Zahr and C. Abou-Rjeily, "Threshold based relay selection for buffer-aided cooperative relaying systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 6210–6223, 2021.
- [40] P. Xu, Z. Yang, Z. Ding, I. Krikidis, and Q. Chen, "A novel probabilistic buffer-aided relay selection scheme in cooperative networks," vol. 69, no. 4, pp. 4548–4552, Apr. 2020.

- [41] P. Xu, G. Chen, Z. Yang, and H. Lei, “Buffer-state-based probabilistic relay selection for cooperative networks with delay constraints,” *IEEE Wireless Commun. Letters*, 2020, Early Access.
- [42] 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.
- [43] C. Dong, L.-L. Yang, and L. Hanzo, “Performance analysis of multihop-diversity-aided multihop links,” vol. 61, no. 6, pp. 2504–2516, July 2012.
- [44] B. Manoj, R. K. Mallik, and M. R. Bhatnagar, “Buffer-aided multi-hop DF cooperative networks: A state-clustering based approach,” vol. 64, no. 12, pp. 4997–5010, 2016.
- [45] V. Jamali, N. Zlatanov, H. Shoukry, and R. Schober, “Achievable rate of the half-duplex multi-hop buffer-aided relay channel with block fading,” vol. 14, no. 11, pp. 6240–6256, Nov. 2015.
- [46] S. M. Kim and M. Bengtsson, “Virtual full-duplex buffer-aided relaying in the presence of inter-relay interference,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2966–2980, 2016.
- [47] N. Nomikos, T. Charalambous, I. Krikidis, D. N. Skoutas, D. Vouyioukas, and M. Johansson, “A buffer-aided successive opportunistic relay selection scheme with power adaptation and inter-relay interference cancellation for cooperative diversity systems,” *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1623–1634, 2015.
- [48] 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.
- [49] M. M. Razlighi and N. Zlatanov, “Buffer-aided relaying for the two-hop full-duplex relay channel with self-interference,” vol. 17, no. 1, pp. 477–491, Jan. 2018.
- [50] G. Yang, M. Haenggi, and M. Xiao, “Traffic allocation for low-latency multi-hop networks with buffers,” vol. 66, no. 9, pp. 3999–4013, May 2018.
- [51] C. Abou-Rjeily and W. Fawaz, “Buffer-aided serial relaying for FSO communications: asymptotic analysis and impact of relay placement,” vol. 17, no. 12, pp. 8299–8313, Dec. 2018.

- [52] C. Huang, G. Chen, and Y. Gong, “Delay constrained buffer-aided relay selection in the internet of things with decision-assisted reinforcement learning,” *IEEE Internet of Things Journal*, 2021.
- [53] C. Huang, G. Chen, Y. Gong, P. Xu, Z. Han, and J. A. Chambers, “Buffer-aided relay selection for cooperative hybrid noma/oma networks with asynchronous deep reinforcement learning,” *IEEE Journal on Selected Areas in Communications*, 2021.
- [54] C. Huang, J. Zhong, Y. Gong, Z. Abdullah, and G. Chen, “Novel deep reinforcement learning-based delay-constrained buffer-aided relay selection in cognitive cooperative networks,” *Electronics Letters*, vol. 56, no. 21, pp. 1148–1150, 2020.
- [55] C. Huang, G. Chen, and K.-K. Wong, “Multi-agent reinforcement learning-based buffer-aided relay selection in irs-assisted secure cooperative networks,” *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4101–4112, 2021.
- [56] C. Huang, G. Chen, Y. Gong, and P. Xu, “Deep reinforcement learning based relay selection in delay-constrained secure buffer-aided crns,” in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.
- [57] C. Huang, G. Chen, Y. Gong, and Z. Han, “Joint buffer-aided hybrid-duplex relay selection and power allocation for secure cognitive networks with double deep q-network,” *IEEE Transactions on Cognitive Communications and Networking*, 2021.
- [58] H. Zhang, D. Zhan, C. J. Zhang, K. Wu, Y. Liu, and S. Luo, “Deep reinforcement learning-based access control for buffer-aided relaying systems with energy harvesting,” *IEEE Access*, vol. 8, pp. 145 006–145 017, 2020.
- [59] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [60] J. D. Little and S. C. Graves, “Little’s law,” in *Building intuition*. Springer, 2008, pp. 81–100.

