
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Nomikos, Nikolaos; Charalambous, Themistoklis; Vouyioukas, Demosthenes; Karagiannidis, George K.

When buffer-aided relaying meets full duplex and NOMA

Published in:
IEEE WIRELESS COMMUNICATIONS MAGAZINE

DOI:
[10.1109/MWC.001.2000223](https://doi.org/10.1109/MWC.001.2000223)

Published: 01/02/2021

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Nomikos, N., Charalambous, T., Vouyioukas, D., & Karagiannidis, G. K. (2021). When buffer-aided relaying meets full duplex and NOMA. *IEEE WIRELESS COMMUNICATIONS MAGAZINE*, 28(1), 68-73. Article 9363027. <https://doi.org/10.1109/MWC.001.2000223>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

© 2021 IEEE. This is the author's version of an article that has been published by IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

When Buffer-Aided Relaying Meets Full-Duplex and NOMA

Nikolaos Nomikos, *Senior Member, IEEE*, Themistoklis Charalambous, *Senior Member, IEEE*, Demosthenes Vouyioukas, *Senior Member, IEEE*, and George K. Karagiannidis, *Fellow, IEEE*

Abstract—Full-duplex (FD) systems are able to transmit and receive signals over the same frequency band simultaneously with the potential of even doubling the spectral efficiency in comparison with the traditional half-duplex (HD) systems. When combined with relaying, FD systems are expected to dramatically increase the throughput of future wireless networks. However, the degrading effect of self-interference (SI) due to the simultaneous transmission and reception at the relay threatens their efficient roll-out in real-world topologies. At the same time, non-orthogonal multiple access (NOMA) can further increase the spectral efficiency of the network. In this article, we present how buffering at the relays can be integrated with FD and NOMA in order to improve the performance of relay networks. More specifically, the key points for the successful integration of buffer-aided relays with spatial sharing paradigms are presented and details on buffer state information-based relay selection and its fully distributed implementation are discussed. Furthermore, a mathematical framework for the analysis of FD and HD buffer-aided relay networks is rigorously discussed. Performance evaluation shows the importance of data buffering towards mitigating SI and improving the sum-rate of the network. Finally, the goal of this work is both to summarize the current state-of-the-art and, by calling attention to open problems, to spark interest towards targeting these and related problems in relay networks.

Index Terms—Full-duplex relaying, buffer-aided relaying, non-orthogonal multiple access, relay selection.

I. INTRODUCTION

The tremendous increase of mobile data traffic stemming from both users and Internet-of-Things (IoT) devices asks for the development of novel transmission algorithms which offer increased capacity and low end-to-end latency. *Frequency*, *time*, and *space* are the three fundamental dimensions for the presence of an electromagnetic (EM) signal. Additionally, *coding* (a combination of frequency and time into a set of codes for signal isolation), while not a unique dimension, is widely used in communication systems. Finally, the *power of a signal* affects the level of that signal in space, thus affecting the spatial dimension.

Recent communication paradigms have been targeting efficient resource usage by departing from the orthogonal alloca-

tion of resources in time and frequency. Among these techniques, full-duplex (FD) communication and non-orthogonal multiple access (NOMA) have attracted significant attention. Another important technique to improve the quality of wireless transmissions is cooperative relaying. In recent years, a flurry of research has spurred, resulting in novel techniques, such as opportunistic relay selection (ORS) using half-duplex (HD) and FD relays [1]. FD, NOMA and ORS are forms of spatial sharing paradigms that invest in interference mitigation techniques (such as coding and power allocation) for reducing the leakage of EM signals to nearby spatial regions.

Focusing on cooperative relay networks, FD relays enable the simultaneous reception and transmission at the cost of self-interference (SI) [2]. Moreover, the use of at least two relays results in the formation of successive relaying (SuR) mimicking FD operation and introducing inter-relay interference (IRI). At the same time, FD NOMA in the power domain allows the concurrent transmission to/from more than one users and reception of the source's signal by the same relay by appropriately allocating the available power to the different user signals exploiting their rate and channel asymmetries [3]–[5]. However, the performance of wireless networks integrating FD relaying, NOMA, and SuR is compromised due to the inherent presence of interference and the inability to completely eliminate or significantly reduce it with existing techniques.