Appendix A

Closed Subset for $1 < s < L$

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 \mathbf{p} that can be expressed as:

$$\mathbf{p} = \sum_{i_1 \geq 0} \cdots \sum_{i_{K+1} \geq 0} c_{i_1, \dots, i_{K+1}} \prod_{n=1}^{K+1} p_n^{i_n}, \quad (\text{A.1})$$

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

$$O(\mathbf{p}) = \min_{i_1 \geq 0 \cdots i_{K+1} \geq 0} \left\{ \sum_{n=1}^{K+1} i_n \mid c_{i_1, \dots, i_{K+1}} = 0 \ \forall \sum_{n=1}^{K+1} i_n < O(\mathbf{p}) \right\}. \quad (\text{A.2})$$

For example, $O(1/3 - p_1 + 2p_2 + \dots) = 0$ and $O(p_1^2 + p_1 p_2 p_3 + \dots) = 2$. From (A.2), it follows that $O(\mathbf{p}\mathbf{p}') = O(\mathbf{p}) + O(\mathbf{p}')$ and $O(\mathbf{p} + \mathbf{p}') = \min\{O(\mathbf{p}), O(\mathbf{p}')\}$.

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 (3.20) is closed by proving that the corresponding transition probabilities satisfy (3.18). 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 $k \in \{1, \dots, K + 1\}$):

$$\forall \mathbf{l} \in \mathcal{S}_{c,1} : \quad t_{\mathbf{l}, \mathbf{l}'} = \begin{cases} 1 - p_k, & \mathbf{l}' \in \mathcal{S}_{c,1}; \\ p_k, & \mathbf{l}' \in \mathcal{S}_{c,2}. \end{cases} \quad (\text{A.3})$$

$$\forall \mathbf{l} \in \mathcal{S}_{c,2} : \quad t_{\mathbf{l}, \mathbf{l}'} = \begin{cases} 1 - p_k, & \mathbf{l}' \in \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}; \\ p_k, & \mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases}, \quad (\text{A.4})$$

where all transition probabilities whose asymptotic orders exceed one were ig-

nored.

Proposition 4. For the subsets $\mathcal{S}_{c,1}$ and $\mathcal{S}_{c,2}$ satisfying (A.3)-(A.4), the asymptotic orders of the steady-state probabilities of the corresponding states satisfy the relation in (A.5):

$$\begin{cases} O(\pi_{\mathbf{l}}) = 0, & \mathbf{l} \in \mathcal{S}_{c,1}; \\ O(\pi_{\mathbf{l}}) = 1, & \mathbf{l} \in \mathcal{S}_{c,2}; \\ O(\pi_{\mathbf{l}}) \geq 2, & \mathbf{l} \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases} \quad (\text{A.5})$$

Proof. We will prove that (A.5) satisfies the asymptotic orders of all balance equations. For any state \mathbf{l} , the balance equation at steady-state is generalized as follows:

$$\pi_{\mathbf{l}} = \sum_{\mathbf{l}' \in \mathcal{S}_{c,1}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} + \sum_{\mathbf{l}' \in \mathcal{S}_{c,2}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} + \sum_{\mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}} t_{\mathbf{l}',\mathbf{l}} \pi_{\mathbf{l}'} \quad (\text{A.6})$$

implying that:

$$O(\pi_{\mathbf{l}}) = \min \left\{ \min_{\mathbf{l}' \in \mathcal{S}_{c,1}} \underbrace{\{O(t_{\mathbf{l}',\mathbf{l}}) + O(\pi_{\mathbf{l}'})\}}_{\triangleq o_1}, \right. \\ \left. \min_{\mathbf{l}' \in \mathcal{S}_{c,2}} \underbrace{\{O(t_{\mathbf{l}',\mathbf{l}}) + O(\pi_{\mathbf{l}'})\}}_{\triangleq o_2}, \min_{\mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}} \underbrace{\{O(t_{\mathbf{l}',\mathbf{l}}) + O(\pi_{\mathbf{l}'})\}}_{\geq 0} \right\} \quad (\text{A.7})$$

$$\triangleq \min \{O_1(\pi_{\mathbf{l}}), O_2(\pi_{\mathbf{l}}), O_3(\pi_{\mathbf{l}})\}, \quad (\text{A.8})$$

where the asymptotic orders from (A.5) were replaced in (A.7). (i): For $\mathbf{l} \in \mathcal{S}_{c,1}$, $o_1 = 0$ and $o_2 = 0$ following from (A.3) and (A.4), respectively. Consequently, $O_1(\pi_{\mathbf{l}}) = 0$, $O_2(\pi_{\mathbf{l}}) = 1$ and $O_3(\pi_{\mathbf{l}}) \geq 2$ implying from (A.8) that $O(\pi_{\mathbf{l}}) = 0$ thus proving the first relation in (A.5). (ii): For $\mathbf{l} \in \mathcal{S}_{c,2}$, $o_1 = 1$ and $o_2 = 0$ following from (A.3) and (A.4), respectively. Consequently, $O_1(\pi_{\mathbf{l}}) = 1$, $O_2(\pi_{\mathbf{l}}) = 1$ and $O_3(\pi_{\mathbf{l}}) \geq 2$ implying from (A.8) that $O(\pi_{\mathbf{l}}) = 1$ thus proving the second relation in (A.5). (iii): For $\mathbf{l} \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$, $o_1 \geq 2$ and $o_2 = 1$ following from (A.3) and (A.4), 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 (A.8) that $O(\pi_{\mathbf{l}}) = 2$ thus proving the third relation in (A.5). \square

Lemma 1. For the subsets $\mathcal{S}_{c,1}$ and $\mathcal{S}_{c,2}$ satisfying (A.3)-(A.4), 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 (A.5) 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. \square

Lemma 2. For the order-1 asymptotic analysis, the transition probabilities in (A.4) can be approximated by:

$$\forall \mathbf{l} \in \mathcal{S}_{c,2} : \quad t_{\mathbf{l},\mathbf{l}'} \approx \begin{cases} 1, & \mathbf{l}' \in \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}; \\ 0, & \mathbf{l}' \notin \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}. \end{cases} \quad (\text{A.9})$$

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 (A.4). In fact, the inclusion of the probability p_k in the transitions appearing in (A.4) will only yield to terms whose asymptotic orders exceed two and, hence, can be ignored. \square

Proposition 5. The following subsets of the set \mathcal{S}_c in (3.20) satisfy the conditions in (A.3)-(A.4):

$$\begin{aligned} \mathcal{S}_{c,1} &= \{\mathbf{s}_1^{(1)}\} \cup \{\mathbf{s}_n^{(2)} ; n = 1, \dots, K\} \\ \mathcal{S}_{c,2} &= \{\mathbf{s}_n^{(1)} ; n = 2, 3\} \cup \{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(4)} ; n = 1, \dots, K-1\}. \end{aligned} \quad (\text{A.10})$$

Proof. We first prove the condition in (A.3). 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 (A.10) are:

$$\begin{aligned} t_{\mathbf{s}_1^{(1)}, \mathbf{s}_1^{(2)}} &= 1 - p_1 \quad ; \quad t_{\mathbf{s}_K^{(2)}, \mathbf{s}_1^{(1)}} = 1 - p_{K+1} \\ t_{\mathbf{s}_n^{(2)}, \mathbf{s}_{n+1}^{(2)}} &= 1 - p_{n+1}, \quad n = 1, \dots, K-1, \end{aligned} \quad (\text{A.11})$$

and:

$$\begin{aligned} t_{\mathbf{s}_1^{(1)}, \mathbf{s}_3^{(1)}} &= p_1 \quad ; \quad t_{\mathbf{s}_1^{(2)}, \mathbf{s}_2^{(1)}} = p_2 \\ t_{\mathbf{s}_n^{(2)}, \mathbf{s}_{n-1}^{(3)}} &= p_{n+1}, \quad n = 2, \dots, K, \end{aligned} \quad (\text{A.12})$$

where the proof of (A.11) and (A.12) follows directly from the relaying strategy in (3.3)-(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 (A.11)-(A.12) satisfy the condition in (A.3).