In recent years, buffer-aided (BA) relays have been shown to improve the wireless conditions through increased diversity, thus resulting in reduced outages and better throughput; see related surveys in [6] and [7]. The exploitation of data buffers in cooperative relay networks decouples the requirement of successful packet reception in the same time-slot, as relays can store previously received packets, thus increasing the degrees of freedom of relay scheduling. Furthermore, low packet delay can be achieved, by considering the delay constraint of the application [8] in the relaying algorithms. Since buffer-aided relays have FD capabilities, several researchers have proposed adaptive algorithms to improve the throughput under statistical delay constraints and source-destination connectivity [9] or by considering the channel state information (CSI) of the receiving, transmitting, and self-interference channels [10]. Focusing on BA ORS algorithms, the survey in [7] presents several cases where HD, FD and SuR modes of operation were combined, showing that significant gains can be harvested when hybrid algorithms are adopted, efficiently combining the benefits of all the available relaying modes.

In this work, we aim at highlighting the benefits and potentials of adopting buffer-aided relays in full-duplex cooperative

N. Nomikos and D. Vouyioukas are with the Department of Information and Communication Systems Engineering, University of the Aegean, Samos, Greece. *Emails:* {nnomikos, dvouyiou}@aegean.gr.

T. Charalambous is with the School of Electrical Engineering, Aalto University, Espoo, Finland. *Emails:* name.surname@aalto.fi.

G. K. Karagiannidis is with the Wireless Communications and Information Processing (WCIP) Group, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece. *Emails:* geokarag@auth.gr.

relay networks and provide a framework for their theoretical analysis using Markov chains. Towards this end, we focus on the following issues:

- i) The various benefits of data buffering in FD relay networks are discussed, as well as the different relay modes of operation.
- ii) Details on the integration of BA relays in FD NOMA networks are provided.
- iii) A general theoretical framework relying on Markov chains is presented aiming to facilitate the analysis of BA relay networks.
- iv) Performance evaluation is presented, highlighting how data buffering enhances FD relay networks concurrently serving multiple users through NOMA.
- v) A wide range of challenging open problems that merit further attention is discussed.

II. BUFFER-AIDED FULL-DUPLEX SCHEMES

The use of relays with FD capabilities can alleviate the main limitation of conventional relays in two-hop networks, i.e. the HD constraint where twice the radio-resources must be spent to transmit the same amount of data, compared to single-hop transmission. On the downside, interference from the relay's output antenna to the relay's input antenna through the SI channel leads to performance degradation, whose level defines whether or not, full-duplexity is beneficial for the network. In the following, the beneficial role of buffering in FD relay networks will be highlighted, compared to networks with relays without buffers.

A. Buffer-aided full-duplex relaying

Consider a two-hop system in which a source S communicates to a destination D via a cluster of relays \mathcal{C} . Each relay $R_k \in \mathcal{C}$ is able to operate in FD mode, by which information can be transmitted to and from a point in two directions simultaneously on the same physical channel. In general, we do not consider ideal FD operation only, but instead we allow for SI at the relay. Various techniques to mitigate SI have been proposed in [2]; these include not only natural antenna isolation but also interference cancellation and suppression. Moreover, connectivity between the source and the destination is only established via the relays, due to severe fading in the direct link. Fig. 1.(a) depicts a network comprising BA relays where both SI and IRI exist. By Q_k we denote the number of packets in the buffer of relay R_k , having a maximum buffer length equal to L packets.

In a conventional FD relay network, when FD operation fails, the network can operate in HD mode in order to avoid a complete network outage and provide robustness. In this case, at least one relay must successfully receive the source's signal in the source-relay ($\{S \rightarrow R\}$) hop. Then, in the next transmission phase, the subset of relays that have received the source's signal form a set of candidate relays and one of them is scheduled to transmit to the destination. It is obvious that there exists a coupling between the reception and transmissions phases when relays are not able to store packets that were received in a phase different than the preceding one.

With the adoption of buffering, this decoupling is broken and different relays can be selected for transmission and reception, thus increasing the degrees of freedom of selection. Buffering is a promising solution for cooperative networks and motivates the investigation of new protocols and transmission schemes, even though it may result in increased delay [6], [7]. In this context, employing BA relays enables the formulation of various hybrid algorithms that increase the performance of FD relay networks. As it can be seen on Fig. 1.(b), independently of whether or not one or more relays are able to decode the source's signal in the $\{S \rightarrow R\}$ hop, there exist packets in their buffers from previous transmission phases. In this way, both hops can be examined for selection, dramatically increasing the diversity while throughput performance is maintained, as packets residing in the buffers can be scheduled for transmission [11].