Next, we prove the condition in (A.4). For elements of $\mathcal{S}_{c,2}$ in (A.10), the transitions that are confined in $\mathcal{S}_c = \mathcal{S}_{c,1} \cup \mathcal{S}_{c,2}$ occur with the following probabilities:

$$\begin{aligned}
t_{\mathbf{s}_2^{(1)}, \mathbf{s}_1^{(3)}} &= 1 - p_2 \quad ; \quad t_{\mathbf{s}_3^{(1)}, \mathbf{s}_1^{(4)}} = 1 - p_1 \\
t_{\mathbf{s}_{K-1}^{(3)}, \mathbf{s}_1^{(2)}} &= 1 - p_{K+1} \quad ; \quad t_{\mathbf{s}_{K-1}^{(4)}, \mathbf{s}_1^{(1)}} = 1 - p_K \\
t_{\mathbf{s}_n^{(3)}, \mathbf{s}_{n+1}^{(3)}} &= 1 - p_{n+2}, \quad n = 1, \dots, K - 2 \\
t_{\mathbf{s}_n^{(4)}, \mathbf{s}_{n+1}^{(4)}} &= 1 - p_{n+1}, \quad n = 1, \dots, K - 2,
\end{aligned} \tag{A.13}$$

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

$$\begin{aligned}
t_{\mathbf{s}_2^{(1)}, \mathbf{s}_2^{(1)} + \mathbf{e}_1} &= p_2 \quad ; \quad t_{\mathbf{s}_3^{(1)}, \mathbf{s}_3^{(1)} - \mathbf{e}_{K-1} + \mathbf{e}_K} = p_1 \\
t_{\mathbf{s}_{K-1}^{(3)}, \mathbf{s}_{K-1}^{(3)} - \mathbf{e}_1 + \mathbf{e}_2} &= p_{K+1} \quad ; \quad t_{\mathbf{s}_{K-1}^{(4)}, \mathbf{s}_{K-1}^{(4)} + \mathbf{e}_1} = p_K \\
t_{\mathbf{s}_n^{(3)}, \mathbf{s}_n^{(3)} - \mathbf{e}_1 + \mathbf{e}_2} &= p_{n+2}, \quad n = 1, \dots, K - 2 \\
t_{\mathbf{s}_n^{(4)}, \mathbf{s}_n^{(4)} + \mathbf{e}_1} &= p_{n+1}, \quad n = 1, \dots, K - 2,
\end{aligned} \tag{A.14}$$

where (A.13)-(A.14) follow directly from the relaying strategy in (3.3)-(3.4). As an illustration, we will provide the proof for the state $\mathbf{s}_n^{(3)}$ with $n = 1, \dots, K - 2$ 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 $n = 1, \dots, K - 2$), $\Delta_1 = \Delta_2 = \Delta_{n+2} = s$ 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 $\mathbf{s}_n^{(3)} \rightarrow \mathbf{s}_n^{(3)} - \mathbf{e}_{n+1} + \mathbf{e}_{n+2} = \mathbf{s}_{n+1}^{(3)}$. 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 $\mathbf{s}_n^{(3)} \rightarrow \mathbf{s}_n^{(3)} - \mathbf{e}_1 + \mathbf{e}_2 \notin \mathcal{S}_c$. As a conclusion, the transition probabilities in (A.13)-(A.14) satisfy the condition in (A.4). \square

Therefore, the union of the sets in (A.10) is closed asymptotically following from lemma 1 and the transition probabilities in (A.13) and (A.14) can be approximated by 1 and 0, respectively, following from lemma 2. This results in the simplified closed state diagram illustrated in Fig. 29.