Another benefit of buffering is related to the flexibility it provides for algorithmic design. When relays have packets in their buffers, relay scheduling can be initiated from the relay-destination ($\{R \rightarrow D\}$) hop while relying on broadcasting for the source's transmission. In this way, centralized CSI acquisition and processing is avoided compared to FD relaying without buffers, where full CSI is a priori necessary.

B. Buffer-aided successive relaying

When concurrent transmission and reception through the same relay is not possible, the efficiency of radio-resource usage can be maintained by adopting SuR. SuR mimics FD relaying, as one relay receives the source's signal, while another forwards a previously received signal to the destination. Nevertheless, SuR, similar to the presence of SI in FD, may suffer from IRI introduced by the simultaneous transmissions. To reduce performance degradation, efficient IRI mitigation techniques must be employed. Towards this end, BA relays can offer tremendous opportunities by harvesting increased diversity, as multiple relays might already store previously received packets.

Equipping relays with buffers increases the chances of mitigating the degrading effect of IRI. More specifically, the potential candidates for relay selection in the $\{R \rightarrow D\}$ hop exceed those of SuR without buffers, where relays can only store one packet from the previous transmission phase. At the same time, the number of available relays for reception in the $\{S \rightarrow R\}$ hop will be equal to $K - 1$, i.e. only the relay that will transmit towards the destination is excluded. It is evident that when relays can only store a single packet, the diversity of the reception in the $\{S \rightarrow R\}$ hop is significantly undermined. So, BA relaying offers enhanced diversity towards combatting IRI as interference avoidance or interference cancellation are facilitated by optimally selecting relay-pairs in each transmission phase.

III. INTEGRATING BUFFERS IN FULL-DUPLEX NOMA RELAY NETWORKS

Buffering can improve the performance of FD relay networks, facilitating multiple access, and especially in NOMA,

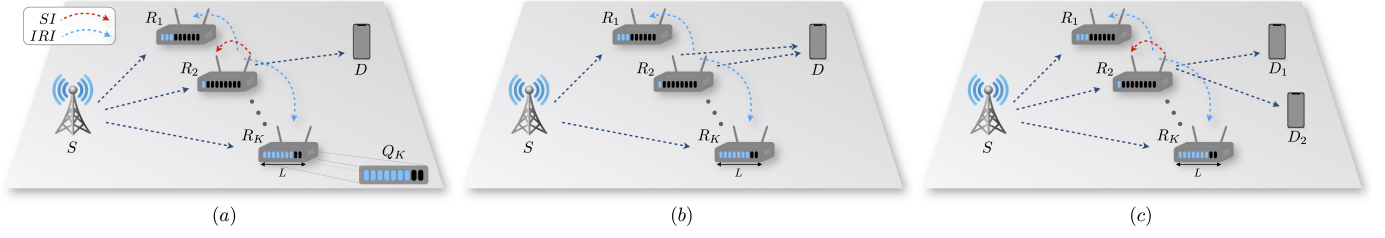


Fig. 1. The topology of a buffer-aided full-duplex relay network: (a) Full-duplex relaying, (b) successive relaying and (c) full-duplex NOMA transmission.

the chances for optimal power allocation and successful reception by multiple users are increased. Moreover, the negative impact of buffers on the end-to-end delay can be minimized through intelligent relay scheduling. Finally, in networks consisting of multiple relays and users, distributed coordination should be guaranteed as overheads might increase.

A. Buffer state-based relay selection

In OMA networks there have been several works where the selection of a relay in each transmission phase depends on both CSI and buffer state information (BSI). In conventional ORS schemes without buffers, such as [?], CSI indicates whether or not a wireless link can support the desired rate, being the main parameter of the selection policy. With the addition of buffers, exploiting BSI has led to advanced algorithms providing important trade-offs, such as promoting delay- or diversity-awareness. More specifically, depending on the target of the selection policy, priority can be given towards emptying the buffers and reducing the end-to-end delay or setting thresholds in terms of the maximum and minimum number of packets in the buffers and guarantee that more links will be available during the selection.