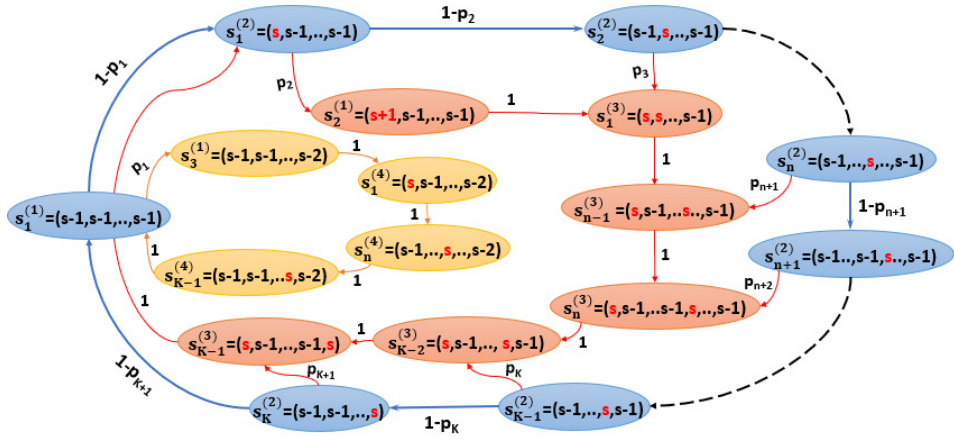


Figure 29: Closed Subset for $1 < s < L$.

From Fig. 29, the $3K + 1$ balance equations in the closed subset \mathcal{S}_c are given by:

$$\pi_{\mathbf{s}_1^{(1)}} = (1 - p_{K+1})\pi_{\mathbf{s}_K^{(2)}} + \pi_{\mathbf{s}_{K-1}^{(4)}} \quad (\text{A.15})$$

$$\pi_{\mathbf{s}_1^{(2)}} = (1 - p_1)\pi_{\mathbf{s}_1^{(1)}} + \pi_{\mathbf{s}_{K-1}^{(3)}} \quad (\text{A.16})$$

$$\pi_{\mathbf{s}_n^{(2)}} = (1 - p_n)\pi_{\mathbf{s}_{n-1}^{(2)}}, \quad n = 2, \dots, K \quad (\text{A.17})$$

$$\pi_{\mathbf{s}_2^{(1)}} = p_2\pi_{\mathbf{s}_1^{(2)}} \quad (\text{A.18})$$

$$\pi_{\mathbf{s}_1^{(3)}} = p_3\pi_{\mathbf{s}_2^{(2)}} + \pi_{\mathbf{s}_2^{(1)}} \quad (\text{A.19})$$

$$\pi_{\mathbf{s}_n^{(3)}} = p_{n+2}\pi_{\mathbf{s}_{n+1}^{(2)}} + \pi_{\mathbf{s}_{n-1}^{(3)}}, \quad n = 2, \dots, K - 1 \quad (\text{A.20})$$

$$\pi_{\mathbf{s}_3^{(1)}} = p_1\pi_{\mathbf{s}_1^{(1)}} \quad (\text{A.21})$$

$$\pi_{\mathbf{s}_1^{(4)}} = \pi_{\mathbf{s}_3^{(1)}} \quad (\text{A.22})$$

$$\pi_{\mathbf{s}_n^{(4)}} = \pi_{\mathbf{s}_{n-1}^{(4)}}, \quad n = 2, \dots, K - 1. \quad (\text{A.23})$$

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 (3.21).

Appendix B

Closed Subset for $s = 1$

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 $n = 1, \dots, K-1$ since these states do not exist for $s = 1$ (since $s - 2$ becomes negative in this case). Removing these states from (3.20) results in the closed subset provided in (3.22). The closed subset is now as presented in Fig. 30 that is obtained by removing the above mentioned K states from the state diagram in Fig. 29.