Recently, buffering has been considered in NOMA networks, relying on HD relays. Most works have proposed variations of equivalent BA OMA policies, leveraging BSI to ensure the diversity of the NOMA transmission or switching among OMA/NOMA in order to maintain the average sum-rate and reduce the average delay [12]. Since BA relays store packets and forward them at a later instant, it is possible to enable seamless switching among HD and FD transmissions while serving multiple users through NOMA. Such a solution can be applied in BA NOMA networks comprising relays with FD capabilities, as the one shown in Fig. 1.(c). In an arbitrary transmission phase, each relay might have a random number of packets residing in its buffer. A simple yet efficient policy exploiting buffering in FD relay networks involves, first examining the $\{R \rightarrow D\}$ hop for possible transmission based on a pre-determined target, e.g. reduced delay, accordingly selecting a relay. Then, this relay might be able to receive a packet from the source, resulting in this way in FD operation. Alternatively, a different relay can be scheduled for reception, e.g. the one with the least number of packets in its buffer, mimicking FD operation through SuR. If an $\{R \rightarrow D\}$ transmission is not possible due to empty buffers or severe fading, HD interference-free $\{S \rightarrow R\}$ reception will occur.

Different policies can be based on data buffers and BSI in NOMA networks in cases where users have different

packet priorities and delay constraints. So, an effective way of leveraging the potential of buffering would be to perform dynamic scheduling of the different users allocating more power to increase the transmission rate of the user having packets that might expire. Furthermore, switching to FD OMA when NOMA is infeasible, dynamically scheduling the user with high priority service can enable scenarios with coexisting users and IoT devices and heterogeneous traffic.

B. Distributed relay selection

In networks with a large number of nodes, such as those supporting device-to-device communication or IoT applications, distributed network operation can safeguard the performance of BA FD relaying by reducing the delay between CSI estimation and relay activation. A popular method for the distributed operation of relay selection is based on the use of synchronized timers at each relay. Fig. 2 depicts the three steps of distributed relay selection:

Step 1 – Preamble and Contention: Both destinations broadcast a pilot and each relay R_j , assuming channel reciprocity, first determines its $\{R \rightarrow D\}$ links. Then, a power allocation factor α_j for NOMA is chosen, adopting the method proposed in [12], based on instantaneous CSI. Next, among the non-empty relays that can transmit towards the destinations, the one having the maximum queue length is selected for $\{R \rightarrow D\}$ NOMA transmission. This process can be implemented in a fully distributed fashion, using timers: each relay R_j competes for channel access by setting its timer value proportional to $(Q_j + 1 + \nu_j)^{-1}$, where ν_j is uniformly distributed in $(-0.5, 0.5)$. As a result, the relay whose timer has the *maximum* buffer size expires first. If a tie, due to equal buffer sizes occurs, random variable ν_j guarantees different expiration times. So, a flag packet is transmitted by the selected relay, notifying the other relays to back-off and remain silent.

Step 2 – Transmission: Source S broadcasts the combined signals for D_1 and D_2 with a rate equal to the sum of their desired, while the transmitting relay R_j (if there is a feasible one) simultaneously transmits with one of its antennas to the destinations, leading to FD operation, whenever S and R_j are activated at the same time. All the remaining $K - 1$ relays use both antennas to receive the source's packets.

Step 3 – ACK/NACK: Once the destinations receive the packets from the transmitting relay and the broadcasting from the source is over, each of them in turn broadcasts an ACK that includes the identity (ID) of the packet received, so that the relays that have stored the same packets in their buffers (due to

the source broadcasting) discard them before the beginning of the next time frame, to avoid retransmitting the same packets unnecessarily and clogging up their buffers with obsolete packets. Note that in imperfect conditions where the ACKs are not received by some relays, the overall performance of the network will degrade.

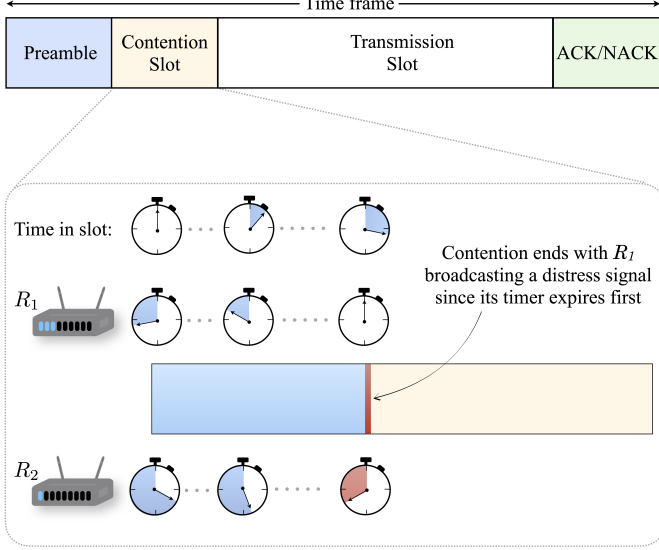


Fig. 2. A simple example of how the timer-based approach works for a network of two relays. The relay with more packets in its buffer has its timer expire first during the contention period.

IV. MARKOV CHAIN-BASED FRAMEWORK

In general, communication networks consisting of nodes equipped with finite (and infinite) buffer sizes have been traditionally modeled using Markov Chains (MCs). The mathematical analysis of cooperative relaying with buffers has also been done using MC extensively in the literature. We provide a brief exposition of the framework because (i) it is a very powerful tool that provides simple ways of analyzing the performance of the overall system, and (ii) a simple modification on the classical construction of the MC, if we additionally have FD transmissions, facilitates the analysis of networks with FD relaying.

Since the buffers are of finite size, one can represent the buffer states with the states of a MC. The theoretical framework introduced in [11] was based on the idea that the transitions between the states are given by the probabilities of successful transmissions of packets. For example, if we consider a network with 2 relays with a buffer size of 2 packets, then there will be 9 possible states. When only one transmission per time-slot takes place, the MC is depicted in Fig. 3(a) with the transition probabilities represented as black edges. When, SuR is facilitated, then transitions among more states are achieved (additional blue lines in Fig. 3(a)). Therefore, if the current state of the buffers (and, therefore, the state of the MC) does not change, then it is evident that no transmission has occurred. Such MCs due their structure are stochastic, irreducible and aperiodic (SIA), and as a result, they have a steady state distribution. Being SIA, the MC has several

properties that allow the extraction of simple expressions for the outage probability, the average throughput and average delay, among others.

The main challenge of the analysis of FD systems, compared to the framework described, is that if we additionally have FD transmissions, the successful transmissions cannot be captured by the aforementioned model, since the buffer state does not change (and the buffer change is the one showing successful transmissions). To capture FD transmissions, a state of the MC is represented by the number of elements at each buffer, combined now with a change of state at the destination, every time a packet is received *via FD operation* from the source. This representation can be easily visualized by adding an extra layer in the MC. For our example of a network with 2 relays with a buffer size of 2 packets, this is illustrated in Fig. 3(b).

V. PERFORMANCE EVALUATION

The performance of BA FD NOMA is evaluated for $K = 3$ relays and varying buffer size L , in a topology with a saturated source S where time is divided into “frames” of one packet duration (e.g., fixed-size packets). For the link $\{i \rightarrow j\}$, the channel gains $g_{ij} \triangleq |h_{ij}|^2$ are exponentially distributed, i.e., $g_{ij} \sim \text{Exp}(\sigma_{ij}^{-2}/2)$. At each time-slot, S and/or one of the relays R_k transmit with fixed power levels P_i , $i \in \{S, R_1, \dots, R_K\}$. In the comparisons, the x-axis values correspond to the maximum transmit power per node.

As FD relaying is supported, SI exists and the instantaneous residual SI channel $h_{R_k R_k}$ between the two antennas of relay R_k follows a complex Gaussian distribution, taking values in the range $(0, \sigma_{R_k R_k}^2)$. The SI channel of the k -th relay is characterized by $\sigma_{R_k R_k}^2 = 10^{-3} \sigma_{S R_k}^2$, unless otherwise stated, while the noise level is equal to 1 mW and the users’ channels are independent non identically distributed (i.n.i.d). Moreover, the two users require equal rates $r_1 = r_2 = r_D = 2$ bits-per-channel-use (BPCU) and their channel asymmetry is defined as $\sigma_{S R_k}^2 = \sigma_{R_k D_1}^2 = 4 \sigma_{R_k D_2}^2$. BA FD NOMA is compared to BA HD NOMA with $\{R \rightarrow D\}$ prioritization (NOMA RD) and the hybrid BA-HD-NOMA/OMA of [12], where if the NOMA transmission fails, a user is selected in the $\{R \rightarrow D\}$ link to be served with its desired rate.