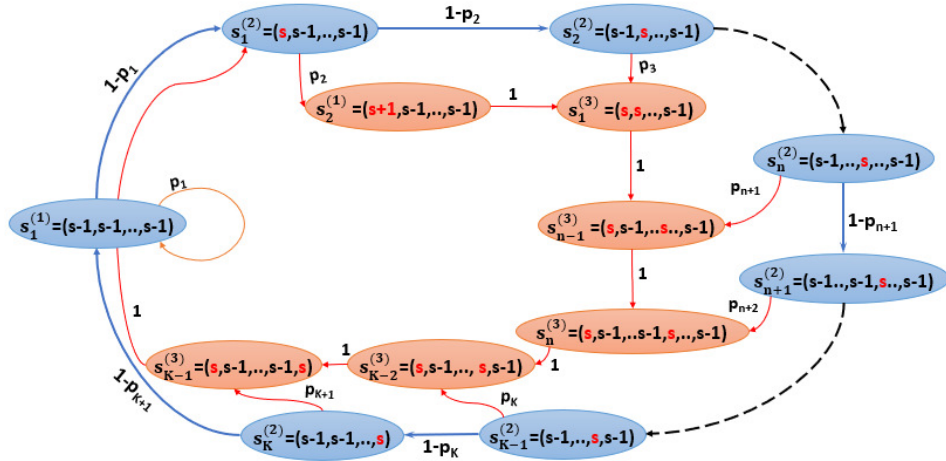


Figure 30: Closed Subset for $s = 1$.

Consequently, in the balance equations (A.15)-(A.23), equations (A.21)-(A.23) must be removed while equation (A.15) must be replaced by:

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

Solving (A.16)-(A.20) and (B.1) along with the relation $\sum_{\mathbf{l} \in \mathcal{S}_c} \pi_{\mathbf{l}} = 1$ generates the steady-state probabilities presented in (3.23).

Appendix D

Transition Probabilities of Scheme 1

Consider an integer $i \in \{1, \dots, K-1\}$. From (4.5):

$$\mathbf{s}_i = \left(\underbrace{s-1, \dots, s-1}_{i-1 \text{ times}}, s, \underbrace{s-1, \dots, s-1}_{K-i \text{ times}} \right), \quad (\text{D.1})$$

implying from (4.2) that $\Delta_k = s$ for $k = 1, i+1$ and $\Delta_k = s-1$ otherwise. As such, link $i+1$ will be activated according to the tie-breaking rule. This implies that the number of packets in B_i will decrease by one while the number of packets in B_{i+1} will increase by one reflecting the flow of a packet from R_i to R_{i+1} along link $i+1$ assuming that all links are not in outage in the asymptotic regime. Therefore, $\mathbf{s}_i(i) \rightarrow \mathbf{s}_i(i) - 1 = s-1$ and $\mathbf{s}_i(i+1) \rightarrow \mathbf{s}_i(i+1) + 1 = s$. The new state has the structure of \mathbf{s}_{i+1} following from (D.1) implying a transition of the MC from state \mathbf{s}_i to state \mathbf{s}_{i+1} .

Appendix E

Derivation of the DO Expression

Full and empty buffers contribute to decreasing the DO since they reduce the number of available links. From (3.2) and (4.8), $\mathbf{s}_i(k) = L \Rightarrow q_k(\mathbf{s}_i) = 1$ and $\mathbf{s}_i(k) = 0 \Rightarrow q_{k+1}(\mathbf{s}_i) = 1$.

From (4.6), for $i = 1, \dots, K$, one element of \mathbf{s}_i is equal to $d - 1$ while the remaining $K - 1$ elements are equal to d and all elements of \mathbf{s}_{K+1} are equal to d . (i): For $1 < d < L$, none of the elements of $\{\mathbf{s}_i\}_{i=1}^K$ and \mathbf{s}_{K+1} is equal to 0 or L . As such, assuming that $p_k \rightarrow 0$ asymptotically for all values of k , all links in the network are available and $DO = K + 1$. (ii): For $d = 1$, each state in $\{\mathbf{s}_i\}_{i=1}^K$ has one zero element corresponding to a single empty buffer while none of the elements of \mathbf{s}_{K+1} is equal to 0 or L . Consequently, for each state in $\{\mathbf{s}_i\}_{i=1}^K$, one link out of the $K + 1$ links is unavailable resulting in $DO = K$. (iii): For $d = L$, all buffers are full except for the buffer at R_{K-i+1} that contains $L - 1$ packets for $i = 1, \dots, K$ from (4.6). Therefore, for each state in $\{\mathbf{s}_i\}_{i=1}^K$, two links are available; namely, link R_K -D and R_{K-i} - R_{K-i+1} implying that the corresponding summation in (4.8) is equal to 2. For \mathbf{s}_{K+1} , all buffers are full and only link R_K -D is available implying that the corresponding summation in (4.8) is equal to 1. As a conclusion, $DO = \min\{1, 2\} = 1$.