Fig. 4 depicts the average sum-rate performance for different schemes. It can be observed that BA FD NOMA can significantly increase the average sum-rate for high transmit power, while for low and medium transmit power values, BA HD NOMA/OMA offers the best performance, due to its ability to switch to OMA transmission in the $\{R \rightarrow D\}$ link, when NOMA fails. At the same time, it does not suffer from SI, as it is an HD scheme. When FD transmissions are performed only through SuR, the average sum-rate is improved after 10 dBm, compared to the HD schemes, but is significantly lower than BA FD NOMA with FD relays. Finally, NOMA with $\{R \rightarrow D\}$ prioritization exhibits identical performance with BA FD NOMA, since for low transmit power the latter’s operation is dominated by HD relaying. Then, as transmit power increases and the chances for successful power allocation through NOMA increase, NOMA RD reaches the HD upper bound.

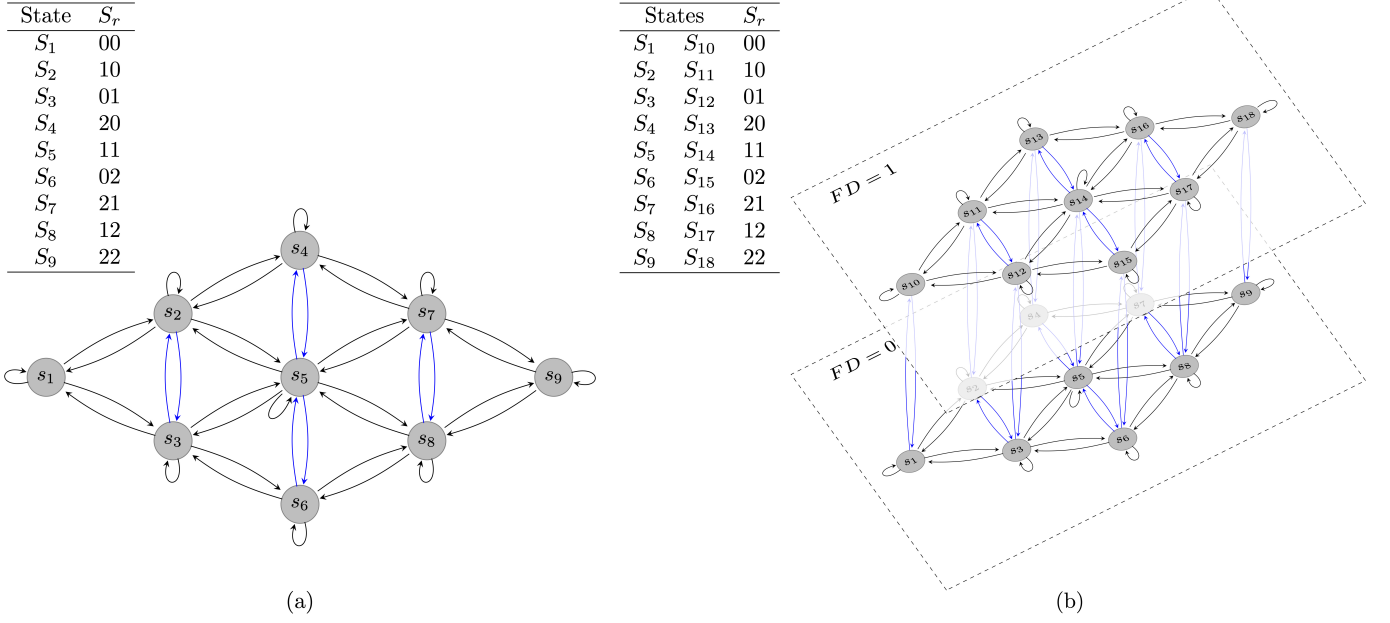


Fig. 3. Markov chain for different scenarios: (a) When no FD transmissions are available (e.g., because of HD relays), single transmissions within a slot are represented by black lines, whereas SuR with blue lines. (b) Full-duplex relaying requires one more level to depict the transmission, due to the fact that the state of the buffers remains unchanged.

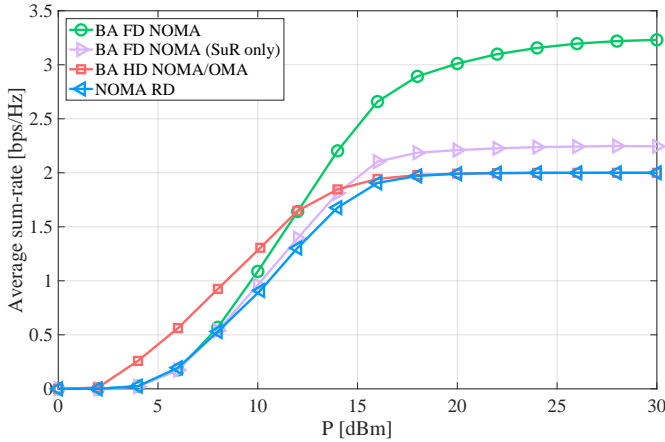


Fig. 4. Average sum-rate for $K = 3$, $L = 4$ and various algorithms.

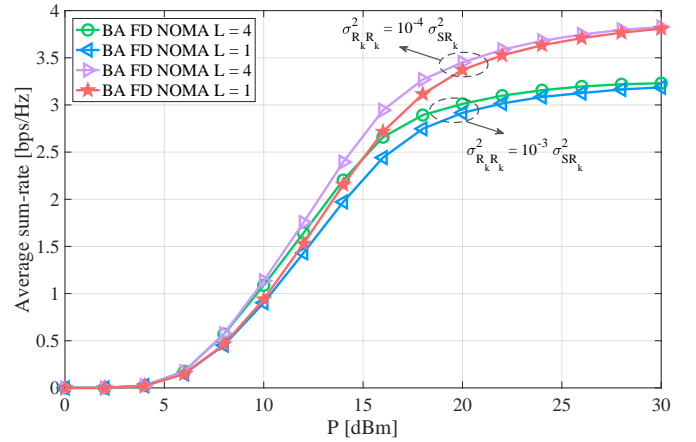


Fig. 5. Average sum-rate for $K = 3$, varying L and SI channel conditions.

Next, Fig. 5 shows the average sum-rate performance for BA FD NOMA for varying buffer sizes and under different SI channel conditions. It is observed that the buffer size plays an important role for medium transmit power values, since for low transmit power increased outages occur due to fading, while for high transmit power, FD operation is dominant and packets are concurrently transmitted and received, thus reducing the impact of buffering. Thus, in practical topologies and operating conditions, buffering can provide an important performance boost in FD NOMA relay networks. Moreover, the effect of reduced SI channel power is obvious, as the upper bound of FD NOMA is almost reached, under weak SI.

VI. OPEN CHALLENGES

Even though buffers have shown their potential in improving the performance and robustness of FD relay networks, still, there exist various challenges that could ignite the interest of the readers.

Distributed coordination: Node density in future networks will be significantly increased, due to the coexistence of users and IoT devices. In such topologies, low-complexity and distributed network coordination and power control will accelerate the integration of BA solutions. Towards this end, the adoption of machine learning can lead to efficient algorithms without the need for full BSI and CSI knowledge. Especially in NOMA cases, learning might provide efficient solutions to the exponential growth of complexity of user pairing and power allocation, as the number of users and devices increases [13].

Interference-aware relay scheduling: The consideration of the interference characteristics in determining relay scheduling is another important issue that can further improve the performance of FD transmissions. In many cases, the SI can be much smaller than the IRI, as countermeasures can be more easily put into place. Further interference mitigation can be achieved by exploiting cloud-driven relay systems where joint signal detection takes place at the nodes and the cloud [14]. If information on the statistics of the interference channel is not available, then, learning algorithms can be used in order to infer this information.

Complex interference scenarios: More advanced cases of interference should be studied, departing from the simplified topologies that are considered in the majority of current studies. There can be cases with multiple sources as well in which the relay node may not be able to handle both the SI and the interference from multiple sources, requiring a precoding decision based solely on statistical knowledge of the channel conditions. In this context, cognitive networks represent important scenarios where buffer-aided relaying can facilitate the operation of secondary networks [15].

Flexible NOMA strategies: As it has been observed, employing hybrid FD/HD relaying can maintain the network's performance in cases of severe interference. Similarly, cases where power allocation does not allow simultaneous service to more than one users must be overcome. Thus, flexible NOMA strategies are required exploiting improved user pairing or OMA transmissions to maintain the average sum-rate of the network. Accordingly, multi-antenna nodes can provide additional degrees of freedom, improving the diversity of the reception/transmission. Additionally, the impact of the direct $\{S \rightarrow D\}$ link should be investigated. Although the provision of this additional link can improve the diversity, signals from multiple links might interfere, thus degrading the successive interference cancellation's performance at the destinations. Finally, the integration of rate-splitting multiple access (RSMA) in the proposed framework can offer additional sum-rate gains by flexibly switching between NOMA and OMA.

Intelligent data caching: Data-hungry multimedia applications will benefit from the enhanced spectral efficiency of BA FD relaying. Aiming to better support enhanced mobile broadband services, content caching can play a major role, since popular contents can be cached at different network nodes, such as user devices or FD relays. When the requested data are not located nearby, FD relays can fetch them at high speeds from a macro base station. In these cases, it is necessary to develop schemes integrating relays with buffering capabilities in FD networks jointly with intelligent data caching and optimal content location for content availability, but also interference mitigation.

VII. CONCLUSIONS

Whether or not wireless networks will be able to cope with the increased capacity requirements of coexisting users and IoT devices will depend upon the development of radical communication paradigms. Among these techniques, non-orthogonal usage of temporal and spectral radio-resources

through full-duplex transmission and reception and the allocation of the available power to simultaneously serve multiple users NOMA are quite promising. In this paper, we aimed at presenting how buffer-aided cooperative relays can significantly improve the efficiency of full-duplex multi-user networks. More specifically, different cases of buffer-aided relays and transmission algorithms were presented, while the details of a theoretical framework relying on Markov chains were given. Furthermore, performance evaluation showed the advantages of buffering in increasing the sum-rate of full-duplex NOMA networks. Finally, several open issues were discussed, highlighting possible ways for further increasing the performance of full-duplex relay networks.

REFERENCES

- [1] C. H. M. de Lima, H. Alves, P. H. J. Nardelli and M. Latva-aho, "Effects of relay selection strategies on the spectral efficiency of wireless systems with half- and full-duplex nodes," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7578–7583, Aug. 2017.
- [2] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.
- [3] M. Mohammadi, X. Shi, B. K. Chalise, Z. Ding, H. A. Suraweera, C. Zhong and J. S. Thompson, "Full-duplex non-orthogonal multiple access for next generation wireless systems," *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 110–116, May 2019.
- [4] X. Chen, G. Liu, Z. Ma, X. Zhang, P. Fan, S. Chen and F. R. Yu, "When full duplex wireless meets non-orthogonal multiple access: Opportunities and challenges," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 148–155, Aug. 2019.
- [5] A. Tregancini, E. E. B. Olivo, D. P. M. Osorio, C. H. M. de Lima and H. Alves, "Performance analysis of full-duplex relay-aided NOMA systems using partial relay selection," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 622–635, Jan. 2020.
- [6] N. Zlatanov, A. Ikhlef, T. Islam and R. Schober, "Buffer-aided cooperative communications: opportunities and challenges," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 146–153, April 2014.
- [7] N. Nomikos, T. Charalambous, I. Krikidis, D. N. Skoutas, D. Vouyioukas, M. Johansson, C. Skianis, "A survey on buffer-aided relay selection," *IEEE Commun. Surv. & Tut.*, vol. 18, no. 2, pp. 1073–1097, Secondquarter 2016.
- [8] D. Qiao and M. C. Gursoy, "Buffer-aided relay systems under delay constraints: Potentials and challenges," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 168–174, Sept. 2017.
- [9] D. Qiao, "Effective capacity of buffer-aided full-duplex relay systems With selection relaying," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 117–129, Jan. 2016.
- [10] M. Mohammadkhani Razlighi and N. Zlatanov, "Buffer-aided relaying for the two-hop full-duplex relay channel with self-interference," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 477–491, Jan. 2018.
- [11] I. Krikidis, T. Charalambous and J. S. Thompson, "Buffer-aided relay selection for cooperative diversity systems without delay constraints," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1957–1967, May 2012.
- [12] N. Nomikos, T. Charalambous, D. Vouyioukas, G. K. Karagiannidis and R. Wichman, "Hybrid NOMA/OMA with buffer-aided relay selection in cooperative networks," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 524–537, June 2019.
- [13] M. Vaezi, R. Schober, Z. Ding and H. V. Poor, "Non-orthogonal multiple access: Common myths and critical questions," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 174–180, Oct. 2019.
- [14] F. L. Duarte and R. C. de Lamare, "Cloud-driven multi-way multiple-antenna relay systems: Joint detection, best-user-link selection and analysis," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3342–3354, June 2020.
- [15] R. Zhang, R. Nakai, K. Sezaki and S. Sugiura, "Generalized buffer-state-based relay selection in cooperative cognitive radio networks," *IEEE Access*, vol. 8, pp. 11644–11657, 2020.